Classification of Finite-Dimensional Simple Lie Algebras over the Complex Numbers

by

Shawn Baland

A thesis submitted to the

Faculty of the University of Colorado

in partial fulfillment of the requirements

for graduation with departmental honors

Department of Mathematics

November 12, 2007

Updated December 6, 2007

Defense Committee

Prof. Arlan Ramsay, Mathematics (Thesis Advisor)

Prof. Richard Clelland, Mathematics

Prof. James Meiss, Applied Mathematics

Baland, Shawn

Classification of Finite-Dimensional Simple Lie Algebras over the Complex Numbers

Thesis directed by Prof. Arlan Ramsay

Lie algebras were first introduced by S. Lie in 1880 as algebraic structures used to study local problems in Lie groups. Specifically, the tangent space of a Lie group at the identity has the natural structure of a Lie algebra. Aside from their many applications, Lie algebras have attracted interest in their own right. This paper deals primarily with the classification of finite-dimensional simple Lie algebras over the complex numbers, first achieved through the independent work of E. Cartan and W. Killing during the decade 1890-1900. Contemporary methods of studying Lie algebras and their representations were later developed by H. Weyl.

The classification theorem begins by studying the decomposition of a semisimple Lie algebra into a direct sum of subspaces with respect to a carefully chosen subalgebra called the Cartan subalgebra. One then forms a Euclidean space from the nonzero linear functionals associated with these subspaces via an inner product called the Killing form. The geometry of this Euclidean space allows us to encode vital information about a Lie algebra's structure into a graph called the Dynkin diagram. It is then shown that there are only a limited number of possible connected Dynkin diagrams, and that these diagrams are in one-to-one correspondence with the finite-dimensional simple Lie algebras over the complex numbers. The paper's structure closely follows that of [Car05]. It is assumed that the reader has a thorough knowledge of linear algebra, although this is the only major prerequisite.

Dedication

To my mother, for supporting me when I was less ambitious.

Acknowledgments

I would first like to thank Professor Peter Elliott for encouraging me to go beyond what is usually expected of an undergraduate student. I would also like to thank Professor Richard Green for suggesting this project, and for pointing me toward many valuable resources. I owe the greatest thanks to Professor Arlan Ramsay, my thesis advisor, for spending a great deal of time answering all of my questions, and for asking many questions that I had not considered. Thanks also to Mike Noyes, my reader, for improving this paper's clarity and accuracy. Finally, I would like to thank my wife, Lyndsie, for her support and understanding throughout my undergraduate career.

Contents

Chapter

1 Basic concepts 1

1.1 Lie algebras, subalgebras, and ideals 1

1.2 Lie homomorphisms and quotient algebras 3

1.3 Nilpotent and solvable Lie algebras 6

2 Representations of solvable and nilpotent Lie algebras 11

2.1 Basic representation theory 11

2.2 A transition to finite-dimensional complex Lie algebras 14

2.3 Representations of solvable Lie algebras 14

2.4 Representations of nilpotent Lie algebras 19

3 Cartan subalgebras 26

3.1 Existence of Cartan subalgebras 26

3.2 Derivations and automorphisms 28

3.3 Concepts from algebraic geometry 30

3.4 Conjugacy of Cartan subalgebras 35

4 The Cartan decomposition 39

4.1 Root spaces 39

4.2 The Killing form 42

4.3 The Cartan decomposition of a semisimple Lie algebra 48

5 Root systems and the Weyl group 57

5.1 Positive systems and fundamental systems of roots 57

5.2 The Weyl group 61

6 The Cartan matrix and the Dynkin diagram 70

6.1 The Cartan matrix 70

6.2 The Dynkin diagram 74

6.3 Classification of Dynkin diagrams 77

6.4 Classification of Cartan matrices 83

7 The existence and uniqueness theorems 94

7.1 Properties of structure constants 94

7.2 The uniqueness theorem 100

7.3 Some generators and relations in a simple Lie algebra 103

7.4 The Lie algebras $\mathfrak{g}(A)$ and $\tilde{\mathfrak{g}}(A)$ 105

7.5 The existence theorem 114

Bibliography 133

Figures

Figure

6.1 Cartan matrices and Dynkin diagrams for $l = 1$ and 2 74

6.2 Dynkin diagrams and quadratic forms for $l = 1$ and 2 76

6.3 Standard list of connected Dynkin diagrams 84

Chapter 1

Basic concepts

1.1 Lie algebras, subalgebras, and ideals

Let k be a field. A ***k*-algebra** is a k-vector space A equipped with a bilinear operation $A \times A \to A$ called **multiplication**. The dimension of A is the dimension of A as a k-vector space.

A **Lie algebra** is a k-algebra $\mathfrak{g}$ with multiplication $(x, y) \mapsto [x, y]$ satisfying the following conditions:

(i) $[x, x] = 0$ for all $x \in \mathfrak{g}$.

(ii) $[[x, y], z] + [[y, z], x] + [[z, x], y] = 0$ for all $x, y, z \in \mathfrak{g}$.

Condition (ii) is called the **Jacobi identity**.

Proposition 1.1. $[y, x] = -[x, y]$ *for all* $x, y \in \mathfrak{g}$*, i.e., multiplication in a Lie algebra is anticommutative.*

Proof. If $x, y \in \mathfrak{g}$, then

$$\begin{aligned} 0 = [x + y, x + y] &= [x, x] + [x, y] + [y, x] + [y, y] \\ &= 0 + [x, y] + [y, x] + 0. \end{aligned}$$

It follows that $[y, x] = -[x, y]$. □

Notice that anticommutativity implies $[x,x] = -[x,x]$ for all $x \in \mathfrak{g}$, yielding $2[x,x] = 0$. Hence if $\operatorname{char} k \neq 2$, then anticommutativity is equivalent to condition (i) for a Lie algebra.

Example 1.2. An **associative algebra** is a k-algebra A satisfying the associative law, that is,

$$(xy)z = x(yz) \qquad \text{for all } x, y, z \in A.$$

Lie algebras are easily obtained from associative algebras in the following way: Given any associative algebra A, we define the **commutator** of $x, y \in A$ by

$$[x,y] = xy - yx.$$

The commutator is clearly bilinear. Moreover, $[x,x] = xx - xx = 0$ and

$$\begin{aligned} [[x,y],z] + [[y,z],x] + [[z,x],y] &= (xy-yx)z - z(xy-yx) + (yz-zy)x - x(yz-zy) \\ &\quad + (zx-xz)y - y(zx-xz) = 0 \end{aligned}$$

for all $x, y, z \in A$. Thus A forms a Lie algebra via the commutator product. We call this algebra the **Lie algebra of $\boldsymbol{A}$** and denote it by $[A]$.

If $\mathfrak{a}$ and $\mathfrak{b}$ are subspaces of a Lie algebra $\mathfrak{g}$, we define $[\mathfrak{a},\mathfrak{b}]$ to be the subspace of $\mathfrak{g}$ generated by all elements of the form $[x,y]$ with $x \in \mathfrak{a}$ and $y \in \mathfrak{b}$.

Proposition 1.3. $[\mathfrak{a},\mathfrak{b}] = [\mathfrak{b},\mathfrak{a}]$ *for all subspaces* $\mathfrak{a}, \mathfrak{b}$ *of* $\mathfrak{g}$.

Proof. Any element in $[\mathfrak{a},\mathfrak{b}]$ is of the form

$$\lambda_1[x_1,y_1] + \cdots + \lambda_r[x_r,y_r]$$

where $x_i \in \mathfrak{a}$, $y_i \in \mathfrak{b}$, $\lambda_i \in k$ for $i = 1,\ldots,r$. Now $[x_i,y_i] = -[y_i,x_i] \in [\mathfrak{b},\mathfrak{a}]$ by Proposition 1.1. It follows that $[\mathfrak{a},\mathfrak{b}] \subset [\mathfrak{b},\mathfrak{a}]$. By a symmetric argument, we also have $[\mathfrak{b},\mathfrak{a}] \subset [\mathfrak{a},\mathfrak{b}]$, and so $[\mathfrak{a},\mathfrak{b}] = [\mathfrak{b},\mathfrak{a}]$. □

A **subalgebra** of a Lie algebra $\mathfrak{g}$ is a subspace $\mathfrak{h} \subset \mathfrak{g}$ such that $[\mathfrak{h},\mathfrak{h}] \subset \mathfrak{h}$. Hence $\mathfrak{h}$ forms a Lie algebra by restricting the multiplication on $\mathfrak{g}$ to the subspace $\mathfrak{h}$.

Proposition 1.4. *If* $\mathfrak{h}$ *and* $\mathfrak{k}$ *are subalgebras of* $\mathfrak{g}$*, then* $\mathfrak{h} \cap \mathfrak{k}$ *is a subalgebra of* $\mathfrak{g}$.

Proof. We know that $\mathfrak{h} \cap \mathfrak{k}$ is a subspace of $\mathfrak{g}$ from linear algebra. Now $[\mathfrak{h} \cap \mathfrak{k}, \mathfrak{h} \cap \mathfrak{k}]$ is generated by elements of the form $[x, y]$ with $x, y \in \mathfrak{h} \cap \mathfrak{k}$. We have $x, y \in \mathfrak{h}$ and $x, y \in \mathfrak{k}$ so that $[x, y] \in [\mathfrak{h}, \mathfrak{h}] \cap [\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{h} \cap \mathfrak{k}$. It follows that $[\mathfrak{h} \cap \mathfrak{k}, \mathfrak{h} \cap \mathfrak{k}] \subset \mathfrak{h} \cap \mathfrak{k}$, and so $\mathfrak{h} \cap \mathfrak{k}$ is a subalgebra of $\mathfrak{g}$. □

An **ideal** of $\mathfrak{g}$ is a subspace $\mathfrak{a} \subset \mathfrak{g}$ such that $[\mathfrak{a}, \mathfrak{g}] \subset \mathfrak{a}$. This condition is equivalent to $[\mathfrak{g}, \mathfrak{a}] \subset \mathfrak{a}$ by Proposition 1.3, and so every ideal of a Lie algebra is two-sided. Note also that $[\mathfrak{a}, \mathfrak{a}] \subset \mathfrak{a}$ since $\mathfrak{a} \subset \mathfrak{g}$. Thus ideals are also subalgebras.

Proposition 1.5. *If* $\mathfrak{a}$ *and* $\mathfrak{b}$ *are ideals of* $\mathfrak{g}$*, then* $\mathfrak{a} \cap \mathfrak{b}$*,* $\mathfrak{a} + \mathfrak{b}$*, and* $[\mathfrak{a}, \mathfrak{b}]$ *are ideals of* $\mathfrak{g}$.

Proof. We know that these sets are subspaces of $\mathfrak{g}$ from linear algebra. For the first part, notice that $[\mathfrak{a} \cap \mathfrak{b}, \mathfrak{g}]$ is generated by elements of the form $[x, y]$ with $x \in \mathfrak{a} \cap \mathfrak{b}$, $y \in \mathfrak{g}$. We have $x \in \mathfrak{a}$ and $x \in \mathfrak{b}$ so that $[x, y] \in [\mathfrak{a}, \mathfrak{g}] \cap [\mathfrak{b}, \mathfrak{g}] \subset \mathfrak{a} \cap \mathfrak{b}$. It follows that $[\mathfrak{a} \cap \mathfrak{b}, \mathfrak{g}] \subset \mathfrak{a} \cap \mathfrak{b}$, and so $\mathfrak{a} \cap \mathfrak{b}$ is an ideal of $\mathfrak{g}$.

For the second part, notice that $[\mathfrak{a} + \mathfrak{b}, \mathfrak{g}]$ is generated by elements of the form $[x+y, z]$ with $x \in \mathfrak{a}$, $y \in \mathfrak{b}$, $z \in \mathfrak{g}$. We have $[x+y, z] = [x, z]+[y, z] \in [\mathfrak{a}, \mathfrak{g}]+[\mathfrak{b}, \mathfrak{g}] \subset \mathfrak{a}+\mathfrak{b}$. It follows that $[\mathfrak{a} + \mathfrak{b}, \mathfrak{g}] \subset \mathfrak{a} + \mathfrak{b}$, and so $\mathfrak{a} + \mathfrak{b}$ is an ideal of $\mathfrak{g}$.

For the third part, notice that $[[\mathfrak{a}, \mathfrak{b}], \mathfrak{g}]$ is generated by elements of the form $[[x, y], z]$ with $x \in \mathfrak{a}$, $y \in \mathfrak{b}$, $z \in \mathfrak{g}$. We have $[[x, y], z] = [[x, z], y] - [[y, z], x]$ by the Jacobi identity. Now $[x, z] \in \mathfrak{a}$ and $[y, z] \in \mathfrak{b}$ since $\mathfrak{a}$ and $\mathfrak{b}$ are ideals. Thus $[[x, y], z] \in [\mathfrak{a}, \mathfrak{b}]$. It follows that $[[\mathfrak{a}, \mathfrak{b}], \mathfrak{g}] \subset [\mathfrak{a}, \mathfrak{b}]$, and so $[\mathfrak{a}, \mathfrak{b}]$ is an ideal of $\mathfrak{g}$. □

1.2 Lie homomorphisms and quotient algebras

As with Lie subalgebras and ideals, the concepts of Lie homomorphisms and quotient algebras are analogous to those of associative algebras. Let $\mathfrak{g}, \mathfrak{h}$ be Lie algebras over a field k. A map $\phi : \mathfrak{g} \to \mathfrak{h}$ is a **homomorphism of Lie algebras** if ϕ is k-linear and

$$\phi([x,y]) = [\phi(x), \phi(y)] \qquad \text{for all } x, y \in \mathfrak{g}.$$

An **isomorphism of Lie algebras** is a bijective homomorphism of Lie algebras. If there exists an isomorphism $\phi : \mathfrak{g} \to \mathfrak{h}$, then we say $\mathfrak{g}$ and $\mathfrak{h}$ are **isomorphic** and write $\mathfrak{g} \cong \mathfrak{h}$.

Proposition 1.6. *If $\mathfrak{g}$ is a Lie algebra and $\mathfrak{a}$ is an ideal of $\mathfrak{g}$, then the quotient space $\mathfrak{g}/\mathfrak{a}$ forms a Lie algebra via the multiplication*

$$[x + \mathfrak{a}, y + \mathfrak{a}] = [x, y] + \mathfrak{a} \qquad \textit{for all } x, y \in \mathfrak{g}.$$

Proof. To prove this multiplication is well defined, suppose $x + \mathfrak{a} = x' + \mathfrak{a}$ and $y + \mathfrak{a} = y' + \mathfrak{a}$. We must show that $[x, y] + \mathfrak{a} = [x', y'] + \mathfrak{a}$. Now $x + \mathfrak{a} = x' + \mathfrak{a}$ implies $x = x' + a_1$ for some $a_1 \in \mathfrak{a}$. Likewise, $y + \mathfrak{a} = y' + \mathfrak{a}$ implies $y = y' + a_2$ for some $a_2 \in \mathfrak{a}$. Thus

$$[x, y] + \mathfrak{a} = [x' + a_1, y' + a_2] + \mathfrak{a} = [x', y'] + [x', a_2] + [a_1, y'] + [a_1, a_2] + \mathfrak{a}.$$

The last three summands are in $\mathfrak{a}$ since $\mathfrak{a}$ is an ideal, and so $[x, y] + \mathfrak{a} = [x', y'] + \mathfrak{a}$. Next, we have

$$[x + \mathfrak{a}, x + \mathfrak{a}] = [x, x] + \mathfrak{a} = 0 + \mathfrak{a} = \mathfrak{a} \qquad \text{for all } x \in \mathfrak{a}.$$

Finally, the Jacobi identity follows from the Jacobi identity in $\mathfrak{g}$. ☐

Proposition 1.7. *Let $\phi : \mathfrak{g} \to \mathfrak{h}$ be a homomorphism of Lie algebras. Then $\operatorname{im}\phi$ is a subalgebra of $\mathfrak{h}$, $\ker\phi$ is an ideal of $\mathfrak{g}$, and $\mathfrak{g}/\ker\phi \cong \operatorname{im}\phi$.*

Proof. We know that $\operatorname{im}\phi$ is a subspace of $\mathfrak{h}$ from linear algebra. If $x, y \in \mathfrak{g}$, then

$$[\phi(x), \phi(y)] = \phi([x, y]) \in \operatorname{im}\phi.$$

This shows that $[\operatorname{im}\phi, \operatorname{im}\phi] \subset \operatorname{im}\phi$, and so $\operatorname{im}\phi$ is a subalgebra of $\mathfrak{h}$. We also know that $\ker\phi$ is a subspace of $\mathfrak{g}$. If $x \in \ker\phi$ and $y \in \mathfrak{g}$, then

$$\phi([x, y]) = [\phi(x), \phi(y)] = [0, \phi(y)] = 0,$$

showing that $[x, y] \in \ker \phi$. It follows that $[\ker \phi, \mathfrak{g}] \subset \ker \phi$, and so $\ker \phi$ is an ideal of $\mathfrak{g}$. To prove the last statement, consider the linear map

$$\mathfrak{g}/\ker \phi \to \operatorname{im} \phi$$

$$x + \ker \phi \mapsto \phi(x).$$

This map is well defined since

$$x + \ker \phi = y + \ker \phi \;\Leftrightarrow\; x - y \in \ker \phi \;\Leftrightarrow\; \phi(x - y) = 0 \;\Leftrightarrow\; \phi(x) = \phi(y).$$

The leftward implication proves that the map is injective. If $\phi(x) \in \operatorname{im} \phi$, then $x + \ker \phi$ maps to $\phi(x)$, showing that the map is surjective. Finally, we prove that the map is a homomorphism of Lie algebras. If $x, y \in \mathfrak{g}$, then

$$[x + \ker \phi, y + \ker \phi] \;=\; [x, y] + \ker \phi \;\mapsto\; \phi([x, y]) \;=\; [\phi(x), \phi(y)].$$

Hence the map respects Lie multiplication, and so it is an isomorphism of Lie algebras.

□

Lemma 1.8. *If $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$ and $\mathfrak{a}$ is an ideal of $\mathfrak{g}$, then $\mathfrak{h} + \mathfrak{a}$ is a subalgebra of $\mathfrak{g}$.*

Proof. We know that $\mathfrak{h} + \mathfrak{a}$ is a subspace of $\mathfrak{g}$ from linear algebra. Now $[\mathfrak{h} + \mathfrak{a}, \mathfrak{h} + \mathfrak{a}]$ is generated by elements of the form $[x + y, x' + y']$ with $x, x' \in \mathfrak{h}$ and $y, y' \in \mathfrak{a}$. We have

$$\begin{aligned} [x + y, x' + y'] &= [x, x'] + [x, y'] + [y, x'] + [y, y'] \\ &\in [\mathfrak{h}, \mathfrak{h}] + ([\mathfrak{h}, \mathfrak{a}] + [\mathfrak{a}, \mathfrak{h}] + [\mathfrak{a}, \mathfrak{a}]) \\ &\subset \mathfrak{h} + \mathfrak{a}. \end{aligned}$$

It follows that $[\mathfrak{h} + \mathfrak{a}, \mathfrak{h} + \mathfrak{a}] \subset \mathfrak{h} + \mathfrak{a}$, and so $\mathfrak{h} + \mathfrak{a}$ is a subalgebra of $\mathfrak{g}$. □

Proposition 1.9. *Let $\mathfrak{h}$ be a subalgebra of $\mathfrak{g}$ and $\mathfrak{a}$ an ideal of $\mathfrak{g}$. Then $\mathfrak{a}$ is an ideal of $\mathfrak{h} + \mathfrak{a}$, $\mathfrak{h} \cap \mathfrak{a}$ is an ideal of $\mathfrak{h}$, and $(\mathfrak{h} + \mathfrak{a})/\mathfrak{a} \cong \mathfrak{h}/(\mathfrak{h} \cap \mathfrak{a})$.*

Proof. We know that $\mathfrak{h}+\mathfrak{a}$ and $\mathfrak{h}\cap\mathfrak{a}$ are subalgebras of $\mathfrak{g}$ by Lemma 1.8 and Proposition 1.4, respectively. We have $[\mathfrak{a}, \mathfrak{h}+\mathfrak{a}] \subset [\mathfrak{a}, \mathfrak{g}] \subset \mathfrak{a}$, and so $\mathfrak{a}$ is an ideal of $\mathfrak{h}+\mathfrak{a}$. Also, $[\mathfrak{h}\cap\mathfrak{a}, \mathfrak{h}] \subset [\mathfrak{h}, \mathfrak{h}]$ and $[\mathfrak{h}\cap\mathfrak{a}, \mathfrak{h}] \subset [\mathfrak{a}, \mathfrak{h}]$ so that $[\mathfrak{h}\cap\mathfrak{a}, \mathfrak{h}] \subset [\mathfrak{h}, \mathfrak{h}] \cap [\mathfrak{a}, \mathfrak{h}] \subset \mathfrak{h}\cap\mathfrak{a}$. Thus $\mathfrak{h}\cap\mathfrak{a}$ is an ideal of $\mathfrak{h}$.

Notice that $(\mathfrak{h}+\mathfrak{a})/\mathfrak{a}$ forms a Lie algebra by Proposition 1.6. Consider the linear map $\phi : \mathfrak{h} \to (\mathfrak{h}+\mathfrak{a})/\mathfrak{a}$ defined by $\phi(x) = x + \mathfrak{a}$. We have

$$\phi([x, y]) = [x, y] + \mathfrak{a} = [x + \mathfrak{a}, y + \mathfrak{a}] = [\phi(x), \phi(y)],$$

and so ϕ is a homomorphism of Lie algebras. It is surjective since each element of $(\mathfrak{h}+\mathfrak{a})/\mathfrak{a}$ can be written as $x + \mathfrak{a}$ with $x \in \mathfrak{h}$. Finally, we have

$$\ker\phi = \{x \in \mathfrak{h} \mid \phi(x) = 0\} = \{x \in \mathfrak{h} \mid x + \mathfrak{a} = \mathfrak{a}\} = \{x \in \mathfrak{h} \mid x \in \mathfrak{a}\} = \mathfrak{h}\cap\mathfrak{a}.$$

Thus $\mathfrak{h}/(\mathfrak{h}\cap\mathfrak{a}) \cong (\mathfrak{h}+\mathfrak{a})/\mathfrak{a}$ by Proposition 1.7. □

1.3 Nilpotent and solvable Lie algebras

The key to studying a Lie algebra's structure is knowing about its ideals. The existence of the right kind of ideals will be useful in the sequel and is introduced in this section.

Let $\mathfrak{g}$ be a Lie algebra. We consider the subspaces of $\mathfrak{g}$ defined recursively by

$$\mathfrak{g}^1 = \mathfrak{g}, \qquad \mathfrak{g}^{n+1} = [\mathfrak{g}^n, \mathfrak{g}] \quad \text{for } n \geq 1.$$

Proposition 1.10. $\mathfrak{g}^n$ *is an ideal of* $\mathfrak{g}$ *and*

$$\mathfrak{g} = \mathfrak{g}^1 \supset \mathfrak{g}^2 \supset \mathfrak{g}^3 \supset \cdots .$$

Proof. We prove the first statement by induction on n. If $n = 1$, then $\mathfrak{g}^1 = \mathfrak{g}$ is an ideal since $[\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{g}$. Now suppose the statement is true for $n = r$, that is, $\mathfrak{g}^r$ is an ideal. Then $\mathfrak{g}^{r+1} = [\mathfrak{g}^r, \mathfrak{g}]$ is an ideal by Proposition 1.5, completing the induction. To prove the second statement, let $n \geq 1$. Then

$$\mathfrak{g}^{n+1} = [\mathfrak{g}^n, \mathfrak{g}] \subset \mathfrak{g}^n$$

since $\mathfrak{g}^n$ is an ideal. □

Given Proposition 1.10, we call the chain of ideals $\mathfrak{g} = \mathfrak{g}^1 \supset \mathfrak{g}^2 \supset \mathfrak{g}^3 \supset \cdots$ the **lower central series** of $\mathfrak{g}$. We say that $\mathfrak{g}$ is **nilpotent** if $\mathfrak{g}^n = 0$ for some $n \geq 1$. If $\mathfrak{g}^2 = 0$, then we say $\mathfrak{g}$ is **abelian**. This is equivalent to the condition $[\mathfrak{g}, \mathfrak{g}] = 0$, i.e., $[x, y] = 0$ for all $x, y \in \mathfrak{g}$.

We now consider the subspaces of $\mathfrak{g}$ defined recursively by

$$\mathfrak{g}^{(0)} = \mathfrak{g}, \qquad \mathfrak{g}^{(n+1)} = \left[\mathfrak{g}^{(n)}, \mathfrak{g}^{(n)}\right] \quad \text{for } n \geq 0.$$

Proposition 1.11. *$\mathfrak{g}^{(n)}$ is an ideal of $\mathfrak{g}$ and*

$$\mathfrak{g} = \mathfrak{g}^{(0)} \supset \mathfrak{g}^{(1)} \supset \mathfrak{g}^{(2)} \supset \cdots.$$

Proof. We prove the first statement by induction on n. If $n = 0$, then $\mathfrak{g}^{(0)} = \mathfrak{g}$ is clearly an ideal. Now suppose the statement is true for $n = r$, that is, $\mathfrak{g}^{(r)}$ is an ideal. Then $\mathfrak{g}^{(r+1)} = \left[\mathfrak{g}^{(r)}, \mathfrak{g}^{(r)}\right]$ is an ideal by Proposition 1.5, completing the induction. To prove the second statement, let $n \geq 0$. Then

$$\mathfrak{g}^{(n+1)} = \left[\mathfrak{g}^{(n)}, \mathfrak{g}^{(n)}\right] \subset \left[\mathfrak{g}^{(n)}, \mathfrak{g}\right] \subset \mathfrak{g}^{(n)}$$

since $\mathfrak{g}^{(n)}$ is an ideal. □

Given Proposition 1.11, we call the chain of ideals $\mathfrak{g} = \mathfrak{g}^{(0)} \supset \mathfrak{g}^{(1)} \supset \mathfrak{g}^{(2)} \supset \cdots$ the **derived series** of $\mathfrak{g}$. We say that $\mathfrak{g}$ is **solvable** if $\mathfrak{g}^{(n)} = 0$ for some $n \geq 0$.

Proposition 1.12.

(i) $[\mathfrak{g}^m, \mathfrak{g}^n] \subset \mathfrak{g}^{m+n}$ *for all $m, n \geq 1$.*

(ii) $\mathfrak{g}^{(n)} \subset \mathfrak{g}^{2^n}$ *for all $n \geq 0$.*

(iii) *Every nilpotent Lie algebra is solvable.*

Proof. For part (i), we fix m and use induction on n. If $n = 1$, then $[\mathfrak{g}^m, \mathfrak{g}] = \mathfrak{g}^{m+1}$. Now suppose the statement is true for $n = r$. By definition, $\left[\mathfrak{g}^m, \mathfrak{g}^{r+1}\right] = [\mathfrak{g}^m, [\mathfrak{g}^r, \mathfrak{g}]]$. This subspace is generated by elements of the form $[x, [y, z]]$ with $x \in \mathfrak{g}^m$, $y \in \mathfrak{g}^r$, $z \in \mathfrak{g}$. We have

$$\begin{aligned}
[x, [y, z]] &= [[z, x], y] + [[x, y], z] && \text{(by the Jacobi identity)} \\
&\in [[\mathfrak{g}, \mathfrak{g}^m], \mathfrak{g}^r] + [[\mathfrak{g}^m, \mathfrak{g}^r], \mathfrak{g}] \\
&\subset \left[\mathfrak{g}^{m+1}, \mathfrak{g}^r\right] + \left[\mathfrak{g}^{m+r}, \mathfrak{g}\right] \\
&\subset \mathfrak{g}^{m+r+1} && \text{(by the inductive hypothesis).}
\end{aligned}$$

It follows that $[\mathfrak{g}^m, \mathfrak{g}^n] \subset \mathfrak{g}^{m+n}$ for all n.

For part (ii), we again use induction on n. For $n = 0$, we have $\mathfrak{g}^{(0)} = \mathfrak{g} = \mathfrak{g}^{2^0}$. Now suppose the statement is true for $n = r$. We have

$$\mathfrak{g}^{(r+1)} = \left[\mathfrak{g}^{(r)}, \mathfrak{g}^{(r)}\right] \subset \left[\mathfrak{g}^{2^r}, \mathfrak{g}^{2^r}\right] \subset \mathfrak{g}^{2 \cdot 2^r} = \mathfrak{g}^{2^{r+1}}$$

by part (i), completing the induction.

For part (iii), if $\mathfrak{g}$ is nilpotent, then $\mathfrak{g}^n = 0$ for some $n \geq 1$. Thus $\mathfrak{g}^{(n)} \subset \mathfrak{g}^{2^n} \subset \mathfrak{g}^n = 0$ by part (ii), and so $\mathfrak{g}$ is solvable. □

Proposition 1.13. *If $\mathfrak{g}$ is solvable, then every subalgebra and every quotient algebra of $\mathfrak{g}$ is solvable.*

Proof. We first show that if $\phi : \mathfrak{g} \to \mathfrak{h}$ is a surjective homomorphism of Lie algebras, then $\phi\left(\mathfrak{g}^{(n)}\right) = \mathfrak{h}^{(n)}$ for all $n \geq 0$. We proceed by induction on n. For $n = 0$, we have $\phi\left(\mathfrak{g}^{(0)}\right) = \phi(\mathfrak{g}) = \mathfrak{h} = \mathfrak{h}^{(0)}$. Now suppose the statement is true for $n = r$. We have

$$\phi\left(\mathfrak{g}^{(r+1)}\right) = \phi\left(\left[\mathfrak{g}^{(r)}, \mathfrak{g}^{(r)}\right]\right) = \left[\phi\left(\mathfrak{g}^{(r)}\right), \phi\left(\mathfrak{g}^{(r)}\right)\right] = \left[\mathfrak{h}^{(r)}, \mathfrak{h}^{(r)}\right] = \mathfrak{h}^{(r+1)},$$

completing the proof of the claim.

To prove the proposition, choose n such that $\mathfrak{g}^{(n)} = 0$. If $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, then $\mathfrak{h}^{(n)} \subset \mathfrak{g}^{(n)} = 0$, and so $\mathfrak{h}$ is solvable. Finally, if $\mathfrak{a}$ is an ideal of $\mathfrak{g}$, then $(\mathfrak{g}/\mathfrak{a})^{(n)}$ is the image of $\mathfrak{g}^{(n)}$ under the canonical homomorphism $\mathfrak{g} \to \mathfrak{g}/\mathfrak{a}$ by the preceding result. We have $\mathfrak{g}^{(n)} = 0$, yielding $(\mathfrak{g}/\mathfrak{a})^{(n)} = 0$, and so $\mathfrak{g}/\mathfrak{a}$ is solvable. □

Proposition 1.14. *If $\mathfrak{a}$ is an ideal of $\mathfrak{g}$ such that $\mathfrak{a}$ and $\mathfrak{g}/\mathfrak{a}$ are solvable, then $\mathfrak{g}$ is solvable.*

Proof. We first show that $\mathfrak{g}^{(m+n)} = \left(\mathfrak{g}^{(m)}\right)^{(n)}$ for all $m, n \geq 0$. We fix m and use induction on n. For $n = 0$, we have $\mathfrak{g}^{(m+0)} = \mathfrak{g}^{(m)} = \left(\mathfrak{g}^{(m)}\right)^{(0)}$. Now suppose the statement is true for $n = r$. We have

$$\mathfrak{g}^{(m+r+1)} = \left[\mathfrak{g}^{(m+r)}, \mathfrak{g}^{(m+r)}\right] = \left[\left(\mathfrak{g}^{(m)}\right)^{(r)}, \left(\mathfrak{g}^{(m)}\right)^{(r)}\right] = \left(\mathfrak{g}^{(m)}\right)^{(r+1)},$$

completing the proof of the claim.

To prove the proposition, notice that since $\mathfrak{g}/\mathfrak{a}$ is solvable, we have $(\mathfrak{g}/\mathfrak{a})^{(m)} = 0$ for some $m \geq 0$, that is, $\mathfrak{g}^{(m)} \subset \mathfrak{a}$. And since $\mathfrak{a}$ is solvable, we have $\mathfrak{a}^{(n)} = 0$ for some $n \geq 0$. By the preceding result, this yields

$$\mathfrak{g}^{(m+n)} = \left(\mathfrak{g}^{(m)}\right)^{(n)} \subset \mathfrak{a}^{(n)} = 0,$$

and so $\mathfrak{g}$ is solvable. □

Proposition 1.15. *Every finite-dimensional Lie algebra $\mathfrak{g}$ contains a unique maximal solvable ideal $\mathfrak{r}$.*

Proof. We first show that if $\mathfrak{a}$ and $\mathfrak{b}$ are solvable ideals of $\mathfrak{g}$, then $\mathfrak{a}+\mathfrak{b}$ is a solvable ideal of $\mathfrak{g}$. Notice that $\mathfrak{a}+\mathfrak{b}$ is an ideal by Proposition 1.5. Because $\mathfrak{a}$ is solvable, we know that $\mathfrak{a}/(\mathfrak{a}\cap\mathfrak{b})$ is solvable by Proposition 1.13. Now $(\mathfrak{a}+\mathfrak{b})/\mathfrak{b} \cong \mathfrak{a}/(\mathfrak{a}\cap\mathfrak{b})$ by Proposition 1.9. Thus $(\mathfrak{a}+\mathfrak{b})/\mathfrak{b}$ is solvable. Since $\mathfrak{b}$ is also solvable, we see that $\mathfrak{a}+\mathfrak{b}$ is solvable by Proposition 1.14.

To prove the proposition, we choose a solvable ideal $\mathfrak{r}$ of $\mathfrak{g}$ having maximal dimension. Let $\mathfrak{a}$ be any solvable ideal of $\mathfrak{g}$. Then $\mathfrak{a}+\mathfrak{r}$ is a solvable ideal by the preceding result. We have $\mathfrak{a}+\mathfrak{r}=\mathfrak{r}$ since $\dim \mathfrak{r}$ is maximal, and so $\mathfrak{a} \subset \mathfrak{r}$. Thus $\mathfrak{r}$ contains every solvable ideal of $\mathfrak{g}$. □

Given Proposition 1.15, we call the maximal solvable ideal $\mathfrak{r}$ the **solvable radical** of $\mathfrak{g}$. A Lie algebra $\mathfrak{g}$ is **semisimple** if $\mathfrak{r}=0$, that is, if $\mathfrak{g}$ contains no nonzero solvable ideal. A Lie algebra $\mathfrak{g}$ is **simple** if it contains no proper ideals, i.e., no ideals other than 0 and $\mathfrak{g}$.

Suppose $\mathfrak{g}$ is 1-dimensional over k. Then for any nonzero $x \in \mathfrak{g}$, $\{x\}$ forms a basis of $\mathfrak{g}$. Since $[x,x]=0$, we see that $\mathfrak{g}^2=0$, and so $\mathfrak{g}$ is abelian. This is true for any Lie algebra of dimension 1. Thus any two 1-dimensional Lie algebras over k are isomorphic. Because the dimension of any subspace is either 0 or 1, we see that $\mathfrak{g}$ contains no proper subspaces. Hence $\mathfrak{g}$ contains no proper ideals, and so $\mathfrak{g}$ is simple. We call the 1-dimensional Lie algebra the **trivial simple Lie algebra**.

Proposition 1.16. *Every non-trivial simple Lie algebra is semisimple.*

Proof. Suppose $\mathfrak{g}$ is simple but not semisimple. Then $\mathfrak{r} \neq 0$, and since $\mathfrak{r}$ is an ideal, we must have $\mathfrak{r}=\mathfrak{g}$. Because $\mathfrak{r}$ is solvable, $\mathfrak{g}$ must be solvable, and so $\mathfrak{g}^{(n)}=0$ for some $n \geq 0$. Now $\mathfrak{g}^{(1)}$ is an ideal. Hence $\mathfrak{g}^{(1)}=\mathfrak{g}$ or $\mathfrak{g}^{(1)}=0$. But $\mathfrak{g}^{(1)}$ cannot equal $\mathfrak{g}$, because this would imply $\mathfrak{g}^{(n)}=\mathfrak{g}$ for all $n \geq 0$, contradicting the solvability of $\mathfrak{g}$. Thus $\mathfrak{g}^{(1)}=0$, i.e., $[\mathfrak{g},\mathfrak{g}]=0$. If $\mathfrak{a}$ is any subspace of $\mathfrak{g}$, then $[\mathfrak{a},\mathfrak{g}] \subset [\mathfrak{g},\mathfrak{g}]=0 \subset \mathfrak{a}$. It follows that every subspace of $\mathfrak{g}$ is an ideal, and so $\mathfrak{g}$ cannot have any proper subspaces. Hence $\dim \mathfrak{g}=1$, and so $\mathfrak{g}$ is the trivial simple Lie algebra. □

Chapter 2

Representations of solvable and nilpotent Lie algebras

2.1 Basic representation theory

Recall from Example 1.2 that any associative algebra A gives rise to a Lie algebra $[A]$ via the commutator product. We consider the case where $A = M_n(k)$, the ring of $n \times n$ matrices over k. One easily checks that this is an associative k-algebra. We define the Lie algebra $\mathfrak{gl}_n(k)$ by

$$\mathfrak{gl}_n(k) = [M_n(k)].$$

A **representation** of a Lie algebra $\mathfrak{g}$ is a homomorphism of Lie algebras

$$\rho : \mathfrak{g} \to \mathfrak{gl}_n(k)$$

for some $n \geq 0$. In this case, ρ is called a representation of degree n. Two representations ρ, ρ' of degree n are **equivalent** if there exists an invertible matrix $A \in M_n(k)$ such that $\rho'(x) = A^{-1}\rho(x)A$ for all $x \in \mathfrak{g}$.

A **left $\mathfrak{g}$-module** is a k-vector space V equipped with a left action

$$\mathfrak{g} \times V \to V$$
$$(x, v) \mapsto xv$$

satisfying the following conditions:

(i) $(x, v) \mapsto xv$ is linear in x and v.

(ii) $[x, y]v = x(yv) - y(xv) \quad$ for all $x, y \in \mathfrak{g}$, $v \in V$.

Note that we can define right $\mathfrak{g}$-modules in a similar way, although we will not have occasion to use them in this paper. Thus from now on, we will assume that all $\mathfrak{g}$-modules are left $\mathfrak{g}$-modules.

There is a one-to-one correspondence between representations of degree n and n-dimensional $\mathfrak{g}$-modules. Suppose ρ is a representation of $\mathfrak{g}$ of degree n. Let V be an n-dimensional vector space and choose a basis $e_1, \ldots, e_n$ of V. For each $x \in \mathfrak{g}$, we define the action of x on an element of V by

$$x\left(\sum_j \lambda_j e_j\right) = \sum_{ij} \lambda_j \rho_{ij}(x) e_i.$$

Thus for each $x \in \mathfrak{g}$, the matrix $\rho(x)$ acts as a linear transformation on V with respect to the basis $e_1, \ldots, e_n$. For all $x, y \in \mathfrak{g}$, $v \in V$, we have

$$\rho([x,y])v = [\rho(x), \rho(y)]v = (\rho(x)\rho(y) - \rho(y)\rho(x))v = \rho(x)\rho(y)v - \rho(y)\rho(x)v,$$

showing that this action transforms V into a $\mathfrak{g}$-module.

Conversely, suppose V is an n-dimensional $\mathfrak{g}$-module and let $e_1, \ldots, e_n$ be a basis of V. For each $x \in \mathfrak{g}$, we write

$$xe_j = \sum_i \rho_{ij}(x) e_i$$

where $\rho_{ij}(x) \in k$. This gives rise to the matrix $\rho(x) = (\rho_{ij}(x))$. We have

$$\begin{aligned}
[x,y]\, e_j &= x(ye_j) - y(xe_j) \\
&= x\left(\sum_k \rho_{kj}(y) e_k\right) - y\left(\sum_k \rho_{kj}(x) e_k\right) \\
&= \sum_k \rho_{kj}(y) x e_k - \sum_k \rho_{kj}(x) y e_k \\
&= \sum_k \rho_{kj}(y)\left(\sum_i \rho_{ik}(x) e_i\right) - \sum_k \rho_{kj}(x)\left(\sum_i \rho_{ik}(y) e_i\right) \\
&= \sum_i \left(\sum_k (\rho_{ik}(x)\rho_{kj}(y) - \rho_{ik}(y)\rho_{kj}(x))\right) e_i \\
&= \sum_i (\rho(x)\rho(y) - \rho(y)\rho(x))_{ij} e_i.
\end{aligned}$$

It follows that $\rho([x,y]) = \rho(x)\rho(y) - \rho(y)\rho(x) = [\rho(x), \rho(y)]$, and so ρ is a representation of $\mathfrak{g}$.

Now let $f_1, \ldots, f_n$ be another basis of V and ρ' the representation obtained from this basis. I claim that ρ' is equivalent to ρ. We know there exists an invertible $n \times n$ matrix A such that $f_j = \sum_i A_{ij} e_i$. We have

$$xf_j = \sum_k A_{kj} x e_k = \sum_k A_{kj} \left(\sum_i \rho_{ik}(x) e_i \right) = \sum_i \left(\sum_k \rho_{ik}(x) A_{kj} \right) e_i.$$

We also have

$$xf_j = \sum_k \rho'_{kj}(x) f_k = \sum_k \rho'_{kj}(x) \left(\sum_i A_{ik} e_i \right) = \sum_i \left(\sum_k A_{ik} \rho'_{kj}(x) \right) e_i.$$

It follows that $\rho(x)A = A\rho'(x)$ so that $\rho'(x) = A^{-1}\rho(x)A$ for all $x \in \mathfrak{g}$. Thus ρ' is equivalent to ρ as desired.

Let V be a $\mathfrak{g}$-module, $\mathfrak{a}$ a subspace of $\mathfrak{g}$, and U a subspace of V. We define $\mathfrak{a}U$ to be the subspace of V spanned by all elements of the form xu with $x \in \mathfrak{a}$ and $u \in U$. A **submodule** of V is a subspace U of V such that $\mathfrak{g}U \subset U$. Notice that both 0 and V are submodules of V. A **proper submodule** of V is a submodule distinct from 0 and V.

A $\mathfrak{g}$-module V is **irreducible** if it contains no proper submodules. We say that V is **completely reducible** if it is a direct sum of irreducible submodules. Finally, V is said to be **indecomposable** if it cannot be written as a sum of two proper submodules. We see that every irreducible $\mathfrak{g}$-module is indecomposable, although the converse need not hold.

Example 2.1. Let V be a $\mathfrak{g}$-module and W a submodule of V. Then the quotient space V/W forms a $\mathfrak{g}$-module via the action

$$(x, v + W) \mapsto xv + W.$$

To prove this action is well defined, let $v + W = v' + W$. We must show that

$$xv + W = xv' + W.$$

Now $v + W = v' + W$ implies $v = v' + w$ for some $w \in W$. Thus

$$xv + W = x(v' + w) + W = xv' + xw + W = xv' + W$$

since $xw \in W$. The module properties in V/W follow easily from those in V.

Example 2.2. The vector space $\mathfrak{g}$ can be made into a $\mathfrak{g}$-module via the action

$$(x, y) \mapsto [x, y].$$

This is certainly linear in x and y. By the Jacobi identity, we also have

$$[[x, y], z] = [x, [y, z]] - [y, [x, z]],$$

showing that $\mathfrak{g}$ is indeed a $\mathfrak{g}$-module. We call this module the **adjoint module** and denote the action of x on the vector y by $\mathrm{ad}\, x \cdot y$. By the module properties, we have

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

Notice that a submodule $\mathfrak{a}$ of $\mathfrak{g}$ is defined by the property $[\mathfrak{g}, \mathfrak{a}] \subset \mathfrak{a}$. Hence the submodules of the adjoint module $\mathfrak{g}$ are precisely the ideals of $\mathfrak{g}$.

2.2 A transition to finite-dimensional complex Lie algebras

With the exception of Propositions 1.15 and 1.16, every definition and result in this paper has been applicable to Lie algebras with arbitrary base fields and dimensions. We wish now to draw a line in the sand. *From this point forward, we shall assume that every Lie algebra* $\mathfrak{g}$ *is finite-dimensional and that the base field* k *is the field* $\mathbb{C}$ *of complex numbers.* These types of Lie algebras have particularly nice representations, as we shall witness in the following sections.

2.3 Representations of solvable Lie algebras

This section will produce some very useful results about solvable Lie algebras.

As one might expect, a **1-dimensional representation** of $\mathfrak{g}$ is a homomorphism of Lie algebras $\rho : \mathfrak{g} \to [\mathbb{C}]$.

Lemma 2.3. *A linear map $\rho : \mathfrak{g} \to \mathbb{C}$ is a 1-dimensional representation of $\mathfrak{g}$ if and only if ρ vanishes on $\mathfrak{g}^2$.*

Proof. Suppose ρ is a 1-dimensional representation of $\mathfrak{g}$ and let $x, y \in \mathfrak{g}$. Then

$$\rho([x,y]) = [\rho(x), \rho(y)] = \rho(x)\rho(y) - \rho(y)\rho(x) = 0,$$

and so ρ vanishes on $\mathfrak{g}^2$. Conversely, if ρ vanishes on $\mathfrak{g}^2$, then

$$\rho([x,y]) = 0 = \rho(x)\rho(y) - \rho(y)\rho(x) = [\rho(x), \rho(y)]$$

so that ρ is a homomorphism from $\mathfrak{g}$ into $[\mathbb{C}]$. □

Theorem 2.4 (Lie's theorem). *Let $\mathfrak{g}$ be a solvable Lie algebra and V a finite-dimensional irreducible $\mathfrak{g}$-module. Then* $\dim V = 1$.

Proof. We proceed by induction on $\dim \mathfrak{g}$. If $\dim \mathfrak{g} = 1$, then $\mathfrak{g} = \mathbb{C}x$ for some $x \in \mathfrak{g}$. Since $\mathbb{C}$ is algebraically closed, we can find an eigenvector v of x in V. Then $xv \in \mathbb{C}v$, and so $\mathbb{C}v$ is a $\mathfrak{g}$-submodule of V. Because V is irreducible, this forces $V = \mathbb{C}v$ and $\dim V = 1$.

Now suppose $\dim \mathfrak{g} > 1$ and that the theorem is true for all solvable Lie algebras of dimension $\dim \mathfrak{g} - 1$. Notice that since $\mathfrak{g}$ is solvable, we have $\mathfrak{g}^2 \neq \mathfrak{g}$, or else we would have $\mathfrak{g}^{(n)} = \mathfrak{g}$ for all $n \geq 0$. Let $\mathfrak{a}$ be a subspace of $\mathfrak{g}$ containing $\mathfrak{g}^2$ such that $\dim \mathfrak{a} = \dim \mathfrak{g} - 1$. Then

$$[\mathfrak{a}, \mathfrak{g}] \subset [\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}^2 \subset \mathfrak{a},$$

and so $\mathfrak{a}$ is an ideal of $\mathfrak{g}$. We may regard V as an $\mathfrak{a}$-module. Let W be an irreducible $\mathfrak{a}$-submodule of V. By induction, we have $\dim W = 1$. Then $yw = \lambda(y)w$ for all $y \in \mathfrak{a}$, where λ is the 1-dimensional representation induced by W. Define the set

$$U = \{u \in V \mid yu = \lambda(y)u \quad \text{for all } y \in \mathfrak{a}\}.$$

We have $0 \neq W \subset U \subset V$. I claim that U is a $\mathfrak{g}$-submodule of V. To prove this, let $x \in \mathfrak{g}$, $y \in \mathfrak{a}$, $u \in U$. Then

$$y(xu) = x(yu) - [x,y]u = x(\lambda(y)u) - \lambda([x,y])u = \lambda(y)xu - \lambda([x,y])u.$$

We shall show that $\lambda([x,y]) = 0$. Having proved this, it follows that $xu \in U$, and so U is a $\mathfrak{g}$-submodule of V. Since V is irreducible, this forces $U = V$, and so $yv = \lambda(y)v$ for all $y \in \mathfrak{a}$, $v \in V$. Now because $\mathfrak{a}$ has codimension 1, we have $\mathfrak{g} = \mathfrak{a} \oplus \mathbb{C}z$ for some $z \in \mathfrak{g}\backslash\mathfrak{a}$. Let v be an eigenvector of z in V. Then $\mathbb{C}v$ is invariant under the actions of both $\mathfrak{a}$ and z, and so $\mathbb{C}v$ is a $\mathfrak{g}$-submodule of V. Because V is irreducible, this forces $V = \mathbb{C}v$ and $\dim V = 1$.

