Gradings of Lie Algebras

Thomas Leenen Meyer
Supervised by Dr Juana Sánchez-Ortega
Dissertation presented for the degree of Master of Science in the
Department of Mathematics and Applied Mathematics
University of Cape Town

October, 2021

Plagiarism Declaration

I know the meaning of plagiarism and declare that all of the work in the dissertation (or thesis), save for that which is properly acknowledged, is my own.

Signature:. Signed by candidate Date: 22/10/2021
…………………………..

Acknowledgements

I would like to thank Dr Juana Sánchez-Ortega for all the time and effort she has spent encouraging, guiding, and helping me. Under difficult circumstances you have been a source of mathematical, mental, and emotional support.

Dr Cristina Draper is owed acknowledgement for all the time she has invested in teaching and helping me this year. You have acted as a second supervisor and I have enjoyed learning and growing under your guidance, thank you.

I am grateful to my loving family for their ongoing emotional, mental, and financial support. My patient partner, Gabbi, has held space for me throughout all the feelings of stress, excitement, frustration, and relief that have arisen in writing this dissertation.

Introduction

The main focus of this dissertation is to present an introduction to gradings of Lie algebras. The aim is twofold: to lay the necessary foundations to become (in the near future) an algebraist working in this area of research, and to tackle the problem of finding and classifying the Lie algebras arising as graded contractions of a specific $\mathbb{Z}_2^3$-grading of the Lie algebra $\mathfrak{g}_2$. As a result, this dissertation consists of six chapters and four appendices which might appear very different, at first glance, but they are indeed connected since they are all very important tools for anyone interested in gradings of Lie algebras.

The first chapter is devoted to introducing Lie algebras and the study of semisimple Lie algebras. This chapter will place into context much of the work of the second chapter. We begin by describing the basic notions of Lie algebras and their representations, and continue with the study of the Killing form of a Lie algebra, as well as, the root space decomposition. The material in this chapter is based on [8, 6].

In the second chapter we study root systems and their bases. This leads to an investigation of the Weyl group associated to a root system. This work allows us to describe how one can uniquely extend isomorphisms between root systems to isomorphisms between Lie algebras to which those root systems correspond. We are then able to describe the special properties of Chevalley bases. The work in this chapter is based on the text and exercises from [8]

Gradings make their first appearance in Chapter three. We quickly shift our focus to group gradings. We describe a process to obtain a universal grading group amongst equivalent gradings. We spend some time preparing and presenting an example of this process. The chapter ends with some results relating to the automorphisms of a grading. Although most of this chapter is based on the work in [4], some of the later results and examples owe acknowledgement to [11].

We present the construction of the exceptional Lie algebra $\mathfrak{g}_2$ in the fourth chapter. This chapter uses some definitions and results which are presented in Appendix F. We start by looking at useful results relating to alternative algebras. Then we introduce upper bounds to the dimension of $\mathfrak{g}_2$. Finally we show that $\mathfrak{g}_2$ is 14-dimensional and we construct an important

$\mathbb{Z}_2^3$-grading of $\mathfrak{g}_2$. Much of the work in this chapter is based on hand-written notes sent in personal correspondence [3].

In the fifth and final chapter we study graded contractions. This work continues into Appendix A, however this is the newest work and as such is still under revision. It is worth mentioning here that the bulk of Chapter 5 and Appendix A is original work under construction. It is the result of an ongoing collaboration with Dr Cristina Draper and Dr Juana Sánchez-Ortega. Although some of the proofs may be shortened in the future, we decided to include them as we are excited about the findings.

After introducing the notions for general Lie algebras and gradings we look specifically at the grading on $\mathfrak{g}_2$ which we constructed in the previous chapter. We are now in a position to attack the problem of finding and classifying the graded contractions relating to the $\mathbb{Z}_2^3$-grading of $\mathfrak{g}_2$ presented in Chapter 4. The definitions in the first section of this chapter come from [2]. The rest of Chapter 6 and Appendix A consist of original work completed for this dissertation.

Tensor products of modules over a commutative ring R are the sole focus of Appendix B. We explicitly construct the tensor product of two R-modules, and see how all multi-linear maps filter through tensor products. This is followed by a collection of results chosen to help build intuition for the structure and workings of the tensor product. Lastly, we examine how tensor products interact with direct sums and how linear maps, between modules, may induce maps between the tensor products of those modules. This chapter is based on [1].

Appendix C is centred around affine group schemes. We introduce the topic as familiarity with this area presents opportunities for future research problems and investigations. Our main aim in this chapter is to describe Hopf algebras. The work in this appendix is based on [10, 12].

Appendix D is focussed on presenting a proof of Weyl's Theorem, used in the third chapter. Such a proof requires results about the Jordan canonical form of a matrix and the Casimir operator of a Lie algebra representation. The work in this appendix is based on [8, 6, 13].

The main goal of Appendix E is to describe the differential of a Lie group homomorphism. We make use of this work in chapter 3. Before we can study the differential of a Lie group homomorphism we need to study matrix Lie groups and the exponential map. The work in this appendix is based on [7, 9].

The last appendix, Appendix F, is a brief summary of important definitions and results related to the octonions. We need this work to accomplish our goals in Chapter six. The material in this appendix is based on [3].

Contents

Chapter 1

The Basics

In this chapter we introduce the fundamentals of Lie algebras and their representations. Throughout the text we will assume that $\mathbb{F}$ denotes a field with $char(\mathbb{F}) \neq 2$. We assume, unless otherwise specified that all vector spaces are finite-dimensional and defined over $\mathbb{F}$.

1.1 What is a Lie algebra?

Definition 1.1: *A* **Lie algebra** *L is a vector space endowed with a bilinear map $[-,-]\colon L \times L \to L$, called the* **Lie bracket**, *satisfying the following two identities for all elements $x, y, z \in L$:*

(L1) $[x, x] = 0$, *and*

(L2) $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0.$

The identity (L2) is called the **Jacobi Identity**.

Remark 1.2: *Notice that the linearity of the Lie bracket implies*

$$\begin{aligned} 0 &\overset{(L1)}{=} [x+y, x+y] = [x, x] + [x, y] + [y, x] + [y, y] \\ &\overset{(L1)}{=} [x, y] + [y, x], \end{aligned}$$

and so $[x, y] = -[y, x]$, for all $x, y \in L$. This property of the Lie bracket is known as **skew-symmetry**.

We close this first section with some examples of Lie algebras.

Example 1.3: *Let V be any vector space (finite- or infinite-dimensional). We may then define a Lie bracket for V by $[u, v] = 0$, for all $u, v \in V$. This is called the* **abelian** *Lie algebra structure on the vector space V. The field $\mathbb{F}$ as a vector space over itself may be regarded as a 1-dimensional abelian Lie algebra.*

Example 1.4: *Every associative algebra A gives rise to a Lie algebra A^- with the same underlying vector space as A. The Lie bracket may be defined as the commutator, that is, $[a,b] = ab - ba$ for all elements $a, b \in A$ (where ab denotes the product in A). This is bilinear because the associative algebra's product is bilinear. For $a, b, c \in A$ we have that $[a,a] = aa - aa = 0$, proving (L1), and*

$$\begin{aligned}
[a,[b,c]] + [b,[c,a]] + [c,[a,b]] &= [a, bc - cb] + [b, ca - ac] + [c, ab - ba] \\
&= a(bc - cb) - (bc - cb)a + b(ca - ac) \\
&- (ca - ac)b + c(ab - ba) - (ab - ba)c \\
&= (abc - abc) + (acb - acb) + (bca - bca) \\
&+ (cba - cba) + (bac - bac) + (cab - cab) = 0.
\end{aligned}$$

Hence, A endowed with the commutator is a Lie algebra.

Example 1.5: *Let V be a vector space of dimension n. We let $End(V)$ denote the vector space of all linear maps from V to V with the usual addition of maps and the scalar multiplication*

$$(\lambda f)(x) = \lambda f(x), \quad \forall\, \lambda \in \mathbb{F},\ f \in End(V),\ x \in V.$$

We may then endow $End(V)$ with an associative algebra structure, with the product being composition of maps. Using the construction in Example 1.4, $End(V)^-$ is a Lie algebra with the Lie bracket defined as

$$[f,g] = f \circ g - g \circ f \quad \forall\, f, g \in End(V).$$

This is known as the **general linear algebra** *of V. We will denote it by $\mathfrak{gl}(V)$ to distinguish the Lie algebra from the vector space.*

Let $\mathfrak{gl}(n,\mathbb{F})$ denote the vector space of $n \times n$ matrices with entries in $\mathbb{F}$. By fixing a basis for V we may identify $\mathfrak{gl}(V)$ with $\mathfrak{gl}(n,\mathbb{F})$. We let e_{ij} denote the $n \times n$ matrix with the ij-th entry 1 and the rest 0, and remark that $\mathfrak{gl}(n,\mathbb{F})$ has dimension n^2 with basis $\{e_{ij} \,|\, 1 \leq i,j \leq n\}$. Its Lie bracket is given by

$$[x,y] = xy - yx, \quad \forall\, x, y \in \mathfrak{gl}(n,\mathbb{F}).$$

Notice that

$$[e_{ij}, e_{kl}] = \delta_{jk} e_{il} - \delta_{il} e_{kj}, \tag{1.1}$$

where $\delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases}$

Subalgebras and Ideals

We now turn our attention to important vector subspaces of Lie algebras: subalgebras and ideals.

Definition 1.6: *Given a Lie algebra L, a vector subspace K of L is a* **Lie subalgebra** *of L if it is closed under the Lie bracket, that is,*

$$[x, y] \in K, \quad \forall\, x, y \in K.$$

Clearly a Lie subalgebra is also a Lie algebra in its own right. Properties (L1) and (L2) are inherited from the original Lie algebra.

A subalgebra of $\mathfrak{gl}(n, F)$ is called a ***linear Lie algebra***. Our next two examples will be linear Lie algebras.

Example 1.7: *Let $\mathfrak{b}(n, \mathbb{F})$ denote the vector subspace of $\mathfrak{gl}(n, \mathbb{F})$ consisting of upper-triangular matrices. It is straightforward to check that $\{e_{ij} \mid i \leq j\}$ is a basis for $\mathfrak{b}(n, \mathbb{F})$. By (1.1) we have that $[e_{ij}, e_{kl}] = 0$ if $l < i$, since $k \leq l < i$ and so, $j \neq k$. Whereas, if $j < k$ then $i \leq j < k \leq l$, thus $i \neq l$ which means that again $[e_{ij}, e_{kl}] = 0$. This shows that $\mathfrak{b}(n, \mathbb{F})$ is a subalgebra of $\mathfrak{gl}(n, \mathbb{F})$.*

Definition 1.8: *Given a Lie algebra L, a vector subspace K of L is an* **ideal** *of L if $[x, y] \in K$ for all $x \in K$ and $y \in L$.*

Remark 1.9: *When talking about ideals of associative rings one distinguishes between left and right ideals. This distinction is not needed in the context of Lie algebras due to the skew-symmetric property of the Lie bracket.*

All ideals are subalgebras but subalgebras need not necessarily be ideals. Example 1.7 provides an example of a subalgebra of $\mathfrak{gl}(n, \mathbb{F})$ which is not an ideal. To see why $\mathfrak{b}(n, \mathbb{F})$ is not an ideal notice that

$$[e_{22}, e_{21}] = \delta_{22}e_{21} - \delta_{21}e_{22} = e_{21} \notin \mathfrak{b}(n, \mathbb{F}).$$

Example 1.10: *Let $\mathfrak{sl}(n, \mathbb{F})$ denote the vector subspace of $\mathfrak{gl}(n, \mathbb{F})$ consisting of all matrices having trace zero. This is an ideal of $\mathfrak{gl}(n, \mathbb{F})$ since $tr(xy) = tr(yx)$ and $tr(x + y) = tr(x) + tr(y)$. This is known as the* **special linear algebra***.*

Example 1.11: *The* **centre** *of a Lie algebra L is defined as*

$$Z(L) = \{z \in L \mid [z, v] = 0, \quad \forall\, v \in L\}$$

The centre is clearly an ideal of L. We can also see that $Z(L) = L$ if and only if L is abelian, that is, $[L, L] = 0$.

Constructing New Lie Algebras

Let us look at two ways in which we may construct new ideals from existing ones. Let L be a Lie algebra with ideals I and J. Then

$$I + J = \{x + y \mid x \in I, y \in J\} \text{ and } [I, J] = \mathrm{Span}\{[x, y] \mid x \in I, y \in J\}$$

are both ideals (and subspaces) of L. To see this, take $x \in I$, $y \in J$ and $v \in L$, then

$$[x + y, v] = [x, v] + [y, v] \in I + J$$

and,

$$[[x, y], v] = -[[y, v], x] - [[v, x], y] \in [I, J]$$

since $[y, v] \in J$ and $[v, x] \in I$.

In particular, for $I = J = L$, we obtain that $[L, L]$ is an ideal of L. The ideal $[L, L]$ is called the ***derived algebra*** of L and denoted L'. Notice that L is abelian if and only if $L' = 0$.

Homomorphisms

The next basic concept we need is that of a map which preserves the Lie algebra structure. As Lie algebras are also vector spaces we would naturally like to work with linear maps but we would also like these maps to 'behave well' with the Lie bracket.

Definition 1.12: *A* **homomorphism** *between two Lie algebras L_1 and L_2 is a linear map $\psi\colon L_1 \to L_2$ such that*

$$\psi([x, y]) = [\psi(x), \psi(y)], \quad \forall\ x, y \in L_1.$$

Monomorphisms, epimorphisms *and* **isomorphisms** *are injective, surjective and bijective homomorphisms, respectively.*

Example 1.13: *Let L be an arbitrary Lie algebra. The map*

$$ad\colon L \to \mathfrak{gl}(L),$$

given by $ad(x) = ad_x$, where $ad_x\colon L \to L$ is defined by $ad_x(y) = [x, y]$ for all $y \in L$, is a homomorphism of Lie algebras.

In fact, for $x, y, z \in L$ and $\alpha, \beta \in \mathbb{F}$, we have that

$$(\alpha ad_x + \beta ad_y)(z) = \alpha[x, z] + \beta[y, z] = [\alpha x + \beta y, z] = ad_{\alpha x+\beta y}(z),$$

which shows that ad is linear. Furthermore:

$$\begin{aligned} ad_{[x,y]}(z) &= [[x, y], z] = -[[y, z], x] - [[z, x], y] \\ &= [x, [y, z]] - [y, [x, z]] = (ad_x \circ ad_y - ad_y \circ ad_x)(z) \\ &= [ad_x, ad_y](z), \end{aligned}$$

which shows that ad preserves the Lie bracket. This map is called the **adjoint homomorphism**.

Proposition 1.14: *Let $\psi\colon L_1 \to L_2$ be a Lie algebra homomorphism. Then* $\mathrm{Ker}(\psi)$ *is an ideal of L_1 and* $\mathrm{Im}(\psi)$ *is a subalgebra of L_2.*

Proof. Since ψ is a linear map we know that $\mathrm{Ker}(\psi)$ and $\mathrm{Im}(\psi)$ are vector subspaces of L_1 and L_2, respectively. For $x \in \mathrm{Ker}(\psi)$ and $y \in L_1$ we have

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

which shows that $[x, y] \in \mathrm{Ker}(\psi)$, as required. Now if $\psi(a), \psi(b) \in \mathrm{Im}(\psi)$,

$$[\psi(a), \psi(b)] = \psi([a, b]) \in \mathrm{Im}(\psi),$$

and so $\mathrm{Im}(\psi)$ is a subalgebra of L_2. □

Proposition 1.15: *Let $\psi\colon L_1 \to L_2$ be a Lie algebra homomorphism.*

(i) If ψ is an epimorphism, then $\psi(L_1') = L_2'$.

(ii) If ψ is an isomorphism, then $\psi(Z(L_1)) = Z(L_2)$.

Proof. The proof of (i) follows directly from ψ being a Lie algebra epimorphism. To see that (ii) holds note that

$$\begin{aligned} x \in \psi(Z(L_1)) &\iff x = \psi(z) \text{ for some } z \in Z(L_1) \\ &\iff [\psi(z), \psi(y)] = \psi([z, y]) = 0, \quad \forall\ \psi(y) \in L_2. \end{aligned}$$

□

1.2 Quotient Lie Algebras

We can use ideals to define quotient algebras for Lie algebras. This will lead us to the familiar Isomorphism Theorems and an important correspondence.

Let L be a Lie algebra and I an ideal of L. For $x \in L$, we may then consider the quotient vector space $L/I = \{x+I \mid x \in L\}$ where the (additive) coset $x + I$ is defined as $\{x + u \mid u \in I\}$. To define a Lie bracket for L/I we set

$$[x + I, y + I] = [x, y] + I, \quad \forall\ x, y \in L.$$

To check that this is well-defined take $x, x', y, y' \in L$ such that $x+I = x'+I$ and $y + I = y' + I$. This tells us that $x - x' = u \in I$ and $y - y' = v \in I$. Hence,

$$\begin{aligned} [x + I, y + I] &= [x, y] + I = [u + x', v + y'] + I \\ &= ([u, v] + [u, y'] + [x', v] + [x', y']) + I \\ &= [x', y'] + I = [x' + I, y' + I]. \end{aligned}$$

Fix $x \in L$, then for $y, z \in L$ and $\alpha, \beta \in \mathbb{F}$:

$$\begin{aligned} [x + I, \alpha(y + I) + \beta(z + I)] &= (\alpha[x, y] + I) + (\beta[x, z] + I) \\ &= \alpha[x + I, y + I] + \beta[x + I, z + I]. \end{aligned}$$

This proves that fixing the first argument induces a linear map. In much the same way one can show that fixing the second argument induces a linear map. This proves bilinearity.

Let us now prove identity (L1). For all $x \in L$ we have that

$$[x + I, x + I] = [x, x] + I = I.$$

Lastly, for $x, y, z \in L$, we have that

$$\begin{aligned}&[x + I, [y + I, z + I]] + [y + I, [z + I, x + I]] + [z + I, [x + I, y + I]\\&= ([x, [y, z]] + [y, [z, x]] + [z, [x, y]]) + I = I,\end{aligned}$$

which shows that identity (L2) holds. Hence, L/I is a Lie algebra over the same field as L, called the ***quotient algebra*** of L by I.

Proposition 1.16: *Let I an ideal of a Lie algebra L. There is a bijective correspondence between the ideals of L/I and the ideals of L containing I. In fact, an ideal $\widetilde{K}$ of L/I corresponds to $K := \{x \in L \,|\, x + I \in \widetilde{K}\}$. While if J is an ideal of L which contains I, then it corresponds to J/I, which is an ideal of L/I.*

Proof. Let us map an ideal $\widetilde{K}$ of L/I to $K := \{x \in L \,|\, x + I \in \widetilde{K}\}$. We must check that K is an ideal of L and that K contains I. Take $x \in K$ and $y \in L$, then $[x, y] + I = [x + I, y + I] \in \widetilde{K}$, since $\widetilde{K}$ is an ideal of L/I. This shows that K is an ideal of L. We can also see that for all $x \in I$ we have that $x + I = I \in K$, which shows that I is contained in J.

On the other hand, let us map an ideal J of L which contains I to J/I. We need to show that J/I is an ideal of L/I. Take $x \in J$, $y \in L$ then $[x, y] \in J$. Thus, $[x + I, y + I] = [x, y] + I \in J/I$.

Lastly, notice that these maps are inverses of one another. □

1.3 Isomorphism Theorems

In this section we focus on the isomorphism theorems for Lie algebras. We go on to introduce a new type of linear maps which will play an important role as we continue our investigation of Lie algebras.

Theorem 1.17 (Isomorphism Theorems)**:**

(i) If $\psi\colon L_1 \to L_2$ is a Lie algebra homomorphism, then

$$L/\operatorname{Ker}(\psi) \cong \operatorname{Im}(\psi).$$

(ii) For arbitrary ideals I and J of a Lie algebra

$$(I + J)/J \cong I/(I \cap J).$$

(iii) Let I and J be ideals of a Lie algebra L with $I \subseteq J$. Then J/I is an ideal of L/I and

$$(L/I)/(J/I) \cong L/J.$$

Proof.

(i): Let $K := \mathrm{Ker}(\psi)$. By the isomorphism theorems for vector spaces we have a vector space isomorphism $L/K \cong \mathrm{Im}(\psi)$, given by

$$\begin{aligned} \widetilde{\psi} \colon L/K &\to \mathrm{Im}(\psi), \\ x + K &\mapsto \psi(x), \end{aligned}$$

for all $x \in L$. To extend this to a Lie algebra isomorphism it suffices to show that $\widetilde{\psi}$ preserves the Lie bracket. Taking $x, y \in L$, we have that

$$\widetilde{\psi}([x+K, y+K]) = \psi([x,y]) = [\psi(x), \psi(y)] = [\widetilde{\psi}(x+K), \widetilde{\psi}(y+K)],$$

as required. Thus, $\widetilde{\psi}$ is a Lie algebra homomorphism.

(ii): Define $\psi \colon I+J \to I/(I \cap J)$ by $\psi(x+y) = x + I \cap J$, for all $x \in I,\ y \in J$. First, we show that this is a linear map. For $x, x' \in I,\ y, y' \in J$ and $\alpha, \beta \in \mathbb{F}$, we have that

$$\begin{aligned} \alpha\psi(x+y) + \beta\psi(x'+y') &= (\alpha x + \beta x') + I \cap J \\ &= \psi(\alpha x + \beta x' + \alpha y + \beta y') \\ &= \psi(\alpha(x+y) + \beta(x'+y')). \end{aligned}$$

Now, we show that ψ is a Lie algebra homomorphism. For all $x, x' \in I$ and $y, y' \in J$, we have that

$$\begin{aligned} \psi([x+y, x'+y']) &= \psi([x,x'] + [x,y'] + [y,x'] + [y,y']) \\ &= [x,x'] + I \cap J \\ &= [\psi(x+y), \psi(x'+y')]. \end{aligned}$$

Let us show that $\mathrm{Ker}(\psi) = J$. If $x \in I$ and $y \in J$ then,

$$x + y \in \mathrm{Ker}(\psi) \iff x + y \in J.$$

Notice that ψ is surjective since if $x + I \cap J \in I/(I \cap J)$, then

$$x + I \cap J = \psi(x + 0).$$

An application of (i) gives the desired result.

(iii) From Proposition 1.16 we have that J/I is an ideal of L/I. Define a map $\psi\colon L/I \to L/J$ by $\psi(x+I) = x+J$. This is well-defined because $x+I = x'+I$ if and only if $x - x' \in I \subseteq J$.

We first prove the linearity of ψ. For $x, x' \in L$ and $\alpha, \beta \in \mathbb{F}$, we have that

$$\begin{aligned}\psi(\alpha(x+I) + \beta(x'+I)) &= \alpha(x+J) + \beta(x'+J) \\ &= \alpha\psi(x+I) + \beta\psi(x'+I)\end{aligned}$$

Next, we show that ψ is a Lie algebra homomorphism. To do so, take $x, x' \in L$

$$\psi([x+I, x'+I]) = [x, x'] + J = [\psi(x+I), \psi(x'+I)].$$

Lastly, to see that $\mathrm{Ker}(\psi) = J/I$, notice that

$$x+I \in \mathrm{Ker}(\psi) \iff x \in J \iff x+I \in J/I.$$

Since $I \subseteq J$, ψ must be surjective. Now (iii) follows from (i).

□

Example 1.18: *We claim that $\mathfrak{gl}(n,\mathbb{F})/\mathfrak{sl}(n,\mathbb{F}) \cong \mathbb{F}$. In fact, consider the map $tr\colon \mathfrak{gl}(n,\mathbb{F}) \to \mathbb{F}$ which sends a matrix to its trace. Well known properties of the trace of a matrix tell us that tr is linear. Furthermore,*

$$tr([x,y]) = tr(xy - yx) = tr(xy) - tr(yx) = 0,$$

which shows that tr is a Lie algebra homomorphism. Clearly, tr is an epimorphism. Lastly, it is clear that $\mathrm{Ker}(tr) = \mathfrak{sl}(n,\mathbb{F})$. An application of Theorem 1.17 gives that $\mathfrak{gl}(n,\mathbb{F})/\mathfrak{sl}(n,\mathbb{F}) \cong \mathbb{F}$, as we claimed.

Definition 1.19: *Let L be a Lie algebra. A linear map $D\colon L \to L$ is called a* **derivation** *of L if it has the property*

$$D([a,b]) = [a, D(b)] + [D(a), b], \quad \forall\, a, b \in L.$$

The set of derivations of L is denoted by $Der(L)$.

Proposition 1.20: *Let L be a Lie algebra, then*

(i) $Der(L)$ is a subalgebra of $\mathfrak{gl}(L)$.

(ii) $ad(L)$ is an ideal of $Der(L)$, whose elements are called inner derivations of L.

(iii) $L/Z(L) \cong ad(L)$.

Proof.

(i): Let $D, E \in Der(L)$ and $\alpha, \beta \in \mathbb{F}$. Then

$$\begin{aligned}(\alpha D + \beta E)([a,b]) &= \alpha D([a,b]) + \beta E([a,b]) \\ &= [a, (\alpha D + \beta)(b)] + [(\alpha D + \beta E)(a), b].\end{aligned}$$

Hence, $\mathrm{Der}(L)$ is closed under linear combinations. The set $\mathrm{Der}(L)$ is non-empty since the zero map is a derivation, so $\mathrm{Der}(L)$ is a vector subspace of $\mathfrak{gl}(L)$. To see that $\mathrm{Der}(L)$ is closed under the Lie bracket take $a, b \in L$ and $D, E \in Der(L)$.

$$\begin{aligned}[D,E]([a,b]) &= (D \circ E)([a,b]) - (E \circ D)([a,b]) \\ &= D([a, E(b)] + [E(a), b]) - E([a, D(b)] + [D(a), b]) \\ &= D([a, E(b)]) + D([E(a), b]) - E([a, D(b)]) \\ &- E([D(a), b]) \\ &= [a, (D \circ E)(b)] + [D(a), E(b)] + [E(a), D(b)] \\ &+ [(D \circ E)(a), b] - [a, (E \circ D)(b)] - [E(a), D(b)] \\ &- [D(a), E(b)] - [(E \circ D)(a), b] \\ &= [a, (D \circ E - E \circ D)(b)] + [(D \circ E - E \circ D)(a), b] \\ &= [a, [D,E](b)] + [[D,E](a), b].\end{aligned}$$

So, $[D, E] \in Der(L)$, as required.

(ii): For $x, a, b \in L$ we have that

$$\begin{aligned}ad_x([a,b]) = [x,[a,b]] &= -[a,[b,x]] - [b,[x,a]] = [a,[x,b]] + [[x,a],b] \\ &= [a, ad_x(b)] + [ad_x(a), b],\end{aligned}$$

which shows that $ad(L) \subseteq Der(L)$. Furthermore, for $x, y \in L$ and $D \in Der(L)$ we have that

$$\begin{aligned}[D, ad_x](y) &= D([x,y]) - [x, D(y)] \\ &= [x, D(y)] + [D(x), y] - [x, D(y)] = [D(x), y] \\ &= ad_{D(x)}(y)\end{aligned}$$

(iii) An application of (i) from Theorem 1.17 (the First Isomorphism Theorem) to the adjoint homomorphism gives the desired result.

□

1.4 Weights

In this section we generalise the notions of eigenvectors and eigenvalues of a linear transformation to families of linear transformations.

Definition 1.21: *Let V be a vector space and L a subalgebra of the Lie algebra $\mathfrak{gl}(V)$. We say that $v \in V$ is an* **eigenvector** *for L if v is an eigenvector for every element of L.*

To generalise the notion of an eigenvalue we need to be a little careful, since the eigenvector v of the subalgebra L of $\mathfrak{gl}(V)$ may have different eigenvalues corresponding to different elements of L. Thus, the generalisation of the notion of an eigenvalue is a map $\lambda\colon L \to \mathbb{F}$, and the corresponding eigenspace is defined as

$$V_\lambda = \{v \in V \mid x(v) = \lambda(x)v, \text{ for all } x \in L\}.$$

We can see that V_λ is non-empty since $0 \in V_\lambda$. Moreover, V_λ is a subspace of V. In fact, for $u, v \in V_\lambda,\ x \in L$, and $\alpha, \beta \in \mathbb{F}$ we have that

$$x(\alpha u + \beta v) = \alpha\lambda(x)u + \beta\lambda(x)v = \lambda(x)(\alpha u + \beta v),$$

which shows that V_λ is a vector subspace of V.

Assume V_λ is non-zero. For $v \in V_\lambda,\ x, y \in L$, and $\alpha, \beta \in \mathbb{F}$ we have that

$$(\alpha x + \beta y)v = \alpha\lambda(x)v + \beta\lambda(y)v = (\alpha\lambda(x) + \beta\lambda(y))v,$$

which shows the linearity of $\lambda\colon L \to \mathbb{F}$, and hence λ belongs to the dual space of L.

Definition 1.22: *Let V be a vector space, and L a subalgebra of the Lie algebra $\mathfrak{gl}(V)$. A linear map $\lambda\colon L \to \mathbb{F}$, such that $V_\lambda \neq 0$, where*

$$V_\lambda = \{v \in V \mid x(v) = \lambda(x)v, \text{ for all } x \in L\},$$

is called a **weight** *for L. We call V_λ the* **weight space** *corresponding to the weight λ.*

Example 1.23: *Consider $L = \mathfrak{b}(n, \mathbb{F})$ as a subalgebra of $\mathfrak{gl}(n, \mathbb{F})$. We will denote by e_1 the column vector with the first entry 1 and the rest 0. Then for any $x = (x_{ij}) \in L$ we have that $xe_1 = x_{11}e_1$, and so e_1 is an eigenvector for L. The associated weight $\lambda\colon L \to \mathbb{F}$ is defined by $\lambda(x) = x_{11}$, for all $x \in L$. We claim that $V_\lambda = Span(e_1)$. To see this we first take $\delta \in \mathbb{F}$ and $x \in L$ and see that $x\delta e_1 = \delta x_{11}e_1 = \lambda(x)\delta e_{11}$, which shows that $Span(e_1) \subseteq V_\lambda$. Now we take $v \in V_\lambda$ and notice that we may write $v = \sum_{i=1}^{n} \beta_i e_i$, for some $\beta_1, \ldots, \beta_n \in \mathbb{F}$. We also have that $xv = x_{11}v$ for all $x \in L$, which implies that $e_{11}v = v$, and hence $\beta_2 = \beta_3 = \ldots = \beta_n = 0$. We then have that $v = \beta_1 e_1 \in Span(e_1)$. Thus, $V_\lambda = Span(e_1)$.*

Lemma 1.24 (Invariance Lemma)**:** *Let V be a vectorspace over a field $\mathbb{F}$ of characteristic 0, and L a subalgebra of the Lie algebra $\mathfrak{gl}(V)$. Let I be an ideal of L, and $\lambda\colon I \to \mathbb{F}$ a weight for I. Then V_λ is an L-invariant vector subspace of V; that is, $y(V_\lambda) \subseteq V_\lambda$, for all $y \in L$.*

Proof. We need to prove that if we take $y \in L$ and $v \in V_\lambda$, then $y(v)$ is an eigenvector of I with $x(y(v)) = \lambda(x)y(v)$, for all $x \in I$. For $x \in I$, we have that

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

which implies that $x(y(v)) = \lambda(x)y(v) + [x, y](v)$. Since I is an ideal of L we have that $[x, y] \in I$, which yields that $[x, y](v) = \lambda([x, y])v$. Hence, the result will follow by showing that $\lambda([x, y]) = 0$.

We now construct a y-invariant subspace of V by considering

$$U = \mathit{Span}(\{y^k(v) \mid k \geq 0\}).$$

Let m be the smallest number such that the vectors $v, y(v), \ldots, y^m(v)$ are linearly dependent. Then $B = \{v, y(v), \ldots, y^{m-1}(v)\}$ is a basis for U. We claim that U is z-invariant for all $z \in I$. Furthermore, we will show that–with respect to B–z acting on U is represented by an upper triangular matrix with the diagonal entries all being $\lambda(z)$. We proceed by induction on the column number. For the first column, $z(v) = \lambda(z)v$. Since $[z, y] = zy - yz$, we can see that in column $k + 1$ we have

$$z(y^k(v)) = zy(y^{k-1}(v)) = ([z, y] + yz)y^{k-1}(v).$$

From the inductive hypothesis we have that $zy^{k-1}(v) = \lambda(z)y^{k-1}(v) + u$, for some $u \in \mathit{Span}(\{v, y(v), \ldots, y^{k-2}(v)\})$. We then have that

$$yzy^{k-1}(v) = \lambda(z)y^k(v) + y(u),$$

where $y(u) \in \mathit{Span}(\{v, y(v), \ldots, y^{r-1}(v)\})$. We may also apply the inductive hypothesis to $[z, y]$ as $[z, y] \in I$. From this we may conclude that $[z, y]y^{k-1}(v) \in \mathit{Span}(\{v, y(v), \ldots, y^{k-1}(v)\})$. These last two results mean that column $k + 1$ indeed has the form we claimed.

Set $z = [x, y]$, from our claim above we know that the trace of z acting on U is $m\lambda(z)$. We also know that U is invariant under the actions of x (we showed this above) and y (this is true because of how U was constructed). Thus, the trace of z acting on U is the same as the trace of $[x, y] = xy - yx$ which is 0. So, $m\lambda([x, y]) = 0$ and since $\mathbb{F}$ has characteristic 0, we must have that $\lambda([x, y]) = 0$. This shows us that $x(y(v)) = \lambda(x)y(v)$ which proves the result. □

1.5 Representations

Given that we have some specific knowledge about linear Lie algebras, it would be nice if we could somehow use it when working with an abstract Lie algebra.

Definition 1.25: *Let L be a Lie algebra. A* **representation** *of L is a Lie algebra homomorphism $\psi\colon L \to \mathfrak{gl}(V)$, where V is some vector space. If ψ is a monomorphism, the representation is said to be* **faithful**.

When referring to a representation we may at times refer to V rather than to the Lie algebra homomorphism ψ.

A representation of the form $\psi\colon L \to \mathfrak{gl}(n, \mathbb{F})$ is called a ***matrix representation***.

Example 1.26: *We have already encountered a representation for any Lie algebra L. Namely, the adjoint homomorphism $ad\colon L \to \mathfrak{gl}(L)$ defined in Example 1.13, is a representation of L. For this reason it may also be called the* **adjoint representation**.

Definition 1.27: *A* **module** *for a Lie algebra L, or L-module, is a vector space V together with a map $L \times V \to V$, denoted $(x, v) \mapsto x \cdot v$, satisfying the following identities for all $x, y \in L,\ v, w \in V$, and $\alpha, \beta \in \mathbb{F}$:*

(M1) $(\alpha x + \beta y) \cdot v = \alpha(x \cdot v) + \beta(y \cdot v)$,

(M2) $x \cdot (\alpha v + \beta w) = \alpha(x \cdot v) + \beta(x \cdot w)$,

(M3) $[x, y] \cdot v = x \cdot (y \cdot v) - y \cdot (x \cdot v)$.

A vector subspace W of V is a **submodule** *of V if W is L-invariant; that is, $x \cdot w \in W$ for all $x \in L,\ w \in W$. In the context of representations the equivalent of a submodule is a* **subrepresentation**.

Proposition 1.28: *Representations of a Lie algebra L, and modules over L are equivalent definitions.*

Proof. Let $\psi\colon L \to \mathfrak{gl}(V)$ be a representation of L. Then we may make V an L-module by defining $x \cdot v = \psi(x)(v)$, for all $x \in L,\ v \in V$. To verify this, take $x, y \in L,\ v, w \in V$, and $\alpha, \beta \in \mathbb{F}$ and have

$$(\alpha x + \beta y) \cdot v = \alpha\psi(x)(v) + \beta\psi(y)(v) = \alpha(x \cdot v) + \beta(y \cdot v),$$
$$x \cdot (\alpha v + \beta w) = \alpha\psi(x)(v) + \beta\psi(x)(v) = \alpha(x \cdot v) + \beta(x \cdot w),$$

and

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

as required.

Conversely, if we have an L-module V, then we may define a representation of L by $\psi(x)(v) = x \cdot v$, for all $x \in L,\ v \in V$. This map is linear by (M1), it is a Lie algebra homomorphism because of (M3), and it is a representation since $\psi(x) \in \mathfrak{gl}(V)$ for all $x \in L$, by (M2). □

Example 1.29: *Let L be a Lie algebra, and I an ideal of L. If we consider L as an L-module via the adjoint representation, then I is a submodule of L. Certainly, I is a vector subspace of L, and I is L-invariant exactly because it is an ideal of L.*

Proposition 1.30: *Let L be a Lie algebra, V an L-module, and W a submodule of V. We may endow the quotient vector space V/W with the structure of an L-module by setting $x \cdot (v+W) = (x \cdot v) + W$ for all $x \in L,\ v \in V$.*

Proof. We need to check that the action of L on V/W above is well-defined. In fact, take $x \in L$ and $v, v' \in V$ such that $v - v' \in W$, then we have $x \cdot (v+W) - x \cdot (v'+W) = x \cdot (v - v') + W = W$, since $x \cdot (v - v') \in W$. We should also check that the three identities (M1), (M2), and (M3) hold. To do so we take $x, y \in L,\ u, v \in V$, and $\alpha, \beta \in \mathbb{F}$

$$(\alpha x + \beta y) \cdot (v+W) = \alpha(x \cdot v) + \beta(y \cdot v) + W = \alpha(x \cdot (v+W)) + \beta(x \cdot (v+W)),$$

$$x \cdot (\alpha(u+W) + \beta(v+W)) = \alpha(x \cdot (u+W)) + \beta(x \cdot (v+W)),$$

and

$$[x, y] \cdot (v+W) = ([x, y] \cdot V) + W = x \cdot (y \cdot (v+W)) - y \cdot (x \cdot (v+W)).$$

This confirms that V/W is in fact an L-module. We call the L-module V/W the ***quotient*** module of V by W. □

Example 1.31: *Let L be a Lie algebra and I an ideal of L. We saw in Example 1.29 that I is a submodule of L via the adjoint representation. Thus, L/I is a quotient L-module with $x \cdot (y+I) = ad_x(y) + I = [x, y] + I$, for all $x, y \in L$.*

Proposition 1.32: *Let L be a Lie algebra, V an L-module, and W a submodule of V. There is a bijective correspondence between the submodules of V which contain W and the submodules of V/W.*

Proof. If U is a submodule of V which contains W, we map U to U/W. We note that U/W is a vector subspace of V/W because U is a subspace of V, and U/W is L-invariant since U is L-invariant. Thus, U/W is a submodule of V/W.

If $\widetilde{U}$ is a submodule of V/W, we map $\widetilde{U}$ to $U = \{u \in V \mid u + W \in \widetilde{U}\}$. By construction $W \subseteq U$. Since $0 \in U$ we know that U is non-empty. Take $x, y \in U,\ \alpha, \beta \in \mathbb{F}$, then $\alpha x + \beta y + W = \alpha(x+W) + \beta(y+W) \in \widetilde{U}$, which implies that $\alpha x + \beta y \in U$. Hence, U is a vector subspace of V. We need to check that U is L-invariant. To do so take $x \in L,\ u \in U$, then $(x \cdot u) + W \in \widetilde{U}$ means that $x \cdot u \in U$. Thus, U is a submodule of V containing W, as required.

Lastly, notice that these two maps are inverses of one another. □

Definition 1.33: *Let L be a Lie algebra and V an L-module. We say V is* **irreducible** *if $V \neq 0$ and its only submodules are V and 0. We say V is* **completely reducible** *if it can be written as a direct sum of irreducible L-modules.*

Example 1.34: *Let L be a* **simple** *Lie algebra, that is a non-abelian Lie algebra which has no proper, non-trivial ideals. Then L is an irreducible L-module via the adjoint representation.*

Example 1.35: *Let $L = \mathfrak{d}(n, \mathbb{F})$ be the subalgebra of $\mathfrak{gl}(n, \mathbb{F})$ consisting of all diagonal matrices. To confirm that L is a linear Lie algebra notice that the product of diagonal matrices is a diagonal matrix. The inclusion homomorphism $\psi : L \to \mathfrak{gl}(n, \mathbb{F})$ is a matrix representation called the* **natural** *representation of L. The natural L-module is then $V = \mathbb{F}^n$. Let E_i denote $Span(e_i)$ for all $1 \leq i \leq n$. Then E_i is a submodule of V and since it is 1-dimensional it must be irreducible. We may then write $V = \bigoplus_{i=1}^{n} E_i$, and thus V is completely reducible.*

Lie Module Homomorphisms

Having introduced L-modules, we will now look at homomorphisms between them. We will see the familiar Isomorphism theorems and the well-known Schur's lemma.

Definition 1.36: *Let L be a Lie algebra with L-modules U and V. A linear map $\varphi : U \to V$ is called an* **L-module homomorphism** *if $\varphi(x \cdot u) = x \cdot \varphi(u)$, for all $x \in L$ and $u \in U$.*

In the equivalent language of representations let $\phi_U : L \to \mathfrak{gl}(U)$ and $\phi_V : L \to \mathfrak{gl}(V)$ denote representations of L corresponding to the L-modules U and V, respectively. Then the equivalent condition for φ to be a homomorphism between representations of L is that $\varphi \circ \phi_U = \phi_V \circ \varphi$.

Lemma 1.37: *Let L be a Lie algebra, U and V two L-modules. For any L-module homomorphism $\varphi : U \to V$, we have that $\mathrm{Ker}(\varphi)$ is a submodule of U while $\mathrm{Im}(\varphi)$ is a submodule of V.*

Proof. Notice that $\mathrm{Ker}(\varphi)$ and $\mathrm{Im}(\varphi)$ are vector subspaces of U and V, respectively. We need only show that they are L-invariant. To do so take $x \in L$, $u \in \mathrm{Ker}(\varphi)$ and $\varphi(v) \in \mathrm{Im}(\varphi)$ Then

$$\varphi(x \cdot u) = x \cdot \varphi(u) = 0, \text{ and}$$
$$x \cdot \varphi(v) = \varphi(x \cdot v) \in \mathrm{Im}(\varphi),$$

so both $\mathrm{Ker}(\varphi)$ and $\mathrm{Im}(\varphi)$ are L-invariant, as required. □

Theorem 1.38 (Isomorphism Theorems)**:**

(i) For an arbitrary L-module homomorphism $\varphi : U \to V$, we have that $U/\mathrm{Ker}(\varphi) \cong \mathrm{Im}(\varphi)$.

(ii) Let U and W be submodules of an L-module V. Then $U + W$ and $U \cap W$ are submodules of V, and $(U + W)/W \cong U/U \cap W$.

(iii) Let U and W be submodules of an L-module V, with $W \subseteq U$. Then U/W is a submodule of V/W and $(V/W)/(U/W) \cong V/U$.

Proof.

(i): By the Isomorphism Theorem for vector spaces we have that a vector space isomorphism between $U/\operatorname{Ker}(\varphi)$ and $\operatorname{Im}(\varphi)$, given by

$$\widetilde{\varphi}\colon U/\operatorname{Ker}(\varphi) \to \operatorname{Im}(\varphi)$$
$$u + \operatorname{Ker}(\varphi) \mapsto \varphi(u),$$

for all $u \in U$. To extend this to an L-module isomorphism it suffices to prove that $\widetilde{\varphi}$ is a L-module homomorphism. To do so take $x \in L$ and $u \in U$

$$x \cdot \widetilde{\varphi}(u + \operatorname{Ker}(\varphi)) = x \cdot \varphi(u) = \varphi(x \cdot u) = \widetilde{\varphi}(x \cdot u + \operatorname{Ker}(\varphi)),$$

Thus, $\widetilde{\varphi}$ is an L-module homomorphism, as required.

(ii): Notice that $U + W$ and $U \cap W$ are vector subspaces of V. For all $x \in L,\ u \in U,\ w \in W$, we have $x \cdot (u + w) = x \cdot u + x \cdot w \in U + W$, which shows that $U + W$ is L-invariant. Since both U and W are L-invariant, $U \cap W$ must be too. Thus, $U + W$ and $U \cap W$ are submodules of V.

Define $\phi\colon U + W \to U/U \cap W$ by $\phi(u + w) = u + U \cap W$, for all $u \in U,\ w \in W$. By the proof for Theorem 1.17 (ii) we know ϕ is linear, $\operatorname{Ker}(\phi) = W$, and ϕ is a surjective. To see that ϕ is an L-module homomorphism take $x \in L,\ u \in U,\ w \in W$, we then have

$$x \cdot \phi(u + w) = (x \cdot u) + U \cap W = \phi(x \cdot (u + w)).$$

An application of (i) gives the desired result.

(iii) From Proposition 1.32 we have that U/W is a submodule of V/W. Define $\phi\colon V/W \to V/U$ by $\phi(v + W) = (v + U)$, for all $u \in U$. By the proof of Theorem 1.17 (iii) we have that ϕ is linear, $\operatorname{Ker}(\phi) = U/W$, and ϕ is surjective. To see that ϕ is an L-module homomorphism we take $x \in L,\ v \in V$, we then have

$$x \cdot \phi(v + W) = (x \cdot v) + U = \phi(x \cdot (v + W)).$$

An application of (i) gives the desired result.

□

Lemma 1.39 (Schur's Lemma)**:** *Let L be a Lie algebra - defined over an algebraically closed field $\mathbb{F}$ of characteristic 0 - and V an irreducible L-module. The only L-module endomorphisms of V are scalar multiples of the identity.*

Proof. Let $\varphi\colon V \to V$ be an L-module endomorphism. As φ is a linear transformation of an $\mathbb{F}$-vector space, it must have some $\lambda \in \mathbb{F}$ as an eigenvalue with corresponding eigenvector $v \in V$. If we denote the identity linear transformation of V by 1_V, we can see that $\varphi - \lambda 1_V$ is a linear transformation of V. For $x \in L,\ u \in V$ we have that

$$x \cdot (\varphi - \lambda 1_V)(u) = x \cdot \varphi(u) - x \cdot (\lambda u) = \varphi(x \cdot u) - \lambda(x \cdot u) = (\varphi - \lambda 1_V)(x \cdot u),$$

which shows that $\varphi - \lambda 1_V$ is an L-module homomorphism. An application of Lemma 1.37 yields that $\mathrm{Ker}(\varphi - \lambda 1_V)$ is a submodule of V. We also know that it contains the eigenvectors of φ corresponding to the eigenvalue λ, so it contains v and thus is non-zero. Since V is irreducible we have that $\mathrm{Ker}(\varphi - \lambda 1_V) = V$, which implies that $\varphi = \lambda 1_V$. Lastly, notice that all the scalar multiples of 1_V are L-module endomorphisms of V. □

1.6 Solvability, Nilpotency, and Semisimplicity

In this section we describe what it means for a Lie algebra to be solvable, nilpotent, or semisimple. These notions will play an important roles in the rest of this chapter where we lay out the basic theory of semisimple Lie algebras over an algebraically closed field of characteristic zero.

Definition 1.40: *A Lie algebra L is called* **simple** *if $[L, L] \neq 0$, and the only ideals of L are itself and 0. For a simple Lie algebra L it is clear that $[L, L] = L$ and $Z(L) = 0$.*

Example 1.41: *Let $L = \mathfrak{sl}(2, \mathbb{F})$. Then L is simple if $char(\mathbb{F}) \neq 2$.*

Proof. Notice that $\{e_{12},\ e_{21},\ h_1 = e_{11} - e_{22}\}$ is a basis for L. We can then see that $[e_{12}, e_{21}] = h_1,\ [e_{12}, h_1] = -2e_{12},\ [e_{21}, h_1] = 2e_{21}$. Thus, $L \subseteq L'$ and so $L = L'$.

Suppose $I \neq L$ is a non-trivial ideal of L. Then we must have that $\dim(I) = 2$ or $\dim(I) = 1$. If $\dim(I) = 2$, then L/I is 1-dimensional and thus abelian. This implies that $[x, y] \in I$, for all $x, y \in L$, which implies that $L = L' \subseteq I$. This contradicts our assumption that I is a proper ideal of L. If $\dim(I) = 1$, then L/I is 2-dimensional. Fix a basis $\{x + I, y + I\}$ for L/I and take any $z \in I$. Then $\{x, y, z\}$ is a basis for L. This means that $L' = Span(\{[x, y], [y, z], [x, z]\})$. However, since $I = Span(\{z\})$ and I is an ideal of L, we have that $[y, z] = \alpha z$ and $[x, z] = \beta z$ for some $\alpha, \beta \in \mathbb{F}$. Hence $L' = Span(\{[x, y], z\})$, which contradicts the fact that L' is 3-dimensional.

We have then shown that the only ideals of L are 0 and itself, as required. □

Let L be a Lie algebra. The *derived series* of L is defined to be the sequence of ideals with terms

$$L = L^{(0)} \subseteq L^{(1)} \subseteq L^{(2)} \subseteq \dots,$$

where $L^{(0)} = L$, and $L^{(k)} = [L^{(k-1)}, L^{(k-1)}]$ for all $k \geq 1$.

Definition 1.42: *A Lie algebra L is called* **solvable** *if $L^{(k)} = 0$ for some $k \in \mathbb{N}$.*

Example 1.43: *Let $L = \mathfrak{b}(n, \mathbb{F})$. Then L is solvable.*

Proof. We first claim that $B_m = \{e_{ij} \,|\, j - i \geq 2^{m-1}\}$ is a basis for $L^{(m)}$, for each $m \geq 1$. We proceed by induction on m.

In the case where $m = 1$ it suffices to show that for all $e_{ij}, e_{kl} \in L$ we have $[e_{ij}, e_{kl}] \in Span(B_1)$. We know that $[e_{ij}, e_{kl}] = \delta_{jk}e_{il} - \delta_{il}e_{kj}$, since $i \leq j$ and $k \leq l$. If $j \neq k$ and $i \neq l$ then $[e_{ij}, e_{kl}] = 0$. If $j = k$ and $i \neq l$ then $l - i > k - j = 0$, so $[e_{ij}, e_{kl}] = e_{il} \in Span(B_1)$. If $j \neq k$ and $i = l$, then $j - k > i - l = 0$ and so $[e_{ij}, e_{kl}] = -e_{kj} \in Span(B_1)$. Finally, if $j = k$ and $i = l$ then $[e_{ij}, e_{kl}] = e_{ii} - e_{jj} \in Span(B_1)$.
Let $m \geq 1$. In this case it suffices to show that for all $e_{ij}, e_{kl} \in B_m$ we have $[e_{ij}, e_{kl}] \in Span(B_{m+1})$. Now we have that $(j - i), (l - k) \geq 2^{m-1}$. If $j \neq k$ and $i \neq l$ then $[e_{ij}, e_{kl}] = 0 \in Span(B_{m+1})$. If $j = k$ then $l - i \geq 2^{m-1} + 2^{m-1} + k - j = 2^m$. Similarly, if $i = l$ then $j - k \geq 2^m$. This proves the claim.

Using the claim we can see that $L^{(m)} = 0$ whenever $2^{m-1} > n - 1$. □

Lemma 1.44: *Let $\psi\colon L_1 \to L_2$ be a Lie algebra homomorphism. Then $\psi(L_1^{(n)}) = \psi(L_1)^{(n)}$, for all $n \geq 0$.*

Proof. We proceed by induction on n. The case $n = 0$ is trivial. Let $n \geq 1$, then $v \in \psi(L_1^{(n)})$ means that $v = \psi\left(\sum_{i=1}^{m} \alpha_i[x_i, y_i]\right) = \sum_{i=1}^{m} \alpha_i[\psi(x_i), \psi(y_i)]$, for some $\alpha_i \in F$, $x_i, y_i \in L_1^{(n-1)}$. By the inductive hypothesis we have that $\psi(x_i), \psi(y_i) \in \psi(L_1)^{(n-1)}$, which implies that $[\psi(x_i), \psi(y_i)] \in \psi(L_1)^{(n)}$. Thus, $v \in \psi(L_1)^{(n)}$. On the other hand, taking $v \in \psi(L_1)^{(n)}$ means that $v = \sum_{i=1}^{m} \alpha_i[x_i, y_i]$, for some $\alpha_i \in \mathbb{F}$, $x_i, y_i \in \psi(L_1)^{(n-1)}$. By the inductive hypothesis $x_i, y_i \in \psi(L_1^{(n-1)})$, which implies that $[x_i, y_i] \in \psi(L_1^{(n)})$. Thus, $v \in \psi(L_1^{(n)})$. □

Proposition 1.45: *Let L be a Lie algebra.*

(i) If L is solvable, then all subalgebras and homomorphic images of L are solvable.

(ii) If I is a solvable ideal of L such that L/I is solvable, then L is also solvable.

(iii) If I and J are solvable ideals of L, then $I + J$ is solvable.

Proof.

(i) Since L is solvable there is a $k \geq 1$ such that $L^{(k)} = 0$. Let A be a subalgebra of L, B a Lie algebra, and $\psi\colon L \to B$ a Lie algebra homomorphism. Then $A^{(k)} \subseteq L^{(k)} = 0$. Using Lemma 1.44, we have that $\psi(L)^{(k)} = \psi(L^{(k)}) = 0$.

(ii) We have that $(L/I)^{(k)} = I$ for some $k \geq 1$. Applying Lemma 1.44 to the canonical homomorphism $\pi\colon L \to L/I$ we have that

$$(L/I)^{(k)} = (L^{(k)} + I)/I = I,$$

so $L^{(k)} \subseteq I$. Next, proceeding by induction on n we show that $(L^{(k)})^{(n)} = L^{k+n}$. For $n = 0$ this is trivial. Let $n \geq 1$, then

$$(L^{(k)})^{(n)} = [(L^{(k)})^{(n-1)}, (L^{(k)})^{(n-1)}] = [L^{(k+n-1)}, L^{(k+n-1)}] = L^{(k+n)}.$$

I is solvable so $I^{(m)} = 0$ for some $m \geq 1$. Putting this all together we get $L^{(k+m)} = (L^{(k)})^{(m)} \subseteq I^{(m)} = 0$.

(iii) Choose $m, n \geq 1$ such that $I^{(m)} = 0$ and $J^{(n)} = 0$. Applying Theorem 1.17 (the Isomorphism Theorems), Lemma 1.44 and the claim from (ii) we have that $((I+J)^{(m)} + J)/J = (I^{(m)} + I \cap J)/I \cap J = 0$. Since I is solvable an application of (ii) gives the desired result.

□

Corollary 1.46: *Let L be a Lie algebra. There is a unique solvable ideal of L which contains every solvable ideal of L.*

Proof. Let R be a maximal solvable ideal of L; that is, a solvable ideal of the largest possible dimension. Let I be any solvable ideal of L. By Proposition 1.45 we have that $R+I$ is a solvable ideal of L. The maximality of R means that $R + I = R$ and so I is contained in R. □

Definition 1.47: *The unique largest solvable ideal of a Lie algebra L is called the* **radical** *of L and is denoted by $rad(L)$.*

Definition 1.48: *Let L be a non-trivial Lie algebra. If $rad(L) = 0$ then L is said to be* **semisimple**.

Notice that any simple Lie algebra is immediately semisimple. Without the condition that a Lie algebra must be non-abelian to be simple, the 1-dimensional abelian Lie algebra would be simple but not semisimple.

Let L be a Lie algebra. The *lower central series* of L is the sequence of ideals with terms

$$L = L^0 \subseteq L^1 \subseteq L^2 \subseteq \dots,$$

where $L^0 = L$, and $L^k = [L, L^{k-1}]$ for all $k \geq 1$.

Remark 1.49: *Let L be a Lie algebra. Then L^k/L^{k+1} is contained in the centre of L/L^{k+1} for all $k \in \mathbb{N}$.*

Proof. Take $x \in L^k$, $y \in L$. By definition of the lower central series of L we have that $[x, y] \in L^{k+1}$. Hence $[x + L^{k+1}, y + L^{k+1}] = L^{k+1}$. □

Definition 1.50: *A Lie algebra L is* **nilpotent** *if there is a $k \geq 1$ such that $L^k = 0$. We note that $L^{(k)} \subseteq L^k$, and so nilpotency implies solvability.*

Example 1.51: *Let $\mathfrak{n}(n, F)$ denote the vector subspace of $\mathfrak{b}(n, F)$ consisting of all strictly upper triangular matrices. By the proof in Example 1.43 we can see that $\mathfrak{n}(n, F) = (\mathfrak{b}(n, F))^{(1)}$ which tells us that $\mathfrak{n}(n, F)$ is a Lie algebra in its own right. We claim that it is a nilpotent Lie algebra.*

Proof. We will show that $B_k = \{e_{ij} \mid j - i > k\}$ is a basis for L^k for all $k \in \mathbb{N}$. We proceed by induction on k. The case $k = 0$ has already been proved in the proof of Claim 1.43
Let $k \geq 0$. Take $e_{ij} \in L$ and $e_{i'j'} \in L^k$. To prove our claim it suffices to show that $[e_{ij}, e_{i'j'}] \in Span(B^{k+1})$. Note that $j - i > 0$ and $j' - i' > k$ which tells us that $j - i \geq 1$ and $j' - i' \geq k + 1$. We can see that $[e_{ij}, e_{i'j'}] = \delta_{ji'} e_{ij'} - \delta_{ij'} e_{i'j}$. If $i' = j$ then $j' - i \geq k + 1 + i' - i = k + 1 + j - i \geq k + 2$ and so $j' - i > k + 1$, whereas if $i = j'$ then $j - i' \geq 1 + i - i' = 1 + j' - i' \geq k + 2$ and so $j - i' > k + 1$. Thus, $[e_{ij}, e_{i'j'}] \in Span(B^{k+1})$, as required.
Now we can see that $L^n = 0$. □

Lemma 1.52: *Let $\psi \colon L_1 \to L_2$ be a Lie algebra homomorphism. Then $\psi({L_1}^n) = \psi(L_1)^n$, for all $n \geq 0$.*

Proof. We proceed by induction on n. The case $n = 0$ is trivial. Let $n \geq 1$, then $v \in \psi({L_1}^n)$ means that $v = \psi\left(\sum_{i=1}^{m} \alpha_i [x_i, y_i]\right) = \sum_{i=1}^{m} \alpha_i [\psi(x_i), \psi(y_i)]$, for some $\alpha_i \in F$, $x_i \in L_1$, $y_i \in {L_1}^{n-1}$. By the inductive hypothesis we have that $\psi(y_i) \in \psi(L_1)^{n-1}$, which implies that $[\psi(x_i), \psi(y_i)] \in \psi(L_1)^n$. Thus, $v \in \psi(L_1)^n$. On the other hand, taking $v \in \psi(L_1)^n$ means that $v = \sum_{i=1}^{m} \alpha_i [x_i, y_i]$, for some $\alpha_i \in F$, $x_i \in \psi(L_1)$, $y_i \in \psi(L_1)^{n-1}$. By the inductive hypothesis $y_i \in \psi({L_1}^{n-1})$, which implies that $[x_i, y_i] \in \psi({L_1}^n)$. Thus, $v \in \psi({L_1}^n)$. □

Proposition 1.53: *Let L be a Lie algebra.*

(i) If L is nilpotent, then all subalgebras and homomorphic images of L are nilpotent.

(ii) If $L/Z(L)$ is nilpotent, then L is nilpotent.

(iii) If $L \neq 0$ is nilpotent, then $Z(L) \neq 0$.

Proof.

(i) Since L is nilpotent there is some $n \geq 1$ such that $L^n = 0$. Let A be a subalgebra of L, and $\psi \colon L \to K$ a Lie algebra homomorphism. Then $A^n \subseteq L^n = 0$, so A is nilpotent. By Lemma 1.52 we have that $\psi(L)^n = \psi(L^n) = 0$, so $\mathrm{Im}(\psi)$ is nilpotent.

(ii) As $L/Z(L)$ is nilpotent there is some $n \geq 1$ such that $(L/Z(L))^n = 0$. By Lemma 1.52 we have that $0 = (L/Z(L))^n = (L^n + Z(L))/Z(L)$ which means that $L^n \subseteq Z(L)$. Thus, $L^{n+1} = 0$ and so L is nilpotent.

(iii) Suppose, to the contrary, that $Z(L) = 0$. Then $L^1 = L$ which implies that $L^n = L$, for all $n \in \mathbb{N}$. This contradicts L being nilpotent.

□

1.7 Engel's Theorem

In prior sections we introduced the notions of nilpotency and solvability, as well as representations. The adjoint representation will allow us to apply some of this knowledge of representations to abstract Lie algebras. This section and the next focus on Engel's and Lie's theorems, respectively. We begin by studying results linking nilpotency for a Lie algebra and ad-nilpotency of its elements.

Definition 1.54: *Let L be a Lie algebra and K a subalgebra of L. The* normaliser *of K in L is defined by*

$$N_L(K) = \{x \in L \,|\, [x, y] \in K, \ \forall \ y \in K\}.$$

Proposition 1.55: *Let L be a Lie algebra and K a subalgebra of L. Then $N_L(K)$ is the largest subalgebra of L in which K is an ideal.*

Proof. We first prove that $N_L(K)$ is a subalgebra of L. To do so we take $x, y \in N_L(K), \ z \in L, \ \alpha, \beta \in \mathbb{F}$, then we have

$$[\alpha x + \beta y, z] = \alpha[x, z] + \beta[y, z] \in K,$$

which shows that $N_L(K)$ is a vector subspace of L. Moreover, by the Jacobi identity we have

$$[[x, y], z] = -[[y, z], x] - [[z, x], y] \in K,$$

which shows that $[x, y] \in N_L(K)$. Thus, $N_L(K)$ is a subalgebra of L.

We note that K is an ideal of $N_L(K)$ by construction. Suppose J is a subalgebra of L in which K is an ideal. Then $[x, y] \in K$ for all $x \in J, \ y \in L$, and so $J \subseteq N_L(K)$, as required. □

Proposition 1.56: *Let L be a subalgebra of $\mathfrak{gl}(V)$ for some vector space V. If $x \in L$ is nilpotent, then $ad_x \colon L \to L$ is nilpotent.*

Proof. Let $y \in L$ and $k \geq 1$, then $(ad_x)^k(y) = [x, [x, \ldots, [x, y] \ldots]]$ and expanding this using the commutator we get a sum of expressions, all of which we claim are of the form $x^j y x^{k-j}$ for some $0 \leq j \leq k$. We proceed by induction on k. The case where $k = 1$ is clear as $[x, y] = xy - yx$. Fix $k \geq 1$ and assume the result holds for k. We have that $(ad_x)^{k+1}(y) = [x, (ad_x)^k(y)]$, which tells us that each summand of $(ad_x)^{k+1}(y)$ is of the the form

$$[x, x^j y x^{k-j}] = x^{j+1} y x^{k+1-(j+1)} - x^j y x^{k+1-j},$$

as claimed.

Since x is nilpotent there is some $n \geq 1$ such that $x^n = 0$. From the claim we can see that $(ad_x)^{2n}(y) = 0$, and so ad_x is nilpotent. □

Lemma 1.57: *Let L be a Lie subalgebra of $\mathfrak{gl}(V)$, where V is a non-zero vector space. If L consists of nilpotent linear transformations, then there is some non-zero $w \in V$ such that $L \cdot w = 0$.*

Proof. We proceed by induction on $\dim(L)$. First, we cover the case where $\dim(L) = 1$. Suppose, to the contrary, that $x \cdot v = 0$ only if $v = 0$, for all $x \in L,\ v \in V$. Then $Im(x) \cong V$, which further implies that $Im(x^n) \cong V$ for all $n \geq 1$. This contradicts x being nilpotent.

Assume now that $\dim(L) > 1$. Let K be a maximal proper subalgebra of L. By Proposition 1.56 we have that the elements of K are ad-nilpotent on L, so for $u \in L$ we have that $ad_u \colon L \to L$ is nilpotent. We may the consider the action of $u + K$ on the quotient vector space L/K, that is,

$$(u + K) \cdot (y + K) = ad_{u+K} \cdot (y + K) = [u, y] + K.$$

Since ad_u is nilpotent we have that ad_{u+K} is also nilpotent. Hence, the elements of $ad(L/K)$ are nilpotent. By the inductive hypothesis there is some $x \in L - K$ such that $[x, y] + K = K$ for all $y \in K$. This implies that $[x, y] \in K$ for all $y \in K$. From the definition of $N_L(K)$ we may conclude that $x \in N_L(K)$. However $x \notin K$, so K is a proper subset of $N_L(K)$. Since K is a maximal proper subalgebra of L, we must have that $N_L(K) = L$. Hence, from Proposition 1.55 we have that K is an ideal of L.

Set $\widetilde{K} = K \oplus Span(\{x\})$. This is a Lie subalgebra of L which properly contains K, so $\widetilde{K} = L$. Thus, K has codimension 1. We now apply the inductive hypothesis to K. This tells us that there is a non-zero $v \in V$ such that $y \cdot v = 0$, for all $y \in K$. Let $W = \{w \in V \mid y(w) = 0,\ \forall\ y \in K\}$. Since $v \in W$ we know that W is non-zero. We apply Lemma 1.24 and see that W is L-invariant, specifically $x(W) \subseteq W$. We remark that in this application of Lemma 1.24 we do not require that $char(\mathbb{F}) = 0$ because the associated weight is the zero map. Now, x is nilpotent so the action of x on W is nilpotent. Thus, there is a non-zero $w \in W$ such that $x \cdot w = 0$. For any $z \in L$ we may write $z = y + \beta x$ for some $y \in K,\ \beta \in \mathbb{F}$. So $z \cdot w = y \cdot w + \beta x \cdot w = 0$, for all $z \in L$. □

Let L be a nilpotent Lie algebra. Then $L^n = 0$ for some $n \geq 1$. We can then conclude that $(ad_x)^n = 0$ for any $x \in L$. So all elements of L are ad-nilpotent. The following important theorem shows that the converse is also true.

Theorem 1.58 (Engel's Theorem)**:** *Let L be a Lie algebra. If every element of L is ad-nilpotent, then L is nilpotent.*

Proof. We proceed by induction on $\dim(L)$. In the case where $\dim(L) = 1$, L must be abelian and thus nilpotent.

Assume now that $\dim(L) > 1$. We consider the image of L under the adjoint representation, which is a linear Lie algebra. We know that each element of $ad(L)$ is a nilpotent linear transformation of L. An application of Lemma 1.57 yields the existence of a non-zero $v \in L$ such that $ad(L) \cdot v = 0$. Thus, $v \in Z(L)$ which implies that $Z(L)$ is non-zero. So the quotient Lie algebra $L/Z(L) \cong ad(L)$ has dimension strictly smaller than L. By the inductive hypothesis we conclude that $L/Z(L)$ is nilpotent and then an application of Proposition 1.53 yields that L is nilpotent. □

There is an equivalent statement of Engel's Theorem which is often useful. We will prove it as a corollary of Lemma 1.57.

Corollary 1.59: *Let L be a subalgebra of $\mathfrak{gl}(V)$, where V is a non-zero vector space. If L consists of nilpotent linear transformations, then there is a basis for V relative to which the matrices of L are all elements of $\mathfrak{n}(n, \mathbb{F})$.*

Proof. We know that there is a non-zero $v \in V$ such that $L \cdot v = 0$. We proceed by induction on $\dim(V)$. If $\dim(V) = 1$, then since $L = Span(\{v\})$ we are done.