To prove that $\lambda([x,y]) = 0$ for all $x \in \mathfrak{g}$, $y \in \mathfrak{a}$, it suffices to consider the element $x = z$ such that $\mathfrak{g} = \mathfrak{a} \oplus \mathbb{C}z$ since λ vanishes on $\mathfrak{a}^2$ by Lemma 2.3. Let u be any nonzero vector in U. We define the vectors

$$v_0 = u, \quad v_1 = zu, \quad v_2 = z(zu), \quad \ldots$$

Let $V_0 = 0$ and $V_i = \langle v_0, v_1, \ldots, v_{i-1}\rangle$ for $i > 0$. We have $V_0 \subset V_1 \subset V_2 \subset \cdots$. I claim that $yv_i = \lambda(y)v_i \pmod{V_i}$ for all $y \in \mathfrak{a}$, $i \geq 0$. We proceed by induction on i. For $i = 0$, we have

$$yv_0 = yu = \lambda(y)u = \lambda(y)v_0.$$

Now suppose $i > 0$ and that the statement is true for $i-1$. We have

$$yv_i = y(zv_{i-1}) = z(yv_{i-1}) - [z,y]v_{i-1}.$$

By induction, $yv_{i-1} = \lambda(y)v_{i-1} + v'$ and $[z,y]v_{i-1} = \lambda([z,y])v_{i-1} + v''$ for some $v', v'' \in V_{i-1}$. Hence

$$\begin{aligned} yv_i &= z(\lambda(y)v_{i-1} + v') - (\lambda([z,y])v_{i-1} + v'') \\ &= \lambda(y)v_i + zv' - \lambda([z,y])v_{i-1} - v'' \\ &\equiv \lambda(y)v_i \pmod{V_i}, \end{aligned}$$

completing the proof of the claim. Now because V is finite-dimensional, there exists a minimal $r > 0$ such that $V_r = V_{r+1} = V_{r+2} = \cdots$. By construction, V_r is invariant under the action of z, and given the above result, V_r is also invariant under the action of $\mathfrak{a}$. Hence V_r is a $\mathfrak{g}$-submodule of V, and since V is irreducible, we have $V = V_r$. Then by the above result, any element $y \in \mathfrak{a}$ acts on V via an upper-triangular matrix with $\lambda(y)$ along the main diagonal. This is certainly true for the element $[z, y]$, and so $\mathrm{tr}_V([z, y]) = r\lambda([z, y])$. But we also have

$$\mathrm{tr}_V([z, y]) = \mathrm{tr}_V(zy - yz) = \mathrm{tr}_V(zy) - \mathrm{tr}_V(yz) = 0.$$

Hence $r\lambda([z, y]) = 0$, and since $\mathbb{C}$ is an integral domain, this yields $\lambda([z, y]) = 0$. □

Corollary 2.5. *Let $\mathfrak{g}$ be a solvable Lie algebra and V a finite-dimensional $\mathfrak{g}$-module. Then a basis can be chosen for V with respect to which we obtain a matrix representation ρ of $\mathfrak{g}$ of the form*

$$\rho(x) = \begin{pmatrix} * & & & * \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & * \end{pmatrix} \quad \textit{for all } x \in \mathfrak{g}.$$

Thus elements of $\mathfrak{g}$ are represented by upper-triangular matrices.

Proof. It suffices to show there exists a chain of submodules

$$0 = V_0 \subset V_1 \subset \cdots \subset V_n = V$$

such that $\dim V_i = i$ for $i = 0, \ldots, n$. We proceed by induction on $\dim V$. The statement is trivial for $\dim V = 1$. So suppose $\dim V = n > 1$ and that the statement is true for all $\mathfrak{g}$-modules of dimension $n - 1$. Choose an irreducible $\mathfrak{g}$-submodule W of V. By Theorem 2.4, we have $\dim W = 1$. Then V/W is a $\mathfrak{g}$-module of dimension $n - 1$, and by the inductive hypothesis, we may assume there exists a chain of submodules

$$0 = \overline{V}_0 \subset \overline{V}_1 \subset \cdots \subset \overline{V}_{n-1} = V/W$$

with the desired properties. For $i = 1, \ldots, n$, let V_i denote the preimage of $\overline{V}_{i-1}$ in V under the canonical homomorphism. We claim that the chain

$$0 = V_0 \subset V_1 \subset \cdots \subset V_n = V$$

has the desired properties for V. Let $\overline{v}_1, \ldots, \overline{v}_{i-1}$ be a basis of $\overline{V}_{i-1}$ where $\overline{v_j} \in \overline{V}_j \backslash \overline{V}_{j-1}$ for $j = 1, \ldots, i-1$. Consider the elements $w, v_1, \ldots, v_{i-1} \in V_i$ where $0 \neq w \in W$ and v_j is in the preimage of $\overline{v}_j$. We shall prove these elements forms a basis of V_i. Let ϕ be the canonical homomorphism $V \to V/W$. If $\lambda_0 w + \sum_{j=1}^{i-1} \lambda_j v_j = 0$ where $\lambda_j \in \mathbb{C}$, then

$$0 = \phi(0) = \phi\left(\lambda_0 w + \sum_{j=1}^{i-1} \lambda_j v_j\right) = \lambda_0 \phi(w) + \sum_{j=1}^{i-1} \lambda_j \phi(v_j) = \sum_{j=1}^{i-1} \lambda_j \overline{v}_j.$$

Because $\overline{v}_1, \ldots, \overline{v}_{i-1}$ are linearly independent, we must have $\lambda_1 = \cdots = \lambda_{i-1} = 0$. It follows that $\lambda_0 w = 0$, and since $w \neq 0$, this yields $\lambda_0 = 0$. Thus $w, v_1, \ldots, v_{i-1}$ are linearly independent, and because

$$\dim V_i = \dim V_i/W + \dim W = (i-1) + 1 = i,$$

we see that $w, v_1, \ldots, v_{i-1}$ is indeed a basis of V_i. Finally, we show that V_i is a submodule of V. Let $x \in \mathfrak{g}$, $v \in V_i$. Then $\phi(v) \in \overline{V}_{i-1}$, and by induction, we have $\rho(x)v + W \in \overline{V}_{i-1}$. Hence $\phi(\rho(x)v) \in \overline{V}_{i-1}$, and so $\rho(x)v \in V_i$. □

Corollary 2.6. *Every solvable Lie algebra* $\mathfrak{g}$ *has a chain of ideals*

$$0 = \mathfrak{a}_0 \subset \mathfrak{a}_1 \subset \cdots \subset \mathfrak{a}_n = \mathfrak{g}$$

where $\dim \mathfrak{a}_i = i$*.*

Proof. We apply the proof of Corollary 2.5 to the adjoint module $\mathfrak{g}$. The statement follows immediately since each submodule of $\mathfrak{g}$ is an ideal by Example 2.2. □

2.4 Representations of nilpotent Lie algebras

Recall from Proposition 1.12 that every nilpotent Lie algebra is solvable. In this section, we will apply our results about solvable Lie algebras to nilpotent Lie algebras in order to gain some very useful representations.

We first introduce some standard results from linear algebra. Let V be a finite-dimensional vector space and let $T : V \to V$ be a linear transformation with eigenvalues $\lambda_1, \dots, \lambda_r$. The **generalized eigenspace** of V with respect to λ_i is the set V_i of all $v \in V$ annihilated by some power of $T - \lambda_i 1$. Let $\chi(t) = (t-\lambda_1)^{m_1}(t-\lambda_2)^{m_2} \cdots (t-\lambda_r)^{m_r}$ be the characteristic polynomial of T. It follows from the Jordan canonical form (see [DF04, §12.3]) that

(i) $V = V_1 \oplus \cdots \oplus V_r$.

(ii) Each V_i is invariant under the action of T.

(iii) $\dim V_i = m_i$, and the characteristic polynomial of $T|_{V_i}$ is $(t - \lambda_i)^{m_i}$.

The relation between generalized eigenspaces and representations of nilpotent Lie algebras will be exhibited in the next theorem. A useful proposition is needed first.

Proposition 2.7. *Let $\mathfrak{g}$ be a Lie algebra and V a $\mathfrak{g}$-module. For each $y \in \mathfrak{g}$, define the linear map $\rho(y) : V \to V$ by $\rho(y)v = yv$. Let $v \in V$, $x, y \in \mathfrak{g}$, $\alpha, \beta \in \mathbb{C}$. Then*

$$(\rho(y) - (\alpha+\beta)1)^n xv = \sum_{i=0}^{n} \binom{n}{i} \left((\operatorname{ad} y - \beta 1)^i x\right) \left((\rho(y) - \alpha 1)^{n-i} v\right).$$

Proof. We proceed by induction on n, the case where $n = 0$ being obvious. Suppose the statement is true for $n = r$. For each i, let $x_i = (\operatorname{ad} y - \beta 1)^i x \in \mathfrak{g}$. Then

$$(\rho(y) - (\alpha+\beta)1)^{r+1} xv = (\rho(y) - (\alpha+\beta)1) \sum_{i=0}^{r} \binom{r}{i} \rho(x_i)(\rho(y) - \alpha 1)^{r-i} v.$$

Now

$$(\rho(y) - (\alpha+\beta)1)\rho(x_i) = \rho([y,x_i]) + \rho(x_i)\rho(y) - (\alpha+\beta)\rho(x_i)$$
$$= \rho((\operatorname{ad} y - \beta 1)x_i) + \rho(x_i)(\rho(y) - \alpha 1)$$
$$= \rho(x_{i+i}) + \rho(x_i)(\rho(y) - \alpha 1).$$

It follows that

$$\begin{aligned}
&(\rho(y) - (\alpha+\beta)1)^{r+1} xv \\
&= \sum_{i=0}^{r} \binom{r}{i} \rho(x_{i+1})(\rho(y) - \alpha 1)^{r-i} v + \sum_{i=0}^{r} \binom{r}{i} \rho(x_i)(\rho(y) - \alpha 1)^{r-i+1} v \\
&= \sum_{i=0}^{r+1} \binom{r}{i-1} \rho(x_i)(\rho(y) - \alpha 1)^{r-(i-1)} v + \sum_{i=0}^{r+1} \binom{r}{i} \rho(x_i)(\rho(y) - \alpha 1)^{r+1-i} v \\
&\qquad \left(\text{since } \binom{r}{-1} = 0 \text{ and } \binom{r}{r+1} = 0\right) \\
&= \sum_{i=0}^{r+1} \binom{r+1}{i} \left((\operatorname{ad} y - \beta 1)^i x\right) \left(\rho(y) - \alpha 1)^{r+1-i} v\right).
\end{aligned}$$

(The last line follows from the identity $\binom{r+1}{i} = \binom{r}{i-1} + \binom{r}{i}$.) This completes the induction. □

Theorem 2.8. *Let $\mathfrak{g}$ be a nilpotent Lie algebra and V a finite-dimensional $\mathfrak{g}$-module. For $y \in \mathfrak{g}$, let $\rho(y)$ be the linear map described in Proposition 2.7. Then the generalized eigenspaces of V associated with $\rho(y)$ are all submodules of V.*

Proof. Let V_i be a generalized eigenspace of V. If $x \in \mathfrak{g}$ and $v \in V_i$, then

$$(\rho(y) - \lambda_i 1)^n xv = \sum_{i=0}^{n} \binom{n}{i} \left((\operatorname{ad} y)^i x\right) \left((\rho(y) - \lambda_i 1)^{n-i} v\right)$$

by Proposition 2.7 with $\alpha = \lambda_i$, $\beta = 0$. Since $\mathfrak{g}$ is nilpotent, we have $(\operatorname{ad} y)^i x = 0$ if i is sufficiently large. And since $v \in V_i$, we have $(\rho(y) - \lambda_i 1)^{n-i} v = 0$ if $n - i$ is sufficiently large. Hence $(\rho(y) - \lambda_i 1)^n xv = 0$ if n is sufficiently large. It follows that $xv \in V_i$, and so V_i is a submodule of V. □

Corollary 2.9. *Let $\mathfrak{g}$ be a nilpotent Lie algebra and V a finite-dimensional indecomposable $\mathfrak{g}$-module. Then a basis can be chosen for V with respect to which we obtain a* *representation ρ of $\mathfrak{g}$ of the form*

$$\rho(x) = \begin{pmatrix} \lambda(x) & & & * \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot & \\ 0 & & & \lambda(x) \end{pmatrix} \qquad \textit{for all } x \in \mathfrak{g}.$$

Proof. By Corollary 2.5, we may choose a basis of V with respect to which $\rho(x)$ is upper-triangular. The generalized eigenspaces of V with respect to $\rho(x)$ are all submodules of V by Theorem 2.8, and V is their direct sum. Since V is indecomposable, only one of the generalized eigenspaces can be nonzero. Thus $\rho(x)$ has only one eigenvalue. Let $\lambda(x)$ be this eigenvalue. Then all of the diagonal entries of the upper-triangular matrix $\rho(x)$ will be equal to $\lambda(x)$. □

One easily checks that the map $x \mapsto \lambda(x)$ is a 1-dimensional representation of $\mathfrak{g}$.

Theorem 2.8 showed us that for any element $y \in \mathfrak{g}$, we obtain a decomposition of V into a direct sum of $\mathfrak{g}$-submodules. We wish now to obtain a more general decomposition of V, i.e., one that does not rely on the choice of y.

Let $\mathfrak{g}$ be a nilpotent Lie algebra and V a finite-dimensional $\mathfrak{g}$-module. For any 1-dimensional representation λ of $\mathfrak{g}$, we define the set

$$V_\lambda = \{v \in V \mid (\forall x \in \mathfrak{g})(\exists N(x) \geq 1) \quad (\rho(x) - \lambda(x)1)^{N(x)}v = 0\}.$$

Theorem 2.10. *$V = \bigoplus_\lambda V_\lambda$, and each V_λ is a submodule of V.*

Proof. We first decompose V into a direct sum of indecomposable submodules. Then by Corollary 2.9, each such submodule will induce a 1-dimensional representation λ of $\mathfrak{g}$. Let W_λ be the direct sum of all indecomposable submodules giving rise to λ. We have

$$V = \bigoplus_\lambda W_\lambda.$$

I shall prove that $W_\lambda = V_\lambda$, and thus W_λ is independent of the way we obtained our indecomposable components. Suppose $W_\lambda \neq V_\lambda$. Observe that $W_\lambda \subset V_\lambda$ by Corollary 2.9. It follows from a standard result in linear algebra that

$$\left(W_\lambda + \bigoplus_{\mu\neq\lambda} W_\mu\right) \cap V_\lambda = W_\lambda + \left(V_\lambda \cap \bigoplus_{\mu\neq\lambda} W_\mu\right).$$

We have $W_\lambda + \bigoplus_{\mu\neq\lambda} W_\mu = V$, and thus $V_\lambda = W_\lambda + \left(V_\lambda \cap \bigoplus_{\mu\neq\lambda} W_\mu\right)$. Since $V_\lambda \neq W_\lambda$, there must exist a nonzero $v \in V_\lambda \cap \bigoplus_{\mu\neq\lambda} W_\mu$. We write $v = \sum_{\mu\neq\lambda} w_\mu$ where $w_\mu \in W_\mu$.

Now the vector space $\mathfrak{g}$ over the infinite field $\mathbb{C}$ cannot be expressed as the union of finitely-many proper subspaces. One easily checks that for each $\mu \neq \lambda$, the set of $x \in \mathfrak{g}$ satisfying $\mu(x) = \lambda(x)$ is a proper subspace. Thus there exists an $x \in \mathfrak{g}$ such that $\mu(x) \neq \lambda(x)$ for all $\mu \neq \lambda$. Now since $v \in V_\lambda$, there exists an $N_\lambda \geq 1$ such that

$$(\rho(x) - \lambda(x)1)^{N_\lambda} v = 0.$$

And since $w_\mu \in W_\mu$, there exists an $N_\mu \geq 1$ such that $(\rho(x) - \mu(x)1)^{N_\mu} w_\mu = 0$. Hence

$$\prod_{\mu\neq\lambda} (\rho(x) - \mu(x)1)^{N_\mu} v = 0.$$

Because $\mu(x) \neq \lambda(x)$ for all $\mu \neq \lambda$, we see that the polynomials

$$(t - \lambda(x))^{N_\lambda}, \qquad \prod_{\mu\neq\lambda} (t - \mu(x))^{N_\mu}$$

are coprime. Thus there exist polynomials $a(t), b(t) \in \mathbb{C}[t]$ such that

$$a(t)(t - \lambda(x))^{N_\lambda} + b(t) \prod_{\mu\neq\lambda} (t - \mu(x))^{N_\mu} = 1.$$

It follows that

$$a(\rho(x))(\rho(x) - \lambda(x)1)^{N_\lambda} v + b(\rho(x)) \prod_{\mu\neq\lambda} (\rho(x) - \mu(x)1)^{N_\mu} v = v.$$

The left-hand side of this equation is zero by the above results, and thus $v = 0$. This is a contradiction, so we must have $W_\lambda = V_\lambda$. It follows that $V = \bigoplus_\lambda V_\lambda$, and each V_λ is a submodule of V. $\square$

If λ is a 1-dimensional representation of $\mathfrak{g}$ and $V_\lambda \neq 0$, then we call λ a **weight** of V, and we call V_λ the **weight space** of λ. Given Theorem 2.10, we call the decomposition $\bigoplus_\lambda V_\lambda$ the **weight space decomposition** of V. Since each V_λ is the direct sum of the indecomposable components giving rise to λ, it follows from Corollary 2.9 that a basis can be chosen for V_λ with respect to which a representation ρ of $\mathfrak{g}$ on V_λ has the form

$$\rho(x) = \begin{pmatrix} \lambda(x) & & & * \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \lambda(x) \end{pmatrix} \qquad \text{for all } x \in \mathfrak{g}.$$

This fact will prove extremely useful throughout the rest of this paper.

The final theorem of this chapter illustrates the connection between nilpotent Lie algebras and nilpotent linear transformations.

Theorem 2.11 (Engel's theorem)**.** *A Lie algebra* $\mathfrak{g}$ *is nilpotent if and only if* $\operatorname{ad} x : \mathfrak{g} \to \mathfrak{g}$ *is nilpotent for all* $x \in \mathfrak{g}$.

Proof. Suppose $\mathfrak{g}$ is nilpotent. Then $\mathfrak{g}^n = 0$ for some n. Let $y \in \mathfrak{g}$. We have

$$\operatorname{ad} x \cdot y \in \mathfrak{g}^2, \quad (\operatorname{ad} x)^2 y \in \mathfrak{g}^3, \quad \ldots, \quad (\operatorname{ad} x)^{i-1} y \in \mathfrak{g}^i, \quad \ldots$$

Thus $(\operatorname{ad} x)^{n-1} y \in \mathfrak{g}^n = 0$. It follows that $(\operatorname{ad} x)^{n-1} = 0$, and so $\operatorname{ad} x$ is a nilpotent linear map.

Now suppose $\operatorname{ad} x$ is nilpotent for all $x \in \mathfrak{g}$. We wish to show that $\mathfrak{g}$ is nilpotent. Suppose $\mathfrak{g}$ is not nilpotent. Choosing a maximal nilpotent subalgebra $\mathfrak{h}$ of $\mathfrak{g}$, we may regard $\mathfrak{g}$ as an $\mathfrak{h}$-module. Then $\mathfrak{h}$ is an $\mathfrak{h}$-submodule of $\mathfrak{g}$. Let $\mathfrak{m}$ be an $\mathfrak{h}$-submodule of $\mathfrak{g}$ properly containing $\mathfrak{h}$ such that $\mathfrak{m}/\mathfrak{h}$ is irreducible. By Theorem 2.4, we have $\dim \mathfrak{m}/\mathfrak{h} = 1$. Then $\mathfrak{m}/\mathfrak{h} = \mathbb{C}(x + \mathfrak{h})$ for some $x \in \mathfrak{m}\backslash\mathfrak{h}$, and $\mathfrak{m} = \mathfrak{h} \oplus \mathbb{C}x$.

The 1-dimensional representation induced by $\mathfrak{m}/\mathfrak{h}$ must be the zero map. To see why, suppose there existed a $y \in \mathfrak{h}$ such that $\operatorname{ad} y \cdot (x + \mathfrak{h}) = \lambda x + \mathfrak{h}$ with $\lambda \neq 0$. Then $\operatorname{ad} y \cdot x = \lambda x + z$ for some $z \in \mathfrak{h}$. Thus $(\operatorname{ad} y)^n x = \lambda^n x + (\operatorname{ad} y)^n z \neq 0$ for all n since the second term is in $\mathfrak{h}$. Hence $\operatorname{ad} y$ would fail to be nilpotent, contrary to our assumption.

Because the 1-dimensional representation induced by $\mathfrak{m}/\mathfrak{h}$ is the zero map, we have $[\mathfrak{h}, \mathfrak{m}] \subset \mathfrak{h}$. It follows that

$$\begin{aligned}[\mathfrak{m}, \mathfrak{m}] &= [\mathfrak{h} \oplus \mathbb{C}x, \mathfrak{h} \oplus \mathbb{C}x] \subset [\mathfrak{h}, \mathfrak{h}] + [\mathfrak{h}, \mathbb{C}x] + [\mathbb{C}x, \mathfrak{h}] + [\mathbb{C}x, \mathbb{C}x] \\ &= [\mathfrak{h}, \mathfrak{h}] + [\mathfrak{h}, \mathbb{C}x] \subset \mathfrak{h}.\end{aligned}$$

Thus $\mathfrak{m}$ is a subalgebra of $\mathfrak{g}$ and $\mathfrak{h}$ is an ideal of $\mathfrak{m}$.

We shall prove that for each $i \geq 1$, there exists a positive integer $e(i)$ such that $\mathfrak{m}^{e(i)} \subset \mathfrak{h}^i$. We proceed by induction on i. The statement is true for $i = 1$ since $\mathfrak{m}^2 \subset \mathfrak{h}$. Now suppose the statement is true for $i = r$. Then

$$\mathfrak{m}^{e(r)+1} \;=\; \left[\mathfrak{m}^{e(r)}, \mathfrak{h} \oplus \mathbb{C}x\right] \;\subset\; [\mathfrak{h}^r, \mathfrak{h}] + \left[\mathfrak{m}^{e(r)}, \mathbb{C}x\right] \;=\; \mathfrak{h}^{r+1} + \operatorname{ad} x \cdot \mathfrak{m}^{e(r)}.$$

I claim that $\mathfrak{m}^{e(r)+j} \subset \mathfrak{h}^{r+1} + (\operatorname{ad} x)^j \mathfrak{m}^{e(r)}$ for all $j \geq 1$. We proceed by induction on j, the case for $j = 1$ being proved. Assuming the statement is true for $j = s$, we have

$$\begin{aligned}\mathfrak{m}^{e(r)+s+1} &\subset \left[\mathfrak{h}^{r+1} + (\operatorname{ad} x)^s \mathfrak{m}^{e(r)}, \mathfrak{m}\right] \\ &\subset \mathfrak{h}^{r+1} + \left[(\operatorname{ad} x)^s \mathfrak{m}^{e(r)}, \mathfrak{h} \oplus \mathbb{C}x\right] \\ &\subset \mathfrak{h}^{r+1} + (\operatorname{ad} x)^{s+1} \mathfrak{m}^{e(r)}\end{aligned}$$

since $\mathfrak{h}^{r+1}$ is an ideal of $\mathfrak{m}$ and $(\operatorname{ad} x)^s \mathfrak{m}^{e(r)} \subset (\operatorname{ad} x)^s \mathfrak{h}^r \subset \mathfrak{h}^r$. Thus we have shown $\mathfrak{m}^{e(r)+j} \subset \mathfrak{h}^{r+1} + (\operatorname{ad} x)^j \mathfrak{m}^{e(r)}$ for all j. Since $\operatorname{ad} x$ is nilpotent, $(\operatorname{ad} x)^j = 0$ if j is sufficiently large. For such j, we have $\mathfrak{m}^{e(r)+j} \subset \mathfrak{h}^{r+1}$. Setting $e(r+1) = e(r) + j$, we obtain $\mathfrak{m}^{e(r+1)} \subset \mathfrak{h}^{r+1}$, completing the induction.

Now because $\mathfrak{h}$ is nilpotent, $\mathfrak{h}^i = 0$ if i is sufficiently large. For such i, we have $\mathfrak{m}^{e(i)} = 0$, showing that $\mathfrak{m}$ is nilpotent. This contradicts the maximality of $\mathfrak{h}$, and so $\mathfrak{g}$ must be nilpotent. $\square$

Note: The traditional statement of Theorem 2.11 applies to arbitrary finite-dimensional $\mathfrak{g}$-modules and linear maps. (See [Ser92, §V.2].) The proof, however, involves embedding $\mathfrak{g}$ into $\mathfrak{gl}_n(\mathbb{C})$ for some n. This requires a result known as Ado's theorem, the proof of which lies beyond the scope of this paper. (For a proof, refer to [Jac71, §VI.2].)

Corollary 2.12. *A Lie algebra $\mathfrak{g}$ is nilpotent if and only if $\mathfrak{g}$ has a basis with respect to which the adjoint representation of $\mathfrak{g}$ has the form*

$$\rho(x) = \begin{pmatrix} 0 & & & * \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & 0 \end{pmatrix} \quad \textit{for all } x \in \mathfrak{g}.$$

Proof. Suppose $\mathfrak{g}$ is nilpotent. Then $\mathfrak{g}$ has a chain of ideals

$$\mathfrak{g} \supset \mathfrak{g}^2 \supset \mathfrak{g}^3 \supset \cdots \supset \mathfrak{g}^r = 0$$

for some r. We make this chain finer by choosing a series of subspaces between consecutive ideals, each having codimension 1 in its predecessor. Let $\mathfrak{a}$ be a subspace such that $\mathfrak{g}^i \supset \mathfrak{a} \supset \mathfrak{g}^{i+1}$. Then

$$[\mathfrak{a}, \mathfrak{g}] \subset \left[\mathfrak{g}^i, \mathfrak{g}\right] = \mathfrak{g}^{i+1} \subset \mathfrak{a},$$

showing that $\mathfrak{a}$ is an ideal. Hence we obtain a chain of ideals

$$\mathfrak{g} = \mathfrak{a}_n \supset \mathfrak{a}_{n-1} \supset \cdots \supset \mathfrak{a}_1 \supset \mathfrak{a}_0 = 0$$

with $\dim \mathfrak{a}_k = k$ and $[\mathfrak{g}, \mathfrak{a}_k] \subset \mathfrak{a}_{k-1}$. Choosing a basis of $\mathfrak{g}$ adapted to this chain of ideals, the representation $\rho(x)$ will be of strictly upper-triangular form for all $x \in \mathfrak{g}$.

Conversely, suppose a basis can be chosen for $\mathfrak{g}$ with respect to which $\rho(x)$ is strictly upper-triangular for all $x \in \mathfrak{g}$. Then $\rho(x)$ is clearly nilpotent, whence $\operatorname{ad} x$ is nilpotent for all $x \in \mathfrak{g}$. Thus $\mathfrak{g}$ is nilpotent by Theorem 2.11. □

Chapter 3

Cartan subalgebras

If $\mathfrak{g}$ is a Lie algebra and $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, then $\mathfrak{g}$ forms an $\mathfrak{h}$-module via the adjoint action. If $\mathfrak{g}$ is semisimple and $\mathfrak{h}$ is chosen carefully, then the action of $\mathfrak{h}$ induces a decomposition of $\mathfrak{g}$ that tells us essentially all we need to know about the structure of $\mathfrak{g}$. In this chapter, we focus on finding the subalgebra $\mathfrak{h}$ yielding such a decomposition.

3.1 Existence of Cartan subalgebras

Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{h}$ a subalgebra of $\mathfrak{g}$. We define the **normalizer** of $\mathfrak{h}$ to be the set

$$N(\mathfrak{h}) = \{x \in \mathfrak{g} \mid [h, x] \in \mathfrak{h} \quad \text{for all } h \in \mathfrak{h}\}.$$

Proposition 3.1. *$N(\mathfrak{h})$ is a subalgebra of $\mathfrak{g}$, $\mathfrak{h}$ is an ideal of $N(\mathfrak{h})$, and $N(\mathfrak{h})$ is the largest subalgebra of $\mathfrak{g}$ containing $\mathfrak{h}$ as an ideal.*

Proof. For the first statement, let $x, y \in N(\mathfrak{h})$, $h \in \mathfrak{h}$. Then

$$[h, [x, y]] = [[y, h], x] + [[h, x], y] \in \mathfrak{h}$$

by the Jacobi identity, and so $N(\mathfrak{h})$ is a subalgebra of $\mathfrak{g}$. The second statement is obvious from the definition of $N(\mathfrak{h})$. For the third statement, notice that if $\mathfrak{h}$ is an ideal of $\mathfrak{m}$, then $[\mathfrak{h}, \mathfrak{m}] \subset \mathfrak{h}$ so that $\mathfrak{m} \subset N(\mathfrak{h})$. □

A subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is a **Cartan subalgebra** if $\mathfrak{h}$ is nilpotent and $N(\mathfrak{h}) = \mathfrak{h}$. We wish to show that every Lie algebra contains a Cartan subalgebra.

Let $x \in \mathfrak{g}$ and consider the map $\operatorname{ad} x : \mathfrak{g} \to \mathfrak{g}$. Define $\mathfrak{g}_{0,x}$ to be the generalized eigenspace of $\operatorname{ad} x$ with eigenvalue 0, that is,

$$\mathfrak{g}_{0,x} = \{y \in \mathfrak{g} \mid (\operatorname{ad} x)^n y = 0 \quad \text{for some } n \geq 1\}.$$

We call $\mathfrak{g}_{0,x}$ the **null component** of $\mathfrak{g}$ with respect to x.

An element $x \in \mathfrak{g}$ is **regular** if $\dim \mathfrak{g}_{0,x}$ is as small as possible. Clearly, any Lie algebra will contain regular elements.

Theorem 3.2. *If x is a regular element of $\mathfrak{g}$, then the null component $\mathfrak{g}_{0,x}$ is a Cartan subalgebra of $\mathfrak{g}$.*

Proof. Let $\mathfrak{h} = \mathfrak{g}_{0,x}$. We first show that $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$. Let $y, z \in \mathfrak{h}$. Then

$$(\operatorname{ad} x)^n [y, z] = \sum_{i=0}^{n} \left[(\operatorname{ad} x)^i y, (\operatorname{ad} x)^{n-i} z\right]$$

by Proposition 2.7 with $V = \mathfrak{g}$ and $\alpha = \beta = 0$. Since $y \in \mathfrak{h}$, we have $(\operatorname{ad} x)^i y = 0$ if i is sufficiently large. And since $z \in \mathfrak{h}$, we have $(\operatorname{ad} x)^{n-i} z = 0$ if $n - i$ is sufficiently large. Hence $(\operatorname{ad} x)^n [y, z] = 0$ if n is sufficiently large. This shows that $[y, z] \in \mathfrak{h}$, and so $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$.

We next show that $\mathfrak{h}$ is nilpotent. Set $\dim \mathfrak{h} = l$ and let $b_1, \ldots, b_l$ be a basis of $\mathfrak{h}$. Let $y = \lambda_1 b_i + \cdots + \lambda_l b_l \in \mathfrak{h}$ where $\lambda_1, \ldots, \lambda_l \in \mathbb{C}$. Since $\mathfrak{h}$ is a subalgebra, the linear map $\operatorname{ad} y : \mathfrak{g} \to \mathfrak{g}$ restricts to $\operatorname{ad} y : \mathfrak{h} \to \mathfrak{h}$. This induces the linear map $\operatorname{ad} y : \mathfrak{g}/\mathfrak{h} \to \mathfrak{g}/\mathfrak{h}$. Let $\chi(t)$ be the characteristic polynomial of $\operatorname{ad} y$ on $\mathfrak{g}$, $\chi_1(t)$ the characteristic polynomial of $\operatorname{ad} y$ on $\mathfrak{h}$, and $\chi_2(t)$ the characteristic polynomial of $\operatorname{ad} y$ on $\mathfrak{g}/\mathfrak{h}$. We have

$$\chi(t) = \chi_1(t)\chi_2(t).$$

Now $\chi(t) = \det(t1 - \operatorname{ad} y)$ and y depends linearly on $\lambda_1, \ldots, \lambda_l$. Hence the coefficients of $\chi(t)$ are polynomial functions of $\lambda_1, \ldots, \lambda_l$. The same applies to $\chi_1(t)$ and $\chi_2(t)$. Let

$$\chi_2(t) = d_0 + d_1 t + d_2 t^2 + \cdots$$

where $d_0, d_1, d_2, \ldots$ are polynomial functions of $\lambda_1, \ldots, \lambda_l$. Note that d_0 is not the zero polynomial, for in the special case $y = x$, the eigenvalues of $\operatorname{ad} y$ on $\mathfrak{g}/\mathfrak{h}$ are nonzero and d_0 is their product. Let

$$\chi_1(t) = t^m \left(c_0 + c_1 t + c_2 t^2 + \cdots\right)$$

where $c_0, c_1, c_2, \ldots$ are polynomial functions of $\lambda_1, \ldots, \lambda_l$ and c_0 is not the zero polynomial. Then $m \leq l = \deg \chi_1(t)$. It follows that

$$\chi(t) = t^m (c_0 d_0 + \text{a sum involving positive powers of } t).$$

Now $c_0 d_0$ is not the zero polynomial, and thus we may choose elements $\lambda_1, \ldots, \lambda_l$ in $\mathbb{C}$ such that $c_0 d_0$ is nonzero. For such an element y we have $\dim \mathfrak{g}_{0,y} = m$. Since x is a regular element with $\dim \mathfrak{g}_{0,x} = l$, we have $l \leq m$. But we also have $m \leq l$, and so $m = l$. Thus t^l divides $\chi_1(t)$, and since $\dim \chi_1(t) = l$, we have $\chi_1(t) = t^l$. It follows by the Cayley-Hamilton theorem that for any $y \in \mathfrak{h}$, $(\operatorname{ad} y)^l : \mathfrak{h} \to \mathfrak{h}$ is the zero map. Thus $\operatorname{ad} y$ is nilpotent for all $y \in \mathfrak{h}$, and so $\mathfrak{h}$ is nilpotent by Theorem 2.11.

Finally, we show that $N(\mathfrak{h}) = \mathfrak{h}$. Clearly $\mathfrak{h} \subset N(\mathfrak{h})$. So let $z \in N(\mathfrak{h})$. Then $[x, z] \in \mathfrak{h}$, and hence $(\operatorname{ad} x)^n [x, z] = 0$ for some $n \geq 1$. But this means $(\operatorname{ad} x)^{n+1} z = 0$, and so $z \in \mathfrak{h}$. It follows that $N(\mathfrak{h}) \subset \mathfrak{h}$, and thus $N(\mathfrak{h}) = \mathfrak{h}$. □

3.2 Derivations and automorphisms

A **derivation** of a Lie algebra $\mathfrak{g}$ is a linear map $D : \mathfrak{g} \to \mathfrak{g}$ satisfying

$$D[x, y] = [Dx, y] + [x, Dy] \qquad \text{for all } x, y \in \mathfrak{g}.$$

Proposition 3.3. $\operatorname{ad} x$ *is a derivation for all* $x \in \mathfrak{g}$.

Proof. If $y, z \in \mathfrak{g}$, then

$$\operatorname{ad} x \cdot [y, z] = [x, [y, z]] = [[x, y], z] + [y, [x, z]] = [\operatorname{ad} x \cdot y, z] + [y, \operatorname{ad} x \cdot z]$$

by the Jacobi identity. Thus $\operatorname{ad} x$ is a derivation of $\mathfrak{g}$. □

An **automorphism** of $\mathfrak{g}$ is an isomorphism $\phi : \mathfrak{g} \to \mathfrak{g}$. The automorphisms of $\mathfrak{g}$ form a group $\mathrm{Aut}(\mathfrak{g})$ under composition of maps.

Proposition 3.4. *If D is a nilpotent derivation of $\mathfrak{g}$, then $\exp(D)$ is an automorphism of $\mathfrak{g}$.*

Proof. Since D is nilpotent, we have $D^n = 0$ for some n. It is easy to check that the Leibniz rule holds for the linear map D, that is,

$$D^r[x, y] = \sum_{i=0}^{r} \binom{r}{i} \left[D^i x, D^{r-i} y\right] \qquad \text{for all } x, y \in \mathfrak{g}.$$

It follows that

$$\begin{aligned}
[\exp(D)x, \exp(D)y] &= \left[\sum_{i=0}^{n-1} \frac{D^i x}{i!}, \sum_{j=0}^{n-1} \frac{D^j y}{j!}\right] \\
&= \sum_{k=0}^{2n-2} \sum_{l=0}^{k} \left[\frac{D^l x}{l!}, \frac{D^{k-l} y}{(k-l)!}\right] \\
&= \sum_{k=0}^{2n-2} \frac{1}{k!} \sum_{l=0}^{k} \binom{k}{l} \left[D^l x, D^{k-l} y\right] \\
&= \sum_{k=0}^{2n-2} \frac{D^k [x, y]}{k!} \qquad \text{(by the Leibniz rule)} \\
&= \exp(D)[x, y].
\end{aligned}$$

Hence $\exp(D) : \mathfrak{g} \to \mathfrak{g}$ is a homomorphism. Likewise, $\exp(-D) : \mathfrak{g} \to \mathfrak{g}$ is a homomorphism, and we see that $\exp(D) \cdot \exp(-D) = 1$. Thus $\exp(D)$ is an automorphism. □

An **inner automorphism** of $\mathfrak{g}$ is an automorphism of the form $\exp(\operatorname{ad} x)$ for $x \in \mathfrak{g}$ with $\operatorname{ad} x$ nilpotent. The **inner automorphism group** $\mathrm{Inn}(\mathfrak{g})$ is the subgroup of $\mathrm{Aut}(\mathfrak{g})$ generated by all inner automorphisms.

Proposition 3.5. *$\mathrm{Inn}(\mathfrak{g})$ is a normal subgroup of $\mathrm{Aut}(\mathfrak{g})$.*

Proof. Let $\phi \in \mathrm{Aut}(\mathfrak{g})$. Every element of $\mathrm{Inn}(\mathfrak{g})$ is of the form

$$\exp(\operatorname{ad} x_1) \cdot \exp(\operatorname{ad} x_2) \cdot \cdots \cdot \exp(\operatorname{ad} x_r)$$

where $x_1, \ldots, x_r \in \mathfrak{g}$ and $\operatorname{ad} x_1, \ldots, \operatorname{ad} x_r$ are nilpotent. We have

$$\begin{aligned}\exp(\operatorname{ad} x_1) \cdot \exp(\operatorname{ad} x_2) \cdot \cdots \cdot \exp(\operatorname{ad} x_r) &= \exp(\operatorname{ad} x_1 + \operatorname{ad} x_2 + \cdots + \operatorname{ad} x_r) \\ &= \exp(\operatorname{ad}(x_1 + x_2 + \cdots + x_r)).\end{aligned}$$

Let $x = x_1 + x_2 + \cdots + x_r$. For all $y \in \mathfrak{g}$, we have

$$\phi(\operatorname{ad} x)\phi^{-1} y = \phi[x, \phi^{-1} y] = [\phi x, y] = \operatorname{ad} \phi x \cdot y.$$

Thus $\phi(\operatorname{ad} x)\phi^{-1} = \operatorname{ad} \phi x$. By linearity, we have $\phi(\exp(\operatorname{ad} x))\phi^{-1} = \exp(\operatorname{ad} \phi x)$. Hence

$$\phi(\exp(\operatorname{ad} x_1) \cdot \exp(\operatorname{ad} x_2) \cdot \cdots \cdot \exp(\operatorname{ad} x_r))\phi^{-1} = \exp(\operatorname{ad} \phi x_1) \cdot \exp(\operatorname{ad} \phi x_2) \cdot \cdots \cdot \exp(\operatorname{ad} \phi x_r).$$

Since ϕ is an automorphism, we see that $\operatorname{ad} \phi x_1, \ldots, \operatorname{ad} \phi x_r$ are nilpotent, and so $\phi(\exp(\operatorname{ad} x_1) \cdot \exp(\operatorname{ad} x_2) \cdot \cdots \cdot \exp(\operatorname{ad} x_r))\phi^{-1} \in \operatorname{Inn}(\mathfrak{g})$. □

Two subalgebras $\mathfrak{h}, \mathfrak{k}$ of $\mathfrak{g}$ are **conjugate** in $\mathfrak{g}$ if there exists a $\phi \in \operatorname{Inn}(\mathfrak{g})$ such that $\phi(\mathfrak{h}) = \mathfrak{k}$. We shall prove that any two Cartan subalgebras of $\mathfrak{g}$ are conjugate in $\mathfrak{g}$. We first require some concepts from algebraic geometry.

3.3 Concepts from algebraic geometry

If $\mathfrak{h}$ is a nilpotent subalgebra of a Lie algebra $\mathfrak{g}$, then $\mathfrak{g}$ forms an $\mathfrak{h}$-module via the restriction of the adjoint action. This action gives rise to the weight space decomposition

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

as in Theorem 2.10 where

$$\mathfrak{g}_\lambda = \{x \in \mathfrak{g} \mid (\forall h \in \mathfrak{h})(\exists n \geq 1) \quad (\operatorname{ad} h - \lambda 1)^n x = 0\}.$$

It follows from Corollary 2.12 that $\mathfrak{h} \subset \mathfrak{g}_0$. Suppose our nilpotent subalgebra $\mathfrak{h}$ is equal to $\mathfrak{g}_0$. Then there exist nonzero 1-dimensional representations $\lambda_1, \ldots, \lambda_r$ of $\mathfrak{h}$ such that

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{g}_{\lambda_1} \oplus \cdots \oplus \mathfrak{g}_{\lambda_r}.$$

Thus if $x \in \mathfrak{g}$, then $x = x_0 + x_1 + \cdots + x_r$ where $x_0 \in \mathfrak{h}$ and $x_i \in \mathfrak{g}_{\lambda_i}$ for $i = 1, \ldots, r$.

Proposition 3.6. $\operatorname{ad} x_i$ *is nilpotent for all* $i \neq 0$.

Proof. Let $\mu : \mathfrak{h} \to \mathbb{C}$ be a weight of the $\mathfrak{h}$-module $\mathfrak{g}$ and let $y \in \mathfrak{g}_\mu$. We have

$$(\operatorname{ad} h - \mu(h)1 - \lambda_i(h)1)^n [x_i, y] = \sum_{j=0}^{n} \left[(\operatorname{ad} h - \lambda_i(h)1)^j x_i, (\operatorname{ad} h - \mu(h)1)^{n-j} y \right]$$

for all $h \in \mathfrak{h}$ by Proposition 2.7. Since $x_i \in \mathfrak{g}_{\lambda_i}$, we have $(\operatorname{ad} h - \lambda_i(h)1)^j x_i = 0$ if j is sufficiently large. And since $y \in \mathfrak{g}_\mu$, we have $(\operatorname{ad} h - \mu(h)1)^{n-j} y = 0$ if $n-j$ is sufficiently large. Hence $(\operatorname{ad} h - \mu(h)1 - \lambda_i(h)1)^n [x_i, y] = 0$ if n is sufficiently large. This shows that $[x_i, y] \in \mathfrak{g}_{\lambda_i + \mu}$, and thus $\operatorname{ad} x_i \cdot \mathfrak{g}_\mu \subset \mathfrak{g}_{\lambda_i + \mu}$. Now because $\lambda_i \neq 0$ and there are only finitely-many $\mu : \mathfrak{h} \to \mathbb{C}$ such that $\mathfrak{g}_\mu \neq 0$, we see that $(\operatorname{ad} x_i)^N = 0$ if N is sufficiently large. Thus $\operatorname{ad} x_i$ is nilpotent. □

It follows immediately that $\exp(\operatorname{ad} x_i) \in \operatorname{Inn}(\mathfrak{g})$ for $i \neq 0$. We define the map $f : \mathfrak{g} \to \mathfrak{g}$ by

$$f(x) = \exp(\operatorname{ad} x_1) \cdot \exp(\operatorname{ad} x_2) \cdot \cdots \cdot \exp(\operatorname{ad} x_r) \cdot x_0.$$

Let $\{b_{ij}\}$ be a basis of $\mathfrak{g}$ where $i = 1, \ldots, r$ and, for a fixed i, the elements b_{ij} form a basis of $\mathfrak{g}_{\lambda_i}$ with respect to which elements of $\mathfrak{h}$ are represented by upper-triangular matrices as in Corollary 2.5. (Here $\lambda_0 = 0$.)

Proposition 3.7. $f : \mathfrak{g} \to \mathfrak{g}$ *is a polynomial function, that is,*

$$f\left(\sum \lambda_{ij} b_{ij} \right) = \sum \mu_{ij} b_{ij}$$

where each μ_{ij} *is a polynomial in the* λ_{kl}.

Proof. Each map $\operatorname{ad} x_i : \mathfrak{g} \to \mathfrak{g}$ is clearly linear. We also have

$$\exp(\operatorname{ad} x_i) = \sum_{k=0}^{N} \frac{(\operatorname{ad} x_i)^k}{k!}$$

for some $N \geq 1$ since $\operatorname{ad} x_i$ is nilpotent. Thus $\exp(\operatorname{ad} x_i) : \mathfrak{g} \to \mathfrak{g}$ is a polynomial function. The map f is a composition of the linear map $x \mapsto x_0$ with the polynomial functions $\exp(\operatorname{ad} x_i)$ and is thus a polynomial function. □

We now write $\mu_{ij} = f_{ij}(\lambda_{kl})$. We define the Jacobian matrix

$$J(f) = (\partial f_{ij}/\partial \lambda_{kl})$$

and the Jacobian determinant $\det J(f)$ of f. Clearly $\det J(f)$ is an element of the polynomial ring $\mathbb{C}[\lambda_{kl}]$.

Proposition 3.8. $\det J(f)$ *is not the zero polynomial.*

Proof. We shall show that $\det J(f) \neq 0$ when evaluated at a carefully chosen element of $\mathfrak{h}$. Let $h \in \mathfrak{h}$ and consider $(\partial f_{ij}/\partial \lambda_{kl})_h$. Suppose $k \neq 0$. We have

$$\begin{aligned}
(\partial f/\partial \lambda_{kl})_h &= \lim_{t\to\infty} \frac{f(h + tb_{kl}) - f(h)}{t} \\
&= \lim_{t\to\infty} \frac{\exp(\operatorname{ad} tb_{kl})h - h}{t} \\
&= \lim_{t\to\infty} \frac{h + t[b_{kl}, h] + \cdots - h}{t} \\
&= [b_{kl}, h] = -[h, b_{kl}] \\
&\equiv -\lambda_k(h) b_{kl} \pmod{\langle b_{k1}, \ldots, b_{kl-1}\rangle}.
\end{aligned}$$

Now suppose $k = 0$. We have

$$\begin{aligned}
(\partial f/\partial \lambda_{0l})_h &= \lim_{t\to\infty} \frac{f(h + tb_{0l}) - f(h)}{t} \\
&= \lim_{t\to\infty} \frac{h + tb_{0l} - h}{t} = b_{0l}.
\end{aligned}$$

Thus $J(f)$ is a block matrix of the form

$$\begin{pmatrix}
J_0 & & & & & 0 \\
& J_1 & & & & \\
& & \cdot & & & \\
& & & \cdot & & \\
& & & & \cdot & \\
0 & & & & & J_r
\end{pmatrix}$$

where J_0 is the identity matrix and J_i is of the form

$$\begin{pmatrix} -\lambda_i(h) & & & & * \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ 0 & & & & -\lambda_i(h) \end{pmatrix}$$

for $i = 1, \ldots, r$. It follows that $(\det J(f))_h = \pm \prod_{i=1}^{r} \lambda_i(h)^{d_i}$ where $d_i = \dim \mathfrak{g}_{\lambda_i}$. Now the vector space $\mathfrak{h}$ over the infinite field $\mathbb{C}$ cannot be expressed as the union of finitely-many proper subspaces. For each i, the set of $h \in \mathfrak{h}$ satisfying $\lambda_i(h) = 0$ is a proper subspace of $\mathfrak{h}$ since the linear maps λ_i are all nonzero. Thus there exists an $h \in \mathfrak{h}$ such that $\lambda_i(h) \neq 0$ for all i. For such an h, we have $(\det J(f))_h \neq 0$, showing that $\det J(f)$ is not the zero polynomial. □

Proposition 3.9. *The polynomial functions f_{ij} are algebraically independent.*

Proof. Suppose there exists a nonzero polynomial $F(x_{ij}) \in \mathbb{C}(x_{ij})$ such that $F(f_{ij}) = 0$. We choose an F with minimal total degree in the variables x_{ij}. For each index k, l, we have

$$\frac{\partial}{\partial \lambda_{kl}} F(f_{ij}) = 0$$

so that

$$\sum_{i,j} \frac{\partial F}{\partial f_{ij}} \frac{\partial f_{ij}}{\partial \lambda_{kl}} = 0.$$

Let v be the vector $(\partial F / \partial f_{ij})$. Then $vJ(f) = (0, \ldots, 0)$. Since $\det J(f) = 0$, we must have $v = (0, \ldots, 0)$, and hence $\partial F / \partial f_{ij} = 0$ for all f_{ij}. Notice that $\partial F / \partial x_{ij}$ is a polynomial in $\mathbb{C}[x_{ij}]$ of smaller total degree than F. By our choice of F, $\partial F / \partial x_{ij}$ must be the zero polynomial, that is, F does not involve the indeterminate x_{ij}. This is true for all x_{ij}, and so F must be a constant polynomial. Since $F(f_{ij}) = 0$, this constant must be zero. This contradicts the fact that F is not the zero polynomial, and so the f_{ij} must be algebraically independent. □

Let $B = \mathbb{C}[f_{ij}]$ be the polynomial ring in the f_{ij} and $A = \mathbb{C}[\lambda_{ij}]$ the polynomial ring in the λ_{ij}. Consider the homomorphism $\theta : B \to A$ uniquely determined by

$$\theta(f_{ij}) = f_{ij}(\lambda_{kl}) \in A.$$

Proposition 3.10. *The homomorphism $\theta : B \to A$ is injective.*

Proof. Let $F \in \ker\theta$, that is, $\theta(F) = 0$. Then $F(f_{ij}) = 0$ as a function of the λ_{kl}. Since the f_{ij} are algebraically independent, this implies $F = 0$. Thus $\ker\theta = 0$, and so θ is injective. □

It follows that θ is an embedding of B into A. The following is a standard result from algebraic geometry.

Proposition 3.11. *Let A and B be integral domains such that $B \subset A$, A, B have common identity element 1, and A is finitely-generated over B. If p is a nonzero element of A, then there exists a nonzero element q of B such that any homomorphism $\phi : B \to \mathbb{C}$ with $\phi(q) \neq 0$ extends to a homomorphism $\psi : A \to \mathbb{C}$ with $\psi(p) \neq 0$.*

Proof. We refer the reader to [Car05, Proposition 3.10]. □

We wish to apply this result to our earlier situation. Let $d = \dim\mathfrak{g}$. We have the polynomial function

$$f : \mathbb{C}^d \to \mathbb{C}^d$$

$$(\lambda_{ij}) \mapsto (f_{ij}(\lambda_{kl})).$$

Let $V = \mathbb{C}^d$. For each polynomial $p \in \mathbb{C}[x_{ij}]$, we define the set

$$V_p = \{v \in V \mid p(v) \neq 0\}.$$

Corollary 3.12. *For each nonzero polynomial $p \in \mathbb{C}[x_{ij}]$, there exists a nonzero polynomial $q \in \mathbb{C}[x_{ij}]$ such that $V_q \subset f(V_p)$.*

Proof. Consider the integral domains $B \subset A$ described earlier, that is, $B = \mathbb{C}[f_{ij}]$ and $A = \mathbb{C}[\lambda_{ij}]$. We may regard p as a polynomial in A. By Proposition 3.11, we know there exists a nonzero $q \in B$ such that any homomorphism $\phi : B \to \mathbb{C}$ with $\phi(q) \neq 0$ extends to a homomorphism $\psi : A \to \mathbb{C}$ with $\psi(p) \neq 0$. Now each $v \in V_q$ induces such a homomorphism ϕ via the relation $\phi(f_{ij}) = v_{ij}$. Let ψ be an extension of ϕ as described above and define the vector $w \in \mathbb{C}^d$ by $w_{ij} = \psi(\lambda_{ij})$. Clearly, $w \in V_p$. By the way we embedded B into A, we have

$$v_{ij} = \phi(f_{ij}) = \psi(f_{ij}(\lambda_{kl})) = f_{ij}(\psi(\lambda_{kl})) = f_{ij}(w),$$

and so $v = f(w)$. Hence $v \in f(V_p)$ and $V_q \subset f(V_p)$ as desired. □

3.4 Conjugacy of Cartan subalgebras

We saw in Theorem 3.2 that the null component of any regular element of $\mathfrak{g}$ is a Cartan subalgebra. In this section we prove the converse: that every Cartan subalgebra is the null component of some regular element of $\mathfrak{g}$. Finally, we prove that the null components of any two regular elements are conjugate, and thus any Cartan subalgebra of $\mathfrak{g}$ is unique up to isomorphism.

Proposition 3.13. *If $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$, then there exists a regular element $x \in \mathfrak{g}$ such that $\mathfrak{h} = \mathfrak{g}_{0,x}$.*

Proof. Since $\mathfrak{h}$ is nilpotent, we may regard $\mathfrak{g}$ as an $\mathfrak{h}$-module and decompose $\mathfrak{g}$ into weight spaces with respect to $\mathfrak{h}$ as in Theorem 2.10. It is clear from Corollary 2.12 that $\mathfrak{h}$ is contained in the zero weight space $\mathfrak{g}_0$. Suppose $\mathfrak{h} \neq \mathfrak{g}_0$. Let $\mathfrak{m}/\mathfrak{h}$ be an irreducible $\mathfrak{h}$-submodule of $\mathfrak{g}_0/\mathfrak{h}$. By Theorem 2.4, we have $\dim \mathfrak{m}/\mathfrak{h} = 1$. As in the proof of Theorem 2.11, the 1-dimensional representation induced by $\mathfrak{m}/\mathfrak{h}$ must be the zero map. Hence $[\mathfrak{h}, \mathfrak{m}] \subset \mathfrak{h}$, and so $\mathfrak{m} \subset N(\mathfrak{h})$. This contradicts $\mathfrak{h} = N(\mathfrak{h})$, so we must have $\mathfrak{h} = \mathfrak{g}_0$. Let

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{g}_{\lambda_1} \oplus \cdots \oplus \mathfrak{g}_{\lambda_r} \qquad \lambda_1, \ldots, \lambda_r \neq 0$$

be the weight space decomposition of $\mathfrak{g}$ with respect to $\mathfrak{h}$. For $x \in \mathfrak{g}$, we write $x = x_0 + x_1 + \cdots + x_r$ where $x_0 \in \mathfrak{h}$ and $x_i \in \mathfrak{g}_{\lambda_i}$ for $i = 1, \ldots, r$. We define the polynomial function $f : \mathfrak{g} \to \mathfrak{g}$ by

$$f(x) = \exp(\operatorname{ad} x_1) \cdot \exp(\operatorname{ad} x_2) \cdot \cdots \cdot \exp(\operatorname{ad} x_r) \cdot x_0$$

as in the previous section. We also define $p : \mathfrak{g} \to \mathbb{C}$ by

$$p(x) = \lambda_1(x_0)\lambda_2(x_0) \cdots \lambda_r(x_0).$$

Then p is a polynomial function on $\mathfrak{g}$. Clearly p is not the zero polynomial since we can find an element $x_0 \in \mathfrak{h}$ such that $\lambda_i(x_0) \neq 0$ for all $i = 1, \ldots, r$ as in the proof of Proposition 3.8. By Corollary 3.12, there exists a nonzero polynomial function $q : \mathfrak{g} \to \mathbb{C}$ such that $\mathfrak{g}_q \subset f(\mathfrak{g}_p)$.

Now consider the set R of regular elements of $\mathfrak{g}$. Let $y \in \mathfrak{g}$ and let

$$\chi(y) = \det(t1 - \operatorname{ad} y) = t^n + \mu_1(y)t^{n-1} + \cdots + \mu_n(y)$$

be the characteristic polynomial of $\operatorname{ad} y$ on $\mathfrak{g}$. Then $\mu_1, \ldots, \mu_n$ are polynomial functions on $\mathfrak{g}$. There exists a unique integer k such that μ_{n-k} is not the zero polynomial but $\mu_{n-k+1}, \ldots, \mu_n$ are identically zero. The generalized eigenspace of $\operatorname{ad} y$ with eigenvalue 0 has dimension k if $\mu_{n-k} = 0$ and dimension greater than k if $\mu_{n-k}(y) = 0$. Hence $y \in R$ if and only if $\mu_{n-k}(y) \neq 0$.

Now q and μ_{n-k} are nonzero polynomials. Thus $q\mu_{n-k}$ is a nonzero polynomial, and so there exists a $y \in \mathfrak{g}$ such that $(q\mu_{n-k})(y) \neq 0$. Then $q(y) \neq 0$ and $\mu_{n-k}(y) \neq 0$ so that $y \in \mathfrak{g}_q \cap R$. Since $\mathfrak{g}_q \subset f(\mathfrak{g}_p)$, there exists an $x \in \mathfrak{g}_p$ such that $y = f(x)$, that is

$$y = \exp(\operatorname{ad} x_1) \cdot \exp(\operatorname{ad} x_2) \cdot \cdots \cdot \exp(\operatorname{ad} x_r) \cdot x_0.$$

Thus y and x_0 are conjugate in $\mathfrak{g}$. Since y is regular, x_0 must also be regular.

I claim that $\mathfrak{h} = \mathfrak{g}_{0,x_0}$. We know that $x_0 \in \mathfrak{h}$ and that $\mathfrak{h}$ is nilpotent. Hence $\mathfrak{h} \subset \mathfrak{g}_{0,x_0}$ by Corollary 2.12. Suppose $\mathfrak{h} \neq \mathfrak{g}_{0,x_0}$ and choose $z \in \mathfrak{g}_{0,x_0} \backslash \mathfrak{h}$. We write

$z = z_0 + z_1 + \cdots + z_r$ where $z_0 \in \mathfrak{h}$ and $z_i \in \mathfrak{g}_{\lambda_i}$ for $i = 1, \ldots, r$. Because $z \in \mathfrak{g}_{0,x_0}$, we have $(\operatorname{ad} x_0)^n z = 0$ for some $n \geq 1$, that is,

$$(\operatorname{ad} x_0)^n z_0 + (\operatorname{ad} x_0)^n z_1 + \cdots + (\operatorname{ad} x_0)^n z_r = 0.$$

For any $i = 1, \ldots, r$, we have

$$(\operatorname{ad} x_0)^n z_i = -(\operatorname{ad} x_0)^n z_0 - \cdots - (\operatorname{ad} x_0)^n z_{i-1} - (\operatorname{ad} x_0)^n z_{i+1} - \cdots - (\operatorname{ad} x_0)^n z_r.$$

Each weight space is invariant under the action of $\operatorname{ad} x_0$, so the left-hand side is in $\mathfrak{g}_{\lambda_i}$ and right-hand side is in its complement. This implies $(\operatorname{ad} x_0)^n z_i = 0$, and so $z_i \in \mathfrak{g}_{0,x_0}$ for all i. We also know that $z \notin \mathfrak{h}$. Thus $z_i \neq 0$ for some $i \geq 1$. Since $z_i \in \mathfrak{g}_{\lambda_i}$, we have $(\operatorname{ad} x_0 - \lambda_i(x_0)1)^m z_i = 0$ for some $m \geq 1$. Hence z_i is in the generalized eigenspace of $\operatorname{ad} x_0$ with eigenvalue $\lambda_i(x_0)$. Because $z_i \in \mathfrak{g}_{0,x_0}$, we must have $\lambda_i(x_0) = 0$. But $x \in \mathfrak{g}_p$, and so

$$\lambda_1(x_0)\lambda_2(x_0)\cdots\lambda_r(x_0) \neq 0.$$

In particular, this implies $\lambda_i(x_0) \neq 0$, a contradiction. It follows that $\mathfrak{h} = \mathfrak{g}_{0,x_0}$. □

Theorem 3.14. *Any two Cartan subalgebras of $\mathfrak{g}$ are conjugate.*

Proof. Let $\mathfrak{h}, \mathfrak{h}'$ be Cartan subalgebras of $\mathfrak{g}$. We regard $\mathfrak{g}$ as an $\mathfrak{h}$-module and decompose $\mathfrak{g}$ into weight spaces with respect to $\mathfrak{h}$. We know from the proof of Proposition 3.13 that $\mathfrak{h} = \mathfrak{g}_0$. Let

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{g}_{\lambda_1} \oplus \cdots \oplus \mathfrak{g}_{\lambda_r} \qquad \lambda_1, \ldots, \lambda_r \neq 0$$

be the weight space decomposition of $\mathfrak{g}$ with respect to $\mathfrak{h}$. For $x \in \mathfrak{g}$, we write $x = x_0 + x_1 + \cdots + x_r$ where $x_0 \in \mathfrak{h}$ and $x_i \in \mathfrak{g}_{\lambda_i}$ for $i = 1, \ldots, r$. For each $x_0 \in \mathfrak{h}$, we have $\mathfrak{h} \subset \mathfrak{g}_{0,x_0}$. By dimension considerations, we see that x_0 is regular if and only if $\mathfrak{h} = \mathfrak{g}_{0,x_0}$. We know from the proof of Proposition 3.13 that the condition

$$\lambda_1(x_0)\lambda_2(x_0)\cdots\lambda_r(x_0) \neq 0$$

implies $\mathfrak{h} = \mathfrak{g}_{0,x_0}$. Now suppose $\mathfrak{h} = \mathfrak{g}_{0,x_0}$ but that $\lambda_i(x_0) = 0$ for some $i \geq 1$. Choosing a nonzero $y \in \mathfrak{g}_{\lambda_i}$, we have $(\text{ad } x_0 - \lambda_i(x_0)1)^n y = 0$ for some $n \geq 1$. Thus $y \in \mathfrak{g}_0$, contradicting the fact that $\lambda_i \neq 0$. Hence $\mathfrak{h} = \mathfrak{g}_{0,x_0}$ if and only if

$$\lambda_1(x_0)\lambda_2(x_0)\cdots\lambda_r(x_0) \neq 0.$$

It follows that this condition is also equivalent to x_0 being regular.

Consider the polynomial function $f : \mathfrak{g} \to \mathfrak{g}$ defined by

$$f(x) = \exp(\text{ad } x_1) \cdot \exp(\text{ad } x_2) \cdot \cdots \cdot \exp(\text{ad } x_r) \cdot x_0.$$

Define the polynomial function $p : \mathfrak{g} \to \mathbb{C}$ by

$$p(x) = \lambda_1(x_0)\lambda_2(x_0)\cdots\lambda_r(x_0).$$

Note that p is not the zero polynomial since $p(x_0) \neq 0$ when x_0 is regular. By Corollary 3.12, there exists a nonzero polynomial $q : \mathfrak{g} \to \mathbb{C}$ such that $\mathfrak{g}_q \subset f(\mathfrak{g}_p)$.

We now consider the second Cartan subalgebra $\mathfrak{h}'$ and define the corresponding functions $f' : \mathfrak{g} \to \mathfrak{g}$ and $p' : \mathfrak{g} \to \mathbb{C}$. We know there exists a nonzero polynomial function $q' : \mathfrak{g} \to \mathbb{C}$ such that $\mathfrak{g}_{q'} \subset f'(\mathfrak{g}_{p'})$.

Since q and q' are nonzero polynomials, qq' is a nonzero polynomial, and so there exists a $z \in \mathfrak{g}$ such that $(qq')(z) \neq 0$. Then $q(z) \neq 0$ and $q'(z) \neq 0$ so that $z \in \mathfrak{g}_q \cap \mathfrak{g}_{q'}$. It follows that $z \in f(\mathfrak{g}_p) \cap f'(\mathfrak{g}_{p'})$. Thus there exists an $x \in \mathfrak{g}$ such that $z = f(x)$ and $p(x) \neq 0$. Similarly, there exists an $x' \in \mathfrak{g}$ such that $z = f'(x')$ and $p'(x') = 0$. We have

$$z = \exp(\text{ad } x_1) \cdot \exp(\text{ad } x_2) \cdot \cdots \cdot \exp(\text{ad } x_r) \cdot x_0.$$

Hence z and x_0 are conjugate in $\mathfrak{g}$. Since $p(x_0) \neq 0$, we see that x_0 is regular. Likewise, z is conjugate to x_0' and x_0' is regular. Thus we have found regular elements $x_0 \in \mathfrak{h}$ and $x_0' \in \mathfrak{h}'$ such that x_0 and x_0' are conjugate in $\mathfrak{g}$. Now $\mathfrak{h} = \mathfrak{g}_{0,x_0}$ and $\mathfrak{h}' = \mathfrak{g}_{0,x_0'}$ since x_0 and x_0' are regular. Any automorphism mapping x_0 to x_0' will transform $\mathfrak{h}$ into $\mathfrak{h}'$. It follows that $\mathfrak{h}$ and $\mathfrak{h}'$ are conjugate in $\mathfrak{g}$. □

The dimension of the Cartan subalgebras of $\mathfrak{g}$ is called the **rank** of $\mathfrak{g}$.

Chapter 4

The Cartan decomposition

4.1 Root spaces

We saw in the last chapter that any two Cartan subalgebras of a Lie algebra $\mathfrak{g}$ are conjugate in $\mathfrak{g}$. Thus any weight space decomposition of $\mathfrak{g}$ with respect to a Cartan subalgebra $\mathfrak{h}$ will be unique up to isomorphism. Let $\bigoplus_\lambda \mathfrak{g}_\lambda$ be such a decomposition. By the proof of Proposition 3.13, we have $\mathfrak{g}_0 = \mathfrak{h}$. Thus

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{g}_{\lambda_1} \oplus \cdots \oplus \mathfrak{g}_{\lambda_r} \qquad \lambda_1, \ldots, \lambda_r \neq 0.$$

A 1-dimensional representation λ of $\mathfrak{h}$ is called a **root** of $\mathfrak{g}$ with respect to $\mathfrak{h}$ if $\lambda \neq 0$ and $\mathfrak{g}_\lambda \neq 0$. We denote the set of all roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$ by Φ. Thus

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

We call this decomposition the **Cartan decomposition** of $\mathfrak{g}$ with respect to $\mathfrak{h}$. Each $\mathfrak{g}_\alpha$ is called the **root space** of α.

Proposition 4.1. *If λ and μ are 1-dimensional representations of $\mathfrak{h}$, then*

$$[\mathfrak{g}_\lambda, \mathfrak{g}_\mu] \subset \mathfrak{g}_{\lambda+\mu}.$$

Proof. We know that $[\mathfrak{g}_\lambda, \mathfrak{g}_\mu]$ is generated by elements of the form $[y, z]$ with $y \in \mathfrak{g}_\lambda$ and $z \in \mathfrak{g}_\mu$. Let $x \in \mathfrak{h}$. By Proposition 2.7, we have

$$(\operatorname{ad} x - \lambda(x)1 - \mu(x)1)^n [y, z] = \sum_{i=0}^{n} \binom{n}{i} \left[(\operatorname{ad} x - \lambda(x)1)^i y, (\operatorname{ad} x - \mu(x)1)^{n-i} z\right].$$

Since $y \in \mathfrak{g}_\lambda$, we have $(\operatorname{ad} x - \lambda(x)1)^i y = 0$ if i is sufficiently large. And since $z \in \mathfrak{g}_\mu$, we have $(\operatorname{ad} x - \mu(x)1)^{n-i} z = 0$ if $n - i$ is sufficiently large. Hence

$$(\operatorname{ad} x - \lambda(x)1 - \mu(x)1)^n [y, z] = 0$$

if n is sufficiently large. It follows that $[y, z] \in \mathfrak{g}_{\lambda+\mu}$, and so $[\mathfrak{g}_\lambda, \mathfrak{g}_\mu] \subset \mathfrak{g}_{\lambda+\mu}$. □

Corollary 4.2. *Let $\alpha, \beta \in \Phi$ be roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$. Then*

$$\begin{aligned} &[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta} && \text{if } \alpha + \beta \in \Phi \\ &[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{h} && \text{if } \beta = -\alpha \\ &[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = 0 && \text{if } \alpha + \beta \neq 0 \quad \text{and} \quad \alpha + \beta \notin \Phi. \end{aligned}$$

Proof. This is immediate from Proposition 4.1 and the fact that $\mathfrak{h} = \mathfrak{g}_0$. □

Proposition 4.3. *Let $\alpha \in \Phi$. Given any $\beta \in \Phi$, there exists a number $r \in \mathbb{Q}$, depending on α and β, such that $\beta = r\alpha$ on the subspace $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ of $\mathfrak{h}$.*

Proof. If $-\alpha$ is not a weight of $\mathfrak{g}$ with respect to $\mathfrak{h}$, then $\mathfrak{g}_{-\alpha} = 0$, and the proof is trivial. So assume $-\alpha$ is a weight. Then since $\alpha \neq 0$, we must have $-\alpha \in \Phi$. For $i \in \mathbb{Z}$, we consider the function $i\alpha + \beta : \mathfrak{h} \to \mathbb{C}$. Since Φ is finite, there exist integers p and q with $p \geq 0$ and $q \geq 0$ such that

$$-p\alpha + \beta, \ldots, \beta, \ldots, q\alpha + \beta$$

are all in Φ but $-(p+1)\alpha + \beta$ and $(q+1)\alpha + \beta$ are not in Φ. If either $-(p+1)\alpha + \beta = 0$ or $(q+1)\alpha + \beta = 0$, then the result is obvious. So assume $-(p+1)\alpha + \beta \neq 0$ and $(q+1)\alpha + \beta \neq 0$. Let $\mathfrak{m}$ be the subspace of $\mathfrak{g}$ given by

$$\mathfrak{m} = \mathfrak{g}_{-p\alpha+\beta} \oplus \cdots \oplus \mathfrak{g}_{q\alpha+\beta}.$$

Now $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ is generated by elements of the form $x = [y, z]$ with $y \in \mathfrak{g}_\alpha$ and $z \in \mathfrak{g}_{-\alpha}$. We have $\operatorname{ad} y \cdot \mathfrak{g}_{q\alpha+\beta} \subset \mathfrak{g}_{(q+1)\alpha+\beta}$ by Proposition 4.1. Because $(q+1)\alpha + \beta \neq 0$ and

$(q+1)\alpha+\beta \notin \Phi$, we must have $\mathfrak{g}_{(q+1)\alpha+\beta} = 0$. Thus $\operatorname{ad} y \cdot \mathfrak{m} \subset \mathfrak{m}$. By a similar argument, we have $\operatorname{ad} z \cdot \mathfrak{m} \subset \mathfrak{m}$, and so

$$\operatorname{ad} x \cdot \mathfrak{m} = (\operatorname{ad} y \ \operatorname{ad} z - \operatorname{ad} z \ \operatorname{ad} y)\mathfrak{m} \subset \mathfrak{m}.$$

We calculate the trace $\operatorname{tr}_{\mathfrak{m}}(\operatorname{ad} x)$. Since $x \in \mathfrak{h}$, each weight space $\mathfrak{g}_{i\alpha+\beta}$ is invariant under $\operatorname{ad} x$. Thus

$$\operatorname{tr}_{\mathfrak{m}}(\operatorname{ad} x) = \sum_{i=-p}^{q} \operatorname{tr}_{\mathfrak{g}_{i\alpha+\beta}}(\operatorname{ad} x).$$

Now $\operatorname{ad} x$ acts on $\mathfrak{g}_{i\alpha+\beta}$ via a matrix of the form

$$\begin{pmatrix} (i\alpha+\beta)(x) & & & * \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & (i\alpha+\beta)(x) \end{pmatrix}.$$

Thus $\operatorname{tr}_{\mathfrak{g}_{i\alpha+\beta}}(\operatorname{ad} x) = \dim \mathfrak{g}_{i\alpha+\beta}(i\alpha+\beta)(x)$. It follows that

$$\begin{aligned} \operatorname{tr}_{\mathfrak{m}}(\operatorname{ad} x) &= \sum_{i=-p}^{q} \dim \mathfrak{g}_{i\alpha+\beta}(i\alpha+\beta)(x) \\ &= \left(\sum_{i=-p}^{q} i \dim \mathfrak{g}_{i\alpha+\beta} \right) \alpha(x) + \left(\sum_{i=-p}^{q} \dim \mathfrak{g}_{i\alpha+\beta} \right) \beta(x). \end{aligned}$$

But we also have

$$\operatorname{tr}_{\mathfrak{m}}(\operatorname{ad} x) = \operatorname{tr}_{\mathfrak{m}}(\operatorname{ad} y \ \operatorname{ad} z - \operatorname{ad} z \ \operatorname{ad} y) = \operatorname{tr}_{\mathfrak{m}}(\operatorname{ad} y \ \operatorname{ad} z) - \operatorname{tr}_{\mathfrak{m}}(\operatorname{ad} z \ \operatorname{ad} y) = 0.$$

Hence

$$\left(\sum_{i=-p}^{q} i \dim \mathfrak{g}_{i\alpha+\beta} \right) \alpha(x) + \left(\sum_{i=-p}^{q} \dim \mathfrak{g}_{i\alpha+\beta} \right) \beta(x) = 0.$$

We know that $\dim \mathfrak{g}_{i\alpha+\beta} > 0$ for all $-p \le i \le q$. Thus

$$\beta(x) = -\frac{\left(\sum_{i=-p}^{q} i \dim \mathfrak{g}_{i\alpha+\beta} \right)}{\left(\sum_{i=-p}^{q} \dim \mathfrak{g}_{i\alpha+\beta} \right)} \alpha(x).$$

Hence $\beta(x) = r\alpha(x)$ for some $r \in \mathbb{Q}$ independent of x. It follows that $\beta = r\alpha$ on $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$. □

4.2 The Killing form

In order to further understand the Cartan decomposition of $\mathfrak{g}$, we define a map

$$\mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$$

$$(x, y) \mapsto \langle x, y \rangle$$

by

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

One easily checks that this map is bilinear. We call this bilinear form the **Killing form** of $\mathfrak{g}$.

Proposition 4.4.

(i) *The Killing form is symmetric, i.e.,* $\langle x, y \rangle = \langle y, x \rangle$ *for all* $x, y \in \mathfrak{g}$.

(ii) *The Killing form is invariant, i.e.,*

$$\langle [x, y], z \rangle = \langle x, [y, z] \rangle \qquad \text{for all } x, y, z \in \mathfrak{g}.$$

Proof. The first statement follows form the fact that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ for all square matrices A, B. To prove the second statement, we calculate

$$\begin{aligned} \langle [x, y], z \rangle &= \operatorname{tr}(\operatorname{ad} [x, y] \ \operatorname{ad} z) = \operatorname{tr}((\operatorname{ad} x \ \operatorname{ad} y - \operatorname{ad} y \ \operatorname{ad} x) \ \operatorname{ad} z) \\ &= \operatorname{tr}(\operatorname{ad} x \ \operatorname{ad} y \ \operatorname{ad} z) - \operatorname{tr}(\operatorname{ad} y \ \operatorname{ad} x \ \operatorname{ad} z) \\ &= \operatorname{tr}(\operatorname{ad} x \ \operatorname{ad} y \ \operatorname{ad} z) - \operatorname{tr}(\operatorname{ad} x \ \operatorname{ad} z \ \operatorname{ad} y) \\ &= \operatorname{tr}((\operatorname{ad} x \ (\operatorname{ad} y \ \operatorname{ad} z - \operatorname{ad} z \ \operatorname{ad} y)) = \operatorname{tr}(\operatorname{ad} x \ \operatorname{ad} [y, z]) = \langle x, [y, z] \rangle. \end{aligned}$$ □

Proposition 4.5. *Let* $\mathfrak{a}$ *be an ideal of* $\mathfrak{g}$ *and let* $x, y \in \mathfrak{a}$. *Then*

$$\langle x, y \rangle_{\mathfrak{a}} = \langle x, y \rangle_{\mathfrak{g}}.$$

Hence the killing form of $\mathfrak{g}$ *restricted to* $\mathfrak{a}$ *is the Killing form of* $\mathfrak{a}$.

Proof. We choose a basis of $\mathfrak{a}$ and extend it to a basis of $\mathfrak{g}$. With respect to this basis, $\operatorname{ad} x : \mathfrak{g} \to \mathfrak{g}$ is represented by a matrix of the form

$$\begin{pmatrix} A_1 & A_2 \\ 0 & 0 \end{pmatrix}$$

since $x \in \mathfrak{a}$. Similarly, $\operatorname{ad} y : \mathfrak{g} \to \mathfrak{g}$ is represented by a matrix of the form

$$\begin{pmatrix} B_1 & B_2 \\ 0 & 0 \end{pmatrix}.$$

Thus $\operatorname{ad} x \operatorname{ad} y : \mathfrak{g} \to \mathfrak{g}$ is represented by the matrix

$$\begin{pmatrix} A_1B_1 & A_1B_2 \\ 0 & 0 \end{pmatrix}.$$

Hence $\operatorname{tr}_{\mathfrak{a}}(\operatorname{ad} x \operatorname{ad} y) = \operatorname{tr}(A_1B_1) = \operatorname{tr}_{\mathfrak{g}}(\operatorname{ad} x \operatorname{ad} y)$, and so $\langle x, y\rangle_{\mathfrak{a}} = \langle x, y\rangle_{\mathfrak{g}}$. □

For any subspace $\mathfrak{m}$ of $\mathfrak{g}$, we define the set

$$\mathfrak{m}^{\perp} = \{x \in \mathfrak{g} \mid \langle x, y\rangle = 0 \quad \text{for all } y \in \mathfrak{m}\}.$$

One easily checks that $\mathfrak{m}^{\perp}$ is a subspace of $\mathfrak{g}$.

Proposition 4.6. *If $\mathfrak{a}$ is an ideal of $\mathfrak{g}$, then $\mathfrak{a}^{\perp}$ is an ideal of $\mathfrak{g}$.*

Proof. The subspace $[\mathfrak{a}^{\perp}, \mathfrak{g}]$ is generated by elements of the form $[x, y]$ with $x \in \mathfrak{a}^{\perp}$ and $y \in \mathfrak{g}$. For all $z \in \mathfrak{a}$, we have

$$\langle [x, y], z\rangle = \langle x, [y, z]\rangle = 0$$

since $[y, z] \in \mathfrak{a}$. This shows that $[x, y] \in \mathfrak{a}^{\perp}$. It follows that $[\mathfrak{a}^{\perp}, \mathfrak{g}] \subset \mathfrak{a}^{\perp}$, and so $\mathfrak{a}^{\perp}$ is an ideal of $\mathfrak{g}$. □

In particular, we see that $\mathfrak{g}^{\perp}$ is an ideal of $\mathfrak{g}$. The Killing form of $\mathfrak{g}$ is said to be **nondegenerate** if $\mathfrak{g}^{\perp} = 0$. This is equivalent to the condition that if $\langle x, y\rangle = 0$ for all $y \in \mathfrak{g}$, then $x = 0$. The Killing form of $\mathfrak{g}$ is **identically zero** if $\mathfrak{g}^{\perp} = \mathfrak{g}$. This means that $\langle x, y\rangle = 0$ for all $x, y \in \mathfrak{g}$.

Proposition 4.7. *Let $\mathfrak{g}$ be a Lie algebra such that $\mathfrak{g} \neq 0$ and $\mathfrak{g}^2 = \mathfrak{g}$. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$. Then there exists an $x \in \mathfrak{h}$ such that $\langle x, x \rangle \neq 0$.*

Proof. Let $\mathfrak{g} = \oplus \mathfrak{g}_\lambda$ be the Cartan decomposition of $\mathfrak{g}$. Then

$$\mathfrak{g}^2 = [\mathfrak{g}, \mathfrak{g}] = \left[\bigoplus_\lambda \mathfrak{g}_\lambda, \bigoplus_\lambda \mathfrak{g}_\lambda \right] = \sum_{\lambda,\mu} [\mathfrak{g}_\lambda, \mathfrak{g}_\mu].$$

We have $[\mathfrak{g}_\lambda, \mathfrak{g}_\mu] \subset \mathfrak{g}_{\lambda+\mu}$ by Proposition 4.1. Thus $[\mathfrak{g}_\lambda, \mathfrak{g}_{-\lambda}] \subset \mathfrak{g}_0$, while $[\mathfrak{g}_\lambda, \mathfrak{g}_\mu]$ is contained in the complement of $\mathfrak{g}_0$ in $\mathfrak{g}$ if $\mu \neq -\lambda$. Since $\mathfrak{g} = \mathfrak{g}^2$, we must have

$$\mathfrak{g}_0 = \sum_\lambda [\mathfrak{g}_\lambda, \mathfrak{g}_{-\lambda}]$$

summed over all weights λ such that $-\lambda$ is also a weight. We also have $\mathfrak{g}_0 = \mathfrak{h}$ by the proof of Proposition 3.13. Thus

$$\mathfrak{h} = [\mathfrak{h}, \mathfrak{h}] + \sum_\alpha [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$$

summed over all roots α such that $-\alpha$ is also a root. Note that $\mathfrak{g}$ is not nilpotent since $\mathfrak{g}^2 = \mathfrak{g} \neq 0$. But we know that $\mathfrak{h}$ is nilpotent, and so $\mathfrak{h} \neq \mathfrak{g}$. Thus there exists at least one root $\beta \in \Phi$. Now β is a 1-dimensional representation of $\mathfrak{h}$, and so β vanishes on $[\mathfrak{h}, \mathfrak{h}]$ by Lemma 2.3. But β does not vanish on $\mathfrak{h}$ since $\beta \neq 0$. Using the above decomposition of $\mathfrak{h}$, we see there exists some root $\alpha \in \Phi$ such that $-\alpha \in \Phi$ and β does not vanish on $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$. Choose an $x \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ such that $\beta(x) \neq 0$. Then

$$\langle x, x \rangle = \operatorname{tr}(\operatorname{ad} x \ \operatorname{ad} x) = \sum_\lambda \dim \mathfrak{g}_\lambda (\lambda(x))^2$$

since $\operatorname{ad} x$ is represented on $\mathfrak{g}_\lambda$ by a matrix of the form

$$\begin{pmatrix} \lambda(x) & & & * \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \lambda(x) \end{pmatrix}.$$

For each λ, there exists an $r_{\lambda,\alpha} \in \mathbb{Q}$ such that $\lambda(x) = r_{\lambda,\alpha}\alpha(x)$ by Proposition 4.3. Thus

$$\langle x, x\rangle = \left(\sum_{\lambda} \dim \mathfrak{g}_\lambda r_{\lambda,\alpha}^2\right) \alpha(x)^2.$$

Now $\beta(x) = r_{\beta,\alpha}\alpha(x)$ and $\beta(x) \neq 0$. Thus $r_{\beta,\alpha} \neq 0$ and $\alpha(x) \neq 0$. It follows that $\langle x, x\rangle \neq 0$. □

Theorem 4.8. *If the Killing form of* $\mathfrak{g}$ *is identically zero, then* $\mathfrak{g}$ *is solvable.*

Proof. We proceed by induction on $\dim \mathfrak{g}$. If $\dim \mathfrak{g} = 1$, then $\mathfrak{g}$ is clearly solvable. So assume $\dim \mathfrak{g} > 1$. By the contrapositive of Proposition 4.7, we see that $\mathfrak{g} \neq \mathfrak{g}^2$. Now $\mathfrak{g}^2$ is an ideal of $\mathfrak{g}$, so the Killing form of $\mathfrak{g}^2$ is the restriction of the Killing form of $\mathfrak{g}$ by Proposition 4.5. Hence the Killing form of $\mathfrak{g}^2$ is identically zero. It follows by induction that $\mathfrak{g}^2$ is solvable. We also have $\left(\mathfrak{g}/\mathfrak{g}^2\right)^2 = 0$, and so $\mathfrak{g}/\mathfrak{g}^2$ is solvable. Thus $\mathfrak{g}$ is solvable by Proposition 1.14. □

Theorem 4.9 (Cartan's criterion)**.** *A Lie algebra* $\mathfrak{g}$ *is semisimple if and only if the Killing form of* $\mathfrak{g}$ *is nondegenerate.*

Proof. We shall prove the contrapositive of the theorem. Suppose the Killing form of $\mathfrak{g}$ is degenerate. Then $\mathfrak{g}^\perp \neq 0$. We know that $\mathfrak{g}^\perp$ is an ideal by Proposition 4.6. Thus the Killing form of $\mathfrak{g}^\perp$ is identically zero by Proposition 4.5. This implies $\mathfrak{g}^\perp$ is solvable by Theorem 4.8. Thus $\mathfrak{g}$ has a nonzero solvable ideal, and so $\mathfrak{g}$ is not semisimple.

Now suppose $\mathfrak{g}$ is not semisimple. Then the solvable radical $\mathfrak{r}$ of $\mathfrak{g}$ is nonzero. Consider the chain of subspaces

$$\mathfrak{r} = \mathfrak{r}^{(0)} \supset \mathfrak{r}^{(1)} \supset \mathfrak{r}^{(2)} \supset \cdots \supset \mathfrak{r}^{(k-1)} \supset \mathfrak{r}^{(k)} = 0.$$

Each subspace $\mathfrak{r}^{(i)}$ is an ideal of $\mathfrak{g}$ since the product of two ideals is an ideal. Let $\mathfrak{a} = \mathfrak{r}^{(k-1)}$. Then $\mathfrak{a}$ is a nonzero ideal such that $\mathfrak{a}^2 = 0$. We choose a basis of $\mathfrak{a}$ and extend it to a basis of $\mathfrak{g}$. Let $x \in \mathfrak{a}$, $y \in \mathfrak{g}$. With respect to our chosen basis, $\operatorname{ad} x$ is represented by a matrix of the form

$$\begin{pmatrix} 0 & A \\ 0 & 0 \end{pmatrix}$$

since $\mathfrak{a}^2 = 0$ and $\mathfrak{a}$ is an ideal of $\mathfrak{g}$, and $\operatorname{ad} y$ is represented by a matrix of the form

$$\begin{pmatrix} B_1 & B_2 \\ 0 & B_3 \end{pmatrix}.$$

Thus $\operatorname{ad} x \operatorname{ad} y$ is represented by the matrix

$$\begin{pmatrix} 0 & AB_3 \\ 0 & 0 \end{pmatrix}.$$

Hence $\langle x, y\rangle = \operatorname{tr}(\operatorname{ad} x \operatorname{ad} y) = 0$. This holds for all $x \in \mathfrak{a}$, $y \in \mathfrak{g}$, and so $\mathfrak{a} \subset \mathfrak{g}^{\perp}$. Thus $\mathfrak{g}^{\perp} \neq 0$, and so the Killing form of $\mathfrak{g}$ is degenerate. □

We now define the direct sum of Lie algebras $\mathfrak{g}_1, \mathfrak{g}_2$ by taking the vector space $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ and defining the multiplication component-wise, that is,

$$[(x_1, x_2), (y_1, y_2)] = ([x_1, y_1], [x_2, y_2]).$$

We also define the subspaces $\mathfrak{a}_1 = \{(x_1, 0) \mid x_1 \in \mathfrak{g}_1\}$ and $\mathfrak{a}_2 = \{(0, x_2) \mid x_2 \in \mathfrak{g}_2\}$. One easily checks that $\mathfrak{a}_1$ and $\mathfrak{a}_2$ are ideals of $\mathfrak{g}_1 \oplus \mathfrak{g}_2$, that $\mathfrak{a}_1 \cap \mathfrak{a}_2 = 0$, and that $\mathfrak{a}_1 + \mathfrak{a}_2 = \mathfrak{g}_1 \oplus \mathfrak{g}_2$. Furthermore, we have $\mathfrak{g}_1 \cong \mathfrak{a}_1$ and $\mathfrak{g}_2 \cong \mathfrak{a}_2$ via the projections $(x_1, 0) \mapsto x_1$ and $(0, x_2) \mapsto x_2$, respectively.

Conversely, let $\mathfrak{g}$ be a Lie algebra containing two ideals $\mathfrak{a}_1, \mathfrak{a}_2$ such that $\mathfrak{a}_1 \cap \mathfrak{a}_2 = 0$ and $\mathfrak{a}_1 + \mathfrak{a}_2 = \mathfrak{g}$. We shall show that the Lie algebra $\mathfrak{a}_1 \oplus \mathfrak{a}_2$ is isomorphic to $\mathfrak{g}$ under the map

$$\phi : \mathfrak{a}_1 \oplus \mathfrak{a}_2 \to \mathfrak{g}$$

$$(x_1, x_2) \mapsto x_1 + x_2.$$

Clearly, ϕ is an isomorphism of vector spaces. We have $[\mathfrak{a}_1, \mathfrak{a}_2] \subset \mathfrak{a}_1 \cap \mathfrak{a}_2 = 0$ so that

$$\begin{aligned} [\phi(x_1, x_2), \phi(y_1, y_2)] &= [x_2 + x_2, y_1 + y_2] = [x_1, y_1] + [x_2, y_2] \\ &= \phi([x_1, y_1], [x_2, y_2]) = \phi[(x_1, x_2), (y_1, y_2)]. \end{aligned}$$

Hence ϕ preserves Lie multiplication. This shows that if a Lie algebra has complementary ideals $\mathfrak{a}_1, \mathfrak{a}_2$, then it is isomorphic to $\mathfrak{a}_1 \oplus \mathfrak{a}_2$.

We may generalize these results to include the direct sum $\mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_n$ of more than two Lie algebras.

Theorem 4.10. *A Lie algebra $\mathfrak{g}$ is semisimple if and only if $\mathfrak{g}$ is isomorphic to a direct sum of non-trivial simple Lie algebras.*

Proof. Suppose $\mathfrak{g}$ is semisimple. If $\mathfrak{g}$ is simple, then $\mathfrak{g}$ must be non-trivial since the trivial simple Lie algebra is not semisimple. So suppose $\mathfrak{g}$ is not simple. Let $\mathfrak{a}$ be a maximal ideal of $\mathfrak{g}$. Then $\mathfrak{a} \neq 0$ and $\mathfrak{a} \neq \mathfrak{g}$. Consider the subspace $\mathfrak{a}^\perp$ of $\mathfrak{g}$. This is an ideal of $\mathfrak{g}$ by Proposition 4.6. The Killing form of $\mathfrak{g}$ is nondegenerate by Theorem 4.9. Thus an element $x \in \mathfrak{g}$ lies in $\mathfrak{a}^\perp$ if and only if the coordinates of x with respect to a basis of $\mathfrak{g}$ satisfy $\dim \mathfrak{a}$ homogeneous linear equations that are linearly independent. Hence

$$\dim \mathfrak{a}^\perp = \dim \mathfrak{g} - \dim \mathfrak{a}.$$

Now consider the subspace $\mathfrak{a} \cap \mathfrak{a}^\perp$. This is an ideal of $\mathfrak{g}$ by Proposition 1.5. Thus the Killing form of $\mathfrak{a} \cap \mathfrak{a}^\perp$ is the restriction of the Killing form of $\mathfrak{g}$ by Proposition 4.5. Hence the Killing form of $\mathfrak{a} \cap \mathfrak{a}^\perp$ is identically zero, and so $\mathfrak{a} \cap \mathfrak{a}^\perp$ is solvable by Theorem 4.8. Since $\mathfrak{g}$ is semisimple, we have $\mathfrak{a} \cap \mathfrak{a}^\perp = 0$. Thus

$$\begin{aligned} \dim(\mathfrak{a} + \mathfrak{a}^\perp) &= \dim \mathfrak{a} + \dim \mathfrak{a}^\perp - \dim(\mathfrak{a} \cap \mathfrak{a}^\perp) \\ &= \dim \mathfrak{a} + \dim \mathfrak{a}^\perp = \dim \mathfrak{g}. \end{aligned}$$

Hence $\mathfrak{a} + \mathfrak{a}^\perp = \mathfrak{g}$, and so $\mathfrak{g}$ is the direct sum of the ideals $\mathfrak{a}$ and $\mathfrak{a}^\perp$. It follows that $\mathfrak{g} \cong \mathfrak{a} \oplus \mathfrak{a}^\perp$.

We shall show that $\mathfrak{a}$ is a simple Lie algebra. Let $\mathfrak{b}$ be an ideal of $\mathfrak{a}$. We have

$$[\mathfrak{b}, \mathfrak{g}] \subset [\mathfrak{b}, \mathfrak{a}] + [\mathfrak{b}, \mathfrak{a}^\perp] = [\mathfrak{b}, \mathfrak{a}] \subset \mathfrak{b}$$

since $[\mathfrak{b}, \mathfrak{a}^\perp] \subset [\mathfrak{a}, \mathfrak{a}^\perp] \subset \mathfrak{a} \cap \mathfrak{a}^\perp = 0$. Thus $\mathfrak{b}$ is an ideal of $\mathfrak{g}$ contained in $\mathfrak{a}$. Since $\mathfrak{a}$ is a maximal ideal of $\mathfrak{g}$, we must have $\mathfrak{b} = 0$ or $\mathfrak{b} = \mathfrak{a}$. Thus $\mathfrak{a}$ is simple.