Assume now that $\dim(V) > 1$. Then the L-module $\widetilde{V} = V/Span(v)$ has dimension $\dim(V) - 1 = n$. By the inductive hypothesis there is a basis, $\{b_1 + Span(\{v\}), \ldots, b_n + Span(\{v\})\}$, of $\widetilde{V}$ relative to which the matrices representing the action of L on $\widetilde{V}$ are elements of $\mathfrak{n}(n, \mathbb{F})$. We claim that $B = \{v, b_1, \ldots, b_n\}$ is a basis for V. Clearly, $Span(B) = V$ and $B - \{v\}$ is linearly independent. Suppose $v = \sum_{i=1}^{n} \beta_i b_i$ for some $\beta_i \in \mathbb{F}$. Then $\sum_{i=1}^{n} \beta_i b_i + Span(\{v\}) = Span(\{v\})$, which contradicts $B - Span(\{v\})$ being linearly independent. Since $L \cdot v = 0$ we may conclude that, with respect to B, the matrices of the elements of L are elements of $\mathfrak{n}(n, \mathbb{F})$, as required. □

1.8 Lie's Theorem

We will now study an analogous result to that of Engel's Theorem in the context of solvability. This will require working with an algebraically closed field of characteristic zero. For the rest of the chapter we will assume that $\mathbb{F}$ denotes such a field.

Lemma 1.60: *Let L be a Lie subalgebra of $\mathfrak{gl}(V)$ where V is a non-zero vector space. If L is solvable, then V contains an eigenvector for L.*

Proof. We proceed by induction on $\dim(L)$. In the case where $\dim(L) = 1$, we note that since every $y \in L$ is a linear transformation and $\mathbb{F}$ is algebraically closed, x must have an eigenvector in V.

Assume now that $\dim(L) > 1$. Since L is solvable we know that L' is a proper ideal of L. Take a subspace $\widetilde{K}$ of L/L', of codimension 1. Since L/L' is abelian we must have that $\widetilde{K}$ is an ideal of L/L'. By Proposition 1.16 there is a corresponding ideal K of L which has codimension 1 and contains L'. Thus, we may take $z \in L - K$ and write $L = K \oplus Span(\{z\})$. Applying the inductive hypothesis we get an eigenvector $v \in V$ for K. This also gives us a weight $\lambda\colon K \to \mathbb{F}$ and a corresponding weight space V_λ, which is non-zero since $v \in V_\lambda$. An application of Lemma 1.24 tells us that V_λ is L-invariant. Specifically, this means that $z \cdot V_\lambda \subseteq V_\lambda$. Thus, z restricted to V_λ is a linear transformation and since $\mathbb{F}$ is algebraically closed there must be a $w \in V_\lambda$ such that w is an eigenvector for z.

Taking $y \in L$ we may write $y = x + \beta z$, for some $x \in K,\ \beta \in \mathbb{F}$. Thus,

$$y(w) = x(w) + \beta z(w) = \lambda(x)w + \lambda(\beta z)w.$$

Hence, w is the required eigenvector for L. □

Theorem 1.61 (Lie's Theorem)**:** *Let L be a subalgebra of $\mathfrak{gl}(V)$, where V is a non-zero vector space. If L is solvable, then there is a basis for V relative to which the matrices of L are all elements of $\mathfrak{b}(n, \mathbb{F})$.*

Proof. From Lemma 1.60 we know that there is a $v \in V$ which is an eigenvector for L. We proceed by induction on $\dim(V)$. In the case where $\dim(V) = 1$ the existence of the eigenvector v makes the result clear.

Assume now that $\dim(V) > 1$. Then the L-module $\widetilde{V} = V/Span(\{v\})$ has dimension $\dim(V) - 1 = n$. By the inductive hypothesis there is a basis $\{b_1 + Span(\{v\}), \ldots, b_n + Span(\{v\})\}$ of $\widetilde{V}$, relative to which the matrices representing the action of L on $\widetilde{V}$ are elements of $\mathfrak{b}(n, \mathbb{F})$. We claim that $B = \{v, b_1, \ldots, b_n\}$ is a basis for V. Clearly, $Span(B) = V$ and $B - \{v\}$ is linearly independent. Suppose $v = \sum_{i=1}^{n} \beta_i b_i$ for some $\beta_i \in \mathbb{F}$. Thus implies that $\sum_{i=1}^{n} \beta_i b_i + Span(\{v\}) = Span(\{v\})$, which contradicts $B - Span(\{v\})$ being linearly independent. Since v is an eigenvector for L we may conclude that, with respect to B, the matrices of the elements of L are elements of $\mathfrak{b}(n, \mathbb{F})$, as required. □

1.9 Killing Form

In this section we look at an important bilinear form on Lie algebras over $\mathbb{F}$. We will later see that this form can give us important information about whether a Lie algebra is solvable or semisimple.

Definition 1.62: *Let L be a Lie algebra. Define $\kappa\colon L \times L \to \mathbb{F}$ by*

$$\kappa(x, y) = tr(ad_x \circ ad_y),$$

for all $x, y \in L$. Then κ is a symmetric bilinear form on L, called the Killing *form.*

We know that the adjoint representation is linear and we saw in Example 1.18 that tr is linear. To see that κ is bilinear we take $\alpha, \beta \in \mathbb{F},\ x, y, z \in L$, and see that

$$\begin{aligned}\kappa(\alpha x + \beta y, ad_z) &= tr(ad_{\alpha x+\beta y} \circ ad_z) = tr((\alpha ad_x + \beta ad_y) \circ ad_z)\\ &= \alpha tr(ad_x \circ ad_z) + \beta tr(ad_y \circ ad_z) = \alpha\kappa(x, z) + \beta\kappa(y, z),\end{aligned}$$

which shows that one argument is linear. Similarly, one can see the other argument is linear and so, κ is bilinear. We note that κ is symmetric because $tr(xy) = tr(yx)$ for linear maps. Another useful property of the Killing form is its *associativity*, which is the property that $\kappa([x, y], z) = \kappa(x, [y, z])$, for all $x, y, z \in L$. To show that this property holds we need the property of tr that $tr((ab)c) = tr(c(ab))$, for linear transformations a, b, c, which implies that

$$tr([a, b]c) = tr(abc) - tr(bac) = tr(abc) - tr(acb) = tr(a[b, c]).$$

From Proposition 1.20, for $x, y, z \in L$ we have that

$$\begin{aligned}\kappa([x, y], z) &= tr(ad_{[x,y]}ad_z) = tr(([x, ad_y] + [ad_x, y])ad_z)\\ &= tr(ad_x ad_y ad_z) - tr(ad_y ad_x ad_z)\\ &= tr(ad_x ad_y ad_z) - tr(ad_x ad_z ad_y) = tr(ad_x(ad_y ad_z - ad_z ad_y))\\ &= tr(ad_x([y, ad_z] + [ad_y, z])) = tr(ad_x ad_{[y,z]}) = \kappa(x, [y, z]),\end{aligned}$$

which shows the κ is associative.

Proposition 1.63: *Let L be a Lie algebra. Then L is solvable if and only if L' is nilpotent.*

Proof. Suppose L is solvable. Applying Proposition 1.45 we have that $ad(L)$ is solvable. By Theorem 1.61 (Lie's Theorem) there is a basis for V, relative to which every element of $ad(L)$ is upper triangular. Then $ad_x(y) = [x, y]$ is strictly upper triangular, and thus nilpotent, for all $x, y \in L$. Thus, every element of L' is ad-nilpotent. Applying Theorem 1.58 (Engel's Theorem) we have that L' is nilpotent.

On the other hand, suppose L' is nilpotent. Then L' is solvable and L/L' is clearly solvable. We may then apply Proposition 1.45 and see that L is solvable. □

Lemma 1.64: *Let $x, y \in \mathfrak{gl}(V)$, where V is some vector space. If x and y commute then for all $k \geq 1$ each summand of $(x-y)^k$ is of the form $x^i y^j$ for natural numbers i and j, where at least one of i or j is greater than $k/2$.*

Proof. We proceed by induction on k. The case $k = 1$ is clear. Assume $k > 1$. Then $(x-y)^k = (x-y)(x-y)^{k-1}$. Applying the inductive hypothesis to $(x-y)^{k-1}$ and using the knowledge that x and y commute yields the desired result. □

Lemma 1.65: *Let L be a subalgebra of $\mathfrak{gl}(V)$, where V is some vector space. If $tr(xy) = 0$ for all $x, y \in L$, then L is solvable.*

Proof. We claim that L' is nilpotent. Take $x \in L'$ with Jordan decomposition $x = d + n$. We may choose a basis for V relative to which d is diagonal and n is strictly upper triangular. We would like to show that $d = 0$. Suppose d has diagonal entries $\lambda_1, \ldots, \lambda_n$. Let $\overline{d}$ be the diagonal matrix with diagonal entries $\overline{\lambda_1}, \ldots, \overline{\lambda_n}$ (where $\overline{a}$ denotes the conjugate of a). Then,

$$\overline{d}x = \overline{d}(d+n) = \begin{pmatrix} \lambda_1\overline{\lambda_1} & & & & \\ & \lambda_2\overline{\lambda_2} & & & 0 \\ & & \ddots & & \\ & & & \lambda_{n-1}\overline{\lambda_{n-1}} & \\ & 0 & & & \lambda_n\overline{\lambda_n} \end{pmatrix} + n',$$

where n' is some strictly upper triangular matrix. Thus, $tr(\overline{d}x) = \sum_{i=1}^{n} \lambda_i\overline{\lambda_i}$. Since $x \in L'$ we may write $x = \sum_{i=1}^{n} \alpha_i[y_i, z_i]$ for $\alpha_i \in \mathbb{F}$, $y_i, z_i \in L$. We have that

$$tr(\overline{d}x) = \sum_{i=1}^{n} \alpha_i tr(\overline{d}[y_i, z_i]) = \sum_{i=1}^{n} \alpha_i tr([\overline{d}, y_i]z_i).$$

If we can show that $[\overline{d}, y_i] \in L$ we may conclude, by our hypothesis, that $tr([\overline{d}, y_i]z_i) = 0$ for all $i \in \{1, \ldots, n\}$. Then we would have $tr(\overline{d}x) = 0$ which would imply that $d = 0$ and then we would be done. So let us show, equivalently, that $ad_{\overline{d}}$ maps L to L. By Corollary D.9 we have that the Jordan decomposition of $ad_{\overline{x}}$ is $ad_{\overline{x}} = ad_{\overline{d}} + ad_{\overline{n}}$. From Theorem D.7 we know that there is a polynomial $p \in \mathbb{F}[X]$ such that $p(ad_{\overline{x}}) = ad_{\overline{d}}$. Since $ad_{\overline{x}}$ maps L to L, $p(ad_{\overline{x}})$ does too. Thus, $[\overline{d}, y_i] \in L$, and so $d = 0$. This shows that x is nilpotent. From Proposition 1.56 we have that x is ad-nilpotent, which means that we may apply Engel's Theorem (which states that if every element of a Lie algebra is ad-nilpotent, the algebra is nilpotent) and see that L' is nilpotent. Finally, since L is solvable if and only if L' is nilpotent, we are done. □

1.10 Cartan's Criteria

We will now see some important application of the Killing form of a Lie algebra.

Theorem 1.66 (Cartan's First Criterion)**:** *Let L be a Lie algebra. Then L is solvable if and only if $\kappa(x, y) = 0$ for all $x \in L,\ y \in L'$.*

Proof. Suppose L is solvable. Then by Proposition 1.45 we have that $ad(L)$ is solvable. By Theorem 1.61 (Lie's Theorem) there is a basis for V relative to which every element of $ad(L)$ is represented by an upper triangular matrix. Since the commutator of two diagonal matrices is zero, we conclude that relative to this basis $ad_{[u,v]} \in ad(L')$ is represented by a strictly upper triangular matrix for all $u, v \in L$. Let $x \in L,\ y \in L'$. Then there is a basis for V relative to which ad_x is upper triangular and ad_y is strictly upper triangular. Then $tr(ad_x \circ ad_y) = 0$.

On the other hand, suppose $\kappa(x, y) = 0$ for all $x \in L,\ y \in L'$. Applying Lemma 1.65 we have that $ad(L')$ is solvable. By Proposition 1.20 we know that $ad(L') \cong L'/Z(L')$ is solvable. Since $L/Z(L')$ and $Z(L')$ are solvable Proposition 1.45 tells us that L' is solvable. Finally, this means that L is solvable. □

Proposition 1.67: *Let L be a Lie algebra and I and ideal of L. Then $\kappa_I(x, y) = \kappa(x, y)$ for all $x, y \in I$.*

Proof. Since $x \in I$ and I is an ideal of L, we know that ad_x maps L into I. So if we take a basis for I and extend it to a basis for L, the matrix of ad_x with respect to this basis will be of the form

$$\begin{pmatrix} X_I & X_* \\ 0 & 0 \end{pmatrix},$$

where X_I is the matrix representing ad_x restricted to I. We may follow similar process for ad_y. Then $ad_x \circ ad_y$ is represented by a matrix of the form

$$\begin{pmatrix} X_I Y_I & X_I Y_* \\ 0 & 0 \end{pmatrix}.$$

We can then see that

$$\kappa(x, y) = tr(X_I Y_I) + tr(0) = tr(X_I Y_I) = \kappa_I(x, y).$$

□

Lemma 1.68: *Let I be an ideal of a Lie algebra L. Then $I^\perp$ is an ideal of L, in particular $L^\perp$ is an ideal of L.*

Proof. We already know that $I^\perp$ is a vector subspace of L. For $u \in I^\perp,\ v \in L$ and $w \in I$ we have

$$\kappa([u,v],w) = \kappa(u,[v,w]) = 0,$$

since $[v,w] \in I$. This shows that $[u,v] \in I^\perp$, as required. □

Lemma 1.69: *Let L be a Lie algebra. Then L is semisimple if and only if L has no non-zero abelian ideals.*

Proof. Suppose L is semisimple. Let A be an abelian ideal of L, so clearly A is solvable. Thus, $A \subseteq rad(L) = 0$.

On the other hand, suppose L has no non-zero abelian ideals. Let $k \geq$ be the smallest positive integer such that $rad(L)^{(k)} = 0$. Then $rad(L)^{(k-1)}$ is an abelian ideal of L, and thus $rad(L)^{(k-1)} = 0$. This implies that $rad(L) = 0$, and so L is semisimple. □

Theorem 1.70 (Cartan's Second Criterion)**:** *Let L be a Lie algebra. Then L is semisimple if and only if the Killing form on L is non-degenerate.*

Proof. Suppose L is semisimple. From Lemma 1.68 we know that $L^\perp$ is an ideal of L. Let $x \in L^\perp$ and $y \in (L^\perp)'$, then it is also true that $y \in L$. Thus, $\kappa(x,y) = 0$ and an application of Theorem 1.66 (Cartan's First Criterion) yields that $L^\perp$ is solvable. Since L is semisimple and $L^\perp$ is a solvable ideal of L, we must have that $L^\perp = 0$, and so κ is non-degenerate.

On the other hand, suppose that $L^\perp = 0$. Let A be an abelian ideal of L. For $x \in A,\ y,z \in L$ we have that

$$(ad_y \circ ad_x)(z) = [y,[x,z]] \in A,$$

since $[x,z] \in A$. This means that $ad_x \circ ad_y \circ ad_x = 0$. Thus, $(ad_x \circ ad_y)^2 = 0$, and so the map $ad_x \circ ad_y$ is nilpotent and has trace zero. We may then conclude that since $\kappa(x,y) = 0$ that A is contained in $L^\perp = 0$. So the only abelian ideal of L is 0. An application of Lemma 1.69 tells us that L is semisimple. □

Direct Sums

In this section we will define direct sums of Lie algebras and show that semisimple Lie algebras are direct sums of simple Lie algebras.

Let L_1 and L_2 be Lie algebras defined over the same field $\mathbb{F}$. If we let $L := L_1 \oplus L_2$ be the direct sum of their underlying vector spaces, we may then endow L with a Lie bracket by applying the Lie brackets of L_1 and L_2

component-wise. Bilinearity follows from the linearity of the brackets of L_1 and L_2, while properties (L1) and (L2) follow directly from L_1 and L_2 being Lie algebras.

Definition 1.71: *Let L_1 and L_2 be Lie algebras. Then the direct sum of their underlying vector spaces $L = L_1 \oplus L_2$ endowed with Lie bracket defined component-wise, that is, for $x_1, y_1 \in L_1,\ x_2, y_2 \in L_2$*

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

forms a Lie algebra.

Lemma 1.72: *Let $L = L_1 \oplus L_2$ be a Lie algebra. Then $Z(L) = Z(L_1) \oplus Z(L_2)$.*

Proof. Let $z \in Z(L)$, then $z = (z_1, z_2)$ for some $z_1 \in L_1,\ z_2 \in L_2$. For $x_1 \in L_1$, we have

$$\begin{aligned}[(z_1, 0), (x_1, 0)] &= [(z_1, z_2), (x_1, 0)] - [(0, z_2), (x_1, 0)] = -([0, x_1], [z_2, 0]) \\ &= (0, 0),\end{aligned}$$

which implies that $z_1 \in Z(L_1)$. Similarly, we find that $z_2 \in Z(L_2)$. Thus, $Z(L) \subseteq Z(L_1) \oplus Z(L_2)$.

Let $z_1 \in Z(L_1)$ and $z_2 \in Z(L_2)$. Then for $(x_1, x_2) \in L$, we have

$$[(z_1, z_2), (x_1, x_2)] = ([z_1, x_1], [z_2, x_2]) = (0, 0),$$

which shows that $Z(L) = Z(L_1) \oplus Z(L_2)$. □

Lemma 1.73: *Let L be a semisimple Lie algebra. If $I \neq L$ is a non-zero ideal of L, then $L = I \oplus I^{\perp}$, and I is semisimple.*

Proof. From Lemma 1.68 we know that $I^{\perp}$ is an ideal of L and thus that $I \cap I^{\perp}$ is also an ideal of L. Then by Proposition 1.67 we have that the Killing form restricted to $I \cap I^{\perp}$ is the zero map. An application of Theorem 1.66 (Cartan's First Criterion) yields that $I \cap I^{\perp}$ is solvable and since L is semisimple we may conclude that $I \cap I^{\perp} = 0$.

Since L is semisimple we apply Theorem 1.70 (Cartan's Second Criterion) and get that the Killing form is non-degenerate. From Appendix D (see Lemma D.13) we have that $\dim(I) + \dim(I^{\perp}) = \dim(L)$, and then we conclude that $L = I \oplus I^{\perp}$ since L is finite-dimensional.

Suppose, to the contrary that I is not semisimple. Applying Theorem 1.70 we see that the Killing form on I must be degenerate. However, since I is an ideal of L, we know from Proposition 1.67 that the Killing form on I is same as the Killing form on L restricted to I. So then there is some non-zero $x \in I$ such that $\kappa(x, z) = 0$, for all $z \in I$. By definition we then also have that $\kappa(x, y) = 0$, for all $y \in I^{\perp}$. Since $L = I \oplus I^{\perp}$, this means that the Killing form on L is degenerate, which contradicts our assumption that L is semisimple. □

Theorem 1.74: *Let L be a Lie algebra. Then L is semisimple if and only if L is a direct sum of simple ideals of L.*

Proof. Suppose L is semisimple, then $\dim(L) \geq 2$ since a 1-dimensional Lie algebra is abelian. We proceed by induction on $\dim(L)$. If $\dim(L) = 2$, then L has no non-zero abelian ideals by Lemma 1.69. Thus, ideals of L can have dimension 0 or 2. So, the only ideals of L are 0 and L, and thus L itself is simple (Notice that $L' \neq 0$ because L is semisimple and therefore non-abelian).

Assume now that $\dim(L) > 2$. Let I be a non-zero ideal of L of the smallest possible dimension. If $I = L$, then L is simple and we are done. Otherwise, we apply Lemma 1.73 and we have that $L = I \oplus I^{\perp}$, we also get that I is semisimple. In fact, we claim that I is a simple ideal of L. Since L is semisimple, an application of Lemma 1.69 gives us that I is not abelian. Let J be a non-zero ideal of I. Then $[J, I^{\perp}] \subseteq [I, I^{\perp}] \subseteq I \cap I^{\perp} = 0$. This implies that $[L, J] = [I \oplus I^{\perp}, J] = [I, J] = J$. So, J is an ideal of L, and then since I is a non-zero ideal of L of smallest possible dimension, we conclude that $J = I$. Thus, I is simple.

By the inductive hypothesis we may write $I^{\perp} = \bigoplus_{i=1}^{k} J_i$, where each J_i is a simple ideal of $I^{\perp}$. Proceeding, as before, to show that each J_i is a simple ideal of L, notice that $[I, J_i] \subseteq [I, I^{\perp}] \subseteq I \cap I^{\perp} = 0$, so

$$[L, J_i] = [I \oplus I^{\perp}, J_i] \subseteq J_i.$$

Thus, each J_i is a simple ideal of L. We then have $L = I \oplus \bigoplus_{i=1}^{k} J_i$, as required.

On the other hand, suppose $L = \bigoplus_{i=1}^{k} I_i$, where each I_i is a simple ideal of L. Then we have that $[rad(L), I_i] \subseteq rad(L) \cap I_i$, for all $i \in \{1, \ldots, k\}$. But $rad(L) \cap I_i$ is a solvable ideal of I_i, which is simple. So, $rad(L) \cap I_i = 0$ which implies that $[rad(L), I_i] = 0$. Thus, $[rad(L), L] = 0$, which tells us that $rad(L) \subseteq Z(L)$. From Lemma 1.72 we get that $Z(L) = \bigoplus_{i=1}^{k} Z(I_i)$. The I_i are simple, so each $Z(I_i) = 0$. So, $Z(L) = 0$, which implies that $rad(L) = 0$, and thus L is semisimple. □

Corollary 1.75: *Let L be a semisimple Lie algebra. Then the image of L under any Lie algebra homomorphism is semisimple.*

Proof. Let $\phi\colon L \to M$ be a Lie algebra homomorphism between two Lie algebras L and M. From Theorem 1.17 (Isomorphism theorems) we know that $L/Ker(\phi) \cong \phi(L)$. If ϕ is a monomorphism the result is clear, so we assume that $Ker(\phi)$ is non-zero. Notice that $L = Ker(\phi) \oplus (Ker(\phi))^{\perp}$ by Lemma 1.73, and $(Ker(\phi))^{\perp}$ is semisimple. However, from this we can see that $L/Ker(\phi) \cong (Ker(\phi))^{\perp}$, and so we are done. □

1.11 Root Space Decomposition

In this we introduce the root space decomposition of a semisimple Lie algebra. We begin by defining maximal toral algebras and study some important properties of these. For the rest of this chapter we assume that L denotes a semisimple Lie algebra over $\mathbb{F}$.

From Theorem 1.58 (Engel's Theorem) we know that if all the elements of L were ad-nilpotent, that L would be nilpotent. On the other hand, suppose there is an element $x \in L$, which is not ad-nilpotent. Let $x = d+n$, be the abstract Jordan decompostion of x. Then, d is a non-zero semisimple element of L. Hence, $Span(\{d\})$ is a non-zero subalgebra of L which consists entirely of semisimple elements of L.

Definition 1.76: *Let L be a Lie algebra, and T a non-zero subalgebra of L which consists entirely of semisimple elements of L. Then, T is a* toral *subalgebra of L. Moreover, if T is a toral subalgebra of L which is not properly contained in any toral subalgebra of L, then T is a* maximal toral *subalgebra of L.*

Lemma 1.77: *Let L be a Lie algebra, and T a toral subalgebra of L. Then T is abelian.*

Proof. Let $x \in T$. Then the action of ad_x on L is diagonalisable. Moreover, $ad_x(T) = [x, T] \subseteq T$, since T is a subalgebra of L. Hence the action of ad_x on T is also diagonalisable. We claim ad_x has no non-zero eigenvalues in T. Suppose to the contrary, that $0 \neq \lambda \in \mathbb{F}$ is an eigenvalue of ad_x, with corresponding eigenvector $y \in T$. So, $ad_x(y) = [x, y] = \lambda y$. This implies that $ad_y(x) = -[x, y] = -\lambda y$. However, similarly to ad_x, we know that ad_y acts diagonalisably on T. We may therefore find a basis of T, consisting of eigenvectors for ad_y, say $\{y, v_2, \ldots, v_n\}$, with corresponding eigenvalues $\{0, \lambda_2, \ldots, \lambda_n\}$.We may then write

$$x = \alpha_1 y + \sum_{k=2}^{n} \alpha_k v_k,$$

for some $\alpha_k \in \mathbb{F}$. And so,

$$ad_y(x) = \alpha_1[y, y] + \sum_{k=2}^{n} \alpha_k [y, v_k] = \sum_{k=2}^{n} \alpha_k \lambda_k v_k \neq -\lambda y = ad_y(x).$$

This is a contradiction. Thus, the only eigenvalue of ad_x, acting on T, is 0. This shows that $ad_x(T) = 0$, and hence T is abelian. □

We note that some texts define *Cartan* subalgebras as maximal abelian toral subalgebras. The above lemma makes it clear that maximal toral algebras are equivalent to Cartan subalgebras.

Let H be a maximal toral subalgebra of a Lie algebra L. Then the elements of $ad(H)$ are commuting linear endomorphisms of L. Since these elements are all semisimple, and $\mathbb{F}$ is algebraically closed, we may simultaneously diagonalise them. We therefore have a basis for L, consisting of eigenvectors for $ad(H)$. By Theorem D.1 (the Primary Decomposition Theorem), we find that L will decompose into the direct sum of the weight spaces of $ad(H)$. We have seen that weights are elements of the dual space (in this case of H). We will denote the weight space corresponding to $\alpha \in H^*$ by

$$L_\alpha = \{x \in L \mid [h, x] = \alpha(h)x, \text{ for all } h \in H\}.$$

We may then write $L = \bigoplus_{\alpha \in H^*} L_\alpha$. When $\alpha = 0$, the corresponding weight space is, in fact, the ***centraliser*** of H; that is, $L_0 = C_L(H)$. We recall from our definition of weight spaces that L_α must be non-zero. We therefore, need only work with a subset of H^*; the non-zero $\alpha \in H^*$ for which the L_α are weight spaces, the *roots* of L with respect to H. We denote the set of roots by Φ. We call this decomposition of L,

$$L = C_L(H) \oplus \bigoplus_{\alpha \in \Phi} L_\alpha, \tag{1.2}$$

the ***root space*** decompostion of L. As L is finite-dimensional, as a vector space, we note that Φ contains only finitely many elements.

Moving forward, our next goal is to show that maximal toral subalgebras are self-centralising. To do so, we first need some preliminary results.

Proposition 1.78: *Let L be a Lie algebra and H a maximal toral subalgebra of L. Let $\alpha, \beta \in H^*$, then $[L_\alpha, L_\beta] \subseteq L_{\alpha+\beta}$. If $\alpha \neq -\beta$, then $\kappa(L_\alpha, L_\beta) = 0$, where κ is the Killing form on L.*

Proof. For $x \in L_\alpha,\ y \in L_\beta$, and $h \in H$, we have

$$\begin{aligned}[h, [x, y]] &= -[x, [y, h]] - [y, [h, x]] = [x, [h, y]] + [[h, x], y] \\ &= \beta(h)[x, y] + \alpha(h)[x, y] = (\alpha + \beta)(h)[x, y],\end{aligned}$$

since $[h, x] = \alpha(h)x$ and $[h, y] = \beta(h)y$. This implies that $[x, y]$ is an eigenvector of H with weight $\alpha + \beta$; that is, $[x, y] \in L_{\alpha+\beta}$. Assume now, that $\alpha \neq -\beta$. Then we may find $h \in H$ such that $(\alpha + \beta)(h) \neq 0$. Then, by the associativity of the Killing form on L, we have

$$\alpha(h)\kappa(x, y) = \kappa([h, x], y) = -\kappa([x, h], y) = -\kappa(x, [h, y]) = -\beta(h)\kappa(x, y),$$

which implies that $(\alpha + \beta)(h)\kappa(x, y) = 0$. Hence, $\kappa(x, y) = 0$. □

Corollary 1.79: *Let L be a Lie algebra and H a maximal toral subalgebra of L. Then the Killing form on L restricted to $C_L(H)$ is non-degenerate.*

Proof. Suppose $z \in C_L(H)$ satisfies that $\kappa(z,x) = 0$, for all $x \in C_L(H)$. Since L is semisimple, Theorem 1.70 (Cartan's Second Criterion) tells us that κ is non-degenerate. From the proposition above we know that $L_0 = C_L(H)$ is orthogonal to L_α, for $\alpha \in \Phi$. If $y \in L$, then from (1.2) we may write $y = y_0 + \sum_{\alpha\in\Phi} y_\alpha$,where $y_\alpha \in L_\alpha$ and $y_0 \in L_0$. Then, we have

$$\kappa(z,y) = \kappa(z,y_0) + \sum_{\alpha\in\Phi} \kappa(z,y_\alpha) = \kappa(z,y_0) = 0,$$

which implies that $z = 0$. □

Theorem 1.80: *Let L be a Lie algebra and H a maximal toral subalgebra of L. Then $H = C_L(H)$.*

Proof. We will make several claims throughout the proof. We will denote $C_L(H)$ by C.

(1) We claim that for $x \in C$, the components of the abstract Jordan decomposition $x = d + n$, are contained in C; that is, $d, n \in C$. Since $x = d+n$ is in the centraliser of H, we have that $ad_x(H) = 0$. Let the Jordan decomposition of ad_x be $ad_x = ad_{d'} + ad_{n'}$. From Theorem D.7, we find polynomials with no constant term, $p, q \in \mathbb{F}[X]$, such that $p(ad_x) = ad_{d'}$ and $q(ad_x) = ad_{n'}$. This implies that $ad_{d'}(H) = ad_{n'}(H) = 0$. An application of Corollary D.11 we see that $d = d'$ and $n = n'$. Thus, $d, n \in C$.

(2) We claim that the semisimple elements of C are exactly the elements of H. Let x be a semisimple element of C. We consider $S = H + Span(\{x\})$. Clearly, S is a vector subspace of L. Furthermore, for $h, h' \in H$ and $\alpha, \beta \in \mathbb{F}$ we have

$$[h + \alpha x, h' + \beta x] = [h, h'] + \beta[h, x] + \alpha[x, h'] + \alpha\beta[x, x] = 0,$$

which shows that S is a subalgebra of L. We also have that h and αx are semisimple. Since we are working over an algebraically closed field, they are diagonalisable. They also commute with one another since $x \in C$, so they may be simultaneously diagonalised, and their sum is still semisimple. Hence, S is a toral subalgebra of L, by the maximality of H we have that $S = H$. This implies that $x \in H$.

(3) We claim that the restriction of the Killing form on L to H is non-degenerate. Let $h \in H$ such that $\kappa(h, h') = 0$, for all $h' \in H$. We want to show that $h = 0$. Let $x \in C$ with abstract Jordan decomposition $x = d + n$. Then ad_n is nilpotent and ad_n and ad_y commute, for all $y \in H$, because by (1) $n \in C$. We then have that $ad_y \circ ad_n$ must be nilpotent. This means that $0 = tr(ad_y \circ ad_n) = \kappa(y, n)$, for all $y \in H$. From (1) we know that $d \in C$ and since d is semisimple (2) tells us that $d \in H$. By hypothesis $\kappa(h, d) = 0$. Together, this means that $\kappa(h, x) = \kappa(h, d) + \kappa(h, n) = 0$, for all $x \in C$.

From Corollary 1.79 we know that the Killing form on L is non-degenerate, and so $h = 0$.

(4) We claim that C is nilpotent. Let $x \in C$, with abstract Jordan decomposition $x = d + n$. From (1) we know that $d, n \in C$ and from (2) we have that $d \in H$. Thus,

$$ad_x|_C = ad_d|_C + ad_n|_C = ad_n|_C,$$

where $ad_x|_C$ denotes ad_x restricted to C. So, each $x \in C$ is ad-nilpotent. An application of Theorem 1.58 (Engel's Theorem) yields that C is nilpotent.

(5) We claim that $H \cap (C)' = 0$. We have that $[H, C] = 0$, by definition of the centraliser of H. By the associativity of the Killing form on L, we have that $0 = \kappa([H, C], C)) = \kappa(H, (C)')$. So, for $h \in H \cap C$, we have that $\kappa(h, H) = 0$. From (3) we may conclude that $h = 0$.

(6) We claim that C is abelian, suppose it is not. Then $(C)'$ must be non-zero. Since L is nilpotent Proposition 1.53 tells us that C is also nilpotent. Hence, the elements of C are all ad-nilpotent. Since $(C)'$ is an ideal of C, we may consider $(C)'$ a C-module via the representation $\phi \colon C \to \mathfrak{gl}((C)')$, which we define by

$$\phi(x) = ad_x \colon (C)' \to (C)',$$

for all $x \in C$. We apply Lemma 1.57 to find a non-zero $z \in (C)'$ such that $[C, z] = 0$. This implies that $z \in Z(C) \cap (C)'$. From (5), and since z is non-zero, we know that $z \notin H$. From (2) we conclude that z is not semisimple. Let $z = d+n$ be the abstract Jordan decomposition of z. We then have that $n \neq 0$, and from (1) we have that $n \in C$. Since the Jordan decomposition coincides with the abstract Jordan decomposition (Corollary D.11) and from Theorem D.7 we may find a polynomial, with no constant term, $q \in \mathbb{F}[X]$ such that $q(z) = n$. We also have that ad_z maps C to 0. Putting these facts together, we conclude that ad_n maps C to 0, and so $n \in Z(C)$. For $x \in C$, we know that both ad_x and ad_n are nilpotent and ad_n commutes with ad_x. Hence, $ad_x \circ ad_n$ is also nilpotent, and thus has trace zero. This contradicts Corollary 1.79. So, C must be abelian.

Suppose $C \neq H$. Then from (1) and (2) we see that there is an element $x \in C$ with Jordan decomposition $x = d+n$, where n is non-zero. For $y \in C$ we can see from (6) that ad_y and ad_n commute and thus that $ad_n \circ ad_y$ is nilpotent, since ad_n is nilpotent. This means that $tr(ad_n \circ ad_y) = 0$, which contradicts Corollary 1.79. Hence, $C_L(H) = C = H$. □

Corollary 1.81: *Let L be a Lie algebra and H a maximal toral subalgebra of L. Then the restriction of the Killing form on L to H is non-degenerate.*

Proof. See part (3) of the proof of the theorem above. □

The theorem above tells us more about the root space decomposition (1.2) which we saw earlier. Let L be a Lie algebra and H a maximal toral subalgebra of L. We may now write

$$L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha, \tag{1.3}$$

where Φ denotes the set of roots of L, with respect to H.

Lemma 1.82: *Let L be a Lie algebra and H a maximal toral subalgebra of L. Then the Killing form on L induces a vector space isomorphism between H and H^*. In particular, for every $\alpha \in H^*$ there is a unique $t_\alpha \in H$ such that $\kappa(t_\alpha, k) = \alpha(k)$, for all $k \in H$.*

Proof. Define $\tau \colon H \to H^*$ by $\tau(h) = \theta_h$ for all $h \in H$, where $\theta_h(k) = \kappa(h, k)$, for all $k \in H$. To see that τ is linear, take $h, h', k \in H$ and $\alpha, \beta \in \mathbb{F}$, we then have

$$\begin{aligned}\tau(\alpha h + \beta h')(k) = \kappa(\alpha h + \beta h', k) &= \alpha\kappa(h, k) + \beta\kappa(h', k)\\ &= \alpha\tau(h)(k) + \beta\tau(h')(k).\end{aligned}$$

It is well-known that the restriction of κ to H is non-degenerate. Thus, $\mathrm{Ker}(\tau) = 0$, and since $\dim(H) = \dim(H^*)$ we find that τ is a vector space isomorphism. □

Lemma 1.83: *If $\alpha \in \Phi$, $x \in L_\alpha$, $y \in L_{-\alpha}$, then $[x, y] = \kappa(x, y)t_\alpha$ (where t_α follows the notation in Lemma 1.82).*

Proof. Let $\alpha \in \Phi$. By the associative property of the Killing form on L, for $x \in L_\alpha$, $y \in L_{-\alpha}$, we have

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

for all $h \in H$. This means that $\kappa(H, [x, y] - \kappa(x, y)t_\alpha) = 0$. So H is orthogonal to $[x, y] - \kappa(x, y)t_\alpha$ and then the reuslt follows from the well-known fact that the restriction of the Killing form to H is non-degenerate. □

Proposition 1.84: *For each $\alpha \in \Phi$ there is a corresponding subalgebra (which we denote by $\mathfrak{sl}(\alpha)$) isomorphic to $\mathfrak{sl}(2, \mathbb{F})$. Furthermore, there is a basis $B_\alpha := \{e_\alpha, f_\alpha, h_\alpha\}$ for $\mathfrak{sl}(\alpha)$ with the following properties:*

1. $e_\alpha \in L_\alpha$,
2. $f_\alpha \in L_{-\alpha}$,
3. $h_\alpha = [e_\alpha, f_\alpha] \in H$, *and*
4. $\alpha(h_\alpha) = 2$.

Proof. Let $\alpha \in \Phi$ and $0 \neq x \in L_\alpha$. We claim there is some $y \in L_{-\alpha}$ such that $[x, y] \neq 0$. By Lemma 1.83 we know that $[x, y] = \kappa(x, y)t_\alpha$. Since $\alpha \in \Phi$ we may conclude from Lemma 1.82 that $t_\alpha \neq 0$ (otherwise we would have that $\alpha = 0$). If $\kappa(x, y) = 0$, for all $y \in L_{-\alpha}$ then we would have that

$$\kappa(L_\alpha, L) = \kappa(L_\alpha, H) + \sum_{\beta \in \Phi} \kappa(L_\alpha, L_\beta) = 0.$$

This would contradict the non-degeneracy of κ. So, we have $y \in L_{-\alpha}$ such that$[x, y] \neq 0$.

We can see that $[x, y] = h \in H$. We claim that $\alpha(h) \neq 0$. Suppose, to the contrary, that $\alpha(h) = 0$, then $[h, x] = [h, y] = 0$. We can then see that $\mathfrak{sl}(\alpha) = Span(\{x, y, h\})$ is a solvable subalgebra of L. Since L is semisimple we know that $Z(L) = 0$, so $ad(S) \cong S$. Then, making use of the well-known fact that a Lie algebra is solvable if and only if its derived algebra is nilpotent, S' is nilpotent. This implies that ad_h is nilpotent. Since $h \in H$ we know that h is semisimple. The only element of L which is semisimple and nilpotent is 0, so $h = 0$, a contradiction. Thus, $\alpha(h) \neq 0$.

Set $e_\alpha = x$ and $f_\alpha = \frac{2}{\alpha(h)} y$. Then, since $Span(B) = \mathfrak{sl}(\alpha)$, is a subalgebra of L. By definition $[e_\alpha, f_\alpha] = h_\alpha$, while

$$\begin{aligned}
[h_\alpha, e_\alpha] &= [[e_\alpha, f_\alpha], e_\alpha] = \frac{2}{\alpha(h)}[[x, y], x] = \frac{2}{\alpha(h)}[h, x] = \frac{2}{\alpha(h)}\alpha(h)x = 2e_\alpha, \\
[h_\alpha, f_\alpha] &= [[e_\alpha, f_\alpha], f_\alpha] = \frac{2^2}{(\alpha(h))^2}[[x, y], y] = \frac{2^2}{(\alpha(h))^2}[h, y] \\
&= \frac{2^2}{(\alpha(h))^2}\alpha(h)y = 2\left(\frac{2}{\alpha(h)}y\right) = 2f_\alpha.
\end{aligned}$$

We may then construct a Lie algebra isomorphism $\tau\colon \mathfrak{sl}(\alpha) \to \mathfrak{sl}(2, \mathbb{F})$ by

$$\tau(e_\alpha) = e_{12},\ \tau(f_\alpha) = e_{21},\ \tau(h_\alpha) = h_1 = e_{11} - e_{22}.$$

By construction τ is a vector space isomorphism. From our earlier calculations we can see that τ agrees with the Lie bracket which extends τ to a Lie algebra isomorphism, as required. □

Chapter 2

Further Structure

In this chapter we will familiarise ourselves with the structure of root systems. A euclidean space is a finite-dimensional real vector space with a positive definite symmetric bilinear form denoted by $(-,-)$. A ***root system*** is a finite subset Φ of a euclidean space E with the following properties:

1. The set Φ spans the euclidean space E, and does not contain 0.

2. For $\alpha \in \Phi$, the only scalar multiples of α contained in Φ are $\pm\alpha$.

3. For each $\alpha \in \Phi$ the corresponding reflection σ_α in E (that is, the invertible linear transformation which fixes pointwise the hyperplane $P_\alpha := \{\nu \in E \,|\, (\alpha, \nu) = 0\}$ and maps any vector orthogonal to P_α to its negative) leaves Φ invariant.

4. If $\alpha, \beta \in \Phi$, then $\langle \beta, \alpha \rangle := \frac{2(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}$.

We let L be a semisimple Lie algebra over an algebraically closed field $\mathbb{F}$ of characteristic 0, H a maximal toral subalgebra of L, and Φ the corresponding root system. In what follows, $(-,-)$ denotes the symmetric bilinear form dual to the Killing form on L.

Since Φ spans H^* we may find a basis $B_\Phi := \{\alpha_i\}_{i=1}^n$ for H^* consisting of roots relative to L. We denote by E_Φ the space consisting of all linear combinations, with rational coefficients, of elements of B_Φ. We note that we may drop the Φ subscript where there is no ambiguity. For each $\alpha \in H^*$ we denote the corresponding reflection by σ_α. Recall that for each $\alpha \in H^*$ there is a unique $t_\alpha \in H$ such that $\kappa(t_\alpha, k) = \alpha(k)$, for all $k \in H$. This correspondence arises as a result of the vector space isomorphism between H and H^* which the Killing form induces (see Lemma 1.82). Then the following defines not only a non-degenerate symmetric bilinear form on H^*

but also an inner product on E.

$$\begin{aligned}(-,-)\colon H^* \times H^* &\to \mathbb{F}\\ (\theta,\phi) &\mapsto \kappa(t_\theta, t_\phi).\end{aligned}$$

We will denote $\dfrac{2(\beta,\alpha)}{(\alpha,\alpha)}$ by $\langle\beta,\alpha\rangle$, as this comes up frequently. For instance, we can then write $\sigma_\alpha(\beta) = \beta - \langle\beta,\alpha\rangle\alpha$. We note that $\langle\beta,\alpha\rangle$ is linear in only the first argument.

The canonical example of a root system is the inner product endowed on the dual space of a maximal toral subalgebra of a semisimple Lie algebra. We study some properties of these roots systems and their roots.

The study of root systems is a key component in the classification of semisimple Lie algebras (which we will not pursue in this text). Part of our interest in root systems lies in their being foundational to the construction of the ***Chevalley Bases*** of semisimple Lie algebras. We will return these in the next section.

2.1 Roots and Bases

This section will serve as an introduction to root systems. We consider the structure of roots and show that any root system has a special type of basis, called a base, with very useful properties

Remark 2.1: *Let $\alpha,\beta \in \Phi$. Then $\langle\beta,\alpha\rangle = 2\dfrac{\|\beta\|}{\|\alpha\|}\cos\theta$, where θ is the angle between the vectors. Hence $\langle\alpha,\beta\rangle\langle\beta,\alpha\rangle = 4\cos^2\theta$, which is a non-negative integer. Since $0 \le \cos^2\theta \le 1$, if we assume without loss of generality that $\|\beta\| \ge \|\alpha\|$ then for $\alpha \ne \pm\beta$ we have the following possibilities:*

$\langle\alpha,\beta\rangle$	$\langle\beta,\alpha\rangle$	θ	$\|\beta\|^2 / \|\alpha\|^2$
0	0	$\pi/2$	*Undetermined*
1	1	$\pi/3$	1
-1	-1	$2\pi/3$	1
1	2	$\pi/4$	2
-1	-2	$3\pi/4$	2
1	3	$\pi/6$	3
-1	-3	$5\pi/6$	3

.

Lemma 2.2: *Let α and β be non-proportional roots. If $(\alpha,\beta) > 0$, then $\alpha - \beta \in \Phi$. If $(\alpha,\beta) < 0$, then $\alpha + \beta \in \Phi$.*

Proof. It suffices to prove the first assertion, because if we apply this to $-\beta$ the second will follow. We know that (α,β) can only be positive if $\langle\alpha,\beta\rangle$ is positive. We then see that at least one of $\langle\alpha,\beta\rangle, \langle\beta,\alpha\rangle$ is 1. If the former is positive, then $\sigma_\beta(\alpha) = \alpha - \beta \in \Phi$. If the latter is positive, a similar argument yields that $\beta - \alpha$ is a root, whence $\alpha - \beta$ is a root. □

Definition 2.3: *A subset Δ of a root system Φ (for a euclidean space E) is called a* **base** *if the following conditions are satisfied:*

1. *Δ is a basis for E,*
2. *For each root $\beta \in \Phi$ there are $k_\alpha \in \mathbb{Z}$ such that we may uniquely express β as $\beta = \sum_{\alpha\in\Delta} k_\alpha \alpha$, where the k_α are all non-negative or all non-positive.*

We call roots in Δ **simple roots***.*

We define the **height** *of a root, relative to Δ, as $ht(\beta) := \sum_{\alpha\in\Delta} k_\alpha$. The height is well-defined since the expression for a root as the sum of simple roots is unique because Δ is a basis for E. Roots whose expression as a sum of simple roots include only non-negative (respectively non-positive) coefficients are called* **positive** *(respectively* **negative***) roots, denoted $\beta \succ 0$ (respectively $\beta \prec 0$). In fact, this defines a partial order on $E : \gamma \prec \delta$ iff $\delta - \gamma$ is a sum of positive roots or $\delta = \gamma$, for all $\delta, \gamma \in E$.*

Remark 2.4: *Our goal now is to show that every root system has a base and moreover, to outline a method with which we can construct every possible base. To this end, we establish some notation. For each $\delta \in E$ we set $\Phi^+(\delta) := \{\alpha \in \Phi \mid (\delta, \alpha) > 0\}$, that is, all the roots which lie on the 'positive' side of the hyperplane P_δ. We will call a vector $\delta \in E$* **regular** *if we have $\delta \in E - \bigcup_{\alpha\in\Phi} P_\alpha$, and* **singular** *otherwise. If $\delta \in E$ is regular, then we can see that $\Phi^+(\delta) \cup -\Phi^+(\delta) = \Phi$. Relative to a regular vector δ, we call a root $\alpha \in \Phi^+(\delta)$* **decomposable** *if there are roots $\beta_1, \beta_2 \in \Phi^+(\delta)$ such that $\alpha = \beta_1 + \beta_2$, and we call α* **indecomposable** *otherwise.*

Lemma 2.5: *Let E be a finite-dimensional vector space over an infinite field. The union, $D := \bigcup_{i=1}^n D_i$, of any finite number of distinct proper subspaces of the same dimension is not a subspace.*

Proof. We proceed by induction on n, the number of proper subspaces. If $n = 2$, we may choose $d_1 \in D_1 - D_2$ and $d_2 \in D_2 - D_1$. We claim $d := d_1 + d_2 \notin D_1 \cup D_2$. If $d \in D_1$, then $d_2 = d - d_1 \in D_1$. Similarly, $d \notin D_2$.

We now consider the case where $n > 1$ and assume the result holds for $n-1$. We must have the D_1 is not contained in $\bigcup_{i=2}^n D_i$, since this would contradict the result holding for $n-1$. Hence, there is some $d_1 \in D_1$ which is not in any other D_i. Similarly, we find $d_2 \in D_2$ which is not in any other D_i. Choose $n-1$ distinct non-zero scalars $\delta_1, \delta_2, \ldots, \delta_{n-1}$. Set $c_i := d_1 + \delta_i d_2$. Each of the c_i are elements of U, we claim they must all lie in pairwise distinct subspaces D_j. To see this, consider that if c_1 and c_2 were both contained in D_3, then their difference would also be an element of D_3. This would imply that $d_2 \in D_3$, a contradiction.

Keeping in mind that D_1 and D_2 cannot contain any of the c_i, an application of the Pigeonhole Principle yields that the at least one of the $n-2$ remaining subspaces must contain two of the c_i. □

Corollary 2.6: *The union of finitely many hyperplanes* $\bigcup_{\alpha \in \Phi} P_\alpha$ *is a proper subset of* E.

Lemma 2.7: *Suppose* Δ *is a base of* Φ. *Then* $(\alpha, \beta) \leq 0$ *for all distinct* $\alpha, \beta \in \Delta$. *Furthermore,* $\alpha - \beta$ *is not a root.*

Proof. Suppose, to the contrary, that $(\alpha, \beta) > 0$, then Lemma 2.2 yields that $\alpha - \beta$ is a root. This contradicts every root being a linear combination of simple roots, wherein the coefficients are all non-negative or all non-positive. □

Lemma 2.8: *Let* Δ *be a set of vectors lying on one side of a hyperplane in a euclidean space* E. *If* $(\alpha, \beta) \leq 0$ *for distinct* $\alpha, \beta \in \Delta$, *then* Δ *is a linearly independent set.*

Proof. Suppose we have $d_\delta \in \mathbb{R}$ such that $\sum_{\delta \in \Delta} d_\delta \delta = 0$. We may denote the non-negative and negative d_δ as $p_\alpha (\geq 0)$ and $n_\beta (> 0)$, respectively. We may then write $\sum_\alpha p_\alpha \alpha = \sum_\beta n_\beta \beta$. We set $\epsilon := \sum_\alpha p_\alpha \alpha$ and find that

$$(\epsilon, \epsilon) = \sum_{\alpha, \beta} p_\alpha n_\beta (\alpha, \beta) \leq 0,$$

by hypothesis. This yields that $\epsilon = 0$.

By assumption all the vectors in Δ lie on one side of a hyperplane in E. There is some vector $\gamma \in E$ which is orthogonal to this hyperplane. We can see that $0 = (\gamma, \epsilon) = \sum_\alpha p_\alpha (\gamma, \alpha)$. This forces that each $p_\alpha = 0$. Similarly, the n_β are all 0. □

Lemma 2.9: *The intersection of the 'positive' open half-spaces, corresponding to some basis of* E, *is non-empty.*

Proof. Fix a basis $B := \{\beta_i\}_{i=1}^n$. We let ρ_i denote the projection of β_i onto the orthogonal complement of the span of the basis elements other than β_i. We let $\beta := \sum_{i=1}^n c_i \rho_i$, where each $c_i > 0$.

For each $k \in \{1, \ldots, n\}$ we have $(\beta_k, \beta) = \sum_{i=1}^n c_i (\beta_k, \rho_i) = c_k (\beta_k, \rho_k)$. Furthermore, we claim that $(\beta_k, \rho_k) > 0$. In general, for a subspace U of E we have that $(\gamma, \delta) \geq 0$, where $\gamma \in E$ and δ denotes the projection of γ onto U. This follows because, if $\{\mu_1, \ldots, \mu_m\}$ is an orthonormal basis for U, then

$$(\gamma, \delta) = \left(\gamma, \sum_{i=1}^m \frac{(\gamma, \mu_i)}{(\mu_i, \mu_i)} \mu_i \right) = \left(\gamma, \sum_{i=1}^m (\gamma, \mu_i) \mu_i \right) = \sum_{i=1}^m (\gamma, \mu_i)^2 \geq 0.$$

In this case, if $(\beta_k, \rho_k) = 0$ then β_k is in the span of $B - \{\beta_k\}$, which contradicts B being a basis. Hence $(\beta_k, \rho_k) > 0$ and so β is in the intersection of the 'positive' open half spaces. □

Theorem 2.10: *Fix a regular vector $\gamma \in E$. The set $\Delta(\gamma)$ consisting of all indecomposable roots in $\Phi^+(\gamma)$ is a base of Φ.*

Proof. We note that a regular $\gamma \in E$ must exist as a result of Corollary 2.6.

We claim that each root in $\Phi^+(\gamma)$ can be written as a linear combination of the elements in $\Delta(\gamma)$ with each coefficient a non-negative integer. Since $\Phi = \Phi^+(\gamma) \cup -\Phi^+(\gamma)$ this is sufficient to prove that $\Delta(\gamma)$ has the second property of a base. Moreover, this would also show that $\Delta(\gamma)$ spans E.

Suppose there is some $\alpha \in \Phi^+(\gamma)$ which cannot be expressed as claimed. We may choose α so as to make (γ, α) minimal. As α cannot be in $\Delta(\gamma)$, there must be $\beta_1, \beta_2 \in \Phi^+(\gamma)$ such that $\alpha = \beta_1 + \beta_2$. Both of the (γ, β_i) must be positive and strictly less than $(\gamma, \alpha) = (\gamma, \beta_1) + (\gamma, \beta_2)$. This means that each of the β_i must be expressible as a linear combination of $\Delta(\gamma)$ with non-negative integer coefficients because (γ, α) is minimal. Since $\alpha = \beta_1 + \beta_2$ we see that α must also be expressible as a linear combination of $\Delta(\gamma)$ with non-negative integer coefficients.

Non-proportional roots $\alpha, \beta \in \Delta(\gamma)$ must have $(\alpha, \beta) < 0$. If not, then from Lemma 2.2 we find that $\alpha - \beta$ is a root. This means that one of $\alpha - \beta$ or $\beta - \alpha$ must be in $\Phi^+(\gamma)$. Then, either $\alpha = \beta + (\alpha - \beta)$ or $\beta = \alpha + (\beta - \alpha)$ is decomposable, a contradiction.

An application of Lemma 2.8 yields that $\Delta(\gamma)$ is linearly independent and therefore a base. □

Corollary 2.11: *Every base Δ of Φ arises in the form $\Delta(\gamma)$ for some regular $\gamma \in E$.*

Proof. Suppose Δ is a base. Lemma 2.9 allows us to choose $\gamma \in E$ such that $(\gamma, \alpha) > 0$, for all $\alpha \in \Delta$. Since any root can be written as a linear combination of elements from Δ with integer coefficients which are all non-negative or all non-positive, we conclude that γ is regular.

We denote by Φ^+ and Φ^- the positive and negative roots, respectively, relative to Δ. For any $\alpha \in \Phi^+$ we may write $\alpha = \sum_{\delta \in \Delta} p_\delta \delta$, where each p_δ is a non-negative integer. Therefore,

$$(\gamma, \alpha) = \sum_{\delta \in \Delta} p_\delta (\gamma, \delta) > 0.$$

Thus, $\Phi^+ \subseteq \Phi^+(\gamma)$, and similarly we find that $\Phi^- \subseteq -\Phi^+(\gamma)$. We know that

$$\Phi^+(\gamma) \cup -\Phi^+(\gamma) = \Phi = \Phi^+ \cup \Phi^-,$$

which allows us to conclude that $\Phi^+ = \Phi^+(\gamma)$ and $\Phi^- = -\Phi^+(\gamma)$.

Therefore Δ must consist enitirely of indecomposable elements, and so $\Delta \subseteq \Delta(\gamma)$. Moreover, each of Δ and $\Delta(\gamma)$ is a basis. Therefore, they have the same cardinality and thus $\Delta = \Delta(\gamma)$. □

2.2 Weyl Group

This section is dedicated to the Weyl group of a root system. We derive some useful results leading up to an important theorem about the Weyl group and Weyl chambers which is key to the results in the following section. In what follows, E denotes a euclidean space, Φ a root system for E. We fix a base Δ for Φ.

Definition 2.12: *The* **Weyl Group** *of Φ, denoted by $\mathcal{W}(\Phi)$ or simply $\mathcal{W}$, is the subgroup of $GL(E)$ generated by the reflections σ_α where $\alpha \in \Phi$. Since such reflections leave Φ invariant, $\mathcal{W}$ may be identified with a subgroup of the symmetric group on Φ. In particular, $\mathcal{W}$ is finite.*

Lemma 2.13: *Suppose D is a finite spanning subset of E such that all reflections corresponding to elements of D permute D, that is $\sigma_\gamma(D) = D$ for all $\gamma \in D$. If $\sigma \in GL(E)$ permutes the elements of D, fixes each point of a hyperplane P of E, and maps some $0 \neq \alpha \in D$ to $-\alpha$, then $\sigma = \sigma_\alpha$ and $P = P_\alpha$.*

Proof. Set $\tau := \sigma\sigma_\alpha$. Then τ leaves D invariant and $\tau(\alpha) = \alpha$. Clearly, τ acts as the identity on $\langle\alpha\rangle$. Moreover, σ fixes P pointwise and $\langle\alpha\rangle$ is a 1-dimensional subspace whose intersection with P is trivial. Hence, τ also acts as the identity on the quotient space $E/\langle\alpha\rangle$. Thus, only eigenvalue of τ is 1. If E has dimension n, then the minimal polynomial of τ must divide the polynomial $(X-1)^n$.

Since D is finite and τ leaves D invariant we may choose k to be the smallest integer such that τ^k fixes each element of D. However, D spans E and so $\tau^k = 1$. Therefore, the minimal polynomial of τ divides $X^k - 1$.

The minimal polynomial of τ must be $X - 1 = g.c.d((X-1)^n, X^k - 1)$. This forces $\tau = 1$, and $\sigma = \sigma_\alpha$, as required. □

Lemma 2.14: *Suppose that $\sigma \in GL(E)$ permutes the elements of Φ. Then for $\alpha, \beta \in \Phi$, then*

1. $\sigma\sigma_\alpha\sigma^{-1} = \sigma_{\sigma(\alpha)}$,
2. $\langle\beta,\alpha\rangle = \langle\sigma(\beta),\sigma(\alpha)\rangle$.

Proof. We set $\tau := \sigma\sigma_\alpha\sigma^{-1}$. By hypothesis we have that $\sigma(\beta) \in \Phi$ and that each root may be uniquely expressed as $\sigma(\beta)$ for some $\beta \in \Phi$. Since $\tau(\sigma(\beta)) = \sigma(\sigma_\alpha(\beta)) \in \Phi$, we can see that τ permutes the elements of Φ. Moreover,

$$\tau(\sigma(\beta)) = \sigma(\beta - \langle\beta,\alpha\rangle\alpha) = \sigma(\beta) - \langle\beta,\alpha\rangle\sigma(\alpha),$$

and $\sigma(P_\alpha) = \{\sigma(\beta) \in E \,|\, (\beta,\alpha) = 0\}$. Hence, τ acts as the identity on the hyperplane $\sigma(P_\alpha)$. Finally $\tau(\sigma(\alpha)) = \sigma(\alpha - \langle\alpha,\alpha\rangle\alpha) = -\sigma(\alpha)$, and (1) follows.

An application of Lemma 2.13 yields that $\tau = \sigma_{\sigma(\alpha)}$ and $\sigma(P_\alpha) = P_{\sigma(\alpha)}$. The second assertion follows because

$$\sigma(\beta) - \langle\sigma(\beta), \sigma(\alpha)\rangle\sigma(\alpha) = \sigma_{\sigma(\alpha)}(\sigma(\beta)) = \tau(\sigma(\beta)) = \sigma(\beta) - \langle\beta, \alpha\rangle\sigma(\alpha).$$

□

Lemma 2.15: *If α is a positive root but not simple, then $\alpha - \beta \in \Phi$ for some simple root β. Moreover, $\alpha - \beta$ is a positive root.*

Proof. We claim that there is some $\beta \in \Delta$ such that $(\alpha, \beta) > 0$. If not, then Lemma 2.8 would apply and then $\Delta \cup \{\alpha\}$ would be linearly independent. This would contradict Δ being a basis.

Since β is not proportional to α Lemma 2.2 yields that $\alpha - \beta$ is a root. We may write α as a linear combination of simple roots with non-negative coefficients, at least one of which (other than β) is positive. Writing $\alpha - \beta$ as a linear combination of simple roots leaves the corresponding coefficient positive and so $\alpha - \beta$ is positive. □

Corollary 2.16: *Each positive root β can be written as a sum of simple roots $\beta = \sum\limits_{i=1}^{m} \alpha_i$ (the α_i are not necessarily distinct) such that each of the partial sums is a root.*

Proof. We proceed by induction on the height of β. If $ht(\beta) = 1$, then β is simple and we are done.

Let $ht(\beta) = n \geq 2$. Using Lemma 2.15 we can see that there is some simple root α such that $\beta - \alpha$ is a positive root. Since $ht(\beta - \alpha) = ht(\beta) - 1$ the hypothesis applies and we may write $\beta - \alpha = \sum_{i=1}^{m} \alpha_i$ such that each partial sum is a root. This allows us to write $\beta = \sum_{i=1}^{m} \alpha_i + \alpha$, and we are done. □

Lemma 2.17: *If α is a simple root, then σ_α permutes the other positive roots.*

Proof. Let β be a positive root with $\beta \neq \alpha$. We may write $\beta = \sum_{\delta \in \Delta} c_\delta \delta$, for some non-negative integers c_δ. Therefore,

$$\sigma_\alpha(\beta) = \sum_{\delta \in \Delta - \{\alpha\}} c_\delta \sigma(\delta) + (c_\alpha - \langle\beta, \alpha\rangle)\alpha.$$

There is at least one $\delta \in \Delta - \{\alpha\}$ such that $c_\delta > 0$. Otherwise $\beta = \alpha$ or $\beta = -\alpha$, which contradicts the assumption that β is positive. Hence, $\sigma_\alpha(\beta)$ must be a positive root. □

Corollary 2.18: *If $\delta := \frac{1}{2} \sum\limits_{\beta \in \Phi^+} \beta$ then $\sigma_\alpha(\delta) = \delta - \alpha$, where α is any simple root.*

Proof. In fact,

$$\sigma_\alpha(\delta) = \frac{1}{2}\sum_{\beta\in\Phi^+}\sigma_\alpha(\beta) = \frac{1}{2}\left(\sum_{\beta\in\Phi^+-\{\alpha\}}\beta - \alpha\right) = \delta - \frac{1}{2}\alpha - \frac{1}{2}\alpha = \delta - \alpha.$$

□

Lemma 2.19: *We let σ_i denote σ_{α_i}. Suppose now, that there is some sequence of (not necessarily pair-wise distinct) simple roots $\alpha_1, \dots, \alpha_n$ such that $\sigma_1 \dots \sigma_{n-1}(\alpha_n) \prec 0$. Then, there is some $1 \le m < n$ such that $\sigma_1 \dots \sigma_n = \sigma_1 \dots \sigma_{m-1}\sigma_{m+1} \dots \sigma_{n-1}$.*

Proof. We set $\tau_i := \sigma_{i+1} \dots \sigma_{n-1}$, for $0 \le i \le n-2$. By hypothesis $\tau_0(\alpha_n)$ is negative and α_n is positive. Let k be the smallest non-negative integer such that $\tau_k(\alpha_n)$ is positive. By construction $\sigma_k\tau_k(\alpha_n)$ is negative. Since $\tau_k(\alpha_n)$ is positive Lemma 2.17 forces $\tau_k(\alpha_n) = \alpha_k$.

Applying Lemma 2.14 we find that

$$\sigma_k = \sigma_{\tau_k(\alpha_n)} = \tau_k\sigma_n\tau_k^{-1} = \sigma_{k+1} \dots \sigma_{n-1}\sigma_n\sigma_{n-1} \dots \sigma_{k+1}.$$

The result follows by multiplying both sides of the equation on the right by τ_k and on the left by $\sigma_1 \dots \sigma_{k-1}$. □

Corollary 2.20: *Suppose the following expression of $\sigma = \sigma_1 \dots \sigma_n \in \mathcal{W}$ in terms of simple reflections has n as small as possible. Then $\sigma(\alpha_n)$ is negative.*

Proof. If $\sigma\sigma_n(\alpha_n)$ were positive the lemma would apply, contradicting our assumption that n is minimal. □

Definition 2.21: *The hyperplanes P_α corresponding to roots $\alpha \in \Phi$ partition E into a finite number of regions. Each connected component of $E - \bigcup_{\alpha\in\Phi} P_\alpha$ is called a* **Weyl chamber** *of E.*

Remark 2.22: *Each regular vector will lie in exactly one Weyl chamber. We will denote the Weyl chamber corresponding to a regular vector $\gamma \in E$ by $\mathcal{C}(\gamma)$. For regular vectors $\gamma, \gamma' \in E$ we find $\mathcal{C}(\gamma) = \mathcal{C}(\gamma')$ if and only if γ and γ' lie on the same side of each hyperplane P_α, corresponding to $\alpha \in \Phi$. Furthermore, Corollary 2.11 shows us that Weyl chambers are in one-to-one correspondence with bases.*

If the base $\Delta = \Delta(\gamma)$ for regular $\gamma \in E$, then we set $\mathcal{C}(\Delta) := \mathcal{C}(\gamma)$ and call this the **fundamental Weyl chamber** *relative to Δ. Since $\mathcal{C}(\Delta)$ is, by construction, the intersection of open 'positive' half spaces relative to the simple roots, it is an open convex set.*

Lemma 2.23: *The Weyl group permutes Weyl chambers and bases. Moreover, these actions of $\mathcal{W}$ preserve the correspondence between Weyl chambers and bases determined by the regular vectors.*

Proof. The result follows because $\mathcal{W}$ is generated by reflections; which are orthogonal, linear, and bijective. □

Theorem 2.24:

1. *The Weyl group acts transitively on the Weyl chambers.*
2. *The Weyl group acts transitively on bases.*
3. *For each root α, there is some $\sigma \in \mathcal{W}$ such that $\sigma(\alpha) \in \Delta$.*
4. *The Weyl group is generated by simple reflections (reflections corresponding to simple roots).*
5. *The Weyl group acts simply transitively on bases.*

Proof. We proceed by proving parts (1), (2), and (3) for the subgroup $\mathcal{W}'$ of $\mathcal{W}$ which is generated by simple reflections. The full result for these parts will then follow from part (4).

(1): It suffices to show that for any regular $\gamma \in E$ there is some $\sigma \in \mathcal{W}'$ such that $\sigma(\gamma)$ lies in the fundamental Weyl chamber, since this shows that any Weyl chamber can be mapped to the fundamental Weyl chamber (and therefore to any other) by some element of $\mathcal{W}'$.

Set $\delta := \frac{1}{2}\sum_{\alpha\in\Phi^+}\alpha$. Since $\mathcal{W}$ is finite (and therefore $\mathcal{W}'$ also) we may choose $\sigma \in \mathcal{W}'$ such that $(\sigma(\gamma), \delta)$ is maximal. Since reflections are orthogonal and using Corollary 2.18 we find that

$$(\sigma(\gamma), \delta) \geq (\sigma_\alpha\sigma(\gamma), \delta) = (\sigma(\gamma), \sigma_\alpha(\delta)) = (\sigma(\gamma), \delta) - (\sigma(\gamma), \alpha),$$

for any $\alpha \in \Phi$. From which we may conclude $(\sigma(\gamma), \alpha) \geq 0$. To see that this inequality is strict we note that if there were some root α for which equality held, then the regular vector γ would be orthogonal to the root $\sigma^{-1}(\alpha)$.

(2): This follows from Lemma 2.23 and part (1).

(3): Let α be a root. We claim that α must belong to at least one base. We set $R := \Phi - \{\pm\alpha\}$. For each $\beta \in R$, we have that $P_\alpha \cap P_\beta$ is a proper subspace of P_α. Applying Lemma 2.5 to $\bigcup_{\beta\in R}(P_\alpha \cap P_\beta)$ we find that there is some $\gamma \in P_\alpha$ which does not lie in any P_β, where $\beta \in R$. We may now choose γ' which is near enough to γ that $(\gamma', \alpha) > 0$ while also $(\gamma', \alpha) < |(\gamma', \beta)|$, for all $\beta \in R$. It follows from the definition of $\Phi^+(\gamma')$ and being indecomposable that $\alpha \in \Delta(\gamma')$.

The result follows from an appropriate application of part (2) to $\Delta(\gamma')$.

(4): Since the Weyl group is generated by the reflections σ_α, where $\alpha \in \Phi$, it suffices to show that each such reflection lies in $\mathcal{W}'$. For any $\alpha \in \Phi$ part (3) allows us to find $\sigma \in \mathcal{W}'$ such that $\sigma(\alpha) \in \Delta$. Since $\sigma \in \mathcal{W}'$ is generated by simple reflections we may apply Lemma 2.14 to find that $\sigma_{\sigma(\alpha)} = \sigma\sigma_\alpha\sigma^{-1}$. Therefore, $\sigma_\alpha = \sigma^{-1}\sigma_{\sigma(\alpha)}\sigma$. Since $\sigma(\alpha) \in \Delta$ we conclude that $\sigma_\alpha \in \mathcal{W}'$ and the result follows.

(5): Suppose we have $\sigma \in \mathcal{W}$ such that $\sigma(\Delta) = \Delta$. Part (4) allows us to express σ as a product of simple reflections. If we do so with the minimal necessary number of simple reflections, then Corollary 2.20 forces σ to be the identity map. Otherwise some simple root would be mapped to it's negative and this would contradict the assumption that Δ is invariant under σ.

□

2.3 Isomorphism Theorem

Following our introduction to root systems and their bases, we now work towards a result showing that Lie algebras with isomorphic root systems will themselves be isomorphic. The structure of roots and bases will allow us to come to a stronger conclusion.

Following from Theorem 2.10 we may fix a base Δ for Φ.

Definition 2.25: *A root system Φ is said to be* **reducible** *if we are able to express $\Phi = \Phi_1 \cup \Phi_2$ as the disjoint union of two proper subsets such that $(\Phi_1, \Phi_2) = 0$, that is, each root in one subset is orthogonal to every root in the other. If Φ cannot be partitioned in this way we say it is* **irreducible**.

Lemma 2.26: *The root system Φ is irreducible if and only if its base Δ cannot be partitioned into two disjoint orthogonal proper subsets.*

Proof. We consider the direct case. Suppose $\Phi = \Phi_1 \cup \Phi_2$ is reducible. We set $\Delta_1 := \Delta \cap \Phi_1$ and $\Delta_2 := \Delta \cap \Phi_2$. Clearly, $\Delta = \Delta_1 \cup \Delta_2$. It will suffice to show that neither Δ_1 nor Δ_2 is empty. Let us suppose, without loss of generality, that Δ_2 is empty, then $\Delta = \Delta_1 \subseteq \Phi_1$. Thus, $(\Delta, \Phi_2) = 0$, which implies that $(\Phi_2, \Phi_2) = 0$ since Δ is a basis. Therefore, Φ_2 is empty, which contradicts our assumption that Φ is reducible.

Conversely, let us suppose that $\Delta = \Delta_1 \cup \Delta_2$ is reducible. If $\alpha, \beta \in \Phi$ are orthogonal, then their corresponding reflections, σ_α and σ_β, commute. Part (4) of Theorem 2.24 then yields that $\mathcal{W} = \mathcal{W}_1 \times \mathcal{W}_2$, (where $\mathcal{W}_i$ is generated by the simple reflections corresponding to simple roots in Δ_i). From part (3) of Theorem 2.24 we know that each root is conjugate to a simple root.

Now, the $\mathcal{W}_i$ fix the orthogonal subspaces of the spans of the Δ_i, respectively. Therefore, the roots which are conjugate to simple roots in Δ_1 cannot be conjugate to those which are conjugate to the simple roots in Δ_2. This shows that Φ is reducible as a partition into the subsets of roots conjugate to Δ_1 and Δ_2, respectively. □

Lemma 2.27: *Suppose Φ is irreducible. Relative to the partial order introduced in Definition 2.3 there is a unique maximal root β. If $\beta = \sum_{\alpha\in\Delta} b_\alpha \alpha$, then each $b_\alpha > 0$.*

Proof. Let $\beta = \sum_{\alpha\in\Delta} b_\alpha \alpha$ be maximal. Set $\Delta_+ := \{\alpha \in \Delta \,|\, b_\alpha > 0\}$ and $\Delta_0 := \{\alpha \in \Delta \,|\, b_\alpha = 0\}$. Notice that $\Delta = \Delta_+ \cup \Delta_0$ is a partition. Suppose there is some $\alpha \in \Delta_0$. From Lemma 2.7 we find that $(\alpha, \beta) \leq 0$. We may assume w.l.o.g. that $\{\alpha\}$ is not orthogonal to Δ_+ since Φ is irreducible. Hence, there is some $\alpha_+ \in \Delta_+$ such that $(\alpha, \alpha_+) < 0$, and so $(\alpha, \beta) < 0$. Lemma 2.2 yields that $\beta + \alpha$ is a root, contradicting the maximality of β. Hence Δ_0 is empty.

The argument above also shows that $(\alpha, \beta) \leq 0$ for all $\alpha \in \Delta$ and maximal β. Suppose β' is also a maximal root, then all the above also apply to β'. Thus, it follows that $(\beta, \beta') > 0$. Now, either $\beta = \beta'$ or $\beta - \beta'$ is a root. However, if $\beta - \beta'$ is a root then either $\beta \prec \beta'$ or $\beta' \prec \beta$, contradicting their mutual maximality. □

Lemma 2.28: *Let Φ be a root system in E and E' a subspace of E. If a reflection σ_α leaves E' invariant then $\alpha \in E'$ or E' is contained in the hyperplane orthogonal to α.*

Proof. Fix a basis $\{e_i\}_{i=1}^k$ for E'. Then $\sigma_\alpha(e_i) = e_i - \langle e_i, \alpha\rangle\alpha$. If $\langle e_i, \alpha\rangle = 0$ for each i, then the reflection fixes E' and thus E' is orthogonal to α. Otherwise there is some $j \in \{1, \ldots, k\}$ for which $\langle e_j, \alpha\rangle$ is non-zero. Then there are coefficients a_i such that

$$e_j - \langle e_j, \alpha\rangle\alpha = \sum_{i=1}^{k} a_i e_i.$$

Rearranging, we may write

$$\alpha = \frac{1}{\langle e_j, \alpha\rangle} e_j - \sum_{i=1}^{k} \frac{a_i}{\langle e_j, \alpha\rangle} e_i \in E'.$$

□

Proposition 2.29: *Let L be a simple Lie algebra. Then Φ is irreducible.*

Proof. Suppose that Φ is reducible, then we may decompose $\Phi = \Phi_1 \cup \Phi_2$, such that Φ_1 is orthogonal to Φ_2. Choose arbitrary $\alpha \in \Phi_1$ and $\beta \in \Phi_2$. Then $(\alpha + \beta, \alpha), (\alpha + \beta, \beta) > 0$, specifically both are non-zero. Therefore, $\alpha + \beta$ cannot be a root, since it would need to be in exactly one of the Φ_i.

Morever, $[L_\alpha, L_\beta] = 0$, keeping in mind that α and β are arbitrary. We denote by K the subalgebra of L generated by the L_α where $\alpha \in \Phi_1$.

We claim that K is a proper ideal of L, which would contradict L being simple. Notice that $[K, L_\beta] = 0$, for any $\beta \in \Phi_2$. K is a proper subalgebra of L since $Z(L) = 0$. Moreover, the normaliser of K clearly contains all the L_α and the L_β. Hence, K is a proper ideal of L. □

Corollary 2.30: *Let $\{L_i\}_{i=1}^k$ be simple Lie algebras and set $L := \bigoplus_{i=1}^k L_i$. Let H be a maximal toral subalgebra of L and Φ the associated root system. Then $H_i := H \cap L_i$ is a maximal toral subalgebra of L_i and the corresponding root system Φ_i is irreducible. Moreover, Φ canonically decomposes into $\coprod_{i=1}^n \Phi_i$.*

Proof. Since L decomposes into the L_i, we have that $H = \bigoplus_{i=1}^k H_i$. Clearly each H_i is toral in each L_i, respectively. We claim each H_i is maximal toral in each L_i. Suppose H_i is contained in some toral subalgebra T of L_i. Notice that T is also toral in L and $[T, H_j] = 0$ $(i \neq j)$. Then $T \cup H$ generates a toral subalgebra of L which properly contains H, contradicting the maximality of H.

Now for each $\alpha \in \Phi_i$ we may view α as a root of L relative to H by setting $\alpha(H_j) = 0$ for $j \neq i$. Clearly, $L_\alpha \subseteq L_i$. On the other hand, if $\alpha \in \Phi$, then there must be some i such that $[L_\alpha, H_i]$ is non-zero, otherwise α would not be a root. Therefore, $L_\alpha \subseteq L_i$ and so $\alpha|_{H_i}$ (that is α restricted to H_i) is a root of L_i relative to H_i.

It now suffices to show that if E is the span of Φ, then Φ uniquely decomposes as the union of irreducible root systems Φ_i in E_i such that $E = \bigoplus_{i=1}^k E_i$. If W is the Weyl group of Φ, then the Weyl groups W_i corresponding to the Φ_i are, respectively, $W \cap GL(E_i)$. Since σ_α acts trivially on E_i when $\alpha \in \Phi_j$ where $j \neq i$, we find that each E_i is W-invariant. Lemma 2.28 then shows that each root lies in exactly one of the E_i and therefore that Φ decomposes into the Φ_i. □

Proposition 2.31: *The semisimple Lie algebra L is generated by the root spaces $L_\alpha, L_{-\alpha}$ corresponding to simple roots ($\alpha \in \Delta$).*

Proof. Let β be some positive root. From the Corollary of Lemma 2.15 we may write $\beta = \sum_{i=1}^k \alpha_i$ as a sum of simple roots such that each partial sum is a root. Remembering that $[L_\gamma, L_\varphi] = L_{\gamma+\varphi}$ for roots γ and φ, we can see that L_β is contained in the subalgebra of L which is generated by $L_{\alpha_i}(1 \leq i \leq k)$. Hence, L_β is certainly contained in the subalgebra of L which is generated by the root spaces corresponding to all the simple roots $L_\alpha(\alpha \in \Delta)$. Similary, if β is negative, then L_β is contained in the $L_{-\alpha}(\alpha \in \Delta)$.

The root space decomposition $L := H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha$ yields the result, since $H = \sum_{\alpha \in \Phi}[L_\alpha, L_{-\alpha}]$, and because all the roots may be written as linear combinations of simple roots. □

Theorem 2.32: *Let L and L' be simple Lie algebras with respective maximal toral subalgebras H and H' and respective root systems Φ and Φ'. Suppose that $\tau_\Phi\colon \Phi \to \Phi'$ is an isomorphism (we will denote $\tau_\Phi(\alpha) = \alpha'$), which induces an isomorphism $\tau\colon H \to H'$. We fix a base Δ for Φ, such that $\Delta' := \{\alpha' \,|\, \alpha \in \Delta\}$ is a base for Φ'. If, for each $\alpha \in \Delta$, we choose some non-zero $x_\alpha \in L_\alpha$ and $x_{\alpha'} \in L_{\alpha'}$ (this is equivalent to choosing some isomorphism $\tau_\alpha\colon L_\alpha \to L_{\alpha'}$), then there is a unique isomorphism $\tau_L\colon L \to L'$ which extends τ and each $\tau_\alpha(\alpha \in \Delta)$.*

Proof. We note that any isomorphism $\tau_\Phi\colon \Phi \to \Phi'$ uniquely induces an isomorphism $\tau_*\colon H^* \to H'^*$ since Φ and Φ' span H^* and H'^*, respectively. Furthermore, the identification between a maximal toral subalgebra and its dual provided by the Killing form, means that τ_* induces an isomorphism $\tau\colon H \to H'$.

We first prove the uniqueness of τ_L before proving its existence. For each simple root α the chosen x_α uniquely determines $y_\alpha \in L_{-\alpha}$ such that $[x_\alpha, y_\alpha] = h_\alpha$. Moreover L is generated by the L_α and $L_{-\alpha}(\alpha \in \Delta)$ by Proposition 2.31.

To prove the existence of τ_L we construct a subalgebra D of $L \oplus L'$ and claim that D is isomorphic to L and L' under the respective projections. As above, for each simple root α the corresponding x_α uniquely determines $y_\alpha \in L_{-\alpha}$, such that $[x_\alpha, y_\alpha] = h_\alpha$. Similarly in L', $[x'_{\alpha'}, y'_{\alpha'}] = h'_{\alpha'}$. We set $\bar{x}_\alpha := (x_\alpha, x'_{\alpha'}) \in L \oplus L'$, and define $\bar{y}_\alpha$ and $\bar{h}_\alpha$ similarly. We then let $D := \langle \bar{x}_\alpha, \bar{y}_\alpha, \bar{h}_\alpha \rangle_{\alpha \in \Delta}$.

We claim D is a proper subalgebra of $L \oplus L'$. Since L and L' are simple, Proposition 2.29 yields that Φ and Φ' are irreducible. From Lemma 2.27 we find that there are unique maximal roots (relative to Δ and Δ') β and β' in Φ and Φ', respectively. These maximal roots correspond via τ_Φ because of the correspondence which τ_Φ establishes between the simple roots in Φ and Φ'. We now choose arbitrary non-zero elements $x_\beta \in L_\beta$ and $x'_{\beta'} \in L'_{\beta'}$ and set $\bar{x}_\beta := (x_\beta, x'_{\beta'})$. Let M be the subspace of $L \oplus L'$ spanned by all the $m := ad_{\bar{y}_{\alpha_1}} ad_{\bar{y}_{\alpha_2}} \dots ad_{\bar{y}_{\alpha_k}}(\bar{x}_\beta)$, where each $\alpha_i \in \Delta$ (or equivalently, M is the submodule of $\langle \bar{y}_\alpha \rangle_{\alpha \in \Delta}$ generated by $\bar{x}_\beta$). Since $\bar{y}_{\alpha_i} \in L_{-\alpha_i} \oplus L'_{-\alpha'_i}$, we have that each $m \in L_{\beta - \sum_i \alpha_i} \oplus L'_{\beta' - \sum_i \alpha'_i}$. Therefore $M \cap (L_\beta \oplus L'_{\beta'}) = \mathbb{F}\bar{x}_\beta$ is 1-dimensional. This makes M a proper subspace of $L \oplus L'$ since it contains no elements of $L \oplus L'$ wherein the the first and second components of the element do not correspond via τ_Φ.

We claim that $[D, M] \subseteq M$. By definition $ad_{\bar{y}_\alpha}(\alpha \in \Delta)$ stabilizes M. Since $[\bar{h}, \bar{y}_\alpha]$ is a multiple of $\bar{y}_\alpha$, for every $\bar{h} \in H \oplus H'$, we may conclude that $ad_{\bar{h}_\alpha}$ also stabilizes M. We know that the difference of simple roots is not a root (Lemma 2.7). This means that $ad_{\bar{x}_\alpha}$ commutes with $ad_{\bar{y}_\gamma}$ for all simple $\gamma \neq \alpha$. Hence, if $\alpha \neq \alpha_1$, then by an easy induction argument on k we find that

$$[\bar{x}_\alpha, [\bar{y}_{\alpha_1}, \ldots [\bar{y}_{\alpha_k}, \bar{x}_\beta] \ldots]] = [\bar{y}_{\alpha_1}, [\bar{x}_\alpha, [\bar{y}_{\alpha_2}, \ldots [\bar{y}_{\alpha_k}, \bar{x}_\beta] \ldots]]]$$

lies in $[\bar{y}_{\alpha_1}, M] \subseteq M$. Whereas, if $\alpha = \alpha_1$, then

$$\begin{aligned}[\bar{x}_\alpha, [\bar{y}_\alpha, [\bar{y}_{\alpha_2}, \ldots [\bar{y}_{\alpha_k}, \bar{x}_\beta] \ldots]]] = {} & [\bar{y}_\alpha, [\bar{x}_\alpha, [\bar{y}_{\alpha_2}, \ldots [\bar{y}_{\alpha_k}, \bar{x}_\beta] \ldots]] \\ & + [\bar{h}_\alpha, [\bar{y}_{\alpha_2}, \ldots [\bar{y}_{\alpha_k}.\bar{x}_\beta] \ldots]]]\end{aligned}$$

The first summand lies in M by induction and the second summand lies in M because, as we have seen, $ad_{\bar{h}_\alpha}$ stabilizes M. This allows us to conclude that D is a proper subalgebra, otherwise M would be a proper ideal of $L \oplus L'$ but M is clearly distinct from both L and L'.

We claim that the projections of D onto L and L' are isomorphisms. The projections are of course Lie algebra homomorphisms. Moreover, by virtue of their construction and Proposition 2.31 they are surjective. Noting that L is the kernel of the projection of D onto L', suppose $D \cap L \neq 0$. Then there is a non-zero $(w, 0) \in D$. Now $\bar{x}_\alpha$ acts on $(w, 0)$ by

$$[\bar{x}_\alpha, (w, 0)] = ([x_\alpha, w], 0),$$

and $\bar{y}_\alpha$ acts similarly on $(w, 0)$. An application of Proposition 2.31 yields that $[D, (w, 0)] = ([L, w], 0) = (L, 0)$. This means that D includes L and by symmetry D would also include L'. Altogether this implies that $D = L \oplus L'$, a contradiction. Therefore, the projections are also injective, and hence are isomorphisms.

We now have that D is isomorphic (via projections) to both L and L'. This allows us to construct an explicit isomorphism $L \to L'$ by $x_\alpha \mapsto x'_{\alpha'}$ $(\alpha \in \Delta)$ and $h_\alpha \mapsto h'_{\alpha'}$. Cleary this coincides with τ on H, and so the result follows. □

Corollary 2.33: *We may generalise the theorem to the case where L and L' are semisimple.*

Proof. Take L and L' as in the hypothesis of the theorem except that they are semisimple. We may decompose L and L' as direct sums of ideals (which are simple Lie algebras in their own right) as follows

$$L = \bigoplus_{i=1}^{k} L_i, \qquad L' = \bigoplus_{i=1}^{m} L'_i.$$

From Corollary 2.30 we have the canonical decomposition $\Phi := \coprod_{i=1}^{k} \Phi_i$ and $\Phi' = \coprod_{i=1}^{m} \Phi'_i$ where the Φ_i and Φ'_i are the (irreducible) root spaces corresponding to their respective maximal toral subalgebras $H_i := H \cap L_i$ and $H'_i := H' \cap L'_i$. Since decompositions of a semisimple Lie algebra into simple Lie algebras is unique (up to order), and also the decomposition of their root systems into irreducible root systems is unique we note that the indices k and m in the decompositions are equal. We may also assume w.l.o.g. that the order of the decomposition matches up isomorphic ideals, i.e. that $L_i \cong L'_i$ for each i.

We may apply the theorem to each of the L_i and L'_i to get isomorphisms $\tau_{L_i} \colon L_i \to L'_i$ which fulfil the criteria in the hypothesis. We can then define an isomorphism $\tau_L \colon L \to L'$ by combining these τ_{L_i}. The result follows. □

Proposition 2.34: *Let L be a semisimple Lie algebra. Fix an arbitrary non-zero $x_\alpha \in L_\alpha (\alpha \in \Delta)$ and choose $y_\alpha \in L_{-\alpha}$ such that $[x_\alpha, y_\alpha] = h_\alpha$. There is an automorphism, σ of L of order 2 which satisfies the following:*

1. $\sigma(x_\alpha) = -y_\alpha$,
2. $\sigma(y_\alpha) = -x_\alpha$,
3. $\sigma(h) = -h$

for all $\alpha \in \Delta$ and $h \in H$.

Proof. We have seen in Theorem 2.32 that each automorphism of Φ induces an automorphism of H which can be extended to L. We therefore begin by defining an automorphism σ_Φ which sends each root to its negative. This clearly defines a vector space automorphism of H^*. Moreover, for $\alpha, \beta \in H^*$ we have that

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

and so this is an automorphism of Φ. Since each root is sent to its negative, it follows that the induced map on H sends each $h \in H$ to its negative. We may label this induced automorphism $\sigma \colon H \to H$. One of the orthogonality properties of roots states that $-h_\alpha = h_{-\alpha}$ for each $\alpha \in \Phi$. For each simple root α we set $\sigma(x_\alpha) = y_\alpha$ ($[x_\alpha, y_\alpha] = h_\alpha$). We may now apply Theorem 2.32 (or its Corollary, whichever is appropriate) to extend σ to L.

Finally, σ has order 2 because σ^2 acts as the identity on a set of generators, the simple roots (see Proposition 2.31). □

2.4 Chevalley Bases

This section is devoted to proving the existence, and special properties, of Chevalley bases. To this end, we will begin with some results concerning representations of semisimple Lie algebras. Specifically, we will need to look at modules of the special linear Lie algebra $\mathfrak{sl}(2, \mathbb{F})$.

Remark 2.35: *For vectors $\alpha \in \Phi$ and $\beta \in \Phi \cup \{0\}$ we define the* α**-string through** β *as*

$$M_{\beta,\alpha} := \bigoplus_{c \in \mathbb{F}} L_{\beta+c\alpha},$$

noting that $L_{\beta+c\alpha}$ is non-zero only if $\beta+c\alpha \in \Phi$. $M_{\beta,\alpha}$ is an $\mathfrak{sl}(\alpha)$-submodule of L.

There exists integers r, q such that $\beta - c\alpha \in \Phi$ if and only if c is an integer and $-r \leq c \leq q$. Since reflections corresponding to roots (σ_α for $\alpha \in \Phi$) leave Φ invariant, and σ_α adds a multiple of α to roots, we find that σ_α commutes the α-strings of β. In fact, σ_α reverses the order of the strings, and so $\sigma_\alpha(\beta + q\alpha) = \beta - r\alpha$. This yields that $r - q = \langle \beta, \alpha \rangle$.

Now we can move forward in describing, and proving the existence of, Chevalley bases.

Lemma 2.36: *Let Φ' be a non-empty subset of Φ and let E' denote the subspace of E spanned by Φ'. If Φ' satisfies the following conditions:*

1. *$\Phi' = -\Phi'$, and*
2. *If $\alpha, \beta \in \Phi'$ and $\alpha + \beta \in \Phi$, then $\alpha + \beta \in \Phi'$,*

then Φ' is a root system in E'.

Proof. It is clear that Φ' is finite, and since E' is the span of Φ' it is certainly true that Φ' spans E'. Clearly $0 \notin \Phi' \subseteq \Phi$.

The condition that $\Phi' = -\Phi'$ implies that for any $\alpha \in \Phi'$ we have that $-\alpha \in \Phi'$. Moreover, as $\Phi' \subseteq \Phi$ the only multiples of α which could reside in Φ' are $\pm\alpha$.

Let $\alpha, \beta \in \Phi'$. We may assume without loss of generality that $\|\beta\| \geq \|\alpha\|$. We divide the proof of the third axiom into cases based on the value of $\langle \beta, \alpha \rangle$. If $\langle \beta, \alpha \rangle = 0$, then $\sigma_\alpha(\beta) = \beta \in \Phi'$. If $\langle \beta, \alpha \rangle = 1$, then $\sigma_\alpha(\beta) = \beta - \alpha$. Since $\alpha, \beta \in \Phi'$ and σ_α leaves Φ invariant, we have that $\beta - \alpha \in \Phi'$ by the second hypothesis condition. The case for $\langle \beta, \alpha \rangle = -1$ is similar.

Now we consider the case $\langle \beta, \alpha \rangle = 2$ (for -2 the argument is similar). Here $\sigma_\alpha(\beta) = \beta - 2\alpha \in \Phi$. Considering the α-string through β we can see that $\beta - \alpha \in \Phi$, therefore $\beta + (-\alpha) \in \Phi'$ (by condition 2 in the hypothesis). However, now $\beta - \alpha, -\alpha \in \Phi'$ and so by condition 2 we find that $\beta - 2\alpha \in \Phi'$. This same argument can be extended to cover the case where $\langle \beta, \alpha \rangle = 3$ and the case for -3 is similar.

Finally, for $\alpha, \beta \in \Phi'$ we have that $\langle \beta, \alpha \rangle \in \mathbb{Z}$ because Φ is a root system. □

Corollary 2.37: *Take $\alpha, \beta \in \Phi$ and let E' denote their span as a subspace of E. Then $\Phi \cap (\mathbb{Z}\alpha + \mathbb{Z}\beta)$ is a root system in E'.*

Proposition 2.38: *Suppose α and β are non-proportional roots and let $\beta - r\alpha, \ldots, \beta, \ldots, \beta + q\alpha$ be the α-string through β.*

1. *At most two roots lengths occur in this string.*

2. *If $\alpha+\beta$ is a root, then*

$$r+1=\frac{q(\alpha+\beta,\alpha+\beta)}{(\beta,\beta)}.$$

Proof.

1. From Corollary 2.37 we have that $\Phi' := \Phi\cap(\mathbb{Z}\alpha+\mathbb{Z}\beta)$ is a root system of rank 2 in E' (using the notation in the corollary).

   If Φ' is reducible, then it must be of type $A_1\times A_1$. This would mean that $\Phi' := \{\pm\alpha,\pm\beta\}$.

   If Φ' is irreducible, we may use the fact that irreducible root systems contain at most two different root lengths.

2. We note (see Remark 2.35) that $\langle\beta,\alpha\rangle = r-q$. This yields that

$$r=\frac{2(\beta,\alpha)}{(\alpha,\alpha)}+q.$$

   Therefore

$$\begin{aligned} r+1-\frac{q(\alpha+\beta,\alpha+\beta)}{(\beta,\beta)} &= q+\frac{2(\beta,\alpha)}{(\alpha,\alpha)}+1-\frac{q(\alpha+\beta,\alpha+\beta)}{(\beta,\beta)} \\ &= \frac{2(\beta,\alpha)}{(\alpha,\alpha)}+1-\frac{q(\alpha,\alpha)}{(\beta,\beta)}-\frac{2q(\alpha,\beta)}{(\alpha,\alpha)} \\ &= (\langle\beta,\alpha\rangle+1)\left(1-\frac{q(\alpha,\alpha)}{(\beta,\beta)}\right), \end{aligned}$$

   noting that

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

   We will use the labels $X := \langle\beta,\alpha\rangle+1$ and $Y := 1-\dfrac{q(\alpha,\alpha)}{(\beta,\beta)}$. It suffices to show that X or Y is 0. We split this into cases.

   Case 1: $(\alpha,\alpha)\geq(\beta,\beta)$. This implies that $|\langle\beta,\alpha\rangle|\leq|\langle\alpha,\beta\rangle|$. Since $\beta\neq\pm\alpha$, the possible values for $\langle\beta,\alpha\rangle\langle\alpha,\beta\rangle$ are $0,1,2$, and 3. We can also see that the possible values for $\langle\beta,\alpha\rangle$ are $-1,0$, and 1.
   If $\langle\beta,\alpha\rangle=-1$ then $X=0$ and we are done. Otherwise, we must have that $\langle\beta,\alpha\rangle\geq 0$, whence $(\beta+\alpha,\beta+\alpha)$ is strictly larger than (β,β) and (α,α). This follows as

$$(\beta+\alpha,\beta+\alpha)=(\beta,\beta)+2(\beta,\alpha)+(\alpha,\alpha).$$

We have that $\alpha + \beta$ is a root and so, using part (1), we find that $(\alpha, \alpha) = (\beta, \beta)$. A similar argument yields that $(\beta + 2\alpha, \beta + 2\alpha) > (\beta + \alpha, \beta + \alpha)$. Again using part (1) we conclude that $\beta + 2\alpha$ is not a root, and hence $q = 1$. This means that $Y = 0$.

Case 2: $(\alpha, \alpha) < (\beta, \beta)$. Using part (1) we conclude that $\alpha + \beta$ has the same length as α or β, since $\alpha + \beta$ is a root. Either way $(\alpha, \beta) < 0$ (this follows from a similar argument to that used in the first case). Therefore $\langle \alpha, \beta \rangle < 0$. We find that

$$(\beta - \alpha, \beta - \alpha) > (\beta, \beta) > (\alpha, \alpha).$$

From which we conclude that $\beta - \alpha$ cannot be a root (again using part (1)). Thus, $r = 0$. In this case, the possible values for $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle$ are $0, 1, 2$, and 3. While the possible values for $\langle \alpha, \beta \rangle$ are $-1, 0$, and 1. We have established that $\langle \alpha, \beta \rangle < 0$ and so $\langle \alpha, \beta \rangle = -1$. Using $\langle \beta, \alpha \rangle = r - q$ we may conclude that

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

Thus, $Y = 0$.

□

We are nearly in a position to prove Chevalley's Theorem, which describes and guarantees the existence of a Chevalley basis for L, first we will state some well-known background results. We present these without proof to avoid this dissertation becoming bloated with basic, well-known results. Thereafter we will provide the final work needed for the theorem.

Remark 2.39: *We denote by* $V_m := Span(X^m, X^{m-1}Y, \ldots, XY^{m-1}, Y)$, *the vector subspace of* $\mathbb{F}[X, Y]$ *consisting of all homogeneous polynomials of degree* m. *Clearly,* $\dim(V_m) = m + 1$. *We may define a representation*

$$\begin{aligned} \varphi \colon \mathfrak{sl}(2, \mathbb{F}) &\to \mathfrak{gl}(V_m) \\ e_{12} &\mapsto X \frac{\partial}{\partial Y} \\ e_{21} &\mapsto Y \frac{\partial}{\partial X} \\ e_{11} - e_{22} &\mapsto X \frac{\partial}{\partial X} - Y \frac{\partial}{\partial Y}. \end{aligned}$$

Proposition 2.40: *The* V_m *are irreducible* $\mathfrak{sl}(2, \mathbb{F})$*-modules.*

Theorem 2.41: *Any (finite) irreducible* $\mathfrak{sl}(2, \mathbb{F})$*-module is isomorphic to one of the* V_m

Remark 2.42: *Suppose V is an irreducible $\mathfrak{sl}(2,\mathbb{F})$-module and $v \in V$ an $(e_{11}-e_{22})$-eigenvector such that $e_{12}\cdot v = 0$ (such a vector is called a* **maximal vector***). There is a non-negative integer m such that $(e_{11} - e_{22}) \cdot v = mv$, and $V \cong V_m$. We call such a vector v, a* **highest weight vector** *for V and m is the* **highest weight** *of V. A highest weight vector is unique up to non-zero scalar multiples exactly because* $\dim(V_m) = m + 1$.

Lemma 2.43: *Let V be an $\mathfrak{sl}(2,\mathbb{F})$-module, then V contains a maximal vector.*

Lemma 2.44: *Let V be an irreducible $\mathfrak{sl}(2,\mathbb{F})$-module, and v_0 a maximal vector of weight λ. Set $v_{-1} := 0$ and $v_i := (1/i!)y^i \cdot v_0$, when $i \geq 0$. Then for $i \geq 0$,*

1. $h \cdot v_i = (\lambda - 2i)v_i$,
2. $y \cdot v_i = (i + 1)v_{i+1}$,
3. $x \cdot v_i = (\lambda - i + 1)v_{i-1}$.

Proof. We have that V decomposes as the direct sum of weight spaces associated to h, which we denote $V_\lambda := \{v \in V \mid h \cdot v = \lambda v\}$. We note that for $v \in V_\lambda$ we have

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

Therefore $x \cdot v \in V_{\lambda+2}$, while a similar argument yields that $y \cdot v \in V_{\lambda-2}$. Part (1) follows from repeatedly applying the latter of these conclusions.

Part (2) we prove by induction on k. If $k = 0$, then $y \cdot v_0 = (0 + 1)v_{0+1}$ by the definition of v_1. Assuming the result holds for some $k \geq 0$ we have that

$$(k + 2)v_{k+2} = (k + 2)\frac{1}{(k + 2)!}y^{k+2} \cdot v_0 = y\frac{1}{(k + 1)!}y^{k+1} \cdot v_0 = y \cdot v_{k+1}.$$

Part (3) we prove by induction on k. If $k = 0$, then the result follows trivially. Assuming the result holds for some $k - 1 \geq 0$ we have

$$\begin{aligned} kx \cdot v_k &= x \cdot y \cdot v_{k-1} = [x, y] \cdot v_{k-1} + y \cdot x \cdot v_{k-1} = h \cdot v_{k-1} + y \cdot x \cdot v_{k-1} \\ &= (\lambda - 2(k - 1))v_{k-1} + (\lambda - k + 2)y \cdot v_{k-2} \\ &= (\lambda - 2k + 2)v_{k-1} + (k - 1)(\lambda - k + 2)v_{k-1} \\ &= k(\lambda - k + 1)v_{k-1}. \end{aligned}$$

Divinding both sides by k we are done. □

Lemma 2.45: *Suppose α and β are non-proportional roots. We choose $x_\alpha \in L_\alpha$ and $y_\alpha := x_{-\alpha} \in L_{-\alpha}$ such that $[x_\alpha, y_\alpha] = h_\alpha$, and let $x_\beta \in L_\beta$. If $\beta - r\alpha, \ldots, \beta, \ldots \beta + q\alpha$ is the α-string through β, then*

$$[y_\alpha, [x_\alpha, x_\beta]] = q(r + 1)x_\beta.$$

Proof. If $\alpha+\beta$ is not a root, then $q=0$ and since $[x_\alpha, x_\beta] \in L_{\alpha+\beta} = 0$, both sides of the equation are 0.

More generally, we will consider the $\mathfrak{sl}(\alpha)$-submodule M_β of L generated by x_β. To be an $\mathfrak{sl}(\alpha)$-module M_β must be invariant under the actions of e_α, f_α, and h_α (see Proposition 1.84). The (repeatedly applied) action of h_α on x_β will give us scalar multiples of x_β. The (repeatedly applied) action of e_α on x_β will yield vectors in $L_{\beta+\alpha}, \ldots, L_{\beta+q\alpha}$. Similarly, the (repeatedly applied) action of f_α on x_β will yield vectors in $L_{\beta-\alpha}, \ldots, L_{\beta-r\alpha}$. Hence, M_β has dimension $(r+q+1)$ which is the numbers of roots in the α-string through β.

As M_β is generated by a single element, it must be irreducible. Moreover, we have established that the highest weight of M_β is $r+q$. From Remark 2.42 we find that M_β is isomorphic to V_{r+q}.

Take a maximal vector $v_0 \in L$. Then x_β is a non-zero scalar multiple of v_q, using the notation from Lemma 2.44. This follows because v_0 is a h_α-eigenvector such that $x_\alpha \cdot v_0 = 0$, so $v_q \in L_{\beta+q\alpha}$. Applying the action of y_α q times leaves us with a vector in L_β which is 1-dimensional and hence must be spanned by $x_\beta \in L_\beta$.

Then using Lemma 2.44 we find that

$$[y_\alpha, [x_\alpha, x_\beta]] = (r+q-q+1)[y_\alpha, x_\beta] = (q-1+1)(r+1)x_\beta = q(r+1)x_\beta.$$

□

Proposition 2.46: *There are vectors $x_\alpha \in L_\alpha (\alpha \in \Delta)$ such that*

1. $[x_\alpha, x_{-\alpha}] = h_\alpha$.
2. *If $\alpha, \beta, \alpha+\beta \in \Phi$ and $[x_\alpha, x_\beta] = c_{\alpha,\beta} x\alpha + \beta$, then $c_{\alpha,\beta} = -c_{-\alpha,-\beta}$.*
3. *Moreover, for any such choice of root vectors if $\beta - r\alpha, \ldots, \beta + q\alpha$ is the α-string through β, then*

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

Proof. We will use the automorphism σ from Proposition 2.34. For a root $\alpha \in \Phi$, take an arbitrary non-zero $x_\alpha \in L_\alpha$, then $x_{-\alpha} := -\sigma(x_\alpha) \in L_{-\alpha}$ is non-zero. Since the Killing form on L is non-degenerate, we have that $\kappa(x_\alpha, x_{-\alpha})$ is non-zero. We note that for $c \in \mathbb{F}$, replacing x_α by cx_α results in $\kappa(cx_\alpha, cx_{-\alpha}) = c^2\kappa(x_\alpha, x_{-\alpha})$. Therefore, as $\mathbb{F}$ is algebraically closed, we can choose to make $\kappa(x_\alpha, x_{-\alpha})$ any non-zero value by scaling x_α (and therefore also $x_{-\alpha}$) appropriately. We may therefore choose to set

$$\kappa(x_\alpha, x_{-\alpha}) = \frac{2}{(\alpha,\alpha)}.$$

Using one of the orthogonality properties of roots, this means that

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

Part (1) follows.

Let $\alpha, \beta, \alpha+\beta \in \Phi$, then $[x_\alpha, x_\beta] = c_{\alpha,\beta} x_{\alpha+\beta}$ for some $c_{\alpha,\beta} \in \mathbb{F}$. We apply σ to this equation and find that $[-x_{-\alpha}, -x_{-\beta}] = -c_{\alpha,\beta} x_{-\alpha-\beta}$. However, we also have that $[x_{-\alpha,}, x_{-\beta}] = c_{-\alpha,-\beta} x_{-\alpha-\beta}$. Part (2) follows.

We may now choose $\{x_\alpha\}_{\alpha\in\Phi}$ which satisfy (1) and (2). Suppose that $\alpha, \beta, \alpha+\beta \in \Phi$. Then α and β are linearly independent and hence t_α and t_β are linearly independent. We recall that t_α is defined by $\kappa(t_\alpha, h) = \alpha(h)$, for all $\alpha \in H^*, h \in H$. This is the correspondence which the Killing form establishes between H and H^*. Therefore,

$$\kappa(t_{\alpha+\beta}, h) = (\alpha+\beta)(h) = \alpha(h) + \beta(h) = \kappa(t_\alpha, h) + \kappa(t_\beta, h),$$

for all $h \in H$, yields that $t_{\alpha+\beta} = t_\alpha + t_\beta$. From (1) it now follows that

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

From (2) and the Jacobi identity we also find that

$$\begin{aligned}
[c_{\alpha,\beta} x_{\alpha+\beta}, c_{\alpha,\beta} x_{-\alpha-\beta}] &= [[x_\alpha, x_\beta], -[x_{-\alpha}, x_{-\beta}]] \\
&= -[[x_\beta, -[x_{-\alpha}, x_{-\beta}]], x_\alpha] - [[-[x_{-\alpha}, x_{-\beta}], x_\alpha], x_\beta] \\
&= -[x_\alpha, [x_\beta, [x_{-\alpha}, x_{-\beta}]]] + [x_\beta, [x_\alpha, [x_{-\alpha}, x_{-\beta}]]] \\
&= [x_\alpha, [x_\beta, [x_{-\beta}, x_{-\alpha}]]] + [x_\beta, [x_\alpha, [x_{-\alpha}, x_{-\beta}]]].
\end{aligned}$$

If the β-string through α is $\alpha - r'\beta, \dots, \alpha + q'\beta$, then (replacing α and β with their negatives does not change r, r', q, q') applying Lemma 2.45 to both terms yields

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

Part (3) follows from comparing these coefficients of t_α and t_β. □

Proposition 2.47: *We define the dual system of Φ as $\Phi^\vee = \{\alpha^* \mid \alpha \in \Phi\}$, where $\alpha^\vee := \frac{2\alpha}{(\alpha,\alpha)}$. $\Phi^\vee$ is a root system in E. Moreover, if Δ is a base for Φ, then $\Delta^\vee$ is a base for $\Phi^\vee$.*

Proof. Clearly, $\Phi^\vee$ is finite and does not contain 0. Since $\alpha^\vee$ is a non-zero scalar multiple of α, for each $\alpha \in \Phi$, we find that $\Phi^\vee$ spans E. Suppose $\alpha^\vee,\ c\alpha^\vee \in \Phi^\vee$, for some $c \in \mathbb{R}$. Then $\alpha,\ c\alpha \in \Phi$ which means that $c = \pm 1$. Let $\alpha, \beta \in \Phi$. Then

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

We claim that $(\beta,\beta) = (\sigma_\alpha(\beta), \sigma_\alpha(\beta))$, from which it would follow that $\sigma_{\alpha^\vee}(\beta^\vee) = \sigma_\alpha(\beta)^\vee \in \Phi^\vee$. To this end, we note that

$$\begin{aligned}(\sigma_\alpha(\beta), \sigma_\alpha(\beta)) &= (\beta,\beta) - 2\langle \beta,\alpha\rangle(\beta,\alpha) + \langle\beta,\alpha\rangle^2(\alpha,\alpha) \\ &= (\beta,\beta) - 4\frac{(\beta,\alpha)^2}{(\alpha,\alpha} + 4\frac{(\beta,\alpha)^2}{(\alpha,\alpha)} = (\beta,\beta).\end{aligned}$$

It remains to show that $\alpha, \beta \in \Phi$ implies that $\langle \beta^\vee, \alpha^\vee \rangle \in \mathbb{Z}$. For $\alpha, \beta, \in \Phi$ we have

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

Therefore, $\Phi^\vee$ is a root system.

The Weyl chambers corresponding to $\Phi^\vee$ are determined by the hyperplanes orthogonal to the roots $\alpha^\vee$. However, since

$$(\gamma, \alpha^\vee) = \frac{2}{(\alpha,\alpha)}(\gamma,\alpha),$$

we find that $P_\alpha = P_{\alpha^\vee}$, for each $\alpha \in \Phi$. Hence, the Weyl chambers corresponding to Φ are the same as those corresponding to $\Phi^\vee$. Therefore, if Δ is a base for Φ, then $\Delta^\vee$ is a base for $\Phi^\vee$. □

Definition 2.48: *A basis* $\{x_\alpha, h_i \,|\, \alpha \in \Phi, 1 \le i \le l\}$ *of* L *with the following properties:*

1. $[x_\alpha, x_{-\alpha}] = h_\alpha$.
2. *If* $\alpha, \beta, \alpha+\beta \in \Phi$ *and* $[x_\alpha, x_\beta] = c_{\alpha,\beta} x\alpha + \beta$, *then* $c_{\alpha,\beta} = -c_{-\alpha,-\beta}$.
3. $h_i = h_{\alpha_i}$, *for some base* $\Delta = \{\alpha_1, \ldots, \alpha_l\}$ *of* Φ,

is called a **Chevalley basis**.

Theorem 2.49 (Chevalley's Theorem)**:** *Suppose we have a Chevalley basis* $C := \{x_\alpha,\ h_i \,|\, \alpha \in \Phi,\ 1 \le i \le l\}$ *for* L. *Then*

1. *$[h_i, h_j] = 0$, for $1 \le i, j \le l$.*

2. *$[h_i, x_\alpha] = \langle \alpha, \alpha_i \rangle x_\alpha$, for $1 \le i \le l$, $\alpha \in \Phi$.*

3. *$[x_\alpha, x_{-\alpha}] = h_\alpha$, is a $\mathbb{Z}$-linear combination of the h_i, $1 \le i \le l$.*

4. *Let α and β be non-proportional roots, and $\beta - r\alpha, \ldots, \beta + q\alpha$ the α-string through β. If $q = 0$, then $[x_\alpha, x_\beta] = 0$; whereas if $\alpha + \beta \in \Phi$, then $[x_\alpha, x_\beta] = \pm(r+1)x_{\alpha+\beta}$.*

Proof.

(1): This is clear.

(2): This follows from $\alpha(h_i) = \langle \alpha, \alpha_i \rangle$.

(3): Applying Proposition 2.47, we may consider the dual root system $\Phi^\vee$ with base $\Delta^\vee$. We also make use of the Killing form identification of H^* with H, described in Lemma 1.82. Using this we find that, for each $k \in H$

$$\alpha^\vee(k) = \frac{2\alpha(k)}{(\alpha, \alpha)} = \frac{2}{(\alpha, \alpha)}\kappa(t_\alpha, k) = \kappa\left(\frac{2t_\alpha}{(\alpha, \alpha)}, k\right) = \kappa\left(\frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}, k\right).$$

Furthermore, we claim that $h_\alpha = 2t_\alpha/\kappa(t_\alpha, t_\alpha)$. From Proposition 1.84 and Lemma 1.83 we find that $h_\alpha = [e_\alpha, f_\alpha] = \kappa(e_\alpha, f_\alpha)t_\alpha$. Now, Proposition 1.84 and Lemma 1.82 yield that

$$\begin{aligned} 2 = \alpha(h_\alpha) &= \kappa(t_\alpha, h_\alpha) = \kappa(t_\alpha, [e_\alpha, f_\alpha]) = \kappa(t_\alpha, \kappa(e_\alpha, f_\alpha)t_\alpha) \\ &= \kappa(e_\alpha, f_\alpha)\kappa(t_\alpha, t_\alpha). \end{aligned}$$

Therefore, $\alpha^\vee(k) = \kappa(h_\alpha, k)$, for all $k \in H$. We conclude that $t_{\alpha^\vee} = h_\alpha$. For any $\alpha \in \Phi$ we know that $\alpha^\vee$ is a $\mathbb{Z}$-linear combination of the elements in $\Delta^\vee$. Thus, each h_α is a $\mathbb{Z}$-linear combination of the h_i.

(4): This follows from part (3) of Proposition 2.46 combined with part (2) of Proposition 2.38.

□

Chapter 3

Gradings

The central focus of this chapter is gradings of Lie algebras. The chapter begins with grading on algebras in general. We then narrow our view to the study of gradings on Lie algebras and finally on simple Lie algebras. We will encounter a result which yields the notion of a universal grading group for gradings equivalent to group gradings. At the end of this chapter we investigate automorphisms of gradings.

For the rest of the text we will assume, unless stated otherwise, that all vector spaces are finite-dimensional and that we are working over an algebraically closed field $\mathbb{F}$ of characteristic 0.

3.1 Basics of Gradings

This section starts with an introduction to the basic notions, notation, and terminology around gradings. Thereafter we encounter some results on - and examples of - equivalences of gradings and the connection between reductive Lie algebras and gradings. The section closes out by showing that any indexing group for a simple Lie algebra must be abelian.

Definition 3.1: *Let A be an algebra and I a nonempty indexing set. A vector space decomposition*

$$\Gamma\colon A = \bigoplus_{i\in I} A_i,$$

is called a **grading** *of A if it satisfies the following:*

(i) Each A_i is a non-zero subspace of A, called the **homogeneous component of degree** *i. Its elements are said to be* **homogeneous of degree** *i.*

(ii) For any $i, j \in I$ there is a $k \in I$ such that $A_iA_j \subseteq A_k$. We note that such a k will be unique since Γ is a vector space decomposition.

Suppose we have another grading of A,

$$\Gamma_X\colon A = \bigoplus_{j\in J} X_j.$$

If each X_j consists of a union of some of the A_i's we say that Γ_X is a **coarsening** *of Γ, or equivalently that Γ is a* **refinement** *of Γ_X. We say that Γ is a* **proper refinement** *of Γ_X if at least one of the A_i is a proper subset of one of the X_j. If Γ has no proper refinement we say it is a* **fine grading**. *In the case that A is a Lie algebra we shall call Γ a* **Lie grading** *of A.*

Example 3.2: *Let $\mathbb{F}[x_1, x_2, \ldots, x_n]$ be the polynomial ring over a field $\mathbb{F}$. Write $\mathbb{F}_i$ to denote the subspace of $\mathbb{F}[x_1, x_2, \ldots, x_n]$ consisting of all homogeneous polynomials of degree i. Then:*

$$\Gamma\colon \mathbb{F}[x_1, x_2, \ldots, x_n] = \bigoplus_{i\in\mathbb{N}} \mathbb{F}_i, \tag{3.1}$$

is a grading. In fact, Γ is a vector space decomposition with $\mathbb{F}_i\mathbb{F}_j \subseteq \mathbb{F}_{i+j}$, for any $i, j \in \mathbb{N}$.

Definition 3.3: *Let $\Gamma\colon A = \bigoplus_{g\in G} A_g$ be a vector space decomposition of an algebra A, where G is a semigroup. If the following conditions are satisfied:*

(i) $A_gA_h \subseteq A_{gh}$ *for all $g, h \in G$ and,*

(ii) $G = \langle\{g \in G \mid A_g \neq 0\}\rangle$,

*then this decomposition is called a G-***grading** *of A. The set*

$$\mathrm{Supp}(\Gamma) := \{g \in G \mid A_g \neq 0\},$$

is said to be the **support** *of Γ. Notice that $\mathrm{Supp}(\Gamma)$ generates G.*

Example 3.4: *Suppose A is an associative algebra and $A = \bigoplus_{g\in G} A_g$ a G-grading with G abelian. Then the Lie algebra A^- has a G-grading via the same vector space decomposition.*

$$A^- = \bigoplus_{g\in G} A_g.$$

This follows because for $x \in A_g$ and $y \in A_h$, we have that $xy \in A_{gh}$ and $yx \in A_{hg} = A_{gh}$ and so

$$[x, y] = xy - yx \in A_{gh}.$$

We now present an S_3-grading of a Lie algebra.

Example 3.5: *Let* $\mathfrak{L} := \mathfrak{sl}(2,\mathbb{F}) \bigoplus \mathfrak{sl}(2,\mathbb{F})$. *Let* $\{e_i, f_i, h_i\}$ *denote the standard basis for* $\mathfrak{sl}(2,\mathbb{F})$, *that is* $[e_i, f_i] = h_i$, $[h_i, e_i] = 2e_i$, $[h_i, f_i] = -2f_i$. *Then* $\mathfrak{L}$ *has the following set as a basis* $\{e_1, e_2, f_1, f_2, h_1, h_2\}$.

We denote the symmetric group on 3 elements by

$$S_3 = \{e, \tau, \tau^2, \sigma, \sigma\tau, \sigma\tau^2\},$$

where $\tau^3 = \sigma^2 = e$, $\tau\sigma = \sigma\tau^2$. *We may then define an* S_3*-grading on* $\mathfrak{L}$ *as follows:*

$$\mathfrak{L}_e = \langle h_1, h_2\rangle,\ \mathfrak{L}_\tau = \langle e_1\rangle,\ \mathfrak{L}_{\tau^2} = \langle f_1\rangle,\ \mathfrak{L}_\sigma = \langle e_2, f_2\rangle,\ \mathfrak{L}_{\sigma\tau} = 0,\ \mathfrak{L}_{\sigma\tau^2} = 0.$$

In fact, $[h_i, e_i] = 2e_i$ *and* $[h_i, f_i] = -2f_i$ *yield that*

$$[\mathfrak{L}_e, \mathfrak{L}_\tau] \subseteq \mathfrak{L}_\tau,\ [\mathfrak{L}_e, \mathfrak{L}_{\tau^2}] \subseteq \mathfrak{L}_{\tau^2},\ [\mathfrak{L}_e, \mathfrak{L}_\sigma] \subseteq \mathfrak{L}_\sigma.$$

While $[e_i, f_i] = h_i$ *gives us that*

$$[\mathfrak{L}_\tau, \mathfrak{L}_{\tau^2}] \subseteq \mathfrak{L}_e,\ [\mathfrak{L}_\sigma, \mathfrak{L}_\sigma] \subseteq \mathfrak{L}_e,\ [\mathfrak{L}_{\tau^2}, \mathfrak{L}_\sigma] = \mathfrak{L}_{\tau^2\sigma} = \mathfrak{L}_{\sigma\tau}.$$

These observations - along with the skew-symmetry of the Lie bracket - show us that the decomposition, indexed by S_3, *agrees with the group operation. Moreover, it is clear that* S_3 *is generated by* $\{\tau, \sigma\} \subset \{e, \tau, \tau^2, \sigma\}$.

Definition 3.6: *We say two gradings*

$$A = \bigoplus_{i\in I} X_i = \bigoplus_{j\in J} Y_j,$$

of an algebra A *are* **equivalent** *if there is a bijection* $\tau\colon I \to J$ *and an algebra automorphism* $f\colon A \to A$ *such that* $f(X_i) = Y_{\tau(i)}$ *for all* $i \in I$.

A G*-grading* $\Gamma_G : A = \bigoplus_{g\in G} Y_g$ *and a grading* $\Gamma : A = \bigoplus_{i\in I} X_i$ *of* A *are equivalent if the grading* $A = \bigoplus_{g\in \mathrm{Supp}(\Gamma_G)} Y_g$ *is equivalent to* Γ.

A G*-grading and a* G'*-grading*

$$A = \bigoplus_{g\in G} X_g = \bigoplus_{g'\in G'} Y_{g'},$$

of A *are* **isomorphic** *if there exist a group isomorphism* $\tau\colon G \to G'$ *and an algebra automorphism* $f\colon A \to A$ *such that* $f(X_g) = Y_{\tau(g)}$, *for all* $g \in G$.

The following example will show that semisimple Lie algebras can be graded by non-abelian groups and that an abelian group grading can be equivalent to a non-abelian grading.

Example 3.7: *We let* $\mathfrak{L} = \mathfrak{sl}(2,\mathbb{F}) \oplus \mathfrak{sl}(2,\mathbb{F})$, *as in Example 3.5. We now give a* $\mathbb{Z}_6$*-grading on* $\mathfrak{L}$ *and then see that this grading is equivalent to the grading in that example.*

Note that we may decompose $\mathfrak{L}$ *as*

$$\mathfrak{L} = \mathfrak{M}_0 \oplus \mathfrak{M}_2 \oplus \mathfrak{M}_3 \oplus \mathfrak{M}_4, \tag{3.2}$$

where

$$\mathfrak{M}_0 = \langle h_1, h_2\rangle,\ \ \mathfrak{M}_2 = \langle e_1\rangle,\ \ \mathfrak{M}_3 = \langle e_2, f_2\rangle,\ \ \mathfrak{M}_4 = \langle f_1\rangle.$$

As $[h_i, e_i] = 2e_i$ *and* $[h_i, f_i] = -2f_i$ *we have that*

$$[\mathfrak{M}_0, \mathfrak{M}_2] \subseteq \mathfrak{M}_2,\ \ [\mathfrak{M}_0, \mathfrak{M}_3] \subseteq \mathfrak{M}_3,\ \ [\mathfrak{M}_0, \mathfrak{M}_4] \subseteq \mathfrak{M}_4.$$

Since $[h_i, e_i] = 2e_i$, $[h_i, f_i] = -2f_i$, *and* $[e_i, f_i] = h_i$ *we find that*

$$[\mathfrak{M}_2, \mathfrak{M}_4] \subseteq \mathfrak{M}_0,\ \ [\mathfrak{M}_3, \mathfrak{M}_3] \subseteq \mathfrak{M}_0,\ \ [\mathfrak{M}_2, \mathfrak{M}_4] \subseteq \mathfrak{M}_0.$$

Furthermore, from the skew-symmetry of the Lie bracket, the fact that $\mathbb{Z}_6$ *is abelian, and that* $\mathbb{Z}_6 = \langle 2, 3\rangle = \langle 0, 2, 3, 4\rangle$ *we conclude that* (3.2) *is a* $\mathbb{Z}_6$*-grading of* $\mathfrak{L}$.

We saw in Example 3.5 that $\{\epsilon, \sigma\} \subset \{\epsilon, \tau, \tau^2, \sigma\}$ *generates* S_3. *We may therefore define an bijection* $\delta\colon \{\epsilon, \tau, \tau^2, \sigma\} \to \{0, 2, 3, 4\}$ *by*

$$\epsilon \mapsto 0, \tau \mapsto 2, \tau^2 \mapsto 4, \sigma \mapsto 3.$$

Taking the identity automorphism on $\mathfrak{L}$ *we see that the* S_3*-grading of* $\mathfrak{L}$ *from Example 3.5 is equivalent to the* $\mathbb{Z}_6$*-grading in this example.*

Definition 3.8: *We say a Lie algebra* $\mathfrak{L}$ *is* **reductive** *if its adjoint representation is completely reducible.*

Remark 3.9: *Any semisimple Lie algebra* $\mathfrak{L}$ *is reductive.*

Proof. By Weyl's Theorem D.21 we see that the adjoint representation of $\mathfrak{L}$ is completely reducible. □

The above remark shows that being reductive is a generalisation of semisimplicity.

Definition 3.10: *We say* $\mathfrak{L}$ *has a* **reductive decomposition** *if we may express it as a direct sum of Lie algebras* $\mathfrak{L} = \mathfrak{K} \oplus \mathfrak{M}$, *where* $\mathfrak{K}$ *is a subalgebra of* $\mathfrak{L}$ *and* $\mathfrak{M}$ *is a* $\mathfrak{K}$*-module under the adjoint representation.*

Example 3.11: *We saw in Example 1.18 that* $\mathfrak{gl}(n, \mathbb{F})/\mathfrak{sl}(n, \mathbb{F}) \cong \mathbb{F}$. *We note that* $\mathbb{F}$ *is isomorphic to* $\mathfrak{s}(n, \mathbb{F})$, *that is, the subalgebra of* $\mathfrak{d}(n, \mathbb{F})$ *consisting of scalar multiples of the identity matrix. Thus,* $\mathfrak{gl}(n, \mathbb{F}) = \mathfrak{sl}(n, \mathbb{F}) \oplus \mathfrak{s}(n, \mathbb{F})$. *We have seen that* $\mathfrak{sl}(n, \mathbb{F})$ *is a subalgebra of* $\mathfrak{gl}(n, \mathbb{F})$, *and that* $\mathfrak{s}(n, \mathbb{F})$ *is an* $\mathfrak{sl}(n, \mathbb{F})$*-module (*$\mathfrak{sl}(n, \mathbb{F})$ *acts trivially on* $\mathfrak{s}(n, \mathbb{F})$*). Hence, this is a reductive decomposition of* $\mathfrak{gl}(n, \mathbb{F})$.

Definition 3.12: *Let $\mathfrak{L}$ be a Lie algebra, $\mathfrak{M}$ an $\mathfrak{L}$-module and $\mathfrak{N}$ a non-empty subset of $\mathfrak{M}$. We define the* **annihilator** *of $\mathfrak{N}$ in $\mathfrak{L}$ as*

$$Ann_{\mathfrak{L}}(\mathfrak{N}) = \{l \in \mathfrak{L} \,|\, l \cdot n = 0 \quad \forall n \in \mathfrak{N}\}.$$

Proposition 3.13: *Let $\mathfrak{L}$ be a Lie algebra, $\mathfrak{M}$ an $\mathfrak{L}$-module and $\mathfrak{N}$ a non-empty subset of $\mathfrak{M}$.*

(i) The annihilator of $\mathfrak{N}$ in $\mathfrak{L}$, $Ann_{\mathfrak{L}}(\mathfrak{N})$, is a subalgebra of $\mathfrak{L}$.

(ii) If $\mathfrak{N}$ is a submodule of $\mathfrak{M}$ then $Ann_{\mathfrak{L}}(\mathfrak{N})$ is an ideal of $\mathfrak{L}$.

Proof. We denote by $A := Ann_{\mathfrak{L}}(\mathfrak{N})$. For $x, y \in A$, $n \in \mathfrak{N}$, and $\alpha, \beta \in \mathbb{F}$ we have that $(\alpha x + \beta y) \cdot n = \alpha x \cdot n + \beta y \cdot n = 0$, and that $0 \in A$. Thus, A is a vector subspace of $\mathfrak{L}$.

Moreover, for $x, y \in A$ and $n \in \mathfrak{N}$ we have that

$$[x, y] \cdot n = x \cdot (y \cdot n) - y \cdot (x \cdot n) = x \cdot 0 + y \cdot 0 = 0,$$

which proves (i).

If $\mathfrak{N}$ is a submodule of $\mathfrak{M}$ then, for $x \in A$, $y \in \mathfrak{L}$, $n \in \mathfrak{N}$, we have that $y \cdot n \in \mathfrak{N}$, and so

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

which proves (ii). □

Lemma 3.14: *Let $\mathfrak{L}$ be a Lie algebra graded by a group G.*

(i) If $\mathfrak{L} = \mathfrak{K} \oplus \mathfrak{M}$ is a reductive decomposition, then $Ann_{\mathfrak{K}}(\mathfrak{M})$ is an ideal of $\mathfrak{L}$. Hence, if $\mathfrak{L}$ is simple, then $Ann_{\mathfrak{K}}(\mathfrak{M}) = 0$ if $\mathfrak{M}$ is not the trivial module.

(ii) Suppose G' is a subgroup of G. Let $\mathfrak{K}$ denote the direct sum of the distinct homogeneous components of degree $g' \in G'$, and let $\mathfrak{M}$ be the direct sum of the remaining homogeneous components. If $\mathfrak{L}$ is simple, then $\mathfrak{L} = \mathfrak{K} \oplus \mathfrak{M}$ is a reductive decomposition.

Proof.

(i): We will denote by $A := Ann_{\mathfrak{K}}(\mathfrak{M})$. Now, for any $l \in \mathfrak{L}$ we may write $l = k + m$ for some $k \in \mathfrak{K}$, $m \in \mathfrak{M}$. Therefore, for $l \in \mathfrak{L}$ and $a \in A$, we have

$$[l, a] \cdot \mathfrak{M} = [k + m, a] \cdot \mathfrak{M} = [k, a] \cdot \mathfrak{M} + [m, a] \cdot \mathfrak{M} = [k, a] \cdot \mathfrak{M} \in [\mathfrak{K}, \mathfrak{K}] \subseteq \mathfrak{K},$$

Furthermore, for $m' \in \mathfrak{M}$ we have $[[l, a], m] = [[l, m], a] + [l, [a, m]]$ but $[a, m] = 0$. Therefore

$$[[l, a], m'] = -[a, [l, m']] = -[a, [k, m']] - [a, [m, m']].$$

But $[k, m'] \in \mathfrak{M}$ and so $-[a, [k, m']] = 0$. Noting that

$$[a, [m, m']] = [[a, m], m'] + [m, [a, m']] = 0 + 0,$$

we see that $[[l, a], m] = 0$ and (i) follows.

(ii): If we have a G-grading $\mathfrak{L} = \bigoplus_{g \in G} \mathfrak{L}_g$ of $\mathfrak{L}$, we define $\mathfrak{K} = \bigoplus_{g \in G'} \mathfrak{L}_g$ and $\mathfrak{M} = \bigoplus_{g \in G - G'} \mathfrak{L}_g$. Clearly, $\mathfrak{L} = \mathfrak{K} \oplus \mathfrak{M}$. Moreover, we claim this is a reductive decomposition.

Since G' is a group, we have that $gh \in G'$ for all $g, h \in G'$. Therefore,

$$[\mathfrak{K}, \mathfrak{K}] = \left[\bigoplus_{g \in G'} \mathfrak{L}_g, \bigoplus_{h \in G'} \mathfrak{L}_h \right] = \bigoplus_{g,h \in G'} [\mathfrak{L}_g, \mathfrak{L}_h] \subseteq \bigoplus_{g,h \in G'} \mathfrak{L}_{gh},$$

and so it follows $\mathfrak{K}$ is a subalgebra of $\mathfrak{L}$. To see that $\mathfrak{M}$ is a $\mathfrak{K}$-module, it will suffice to show that $[\mathfrak{K}, \mathfrak{M}] \subseteq \mathfrak{M}$. To do so we take $g \in G'$ and $h \in G - G'$ and show that $[\mathfrak{L}_g, \mathfrak{L}_h] \subseteq \mathfrak{L}_{gh} \subseteq \mathfrak{M}$. We have that $gh \in G'$ or $gh \in G - G'$. In the latter case we are done. In the former case we would have that $g^{-1}(gh) = h \in G'$, which is a contradiction. Therefore, $[\mathfrak{L}_g, \mathfrak{L}_h] \subseteq \mathfrak{L}_{gh} \subseteq \mathfrak{M}$.

□

Proposition 3.15: *For any group grading* $\Gamma \colon \mathfrak{L} = \bigoplus_{g \in G} \mathfrak{L}_g$ *of a simple Lie algebra* $\mathfrak{L}$*, the group* G *is abelian.*

Proof. We know that the support of Γ generates G. For every $g_1 \in \mathrm{Supp}(\Gamma)$ we may consider C_{g_1} the centralizer of g_1 in G. If we have $C_{g_1} = G$ for every $g_1 \in \mathrm{Supp}(\Gamma)$, then we have a generating set for G whose elements commute with all of G, which would make G abelian.

We suppose then that there is a $g_1 \in \mathrm{Supp}(\Gamma)$ such that C_{g_1} is a proper subgroup of G. An application of Lemma 3.14 yields a reductive decomposition $\mathfrak{L} = \mathfrak{K} \oplus \mathfrak{M}$ for $\mathfrak{L}$, where $\mathfrak{K} = \bigoplus_{g \in C_{g_1}} \mathfrak{L}_g$ and $\mathfrak{M} = \bigoplus_{g \in G - C_{g_1}} \mathfrak{L}_g$. Again using Lemma 3.14 we find that $Ann_{\mathfrak{K}}(\mathfrak{M}) = 0$. We note that for non-commuting elements $h_1, h_2 \in G$ we must have that $[\mathfrak{L}_{h_1}, \mathfrak{L}_{h_2}] = 0$. If not, then skew-symmetry yields that $[\mathfrak{L}_{h_1}, \mathfrak{L}_{h_2}] = [\mathfrak{L}_{h_2}, \mathfrak{L}_{h_1}]$. From which follows $0 \neq [\mathfrak{L}_{h_1}, \mathfrak{L}_{h_2}] \subseteq \mathfrak{L}_{h_1 h_2} \cap \mathfrak{L}_{h_2 h_1} = 0$, a contradiction.

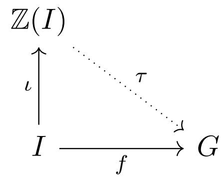
From there we find that $[\mathfrak{L}_{g_1}, \mathfrak{M}] = 0$, since g_1 does not commute with any element of $G - C_{g_1} \neq \emptyset$. This means that $\mathfrak{L}_{g_1} \subseteq Ann_{\mathfrak{K}}(\mathfrak{M}) = 0$, which contradicts the fact that $g_1 \in \mathrm{Supp}(\Gamma)$. Thus, $C_{g_1} = G$ for all $g_1 \in \mathrm{Supp}(\Gamma)$, and so G is abelian. □

3.2 Universal Grading Group

In the previous section we saw that different groups may produce equivalent gradings. We would like to have a procedure which would allow us to produce a unique group among all those which are equivalent. We start this section by describing a procedure from which one can find a universal grading group. The rest of the section is used to provide an example this procedure.

From here onwards we use the convention that all gradings will be equivalent to group gradings.

Remark 3.16: *Let I be a finite set. We denote by $\mathbb{Z}(I)$ the free $\mathbb{Z}$-module generated by I, which is the free abelian group generated by I. $\mathbb{Z}(I)$ has the following universal property. For any function $f\colon I \to G$, where G is a $\mathbb{Z}$-module (an abelian group), there is a unique $\mathbb{Z}$-module homomorphism $\tau\colon \mathbb{Z}(I) \to G$ such that the following diagram, where $\iota\colon I \to \mathbb{Z}(I)$ denotes the canonical injection, commutes.*



Proposition 3.17: *Let $\Gamma\colon \mathfrak{L} = \bigoplus_{i\in I} \mathfrak{L}_I$ be a grading of a simple Lie algebra $\mathfrak{L}$. There is a finitely generated abelian group G_I which satisfies the following properties:*

(i) $I \subseteq G_I$.