We now show that $\mathfrak{a}^\perp$ is semisimple. Let $\mathfrak{b}$ be a solvable ideal of $\mathfrak{a}^\perp$. Then

$$[\mathfrak{b}, \mathfrak{g}] \subset [\mathfrak{b}, \mathfrak{a}] + [\mathfrak{b}, \mathfrak{a}^\perp] = [\mathfrak{b}, \mathfrak{a}^\perp] \subset \mathfrak{b}$$

since $[\mathfrak{b}, \mathfrak{a}] \subset [\mathfrak{a}^\perp, \mathfrak{a}] \subset \mathfrak{a} \cap \mathfrak{a}^\perp = 0$. Thus $\mathfrak{b}$ is an ideal of $\mathfrak{g}$. Since $\mathfrak{g}$ is semisimple and $\mathfrak{b}$ is solvable, we must have $\mathfrak{b} = 0$. Thus $\mathfrak{a}^\perp$ is semisimple.

Because $\mathfrak{a} \neq 0$, we see that $\dim \mathfrak{a}^\perp < \dim \mathfrak{g}$. By induction, we may assume $\mathfrak{a}^\perp$ is a direct sum of non-trivial simple Lie algebras. Since $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$ and $\mathfrak{a}$ is simple and non-trivial, $\mathfrak{g}$ is also a direct sum of non-trivial simple Lie algebras.

Conversely, suppose

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_r$$

where each $\mathfrak{g}_i$ is a non-trivial simple Lie algebra. Then each $\mathfrak{g}_i$ is semisimple by Proposition 1.16, and thus has a nondegenerate Killing form by Theorem 4.9. Each $\mathfrak{g}_i$ is also an ideal of $\mathfrak{g}$. Suppose $i \neq j$ and let $x_i \in \mathfrak{g}_i$, $x_j \in \mathfrak{g}_j$. Then

$$\operatorname{ad} x_i \operatorname{ad} x_j \cdot y \in \mathfrak{g}_i \cap \mathfrak{g}_j = 0 \qquad \text{for all } y \in \mathfrak{g},$$

and so $\langle x_i, x_j \rangle = \operatorname{tr}(\operatorname{ad} x_i \operatorname{ad} x_j) = 0$. Now let $x = x_1 + x_2 + \cdots + x_r \in \mathfrak{g}^\perp$ with $x_i \in \mathfrak{g}_i$. If $y_i \in \mathfrak{g}_i$, then by the above result, we have

$$0 = \langle x, y_i \rangle = \sum_{j=1}^{r} \langle x_j, y_i \rangle = \langle x_i, y_i \rangle.$$

Since this holds for all $y_i \in \mathfrak{g}_i$, we must have $x_i = 0$. This holds for all i, and so $x = 0$. Thus the Killing form of $\mathfrak{g}$ is nondegenerate, and so $\mathfrak{g}$ is semisimple. □

4.3 The Cartan decomposition of a semisimple Lie algebra

In this section, without further comment, we shall assume $\mathfrak{g}$ is a semisimple Lie algebra. The Cartan decomposition of a semisimple Lie algebra gives us more information than that of the general case. This information will be useful in studying the simple case, which is the goal of this paper.

As usual, we denote the Cartan decomposition of $\mathfrak{g}$ with respect to $\mathfrak{h}$ by $\mathfrak{g} = \bigoplus_\lambda \mathfrak{g}_\lambda$ where $\mathfrak{g}_0 = \mathfrak{h}$.

Proposition 4.11. *If $\mu \neq -\lambda$, then $\mathfrak{g}_\lambda$ and $\mathfrak{g}_\mu$ are orthogonal with respect to the Killing form.*

Proof. Let $x \in \mathfrak{g}_\lambda$, $y \in \mathfrak{g}_\mu$. For every weight space $\mathfrak{g}_\nu$, we have

$$\operatorname{ad} x \operatorname{ad} y \cdot \mathfrak{g}_\nu \subset \mathfrak{g}_{\lambda+\mu+\nu}$$

by Proposition 4.1. We choose a basis of $\mathfrak{g}$ adapted to the Cartan decomposition. With respect to such a basis, ad x ad y is represented by a block matrix of the form

$$\begin{pmatrix} 0 & & & * \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ * & & & 0 \end{pmatrix}$$

since $\lambda + \mu + \nu \neq \nu$. It follows that $\langle x, y\rangle = \operatorname{tr}(\operatorname{ad} x \operatorname{ad} y) = 0$, and so $\mathfrak{g}_\lambda$ is orthogonal to $\mathfrak{g}_\mu$. □

Proposition 4.12. *If α is a root of $\mathfrak{g}$ with respect to $\mathfrak{h}$, then $-\alpha$ is also a root.*

Proof. Recall that α is a root if $\alpha \neq 0$ and $\mathfrak{g}_\alpha \neq 0$. Suppose $-\alpha$ is not a root. Then since $-\alpha \neq 0$, we must have $\mathfrak{g}_{-\alpha} = 0$. By Proposition 4.11, this implies that $\mathfrak{g}_\alpha$ is orthogonal to all $\mathfrak{g}_\lambda$, and thus $\mathfrak{g}_\alpha \subset \mathfrak{g}^\perp$. But $\mathfrak{g}$ is semisimple, and so $\mathfrak{g}^\perp = 0$ by Theorem 4.9. Thus $\mathfrak{g}_\alpha = 0$, contradicting the fact that α is a root. □

Proposition 4.13. *The Killing form of $\mathfrak{g}$ remains nondegenerate on restriction to $\mathfrak{h}$.*

Proof. Let $x \in \mathfrak{h}$ and suppose $\langle x, y\rangle = 0$ for all $y \in \mathfrak{h}$. We also have $\langle x, y\rangle = 0$ for all $y \in \mathfrak{g}_\alpha$ where $\alpha \in \Phi$ by Proposition 4.11. Thus $\langle x, y\rangle = 0$ for all $y \in \mathfrak{g}$, and so $x \in \mathfrak{g}^\perp$. But $\mathfrak{g}^\perp = 0$ since $\mathfrak{g}$ is semisimple, and so $x = 0$. □

Note that the Killing form of $\mathfrak{g}$ restricted to $\mathfrak{h}$ is not the same as the Killing form of $\mathfrak{h}$, because $\mathfrak{h}$ is solvable and thus not semisimple.

Theorem 4.14. *The Cartan subalgebras of a semisimple Lie algebra are abelian.*

Proof. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$. For all $x \in [\mathfrak{h}, \mathfrak{h}]$, $y \in \mathfrak{h}$, we have

$$\langle x, y \rangle = \operatorname{tr}(\operatorname{ad} x \operatorname{ad} y) = \sum_{\lambda} \dim \mathfrak{g}_\lambda \, \lambda(x)\lambda(y)$$

since $\operatorname{ad} x \operatorname{ad} y$ is represented on $\mathfrak{g}_\lambda$ by a matrix of the form

$$\begin{pmatrix} \lambda(x)\lambda(y) & & & * \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \lambda(x)\lambda(y) \end{pmatrix}.$$

But λ is a 1-dimensional representation of $\mathfrak{h}$, and so λ vanishes on $[\mathfrak{h}, \mathfrak{h}]$. Thus $\lambda(x) = 0$. It follows that $\langle x, y \rangle = 0$ for all $y \in \mathfrak{h}$. Since the Killing form of $\mathfrak{g}$ restricted to $\mathfrak{h}$ is nondegenerate, this implies $x = 0$. Hence $[\mathfrak{h}, \mathfrak{h}] = 0$, and so $\mathfrak{h}$ is abelian. □

Let $\mathfrak{h}^* = \operatorname{Hom}(\mathfrak{h}, \mathbb{C})$ be the dual space of $\mathfrak{h}$. It follows from [Hal74, Theorem 15.2] that this is a vector space of linear maps from $\mathfrak{h}$ into $\mathbb{C}$ and that $\dim \mathfrak{h}^* = \dim \mathfrak{h}$. We define a map $\mathfrak{h} \to \mathfrak{h}^*$ using the Killing form of $\mathfrak{g}$. Given $h \in \mathfrak{h}$, we define $h^* \in \mathfrak{h}^*$ by

$$h^*(x) = \langle h, x \rangle \qquad \text{for all } x \in \mathfrak{h}.$$

Lemma 4.15. *The map $h \mapsto h^*$ is an isomorphism of vector spaces between $\mathfrak{h}$ and $\mathfrak{h}^*$.*

Proof. This map is clearly linear. Suppose h lies in the kernel. Then h^* is the zero map, that is, $\langle h, x \rangle = 0$ for all $x \in \mathfrak{h}$. It follows that $h = 0$ by Proposition 4.13, showing that the map is injective. By the rank-nullity theorem, this also proves that the map is surjective. □

Notice that Φ is a finite subset of $\mathfrak{h}^*$. Because the map $h \mapsto h^*$ is bijective, we know that for each $\alpha \in \Phi$, there exists a unique element $h'_\alpha \in \mathfrak{h}$ such that $h'^*_\alpha(x) = \alpha(x)$ for all $x \in \mathfrak{h}$, that is,

$$\alpha(x) = \left\langle h'_\alpha, x \right\rangle \qquad \text{for all } x \in \mathfrak{h}.$$

(The notation h_α would seem more natural, but this will be reserved for the coroot of α, which will be introduced in Chapter 7.)

Proposition 4.16. *The vectors h'_α for $\alpha \in \Phi$ span $\mathfrak{h}$.*

Proof. Suppose the vectors h'_α are contained in a proper subspace of $\mathfrak{h}$. Then the annihilator of this subspace is nonzero by [Hal74, Theorem 17.1]. Thus there exists a nonzero $x \in \mathfrak{h}$ such that $x^*\left(h'_\alpha\right) = 0$ for all $\alpha \in \Phi$, that is, $\left\langle h'_\alpha, x \right\rangle = 0$. Hence $\alpha(x) = 0$ for all $\alpha \in \Phi$. Let $y \in \mathfrak{h}$. Then

$$\langle x, y \rangle = \operatorname{tr}(\operatorname{ad} x \operatorname{ad} y) = \sum_\lambda \dim \mathfrak{g}_\lambda \, \lambda(x)\lambda(y) = 0$$

since $\lambda(x) = 0$ for all weights λ. Thus $\langle x, y \rangle = 0$ for all $y \in \mathfrak{h}$. Since the Killing form of $\mathfrak{g}$ restricted to $\mathfrak{h}$ is nondegenerate, this implies $x = 0$, a contradiction. $\square$

Proposition 4.17. *$h'_\alpha \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ for all $\alpha \in \Phi$.*

Proof. We know that $\mathfrak{g}_\alpha$ is an $\mathfrak{h}$-module by Theorem 2.10. Since all irreducible $\mathfrak{h}$-modules are 1-dimensional, $\mathfrak{g}_\alpha$ contains a 1-dimensional $\mathfrak{h}$-submodule $\mathbb{C}e_\alpha$. We have $[x, e_\alpha] = \alpha(x)e_\alpha$ for all $x \in \mathfrak{h}$. Let $y \in \mathfrak{g}_{-\alpha}$. Then $[e_\alpha, y] \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subset \mathfrak{h}$. I claim that $[e_\alpha, y] = \langle e_\alpha, y \rangle h'_\alpha$. We define the element

$$z = [e_\alpha, y] - \langle e_\alpha, y \rangle h'_\alpha \in \mathfrak{h}.$$

Let $x \in \mathfrak{h}$. Then

$$\begin{aligned} \langle x, z \rangle &= \langle x, [e_\alpha, y] \rangle - \langle e_\alpha, y \rangle \left\langle x, h'_\alpha \right\rangle \\ &= \langle [x, e_\alpha], y \rangle - \langle e_\alpha, y \rangle \alpha(x) \\ &= \alpha(x) \langle e_\alpha, y \rangle - \langle e_\alpha, y \rangle \alpha(x) = 0. \end{aligned}$$

Thus $\langle x, z\rangle = 0$ for all $x \in \mathfrak{h}$. Since the Killing form of $\mathfrak{g}$ restricted to $\mathfrak{h}$ is nondegenerate, this implies $z = 0$. Hence $[e_\alpha, y] = \langle e_\alpha, y\rangle h'_\alpha$ for all $y \in \mathfrak{g}_{-\alpha}$.

Now there exists a $y \in \mathfrak{g}_{-\alpha}$ such that $\langle e_\alpha, y\rangle \neq 0$. For otherwise e_α would be orthogonal to $\mathfrak{g}_{-\alpha}$, and thus to the whole of $\mathfrak{g}$ by Proposition 4.11. This would imply $e_\alpha \in \mathfrak{g}^\perp$. But $\mathfrak{g}^\perp = 0$ since $\mathfrak{g}$ is semisimple, and so $e_\alpha = 0$, a contradiction. Choosing a $y \in \mathfrak{g}_{-\alpha}$ such that $\langle e_\alpha, y\rangle \neq 0$, we have

$$h'_\alpha = \frac{1}{\langle e_\alpha, y\rangle}[e_\alpha, y] \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]. \qquad \square$$

Proposition 4.18. $\left\langle h'_\alpha, h'_\alpha\right\rangle \neq 0$ *for all* $\alpha \in \Phi$.

Proof. Suppose $\left\langle h'_\alpha, h'_\alpha\right\rangle = 0$ for some $\alpha \in \Phi$. Let β be any element of Φ. By Proposition 4.3, there exists an $r_{\beta,\alpha} \in \mathbb{Q}$ such that $\beta = r_{\beta,\alpha}\alpha$ on $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$. Now $h'_\alpha \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ by Proposition 4.17. Thus

$$\beta\left(h'_\alpha\right) = r_{\beta,\alpha}\alpha\left(h'_\alpha\right),$$

that is, $\left\langle h'_\beta, h'_\alpha\right\rangle = r_{\beta,\alpha}\left\langle h'_\alpha, h'_\alpha\right\rangle = 0$. This holds for all $\beta \in \Phi$. But the vectors h'_α for $\alpha \in \Phi$ span $\mathfrak{h}$ by Proposition 4.16, and so $\left\langle x, h'_\alpha\right\rangle = 0$ for all $x \in \mathfrak{h}$. Since the Killing form of $\mathfrak{g}$ restricted to $\mathfrak{h}$ is nondegenerate, this implies $h'_\alpha = 0$. Thus $\alpha = 0$, contradicting the fact that $\alpha \in \Phi$. $\square$

Having obtained many useful results, we are now in a position to prove one of the most important theorems about the Cartan decomposition of a semisimple Lie algebra.

Theorem 4.19. $\dim \mathfrak{g}_\alpha = 1$ *for all* $\alpha \in \Phi$.

Proof. Choose a 1-dimensional $\mathfrak{h}$-submodule $\mathbb{C}e_\alpha$ of $\mathfrak{g}_\alpha$ as in the proof of Proposition 4.17. By the same proof, we can find an $e_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that $[e_\alpha, e_{-\alpha}] = h'_\alpha$. Consider the subspace $\mathfrak{m}$ of $\mathfrak{g}$ given by

$$\mathfrak{m} = \mathbb{C}e_\alpha \oplus \mathbb{C}h'_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus \mathfrak{g}_{-2\alpha} \oplus \cdots.$$

There are only finitely-many summands of $\mathfrak{m}$ since Φ is finite. Thus there are only finitely-many non-negative integers r such that $\mathfrak{g}_{-r\alpha} \neq 0$. Observe that $\operatorname{ad} e_\alpha \cdot \mathfrak{m} \subset \mathfrak{m}$ because

$$[e_\alpha, e_\alpha] = 0,$$
$$\left[e_\alpha, h'_\alpha\right] = -\alpha\left(h'_\alpha\right) e_\alpha,$$
$$[e_\alpha, y] = \langle e_\alpha, y\rangle h'_\alpha \qquad \text{for all } y \in \mathfrak{g}_{-\alpha}$$

by the proof of Proposition 4.17, and

$$\operatorname{ad} e_\alpha \cdot \mathfrak{g}_{-r\alpha} \subset \mathfrak{g}_{-(r-1)\alpha} \qquad \text{for all } r \geq 2$$

by Proposition 4.1. Similarly, $\operatorname{ad} e_{-\alpha} \cdot \mathfrak{m} \subset \mathfrak{m}$ because

$$[e_{-\alpha}, e_\alpha] = h'_\alpha,$$
$$\left[e_{-\alpha}, h'_\alpha\right] = \alpha\left(h'_\alpha\right) e_{-\alpha},$$

and $\operatorname{ad} e_\alpha \cdot \mathfrak{g}_{-r\alpha} \subset \mathfrak{g}_{-(r+1)\alpha}$ for all $r \geq 1$. Now $h'_\alpha = [e_\alpha, e_{-\alpha}]$, and so

$$\operatorname{ad} h'_\alpha = \operatorname{ad} e_\alpha \operatorname{ad} e_{-\alpha} - \operatorname{ad} e_{-\alpha} \operatorname{ad} e_\alpha.$$

Thus $\operatorname{ad} h'_\alpha \cdot \mathfrak{m} \subset \mathfrak{m}$. We calculate the trace of $\operatorname{ad} h'_\alpha$ on $\mathfrak{m}$ in two different ways. First, we have

$$\begin{aligned}\operatorname{tr}_{\mathfrak{m}}\left(\operatorname{ad} h'_\alpha\right) &= \alpha\left(h'_\alpha\right) + \dim \mathfrak{g}_{-\alpha}\left(-\alpha\left(h'_\alpha\right)\right) + \dim \mathfrak{g}_{-2\alpha}\left(-2\alpha\left(h'_\alpha\right)\right) + \cdots \\ &= \alpha\left(h'_\alpha\right)(1 - \dim \mathfrak{g}_{-\alpha} - 2\dim \mathfrak{g}_{-2\alpha} - \cdots).\end{aligned}$$

Second, we have

$$\operatorname{tr}_{\mathfrak{m}}\left(h'_\alpha\right) = \operatorname{tr}_{\mathfrak{m}}(\operatorname{ad} e_\alpha \operatorname{ad} e_{-\alpha}) - \operatorname{tr}_{\mathfrak{m}}(\operatorname{ad} e_{-\alpha} \operatorname{ad} e_\alpha) = 0.$$

Thus

$$\alpha\left(h'_\alpha\right)(1 - \dim \mathfrak{g}_{-\alpha} - 2\dim \mathfrak{g}_{-2\alpha} - \cdots) = 0.$$

Now $\alpha\left(h'_\alpha\right) = \left\langle h'_\alpha, h'_\alpha\right\rangle \neq 0$ by Proposition 4.18, and so

$$1 - \dim \mathfrak{g}_{-\alpha} - 2\dim \mathfrak{g}_{-2\alpha} - \cdots = 0.$$

This can happen only if $\dim \mathfrak{g}_{-\alpha} = 1$ and $\dim \mathfrak{g}_{-r\alpha} = 0$ for all $r \geq 2$. Now $\alpha \in \Phi$ if and only if $-\alpha \in \Phi$ by Proposition 4.12. Thus $\dim \mathfrak{g}_\alpha = 1$ for all $\alpha \in \Phi$. □

Note that while all of the root spaces $\mathfrak{g}_\alpha$ are 1-dimensional, the space $\mathfrak{g}_0 = \mathfrak{h}$ need not be 1-dimensional.

Proposition 4.20. *If $\alpha \in \Phi$ and $r\alpha \in \Phi$ where $r \in \mathbb{Z}$, then $r = 1$ or $r = -1$.*

Proof. By the proof of Theorem 4.19, we have $\dim \mathfrak{g}_{-r\alpha} = 0$ for all $r \geq 2$, that is, $-r\alpha$ is not a root. Now $r\alpha \in \Phi$ if and only if $-r\alpha \in \Phi$ by Proposition 4.12. Thus only α and $-\alpha$ can be roots. □

We are now ready to examine some stronger properties of the set Φ of roots. Let $\alpha, \beta \in \Phi$ be roots such that $\beta \neq \alpha$ and $\beta \neq -\alpha$. Then by Proposition 4.20, β is not an integer multiple of α. There do, however, exist integers $p \geq 0, q \geq 0$ such that the elements

$$-p\alpha + \beta, \ldots, -\alpha + \beta, \beta, \alpha + \beta, \ldots, q\alpha + \beta$$

all lie in Φ but $-(p+1)\alpha + \beta$ and $(q+1)\alpha + \beta$ do not. The set of roots

$$-p\alpha + \beta, \ldots, q\alpha + \beta$$

is called the ***α*-chain** of roots through β.

Proposition 4.21. *Let α, β be roots such that $\beta \neq \alpha$ and $\beta \neq -\alpha$. Let*

$$-p\alpha + \beta, \ldots, \beta, \ldots, q\alpha + \beta$$

be the α-chain of roots through β. Then

$$2\frac{\langle h'_\alpha, h'_\beta \rangle}{\langle h'_\alpha, h'_\alpha \rangle} = p - q.$$

Proof. Consider the subspace $\mathfrak{m}$ of $\mathfrak{g}$ given by

$$\mathfrak{m} = \mathfrak{g}_{-p\alpha+\beta} \oplus \cdots \oplus \mathfrak{g}_{q\alpha+\beta}.$$

Recall that $h'_\alpha = [e_\alpha, e_{-\alpha}] \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ by the proof of Proposition 4.17. Now β is not an integer multiple of α, and so $-(p+1)\alpha + \beta \neq 0$ and $(q+1)\alpha + \beta \neq 0$. We have ad $e_\alpha \cdot \mathfrak{g}_{q\alpha+\beta} \subset \mathfrak{g}_{(q+1)\alpha+\beta}$ by Proposition 4.1. Because $(q+1)\alpha + \beta \neq 0$ and $(q+1)\alpha + \beta \notin \Phi$, we must have $\mathfrak{g}_{(q+1)\alpha+\beta} = 0$. Thus ad $e_\alpha \cdot \mathfrak{m} \subset \mathfrak{m}$. By a similar argument, we have ad $e_{-\alpha} \cdot \mathfrak{m} \subset \mathfrak{m}$, and so

$$\operatorname{ad} h'_\alpha \cdot \mathfrak{m} = (\operatorname{ad} e_\alpha \ \operatorname{ad} e_{-\alpha} - \operatorname{ad} e_{-\alpha} \ \operatorname{ad} e_\alpha)\mathfrak{m} \subset \mathfrak{m}.$$

We calculate the trace of ad h'_α on $\mathfrak{m}$ in two different ways. Following the reasoning of Proposition 4.3, we have

$$\operatorname{tr}_{\mathfrak{m}}\left(\operatorname{ad} h'_\alpha\right) = \sum_{i=-p}^{q} (i\alpha + \beta)\left(h'_\alpha\right)$$

since $\dim \mathfrak{g}_{i\alpha+\beta} = 1$ by Theorem 4.19. Second, we have

$$\operatorname{tr}_{\mathfrak{m}}\left(\operatorname{ad} h'_\alpha\right) = \operatorname{tr}_{\mathfrak{m}}(\operatorname{ad} e_\alpha \ \operatorname{ad} e_{-\alpha}) - \operatorname{tr}_{\mathfrak{m}}(\operatorname{ad} e_{-\alpha} \ \operatorname{ad} e_\alpha) = 0.$$

Thus

$$\sum_{i=-p}^{q} (i\alpha + \beta)\left(h'_\alpha\right) = 0,$$

that is,

$$\left(\frac{q(q+1)}{2} - \frac{p(p+1)}{2}\right)\alpha\left(h'_\alpha\right) + (p+q+1)\beta\left(h'_\alpha\right) = 0.$$

Since $p+q+1 \neq 0$, this yields

$$\frac{(q-p)}{2}\left\langle h'_\alpha, h'_\alpha\right\rangle + \left\langle h'_\alpha, h'_\beta\right\rangle = 0,$$

and so

$$2\frac{\left\langle h'_\alpha, h'_\beta\right\rangle}{\left\langle h'_\alpha, h'_\alpha\right\rangle} = p - q$$

since $\left\langle h'_\alpha, h'_\alpha\right\rangle \neq 0$ by Proposition 4.18. □

This proposition has some very useful corollaries. The first is a strengthening of Proposition 4.20.

Corollary 4.22. *If $\alpha \in \Phi$ and $\zeta\alpha \in \Phi$ where $\zeta \in \mathbb{C}$, then $\zeta = 1$ or $\zeta = -1$.*

Proof. Suppose $\zeta \neq \pm 1$ and let $\beta = \zeta\alpha$. Then $\beta(h'_\alpha) = \zeta\alpha(h'_\alpha)$, that is,

$$\langle h'_\alpha, h'_\beta \rangle = \zeta \langle h'_\alpha, h'_\alpha \rangle.$$

Applying Proposition 4.21, this yields

$$2\zeta = 2\frac{\langle h'_\alpha, h'_\beta \rangle}{\langle h'_\alpha, h'_\alpha \rangle} = p - q.$$

Hence $2\zeta \in \mathbb{Z}$. If $\zeta \in \mathbb{Z}$, then $\zeta = \pm 1$ by Proposition 4.20. Thus $\zeta \notin \mathbb{Z}$. It follows that $p - q$ is odd. The α-chain of roots through β is

$$-\left(\frac{p+q}{2}\right)\alpha, \ldots, \beta = \left(\frac{p-q}{2}\right)\alpha, \ldots, \left(\frac{p+q}{2}\right)\alpha.$$

Since $p - q$ is odd and consecutive roots differ by α, we see that all roots in the α-chain are odd multiples of $\frac{1}{2}\alpha$. Also, $p - q \neq 0$, and so p and q cannot both be zero. Thus $p + q \neq 0$. Because the first and last roots are negatives of one another, $\frac{1}{2}\alpha$ must lie in the α-chain. Thus $\frac{1}{2}\alpha \in \Phi$. But $\alpha \in \Phi$, and so $2\left(\frac{1}{2}\alpha\right) \in \Phi$, contradicting Proposition 4.20. It follows that ζ must be 1 or -1. □

Hence the only roots that are scalar multiples of a root α are α and $-\alpha$.

Corollary 4.23. *$\langle h'_\alpha, h'_\beta \rangle \in \mathbb{Q}$ for all $\alpha, \beta \in \Phi$.*

Proof. We already know that $\langle h'_\alpha, h'_\beta \rangle \in \mathbb{C}$. We also have

$$2\frac{\langle h'_\alpha, h'_\beta \rangle}{\langle h'_\alpha, h'_\alpha \rangle} \in \mathbb{Z} \qquad \text{by Proposition 4.21.}$$

Thus $\frac{\langle h'_\alpha, h'_\beta \rangle}{\langle h'_\alpha, h'_\alpha \rangle} \in \mathbb{Q}$. It is therefore sufficient to show that $\langle h'_\alpha, h'_\alpha \rangle \in \mathbb{Q}$. We have

$$\langle h'_\alpha, h'_\alpha \rangle = \operatorname{tr}\left(\operatorname{ad} h'_\alpha \operatorname{ad} h'_\alpha\right) = \sum_{\beta \in \Phi} \left(\beta\left(h'_\alpha\right)\right)^2 = \sum_{\beta \in \Phi} \langle h'_\alpha, h'_\beta \rangle^2.$$

Dividing by $\langle h'_\alpha, h'_\alpha \rangle^2$, this yields

$$\frac{1}{\langle h'_\alpha, h'_\alpha \rangle} = \sum_{\beta \in \Phi} \left(\frac{\langle h'_\alpha, h'_\beta \rangle}{\langle h'_\alpha, h'_\alpha \rangle}\right)^2 \in \mathbb{Z}.$$

Hence $\langle h'_\alpha, h'_\alpha \rangle \in \mathbb{Q}$, completing the proof. □

Chapter 5

Root systems and the Weyl group

5.1 Positive systems and fundamental systems of roots

As before, let $\mathfrak{g}$ be a semisimple Lie algebra, $\mathfrak{h}$ a Cartan subalgebra of $\mathfrak{g}$, and Φ the set of roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$. Recall from Proposition 4.16 that the elements h'_α for $\alpha \in \Phi$ span $\mathfrak{h}$. We can thus find a subset that forms a basis of $\mathfrak{h}$. Let $h'_{\alpha_1}, \dots, h'_{\alpha_l}$ be such a basis.

Proposition 5.1. *If $\alpha \in \Phi$, then $h'_\alpha = \sum_{i=1}^{l} \mu_i h'_{\alpha_i}$ where each μ_i lies in $\mathbb{Q}$.*

Proof. We know that $h'_\alpha = \sum_{i=1}^{l} \mu_i h'_{\alpha_i}$ for uniquely determined elements $\mu_i \in \mathbb{C}$. Let $\langle h'_{\alpha_i}, h'_{\alpha_j} \rangle = \xi_{ij}$. Then $\xi_{ij} \in \mathbb{Q}$ by Corollary 4.23. Consider the system of equations

$$\begin{aligned}
\langle h'_\alpha, h'_{\alpha_1} \rangle &= \mu_1 \xi_{11} + \mu_2 \xi_{21} + \cdots + \mu_l \xi_{l1} \\
\langle h'_\alpha, h'_{\alpha_2} \rangle &= \mu_1 \xi_{12} + \mu_2 \xi_{22} + \cdots + \mu_l \xi_{l2} \\
&\vdots \\
\langle h'_\alpha, h'_{\alpha_l} \rangle &= \mu_1 \xi_{1l} + \mu_2 \xi_{2l} + \cdots + \mu_l \xi_{ll}.
\end{aligned}$$

This is a system of l linear equations in l variables $\mu_1, \dots, \mu_l$. Since the Killing form of $\mathfrak{g}$ restricted to $\mathfrak{h}$ is nondegenerate, it follows from [HK71, Corollary 10.2] that $\det(\xi_{ij}) \neq 0$. Thus we may solve this system of equations for $\mu_1, \dots, \mu_l$ in terms of the $\langle h'_\alpha, h'_{\alpha_i} \rangle$ and ξ_{ij}. Because $\langle h'_\alpha, h'_{\alpha_i} \rangle \in \mathbb{Q}$ and $\xi_{ij} \in \mathbb{Q}$, it follows that $\mu_i \in \mathbb{Q}$ for $i = 1, \dots, l$. □

We define $\mathfrak{h}_\mathbb{Q}$ to be the set of all elements of the form $\sum_{i=1}^{l} \mu_i h'_{\alpha_i}$ with $\mu_i \in \mathbb{Q}$. Similarly, we define $\mathfrak{h}_\mathbb{R}$ to be the set of all such elements with $\mu_i \in \mathbb{R}$. Proposition 5.1

shows that $\mathfrak{h}_{\mathbb{Q}}$ and $\mathfrak{h}_{\mathbb{R}}$ are independent of the choice of basis $h'_{\alpha_1}, \ldots, h'_{\alpha_l}$. Since the elements h'_α for $\alpha \in \Phi$ span $\mathfrak{h}$, this also shows that $\mathfrak{h}_{\mathbb{Q}}$ is the set of all rational linear combinations of the h'_α, while $\mathfrak{h}_{\mathbb{R}}$ is the set of real linear combinations of such elements.

We now show that the Killing form of $\mathfrak{g}$ behaves favorably when restricted to $\mathfrak{h}_{\mathbb{R}}$.

Proposition 5.2. *Let $x \in \mathfrak{h}_{\mathbb{R}}$. Then $\langle x, x\rangle \in \mathbb{R}$ and $\langle x, x\rangle \geq 0$. If $\langle x, x\rangle = 0$, then $x = 0$.*

Proof. Let $x = \sum_{i=1}^{l} \mu_i h'_{\alpha_i}$ where $\mu_i \in \mathbb{R}$. Then

$$\begin{aligned}
\langle x, x\rangle &= \sum_{i=1}^{l}\sum_{j=1}^{l} \mu_i\mu_j \big\langle h'_{\alpha_i}, h'_{\alpha_j}\big\rangle \\
&= \sum_{i}\sum_{j} \mu_i\mu_j \operatorname{tr}\big(\operatorname{ad} h'_{\alpha_i}\ \operatorname{ad} h'_{\alpha_j}\big) \\
&= \sum_{i}\sum_{j} \mu_i\mu_j \sum_{\lambda\in\Phi} \lambda\left(h'_{\alpha_i}\right)\lambda\big(h'_{\alpha_j}\big) \\
&= \sum_{\lambda\in\Phi}\sum_{i}\sum_{j} \mu_i\mu_j \lambda\left(h'_{\alpha_i}\right)\lambda\big(h'_{\alpha_j}\big) \\
&= \sum_{\lambda\in\Phi}\left(\sum_{i} \mu_i \lambda\left(h'_{\alpha_i}\right)\right)^2 .
\end{aligned}$$

Now $\lambda\left(h'_{\alpha_i}\right) = \left\langle h'_\lambda, h'_{\alpha_i}\right\rangle \in \mathbb{Q}$ by Corollary 4.23. Thus $\langle x, x\rangle \in \mathbb{R}$, and also $\langle x, x\rangle \geq 0$. Suppose now that $\langle x, x\rangle = 0$. Then $\sum_i \mu_i \lambda\left(h'_{\alpha_i}\right) = 0$ for all $\lambda \in \Phi$. In particular, $\sum_i \mu_i \alpha_j\left(h'_{\alpha_i}\right) = 0$ for $j = 1, \ldots, l$. This yields $\sum_i \mu_i \big\langle h'_{\alpha_i}, h'_{\alpha_j}\big\rangle = 0$, that is, $\sum_i \mu_i \xi_{ij} = 0$. Since the matrix (ξ_{ij}) is non-singular, we must have $\mu_i = 0$ for all i. Hence $x = 0$. $\square$

This proposition shows that the Killing form of $\mathfrak{g}$ restricted to $\mathfrak{h}_{\mathbb{R}}$ is a map

$$\mathfrak{h}_{\mathbb{R}} \times \mathfrak{h}_{\mathbb{R}} \to \mathbb{R}$$

that is a symmetric positive definite bilinear form. The vector space $\mathfrak{h}_{\mathbb{R}}$ endowed with this positive definite form is thus a Euclidean space containing all vectors h'_α for $\alpha \in \Phi$.

Recall from Lemma 4.15 that the map $h \mapsto h^*$ given by $h^*(x) = \langle h, x\rangle$ is an isomorphism from $\mathfrak{h}$ into $\mathfrak{h}^*$. We define $\mathfrak{h}^*_{\mathbb{R}}$ to be the image of $\mathfrak{h}_{\mathbb{R}}$ under this isomorphism.

Thus $\mathfrak{h}_{\mathbb{R}}^*$ is the real subspace of $\mathfrak{h}^*$ spanned by Φ. We may also define a symmetric bilinear form on $\mathfrak{h}_{\mathbb{R}}^*$ by

$$\langle h_1^*, h_2^* \rangle = \langle h_1, h_2 \rangle \in \mathbb{R}.$$

Thus $\mathfrak{h}_{\mathbb{R}}^*$ becomes a Euclidean space containing all roots $\alpha \in \Phi$. We wish to investigate the structure of the roots in the Euclidean space $\mathfrak{h}_{\mathbb{R}}^*$. For convenience, we write $V = \mathfrak{h}_{\mathbb{R}}^*$.

A **total ordering** on V is a relation $\prec$ satisfying the following conditions:

(i) If $\lambda \prec \mu$ and $\mu \prec \nu$, then $\lambda \prec \nu$.

(ii) If $\lambda, \mu \in V$, then exactly one of the following holds: $\lambda \prec \mu$, $\lambda = \mu$, or $\mu \prec \lambda$.

(iii) If $\lambda \prec \mu$, then $\lambda + \nu \prec \mu + \nu$.

(iv) If $\lambda \prec \mu$ and $\xi \in \mathbb{R}$ with $\xi > 0$, then $\xi\lambda \prec \xi\mu$.

Every real vector space has such orderings. For example, let $v_1, \ldots, v_l$ be a basis of V and let $\lambda = \sum \lambda_i v_i$, $\mu = \sum \mu_i v_i$ with $\lambda \neq \mu$. We define $\lambda \prec \mu$ if the first nonzero coefficient $\mu_i - \lambda_i$ is positive. One easily checks that this is a total ordering on V.

A **positive system** $\Phi^+ \subset \Phi$ is the set of all roots $\alpha \in \Phi$ satisfying $0 \prec \alpha$ for some total ordering on V. Given a positive system Φ^+, we define a **fundamental system** $\Pi \subset \Phi^+$ by

$$\Pi = \{\alpha \in \Phi^+ \mid \alpha \text{ cannot be expressed as the sum of two elements of } \Phi^+\}.$$

We define Φ^- to be the corresponding set of negative roots.

Proposition 5.3. *Every root in Φ^+ is a sum of roots in Π.*

Proof. Let $\alpha \in \Phi^+$ and suppose $\alpha \notin \Pi$. Then $\alpha = \beta + \gamma$ where $\beta, \gamma \in \Phi^+$ and $\beta \prec \alpha, \gamma \prec \alpha$. Since Φ^+ is finite, we may continue decomposing the summands of α into lesser summands until each summand of α is in Π. □

Proposition 5.4. *If $\alpha, \beta \in \Pi$ and $\alpha \neq \beta$, then $\langle \alpha, \beta \rangle \leq 0$.*

Proof. We first show that $\alpha - \beta \notin \Phi$. Suppose $\alpha - \beta \in \Phi$. Then $\alpha - \beta \in \Phi^+$ or $\beta - \alpha \in \Phi^+$. If $\alpha - \beta \in \Phi^+$, then $\alpha = (\alpha - \beta) + \beta$, contradicting $\alpha \in \Pi$. Similarly, if $\beta - \alpha \in \Phi^+$, then $\beta = (\beta - \alpha) + \alpha$, contradicting $\beta \in \Pi$. It follows that $\alpha - \beta \notin \Phi$.

Now consider the α-chain of roots through β. This is of the form

$$\beta, \alpha + \beta, \dots, q\alpha + \beta$$

since $-\alpha + \beta \notin \Phi$. By Proposition 4.21, we deduce that

$$2\frac{\langle h'_\alpha, h'_\beta \rangle}{\langle h'_\alpha, h'_\alpha \rangle} = -q.$$

Now $\langle h'_\alpha, h'_\alpha \rangle > 0$ by Proposition 5.2, so we must have $\langle h'_\alpha, h'_\beta \rangle \leq 0$. It follows that $\langle \alpha, \beta \rangle \leq 0$ by the way we defined the bilinear form on V. □

Thus any two distinct roots in the fundamental system Π are inclined at an obtuse angle. The next theorem illustrates the importance of the concept of a fundamental system.

Theorem 5.5. *A fundamental system* Π *forms a basis of* $V = \mathfrak{h}^*_{\mathbb{R}}$.

Proof. We first show that Π spans V. We know that Φ spans V. Since $\alpha \in \Phi$ if and only if $-\alpha \in \Phi$, we see that Φ^+ spans V. Thus Π spans V by Proposition 5.3.

We now show that the set Π is linearly independent. Suppose it is not. Then there exists a non-trivial linear combination of roots $\alpha_i \in \Pi$ that is equal to zero. Taking the terms with positive coefficients to one side of the equation, we obtain

$$\mu_{i_1}\alpha_{i_1} + \cdots + \mu_{i_r}\alpha_{i_r} = \mu_{j_1}\alpha_{j_1} + \cdots + \mu_{j_r}\alpha_{j_r}$$

where $\mu_{i_1}, \dots, \mu_{i_r}, \mu_{j_1}, \dots, \mu_{j_s} > 0$ and $\alpha_{i_1}, \dots, \alpha_{i_r}, \alpha_{j_1}, \dots, \alpha_{j_s}$ are distinct elements in Π. Writing

$$v = \mu_{i_1}\alpha_{i_1} + \cdots + \mu_{i_r}\alpha_{i_r} = \mu_{j_1}\alpha_{j_1} + \cdots + \mu_{j_r}\alpha_{j_r},$$

we have $\langle v, v \rangle = \langle \mu_{i_1}\alpha_{i_1} + \cdots + \mu_{i_r}\alpha_{i_r}, \mu_{j_1}\alpha_{j_1} + \cdots + \mu_{j_r}\alpha_{j_r} \rangle$. Expanding this linearly, we see that $\langle v, v \rangle \leq 0$ by Proposition 5.4. Since the bilinear form is positive definite, this implies $v = 0$. But $0 \prec v$ since we have $0 \prec \alpha_i$ and $\mu_i > 0$. This is a contradiction, and so Π must be linearly independent. □

Now let Π be a fundamental system in Φ. By Theorem 5.5, we have $|\Pi| = \dim \mathfrak{h}_{\mathbb{R}}^*$. But $\mathfrak{h}_{\mathbb{R}}^*$ is isomorphic to $\mathfrak{h}_{\mathbb{R}}$, and $\dim \mathfrak{h}_{\mathbb{R}} = \dim \mathfrak{h}$. Thus $|\Pi| = l = \dim \mathfrak{h}$, that is, the number of roots in a fundamental system is equal to the rank of the Lie algebra $\mathfrak{g}$.

We may now write $\Pi = \{\alpha_1, \ldots, \alpha_l\}$. Since the roots $\alpha_1, \ldots, \alpha_l$ span $\mathfrak{h}_{\mathbb{R}}^*$, the corresponding vectors $h'_{\alpha_1}, \ldots, h'_{\alpha_l}$ must span $\mathfrak{h}_{\mathbb{R}}$. By our construction of $\mathfrak{h}_{\mathbb{R}}$, we also know that $\mathfrak{h}_{\mathbb{R}}$ spans $\mathfrak{h}$. Thus $h'_{\alpha_1}, \ldots, h'_{\alpha_l}$ are l vectors that span $\mathfrak{h}$, and hence form a basis of $\mathfrak{h}$. This fact will become useful as we enter Chapter 7.

Corollary 5.6. *Let Π be a fundamental system of roots. Then each $\alpha \in \Phi$ can be expressed in the form $\alpha = \sum n_i \alpha_i$ where $\alpha_i \in \Pi$, $n_i \in \mathbb{Z}$, and $n_i \geq 0$ for all i or $n_i \leq 0$ for all i.*

Proof. If $\alpha \in \Phi^+$, then α is a positive integer sum of elements of Π by Proposition 5.3. If $\alpha \in \Phi^-$, then α is the negative of an integer sum of elements of Π. □

5.2 The Weyl group

Inside the root system Φ, a positive system Φ^+ can be chosen in many different ways. In this section, we will show that any two positive systems in Φ can be transformed into one another by an element of a certain finite group W that acts on Φ.

For each $\alpha \in V$, we define a linear map $s_\alpha : V \to V$ by

$$s_\alpha(x) = x - 2\frac{\langle \alpha, x \rangle}{\langle \alpha, \alpha \rangle}\alpha \qquad \text{for all } x \in V$$

where, as before, $V = \mathfrak{h}_{\mathbb{R}}^*$. One sees immediately that s_α satisfies

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

$$s_\alpha(x) = x \qquad \text{if } \langle \alpha, x \rangle = 0.$$

There is a unique linear map satisfying these conditions – the reflection in the hyperplane of V orthogonal to α. Thus s_α is this reflection.

The group W of all non-singular maps on V generated by the s_α for all $\alpha \in \Phi$ is called the **Weyl group**. This group plays a vital role in the Lie theory.

Proposition 5.7. *The Weyl group is a group of isometries on V, that is,*

$$\langle w(x), w(y)\rangle = \langle x, y\rangle \qquad \text{for all } x, y \in V,\ w \in W.$$

Proof. Let $\alpha \in \Phi$. We have

$$\begin{aligned}
\langle s_\alpha(x), s_\alpha(y)\rangle &= \left\langle x - 2\frac{\langle \alpha, x\rangle}{\langle \alpha, \alpha\rangle}\alpha, y - 2\frac{\langle \alpha, y\rangle}{\langle \alpha, \alpha\rangle}\alpha \right\rangle \\
&= \langle x, y\rangle - \left\langle x, 2\frac{\langle \alpha, y\rangle}{\langle \alpha, \alpha\rangle}\alpha \right\rangle - \left\langle 2\frac{\langle \alpha, x\rangle}{\langle \alpha, \alpha\rangle}\alpha, y \right\rangle + \left\langle 2\frac{\langle \alpha, x\rangle}{\langle \alpha, \alpha\rangle}\alpha, 2\frac{\langle \alpha, y\rangle}{\langle \alpha, \alpha\rangle}\alpha \right\rangle \\
&= \langle x, y\rangle - 2\frac{\langle \alpha, y\rangle}{\langle \alpha, \alpha\rangle}\langle x, \alpha\rangle - 2\frac{\langle \alpha, x\rangle}{\langle \alpha, \alpha\rangle}\langle \alpha, y\rangle + 4\frac{\langle \alpha, x\rangle\langle \alpha, y\rangle}{\langle \alpha, \alpha\rangle\langle \alpha, \alpha\rangle}\langle \alpha, \alpha\rangle \\
&= \langle x, y\rangle.
\end{aligned}$$

Thus s_α is an isometry for all $\alpha \in \Phi$. The proof follows immediately since W is generated by the elements s_α. ∎

Proposition 5.8. *If $\alpha \in \Phi$ and $w \in W$, then $w(\alpha) \in \Phi$, i.e., W permutes the roots.*

Proof. It suffices to show that $s_\alpha(\beta) \in \Phi$ for all $\alpha, \beta \in \Phi$ since the elements s_α generate W. The statement is clear if β is either α or $-\alpha$. So suppose $\beta \neq \pm\alpha$ and let

$$-p\alpha + \beta, \ldots, \beta, \ldots, q\alpha + \beta$$

be the α-chain of roots through β. Then we have

$$s_\alpha(\beta) = \beta - 2\frac{\langle \alpha, \beta\rangle}{\langle \alpha, \alpha\rangle}\alpha = \beta - (p - q)\alpha$$

by Proposition 4.21. Now $\beta - (p - q)\alpha$ lies in the α-chain, and so $s_\alpha(\beta) \in \Phi$.

In fact we observe that s_α inverts the above α-chain. In particular, we have $s_\alpha(q\alpha + \beta) = -p\alpha + \beta$ and $s_\alpha(-p\alpha + \beta) = q\alpha + \beta$. ∎

Proposition 5.9. *The Weyl group W is finite.*

Proof. Each element of W induces a permutation of Φ. If two elements of W induce the same permutation of Φ, then they must be equal since Φ spans V. Thus there is a one-to-one correspondence between the elements of W and the permutations of Φ. Since Φ is finite, there can be only finitely-many permutations of Φ. Hence W is finite. $\square$

Now suppose Φ^+ is a positive system in Φ and let Π be the corresponding fundamental system.

Lemma 5.10. *Let $\alpha \in \Pi$. If $\beta \in \Phi^+$ and $\beta \neq \alpha$, then $s_\alpha(\beta) \in \Phi^+$.*

Proof. We can express β in the form

$$\beta = \sum_i n_i \alpha_i \qquad \alpha_i \in \Pi, \qquad n_i \in \mathbb{Z}, \qquad n_i \geq 0$$

by Corollary 5.6. Since $\beta \neq \alpha$, there must be some $n_i \neq 0$ with $\alpha_i \neq \alpha$. We have

$$s_\alpha(\beta) = \beta - 2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}\alpha.$$

We express $s_\alpha(\beta)$ as a linear combination of the elements of Π. Since $\beta \neq 0$, we have $s_\alpha(\beta) \in \Phi^+$ or $s_\alpha(\beta) \in \Phi^-$. It follows from Corollary 5.6 that if $s_\alpha(\beta) \in \Phi^+$, then the integer coefficients of the elements of Π will all be nonnegative, whereas if $s_\alpha(\beta) \in \Phi^-$, then the integer coefficients of the elements of Π will all be nonpositive. Since the coefficient of α_i remains n_i, we must have $s_\alpha(\beta) \in \Phi^+$. $\square$

Theorem 5.11. *If Φ_1^+, Φ_2^+ are two positive systems in Φ, then there exists a $w \in W$ such that $w\left(\Phi_1^+\right) = \Phi_2^+$.*

Proof. Let $m = \left|\Phi_1^+ \cap \Phi_2^-\right|$. We proceed by induction on m. If $m = 0$, then $\Phi_1^+ = \Phi_2^+$, and so $w = 1$ satisfies the theorem. Now assume $m > 0$, that is, $\Phi_1^+ \cap \Phi_2^- \neq \varnothing$. We cannot have $\Pi_1 \subset \Phi_2^+$, as this would imply $\Phi_1^+ \subset \Phi_2^+$, contrary to $m > 0$. Thus there exists an $\alpha \in \Pi_1 \cap \Phi_2^-$. Consider the positive system $s_\alpha\left(\Phi_1^+\right)$ in Φ. By Lemma 5.10, $s_\alpha\left(\Phi_1^+\right)$ contains all roots in Φ_1^+ except α, together with $-\alpha$. It follows that

$$\left|s_\alpha\left(\Phi_1^+\right) \cap \Phi_2^-\right| = m - 1.$$

By induction, there exists a $w' \in W$ such that $w's_\alpha\left(\Phi_1^+\right) = \Phi_2^+$. Letting $w = w's_\alpha$, we have $w\left(\Phi_1^+\right) = \Phi_2^+$ as desired. □

Corollary 5.12. *If Π_1, Π_2 are two fundamental systems in Φ, then there exists a $w \in W$ such that $w(\Pi_1) = \Pi_2$.*

Proof. Let Φ_1^+, Φ_2^+ be positive systems containing Π_1, Π_2, respectively. By Theorem 5.11, there exists a $w \in W$ such that $w\left(\Phi_1^+\right) = \Phi_2^+$. Then $w(\Pi_1)$ is a fundamental system contained in Φ_2^+, and so $w(\Pi_1) = \Pi_2$. □

Proposition 5.13. *Let Π be a fundamental system in Φ. Then for each $\alpha \in \Phi$, there exist $\alpha_i \in \Pi$ and $w \in W$ such that $\alpha = w(\alpha_i)$.*

Proof. Let Φ^+ be the positive system with fundamental system Π. Suppose first that $\alpha \in \Phi^+$. Then we have

$$\alpha = \sum_i n_i \alpha_i \qquad \alpha_i \in \Pi, \qquad n_i \in \mathbb{Z}, \qquad n_i \geq 0$$

by Corollary 5.6. We define the **height** of α by

$$\operatorname{ht} \alpha = \sum_i n_i.$$

We proceed by induction on $\operatorname{ht} \alpha$. If $\operatorname{ht} \alpha = 1$, then $\alpha = \alpha_i$ for some i, and so $\alpha \in \Pi$. Thus $\alpha_i = \alpha$ and $w = 1$ satisfy the proposition. Now suppose $\operatorname{ht} \alpha > 1$. Then $n_i > 0$ for at least two values of i by Proposition 4.20. We have

$$\langle \alpha, \alpha \rangle = \sum_i n_i \langle \alpha, \alpha_i \rangle.$$

Since $\langle \alpha, \alpha \rangle > 0$ and each $n_i > 0$, there exists an $\alpha_i \in \Pi$ such that $\langle \alpha, \alpha_i \rangle > 0$. Let $\beta = s_{\alpha_i}(\alpha)$. Then $\beta \in \Phi$ and

$$\beta = \alpha - 2\frac{\langle \alpha, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle}\alpha_i.$$

Since $\langle \alpha, \alpha_i \rangle > 0$, we see that the coefficient of α_i in β is less than the coefficient of α_i in α. Thus $\operatorname{ht} \beta < \operatorname{ht} \alpha$. On the other hand, we know that $\beta \in \Phi^+$ since there are at least two positive coefficients in α and only one coefficient is reduced in passing from α to β. By induction, there exist $\alpha_j \in \Pi$ and $w' \in W$ such that $\beta = w'(\alpha_j)$. This yields

$$\alpha = s_{\alpha_i}(\beta) = s_{\alpha_i} w'(\alpha_j)$$

as desired.

Now suppose $\alpha \in \Phi^-$. Then $\alpha = s_\alpha(-\alpha)$ and $-\alpha \in \Phi^+$. Thus there exist $w' \in W$ and $\alpha_i \in \Pi$ such that $-\alpha = w'(\alpha_i)$. Hence $\alpha = s_\alpha w'(\alpha_i)$, completing the proof. □

Theorem 5.14. *Let $\Pi = \{\alpha_1, \dots, \alpha_l\}$ be a fundamental system in Φ. Then the reflections $s_{\alpha_1}, \dots, s_{\alpha_l}$ generate W.*

Proof. Let W_0 be the subgroup of W generated by $s_{\alpha_1}, \dots, s_{\alpha_l}$. Since the elements s_α for $\alpha \in \Phi$ generate W, it is sufficient to prove that each s_α lies in W_0. We may assume $\alpha \in \Phi^+$ since $s_\alpha = s_{-\alpha}$. Applying Proposition 5.13 to the subgroup W_0 (each time we used the inductive hypothesis, we could have assumed that w was in W_0), we see that there exist $\alpha_i \in \Pi$ and $w \in W_0$ such that $\alpha = w'(\alpha_i)$. Consider the element $w s_{\alpha_i} w^{-1} \in W_0$. We have

$$w s_{\alpha_i} w^{-1}(\alpha) = w s_{\alpha_i}(\alpha_i) = w(-\alpha_i) = -\alpha.$$

We shall show that $w s_{\alpha_i} w^{-1}(x) = x$ if $\langle \alpha, x \rangle = 0$. Now $\langle \alpha, x \rangle = 0$ implies

$$\left\langle w^{-1}(\alpha), w^{-1}(x) \right\rangle = 0$$

since w^{-1} is an isometry. Hence $\left\langle \alpha_i, w^{-1}(x) \right\rangle = 0$. It follows that $w^{-1}(x)$ is orthogonal to α_i, and so $s_{\alpha_i} w^{-1}(x) = w^{-1}(x)$, that is, $w s_{\alpha_i} w^{-1}(x) = x$. Thus $w s_{\alpha_i} w^{-1}$ is the reflection in the hyperplane orthogonal to α, and so $ws_{\alpha_i}w^{-1} = s_\alpha$. This shows that $s_\alpha \in W_0$, and thus $W_0 = W$. □

We wish to further study the way in which the Weyl group is generated by the reflections of the fundamental roots. As before, let $\Pi = \{\alpha_1, \ldots, \alpha_l\}$ be a fundamental system of roots. For convenience, we write

$$s_1 = s_{\alpha_1}, \quad s_2 = s_{\alpha_2}, \quad \ldots, \quad s_l = s_{\alpha_l}.$$

Each element of W can be expressed as a product of the s_i. (We do not need to worry about inverses since $s_i^{-1} = s_i$ for all i.) For each $w \in W$, we define $l(w)$ to be the least nonnegative integer m such that w can be expressed as a product of m fundamental reflections. We call $l(w)$ the **length** of w. We have $l(1) = 0$ and $l(s_i) = 1$ for all i. An expression of w as a product of $l(w)$ fundamental reflections is called a **reduced expression** for w.

Recall that each element of W permutes the roots in Φ by Proposition 5.8. For each $w \in W$, we define $n(w)$ to be the number of roots $\alpha \in \Phi^+$ such that $w(\alpha) \in \Phi^-$. Thus $n(w)$ is the number of positive roots made negative by w. We wish to show that $l(w) = n(w)$.

Proposition 5.15. *$n(w) \leq l(w)$ for all $w \in W$.*

Proof. We fix i and compare $n(w)$ with $n(ws_i)$. We know that s_i transforms α_i to $-\alpha_i$ and all other positive roots to positive roots by Lemma 5.10. Hence

$$n(ws_i) = n(w) \pm 1.$$

In order to determine the sign, we study the effect of w and ws_i on α_i. If $w(\alpha_i) \in \Phi^+$, then w transforms α_i into a positive root, and so ws_i transforms α_i into a negative root. Hence $n(ws_i) = n(w) + 1$. If, on the other hand, $w(\alpha_i) \in \Phi^-$, then ws_i transforms α_i into a positive root so that $n(ws_i) = n(w) - 1$.

Now let $w = s_{i_1} s_{i_2} \cdots s_{i_r}$ be a reduced expression for w where $r = l(w)$. Then

$$n(w) \leq n(s_{i_1} \cdots s_{i_{r-1}}) + 1 \leq n(s_{i_1} \cdots s_{i_{r-2}}) + 1 \leq \cdots \leq r.$$

Thus $n(w) \leq l(w)$ as desired. □

In order to prove the converse, that $l(w) \leq n(w)$, we need a very useful theorem called the **deletion condition**.

Theorem 5.16. *Let $w = s_{i_1} \cdots s_{i_r}$ be any expression of $w \in W$ as a product of the fundamental reflections. If $n(w) < r$, then there exist integers j, k with $1 \leq j < k \leq r$ such that*

$$w = s_{i_1} \cdots \hat{s}_{i_j} \cdots \hat{s}_{i_k} \cdots s_{i_r}$$

where $\hat{}$ denotes omission.

Proof. Recall from Proposition 5.15 that for all $w \in W$, we have $n(ws_i) = n(w) \pm 1$. Consider the given expression $w = s_{i_1} \cdots s_{i_r}$. Looking at the chain

$$n(w) \leq n(s_{i_1} \cdots s_{i_{r-1}}) + 1 \leq n(s_{i_1} \cdots s_{i_{r-2}}) + 1 \leq \cdots \leq r,$$

we see that since $n(w) < r$, we must have

$$n(s_{i_1} \cdots s_{i_k}) = n(s_{i_1} \cdots s_{i_{k-1}}) - 1$$

for some k with $1 < k \leq r$. This implies $s_{i_1} \cdots s_{i_{k-1}}(\alpha_{i_k}) \in \Phi^-$ as in the proof of Proposition 5.15. Since $\alpha_{i_k} \in \Phi^+$, there exists a j with $1 \leq j < k$ such that

$$s_{i_{j+1}} \cdots s_{i_{k-1}}(\alpha_{i_k}) \in \Phi^+$$

$$s_{i_j} s_{i_{j+1}} \cdots s_{i_{k-1}}(\alpha_{i_k}) \in \Phi^-.$$

Now s_j transforms only one positive root into a negative root by Lemma 5.10, namely α_{i_j}. It follows that

$$s_{i_{j+1}} \cdots s_{i_{k-1}}(\alpha_{i_k}) = \alpha_{i_j}. \tag{1}$$

Applying $s_{i_k}s_{i_{k-1}}\cdots s_{i_{j+1}}$ to both sides of (1) yields

$$-\alpha_{i_k} = s_{i_k}s_{i_{k-1}}\cdots s_{i_{j+1}}(\alpha_{i_j}). \tag{2}$$

But (1) also gives us the relation

$$s_{i_{j+1}}\cdots s_{i_{k-1}}(-\alpha_{i_k}) = -\alpha_{i_j}. \tag{3}$$

Plugging (2) into (3), we get

$$s_{i_{j+1}}\cdots s_{i_{k-1}}s_{i_k}s_{i_{k-1}}\cdots s_{i_{j+1}}(\alpha_{i_j}) = -\alpha_{i_j}.$$

Since $s_{i_{j+1}}\cdots s_{i_{k-1}}s_{i_k}s_{i_{k-1}}\cdots s_{i_{j+1}}$ transforms α_{i_j} into $-\alpha_{i_j}$, we must have

$$s_{i_j} = s_{i_{j+1}}\cdots s_{i_{k-1}}s_{i_k}s_{i_{k-1}}\cdots s_{i_{j+1}}.$$

It follows that

$$s_{i_j}s_{i_{j+1}}\cdots s_{i_{k-1}} = s_{i_{j+1}}\cdots s_{i_{k-1}}s_{i_k}.$$

Plugging this relation into our original expression for w yields

$$s_{i_1}\cdots s_{i_r} = s_{i_1}\cdots s_{i_{j-1}}s_{i_{j+1}}\cdots s_{i_{k-1}}s_{i_{k+1}}\cdots s_{i_r},$$

and so $w = s_{i_1}\cdots \hat{s}_{i_j}\cdots \hat{s}_{i_k}\cdots s_{i_r}$ as desired. □

Corollary 5.17. $n(w) = l(w)$

Proof. We know that $n(w) \leq l(w)$ by Proposition 5.15. So suppose $n(w) < l(w)$. Let $w = s_{i_i}\cdots s_{i_r}$ be a reduced expression for w, that is, $r = l(w)$. Since $n(w) < r$, we may use the deletion condition and obtain an expression for w that is a product of $r-2$ fundamental reflections. This contradicts the definition of $l(w)$, so we must have $n(w) = l(w)$. □

This shows that the length of w is equal to the number of positive roots made negative by w.

Proposition 5.18.

(i) *The maximal length of any element of* W *is* $|\Phi^+|$.

(ii) *There is a unique element* $w_0 \in W$ *with* $l(w_0) = |\Phi^+|$.

(iii) $w_0\left(\Phi^+\right) = \Phi^-$.

(iv) $w_0^2 = 1$.

Proof. Part (i) follows from the fact that $l(w) = n(w)$ and $n(w)$ is bounded by $|\Phi^+|$. To prove part (ii), notice that for each fundamental system Π, $-\Pi$ is also a fundamental system, arising from the opposite total ordering on $\mathfrak{h}_{\mathbb{R}}^*$. By Corollary 5.12, there exists a $w_0 \in W$ such that $w_0(\Pi) = -\Pi$. Hence $w_0\left(\Phi^+\right) = \Phi^-$, and so $n(w_0) = |\Phi^+|$. Thus $l(w) = |\Phi^+|$, that is, w_0 is an element of W of maximal length. Let w_0' be another element of W such that $l\left(w_0'\right) = |\Phi^+|$. Then $n\left(w_0'\right) = |\Phi^+|$, that is, $w_0'\left(\Phi^+\right) = \Phi^-$. Let $w = w_0'^{-1} w_0$. Then $w\left(\Phi^+\right) = \Phi^+$ so that $n(w) = 0$. It follows that $l(w) = 0$, and so $w = 1$. Thus $w_0' = w_0$, showing that w_0 is unique. This proves part (ii), and part (iii) follows immediately. Finally, we have $w_0^2\left(\Phi^+\right) = w_0\left(\Phi^-\right) = \Phi^+$, and so $n\left(w_0^2\right) = 0$. Hence $l\left(w_0^2\right) = 0$ and $w_0^2 = 1$. □

Now consider a group having presentation

$$\left\langle s_1, \ldots, s_l \mid (s_i s_j)^{m_{ij}} = 1 \right\rangle$$

where $m_{ii} = 1$ for all i and $m_{ij} \geq 2$ for $i \neq j$. A group having such a presentation is called a **Coxeter group**. One can show that every Weyl group is isomorphic to a finite Coxeter group. For a proof, we refer the reader to [Car05, §5.3].

Chapter 6

The Cartan matrix and the Dynkin diagram

6.1 The Cartan matrix

We recall from Proposition 5.2 that the vector space $V = \mathfrak{h}_{\mathbb{R}}^*$ forms a Euclidean space under the scalar product $\langle , \rangle$. We also recall that the roots Φ span V but are not linearly independent. Any fundamental system $\Pi \subset \Phi$ does, however, form a basis of V by Theorem 5.5. The goal of this section is to further study the geometry of the root system Φ in the Euclidean space V.

We shall first find the possible angles between any two roots $\alpha, \beta \in \Phi$ and the relation between the lengths of the roots α, β. We consider only the angles θ satisfying $0 \leq \theta \leq \pi$.

Proposition 6.1. *Let $\alpha, \beta \in \Phi$ such that $\beta \neq \pm\alpha$. Then*

(i) *The angle between α, β is one of $\pi/6$, $\pi/4$, $\pi/3$, $\pi/2$, $2\pi/3$, $3\pi/4$, $5\pi/6$.*

(ii) *If α, β are inclined at $\pi/3$ or $2\pi/3$, then α, β have the same length.*

(iii) *If α, β are inclined at $\pi/4$ or $3\pi/4$, then the ratio of their lengths is $\sqrt{2}$.*

(iv) *If α, β are inclined at $\pi/6$ or $5\pi/6$, then the ratio of their lengths is $\sqrt{3}$.*

Proof. Let θ be the angle between α, β. Then

$$\langle \alpha, \beta \rangle = |\alpha||\beta| \cos\theta$$

where, as usual, $|\alpha| = \sqrt{\langle \alpha, \alpha \rangle}$ and $|\beta| = \sqrt{\langle \beta, \beta \rangle}$. Thus

$$\cos^2 \theta = \frac{\langle \alpha, \beta \rangle^2}{\langle \alpha, \alpha \rangle \langle \beta, \beta \rangle} = \frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} \cdot \frac{\langle \beta, \alpha \rangle}{\langle \beta, \beta \rangle}$$

so that

$$4 \cos^2 \theta = 2 \frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} \cdot 2 \frac{\langle \beta, \alpha \rangle}{\langle \beta, \beta \rangle}.$$

Recall from Proposition 4.21 that $2\frac{\langle \alpha,\beta \rangle}{\langle \alpha,\alpha \rangle}$ and $2\frac{\langle \beta,\alpha \rangle}{\langle \beta,\beta \rangle}$ are integers. Hence $4\cos^2\theta \in \mathbb{Z}$. Since $0 \leq 4\cos^2\theta \leq 4$ and $\beta \neq \pm\alpha$, we must have $4\cos^2\theta \in \{0, 1, 2, 3\}$. In each case, we consider the possible factorization of $4\cos^2\theta$ into the product of two integers. Without loss of generality, we assume that $|\alpha| \leq |\beta|$.

First, suppose $4\cos^2\theta = 0$. Then $\cos\theta = 0$ so that $\theta = \pi/2$.

Second, suppose $4\cos^2\theta = 1$. Then $\cos\theta = 1/2$ or $-1/2$ so that $\theta = \pi/3$ or $2\pi/3$. The possible factorizations of $4\cos^2\theta$ are

$$1 = 1 \cdot 1 \quad \text{or} \quad 1 = (-1)(-1).$$

In either case, we have

$$2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} = 2\frac{\langle \beta, \alpha \rangle}{\langle \beta, \beta \rangle}$$

so that $\langle \alpha, \alpha \rangle = \langle \beta, \beta \rangle$, that is, α and β have the same length.

Now suppose $4\cos^2\theta = 2$. Then $\cos\theta = 1/\sqrt{2}$ or $-1/\sqrt{2}$ so that $\theta = \pi/4$ or $3\pi/4$. The possible factorizations of $4\cos^2\theta$ are

$$2 = 1 \cdot 2 \quad \text{or} \quad 2 = (-1)(-2).$$

In either case, since $|\alpha| \leq |\beta|$, we must have

$$2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} = 2 \cdot 2\frac{\langle \beta, \alpha \rangle}{\langle \beta, \beta \rangle}.$$

Then $\langle \beta, \beta \rangle = 2\langle \alpha, \alpha \rangle$ and $|\beta| = \sqrt{2}|\alpha|$. Hence the ratio of the lengths of α, β is $\sqrt{2}$.

Finally, suppose $4\cos^2\theta = 3$. Then $\cos\theta = \sqrt{3}/2$ or $-\sqrt{3}/2$ so that $\theta = \pi/6$ or $5\pi/6$. The possible factorizations of $4\cos^2\theta$ are

$$3 = 1 \cdot 3 \quad \text{or} \quad 3 = (-1)(-3).$$

In either case, since $|\alpha| \leq |\beta|$, we must have

$$2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} = 3 \cdot 2\frac{\langle \beta, \alpha \rangle}{\langle \beta, \beta \rangle}.$$

Then $\langle \beta, \beta \rangle = 3\langle \alpha, \alpha \rangle$ and $|\beta| = \sqrt{3}|\alpha|$. Hence the ratio of the lengths of α, β is $\sqrt{3}$.

This completes the proof. Notice that we obtain no information about the ratio of the lengths of α, β in the case where $\theta = \pi/2$. □

Corollary 6.2. *Let* Π *be a fundamental system of roots in* Φ *and let* $\alpha, \beta \in \Pi$ *with* $\beta \neq \alpha$. *Then the angle between* α, β *is one of* $\pi/2$, $2\pi/3$, $3\pi/4$, $5\pi/6$.

Proof. We saw in Proposition 5.4 that the angle θ between two distinct fundamental roots satisfies $\pi/2 \leq \theta < \pi$. Looking at the list of possible angles in Proposition 6.1, $\pi/2$, $2\pi/3$, $3\pi/4$, $5\pi/6$ are the only angles satisfying this property. □

Now let $\Pi = \{\alpha_1, \ldots, \alpha_l\}$ be a fundamental system. We wish to encode the information about the angles between the α_i and the ratios of their lengths into a matrix. For $i, j = 1, \ldots, l$, we define the elements

$$A_{ij} = 2\frac{\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_i, \alpha_i \rangle}.$$

We have $A_{ij} \in \mathbb{Z}$ for all i, j. The $l \times l$ matrix $A = (A_{ij})$ is called the **Cartan matrix**.

Proposition 6.3. *The Cartan matrix* A *has the following properties:*

(i) $A_{ii} = 2$ *for all* i.

(ii) $A_{ij} \in \{0, -1, -2, -3\}$ *if* $i \neq j$.

(iii) *If* $A_{ij} = -2$ *or* -3, *then* $A_{ji} = -1$.

(iv) $A_{ij} = 0$ *if and only if* $A_{ji} = 0$.

Proof. Properties (i) and (iv) follow from the definition of A_{ij}. Properties (ii) and (iii) follow from the proof of Proposition 6.1. □

If we permute the indices of the roots in Π, we will not, in general, get the same Cartan matrix A. But aside from the ambiguity of indexing the roots, the Cartan matrix is uniquely determined by the semisimple Lie algebra $\mathfrak{g}$.

Proposition 6.4. *The Cartan matrix of* $\mathfrak{g}$ *depends only on the indexing of the fundamental roots. It is independent of the choice of Cartan subalgebra* $\mathfrak{h}$ *and fundamental system* Π.

Proof. The independence of the choice of Cartan subalgebra follows from the conjugacy of Cartan subalgebras, which was Theorem 3.14.

Now let Π' be another fundamental system. By Corollary 5.12, there exists a $w \in W$ such that $w(\Pi) = \Pi'$. For each i, we write $w(\alpha_i) = \alpha_i'$. Because w is an isometry of V, we have

$$2\frac{\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_i, \alpha_i \rangle} = 2\frac{\langle \alpha_i', \alpha_j' \rangle}{\langle \alpha_i', \alpha_i' \rangle}.$$

It follows that the Cartan matrices defined by Π and Π' with respect to the chosen labellings are the same. □

We now determine the Cartan matrices for $l = 1$ and 2 using Proposition 6.3. The only possible 1×1 Cartan matrix is (2). We also see that any 2×2 Cartan matrix must be one of the following:

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \quad \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad \begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix} \quad \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix} \quad \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \quad \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}.$$

The pair

$$\begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix} \qquad \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}$$

are obtained by switching the indices 1 and 2, as are the pair

$$\begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \qquad \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}.$$

6.2 The Dynkin diagram

In order to determine the possible Cartan matrices for larger values of l, it is useful to introduce a combinatorial structure called the **Dynkin diagram**. The Dynkin diagram is determined by the Cartan matrix. It is a graph with vertices labeled $1, \ldots, l$. For $i \neq j$, the vertices i, j are joined by n_{ij} edges where

$$n_{ij} = A_{ij} A_{ji}.$$

By Proposition 6.4, the Dynkin diagram is uniquely determined by the semisimple Lie algebra $\mathfrak{g}$. The Dynkin diagrams of the Cartan matrices for $l = 1$ and 2 are given in Figure 6.1.

Cartan matrix

$$(2)$$

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}$$

Dynkin diagram

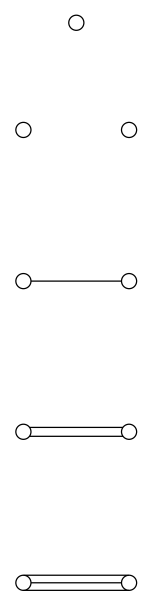


Figure 6.1: Cartan matrices and Dynkin diagrams for $l = 1$ and 2.

Proposition 6.5. $n_{ij} \in \{0, 1, 2, 3\}$ *for all* $i \neq j$.

Proof. This follows immediately from Proposition 6.3 since $n_{ij} = A_{ij} A_{ji}$. $\square$

This shows that the number of edges joining any two vertices of the Dynkin diagram is either 0, 1, 2, or 3. Notice that the Dynkin diagram need not be a connected graph. If it is connected, however, then it will split into connected components. If we renumber the vertices so that those in each connected component are numbered consecutively, then the Cartan matrix will split into blocks of the form

$$A = \left(\begin{array}{c|c|c|c} * & 0 & 0 & 0 \\ \hline 0 & * & 0 & 0 \\ \hline 0 & 0 & * & 0 \\ \hline 0 & 0 & 0 & * \end{array}\right)$$

where each block corresponds to a connected component. Each block will, in fact, be the Cartan matrix for its corresponding connected component. The set $\Pi = \{\alpha_1, \ldots, \alpha_l\}$ will be partitioned into subsets in a corresponding way. One easily checks that roots in different subsets are mutually orthogonal.

At first glance, it may appear that there are many ways to arrange vertices and edges to form connected components. But the set of graphs that can occur as connected Dynkin diagrams turns out to be quite restricted. In order to reduce our list of possibilities, it is useful to introduce a quadratic form $\mathrm{Q}(x_1, \ldots, x_l)$ defined in terms of the Dynkin diagram. We define

$$\mathrm{Q}(x_1, \ldots, x_l) = 2\sum_{i=1}^{l} x_i^2 - \sum_{\substack{i,j=1 \\ i\neq j}}^{l} \sqrt{n_{ij}}\, x_i x_j.$$

Figure 6.2 gives the quadratic forms of the Dynkin diagrams for $l = 1$ and 2.

Proposition 6.6. *The quadratic form* $\mathrm{Q}(x_1, \ldots, x_l)$ *is positive definite.*

Proof. For $i \neq j$, we have

$$n_{ij} = A_{ij}A_{ji} = 2\frac{\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_i, \alpha_i \rangle} \cdot 2\frac{\langle \alpha_j, \alpha_i \rangle}{\langle \alpha_j, \alpha_j \rangle}.$$

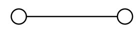
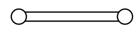
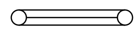
Dynkin diagram	**Quadratic form**
	$2x_1^2$
	$2x_1^2 + 2x_2^2$
	$2x_1^2 + 2x_1x_2 + 2x_2^2$
	$2x_1^2 + 2\sqrt{2}\,x_1x_1 + 2x_2^2$
	$2x_1^2 + 2\sqrt{3}\,x_1x_1 + 2x_2^2$

Figure 6.2: Dynkin diagrams and quadratic forms for $l = 1$ and 2.

By Proposition 5.4, we know that $\langle \alpha_i, \alpha_j \rangle \leq 0$. Hence $-\sqrt{n_{ij}} = 2\frac{\langle \alpha_i, \alpha_j \rangle}{|\alpha_i||\alpha_j|}$. For $i = j$, we have $2\frac{\langle \alpha_i, \alpha_j \rangle}{|\alpha_i||\alpha_j|} = 2$. Thus the quadratic form is

$$\mathrm{Q}(x_1, \ldots, x_l) = \sum_{i,j=1}^{l} 2\frac{\langle \alpha_i, \alpha_j \rangle}{|\alpha_i||\alpha_j|} x_i x_j = 2\left\langle \sum_{i=1}^{l} \frac{x_i \alpha_i}{|\alpha_i|}, \sum_{j=1}^{l} \frac{x_j \alpha_j}{|\alpha_j|} \right\rangle = 2\langle y, y \rangle$$

where $y = \sum_{i=1}^{l} \frac{x_i \alpha_i}{|\alpha_i|}$. Thus $\mathrm{Q}(x_1, \ldots, x_l) \geq 0$ since the bilinear form $\langle , \rangle$ is positive definite. Furthermore, if $\mathrm{Q}(x_1, \ldots, x_l) = 0$, then $y = 0$. In this case, since $\alpha_1, \ldots, \alpha_l$ are linearly independent, we must have $x_i = 0$ for all i. Thus the quadratic form is positive definite. □

Thus we have shown that the the connected components of the Dynkin diagram of a semisimple Lie algebra have the following properties:

(A) The graph is connected.

(B) Any pair of distinct vertices are joined by 0, 1, 2, or 3 edges.

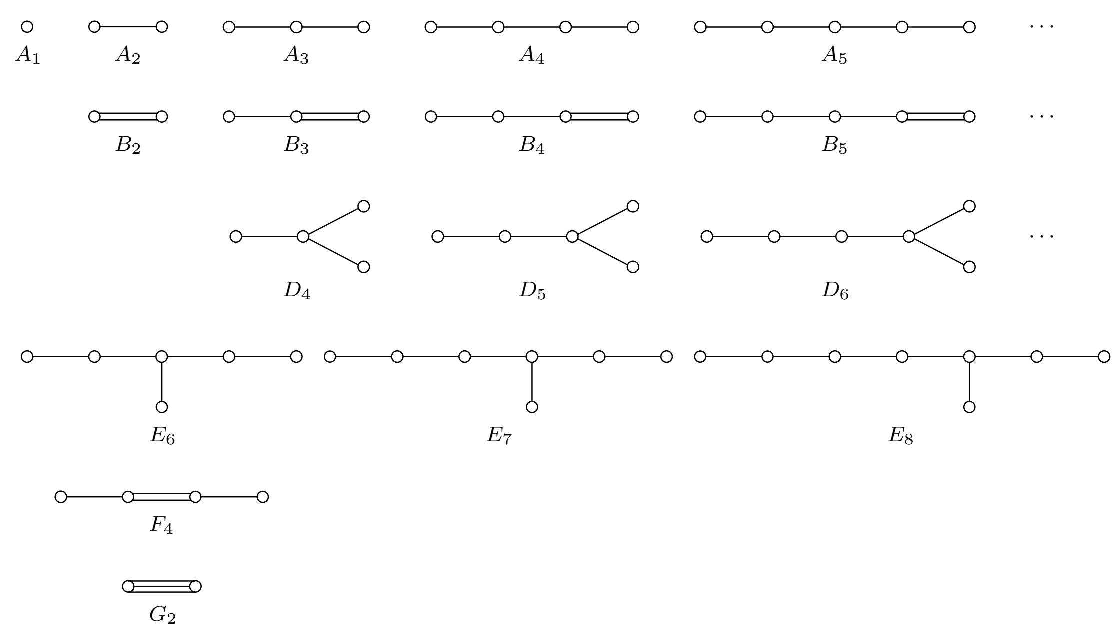
(C) The corresponding quadratic form $\mathrm{Q}(x_1, \ldots, x_l)$ is positive definite.

In order to find all possible Dynkin diagrams, we are going to to determine all graphs satisfying conditions (A), (B), (C). Then, having found all such graphs, we will determine which ones occur as Dynkin diagrams.

6.3 Classification of Dynkin diagrams

In this section, we will focus on proving the following theorem.

Theorem 6.7. *The graphs satisfying conditions (A), (B), (C) from the previous section are precisely those in the following list:*



Proof. These graphs clearly satisfy conditions (A) and (B). We now concentrate on proving that they satisfy condition (C). It is a standard result in linear algebra that a quadratic form $\sum a_{ij}x_ix_j$ is positive definite if and only if the leading minors of its symmetric matrix (a_{ij}) have positive determinant, that is,

$$|a_{11}| > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \ldots, \quad \det(a_{ij}) > 0.$$

Given a graph Γ on the list with l vertices, we shall show that $\mathrm{Q}(x_1, \ldots, x_l)$ is positive definite by induction on l.

If $l = 1$, then $\Gamma = A_1$ and $\mathrm{Q}(x_1) = 2x_1^2$ is positive definite. If $l = 2$, then $\Gamma = A_2$,

B_2, or G_2. In each case, the symmetric matrix corresponding to $\mathrm{Q}(x_1, x_2)$ is

$$A_2 : \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \qquad B_2 : \begin{pmatrix} 2 & -\sqrt{2} \\ -\sqrt{2} & 2 \end{pmatrix} \qquad G_2 : \begin{pmatrix} 2 & -\sqrt{3} \\ -\sqrt{3} & 2 \end{pmatrix}.$$

One sees that the leading minors of these matrices have positive determinant. Now suppose $l \geq 3$. Looking at the above list, we see that Γ contains at least one vertex that is joined to just one other vertex, and joined to it by a single edge. Label such a vertex l and label the vertex it is joined to $l-1$. We write $\Gamma = \Gamma_l$. We denote the graph obtained from Γ_l by removing the vertex l by Γ_{l-1} and the graph obtained from Γ_{l-1} by removing the vertex $l-1$ by Γ_{l-2}. Observe that Γ_{l-1} and Γ_{l-2} are also on the list. Let $\det \Gamma_l$ be the determinant of the symmetric matrix representing the quadratic form $\mathrm{Q}(x_1, \ldots, x_l)$ associated with Γ_l. We obtain the equality

$$\det \Gamma_l = \begin{vmatrix} & & & & 0 \\ & & & & \vdots \\ & & & & 0 \\ & & & 2 & -1 \\ 0 & \ldots & 0 & -1 & 2 \end{vmatrix} = 2 \det \Gamma_{l-1} - \det \Gamma_{l-2}$$

by expanding the determinant by its last row. This gives an inductive way to calculate $\det \Gamma_l$. We perform this calculation for each Γ_l on the list.

We have

$$\det A_1 = 2, \quad \det A_2 = 3, \quad \ldots, \quad \det A_l = 2 \det A_{l-1} - \det A_{l-2}.$$

Thus $\det A_l = l + 1$. We also have

$$\det A_1 = 2, \quad \det B_2 = 2, \quad \det B_3 = 2, \quad \ldots, \quad \det B_l = 2 \det B_{l-1} - \det B_{l-2}.$$

Thus $\det B_l = 2$. Next, we have

$$\det A_3 = 4, \quad \det D_4 = 4, \quad \det D_5 = 4, \quad \ldots, \quad \det D_l = 2 \det D_{l-1} - \det D_{l-2}.$$

Thus $\det D_l = 4$. Finally, we have

$$\det E_6 = 2\det D_5 - \det A_4 = 3$$

$$\det E_7 = 2\det D_6 - \det A_5 = 2$$

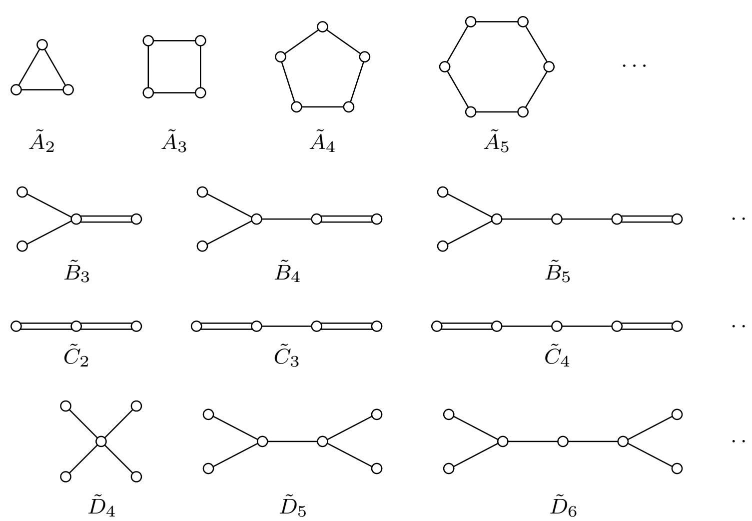
$$\det E_8 = 2\det E_7 - \det A_6 = 1$$

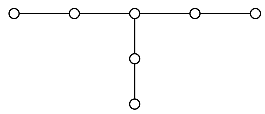
$$\det F_4 = 2\det B_3 - \det A_2 = 1$$

and $\det G_2 = 1$ from the previous list of matrices. Thus we have shown that $\det \Gamma_l > 0$ for all Γ_l. The leading minors of the symmetric matrix associated with Γ_l are the symmetric matrices associated with certain subgraphs of Γ_l. We may number the vertices of Γ_l such that every subgraph is connected. But the list of graphs has the property that every subgraph of any graph on the list is also on the list. Thus every leading minor of the symmetric matrix associated with Γ_l is positive, and so the quadratic form $\mathrm{Q}(x_1, \ldots, x_l)$ associated with Γ_l is positive definite.

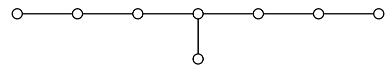
So far, we have shown that the graphs on the list satisfy conditions (A), (B), (C). The following lemmas will prove the converse, that is, that any graph satisfying conditions (A), (B), (C) is on the list.

Lemma 6.8. *For each of the graphs on the following list, the corresponding quadratic form* $\mathrm{Q}(x_1, \ldots, x_l)$ *has determinant 0:*

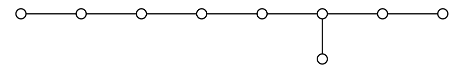




$\tilde{E}_6$



$\tilde{E}_7$



$\tilde{E}_8$

$\tilde{F}_4$

$\tilde{G}_2$

Proof. For the graphs $\Gamma = \tilde{A}_l$, each row of the symmetric matrix associated with the given quadratic form contains one entry 2 and two entries -1. The remaining entries are 0. Thus the sum of the columns is zero, and so $\det \tilde{A}_l = 0$.

In all of the other graphs Γ on the list, we can find a vertex l that is joined to just one vertex $l-1$. Moreover, we may choose l such that it is connected to $l-1$ by either a single or a double edge. If there is a single edge, then we may use the formula

$$\det \Gamma_l = 2 \det \Gamma_{l-1} - \det \Gamma_{l-2}$$

as before. If there is a double edge, then we may modify the way we obtained the previous formula to get the formula

$$\det \Gamma_l = 2 \det \Gamma_{l-1} - 2 \det \Gamma_{l-2}.$$

As before, we calculate the determinants of the above graphs inductively. We have

$$\begin{aligned}
&\det \tilde{B}_3 = 2 \det A_3 - 2(\det A_1)^2 = 0 && \\
&\det \tilde{B}_l = 2 \det D_l - 2 \det D_{l-1} = 0 && \text{for } l \geq 4 \\
&\det \tilde{C}_2 = 2 \det B_2 - 2 \det A_1 = 0 && \\
&\det \tilde{C}_l = 2 \det B_l - 2 \det B_{l-1} = 0 && \text{for } l \geq 3 \\
&\det \tilde{D}_4 = 2 \det D_4 - (\det A_1)^3 = 0 && \\
&\det \tilde{D}_l = 2 \det D_l - \det D_{l-2} \cdot \det A_1 = 0 && \text{for } l \geq 5 \\
&\det \tilde{E}_6 = 2 \det E_6 - \det A_5 = 0 && \\
&\det \tilde{E}_7 = 2 \det E_7 - \det D_6 = 0 && \\
&\det \tilde{E}_8 = 2 \det E_8 - \det E_7 = 0 &&
\end{aligned}$$

$$\det \tilde{F}_4 = 2\det F_4 - \det B_3 = 0$$

$$\det \tilde{G}_2 = 2\det G_2 - \det A_1 = 0.$$

This exhausts all cases. □

Lemma 6.9. *Let Γ be a graph satisfying conditions (A), (B), (C). Let Γ′ be a connected subgraph of Γ, that is, a connected graph obtained from Γ by omitting vertices or decreasing the number of edges between vertices or both. Then Γ′ satisfies conditions (A), (B), (C), also.*

Proof. The subgraph Γ′ clearly satisfies conditions (A) and (B). We need to show that it satisfies (C). Let $\mathrm{Q}(x_1, \dots, x_l)$ be the quadratic form of Γ and $\mathrm{Q}'(x_1, \dots, x_m)$ the quadratic form of Γ′ where $m \leq l$. We have

$$\mathrm{Q}(x_1, \dots, x_l) = 2\sum_{i=1}^{l} x_i^2 - \sum_{\substack{i,j=1 \\ i\neq j}}^{l} \sqrt{n_{ij}}\, x_i x_j$$
$$\mathrm{Q}'(x_1, \dots, x_m) = 2\sum_{i=1}^{m} x_i^2 - \sum_{\substack{i,j=1 \\ i\neq j}}^{m} \sqrt{n'_{ij}}\, x_i x_j$$

where $n'_{ij} \leq n_{ij}$ for $i, j = 1, \dots, m$. Suppose Q′ is not positive definite. Then there exist $y_1, \dots, y_m \in \mathbb{R}$, not all zero, such that $\mathrm{Q}'(y_1, \dots, y_m) \leq 0$. Now $(|y_1|, \dots, |y_m|, 0, \dots, 0)$ is not the zero vector in $\mathbb{R}^l$, but

$$\begin{aligned}\mathrm{Q}(|y_1|, \dots, |y_m|, 0, \dots, 0) &= 2\sum_{i=1}^{m} |y_i|^2 - \sum_{\substack{i,j=1 \\ i\neq j}}^{m} \sqrt{n_{ij}}\, |y_i||y_j| \\ &\leq 2\sum_{i=1}^{m} y_i^2 - \sum_{\substack{i,j=1 \\ i\neq j}}^{m} \sqrt{n'_{ij}}\, |y_i||y_j| \\ &\leq 2\sum_{i=1}^{m} y_i^2 - \sum_{\substack{i,j=1 \\ i\neq j}}^{m} \sqrt{n'_{ij}}\, y_i y_j \\ &= \mathrm{Q}'(y_1, \dots, y_m) \leq 0.\end{aligned}$$

This contradicts the fact that $\mathrm{Q}(x_1, \dots, x_l)$ is positive definite, so $\mathrm{Q}'(x_1, \dots, x_m)$ must be positive definite, also. □

Having proved Lemmas 6.8 and 6.9, we are now ready to complete the proof of Theorem 6.7.

Let Γ be a graph satisfying conditions (A), (B), (C). Then by Lemmas 6.8 and 6.9, Γ can have no subgraph of type $\tilde{A}_l$, $\tilde{B}_l$, $\tilde{C}_l$, $\tilde{D}_l$, $\tilde{E}_6$, $\tilde{E}_7$, $\tilde{E}_8$, $\tilde{F}_4$, or $\tilde{G}_2$. We will use this information to show that Γ must be one of the graphs on the list in Theorem 6.7.

We first notice that Γ contains no cycles, for otherwise $\tilde{A}_l$ would be a subgraph of Γ for some $l \geq 2$.

Suppose Γ contains a triple edge. Then Γ must be the graph G_2, as otherwise $\tilde{G}_2$ would be a subgraph of Γ.

Thus we may assume Γ contains no triple edge. Suppose Γ contains a double edge. Then Γ cannot contain more than one double edge, for otherwise $\tilde{C}_l$ would be a subgraph of Γ for some $l \geq 2$. Now Γ cannot contain a branch point in addition to a double edge, as otherwise $\tilde{B}_l$ would be a subgraph of Γ for some $l \geq 3$. Thus Γ is a chain containing only one double edge. If the double edge lies at the end of the chain, then $\Gamma = B_l$ for some $l \geq 2$. If the double edge lies elsewhere, then $\Gamma = F_4$, for otherwise $\tilde{F}_4$ would be a subgraph of Γ.

Now suppose Γ contains no double or triple edges. If Γ contains no branch point, then $\Gamma = A_l$ for some $l \geq 1$. Thus we assume that Γ has at least one branch point. We see that Γ cannot contain more than one branch point, for otherwise $\tilde{D}_l$ would be a subgraph of Γ for some $l \geq 5$. Thus Γ contains exactly one branch point. There can be at most three branches stemming from this branch point, since otherwise $\tilde{D}_4$ would be a subgraph of Γ. Let l_1, l_2, l_3 be the number of vertices on the three branches where $l_1 \geq l_2 \geq l_3$. Then the total number of vertices of Γ is $l_1 + l_2 + l_3 + 1$.

We observe that l_3 must equal 1, for otherwise we would have $l_i \geq 2$ for $i = 1, 2, 3$, and so $\tilde{E}_6$ would be a subgraph of Γ. If $l_2 = 1$, then $\Gamma = D_l$ for some $l \geq 4$. Thus we may assume $l_2 \geq 2$. But we cannot have $l_2 > 2$, as otherwise we would have $l_1 \geq 3, l_2 \geq 3$, and so $\tilde{E}_7$ would be a subgraph of Γ. Hence $l_3 = 1$ and $l_2 = 2$. We must have $l_1 \leq 4$, for otherwise $\tilde{E}_8$ would be a subgraph of Γ. Thus Γ is either E_6, E_7, or E_8.

This exhausts all possibilities for Γ, and in each case Γ is a graph on the list in Theorem 6.7. This completes the proof. □

Corollary 6.10. *Let Δ be the Dynkin diagram of a semisimple Lie algebra. Then each connected component of Δ must be one of the following graphs:*

$$A_l, \quad l \geq 1; \quad B_l, \quad l \geq 2; \quad D_l, \quad l \geq 4; \quad E_6; \quad E_7; \quad E_8; \quad F_4; \quad G_2.$$

Proof. This is immediate from Theorem 6.7 since the connected components of the Dynkin diagram of a semisimple Lie algebra satisfy conditions (A), (B), (C). □

We will consider whether all of the graphs listed above actually occur as Dynkin diagrams in the following chapter.

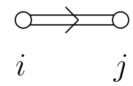
6.4 Classification of Cartan matrices

Recall that the Dynkin diagram is uniquely determined by the Cartan matrix via the relation

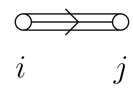
$$n_{ij} = A_{ij}A_{ji} \qquad i \neq j.$$

But the Cartan matrix is not necessarily determined by the Dynkin diagram. Given the integers $n_{ij} \in \{0, 1, 2, 3\}$ for all i, j with $i \neq j$, we now consider the extent to which the A_{ij} are determined. If $n_{ij} = 0$, then we must have $A_{ij} = A_{ji} = 0$ since $\mathbb{Z}$ is an integral domain. If $n_{ij} = 1$, then we must have $A_{ij} = A_{ji} = -1$ since $A_{ij}, A_{ji} \in \mathbb{Z}$ and $A_{ij}, A_{ji} \leq 0$. However, for $n_{ij} = 2$, there are two ways to factor $n_{ij} = A_{ij}A_{ji}$, namely $A_{ij} = -1, A_{ji} = -2$ and $A_{ij} = -2, A_{ji} = -1$. Similarly, if $n_{ij} = 3$, then we have $A_{ij} = -1, A_{ji} = -3$ or $A_{ij} = -3, A_{ji} = -1$.

Out of the list in Corollary 6.10, the only graphs giving rise to such an ambiguity are B_l, $l \geq 2$, F_4, and G_2. In order to remove this ambiguity, we place an arrow on the double and triple edges. The direction of the arrow is determined as follows: The arrow points from vertex i to vertex j if and only if $|\alpha_i| > |\alpha_j|$ in the Euclidean space V. This is equivalent to the condition $|A_{ji}| > |A_{ij}|$. Thus in the situation



we have $|\alpha_i| = \sqrt{2}|\alpha_j|$ and $A_{ij} = -1, A_{ji} = -2$. In the situation



we have we have $|\alpha_i| = \sqrt{3}|\alpha_j|$ and $A_{ij} = -1, A_{ji} = -3$. Thus we may regard the arrow as an inequality sign on the lengths of the fundamental roots at the vertices.