(ii) $\mathfrak{L} = \bigoplus\limits_{i\in I} \mathfrak{L}_i = \bigoplus\limits_{j\in G_I} \mathfrak{L}'_j$, *where* $\mathfrak{L}'_j = \begin{cases} \mathfrak{L}_j, & \text{for all } j \in I \\ 0, & \text{for all } j \in G_I - I \end{cases}$.

(iii) For any group grading $\Gamma^*\colon \mathfrak{L} = \bigoplus_{g\in G} \mathfrak{M}_g$, which is a coarsening of the G_I-grading on $\mathfrak{L}$, there is a unique group epimorphism $\tau\colon G_I \to G$ such that $\mathfrak{M}_g = \bigoplus\limits_{j\in J} \mathfrak{L}'_j$ for any $g \in G$, where $J := \{j \in G_I \,|\, \tau(j) = g\}$.

Proof. We will let $\mathbb{Z}(I)$ denote the free $\mathbb{Z}$-module generated by I and define $N := \{i_1 + i_2 - i_3 \,|\, i_j \in I \text{ and } 0 \neq [\mathfrak{L}_{i_1}, \mathfrak{L}_{i_2}] \subseteq \mathfrak{L}_{i_3}\}$. We then define $M = \langle N\rangle$, that is, the $\mathbb{Z}$-submodule of $\mathbb{Z}(I)$ generated by N. We then claim that the $\mathbb{Z}$-module $G_I = \mathbb{Z}(I)/M$ has the properties described in the hypothesis.

For any $i \in I$, we have that $i \notin N$. This means that $i \notin M$. If $i, j \in N$ such that $i - j \in M$. then $0 \neq [\mathfrak{L}_i, \mathfrak{L}_0] \subseteq \mathfrak{L}_j$. Thus, $\mathfrak{L}_i \subseteq \mathfrak{L}_j$ and so $i = j$. Hence we may conclude that $I \subseteq G_I$. Since $\mathfrak{L}$ is finite-dimensional we find that I is finite, and so G_I is finitely generated. Moreover, as a $\mathbb{Z}$-module G_I is an abelian group. Parts (i) and (ii) thus follow.

Suppose G is a finitely generated abelian group and $\Gamma_G\colon \mathfrak{L} = \bigoplus_{g\in G}\mathfrak{M}_g$ is a coarsening of the G_I-grading Γ. We denote by $S := \mathrm{Supp}(\Gamma_G)$. Then there is a surjective map $f\colon I \to S$, with $\mathfrak{L}'_i \subseteq \mathfrak{M}_{f(i)}$ for each $i \in I$. In fact, for each $g \in S$ we may write $\mathfrak{M}_g = \bigoplus_{i=1}^{n_g}\mathfrak{L}_{g_i}$ and then we define $f(g_i) = g$ for each $i \in \{1, 2, \ldots, n_g\}$. We may also view f as having codomain G, and so by Remark 3.16 there is a unique $\mathbb{Z}$-module homomorphism $\varphi\colon \mathbb{Z}(I) \to G$ such that $\varphi \circ j = f$, where $j\colon I \to \mathbb{Z}(I)$ is the natural injection, $i \mapsto i$ for each $i \in I$. This gives us that $\varphi(i) = f(i)$ for all $i \in I$.

We note that φ induces a unique group homomorphism $\widetilde{\varphi}\colon \mathbb{Z}(I) \to G$. Moreover, if $\varphi\colon \mathbb{Z}(I) \to G$ is a group homomorphism such that $\tau \circ j = f$, then for $z \in \mathbb{Z}$ and $i \in I$ we have $\tau(zi) = \tau(i + i + \ldots + i) = z\tau(i)$. By uniqueness we have that $\tau = \varphi = \widetilde{\varphi}$.

We claim that $M \subseteq \mathrm{Ker}(\varphi)$. It suffices to show that $N \subseteq \mathrm{Ker}(\varphi)$. For $i_1+i_2-i_3 \in N$ we have that $0 \neq [\mathfrak{L}_{i_1}, \mathfrak{L}_{i_2}] \subseteq \mathfrak{L}_{i_3}$, and so $0 \neq [\mathfrak{L}'_{i_1}, \mathfrak{L}'_{i_2}] \subseteq \mathfrak{L}'_{i_3}$. Hence, $0 \neq [\mathfrak{L}_{i_1}, \mathfrak{L}_{i_2}] = [\mathfrak{L}'_{i_1}, \mathfrak{L}'_{i_2}] \subseteq \mathfrak{L}'_{i_1+i_2}$, but this means that $i_1 + i_2 \in I$, as otherwise $\mathfrak{L}'_{i_1+i_2} = 0$. Therefore,

$$\varphi(i_1 + i_2 - i_3) = \varphi(i_1 + i_2) - \varphi(i_3) = f(i_1 + i_2) - f(i_3).$$

Since $[\mathfrak{L}'_{i_1}, \mathfrak{L}'_{i_2}] \subseteq \mathfrak{L}'_{i_3}$ we have that $\mathfrak{L}'_{i_1+i_2} \subseteq \mathfrak{L}'_{i_3}$ and so $f(i_1+i_2) = f(i_3)$. It follows that $\varphi(i_1+i_2-i_3) = 0$, and since $M = \langle N\rangle$ we have that $M \subseteq \mathrm{Ker}(\varphi)$.

We now find that φ induces a group homomorphism $\overline{\varphi}\colon G_I \to G$ defined by $\sum\limits_{i\in I} z_i i + M \mapsto \sum\limits_{i\in I} z_i\varphi(i)$. Then $\overline{\varphi}(i + M) = f(i)$ for each $i \in I$. As $G = \langle S\rangle$ and f is surjective, we find that $\overline{\varphi}$ is an epimorphism. Then $\tau = \overline{\varphi}$ and uniqueness follows since φ is unique. □

Before we give an example of finding the universal grading group of a grading, we will need some definitions and results.

Remark 3.18: *Let $M = \langle m_1, \ldots, m_s\rangle$ be a finitely generated module of a principal ideal domain R. Then we may define an R-module epimorphism*

$$\theta\colon R^s \to M$$
$$(r_1, \ldots, r_s) \mapsto \sum_{i=1}^{s} r_i m_i.$$

Conversely, if there is such an R-module epimorphism then M is finitely generated.

If K is the kernel of θ, then $M \cong R^s/K$. Moreover, for $(r_1, \ldots, r_s) \in K$ we find that $\sum_{i=1}^{s} r_i m_i = 0$. So every element of K defines a relation on the generators of M. We therefore call K the **relation submodule** *of R^s for M relative to $\{m_1, \ldots, m_s\}$. If $K = \langle k_1, \ldots, k_t\rangle$ is finitely generated, we call the matrix whose rows consist of the generators of K the* **relation matrix** *of R^s for M relative to $\{m_1, \ldots, m_s\}$ and $\{k_1, \ldots, k_t\}$.*

Lemma 3.19: *Let $M = \langle m_1, \ldots, m_s \rangle$ be a finitely generated module of a principal ideal domain R. If $K = \langle k_1, \ldots, k_t \rangle$ is the corresponding relation module for M and we let A denote the relation matrix for M, that is, if each $k_i = (k_{i1}, \ldots, k_{is})$, then*

$$A = \begin{pmatrix} k_{11} & k_{12} & \ldots & k_{1s} \\ k_{21} & k_{22} & \ldots & k_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ k_{t1} & k_{t2} & \ldots & k_{ts} \end{pmatrix}.$$

(i) If P is an invertible $t \times t$ matrix with entries in R, then the rows of PA generate K. So PA is a relation matrix for M relative to $\{m_1, \ldots, m_s\}$ and the rows of PA (the rows are taken as elements of R^s).

(ii) Let Q be a $s \times s$ invertible matrix with entries in R. For each $1 \leq i \leq s$ we define $m'_i := \sum_{j=1}^{s} (Q^{-1})_{ji} m_j$. Then $M = \langle m'_1, \ldots, m'_s \rangle$ and AQ is a relation matrix for M relative to $\{m'_1, \ldots, m'_s\}$.

(iii) Let P and Q be $t \times t$ and $s \times s$ invertible matrices, respectively, with entries in R. For each $1 \leq i \leq s$ we define $m''_i := \sum_j ((AQ)^{-1})_{ji} m_j$. Then PAQ is the relation matrix relative to $\{m''_1, \ldots, m''_s\}$ and the rows of PAQ (the rows are taken as elements of R^s).

Proof.

(i): We find that

$$PA = \begin{pmatrix} \sum_{i=1}^{t} (P)_{1i} k_{i1} & \sum_{i=1}^{t} (P)_{1i} k_{i2} & \ldots & \sum_{i=1}^{t} (P)_{1i} k_{is} \\ \sum_{i=1}^{t} (P)_{2i} k_{i1} & \sum_{i=1}^{t} (P)_{2i} k_{i2} & \ldots & \sum_{i=1}^{t} (P)_{2i} k_{is} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{t} (P)_{ti} k_{i1} & \sum_{i=1}^{l} (P)_{ti} k_{i2} & \ldots & \sum_{i=1}^{t} (P)_{ti} k_{is} \end{pmatrix}.$$

From this we can see that for each $1 \leq l \leq t$ the l-th row of PA is

$$\begin{aligned}
r_l &= \sum_{i=1}^{t} (P)_{li} k_{i1} + \sum_{i=1}^{t} (P)_{li} k_{i2} + \ldots + \sum_{i=1}^{t} (P)_{li} k_{is} \\
&= \sum_{j=1}^{s} \sum_{i=1}^{t} (P)_{li} k_{ij} \\
&= (P)_{l1} \sum_{j=1}^{s} k_{1j} + (P)_{l2} \sum_{j=1}^{s} k_{2j} + \ldots + (P)_{lt} \sum_{j=1}^{s} k_{tj} \\
&= (P)_{l1} k_1 + (P)_{l2} k_2 + \ldots + (P)_{lt} k_t \\
&= \sum_{i=1}^{t} (P)_{li} k_i.
\end{aligned}$$

As the k_i are generators for K, we have that the r_i are elements of K. Moreover,

$$\begin{aligned}
\sum_{i=1}^{t} (P^{-1})_{li} r_i &= \sum_{i=1}^{t} (P^{-1})_{li} \left(\sum_{j=1}^{t} (P)_{ij} k_j \right) \\
&= \sum_{i,j=1}^{t} (P^{-1})_{li} (P)_{ij} k_j \\
&= (P^{-1})_{l1} (P)_{11} k_1 + \ldots + (P^{-1})_{l1} (P)_{1t} k_t \\
&+ (P^{-1})_{l2} (P)_{21} k_1 + \ldots + (P^{-1})_{l2} (P)_{2t} k_t \\
&+ \ldots \\
&+ (P^{-1})_{lt} (P)_{t1} k_1 + \ldots + (P^{-1})_{lt} (P)_{tt} k_t \\
&= \left(\sum_{j=1}^{t} (P^{-1})_{lj} (P)_{j1} \right) k_1 \\
&+ \left(\sum_{j=1}^{t} (P^{-1})_{lj} (P)_{j2} \right) k_2 \\
&+ \ldots \\
&+ \left(\sum_{j=1}^{t} (P^{-1})_{lj} (P)_{jt} \right) k_t \\
&= \sum_{i=1}^{t} (I_t)_{li} k_i \\
&= k_l.
\end{aligned}$$

Therefore we may recover the k_i from linear combinations of the r_j. This means that the r_j also generate K and so PA is a relation matrix for M relative to $\{m_1, \ldots, m_s\}$ and $\{r_1, \ldots, r_t\}$.

(ii): The m'_i are linear combinations of the m_j and so it immediately follows that $M = \langle m'_1, \ldots, m'_s \rangle$. From $m'_j = \sum_{i=1}^{s} (Q^{-1})_{ij} m_i$ and a calculation similar to one in (i) we see that $m_i = \sum_{j=1}^{s} (Q)_{ij} m'_j$ for each $1 \leq i \leq s$. Since A is a relation matrix for M relative to $\{m_1, \ldots, m_s\}$ we have that

$$A \begin{pmatrix} m_1 \\ \vdots \\ m_s \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Therefore,

$$(AQ)Q^{-1} \begin{pmatrix} m_1 \\ \vdots \\ m_s \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

and so from $m'_j = \sum\limits_{i=1}^{s} (Q^{-1})_{ij} m_i$ we find that

$$AQ \begin{pmatrix} m'_1 \\ \vdots \\ m'_s \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

This shows that AQ is a relation matrix for M relative to the set $\{m'_1, \ldots, m'_s\}$.

(iii): This follows directly from (i) and (ii).

□

The above lemma shows us that elementary row operations performed on a relation matrix A are akin to changing generators for the relation module K, while elementary column operations correspond to changing generators for the module M.

Example 3.20: *Let $M = \langle m_1, m_2 \rangle$ be a $\mathbb{Z}$-module with relations*

$$m_1 + 4m_2 = 0 \quad \textit{and} \quad -3m_1 + 5m_2 = 0.$$

This gives the relation matrix $A = \begin{pmatrix} 1 & 4 \\ -3 & 5 \end{pmatrix}$. Adding 3 times row 1 to row 2 then switches our generators of K from $\begin{pmatrix} 1 \\ 4 \end{pmatrix}, \begin{pmatrix} -3 \\ 5 \end{pmatrix}$ to $\begin{pmatrix} 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ 17 \end{pmatrix}$. Subtracting 4 times column 1 from column 2 leaves us with the matrix

$$A' = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ -3 & 5 \end{pmatrix} \begin{pmatrix} 1 & -4 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 17 \end{pmatrix}$$

and using our algorithm from the lemma we see that our generators for M switch from $\{m_1, m_2\}$ to $\{m_1, 4m_1 + m_2\}$. Now, A' is equivalent to A and in Smith normal form. Using the new generators for K we see that $K = \langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 17 \end{pmatrix} \rangle = \mathbb{Z} \oplus 17\mathbb{Z}$*. Since $M \cong \mathbb{Z}^2/K$ from Remark 3.18 we see*

$$M \cong (\mathbb{Z}/\mathbb{Z}) \oplus (\mathbb{Z}/17\mathbb{Z}) \cong \mathbb{Z}_{17}.$$

Proposition 3.21: *Let $M = \langle m_1, \ldots, m_s \rangle$ be a module of a principal ideal domain R, and A a relation matrix for M. If there are invertible matrices P and Q such that PAQ is diagonal with non-zero diagonal entries $d_1, \ldots, d_p$, then*

$$M \cong R/\langle d_1 \rangle \oplus \ldots \oplus R/\langle d_p \rangle \oplus R^{s-p}.$$

Proof. We first note that by Lemma 3.19 we have that PAQ is the relation matrix for M relative to the corresponding generators of M and the rows of PAQ. As in Remark 3.18 we have a corresponding R-module homomorphism $\theta \colon R^s \to M$ with $\mathrm{Ker}(\theta) = K$ generated by the rows of A. Then we have that

$$\begin{aligned} M \cong R^s/K = R^s/\langle \begin{pmatrix} d_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ d_p \end{pmatrix} \rangle \\ \cong R/\langle d_1 \rangle \oplus \ldots \oplus R/\langle d_p \rangle \oplus R/\langle 0 \rangle \oplus \ldots \oplus R/\langle 0 \rangle \\ \cong R/\langle d_1 \rangle \oplus \ldots \oplus R/\langle d_p \rangle \oplus R^{s-p}. \end{aligned}$$

□

Example 3.22: *Let $\mathfrak{L} = \mathfrak{sl}(3, \mathbb{F})$ with basis*

$$B := \{e_{12}, e_{13}, e_{21}, e_{23}, e_{31}, e_{32}, h_1 := e_{11} - e_{22}, h_2 := e_{22} - e_{33}\}.$$

Then,

$$\mathfrak{L} = \mathfrak{L}_i \oplus \mathfrak{L}_j \oplus \mathfrak{L}_k \oplus \mathfrak{L}_m \oplus \mathfrak{L}_n,$$

where

$$\mathfrak{L}_i = \langle h_1, h_2 \rangle,\ \mathfrak{L}_j = \langle e_{21}, e_{32} \rangle,\ \mathfrak{L}_k = \langle e_{31} \rangle,\ \mathfrak{L}_m = \langle e_{12}, e_{23} \rangle,\ \mathfrak{L}_n = \langle e_{13} \rangle,$$

is a grading on $\mathfrak{L}$. We note that $\mathfrak{L}_i$ is a maximal toral subalgebra of $\mathfrak{L}$.

We have that $[h_1, h_2] = 0$ so $[\mathfrak{L}_i, \mathfrak{L}_i] = 0$, whereas $[h_1, e_{21}] = -2e_{21}$, $[h_1, e_{32}] = -e_{32}$, $[h_2, e_{21}] = e_{21}$, and $[h_2, e_{32}] = -2e_{32}$ mean $[\mathfrak{L}_i, \mathfrak{L}_j] \subseteq \mathfrak{L}_j$. We find that $[\mathfrak{L}_i, \mathfrak{L}_k] \subseteq \mathfrak{L}_k$ because $[h_1, e_{31}] = -e_{31} = [h_2, e_{31}]$. Now, from $[h_1, e_{12}] = 2e_{12}$, $[h_1, e_{23}] = -e_{23}$. $[h_2, e_{12}] = -e_{12}$, and $[h_2, e_{23}] = 2e_{23}$ we get that $[\mathfrak{L}_i, \mathfrak{L}_m] \subseteq \mathfrak{L}_m$. Since $[h_1, e_{13}] = e_{13} = [h_2, e_{13}]$ we must have that

$[\mathfrak{L}_i, \mathfrak{L}_n] \subseteq \mathfrak{L}_n$. *Noting that* $[e_{21}, e_{13}] = e_{23}$ *and* $[e_{32}, e_{13}] = -e_{12}$ *we find that* $[\mathfrak{L}_j, \mathfrak{L}_n] \subseteq \mathfrak{L}_m$.

Seeing that $[e_{21}, e_{32}] = -e_{31}$ *gives that* $[\mathfrak{L}_j, \mathfrak{L}_j] \subseteq \mathfrak{L}_k$. *Then, since* $[e_{21}, e_{31}] = 0 = [e_{32}, e_{31}]$ *we have that* $[\mathfrak{L}_j, \mathfrak{L}_k] = 0$. *We can also see that* $[e_{21}, e_{12}] = -h_1$, $[e_{21}, e_{23}] = 0 = [e_{32}, e_{12}]$, *and* $[e_{32}, e_{23}] - h_2$ *and therefore* $[\mathfrak{L}_j, \mathfrak{L}_m] \subseteq \mathfrak{L}_i$.

We have that $[e_{31}, e_{12}] = 0$ *and* $[e_{31}, e_{23}] = -e_{21}$ *which shows that* $[\mathfrak{L}_k, \mathfrak{L}_m] \subseteq \mathfrak{L}_j$. *We can see that* $[e_{31}.e_{13}] = -(h_1 - h_2)$ *and so* $[\mathfrak{L}_k, \mathfrak{L}_n] \subseteq \mathfrak{L}_i$.

Since $[e_{12}, e_{23}] = e_{13}$ *we see that* $[\mathfrak{L}_m, \mathfrak{L}_m] \subseteq \mathfrak{L}_n$. *Finally, seeing that* $[e_{12}, e_{13}] = 0 = [e_{23}, e_{13}]$ *we conclude that* $[\mathfrak{L}_m, \mathfrak{L}_n] = 0$.

From our conclusions above, and from the properties of the Lie bracket we may summarize the Lie bracket relations on the gradings subspaces in the following table:

$[\,,\,]$	$\mathfrak{L}_i$	$\mathfrak{L}_j$	$\mathfrak{L}_k$	$\mathfrak{L}_m$	$\mathfrak{L}_n$
$\mathfrak{L}_i$	0	$\mathfrak{L}_j$	$\mathfrak{L}_k$	$\mathfrak{L}_m$	$\mathfrak{L}_n$
$\mathfrak{L}_j$		$\mathfrak{L}_k$	0	$\mathfrak{L}_i$	$\mathfrak{L}_m$
$\mathfrak{L}_k$			0	$\mathfrak{L}_j$	$\mathfrak{L}_i$
$\mathfrak{L}_m$				$\mathfrak{L}_n$	0
$\mathfrak{L}_n$					0

We need only consider the section above the main diagonal because the Lie bracket is skew-symmetric.

To continue we would like to find the universal grading group G_I *for the set* $I = \{i, j, k, l, m, n\}$. *To do so we follow the process described in the proof of Proposition 3.17. We must therefore consider the submodule* M *of the free* $\mathbb{Z}$*-module* $\mathbb{Z}(I)$, *where* $M = \langle N \rangle$ *and*

$$N := \{x + y - z \,|\, x, y, z \in I \text{ and } 0 \neq [\mathfrak{L}_x, \mathfrak{L}_y] \subseteq \mathfrak{L}_z\}.$$

By consulting the table of Lie bracket relations we find that

$$N = \{i + j - j, 2j - k, j + m - i, j + n - m, k + m - j, k + n - i, 2m - n\}$$

Next, we must find $\mathbb{Z}(I)/M$. *First, we note that we may identify* $\mathbb{Z}(I)$ *with* $\mathbb{Z}^5$ *by mapping*

$$i \mapsto \begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}, \; j \mapsto \begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix}, \; k \mapsto \begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix}, \; m \mapsto \begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix}, \; n \mapsto \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix},$$

and $M = \langle i, 2j - k, j + m - i, j + n - m, k + m - j, k + n - i, 2m - n\rangle$ *we* *may identify with*

$$M' = \langle \begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\2\\-1\\0\\0\end{pmatrix}, \begin{pmatrix}-1\\1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\-1\\1\end{pmatrix}, \begin{pmatrix}0\\-1\\1\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\0\\1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\0\\0\\2\\-1\end{pmatrix} \rangle.$$

Therefore we conclude that $G_I = \mathbb{Z}(I)/M \cong \mathbb{Z}^5/M'$. *We now represent the generators of* M' *as rows in a matrix*

$$A' = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & -1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 1 \\ 0 & -1 & 1 & 1 & 0 \\ -1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 & -1 \end{pmatrix}.$$

Elementary row operations correspond to multiplying A' *on the left by invertible* 7×7 *matrices, and elementary column operations correspond to multiplying* A' *on the right by invertible* 5×5 *matrices. We denote elementary row operations as follows:* $R_i \mapsto R_i + kR_j$ *indicates replacement of the* i*-th row by the* i*-th row added to* k *times the* j*-th row (where* k *is a non-zero constant), and* $R_i \leftrightarrow R_j$ *means that we swap the* i*-th and* j*-th rows. We denote elementary column operations similarly, replacing* R *by* C. *We perform the following elementary row and column operations (ordered left to right and top to bottom):*

$R_3 \mapsto R_3 + R_1$	$R_5 \mapsto R_5 + R_2$	$R_4 \mapsto -\frac{1}{2}R_4$	$R_4 \mapsto R_4 + R_2$	$R_4 \mapsto 2R_4$
$R_6 \mapsto R_6 + R_1$	$R_3 \mapsto 2R_3$	$R_6 \mapsto R_6 + 2R_4$	$C_5 \mapsto C_5 - C_3$	$C_4 \leftrightarrow C_5$
$R_2 \mapsto \frac{1}{2}R_2$	$R_4 \mapsto R_4 - \frac{1}{2}R_3$	$R_7 \mapsto R_7 - 2R_4$	$C_2 \mapsto C_2 - 2C_5$	
$R_3 \mapsto R_3 - R_2$	$R_5 \mapsto R_5 - \frac{1}{2}R_3$	$R_3 \mapsto R_3 - 2R_4$	$R_2 \leftrightarrow R_4$	
$R_4 \mapsto R_4 - R_2$	$R_6 \mapsto R_6 - R_3$	$R_2 \mapsto R_2 + \frac{1}{2}R_3$	$C_4 \mapsto C_4 - C_2$	

and get a row equivalent matrix

$$PA'Q = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

where P *and* Q *are invertible matrices. From Proposition 3.21 we may conclude that* $M' \cong \mathbb{Z}^4$. *Therefore* $G_I = \mathbb{Z}(I)/M \cong \mathbb{Z}^5/\mathbb{Z}^4 \cong \mathbb{Z}$. *In fact we have that the grading of* $\mathfrak{L}$ *with indexing set* I *is a* $\mathbb{Z}$*-grading for* $\mathfrak{L}$

$$\mathfrak{L} = \mathfrak{L}_{-2} \oplus \mathfrak{L}_{-1} \oplus \mathfrak{L}_0 \oplus \mathfrak{L}_1 \oplus \mathfrak{L}_2,$$

for

$$\mathfrak{L}_{-2} = \mathfrak{L}_n, \ \ \mathfrak{L}_{-1} = \mathfrak{L}_m, \ \ \mathfrak{L}_0 = \mathfrak{L}_i, \ \ \mathfrak{L}_1 = \mathfrak{L}_j, \ \ \mathfrak{L}_2 = \mathfrak{L}_k.$$

Moreover, this grading may also be seen as a $\mathbb{Z}^5$ *grading. First we identify* $\mathbb{Z}^5$ *with* $\{\omega^z = exp(\frac{2\pi i}{5})^z \mid z \in \mathbb{Z}\}$*. Then,*

$$\mathfrak{L} = \bigoplus_{j=0}^{4} \mathfrak{K}_{\omega^z},$$

is a $\mathbb{Z}^5$*-grading for*

$$\mathfrak{K}_1 = \mathfrak{L}_i, \ \ \mathfrak{K}_\omega = \mathfrak{L}_j, \ \ \mathfrak{K}_{\omega^2} = \mathfrak{L}_k, \ \ \mathfrak{K}_{\omega^3} = \mathfrak{L}_n, \ \ \mathfrak{K}_{\omega^4} = \mathfrak{L}_m.$$

By the universal property of G_I *there is a unique epimorphism* $\tau \colon \mathbb{Z} \to \mathbb{Z}^5$ *such that* $\mathfrak{L}_z = \mathfrak{K}_{\tau(z)}$ *for each* $z \in \{-2, -1, 0, 1, 2\}$*. This epimorphism may be defined by* $\tau(1) := \omega$*.*

3.3 Automorphisms

Automorphisms of Lie gradings of complex algebras are the focus of this section. This will lead us to some important results concerning complex Lie gradings. We will see an important connection between group gradings and homomorphisms from the character group to the automorphism group. Some knowledge of matrix Lie groups is necessary later in this section. Appendix E covers all that is required.

In this section any group is assumed to be finitely generated and abelian, unless stated otherwise.

Definition 3.23: *Suppose* G *is a group. The* **character group** *of* G, $\mathfrak{X}(G) := \mathrm{Hom}(G, \mathbb{C}^*)$, *is the group of representations of* G *by complex-valued functions. That is,* $f \colon G \to \mathbb{C}^*$ *is an element of* $\mathfrak{X}(G)$ *if* f *is a group homomorphism. The operation on* $\mathfrak{X}(G)$ *is defined by*

$$(f_i \cdot f_j)(g) := f_i(g) f_j(g),$$

for all $g \in G$*. For the sake of convenience, when there is no ambiguity we will denote this operation simply by juxtaposition. That is, we may write* $f_i \cdot f_j$ *as* $f_i f_j$ *when it is clear what we mean. Extending this notation, as clarity and convenience allow, we may write* $f_i \cdot \ldots \cdot f_i = f_i^k = f_i^{\odot k}$*.*

Proposition 3.24: *Let* $S = \mathbb{Z}_{n_1} \times \ldots \times \mathbb{Z}_{n_k} \times \mathbb{Z}^m$ *be a finitely generated abelian group. Then* $\mathfrak{X}(S) \cong \mathbb{Z}_{n_1} \times \ldots \times \mathbb{Z}_{n_k} \times (\mathbb{C}^*)^m$*.*

Proof. We first show that $\mathfrak{X}(\mathbb{Z}) = \mathbb{C}^*$ and $\mathfrak{X}(\mathbb{Z}_n) = \mathbb{Z}_n$. Define $\theta \colon \mathfrak{X}(\mathbb{Z}) \to \mathbb{C}^*$ by $\theta(f) = f(1)$. Then, for $f, g \in \mathfrak{X}(\mathbb{Z})$ we have that

$$\theta(fg) = fg(1) = f(1)g(1) = \theta(f)\theta(g),$$

which shows that θ is a group homomorphism. If $f, g \in \mathfrak{X}(\mathbb{Z})$ such that $\theta(f) = \theta(g)$, then $f(1) = g(1)$ which means that $f = g$. Moreover, for $c \in \mathbb{C}^*$ we see that $f\colon \mathbb{Z} \to \mathbb{C}^*$ defined by $f(z) = c^z$ is an element of $\mathfrak{X}(\mathbb{Z})$. Therefore θ is an isomorphism.

We now define $\phi\colon \mathfrak{X}(\mathbb{Z}_n) \to U$, where $U := \{\omega^j = exp(\frac{2\pi i}{n})^j \mid j \in \mathbb{Z}\}$. For each $f \in \mathfrak{X}(\mathbb{Z}_n)$ we have that $f(z) = f(1)^z$, for each $z \in \mathbb{Z}_n$. In particular, $f(1)^n = 1$. Therefore, we may define

$$\phi(f) := f(1) \in U.$$

A similar argument to that of θ above shows that ϕ is injective. For $\omega^j \in U$ we may assume without loss of generality that $0 \leq j \leq n$. We may induce $f_j \in \mathfrak{X}(\mathbb{Z}_n)$ by $f_j(1) = \omega^j$. Thus, ϕ is an isomorphism. We may identify U with $\mathbb{Z}_n$ and so $\mathfrak{X}(\mathbb{Z}_n) \cong \mathbb{Z}_n$.

Now we may define $\varphi\colon \mathfrak{X}(S) \to \mathbb{Z}_{n_1} \times \ldots \times \mathbb{Z}_{n_k} \times (\mathbb{C}^*)^m$, by

$$\varphi(f) = (f(1, 0, \ldots, 0), f(0, 1, 0, \ldots, 0), \ldots, f(0, \ldots, 0, 1)).$$

The result then follows from $\mathfrak{X}(\mathbb{Z}) \cong \mathbb{C}^*$ and $\mathfrak{X}(\mathbb{Z}_n) \cong \mathbb{Z}_n$. □

Proposition 3.25: *Let S be a finitely generated abelian group and A a complex algebra.*

(i) If $\tau\colon \mathfrak{X}(S) \to Aut(A)$ is a homomorphism, then τ induces an S-grading of A

$$\Gamma_\tau\colon A = \bigoplus_{i \in S} A_i,$$

where $A_i = \{a \in A \mid \tau(f)(a) = f(i)a \quad \forall\, f \in \mathfrak{X}(S)\}$.

(ii) If $\Gamma_\tau\colon A = \bigoplus_{i \in S} A_i$ is an S-grading of A, then it is induced by a homomorphism

$$\begin{aligned} \tau\colon \mathfrak{X}(S) &\to Aut(A) \\ f &\mapsto \tau(f)\colon \quad A \to A, \\ &\qquad\qquad \alpha_i \mapsto f(i)\alpha_i, \end{aligned}$$

for $\alpha_i \in A_i$. That is, $\tau(f)(\alpha_i) = f(i)\alpha_i$, for $\alpha_i \in A_i$.

Proof.

(i): We note first that the A_i are vector subspaces of A. To see this take $x, y \in A_i,\ \alpha, \beta \in \mathbb{C},\ f \in \mathfrak{X}(S)$ and we have that

$$\begin{aligned} \tau(f)(\alpha x + \beta y) &= \alpha\tau(f)(x) + \beta\tau(f)(y) = \alpha f(i)x + \beta f(i)y \\ &= f(i)(\alpha x + \beta y). \end{aligned}$$

We claim that the A_i are linearly independent. We remark that if $a \in A_i \cap A_j$ (for $i \neq j$) then $f(i)a = f(j)a$ for all $f \in \mathfrak{X}(S)$. So $A_i \cap A_j = 0$ (for $i \neq j$). For $a_i \in A_i,\ \alpha_i \in \mathbb{C}$ suppose that $\sum_{i \in S} \alpha_i a_i = 0$. Then,

$$0 = \tau(f)\left(\sum_{i \in S} \alpha_i a_i\right) = \sum_{i \in S} \alpha_i \tau(f)(a_i) = \sum_{i \in S} \alpha_i f(i) a_i,$$

for all $f \in \mathfrak{X}(S)$. Therefore, $\sum_{i \in S} (\alpha_i - \alpha_i f(i)) a_i = 0$ for all $f \in \mathfrak{X}(S)$. Considering only the i for which $a_i \neq 0$, we see that $\alpha_i = \alpha_i - \alpha_i f(i)$ for all $f \in \mathfrak{X}(S)$. This tells us that either $\alpha_i = 0$ for all such i, or $f(i) = 1$ for all $f \in \mathfrak{X}(S)$ and for all such i. In the former case we are done. The latter case implies that $S = 0$ and so $A_i = A$.

Since the A_i are linearly independent and A is finite-dimensional, we find that $I := \mathrm{Supp}(\Gamma_\tau)$ is finite. We proceed by induction on $|I|$. If $|I| = 1$ then $A = A_i$ and we are done.

Suppose now that $|I| = n > 1$. Fix $j \in I$ and set $Q := A/A_j$, as a vector space (the algebra structure is not important as we need only demonstrate that A has a vector space decomposition). We define

$$\begin{aligned} \tilde{\tau} \colon \mathfrak{X}(S) &\to Aut(Q) \\ f &\mapsto \tau(f) \colon \qquad Q \to Q \\ &\qquad\qquad x + A_j \mapsto \tau(f)(x) + A_j, \end{aligned}$$

that is, $\tilde{\tau}(f)(x + A_j) = \tau(f)(x) + A_j$. To see that $\tilde{\tau}$ is well-defined we note that if $x, y \in A$ such that $x - y \in A_j$, we have that $f(j)(x-y) \in A_j$ for all $f \in \mathfrak{X}(S)$ since A_j is a vector subspace of A. For $f, g \in \mathfrak{X}(S)$ we have that

$$\begin{aligned} \tilde{\tau}(fg)(x + A_j) &= \tau(fg)(x + A_j) = (\tau(f) \circ \tau(g))(x + A_j) \\ &= f(i)g(i)(x + A_j) = \tilde{\tau}(f)(g(i)x + A_j) \\ &= (\tilde{\tau}(f) \circ \tilde{\tau}(g))(x + A_j), \end{aligned}$$

which shows that $\tilde{\tau}$ is a homomorphism. By the inductive hypothesis we find that $\tilde{\tau}$ induces an S-grading

$$A/A_j = Q = \bigoplus_{i \in I \setminus \{j\}} Q_i = \bigoplus_{i \in S} Q_i,$$

where $Q_i = \{x + A_j \in Q \,|\, \tilde{\tau}(f)(x + A_j) = f(i)x + A_j,\ \forall\ f \in \mathfrak{X}(S)\}$.

Take $v \in A$, then

$$v + A_j = \sum_{i \in I \setminus \{j\}} \alpha_i a_i + A_j,$$

for some $a_i \in A_i,\ \alpha_i \in \mathbb{C}$. Therefore, $v = \sum_{i \in I \setminus \{j\}} a_i + a_j$ for some $a_j \in A_j$. This shows that

$$A = \bigoplus_{i \in I \setminus \{j\}} Q_i \oplus A_j.$$

We now claim that $\bigoplus_{i \in I \setminus \{j\}} Q_i \cong \bigoplus_{i \in I \setminus \{j\}} A_i$. We define a linear map

$$\begin{aligned} \phi \colon \bigoplus_{i \in I \setminus \{j\}} Q_i &\to \bigoplus_{i \in I \setminus \{j\}} A_i \\ \sum_i a_i + A_j &\mapsto \sum_i a_i. \end{aligned}$$

To see that ϕ is well-defined we note that for $\sum_i a_i + A_j = \sum_i b_i + A_j$ we have that $\sum_i a_i - b_i \in A_j$. If $\sum_i a_i = \sum_i b_i$ then we must also have $\sum_i a_i + A_j = \sum_i b_i + A_j$ and so ϕ is injective. Take an element $\sum_i a_i$ of the codomain. Then $\phi\left(\sum_i a_i + A_j\right) = \sum_i a_i$ and hence ϕ is surjective. This gives us the desired vector space decomposition.

Moreover for $a_i \in A_i,\ a_j \in A_j$, and $f \in \mathfrak{X}(S)$ we have that

$$\begin{aligned} \tau(f)(a_i a_j) &= \tau(f)(a_i)\tau(f)(a_j) = f(i)a_i f(j)a_j = f(i)f(j)a_i a_j \\ &= f(i+j)a_i a_j \end{aligned}$$

Which means that $A_i A_j \subseteq A_{i+j}$, and so the decomposition is an S-grading, except that it remains to show that I generates S.

We let $\pi \colon S \to S_1 := S/\langle I \rangle$ be the canonical quotient epimorphism. We may then define a monomorphism

$$\begin{aligned} \iota \colon \mathfrak{X}(S_1) &\to \mathfrak{X}(S) \\ \tilde{f} &\mapsto f \colon \quad S \to \mathbb{C}^* \\ &\qquad\qquad s \mapsto \tilde{f}(s + \langle I \rangle), \end{aligned}$$

that is, $\iota(\tilde{f})(s) = \tilde{f}(s + \langle I \rangle)$. To see that ι is a homomorphism we note that

$$\begin{aligned} \iota(\tilde{f})\iota(\tilde{g})(s) &= \iota(\tilde{f})(s)\iota(\tilde{g})(s) = \tilde{f}(s + \langle I \rangle)\tilde{g}(s + \langle I \rangle) = \tilde{f}\tilde{g}(s + \langle I \rangle) \\ &= \iota(\tilde{f}\tilde{g})(s), \end{aligned}$$

for all $s \in S,\ \tilde{f}, \tilde{g} \in \mathfrak{X}(S_1)$. To see that ι is injective we note that if $\iota(\tilde{f}) = \iota(\tilde{g})$ for $\tilde{f}, \tilde{g} \in \mathfrak{X}(S_1)$ then $\tilde{f}(s + \langle I \rangle) = \tilde{g}(s + \langle I \rangle)$ for all $s \in S$, and so $\tilde{f} = \tilde{g}$.

We then have the $\tau_1 := \tau \circ \iota \colon \mathfrak{X}(S_1) \to Aut(A)$ is a homomorphism. This induces a decomposition

$$A = \bigoplus_{i \in S_1} A_i^1,$$

where $A_i^1 := \{a \in A \,|\, \tau_1(\tilde{f})(a) = \tilde{f}(i)a,\ \forall \tilde{f} \in \mathfrak{X}(S_1)\}$. We now consider the set $I_1 := \{i \in S_1 \,|\, A_i^1 \neq 0\}$. Take $i_1 = s_1 + \langle I \rangle \in I_1$ and $0 \neq a \in A_i^1$. Then

$$\tau_1(\tilde{f})a = (\tau \circ \iota)(\tilde{f})(a) = \iota(\tilde{f})(i_1)a = \iota(\tilde{f})(s_1 + \langle I \rangle)a,$$

for all $\tilde{f} \in \mathfrak{X}(S_1)$.

We set $S_2 := S_1/\langle I_1 \rangle$. Similar to the above, this defines a homomorphism $\tau_2 \colon \mathfrak{X}(S_2) \to Aut(A)$ which induces a decomposition

$$A = \bigoplus_{i \in S_2} A_i^2,$$

where $A_i^2 := \{a \in A \,|\, \tau_2(\tilde{f})(a) = \tilde{f}(i)a,\ \forall \tilde{f} \in \mathfrak{X}(S_2)\}$. We may now consider the set $I_2 := \{i \in S_2 \,|\, A_i^2 \neq 0\}$. Take $i_2 = s_2 + \langle I_1 \rangle \in I_2$ and $0 \neq a \in A_i^2$. Then

$$\tau_2(\tilde{f})a = (\tau_1 \circ \iota_1)(\tilde{f})(a) = \iota_1(\tilde{f})(i_2)a = \iota_1(\tilde{f})(s_2 + \langle I_1 \rangle)a,$$

for all $\tilde{f} \in \mathfrak{X}(S_2)$.

Continuing in this fashion, if the process ever reaches an $n \in \mathbb{Z}^+$ such that s_n must be in $\langle I_{n-1} \rangle$ then $S_n = \langle I_{n-1} \rangle$ and so (cascading backwards) we find that $S_1 = \langle I \rangle$. This means that $S = \langle I \rangle$, and so we are done. However, if the process does not terminate, then we have $s_1 \notin \langle I \rangle$, $s_2 \notin \langle I_1 \rangle$, $\ldots$, and so we have that

$$\langle I \rangle \subset \langle I \cup \{s_1\} \rangle \subset \langle I \cup \{s_1\} \cup \{s_2\} \rangle \subset \ldots \subset \langle I \cup \bigcup_{i \in \mathbb{Z}^+} \{s_i\} \rangle \subset S,$$

where each is a proper subset of the next. This is true for arbitrary $s_1, s_2, \ldots$. We can therefore choose a finite generating set J for S and choose the s_i such that $J \subset \bigcup_{i \in \mathbb{Z}^+} s_i$. But this contradicts J being a generating set for S.

(ii): For $f_1, f_2 \in \mathfrak{X}(S)$ and $\alpha_i \in A_i$ we have that

$$\begin{aligned}(\tau(f_1) \circ \tau(f_2))(\alpha_i) &= \tau(f_1)(f_2(i)\alpha_i) = f_2(i)\tau(f_1)(\alpha_i) = f_2(i)f_1(i)\alpha_i \\ &= f_1(i)f_2(i)\alpha_i = f_1f_2(i)\alpha_i = \tau(f_1f_2)(\alpha_i).\end{aligned}$$

Thus, τ is a homomorphism. By construction, τ induces a S-grading of A as in (i). The decomposition follows in the same way. The conditions that $A_iA_j \subseteq A_{i+j}$ and that S is generated by $I := \text{Supp}(\Gamma_\tau)$, follow from the original grading.

□

Lemma 3.26: *Set $\phi := \theta^{-1}$, where $\theta\colon \mathfrak{X}(\mathbb{Z}) \to \mathbb{C}^*$ is the isomorphism from Proposition 3.24. If $\tau\colon \mathfrak{X}(\mathbb{Z}) \to Aut(A)$ is a homomorphism arising from a $\mathbb{Z}$-grading of a complex algebra A, then $\psi := \tau \circ \phi$ is continuous.*

Proof. We note that by Remark E.2 and Proposition E.22 we have that $\mathbb{C}^*$ and $Aut(A)$ are matrix Lie groups. Let F be an open set in $Aut(A)$. We claim that $C := \mathbb{C}^* \backslash \psi^{-1}(F)$ is closed. Let $(z_m) \subseteq C$ such that $\lim_{m\to\infty} z_m = z$. Suppose $\psi(z) \in F$. Then there is an open ball B around z contained in F. Hence, there is some M such that $M \leq m$ implies that $\psi(z_m) \in F$ which contradicts $z_m \in C$. Thus, C is closed and ψ is continuous. □

Proposition 3.27: *If $\Gamma\colon A = \bigoplus_{g\in G} A_g$ is a G-grading of a complex algebra A and G is a cyclic group, then Γ is a decomposition of A into a direct sum of eigenspaces of a diagonalisable automorphism of A.*

Proof. There are two cases: $G = \mathbb{Z}$ or $G = \mathbb{Z}_n$ for some n.

Suppose $G = \mathbb{Z}_n$. By Proposition 3.24 we have that $\mathfrak{X}(\mathbb{Z}_n) \cong \mathbb{Z}_n$ and so, Proposition 3.25 yields a homomorphism $\tau\colon \mathfrak{X}(\mathbb{Z}_n) \to Aut(A)$ which induces the grading

$$\Gamma\colon A = \bigoplus_{z\in\mathbb{Z}_n} A_z, \text{ where}$$

$$A_z = \{a \in A \,|\, \tau(f^m)(a) = f^m(z)a,\ \forall 0 \leq m \leq n-1\},$$

(Here we have taken f as a generator for $\mathfrak{X}(\mathbb{Z}_n)$ whose existence is guaranteed because $\mathfrak{X}(\mathbb{Z}_n) \cong \mathbb{Z}_n$). The homomorphism is completely determined by $\tau(1) \in Aut(A)$ because $\tau(z) = \tau(1 + \ldots + 1) = \tau(1)^z$. Therefore, we also have that $\tau(n) = \tau(1)^n = 1_{Aut(A)}$. Let $\phi\colon \mathbb{Z}_n \to \mathfrak{X}(\mathbb{Z}_n)$ be an isomorphism. Then, $\psi := \tau \circ \phi\colon \mathbb{Z}_n \to Aut(A)$ is a homomorphism. We may set $f := \phi(1)$. Let us show that $\psi(1)$ is diagonalisable.

Set $S := \text{Supp}(\Gamma)$. We proceed by induction on $|S|$. If $|S| = 1$ then $A = A_i$ for some $i \in I$. Therefore every element of A is an eigenvector of $\psi(1)$, so we can find a basis of eigenvectors of $\psi(1)$. Thus, $\psi(1)$ is diagonalisable.

Assume now that $|S| = k + 1 > 1$. Fix $j \in S$ and set $Q := A/A_j$. Then τ induces a homomorphism

$$\begin{aligned} \tilde{\tau}\colon \mathfrak{X}(\mathbb{Z}_n) &\to Aut(Q) \\ f &\mapsto \tilde{\tau}(f)\colon \qquad Q \to Q \\ &\qquad\qquad a + A_j \mapsto \tau(f)(a) + A_j. \end{aligned}$$

That is, $\tilde{\tau}(f)(a + A_j) = \tau(f)(a) + A_j$. To see that $\tilde{\tau}$ is well-defined we note that for $x, y \in A$ such that $x - y \in A_j$, we know that

$$\tau(f)(x - y) = f(j)(x - y) \in A_j.$$

We may then set $\tilde{\psi} := \tilde{\tau} \circ \phi$. By the inductive hypothesis $\tilde{\psi}(1)$ is diagonalisable, and so there is a basis $B = \{e_1, \ldots, e_d\}$ for Q consisting of eigenvectors of $\tilde{\psi}(1)$. Choose $0 \neq a_j \in A_j$. Now we know that each e_i is an eigenvector for $\psi(1)$. Furthermore, a_j is also an eigenvector of $\psi(1)$. We claim that $B' := B \cup \{a_j\}$ is a basis for A. Clearly B' spans A, so it suffices to show that B' is linearly independent. Suppose then that $\sum_{i=1}^{d} \eta_i e_i = a_j$. Then $\sum_{i=1}^{d} \eta_i e_i + A_j \in A_j$. Since B is a basis for Q this means that $\eta_i = 0$ for each i. Thus, $a_j = 0$, a contradiction. Hence, B' is a basis for A consisting of eigenvectors of $\psi(1)$and so $\psi(1)$ is diagonalisable.

From ψ we get a grading

$$A = \bigoplus_{z \in \mathbb{Z}_n} A'_z, \text{ where}$$

$$A'_z = \{a \in A \,|\, \psi(m)(a) = (f^{\odot m}(1))^z a, \ \forall 0 \leq m \leq n-1\},$$

and $A_z = A'_z$ for each $z \in \mathbb{Z}_n$. We also have that

$$\begin{aligned}\psi(m) &= (\tau \circ \phi)(m) = \tau(\phi(1)^{\odot m}) = \tau(f \cdot \ldots \cdot f) \\ &= \tau(f) \circ \ldots \circ \tau(f) = [\tau(f)]^m = \psi(1)^m.\end{aligned}$$

Since $\psi(1)$ is diagonalisable its eigenspaces are exactly equal to its generalised eigenspaces. From here it follows that the $\mathbb{Z}_n$-grading of A is a decomposition into the eigenspaces of $\psi(1)$.

Now we move to the case $G = \mathbb{Z}$. The grading is induced by a homomorphism $\tau\colon \mathfrak{X}(\mathbb{Z}) \to Aut(A)$. Let $\phi := \theta^{-1}$ where $\theta\colon \mathfrak{X}(\mathbb{Z}) \to \mathbb{C}^*$ is the isomorphism from Proposition 3.24. Then we may define a homomorphism $\psi := \tau \circ \phi\colon \mathbb{C}^* \to Aut(A)$, which induces a grading equivalent to the original grading. For $f \in \mathfrak{X}(\mathbb{Z})$ we have $f(m) = f(1)^m = z^m$, where $z := f(1)$. Since, for each $f \in \mathfrak{X}(\mathbb{Z})$, this holds for some $z \in \mathbb{C}^*$ we have that the grading is

$$A = \bigoplus_{z \in \mathbb{Z}} A_z,$$

where $A_z = \{a \in A \,|\, \psi(w)a = w^z a, \ \forall w \in \mathbb{C}^*\}$.

From Lemma 3.26 we know that ψ is continuous and thus a matrix Lie group homomorphism. An application of Theorem E.20 yields the differential $d\psi(1)\colon Lie(\mathbb{C}^*) \to Lie(Aut(A))$. By Remark E.16 and Proposition E.24 we have that $Lie(\mathbb{C}^*) = \mathbb{C}$ and $Lie(Aut(A)) = Der(A)$. Hence, $d\psi(1)\colon \mathbb{C} \to Der(A)$. We then have that

$$e^{d\psi(1)(tz)} = \psi(e^{tz}),$$

for all $t \in \mathbb{R}$. Hence, for $a \in A_m$ we have that

$$e^{td\psi(1)(z)}(a) = \psi(e^{tz})(a) = (e^{(tz)})^m(a) = e^{tmz}(a).$$

We will denote $D := d\psi(1)$. We know that $D \in Der(A)$ is linear, therefore $D(z) = zD(1)$ for all $z \in \mathbb{Z}$. Taking the derivative at $t = 0$ we find that

$$D(z)(a) = mza,$$

and taking this at $z = 1$ yields that $a \in A_m$ if and only if $D(1)(a) = ma$. This means that the the grading is a decomposition of A into the eigenspaces of $D(1)$, which is then a diagonalizable derivation of A. Consider the map $\eta := e^{D(1)}$. This is a diagonalizable automorphism of A. Moreover, e^m is an eigenvalue of η if and only if m is an eigenvalue of $D(1)$. Hence, we find that the grading spaces are each of the form

$$A_z = \{a \in A \,|\, \eta(a) = e^z a\}.$$

□

Remark 3.28: *Let $G = \prod_{i=1}^n G_i$ be the direct product of groups G_i. Each projection $\pi\colon G \to G_i$ induces a monomorphism*

$$\begin{aligned} m_i\colon \mathfrak{X}(G_i) &\to X(G) \\ f\colon G_i \to \mathbb{C}^* &\mapsto \tilde{f}\colon G \to \mathbb{C}^*, \end{aligned}$$

where $\tilde{f}(g) = \begin{cases} f(g), & \text{if } g \in G_i \\ 1, & \text{if } g \in G \backslash G_i \end{cases}$.

The m_i are monomorphisms because if $m_i(f) = m_i(f')$ then f and f' agree on G_i and so $f = f'$.

Therefore, any homomorphism $\phi\colon \mathfrak{X}(G) \to Aut(A)$ induces, not only a G-grading of A but also, G_i-gradings via the homomorphisms

$$\phi_i = \phi \circ m_i\colon \mathfrak{X}(G_i) \to Aut(A).$$

Moreover, for any $t_i \in \mathfrak{X}(G_i)$ and $t_j \in \mathfrak{X}(G_j)$ we have that

$$\begin{aligned} \phi_i(t_i) \circ \phi_j(t_j) &= (\phi \circ m_i)(t_i) \circ (\phi \circ m_j)(t_j) = \phi(\tilde{t_i}) \circ \phi(\tilde{t_j}) = \phi(\tilde{t_i} \cdot \tilde{t_j}) \\ &= \phi(\tilde{t_j} \cdot \tilde{t_i}) = \phi_j(t_j) \circ \phi_i(t_i). \end{aligned} \tag{3.3}$$

That is, $\phi_i(t_i)$ and $\phi_j(t_j)$ commute.

Lemma 3.29: *A family $\{\alpha_1, \ldots, \alpha_k\}$ of commuting diagonalizable automorphisms of a complex algebra, A, induces a grading $A = \bigoplus_{i \in I} A_i$, where*

(i) The A_i are α_j-invariant for all $i \in I,\ j \in \{1, \ldots, k\}$.

(ii) For any $i \in I$ the restriction $\alpha_j|_{A_i}$ acts as a scalar multiple of the identity, for every $j \in \{1, \ldots, k\}$.

Moreover, every G-grading of A takes the form described above.

Proof. We first focus on the direct part of the lemma. Let $L := \{\lambda_1, \ldots, \lambda_r\}$ be the set of eigenvalues of α_k, and $A_j := E_{\alpha_k}(\lambda_j)$ the eigenspace of α_k for λ_j. Since the α_j commute we may find a basis for A consisting of simultaneous eigenvectors for the α_j. We then have that A decomposes into the direct sum of the A_j and these eigenspaces agree with each of the α_i because we may simultaneously diagonalize them. For $v \in A_i$ we find that

$$\alpha_k \alpha_j(v) = \alpha_j \alpha_k(v) = \lambda_i \alpha_j(v),$$

which shows that A_i is α_j-invariant. As the α_j are simultaneously diagonalizable, (i) follows.

Now, (ii) follows because the A_j are eigenspaces and the decomposition into eigenspaces agrees with each α_i. We note that for $v_i \in A_i,\ v_j \in A_j$ we have

$$\alpha_k(v_i v_j) = \alpha_k(v_i)\alpha_k(v_j) = \lambda_i \lambda_j v_i v_j,$$

which shows that the decomposition of A into the A_j is a grading.

Suppose now that $\Gamma\colon A = \bigoplus_{g\in G} A_g$ is a G-grading of A, where

$$G = \mathbb{Z}_{n_1} \times \ldots \times \mathbb{Z}_{n_r} \times \mathbb{Z}^m = \prod_{i=1}^{r+1} G_i,$$

for some $r, m \geq 0$. There is some homomorphism $\tau\colon \mathfrak{X}(G) \to Aut(A)$ which induces Γ. As in Remark 3.28 τ induces G_i-gradings via $\tau_i\colon \mathfrak{X}(G_i) \to Aut(A)$. An application of Proposition 3.27 yields that each grading, τ_i, is a decomposition of A into eigenspaces of a diagonalizable automorphism of A, say $\alpha_i := \tau_i(t_i)$. From (3.3) we see that the α_i commute. The result then follows from the direct part of the lemma. □

Corollary 3.30: *Let $\mathfrak{L}$ be a complex semisimple Lie algebra with a $G := \mathbb{Z}^n$ grading $\psi\colon \mathfrak{X}(G) \to Aut(\mathfrak{L})$. Then $d\psi(1)\colon \mathbb{C}^n \to Der(\mathfrak{L})$ is a monomorphism.*

Proof. Let $\mathfrak{L} = \bigoplus_{g\in G} \mathfrak{L}_g$ be the grading and, applying Proposition 3.25, let $\psi\colon \mathfrak{X}(G) \to Aut(\mathfrak{L})$ be the homomorphism which induces the grading. As in Remark 3.28 we may consider the gradings $\psi_i\colon \mathfrak{X}(\mathbb{Z}) \to Aut(\mathfrak{L})$ for $1 \leq i \leq n$. The proof of Proposition 3.27 yields that each of the ψ_i comes from a diagonalizable derivation of $\mathfrak{L}$ with integer eigenvalues, we will denote these as $D_i := d\psi_i(1)(1)\colon \mathfrak{L} \to \mathfrak{L}$. Since each of the D_i are linear we have that

$$d\psi(1)(z_1, \ldots, z_n) = \sum_{i=1}^{n} z_i D_i.$$

Therefore, to prove $\mathrm{Ker}(d\psi(1)) = 0$ it suffices to show that the D_i are linearly independent.

Suppose that $\sum_{i=1}^{n} z_i D_i = 0$. Now, for $\lambda = (\lambda_1, \ldots, \lambda_n) \in G$ we have $x \in \mathfrak{L}_g$ if and only if $D_i(x) = d\psi_i(1)(1)(x) = \lambda_i x$ for each i. In such a case we have that

$$0 = \sum_{i=1}^{n} z_i D_i(x) = \left(\sum_{i=1}^{n} z_i \lambda_i\right) x.$$

We set $g := (g_1, \ldots, g_n)$ for each $g \in I := \{g \in G \,|\, \mathfrak{L}_g \neq 0\}$. Hence, for $v := \sum_{g \in G} v_g$ with $v_g \in \mathfrak{L}_g$ we have

$$\begin{aligned} d\psi(1)(z_1, \ldots, z_n)(v) &= \sum_{g \in G} d\psi(1)(z_1, \ldots, z_n)(v_g) = \sum_{g \in G} \sum_{i=1}^{n} z_i D_i(v_g) \\ &= \sum_{i=1}^{n} \sum_{g \in G} z_i D_i(v_g) = \sum_{i=1}^{n} \sum_{g \in G} z_i g_i v_g \\ &= \sum_{i=1}^{n} z_i g_i v. \end{aligned}$$

Therefore, if $\lambda = (\lambda_1, \ldots, \lambda_n) \in I$ then

$$\sum_{i=1}^{n} z_i \lambda_i = 0. \tag{3.4}$$

We will now show that $z_1 = 0$ but a similar argument holds for each i. Noting that $G = \langle I \rangle$, we choose integers m_i and l such that

$$(1, 0, \ldots, 0) = \sum_{i=1}^{l} m_i g_i,$$

where $g_i = (\lambda_{i1}, \ldots, \lambda_{in}) \in I$. Therefore by (3.4) we have that $\sum\limits_{k=1}^{n} z_k \lambda_{ik} = 0$ for each i. Moreover, $1 = \sum\limits_{i=1}^{l} m_i \lambda_{i1}$ and $0 = \sum\limits_{i=1}^{l} m_i \lambda_{ik}$ for each $k \neq 1$. Hence,

$$\begin{aligned} z_1 &= z_1 \sum_{i=1}^{l} m_i \lambda_{i1} + \sum_{k=2}^{n} z_k \sum_{i=1}^{l} m_i \lambda_{ik} = \sum_{k=1}^{n} z_k \sum_{i=1}^{l} m_i \lambda_{ik} = \sum_{i=1}^{l} m_i \sum_{k=1}^{n} z_k \lambda_{ik} \\ &= \sum_{i=1}^{l} m_i \cdot 0 = 0. \end{aligned}$$

$\square$

Chapter 4

Derivations of the Octonions

This chapter is devoted to constructing the Lie algebra $\mathfrak{g}_2$. We will construct it as the algebra of derivations of the octonions. Some prerequisite definitions and results relating to the octonions can be found in Appendix F. This includes definitions of some maps we will use in this chapter. We begin by looking at alternative algebras. We then shift our attention to $\mathfrak{g}_2$ and look at results bounding its dimension. Finally, we find its exact dimension and define a specific grading on $\mathfrak{g}_2$. Throughout this chapter (unless otherwise specified) all vector spaces are assumed to be finite-dimensional and defined over the field of real numbers.

4.1 Alternative Algebras

Alternative algebras are the focus of this section. We also look at some properties of maps which we will use frequently in later sections of this chapter.

Remark 4.1: *Suppose A is an algebra. We may define a tri-linear map, called the* **associator***, as follows*

$$\begin{aligned}(-,-,-)\colon A \times A \times A &\to A \\ (x,y,z) &:= (xy)z - x(yz),\end{aligned}$$

for all $x, y, z \in A$. We say A is **alternative** *if the associator is an alternative map. That is, if*

$$(x,y,z) = -(y,x,z) = -(x,z,y),$$

for all $x, y, z \in A$.

Definition 4.2: *We will now define the left and right multiplciation operators on an algebra A. Respectively, these are $L_x\colon A \to A$ and $R_x\colon A \to A$. We define them for $x, y \in A$, as follows: $L_x(y) := xy$ and $R_x(y) := yx$.*

Lemma 4.3: *An algebra A is alternative if and only if the following equalities hold for all $x, z \in A$.*

$$\begin{aligned} R_{xz} - R_z R_x &= R_x R_z - R_{zx} \\ &= L_{zx} - L_z L_x = L_x L_z - L_{xz} \\ &= [R_z, L_x] = [L_z, R_x]. \end{aligned}$$

Proof. Note that,

$$\begin{aligned} (R_{xz} - R_z R_x)(y) &= y(xz) - (yx)z = -(y, x, z), \\ (R_x R_z - R_{zx})(y) &= (yz)x - y(zx) = (y, z, x), \\ (L_{zx} - L_z L_x)(y) &= (zx)y - z(xy) = (z, x, y), \\ (L_x L_z - L_{xz})(y) &= x(zy) - (xz)y = -(x, z, y), \\ [R_z, L_x](y) &= R_z L_x(y) - L_x R_z(y) = (xy)z - x(yz) = (x, y, z), \\ [L_z, R_x](y) &= z(yx) - (zy)x = -(z, y, x), \end{aligned}$$

for any $y \in A$. Now it follows that all of the above are equal if and only if A is alternative. □

Lemma 4.4: *Suppose A is an alternative algebra. For all $x, z \in A$,*

1. $R_x R_z + R_z R_x = R_{xz+zx}$,
2. $[R_x, R_z] = -R_{[x,z]} - 2[L_x, R_z]$, *and*
3. $[L_x, L_z] = L_{[x,z]} - 2[L_x, R_z]$.

Proof.

(1): From Lemma 4.3 we have

$$R_{xz} - R_z R_x = R_x R_z - R_{zx},$$

for all $x, z \in A$. Therefore,

$$R_x R_z + R_z R_x = R_{xz} + R_{zx} = R_{xz+zx}.$$

(2): For all $x, z \in A$ we have

$$\begin{aligned} &[R_x, R_z] + R_{[x,z]} + 2[L_x, R_z] = R_x R_z - R_z R_x + R_{xz} - R_{zx} + 2[L_x, R_z] \\ &= (R_{xz} - R_z R_x) + (R_x R_z - R_{zx}) + 2[L_x, R_z] \\ &= [R_z, L_x] + [R_z, L_x] + 2[L_x, R_z] = 0. \end{aligned}$$

(3): This follows similarly to part (2).

□

Lemma 4.5: *If A is an alternative algebra, then for all $x, y, z \in A$*

$$[R_y, [R_x, R_z]] = R_{[y,[x,z]]-2(x,y,z)}.$$

Proof. For all $x, y, z \in A$ we have

$$\begin{aligned}
&[R_y, [R_x, R_z]] = [R_y, R_xR_z - R_zR_x] \\
&= R_yR_xR_z - R_yR_zR_x - R_xR_zR_y + R_zR_xR_y \\
&+ (R_xR_yR_z - R_xR_yR_z) + (R_zR_yR_x - R_zR_yR_x) \\
&= (R_yR_x + R_xR_y)R_z + R_z(R_xR_y + R_yR_x) \\
&- R_x(R_zR_y + R_yR_z) - (R_yR_z + R_zR_y)R_x \\
&= R_{yx+xy}R_z + R_zR_{xy+yx} - R_xR_{zy+yz} - R_{yz+zy}R_x \\
&= R_{z(yx+xy)+(yx+xy)z} - R_{x(yz+zy)+(zy+yz)x}.
\end{aligned}$$

Furthermore, we can see

$$\begin{aligned}
&z(xy + yx) + (yx)z + (xy)z - x(yz + zy) - (zy + yz)x \\
&= z(xy) + [z(yx) - (zy)x] + (yx)z + [(xy)z - x(yz)] - x(zy) - (yz)x \\
&= [(zx)y - (z, x, y)] + [(y, x, z) + y(xz)] - (z, y, x) + (x, y, z) \\
&+ [(x, z, y) - (xz)y] + [-(y, z, x) - y(zx)] \\
&= -2(x, y, z) + (zx)y + y(xz) - (xz)y - y(zx) \\
&= [(zx) - (xz)]y + y[xz - zx] - 2(x, y, z) \\
&= [y, [x, z]] - 2(x, y, z).
\end{aligned}$$

□

Definition 4.6: *We can now use the left and right multiplication operators to define a new bilinear map $D_{x,y} \colon A \to A$, by*

$$D_{x,y} := [L_x, L_y] + [L_x, R_y] + [R_x, R_y],$$

for $x, y \in A$.

Lemma 4.7: *Suppose A is an alternative algebra. For all $x, y, z \in A$, $D_{x,y}(z) = ad_{[x,y]}(z) - 3(x, y, z)$.*

Proof. Using Lemma 4.4 for all $x, y \in A$ we have

$$\begin{aligned}
D_{x,y} &= [L_x, L_y] + [L_x, R_y] + [R_x, R_y] \\
&= L_{[x,y]} - 2[L_x, R_y] + [L_x, R_y] - R_{[x,y]} - 2[L_x, R_y] \\
&= L_{[x,y]} - R_{[x,y]} - 3[L_x, R_y]. \qquad (4.1)
\end{aligned}$$

Therefore, for all $x, y, z \in A$ we have

$$\begin{aligned}
D_{x,y}(z) &= [x, y]z - z[x, y] - 3[L_xR_y - R_yL_x](z) \\
&= [[x, y], z] - 3[x(zy) - (xz)y] = [[x, y], z] - 3(x, y, z).
\end{aligned}$$

□

Lemma 4.8: *For any algebra A, the following are equivalent for $d \in \mathfrak{gl}(A)$ and $y \in A$.*

1. $d \in Der(A)$,
2. $[d, R_y] = R_{d(y)}$, and
3. $[d, L_y] = L_{d(y)}$.

Proof. For $x \in A$ we note that $R_{d(y)} = xd(y)$ and $[d, R_y](x) = d(xy) - d(x)y$. Hence, parts (1) and (2) are equivalent. It follows similarly that parts (1) and (3) are equivalent. □

Proposition 4.9: *Suppose A is an alternative algebra. If $x, z \in A$, then $D_{x,z} \in Der(A)$.*

Proof. Lemma 4.4 part (2) implies that

$$R_{[x,z]} = [R_z, R_x] + 2[R_z, L_x] = [R_z, R_x + 2L_x], \tag{4.2}$$

for all $x, z \in A$. Therefore, we find that, for any $x, y, z \in A$ we have

$$\begin{aligned}
2R_{D_{x,z}(y)} &= R_{2[[x,z],y]-6(x,z,y)} = R_{3([[x,z],y]-2(x,z,y))-[[x,z],y]} \\
&= 3R_{[[x,z],y]-2(x,z,y)} - R_{[[x,z],y]} = -3[R_y, [R_x, R_z]] - R_{[[x,z],y]} \\
&= -3[R_y, [R_x, R_z]] - [R_y, R_{[x,z]} + 2L_{[x,z]}] \\
&= -3[R_y, -R_{[x,z]} - 2[R_x, L_z]] - [R_y, R_{[x,z]} + 2L_{[x,z]}] \\
&= [R_y, 3R_{[x,z]} + 6[R_x, L_z] - R_{[x,z]} - 2L_{[x,z]}] \\
&= -2[R_y, L_{[x,z]} - R_{[x,z]} - 3[L_x, R_z]] \\
&= -2[R_y, D_{x,z}] = 2[D_{x,z}, R_y].
\end{aligned}$$

Hence, $R_{D_{x,z}(y)} = [D_{x,z}, R_y]$. The result follows from an application of Lemma 4.8. □

Lemma 4.10: *Suppose A is an alternative algebra. If $d \in Der(A)$ then $[d, D_{x,y}] = D_{d(x),y} + D_{x,d(y)}$, for all $x, y \in A$.*

Proof. From (4.1) we find that

$$[d, D_{x,y}] = [d, L_{[x,y]} - R_{[x,y]}] - 3[d, [L_x, R_y]]. \tag{4.3}$$

Using the Jacobi identity and Lemma 4.8 we can see that

$$[d, [L_x, R_y]] = [[d, L_x], R_y] + [L_x, [d, R_y]] = [L_{d(x)}, R_y] + [L_x, R_{d(y)}]. \tag{4.4}$$

Moreover, we note that

$$[d, L_{[x,y]}] = L_{d([x,y])} = L_{[d(x),y]+[x,d(y)]}. \tag{4.5}$$

Furthermore, we have that

$$[d, R_{[x,y]}] = R_{d([x,y])} = R_{[d(x),y]+[x,d(y)]} \tag{4.6}$$

Now, applications of (4.3), (4.4), (4.5), (4.6), and Lemma 4.4 yield that

$$\begin{aligned}
&D_{d(x),y} + D_{x,d(y)} = [L_{d(x)}, L_y] + [L_{d(x)}, R_y] \\
&+ [R_{d(x)}, R_y] + [L_x, L_{d(y)}] + [L_x, R_{d(y)}] + [R_x, R_{d(y)}] \\
&= L_{[d(x),y]} - 2[L_{d(x)}, R_y] + [L_{d(x)}, R_y] - R_{[d(x),y]} - 2[L_{d(x)}, R_y] + L_{[x,d(y)]} \\
&- 2[L_x, R_{d(y)}] + [L_x, R_{d(y)}] - R_{[x,d(y)]} - 2[L_x, R_{d(y)}] \\
&= [d, L_{[x,y]}] - [d, R_{[x,y]}] - 3[d, [L_x, R_y]] \\
&= [d, L_{[x,y]} - R_{[x,y]} - 3[L_x, R_y]] = [d, D_{x,y}].
\end{aligned}$$

□

4.2 Dimension Bounds

In this section we look at the Lie algebra $\mathfrak{g}_2$ and specifically we introduce upper bounds for its dimension. The results we cover will also be useful in the next section when we look at a specific grading of $\mathfrak{g}_2$.