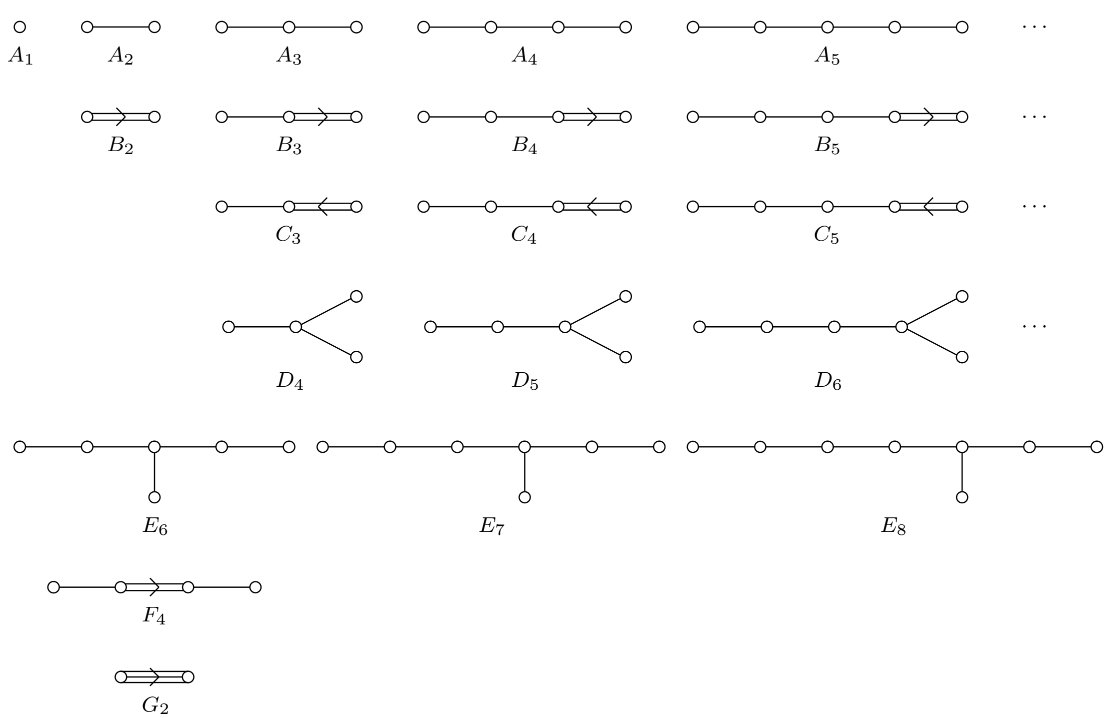


Figure 6.3: Standard list of connected Dynkin diagrams.

The set of possible Dynkin diagrams, including arrows, is shown in Figure 6.3.

Note that, since the diagrams B_2, F_4, G_2 are symmetric, it does not matter in which direction the arrow is drawn in these graphs. By the preceding remarks, the connected components of the Dynkin diagram of any semisimple Lie algebra must appear on this list.

From this list, we obtain a standard list of corresponding Cartan matrices. Two Cartan matrices $(A_{ij}), \left(A'_{ij}\right)$ are called **equivalent** if there exists a permutation σ of $1, \ldots, l$ such that

$$A'_{ij} = A_{\sigma(i)\sigma(j)} \qquad \text{for all } i, j.$$

We see that equivalent Cartan matrices arise from different labellings of the vertices of the same Dynkin diagram. By choosing a suitable labeling of vertices of the Dynkin diagrams in Figure 6.3, we obtain the following list of corresponding Cartan matrices determined uniquely up to equivalence:

$$A_l = \begin{pmatrix} 2 & -1 & & & & & & & \\ -1 & 2 & -1 & & & & & & \\ & -1 & 2 & -1 & & & & & \\ & & -1 & \cdot & \cdot & & & & \\ & & & \cdot & \cdot & \cdot & & & \\ & & & & \cdot & \cdot & -1 & & \\ & & & & & -1 & 2 & -1 & \\ & & & & & & -1 & 2 & -1 \\ & & & & & & & -1 & 2 \end{pmatrix} \qquad l \geq 1$$

$$B_l = \begin{pmatrix} 2 & -1 & & & & & & & \\ -1 & 2 & -1 & & & & & & \\ & -1 & 2 & -1 & & & & & \\ & & -1 & \cdot & \cdot & & & & \\ & & & \cdot & \cdot & \cdot & & & \\ & & & & \cdot & \cdot & -1 & & \\ & & & & & -1 & 2 & -1 & \\ & & & & & & -1 & 2 & -1 \\ & & & & & & & -2 & 2 \end{pmatrix} \qquad l \geq 2$$

$$C_l = \begin{pmatrix} 2 & -1 & & & & & & & \\ -1 & 2 & -1 & & & & & & \\ & -1 & 2 & -1 & & & & & \\ & & -1 & \cdot & \cdot & & & & \\ & & & \cdot & \cdot & \cdot & & & \\ & & & & \cdot & \cdot & -1 & & \\ & & & & & -1 & 2 & -1 & \\ & & & & & & -1 & 2 & -2 \\ & & & & & & & -1 & 2 \end{pmatrix} \qquad l \geq 3$$

$$D_l = \begin{pmatrix} 2 & -1 & & & & & & & & \\ -1 & 2 & -1 & & & & & & & \\ & -1 & 2 & -1 & & & & & & \\ & & -1 & \cdot & \cdot & & & & & \\ & & & \cdot & \cdot & \cdot & & & & \\ & & & & \cdot & \cdot & -1 & & & \\ & & & & & -1 & 2 & -1 & & \\ & & & & & & -1 & 2 & -1 & -1 \\ & & & & & & & -1 & 2 & 0 \\ & & & & & & & -1 & 0 & 2 \end{pmatrix} \qquad l \geq 4$$

$$E_6 = \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & -1 & \\ & & -1 & 2 & & \\ & & -1 & & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix}$$

$$E_7 = \begin{pmatrix} 2 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & -1 & 2 & -1 & & & \\ & & -1 & 2 & -1 & -1 & \\ & & & -1 & 2 & & \\ & & & -1 & & 2 & -1 \\ & & & & & -1 & 2 \end{pmatrix}$$

$$E_8 = \begin{pmatrix} 2 & -1 & & & & & & \\ -1 & 2 & -1 & & & & & \\ & -1 & 2 & -1 & & & & \\ & & -1 & 2 & -1 & & & \\ & & & -1 & 2 & -1 & -1 & \\ & & & & -1 & 2 & & \\ & & & & -1 & & 2 & -1 \\ & & & & & & -1 & 2 \end{pmatrix}$$

$$F_4 = \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & -2 & 2 & -1 \\ & & -1 & 2 \end{pmatrix}$$

$$G_2 = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}$$

A Cartan matrix is **indecomposable** if its Dynkin diagram is connected. We know that every Cartan matrix uniquely determines a Dynkin diagram, and that this diagram decomposes into connected components. Each connected component corresponds to an indecomposable Cartan matrix, and thus every Cartan matrix can be expressed as the direct sum of indecomposable Cartan matrices.

If A is the Cartan matrix of any semisimple Lie algebra, then each indecomposable component of A will be equivalent to some Cartan matrix from the previous list.

Proposition 6.11. *If a semisimple Lie algebra $\mathfrak{g}$ has a connected Dynkin diagram, then $\mathfrak{g}$ is simple.*

Proof. Let $\mathfrak{h} \oplus \sum_{\alpha\in\Phi} \mathfrak{g}_\alpha$ be a Cartan decomposition of $\mathfrak{g}$ giving rise to the Dynkin diagram Δ. Let $\mathfrak{a}$ be a nonzero ideal of $\mathfrak{g}$. We will show that $\mathfrak{a} = \mathfrak{g}$, thus proving that $\mathfrak{g}$ is simple.

We first prove that $\mathfrak{a} \cap \mathfrak{h} \neq 0$. Suppose $\mathfrak{a} \cap \mathfrak{h} = 0$. For each $\alpha \in \Phi$, let e_α be a nonzero element in $\mathfrak{g}_\alpha$. We choose a nonzero element $x \in \mathfrak{a}$ and write

$$x = h + \sum_{\alpha\in\Phi} \mu_\alpha e_\alpha \qquad h \in \mathfrak{h},\ \mu_\alpha \in \mathbb{C}$$

such that the number of nonzero μ_α is as small as possible. Now since $\mathfrak{a} \cap \mathfrak{h} = 0$, there exists some $\beta \in \Phi$ with $\mu_\beta \neq 0$. We have

$$[h'_\beta, x] = \sum_{\alpha\in\Phi} \mu_\alpha \left[h'_\beta, e_\alpha\right] = \sum_{\alpha\in\Phi} \mu_\alpha \alpha\left(h'_\beta\right) e_\alpha$$

since $\mathfrak{h}$ is abelian. By Proposition 4.17, there exist elements $e_\beta \in \mathfrak{g}_\beta$ and $e_{-\beta} \in \mathfrak{g}_{-\beta}$ such that $[e_\beta, e_{-\beta}] = h'_\beta$. Thus

$$\left[\left[h'_\beta, x\right], e_{-\beta}\right] = \sum_{\alpha\in\Phi} \mu_\alpha \alpha\left(h'_\beta\right)\left[e_\alpha, e_{-\beta}\right] = \mu_\beta \beta\left(h'_\beta\right) h'_\beta + \sum_{\substack{\alpha\in\Phi \\ \alpha\neq\beta}} \mu_\alpha \alpha\left(h'_\beta\right) N_{\alpha,-\beta} e_{\alpha-\beta}$$

where $[e_\alpha, e_{-\beta}] = N_{\alpha,-\beta} e_{\alpha-\beta}$. Now $[[h'_\beta, x], e_{-\beta}] \in \mathfrak{a}$ since $x \in \mathfrak{a}$. Also, $[[h'_\beta, x], e_{-\beta}] \neq 0$ since $\mu_\beta \neq 0$ and $\beta\big(h'_\beta\big) = \left\langle h'_\beta, h'_\beta\right\rangle \neq 0$ by Proposition 4.18. Notice that the number of nonzero components from the root spaces $\mathfrak{g}_\alpha$ is less for $[[h'_\beta, x], e_{-\beta}]$ than it was for x. This contradicts our choice of x, so we must have $\mathfrak{a} \cap \mathfrak{h} \neq 0$.

We next show that $\mathfrak{h} \subset \mathfrak{a}$. Suppose this is not the case. Then $0 \neq \mathfrak{a} \cap \mathfrak{h} \neq \mathfrak{h}$. Now there exist $\alpha_i \in \Pi$ and $x \in \mathfrak{a} \cap \mathfrak{h}$ such that $\left\langle h'_{\alpha_i}, x \right\rangle \neq 0$, for otherwise $\mathfrak{a} \cap \mathfrak{h}$ would be orthogonal to each h'_{α_i}, hence to the whole of $\mathfrak{h}$, and would thus be 0. We choose $e_{\alpha_i} \in \mathfrak{g}_{\alpha_i}$ and $e_{-\alpha_i} \in \mathfrak{g}_{-\alpha_i}$ such that $[e_{\alpha_i}, e_{-\alpha_i}] = h'_{\alpha_i}$. We have

$$[x, e_{\alpha_i}] = \alpha_i(x) e_{\alpha_i} = \left\langle h'_{\alpha_i}, x \right\rangle e_{\alpha_i} \in \mathfrak{a}.$$

Since $\left\langle h'_{\alpha_i}, x \right\rangle \neq 0$, we see that $e_{\alpha_i} \in \mathfrak{a}$. It follows that $[e_{\alpha_i}, e_{-\alpha_i}] = h'_{\alpha_i} \in \mathfrak{a}$.

We may now divide the $\alpha_i \in \Pi$ into two classes: those with $h'_{\alpha_i} \in \mathfrak{a}$, and those with $h'_{\alpha_i} \notin \mathfrak{a}$. Since we assumed $\mathfrak{h}$ was not contained in $\mathfrak{a}$, we see that both classes are nonempty. I claim that if $h'_{\alpha_i} \in \mathfrak{a}$ and $h'_{\alpha_j} \notin \mathfrak{a}$, then $\left\langle h'_{\alpha_j}, h'_{\alpha_i} \right\rangle = 0$. Suppose this is not true. We choose $e_{\alpha_j} \in \mathfrak{g}_{\alpha_j}$ and $e_{-\alpha_j} \in \mathfrak{g}_{-\alpha_j}$ such that $[e_{\alpha_j}, e_{-\alpha_j}] = h'_{\alpha_j}$. Then

$$\left[h'_{\alpha_i}, e_j\right] = \alpha_j \left(h'_{\alpha_i}\right) e_j = \left\langle h'_{\alpha_i}, h'_{\alpha_j} \right\rangle e_j \in \mathfrak{a}.$$

Since $\left\langle h'_{\alpha_i}, h'_{\alpha_j} \right\rangle \neq 0$, we must have $e_{\alpha_j} \in \mathfrak{a}$. But then $[e_{\alpha_j}, e_{-\alpha_j}] = h'_{\alpha_j} \in \mathfrak{a}$, a contradiction. This completes the proof of the claim. Thus if $h'_{\alpha_i} \in \mathfrak{a}$ and $h'_{\alpha_j} \notin \mathfrak{a}$, then $\left\langle h'_{\alpha_j}, h'_{\alpha_i} \right\rangle = 0$, that is, $\langle \alpha_i, \alpha_j \rangle = 0$. This means that the vertices corresponding to the roots in both classes are not joined in the Dynkin diagram Δ, and so Δ is disconnected. This is a contradiction, so we must have $\mathfrak{h} \subset \mathfrak{a}$.

We are now ready to prove that $\mathfrak{a} = \mathfrak{g}$. Let $\alpha \in \Phi$. We have

$$\left[h'_{\alpha}, e_{\alpha}\right] = \alpha \left(h'_{\alpha}\right) e_{\alpha} = \left\langle h'_{\alpha}, h'_{\alpha} \right\rangle e_{\alpha}.$$

Since $\mathfrak{h} \subset \mathfrak{a}$, we have $h'_{\alpha} \in \mathfrak{a}$, and so $[h'_{\alpha}, e_{\alpha}] \in \mathfrak{a}$. Because $\left\langle h'_{\alpha}, h'_{\alpha} \right\rangle \neq 0$, this implies that $e_{\alpha} \in \mathfrak{a}$. This is true for all $\alpha \in \Phi$, and so $\mathfrak{a} = \mathfrak{g}$. Thus $\mathfrak{g}$ is simple. □

Our next goal is to prove the converse of Proposition 6.11. We first define an action of the Weyl group W on the Cartan subalgebra $\mathfrak{h}$. Recall that the Weyl group is a group of non-singular linear transformations on the real vector space $\mathfrak{h}^*_{\mathbb{R}}$. We can extend the action of W on $\mathfrak{h}^*_{\mathbb{R}}$ by linearity to give an action of W on $\mathfrak{h}^*$ by $\mathbb{C}$-linear transformations. We now define the action of W on $\mathfrak{h}$ by $h \mapsto wh$ where

$$\lambda(wh) = \left(w^{-1}\lambda\right) h \qquad \text{for all } h \in \mathfrak{h},\ \lambda \in \mathfrak{h}^*,\ w \in W.$$

To prove this action is well defined, suppose $x_1, x_2 \in \mathfrak{h}$ satisfy $\lambda(x_1) = \left(w^{-1}\lambda\right) h$ and $\lambda(x_2) = \left(w^{-1}\lambda\right) h$ for all $\lambda \in \mathfrak{h}^*$. Then $x_1 = x_2$. For otherwise $x_1 - x_2 \neq 0$, and so there would exist a $\lambda \in \mathfrak{h}^*$ such that $\lambda(x_1 - x_2) \neq 0$. But $\lambda(x_1) = \lambda(x_2)$ since x_1, x_2 satisfy the above conditions. Thus $\lambda(x_1 - x_2) = 0$, a contradiction.

The action $h \mapsto wh$ has some very nice properties. We observe that

$$w_1(w_2h) = (w_1w_2)h \qquad \text{for all } h \in \mathfrak{h},\ w_1, w_2 \in W.$$

This follows from the fact that

$$\begin{aligned}\lambda(w_1(w_2h)) &= \left(w_1^{-1}\lambda\right) w_2h = \left(w_2^{-1}\left(w_1^{-1}\lambda\right)\right) h = \left(\left(w_2^{-1}w_1^{-1}\right)\lambda\right) h \\ &= \left((w_1w_2)^{-1}\lambda\right) h = \lambda((w_1w_2)h)\end{aligned}$$

for all $\lambda \in \mathfrak{h}^*$. Moreover, the actions of W on $\mathfrak{h}^*$ and $\mathfrak{h}$ are compatible with the isomorphism $\mathfrak{h}^* \to \mathfrak{h}$ given by $\lambda \mapsto h'_\lambda$ where $\lambda(x) = \left\langle h'_\lambda, x\right\rangle$ for all $x \in \mathfrak{h}$. For suppose $w(\lambda) = \mu$ for $\lambda, \mu \in \mathfrak{h}^*$. Then

$$\left\langle w\left(h'_\lambda\right), x\right\rangle = \left\langle h'_\lambda, w^{-1}(x)\right\rangle = \lambda\left(w^{-1}(x)\right) = (w\lambda)x = \mu(x) = \left\langle h'_\mu, x\right\rangle \qquad \text{for all } x \in \mathfrak{h}.$$

Hence $w(\lambda) = \mu$ implies $w\left(h'_\lambda\right) = h'_\mu$. Having this relation, we deduce that since

$$s_\alpha(\lambda) = \lambda - 2\frac{\langle \alpha, \lambda\rangle}{\langle \alpha, \alpha\rangle}\alpha \qquad \text{for } \alpha \in \Phi,\ \lambda \in \mathfrak{h}^*,$$

we must have

$$s_\alpha(x) = x - 2\frac{\left\langle h'_\alpha, x\right\rangle}{\left\langle h'_\alpha, h'_\alpha\right\rangle} h'_\alpha \qquad \text{for } x \in \mathfrak{h}.$$

Proposition 6.12. *Let $\mathfrak{g}$ be a semisimple Lie algebra whose Dynkin diagram Δ splits into connected components $\Delta_1, \ldots, \Delta_r$. Then we have*

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r$$

where $\mathfrak{g}_i$ is a simple Lie algebra with Dynkin diagram Δ_i.

Proof. We have $\Delta = \Delta_1 \dot{\cup} \Delta_2 \dot{\cup} \cdots \dot{\cup} \Delta_r$. For each i, let Π_i be the subset of Π corresponding to the vertices in Δ_i. We have

$$\Pi = \Pi_1 \dot{\cup} \Pi_2 \dot{\cup} \cdots \dot{\cup} \Pi_r.$$

We also have $\langle \alpha, \beta \rangle = 0$ if $\alpha \in \Pi_i$, $\beta \in \Pi_j$, and $i \neq j$. Let $\mathfrak{h}_i$ be the subspace of $\mathfrak{h}$ spanned by the elements h'_α with $\alpha \in \Pi_i$. Then

$$\mathfrak{h} = \mathfrak{h}_1 \oplus \mathfrak{h}_2 \oplus \cdots \oplus \mathfrak{h}_r$$

where $\left\langle h, h' \right\rangle = 0$ if $h \in \mathfrak{h}_i$, $h' \in \mathfrak{h}_j$, and $i \neq j$.

Now let $\alpha \in \Pi_i$ and consider the fundamental reflection $s_\alpha \in W$. One easily checks that s_α transforms $\mathfrak{h}_i$ into itself and fixes every vector in $\mathfrak{h}_j$ for all $j \neq i$. Thus

$$s_\alpha (\mathfrak{h}_j) = \mathfrak{h}_j \qquad j = 1, \ldots, r.$$

Because the elements s_α generate W, we see that

$$w (\mathfrak{h}_j) = \mathfrak{h}_j \qquad j = 1, \ldots, r \quad w \in W.$$

Now for all $\alpha \in \Phi$, we have $h'_\alpha = w \left(h'_{\alpha_i} \right)$ for some $\alpha_i \in \Pi$ and $w \in W$ by Proposition 5.13 and the way we defined the W-action on $\mathfrak{h}$. It follows that for each $\alpha \in \Phi$, h'_α lies in $\mathfrak{h}_i$ for some i. For each i, let Φ_i be the set of all $\alpha \in \Phi$ such that $h'_\alpha \in \mathfrak{h}_i$. Then

$$\Phi = \Phi_1 \dot{\cup} \Phi_2 \dot{\cup} \cdots \dot{\cup} \Phi_r.$$

We now define $\mathfrak{g}_i$ to be the subspace of $\mathfrak{g}$ spanned by $\mathfrak{h}_i$ and the e_α for all $\alpha \in \Phi_i$. Given the Cartan decomposition $\mathfrak{g} = \mathfrak{h} \oplus \sum_{\alpha \in \Phi} \mathbb{C} e_\alpha$ and the above decomposition of $\mathfrak{h}$, we see that

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_r,$$

a direct sum of subspaces. I claim that each $\mathfrak{g}_i$ is a subalgebra of $\mathfrak{g}$. Since $\mathfrak{h}$ is abelian and $[h'_\alpha, e_\beta] = \beta (h'_\alpha) e_\beta$ for all $\alpha \in \Pi_i$, $\beta \in \Phi_i$, we need only verify that $[e_\alpha, e_\beta] \in \mathfrak{g}_i$ if

$\alpha, \beta \in \Phi_i$. Now if $\alpha + \beta \in \Phi$, then $\alpha + \beta \in \Phi_i$ since $h'_{\alpha+\beta} = h'_\alpha + h'_\beta \in \mathfrak{h}_i$. If $\alpha + \beta = 0$, then $[e_\alpha, e_\beta]$ is a scalar multiple of h'_α and thus lies in $\mathfrak{h}_i$, and hence in $\mathfrak{g}_i$. Finally, if $\alpha + \beta \neq 0$ but is not a root, then $[e_\alpha, e_\beta] = 0$. In all cases, we have $[e_\alpha, e_\beta] \in \mathfrak{g}_i$, and so $\mathfrak{g}_i$ is a subalgebra of $\mathfrak{g}$.

Next, we show that $[\mathfrak{g}_i, \mathfrak{g}_j] = 0$ if $i \neq j$. Let $\alpha \in \Phi_i$ and $\beta \in \Phi_j$. We have

$$\left[h'_\alpha, e_\beta\right] = \beta\left(h'_\alpha\right) e_\beta = \langle h'_\beta, h'_\alpha \rangle e_\beta = 0.$$

Similarly, we have $[e_\alpha, h'_\beta] = 0$. We also have $[e_\alpha, e_\beta] = 0$, because $\alpha + \beta \notin \Phi$ since $h'_\alpha + h'_\beta$ does not lie in any subspace $\mathfrak{h}_k$ of $\mathfrak{h}$. It follows that $[\mathfrak{g}_i, \mathfrak{g}_j] = 0$.

Now each $\mathfrak{g}_i$ is an ideal of $\mathfrak{g}$ since

$$[\mathfrak{g}_i, \mathfrak{g}] = \sum_j [\mathfrak{g}_i, \mathfrak{g}_j] = [\mathfrak{g}_i, \mathfrak{g}_i] \subset \mathfrak{g}_i.$$

This implies that

$$[x_1 + \cdots + x_r, y_1 + \cdots + y_r] = [x_1, y_1] + \cdots + [x_r, y_r]$$

where $x_i, y_i \in \mathfrak{g}_i$. Thus

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_r$$

is a direct sum of Lie algebras.

We also observe that each $\mathfrak{g}_i$ is a semisimple Lie algebra. Let $\mathfrak{a}$ be a solvable ideal of $\mathfrak{g}_i$. Since $[\mathfrak{a}, \mathfrak{g}_j] = 0$ for all $j \neq i$, we see that $\mathfrak{a}$ is an ideal of $\mathfrak{g}$. Because $\mathfrak{g}$ is semisimple, we have $\mathfrak{a} = 0$. Hence $\mathfrak{g}_i$ is semisimple.

Next, we show that $\mathfrak{h}_i$ is a Cartan subalgebra of $\mathfrak{g}_i$. We know that $\mathfrak{h}_i$ is abelian, and hence nilpotent. Let $x \in N(\mathfrak{h}_i)$, the normalizer of $\mathfrak{h}_i$ in $\mathfrak{g}_i$. Then $[x, h] \in \mathfrak{h}_i$ for all $h \in \mathfrak{h}_i$. We also have $[x, h] = 0$ for all $h \in \mathfrak{h}_j$ with $j \neq i$. It follows that $[x, h] \in \mathfrak{h}$ for all $h \in \mathfrak{h}$. But $N(\mathfrak{h}) = \mathfrak{h}$ since $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$. Thus $x \in \mathfrak{h}$. It follows that $x \in \mathfrak{h} \cap \mathfrak{g}_i = \mathfrak{h}_i$. This shows that $N(\mathfrak{h}_i) = \mathfrak{h}_i$, and so $\mathfrak{h}_i$ is a Cartan subalgebra of $\mathfrak{g}_i$.

We now consider the Cartan decomposition

$$\mathfrak{g}_i = \mathfrak{h}_i \oplus \sum_{\alpha \in \Phi_i} \mathbb{C} e_\alpha$$

of $\mathfrak{g}_i$ with respect to $\mathfrak{h}_i$. We see that Φ_i is the root system of $\mathfrak{g}_i$, that Π_i is a fundamental system in Φ_i, and that Δ_i is the Dynkin diagram of $\mathfrak{g}_i$. Now Δ_i is connected, and so $\mathfrak{g}_i$ is simple by Proposition 6.11. Thus we have obtained a decomposition of $\mathfrak{g}$ as a direct sum of Lie algebras $\mathfrak{g}_i$ whose respective Dynkin diagrams are the connected components Δ_i of Δ. □

Corollary 6.13. *A semisimple Lie algebra $\mathfrak{g}$ has a connected Dynkin diagram if and only if $\mathfrak{g}$ is simple.*

Proof. This follows immediately from Propositions 6.11 and 6.12. □

Chapter 7

The existence and uniqueness theorems

In the last chapter, we saw that each non-trivial simple Lie algebra $\mathfrak{g}$ has a Dynkin diagram Δ that appears on the standard list of connected Dynkin diagrams in Figure 6.3. But does the converse hold? That is, given a Dynkin diagram Δ on the standard list, is there a simple Lie algebra $\mathfrak{g}$ with Dynkin diagram Δ? If so, does Δ uniquely determine $\mathfrak{g}$ up to isomorphism? In this chapter, we will show that both existence and uniqueness hold. The proof of the uniqueness property is somewhat easier, so we will prove this first. In order to do so, we first study some properties of the structure constants of the Lie algebra $\mathfrak{g}$.

7.1 Properties of structure constants

Let $\mathfrak{g}$ be a simple Lie algebra with Dynkin diagram Δ. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$ and

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

the Cartan decomposition of $\mathfrak{g}$ with respect to $\mathfrak{h}$. We know form Theorem 4.19 that $\dim \mathfrak{g}_\alpha = 1$ for all $\alpha \in \Phi$. For each α, we choose a nonzero element $e_\alpha \in \mathfrak{g}_\alpha$. Let $\Pi = \{\alpha_1, \ldots, \alpha_l\}$ be a fundamental system of roots in Φ. Then the elements h'_{α_i} form a basis of $\mathfrak{h}$. We wish now to choose a more convenient basis consisting of scalar multiples of the h'_{α_i}. For $1 = 1, \ldots, l$, we define the elements $h_i \in \mathfrak{h}$ by

$$h_i = \frac{2h'_{\alpha_i}}{\langle h'_{\alpha_i}, h'_{\alpha_i} \rangle}.$$

Notice that $\alpha_i(h_i) = 2$ for all i. We see that $\{h_i, \quad i = 1, \ldots, l\,;\, e_\alpha, \quad \alpha \in \Phi\}$ is a basis of $\mathfrak{g}$. We know from Proposition 4.17 that $h'_\alpha \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ for all $\alpha \in \Phi$. Thus, having chosen the elements e_α for all $\alpha \in \Phi^+$, we may choose $e_{-\alpha}$ for all $\alpha \in \Phi^+$ to satisfy

$$[e_\alpha, e_{-\alpha}] = \frac{2h'_\alpha}{\langle h'_\alpha, h'_\alpha \rangle}.$$

This relation will then be automatically satisfied for all $\alpha \in \Phi^-$.

For each $\alpha \in \Phi$, we define the element $h_\alpha \in \mathfrak{h}$ by

$$h_\alpha = \frac{2h'_\alpha}{\langle h'_\alpha, h'_\alpha \rangle}.$$

The element h_α is called the **coroot** corresponding to the root α. In particular, we have $h_i = h_{\alpha_i}$. By construction, we have

$$[e_\alpha, e_{-\alpha}] = h_\alpha \qquad \text{for all } \alpha \in \Phi.$$

We now consider the product $[e_\alpha, e_\beta]$ when $\alpha \neq \beta$. We know that $[e_\alpha, e_\beta] = 0$ if $\alpha + \beta \neq 0$ and $\alpha + \beta \notin \Phi$. If $\alpha + \beta \in \Phi$, then $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$. We define the elements $N_{\alpha,\beta} \in \mathbb{C}$ by the relation

$$[e_\alpha, e_\beta] = N_{\alpha,\beta} e_{\alpha+\beta}.$$

The numbers $N_{\alpha,\beta}$ for $\alpha, \beta, \alpha + \beta \in \Phi$ are called **structure constants** of $\mathfrak{g}$. Notice that the structure constants of $\mathfrak{g}$ depend on our choice of elements $e_\alpha \in \mathfrak{g}_\alpha$.

Next, we consider the multiplication of the basis vectors $\{h_i, e_\alpha\}$ of $\mathfrak{g}$. We have

$$\begin{aligned}
&[h_i, h_j] = 0 \\
&[h_i, e_\alpha] = \alpha(h_i) e_\alpha \\
&[e_\alpha, e_{-\alpha}] = h_\alpha \\
&[e_\alpha, e_\beta] = N_{\alpha,\beta} e_{\alpha+\beta} \quad && \text{if } \alpha, \beta, \alpha + \beta \in \Phi \\
&[e_\alpha, e_\beta] = 0 && \text{if } \alpha + \beta \neq 0 \text{ and } \alpha + \beta \notin \Phi.
\end{aligned}$$

In order to express $[e_\alpha, e_{-\alpha}]$ as a linear combination of the basis elements, we simply express h_α as a linear combination of the h_i.

We now derive some relations among the structure constants $N_{\alpha,\beta}$.

Proposition 7.1. *The structure constants $N_{\alpha,\beta}$ satisfy the following relations:*

(i) $N_{\beta,\alpha} = -N_{\alpha,\beta}$.

(ii) *If $\alpha, \beta, \gamma \in \Phi$ satisfy $\alpha + \beta + \gamma = 0$, then*

$$\frac{N_{\alpha,\beta}}{\langle \gamma, \gamma \rangle} = \frac{N_{\beta,\gamma}}{\langle \alpha, \alpha \rangle} = \frac{N_{\gamma,\alpha}}{\langle \beta, \beta \rangle}.$$

(iii) $N_{\alpha,\beta} N_{-\alpha,-\beta} = -(p+1)^2$ *where $-p\alpha + \beta, \ldots, \beta, \ldots, q\alpha + \beta$ is the α-chain of roots through β.*

(iv) *If $\alpha, \beta, \gamma, \delta \in \Phi$ satisfy $\alpha + \beta + \gamma + \delta = 0$, and no pair are negatives of one another, then*

$$\frac{N_{\alpha,\beta} N_{\gamma,\delta}}{\langle \alpha + \beta, \alpha + \beta \rangle} + \frac{N_{\beta,\gamma} N_{\alpha,\delta}}{\langle \beta + \gamma, \beta + \gamma \rangle} + \frac{N_{\gamma,\alpha} N_{\beta,\delta}}{\langle \gamma + \alpha, \gamma + \alpha \rangle} = 0.$$

Proof. Part (i) follows immediately from the definition. To prove part (ii), suppose $\alpha + \beta + \gamma = 0$. We consider the Jacobi identity

$$[[e_\alpha, e_\beta], e_\gamma] + [[e_\beta, e_\gamma], e_\alpha] + [[e_\gamma, e_\alpha], e_\beta] = 0.$$

This yields

$$N_{\alpha,\beta}[e_{\alpha+\beta}, e_{-(\alpha+\beta)}] + N_{\beta,\gamma}[e_{-\alpha}, e_\alpha] + N_{\gamma,\alpha}[e_{-\beta}, e_\beta] = 0,$$

that is,

$$2N_{\alpha,\beta} \frac{h'_{\alpha+\beta}}{\langle h'_{\alpha+\beta}, h'_{\alpha+\beta} \rangle} = 2N_{\beta,\gamma} \frac{h'_\alpha}{\langle h'_\alpha, h'_\alpha \rangle} + 2N_{\gamma,\alpha} \frac{h'_\beta}{\langle h'_\beta, h'_\beta \rangle}.$$

Now the roots α, β are linearly independent, for if they were not, then $\alpha + \beta = -\gamma$ would fail to be a root. Thus h'_α, h'_β are linearly independent, and $h'_{\alpha+\beta} = h'_\alpha + h'_\beta$. It follows that

$$\frac{N_{\alpha,\beta}}{\langle h'_{\alpha+\beta}, h'_{\alpha+\beta} \rangle} = \frac{N_{\beta,\gamma}}{\langle h'_\alpha, h'_\alpha \rangle} = \frac{N_{\gamma,\alpha}}{\langle h'_\beta, h'_\beta \rangle},$$

that is,

$$\frac{N_{\alpha,\beta}}{\langle\gamma,\gamma\rangle} = \frac{N_{\beta,\gamma}}{\langle\alpha,\alpha\rangle} = \frac{N_{\gamma,\alpha}}{\langle\beta,\beta\rangle}.$$

To prove part (iii), suppose $\alpha, \beta \in \Phi$ are linearly independent. Without loss of generality, we may assume that $|\alpha| \leq |\beta|$. Consider the Jacobi identity

$$[[e_\alpha, e_{-\alpha}], e_\beta] + [[e_{-\alpha}, e_\beta], e_\alpha] + [[e_\beta, e_\alpha], e_{-\alpha}] = 0.$$

This yields

$$2\frac{[h'_\alpha, e_\beta]}{\langle h'_\alpha, h'_\alpha\rangle} + N_{-\alpha,\beta}N_{-\alpha+\beta,\alpha}e_\beta + N_{\beta,\alpha}N_{\alpha+\beta,-\alpha}e_\beta = 0.$$

Since $[h'_\alpha, e_\beta] = \beta(h'_\alpha)\, e_\beta = \langle h'_\alpha, h'_\beta\rangle e_\beta$ and e_β is not the zero vector, it follows that

$$2\frac{\langle h'_\alpha, h'_\beta\rangle}{\langle h'_\alpha, h'_\alpha\rangle} + N_{-\alpha,\beta}N_{-\alpha+\beta,\alpha} + N_{\beta,\alpha}N_{\alpha+\beta,-\alpha} = 0.$$

Rearranging terms and using part (i), this expression becomes

$$N_{\alpha,\beta}N_{\alpha+\beta,-\alpha} + N_{-\alpha,\beta}N_{\alpha,-\alpha+\beta} = 2\frac{\langle\alpha,\beta\rangle}{\langle\alpha,\alpha\rangle}. \tag{1}$$

From part (ii), we can derive the relations

$$\frac{N_{\alpha+\beta,-\alpha}}{\langle\beta,\beta\rangle} = \frac{N_{-\alpha,-\beta}}{\langle\alpha+\beta,\alpha+\beta\rangle}, \qquad \frac{N_{-\alpha,\beta}}{\langle-\alpha+\beta,-\alpha+\beta\rangle} = \frac{N_{\alpha-\beta,-\alpha}}{\langle\beta,\beta\rangle}.$$

Plugging these relations into (1) gives us

$$N_{\alpha,\beta}N_{-\alpha,-\beta}\frac{\langle\beta,\beta\rangle}{\langle\alpha+\beta,\alpha+\beta\rangle} - N_{\alpha,-\alpha+\beta}N_{-\alpha,\alpha-\beta}\frac{\langle-\alpha+\beta,-\alpha+\beta\rangle}{\langle\beta,\beta\rangle} = 2\frac{\langle\alpha,\beta\rangle}{\langle\alpha,\alpha\rangle}. \tag{2}$$

(Note: If $-\alpha+\beta$ is not a root, then $N_{-\alpha,\beta}$ is interpreted as 0, so the second term on the left-hand side of (1) disappears.) We next consider the α-chain of roots through β,

$$-p\alpha+\beta, \ldots, \beta, \ldots, q\alpha+\beta.$$

We plug the pairs $(\alpha,\beta), (\alpha,-\alpha+\beta), \ldots, (\alpha,-p\alpha+\beta)$ into (2) to get the system of equations

$$\begin{aligned}
N_{\alpha,\beta}N_{-\alpha,-\beta}\frac{\langle\beta,\beta\rangle}{\langle\alpha+\beta,\alpha+\beta\rangle} - N_{\alpha,-\alpha+\beta}N_{-\alpha,\alpha-\beta}\frac{\langle-\alpha+\beta,-\alpha+\beta\rangle}{\langle\beta,\beta\rangle} &= 2\frac{\langle\alpha,\beta\rangle}{\langle\alpha,\alpha\rangle}\\
N_{\alpha,-\alpha+\beta}N_{-\alpha,\alpha-\beta}\frac{\langle-\alpha+\beta,-\alpha+\beta\rangle}{\langle\beta,\beta\rangle} - N_{\alpha,-2\alpha+\beta}N_{-\alpha,2\alpha-\beta}\frac{\langle-2\alpha+\beta,-2\alpha+\beta\rangle}{\langle-\alpha+\beta,-\alpha+\beta\rangle} &= 2\frac{\langle\alpha,-\alpha+\beta\rangle}{\alpha,\alpha}\\
&\vdots\\
N_{\alpha,-p\alpha+\beta}N_{-\alpha,p\alpha-\beta}\frac{\langle-p\alpha+\beta,-p\alpha+\beta\rangle}{\langle-(p-1)\alpha+\beta,-(p-1)\alpha+\beta\rangle} &= 2\frac{\langle\alpha,-p\alpha+\beta\rangle}{\langle\alpha,\alpha\rangle}.
\end{aligned}$$

(The last equation has only one term on the left since $-(p+1)\alpha+\beta$ is not a root.) Adding these equations and using the bilinearity of $\langle , \rangle$, we get

$$N_{\alpha,\beta}N_{-\alpha,-\beta}\frac{\langle\beta,\beta\rangle}{\langle\alpha+\beta,\alpha+\beta\rangle} = 2(p+1)\frac{\langle\alpha,\beta\rangle}{\langle\alpha,\alpha\rangle} - 2\frac{p(p+1)}{2}.$$

But we know from Proposition 4.21 that $2\frac{\langle\alpha,\beta\rangle}{\langle\alpha,\alpha\rangle} = p-q$. It follows that

$$N_{\alpha,\beta}N_{-\alpha,-\beta}\frac{\langle\beta,\beta\rangle}{\langle\alpha+\beta,\alpha+\beta\rangle} = -(p+1)q.$$

In order to obtain the desired result, $N_{\alpha,\beta}N_{-\alpha,-\beta} = -(p+1)^2$, we must show that

$$\frac{\langle\alpha+\beta,\alpha+\beta\rangle}{\langle\beta,\beta\rangle} = \frac{p+1}{q}.$$

Let $\gamma = -p\alpha+\beta$, the initial root in the α-chain through β. Let r, s be non-negative integers such that $-r\alpha+\gamma, \dots, \gamma, \dots, s\alpha+\gamma$ are all roots but $-(r+1)\alpha+\gamma$ and $(s+1)\alpha+\gamma$ are not. Recall from the proof of Proposition 6.1 that

$$2\frac{\langle\alpha,\gamma\rangle}{\langle\alpha,\alpha\rangle}\cdot 2\frac{\langle\gamma,\alpha\rangle}{\langle\gamma,\gamma\rangle} = 4\cos^2\theta$$

where θ is the angle between α, γ, and hence that $2\frac{\langle\alpha,\gamma\rangle}{\langle\alpha,\alpha\rangle} \in \{0, \pm1, \pm2, \pm3\}$. We also have $2\frac{\langle\alpha,\gamma\rangle}{\langle\alpha,\alpha\rangle} = r-s$ by Proposition 4.21. But γ is the initial root in the α-chain, so $r=0$, and hence $s = p+q \leq 3$. This shows that any α-chain contains at most four roots. Thus the possible positions of β in its α-chain are

β	$\alpha+\beta$			$p=0$	$q=1$
β	$\alpha+\beta$	$-2\alpha+\beta$		$p=0$	$q=2$
$-\alpha+\beta$	β	$\alpha+\beta$		$p=1$	$q=1$
β	$\alpha+\beta$	$-2\alpha+\beta$	$-3\alpha+\beta$	$p=0$	$q=3$
$-\alpha+\beta$	β	$\alpha+\beta$	$-2\alpha+\beta$	$p=1$	$q=2$
$-2\alpha+\beta$	$-\alpha+\beta$	β	$\alpha+\beta$	$p=2$	$q=1$

Since $|\alpha| \leq |\beta|$, we can compute the ratio of $\langle \alpha+\beta, \alpha+\beta \rangle$ to $\langle \beta, \beta \rangle$ in each case using the proof of Proposition 6.1. The only case that is not straightforward is that in which $p = q = 1$. In this case, we have

$$2\frac{\langle \alpha, \alpha+\beta \rangle}{\langle \alpha, \alpha \rangle} = 2.$$

We see from Proposition 6.1 that this is possible only when $\langle \alpha+\beta, \alpha+\beta \rangle = 2\langle \alpha, \alpha \rangle$. But we also have $\langle \alpha+\beta, \alpha+\beta \rangle = \langle \alpha, \alpha \rangle + \langle \beta, \beta \rangle$ since $\langle \alpha, \beta \rangle = 0$. It follows that $\langle \alpha, \alpha \rangle = \langle \beta, \beta \rangle$ and $\langle \alpha+\beta, \alpha+\beta \rangle = 2\langle \beta, \beta \rangle$. The remaining computations give us

$$\frac{\langle \alpha+\beta, \alpha+\beta \rangle}{\langle \beta, \beta \rangle} = 1, \tfrac{1}{2}, 2, \tfrac{1}{3}, 1, 3$$

in our six cases, respectively. In each case, we see that

$$\frac{\langle \alpha+\beta, \alpha+\beta \rangle}{\langle \beta, \beta \rangle} = \frac{p+1}{q},$$

and so $N_{\alpha,\beta}N_{-\alpha,-\beta} = -(p+1)^2$.

To prove part (iv), we assume that $\alpha, \beta, \gamma, \delta \in \Phi$ satisfy $\alpha+\beta+\gamma+\delta = 0$ with no pair being negatives. Consider the Jacobi identity

$$[[e_\alpha, e_\beta], e_\gamma] + [[e_\beta, e_\gamma], e_\alpha] + [[e_\gamma, e_\alpha], e_\beta] = 0.$$

This yields

$$N_{\alpha,\beta}N_{\alpha+\beta,\gamma} + N_{\beta,\gamma}N_{\beta+\gamma,\alpha} + N_{\gamma,\alpha}N_{\gamma+\alpha,\beta} = 0. \tag{1}$$

From part (ii), we have the relations

$$\frac{N_{\alpha+\beta,\gamma}}{\langle \delta, \delta \rangle} = \frac{N_{\gamma,\delta}}{\langle \alpha+\beta, \alpha+\beta \rangle}, \qquad \frac{N_{\beta+\gamma,\alpha}}{\langle \delta, \delta \rangle} = \frac{N_{\alpha,\delta}}{\langle \beta+\gamma, \beta+\gamma \rangle}, \qquad \frac{N_{\gamma+\alpha,\beta}}{\langle \delta, \delta \rangle} = \frac{N_{\beta,\delta}}{\langle \gamma+\alpha, \gamma+\alpha \rangle}.$$

Plugging these expressions into (1) gives us

$$\frac{N_{\alpha,\beta}N_{\gamma,\delta}}{\langle \alpha+\beta, \alpha+\beta \rangle} + \frac{N_{\beta,\gamma}N_{\alpha,\delta}}{\langle \beta+\gamma, \beta+\gamma \rangle} + \frac{N_{\gamma,\alpha}N_{\beta,\delta}}{\langle \gamma+\alpha, \gamma+\alpha \rangle} = 0.$$

(As usual, we interpret $N_{\theta,\phi}$ as 0 if $\theta+\phi$ is not a root.) □

Proposition 7.1(iii) has a very useful corollary.

Corollary 7.2. *If $\alpha, \beta, \alpha+\beta \in \Phi$, then $N_{\alpha,\beta} \neq 0$, that is, $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$.*

7.2 The uniqueness theorem

We now use these relations between the structure constants to prove that the Lie algebra $\mathfrak{g}$ is uniquely determined (up to isomorphism) by its Dynkin diagram. Given a Dynkin diagram on the standard list in Figure 6.3, this diagram uniquely determines a Cartan matrix $A = (A_{ij})$ on the standard list in Section 6.4. Each Cartan matrix, in turn, determines the set Φ of roots as linear combinations of the fundamental roots $\Pi = \{\alpha_1, \ldots, \alpha_l\}$. To see this, recall from Proposition 5.13 that each root $\alpha \in \Phi$ is of the form $\alpha = w(\alpha_i)$ for some $\alpha_i \in \Pi$, $w \in W$. Moreover, each element $w \in W$ is a product of the elements $s_1, \ldots, s_l$ by Theorem 5.14. The action of $s_1, \ldots, s_l$ on the fundamental roots $\alpha_1, \ldots, \alpha_l$ are given in terms of the Cartan matrix by

$$s_i(\alpha_j) = \alpha_j - A_{ij}\alpha_j.$$

Thus by applying the fundamental reflections successively to the fundamental roots, we obtain all roots as linear combinations of the fundamental roots.

Next, we show that the scalar products $\langle h'_\alpha, h'_\beta \rangle$ for $\alpha, \beta \in \Phi$ are uniquely determined by the Cartan matrix. By Proposition 4.21, the scalars $2\frac{\langle h'_\alpha, h'_\beta \rangle}{\langle h'_\alpha, h'_\alpha \rangle}$ are determined by the root system, and hence by the Cartan matrix as illustrated above. Then $\langle h'_\alpha, h'_\alpha \rangle$ is determined by the formula

$$\frac{1}{\langle h'_\alpha, h'_\alpha \rangle} = \sum_{\beta \in \Phi} \left(\frac{\langle h'_\alpha, h'_\beta \rangle}{\langle h'_\alpha, h'_\alpha \rangle} \right)^2$$

as in the proof of Corollary 4.23. It follows that $\langle h'_\alpha, h'_\beta \rangle$ is also determined by the Cartan matrix. Thus we see that if the structure constants $N_{\alpha,\beta}$ are known, then the multiplication of the basis elements

$$[h_i, h_j] = 0$$

$$[h_i, e_\alpha] = \alpha(h_i)e_\alpha$$

$$[e_\alpha, e_{-\alpha}] = h_\alpha$$

$$[e_\alpha, e_\beta] = N_{\alpha,\beta} e_{\alpha+\beta} \quad \text{if } \alpha, \beta, \alpha+\beta \in \Phi$$

$$[e_\alpha, e_\beta] = 0 \qquad\qquad \text{if } \alpha+\beta \neq 0 \text{ and } \alpha+\beta \notin \Phi$$

will be completely determined by the Cartan matrix.

Our next goal is to show that for certain pairs (α, β) of roots, the structure constants $N_{\alpha,\beta}$ can be arbitrarily chosen, and that the remaining structure constants are uniquely determined in terms of these by the relations in Proposition 7.1.

We choose a total ordering $\prec$ on the vector space $V = \mathfrak{h}^*_{\mathbb{R}}$ as in Section 5.1. This ordering will give rise to a positive system Φ^+ and fundamental system Π^+ of roots. An ordered pair (α, β) of roots is called **special** if $\alpha + \beta \in \Phi$ and $0 \prec \alpha \prec \beta$. A pair (α, β) is called **extraspecial** if (α, β) is special and, in addition, if for all pairs (γ, δ) such that $\alpha + \beta = \gamma + \delta$, we have $\alpha \preceq \gamma$.

Lemma 7.3. *The structure constants $N_{\alpha,\beta}$ for extraspecial pairs (α, β) can be chosen as arbitrary nonzero elements of $\mathbb{C}$ by appropriate choice of the elements e_α.*

Proof. We begin by choosing the e_α for $\alpha \in \Phi^+$ in the order given by $\prec$. Suppose the pair of roots (α, β) is extraspecial. Then we have

$$[e_\alpha, e_\beta] = N_{\alpha,\beta} e_{\alpha+\beta}$$

where e_α, e_β have already been chosen. Now there is only one extraspecial pair giving rise to the sum $\alpha + \beta$. Thus $e_{\alpha+\beta}$ can be chosen to give any nonzero value of $N_{\alpha,\beta}$. □

Proposition 7.4. *All of the structure constants $N_{\alpha,\beta}$ are determined by the structure constants for extraspecial pairs.*

Proof. We consider the set of all pairs of roots (α, β) such that $\alpha + \beta$ is a root. Let (α, β) be such a pair and let $\gamma = -\alpha - \beta$. One easily checks that the following twelve pairs are of the given type:

$$(\alpha, \beta),\ (\beta, \gamma),\ (\gamma, \alpha),\ (\beta, \alpha),\ (\gamma, \beta),\ (\alpha, \gamma)$$

$$(-\alpha, -\beta),\ (-\beta, -\gamma),\ (-\gamma, -\alpha),\ (-\beta, -\alpha),\ (-\gamma, -\beta),\ (-\alpha, -\gamma).$$

Since $\alpha+\beta+\gamma=0$, we see that either two or one of α,β,γ are positive. Hence either two of α,β,γ are positive or two of $-\alpha,-\beta,-\gamma$ are positive. By choosing two positive roots from α,β,γ or from $-\alpha,-\beta,-\gamma$ and writing them in an appropriate order, we obtain a special pair. Thus only one of the above twelve pairs is a special pair.

Now the relations in Proposition 7.1 (i), (ii), (iii) enable us to express $N_{\beta,\alpha}$, $N_{\beta,\gamma}$, $N_{\gamma,\alpha}$, and $N_{-\alpha,-\beta}$ in terms of $N_{\alpha,\beta}$. Thus, regardless of which pair (θ,ϕ) on the above list is special, we can express all of the structure constants corresponding to these pairs in terms of the $N_{\theta,\phi}$.

Next, we show that the $N_{\alpha,\beta}$ for all special pairs (α,β) can be written in terms of the $N_{\alpha,\beta}$ for extraspecial pairs. Suppose (α,β) is special but not extraspecial. Then there exists an extraspecial pair (γ,δ) such that $\alpha+\beta=\gamma+\delta$. Then

$$\alpha+\beta+(-\gamma)+(-\delta)=0$$

and no pair of $\alpha,\beta,-\gamma,-\delta$ are negatives. By Proposition 7.1(iv), we have

$$\frac{N_{\alpha,\beta}N_{-\gamma,-\delta}}{\langle\alpha+\beta,\alpha+\beta\rangle}+\frac{N_{\beta,-\gamma}N_{\alpha,-\delta}}{\langle\beta-\gamma,\beta-\gamma\rangle}+\frac{N_{-\gamma,\alpha}N_{\beta,-\delta}}{\langle-\gamma+\alpha,-\gamma+\alpha\rangle}=0.$$

Now the roots $\alpha,\beta,\gamma,\delta$ are ordered by

$$0\prec\gamma\prec\alpha\prec\beta\prec\delta.$$

Thus we may use relations (i), (ii), and (iii) from Proposition 7.1 to express $N_{-\gamma,-\delta}$ in terms of $N_{\gamma,\delta}$, $N_{\beta,-\gamma}$ in terms of $N_{\gamma,\beta-\gamma}$, $N_{\alpha,-\delta}$ in terms of $N_{\alpha,\delta-\alpha}$, $N_{-\gamma,\alpha}$ in terms of $N_{\gamma,\alpha-\gamma}$, and $N_{\beta,-\delta}$ in terms of $N_{\beta,\delta-\beta}$. Thus $N_{\alpha,\beta}$ is expressed in terms of

$$N_{\gamma,\delta},\ N_{\gamma,\beta-\gamma},\ N_{\alpha,\delta-\alpha},\ N_{\gamma,\alpha-\gamma},\ N_{\beta,\delta-\beta}.$$

Now (γ,δ) is an extraspecial pair and $(\gamma,\beta-\gamma)$, $(\alpha,\delta-\alpha)$, $(\gamma,\alpha-\gamma)$, and $(\beta,\delta-\beta)$ are all pairs of positive roots whose sums are roots less than $\alpha+\beta=\gamma+\delta$ in the given ordering. We may therefore argue by induction on $\alpha+\beta$ that $N_{\alpha,\beta}$ can be expressed in terms of $N_{\theta,\phi}$ for extraspecial pairs (θ,ϕ). □

We are now ready to prove the uniqueness theorem.

Theorem 7.5. *If two simple Lie algebras have the same Cartan matrix, then they are isomorphic.*

Proof. By Lemma 7.3, we may choose the basis elements $\{h_i, e_\alpha\}$ of such a Lie algebra $\mathfrak{g}$ such that $N_{\alpha,\beta} = 1$ for all extraspecial pairs of roots (α, β). The remaining structure constants $N_{\alpha,\beta}$ are then uniquely determined by Proposition 7.4. Thus the formulae expressing a Lie product of basis elements as a linear combination of basis elements are completely determined by the Cartan matrix. It follows that the Lie algebra $\mathfrak{g}$ is uniquely determined up to isomorphism. □

7.3 Some generators and relations in a simple Lie algebra

We now turn to the question of the existence of a Lie algebra having a Cartan matrix on the standard list in Section 6.4. Let $\mathfrak{g}$ be a simple Lie algebra with Cartan matrix A. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$ and

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

be the Cartan decomposition of $\mathfrak{h}$ with respect to $\mathfrak{h}$. As before, we consider the coroots $h_i \in \mathfrak{h}$ given by

$$h_i = \frac{2h'_{\alpha_i}}{\langle h'_{\alpha_i}, h'_{\alpha_i} \rangle}$$

where $\Pi = \{\alpha_1, \dots, \alpha_l\}$ is a fundamental system in Φ. For each i, we can choose elements $e_i \in \mathfrak{g}_{\alpha_i}$, $f_i \in \mathfrak{g}_{-\alpha_i}$ such that $[e_i, f_i] = h_i$ as in Section 7.1.

We will show that the elements $e_1, \dots, e_l, f_1, \dots, f_l, h_1, \dots, h_l$ generate $\mathfrak{g}$. (This is equivalent to saying that $e_1, \dots, e_l, f_1, \dots, f_l$ generate $\mathfrak{g}$, but it will be useful to include $h_1, \dots, h_l$ in the generating set.)

Lemma 7.6. *If $\alpha \in \Phi^+$ and $\alpha \notin \Pi$, then there exists an $\alpha_i \in \Pi$ such that $\alpha - \alpha_i \in \Phi^+$. Thus every non-fundamental root is the sum of a fundamental root with a positive root.*

Proof. Suppose the statement is false. Then $\alpha - \alpha_i$ is not a root and is nonzero for each i. (We can use Corollary 5.6 to show that $\alpha - \alpha_i$ is not a negative root.) For each i, consider the α_i-chain of roots through α. This is of the form

$$\alpha, \alpha_i + \alpha, \ldots, q\alpha_i + \alpha.$$

By Proposition 4.21, we have

$$2\frac{\langle \alpha_i, \alpha \rangle}{\langle \alpha_i, \alpha_i \rangle} = -q.$$

It follows that $\langle \alpha_i, \alpha \rangle \leq 0$. Now $\alpha \in \Phi^+$ has the form $\alpha = \sum_i n_i \alpha_i$ where $n_i \geq 0$ for all i. Thus

$$\langle \alpha, \alpha \rangle = \sum_i n_i \langle \alpha_i, \alpha \rangle \leq 0.$$

This is a contradiction since we know that $\langle \alpha, \alpha \rangle > 0$. □

Proposition 7.7. *The elements* $e_1, \ldots, e_l, f_1, \ldots, f_l, h_1, \ldots, h_l$ *generate* $\mathfrak{g}$.

Proof. Since $h_1, \ldots, h_l$ generate $\mathfrak{h}$, it suffices to show that each $\mathfrak{g}_\alpha$ for $\alpha \in \Phi^+$ is contained in the subalgebra generated by $e_1, \ldots, e_l$, and that each $\mathfrak{g}_\alpha$ for $\alpha \in \Phi^-$ is contained in the subalgebra generated by $f_1, \ldots, f_l$.

Let $\alpha \in \phi^+$. If $\alpha = \alpha_i$ for some i, then we have $\mathfrak{g}_\alpha = \mathbb{C}e_\alpha$. If $\alpha \notin \Pi$, then we may write $\alpha = \alpha_i + \beta$ for some $\alpha_i \in \Pi$, $\beta \in \Phi^+$ by Lemma 7.6. We then have $[\mathfrak{g}_{\alpha_i}, \mathfrak{g}_\beta] = \mathfrak{g}_\alpha$ by Corollary 7.2. We may thus choose $e_\alpha = [e_i, e_\beta]$ for some nonzero $e_\beta \in \mathfrak{g}_\beta$. By repeating this process, we obtain

$$e_\alpha = [[e_{i_1}, e_{i_2}], \cdots e_{i_k}]$$

for some sequence $i_1, \ldots, i_k$. Thus each $\mathfrak{g}_\alpha$ for $\alpha \in \Phi^+$ is contained in the subalgebra generated by $e_1, \ldots, e_l$. A similar argument proves that each $\mathfrak{g}_\alpha$ for $\alpha \in \Phi^-$ is contained in the subalgebra generated by $f_1, \ldots, f_l$. □

Proposition 7.8. *The generators* $e_1, \ldots, e_l, f_1, \ldots, f_l, h_1, \ldots, h_l$ *of* $\mathfrak{g}$ *satisfy the following relations:*

(a) $[h_i, h_j] = 0$

(b) $[h_i, e_j] = A_{ij} e_j$

(c) $[h_i, f_j] = -A_{ij} f_j$

(d) $[e_i, f_i] = h_i$

(e) $[e_i, f_j] = 0 \quad \textit{if } i \neq j$

(f) $\underbrace{[e_i, [e_i, \cdots [e_i,}_{1-A_{ij} \text{ times}} e_j] \cdots]] = 0 \quad \textit{if } i \neq j$

(g) $\underbrace{[f_i, [f_i, \cdots [f_i,}_{1-A_{ij} \text{ times}} f_j] \cdots]] = 0 \quad \textit{if } i \neq j$

Note that since $A_{ij} \leq 0$ for all $i \neq j$, the number $1 - A_{ij}$ is a positive integer.

Proof. Relation (a) follows from the fact that $[\mathfrak{h}, \mathfrak{h}] = 0$. For relation (b), we have

$$[h_i, e_j] = 2\frac{[h'_{\alpha_i}, e_j]}{\langle h'_{\alpha_i}, h'_{\alpha_i} \rangle} = 2\frac{\alpha_j\left(h'_{\alpha_i}\right)}{\langle h'_{\alpha_i}, h'_{\alpha_i} \rangle} e_j = 2\frac{\langle h'_{\alpha_i}, h'_{\alpha_j} \rangle}{\langle h'_{\alpha_i}, h'_{\alpha_i} \rangle} e_j = 2\frac{\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_i, \alpha_i \rangle} e_j = A_{ij} e_j.$$

A similar argument proves relation (c). Relation (d) is just the definition of e_i, f_i. Relation (e) holds because $[e_i, f_i] \in \mathfrak{g}_{\alpha_i - \alpha_j}$ and $\alpha_i - \alpha_j$ is not a root when $i \neq j$, as follows from Corollary 5.6. To prove relation (f), we consider the α_i-chain of roots through α_j. Since $-\alpha_i + \alpha_j$ is not a root, this chain is of the form

$$\alpha_j, \alpha_i + \alpha_j, \ldots, q\alpha_i + \alpha_j.$$

By Proposition 4.21, we have $A_{ij} = -q$. Thus $(1 - A_{ij})\alpha_i + \alpha_j$ is not a root. Since the element $[e_i, [e_i, \cdots [e_i, e_j]]]$ lies in $\mathfrak{g}_{(1-A_{ij})\alpha_i + \alpha_j}$, this element must be 0. A similar argument proves relation (g). □

7.4 The Lie algebras $\mathfrak{g}(A)$ and $\tilde{\mathfrak{g}}(A)$

Let A be a Cartan matrix on the standard list in Section 6.4. Motivated by Propositions 7.7 and 7.8, we now seek to construct a Lie algebra $\mathfrak{g}(A)$ that will be shown to be a finite-dimensional simple Lie algebra with Cartan matrix A.

Suppose A is an $l \times l$ matrix. Let $\mathfrak{F}$ be the free associative algebra over $\mathbb{C}$ on the $3l$ generators $e_1, \ldots, e_l, h_1, \ldots, h_l, f_1, \ldots, f_l$. The set of all monomials in these generators form a basis of $\mathfrak{F}$. Let $[\mathfrak{F}]$ be the Lie algebra formed from $\mathfrak{F}$ via the commutator product, and let $\mathfrak{L}$ be the subalgebra of $[\mathfrak{F}]$ generated by the elements $e_1, \ldots, e_l, h_1, \ldots, h_l, f_1, \ldots, f_l$. Let $\mathfrak{v}$ be the ideal of $\mathfrak{L}$ generated by the elements

$$[h_i, h_j]$$

$$[h_i, e_j] - A_{ij} e_j$$

$$[h_i, f_j] + A_{ij} f_j$$

$$[e_i, f_i] - h_i$$

$$[e_i, f_j] \quad \text{for } i \neq j$$

$$[e_i, [e_i, \cdots [e_i, e_j]]] \quad \text{for } i \neq j$$

$$[f_i, [f_i, \cdots [f_i, f_j]]] \quad \text{for } i \neq j$$

where the number of occurrences of e_i, f_i in the last two elements, respectively, is $1 - A_{ij}$.

We define $\mathfrak{g}(A) = \mathfrak{L}/\mathfrak{v}$. We will eventually be able to show that $\mathfrak{g}(A)$ is the Lie algebra we require to prove the existence theorem. This construction of $\mathfrak{g}(A)$ by generators and relations is due to J.P. Serre.

In order to more easily study the Lie algebra $\mathfrak{g}(A)$, we define a second, larger Lie algebra $\tilde{\mathfrak{g}}(A)$. Let $\tilde{\mathfrak{v}}$ be the ideal of $\mathfrak{L}$ generated by the elements

$$[h_i, h_j]$$

$$[h_i, e_j] - A_{ij} e_j$$

$$[h_i, f_j] + A_{ij} f_j$$

$$[e_i, f_i] - h_i$$

$$[e_i, f_j] \quad \text{for } i \neq j.$$

Let $\tilde{\mathfrak{g}}(A) = \mathfrak{L}/\tilde{\mathfrak{v}}$. Since $\tilde{\mathfrak{v}} \subset \mathfrak{v}$, we have surjective Lie algebra homomorphisms

$$\mathfrak{L} \to \tilde{\mathfrak{g}}(A) \to \mathfrak{g}(A).$$

We will investigate the properties of the Lie algebra $\tilde{\mathfrak{g}}(A)$. This is generated by the images of the generators of $\mathfrak{L}$ under the above homomorphism. For convenience, we continue to write these images as $e_1, \ldots, e_l, h_1, \ldots, h_l, f_1, \ldots, f_l$. These elements satisfy the relations

$$[h_i, h_j] = 0$$

$$[h_i, e_j] = A_{ij} e_j$$

$$[h_i, f_j] = -A_{ij} f_j$$

$$[e_i, f_i] = h_i$$

$$[e_i, f_j] = 0 \quad \text{for } i \neq j.$$

Proposition 7.9. *Let $\mathfrak{F}^-$ be the free associative algebra over $\mathbb{C}$ with generators $f_1, \ldots, f_l$. Then $\mathfrak{F}^-$ may be made into a $\tilde{\mathfrak{g}}(A)$-module inducing a representation $\rho : \tilde{\mathfrak{g}}(A) \to [\mathrm{End}\mathfrak{F}^-]$ defined by*

$$\rho(f_i) f_{i_1} \cdots f_{i_r} = f_i f_{i_1} \cdots f_{i_r}$$

$$\rho(h_i) f_{i_1} \cdots f_{i_r} = -\left(\sum_{k=1}^{r} A_{i i_k} \right) f_{i_1} \cdots f_{i_r}$$

$$\rho(e_i) f_{i_1} \cdots f_{i_r} = -\sum_{k=1}^{r} \delta_{i i_k} \left(\sum_{h=k+1}^{r} A_{i i_h} \right) f_{i_1} \cdots \hat{f}_{i_k} \cdots f_{i_r}$$

where, as usual, the symbol $\hat{f}_{i_k}$ means that f_{i_k} is omitted from the product.

Proof. Since the monomials $f_{i_1} \cdots f_{i_r}$ form a basis of $\mathfrak{F}^-$, the endomorphisms $\rho(f_i)$, $\rho(h_i)$, $\rho(e_i)$ are uniquely determined by the above formulae. Thus there is a unique homomorphism $\mathfrak{F} \to \mathrm{End}\,\mathfrak{F}^-$ mapping e_i, h_i, f_i to $\rho(e_i), \rho(h_i), \rho(f_i)$, respectively. This induces a Lie algebra homomorphism $[\mathfrak{F}] \to [\mathrm{End}\,\mathfrak{F}^-]$ and hence, by restriction, a Lie algebra homomorphism $\mathfrak{L} \to [\mathrm{End}\,\mathfrak{F}^-]$. In order to obtain a Lie algebra homomorphism $\tilde{\mathfrak{g}}(A) \to [\mathrm{End}\,\mathfrak{F}^-]$, we must verify the following relations:

(a) $[\rho(h_i), \rho(h_j)] = 0$

(b) $[\rho(h_i), \rho(e_j)] = A_{ij} \rho(e_j)$

(c) $[\rho(h_i), \rho(f_j)] = -A_{ij}\rho(f_j)$

(d) $[\rho(e_i), \rho(f_i)] = \rho(h_i)$

(e) $[\rho(e_i), \rho(f_j)] = 0 \quad \text{if } i \neq j.$

Relation (a) is automatic since $\rho(h_i)$ multiplies each basis element of $\mathfrak{F}^-$ by a scalar.

To prove relation (b), we have

$$\rho(h_i)\rho(e_j)f_{i_1}\cdots f_{i_r} = -\sum_{k=1}^{r}\delta_{ji_k}\left(\sum_{h=k+1}^{r}A_{ji_h}\right)\left(-\sum_{g\neq k}A_{ii_g}\right)f_{i_1}\cdots\hat{f}_{i_k}\cdots f_{i_r},$$

$$\rho(e_j)\rho(h_i)f_{i_1}\cdots f_{i_r} = -\sum_{k=1}^{r}\delta_{ji_k}\left(\sum_{h=k+1}^{r}A_{ji_h}\right)\left(-\sum_{g=1}^{r}A_{ii_g}\right)f_{i_1}\cdots\hat{f}_{i_k}\cdots f_{i_r}.$$

Thus

$$(\rho(h_i)\rho(e_j)-\rho(e_j)\rho(h_i))f_{i_1}\cdots f_{i_r} = -\sum_{k=1}^{r}\delta_{ji_k}\left(\sum_{h=k+1}^{r}A_{ji_h}\right)A_{ii_k}f_{i_1}\cdots\hat{f}_{i_k}\cdots f_{i_r}$$

$$= A_{ij}\left(-\sum_{k=1}^{r}\delta_{ji_k}\left(\sum_{h=k+1}^{r}A_{ji_h}\right)\right)f_{i_1}\cdots\hat{f}_{i_k}\cdots f_{i_r}$$

$$= A_{ij}\rho(e_j)f_{i_1}\cdots f_{i_r}.$$

To prove relation (c), we have

$$\rho(h_i)\rho(f_j)f_{i_1}\cdots f_{i_r} = -\left(A_{ij}+\sum_{k=1}^{r}A_{ii_k}\right)f_jf_{i_1}\cdots f_{i_r},$$

$$\rho(f_j)\rho(h_i)f_{i_1}\cdots f_{i_r} = -\left(\sum_{k=1}^{r}A_{ii_k}\right)f_jf_{i_1}\cdots f_{i_r}.$$

Thus

$$(\rho(h_i)\rho(f_j)-\rho(f_j)\rho(h_i))f_{i_1}\cdots f_{i_r} = -A_{ij}f_jf_{i_1}\cdots f_{i_r} = -A_{ij}\rho(f_j)f_{i_1}\cdots f_{i_r}.$$

Next, we prove relation (d). We have

$$\rho(e_i)\rho(f_i)f_{i_1}\cdots f_{i_r} = -\left(\sum_{h=1}^{r}A_{ii_h}\right)f_{i_1}\cdots f_{i_r}$$
$$-\sum_{k=1}^{r}\delta_{ii_k}\left(\sum_{h=k+1}^{r}A_{ii_h}\right)f_if_{i_1}\cdots\hat{f}_{i_k}\cdots f_{i_r},$$

$$\rho(f_i)\rho(e_i)f_{i_1}\cdots f_{i_r} = -\sum_{k=1}^{r}\delta_{ii_k}\left(\sum_{h=k+1}^{r}A_{ii_h}\right)f_if_{i_1}\cdots\hat{f}_{i_k}\cdots f_{i_r}.$$

Thus

$$(\rho(e_i)\rho(f_i) - \rho(f_i)\rho(e_i))f_{i_1}\cdots f_{i_r} = -\left(\sum_{h=1}^{r} A_{ii_h}\right) f_i f_{i_1}\cdots f_{i_r} = \rho(h_i)f_{i_1}\cdots f_{i_r}.$$

Finally, we prove relation (e). Suppose $i \neq j$. Then

$$\rho(e_i)\rho(f_j)f_{i_1}\cdots f_{i_r} = -\sum_{k=1}^{r}\delta_{ii_k}\left(\sum_{h=k+1}^{r} A_{ii_h}\right) f_j f_{i_1}\cdots f_{i_r}$$
$$= \rho(f_j)\rho(e_i)f_{i_1}\cdots f_{i_r}.$$

Thus ρ preserves all relations, and so $\rho : \tilde{\mathfrak{g}}(A) \to [\mathrm{End}\,\mathfrak{F}^-]$ is a homomorphism of Lie algebras. □

This homomorphism will allow us to deduce some useful information about the Lie algebra $\tilde{\mathfrak{g}}(A)$.

Proposition 7.10. *The elements $h_1, \ldots, h_l$ of $\tilde{\mathfrak{g}}(A)$ are linearly independent.*

Proof. We will show that the elements $\rho(h_1), \ldots, \rho(h_l)$ of $\mathrm{End}\,\mathfrak{F}^-$ are linearly independent. We have

$$\rho(h_i)f_j = -A_{ij}f_j.$$

Thus if $\sum \lambda_i \rho(h_i) = 0$, then we have $\sum \lambda_i A_{ij} = 0$ for all $j = 1, \ldots, l$. Since the Cartan matrix $A = (A_{ij})$ is non-singular, This implies that $\lambda_i = 0$ for each i. Hence $\rho(h_1), \ldots, \rho(h_l)$ are linearly independent, and so $h_1, \ldots, h_l$ must be linearly independent, also. □

Let $\tilde{\mathfrak{h}}$ be the subspace of $\tilde{\mathfrak{g}}(A)$ spanned by $h_1, \ldots, h_l$. Then we have $\dim \tilde{\mathfrak{h}} = l$. We also have $[\tilde{\mathfrak{h}}, \tilde{\mathfrak{h}}] = 0$ by our relations, and so $\tilde{\mathfrak{h}}$ is an abelian subalgebra of $\tilde{\mathfrak{g}}(A)$. We consider the weight spaces of $\tilde{\mathfrak{g}}(A)$ with respect to $\tilde{\mathfrak{h}}$. We are no longer dealing with finite-dimensional $\tilde{\mathfrak{h}}$-modules as in Theorem 2.10, but analogous ideas apply in this situation. For each $\lambda \in \mathrm{Hom}(\tilde{\mathfrak{h}}, \mathbb{C})$, we define the set

$$\tilde{\mathfrak{g}}(A)_\lambda = \{x \in \tilde{\mathfrak{g}}(A) \mid [h, x] = \lambda(h)x \quad \text{for all } h \in \tilde{\mathfrak{h}}\}.$$

If $\tilde{\mathfrak{g}}(A)_\lambda \neq 0$, then we call λ a **weight**, and we call $\tilde{\mathfrak{g}}(A)_\lambda$ the **weight space** of λ.

Proposition 7.11. $\tilde{\mathfrak{g}}(A) = \bigoplus_\lambda \tilde{\mathfrak{g}}(A)_\lambda$. *Thus* $\tilde{\mathfrak{g}}(A)$ *is the direct sum of its weight spaces.*

Proof. We first show that $\tilde{\mathfrak{g}}(A) = \sum_\lambda \tilde{\mathfrak{g}}(A)_\lambda$. A vector that lies in a weight space will be called a weight vector. Let $x, y \in \tilde{\mathfrak{g}}(A)$ be weight vectors of weights λ, μ, respectively. Then since

$$[h,[x,y]] = [[h,x],y] + [x,[h,y]] = \lambda(h)[x,y] + \mu(h)[x,y]$$
$$= (\lambda+\mu)(h)[x,y] \qquad \text{for all } h \in \tilde{\mathfrak{h}},$$

we see that $[x, y]$ is a weight vector of weight $\lambda + \mu$.

Now $\tilde{\mathfrak{g}}(A)$ is generated by the elements e_i, h_i, f_i. Let $\alpha_i \in \mathrm{Hom}(\tilde{\mathfrak{h}}, \mathbb{C})$ be the map defined by

$$\alpha_i(h_j) = A_{ji}.$$

One easily checks that e_i is a weight vector of weight α_i, f_i is a weight vector of weight $-\alpha_i$, and h_i is a weight vector of weight 0. Thus all Lie products of the generators e_i, h_i, f_i are weight vectors. It follows that

$$\tilde{\mathfrak{g}}(A) = \sum_\lambda \tilde{\mathfrak{g}}(A)_\lambda.$$

We next show that this is a direct sum. Suppose this is not the case. Then there exists a λ and a nonzero $x \in \tilde{\mathfrak{g}}(A)_\lambda$ such that $x = \sum_\mu x_\mu$ where $x_\mu \in \tilde{\mathfrak{g}}(A)_\mu$ and μ runs over a finite set of weights, all of which are distinct from λ.

Now the vector space $\tilde{\mathfrak{h}}$ over the infinite field $\mathbb{C}$ cannot be expressed as the union of finitely-many proper subspaces. One easily checks that for each μ, the set of $h \in \tilde{\mathfrak{h}}$ satisfying $\mu(h) = \lambda(h)$ is a proper subspace. Thus there exists an $h \in \tilde{\mathfrak{h}}$ such that $\mu(h) \neq \lambda(h)$ for all μ. Now since $x \in \tilde{\mathfrak{g}}(A)_\lambda$, we have

$$(\mathrm{ad}\, h - \lambda(h)1)x = 0.$$

And since $x = \sum_\mu x_\mu$ with $x_\mu \in \tilde{\mathfrak{g}}(A)_\mu$, we have

$$\prod_\mu (\mathrm{ad}\, h - \mu(h)1)x = 0.$$

Because $\mu(h) \neq \lambda(h)$ for all μ, we see that the polynomials

$$t - \lambda(h), \qquad \prod_{\mu}(t - \mu(h))$$

are coprime. Thus there exist polynomials $a(t), b(t) \in \mathbb{C}[t]$ such that

$$a(t)(t - \lambda(h)) + b(t)\prod_{\mu}(t - \mu(h)) = 1.$$

It follows that

$$a(\operatorname{ad} h)(\operatorname{ad} h - \lambda(h)1)x + b(\operatorname{ad} h)\prod_{\mu}(\operatorname{ad} h - \mu(h)1)x = x.$$

The left-hand side of this equation is zero by the preceding results, and thus $x = 0$. This is a contradiction, so the sum $\tilde{\mathfrak{g}}(A) = \sum_{\lambda} \tilde{\mathfrak{g}}(A)_{\lambda}$ must be direct. □

We next obtain some useful information about the weights of $\tilde{\mathfrak{g}}(A)$. The weights $\alpha_1, \ldots, \alpha_l \in \operatorname{Hom}(\tilde{\mathfrak{h}}, \mathbb{C})$ are linearly independent since the Cartan matrix A is non-singular. Because $\operatorname{Hom}(\tilde{\mathfrak{h}}, \mathbb{C})$ has the same dimension as $\tilde{\mathfrak{h}}$, we see that $\alpha_1, \ldots, \alpha_l$ form a basis of $\operatorname{Hom}(\tilde{\mathfrak{h}}, \mathbb{C})$. Thus any weight is of the form $n_1\alpha_1 + \cdots + n_l\alpha_l$ with $n_i \in \mathbb{C}$. Our goal is to show that all weights of $\tilde{\mathfrak{g}}(A)$ have this form with $n_i \in \mathbb{Z}$ and $n_i \geq 0$ for all i or $n_i \leq 0$ for all i.

Let

$$\begin{aligned} Q &= \{n_1\alpha_1 + \cdots + n_l\alpha_l \mid n_i \in \mathbb{Z}\} \\ Q^+ &= \{n_1\alpha_1 + \cdots + n_l\alpha_l \mid n_i \geq 0 \quad \text{for all } i\} \\ Q^- &= \{n_1\alpha_1 + \cdots + n_l\alpha_l \mid n_i \leq 0 \quad \text{for all } i\}. \end{aligned}$$

Let

$$\begin{aligned} \tilde{\mathfrak{g}}(A)^+ &= \sum_{\lambda \in Q^+} \tilde{\mathfrak{g}}(A)_{\lambda} \\ \tilde{\mathfrak{g}}(A)^- &= \sum_{\lambda \in Q^-} \tilde{\mathfrak{g}}(A)_{\lambda}. \end{aligned}$$

It follows from Proposition 7.11 that the sum $\tilde{\mathfrak{g}}(A)^- + \tilde{\mathfrak{h}} + \tilde{\mathfrak{g}}(A)^+$ is direct. We will show that, in fact,

$$\tilde{\mathfrak{g}}(A) = \tilde{\mathfrak{g}}(A)^- \oplus \tilde{\mathfrak{h}} \oplus \tilde{\mathfrak{g}}(A)^+.$$

Let $\tilde{\mathfrak{n}}$ be the subalgebra of $\tilde{\mathfrak{g}}(A)$ generated by $e_1, \ldots, e_l$ and $\tilde{\mathfrak{n}}^-$ the subalgebra generated by $f_1, \ldots, f_l$. Since e_i has weight α_i and f_i has weight $-\alpha_i$, we see that $\tilde{\mathfrak{n}} \subset \tilde{\mathfrak{g}}(A)^+$ and $\tilde{\mathfrak{n}}^- \subset \tilde{\mathfrak{g}}(A)^-$. Thus the sum $\tilde{\mathfrak{n}} + \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}^-$ is direct.

Proposition 7.12.

(i) $\tilde{\mathfrak{g}}(A) = \tilde{\mathfrak{n}}^- \oplus \tilde{\mathfrak{h}} \oplus \tilde{\mathfrak{n}}$.

(ii) $\tilde{\mathfrak{n}} = \tilde{\mathfrak{g}}(A)^+$, $\tilde{\mathfrak{n}}^- = \tilde{\mathfrak{g}}(A)^-$, *and* $\tilde{\mathfrak{h}} = \tilde{\mathfrak{g}}(A)_0$.

(iii) *Every nonzero weight of* $\tilde{\mathfrak{g}}(A)$ *lies in* Q^+ *or in* Q^-.

Proof. We first show that $[h_i, [[e_{i_1}, e_{i_2}], \cdots, e_{i_r}]] \in \tilde{\mathfrak{n}}$ for all products $[[e_{i_1}, e_{i_2}], \cdots, e_{i_r}]$ using induction on r. The relations $[h_i, e_j] = A_{ij} e_j$ prove the statement for $r = 1$. Now let $r > 1$ and write $y = [[e_{i_1}, e_{i_2}], \cdots e_{i_r}]$, $x = [[e_{i_1}, e_{i_2}], \cdots e_{i_{r-1}}]$. We assume by induction that $[h_i, x] \in \tilde{\mathfrak{n}}$. Then

$$[h_i, y] = [h_i, [x, e_{i_r}]] = [[h_i, x], e_{i_r}] + [x, [h_i, e_{i_r}]]$$

by the Jacobi identity. The right-hand side is a sum of elements in $[\tilde{\mathfrak{n}}, \tilde{\mathfrak{n}}] \subset \tilde{\mathfrak{n}}$, and so $[h_i, y] \in \tilde{\mathfrak{n}}$. This completes the induction. It follows that $[\tilde{\mathfrak{h}}, \tilde{\mathfrak{n}}] \subset \tilde{\mathfrak{n}}$ since the elements h_i generate $\tilde{\mathfrak{h}}$ and $[\tilde{\mathfrak{h}}, \tilde{\mathfrak{h}}] = 0$. Having this result, we see that

$$[\tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}, \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}] = [\tilde{\mathfrak{h}}, \tilde{\mathfrak{h}}] + [\tilde{\mathfrak{h}}, \tilde{\mathfrak{n}}] + [\tilde{\mathfrak{n}}, \tilde{\mathfrak{n}}] \subset \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}.$$

Thus $\tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}$ is a subalgebra of $\tilde{\mathfrak{g}}(A)$. Similarly, $\tilde{\mathfrak{n}}^- + \tilde{\mathfrak{h}}$ is a subalgebra of $\tilde{\mathfrak{g}}(A)$.

Next, we show that $[e_i, [[f_{i_1}, f_{i_2}], \cdots f_{i_r}]] \in \tilde{\mathfrak{n}}^- + \tilde{\mathfrak{h}}$ for all products $[[f_{i_1}, f_{i_2}], \cdots f_{i_r}]$. The relations $[e_i, f_i] = h_i$ and $[e_i, f_j] = 0$ for $i \neq j$ prove the statement for the generators of $\tilde{\mathfrak{n}}^-$, and an inductive argument similar to that given above proves that it is true for all such products in $\tilde{\mathfrak{n}}$. It follows that $[e_i, \tilde{\mathfrak{n}}^-] \subset \tilde{\mathfrak{n}}^- + \tilde{\mathfrak{h}}$. Similarly, $[f_i, \tilde{\mathfrak{n}}] \subset \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}$. These relations give us

$$[e_i, \tilde{\mathfrak{n}}^- + \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}] \subset \tilde{\mathfrak{n}}^- + \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}$$

since $\tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}$ is a subalgebra of $\tilde{\mathfrak{g}}(A)$,

$$[f_i, \tilde{\mathfrak{n}}^- + \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}] \subset \tilde{\mathfrak{n}}^- + \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}$$

since $\tilde{\mathfrak{n}}^- + \tilde{\mathfrak{h}}$ is a subalgebra of $\tilde{\mathfrak{g}}(A)$, and

$$[h_i, \tilde{\mathfrak{n}}^- + \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}] \subset \tilde{\mathfrak{n}}^- + \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}$$

since $\tilde{\mathfrak{n}}^- + \tilde{\mathfrak{h}}$ and $\tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}$ are both subalgebras. It follows that the subspace consisting of all $x \in \tilde{\mathfrak{g}}(A)$ such that

$$[x, \tilde{\mathfrak{n}}^- + \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}] \subset \tilde{\mathfrak{n}}^- + \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}$$

contains the generators e_i, h_i, f_i of $\tilde{\mathfrak{g}}(A)$. However, the relation

$$[[x, y], z] = [[x, z], y] + [x, [y, z]]$$

for x, y in the subspace, $z \in \tilde{\mathfrak{n}}^- + \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}$ shows that this subspace is also a subalgebra. Since it contains the generators of $\tilde{\mathfrak{g}}(A)$, this subalgebra must be the whole of $\tilde{\mathfrak{g}}(A)$. Thus $\tilde{\mathfrak{n}}^- + \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}$ is an ideal of $\tilde{\mathfrak{g}}(A)$. Since $\tilde{\mathfrak{g}}(A)$ is generated by the elements e_i, h_i, f_i, it follows that $\tilde{\mathfrak{n}}^- + \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}} = \tilde{\mathfrak{g}}(A)$. We know this sum is direct, so we have

$$\tilde{\mathfrak{g}}(A) = \tilde{\mathfrak{n}}^- \oplus \tilde{\mathfrak{h}} \oplus \tilde{\mathfrak{n}}.$$

Because $\tilde{\mathfrak{n}}^- \subset \tilde{\mathfrak{g}}(A)^-$, $\tilde{\mathfrak{n}} \subset \tilde{\mathfrak{g}}(A)^+$, and the sum $\tilde{\mathfrak{g}}(A)^- + \tilde{\mathfrak{h}} + \tilde{\mathfrak{g}}(A)^+$ is direct, we deduce that $\tilde{\mathfrak{n}}^- = \tilde{\mathfrak{g}}(A)^-$ and $\tilde{\mathfrak{n}} = \tilde{\mathfrak{g}}(A)^+$. Moreover, since $\tilde{\mathfrak{g}}(A) = \tilde{\mathfrak{g}}(A)^- \oplus \tilde{\mathfrak{h}} \oplus \tilde{\mathfrak{g}}(A)^+$, $\tilde{\mathfrak{h}} \subset \tilde{\mathfrak{g}}(A)_0$, and the weights occurring in $\tilde{\mathfrak{g}}(A)^-$ and $\tilde{\mathfrak{g}}(A)^+$ are all nonzero, we deduce from Proposition 7.11 that $\tilde{\mathfrak{h}} = \tilde{\mathfrak{g}}(A)_0$. Thus all parts of the proposition have been proved. □

Proposition 7.13. $\dim \tilde{\mathfrak{g}}(A)_{\alpha_i} = 1$ *and* $\dim \tilde{\mathfrak{g}}(A)_{-\alpha_i} = 1$.

Proof. We know that $e_i \in \tilde{\mathfrak{g}}(A)_{\alpha_i}$. Also, the element $e_i \in \tilde{\mathfrak{g}}(A)$ is nonzero since it induces a nonzero endomorphism $\rho(e_i)$ on the $\tilde{\mathfrak{g}}(A)$-module $\mathfrak{F}^-$ introduced in Proposition 7.9. Hence $\dim \tilde{\mathfrak{g}}(A)_{\alpha_i} \geq 1$. On the other hand, we have

$$\tilde{\mathfrak{g}}(A)_{\alpha_i} \subset \tilde{\mathfrak{g}}(A)^+ = \tilde{\mathfrak{n}}.$$

Now $\tilde{\mathfrak{n}}$ is generated by $e_1, \ldots, e_l$ and is thus spanned by monomials in these elements. All such monomials are weight vectors. The only monomial having weight α_i is e_i since the α_i are linearly independent. Thus we have $\dim \tilde{\mathfrak{g}}(A)_{\alpha_i} = 1$. A similar argument yields $\dim \tilde{\mathfrak{g}}(A)_{-\alpha_i} = 1$. □

7.5 The existence theorem

We now turn to a study of the Lie algebra $\mathfrak{g}(A)$ in order to show that it is a finite-dimensional Lie algebra with Cartan matrix A. From the definitions of $\mathfrak{g}(A)$ and $\tilde{\mathfrak{g}}(A)$, we see that $\mathfrak{g}(A)$ is isomorphic to $\tilde{\mathfrak{g}}(A)/\mathfrak{u}$ where $\mathfrak{u}$ is the ideal of $\tilde{\mathfrak{g}}(A)$ generated by the elements

$$[e_i, [e_i, \cdots [e_i, e_j]]]$$

$$[f_i, [f_i, \cdots [f_i, f_j]]]$$

for all $i \neq j$. As usual, we have $1 - A_{ij}$ factors of e_i and f_i.

Proposition 7.14.

(i) *If* $\mathfrak{u}^+$ *is the ideal of* $\tilde{\mathfrak{n}}$ *generated by the elements* $[e_i, [e_i, \cdots [e_i, e_j]]]$ *for all* $i \neq j$, *then* $\mathfrak{u}^+$ *is an ideal of* $\tilde{\mathfrak{g}}(A)$.

(ii) *If* $\mathfrak{u}^-$ *is the ideal of* $\tilde{\mathfrak{n}}^-$ *generated by the elements* $[f_i, [f_i, \cdots [f_i, f_j]]]$ *for all* $i \neq j$, *then* $\mathfrak{u}^-$ *is an ideal of* $\tilde{\mathfrak{g}}(A)$.

(iii) $\mathfrak{u} = \mathfrak{u}^+ \oplus \mathfrak{u}^-$.

Proof. We write $X_{ij} = [e_i, [e_i, \cdots [e_i, e_j]]]$ and $Y_{ij} = [f_i, [f_i, \cdots [f_i, f_j]]]$. Then $\mathfrak{u}^+$ is the set of all linear combinations of elements

$$[[X_{ij}, e_{k_1}], \cdots e_{k_r}]$$

for all $i \neq j$ and $k_i, \ldots, k_r \in \{1, \ldots, l\}$ since all such linear combinations lie in $\mathfrak{u}^+$ and form an ideal of $\tilde{\mathfrak{n}}$.

Now X_{ij} is a weight vector, being the Lie product of the weight vectors e_i, e_j. Similarly, each $[[X_{ij}, e_{k_1}], \cdots e_{k_r}]$ is a weight vector. It is therefore transformed by each of the elements $h_1, \ldots, h_l$ into a scalar multiple of itself. In order to show that $\mathfrak{u}^+$ is an ideal of $\tilde{\mathfrak{g}}(A)$, it is thus sufficient to show that

$$[f_k, [[X_{ij}, e_{k_1}], \cdots e_{k_r}]] \in \mathfrak{u}^+$$

for all $i, j, k, k_1, \ldots, k_r$. We prove this by induction on r, beginning with $r = 0$. In the following lemma, we will show that $[f_k, X_{ij}] = 0$, thus beginning the induction. Now let $r > 0$ and write $y = [[X_{ij}, e_{k_1}], \cdots e_{k_r}]$, $x = [[X_{ij}, e_{k_1}], \cdots e_{k_{r-1}}]$. We assume by induction that $[f_k, x] \in \mathfrak{u}^+$. Then

$$[f_k, y] = [f_k, [x, e_{k_r}]] = [[f_k, x], e_{k_r}] + [x, [f_k, e_{k_r}]]$$

by the Jacobi identity. If $k_r \neq k$, then $[f_k, y] = [[f_k, x], e_{k_r}] \in \mathfrak{u}^+$. On the other hand, if $k_r = k$, then we have

$$[f_k, y] = [[f_k, x], e_{k_r}] + [h_k, x] \in \mathfrak{u}^+.$$

This completes the induction. Thus $\mathfrak{u}^+$ is an ideal of $\tilde{\mathfrak{g}}(A)$. Similarly, $\mathfrak{u}^-$ is an ideal of $\tilde{\mathfrak{g}}(A)$. It follows that $\mathfrak{u}^+ \oplus \mathfrak{u}^-$ is an ideal of $\tilde{\mathfrak{g}}(A)$ containing the elements X_{ij} and Y_{ij}. Moreover, any ideal of $\tilde{\mathfrak{g}}(A)$ containing the X_{ij} and Y_{ij} must contain $\mathfrak{u}^+$ and $\mathfrak{u}^-$. Hence $\mathfrak{u}^+ \oplus \mathfrak{u}^- = \mathfrak{u}$. □

To complete the proof of Proposition 7.14, we require the following lemma.

Lemma 7.15. $[f_k, X_{ij}] = 0$ *for all* i, j, k *with* $i \neq j$.

Proof. Suppose $k \notin \{i, j\}$. Then $[f_k, e_i] = 0$ and $[f_k, e_j] = 0$ so that

$$[f_k, [e_i, e_j]] = [[f_k, e_i], e_j] + [e_i, [f_k, e_j]] = 0.$$

Using induction, we deduce that $[f_k, X_{ij}] = 0$ for all i, j.

Now suppose $k = j$. Then

$$[f_j, [e_i, e_j]] = [[f_j, e_i], e_j] + [e_i, [f_j, e_j]] = [h_j, e_i] = A_{ji}e_i,$$

$$[f_j, [e_i, [e_i, e_j]]] = [[f_j, e_i], [e_i, e_j]] + [e_i, [f_j, [e_i, e_j]]] = A_{ji}[e_i, e_i] = 0,$$

$$[f_j, \underbrace{[e_i, [e_i, \cdots [e_i,}_{r \text{ times}} e_j]]]] = 0 \qquad \text{for } r \geq 2$$

by induction on r. Hence $[f_j, X_{ij}] = 0$ if $1 - A_{ij} \geq 2$, that is, $A_{ij} \leq -1$. If $1 - A_{ij} = 1$, then $A_{ji} = 0$, and so $[f_k, X_{ij}] = 0$ in this case, also.