Remark 4.11: *We know that* $\mathfrak{g}_2$ *is a subalgebra of* $\mathfrak{gl}(\mathbb{O})$, *which tells us that* $\dim(\mathfrak{g}_2) \leq \dim(\mathfrak{gl}(\mathbb{O})) = 64$.

Lemma 4.12: *If* $d \in \mathfrak{g}_2$, *then*

1. $d(1) = 0$ *and* $d(\mathbb{O}_0) \subseteq \mathbb{O}_0$. *Hence,* $\mathfrak{g}_2$ *is a subalgebra of* $\mathfrak{gl}(\mathbb{O}_0)$ *and so* $\dim(\mathfrak{g}_2) \leq \dim(\mathfrak{gl}(\mathbb{O}_0) = 49$.
2. $d|_{\mathbb{O}_0} \in \mathfrak{so}(\mathbb{O}_0, n)$. *Therefore,* $\mathfrak{g}_2$ *is a subalgebra of* $\mathfrak{so}(\mathbb{O}_0, n)$ *and so* $\dim(\mathfrak{g}_2) \leq \dim(\mathfrak{so}(7)) = 21$.

Proof. $d(1) = d(1^2) = d(1)1 + 1d(1) = 2d(1)$. This implies that $d(1) = 0$. Take $x \in \mathbb{O}_0$, then $\overline{x} = -x$. Therefore, $n(x) = -x^2$. This means that

$$0 = d(-n(x)) = d(x^2) = xd(x) + d(x)x = (x + d(x))^2 - x^2 - d(x)^2.$$

For any $y \in \mathbb{O}$ we have

$$tr(y)y - n(y) = (y + \overline{y})y - y\overline{y} = y^2 + \overline{y}y - y\overline{y} = y^2,$$

since $\overline{y}y = y\overline{y}$. Therefore, we find that

$$\begin{aligned}
&0 = d(x^2) = (x + d(x))^2 - x^2 - d(x)^2 = tr(x + d(x))(x + d(x)) \\
&\quad - n(x + d(x)) - tr(x)x + n(x) - tr(d(x))d(x) + n(d(x)) \\
&\quad = tr(x)[x + d(x)] + tr(d(x))x + tr(d(x))d(x) - n(x + d(x)) \\
&\quad - tr(x)x + n(x) - tr(d(x))d(x) + n(d(x)) \\
&\quad = 0 + tr(d(x))x - n(x + d(x)) - 0 + n(x) + n(d(x)) \\
&\quad = tr(d(x))x - n(x + d(x)) + n(x) + n(d(x)) = tr(d(x))x - n(x, d(x)).
\end{aligned}$$

Altogether, this yields that $tr(d(x)) = 0$ and $n(x, d(x)) = 0$ since $x \in \mathbb{O}_0$ and 1 are linearly independent. Hence, we find $d(x) \in \mathbb{O}_0$, for all $x \in \mathbb{O}_0$.

Finally, for all $x, y \in \mathbb{O}_0$ we have

$$\begin{aligned} 0 &= n(x+y, d(x+y)) = n(x, d(x)) + n(x, d(y)) + n(y, d(x)) + n(y, d(y)) \\ &= n(x, d(y)) + n(y, d(x)). \end{aligned}$$

The result follows. □

Lemma 4.13:

1. $ad(\mathbb{O}_0) \subseteq \mathfrak{so}(\mathbb{O}_0, n)$
2. $ad(\mathbb{O}_0) \cap \mathfrak{g}_2 = 0$
3. $\dim(\mathfrak{g}_2) \leq 14$

Proof.

(1): Take $x \in \mathbb{O}_0$ and $y, z \in \mathbb{O}$. Then Proposition F.3 yields that

$$n(L_x(y), z) = n(y, L_{\overline{x}}(z)) = -n(y, L_x(z)).$$

Therefore, $L_x \in \mathfrak{so}(\mathbb{O}, n)$. A similar argument shows $R_x \in \mathfrak{so}(\mathbb{O}, n)$. Since $ad_x = L_x - R_x$ we may conclude $ad_x \in \mathfrak{so}(\mathbb{O}, n)$. Moreover, for $x, y \in \mathbb{O}_0$ we find that

$$\begin{aligned} tr(ad_x(y)) &= tr(xy - yx) = n(xy - yx, 1) = tr(xy) - tr(yx) \\ &= n(x, \overline{y}) - n(y, \overline{x}) = -n(x, y) + n(y, x) = 0. \end{aligned}$$

From this and the fact that $[\mathbb{O}, \mathbb{O}] = \mathbb{O}_0$, (see the Fano plane in appendix F) we may conclude $ad_x(\mathbb{O}_0) \subseteq \mathbb{O}_0$. The result follows.

(2): Suppose $x \in \mathbb{O}_0$ such that $ad_x \in \mathfrak{g}_2$. For $y, z \in \mathbb{O}$ we have

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

Therefore, we find that

$$\begin{aligned} 0 &= (xy)z - (yx)z + y(xz) - y(zx) - x(yz) + (yz)x \\ &= ((xy)z - x(yz)) + (y(xz) - (yx)z) + ((yz)x - y(zx)) \\ &= (x, y, z) - (y, x, z) + (y, z, x) = 3(x, y, z). \end{aligned}$$

This implies that $(x, \mathbb{O}, \mathbb{O}) = 0$ which implies that $x \in \mathbb{R}1 \cap \mathbb{O}_0 = 0$.

(3): Using part (2) of Lemma 4.12 and parts (1) and (2) of this lemma we find

$$ad(\mathbb{O}_0) \oplus \mathfrak{g}_2 \subseteq \mathfrak{so}(\mathbb{O}_0, n).$$

We know that $\dim(\mathfrak{so}(\mathbb{O}_0, n)) = 21$ from Example F.4. Furthermore, we claim $\dim(ad(\mathbb{O}_0)) = 7$. It suffices to show that the homomorphism $\mathbb{O}_0 \to ad(\mathbb{O}_0)$ given by $x \mapsto ad_x$ is injective. To see this, we note that if $x \in \mathbb{O}_0$ commutes with $\mathbb{O}$, then $x = 0$. Hence, $\dim(\mathfrak{g}_2) \leq 14$.

□

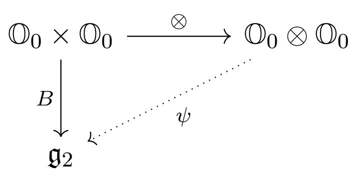
4.3 $\mathbb{Z}_2^3$-Grading

We will now build up to the construction of a fine $\mathbb{Z}_2^3$-grading of $\mathfrak{g}_2$. In fact, this is the only fine grading of $\mathfrak{g}_2$ which is not equivalent to the root space decomposition. This grading will be induced by a grading on the octonions. We will see how gradings on an algebra induce gradings on linear Lie algebras over that algebra. We will also see that $\mathfrak{g}_2$ is 14-dimensional.

Definition 4.14: *We may define a bilinear map by*

$$\begin{aligned} B\colon \mathbb{O}_0 \times \mathbb{O}_0 &\to \mathfrak{g}_2 \\ (x, y) &\mapsto D_{x,y}, \end{aligned}$$

for all $x, y \in \mathbb{O}_0$. Since B is a bilinear map, the universal property of tensor products yields a unique linear map $\psi\colon \mathbb{O}_0 \otimes \mathbb{O}_0 \to \mathfrak{g}_2$ which makes the following diagram commute.



Lemma 4.15: *Let V and W be L-modules, for some Lie algebra L. Then $V \otimes W$ is an L-module with the following action*

$$l \cdot (v \otimes w) := (l \cdot v) \otimes w + v \otimes (l \cdot w).$$

This extends linearly to sums and so determines the action for any tensors. To see this defines a module, take $\alpha, \beta \in \mathbb{R}$, $l_1, l_2 \in L$, $v_1, v_2 \in V$, and $w_1, w_2 \in \mathbb{O}_0$ and then we have the following.

$$\begin{aligned} &(\alpha l_1 + \beta l_2) \cdot (v_1 \otimes w_1) \\ &= \alpha(l_1 \cdot v_1) \otimes v_2 + \beta(l_2 \cdot v_1) \otimes w_1 + v_1 \otimes \alpha(l_1 \cdot w_1) + v_1 \otimes \beta(l_2 \cdot w_1) \\ &= \alpha(l_1 \cdot v_1) \otimes w_1 + v_1 \otimes \alpha(l_1 \cdot y) + \beta(l_2 \cdot v_1) \otimes w_1 + v_1 \otimes \beta(l_2 \cdot w_1) \\ &= \alpha l_1 \cdot (v_1 \otimes w_1) + \beta l_2 \cdot (v_1 \otimes w_1), \end{aligned}$$

and $l_1 \cdot (\alpha v_1 \otimes w_1 + \beta v_2 \otimes w_2) = \alpha l_1 \cdot (v_1 \otimes w_1) + \beta l_1 \cdot (v_2 \otimes w_2)$. *Furthermore,*

$$\begin{aligned}
&[l_1, l_2] \cdot (v_1 \otimes w_1) = (l_1 \cdot l_2 - l_2 \cdot l_1) \cdot (v_1 \otimes w_1) \\
&= l_1 \cdot (l_2 \cdot v_1) \otimes w_1 + v_1 \otimes l_1 \cdot (l_2 \cdot w_1) - l_2 \cdot (l_1 \cdot v_1) \otimes w_1 \\
&- v_1 \otimes l_2 \cdot (l_1 \cdot w_1) + l_2 \cdot v_1 \otimes l_1 \cdot w_1 - l_2 \cdot v_1 \otimes l_1 \cdot w_1 + l_1 \cdot v_1 \otimes l_2 \cdot w_1 \\
&- l_1 \cdot v_1 \otimes l_2 \cdot w_1 \\
&= l_1 \cdot (l_2 \cdot (v_1 \otimes w_1)) - l_2 \cdot (l_1 \cdot (v_1 \otimes w_1)).
\end{aligned}$$

Therefore, $V \otimes W$ is an L-module.

Proposition 4.16: *The linear map $\psi \colon \mathbb{O}_0 \otimes \mathbb{O}_0 \to \mathfrak{g}_2$ is a $\mathfrak{g}_2$-module homomorphism.*

Proof. An application of Lemma 4.15 yields that $\mathbb{O}_0 \otimes \mathbb{O}_0$ is a $\mathfrak{g}_2$-module with the following action: for $d \in \mathfrak{g}_2$ and an elementary tensor $x \otimes y \in \mathbb{O}_0 \otimes \mathbb{O}_0$,

$$d \cdot (x \otimes y) := d(x) \otimes y + x \otimes d(y).$$

This extends linearly to sums and so determines the action for any tensors.

Now, to see that ψ is a homomorphism, take $d \in \mathfrak{g}_2$ and $x, y \in \mathbb{O}_0$ then

$$\begin{aligned}
\psi(d \cdot (x \otimes y)) = \psi(d(x) \otimes y + x \otimes d(y)) &= D_{d(x),y} + D_{x,d(y)} \\
&= [d, D_{x,y}] = d \cdot D_{x,y} = d \cdot \psi(x \otimes y).
\end{aligned}$$

□

Lemma 4.17: *Suppose A is any algebra, then $Der(A)$ is a subalgebra of $\mathfrak{gl}(A)$.*

Proof. If $d_1, d_2 \in Der(A)$ and $x, y \in A$, then

$$\begin{aligned}
[d_1, d_2](x) &= d_1 d_2(x) y + x d_1 d_2(y) - d_2 d_1(x) y - x d_2 d_1(y) \\
&= x(d_1 d_2(y) - d_2 d_1(y)) + (d_1 d_2(x) - d_2 d_1(x)) y \\
&= x[d_1, d_2](y) + [d_1, d_2](x) y.
\end{aligned}$$

□

Lemma 4.18: *Suppose A is an algebra and $\Gamma \colon \bigoplus_{g \in G} A_g$ a G-grading for some group G.*

1. *Then $\mathfrak{gl}(A)$ has a G-grading given by*

$$\bigoplus_{g \in G} \mathfrak{gl}(A)_g,$$

where $\mathfrak{gl}(A)_g := \{f \in \mathfrak{gl}(A) \mid f(A_h) \subseteq A_{g+h},\ \forall h \in G\}$, for each $g \in G$.

2. *Then $Der(A)$ has a G-grading given by*

$$\bigoplus_{g\in G} Der(A)\cap \mathfrak{gl}(A)_g = \bigoplus_{g\in G} D_g,$$

where $D_g := Der(A)\cap \mathfrak{gl}(A)_g$ for each $g\in G$.

3. *Then $\mathfrak{so}(\mathbb{O}_0, n)$ has a G-grading given by*

$$\bigoplus_{g\in G} \mathfrak{so}(\mathbb{O}_0, n)\cap \mathfrak{gl}(A)_g = \bigoplus_{g\in G} S_g,$$

where $S_g := \mathfrak{so}(\mathbb{O}_0, n)\cap \mathfrak{gl}(A)_g$ for each $g\in G$.

Proof.

1. Let us first show that this is, in fact, a vector space decomposition. Take $f\in \mathfrak{gl}(A)$, we need to show that $f\in \bigoplus_{g\in G}\mathfrak{gl}(A)_g$. Consider the product projection maps $\pi_k\colon A\to A_k$ for each $k\in G$. For each $g\in G$ we define
$$f_g := \sum_{k-h=g} \pi_k f \pi_h.$$
We claim that $f_g\in \mathfrak{gl}(A)_g$, for each $g\in G$. To see this, we take $a_i\in A_i$. Then we have
$$f_g(a_i) = \sum_{k-h=g} \pi_k f\pi_h(a_i) = \begin{cases} \pi_k f(a_i)\in A_k = A_{k-h+i}, & \text{if } i=h \\ 0\in A_{k+h+i}, & \text{if } i\neq h \end{cases}.$$
It follows that $f=\sum_{g\in G} f_g \in \bigoplus_{g\in G}\mathfrak{gl}(A)_g$.
It remains for us to show $[\mathfrak{gl}(A)_g, \mathfrak{gl}(A)_h]\subseteq \mathfrak{gl}(A)_{g+h}$, for all $g,h\in G$. In fact, for $g,h,k\in G$, $f_g\in \mathfrak{gl}(A)_g$, and $f_h\in\mathfrak{gl}(A)_h$ we have
$$[f_g, f_h](\mathfrak{gl}(A)_k) = f_g f_h(\mathfrak{gl}(A)_k) - f_h f_g(\mathfrak{gl}(A)_k)\subseteq A_{(g+h)+k},$$
from which the result follows.

2. We first note that
$$\begin{aligned} Der(A) = \mathfrak{gl}(A)\cap Der(A) &= \left(\bigoplus_{g\in G}\mathfrak{gl}(A)_g\right)\cap Der(A) \\ &= \bigoplus_{g\in G}\mathfrak{gl}(A)_g\cap Der(A) = \bigoplus_{g\in G} D_g \end{aligned}$$
To see that the penultimate equality holds we must show that when we write $d\in Der(A)$ as a sum $d := \sum_{g\in G} d_g$, where each $d_g\in\mathfrak{gl}(A)_g$,

we have the each $d_g \in Der(A)$. We know that for $x \in A_h$ and $y \in A_k$ we have

$$d(xy) = d(x)y + xd(y) = \sum_{g \in G} (d_g(x)y + xd_g(y)) \in A_{g+h+k},$$

and also

$$d(xy) = \sum_{g \in G} d_g(xy) \in A_{g+h+k}.$$

Thus, $d_g(x)y + xd_g(y) = d_g(xy)$ for each $g \in G$. It follows that each $d_g \in Der(A)$.

Furthermore, if $d_g \in D_g$ and $d_h \in D_h$ we have that $[d_g, d_h] \in \mathfrak{gl}(A)_{g+h}$ and $[d_g, d_h] \in Der(A)$. Hence, $[d_g, d_h] \in D_{g+h}$.

3. This follows from a similar argument to that in part (2).

□

Example 4.19:

1. *We may describe a $\mathbb{Z}_2^3$-grading of $\mathbb{O}$ by setting:*

$$\begin{aligned} \mathbb{O}_{(1,0,0)} &= \mathbb{R}\mathbf{i} & \mathbb{O}_{(0,1,0)} &= \mathbb{R}\mathbf{j} \\ \mathbb{O}_{(0,0,1)} &= \mathbb{R}\mathbf{l} & \mathbb{O}_{(1,1,0)} &= \mathbb{R}\mathbf{k} \\ \mathbb{O}_{(1,0,1)} &= \mathbb{R}\mathbf{il} & \mathbb{O}_{(0,1,1)} &= \mathbb{R}\mathbf{jl} \\ \mathbb{O}_{(1,1,1)} &= \mathbb{R}\mathbf{kl} & \mathbb{O}_{(0,0,0)} &= \mathbb{R}1. \end{aligned}$$

It is clear to see that this decomposition is in fact a grading by considering the Fano plane description of the product on the octonions (see F.1).

2. *An application of Lemma 4.18 yields a $\mathbb{Z}_2^3$-grading of $\mathfrak{g}_2$ induced by the grading of $\mathbb{O}$ described above. For the remainder of the text we will denote this grading by $\Gamma_{\mathfrak{g}_2}$.*

Lemma 4.20: *Consider the $\mathbb{Z}_2^3$ grading of $\mathbb{O}$ described in Example 4.19 and the grading on $\mathfrak{gl}(\mathbb{O})$ which this induces by Lemma 4.18. Suppose $x \in \mathbb{O}_g$ and $y \in \mathbb{O}_h$, then*

1. *$D_{x,y} \in (\mathfrak{g}_2)_{g+h}$, and*

2. *$\phi_{x,y} := n(x, -)y - n(y, -)x \in \mathfrak{so}(\mathbb{O}_0, n)_{g+h}$.*

Proof. For $z \in \mathbb{O}_k$ we have

$$\begin{aligned} D_{x,y}(z) &= [L_x, L_y](z) + [L_x, R_y](z) + [R_x, R_y](z) \\ &= x(yz) - y(xz) + x(zy) - (xz)y + (zy)x - (zx)y. \end{aligned}$$

Each of the summands is an element of $\mathbb{O}_{(g+h)+k}$. Hence, $D_{x,y} \in (\mathfrak{g}_2)_{g+h}$.

For $u, v \in \mathbb{O}_k$ we have

$$\begin{aligned} n(\phi_{x,y}(u), v) &= n(n(x,u)y - n(y,u)x, v) = n(x,u)n(y,v) - n(y,u)n(x,v) \\ &= -[n(x,v)n(u,y) - n(y,v)n(u,x)] = -n(u, \phi_{x,y}(v)), \end{aligned}$$

which shows that $\phi_{x,y} \in \mathfrak{so}(\mathbb{O}_0, n)$.

It remains to show that $\phi_{x,y} \in \mathfrak{gl}(\mathbb{O})_{g+h}$. Since $\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}, \mathbf{l}, \mathbf{il}, \mathbf{jl}, \mathbf{kl}\}$ is an orthonormal basis for the octonions, we have that $n(\mathbb{O}_g, \mathbb{O}_h) \neq 0$ implies that $g + h = 0$. Now for $x \in \mathbb{O}_g, y \in \mathbb{O}_h$, and $z \in \mathbb{O}_k$ we find the following. If $n(x,z),\ n(y,z) \neq 0$ then $g + k = 0$ and $h + k = 0$ and so $g = h = k$. Therefore, $\phi_{x,y}(z) \in \mathbb{O}_k$. If $n(x,z) \neq 0$ but $n(y,z) = 0$ then $g = k$ and so $\phi_{x,y}(z) = n(x,z)y \in \mathbb{O}_h$. Similarly, if $n(x,z) = 0$ but $n(y,z) \neq 0$ then $\phi_{x,y}(z) \in \mathbb{O}_g$. Finally, if $n(x,z) = n(y,z) = 0$ then $\phi_{x,y}(z) = 0$. In each case it follows that $\phi_{x,y}(z) \in \mathbb{O}_{g+h+k}$. □

Proposition 4.21: *There are 14 linearly independent elements of* $\mathfrak{g}_2$ *and so* $\dim(\mathfrak{g}_2) = 14$.

Proof. An application of Lemma 4.20 part (1) allows us to conclude the following.

$$\begin{aligned} &D_{\mathbf{i},\mathbf{j}}, D_{\mathbf{l},\mathbf{kl}} \in (\mathfrak{g}_2)_{(1,1,0)} && D_{\mathbf{i},\mathbf{l}}, D_{\mathbf{j},\mathbf{kl}} \in (\mathfrak{g}_2)_{(1,0,1)} \\ &D_{\mathbf{i},\mathbf{k}}, D_{\mathbf{l},\mathbf{jl}} \in (\mathfrak{g}_2)_{(0,1,0)} && D_{\mathbf{i},\mathbf{il}}, D_{\mathbf{j},\mathbf{jl}} \in (\mathfrak{g}_2)_{(0,0,1)} \\ &D_{\mathbf{l},\mathbf{k}}, D_{\mathbf{i},\mathbf{jl}} \in (\mathfrak{g}_2)_{(1,1,1)} && D_{\mathbf{i},\mathbf{kl}}, D_{\mathbf{j},\mathbf{l}} \in (\mathfrak{g}_2)_{(0,1,1)} \\ &D_{\mathbf{j},\mathbf{k}}, D_{\mathbf{l},\mathbf{il}} \in (\mathfrak{g}_2)_{(1,0,0)} \end{aligned}$$

Moreover, we have the following.

$$\begin{aligned} D_{\mathbf{i},\mathbf{j}}(\mathbf{l}) &= 4\mathbf{kl} & D_{\mathbf{i},\mathbf{j}}(\mathbf{i}) &= 4\mathbf{j} \\ D_{\mathbf{l},\mathbf{kl}}(\mathbf{l}) &= 4\mathbf{kl} & D_{\mathbf{l},\mathbf{kl}}(\mathbf{i}) &= -2\mathbf{j}, \end{aligned}$$

which means that $D_{\mathbf{i},\mathbf{j}}$ and $D_{\mathbf{l},\mathbf{kl}}$ are linearly independent;

$$\begin{aligned} D_{\mathbf{i},\mathbf{l}}(\mathbf{i}) &= 4\mathbf{l} & D_{\mathbf{i},\mathbf{l}}(\mathbf{j}) &= 2\mathbf{kl} \\ D_{\mathbf{j},\mathbf{kl}}(\mathbf{i}) &= 4\mathbf{l} & D_{\mathbf{j},\mathbf{kl}}(\mathbf{j}) &= -10\mathbf{kl}, \end{aligned}$$

which means that $D_{\mathbf{i},\mathbf{l}}$ and $D_{\mathbf{j},\mathbf{kl}}$ are linearly independent;

$$\begin{aligned} D_{\mathbf{i},\mathbf{k}}(\mathbf{i}) &= 4\mathbf{k} & D_{\mathbf{i},\mathbf{k}}(\mathbf{l}) &= -10\mathbf{jl} \\ D_{\mathbf{l},\mathbf{jl}}(\mathbf{i}) &= 2\mathbf{k} & D_{\mathbf{l},\mathbf{jl}}(\mathbf{l}) &= 4\mathbf{jl}, \end{aligned}$$

which means that $D_{\mathbf{i},\mathbf{k}}$ and $D_{\mathbf{l},\mathbf{jl}}$ are linearly independent;

$$\begin{aligned} D_{\mathbf{i},\mathbf{il}}(\mathbf{i}) &= 4\mathbf{il} & D_{\mathbf{i},\mathbf{il}}(\mathbf{k}) &= 4\mathbf{kl} \\ D_{\mathbf{j},\mathbf{jl}}(\mathbf{i}) &= 4\mathbf{il} & D_{\mathbf{j},\mathbf{jl}}(\mathbf{k}) &= 2\mathbf{kl}, \end{aligned}$$

which means that $D_{\mathbf{i},\mathbf{il}}$ and $D_{\mathbf{j},\mathbf{jl}}$ are linearly independent;

$$D_{\mathbf{l},\mathbf{k}}(\mathbf{i}) = -2\mathbf{jl} \qquad D_{\mathbf{l},\mathbf{k}}(\mathbf{j}) = -10\mathbf{il}$$
$$D_{\mathbf{i},\mathbf{jl}}(\mathbf{i}) = 4\mathbf{jl} \qquad D_{\mathbf{i},\mathbf{jl}}(\mathbf{j}) = 2\mathbf{il},$$

which means that $D_{\mathbf{l},\mathbf{k}}$ and $D_{\mathbf{i},\mathbf{jl}}$ are linearly independent;

$$D_{\mathbf{i},\mathbf{kl}}(\mathbf{i}) = 4\mathbf{kl} \qquad D_{\mathbf{i},\mathbf{kl}}(\mathbf{j}) = -2\mathbf{l}$$
$$D_{\mathbf{j},\mathbf{l}}(\mathbf{i}) = -2\mathbf{kl} \qquad D_{\mathbf{j},\mathbf{l}}(\mathbf{j}) = 4\mathbf{l},$$

which means that $D_{\mathbf{i},\mathbf{kl}}$ and $D_{\mathbf{j},\mathbf{l}}$ are linearly independent;

$$D_{\mathbf{j},\mathbf{k}}(\mathbf{j}) = 4\mathbf{k} \qquad D_{\mathbf{j},\mathbf{k}}(\mathbf{k}) = -4\mathbf{j}$$
$$D_{\mathbf{l},\mathbf{il}}(\mathbf{j}) = -2\mathbf{k} \qquad D_{\mathbf{l},\mathbf{il}}(\mathbf{k}) = -4\mathbf{j},$$

which means that $D_{\mathbf{j},\mathbf{k}}$ and $D_{\mathbf{l},\mathbf{il}}$ are linearly independent.

The elements in the distinct homogeneous components of the $\mathbb{Z}_2^3$-grading are linearly independent. Now we also see that the pairs of elements in each homogeneous component are linearly dependent, we may conclude that the 14 elements of $\mathfrak{g}_2$ which we have presented are linearly independent. □

Corollary 4.22: *Consider the grading $\Gamma_{\mathfrak{g}_2}$ from Example4.19. For any $g, h \in \mathbb{Z}_2^3$ such that $e \notin \{g, h, g+h\}$ we have*

$$[(\mathfrak{g}_2)_g, (\mathfrak{g}_2)_h] = (\mathfrak{g}_2)_{g+h}.$$

Proof. This follows from the Fano plane (see Appendix F) and the proposition above, since $[(\mathfrak{g}_2)_g, (\mathfrak{g}_2)_h]$ has dimension 2. □

Chapter 5

Graded Contractions

Having constructed the $\mathbb{Z}_2^3$-grading $\Gamma_{\mathfrak{g}_2}$ in the previous chapter, we now look at graded contractions in general before looking more specifically at graded contractions relating to $\Gamma_{\mathfrak{g}_2}$. We will introduce the problem of finding and classifying all the Lie algebras arising as graded contractions of $\Gamma_{\mathfrak{g}_2}$. These algebras are not simple any more, and they seem to be a source for finding new solvable and nilpotent graded Lie algebras. Although the solution to this problem is not completely found here, we give some important steps to solve the problem. Moreover, we have enough work to know a considerable amount of interesting Lie algebras arise from modifying $\mathfrak{g}_2$.

Throughout this chapter (unless specified otherwise) all vector spaces are assumed to be finite-dimensional and defined over the field of real numbers.

5.1 Introduction

We first introduce and become acquainted with the notion of a graded contraction, taken from [2]. These are essentially deformations of the Lie bracket from which we may construct new Lie algebras. Throughout this section we will (unless stated otherwise) denote by: $\mathfrak{L}$ a Lie algebra, G an abelian group, and Γ a G-grading of $\mathfrak{L}$.

Definition 5.1: *Let $\Gamma\colon \mathfrak{L} = \bigoplus_{g\in G} \mathfrak{L}_g$ be a G-grading of a Lie algebra $\mathfrak{L}$, where G is some abelian group. A* **graded contraction** *of Γ is a map $\varepsilon\colon G\times G \to \mathbb{R}$ such that $(\mathfrak{L}, [,]^{\epsilon})$ is a Lie algebra with Lie bracket*

$$[x,y]^{\varepsilon} := \varepsilon(g,h)[x,y],$$

for all $x\in\mathfrak{L}_g,\ y\in\mathfrak{L}_h$. We may also write $\varepsilon(g,h) = \varepsilon_{gh} = \varepsilon_{g,h}$.

Example 5.2: *Let $\mathfrak{L}$ be a Lie algebra and Γ any G-grading of $\mathfrak{L}$.*

1. *If we set $\varepsilon(g,h) := 1$ for all $g,h\in G$, then ε is a graded contraction of Γ with $\mathfrak{L}^{\varepsilon} = \mathfrak{L}$.*

2. *If we set* $\varepsilon(g,h) := 0$ *for all* $g, h \in G$*, then* ε *is a graded contraction of* Γ *too, with* $\mathfrak{L}^\varepsilon$ *an abelian Lie algebra.*

Given an arbitrary map $\varepsilon\colon G \times G \to \mathbb{R}$, it is natural to ask under what conditions ε will be a graded contraction of Γ. The answer is not trivial if we do not know some properties of the grading under study. Some general facts are the following.

Remark 5.3:

1. *From the anti-commutativity of the Lie bracket,* $[,]$*, we find that* ε *needs to be nearly symmetric. In fact, for* $x \in \mathfrak{L}_g,\ y \in \mathfrak{L}_h$*, we have*

$$[x, y]^\varepsilon = \varepsilon(g,h)[x,y] = -\varepsilon(g,h)[y,x],$$

which coincides with $-[y,x]^\varepsilon$ *if, and only if,* $\varepsilon(g,h) = \varepsilon(h,g)$ *for all* $g, h \in G$ *such that* $[\mathfrak{L}_g, \mathfrak{L}_h] \neq 0$*.*

For $\mathfrak{L}^\varepsilon$ *to be a Lie algebra we also need the product* $[,]^\varepsilon$ *to satisfy the Jacobi identity, that is*

$$[x,[y,z]^\varepsilon]^\varepsilon + [y,[z,x]^\varepsilon]^\varepsilon + [z,[x,y]^\varepsilon]^\varepsilon = 0, \tag{5.1}$$

for all $x \in \mathfrak{L}_g,\ y \in \mathfrak{L}_g,\ z \in \mathfrak{L}_k$*. This is equivalent to the following:*

$$\begin{aligned}&\varepsilon(g,h+k)\varepsilon(h,k)[x,[y,z]] + \varepsilon(h,k+g)\varepsilon(k,g)[y,[z,x]]\\ &+ \varepsilon(k,g+h)\varepsilon(g,h)[z,[x,y]] = 0,\end{aligned}$$

for all $x \in \mathfrak{L}_g, y \in \mathfrak{L}_h, z \in \mathfrak{L}_k$*. By the Jacobi identity of* $[,]$*, this holds provided* ε *satisfies the following identity:*

$$\varepsilon(g,h+k)\varepsilon(h,k) = \varepsilon(h,k+g)\varepsilon(k,g), \quad \forall g,h,k \in G, \tag{5.2}$$

which a priori is a sufficient but not necessary condition to guarantee (5.1)*. Define a ternary map* $\varepsilon\colon G \times G \times G \to \mathbb{R}$ *by*

$$\varepsilon(g,h,k) := \varepsilon(g,h+k)\varepsilon(h,k),$$

which we denote with the same letter because there is no confusion. So we can simplify (5.2) *to:*

$$\varepsilon(g,h,k) = \varepsilon(h,k,g),$$

for all $g,h,k \in G$*.*

2. *Computations in part (1) yield that a map* $\varepsilon\colon G \times G \to \mathbb{R}$ *is a graded contraction of a grading* Γ *if, and only if, the following conditions hold*

(a1) $(\varepsilon(g,h) - \varepsilon(h,g))[x,y] = 0,$

(a2) $(\varepsilon(g,h,k)-\varepsilon(k,g,h))[x,[y,z]]+(\varepsilon(h,k,g)-\varepsilon(k,g,h))[y,[z,x]] = 0,$

for all $x \in \mathfrak{L}_g, y \in \mathfrak{L}_h, z \in \mathfrak{L}_k$ *and for all* $g,h,k \in G$.

For some concrete gradings, the conditions in the above remark can be improved. At the moment, we are using them for finding new graded contractions starting from some others.

Example 5.4: *Suppose* ε *is a graded contraction of* Γ *which satisfies the condition* (5.2)*. Then any map* $\phi\colon G \to \mathbb{R}-\{0\}$ *will give rise to a new graded contraction* $\varepsilon^\phi\colon G\times G \to \mathbb{R}-\{0\}$*, defined by*

$$\varepsilon^\phi(g,h)=\varepsilon(g,h)\frac{\phi(g)\phi(h)}{\phi(g+h)}, \quad \forall g,h \in G.$$

Indeed, let us show that ε^ϕ *is a graded contraction. Take* $g,h \in G$ *such that* $[\mathfrak{L}_g,\mathfrak{L}_h]\neq 0$*. As* ε *is a graded contraction, we have*

$$\varepsilon^\phi(g,h)=\varepsilon(g,h)\frac{\phi(g)\phi(h)}{\phi(g+h)}=\varepsilon(h,g)\frac{\phi(h)\phi(g)}{\phi(h+g)}=\varepsilon^\phi(h,g).$$

Moreover, as ε *satisfies* (5.2)*, then, for* $g,h,k \in G$*, we have*

$$\begin{aligned}\varepsilon^\phi(g,h,k)&=\varepsilon^\phi(g,h+k)\varepsilon^\phi(h,k)=\varepsilon(g,h,k)\frac{\phi(g)\phi(h+k)}{\phi(g+h+k)}\frac{\phi(h)\phi(k)}{\phi(h+k)}\\&=\varepsilon(g,h,k)\frac{\phi(g)\phi(h)\phi(k)}{\phi(g+h+k)}=\varepsilon(h,k,g)\frac{\phi(h)\phi(k+g)}{\phi(h+k+g)}\frac{\phi(k)\phi(g)}{\phi(k+g)}\\&=\varepsilon^\phi(h,k,g).\end{aligned}$$

Hence, ε *is a graded contraction by Remark 5.3.*

Although $\varepsilon^\phi\neq\varepsilon$*, the induced Lie algebras are isomorphic (even more, graded isomorphic):*

Consider the bijective linear map $\theta\colon \mathfrak{L}^\varepsilon \to \mathfrak{L}^{\varepsilon^\phi}$*, defined by*

$$\theta(x):=\phi(g)^{-1}x, \quad \forall x \in \mathfrak{L}_g^\varepsilon.$$

For $x \in \mathfrak{L}_g, y \in \mathfrak{L}_h$*, we have*

$$\begin{aligned}[\theta(x),\theta(y)]^{\varepsilon^\phi}&=\varepsilon(g,h)\frac{\phi(g)\phi(h)}{\phi(g+h)}[\theta(x),\theta(y)]\\&=\varepsilon(g,h)\frac{\phi(g)\phi(h)}{\phi(g+h)}\phi(g)^{-1}\phi(h)^{-1}[x,y]\\&=\varepsilon(g,h)\phi(g+h)^{-1}[x,y]=\varepsilon(g,h)\theta([x,y])=\theta([x,y]^\varepsilon).\end{aligned}$$

Hence, θ *is a Lie algebra isomorphism which preserves* Γ*.*

The previous example shows that new graded contractions do not always give rise to *new* Lie algebras. The following definition is therefore convenient.

Definition 5.5: *Two graded contractions ε and ε' are said to be* **equivalent** *(and we write $\varepsilon \sim \varepsilon'$) if the Lie algebras $\mathfrak{L}^{\varepsilon}$ and $\mathfrak{L}^{\varepsilon'}$ are isomorphic as graded Lie algebras.*

Graded contractions ε and ε' are said to be **equivalent via normalization** *(we write $\varepsilon \sim_n \varepsilon'$) if there is $\phi\colon G \to \mathbb{R} - \{0\}$ such that $\varepsilon' = \varepsilon^{\phi}$.*

It is quite clear that both $\sim$ and $\sim_n$ are equivalence relations. And, as shown in the above example, if $\varepsilon \sim_n \varepsilon'$, in particular $\varepsilon \sim \varepsilon'$.

5.2 Admissible Maps

Having gained some familiarity with graded contractions in the previous section, we now narrow our focus. We look at graded contractions of the grading $\Gamma_{\mathfrak{g}_2}$ and introduce the problem of finding and classifying its graded contractions up to equivalence. Throughout this section we will denote by $G := \mathbb{Z}_2^3$ and by $\Gamma_{\mathfrak{g}_2}$ the $\mathbb{Z}_2^3$-grading defined in Example 4.19.

Definition 5.6: *A map $\varepsilon\colon G \times G \to \mathbb{R}$ is said to be $\Gamma_{\mathfrak{g}_2}$-admissible, or simply* **admissible** *if ε satisfies*

$$\varepsilon(g, g) = \varepsilon(e, g) = \varepsilon(g, e) = 0,$$

where $e \in G$ is the identity and $g \in G$ is arbitrary.

Our idea is that, for each graded contraction there will exist another graded contraction equivalent to the first one which is at the same time an admissible map. For these admissible maps the conditions in Remark 5.3 can be improved.

Lemma 5.7: *If ε is a graded contraction of $\Gamma_{\mathfrak{g}_2}$, there exists another graded contraction, ε', of $\Gamma_{\mathfrak{g}_2}$ which is an admissible map and equivalent to ε.*

Proof. For all $g, h \in G$ we set

$$\varepsilon'(g,h) := \begin{cases} \varepsilon(g,h), & \text{if } g \neq h \text{ and } g, h \neq e, \\ 0, & \text{otherwise.} \end{cases}$$

By construction ε' is an admissible map. Now note that it is a graded contraction of $\Gamma_{\mathfrak{g}_2}$ equivalent to ε, that is, $(\mathfrak{g}_2, [,]^{\varepsilon'})$ is a Lie algebra isomorphic to $(\mathfrak{g}_2, [,]^{\varepsilon})$. In fact, both Lie algebras are equal (the identity is an isomorphism) since, for $x \in (\mathfrak{g}_2)_g$, $y \in (\mathfrak{g}_2)_h$,

- If $g \neq h$ and $g, h \neq e$, then $\varepsilon'(g,h) = \varepsilon(g,h)$ and so

$$[x,y]^{\varepsilon'} = \varepsilon'(g,h)[x,y] = \varepsilon(g,h)[x,y] = [x,y]^{\varepsilon};$$

- If $g = h$, (so $g + h = e$) $g = e$, or $h = e$ then $[(\mathfrak{g}_2)_g, (\mathfrak{g}_2)_h] = 0$ since $(\mathfrak{g}_2)_e = 0$. Then we also have $[x, y] = 0$ and hence $[x, y]^{\varepsilon'} = 0 = [x, y]^{\varepsilon}$.

□

The key fact on the grading $\Gamma_{\mathfrak{g}_2}$ which contributes to make things easier, is that

$$[\mathfrak{L}_g, \mathfrak{L}_h] = \mathfrak{L}_{g+h} \tag{5.3}$$

whenever $e \notin \{g, h, g + h\}$. Thus, we have the following lemma.

Lemma 5.8: *Take an admissible graded contraction ε of $\Gamma_{\mathfrak{g}_2}$ and $g, h, k \in G$. Then we have*

1. $\varepsilon(g, h) = \varepsilon(h, g)$;
2. $\varepsilon(e, -, -) = \varepsilon(-, e, -) = \varepsilon(-, -, e) = 0$;
3. $\varepsilon(g, k, k) = 0$;
4. $\varepsilon(g, h, k) = \varepsilon(g, k, h)$;
5. $\varepsilon(h + k, h, k) = 0$.

Proof.

1. We know that $\varepsilon(g, h) = \varepsilon(h, g)$ if $[\mathfrak{L}_g, \mathfrak{L}_h] \neq 0$, since ε is a graded contraction. But if $[\mathfrak{L}_g, \mathfrak{L}_h] = 0$, then either $g = h$ or $g = e$ or $h = e$ by (5.3); so the admissibility of ε gives that both $\varepsilon(g, h)$ and $\varepsilon(h, g)$ are 0, in particular they are equal.
2. This follows from ε being admissible; any admissible map with one of its arguments being e evaluates to 0.
3. $\varepsilon(g, k, k) = \varepsilon(g, k + k)\varepsilon(g, k) = \varepsilon(g, e)\varepsilon(g, k) = 0$.
4. $\varepsilon(g, h, k) = \varepsilon(g, h + k)\varepsilon(h, k) = \varepsilon(g, k + h)\varepsilon(k, h) = \varepsilon(g, k, h)$.
5. $\varepsilon(h + k, h, k) = \varepsilon(h + k, h + k)\varepsilon(h, k) = 0$.

□

The next aim is to find conditions on admissible maps which guarantee that they are graded contractions of $\Gamma_{\mathfrak{g}_2}$. To achieve this we must first dive into some specific properties of the grading $\Gamma_{\mathfrak{g}_2}$.

Lemma 5.9: *If we have three elements $g, h, k \in G$ that generate the whole group, that is, $\langle g, h, k\rangle = G$, then there exist $x \in (\mathfrak{g}_2)_g, y \in (\mathfrak{g}_2)_h, z \in (\mathfrak{g}_2)_k$ such that $[x, [y, z]]$ and $[y, [z, x]]$ are linearly independent.*

Proof. For any triplet of elements with $\langle g, h, k\rangle = G$, there is an algebra automorphism $f\colon \mathbb{O} \to \mathbb{O}$ such that

$$f(\mathbb{O}_g) = \mathbb{R}\mathbf{i}, \quad f(\mathbb{O}_h) = \mathbb{R}\mathbf{j}, \quad f(\mathbb{O}_k) = \mathbb{R}\mathbf{l}.$$

The induced map $\tilde{f}\colon \mathfrak{g}_2 \to \mathfrak{g}_2$ defined by $\tilde{f}(d) := f^{-1}df$ for any derivation $d \in \mathfrak{g}_2 = Der(\mathbb{O})$ is an automorphism of the Lie algebra, which satisfies $\tilde{f}(D_{x,y}) = D_{f(x),f(y)}$ for any $x, y \in \mathbb{O}$. Thus there is no loss of generality in assuming that $g = (1, 1, 0)$, $k = (0, 1, 0)$ and $h = (0, 1, 1)$, since there is an automorphism of the graded Lie algebra $\mathfrak{g}_2 = \mathfrak{L}$ which sends $\mathfrak{L}_g$, $\mathfrak{L}_h$ and $\mathfrak{L}_k$ to $\mathfrak{L}_{(1,1,0)}$, $\mathfrak{L}_{(0,1,0)}$ and $\mathfrak{L}_{(0,1,1)}$, respectively.

Under such assumptions, recall that $\mathbf{i} \in \mathbb{O}_{g+k}$, $\mathbf{j} \in \mathbb{O}_k$, and $\mathbf{l} \in \mathbb{O}_{k+h}$. We note that $D_{\mathbb{O}_i,\mathbb{O}_j} \in (\mathfrak{g}_2)_{i,j}$, for any $i, j \in G$. This allows us to take $x := D_{\mathbf{i},\mathbf{j}} \in (\mathfrak{g}_2)_g$, $y := D_{\mathbf{j},\mathbf{l}} \in (\mathfrak{g}_2)_h$, and $z := D_{\mathbf{i},\mathbf{k}} \in (\mathfrak{g}_2)_{g+k+g+k+k} = (\mathfrak{g}_2)_k$. We will now directly compute $[x, [y, z]]$ and $[y, [z, x]]$ and show that they are linearly independent. We will do these computations carefully.

Recall that $[d, D_{a,b}] = D_{d(a),b} + D_{a,d(b)}$ for any $a, b \in \mathbb{O}$ and $d \in Der(\mathbb{O})$. Therefore,

$$[z, x] = [D_{\mathbf{i},\mathbf{k}}, D_{\mathbf{i},\mathbf{j}}] = D_{D_{\mathbf{i},\mathbf{k}}(\mathbf{i}),\mathbf{j}} + D_{\mathbf{i},D_{\mathbf{i},\mathbf{k}}(\mathbf{j})}.$$

As $D_{x,y} = ad_{[x,y]} - 3(x, y, -)$, we find that

$$D_{\mathbf{i},\mathbf{k}}(\mathbf{i}) = [[\mathbf{i}, \mathbf{k}], \mathbf{i}] - 3(\mathbf{i}, \mathbf{k}, \mathbf{i}) = [-2\mathbf{j}, \mathbf{i}] - 0 = 4\mathbf{k},$$

and

$$D_{\mathbf{i},\mathbf{k}}(\mathbf{j}) = [[\mathbf{i}, \mathbf{k}], \mathbf{j}] - 3(\mathbf{i}, \mathbf{k}, \mathbf{j}) = -2[\mathbf{j}, \mathbf{j}] = 0.$$

Here we are taking into account that any three elements in a quaternion subalgebra associate with each other, as is the case for $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$. Joining the previous results,

$$[z, x] = 4D_{\mathbf{k},\mathbf{j}} = -4D_{\mathbf{j},\mathbf{k}}.$$

Hence,

$$-\frac{1}{4}[y, [z, x]] = [D_{\mathbf{j},\mathbf{l}}, D_{\mathbf{j},\mathbf{k}}] = D_{D_{\mathbf{j},\mathbf{l}}(\mathbf{j}),\mathbf{k}} + D_{\mathbf{j},D_{\mathbf{j},\mathbf{l}}(\mathbf{k})}.$$

Now, since

$$D_{\mathbf{j},\mathbf{l}}(\mathbf{j}) = [[\mathbf{j}, \mathbf{l}], \mathbf{j}] - 3(\mathbf{j}, \mathbf{l}, \mathbf{j}) = [2\mathbf{j}\mathbf{l}, \mathbf{j}] - 0 = 2((\mathbf{j}\mathbf{l})\mathbf{j} - \mathbf{j}(\mathbf{l}\mathbf{j})) = 4\mathbf{l},$$

and

$$\begin{aligned} D_{\mathbf{j},\mathbf{l}}(\mathbf{k}) &= 2[\mathbf{j}\mathbf{l}, \mathbf{k}] - 3(\mathbf{j}, \mathbf{l}, \mathbf{k}) = 4(\mathbf{j}\mathbf{l})\mathbf{k} - 3((\mathbf{j}\mathbf{l})\mathbf{k} - \mathbf{j}(\mathbf{l}\mathbf{k})) \\ &= -4\mathbf{i}\mathbf{l} - 3(-\mathbf{i}\mathbf{l} + \mathbf{j}(\mathbf{k}\mathbf{l})) = -4\mathbf{i}\mathbf{l} - 3(-\mathbf{i}\mathbf{l} - \mathbf{i}\mathbf{l}) = 2\mathbf{i}\mathbf{l}, \end{aligned}$$

we can see that

$$-\frac{1}{4}[y, [z, x]] = D_{4\mathbf{l},\mathbf{k}} + D_{\mathbf{j},2\mathbf{i}\mathbf{l}},$$

thus getting our first expression

$$[y,[z,x]] = -16D_{\mathbf{l},\mathbf{k}} - 8D_{\mathbf{j},\mathbf{il}}. \tag{5.4}$$

Now we will compute $[x,[y,z]]$. We begin by computing

$$[y,z] = [D_{\mathbf{j},\mathbf{l}}, D_{\mathbf{i},\mathbf{k}}] = D_{D_{\mathbf{j},\mathbf{l}}(\mathbf{i}),\mathbf{k}} + D_{\mathbf{i},D_{\mathbf{j},\mathbf{l}}(\mathbf{k})}.$$

Then, since

$$\begin{aligned} D_{\mathbf{j},\mathbf{l}}(\mathbf{i}) =& [[\mathbf{j},\mathbf{l}],\mathbf{i}] - 3(\mathbf{j},\mathbf{l},\mathbf{i}) = 2[\mathbf{jl},\mathbf{i}] - 3((\mathbf{jl})\mathbf{i} - \mathbf{j}(\mathbf{li})) \\ &= 4(\mathbf{jl})\mathbf{i} - 3(\mathbf{kl} + \mathbf{j}(\mathbf{il})) = 4\mathbf{kl} - 3(\mathbf{kl} + \mathbf{kl}) = -2\mathbf{kl}, \end{aligned}$$

and, as we saw earlier, $D_{\mathbf{j},\mathbf{l}}(\mathbf{k}) = 2\mathbf{il}$, we find that

$$[y,z] = -2D_{\mathbf{kl},\mathbf{k}} + 2D_{\mathbf{i},\mathbf{il}},$$

and hence,

$$\begin{aligned} [x,[y,z]] =& [x, -2D_{\mathbf{kl},\mathbf{k}} + 2D_{\mathbf{i},\mathbf{il}}] = -2[D_{\mathbf{i},\mathbf{j}}, D_{\mathbf{kl},\mathbf{k}}] + 2[D_{\mathbf{i},\mathbf{j}}, D_{\mathbf{i},\mathbf{il}}] \\ &= -2(D_{D_{\mathbf{i},\mathbf{j}}(\mathbf{kl}),\mathbf{k}} + D_{\mathbf{kl},D_{\mathbf{i},\mathbf{j}}(\mathbf{k})}) + 2(D_{D_{\mathbf{i},\mathbf{j}}(\mathbf{i}),\mathbf{il}} + D_{\mathbf{i},D_{\mathbf{i},\mathbf{j}}(\mathbf{il})}). \end{aligned}$$

We also have that

$$\begin{aligned} D_{\mathbf{i},\mathbf{j}}(\mathbf{kl}) &= [[\mathbf{i},\mathbf{j}],\mathbf{kl}] - 3(\mathbf{i},\mathbf{j},\mathbf{kl}) = 4\mathbf{k}(\mathbf{kl}) - 3((\mathbf{ij})(\mathbf{kl}) - \mathbf{i}(\mathbf{j}(\mathbf{kl}))) \\ &= -4\mathbf{l} - 3(\mathbf{k}(\mathbf{kl}) - \mathbf{i}(-\mathbf{il})) = -4\mathbf{l} - 3(-\mathbf{l} - \mathbf{l}) = 2\mathbf{l}, \end{aligned}$$

and $D_{\mathbf{i},\mathbf{j}}(\mathbf{k}) = [2\mathbf{k},\mathbf{k}] - 3(\mathbf{i},\mathbf{j},\mathbf{k}) = 0$. Moreover, $D_{\mathbf{i},\mathbf{j}}(\mathbf{i}) = 4\mathbf{ki} - 3(\mathbf{i},\mathbf{j},\mathbf{i}) = 4\mathbf{j}$, and

$$\begin{aligned} D_{\mathbf{i},\mathbf{j}}(\mathbf{il}) &= 4\mathbf{k}(\mathbf{il}) - 3(\mathbf{i},\mathbf{j},\mathbf{il}) = -4\mathbf{jl} - 3(\mathbf{k}(\mathbf{il}) - \mathbf{i}(\mathbf{j}(\mathbf{il}))) \\ &= -4\mathbf{jl} - 3(-\mathbf{jl} - \mathbf{i}(\mathbf{kl})) = -4\mathbf{jl} - 3(-\mathbf{jl} - \mathbf{jl}) = 2\mathbf{jl}. \end{aligned}$$

Putting this all together, we find that

$$\begin{aligned} [x,[y,z]] &= -2(D_{2\mathbf{l},\mathbf{k}} + 0) + 2(D_{4\mathbf{j},\mathbf{il}} + D_{\mathbf{i},2\mathbf{jl}}) \\ &= -4D_{\mathbf{l},\mathbf{k}} + 8D_{\mathbf{j},\mathbf{il}} + 4D_{\mathbf{i},\mathbf{jl}}. \end{aligned} \tag{5.5}$$

Take into account that $D_{a,bc} + D_{b,ca} + D_{c,ab} = 0$ for all $a,b,c \in \mathbb{O}$, so

$$D_{\mathbf{l},\mathbf{k}} = D_{\mathbf{l},\mathbf{ij}} = -D_{\mathbf{j},\mathbf{li}} - D_{\mathbf{i},\mathbf{jl}} = -D_{\mathbf{i},\mathbf{jl}} + D_{\mathbf{j},\mathbf{il}}.$$

Now we can compare (5.4) with (5.5), since:

$$\begin{aligned} [y,[z,x]] &= -16D_{\mathbf{l},\mathbf{k}} - 8D_{\mathbf{j},\mathbf{il}} = 8(-3D_{\mathbf{j},\mathbf{il}} + 2D_{\mathbf{i},\mathbf{jl}}), \\ [x,[y,z]] &= -4D_{\mathbf{l},\mathbf{k}} + 8D_{\mathbf{j},\mathbf{il}} + 4D_{\mathbf{i},\mathbf{jl}} = 4(D_{\mathbf{j},\mathbf{il}} + 2D_{\mathbf{i},\mathbf{jl}}). \end{aligned}$$

And now the independence is clear. If $0 = \frac{\alpha}{4}[x,[y,z]] + \frac{\beta}{8}[y,[z,x]]$ for some $\alpha,\beta \in \mathbb{R}$, then

$$0 = (\alpha - 3\beta)D_{\mathbf{j},\mathbf{il}} + (2\alpha + 2\beta)D_{\mathbf{i},\mathbf{jl}}$$

what happens if, and only if, $\alpha - 3\beta = 0 = 2\alpha + 2\beta$; so that $\alpha = \beta = 0$. □

Proposition 5.10: *An admissible map $\varepsilon\colon G \times G \to \mathbb{R}$ is a graded contraction of $\Gamma_{\mathfrak{g}_2}$ if, and only if, the following conditions hold for all $g, h, k \in G$:*

(b1) $\varepsilon(g, h) = \varepsilon(h, g)$,

(b2) $\varepsilon(g, h, k) = \varepsilon(k, g, h)$ *if* $\langle g, h, k\rangle = G$.

Proof. Suppose conditions (b1) and (b2) hold for ε an admissible map. We are going to see that (a1) and (a2) hold in order to use the characterization of graded contractions in Remark 5.3 (2). Of course (a1) is an immediate consequence of (b1). Recall that (a2) says that

$$(\varepsilon(g, h, k) - \varepsilon(k, g, h))[x, [y, z]] + (\varepsilon(h, k, g) - \varepsilon(k, g, h))[y, [z, x]] = 0,$$

for all $g, h, k \in G$ and all homogeneous elements $x \in \mathfrak{L}_g$, $y \in \mathfrak{L}_h$ and $z \in \mathfrak{L}_k$. If g, h, and k generate G we are done by (b2). Otherwise we would have $k \in \langle g, h\rangle = \{g, h, g + h, e\}$:

- If $k = e$: then $y \in (\mathfrak{g}_2)_e = 0$ and (a2) holds.
- If $k = g$: then $[z, x] \in (\mathfrak{g}_2)_{g+g} = (\mathfrak{g}_2)_e = 0$, and

  $$\varepsilon(g, h, k) - \varepsilon(k, g, h) = \varepsilon(g, h, g) - \varepsilon(g, g, h) = 0$$

  by Lemma 5.8, part 4. Hence, (a2) holds.
- If $k = h$: then $[y, z] \in (\mathfrak{g}_2)_{h+h} = (\mathfrak{g}_2)_e = 0$, and

  $$\varepsilon(h, k, g) - \varepsilon(k, g, h) = \varepsilon(h, h, g) - \varepsilon(h, g, h) = 0,$$

  again by Lemma 5.8, part (4).
- If $k = g + h$: then we have $[x, [y, z]] \in (\mathfrak{g}_2)_{g+(h+(g+h))} = (\mathfrak{g}_2)_e = 0$ and $[y, [z, x]] \in (\mathfrak{g}_2)_{h+(g+h)+g} = (\mathfrak{g}_2)_e = 0$.

In any case, (a2) holds.

Conversely, suppose that ε is an admissible graded contraction, in which case it satisfies (a1) and (a2). We know that for $g \neq h \in G - \{e\}$ we have $[(\mathfrak{g}_2)_g, (\mathfrak{g}_2)_h] = (\mathfrak{g}_2)_{g+h} \neq 0$. Hence, there exist $x \in (\mathfrak{g}_2)_g$ and $y \in (\mathfrak{g}_2)_h$ such that $[x, y] \neq 0$ and thus $\varepsilon(g, h) = \varepsilon(h, g)$. But in the remaining cases $g = h$ or $g = e$ or $h = e$, clearly $\varepsilon(g, h) = 0 = \varepsilon(h, g)$, so that (b1) follows. If now $\langle g, h, k\rangle = G$, Lemma 5.9 provides elements $x \in (\mathfrak{g}_2)_g, y \in (\mathfrak{g}_2)_h, z \in (\mathfrak{g}_2)_k$ such that $[x, [y, z]]$ and $[y, [z, x]]$ are linearly independent. So (a2) implies $\varepsilon(g, h, k) - \varepsilon(k, g, h) = \varepsilon(h, k, g) - \varepsilon(k, g, h) = 0$. □

Consequently we can forget about the Lie algebra and its specific grading and deal only with admissible maps satisfying (b1) and (b2). We can further simplify the situation. The rest of this work is original and still under revision, we therefore continue this work in Appendix A.

Appendix A

Graded Contractions of $\Gamma_{\mathfrak{g}_2}$

This chapter is a direct continuation of the work from Chapter 5. We go on to reduce the problem of classifying specific admissible maps to a combinatorial problem and describe a classification. We then investigate the algebras, and their properties, which arise from the graded contractions of $\Gamma_{\mathfrak{g}_2}$.

A.1 Reducing the Problem

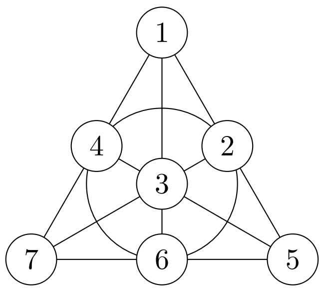
Throughout this section we will denote by $G := \mathbb{Z}_2^3$, $\mathfrak{D} := \mathfrak{g}_2 = Der(\mathbb{O})$.

Our aim is to classify, up to equivalence, all graded contractions which are also admissible maps. We will reduce this problem to a combinatorial problem.

First we set the following notation for elements of $G = \mathbb{Z}_2^3$:

$$g_0 := (0,0,0), \quad g_1 := (1,0,0), \quad g_2 := (0,1,0), \quad g_3 := (0,0,1),$$
$$g_4 := (1,1,1), \quad g_5 := (1,1,0), \quad g_6 := (1,0,1), \quad g_7 := (0,1,1).$$

We will denote $I := \{1, 2, \dots 7\}$ and $I_0 := I \cup \{0\}$. The operation in the group G allows us to define a binary operation $*\colon I_0 \times I_0 \to I_0$ by $(i,j) \mapsto i * j$ such that $g_{i*j} = g_i + g_j$, for any $i, j \in I_0$. This operation, when restricted to different elements from I, can be summarized by the diagram below: The operation applied to any two distinct elements in I yields the third element in the same line and thus $\{i, j, i * j\}$ is one of the lines of the so called Fano plane (symbolizing $P(\mathbb{Z}_2^3)$) for each $i \neq j \in I$.



We denote by

$$X := \{(i\ j) \,|\, i, j \in \{1, 2, \ldots, 7\},\ i \neq j\}.$$

Here we are using the notation $(i\ j)$ for an unordered pair $\{i, j\}$. This notation is unusual but it will allow us to avoid hundreds of braces when dealing with subsets of X throughout rest of this chapter. We will relate to each graded contraction a map from X to $\mathbb{R}$.

Definition A.1: *We will call pairwise distinct $i, j, k \in I$, satisfying $k \neq i * j$, a* **generating triplet** *or simply say they are* **generative**. *This is equivalent to saying that $\langle g_i, g_j, g_k \rangle = G$. In particular this concept does not depend on the order of the considered indices i, j, k.*

For any map $\eta\colon X \to \mathbb{R}$, we can denote by

$$\begin{aligned} &\eta_{ij} = \eta_{i,j} := \eta((i\ j)) && \text{if } i \neq j \in I, \\ &\eta_{ijk} := \eta_{i,j*k}\eta_{j,k} && \text{if } \{i, j, k\} \text{ is a generating triplet.} \end{aligned}$$

(Both expressions are well defined but observe that η_{ijk} is not defined if $\{i, j, k\}$ is not a generating triplet.) These maps give a first reduction of our problem.

Lemma A.2: *There is a bijection between the set of admissible graded contractions of Γ and the set of maps*

$$\mathcal{A} := \{\eta\colon X \to \mathbb{R} \,|\, \eta_{ijk} = \eta_{jki},\ \textit{for all generating triplet } i, j, k \in I\}.$$

Proof. Denote $\mathcal{A}' := \{\varepsilon\colon G \times G \to \mathbb{R} \,|\, \varepsilon \text{ is an admissible graded contraction}\}$. If $\varepsilon \in \mathcal{A}'$, let us check that $\eta\colon X \to \mathbb{R}$ defined by $\eta((i\ j)) = \varepsilon(g_i, g_j)$ belongs to $\mathcal{A}$: first note that η is well defined since $\varepsilon(g_i, g_j) = \varepsilon(g_j, g_i)$ if $(i\ j) \in X$. Second, for a generating triplet $i, j, k \in I$, we have that $\eta_{ijk} = \eta_{i,j*k}\eta_{j,k} = \varepsilon(g_i, g_{j*k})\varepsilon(g_j, g_k) = \varepsilon(g_i, g_j + g_k)\varepsilon(g_j, g_k) = \varepsilon(g_i, g_j, g_k)$ coincides with $\varepsilon(g_j, g_k, g_i) = \eta_{jki}$ by Proposition 5.10, so that $\eta \in \mathcal{A}$. Conversely, if $\eta \in \mathcal{A}$ take $\varepsilon\colon G \times G \to \mathbb{R}$ defined by

$$\varepsilon(g_i, g_j) = \begin{cases} \eta((i\ j)), & \text{if } i \neq j, \\ 0, & \text{if } i = j, i = 0, \text{ or } j = 0. \end{cases}$$

The map ε is admissible by definition, and clearly $\varepsilon(g_i, g_j) = \varepsilon(g_j, g_i)$ for any $i, j \in I_0$. So ε will be a graded contraction if we check that $\langle g_i, g_j, g_k \rangle = G$ implies

$$\varepsilon(g_i, g_j, g_k) = \varepsilon(g_k, g_i, g_j).$$

But in this case $i, j, k \in I$ is a generating triplet and we know $\eta_{ijk} = \eta_{jki}$. Now simply note that

$$\varepsilon(g_i, g_j, g_k) = \varepsilon(g_i, g_j + g_k)\varepsilon(g_j, g_k) = \varepsilon(g_i, g_{j*k})\varepsilon(g_j, g_k) = \eta_{i,j*k}\eta_{j,k} = \eta_{ijk}.$$

Finally observe that the maps $\mathcal{A}' \to \mathcal{A}$, $\varepsilon \mapsto \eta$, and $\mathcal{A} \to \mathcal{A}'$, $\eta \mapsto \varepsilon$, are inverse one of each other.

□

Now, as there is no ambiguity, we also refer to the elements in $\mathcal{A}$ as $\varepsilon\colon X \to \mathbb{R}$. The above lemma says that we only have to worry about the $\binom{7}{2} = 21$ possible images ε_{ij} of the elements in X. We will observe that the set of elements of X with nonzero image, cannot be an arbitrary subset of X.

Definition A.3:

If $\varepsilon \in \mathcal{A}$ we call the following set

$$T^{\varepsilon} := \{(i\ j) \in I \mid \varepsilon_{ij} \neq 0\} \subset X,$$

the **support** *of $\varepsilon\colon X \to \mathbb{R}$.*

These subsets will be *nice* in a very concrete sense: in the sense that they will have some properties related to the subsets of X.

Definition A.4:

1. *If $i, j, k \in I$ is a generating triplet, then we denote by*

$$P(i\ j\ k) := \{(i\ j), (i\ k), (i\ j * k), (j\ k), (j\ i * k), (k\ i * j)\} \subset X.$$

(Clearly $P(i\ j\ k)$ does not depend on the order of the three elements i, j and k.)

2. *A subset $T \subseteq X$ is* **nice** *if for all generative $i, j, k \in I$ we have that $(i\ j), (i * j\ k) \in T$ implies that $P(i\ j\ k) \subseteq T$.*

Proposition A.5:

1. *If $\varepsilon \in \mathcal{A}$ then the support T^ε is nice.*

2. *Suppose T is a nice subset of X. Consider the induced map $\varepsilon^T : X \to \mathbb{R}$ defined as*

$$\varepsilon^T(t) := \begin{cases} 1, & \text{if } t \in T, \\ 0, & \text{if } t \notin T. \end{cases}$$

Then we have that $\varepsilon^T \in \mathcal{A}$ and $T = T^{\varepsilon^T}$.

Thus, classifying nice sets give us not only the possible supports but also there is at least one graded contraction up to equivalence corresponding to such support.

Proof. We begin with (1). Let $\varepsilon \in \mathcal{A}$, let $i, j, k \in I$ be generative and suppose $(i\ j), (k\ i * j) \in T^\varepsilon$. Since $\varepsilon_{k,i*j} \neq 0 \neq \varepsilon_{ij}$ we must have that $\varepsilon_{kij} = \varepsilon_{k,i*j}\varepsilon_{ij} \neq 0$. Since $\varepsilon \in \mathcal{A}$, so $\varepsilon_{ijk} = \varepsilon_{kij}$ what implies that $\varepsilon_{ijk} \neq 0$ too. But $\varepsilon_{ijk} = \varepsilon_{i,j*k}\varepsilon_{jk} \neq 0$ means that $(i\ j * k), (j\ k) \in T^\varepsilon$. Similarly, $\varepsilon_{jki} \neq 0$ and so $(j\ k * i), (i\ k) \in T^\varepsilon$. Thus, $P(i\ j\ k) \subseteq T^\varepsilon$, which allows us to conclude that T^ε is nice.

Now we move on to (2). Let T be nice. We aim to show that $\varepsilon^T_{ijk} = \varepsilon^T_{jki}$ for all generative $i, j, k \in I$. We consider the possible values of ε^T_{ijk} and ε^T_{jki}. If $P(i\ j\ k) \subset T$, then $(j\ k), (i\ j * k) \in T$, so that $\varepsilon^T_{i,j*k} = \varepsilon^T_{jk} = 1$ and hence $\varepsilon^T_{ijk} = \varepsilon^T_{i,j*k}\varepsilon^T_{jk} = 1$. But also $(k\ i), (j\ k * i) \in T$, so that $\varepsilon^T_{jki} = 1$. In particular $\varepsilon^T_{ijk} = \varepsilon^T_{jki} = 1$. Let us check that, if $P(i\ j\ k)$ is not contained in T, then $\varepsilon^T_{ijk} = \varepsilon^T_{jki} = 0$. If $\varepsilon^T_{ijk} \neq 0$, then both $\varepsilon^T_{i,j*k}$ and ε^T_{jk} are nonzero, so that both $(i\ j * k)$ and $(j\ k)$ belong to T. As we are assuming that T is nice, then $P(i\ j\ k) \subset T$, a contradiction with the situation considered.

□

So, we have a first important problem to be solved: finding the nice sets. Perhaps afterwards there will be (or not) two not equivalent admissible graded contractions with the same support, so that both classification problems are not exactly equivalent ones. In any case it is clear that the classification of the admissible graded contractions has to begin by classifying nice subsets of X. A first observation is that collineations of X give equivalent graded contractions. We are going to make these concepts precise.

Definition A.6:

1. *A* **line** *in I is a subset $L \subset I$ of the form $L = \{i, j, i * j\}$, where $i \neq j$. We denote by L_{ij} the line containing i and j (if $i \neq j$). (This means that $L_{ij} = L_{j,i*j} = ...$ and so on.)*

2. *A collineation of I (also called collinear permutation) is a bijective map $\sigma \colon I \to I$ such that*

$$\sigma(i) * \sigma(j) = \sigma(i * j).$$

Observe that $\sigma(L_{ij}) = L_{\sigma(i)\sigma(j)}$, so these maps are those ones sending lines to lines. We will denote the set of collineations on I by $S_*(I)$.

Lemma A.7: *The set of collinear permutations, $S_*(I)$, is a subgroup of the symmetric group on I.*

Proof. Clearly, the identity permutation is a collinear permutation. Moreover, if $\sigma \in S_*(I)$, then for $i, j \in I$,

$$\begin{aligned}\sigma^{-1}(i) * \sigma^{-1}(j) &= \sigma^{-1}\sigma(\sigma^{-1}(i) * \sigma^{-1}(j)) = \sigma^{-1}(\sigma\sigma^{-1}(i) * \sigma\sigma^{-1}(j)) \\ &= \sigma^{-1}(i * j).\end{aligned}$$

□

Remark A.8: *If T is a nice set and σ is a collineation, then $\sigma(T)$ is also nice.*

Proof. Let $i, j, k \in I$ be generative. Suppose $(i\ j), (k\ i*j) \in \sigma(T)$. We know that $\sigma^{-1}(i\ j), \sigma^{-1}(k\ i*j) \in T$ implies that $\sigma^{-1}(P(i\ j\ k)) \subseteq T$ which is equivalent to $\sigma(T)$ being nice. This follows because $S_*(I)$ is a subgroup of the symmetric group. □

Lemma A.9: *If $\sigma \in S_*(I)$ and $\varepsilon \colon X \to \mathbb{R}$ belongs to $\mathcal{A}$, then $\varepsilon_\sigma \in \mathcal{A}$ for the map $\varepsilon_\sigma \colon X \to \mathbb{R}$ defined by $\varepsilon_\sigma((i\ j)) =: \varepsilon(\sigma(i)\ \sigma(j))$. The corresponding graded contraction ε_σ is equivalent to ε. Furthermore, the support satisfies $T^{\varepsilon_\sigma} = \sigma(T^\varepsilon)$.*

The proof is not difficult taking into account the following ideas. If σ is a collineation, the map $\tilde{\sigma} \colon G \to G$ defined by $\tilde{\sigma}(g_i) = g_{\sigma(i)}$ if $i \in I$ and $\tilde{\sigma}(g_0) = g_0$, is a group automorphism. Indeed, if $i \neq j$, then

$$\tilde{\sigma}(g_i + g_j) = \tilde{\sigma}(g_{i*j}) = g_{\sigma(i*j)} = g_{\sigma(i)*\sigma(j)} = g_{\sigma(i)} + g_{\sigma(j)} = \tilde{\sigma}(g_i) + \tilde{\sigma}(g_j).$$

Furthermore, if $i = j$, then $\tilde{\sigma}(g_i + g_i) = \tilde{\sigma}(g_0) = g_0 = \tilde{\sigma}(g_i) + \tilde{\sigma}(g_i)$. Now we consider the properties of the Weyl group of the $\mathbb{Z}_2^3$-grading on the octonions, studied in [5], according to which there is an automorphism $f\colon \mathbb{O} \to \mathbb{O}$ with $f(\mathbb{O}_g) = \mathbb{O}_{\sigma(g)}$ for all $g \in G$; which easily implies that the induced map $\tilde{f}\colon \mathfrak{g}_2 \to \mathfrak{g}_2$ given by $\tilde{f}(d) = f^{-1}df$ is also an automorphism of graded Lie algebras with $\tilde{f}((\mathfrak{g}_2)_g) \subseteq (\mathfrak{g}_2)_{\sigma(g)}$ for all $g \in G$.

What the above lemmas say is that we should try to classify nice sets up to collinearity, since we are finally interested in the Lie algebra classification. Another to note is that, if $\varepsilon, \varepsilon' \in \mathcal{A}$ have different supports up to collineations, that is $T^{\varepsilon} \neq \sigma(T^{\varepsilon'})$ for all $\sigma \in S_*(I)$, then ε and ε' can not be equivalent.

In continuing our task of finding the nice sets of X up to collineations, let us consider some distinguished sets which will also prove to be nice sets.

Definition A.10: *Suppose L is some line in I and $m \in I$.*

1. $X_L := \{(i\ j) \,|\, i, j \in L\}$.
2. $X_{L^C} := \{(i\ j) \,|\, i, j \in L^C := I - L\}$.
3. $X_{(m)} := \{(m\ i) \,|\, i \neq m\}$.
4. $X^{(m)} := \{(i\ j) \,|\, i * j = m\}$.
5. $T(i\ j\ k) := P(i\ j\ k) \cup P(i\ j\ i*k) \cup P(i\ k\ i*j) \cup P(i\ i*j\ i*k)$.

For instance, let us describe them in concrete cases.

Example A.11:

1. $X_{L_{12}} = \{(1\ 2), (1\ 5), (2\ 5)\}$.
2. $X_{L_{12}^C} = \{(3\ 4), (3\ 6), (3\ 7), (4\ 6), (4\ 7), (6\ 7)\}$.
3. $X_{(1)} = \{(1\ 2), (1\ 3), (1\ 4), (1\ 5), (1\ 6), (1\ 7)\}$.
4. $X^{(1)} = \{(2\ 5), (3\ 6), (4\ 7)\}$.
5. $T(1\ 2\ 3) = \{(1\ 2), (1\ 3), (1\ 4), (1\ 5), (1\ 6), (1\ 7), (2\ 3), (2\ 6), (3\ 5), (5\ 6)\}$.

Remark A.12: *For $i, j, k \in I$ we have $P(i\ j\ k) = P(i\ k\ j) = P(j\ i\ k)$. This follows from the definition of $P(i\ j\ k)$.*

To handle these subsets, it is convenient to have a preliminary result.

Remark A.13: *Suppose T is a nice set which does not contain any subset of type $P(l\ m\ n)$. Then, for each $(i\ j) \in T$, if $(k\ i*j) \in T$ then $k \in \{i, j\}$.*

Proof. If T is empty there is nothing to prove. Let $(i\ j) \in T$. Suppose $(i*j\ k) \in T$ for some $k \in I$ (necessarily $k \neq i*j$). If i, j, k are generative, then $(i\ j), (k\ i{*}j) \in T$ implies that $P(i\ j\ k) \subseteq T$, a contradiction. Otherwise, i, j, k are not generative, so that $k \in \{i, j, i*j\}$. But $k \neq i*j$, so that k is either i or j. □

As we previously indicated, there are many nice sets of importance.

Proposition A.14: *Let $i, j, k \in I$ and L be some line. The following subsets of X are nice:*

1. X,
2. X_L,
3. X_{L^C},
4. $X_{(i)}$,
5. $X^{(i)}$,
6. $X - X_{L^C}$,
7. $P(i\ j\ k)$,
8. $T(i\ j\ k)$,
9. *Any subset of* $X_L,\ X_{L^C}, X_{(i)}, X^{(i)}$.

Proof.

1. It is clear that X is nice.
2. Suppose $(i\ j), (k\ i*j) \in X_L$. Then $i, j, k, i*j \in L$, therefore, without loss of generality, we may assume $k = i$. Therefore there are no generative triplets $i, j, k \in I$ such that each $(i\ j), (k\ i*j) \in X_L$. Thus, the result holds vacuously.
3. Suppose $(i\ j), (k\ i*j) \in X_{L^C}$. Then $i, j, k, i*j \in L^C$. It suffices to show that $j*k,\ i*k \in L^C$. If $j*k \in L$ then $j, k \in L$, a contradiction. Similarly, $i*k \in L^C$.
4. Let us show that $X_{(l)}$ is nice for any $l \in I$. Let us suppose that $(i\ j), (k\ i*j) \in X_{(l)}$. Then $(i\ j) \cap (k\ i*j) = \{l\}$. Then we may assume, without loss of generality, that $k = i$. Therefore there are no generative triplets $i, j, k \in I$ such that each $(i\ j), (k\ i*j) \in X_L$. Thus, the result holds vacuously.
5. Let us show that $X^{(l)}$ is nice for any $l \in I$. Suppose $(i\ j), (k\ i*j) \in X^{(l)}$. Then $i*j = l = k*(i*j)$. This implies that $l = k*l$, a contradiction.
6. We note that $X - X_{L^C} ;= \{(m\ n) \,|\, m \in L \text{ or } n \in L\}$. Suppose that $(i\ j), (k\ i*j) \in X - X_{L^C}$. This implies that $i \in L$ or $j \in L$, and it implies that $k \in L$ or $i*j \in L$. We may assume, without loss of generality, that $i \in L$. This immediately implies that $(i\ k), (i\ j*k) \in X - X_{L^C}$. Now we must consider two cases: $k \in L$ and $i*j \in L$. If $k \in L$, then

$(j\ k) \in X - X_{L^C}$. Moreover, since both $i, k \in L$ we know that $i * k \in L$ and so $(j\ i * k) \in X - X_{L^C}$. This covers the first case. On the other hand, if $i * j \in L$, then $j \in L$ and so $(j\ k), (j\ i * k) \in X - X_{L^C}$.

7. It will suffice to show that $P(1\ 2\ 3)$ is nice. Let us suppose that $(i\ j), (k\ i * j) \in P(1\ 2\ 3)$. The possible pairs of elements of X which match up with $(i\ j)$ and $(k\ i * j)$ are: $(1\ 2)$ and $(3\ 5)$, $(1\ 3)$ and $(2\ 6)$, or $(2\ 3)$ and $(1\ 7)$. In all of those possible cases $P(i\ j\ k) = P(1\ 2\ 3)$ means that $P(1\ 2\ 3)$ is nice.

8. We note that $T(i\ j\ k) := \{(i\ j), (i\ k), (i\ j * k), (i\ i * j), (i\ i * k), (i\ i * j * k), (j\ k), (j\ i * k), (k\ i * j), (i * j\ i * k)\}$. Take generative $x, y, z \in I$. Suppose $(x\ y), (z\ x * y) \in T(i\ j\ k)$. Since $P(x\ y\ z) = \{(x\ y), (x\ z), (x\ y * z), (y\ z), (y\ x * z), (z\ x * y)\}$, we need to show that the remaining elements are contained in $T(i\ j\ k)$. We note that x is contained in three elements of $P(x\ y\ z)$. Since the only elements of I which appear in at least three elements of $T(i\ j\ k)$ are $\{i, j, k\}$ we conclude that $x \in \{i, j, k\}$. Similarly, we find $y, z \in \{i, j, k\}$. From here we conclude that

$$P(x\ y\ z) = P(\sigma(i)\ \sigma(j)\ \sigma(k)) = P(i\ j\ k) \subseteq T(i\ j\ k),$$

where σ is some permutation on $\{i, j, k\}$.

9. For subsets of $X_{L^C}, X_{(i)}$, and $X^{(i)}$ this follows from Remark A.13 because none of these sets contain elements of the form $(i\ j)$ and $(k\ i * j)$, where $i, j, k \in I$ are generative. Suppose $T \subset X_L$. It suffices to prove the result for $T \subset X_{L_{12}}$. If T is empty the result is clear. If T contains exactly one element, then clearly T is nice. If T contains exactly two elements we may assume, without loss of generality, that $T = \{(1\ 2), (2\ 5)\}$. Since $1, 2, 5$ is not a generative triplet the result holds vacuously.

□

Our purpose is to prove that Proposition A.14 contains the list of all nice sets. This is a difficult proof so that we will divide in pieces. First we consider the nice sets which are sufficiently big.

Proposition A.15: *Suppose T is nice and there exist generative $i, j, k \in I$, such that $P(i\ j\ k) \subseteq T$. Then there is some collinear permutation σ such that*

$$\sigma(T) \in \{X,\ X - X_{L_{12}^C},\ P(1\ 2\ 3),\ T(1\ 2\ 3)\}.$$

Proof. We divide the proof into six claims. Assume $P(1\ 2\ 3) \subset T \subset X$.

Claim 1: If $X - X_{L_{12}^C} \subset T$, then $T = X$.

From the hypothesis, T contains some element from $X - X_{L_{12}^C}$.

$$X_{L_{12}^C} = \{(3\ 4), (3\ 6), (3\ 7), (4\ 6), (4\ 7), (6\ 7)\}.$$

If $(3\ 4) \in T$, then $(3\ 4), (1\ 5) \in T$ implies that $P(3\ 4\ 1) \subseteq T$. Specifically, $(3\ 7), (4\ 6) \in T$. We also find that, $(4\ 6), (7\ 2) \in T$ implies that $P(4\ 6\ 7) \subseteq T$. In particular, $(4\ 7), (6\ 7) \in T$. Furthermore, we can see that $(6\ 7), (3\ 5) \in T$ implies that $P(6\ 7\ 3) \subseteq T$. This means that $(3\ 7) \in T$ and we conclude $T = X$.

If, instead of $(3\ 4)$, we assumed that some other element of $X - X_{L_{12}^C}$ were in T, then by a similar argument we would find $T = X$.

Claim 2: If $P(1\ 2\ 3) \subset T \subset X$ and $P(1\ 2\ 4) \subset T$, then $T = X - X_{L_{12}^C}$. We have that $(3\ 5), (1\ 4) \in T$ implies that $P(3\ 5\ 1) \subseteq T$. Specifically, $(1\ 5), (5\ 6) \in T$. We also find that $(1\ 5), (7\ 2) \in T$ implies that $P(1\ 5\ 7) \subseteq T$. In particular, $(5\ 7) \in T$. Finally, $(5\ 7), (2\ 6) \in T$ implies that $P(5\ 7\ 2) \subseteq T$ and this means that $(2\ 5) \in T$. Altogether, we then conclude that $X - X_{L_{12}^C} \subseteq T$. Claim 1 now yields that $T = X - X_{L_{12}^C}$.

Claim 3: If $P(1\ 2\ 3) \subseteq T$ and $P(1\ 2\ 6) \subseteq T$, then $T(1\ 2\ 3) \subseteq T$.

We note first that

$$T(1\,2\,3) := \{(1\,2), (1\,3), (1\,7), (2\,3), (2\,6), (3\,5), (1\,6), (1\,4), (5\,6), (1\,5)\}.$$

Now, $(3\ 5), (1\ 4) \in T$ implies that $P(3\ 5\ 1) \subseteq T$, and so $(1\ 5) \in T$ which allows us to conclude that $T(1\ 2\ 3) \subseteq T$.

Claim 4: If $T(1\ 2\ 3) \subset T$, then $T \in \{X - X_{L^C} \mid L \in \{L_{12},\ L_{13},\ L_{23}\}\}$.

From the hypothesis, $T - T(1\ 2\ 3)$ is non-empty. The following set consists of all the elements which could be in $T - T(1\ 2\ 3)$:

$$\{(2\ 4), (2\ 5), (2\ 7), (3\ 4), (3\ 6), (3\ 7), (4\ 5), (4\ 6), (4\ 7), (5\ 7), (6\ 7)\}.$$

If $(2\,4) \in T$, then $(2\,4), (1\,6) \in T$ implies that $P(1\,2\,4) \in T$. Specifically, $(4\ 5) \in T$. Now, $(4\ 5), (1\ 3) \in T$ implies that $P(1\ 4\ 5) \subseteq T$. In particular, $(5\ 7) \in T$. This, in turn, means that $(5\ 7), (1\ 6) \in T$ implies that $P(1\ 5\ 7) \subseteq T$ which means that $(2\ 7), (4\ 5) \in T$. Finally, this means that $(2\,7), (5\,3) \in T$ implies that $P(2\,5\,7) \subseteq T$. We conclude that $(2\ 5) \in T$ and therefore $P(1\ 2\ 4) \subseteq T$ implies that $X - X_{L_{12}^C} \subseteq T$.

If $(2\ 5) \in T$, then $(2\ 5), (4\ 1) \in T$ implies that $P(2\ 4\ 5) \subseteq T$. Specifically, $(2\ 4), (4\ 5) \in T$. Therefore, $(2\ 4), (1\ 6) \in T$ implies that $P(1\ 2\ 4) \subseteq T$. Hence, $X - X_{L_{12}^C} \subseteq T$.

If $(2\ 7) \in T$, then $(2\ 7), (1\ 3) \in T$ implies that $P(1\ 2\ 7) \subseteq T$. In particular, $(2\ 4) \in T$. Similar to previous case, we can now see that $X - X_{L_{12}^C} \subseteq T$.

If $(3\ 4) \in T$, then $(3\ 4), (1\ 5) \in T$ implies that $P(1\ 3\ 4) \subseteq T$. This means that $(3\ 7), (4\ 6) \in T$. Now, $(1\ 3), (2\ 6) \in T$ implies that $P(1\ 3\ 2) \subseteq T$. Furthermore, $(1\ 6), ((7\ 3) \in T$ implies that $P(1\ 6\ 7) \subseteq T$. Specifically, $(6\ 7) \in T$. Finally, $(3\ 5), (4\ 6) \in T$ implies that $P(3\ 5\ 6) \subset T$ and so $(3\ 6) \in T$. We find that $X - X_{L_{13}^C} \subseteq T$.

If $(3\ 6) \in T$, then $(3\ 6), (2\ 1) \in T$ implies that $P(2\ 3\ 6) \subseteq T$. Hence, $(3\ 4) \in T$. The previous case covers this.

If $(3\ 7) \in T$, then $(3\ 7), (1\ 2) \in T$ implies that $P(1\ 3\ 7) \subseteq T$. Therefore $(3\ 4) \in T$. We have covered this case.

If $(4\ 5) \in T$, then $(4\ 5), (1\ 3) \in T$ implies that $P(1\ 4\ 5) \subseteq T$. In particular, $(2\ 4), (5\ 7) \in T$. Now, $(2\ 4), (1\ 6) \in T$ implies that $P(1\ 2\ 4) \subseteq T$. This means that $(2\ 7) \in T$. We have covered this case earlier.

If $(4\ 6) \in T$, then $(4\ 6), (1\ 2) \in T$ implies that $P(1\ 4\ 6) \subseteq T$. Hence, $(3\ 4) \in T$ and so we are in a previously covered case.

If $(4\ 7) \in T$, then $(4\ 7), (2\ 1) \in T$ implies that $P(2\ 4\ 7) \in T$. In particular, $(3\ 4) \in T$, which leaves us in a case we have already covered.

If $(5\ 7) \in T$, then $(5\ 7), (1\ 6) \in T$ implies that $P(1\ 5\ 7) \subseteq T$. This means that $(2\ 7) \in T$ and so we return to a case we have already covered.

If $(6\ 7) \in T$, then $(6\ 7), (1\ 5) \in T$ implies that $P(1\ 6\ 7) \subseteq T$. Therefore, $(4\ 7) \in T$ and we have covered this case.

Claim 5: If $P(1\ 2\ 3) \cup P(i\ j\ k) \subseteq T$ for some generative i, j, k different from $1, 2, 3$, then claim 2 or claim 3 applies.

If $\{1, 2, 3\} \cap \{i, j, k\}$ contains two elements, then (up to equivalence) there are three cases. In the first case $P(i\ j\ k) = P(1\ 2\ 4)$, and then claim 2 is applicable. In the second case $P(i\ j\ k) = P(1\ 2\ 6)$, and then claim 3 is applicable. In the third case $P(i\ j\ k) = P(1\ 2\ 7)$. Then we have that $P(2\ 1\ 7) \cup P(2\ 1\ 3) \subseteq T$ and since $7 = 2 * 3$, claim 3 applies (up to collinearity).

If $\{1, 2, 3\} \cap \{i, j, k\}$ contains one element we may assume, without loss of generality, that $\{i, j, k\}$ is equal to one of $\{1, 4, 5\}$ or $\{1, 4, 6\}$. If $\{i, j, k\} = \{1, 4, 5\}$, then $P(1\ 2\ 3) \cup P(1\ 4\ 5) \subseteq T$ and $(1\ 2), (4\ 5) \in T$ implies that $P(1\ 2\ 4) \subseteq T$. Hence, claim 2 is applicable. If, instead, $\{i, j, k\} = \{1, 4, 6\}$, then $P(1\ 2\ 3) \cup P(1\ 4\ 6) \subseteq T$ and $(2\ 6), (1\ 4) \in T$ implies that $P(1\ 2\ 6) \subseteq T$. Thus, claim 3 would be applicable.

Finally, if $\{1, 2, 3\} \cap \{i, j, k\}$ is empty, then we may assume, without loss of generality, that $\{i, j, k\} = \{4, 5, 6\}$. We have that $P(1\ 2\ 3) \cup$

$P(4\ 5\ 6) \subseteq T$. Moreover, $(1\ 2), (4\ 5) \in T$ implies that $P(1\ 2\ 4) \subseteq T$, and so claim 2 applies.

Claim 6: If $P(1\ 2\ 3) \subset T$, then there exist generative $i, j, k \in I$ with $\{i, j, k\} \neq \{1, 2, 3\}$ such that $P(i\ j\ k) \subset T$.

We note that since $P(1\ 2\ 3) \subset T$ there must be some element of $X - P(1\ 2\ 3)$ in T. If $(i\ 5) \in T$ for $i \neq 3$, then $P(1\ 2\ i) \subset T$. If $(i\ 6) \in T$ for $i \neq 2$, then $P(1\ 3\ i) \subset T$. If $(i\ 4) \in T$, then $P(1\ 7\ i) \subset T$. If $(i\ 7) \in T$, for $i \neq 1$, then $P(2\ 3\ i) \subset T$.

The only element of $X - P(1\ 2\ 3)$ not yet taken into account is $(3\ 7)$. If $(3\ 7) \in T$, then $(3\ 7), (1\ 2) \in T$ implies that $P(1\ 3\ 7) \subset T$.

□

And now we deal with the nice sets with few elements.

Proposition A.16: *Suppose $T \neq \emptyset$ is nice and there is no $P(i\ j\ k) \subseteq T$.*

1. *If $X_L \subseteq T$ for some line L, then $X_L = T$.*
2. *If $T \not\subseteq X_L$ for any line L and $T \not\subseteq X_{(l)}$ for any $l \in I$, then $(i\ j) \in T$ implies that $(i * j, m) \notin T$ for all $m \in I$.*
3. *If $T \not\subseteq X_L$ for any line L and $T \not\subseteq X_{(l)}$ for any $l \in I$, then $T \subseteq X_{L^C}$ for some line L or $T \subseteq X^{(l)}$ for some $l \in I$.*

Proof. We note that if there is no $P(i\ j\ k) \subseteq T$, then the following property, denoted (V1), holds for T: if there exist $i, j, m \in I$ such that $(i\ j), (m\ i{*}j) \in T$, then $m \in \{i, j\}$.

1. We assume $X_{L_{ij}} \subseteq T$. Suppose there is some $t = (k\ l) \in T \cap (X - X_L)$, with $k \in L^C$. First we consider the possibility that $l \in L$. If $l = i * j$, then $(i\ j), (k\ i{*}j) \in T$ implies that $P(i\ j\ k) \subseteq T$, a contradiction. If $l = i$, then $(j\ i{*}j), (i\ k) \in T$ implies that $P(j\ i{*}j\ k) \subseteq T$, a contradiction. If $l = j$, then $(i\ i * j), (j\ k) \in T$ implies that $P(i\ i * j\ k) \subseteq T$, a contradiction. Therefore $l \in L^C$.

   Since $k, l \in L^C$ we know that $k * l \in L$. If $k * l = i$, then $(k\ l), (j\ i) = (j\ k{*}l) \in T$ implies that $P(k\ l\ j) \subseteq T$, a contradiction. If $k{*}l = j$, then $(k\ l), (i\ j) = (i\ k{*}l) \in T$ implies that $P(k\ l\ i) \subseteq T$, a contradiction. If $k{*}l = i{*}j$, then $(k\ l), (i\ i{*}j) = (i\ k{*}l) \in T$ implies that $P(k\ l\ i) \subseteq T$, a contradiction. Hence, we conclude $T = X_{L_{ij}}$.

2. It suffices to assume $(1\ 2) \in T$ and to show that there is no $m \in I$ such that $(5\ m) \in T$. Suppose $(5\ m) \in T$ for some $m \in I$. From Remark A.13 we find that $m \in \{1, 2\}$. If $(1\ 5), (2\ 5) \in T$ then $X_{L_{12}} \subseteq T$ and then part (1) yields that $T = X_{L_{12}}$, a contradiction. Therefore we may assume, without loss of generality, that $(1\ 5) \in T$ and $(2\ 5) \notin T$. By

hypothesis $T \not\subseteq X_{L_{12}}$. Hence, there exists $k \in I$ and $l \in L_{12}^C$ such that $(k\ l) \in T$. We will now consider the possible cases for k.

If $k \in L_{12}^C$, then $k * l \in L_{12}$. Then since $(k\ l), (1\ 2), (1\ 5) \in T$ and $k * l \in L_{12}$ means that $P(k\ l\ n) \subset T$ where $n \in \{1, 2, 5\}$ (depending on whether $k * l$ is $1, 2$, or 5).

If $k \in L_{12}$, then $k * l \in L_{12}^C$. In this case $k =\in \{1, 2, 5\}$ If $k = 2$, then $(1\ 5), (2\ l) = (k\ l) \in T$ implies that $P(1\ 5\ l) \subseteq P$, a contradiction. If $k = 5$, then $(1\ 2), (5\ l) = (k\ l) \in T$ implies that $P(1\ 2\ 5) \in T$, a contradiction.

If $k = 1$, then $(1\ j), (1\ 2), (1\ 5) \in T$. By hypothesis, $T \not\subset X_{(1)}$. Thus there is $(a\ b) \in T$, with $a, b \in I - \{1\}$. If $a = 2$, then $(2\ b), (1\ 5) \in T$ implies that $P(1\ 5\ b) \subseteq T$, a contradiction. If $a = 5$, then $(1\ 2), (5\ b) \in T$ implies $P(1\ 2\ b) \subseteq T$, a contradiction. Therefore, the only remaining possibility is that both $a, b \in L_{12}^C$. This means that $a * b \in L_{12}$. Moreover, this means that $P(a\ b\ n) \subseteq T$, for some $n \in \{1, 2, 5\}$ depending on exactly which values a and b take.

3. We may assume, without loss of generality, that $(1\ 2) \in T$. From (2) we know that for all $m \in I$, $(5\ m) \notin T$. By hypothesis, $T \not\subseteq X_{(1)}, X_{(2)}$. Therefore, there are $j, k \in I - \{1, 2\}$ such that $(j\ k) \in T$, with $j < k$.

   If $j \neq 2$, then $j \in \{3, 4, 6, 7\}$ and $k \in \{4, 6, 7\}$. We may assume, without loss of generality, that $j = 3$. Part (2) yields that we cannot have $3 * k = 1$ or $3 * k = 2$. Therefore, $k \notin \{6, 7\}$. Hence, $k = 4$. Now we know that $(1\ 2), (3\ 4) \in T$.

   Since $(1\ 2) \in T$, part (2) implies that $(4\ 7), (3\ 6), (4\ 6), (3\ 7) \notin T$. Since $(3\ 4) \in T$, part (2) implies that $(1\ 6), (2\ 7), (2\ 6), (1\ 7) \notin T$. We also know that $(5\ m) \notin T$ for all $m \in I$. This means that the following elements could be in T (in addition to $(1\ 2)$ and $(3\ 4)$:

   $$\{(1\ 3), (1\ 4), (2\ 3), (2\ 4), (6\ 7)\}.$$

   However, the following elements are the same up to collinearity: $(1\ 3)$, $(1\ 4)$, $(2\ 3)$, and $(2\ 4)$. This means we need only consider two cases: $(1\ 3) \in T$, and $(6\ 7) \in T$.

   If $(1\ 3) \in T$, then from part (2) we find that elements of the form $(6\ m)$ cannot be in T. Hence, $T = \{(1\ 2), (1\ 3), (3\ 4)\} \subseteq X_{L_{67}^C}$.

   If $(6\ 7) \in T$, then part (2) implies that $(2\ 3), (1\ 4), (1\ 3), (2\ 4) \notin T$. Therefore, $T = \{(1\ 2), (3\ 4), (6\ 7)\} \subseteq X^{(5)}$.

   Now, let us consider the case where $j = 2$. This implies $k \in \{3, 4, 6, 7\}$. We may assume, without loss of generality, that $k = 3$. However since $(1\ 2), (2\ 3) \in T$, part (2) yields that $(3\ 6), (4\ 6), (1\ 6) \notin T$. Furthermore, this also implies that elements of the form $(5\ m)$ or $(7\ m)$ are not in

T. By hypothesis $T \not\subseteq X_{(2)}$, and so there exist $x, y \in I - \{2, 5, 7\}$, such that $(x\ y) \in T$. Keeping mind the restrictions we have seen on elements of T, we are left with two cases: $(x\ y) = (1\ 3)$ or $(x\ y) = (1\ 4)$.

If $(x\ y) = (1\ 3)$, then part (2) implies that elements of the form $(6\ m)$ are not in T. Therefore, the following elements could also be in T : $(1\ 4), (2\ 4), (3\ 4)$. Thus, $T \subseteq X_{L_{67}^C}$.

If $(x\ y) = (1\ 4)$, then part (2) implies that $(2\ 6) \notin T$. The following elements could also be in T : $(1\ 3), (2\ 4), (3\ 4)$. Hence, we find that $T \subseteq X_{L_{67}^C}$.

□

A.2 Classifying Nice Sets

We are now in a position to achieve one of our main aims: classifying nice sets.

Theorem A.17: *Suppose T is a nice set. We can classify the possible values for T, up to collineation, according to the cardinality of T.*

$|T| = 0$: *Then* $T = \emptyset$.

$|T| = 1$: *Then* $T = \{(1\ 2)\}$.

$|T| = 2$: *Then T is one of the following sets:* $\{(1\ 2), (1\ 5)\}$, $\{(3\ 4), (3\ 6)\}$, $\{(3\ 4), (6\ 7)\}$.

$|T| = 3$: *Then T is one of the following sets:* $X_{L_{34}}$, $X^{(5)}$, $\{(3\ 4), (3\ 6), (3\ 7)\}$, $\{(3\ 4), (3\ 6), (4\ 6)\}$, $\{(3\ 4), (3\ 6), (4\ 7)\}$, $\{(3\ 4), (3\ 5), (3\ 6)$.

$|T| = 4$: *Then T is one of the following sets:* $\{(3\ 4), (3\ 6), (3\ 7), (4\ 6)\}$, $\{(3\ 4), (3\ 6), (4\ 7), (6\ 7)\}$, $\{(1\ 2), (1\ 3), (1\ 4), (1\ 5)\}$.

$|T| = 5$: *Then T is one of the following sets:* $\{(3\ 4), (3\ 6), (3\ 7), (4\ 6), (4\ 7)\}$, $\{(1\ 2), (1\ 3), (1\ 4), (1\ 5), (1\ 6)\}$.

$|T| = 6$: *Then T is one of the following sets:* $X_{L_{12}^C}$, $P(1\ 2\ 3)$, $X_{(1)}$.

$|T| = 10$: *Then* $T = T(1\ 2\ 3)$.

$|T| = 15$: *Then* $T = X - X_{L_{12}^C}$.

$|T| = 21$: *Then* $T = X$.

Proof. We split the proof up by the cardinality of T. If $|T| = 0$ the result is trivial. If $|T| = 1$, then we can easily find $\sigma \in S_*(I)$ such that $\sigma(T) = \{(1\ 2)\}$.

We assume $|T| = 2$. From Proposition A.16 we find that T is a subset of X_L, X_{L^C}, $X_{(l)}$, or $X^{(l)}$. If $T \subset X_L$, then, up to collineation, $T = \{(1\ 2), (1\ 5)\} \subset X_{L_{12}}$. This is the unique possibility because this is collinear to $\{(i\ j), (i\ i * j)\}$ for any $i \neq j \in I$. Let us consider nice subsets of $X_{L_{12}^C}$. The two possible sets are $\{(3\ 4), (3\ 6)\} \subset X_{(3)}$ and $\{(3\ 4), (6\ 7)\} \subset X^{(5)}$. These are the only two options because the following three sets $\{(3\,4), (3\,7)\}$, $\{(3\,4), (4\,6)\}$, $\{(3\,4), (4\,7)\}$ are collinear to $\{(34), (3\,6)\}$. Now we consider the nice subsets of $X_{(1)}$. We may assume that $(1\,2) \in T$. Then, up to collineation, we find two nice subsets: $\{(1\,2), (1\,3)\}$ and $\{(1\,2), (1\,5)\}$. Finally, from subsets of $X^{(1)}$ we find $\{(2\ 5), (3\ 6)\}$, to which the remaining subsets of $X^{(1)}$ are collinear.

We can also see that $\{(3\ 4), (3\ 6)\} \sim \{(1\ 2), (1\ 3)\}$ and $\{(3\ 4), (6\ 7)\} \sim \{(2\ 5), (3\ 6)\}$. The result follows.

We assume now that $|T| = 3$. From Proposition A.16 we find that T is a subset of X_L, X_{L^C}, $X_{(l)}$, or $X^{(l)}$. If $T = X_L$ we may assume that $T = X_{L_{34}}$. Similarly, if $T = X^{(l)}$, we may assume $T = X^{(5)}$.

We now consider possible subsets of $X_{L_{12}^C}$. The sets $\{(3\ 4), (3\ 6), (3\ 7)\}$, $\{(3\ 4), (3\ 6), (4\ 6)\}$, and $\{(3\ 4), (3\ 6), (4\ 7)\}$ are not collinear. The set $\{(3\,4), (4\,6), (4\,7)\}$ is collinear to $\{(3\,4), (3\,6), (3\,7)\}$, while $\{(3\,4), (4\,6), (6\,7)\}$ is collinear to $\{(3\ 4), (3\ 6), (4\ 7)\}$ via $(1\ 2\ 5)(3\ 7\ 6) \in S_*(I)$. Furthermore, $\{(3\ 4), (4\ 7), (6\ 7)\}$ is collinear to $\{(3\ 4), (3\ 6), (4\ 7)\}$ via $(1\ 5)(3\ 6\ 7\ 4) \in S_*(I)$.

Let us consider subsets of $X_{(1)}$. We may assume, up to collineation, that $(1\ 2), (1\ 3) \in T$. The sets $\{(1\ 2), (1\ 3), (1\ 4)\}$ and $\{(1\ 2), (1\ 3), (1\ 5)\}$ are not collinear. $\{(1\ 2), (1\ 3), (1\ 6)\}$ is collinear to $\{(1\ 2), (1\ 3), (1\ 5)\}$ by $(2\,3)(5\,6) \in S_*(I)$. Noticing that $2*3 = 7$, we can see that $\{(1\,2), (1\,3), (1\,7)\}$ is not collinear to these sets.

Via $(1\,6\,3)(2\,7\,4) \in S_*(I)$ we see $\{(3\,4), (3\,6), (3\,7)\} \sim \{(1\,2), (1\,3), (1\,4)\}$. We use $(1\ 3\ 6)(2\ 4\ 7) \in S_*(I)$ to write $\{(1\ 2), (1\ 3), (1\ 5)\}$ in the equivalent form $\{(3\ 4), (3\ 5), (3\ 6)\}$, for the sake of convenience. Similarly, we write $X^{(5)}$ rather than $X^{(1)}$, and $X_{L_{34}}$ rather than $X_{L_{12}}$.

This leaves us with six possibilities: $X_{L_{34}}$, $X^{(5)}$, $\{(3\ 4), (3\ 6), (3\ 7)\}$, $\{3\,4), (3\,6), (4\,6)\}$, $\{(3\,4), (3\,6), (4\,7)\}$, and $\{(3\,4), (3\,5), (3\,6)\}$. The set $X_{L_{34}}$ is not collinear to any of the others because it is the only one in the form X_L. Similarly, $X^{(5)}$ is not collinear to the others. The sets $\{(3\ 4), (3\ 6), (3\ 7)\}$ and $\{(3\ 4), (3\ 5), (3\ 6)$ are the only options which contain some number thrice, and so they could only possibly be collinear to each other. However, $3 * 4 = 5$, means that they are not collinear. The last two options are not collinear because one contains exactly three numbers amongst its elements, while the other contains exactly four.

We assume now that $|T| = 4$. From Proposition A.16 we find that T is a subset of X_{L^C}, or $X_{(l)}$. We consider the subsets of $X_{L_{12}^C}$. We may assume, up to collineation, that $(3\ 4), (3\ 6) \in T$. This leaves the following six sets:

$$\{(3\ 4), (3\ 6), (3\ 7), (4\ 6)\},\ \{(3\ 4), (3\ 6), (3\ 7), (4\ 7)\},$$
$$\{(3\ 4), (3\ 6), (3\ 7), (6\ 7)\},\ \{(3\ 4), (3\ 6), (4\ 6), (4\ 7)\},$$
$$\{(3\ 4), (3\ 6), (4\ 6), (6\ 7)\},\ \{(3\ 4), (3\ 6), (4\ 7), (6\ 7)\}.$$

We have the following equivalences:

$$\{(3\ 4), (3\ 6), (3\ 7), (4\ 6)\} \sim \{(3\ 4), (3\ 6), (3\ 7), (4\ 7)\} \text{ by } (6\ 7) \in S_*(I),$$
$$\{(3\ 4), (3\ 6), (3\ 7), (4\ 6)\} \sim \{(3\ 4), (3\ 6), (3\ 7), (6\ 7)\} \text{ via } (4\ 6\ 7) \in S_*(I),$$
$$\{(3\ 4), (3\ 6), (3\ 7), (4\ 6)\} \sim \{(3\ 4), (3\ 6), (4\ 6), (4\ 7)\} \text{ by } (1\ 2)(3\ 4) \in S_*(I),$$
$$\{(3\ 4), (3\ 6), (3\ 7), (4\ 6)\} \sim \{(3\ 4), (3\ 6), (4\ 6), (6\ 7)\} \text{ by } (2\ 5)(3\ 6) \in S_*(I).$$

We are left with only $\{(3\ 4), (3\ 6), (3\ 7), (4\ 6)\}$ and $\{(3\ 4), (3\ 6), (4\ 7), (6\ 7)\}$. These two sets cannot be collinear as the former contains 3 in three of its elements while the latter does not contain any number in three of its elements.

We now look at subsets of $X_{(1)}$. We may assume, up to collineation, that $(1\ 2), (1\ 3) \in T$. This leaves the following sets: $\{(1\ 2), (1\ 3), (1\ 4), (1\ 5)\}$, $\{(1\ 2), (1\ 3), (1\ 4), (1\ 6)\}$, $\{(1\ 2), (1\ 3), (1\ 4), (1\ 7)\}$, $\{(1\ 2), (1\ 3), (1\ 5), (1\ 6)\}$, $\{(1\ 2), (1\ 3), (1\ 5), (1\ 7)\}$, $\{(1\ 2), (1\ 3), (1\ 6), (1\ 7)\}$. The collinear permutations $(5\ 6)$, $(5\ 7)$, $(4\ 6)$, $(4\ 7)$, $(2\ 6\ 5\ 3)(7\ 4)(\in S_*(I)$, respectively, allow us to conclude that $\{(1\ 2), (1\ 3), (1\ 4), (1\ 5)\}$ is collinear to each of $\{(1\ 2), (1\ 3), (1\ 4), (1\ 6)\}$, $\{(1\ 2), (1\ 3), (1\ 4), (1\ 7)\}$, $\{(1\ 2), (1\ 3), (1\ 5), (1\ 6)\}$, $\{(1\ 2), (1\ 3), (1\ 5), (1\ 7)\}$, and $\{(1\ 2), (1\ 3), (1\ 6), (1\ 7)\}$. This leaves us with only $\{(1\ 2), (1\ 3), (1\ 4), (1\ 5)\}$ which cannot be collinear to any of the subsets of $X_{L_{12}^C}$ since none of those sets are a subset of a set of the form $X_{(l)}$.

We assume now that $|T| = 5$. From Proposition A.16 we find that T is a subset of X_{L^C}, or $X_{(l)}$. We first consider the subsets of $X_{(1)}$. They are $X_{(1)} - \{(1\ 7)\}$, $X_{(1)} - \{(1\ 6)\}$, $X_{(1)} - \{(1\ 5)\}$, $X_{(1)} - \{(1\ 4)\}$, $X_{(1)} - \{(1\ 3)\}$, $X_{(1)} - \{(1\ 2)\}$. We can see that the first is collinear to the rest from, respectively, the permutations $(6\ 7)$, $(5\ 7)$, $(4\ 7)$, $(3\ 7)$, $(2\ 7) \in S_*(I)$.

We now consider the subsets of $X_{L_{12}^C}$. They are: $X_{L_{12}^C} - \{(6\ 7)\}$, $X_{L_{12}^C} - \{(4\ 7)\}$, $X_{L_{12}^C} - \{(4\ 6)\}$, $X_{L_{12}^C} - \{(3\ 7)\}$, $X_{L_{12}^C} - \{(3\ 6)\}$, $X_{L_{12}^C} - \{(3\ 4)\}$. We can see that the first is collinear to the rest from, respectively, the permutations $(1\ 5)(4\ 6)$, $(4\ 7\ 6)$, $(3\ 6\ 7)$, $(3\ 7\ 6)$, $(3\ 6\ 4\ 7) \in S_*(I)$.

This leaves us with $X_{(1)} - \{(1\ 7)\}$ and $X_{L_{12}^C} - \{(6\ 7)\}$. These cannot be collinear as the first is a subset of $X_{(1)}$ but the second is not the subset of any set of the form $X_{(l)}$.

We assume now that $|T| = 6$. From Proposition A.15 and Proposition A.16 we find that $T \in \{X_{L_{12}^C},\ P(1\ 2\ 3),\ X_{(1)}\}$.

From Proposition A.15 and Proposition A.16 we find that the remaining nice sets are: $T(1\ 2\ 3)$ ($|T| = 10$), $X - X_{L_{12}^C}$ ($|T| = 15$), and X ($|T| = 21$). Their uniqueness, up to collineation, is a result of the following: if $i, j, k \in I$ are generative, then $(1\ i)(2\ j)(3\ k) \in S_*(I)$. □

A.3 Algebra Properties

In the previous section we classified nice sets up to collineation. In this section we explore some properties of the algebras, $\mathfrak{D}^{\varepsilon^T}$, which correspond to each of our nice sets T. For convenience' sake we will denote $\mathfrak{D}^{\varepsilon^T}$ by $\mathfrak{D}^{\varepsilon}$. This covers graded contractions arising from admissible maps.

Lemma A.18: *Set* $M := \{i \in I \mid \nexists t \in T,\ i \in t\}$. *Then*

$$Z(\mathfrak{D}^{\varepsilon}) = Z(\mathfrak{D}) + \bigoplus_{i \in M} \mathfrak{D}_i.$$

Proof. It is immediately clear that $Z(\mathfrak{D})$, $\bigoplus_{i \in M} \mathfrak{D}_i \subseteq Z(\mathfrak{D}^{\varepsilon})$.

Take $x \in Z(\mathfrak{D}^{\varepsilon})$. We may write $x = \sum_{i \in I} x_i$, with $x_i \in \mathfrak{D}_i$. For all $y = \sum_{j \in I} y_j$, $y_j \in \mathfrak{D}_j$, we have that

$$0 = [x, y]^{\varepsilon^T} = \sum_{i,j \in I} \varepsilon_{ij}[x_i, y_j].$$

This means that, for each $i, j \in I$, either $\varepsilon_i j = 0$ or $[x_i, y_j] = 0$. Therefore, $x \in Z(\mathfrak{D}) + \bigoplus_{i \in M} \mathfrak{D}_i$. □

$\underline{T = \emptyset}$: $\mathfrak{D}^{\varepsilon}$ is abelian.

$\underline{T = \{(1\ 2)\}:}$ $Z(\mathfrak{D}^{\varepsilon}) = \bigoplus_{i \in I - \{1,2\}} \mathfrak{D}_i$. $\mathfrak{D}^{\varepsilon\prime} = \mathfrak{D}_5$. $\mathfrak{D}^{\varepsilon}$ is 2-step nilpotent and solvable. As a Lie algebra

$$\mathfrak{D}^{\varepsilon} = (\mathfrak{D}_1 \oplus \mathfrak{D}_2 \oplus \mathfrak{D}_5) \oplus (\mathfrak{D}_3 \oplus \mathfrak{D}_4 \oplus \mathfrak{D}_6 \oplus \mathfrak{D}_7).$$

$\underline{T = \{(1\ 2), (1\ 5)\}:}$ $Z(\mathfrak{D}^{\varepsilon}) = \bigoplus_{i \in I - \{1,2,5\}} \mathfrak{D}_i$. $\mathfrak{D}^{\varepsilon\prime} = \mathfrak{D}_2 \oplus \mathfrak{D}_5$. $\mathfrak{D}^{\varepsilon}$ is not nilpotent but is 2-step solvable. As a Lie algebra

$$\mathfrak{D}^{\varepsilon} = Z(\mathfrak{D}^{\varepsilon}) \oplus (\mathfrak{D}_1 \oplus \mathfrak{D}_2 \oplus \mathfrak{D}_5).$$

$\underline{T = \{(3\ 4), (3\ 6)\}:}$ $Z(\mathfrak{D}^{\varepsilon}) = \bigoplus_{i \in I - \{3,4,6\}} \mathfrak{D}_i$. $\mathfrak{D}^{\varepsilon\prime} = \mathfrak{D}_1 \oplus \mathfrak{D}_5$. $\mathfrak{D}^{\varepsilon}$ is 2-step nilpotent and solvable. As a Lie algebra

$$\mathfrak{D}^{\varepsilon} = (\mathfrak{D}_1 \oplus \mathfrak{D}_3 \oplus \mathfrak{D}_4 \oplus \mathfrak{D}_5 \oplus \mathfrak{D}_6) \oplus (\mathfrak{D}_2 \oplus \mathfrak{D}_7).$$

$\underline{T = \{(3\ 4), (6\ 7)\}:}$ $Z(\mathfrak{D}^{\varepsilon}) = \mathfrak{D}_1 \oplus \mathfrak{D}_2 \oplus \mathfrak{D}_5$. $\mathfrak{D}^{\varepsilon\prime} = \mathfrak{D}_5$. $\mathfrak{D}^{\varepsilon}$ is 2-step nilpotent and solvable. As a Lie algebra

$$\mathfrak{D}^{\varepsilon} = (\mathfrak{D}_1 \oplus \mathfrak{D}_2) \oplus (\mathfrak{D}_3 \oplus \mathfrak{D}_4 \oplus \mathfrak{D}_5 \oplus \mathfrak{D}_6 \oplus \mathfrak{D}_7).$$

$\underline{T = X_{L_{34}}:}$ $Z(\mathfrak{D}^{\varepsilon}) = \mathfrak{D}_1 \oplus \mathfrak{D}_2 \oplus \mathfrak{D}_6 \oplus \mathfrak{D}_7$. $\mathfrak{D}^{\varepsilon\prime} = \mathfrak{D}_3 \oplus \mathfrak{D}_4 \oplus \mathfrak{D}_5 = \mathfrak{D}^{\varepsilon 2} = \mathfrak{D}^{\varepsilon(2)}$. $\mathfrak{D}^{\varepsilon}$ is neither nilpotent nor solvable. $\mathfrak{D}^{\varepsilon}$ is not simple or semisimple. As Lie algebra

$$\mathfrak{D}^{\varepsilon} = Z(\mathfrak{D}^{\varepsilon}) \oplus \mathfrak{D}^{\varepsilon\prime}.$$

$\underline{T = X^{(5)} :}$ $Z(\mathfrak{D}^\varepsilon) = \mathfrak{D}_5 = \mathfrak{D}^{\varepsilon\prime}$. $\mathfrak{D}^\varepsilon$ is 2-step nilpotent and solvable.

$\underline{T = \{(3\ 4), (3\ 6), (3\ 7)\} :}$ $Z(\mathfrak{D}^\varepsilon) = \mathfrak{D}_1 \oplus \mathfrak{D}_2 \oplus \mathfrak{D}_5 = \mathfrak{D}^{\varepsilon\prime}$. $\mathfrak{D}^\varepsilon$ is 2-step nilpotent and solvable.

$\underline{T = \{(3\ 4), (3\ 6), (4\ 6)\} :}$ $Z(\mathfrak{D}^\varepsilon) = \mathfrak{D}_1 \oplus \mathfrak{D}_2 \oplus \mathfrak{D}_5 \oplus \mathfrak{D}_7$. $\mathfrak{D}^{\varepsilon\prime} = \mathfrak{D}_1 \oplus \mathfrak{D}_2 \oplus \mathfrak{D}_5$. $\mathfrak{D}^\varepsilon$ is 2-step nilpotent and solvable. As a Lie algebra

$$\mathfrak{D}^\varepsilon = (\mathfrak{D}_1 \oplus \mathfrak{D}_2 \oplus \mathfrak{D}_3 \oplus \mathfrak{D}_4 \oplus \mathfrak{D}_5 \oplus \mathfrak{D}_6) \oplus \mathfrak{D}_7.$$

$\underline{T = \{(3\ 4), (3\ 6), (4\ 7)\} :}$ $Z(\mathfrak{D}^\varepsilon) = \mathfrak{D}_1 \oplus \mathfrak{D}_2 \oplus \mathfrak{D}_5$. $\mathfrak{D}^{\varepsilon\prime} = \mathfrak{D}_1 \oplus \mathfrak{D}_5$. $\mathfrak{D}^\varepsilon$ is 2-step nilpotent and solvable. As a Lie algebra

$$\mathfrak{D}^\varepsilon = (\mathfrak{D}_1 \oplus \mathfrak{D}_3 \oplus \mathfrak{D}_4 \oplus \mathfrak{D}_5 \oplus \mathfrak{D}_6 \oplus \mathfrak{D}_7) \oplus \mathfrak{D}_2.$$

$\underline{T = \{(3\ 4), (3\ 5), (3\ 6)\} :}$ $Z(\mathfrak{D}^\varepsilon) = \mathfrak{D}_1 \oplus \mathfrak{D}_2 \oplus \mathfrak{D}_7$. $\mathfrak{D}^{\varepsilon\prime} = \mathfrak{D}_1 \oplus \mathfrak{D}_4 \oplus \mathfrak{D}_5$. $\mathfrak{D}^{\varepsilon 2} = \mathfrak{D}^{\varepsilon 3} = \mathfrak{D}_4 \oplus \mathfrak{D}_5$. $\mathfrak{D}^\varepsilon$ is not nilpotent but is 2-step solvable. As a Lie algebra

$$\mathfrak{D}^\varepsilon = (\mathfrak{D}_1 \oplus \mathfrak{D}_3 \oplus \mathfrak{D}_4 \oplus \mathfrak{D}_5 \oplus \mathfrak{D}_6) \oplus (\mathfrak{D}_2 \oplus \mathfrak{D}_7).$$

$\underline{T = \{(3\ 4), (3\ 6), (3\ 7), (4\ 6)\} :}$ $Z(\mathfrak{D}^\varepsilon) = \mathfrak{D}_1 \oplus \mathfrak{D}_2 \oplus \mathfrak{D}_5 = \mathfrak{D}^{\varepsilon\prime}$. $\mathfrak{D}^\varepsilon$ is 2-step solvable and nilpotent.

$\underline{T = \{(3\ 4), (3\ 6), (4\ 7), (6\ 7)\} :}$ $Z(\mathfrak{D}^\varepsilon) = \mathfrak{D}_1 \oplus \mathfrak{D}_2 \oplus \mathfrak{D}_5$. $\mathfrak{D}^{\varepsilon\prime} = \mathfrak{D}_1 \oplus \mathfrak{D}_5$. $\mathfrak{D}^\varepsilon$ is 2-step nilpotent and solvable. As a Lie algebra

$$\mathfrak{D}^\varepsilon = (\mathfrak{D}_1 \oplus \mathfrak{D}_3 \oplus \mathfrak{D}_4 \oplus \mathfrak{D}_5 \oplus \mathfrak{D}_6 \oplus \mathfrak{D}_7) \oplus \mathfrak{D}_2.$$

$\underline{T = \{(1\ 2), (1\ 3), (1\ 4), (1\ 5)\} :}$ $Z(\mathfrak{D}^\varepsilon) = \mathfrak{D}_6 \oplus \mathfrak{D}_7$. $\mathfrak{D}^{\varepsilon\prime} = \mathfrak{D}_2 \oplus \mathfrak{D}_5 \oplus \mathfrak{D}_6 \oplus \mathfrak{D}_7$. $\mathfrak{D}^{\varepsilon 2} = \mathfrak{D}^{\varepsilon 3} = \mathfrak{D}_2 \oplus \mathfrak{D}_5$. $\mathfrak{D}^\varepsilon$ is not nilpotent but is 2-step solvable.

$\underline{T = \{(3\ 4), (3\ 6), (3\ 7), (4\ 6), (4\ 7)\} :}$ $Z(\mathfrak{D}^\varepsilon) = \mathfrak{D}_1 \oplus \mathfrak{D}_2 \oplus \mathfrak{D}_5 = \mathfrak{D}^{\varepsilon\prime}$. $\mathfrak{D}^\varepsilon$ is 2-step nilpotent and solvable.

$\underline{\{(1\ 2), (1\ 3), (1\ 4), (1\ 5), (1\ 6)\} :}$ $Z(\mathfrak{D}^\varepsilon) = \mathfrak{D}_7$. $\mathfrak{D}^{\varepsilon\prime} = \mathfrak{D}_2 \oplus \mathfrak{D}_3 \oplus \mathfrak{D}_5 \oplus \mathfrak{D}_6 \oplus \mathfrak{D}_7$. $\mathfrak{D}^{\varepsilon 2} = \mathfrak{D}^{\varepsilon 3} = \mathfrak{D}_2 \oplus \mathfrak{D}_3 \oplus \mathfrak{D}_5 \oplus \mathfrak{D}_6$. $\mathfrak{D}^\varepsilon$ is not nilpotent but is 2-step solvable.

$\underline{T = X_{L^C_{12}} :}$ $Z(\mathfrak{D}^\varepsilon) = \mathfrak{D}_1 \oplus \mathfrak{D}_2 \oplus \mathfrak{D}_5 = \mathfrak{D}^{\varepsilon\prime}$. $|dae$ is 2-step nilpotent and solvable.

$\underline{T = P(1\ 2\ 3) :}$ $Z(\mathfrak{D}^\varepsilon) = \mathfrak{D}_4$. $\mathfrak{D}^{\varepsilon\prime} = \mathfrak{D}_4 \oplus \mathfrak{D}_5 \oplus \mathfrak{D}_6 \oplus \mathfrak{D}_7$. $\mathfrak{D}^{\varepsilon 2} = \mathfrak{D}_4$. $\mathfrak{D}^\varepsilon$ is 3-step nilpotent and 2-step solvable.

$\underline{T = X_{(1)} :}$ $Z(\mathfrak{D}^\varepsilon) = 0$. $\mathfrak{D}^{\varepsilon\prime} = \mathfrak{D}^{\varepsilon 2} = \mathfrak{D}_2 \oplus \mathfrak{D}_3 \oplus \mathfrak{D}_4 \oplus \mathfrak{D}_5 \oplus \mathfrak{D}_6 \oplus \mathfrak{D}_7$. $\mathfrak{D}^\varepsilon$ is not nilpotent but is 2-step solvable.

$\underline{T = T(1\ 2\ 3) :}$ $Z(\mathfrak{D}^\varepsilon) = 0$. $\mathfrak{D}^{\varepsilon\prime} = \mathfrak{D}^{\varepsilon 2} = \mathfrak{D}_2 \oplus \mathfrak{D}_3 \oplus \mathfrak{D}_4 \oplus \mathfrak{D}_5 \oplus \mathfrak{D}_6 \oplus \mathfrak{D}_7$. $\mathfrak{D}^{\varepsilon(2)} = \mathfrak{D}_4 \oplus \mathfrak{D}_7$. $\mathfrak{D}^\varepsilon$ is not nilpotent but is 3-step solvable.

$\underline{T = X - X_{L^C_{12}} :}$ $Z(\mathfrak{D}^\varepsilon) = 0$. $\mathfrak{D}^{\varepsilon\prime} = \mathfrak{D}^\varepsilon$. $\mathfrak{D}^\varepsilon$ is neither nilpotent nor solvable. $\mathfrak{D}^\varepsilon$ is neither simple nor semisimple. The only proper ideal of $\mathfrak{D}^\varepsilon$ which is a direct sum of homogeneous components of the grading is $\mathfrak{D}_3 \oplus \mathfrak{D}_4 \oplus \mathfrak{D}_6 \oplus \mathfrak{D}_7$.

$\underline{T = X :}$ $\mathfrak{D}^\varepsilon = \mathfrak{D}$.

A.4 Scaling homogeneous Components

In the previous section we considered admissible maps whose image is $\{0,1\}$. In this section we consider admissible maps which have the same kernel as the admissible graded contractions but for which the non-zero elements need not necessarily have value 1. We may use the following notation to write a graded contraction when its support is clear from context. We write only the relevant parameters (those corresponding to an element in their support). These relevant parameters are unambiguously totally ordered in the following way: we consider the numbers in the pair from X to which they correspond (ε_{15} would correspond to the pari (1 5)) and we first compare the smallest number from each parameter, and then the larger number from each pair. For instance, if ε is a graded contraction with support $\{(3\,4),(3\,6),(3\,7),(4\,6)\}$ then we may write $\varepsilon := (\varepsilon_{34}\,\varepsilon_{36}\,\varepsilon_{37}\,\varepsilon_{46}) = (a\,b\,c\,d)$.

Recall that for an admissible map ε arising from the nice set T, then T is the ***support*** of ε.

We define a matrix $\alpha_{ij} = \dfrac{a_i a_j}{a_{i*j}}$ for $i,j \in I$.

$\underline{T = \{(1\ 2)\}}$: If we set

$$\varepsilon_{ij} = \begin{cases} c, & \text{if } (i\ j) = (1\ 2), \\ 0, & \text{otherwise,} \end{cases}$$

where $c \in \mathbb{C}^*$. Then, setting $a_1 = a_2 = 2$, $a_5 = c$, and $a_i = 0$ for $i \in I - \{1,2,5\}$ we have that $\alpha \bullet \varepsilon = \varepsilon^T$.

$\underline{T = \{(1\ 2),(1\ 5)\}}$: If we set

$$\varepsilon_{ij} = \begin{cases} c, & \text{if } (i\ j) = (1\ 2) \\ d, & \text{if } (i\ j) = (1\ 5) \\ 0, & \text{otherwise} \end{cases},$$

where $c,d \in \mathbb{C}^*$. Then, setting $a_1 = a_2 = \frac{1}{\sqrt{cd}}$, $a_5 = \frac{1}{d}$, and the other $a_i = 1$, we have that $\alpha \bullet \varepsilon = \varepsilon^T$.

$\underline{T = \{(3\ 4),(3\ 6)\}}$: If we set

$$\varepsilon_{ij} = \begin{cases} c, & \text{if } (i\ j) = (3\ 4) \\ d, & \text{if } (i\ j) = (3\ 6) \\ 0, & \text{otherwise} \end{cases},$$

where $c,d \in \mathbb{C}^*$. Then, setting $a_1 = d$, $a_5 = c$, $a_3 = a_4 = a_6 = 1$, and $a_7 = 1$, we have $\alpha \bullet \varepsilon = \varepsilon^T$.

$\underline{T = \{(3\ 4), (6\ 7)\} :}$ If we set

$$\varepsilon_{ij} = \begin{cases} c, & \text{if } (i\ j) = (3\ 4) \\ d, & \text{if } (i\ j) = (3\ 6) \\ 0, & \text{otherwise} \end{cases},$$

where $c, d \in \mathbb{C}^*$. Setting $a_3 = 1/c$, $a_6 = 1/d$, $a_4 = a_5 = a_7 = 1$, and $a_1 = a_2 = 1$, we have $\alpha \bullet \varepsilon = \varepsilon^T$.

$\underline{T = X_{L_{34}}}$:If we set

$$\varepsilon_{ij} = \begin{cases} c, & \text{if } (i\ j) = (3\ 4) \\ d, & \text{if } (i\ j) = (3\ 5) \\ f, & \text{if } (i\ j) = (4\ 5) \\ 0, & \text{otherwise} \end{cases},$$

where $c, d, f \in \mathbb{C}^*$. Then, setting $a_3 = \frac{1}{\sqrt{cd}}$, $a_4 = \frac{1}{\sqrt{cf}}$, $a_5 = \frac{1}{\sqrt{df}}$, and the other $a_i = 1$, we have that $\alpha \bullet \varepsilon = \varepsilon^T$.

$\underline{T = X^{(5)} :}$ If we set

$$\varepsilon_{ij} = \begin{cases} c, & \text{if } (i\ j) = (1\ 2) \\ d, & \text{if } (i\ j) = (3\ 4) \\ f, & \text{if } (i\ j) = (6\ 7) \\ 0, & \text{otherwise} \end{cases},$$

where $c, d, f \in \mathbb{C}^*$. Then, setting $a_1 = 1/c$, $a_3 = 1/d$, $a_6 = 1/f$, and the other $a_i = 1$, we have $\alpha \bullet \varepsilon = \varepsilon^T$.

$\underline{T = \{(3\ 4), (3\ 6), (3\ 7)\} :}$ If we set

$$\varepsilon_{ij} = \begin{cases} c, & \text{if } (i\ j) = (3\ 4) \\ d, & \text{if } (i\ j) = (3\ 6) \\ f, & \text{if } (i\ j) = (3\ 7) \\ 0, & \text{otherwise} \end{cases},$$

where $c, d, f \in \mathbb{C}^*$. Then, setting $a_1 = d$, $a_2 = f$, $a_5 = c$, and the other $a_i = 1$, we have $\alpha \bullet \varepsilon = \varepsilon^T$.

$\underline{T = \{(3\ 4), (3\ 6), (4\ 6)\} :}$ If we set

$$\varepsilon_{ij} = \begin{cases} c, & \text{if } (i\ j) = (3\ 4) \\ d, & \text{if } (i\ j) = (3\ 6) \\ f, & \text{if } (i\ j) = (4\ 6) \\ 0, & \text{otherwise} \end{cases},$$

where $c, d, f \in \mathbb{C}^*$. Then, setting $a_1 = d$, $a_2 = f$, $a_5 = c$, and the other $a_i = 1$, we have $\alpha \bullet \varepsilon = \varepsilon^T$.

$\underline{T = \{(3\ 4), (3\ 6), (4\ 7)\} :}$ If we set

$$\varepsilon_{ij} = \begin{cases} c, & \text{if } (i\ j) = (3\ 4) \\ d, & \text{if } (i\ j) = (3\ 6) \\ f, & \text{if } (i\ j) = (4\ 7) \\ 0, & \text{otherwise} \end{cases},$$

where $c, d, f \in \mathbb{C}^*$. Then, setting $a_5 = c$, $a_6 = 1/d$, $a_7 = 1/f$, and the other $a_i = 1$, we have $\alpha \bullet \varepsilon = \varepsilon^T$.

$\underline{T = \{(3\ 4), (3\ 5), (3\ 6)\} :}$ If we set

$$\varepsilon_{ij} = \begin{cases} c, & \text{if } (i\ j) = (3\ 4) \\ d, & \text{if } (i\ j) = (3\ 5) \\ f, & \text{if } (i\ j) = (3\ 6) \\ 0, & \text{otherwise} \end{cases},$$

where $c, d, f \in \mathbb{C}^*$. Then, setting $a_1 = f$, $a_3 = \frac{1}{\sqrt{cd}}$, $a_4 = \sqrt{d}$, $a_5 = \sqrt{c}$, $a_6 = \sqrt{cd}$, and the other $a_i = 1$, we have $\alpha \bullet \varepsilon = \varepsilon^T$.

$\underline{T = \{(3\ 4), (3\ 6), (3\ 7), (4\ 6)\} :}$ If we set

$$\varepsilon_{ij} = \begin{cases} c, & \text{if } (i\ j) = (3\ 4) \\ d, & \text{if } (i\ j) = (3\ 6) \\ f, & \text{if } (i\ j) = (3\ 7) \\ g, & \text{if } (i\ j) = (4\ 6)\} \\ 0, & \text{otherwise} \end{cases},$$

where $c, d, f, g \in \mathbb{C}^*$. Then, setting $a_1 = d$, $a_2 = g$, $a_3 = a_4 = a_5 = 1/c$, $a_6 = c$ and $a_7 = \frac{cg}{f}$, we have $\alpha \bullet \varepsilon = \varepsilon^T$.

$\underline{T = \{(3\ 4), (3\ 6), (4\ 7), (6\ 7)\} :}$ If we set

$$\varepsilon_{ij} = \begin{cases} c, & \text{if } (i\ j) = (3\ 4) \\ d, & \text{if } (i\ j) = (3\ 6) \\ f, & \text{if } (i\ j) = (4\ 7) \\ g, & \text{if } (i\ j) = (6\ 7)\} \\ 0, & \text{otherwise} \end{cases},$$

where $c, d, f, g \in \mathbb{C}^*$. Then, setting $a_1 = a_4 = d$, $a_3 = \frac{\sqrt{g}}{\sqrt{cdf}}$, $a_5 = \frac{\sqrt{cdg}}{\sqrt{f}}$, $a_6 = \frac{\sqrt{cdf}}{\sqrt{g}}$ and $a_7 = 1/f$, we have $\alpha \bullet \varepsilon = \varepsilon^T$.