Finally, suppose $k = i$. We have

$$[f_i, \underbrace{[e_i, [e_i, \cdots [e_i,}_{r \text{ times}} e_j]]]] = -[h_i, \underbrace{[e_i, [e_i, \cdots [e_i,}_{r-1 \text{ times}} e_j]]]] + [e_i, [f_i, \underbrace{[e_i, [e_i, \cdots [e_i,}_{r-1 \text{ times}} e_j]]]]] \tag{1}$$

by the Jacobi identity. I claim that

$$[e_i, [f_i, \underbrace{[e_i, [e_i, \cdots [e_i,}_{s \text{ times}} e_j]]]]] = -s(A_{ij} + s - 1)\underbrace{[e_i, [e_i, \cdots [e_i,}_{s \text{ times}} e_j]]]$$

for all $s \geq 1$. For $s = 1$, we have

$$[e_i, [f_i, [e_i, e_j]]] = [e_i, [[f_i, e_i], e_j] + [e_i, [f_i, e_j]]] = [e_i, [-h_i, e_j]] = -A_{ij}[e_i, e_j].$$

Now let $s > 1$. For convenience, we write

$$\underbrace{[e_i, [e_i, \cdots [e_i,}_{t \text{ times}} e_j]]] = X_{ij}^t$$

for all $t \geq 1$. We assume by induction that $\left[e_i, \left[f_i, X_{ij}^{s-1}\right]\right] = -(s-1)(A_{ij} + s - 2)X_{ij}^{s-1}$. Then

$$\begin{aligned}
\left[e_i, \left[f_i, X_{ij}^s\right]\right] &= \left[e_i, \left[f_i, \left[e_i, X_{ij}^{s-1}\right]\right]\right] \\
&= \left[e_i, \left[[f_i, e_i], X_{ij}^{s-1}\right] + \left[e_i, \left[f_i, X_{ij}^{s-1}\right]\right]\right] \\
&= \left[e_i, \left[[f_i, e_i], X_{ij}^{s-1}\right]\right] + \left[e_i, \left[e_i, \left[f_i, X_{ij}^{s-1}\right]\right]\right] \\
&= \left[e_i, \left[-h_i, X_{ij}^{s-1}\right]\right] + \left[e_i, -(s-1)(A_{ij} + s - 2)X_{ij}^{s-1}\right] \\
&= -\left[[e_i, h_i], X_{ij}^{s-1}\right] - \left[h_i, \left[e_i, X_{ij}^{s-1}\right]\right] - (s-1)(A_{ij} + s - 2)\left[e_i, X_{ij}^{s-1}\right] \\
&= \left[A_{ii}e_i, X_{ij}^{s-1}\right] - \left[h_i, X_{ij}^s\right] - (s-1)(A_{ij} + s - 2)X_{ij}^s \\
&= 2X_{ij}^s - \left[h_i, X_{ij}^s\right] - (s-1)(A_{ij} + s - 2)X_{ij}^s. \qquad (2)
\end{aligned}$$

One easily verifies by induction that $\left[h_i, X_{ij}^t\right] = (A_{ij} + 2t)X_{ij}^t$ for all $t \geq 1$. Plugging this into (2) for $t = s$ gives us

$$\left[e_i, \left[f_i, X_{ij}^s\right]\right] = -s(A_{ij} + s - 1)X_{ij}^s$$

as desired. Inserting this result into (1) for $s = r - 1$ then yields

$$\begin{aligned}
&[f_i, \underbrace{[e_i, [e_i, \cdots [e_i,}_{r \text{ times}} e_j]]]] \\
&\quad = -[h_i, \underbrace{[e_i, [e_i, \cdots [e_i,}_{r-1 \text{ times}} e_j]]]] - (r-1)(A_{ij} + r - 2) \underbrace{[e_i, [e_i, \cdots [e_i,}_{r-1 \text{ times}} e_j]]] \\
&\quad = (-(A_{ij} + 2r - 2) - (r-1)(A_{ij} + r - 2)) \underbrace{[e_i, [e_i, \cdots [e_i,}_{r-1 \text{ times}} e_j]]] \\
&\quad = -r(A_{ij} + r - 1) \underbrace{[e_i, [e_i, \cdots [e_i,}_{r-1 \text{ times}} e_j]]].
\end{aligned}$$

We now set $r = 1 - A_{ij}$ and obtain $[f_i, X_{ij}] = 0$. □

Corollary 7.16. *$\mathfrak{g}(A) = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}$ where $\mathfrak{h}$ is isomorphic to $\tilde{\mathfrak{h}}$, $\mathfrak{n}$ is isomorphic to $\tilde{\mathfrak{n}}^-/\mathfrak{u}^-$, and $\mathfrak{n}$ is isomorphic to $\tilde{\mathfrak{n}}/\mathfrak{u}^+$.*

Proof. This follows from the fact that $\mathfrak{g}(A)$ is isomorphic to $\tilde{\mathfrak{g}}(A)/\mathfrak{u}$, $\tilde{\mathfrak{g}}(A) = \tilde{\mathfrak{n}}^- \oplus \tilde{\mathfrak{h}} \oplus \tilde{\mathfrak{n}}$, and $\mathfrak{u} = \mathfrak{u}^+ \oplus \mathfrak{u}^-$. □

We will continue to denote the generators of $\mathfrak{g}(A)$ by e_i, h_i, f_i. These are the images of the generators of $\mathfrak{L}$ under the natural homomorphism $\mathfrak{L} \to \mathfrak{g}(A)$.

Proposition 7.17. *The maps $\operatorname{ad} e_i : \mathfrak{g}(A) \to \mathfrak{g}(A)$ and $\operatorname{ad} f_i : \mathfrak{g}(A) \to \mathfrak{g}(A)$ are locally nilpotent.*

Proof. To show that $\operatorname{ad} e_i$ is locally nilpotent, we must show that for each $x \in \mathfrak{g}(A)$, there exists an integer $n(x) \geq 1$ such that $(\operatorname{ad} e_i)^{n(x)}x = 0$. Now if $\operatorname{ad} e_i$ acts locally nilpotently on x and y, then $\operatorname{ad} e_i$ acts locally nilpotently on $[x, y]$. This follows from the relation

$$(\operatorname{ad} e_i)^n[x, y] = \sum_{r=0}^{n} \binom{n}{r} \left[(\operatorname{ad} e_i)^r x, (\operatorname{ad} e_i)^{n-r} y\right].$$

We have $(\mathrm{ad}\, e_i)^r x = 0$ if r is sufficiently large and $(\mathrm{ad}\, e_i)^{n-r} y = 0$ if $n - r$ is sufficiently large. Thus $(\mathrm{ad}\, e_i)^n [x, y] = 0$ if n is sufficiently large.

It follows that the set of elements in $\mathfrak{g}(A)$ on which $\mathrm{ad}\, e_i$ acts locally nilpotently is a subalgebra of $\mathfrak{g}(A)$. However, we have

$$\begin{aligned}
&\mathrm{ad}\, e_i \cdot e_i = 0 \\
&(\mathrm{ad}\, e_i)^{1-A_{ij}} e_j = 0 \quad \text{if } i \neq j \\
&(\mathrm{ad}\, e_i)^2 h_j = 0 \quad \text{for all } j \\
&(\mathrm{ad}\, e_i)^3 f_i = 0 \\
&\mathrm{ad}\, e_i \cdot f_j = 0 \quad \text{if } i \neq j.
\end{aligned}$$

Thus this subalgebra contains all generators of $\mathfrak{g}(A)$, so it must be the whole of $\mathfrak{g}(A)$. A similar argument shows that $\mathrm{ad}\, f_i$ is locally nilpotent on $\mathfrak{g}(A)$. □

Now the proof of Proposition 3.4 shows that if $D : \mathfrak{g} \to \mathfrak{g}$ is a locally nilpotent derivation of a Lie algebra $\mathfrak{g}$, then $\exp(D)$ is an automorphism of $\mathfrak{g}$. Thus $\exp(\mathrm{ad}\, e_i)$ and $\exp(\mathrm{ad}\, f_i)$ are automorphisms of $\mathfrak{g}(A)$. For each i, we define $\phi_i \in \mathrm{Aut}(\mathfrak{g}(A))$ by

$$\phi_i = \exp(\mathrm{ad}\, e_i) \cdot \exp(\mathrm{ad}\, (-f_i)) \cdot \exp(\mathrm{ad}\, e_i).$$

Proposition 7.18.

(i) $\phi_i(\mathfrak{h}) = \mathfrak{h}$.

(ii) $\phi_i(\mathfrak{h}) = s_i(\mathfrak{h})$ *where* $s_i : \mathfrak{h} \to \mathfrak{h}$ *is the linear map defined by*

$$s_i(h_j) = h_j - A_{ji} h_i.$$

Proof. We have

$$\exp(\mathrm{ad}\, e_i) \cdot h_j = (1 + \mathrm{ad}\, e_i) h_j = h_j - A_{ji} e_i,$$

$$\begin{aligned}
\exp(\operatorname{ad}(-f_i)) \cdot \exp(\operatorname{ad} e_i) \cdot h_j &= \exp(\operatorname{ad}(-f_i)) \cdot (h_j - A_{ji} e_i) \\
&= \left(1 - \operatorname{ad} f_i + \frac{(\operatorname{ad} f_i)^2}{2}\right)(h_j - A_{ji} e_i) \\
&= h_j - A_{ji} e_i - A_{ji} f_i - A_{ji} h_i + A_{ji} f_i \\
&= h_j - A_{ji} h_i - A_{ji} e_i,
\end{aligned}$$

$$\begin{aligned}
\exp(\operatorname{ad} e_i) \cdot \exp(\operatorname{ad}(-f_i)) \cdot \exp(\operatorname{ad} e_i) \cdot h_j &= \exp(\operatorname{ad} e_i)(h_j - A_{ji} h_i - A_{ji} e_i) \\
&= (1 + \operatorname{ad} e_i)(h_j - A_{ji} h_i - A_{ji} e_i) \\
&= h_j - A_{ji} h_i - A_{ji} e_i - A_{ji} e_i + 2A_{ji} e_i \\
&= h_j - A_{ji} h_i.
\end{aligned}$$

Thus ϕ_i restricted to the subalgebra $\mathfrak{h}$ is the map s_i described in the second statement. Since $\phi_i(\mathfrak{h}) \subset \mathfrak{h}$ and ϕ_i is an automorphism of $\mathfrak{g}(A)$, we must have $\phi_i(\mathfrak{h}) = \mathfrak{h}$. □

Now the action of s_i on $\mathfrak{h}$ is the same as that of the fundamental reflection $s_i = s_{\alpha_i}$ defined in Section 6.4. We recall that

$$\begin{aligned}
s_i(h) &= h - 2\frac{\langle h'_{\alpha_i}, h \rangle}{\langle h'_{\alpha_i}, h'_{\alpha_i} \rangle} h'_{\alpha_i} \qquad \text{for } h \in \mathfrak{h} \\
&= h - \langle h'_{\alpha_i}, h \rangle h_i.
\end{aligned}$$

In particular, we have

$$s_i(h_j) = h_j - \langle h'_{\alpha_i}, h_j \rangle h_i = h_j - 2\frac{\langle h'_{\alpha_i}, h'_{\alpha_j} \rangle}{\langle h'_{\alpha_j}, h'_{\alpha_j} \rangle} h_i = h_j - A_{ji} h_i.$$

Thus Proposition 7.18 shows that the automorphism ϕ_i induces the fundamental reflection s_i on $\mathfrak{h}$.

We now consider the decomposition of $\mathfrak{g}(A)$ into weight spaces with respect to $\mathfrak{h}$. This time, the weights will be the elements $\lambda \in \operatorname{Hom}(\mathfrak{h}, \mathbb{C})$ for which the set

$$\mathfrak{g}(A)_\lambda = \{x \in \mathfrak{g}(A) \mid [h, x] = \lambda(h)x \quad \text{for all } h \in \mathfrak{h}\}$$

is nonzero.

Proposition 7.19. $\mathfrak{g}(A) = \bigoplus_\lambda \mathfrak{g}(A)_\lambda$.

Proof. The Lie algebra $\mathfrak{g}(A)$ is the sum of its weight spaces since its generators e_i, h_i, f_i are weight vectors. Moreover, the sum of weight spaces is direct, just as in the proof of Proposition 7.11. □

It also follows from Proposition 7.12 that $\mathfrak{g}(A) = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}$ where all weights coming from $\mathfrak{n}$ are in Q^+ and all weights coming from $\mathfrak{n}^-$ are in Q^-.

Proposition 7.20. $\dim \mathfrak{g}(A)_{\alpha_i} = 1$ *and* $\dim \mathfrak{g}(A)_{-\alpha_i} = 1$.

Proof. By Proposition 7.13, we know that $\dim \mathfrak{g}(A)_{\alpha_i} \leq 1$. Now the ideal $\mathfrak{u}^+$ of $\tilde{\mathfrak{n}}$ such that $\mathfrak{n} \cong \tilde{\mathfrak{n}}/\mathfrak{u}^+$ has the property that $\mathfrak{u}^+$ is a sum of weight spaces, and all weights occurring in $\mathfrak{u}^+$ are sums of $\alpha_1, \ldots, \alpha_l$ involving at least two terms. This can be deduced from the proof of Proposition 7.14. Thus α_i is not a weight of $\mathfrak{u}^+$. Hence

$$\dim \mathfrak{g}(A)_{\alpha_i} = \dim \tilde{\mathfrak{g}}(A)_{\alpha_i} = 1.$$

A similar argument shows that $\dim \mathfrak{g}(A)_{-\alpha_i} = 1$. □

Proposition 7.21. *The automorphism ϕ_i of $\mathfrak{g}(A)$ transforms $\mathfrak{g}(A)_\lambda$ to $\mathfrak{g}(A)_{s_i\lambda}$. Hence* $\dim \mathfrak{g}(A)_\lambda = \dim \mathfrak{g}(A)_{s_i\lambda}$.

Proof. Let $x \in \mathfrak{g}(A)_\lambda$. Then $[h, x] = \lambda(h)x$ for all $h \in \mathfrak{h}$. We now apply the automorphism ϕ_i. This fixes $\mathfrak{h}$ by Proposition 7.18. We thus have

$$[\phi_i h, \phi_i x] = \lambda(h)\phi_i x.$$

Hence

$$[h, \phi_i x] = \lambda\left(\phi_i^{-1} h\right)\phi_i x = \lambda\left(s_i^{-1} h\right)\phi_i x = (s_i\lambda(h))\phi_i x,$$

again by Proposition 7.18. Thus we have $\phi_i x \in \mathfrak{g}(A)_{s_i\lambda}$. Hence

$$\phi_i(\mathfrak{g}(A)_\lambda) \subset \mathfrak{g}(A)_{s_i\lambda}.$$

Replacing ϕ_i by ϕ_i^{-1}, λ by $s_i\lambda$, and recalling that $s_i^2 = 1$, we also have

$$\phi_i^{-1}(\mathfrak{g}(A)_{s_i\lambda}) \subset \mathfrak{g}(A)_\lambda.$$

Hence $\mathfrak{g}(A)_{s_i\lambda} \subset \phi_i(\mathfrak{g}(A)_\lambda)$, giving us $\phi_i(\mathfrak{g}(A)_\lambda) = \mathfrak{g}(A)_{s_i\lambda}$. □

We now define W to be the group of non-singular transformations on $\mathfrak{h}^* = \mathrm{Hom}(\mathfrak{h}, \mathbb{C})$ generated by $s_1, \ldots, s_l$ and define Φ to be the set of elements $w(\alpha_i)$ for $w \in W$ and $i = 1, \ldots, l$. Then Φ is the root system determined by the given Cartan matrix A and W is the Weyl group.

Proposition 7.22. $\dim \mathfrak{g}(A)_\alpha = 1$ *for all* $\alpha \in \Phi$.

Proof. We have $\alpha = w(\alpha_i)$ for some i and some $w \in W$. Since W is generated by $s_1, \ldots, s_l$, w is a product of such elements. It follows from Proposition 7.21 that

$$\dim \mathfrak{g}(A)_\alpha = \dim \mathfrak{g}(A)_{\alpha_i} = 1.$$ □

Our goal is to show that $\mathfrak{g}(A)$ is finite-dimensional. As a step in this direction, we will show that the Weyl group W is finite. We know that W is isomorphic to the group of non-singular linear transformations on $\mathfrak{h}_{\mathbb{R}}$ generated by $s_1, \ldots, s_l$ where $\mathfrak{h}_{\mathbb{R}} = \mathbb{R}h_1 + \cdots + \mathbb{R}h_l$. We have $\dim \mathfrak{h}_{\mathbb{R}} = l$. We do not have a scalar product on $\mathfrak{h}_{\mathbb{R}}$ available from the Killing form, so we define a scalar product directly from the Cartan matrix A.

Proposition 7.23. *The Cartan matrix can be factorized as* $A = DB$ *where* D *is diagonal and* B *is symmetric. Moreover,* D *is the diagonal matrix with entries* $d_1, \ldots, d_l$ *defined as follows:*

- *If the Dynkin diagram has only single edges, then* $d_i = 1$ *for all* i.
- *If the Dynkin diagram has a double edge, then* $d_i = 1$ *if* α_i *is a short root and* $d_i = 2$ *if* α_i *is a long root.*
- *If the Dynkin diagram has a triple edge, then* $d_i = 1$ *if* α_i *is a short root and* $d_i = 3$ *if* α_i *is a long root.*

Proof. This can be checked from the standard list of Cartan matrices in Section 6.4. □

For example, for the matrix of type G_2, we have

$$\begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & \frac{2}{3} \end{pmatrix}.$$

We now define a bilinear form on $\mathfrak{h}_{\mathbb{R}}$ by $\langle h_i, h_j \rangle = d_i d_j B_{ij}$. This form is symmetric since B is a symmetric matrix.

Proposition 7.24. *This scalar product is positive definite.*

Proof. We have $n_{ij} = A_{ij}A_{ji} = d_i d_j B_{ij}^2$. Thus $-\sqrt{n_{ij}} = \sqrt{d_i}\sqrt{d_j}B_{ij}$. The matrix of our scalar product is

$$DBD = \begin{pmatrix} \sqrt{d_1} & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \sqrt{d_l} \end{pmatrix} \begin{pmatrix} 2 & & & -\sqrt{n_{ij}} \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ -\sqrt{n_{ij}} & & & 2 \end{pmatrix} \begin{pmatrix} \sqrt{d_1} & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \sqrt{d_l} \end{pmatrix}.$$

This matrix is congruent to the matrix

$$\begin{pmatrix} 2 & & & -\sqrt{n_{ij}} \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ -\sqrt{n_{ij}} & & & 2 \end{pmatrix}$$

of the quadratic form $\mathrm{Q}(x_1, \ldots, x_l)$ of Proposition 6.6, which is positive definite. Thus DBD is also positive definite. □

Proposition 7.25. *Our scalar product on $\mathfrak{h}_{\mathbb{R}}$ is invariant under W.*

Proof. We first observe that $\langle h_i, x \rangle = d_i \alpha_i(x)$ for all $h \in \mathfrak{h}_{\mathbb{R}}$. This follows from the fact that

$$\langle h_i, h_j \rangle = d_i d_j B_{ij} = d_i A_{ji} = d_i \alpha_i(h_j)$$

for all h_i, h_j.

To prove the proposition, it suffices to show that $\langle s_i x, s_i y\rangle = \langle x, y\rangle$ for all $x, y \in \mathfrak{h}_{\mathbb{R}}$ since the reflections s_i generate W. We note that $s_i(x) = x - \alpha_i(x)h_i$ since $s_i(h_j) = h_j - A_{ji}h_i$. Thus

$$\begin{aligned}\langle s_i x, s_i y\rangle &= \langle x - \alpha_i(x)h_i, y - \alpha_i(y)h_i\rangle \\ &= \langle x, y\rangle - \alpha_i(y)\langle h_i, x\rangle - \alpha_i(x)\langle h_i, y\rangle + \alpha_i(x)\alpha_i(y)\langle h_i, h_i\rangle \\ &= \langle x, y\rangle - d_i\alpha_i(y)\alpha_i(x) - d_i\alpha_i(x)\alpha_i(y) + 2d_i\alpha_i(x)\alpha_i(y) \\ &= \langle x, y\rangle. \end{aligned}$$

□

Thus the Weyl group acts as a group of isometries on the Euclidean space $\mathfrak{h}_{\mathbb{R}}$. We now define certain subsets of $\mathfrak{h}_{\mathbb{R}}$ in terms of the scalar product $\langle , \rangle$ as follows:

$$\begin{aligned}\mathfrak{h}_i &= \{x \in \mathfrak{h}_{\mathbb{R}} \mid \langle h_i, x\rangle = 0\} \\ \mathfrak{h}_i^+ &= \{x \in \mathfrak{h}_{\mathbb{R}} \mid \langle h_i, x\rangle > 0\} \\ \mathfrak{h}_i^- &= \{x \in \mathfrak{h}_{\mathbb{R}} \mid \langle h_i, x\rangle < 0\} \\ C &= \mathfrak{h}_1^+ \cap \cdots \cap \mathfrak{h}_l^+. \end{aligned}$$

The set C is called the **fundamental chamber**.

Let W_{ij} be the subgroup of W generated by s_i, s_j where $i \neq j$. I claim that each element $s_i s_j$ has finite order m_{ij} given in terms of the Cartan matrix by $2\cos(\pi/m_{ij}) = \sqrt{n_{ij}}$. To prove this, recall that

$$n_{ij} = A_{ij}A_{ji} = 4\cos^2\theta$$

where θ is the angle between α_i, α_j. It follows that

$$s_i(\alpha_j) = \alpha_j - 2\frac{\langle \alpha_i, \alpha_j\rangle}{\langle \alpha_i, \alpha_i\rangle}\alpha_i = \alpha_j - 2\cos\theta\frac{|\alpha_j|}{|\alpha_i|}\alpha_i = \alpha_j - \sqrt{n_{ij}}\frac{|\alpha_j|}{|\alpha_i|}\alpha_i.$$

Now consider the map s_i' defined by

$$s_i'(\alpha_j) = \alpha_j - 2\cos(\pi/m_{ij})\frac{|\alpha_j|}{|\alpha_i|}\alpha_i.$$

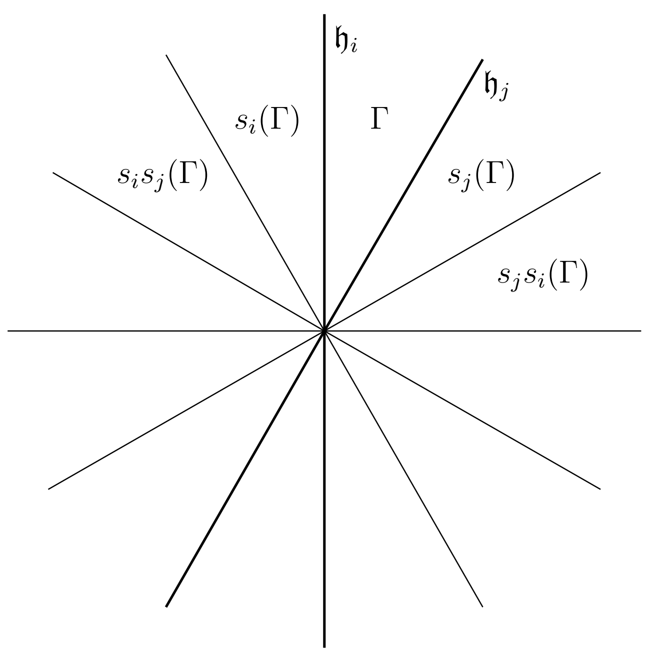
For $i = j$, we have $m_{ii} = 1$ so that $s'_i(\alpha_i) = -\alpha_i$. If α_j is orthogonal to α_i under our inner product, then an interesting calculation shows that $m_{ij} = 2$ so that $s_i(\alpha_j) = \alpha_j$. It follows that s'_i is an orthogonal reflection about a hyperplane that intersects the origin. Thus s'_i is a linear transformation. We see, in fact, that s'_i is the reflection in the hyperplane orthogonal to α_i. It follows that $s'_i = s_i$. Hence for all i, j, we have

$$\alpha_j - 2\cos(\pi/m_{ij})\frac{|\alpha_j|}{|\alpha_i|}\alpha_i = \alpha_j - \sqrt{n_{ij}}\frac{|\alpha_j|}{|\alpha_i|}\alpha_i.$$

Thus $2\cos(\pi/m_{ij}) = \sqrt{n_{ij}}$ as desired. This shows that W_{ij} is a finite dihedral group.

Lemma 7.26. *Let $w \in W_{ij}$ with $i \neq j$. Then either $w(\mathfrak{h}_i^+ \cap \mathfrak{h}_j^+) \subset \mathfrak{h}_i^+$, or $w(\mathfrak{h}_i^+ \cap \mathfrak{h}_j^+) \subset \mathfrak{h}_i^-$ and $l(s_i w) = l(w) - 1$.*

Proof. Let U be the 2-dimensional subspace of $\mathfrak{h}_{\mathbb{R}}$ spanned by h_i, h_j, and let $U^\perp$ be the orthogonal subspace. Then $\mathfrak{h}_{\mathbb{R}} = U \oplus U^\perp$ and the elements in W_{ij} act trivially on $U^\perp$. It is therefore sufficient to prove the result in U. Let $\Gamma = U \cap \mathfrak{h}_i^+ \cap \mathfrak{h}_j^+$. Because U is 2-dimensional, we obtain the following configuration of chambers in U:



From this diagram, we see that the chambers

$$\Gamma, \quad s_j(\Gamma), \quad s_js_i(\Gamma), \quad \ldots, \quad \underbrace{s_js_i\cdots}_{m_{ij}-1\text{ times}}(\Gamma)$$

all lie in $\mathfrak{h}_i^+$, whereas

$$s_i(\Gamma), \quad s_is_j(\Gamma), \quad \ldots, \quad \underbrace{s_is_j\cdots}_{m_{ij}\text{ times}}(\Gamma)$$

all lie in $\mathfrak{h}_i^-$. The elements

$$s_i, \quad s_is_j, \quad \ldots, \quad \underbrace{s_is_j\cdots}_{m_{ij}\text{ times}}$$

all satisfy $l(s_iw) = l(w) - 1$ since $s_i^2 = 1$. Thus for each $w \in W_{ij}$, we have either $w\big(\mathfrak{h}_i^+ \cap \mathfrak{h}_j^+\big) \subset \mathfrak{h}_i^+$, or $w\big(\mathfrak{h}_i^+ \cap \mathfrak{h}_j^+\big) \subset \mathfrak{h}_i^-$ and $l(s_iw) = l(w) - 1$. □

Proposition 7.27.

(i) *Let $w \in W$. Then either $w(C) \subset \mathfrak{h}_i^+$, or $w(C) \subset \mathfrak{h}_i^-$ and $l(s_iw) = l(w) - 1$.*

(ii) *Let $w \in W$ and $i \neq j$. Then there exists a $w' \in W_{ij}$ such that $w(C) \subset w'\big(\mathfrak{h}_i^+ \cap \mathfrak{h}_j^+\big)$ and $l(w) = l(w') + l\left(w'^{-1}w\right)$.*

Note. Part (i) is the result we will need. To prove it, we must also prove part (ii) at the same time.

Proof. We prove both statements together by induction on $l(w)$. If $l(w) = 0$, then $w = 1$, and so (i) and (ii) are true with $w' = 1$. So suppose $l(w) > 0$. Then $w = s_jw'$ with $l(w') = l(w) - 1$ for some j and $w' \in W$. We prove part (i).

First suppose $j = i$. By induction, we have $w'(C) \subset \mathfrak{h}_i^+$, or $w'(C) \subset \mathfrak{h}_i^-$ and $l(s_iw') = l(w') - 1$. But $l(s_iw') = l(w') + 1$, and so $w'(C) \subset \mathfrak{h}_i^+$. Thus $w(C) \subset \mathfrak{h}_i^-$ and $l(s_iw) = l(w) - 1$.

Now suppose $j \neq i$. By induction, there exists a $w'' \in W_{ij}$ such that $w'(C) \subset w''\big(\mathfrak{h}_i^+ \cap \mathfrak{h}_j^+\big)$ and $l(w') = l(w'') + l\left(w''^{-1}w'\right)$. Thus $w(C) \subset s_jw''\big(\mathfrak{h}_i^+ \cap \mathfrak{h}_j^+\big)$. By Lemma

7.26, we have either $s_j w''\left(\mathfrak{h}_i^+ \cap \mathfrak{h}_j^+\right) \subset \mathfrak{h}_i^+$, or $s_j w''\left(\mathfrak{h}_i^+ \cap \mathfrak{h}_j^+\right) \subset \mathfrak{h}_i^-$ and $l\left(s_i s_j w''\right) = l\left(s_j w''\right) - 1$. In the first case, we have $w(C) \subset \mathfrak{h}_i^+$. In the second case, we have $w(C) \subset \mathfrak{h}_i^-$ and

$$
\begin{aligned}
l(s_i w) = l\left(s_i s_j w'\right) &= l\left(s_i s_j w'' w''^{-1} w'\right) \\
&\leq l\left(s_i s_j w''\right) + l\left(w''^{-1} w'\right) \\
&= l\left(s_j w''\right) - 1 + l\left(w''^{-1} w'\right) \\
&\leq l\left(w''\right) + l\left(w''^{-1} w'\right) = l\left(w'\right) = l(w) - 1.
\end{aligned}
$$

Thus $l(s_i w) = l(w) - 1$, and (i) is proved.

We now prove (ii). If $w(C) \subset \mathfrak{h}_i^+ \cap \mathfrak{h}_j^+$, then (ii) is true with $w' = 1$. Without loss of generality, we now assume that $w(C) \not\subset \mathfrak{h}_i^+$. Hence by (i), which is now proved under the assumption of the inductive hypothesis, $w(C) \subset \mathfrak{h}_i^-$ and $l(s_i w) = l(w) - 1$. By induction, there exists a $w' \in W_{ij}$ such that $s_i w(C) \subset w'\left(\mathfrak{h}_i^+ \cap \mathfrak{h}_j^+\right)$ and $l(s_i w) = l\left(w'\right) + l\left(w'^{-1} s_i w\right)$. Thus $w(C) = s_i s_i w(C) \subset s_i w'\left(\mathfrak{h}_i^+ \cap \mathfrak{h}_j^+\right)$ and

$$
l(w) = 1 + l(s_i w) = 1 + l\left(w'\right) + l\left(w'^{-1} s_i w\right) \geq l\left(s_i w'\right) + l\left(\left(s_i w'\right)^{-1} w\right) \geq l(w).
$$

Thus we have equality throughout, and so $l(w) = l\left(s_i w'\right) + l\left(\left(s_i w'\right)^{-1} w\right)$. Hence $s_i w' \in W_{ij}$ is the required element, and (ii) is proved. □

Proposition 7.28. *If $w \in W$ satisfies $C \cap w(C) \neq \varnothing$, then $w = 1$.*

Proof. We prove the contrapositive. Suppose $w \neq 1$. Then $w = s_i w'$ with $l(w') = l(w) - 1$ for some i and $w' \in W$. By Proposition 7.27(i), we have $w'(C) \subset \mathfrak{h}_i^+$. Thus $w(C) \subset \mathfrak{h}_i^-$, and so

$$
C \cap w(C) \subset \mathfrak{h}_i^+ \cap \mathfrak{h}_i^- = \varnothing.
$$

□

Now the Euclidean space $\mathfrak{h}_{\mathbb{R}}$ has an orthonormal basis, and the isometries of $\mathfrak{h}_{\mathbb{R}}$ are represented by orthogonal matrices with respect to this basis. Thus $W \subset O_l$ where O_l is the group of $l \times l$ orthogonal matrices. We have $O_l \subset M_l$, the set of all $l \times l$ matrices over $\mathbb{R}$.

For any matrix $M = (m_{ij}) \in M_l$, we define $\|M\| = \sqrt{\sum_{i,j} m_{ij}^2}$. Similarly, for any column vector $v = (\lambda_1, \ldots, \lambda_l)^t \in \mathbb{R}^l$, we define $\|v\| = \sqrt{\sum_i \lambda_i^2}$.

Lemma 7.29.

(i) *If* $M \in O_l$, $v \in \mathbb{R}^l$, *then* $\|Mv\| = \|v\|$.

(ii) *If* $M \in O_l, N \in M_l$, *then* $\|MN\| = \|N\|$.

(iii) *If* $M \in M_l$, $v \in \mathbb{R}^l$, *then* $\|Mv\| \leq \|M\| \, \|v\|$.

Proof. Parts (i) and (ii) are straightforward calculations using the fact that the columns of M all have length 1. Part (iii) follows from a similarly easy calculation. □

Proposition 7.30.

(i) W *is finite.*

(ii) Φ *is finite.*

Proof. Since $\Phi = W(\Pi)$ where $\Pi = \{\alpha_1, \ldots, \alpha_l\}$, it is clear that (i) implies (ii). Thus we show that W is finite. We consider the W-action on the Euclidean space $\mathfrak{h}_{\mathbb{R}}$. We assign coordinates to elements of $\mathfrak{h}_{\mathbb{R}}$ relative to our orthonormal basis. Let $v = (\lambda_1, \ldots, \lambda_l)^t \in C$. By the definition of C, there exists an $r > 0$ such that $B_r(v) \subset C$ where

$$B_r(v) = \{x \in \mathbb{R}^l \mid \|x - v\| < r\}.$$

Let $w \in W$ with $w \neq 1$. Then $C \cap w(C) = \varnothing$ by Proposition 7.28. Thus $w(v) \notin C$, and so

$$\|w(v) - v\| \geq r.$$

Hence

$$\|w - 1\| \, \|v\| \geq \|w(v) - v\| \geq r$$

by Lemma 7.29, yielding $\|w-1\| \geq \frac{r}{\|v\|}$. Let $\epsilon = \frac{r}{\|v\|}$. Then $\|w-1\| \geq \epsilon$ for all $w \neq 1$ in W. Now let $w, w' \in W$ such that $w \neq w'$. Then

$$\left\|w - w'\right\| = \left\|w'\left(w'^{-1}w - 1\right)\right\| = \left\|w'^{-1}w - 1\right\| \geq \epsilon$$

since $w' \in O_l$. Thus distinct elements of W are separated by a distance of at least ϵ. Since O_l, and hence W, is bounded, it follows that W is finite. □

We now return to our Lie algebra $\mathfrak{g}(A)$. We know that $\dim \mathfrak{g}(A)_0 = l$ and $\dim \mathfrak{g}(A)_\alpha = 1$ for all $\alpha \in \Phi$. If we can prove that every weight of $\mathfrak{g}(A)$ lies in $\Phi \cup \{0\}$, then we will be able to deduce that $\mathfrak{g}(A)$ is finite-dimensional.

Proposition 7.31. *Every nonzero weight of $\mathfrak{g}(A)$ lies in Φ.*

Proof. Let λ be a nonzero weight of $\mathfrak{g}(A)$. Since $\dim \mathfrak{g}(A)_\lambda \leq \dim \tilde{\mathfrak{g}}(A)_\lambda$, we see from Proposition 7.12(iii) that $\lambda \in Q^+$ or $\lambda \in Q^-$. In particular, λ lies in the vector space $\mathfrak{h}_{\mathbb{R}}^*$ of real linear combinations of $\alpha_1, \ldots, \alpha_l$.

Suppose first that λ is a multiple of a root $\alpha \in \Phi$. Then $\lambda = n\alpha$ or $\lambda = -n\alpha$ for some $n > 0$ and $\alpha \in \Phi^+$. We know that $\alpha = w(\alpha_i)$ for some $w \in W$ and $\alpha_i \in \Pi$, and we have $\dim \mathfrak{g}(A)_{n\alpha} = \dim \mathfrak{g}(A)_{n\alpha_i}$ by Proposition 7.21. Hence $\mathfrak{g}(A)_{n\alpha_i} \neq 0$. Now $\mathfrak{n}$ is generated by the elements $e_1, \ldots, e_l$, and no nonzero Lie product of these can have weight $n\alpha_i$ unless $n = 1$. Thus $\lambda = \alpha$ or $\lambda = -\alpha$, that is, $\lambda \in \Phi$.

Now suppose λ is not a multiple of a root. Let

$$\mathfrak{h}_\lambda = \{h \in \mathfrak{h}_{\mathbb{R}} \mid \lambda(h) = 0\}$$

$$\mathfrak{h}_\alpha = \{h \in \mathfrak{h}_{\mathbb{R}} \mid \alpha(h) = 0\}.$$

Then $\mathfrak{h}_\lambda$ is distinct from all the $\mathfrak{h}_\alpha$, $\alpha \in \Phi$. Thus $\mathfrak{h}_\lambda \backslash \bigcup_{\alpha\in\Phi}(\mathfrak{h}_\lambda \cap \mathfrak{h}_\alpha) \neq \varnothing$ since Φ is finite and the vector space $\mathfrak{h}_\lambda$ over $\mathbb{R}$ cannot be expressed as the union of finitely-many proper subspaces. Hence there exists an $h \in \mathfrak{h}_\lambda$ such that $h \notin \mathfrak{h}_\alpha$ for all $\alpha \in \Phi$. Then $w(h) \notin \mathfrak{h}_\alpha$ for all $\alpha \in \Phi$ since W permutes the $\mathfrak{h}_\alpha$.

I claim there exists a $w \in W$ such that $\alpha_i(w(h)) > 0$ for all $i = 1, \ldots, l$. To prove this, we define the height of an element of $\mathfrak{h}_{\mathbb{R}}$ by

$$\operatorname{ht}\left(\sum n_i h_i\right) = \sum n_i.$$

We choose an element $w \in W$ such that $\operatorname{ht} w(h)$ is maximum. This is possible since W is finite. Then

$$s_i(w(h)) = w(h) - \alpha_i(w(h))h_i.$$

Since $\operatorname{ht} s_i(w(h)) \leq \operatorname{ht} w(h)$, we have $\alpha_i(w(h)) \geq 0$. But $\alpha_i(w(h)) = 0$ would imply $w(h) \in \mathfrak{h}_{\alpha_i}$, a contradiction. Thus $\alpha_i(w(h)) > 0$ for all i, and so $w(h) \in C$.

Now we have $(w(\lambda))(w(h)) = \lambda(h) = 0$ by the way we defined the W-action on $\mathfrak{h}$ in Section 6.4. We write $w(\lambda) = \sum_{i=1}^{l} m_i \alpha_i$. Then

$$\sum_{i=1}^{l} m_i \alpha_i(w(h)) = 0.$$

Since $\alpha_i(w(h)) > 0$ for all i, we must have some $m_i > 0$ and some $m_j < 0$ in this sum. Thus $w(\lambda) \notin Q^+$ and $w(\lambda) \notin Q^-$. Hence $\tilde{\mathfrak{g}}(A)_{w(\lambda)} = 0$. By Proposition 7.21, this implies that $\tilde{\mathfrak{g}}(A)_\lambda = 0$. Thus $\mathfrak{g}(A)_\lambda = 0$, a contradiction. □

Corollary 7.32.

(i) $\mathfrak{g}(A) = \mathfrak{h} \oplus \sum_{\alpha \in \Phi} \mathfrak{g}(A)_\alpha$.

(ii) $\dim \mathfrak{g}(A) = l + |\Phi|$.

Proof. Part (i) is clear since $\mathfrak{g}(A)$ is the direct sum of its weight spaces. The 0-weight space is $\mathfrak{h}$, and this has dimension l. The only nonzero weights are those belonging to Φ, and the corresponding weight spaces are 1-dimensional by Proposition 7.22. Thus (ii) is the correct formula for the dimension of $\mathfrak{g}(A)$. □

We have thus shown that $\mathfrak{g}(A)$ is a finite-dimensional Lie algebra – indeed, it has the dimension required for a simple Lie algebra with Cartan matrix A. We next show that $\mathfrak{g}(A)$ has the other required properties.

Proposition 7.33. *The Lie algebra* $\mathfrak{g}(A)$ *is semisimple.*

Proof. Let $\mathfrak{r}$ be the solvable radical of $\mathfrak{g}(A)$ and consider its derived series

$$\mathfrak{r} = \mathfrak{r}^{(0)} \supset \mathfrak{r}^{(1)} \supset \cdots \supset \mathfrak{r}^{(n-1)} \supset \mathfrak{r}^{(n)} = 0$$

where, as usual, $\mathfrak{r}^{(i+1)} = \left[\mathfrak{r}^{(i)}, \mathfrak{r}^{(i)}\right]$. Let $\mathfrak{a} = \mathfrak{r}^{(n-1)}$. Suppose $\mathfrak{r} \neq 0$. Then $\mathfrak{a}$ is a nonzero abelian ideal of $\mathfrak{g}(A)$. We know that $\mathfrak{r}$ is unique, and hence it is invariant under the automorphisms of $\mathfrak{g}(A)$. Thus $[\mathfrak{r}, \mathfrak{r}]$ is invariant under the automorphisms of $\mathfrak{g}(A)$, and an inductive argument shows that $\mathfrak{r}^{(i)}$ is invariant under these automorphisms for all i. Thus $\mathfrak{a}$ is invariant under all automorphisms of $\mathfrak{g}(A)$.

Since $[\mathfrak{h}, \mathfrak{a}] \subset \mathfrak{a}$, we may regard $\mathfrak{a}$ as an $\mathfrak{h}$-submodule. We decompose $\mathfrak{a}$ into its weight spaces with respect to $\mathfrak{h}$. This weight space decomposition is

$$\mathfrak{a} = (\mathfrak{h} \cap \mathfrak{a}) \oplus \sum_{\alpha \in \Phi} (\mathfrak{g}(A)_\alpha \cap \mathfrak{a}).$$

For let $x \in \mathfrak{a}$ with $x = x_0 + \sum_{\alpha \in \Phi} x_\alpha$ where $x_0 \in \mathfrak{h}$ and $x_\alpha \in \mathfrak{g}(A)_\alpha$. We show that $x_0 \in \mathfrak{a}$ and each $x_\alpha \in \mathfrak{a}$. Now the vector space $\mathfrak{h}$ over the infinite field $\mathbb{C}$ cannot be expressed as the union of finitely-many proper subspaces. Fixing α, one easily checks that for each $\beta \in \Phi$ with $\beta \neq \alpha$, the set of $h \in \mathfrak{h}$ satisfying $\beta(h) = \alpha(h)$ is a proper subspace. We also know that $\ker \alpha$ is a proper subspace of $\mathfrak{h}$. Thus there exists an $h \in \mathfrak{h}$ such that $\alpha(h) \neq 0$ and $\beta(h) \neq \alpha(h)$ for all $\beta \in \Phi$ with $\beta \neq \alpha$. Then

$$\operatorname{ad} h \prod_{\substack{\beta \in \Phi \\ \beta \neq \alpha}} (\operatorname{ad} h - \beta(h)1)x = \alpha(h) \prod_{\substack{\beta \in \Phi \\ \beta \neq \alpha}} (\alpha(h) - \beta(h))x_\alpha.$$

The left-hand side is in $\mathfrak{a}$ since $\mathfrak{a}$ is invariant under $\operatorname{ad} h$. Hence $x_\alpha \in \mathfrak{a}$. Since this is true for all $\alpha \in \Phi$, we also have $x_0 \in \mathfrak{a}$. It follows that

$$\mathfrak{a} = (\mathfrak{h} \cap \mathfrak{a}) \oplus \sum_{\alpha \in \Phi} (\mathfrak{g}(A)_\alpha \cap \mathfrak{a}).$$

I claim that $\mathfrak{g}(A)_\alpha \cap \mathfrak{a} = 0$ for each $\alpha \in \Phi$. For otherwise we would have $\mathfrak{g}(A)_\alpha \subset \mathfrak{a}$. Now $\alpha = w(\alpha_i)$ for some $w \in W$ and $i \in \{1, \ldots, l\}$. By Proposition 7.21, we can find an automorphism of $\mathfrak{g}(A)$ that transforms $\mathfrak{g}(A)_\alpha$ into $\mathfrak{g}(A)_{\alpha_i}$. Since $\mathfrak{a}$ is invariant under all automorphisms, we would have $\mathfrak{g}(A)_{\alpha_i} \subset \mathfrak{a}$. Hence $e_i \in \mathfrak{a}$. But then $[e_i, f_i] = h_i \in \mathfrak{a}$, and we would have $[h_i, e_i] = 2e_i \neq 0$, contradicting the fact that $\mathfrak{a}$ is abelian. Thus $\mathfrak{g}(A)_\alpha \cap \mathfrak{a} = 0$ for all $\alpha \in \Phi$, and so $\mathfrak{a} \subset \mathfrak{h}$. Now let $x \in \mathfrak{a}$. Then for $i = 1, \ldots, l$, we have $[x, e_i] = \alpha_i(x) e_i \in \mathfrak{a}$, yielding $\alpha_i(x) = 0$. Since $\alpha_1, \ldots, \alpha_l$ span $\mathfrak{h}^*$, we have $\lambda(x) = 0$ for all $\lambda \in \mathfrak{h}^*$. Thus $x = 0$. It follows that $\mathfrak{a} = 0$, a contradiction. □

Proposition 7.34. *$\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}(A)$.*

Proof. Since we already know that $\mathfrak{h}$ is abelian, it suffices to show that $\mathfrak{h} = N(\mathfrak{h})$. Let $x \in N(\mathfrak{h})$. Then $x = h' + \sum_{\alpha \in \Phi} \lambda_\alpha e_\alpha$ for $h' \in \mathfrak{h}$, $e_\alpha \in \mathfrak{g}(A)_\alpha$. Then for all $h \in \mathfrak{h}$, we have

$$[h, x] = \sum_{\alpha \in \Phi} \lambda_\alpha \alpha(h) e_\alpha \in \mathfrak{h}.$$

But we can find an $h \in \mathfrak{h}$ such that $\alpha(h) \neq 0$ for all $\alpha \in \Phi$. We deduce that $\lambda_\alpha = 0$ for all $\alpha \in \Phi$, and hence $x \in \mathfrak{h}$. □

Proposition 7.35. *$\mathfrak{g}(A)$ is a simple Lie algebra with Cartan matrix A.*

Proof. The Cartan decomposition of $\mathfrak{g}(A)$ with respect to $\mathfrak{h}$ is

$$\mathfrak{g}(A) = \mathfrak{h} \oplus \sum_{\alpha \in \Phi} \mathfrak{g}(A)_\alpha.$$

Thus Φ is the root system of $\mathfrak{g}(A)$. The Cartan matrix $A' = \left(A'_{ij}\right)$ of $\mathfrak{g}(A)$ is determined by the relations

$$s_i(\alpha_j) = \alpha_j - A'_{ij}\alpha_i.$$

But we also have

$$s_i(h_j) = h_j - A_{ji} h_i \qquad \text{by Proposition 7.18.}$$

Thus for any $k = 1, \ldots, l$, we have

$$\alpha_j(s_i h_k) = \alpha_j(h_k) - A_{ki}\alpha_j(h_i) = \alpha_j(h_k) - A_{ij}\alpha_i(h_k) = (\alpha_j - A_{ij}\alpha_i) h_k.$$

Since $\alpha_j(s_i h_k) = (s_i \alpha_j) h_k$, we deduce that

$$s_i(\alpha_j) = \alpha_j - A_{ij}\alpha_i.$$

Hence $A' = A$, and so the Cartan matrix of $\mathfrak{g}(A)$ is A.

Since the Dynkin diagram that determined A was assumed to be connected, $\mathfrak{g}(A)$ must be a simple Lie algebra by Proposition 6.11. □

Thus for each Cartan matrix on the standard list in Section 6.4, we have constructed a finite-dimensional simple Lie algebra $\mathfrak{g}(A)$ with Cartan matrix A. We summarize our results in the final theorem.

Main Theorem. *The finite-dimensional non-trivial simple Lie algebras over $\mathbb{C}$ are*

$$\begin{array}{ll} A_l & l \geq 1 \\ B_l & l \geq 2 \\ C_l & l \geq 3 \\ D_l & l \geq 4 \\ E_6,\ E_7,\ E_8 & \\ F_4 & \\ G_3 & \end{array}$$

These Lie algebras are pairwise non-isomorphic.

Proof. For each Cartan matrix on the standard list in Section 6.4, there exists a finite-dimensional simple Lie algebra which, by Theorem 7.5, is determined up to isomorphism. Simple Lie algebras with different Cartan matrices cannot be isomorphic since, by Proposition 6.4, the Cartan matrix on the standard list is uniquely determined by the Lie algebra. □

Bibliography

[Car05] Roger W. Carter. *Lie Algebras of Finite and Affine Type*. Cambridge University Press, Cambridge, 2005.

[DF04] David S. Dummit and Richard M. Foote. *Abstract Algebra (Third Edition)*. Wiley, Hoboken, NJ, 2004.

[Hal74] Paul R. Halmos. *Finite-Dimensional Vector Spaces*. Springer-Verlag, New York, 1974.

[HK71] Kenneth Hoffman and Ray Kunze. *Linear Algebra (Second Edition)*. Prentice-Hall, Upper Saddle River, NJ, 1971.

[Jac71] Nathan Jacobson. *Lie Algebras*. Dover, Mineola, NY, 1971.

[Ser92] Jean-Pierre Serre. *Lie Algebras and Lie Groups: 1964 Lectures given at Harvard University*. Springer-Verlag, Berlin, 1992.