$\underline{T = \{(1\ 2), (1\ 3), (1\ 4), (1\ 5)\}:}$ If we set

$$\varepsilon_{ij} = \begin{cases} c, & \text{if } (i\ j) = (1\ 2) \\ d, & \text{if } (i\ j) = (1\ 3) \\ f, & \text{if } (i\ j) = (1\ 4) \\ g, & \text{if } (i\ j) = (1\ 5) \\ 0, & \text{otherwise} \end{cases},$$

where $c, d, f, g \in \mathbb{C}^*$. Then, setting $a_1 = \frac{1}{\sqrt{cg}}$, $a_2 = \sqrt{g}$, $a_3 = a_4 = \sqrt{cg}$, $a_5 = \sqrt{c}$, $a_6 = d$, and $a_7 = f$, we have $\alpha \bullet \varepsilon = \varepsilon^T$.

$\underline{T = \{(3\ 4), (3\ 6), (3\ 7), (4\ 6), (4\ 7)\}:}$ If we set

$$\varepsilon_{ij} = \begin{cases} c, & \text{if } (i\ j) = (3\ 4) \\ d, & \text{if } (i\ j) = (3\ 6) \\ f, & \text{if } (i\ j) = (3\ 7) \\ g, & \text{if } (i\ j) = (4\ 6) \\ h, & \text{if } (i\ j) = (4\ 7) \\ 0, & \text{otherwise} \end{cases},$$

where $c, d, f, g, h \in \mathbb{C}^*$. Then, setting $a_1 = d\sqrt{gh}$, $a_2 = g\sqrt{fd}$, $a_3 = \sqrt{gh}$, $a_4 = \sqrt{fd}$, $a_5 = c\sqrt{dfgh}$, $a_6 = 1$, and $a_7 = \frac{\sqrt{dg}}{\sqrt{fh}}$, we have $\alpha \bullet \varepsilon = \varepsilon^T$.

$\underline{T = \{(1\ 2), (1\ 3), (1\ 4), (1\ 5), (1\ 6)\}:}$ If we fix a graded contraction ε with support T, we consider the graded contractions which are equivalent to ε by normalization. That is, we consider all graded contractions $\varepsilon' = \alpha \bullet \varepsilon$, where α is a normalization matrix. We may always normalize $\varepsilon'_{12} = \varepsilon'_{15} = 1$. To see this, we use the (proposed) equalities $\alpha_{12}\varepsilon_{12} = 1$ and $\alpha_{15}\varepsilon_{15} = 1$. We may therefore set $a_1 := \frac{1}{\sqrt{\varepsilon_{12}\varepsilon_{15}}}$, $a_2 := \sqrt{\varepsilon_{15}}$, $a_5 := \sqrt{\varepsilon_{15}}$. This forces $\varepsilon'_{13} = \frac{a_3\varepsilon_{13}}{a_6\sqrt{\varepsilon_{12}\varepsilon_{15}}}$ and $\varepsilon'_{16} = \frac{a_6\varepsilon_{16}}{a_3\sqrt{\varepsilon_{12}\varepsilon_{15}}}$. Finally, $\varepsilon'_{14} = \frac{a_4\varepsilon_{14}}{a_7\sqrt{\varepsilon_{12}\varepsilon_{15}}}$. Altogether, this means that any graded contraction with support T is equivalent (via normalization) to a graded contraction ε with $\varepsilon_{12} = \varepsilon_{15} = 1$, $\varepsilon_{13} = \varepsilon_{16} = c$, $\varepsilon_{14} = d$, where $c, d \in \mathbb{C}^*$. We will specify the five relevant parameters of such a graded contraction by $\varepsilon = (1\ c\ d\ 1\ c)$.

Graded contractions $\varepsilon := (1\ c\ d\ 1\ c)$ and $\varepsilon' := (1\ c'\ d\ 1\ c')$ are equivalent via normalization if, and only if, $c' = \pm c$. To see this, we first note that $\alpha\varepsilon = \varepsilon'$ if, and only if, there exist $a_1, \ldots, a_7 \in \mathbb{C}$ such that the following hold:

$$\alpha_{12} = 1, \quad \alpha_{13} = \frac{c'}{c}, \quad \alpha_{14} = \frac{d'}{d},$$
$$\alpha_{15} = 1, \quad \alpha_{16} = \frac{c'}{c}.$$

These conditions hold if, and only if, the following holds: $a_1^2 = 1 = \left(\frac{c'}{c}\right)^2$. Our claim follows.

$\underline{T = X_{(1)}}$: If we fix a graded contraction ε with support T, we consider the graded contractions which are equivalent to ε by normalization. That is, we consider all graded contractions $\varepsilon' = \alpha \bullet \varepsilon$, where α is a normalization matrix. We may always normalize $\varepsilon'_{12} = \varepsilon'_{15} = 1$. To see this, we use the (proposed) equalities $\alpha_{12}\varepsilon_{12} = 1$ and $\alpha_{15}\varepsilon_{15} = 1$. We may therefore set $a_1 := \frac{1}{\sqrt{\varepsilon_{12}\varepsilon_{15}}}$, $a_2 := \sqrt{\varepsilon_{15}}$, $a_5 := \sqrt{\varepsilon_{15}}$. This forces $\varepsilon'_{13} = \frac{a_3\varepsilon_{13}}{a_6\sqrt{\varepsilon_{12}\varepsilon_{15}}}$, $\varepsilon'_{16} = \frac{a_6\varepsilon_{16}}{a_3\sqrt{\varepsilon_{12}\varepsilon_{15}}}$, $\varepsilon'_{14} = \frac{a_4\varepsilon_{14}}{a_7\sqrt{\varepsilon_{12}\varepsilon_{15}}}$, and $\varepsilon'_{17} = \frac{a_7\varepsilon_{17}}{a_4\sqrt{\varepsilon_{12}\varepsilon_{15}}}$. Altogether, this means that any graded contraction with support T is equivalent (via normalization) to a graded contraction ε with $\varepsilon_{12} = \varepsilon_{15} = 1$, $\varepsilon_{13} = \varepsilon_{16} = c$, $\varepsilon_{14} = \varepsilon_{17} = d$, where $c, d \in \mathbb{C}^*$. We will specify the six relevant parameters of such a graded contraction by $\varepsilon = (1\ c\ d\ 1\ c\ d)$.

Graded contractions $\varepsilon := (1\ c\ d\ 1\ c\ d)$ and $\varepsilon' := (1\ c'\ d'\ 1\ c'\ d')$ are equivalent via normalization if, and only if, $c' = \pm c$ and $d' = \pm d$. To see this, we first note that $\alpha\varepsilon = \varepsilon'$ if, and only if, there exist $a_1, \dots, a_7 \in \mathbb{C}$ such that the following hold:

$$\alpha_{12} = 1, \quad \alpha_{13} = \frac{c'}{c}, \quad \alpha_{14} = \frac{d'}{d},$$
$$\alpha_{15} = 1, \quad \alpha_{16} = \frac{c'}{c}, \quad \alpha_{17} = \frac{d'}{d}.$$

These conditions hold if, and only if, the following holds: $a_1^2 = 1 = \left(\frac{c'}{c}\right)^2 = \left(\frac{d'}{d}\right)^2$. Our claim follows.

$\underline{T = X_{L^C_{12}}}$: Any graded contraction $\varepsilon := (\varepsilon_{34}\ \varepsilon_{36}\ \varepsilon_{37}\ \varepsilon_{46}\ \varepsilon_{47}\ \varepsilon_{67})$, with support T is equivalent via normalization to the graded contraction ε' : $(1\ 1\ 1\ 1\ 1\ 1\ 1)$. To see this, set

$$a_1 := \frac{1}{\sqrt{\varepsilon_{34}\varepsilon_{37}\varepsilon_{46}\varepsilon_{67}}},\ a_2 := \frac{1}{\sqrt{\varepsilon_{34}\varepsilon_{36}\varepsilon_{47}\varepsilon_{67}}},$$
$$a_3 := \frac{1}{\sqrt{\varepsilon_{34}\varepsilon_{36}\varepsilon_{37}}},\ a_4 := \frac{1}{\sqrt{\varepsilon_{34}\varepsilon_{46}\varepsilon_{47}}},$$
$$a_5 := \frac{1}{\sqrt{\varepsilon_{36}\varepsilon_{37}\varepsilon_{46}\varepsilon_{47}}},\ a_6 := \frac{1}{\sqrt{\varepsilon_{36}\varepsilon_{46}\varepsilon_{67}}},$$
$$a_7 := \frac{1}{\sqrt{\varepsilon_{37}\varepsilon_{47}\varepsilon_{67}}}.$$

Then, $\varepsilon' = \alpha \bullet \varepsilon$.

$\underline{T = P(1\ 2\ 3)}$: First we claim that any graded contraction of the form $\varepsilon = (\varepsilon_{12}\ \varepsilon_{13}\ \varepsilon_{17}\ \varepsilon_{23}\ \varepsilon_{26}\ \varepsilon_{35})$, with support T, is equivalent via normalization to a graded contraction, supported by T, of the form $\varepsilon' = (1\ 1\ 1\ 1\ \beta\ 1\ \gamma)$, with $\beta, \gamma \in \mathbb{C}^*$. To see this, set

$$a_1 = a_4 = a_6 := \frac{1}{\varepsilon_{12}},\ a_2 = a_5 := \frac{1}{\varepsilon_{26}},$$
$$a_3 := \frac{1}{\varepsilon_{13}},\ a_7 := \frac{1}{\varepsilon_{17}}.$$

Then,

$$\beta = \frac{\varepsilon_{17}\varepsilon_{23}}{\varepsilon_{13}\varepsilon_{26}},\ \gamma = \frac{\varepsilon_{12}\varepsilon_{35}}{\varepsilon_{13}\varepsilon_{26}}.$$

We now have $\varepsilon' = \alpha \bullet \varepsilon$. Further we claim that graded contractions, with support T, $\varepsilon = (1\ 1\ 1\ \beta\ 1\ \gamma)$ and $\varepsilon' = (1\ 1\ 1\ \beta'\ 1\ \gamma')$ are equivalent via normalization if, and only if, $\beta = \beta'$ and $\gamma = \gamma'$. To see this we note that for us to have $\varepsilon' = \alpha \bullet \varepsilon$ we need the following to hold:

$$\alpha_{12} = 1,\ \alpha_{13} = 1,\ \alpha_{17} = 1,$$
$$\alpha_{23} = \frac{\beta'}{\beta},\ \alpha_{26} = 1,\ \alpha_{35} = \frac{\gamma'}{\gamma}.$$

From these equalities we find that the following must hold:

$$a_5 = a_1a_2,\ a_6 = a_1a_3,\ a_4 = a_1a_2a_3\frac{\beta}{\beta'},$$
$$a_4 = a_1a_2a_3,\ a_7 = a_2a_3\frac{\beta}{\beta'},\ a_4 = a_1a_2a_3\frac{\gamma}{\gamma'}.$$

Our second claim follows.

$\underline{T = T(1\ 2\ 3):}$ If ε is a graded contraction with support T then ε is equivalent via normalization to a graded contraction ε' defined as $\varepsilon' = (1\ c\ d\ 1\ c\ d\ f\ g\ h\ k)$. To see this we set

$$a_1 := \frac{1}{\sqrt{\varepsilon_{12}\varepsilon_{15}}},\ a_2 := \sqrt{\varepsilon_{15}},\ a_3 := \sqrt{\varepsilon_{16}},\ a_4 := \sqrt{\varepsilon_{17}},$$
$$a_5 := \sqrt{\varepsilon_{12}},\ a_6 := \sqrt{\varepsilon_{13}},\ a_7 := \sqrt{\varepsilon_{14}}.$$

This gives $\alpha_{12}\varepsilon_{12} = \alpha_{15}\varepsilon_{15} = 1,\ \alpha_{13}\varepsilon_{13} = \alpha_{16}\varepsilon_{16} = c,\ \alpha_{14}\varepsilon_{14} = \alpha_{17}\varepsilon_{17} = d,$ $f = \alpha_{23}\varepsilon_{23},\ g = \alpha_{26}\varepsilon_{26},\ h = \alpha_{35}\varepsilon_{35},\ k = \alpha_{56}\varepsilon_{56}$.

We will now consider when graded contractions $\varepsilon := (1\ c\ d\ 1\ c\ d\ f\ g\ h\ k)$ and $\varepsilon' := (1\ c'\ d'\ 1\ c'\ d'\ f'\ g'\ h'\ k')$, both with support T, are equivalent via normalization. We therefore consider the conditions under which $\varepsilon' = \alpha \bullet \varepsilon$. This is equivalent to the following holding:

$$\alpha_{12} = \alpha_{15} = 1,\ \alpha_{13} = \alpha_{16} = \frac{c'}{c},\ \alpha_{14} = \alpha_{17} = \frac{d'}{d},$$
$$\alpha_{23} = \frac{f'}{f},\ \alpha_{26} = \frac{g'}{g},\ \alpha_{35} = \frac{h'}{h},\ \alpha_{56} = \frac{k'}{k}.$$

These equalities yield the following: $\alpha_{12}\alpha_{15} = 1 = a_1^2,\ \alpha_{13}/\alpha_{16} = 1 = (a_3/a_6)^2,\ \alpha_{14}/\alpha_{17} = 1 = (a_4/a_7)^2$. This means that $c' = \pm c,\ d' = \pm d,\ a_5 = \pm a_2,\ a_6 = \pm a_3,\ a_7 = \pm a_4$. Substituting this all back into the original equalities allows us to conclude that

$$f' = \pm f\frac{a_2a_3}{a_4},\ g' = \pm g\frac{a_2a_3}{a_4},$$
$$h' = \pm h\frac{a_2a_3}{a_4},\ k' = \pm k\frac{a_2a_3}{a_4}.$$

In summary, ε is equivalent to ε' if, and only if, the following conditions hold: $c' = \pm c,\ d' = \pm d$, and there exists $\beta \in \mathbb{C}^*$ such that $f' = \pm\beta f,\ g' = \pm\beta g,\ h' = \pm\beta h,\ k' = \pm\beta k$.

$\underline{T = X - X_{L_{12}^C}}$: Since $X_{(1)} \subset T$, we can use the argument from the $X_{(1)}$ case to see that any graded contraction with support T is equivalent via normalization to a graded contraction of the form

$$\varepsilon := (1\ c_1\ c_2\ 1\ c_1\ c_2\ c_3\ c_4\ c_5\ c_6\ c_7\ c_8\ c_9\ c_{10}\ c_{11}).$$

We now consider when such a graded contraction is equivalent via normalization to another graded contraction of this form:

$$\varepsilon' := (1\ c_1'\ c_2'\ 1\ c_1'\ c_2'\ c_3'\ \ldots\ c_{11}').$$

We therefore consider under what conditions we can find $a_1, \ldots, a_7 \in \mathbb{C}$ such that $\varepsilon' = \alpha \bullet \varepsilon'$. From the equalities $\alpha_{12} = 1 = \alpha_{15}$ we find that $a_1^2 = 1$. This along with the equalities $\alpha_{13} = c_1'/c_1 = \alpha_{16}$ and $\alpha_{14} = c_2'/c_2 = \alpha_{16}$ we find that $a_5 = \pm a_2,\ a_6 = \pm a_3,\ a_7 = \pm a_4$. This means that $c_1' = \pm c_1$ and $c_2' = \pm c_2$. This also allows us to rewrite all the equalities $\alpha_{ij} = c_k'/c_k$ in a form using only a_2, a_3, and a_4. Then $\alpha_{25} = c_5'/c_5$ becomes $\pm a_2^2 = c_5'/c_5$. Using all this we arrive at the following equalities:

$$\pm\frac{a_3}{a_4} = \frac{c_3'c_4}{c_3c_4'},\ \pm 1 = \frac{c_3'c_6}{c_3c_6'},\ \pm\frac{a_3}{a_4} = \frac{c_3'c_7}{c_3c_7'},$$
$$\pm 1 = \frac{c_3'c_8}{c_3c_8'},\ \pm\frac{a_3}{a_4} = \frac{c_3'c_9}{c_3c_9'},\ \pm 1 = \frac{c_3'c_{10}}{c_3c_{10}'},$$
$$\pm\frac{a_3}{a_4} = \frac{c_3'c_{11}}{c_3c_{11}'},$$

and

$$\pm\frac{a_4}{a_3} = \frac{c_4'c_6}{c_4c_6'},\ \pm 1 = \frac{c_4'c_7}{c_4c_7'},\ \pm\frac{a_4}{a_3} = \frac{c_4'c_8}{c_4c_8'},$$
$$\pm 1 = \frac{c_4'c_9}{c_4c_9'},\ \pm\frac{a_4}{a_3} = \frac{c_4'c_{10}}{c_4c_{10}'},\ \pm 1 = \frac{c_4'c_{11}}{c_4c_{11}'},$$

These allow us to write the following:

$$
\begin{aligned}
&c_5' = \pm c_5 \frac{c_3' c_4'}{c_3 c_4}, \\
&c_6' = \pm c_6 \frac{c_3'}{c_3},\ c_7' = \pm c_7 \frac{c_4'}{c_4},\ c_8' = \pm c_8 \frac{c_3'}{c_3}, \\
&c_9' = c_9 \frac{c_4'}{c_4},\ c_{10}' = \pm c_{10} \frac{c_3'}{c_3},\ c_{11}' = \pm c_{11} \frac{c_4'}{c_4}.
\end{aligned}
$$

Therefore, we conclude that ε is equivalent to ε' via normalization if, and only if, the above holds and $c_1' = \pm c_1$ and $c_2' = \pm c_2$.

Appendix B

Tensor Product

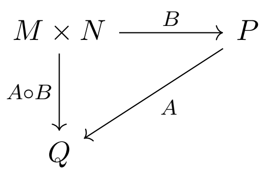
This chapter is devoted to the tensor product of R-modules, for a commutative unital ring R. Further study of graded contractions of Lie algebras requires familiarity with this topic. Moreover, this knowledge is essential is modern algebra research. The tensor product is an algebraic structure which has the universal property that all multi-linear maps filter through it.

Throughout this chapter R will denote a commutative (unital) ring.

B.1 Definition and Construction

We begin with a definition of the tensor products for vector spaces and then generalise this to a definition for R-modules. First we see a definition of the tensor product in terms of a universal property. We end this section with an explicit construction of the tensor product

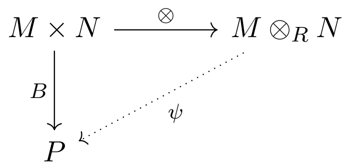
Remark B.1: *Let R be a commutative ring with R-modules M, N, P, and Q. If $B\colon M \times N \to P$ is a bilinear map and $A\colon P \to Q$ a linear map, then clearly $A \circ B$ is a bilinear map.*



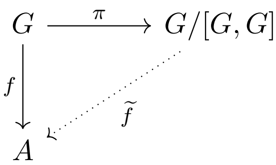
We will now define the tensor product of R-modules M and N by a universal mapping property.

Definition B.2: *Let M and N be R-modules, for a commutative ring R. Then we define the* **tensor product** *$M \otimes_R N$ as an R-module along with a bilinear map $\otimes\colon M \times N \to M \otimes_R N$ such that for any bilinear map*

$B\colon M \times N \to P$ there is a unique R-linear map $\psi : M \otimes_R N \to P$ such that the following diagram commutes.



Example B.3: *Let G be a group, A an abelian group, and $f\colon G \to A$ a homomorphism. Then there is a unique homomorphism $\widetilde{f}\colon G/[G,G] \to A$ such that the following diagram commutes.*



where π is the canonical homomorphism.

We define $\widetilde{f}(g[G,G]) := f(g)$. This is well defined because for $g, h \in G$ such that $gh^{-1} \in [G,G]$ we have that $gh^{-1} = [x,y]$ for some $x, y \in G$ and so

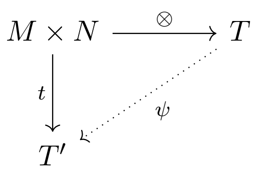
$$f(g) - f(h) = f(gh^{-1}) = f(x^{-1}y^{-1}xy) = -f(x) - f(y) + f(x) + f(y) = 0.$$

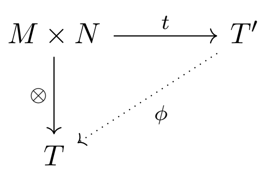
Moreover, if $\phi\colon G/[G,G] \to A$ is a homomorphism such that $\phi \circ \pi = f$ then $\phi \circ \pi = \widetilde{f} \circ \pi$ and since π is an epimorphism we have that $\phi = \widetilde{f}$.

This universal mapping property of the abelianization $G/[G,G]$ of G is analogous to that of the tensor product of R-modules.

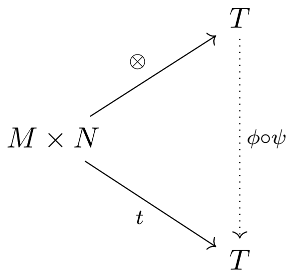
Proposition B.4: *The tensor product of R-modules M and N is unique up to isomorphism.*

Proof. We denote $T := M \otimes_R N$, and let T' along with $t\colon M \times N \to T$ satisfy the universal mapping property of the tensor product. We then get unique linear maps $\psi\colon T \to T'$ and $\phi\colon T' \to T$ such that the following two diagrams commute.





We may combine these into the following commutative diagram.



The linear map $\phi \circ \psi$ is unique and since the identity map $1_T \colon T \to T$ would also fit, we find that $\phi \circ \psi = 1_T$. Therefore, T is unique up to isomorphism. □

Proposition B.5: *For any R-modules M and N, where R is a commutative ring, their tensor product exists.*

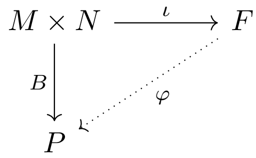
Proof. We begin by looking at $F = F_R(M \times N) := \bigoplus_{(m,n) \in M \times N} R(m,n)$, the free R-module on $M \times N$. We note that elements of F take the form $\sum_{i=1}^{k} r_i(m_i, n_i)$, where $r_i \in R,\ m \in M,\ n \in N$. Note also that the generating elements $(m,n) \in M \times N$ are not elements of $M \oplus N$. Therefore, we see $(m,n)+(m',n) \neq (m+m',n)$ as the left hand side cannot be further simplified. Similarly, $(rm,n) \neq r(m,n) \neq (rm,rn)$. We define D as the submodule of R spanned by the following elements, where $m, m' \in M,\ n, n' \in N, r \in R$

$$(m+m',n)-(m,n)-(m',n), \quad (m,n+n')-(m,n)-(m,n'),$$
$$(rm,n)-(m,rn), \quad r(m,n)-(rm,n), \quad r(m,n)-(m,rn).$$

We may then define the tensor product of M and N as $M \otimes_R N := F/D$, and we write the cosets as $m \otimes n := (m,n) + D$. We may also define the map

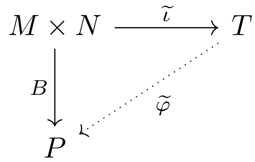
$$\otimes \colon M \times N \to M \otimes_R N$$
$$(m,n) \mapsto m \otimes n.$$

The definition of D defines relations which mean that the map $\otimes$ is bilinear. We denote $T := M \otimes_R N$, and we will show that the pair $(T, \otimes)$ satisfy the universal mapping property of the tensor product. Let P be an R-module and $B \colon M \times N \to P$ a bilinear map. Looking at $M \times N$ as a set and B simply as a function on the set, from the universal mapping property of free modules we have a linear map $\varphi \colon F \to P$ such that the following diagram commutes.



, (B.1)

Where $\iota\colon M \times N \to F$ is the natural injection $(m,n) \mapsto (m,n)$. We would like to show that φ induces a linear map $\widetilde{\varphi}\colon T \to P$, such that the following diagram commutes.



(B.2)

Where $\widetilde{\iota}\colon M \times N \to T$ is given by $(m,n) \mapsto (m,n) + D = m \otimes n$.

From the bilinearity of the map B we have that the following relations hold for all $m, m' \in M,\ n, n' \in N,\ r \in R$

$$B(m+m',n) = B(m,n) + B(m',n), \quad B(m,n+n') = B(m,n) + B(m,n'),$$
$$B(rm,n) = B(m,rn), \quad rB(m,n) = B(rm,n), \quad rB(m,n) = B(m,rn).$$

We also have that $B = \varphi \circ \iota$, and so all the corresponding relations hold for φ on elements of $M \times N$. Since φ is linear, for all $m, m' \in M,\ n, n' \in N,\ r \in R$ we have

$$\varphi(m+m',n) = \varphi((m,n) + (m',n)), \quad \varphi(m,n+n') = \varphi((m,n) - (m,n')),$$
$$\varphi(rm,n) = \varphi(m,rn), \quad r\varphi(m,n) = \varphi(rm,n), \quad r\varphi(m,n) = \varphi(m,rn).$$

Hence, all the generators of D are elements of $\mathrm{Ker}(\varphi)$, and so

$$\widetilde{\varphi}\colon T \to P,$$
$$\sum_{i=1}^{k} r_i m_i \otimes n_i \mapsto \varphi(m_i, n_i),$$

is a well-defined map. Linearity of $\widetilde{\varphi}$ follows from the linearity of φ and similarly the commutativity of (B.2) follows from the commutativity of (B.1).

It only remains to show that $\widetilde{\varphi}$ is the unique linear map which makes (B.2) commute. This follows because the elements $m \otimes n$ for $m \in M$ and $n \in N$ span T. Therefore, linear maps on T are completely determined by their values on these spanning elements. □

B.2 Building Intuition

We now consider some results which will help us to build some intuition for working with tensor products and their elements. We will call elements of $M \otimes N$ ***tensors***, and elements of the form $m \otimes n$ ***elementary tensors***.

In the proof of Proposition B.5 we showed that that tensors are indeed finite linear combinations of elementary tensors. Since $r(m \otimes n) = rm \otimes n$ for all $r \in R$, $m \in M$, $n \in N$ we see that tensors are in fact finite sums of elementary tensors.

Remark B.6: *Let M and N be R-modules. Then $m \otimes 0 = 0 = 0 \otimes n$.*

Proof. For any $m \in M$ we have that $m \otimes 0 = m \otimes (0+0) = m \otimes 0 + m \otimes 0$. Now $0 \otimes n = 0$ follows similarly. ☐

Proposition B.7: *Let M and N be R-modules, for a commutative ring R. If $\{m_i\}_{i \in I}$ and $\{n_j\}_{j \in J}$ are spanning sets for M and N, respectively, then $M \otimes N$ is spanned by $S := \{m_i \otimes n_j \mid i \in I,\ j \in J\}$.*

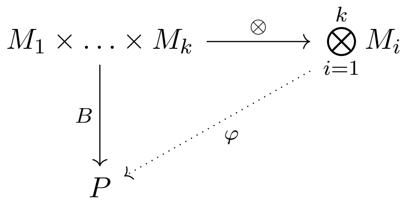
Proof. Since tensors are sums of elementary tensors, we need only show that all elementary tensors are spanned by S. Let $m \otimes n$ be an elementary tensor with $m = \sum\limits_{i \in I} a_i m_i$ and $n = \sum\limits_{j \in J} b_j n_j$. Then by the bilinearity of $\otimes$ we have that

$$m \otimes n = \left(\sum_{i \in I} a_i m_i \right) \otimes \left(\sum_{j \in J} b_j n_j \right) = \sum_{i \in I} \sum_{j \in J} a_i b_j m_i \otimes n_j.$$

☐

Example B.8: *We have seen that linear maps from tensor products are determined by their values on elementary tensors, since tensors are finite sums of elementary tensors. We have to be careful in constructing linear maps from tensor products. For example, if we hoped to construct a function $m \otimes n \mapsto m+n$, then since $-m \otimes -n = m \otimes n$ but (generally) $m+n \neq -m-n$ we could run into trouble. To construct linear maps from tensor products we should use the universal mapping property.*

The 2-fold tensor product may be extended to a k-fold tensor product ($k \in \mathbb{N}$) by the following diagram.



Where $\bigotimes\limits_{i=1}^{k} M_i$ and P are R-modules, B is a multi-linear map, and φ is the unique linear map making the diagram commute.

Example B.9: *Let $a, b \in \mathbb{Z}^+$ and $d = gcd(a, b)$. Then $\mathbb{Z}_a \otimes \mathbb{Z}_b \cong \mathbb{Z}_d$, as abelian groups.*

We will denote $T := \mathbb{Z}_a \otimes \mathbb{Z}_b$. Since 1 spans $\mathbb{Z}_a$ and $\mathbb{Z}_b$ we find from Proposition B.7 that $1 \otimes 1$ spans T. The additive order of $1 \otimes 1$ divides both a and b, and therefore divides d. This follows as a result of the following $a(1 \otimes 1) = a \otimes 1 = 0 \otimes 1 = 0$. Similarly $b(1 \otimes 1) = 0$. Therefore, $|T| \leq d$.

We define $B\colon \mathbb{Z}_a \times \mathbb{Z}_b \to \mathbb{Z}_d$ by $(x \bmod a,\ y \bmod b) \mapsto xy \bmod d$. This is clearly a well-defined bilinear map. By the universal mapping property there is a unique linear map $\varphi\colon T \to \mathbb{Z}_d$ such that $\varphi \circ \otimes = B$. Therefore, for $x \in \mathbb{Z}_d$ we have that $\varphi(x \otimes 1) = x$. Thus φ is an epimorphism, from which we may conclude that $|T| = d$. As we are working with abelian groups we are done.

Remark B.10: *We have seen that tensors are finite sums of elementary tensors. This may tempt us into thinking that we can prove that $f(g(t)) = t$ for all tensors t by showing that $f(g(e)) = e$ for all elementary tensors (for suitable additive maps f and g). It is however wrong to think that proving injectivity on elementary tensors automatically translates to having shown injectivity for all tensors. This comes from the fact that injectivity of a linear map is not an additive identity.*

We demonstrate this point with an example. Consider the additive map $f\colon \mathbb{C} \otimes_{\mathbb{R}} \mathbb{C} \to \mathbb{C}$ given by $z \otimes w \mapsto zw$. We can see that $\{z \otimes w \mid z, w \in \{1, i\}\}$ is a basis for $\mathbb{C} \otimes_R \mathbb{C}$. We also have that $f(1 \otimes i) = f(i \otimes 1)$ but by linear independence we have that $1 \otimes i - i \otimes 1 \neq 0$.

Proposition B.11: *Let M and N be R-modules. Then*

$$\begin{aligned} \psi\colon M \otimes_R N &\to N \otimes_R M \\ m \otimes n &\mapsto n \otimes m, \end{aligned}$$

is an isomorphism, and ψ is unique.

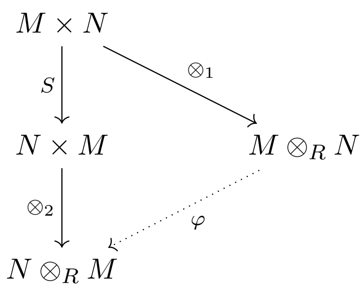
Proof. We define $S\colon M \times N \to N \times M$ by $(m, n) \mapsto (n, m)$. We denote the tensor product maps $\otimes_1\colon M \times N \to M \otimes_R N$ and $\otimes_2\colon N \times M \to N \otimes_R M$. We claim that $\otimes_2 \circ S$ is a bilinear map. Take $m, m' \in M$, $n \in N$, $r, \in R$ and then

$$\otimes_2 \circ S(m + m', n) = n \otimes_2 m + n \otimes_2 m' = (\otimes_2 \circ S)(m, n) + (\otimes_2 \circ S)(m', n),$$

and

$$r(\otimes_2 \circ S)(m, n) = rn \otimes_2 m = n \otimes_2 rm.$$

Thus, $\otimes_2 \circ S$ is bilinear. By the universal mapping property of the tensor product there is a unique linear map $\varphi\colon M \otimes_R N \to N \otimes_R M$ such that the following diagram commutes.



Thus, for $(m, n) \in M \times N$ we have that

$$\varphi(m \otimes_1 n) = (\varphi \circ \otimes_1)(m, n) = (\otimes_2 \circ S)(m, n) = \otimes_2(n, m) = n \otimes_2 m.$$

Which describes the action of φ on elementary tensors, and therefore completely determines φ.

A similar argument yields a unique linear map $\tau\colon N \otimes_R M \to M \otimes_R N$ given by $n \otimes m \mapsto m \otimes n$. We claim these maps are inverses. To this end we take $r_i \in R,\ m_i \in M,\ n_i \in N$ and find that

$$\begin{aligned}(\tau \circ \varphi)\left(\sum_{i=1}^{k} r_i m_i \otimes n_i\right) &= \sum_{i=1}^{k} r_i (\tau \circ \varphi)(m_i \otimes n_i) \\ &= \sum_{i=1}^{k} r_i \tau(n_i \otimes m_i) \\ &= \sum_{i=1}^{k} r_i m_i \otimes n_i.\end{aligned}$$

It follows similarly that $(\varphi \circ \tau)\left(\sum_{i=1}^{k} r_i n_i \otimes m_i\right) = \sum_{i=1}^{k} r_i n_i \otimes n_i$. Hence, we conclude $\tau = \varphi^{-1}$. Uniqueness follows since both φ and τ are unique. □

Proposition B.12: *For R-modules M, N, and P,*

$$\begin{aligned}\varphi\colon (M \otimes N) \otimes P &\to M \otimes (N \otimes P) \\ (m \otimes n) \otimes p &\mapsto m \otimes (n \otimes p),\end{aligned}$$

is an isomorphism, and it is unique.

Proof. We will denote $S := (M \otimes N) \otimes P$ and $T := M \otimes (N \otimes P)$. We define $D\colon M \times N \times P \to T$ by $(m, n, p) \mapsto m \otimes (n \otimes p)$. Since this is tri-linear, each $p \in P$ induces a bilinear map $B_p\colon M \times N \to T$ given by $(m, n) \mapsto m \otimes (n \otimes p)$. By the universal mapping property of tensor products, for each B_p there is a unique linear map $\varphi_p\colon M \otimes N \to T$ such that the following diagram commutes.

$$\begin{array}{ccc} M \times N & \xrightarrow{\otimes} & M \otimes N \\ {\scriptstyle B_p}\downarrow & \swarrow{\scriptstyle \varphi_p} & \\ T & & \end{array}$$

Therefore, $\varphi_p(m \otimes n) = (\varphi_p \circ \otimes)(m, n) = B_p(m, n) = m \otimes (n \otimes p)$. This fully describes φ_p, as it describes the action of φ_p on elementary tensors.

We now define $\Phi\colon (M \otimes N) \times P \to T$ by $\Phi(m \otimes n, p) := \varphi_p(m \otimes n)$. We claim this is biliinear. Since φ_p is linear for each $p \in P$ we know that Φ is linear in the first argument. Taking $m \in M,\ n \in N,\ p, p' \in P,\ r \in R$ we find that

$$\begin{aligned} \varphi_p(m \otimes n) + \varphi_{p'}(m \otimes n) &= m \otimes (n \otimes p) + m \otimes (n \otimes p') \\ &= m \otimes (n \otimes p + n \otimes p') \\ &= m \otimes (n \otimes (p + p')) = \varphi_{p+p'}(m \otimes n), \end{aligned}$$

and

$$\begin{aligned} r\varphi_p(m \otimes n) &= rm \otimes (n \otimes p) = m \otimes (r(n \otimes p)) = m \otimes (n \otimes rp) \\ &= \varphi_{rp}(m \otimes n). \end{aligned}$$

Thus, Φ is also linear in the second argument and therefore bilinear. By the universal mapping property of tensor products there is a unique linear map $\tau\colon S \to T$ such that the following diagram commutes.

$$\begin{array}{ccc} (M \otimes N) \times P & \xrightarrow{\otimes} & S \\ {\scriptstyle \Phi}\downarrow & \swarrow{\scriptstyle \tau} & \\ T & & \end{array}$$

Hence, we can see that the action of τ on elementary tensors is

$$\begin{aligned} \tau\left(\left(\sum_{i=1}^{k} r_i m_i \otimes n_i\right) \otimes p\right) &= (\tau \circ \otimes)\left(\sum_{i=1}^{k} r_i m_i \otimes n_i, p\right) \\ &= \Phi\left(\sum_{i=1}^{k} r_i m_i \otimes n_i, p\right) \\ &= \sum_{i=1}^{k} r_i \Phi(m_i \otimes n_i, p) = \sum_{i=1}^{k} r_i \varphi_p(m_i \otimes n_i) \\ &= \sum_{i=1}^{k} r_i m_i \otimes (n_i \otimes p). \end{aligned}$$

A similar argument yields a unique linear map $\psi\colon T \to S$ whose action on elementary tensors is

$$\psi\left(m \otimes \left(\sum_{i=1}^{k} r_i n_i \otimes p_i\right)\right) = \sum_{i=1}^{k} r_i (m \otimes n_i) \otimes p_i.$$

Since S and T are spanned by elementary tensors, and τ and ψ are linear and unique the result will follow if we can show that the actions of $\tau \circ \psi$ and $\psi \circ \tau$ on elementary tensors are trivial. By taking $r_i \in R,\ m_i \in M,\ n_i \in N,\ p \in P$ we can see

$$\begin{aligned}(\psi \circ \tau)\left(\left(\sum_{i=1}^{k} r_i m_i \otimes n_i\right) \otimes p\right) &= \psi\left(\sum_{i=1}^{k} r_i m_i \otimes (n_i \otimes p)\right)\\ &= \sum_{i=1}^{k} r_i \psi(m_i \otimes (n_i \otimes p))\\ &= \sum_{i=1}^{k} r_i (m_i \otimes n_i) \otimes p\\ &= \left(\sum_{i=1}^{k} r_i m_i \otimes n_i\right) \otimes p.\end{aligned}$$

A similar argument holds for $\tau \circ \psi$, and so we are done. □

B.3 Direct Sums and Tensor Maps

Having built up an idea of tensors in the previous section we move onto their interactions with direct sums. We will see that tensor products' interactions with direct sums of R-modules follow much as we would expect and hope. Thereafter we will see how tensor products may induce R-linear maps from existing R-linear maps between their 'component' modules.

Proposition B.13: *Let M, N, and P be R-modules. Then*

$$\begin{aligned}\varphi\colon M \otimes (N \oplus P) &\to (M \otimes N) \oplus (M \otimes P)\\ m \otimes (n \oplus p) &\mapsto (m \otimes n) + (m \otimes p)\end{aligned}$$

is a unique isomorphism.

Proof. There is a bilinear map $\psi\colon M \times (N \oplus P) \to (M \otimes N) \oplus (M \otimes N)$ which endows the pair $((M \otimes N) \oplus (M \otimes P), \psi)$ with the universal mapping property of the tensor product. We define this map by $\psi(m, n+p) := (m \otimes n) + (m \otimes p)$. To see that ψ is bilinear we take $r, r' \in R,\ m, m' \in M,\ n, n' \in N,\ p, p' \in P$

and note that

$$\begin{aligned}\psi(rm + r'm', n + p) &= ((rm + r'm') \otimes n) + ((rm + r'm') \otimes p) \\ &= r((m \otimes n) + (m \otimes p)) + r'((m' \otimes n) + (m' \otimes p)) \\ &= r\psi(m, n + p) + r'\psi(m', n + p),\end{aligned}$$

which shows that B is linear in the first argument. Moreover, we have that

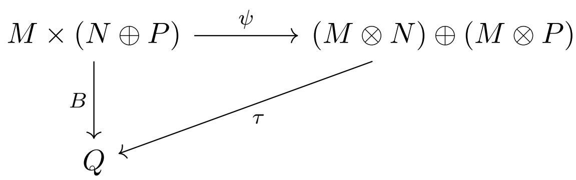
$$\begin{aligned}&\psi(m, (n + p) + (n' + p')) = \psi(m, n + n' + p + p') \\ &= (m \otimes (n + n')) + (m \otimes (p + p')) = (m \otimes n + m \otimes n') + (m \otimes p + m \otimes p') \\ &= ((m \otimes n) + (m \otimes p)) + ((m \otimes n') + (m \otimes p')) \\ &= \psi(m, n + p) + \psi(m, n' + p'),\end{aligned}$$

and,

$$\begin{aligned}\psi(m, r(n + p)) &= \psi(m, rn + rp) = (m \otimes rn) + (m \otimes rp) \\ &= r((m \otimes n) + (m \otimes p)) = r\psi(m, n + p),\end{aligned}$$

which shows that ψ is linear in the second argument, and therefore bilinear.

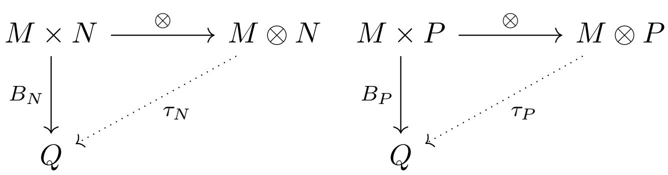
Now we would like to show that the universal mapping property holds. Suppose $B\colon M \times (N \oplus P) \to Q$ is a bilinear map. We need to construct a linear map $\tau\colon (M \otimes N) \oplus (M \otimes P) \to Q$ which makes the following diagram commute.



Since we are trying to construct a linear map, τ would be determined by its values on elements of the form $m \otimes n + 0$ and $0 + m \otimes p$. By the commutativity of the above diagram this forces $\tau(m \otimes n) = (\tau \circ \psi)(m, n + 0) = B(m, n + 0)$, and $\tau(m \otimes p) = (\tau \circ \psi)(m, 0 + p) = B(m, 0 + p)$. This induces two bilinear maps

$$\begin{aligned}B_N\colon M \times N &\to Q \\ (m, n) &\mapsto B(m, n + 0)\end{aligned} \qquad \begin{aligned}B_P\colon M \times P &\to Q \\ (m, p) &\mapsto B(m, 0 + p).\end{aligned}$$

Using the universal mapping property of tensor products twice, we find two unique linear maps $\tau_N\colon M \otimes N \to Q$ and $\tau_P\colon M \otimes P \to Q$ such that the following diagrams commute.



We then define

$$\tau\colon (M \otimes N) \oplus (M \otimes P) \to Q$$
$$\sum_i m_i \otimes n_i + \sum_j m'_j \otimes p_j \mapsto \sum_i \tau_N(m_i \otimes n_i) + \sum_j \tau_P(m'_j \otimes p_j),$$

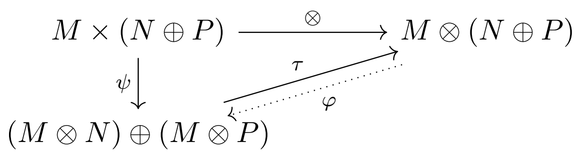
that is, $\tau = \tau_N + \tau_P$. This shows that τ is linear.

Now we can see that for $(m, rn + sp) \in M \times (N \oplus P)$ we have

$$\begin{aligned}(\tau \circ \psi)(m, rn + sp) &= \tau((m \otimes rn) + (m \otimes sp)) = \tau_N(m \otimes rn) + \tau_P(m \otimes sp) \\ &= B_N(m, rn) + B_P(m, sp) \\ &= B(m, rn) + B(m, sp) = B(m, rn + sp).\end{aligned}$$

Uniqueness follows from τ_N and τ_P being unique. This shows that the pair $((M \otimes N) \oplus (M \otimes P), \tau)$ has the universal mapping property, and therefore is isomorphic to $M \otimes (N \oplus P)$.

In the special case that $Q = M \otimes (N \oplus P)$ and $B = \otimes$ we see that the universal mapping property guarantees the existence of a unique linear map φ which makes the following diagram commute.



By uniqueness we may conclude that φ is an isomorphism and that $\varphi = \tau^{-1}$. Moreover, since $\varphi \circ \otimes = \psi$, we can see that φ acts on elementary tensors as

$$\begin{aligned}\varphi(m \otimes (rn + sp)) &= (\varphi \circ \otimes)(m, rn + sp) = \psi(m, rn + sp) \\ &= (m \otimes rn) + (m \otimes sp).\end{aligned}$$

□

Proposition B.14: *Let M and N_i, for $i \in I$ be R-modules, where I is an arbitrary indexing set. Then,*

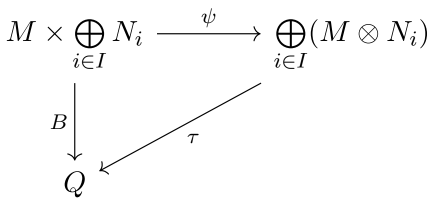
$$\varphi\colon M \otimes \bigoplus_{i \in I} N_i \to \bigoplus_{i \in I} (M \otimes N_i)$$
$$m \otimes (n_i)_{i \in I} \mapsto (m \otimes n_i)_{i \in I},$$

is an isomorphism. We note that there are only finitely many non-zero components of any element in the codomain of ψ.

Proof. We will construct a bilinear map $\psi\colon M \times \bigoplus_{i\in I} N_i \to \bigoplus_{i\in I}(M \otimes N_i)$ which satisfies the universal mapping property of the tensor product. We define

$$\begin{aligned}\psi\colon M \times \bigoplus_{i\in I} N_i &\to \bigoplus_{i\in I}(M \otimes N_i)\\ (m, (n_i)_{i\in I}) &\mapsto (m \otimes n_i)_{i\in I}.\end{aligned}$$

Bilinearity follows from a similar argument to that in the proof of Proposition B.13. We now like to show that ψ satisfies the universal mapping property, that is, for any bilinear map $B\colon M \times \bigoplus_{i\in I} N_i \to Q$ we want to show that there exists a unique linear map τ which makes the following diagram commute.



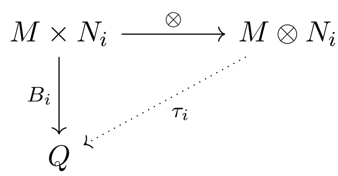
Since we need τ to be a linear map, it would be determined by its action on elements of the form $(m \otimes n_i)$. We require that

$$\tau(m \otimes n_i) = (\tau \circ \psi)(m, n_i) = B(m, n_i),$$

for each $i \in I$. This leads to the construction of bilinear maps which, by a slight abuse of notation, we will write as

$$\begin{aligned}B_i\colon M \times N_i &\to Q\\ (m, n_i) &\mapsto B(m, n_i),\end{aligned}$$

for each $i \in I$. The universal mapping property then gives a unique linear map τ_i, for each $i \in I$, such that the following diagram commutes.



That is, by a slight abuse of notation, $\tau_i(m \otimes n_i) = B_i(m, n_i)$. We can then define

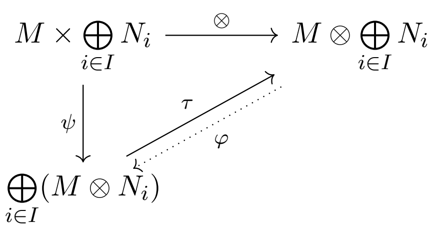
$$\begin{aligned}\tau\colon \bigoplus_{i\in I}(M \otimes N_i) &\to Q\\ (m \otimes n_i)_{i\in I} &\mapsto \sum_{i\in I} \tau_i(m, n_i).\end{aligned}$$

We note that since only finitely many of the n_i are non-zero the sum makes sense. By construction, τ is linear. Now we can see that

$$(\tau \circ \psi)(m, (n_i)_{i\in I}) = \tau((m \otimes n_i)_{i\in I}) = \sum_{i\in I} \tau_i(m \otimes n_i) = \sum_{i\in I} B_i(m, n_i)$$
$$= B(m, (n_i)_{i\in I}).$$

Uniqueness follows from the uniqueness of each of the τ_i. Thus, τ satisfies the universal mapping property.

In the special case that $Q = M \otimes \bigoplus_{i\in I} N_i$ and $B = \otimes$ we see that the universal mapping property guarantees the existence of a unique linear map φ which makes the following diagram commute.



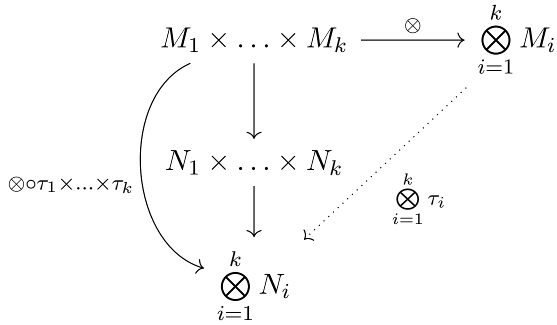
By uniqueness we may conclude that φ is an isomorphism and $\tau = \varphi^{-1}$. Moreover, since $\varphi \circ \otimes = \psi$ we can see that φ acts on elementary tensors as

$$\varphi(m \otimes (n_i)_{i\in I}) = (\varphi \circ \otimes)(m, (n_i)_{i\in I}) = \psi(m, (n_i)_{i\in I}) = (m \otimes n_i)_{i\in I}.$$

□

Remark B.15: *We note (without proof) that the analogue of φ in the above proposition with direct products is a linear map but it may not be an isomorphism.*

Proposition B.16: *Let $(\tau_i \colon M_i \to N_i)_{i=1}^k$ be a family of linear maps between modules. Then, there is a unique linear map $\bigotimes_{i=1}^k \tau_i \colon \bigotimes_{i=1}^k M_i \to \bigotimes_{i=1}^k N_i$, which makes the following diagram commute.*



Moreover, $\bigotimes_{i=1}^{k} \tau_i$ *is completely determined by*

$$\begin{aligned} \bigotimes_{i=1}^{k} \tau_i \colon \bigotimes_{i=1}^{k} M_i &\to \bigotimes_{i=1}^{k} N_i \\ m_1 \otimes \ldots \otimes m_k &\mapsto \tau_1(m_1) \otimes \ldots \otimes \tau_k(m_k). \end{aligned}$$

Proof. We note that the map

$$\begin{aligned} \tau_1 \times \ldots \times \tau_k \colon M_1 \times \ldots \times M_k &\to N_1 \times \ldots \times N_k \\ (m_1, \ldots, m_k) &\mapsto (\tau_1(m_1), \ldots, \tau_k(m_k)), \end{aligned}$$

is multilinear by construction. Since $\otimes$ is linear their composition is bilinear. By the universal mapping property of the tensor product there is a unique linear map $\bigotimes_{i=1}^{k} \tau_i \colon \bigotimes_{i=1}^{k} M_i \to \bigotimes_{i=1}^{k} N_i$, such that

$$\bigotimes_{i=1}^{k} \tau_i \circ \otimes = \otimes \circ \tau_1 \times \ldots \times \tau_k.$$

Therefore, the action of $\bigotimes_{i=1}^{k} \tau_i$ on elementary tensors is

$$\begin{aligned} \bigotimes_{i=1}^{k} \tau_i(m_1 \otimes \ldots \otimes m_k) &= (\bigotimes_{i=1}^{k} \tau_i \circ \otimes)(m_1, \ldots, m_k) \\ &= (\otimes \circ \tau_1 \times \ldots \times \tau_k)(m_1, \ldots, m_k) \\ &= \tau_1(m_1) \otimes \ldots \otimes \tau_k(m_k). \end{aligned}$$

□

Corollary B.17: *Let M, N, and P be R-modules.*

(i) If $\tau_i \colon M_i \to N_i$ and $\psi_i \colon N_i \to P_i$ are linear maps for all $1 \leq i \leq k$, then $\bigotimes_{i=1}^{k} \psi_i \circ \bigotimes_{i=1}^{k} \tau_i = \bigotimes_{i=1}^{k} (\psi_i \circ \tau_i)$.

(ii) $\bigotimes_{i=1}^{k} id_{M_i} = id_{\bigotimes_{i=1}^{k} M_i}$.

Proof.

(i): It suffices to prove the result for elementary tensors. We have that

$$\begin{aligned}
&\left(\bigotimes_{i=1}^{k}\psi_i \circ \bigotimes_{i=1}^{k}\tau_i\right)(m_1 \otimes \ldots \otimes m_k) \\
&= \bigotimes_{i=1}^{k}\psi_i(\tau_1(m_1) \otimes \ldots \otimes \tau_k(m_k)) \\
&= (\psi_1 \circ \tau_1)(m_1) \otimes \ldots \otimes (\psi_k \circ \tau_k)(m_k) \\
&= \bigotimes_{i=1}^{k}(\psi_i \circ \tau_i)(m_1 \otimes \ldots \otimes m_k).
\end{aligned}$$

(ii): Since $\bigotimes_{i=1}^{k} id_{M_i}$ fixes the elementary tensors, it fixes all tensors.

□

Corollary B.18: *Let $\tau_i \colon M_i \to N_i$ be linear maps of modules for $1 \leq i \leq k$.*

(i) If each τ_i is an isomorphism, then $\bigotimes_{i=1}^{k} \tau_i$ is an isomorphism.

(ii) If each τ_i is surjective, then $\bigotimes_{i=1}^{k} \tau_i$ is surjective.

Proof.

(i): We claim that $\bigotimes_{i=1}^{k}(\tau_i^{-1}) = \left(\bigotimes_{i=1}^{k} \tau_i\right)^{-1}$. From the proposition we have that

$$\bigotimes_{i=1}^{k}(\tau_i^{-1}) \circ \bigotimes_{i=1}^{k}\tau_i = \bigotimes_{i=1}^{k}(\tau_i^{-1} \circ \tau_i) = \bigotimes_{i=1}^{k} id_{M_i} = id_{\bigotimes_{i=1}^{k} M_i}.$$

Similarly we find that

$$\bigotimes_{i=1}^{k}\tau_i \circ \bigotimes_{i=1}^{k}(\tau_i^{-1}) = id_{\bigotimes_{i=1}^{k}}.$$

(ii): It suffices to show that $\bigotimes_{i=1}^{k} \tau_i$ is surjective on elementary tensors. Fix $n_1 \otimes \ldots \otimes n_k \in \bigotimes_{i=1}^{k} N_i$. Since each τ_i is surjective, for each i there is an $m_i \in M_i$ such that $\tau_i(m_i) = n_i$. Therefore,

$$\bigotimes_{i=1}^{k}\tau_i(m_1 \otimes \ldots \otimes m_k) = n_1 \otimes \ldots \otimes n_k.$$

□

Appendix C

Affine Group Schemes

Affine group schemes and their associated Lie algebras (much like Lie groups and their Lie algebras). Moreover, there is a strong connection between affine group schemes and G-gradings of a Lie algebra. This also gives a connection between Lie algebras and associative algebras. We will not pursue these connections directly. Rather, present an introduction to affine group schemes which serve as a base for future investigations.

In this chapter we will assume all vector spaces are defined over a field k, unless specified otherwise. We denote by $\mathbf{Alg_k}$ the category of commutative unital k-algebras.

C.1 Representable Functor

We begin by looking at some results concerning representations of polynomial functors.

Definition C.1: *A functor* $F\colon \mathscr{C} \to \mathbf{Set}$ *is* **representable** *if there is an object* $C \in \mathscr{C}$ *such that there is a natural isomorphism*

$$\tau\colon F \xrightarrow{\sim} \operatorname{Hom}_{\mathscr{C}}(C, -).$$

In such a case we call C *a* **representation** *of* F.

Remark C.2: *We show here how a family of polynomials may define a functor. Let* $S \subseteq k[X_1, \ldots, X_n]$*. We will use* S *to construct a functor* $F_S\colon \mathbf{Alg_k} \to \mathbf{Set}$.

Take $R \in \mathbf{Alg_k}$*. We set*

$$F_S(R) := \{r = (r_1, \ldots, r_n) \in R^n \mid f(r) = 0 \quad \forall f \in S\},$$

which is called the **zero-set** *of* S *in* R^n*. Now we take an algebra homomorphism* $f\colon R \to R'$*, and define*

$$\begin{aligned} F_S(f)\colon F_S(R) &\to F_S(R') \\ r = (r_1, \ldots, r_n) &\mapsto (f(r_1), \ldots, f(r_n)). \end{aligned}$$

Clearly $F_S(1_R) = 1_{F_S(R)}$. Moreover, we can see that F_S preserves composition of algebra homomorphisms and hence F_S is a functor. We will call F_S the **polynomial functor relative to** *S.*

Definition C.3: *Let S be a finite subset of $k[X_1, \dots, X_n]$. We define the* **polynomial algebra relative to** *S as*

$$A_s := k[X_1, \dots, X_n]/\langle S\rangle,$$

where $\langle S\rangle = \{\sum_{i\in I} p_i s_i \,|\, p_i \in k[X_1, \dots, X_n],\ s_i \in S,\ |I| < \infty\}$ is the ideal generated by S.

Definition C.4: *An algebra A is* **finitely presented** *if A is isomorphic to some algebra of the form $k[X_1, \dots, X_n]/\langle S\rangle$, where S is a finite subset of $k[X_1, \dots, X_n]$.*

Lemma C.5: *Let $\tau\colon \mathrm{Hom}(A, -) \to \mathrm{Hom}(B, -)$ be a natural transformation, for objects A and B in a category $\mathscr{C}$. Then $\tau_A(1_A)\colon B \to A$ is the unique arrow in $\mathscr{C}$ such that $\tau = \mathrm{Hom}(\tau_A(1_A), -)$.*

Proof. For any $X \in \mathscr{C}$ and $f\colon A \to X$ we have that the following diagram commutes.

$$\begin{array}{ccc} \mathrm{Hom}(A, A) & \xrightarrow{\mathrm{Hom}(A,f)} & \mathrm{Hom}(A, X) \\ \downarrow{\scriptstyle \tau_A} & & \downarrow{\scriptstyle \tau_X} \\ \mathrm{Hom}(B, A) & \xrightarrow[\mathrm{Hom}(B,f)]{} & \mathrm{Hom}(B, X) \end{array}$$

Therefore we can see that

$$\begin{aligned} \tau_X(f) = \tau_X(f1_A) &= \tau_X \mathrm{Hom}(A, f)(1_A) = \mathrm{Hom}(B, f)\tau_A(1_A) \\ &= f\tau_A(1_A) = \mathrm{Hom}(\tau_A(1_A), X)(f). \end{aligned}$$

Since f and X are arbitrary we find that $\tau = \mathrm{Hom}(\tau_A(1_A), -)$.

To show uniqueness, suppose $\mathrm{Hom}(g, -) = \mathrm{Hom}(h, -)$ for some $\mathscr{C}$-arrows $g, h\colon B \to A$. Then

$$g = \mathrm{Hom}(g, A)(1_A) = \mathrm{Hom}(h, A)(1_A) = h.$$

□

Lemma C.6: *Let R and R' be representations of a functor $F\colon \mathscr{C} \to$ **Set**. Then $R \cong R'$.*

Proof. There is a natural isomorphism $\tau\colon \mathrm{Hom}(R, -) \to \mathrm{Hom}(R', -)$. An application of Lemma C.5 yields that $\tau = \mathrm{Hom}(\tau_R(1_R), -)$. The Yoneda embedding $\mathscr{Y}: \mathscr{C} \to \mathbf{Set}^{\mathscr{C}}$ is fully faithful and dually its inverse is also fully faithful (the fact that $\mathscr{Y}^{-1}$ is faithful also follows from the uniqueness claim in Lemma C.5). Fully faithful functors reflect isomorphisms. Therefore, since we have that $\mathscr{Y}^{-1}(R) = \mathrm{Hom}(R, -)$ and $\mathscr{Y}^{-1}(R') = \mathrm{Hom}(R', -)$ and a natural isomorphism τ, it follows that $R \cong R'$. □

Proposition C.7: *Let S be a finite subset of $k[X_1, \ldots, X_n]$. Then F_S is naturally isomorphic to $\mathrm{Hom}(A_S, -)$, where F_S and A_S are the polynomial-functor and -algebra relative to S, respectively. Morover, each finitely presented algebra arises from such a family of polynomials.*

Proof. We first consider the direct case. We define a natural transformation $\eta\colon F_S \to \mathrm{Hom}(A_s, -)$ where the R-component η_R maps $(z_i)_{1\leq i\leq n} \in F_S(R)$ to the homomorphism $\phi\colon A_S \to R$ such that $\phi(X_i) = z_i$ for all $1 \leq i \leq n$. It is clear that η_R is injective.

On the other hand we may also define a related natural transformation $\tau\colon \mathrm{Hom}(A_S, -) \to F_S$ by

$$\begin{aligned} \tau_R\colon\ & \mathrm{Hom}(A_S, R) \to F_S(R) \\ & \phi\colon A_S \to R \mapsto (\phi(X_1), \ldots, \phi(X_n)), \end{aligned}$$

that is, $\tau_R(\phi) = (\phi(X_i))_{1\leq i\leq n} \in R^n$. We remark that $F_{\langle S\rangle}(R) = F_S(R)$. By construction each η_R is the inverse of each τ_R. This also shows that the image of τ_R lies entirely in $F_S(R)$.

Through diagram chasing we can see that the following diagram is commutative, and thus we have a natural isomorphism.

$$\begin{array}{ccc} \mathrm{Hom}(A_S, R) & \xrightarrow{\mathrm{Hom}(A_S, f)} & \mathrm{Hom}(A_S, R') \\ \Big\downarrow{\scriptstyle \tau_R} & & \Big\downarrow{\scriptstyle \tau_{R'}} \\ F_S(R) & \xrightarrow[F_S(f)]{} & F_S(R') \end{array}$$

We move our attention to the converse case. If A is a finitely presented algebra, then A is isomorphic to an algebra of the form $k[X_1, \ldots, X_n]/\langle S\rangle$, where S is a finite subset of $k[X_1, \ldots, X_n]$. □

Definition C.8: *Let $G\colon$ **Alg$_\mathbf{k}$** $\to$ **Grp** be a functor. If G composed with the forgetful functor,* **Grp** $\to$ **Set**, *is representable by a finitely presented algebra, then G is an* **affine group scheme** *or* **affine algebraic group**.

Example C.9:

1. *SL_2 is represented by $k[X_{11}, X_{12}, X_{21}, X_{22}]/\langle X_{11}X_{22} - X_{12}X_{21} - 1\rangle$.*

2. *GL_1 (the multiplicative group) is represented by $k[X, Y]/\langle XY - 1\rangle$.*

3. GL_2 is represented by

$$k[X_{11}, X_{12}, X_{21}, X_{22}, Y]/\langle (X_{11}X_{22} - X_{12}X_{21})Y - 1\rangle.$$

4. For any commutative unital ring we may consider the nth roots of unity $\mu_n(R) = \{x \in R \mid x^n = 1\}$. Then μ_n is represented by $k[X]/\langle X^n - 1\rangle$.

Remark C.10: *The functor which assigns each $R \in \mathbf{Alg_k}$ to a one-element set (for example, $R \mapsto \{R\}$) is represented by k. This follows from k being initial in $\mathbf{Alg_k}$.*

Remark C.11: *Suppose the functors $E, F\colon \mathbf{Alg_k} \to \mathbf{Set}$ are represented by $A, B \in \mathbf{Alg_k}$, respectively. The product of E and F is the functor*

$$\begin{aligned} E \times F\colon \mathbf{Alg_k} &\to \mathbf{Set} \\ R &\mapsto E(R) \times F(R) \\ g\colon R \to R' &\mapsto E(g) \times F(g), \end{aligned}$$

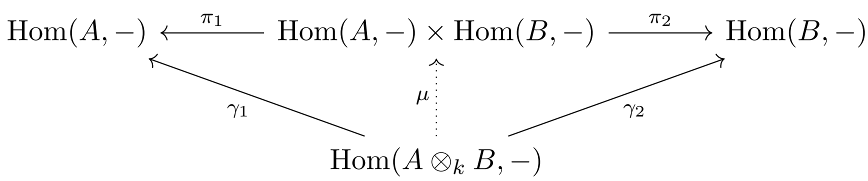
where $(E(g) \times F(g))(e, f) = (g(e), g(f))$, for all $e \in E(R)$ and $f \in F(R)$. We claim $E \times F$ is represented by $A \otimes_k B$.

The assumption that $E \cong \mathrm{Hom}(A, -)$ and $F \cong \mathrm{Hom}(B, -)$ yields that $E \times F \cong \mathrm{Hom}(A, -) \times \mathrm{Hom}(B, -)$. Therefore, it suffices to show that $\mathrm{Hom}(A, -) \times \mathrm{Hom}(B, -) \cong \mathrm{Hom}(A \otimes_k B, -)$. For the sake of convenience we will denote $A' := \mathrm{Hom}(A-)$, $B' := \mathrm{Hom}(B-)$, and $P := \mathrm{Hom}(A \otimes_k B, -)$.

We define natural transformations $\gamma_1\colon P \to A'$ and $\gamma_2\colon P \to B'$ as follows: for $(g\colon A \otimes_k B \to R) \in P$ we set

$$\gamma_{1,R}(g)(a) := g(a \otimes 1), \quad \gamma_{2,R}(g)(b) = g(1 \otimes b),$$

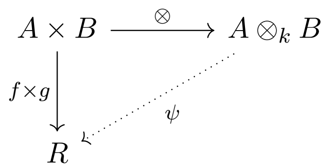
for all $a \in A$, $b \in B$. By the universal property of products there is a unique natural transformation $\mu\colon P \to A' \times B'$ such that the following diagram commutes.



We will now define a natural transformation $\nu\colon A' \times B' \to P$ which we claim is the inverse of μ. For $R \in \mathbf{Alg_k}$ we define the R-component of ν as follows: for a pair of morphisms $(f, g) \in A'(R) \times B'(R)$ we may consider the bilinear map

$$\begin{aligned} f \times g\colon A \times B &\to R \\ (a, b) &\mapsto f(a)g(b). \end{aligned}$$

The universal property of the tensor product yields the existence of a unique k-linear map $\psi\colon A \otimes_k B \to R$ such that the following diagram commutes.



We define $\nu_R(f,g) := \psi$. By uniqueness we have that $\mu\nu = 1$. For $f \in P(R)$ we have that $f = \nu_R(\gamma_1 f, \gamma_2 f)$. Therefore, we find that

$$\nu_R\mu_R(f) = \nu_R\mu_R\nu_R(\gamma_1 f, \gamma_2 f) = \nu_R(\gamma_1 f, \gamma_2 f) = f.$$

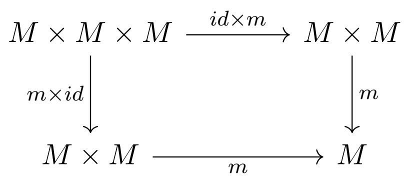
Hence μ is a natural isomorphism.

C.2 Monoid and Group Objects

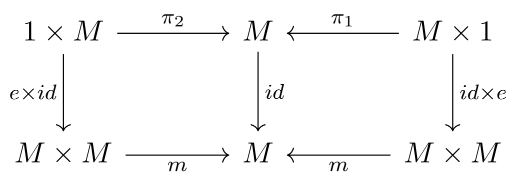
This section is centred around monoid and group objects in a category. We will link this to the work on representations of polynomial functors from the previous section.

Definition C.12: *Let $\mathscr{C}$ be a category with all finite limits, specifically $\mathscr{C}$ has a terminal object, 1. A* **monoid** *in $\mathscr{C}$ is a triple (M, m, e) consisting of an object M in $\mathscr{C}$, and arrows $m\colon M \times M \to M$ and $e\colon 1 \to M$ satisfying the following conditions.*

1. *The following diagram commutes (associativity).*



2. *The following diagram commutes (identity).*



Where π_2 and π_1 are product projections. It is clear to see that π_1 and π_2 are isomorphisms in this case due to the properties of the product.

Definition C.13: *Let $\mathscr{C}$ be a category with all finite limits, specifically $\mathscr{C}$ has a terminal object, 1. A* **group** *in $\mathscr{C}$ is a pair (G, m) consisting of an object G in $\mathscr{C}$ and an arrow $m\colon G \times G \to G$ such that there are arrows $e\colon 1 \to G$ and $i\colon G \to G$ which make (G, m, e) a monoid in $\mathscr{C}$ and for which the following diagram commutes.*

$$\begin{array}{ccccc} G & \xrightarrow{(i,id)} & G \times G & \xleftarrow{(id,i)} & G \\ \downarrow & & \downarrow{\scriptstyle m} & & \downarrow \\ 1 & \xrightarrow[e]{} & G & \xleftarrow[e]{} & 1 \end{array}$$

Definition C.14: *A* **group functor** *is a functor* $G\colon \mathbf{Alg_k} \to \mathbf{Grp}$.

Remark C.15: *Suppose* $G\colon \mathbf{Alg_k} \to \mathbf{Grp}$ *is a group functor. For each* $(f\colon R \to R') \in \mathbf{Alg_k}$ *the associated map* $G(f)\colon G(R) \to G(R')$ *is a homomorphism. This is equivalent to the following diagram being commutative.*

$$\begin{array}{ccc} G(R) \times G(R) & \xrightarrow{G(f)\times G(f)} & G(R') \times G(R') \\ {\scriptstyle m_R}\downarrow & & \downarrow{\scriptstyle m_{R'}} \\ G(R) & \xrightarrow[G(f)]{} & G(R') \end{array}$$

This shows that $m\colon G \times G \to G$ *is a natural transformation. Similarly, we find that* $e\colon 1_{\mathbf{Alg_k}} \to G$ *and* $i\colon G \to G$ *are natural transformations. Therefore a group functor is simply a set functor with natural transformations* m, e, *and* i *which make* G *a group object.*

Proposition C.16: *Suppose* $E, F, G\colon \mathbf{Alg_k} \to \mathbf{Grp}$ *are functors represented by* A, B, *and* C, *respectively. For natural transformations* $\alpha\colon E \to G$ *and* $\beta\colon F \to G$ *the pullback is*

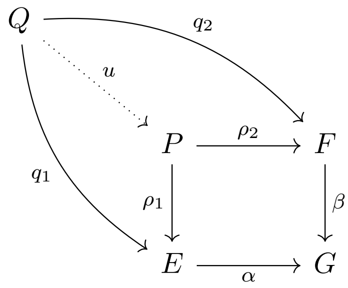
$$\begin{aligned} (E \times_G F)(R) &= \{(e,f) \mid e \in E(R),\ f \in F(R),\ \alpha_R(e) = \beta_R(f)\}, \\ (E \times_G F)(g)(e,f) &= (E(g)(e), F(g)(f)), \end{aligned}$$

for each $R \in \mathbf{Alg_k}$ *and* $(g\colon R \to R') \in \mathbf{Alg_k}$.

Proof. We denote by $P := E \times_G F$. Take $(e,f) \in P(R)$ and a morphism $(g\colon R \to R') \in \mathbf{Alg_k}$, we claim that $P(g)(e,f) \in P(R')$. Since α and β are natural transformations we find that

$$\alpha_{R'} E(g)(e) = G(g)\alpha_R(e) = G(g)\beta_R(f) = \beta_{R'} F(g)(f).$$

Now it follows that P is a functor exactly because E and F are functors. We denote by $\rho_1\colon P \to E$ and $\rho_2\colon P \to F$ the obvious projections. To show that P is a pullback we need to show that for any triple (Q, q_1, q_2), where $\alpha q_1 = \beta q_2$, there is a unique natural transformation $u\colon Q \to P$ such that following diagram commutes.



We define the R-component of u as

$$\begin{aligned} u_R \colon Q(R) &\to P(R) \\ x &\mapsto (q_{1,R}(x), q_{2,R}(x)). \end{aligned}$$

Then, for $g \colon R \to R' \in \mathbf{Alg_k}$ and $x \in Q(R)$ we have

$$\begin{aligned} P(g)u_R(x) = P(g)(q_{1,R}(x), q_{2,R}(x)) &= (E(g)q_{1,R}(x), F(g)q_{2,R}(x)) \\ &= (q_{1,R'}Q(g)(x), q_{2,R'}Q(g)(x)) = u_{R'}Q(g)(x), \end{aligned}$$

where the penultimate equality follows from q_1 and q_2 being natural transformations. Hence, it follows that u is a natural transformation. By construction we have that $\rho_i u = q_i$. It remains to show that u is unique in this context. If $\tau \colon Q \to P$ is a natural transformation which would also make the above diagram commute, then $\tau_R(x) = (q_{1,R}(x), q_{2,R}(x)) = u_R(x)$. □

Lemma C.17: *Let $A, B, C \in \mathbf{Alg_k}$ and let $\alpha \colon \mathrm{Hom}(A, -) \to \mathrm{Hom}(C, -)$ and $\beta \colon \mathrm{Hom}(B, -) \to \mathrm{Hom}(C, -)$ be natural transformations. If we set $Q := \{(\alpha_A(1_A)(c) \otimes 1 - 1 \otimes \beta_B(1_B)(c)) \in A \otimes_k B \mid c \in C\}$ then,*

$$A \otimes_C B := \frac{(A \otimes_k B)}{\langle Q \rangle},$$

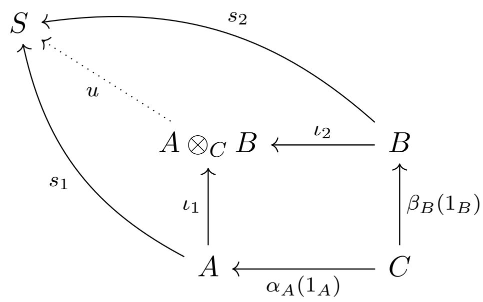
is the pushout of the morphisms $\alpha_A(1_A)$ and $\beta_B(1_B)$.

Proof. An application of Lemma C.5 yields that $\alpha_A(1_A)$ and $\beta_B(1_B)$ are the unique arrows such that $\alpha = \mathrm{Hom}(\alpha_A(1_A), -)$ and $\beta = \mathrm{Hom}(\beta_B(1_B), -)$. We note here that A is a C-module via the following map

$$\begin{aligned} C \times A &\to A \\ (c, a) &\mapsto \alpha_A(1_A)(c)a. \end{aligned}$$

Similarly, B is a C-module.

We need to prove that given any triple (S, s_1, s_2), for which it holds that $s_1\alpha_A(1_A) = s_2\beta_B(1_B)$, there is a unique arrow $u \colon A \otimes_C B \to S$ such that the following diagram commutes.



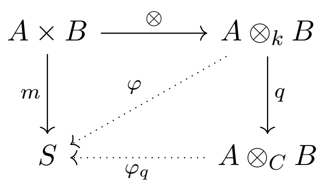
We define u as follows

$$\begin{aligned} u\colon A \otimes_C B &\to S \\ (a \otimes b) &\mapsto s_1(a) + s_2(b). \end{aligned}$$

Since s_1 and s_2 are k-linear maps the map

$$\begin{aligned} m\colon A \times B &\to S \\ (a, b) &\mapsto s_1(a)s_2(b), \end{aligned}$$

is bilinear. By the universal property of the tensor product there is a unique k-linear map $\varphi\colon A \otimes_k B \to S$ such that $\varphi\otimes = d$. Moreover, this induces a k-linear map $\varphi_q\colon A \otimes_C B$ defined as $\varphi_q([a \otimes b]) = \varphi(a \otimes b)$, which is well-defined exactly because (S, s_1, s_2) is a cocone.



We set $u := \varphi_q$ and the result follows. □

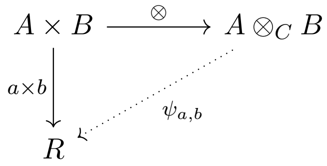
Proposition C.18: *Suppose $E, F, G\colon$ **Alg$_\mathbf{k}$** $\to$ **Grp** are functors represented by A, B, and C, respectively. Let $\alpha\colon E \to G$ and $\beta\colon F \to G$ be natural transformations corresponding to morphisms $C \to A$ and $C \to B$. Then the pullback $E \times_G F$ is represented by the pushout $A \otimes_C B$.*

Proof. Let $\epsilon\colon \mathrm{Hom}(A, -) \to E$, $\delta\colon \mathrm{Hom}(B, -) \to F$, $\gamma\colon \mathrm{Hom}(C, -) \to G$ be isomorphisms. From Lemma C.5 we find that α and β correspond to the arrows $x := (\gamma^{-1}\alpha\epsilon)_A(1_A)$ and $y := (\gamma^{-1}\beta\delta)_B(1_B)$, respectively.

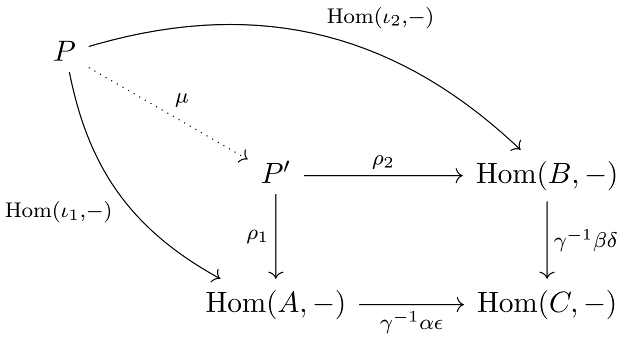
We know that $E \times_G F \cong \mathrm{Hom}(A, -) \times_{\mathrm{Hom}(C,-)} \mathrm{Hom}(B, -) =: P'$. We want to show that P' is naturally isomorphic to $P := \mathrm{Hom}(A \otimes_C B, -)$. We construct a natural transformation $\tau\colon P' \to P$. The R-component is defined as follows

$$\begin{aligned} \tau_R\colon \mathrm{Hom}(A, R) \times_{\mathrm{Hom}(C,R)} \mathrm{Hom}(B, R) &\to \mathrm{Hom}(A \otimes_C B, R) \\ (a, b) &\mapsto \psi_{a,b}. \end{aligned}$$

For any pair of morphisms $a\colon A \to R$ and $b\colon B \to R$, the morphism $\psi_{a,b}$ is the unique C-linear morphism which makes the following diagram commute.



Moreover, if $(\gamma^{-1}\alpha\epsilon)\operatorname{Hom}(\iota_1, -) = (\gamma^{-1}\beta\delta)\operatorname{Hom}(\iota_2, -)$ the universal property of pullbacks guarantees the existence of a unique map μ such that the following diagram commutes.



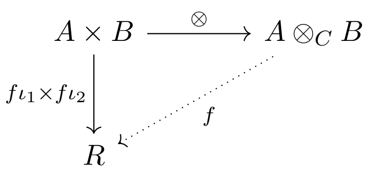
For any $R \in \mathbf{Alg_k}$ and morphism $f\colon A \otimes_C B \to R$ we have that

$$\begin{aligned}(\gamma^{-1}\alpha\epsilon)_R \operatorname{Hom}(\iota_1, R)(f) &= (\gamma^{-1}\alpha\epsilon)_R(f\iota_1) = (\gamma^{-1}\alpha\epsilon)_R\rho_{1,R}(f\iota_1, f\iota_2)\\ &= (\gamma^{-1}\beta\delta)_R\rho_{2,R}(f\iota_1, f\iota_2) = (\gamma^{-1}\beta\delta)_R(f\iota_2)\\ &= (\gamma^{-1}\beta\delta)_R \operatorname{Hom}(\iota_2, R)(f).\end{aligned}$$

We claim that μ and τ are inverse natural transformations. For any morphism $f\colon A \otimes_C B \to R$ it follows that

$$\begin{aligned}\tau_R\mu_R(f) &= \tau_R(\rho_{1,R}\mu_R(f), \rho_{2,R}\mu(f)) = \tau_R(\operatorname{Hom}(\iota_1, R)(f), \operatorname{Hom}(\iota_2, R)(f))\\ &= \tau_R(f\iota_1, f\iota_2) = \psi_{f\iota_1, f\iota_2}.\end{aligned}$$

However, the following diagram commutes.



By the uniqueness of the arrow $\psi_{f\iota_1, f\iota_2}$ we may conclude $f = \psi_{f\iota_1, f\iota_2}$.

For morphisms $f\colon A \to R$ and $g\colon B \to R$ we have

$$\begin{aligned}\mu_R\tau_R(f, g) &= \mu_R(\psi_{f,g}) = (\rho_{1,R}\mu_R(\psi_{f,g}), \rho_{2,R}\mu_R(\psi_{f,g}))\\ &= (\operatorname{Hom}(\iota_1, R)(\psi_{f,g}), \operatorname{Hom}(\iota_2, R)(\psi_{f,g})) = (\psi_{f,g}\iota_1, \psi_{f,g}\iota_2).\end{aligned}$$

However, for $a \in A$

$$\psi_{f,g}\iota_1(a) = \psi_{f,g}(a \otimes 1) = \psi_{f,g} \otimes (a, 1) = (f \times g)(a, 1) = f(a)g(1) = f(a).$$

Therefore, $\psi_{f,g}\iota_1 = f$ and similarly $\psi_{f,g}\iota_2 = g$. □

C.3 Hopf Algebras

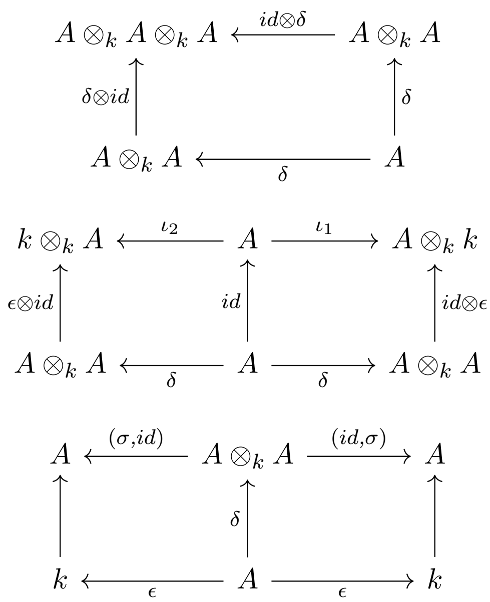
Hopf algebras are introduced in this section. These are associative unital algebras with comultiplication, counit, and coinverse morphisms arising from affine group schemes. We will see some basic constructions and examples.

Definition C.19: *Let G be a representable group functor, then G is called an* **affine group**. *Moreover, if G is representable by a finitely presented algebra, then G is called an* **affine algebraic group** *or* **affine group scheme**.

Remark C.20: *Suppose $G\colon \mathbf{Alg_k} \to \mathbf{Set}$ is represented by A. Then Remark C.11 yields that $G \times G$ is represented by $A \otimes_k A$. We may apply the Yoneda Lemma to find that*

$$Nat(\mathrm{Hom}(A \otimes_k A, -), G) \cong G(A \otimes_k A).$$

Therefore, from Remark C.15 we find that G being a group functor is equivalent to having k-algebra maps which correspond to the multiplication, unit, and inverse natural transformations which make G a group functor. That is, having a comultiplication morphism δ, a counit morphism ϵ, and a coinverse morphism σ, such that the following diagrams commute.



Definition C.21: *A k-algebra with comultiplication, counit, and coinverse morphisms is a* **Hopf Algebra over** k.

Corollary C.22: *Affine group schemes over k correspond to Hopf Algebras over k.*

Remark C.23: *For a Hopf algebra C we may use* **Sweedler** *notation which, for any $c \in C$, expresses the comultiplicative image of c as*

$$\delta(c) = \sum_{(c)} c_{(1)} \otimes c_{(2)}.$$

Anywhere we see a parenthesised subscript we will know there is an implicit summation. This allows us to write the properties as follows:

1. *(Coassociativity):*

$$\begin{aligned}
(\delta \otimes id)\delta(c) &= (\delta \otimes id) \left(\sum_{(c)} c_{(1)} \otimes c_{(2)} \right) \\
&= \sum_{(c)} \left(\sum_{(c_{(1)})} c_{(1)_{(1)}} \otimes c_{(1)_{(2)}} \right) \otimes c_{(2)} \\
&= \sum_{(c)} c_{(1)} \otimes \left(\sum_{(c_{(2)})} c_{(2)_{(1)}} \otimes c_{(2)_{(2)}} \right) \\
&= (id \otimes \delta)\delta(c) = \sum_{(c)} c_{(1)} \otimes c_{(2)} \otimes c_{(3)}.
\end{aligned}$$

2. *(Counit): We note that $A \otimes_k k \cong A$ by the map $a \otimes k \mapsto ka$.*

$$\begin{aligned}
(id \otimes \epsilon)\delta(c) &= (id \otimes \epsilon) \left(\sum_{(c)} c_{(1)} \otimes c_{(2)} \right) \\
&= \sum_{(c)} c_{(1)} \otimes \epsilon(c_{(2)}) \quad \left(\mapsto \sum_{(c)} \epsilon(c_{(2)})c_{(1)} \right) \\
&= \sum_{(c)} \epsilon(c_{(1)}) \otimes c_{(2)} \quad \left(\mapsto \sum_{(c)} \epsilon(c_1)c_{(2)} \right) \\
&= \iota_1 id(c) = c \otimes 1 \ (\mapsto c) \\
&= \iota_2 id(c) = 1 \otimes c \ (\mapsto c).
\end{aligned}$$

We may express this as $\sum_{(c)} \epsilon(c_{(1)})c_{(2)} = \sum_{(c)} \epsilon(c_{(2)})c_{(1)} = c$.

3. *(Coinverse):*

$$(\sigma, id)\delta(c) = (\sigma, id)\left(\sum_{(c)} c_{(1)} \otimes c_{(2)}\right) = \sum_{(c)} \sigma(c_{(1)})c_{(2)} = !_A\epsilon(c)$$
$$= \epsilon(c) = \sum_{(c)} c_{(1)}\sigma(c_{(2)}) = (id, \sigma)\delta(c).$$

Example C.24: *Let G be the affine group scheme represented by $A := k[X]$ and $g, h \in \mathrm{Hom}_{\mathbf{Alg_k}}(A, R)$. From Remark C.15 we have natural transformations m, e, and i making G a group object. From the Yoneda Lemma we find that*

$$Nat(\mathrm{Hom}(A \otimes A, -), \mathrm{Hom}(A, -)) \cong \mathrm{Hom}(A, A \otimes A).$$

Thus, m corresponds to $\delta := m_{A\otimes A}(1_{A\otimes A})\colon A \to A \otimes A$.

We need to construct a map $\delta\colon A \to A \otimes A$ such the the order of 'splitting' does not matter, and which agrees with the Hopf algebra axioms. For instance, if we set $\delta(X) := X \otimes 1$, then we would need $\epsilon(X) = X$ which does not work as the codomain of ϵ is k.

We set $\delta(X) := X \otimes 1 + 1 \otimes X$. We note that as $A = k[X]$ we need only specify the image of X to construct a k-linear map. Furthermore, we may identify $A \otimes A$ with $k[X, Y]$, in which case we are mapping X to $X + Y$. This satisfies the coassociativity axiom.

Now we need find $\epsilon\colon A \to k$. We know that $\iota_2 \circ id(X) = 1 \otimes X$ and identifying $k \otimes A$ with A, we know that $(\epsilon \otimes id)\delta(X)$ must be identified with X. Thus, we have $\epsilon(X) + X = X$ and so $\epsilon(X) := 0$.

Finally, the coinverse axioms require that $(\sigma, id)\delta(X) = \epsilon(X) = 0$. Hence, we must have that $\sigma(X) + \sigma(1)X = 0$, and so $\sigma(X) := -X$.

Appendix D

Weyl's Theorem

Throughout this chapter, unless otherwise specified, we will assume all vector spaces are finite-dimensional and defined over an algebraically closed field, $\mathbb{F}$, of characteristic 0.

D.1 Jordan Canonical Form

Theorem D.1 (Primary Decomposition Theorem)**:** *Let $\tau \colon V \to V$ be a linear transformation, for some vector space V. If $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ are the distinct eigenvalues of τ and $k_1, \ldots, k_n$ their indices, respectively, then*

$$V = \bigoplus_{i=1}^{n} \mathrm{Ker}(\tau - \lambda_i 1_V)^{k_i}.$$

Theorem D.2 (Chinese Remainder Theorem)**:** *Let $p_1, \ldots, p_n$ be nonzero polynomials in $\mathbb{F}[X]$, for $n \geq 2$. If the polynomials are pairwise relatively prime - that is, pairs of distinct polynomials share no polynomial factors - then for any polynomials $q_1, \ldots, q_n \in \mathbb{F}[X]$ there is a polynomial $f \in \mathbb{F}[X]$ such that $f(X) \equiv q_i(X)(mod\ p_i(X))$, for all $i \in \{1, \ldots, n\}$. Moreover, any two such polynomials f, are congruent modulo $p_1(X)) \cdots (p_n(X)$.*

Definition D.3: *A matrix M is said to be in* **Jordan block** *form if it is of the form*

$$M = \begin{pmatrix} m & 1 & & & \\ & m & 1 & & 0 \\ & & \ddots & \ddots & \\ & & & m & 1 \\ & 0 & & & m \end{pmatrix}.$$

For a matrix to be in Jordan block form means that its diagonal entries are all the same and its entries on the superdiagonal (those immediately above the main diagonal) are all 1.

Definition D.4: *An $n \times n$ matrix M is in* **Jordan canonical** *form if it is of the form*

$$M = \begin{pmatrix} M_1 & & & & \\ & M_2 & & & 0 \\ & & \ddots & & \\ & & & M_{n-1} & \\ & 0 & & & M_n \end{pmatrix},$$

where each M_i is a Jordan block matrix for each $i \in \{1, \ldots, n\}$.

Definition D.5: *Let $\tau\colon V \to V$ be a linear transformation with corresponding matrix T, where V is an n-dimensional vector space. The* **characteristic polynomial** *of T, or equivalently of τ, is the polynomial of degree n defined by $c(x) = det(xI_n - T)$.*

Theorem D.6: *Let $\tau\colon V \to V$ be a linear transformation, T its corresponding matrix, and $c(x) = (x-\lambda_1)^{k_1} \cdots (x-\lambda_n)^{k_n}$ its characteristic polynomial. Then there is a basis for V relative to which T is in Jordan canonical form with Jordan block matrices $T_1, \ldots, T_p$. Each block matrix T_i has as its diagonal entries one of the λ_j. Moreover, for each i there are a total of k_i entries of λ_i across the whole of T, that is, the number of times λ_i appears as an entry is equal to its multiplicity as a root of the characteristic polynomial.*

Theorem D.7: *Let $L = \mathfrak{gl}(V)$, for some vector space V, and $x \in L$. There exist unique $x_d, x_n \in \mathfrak{gl}(V)$ such that $x = x_d + x_n$, x_d is diagonalisable, x_n is nilpotent, and x_n and x_d commute with one another. Furthermore, there are polynomials $p, q \in \mathbb{F}[X]$ with no constant term, such that $p(x) = x_d$ and $q(x) = x_n$.*

Proof. Let $c(X) = (X-\lambda_1)^{k_1} \cdots (X-\lambda_n)^{k_n}$ be the characteristic polynomial of x, with the λ_i distinct. From Theorem D.1 (the Primary Decomposition Theorem) we know that

$$V = \bigoplus_{i=1}^{n} \mathrm{Ker}(x - \lambda_i 1_V)^{k_i} := \bigoplus_{i=1}^{n} V_{k_i}(\lambda_i), \tag{D.1}$$

where each $V_{k_i}(\lambda_i)$ is x-invariant. We apply Theorem D.2 (the Chinese Remainder Theorem) to find $p \in \mathbb{F}[X]$ such that $p(X) \equiv \lambda_i \ (mod\ (X-\lambda_i)^{k_i})$, for all $i \in \{1, \ldots, n\}$ and $p(X) \equiv 0 \ (mod\ X)$. We remark that if 0 is an eigenvalue of x the condition $p(X) \equiv 0 \ (mod\ X)$ is redundant, otherwise $X - \lambda_i$ shares no polynomial factors with X, for all $i \in \{1, \ldots, n\}$. Now, set $q(X) = X - p(X)$. We then have that $q(X) \equiv 0 \ (mod\ X)$. Since both p and q are congruent to 0 modulo X, their constant terms must be 0.

Now that we have polynomials p and q we set $x_d := p(x)$, $x_n := q(x)$, and we must now prove that the claimed properties hold. By our definition of q we have that $x = p(x)+q(x) = x_d+x_n$. Being elements of $\mathbb{F}[X]$, we know that $p(x)$ and $q(x)$ (and thus x_d and x_n) commute with one another. Since the $V_{k_i}(\lambda_i)$ are x-invariant they must also be invariant under the actions of x_d and x_n. Thus, x_d restricted to $V_{k_i}(\lambda_i)$ is an element of $\mathfrak{gl}(V_{k_i}(\lambda_i))$. We also have that $p(x) \equiv \lambda_i \ (mod\ (x-\lambda_i)^{k_i})$. Thus, for any $i \in \{1, \ldots, n\}$ we have $p(x)(v) = \lambda_i(v) + r(x)(x-\lambda_i)^{k_i}(v) = \lambda_i(v)$, for some $r(X) \in \mathbb{F}[X]$ and thus we can see

$$\begin{aligned}(x_d - \lambda_i 1_{V_i})(v) &= p(x)(v) - \lambda_i v = (\lambda_i v + r(x)(x-\lambda_i)^{k_i}(v)) - \lambda_i(v) \\ &= \lambda_i(v) - \lambda_i(v) = 0,\end{aligned}$$

for all $v \in V_{k_i}(\lambda_i)$. So $x_d(v) = \lambda_i$ for any $v \in V_{k_i}(\lambda_i)$. This means that the action of x_d on $V_{k_i}(\lambda_i)$ is diagonalisable with the only eigenvalue being λ_i. From (D.1) we conclude that x_d acting on V is diagonalisable.

We know that $x_n = x - x_d$. So, for $v \in V_{k_i}(\lambda_i)$, we have

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

since $V_{k_i}(\lambda_i) = \mathrm{Ker}(x - \lambda_i 1_V)^{k_i}$. Thus, x_n acting on $V_{k_i}(\lambda_i)$ is nilpotent and by (D.1) we may conclude that x_n acting on V is nilpotent.

It remains to prove that this decomposition is unique. Let $x = x_d' + x_n'$ be another decomposition as described. Then, there must be polynomials $p', q' \in \mathbb{F}[X]$ such that $p'(x) = x_d'$ and $q'(x) = x_n'$. Since x_d, x_n, x_d', and x_n' may all be expressed as polynomials as above, they all commute with one another. Thus, $x_d' + x_n' = x = x_d + x_n$ implies that $x_d' - x_d = x_n - x_n'$. We then have that x_d and x_d' are both diagonalisable and they commute with one another, which means that there is a basis of simultaneous eigenvectors for the linear transformations. Hence, $x_d' - x_d$ is diagonalisable. We know that x_n and x_n' are both nilpotent and commute with one another. Applying Lemma 1.64 we have that $x_n - x_n'$ is nilpotent. A linear transformation which is both diagonalisable and nilpotent must be the zero map. This means that $x_d = x_d'$ and $x_n = x_n'$, which proves uniqueness. □

Definition D.8: *Let x be a linear transformation of a vector space V. Fix a basis relative to which x is in Jordan canonical form (see Appendix D). We may then write $x = x_d + x_n$, where x_d is diagonalisable, x_n is nilpotent, and x_d and x_n commute with one another. This is the* **Jordan decomposition** *of x.*

Corollary D.9: *Let $x \in \mathfrak{gl}(V)$, for some vector space V. If x has Jordan decomposition $x = d + n$, then $ad_x\colon \mathfrak{gl}(V) \to \mathfrak{gl}(V)$ has Jordan decomposition $ad_x = ad_d + ad_n$.*

Proof. For any $y \in \mathfrak{gl}(V)$ we have

$$ad_x(y) = ad_{d+n}(y) = [d, y] + [n, y] = (ad_d + ad_n)(y),$$

and

$$\begin{aligned} ad_d \circ ad_n(y) - ad_n \circ ad_d(y) &= [d,[n,y]] - [n,[d,y]] \\ &= [d,[n,y]] + [d,[y,n]] + [y,[n,d]] \\ &= [d,[n,y]] - [d,[n,y]] + [y, nd - dn] \\ &= [y, nd - dn] = 0. \end{aligned}$$

So, $ad_x = ad_d + ad_n$, and ad_d and ad_n commute with one another. It remains to show that ad_d and ad_n are diagonalisable and nilpotent, respectively. Applying Proposition 1.56 we have that ad_n is nilpotent. Choose a basis $\{b_1, \dots, b_k\}$ for V, relative to which d is a diagonal matrix with diagonal entries $\lambda_1, \dots, \lambda_k \in \mathbb{F}$. Notice that $B = \{e_{ij} \mid 1 \le i, j \le k\}$ is a basis for $\mathfrak{gl}(V)$. From (1.1) we have

$$ad_d(e_{ij}) = [d, e_{ij}] = \sum_{l=1}^{k} \lambda_l [e_{ll}, e_{ij}] = \sum_{l=1}^{k} \lambda_l (\delta_{li} e_{lj} - \delta_{lj} e_{il}) = (\lambda_i - \lambda_j) e_{ij},$$

which shows that ad_d acts diagonally on $\mathfrak{gl}(V)$. □

Theorem D.10: *Let $L \subseteq \mathfrak{gl}(V)$, where V is some vector space, be a Lie algebra. For any $x \in L$, with Jordan decomposition $x = d + n$, we have that $d, n \in L$.*

Proof. Let $x \in L$ with Jordan decomposition $x = d + n$. Notice that $ad_x(L) \subseteq L$. From Theorem D.7, we get polynomials (each with constant term 0) $p, q \in \mathbb{F}[X]$, such that $p(x) = d$ and $q(x) = n$. This implies that $ad_d(L), ad_n(L) \subseteq L$. We now consider the normaliser of L in $\mathfrak{gl}(V)$, that is

$$N_{\mathfrak{gl}(V)}(L) = \{y \in \mathfrak{gl}(V) \mid [y, L] \subseteq L\}.$$

For $z \in L$, we have that $[d, z] = ad_d(z) \in L$, and $[n, z] = ad_n(z) \in L$. Hence, $d, n \in N_{\mathfrak{gl}(V)}(L)$. From Proposition 1.55 we know that $N_{\mathfrak{gl}(V)}(L)$ is the largest subalgebra of $\mathfrak{gl}(V)$, in which L is an ideal.

Since the inclusion map $\iota \colon L \to \mathfrak{gl}(V)$, is a representation of L, we have that V is an L-module. Let $\mathscr{F}$ denote the set consisting of all L-submodules of V. For each $W \in \mathscr{F}$, we set

$$L_W = \{y \in \mathfrak{gl}(V) \mid y(W) \subseteq W, \text{ and } tr(y|_W) = 0\},$$

where $y|_W$ denotes y restricted to W. Since W is an L-module, we know that $x(W) \subseteq W$. Moreover, since L is semisimple we have that $L = L'$, so $x = \sum_i \beta_i [a_i, b_i]$, for some $\beta_i \in \mathbb{F}$, $a_i, b_i \in L$. For all $a, b \in L$ we have

$$\begin{aligned} tr([a,b]|_W) &= tr((ab - ba)|_W) = tr((ab)|_W - (ba)|_W) \\ &= tr((ab)|_W) - tr((ba)|_W) = 0. \end{aligned}$$

This implies that $tr(x|_W) = 0$, and so $x \in L_W$. This implies that $L \subseteq L_W$ for all $W \in \mathscr{F}$. We have that $d(W), n(W) \subseteq W$, since $p(x) = d$ and $q(x) = n$. The action of n on V is nilpotent, so it is also nilpotent on W, and thus $tr(n|_W) = 0$. Hence, we have that

$$0 = tr(x|_W) = tr((d + n)|_W) = tr(d|_W) + tr(n|_W) = tr(d|_W).$$

Therefore $d, n \in L_W$ for all $W \in \mathscr{F}$. We set $L_\cap = \bigcap_{W \in \mathscr{F}} L_W$. Then clearly, $L \subseteq L_\cap$. Moreover, since each L_W is a subalgebra of $N_{\mathfrak{gl}(V)}(L)$, we get that $L_\cap$ is a subalgebra of $N_{\mathfrak{gl}(V)}(L)$.

Since $L \subseteq L_\cap$, we know that

$$ad_x(L_\cap) = [x, L_\cap] \subseteq (L_\cap)' \subseteq L_\cap.$$

Hence, we may define a representation $\phi\colon L \to \mathfrak{gl}(L_\cap)$ of L, by

$$\phi(x) = ad_x\colon L_\cap \to L_\cap.$$

This shows that $L_\cap$ is an L-module. This allows us to apply Theorem D.21 (Weyl's Theorem) and find that $L_\cap$ is completely reducible as an L-module. Since L is an L-submodule of $L_\cap$, an application of Lemma D.18 tells us that we may write $L_\cap = L \oplus M$, where M is some submodule of $L_\cap$. Notice that $d, n \in L_\cap$. Our goal is to show that $M = 0$, we will then be done.

We have that

$$[L, L \oplus M] = [L, L_\cap] \subseteq [L, N_{\mathfrak{gl}(V)}(L)] \subseteq L,$$

since $L_\cap$ is a subalgebra of $N_{\mathfrak{gl}(V)}(L)$, and L is an ideal of $N_{\mathfrak{gl}(V)}(L)$. This implies that $[L, M] \subseteq L \cap M = 0$.

By Theorem D.21 (Weyl's Theorem), we may write $V = \bigoplus_{k=1}^{n} V_k$, where each V_k is an irreducible submodule of V. Since $L_\cap \subseteq \mathfrak{gl}(V)$, the inclusion map $\iota\colon L_\cap \to \mathfrak{gl}(V)$, is a representation of $L_\cap$. Hence, for every k we have that V_k is an irreducible $L_\cap$-module. Let $y \in M \subseteq L_\cap$, we claim that $y|_{V_k}$ is a L-module homomorphism. Since $y \in L_\cap$, we know that $y(V_k) \subseteq V_k$. Moreover, we have that $[L, y] = 0$, so $0 = [y, x] \cdot v = y \cdot (x \cdot v) - x \cdot (y \cdot v)$, for all $x \in L,\ v \in V_k$. This means that $y(x \cdot v) = x \cdot y(v)$ for all $x \in L,\ v \in V_k$. Therefore, we have that $y|_{V_k}$ is an L-module homomorphism, and V_k an irreducible L-module. An application of Lemma 1.39 (Schur's Lemma) yields that $y|_{V_k}$ is a scalar multiple of the identity; that is, $y|_{V_k} = \alpha 1_{V_k}$, for some $\alpha \in \mathbb{F}$. Furthermore, since $tr(y|_{V_k}) = 0$, (because $y \in L_\cap$) we have that $\alpha = 0$. So the action of y on each V_k is trivial. This means that $y = 0$, and thus that $M = 0$. Thus $L_\cap = L$. □

Corollary D.11: *Let $L \subseteq \mathfrak{gl}(V)$, where V is some vector space, be a semisimple Lie algebra. Then for $x \in L$, the usual Jordan decomposition of x coincides with the abstract Jordan decomposition of x.*

Proof. Let $x \in L$. Suppose we denote the usual Jordan decomposition of x by $x = d + n$. By the previous theorem we have that $d, n \in L$. By Corollary D.9 we know that ad_x has Jordan decomposition $ad_d + ad_n$. Then, since the abstract Jordan decomposition is unique we are done. □

D.2 Casimir Operator

In this section we will introduce the Casimir operator of a Lie algebra representation. To do so we will first need some further results on representations and modules.

Definition D.12: *Let L be a Lie algebra and $\phi\colon L \to \mathfrak{gl}(V)$ a representation of L. Define the* **trace** *form on L by*

$$\begin{aligned} \beta_V\colon L \times L &\to \mathbb{F} \\ (x, y) &\mapsto tr(\phi(x)\phi(y)), \end{aligned}$$

for all $x, y \in L$.

The trace form is an associative, symmetric bilinear form on a Lie algebra L. It is bilinear because tr is linear and representations are linear. The trace form is linear for the same reason that the Killing form is symmetric, that is, $tr(xy) = tr(yx)$ for linear maps. We will denote the perpendicular space to L, with respect to the trace form, by

$$L^{\beta_V} = \{x \in L \mid \beta_V(x, y) = 0,\ \forall\ y \in L\}.$$

Then L^{β_V} is a vector subspace of L. Furthermore, for $x \in L^{\beta_V},\ y, z \in L$ we have $[[x, y], z] = [0, z] = 0$, which shows that L^{β_V} is an ideal of L.

Lemma D.13: *Suppose that U is a subspace of a vector space V and that $(-,-)\colon V \times V \to \mathbb{F}$ is a non-degenerate bilinear form on V. Then*

$$\dim(U) + \dim(U^{\perp}) = \dim(V).$$

Proof. Define $\phi\colon V \to U^*$ (where U^* denotes the dual space of U) by $\phi(v)(u) = (u, v)$ for all $v \in V,\ u \in U$. To see that ϕ is linear, take $v, v' \in V,\ u \in U$ and $\alpha, \beta \in F$. We then have

$$\phi(\alpha v + \beta v')(u) = (u, \alpha v + \beta v') = \alpha(u, v) + \beta(u, v') = \alpha\phi(v)(u) + \beta\phi(v')(u).$$

Notice that $v \in \mathrm{Ker}(\phi)$ if and only if $\phi(v)(u) = (u, v) = 0$, for all $u \in U$. So, $\mathrm{Ker}(\phi) = U^{\perp}$. We also know that $\dim(U) = \dim(U^*)$. By the Rank-Nullity theorem we have that $\dim(U) + \dim(U^*) = \dim(V)$. □

Proposition D.14: *Let L be a semisimple Lie algebra and $\phi\colon L \to \mathfrak{gl}(V)$ a faithful representation of L. Then the trace form on L is non-degenerate.*

Proof. Let $x, y \in L^{\beta_V}$, then since $x \in L$ we have that

$$\beta_V(x, y) = tr(\phi(x)\phi(y)) = 0.$$

Since $\phi(L^{\beta_V})$ is a subalgebra of $\mathfrak{gl}(V)$, an application of Lemma 1.65 yields that $\phi(L^{\beta_V})$ is solvable. Since ϕ is faithful, an application of Theorem 1.17 (Isomorphism Theorems) yields that $L^{\beta_V} \cong \phi(L^{\beta_V})$ is solvable. By hypothesis L is semisimple, so $L^{\beta_V} = 0$. Hence, the trace form on L is non-degenerate. □

Corollary D.15: *Let L be a semisimple Lie algebra and $\phi\colon L \to \mathfrak{gl}(V)$ a faithful representation of L. If $B_L = \{b_1, \ldots, b_n\}$ is a basis for L, then there are elements $c_1, \ldots, c_n \in L$ such that the c_i form a basis for L and $\beta_V(b_i, c_j) = \delta_{ij}$.*

Proof. We know that for any $\theta \in L^*$ there is a unique $y \in L$ such that $\beta_V(x, y) = \theta(x)$ for every $x \in L$ (see Lemma D.13). Now, B_L induces a basis $B_{L^*} = \{\theta_1, \ldots, \theta_n\}$ for L^*, where each θ_i is defined by

$$\theta_i(b_j) = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases}$$

We may then find $c_1, \ldots, c_n \in L$ such that $\beta_V(x, c_j) = \theta_j(x)$ for every $x \in L$. Thus, we have that $\beta_V(b_i, c_j) = \theta_j(b_i) = \delta_{ij}$.

Notice that mapping the θ_i (which form a basis for L^*) to the c_i, respectively, is a linear map from L^* to L. We claim such a map is injective, which would imply that the c_i are linearly independent. Suppose $c_i = c_j$, then $\beta_V(x, c_i) = \beta_V(x, c_j)$ for all $x \in L$. This means that $\theta_i(x) = \theta_j(x)$ for all $x \in L$. This implies that $i = j$. Hence, the c_i are linearly independent. Since $\dim(L) = \dim(L^*)$, the set $\{c_1, \ldots, c_n\}$ is a maximal linearly independent set and is thus a basis for L. □

Let L be a semisimple Lie algebra and $\phi\colon L \to \mathfrak{gl}(V)$ a faithful representation of L. Choose a basis $\{b_1, \ldots, b_n\}$ for L and apply Corollary D.15 to find a basis $\{c_1, \ldots, c_n\}$ for L such that $\beta_V(b_i, c_j) = \delta_{ij}$. Then, for $x \in L$ we may write $[b_i, x] = \sum_{k=1}^{n} \alpha_{ik} b_k$, for some $\alpha_{ik} \in \mathbb{F}$. For $j \in \{1, \ldots, n\}$ we have

$$\alpha_{ij} = \sum_{k=1}^{n} \alpha_{ik}\delta_{kj} = \sum_{k=1}^{n} \alpha_{ik}\beta_V(b_k, c_j) = \beta_V([b_i, x], c_j) = \beta_V(b_i, [x, c_j]).$$

Since the c_i form a basis for L, we may write $[x, c_k] = \sum_{i=1}^{n} \gamma_{ki} c_i$, for some $\gamma_{ki} \in \mathbb{F}$. Then, we have

$$\alpha_{ij} = \beta_V(b_i, [x, c_j]) = \sum_{k=1}^{n} \gamma_{jk}\beta_V(b_i, c_k) = \gamma_{ji}.$$

Then we may conclude that for any $j \in \{1, \dots, n\}$ we have

$$[x, c_j] = \sum_{i=1}^{n} \gamma_{ji} c_i = \sum_{i=1}^{n} \alpha_{ij} c_i \tag{D.2}$$

Definition D.16: *Let L be a semisimple Lie algebra, $\phi\colon L \to \mathfrak{gl}(V)$ a faithful representation of L, and $B = \{b_1, \dots, b_n\}$ a basis for L. Let $B' = \{c_1, \dots, c_n\}$ be a basis for L such that $\beta_V(b_i, c_j) = \delta_{ij}$ (such a basis exists by Corollary D.15). We define a linear map $c_\phi\colon V \to V$, called the* **Casimir operator** *of ϕ, by*

$$c_\phi(v) = \sum_{i=1}^{n} \phi(b_i)\phi(c_i(v)),$$

in terms of the representation ϕ of L. In terms of the equivalent L-module, that is,

$$c_\phi(v) = \sum_{i=1}^{n} b_i \cdot (c_i \cdot v).$$

We note that c_ϕ is linear because the action of L on V is linear.

Lemma D.17: *Let L be a semisimple Lie algebra and $\phi\colon L \to \mathfrak{gl}(V)$ a faithful representation of L. Let $B = \{b_1, \dots, b_n\}$ and $B' = \{c_1, \dots, c_n\}$ be bases for L as described in Corollary D.15. Then the Casimir operator of ϕ is an L-module homomorphism, and $tr(c_\phi) = \dim(L)$.*

Proof. To show that c_ϕ is an L-module homomorphism it suffices to prove that $c_\phi(x \cdot v) = x \cdot \phi(v)$ for all $x \in L$, $v \in V$. For $x \in L$, we may write $[b_i, x] = \sum_{k=1}^{n} \alpha_{ik} b_k$, and $[x, c_i] = \sum_{k=1}^{n} \gamma_{ik} c_k$, for some $\alpha_{ik}, \gamma_{ik} \in \mathbb{C}$. Then for $v \in V$ we have

$$\begin{aligned}
c_\phi(x \cdot v) &= \sum_{j=1}^{n} b_j \cdot (c_j \cdot (x \cdot v)) \\
&= \sum_{j=1}^{n} b_j \cdot (c_j \cdot (x \cdot v) - x \cdot (c_j \cdot v) + x \cdot (c_j \cdot v)) \\
&= \sum_{j=1}^{n} b_j \cdot ([c_j, x] \cdot v + x \cdot (c_j \cdot v)) \\
&\overset{(D.2)}{=} \sum_{j=1}^{n} \sum_{i=1}^{n} -\alpha_{ij} b_j \cdot (c_i \cdot v) + \sum_{j=1}^{n} b_j \cdot (x \cdot (c_j \cdot v)) \\
&= \sum_{j,i=1}^{n} -\alpha_{ij} b_j \cdot (c_i \cdot v) + \sum_{j=1}^{n} ([b_j, x] + x \cdot b_j) \cdot (c_j \cdot v)
\end{aligned}$$

$$= \sum_{j,i=1}^{n} -\alpha_{ij} b_j \cdot (c_i \cdot v) + \sum_{j,i=1}^{n} \alpha_{ji} b_i \cdot (c_j \cdot v) + x \cdot c_\phi(v)$$
$$= \sum_{j,i=1}^{n} (-\alpha_{ij} b_j \cdot (c_i \cdot v) + \alpha_{ij} b_j \cdot (c_i \cdot v)) + x \cdot c_\phi(v)$$
$$= x \cdot c_\phi(v).$$

Furthermore, from Corollary D.15 we have

$$tr(c_\phi) = tr\left(\sum_{i=1}^{n} \phi(b_i)\phi(c_i)\right) = \sum_{i=1}^{n} tr(\phi(b_i)\phi(c_i)) = \sum_{i=1}^{n} \beta_V(b_i, y_i) = \sum_{i=1}^{n} \delta_{ii}$$
$$= n,$$

as required. □

Lemma D.18: *Let L be a Lie algebra and V an L-module. Then V is completely reducible if and only if, every submodule W of V has a complement. That is to say, for each submodule W of V there exists a submodule U of V such that $V = W \oplus U$.*

Proof. Suppose $V = \bigoplus_{k=1}^{n} S_k$, for irreducible submodules S_k of V. Let W be a non-zero submodule of V. Then we have

$$W = W \cap V = W \cap \left(\bigoplus_{k=1}^{n} S_k\right) = \bigoplus_{k=1}^{n} (W \cap S_k).$$

From Theorem 1.38 we have that each $W \cap S_k$ is a submodule of V. But, the S_k are irreducible, so for each $k \in \{1, \ldots, n\}$ we have that either $W \cap S_k = 0$, or $W \cap S_k = S_k$. For $i \neq j$ we know that $S_i \cap S_j = 0$, so there is exactly one $i \in \{1, \ldots, n\}$ such that $W \cap S_i = S_i$, and so $W = S_i$. Without loss of generality let us say $i = 1$. Then for $U = \bigoplus_{k=2}^{n} S_k$, we have $V = W \oplus U$.

On the other hand, suppose that every submodule W of V has a complement. We proceed by induction on $\dim(V)$. If $\dim(V) = 1$, then V is irreducible and we are done.

Assume now that $\dim(V) > 1$ and that the result holds for all L-modules of dimension strictly less than $\dim(V)$. If V is irreducible we are done. Otherwise, let W be a non-zero proper submodule of V. Then there is some submodule U of V such that $V = W \oplus U$. Let M be a submodule of W, then M is also a submodule of V. Hence, $V = M \oplus N$, for some submodule N of V. Then,

$$W = W \cap V = W \cap (M \oplus N) = (W \cap M) \oplus (W \cap N).$$

Similarly, this holds for U. We may thus apply the inductive hypothesis to U and W. This yields that U and W are completely reducible, say $W = \bigoplus_{k=1}^{i} W_k$ and $U = \bigoplus_{k=1}^{j} U_k$, where each of the W_k and U_k are irreducible submodules of W and U, respectively. This implies that each of the W_k and U_k are irreducible submodules of V. We may then write $V = (\bigoplus_{k=1}^{i} W_k) \oplus (\bigoplus_{k=1}^{j} U_k)$, and so we are done. $\square$

D.3 Weyl's Theorem

We now have the machinery we need to prove the important Weyl's Theorem, which asserts that all representations of a semisimple Lie algebra are completely reducible. We will first prove a special case of Weyl's Theorem and then use this special case to prove the general version.

Theorem D.19: *Let L be a semisimple Lie algebra, and $\phi \colon L \to \mathfrak{gl}(V)$ a faithful representation of L. For any submodule W of V of codimension 1, there is a submodule U of V such that $V = W \oplus U$.*

Proof. Notice that the quotient module V/W has dimension 1. We claim that L acts trivially on V/W. Since L is semisimple, we have that $L = L'$. For $x \in L = L'$ we may write $x = \sum_i \alpha_i [y_i, z_i]$, for some $\alpha_i \in \mathbb{F}$, $y_i, z_i \in L$. Since V/W is 1-dimensional, for any $v \in V - W$ we have $V/W = Span(v + W)$. Thus, for $y, z \in L$ we have that $y \cdot v + W = \alpha v + W$ and $z \cdot v + W = \beta v + W$, for some $\alpha, \beta \in \mathbb{C}$. Therefore,

$$[y, z] \cdot v + W = y \cdot (z \cdot v) - z \cdot (y \cdot v) + W = \alpha\beta v - \beta\alpha v + W = W.$$

This implies that $L \cdot V \subseteq W$. We proceed by induction on $\dim(V)$. If $\dim(V) = 1$, then V is irreducible and we are done.

Assume now that $\dim(V) > 1$. If W is irreducible, then let $c_\phi \colon V \to V$ be the Casimir operator of ϕ. By Lemma D.17, we know that c_ϕ is a Lie module homomorphism. An application of Lemma 1.37 yields that $\mathrm{Ker}(\phi)$ is a submodule of V. We have seen that $L \cdot V \subseteq W$; from this we find that $c_\phi \cdot V \subseteq W$ from the definition of c_ϕ. This means that c_ϕ is not surjective. The Rank-Nullity Theorem then tells us that $\dim(\mathrm{Ker}(c_\phi)) \geq 1$. We conclude that the restriction of c_ϕ to W, which we will denote by $c_\phi|_W \colon W \to W$, is a Lie module homomorphism. Since W is an irreducible L-module we may apply Lemma 1.39 (Schur's Lemma) to see that $c_\phi|_W = \lambda 1_W$, for some $\lambda \in \mathbb{C}$. From Lemma D.17 we know that $tr(c_\phi) = \dim(L)$, and since $c_\phi \cdot V \subseteq W$ and $c_\phi|_W = \lambda 1_W$ we can see that $tr(c_\phi) = \lambda \dim(W)$. We then conclude that $0 \neq \dim(L) = \lambda \dim(W)$, which implies that $\lambda \neq 0$. This tells us that

$\text{Ker}(c_\phi) \cap W = 0$. Finally, $\dim(W) = \dim(V) - 1$ and $\dim(\text{Ker}(c_\phi)) \geq 1$, allow us to conclude that $V = W \oplus \text{Ker}(c_\phi)$.

Assume now that W is reducible. Let W' be a non-zero proper submodule of W. Applying Theorem 1.38 we have

$$(V/W')/(W/W') \cong V/W.$$

Since $\dim(V/W) < \dim(V)$ (as $\dim(W) > 1$) we apply the inductive hypothesis to V/W' to find a submodule $\widetilde{M}$ of V/W' such that

$$V/W' = W/W' \oplus \widetilde{M}.$$

This implies that $\dim(\widetilde{M}) = 1$. By Proposition 1.32, there is a submodule M of V, which contains W', such that $\widetilde{M} = M/W'$. Therefore, we have $\dim(M) = \dim(W') + 1$. Moreover, $\dim(M) < \dim(V)$ since

$$V/W' = W/W' \oplus M/W'.$$

We apply the inductive hypothesis to M and find a submodule C, of M, such that $M = W' \oplus C$. We note that $\dim(C) = 1$. Since $V/W' = W/W' \oplus M/W'$, we find that under the canonical homomorphism $\pi\colon V \to V/W'$, which maps elements of V to their cosets under W', we have $\pi(W \cap M) = W'$. This implies that $W \cap M \subseteq W'$. Thus, we have

$$W \cap C \subseteq W \cap (W' \oplus C) = (W \cap W') \oplus (W \cap C) = W' \cap C = 0.$$

This implies that $V = W \oplus C$, and so we are done. □

Lemma D.20: *Let L be a Lie algebra with L-modules V and W. If we denote by $Hom(V, W)$ the vector space of all linear maps from V to W, then $Hom(V, W)$ is an L-module with the action*

$$(x \cdot \alpha)(v) = x \cdot \alpha(v) - \alpha(x \cdot v),$$

for all $x \in L,\ \alpha \in Hom(V, W),\ v \in V$.

Proof. For $x, y \in L,\ \theta, \varphi \in Hom(V, W),\ \lambda, \mu \in \mathbb{F},\ v \in V$, we have

$$\begin{aligned}((\lambda x + \mu y) \cdot \theta)(v) &= (\lambda x + \mu y) \cdot \theta(v) - \theta((\lambda x + \mu y) \cdot v)\\ &= \lambda(x \cdot \theta(v)) + \mu(y \cdot \theta(v)) - \theta(\lambda(x \cdot v)) - \theta(\mu(y \cdot v))\\ &= \lambda(x \cdot \theta(v) - \theta(x \cdot v)) + \mu(y \cdot \theta(v) - \theta(y \cdot v))\\ &= \lambda(x \cdot \theta)(v) + \mu(y \cdot \theta)(v),\end{aligned}$$

which proves identity (M1), and

$$\begin{aligned}x \cdot (\lambda\theta + \mu\varphi)(v) &= x \cdot (\lambda\theta + \mu\varphi)(v) - (\lambda\theta + \mu\varphi)(x \cdot v)\\ &= x \cdot (\lambda\theta(v)) + x \cdot (\mu\varphi(v)) - \lambda\theta(x \cdot v) - \mu\varphi(x \cdot v)\\ &= \lambda((x \cdot \theta(v)) - \theta(x \cdot v)) + \mu((x \cdot \varphi(v) - \varphi(x \cdot v))\\ &= \lambda(x \cdot \theta)(v) + \mu(x \cdot \varphi)(v),\end{aligned}$$

which proves identity (M2), and

$$\begin{aligned}([x,y]\cdot\theta)(v) &= [x,y]\cdot\theta(v)-\theta([x,y]\cdot v)\\ &= x\cdot(y\cdot\theta(v))-y\cdot(x\cdot\theta(v))-\theta(x\cdot(y\cdot v)-y\cdot(x\cdot v))\\ &= x\cdot(y\cdot\theta(v))-y\cdot(x\cdot\theta(v))-\theta(x\cdot(y\cdot v))+\theta(y\cdot(x\cdot v))\\ &= [x\cdot(y\cdot\theta(v)-\theta(x\cdot(y\cdot v))]-[y\cdot(x\cdot\theta(v))-\theta(y\cdot(x\cdot v))]\\ &= x\cdot(y\cdot\theta)(v)-y\cdot(x\cdot\theta)(v)=(x\cdot(y\cdot\theta)-y\cdot(x\cdot\theta))(v)\end{aligned}$$

which proves identity (M3). □

Theorem D.21 (Weyl's Theorem)**:** *Let L be a semisimple Lie algebra, and $\phi\colon L\to\mathfrak{gl}(V)$ a representation of L. Then V is completely reducible.*

Proof. We claim that we may assume ϕ is faithful. If it is not, then we consider the representation $\widetilde{\phi}\colon L/\operatorname{Ker}(\phi)\to\mathfrak{gl}(V)$ of L, defined by

$$\widetilde{\phi}(x+\operatorname{Ker}(\phi))=\phi(x),$$

for all $x\in L$. In the context of Lie modules, we claim that V is an $L/\operatorname{Ker}(\phi)$-module, with the action $(x+\operatorname{Ker}(\phi))\cdot v=x\cdot v$, for all $x\in L,\ v\in V$. We need to show that $\widetilde{\phi}$ is a well-defined map. Let $x,y\in L$ such that $x+\operatorname{Ker}(\phi)=y+\operatorname{Ker}(\phi)$, then $x-y\in\operatorname{Ker}(\phi)$. So, for $v\in V$, we have

$$\begin{aligned}\widetilde{\phi}(x+\operatorname{Ker}(\phi))(v) &= \phi(x)(v)=\phi(x+y-x)(v)=\phi(y)(v)\\ &= \widetilde{\phi}(y+\operatorname{Ker}(\phi))(v).\end{aligned}$$

In fact, V is an $L/\operatorname{Ker}(\phi)$-module because V is an L-module. Clearly, $\operatorname{Ker}(\widetilde{\phi})=0$. We may therefore assume, without loss of generality, that ϕ is a faithful representation.

Let W be a submodule of V. We shall denote by $H:=Hom(V,W)$, the vector space of all linear maps from V to W. By Lemma D.20 H is an L-module. Set

$$\begin{aligned}&H_S=\{\alpha\in H:\alpha|_W=\lambda 1_W,\ \text{for some }\lambda\in\mathbb{F}\},\ \text{and}\\ &H_0=\{\alpha\in H:\alpha|_W=0\}\end{aligned}$$

where $\alpha|_W$, denotes α restricted to W. Clearly, both are vector subspaces of H, and $H_0\subseteq H_S$. We claim that H_S and H_0 are L-invariant, and thus submodules of H. In fact, for $\alpha\in H_S,\ x\in L,\ w\in W$, we have

$$(x\cdot\alpha)(w)=x\cdot\alpha(w)-\alpha(x\cdot w)=x\cdot(\lambda w)-\lambda(x\cdot w)=\lambda(x\cdot w)-\lambda(x\cdot w)=0,$$

and for $\alpha\in H_0,\ x\in L,\ w\in W$, we have

$$(x\cdot\alpha)(w)=x\cdot\alpha(w)-\alpha(x\cdot w)=0.$$

We now consider the quotient module H_S/H_0. Let $U = Span(V - W)$. Note that U is a vector subspace of V but it need not be a submodule of V. We claim that $V = W \oplus U$, as vector spaces. Clearly $V = W + U$. It remains to show that $U \cap W = 0$. Let $u \in U \cap W$, then we may write $u = \sum_i^k \alpha_i v_i$, for some $\alpha_i \in \mathbb{F}$, $v_i \in V - W$. Suppose that u is non-zero. If we consider the quotient vector space V/W we have

$$0 = u + W = \sum_{i=1}^{k} \alpha_i v_i + W = \sum_{i=1}^{k} \alpha_i (v_i + W),$$

which implies that each $v_i \in W$, a contradiction. Hence, $u = 0$ and then $V = W \oplus U$, as vector spaces. Let $g \colon V \to V$ be defined as $g(w) = w$, for all $w \in W$, and $g(u) = 0$, for all $u \in U$. Then, $g + H_0$ is a non-zero element of H_S/H_0, and thus $H_S/H_0 \neq H_0$. Let $\alpha \in H_S$, then $\alpha|_W = \lambda 1_W$, for some $\lambda \in \mathbb{F}$. Hence, $\alpha - \lambda 1_V \in H_0$, so $\alpha + H_0 = \lambda(1_V + H_0)$. This implies that $\dim(H_S/H_0) = 1$.

We may thus apply Theorem D.19, and conclude that $H_S = H_0 \oplus C$, for some submodule C of H_S. We have that $\dim(C) = 1$. Since L is semisimple, $L = L'$ and C is a 1-dimensional L-module, we have that $L \cdot C = 0$ (we have proved this claim more carefully in the proof of Theorem D.19). This means there is a non-zero $\alpha \in C$, such that $x \cdot \alpha = 0$ for all $x \in L$. Since $x \cdot \alpha = 0$, we have that

$$0 = (x \cdot \alpha)(v) = x \cdot \alpha(v) - \alpha(x \cdot v),$$

which implies that $\alpha \colon V \to W$ is a Lie module homomorphism.

We return our focus to the L-module V. From Lemma 1.37, we see that $\mathrm{Ker}(\alpha)$ is a submodule of V. Let $v \in \mathrm{Ker}(\alpha) \cap W$, then $\alpha(v) = 0$. However, $\alpha \in H_S - H_0$ implies that $\alpha|_W = \lambda 1_W$, for some $0 \neq \lambda \in \mathbb{F}$. Since $v \in W$, this means that $\alpha(v) = \lambda v$. We conclude that $v = 0$, which implies that $\mathrm{Ker}(\alpha) \cap W = 0$. We have that $\mathrm{Im}(\alpha) \subseteq W$, so by the Rank-Nullity Theorem, we have that $Null(\alpha) = \dim(V) - \dim(W)$. Thus, we have $V = \mathrm{Ker}(\alpha) \oplus W$. An application of Lemma D.18 yields that V is completely reducible. □

Appendix E

Matrix Lie Groups and the Exponential Map

We use this chapter to provide some knowledge on the study of matrix Lie groups. Specifically we describe the differential of a Lie group homomorphism and some properties thereof. This is essential for the work in Section 3.3.

E.1 Introduction

This section defines matrix Lie groups and gives an illustrative example. We denote by $GL_n(\mathbb{R})$ the group of all invertible $n \times n$ matrices with real entries, and $M_n(\mathbb{R})$ the algebra of all $n \times n$ matrices. This is a normed algebra relative to the norm topology $\|(a_{ij})\| := \sqrt{\sum_{ij} a_{ij}^2}$. We may then give $GL_n(\mathbb{R})$ the induced topology. Then a sequence $(A_n)_{n\geq 1}$ of matrices in $GL_n(\mathbb{R})$ converges to A if and only if the sequence $(A_n)_{ij}$ consisting of the ij-th entries of the matrices A_n converges to $(A)_{ij}$ (the ij-th entry of A) as $n \to \infty$.

Definition E.1: *A* **matrix Lie group** *is a closed subgroup G (up to isomorphism) of $GL_n(\mathbb{R})$. That is equivalent to saying that G is closed under nonsingular limits. So for any convergent sequence (A_n) of matrices in G, if $A := \lim_{n\to\infty} A_n$ is invertible then $A \in G$.*

Remark E.2: *We note here that $\mathbb{C}^*$ is a matrix Lie group. Let*

$$G_{\mathbb{C}} := \left\{ \begin{pmatrix} a & b & 0 \\ -b & a & 0 \\ 0 & 0 & z \end{pmatrix} : a, b, z \in \mathbb{R},\ z(a^2 + b^2) = 1 \right\} \subset GL_3(\mathbb{R}).$$

We claim that

$$\begin{aligned}\phi\colon \mathbb{C}^* &\to G_{\mathbb{C}}\\ a+ib &\mapsto \begin{pmatrix} a & b & 0\\ -b & a & 0\\ 0 & 0 & \frac{1}{a^2+b^2}\end{pmatrix},\end{aligned}$$

is an isomorphism. To see that ϕ is a homomorphism it suffices to note that for $a, b, c, d \in \mathbb{R}$ we have

$$\begin{aligned}(ac-bd)^2+(ad+cb)^2 &= a^2c^2+b^2d^2+a^2d^2+b^2c^2+2abcd-2abcd\\ &= a^2c^2+a^2d^2+b^2c^2+b^2d^2=(a^2+b^2)(c^2+d^2).\end{aligned}$$

That ϕ is a bijection is clear. It remains to see that $G_{\mathbb{C}}$ is closed in $GL_3(\mathbb{R})$. To do so we define a map $f\colon M_3(\mathbb{R}) \to \mathbb{R}^7$, by

$$f\begin{pmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix} = (a_{11}-a_{22}, a_{12}+a_{21}, a_{13}, a_{23}, a_{31}, a_{32}, a_{33}(a_{11}^2+a_{12}^2)-1).$$

As f involves polynomials it is clearly continuous. Since $G_{\mathbb{C}} = f^{-1}(0)$ we see that $G_{\mathbb{C}}$ is closed.

E.2 Exponential Map

Here we describe the important exponential map. This is essential in studying matrix Lie group and their connection to Lie algebras. We go on to show how Lie algebras arise from Lie groups. Finally we look at Lie group homomorphisms.

Definition E.3: *We define the* **expontential** *map as follows*

$$\begin{aligned}exp\colon \mathfrak{gl}(n,\mathbb{R}) &\to GL_n(\mathbb{R})\\ M &\mapsto \sum_{i=0}^{\infty}\frac{M^i}{i!}.\end{aligned}$$

We will see below that this converges for all $M \in \mathfrak{gl}(n,\mathbb{R})$. Moreover, since $M^0 = I$ we have that $exp(0) = I$. When it is convenient to do so, we may use the equivalent notation $exp(X) = e^X$.

Proposition E.4: *If $M \in \mathfrak{gl}(n,\mathbb{R})$, then $exp(M)$ converges. In fact, $exp(M)$ converges absolutely and exp is continuous.*

Proof. We will make use of the following two properties (the first is the triangle inequality while the second follows from the Cauchy-Schwarz inequality) which hold for all $A, B \in \mathfrak{gl}(n, \mathbb{R})$,

$$\|A + B\| \leq \|A\| + \|B\|,$$
$$\|AB\| \leq \|A\| \|B\|.$$

We have that
$$\sum_{i=0}^{\infty} \left\| \frac{M^i}{i!} \right\| \leq \|I_n\| + \sum_{i=1}^{\infty} \frac{\|M\|^i}{i!} < \infty.$$

Therefore $exp(M)$ converges absolutely. Since M^n is a continuous function of M, we can see that the partial sums of $exp(M)$ are continuous and therefore we may apply the Weierstrass M-test. We find that $exp(M)$ converges uniformly on each set $U_S := \{X \in \mathfrak{gl}(n, \mathbb{R}) \mid \|X\| \leq S\}$. Thus, $exp(M)$ is continuous on each U_S and hence continuous on $\mathfrak{gl}(n, \mathbb{R})$. □

Proposition E.5: *If $A, B \in \mathfrak{gl}(n, \mathbb{R})$ commute, then $e^A e^B = e^{A+B}$.*

Proof. We note that since $exp(A)$ and $exp(B)$ both converge absolutely we may multiply the series term by term. Hence, we have

$$\begin{aligned}
exp(A)exp(B) &= (I_n + A + \frac{A^2}{2!} + \ldots)(I_n + B + \frac{B^2}{2!} + \ldots) \\
&= exp(B) + Aexp(B) + \frac{A^2}{2!}exp(B) + \ldots \\
&+ exp(A) + Bexp(A) + \frac{B^2}{2!}exp(A) + \ldots \\
&= \sum_{i=0}^{\infty} \sum_{j=0}^{i} \frac{A^j}{j!} \frac{B^{i-j}}{(i-j)!} = \sum_{i=0}^{\infty} \frac{1}{i!} \sum_{j=0}^{i} \binom{i}{j} A^i B^{i-j} \\
&= \sum_{i=0}^{\infty} \frac{(A+B)^i}{i!} = exp(A + B).
\end{aligned}$$

We remark that $\sum_{j=0}^{i} \binom{i}{j} A^i B^{i-j} = (A + B)^i$ holds exactly because A and B commute. □

Corollary E.6: *If $X \in \mathfrak{gl}(n, \mathbb{R})$, then e^X is invertible with $(e^X)^{-1} = e^{-X}$.*

Proof. Using the proposition above we have that

$$exp(X)exp(-X) = exp(X - X) = exp(0) = I_n.$$

□

Proposition E.7: *Let $X \in \mathfrak{gl}(n, \mathbb{R})$. Then $det(exp(X)) = e^{trace(X)}$.*

Proof. We may assume without loss of generality that X is in Jordan canonical form, with eigenvalues $E := \{\lambda_1, \ldots, \lambda_n\}$. If $v \in \mathrm{Ker}(\lambda_i I_n - X)$, then

$$exp(X)v = \sum_{m=1}^{\infty} \frac{X^m v}{m!} = \sum_{m=0}^{\infty} \frac{\lambda_i^m}{m!} v = e^{\lambda_i} v.$$

Hence e_i^{λ} is an eigenvalue of $exp(X)$ if and only if λ_i is an eigenvalue of X. Therefore,

$$det(exp(X)) = \prod_{i=1}^{n} e^{\lambda_i} = e^{\sum_{i=0}^{n} \lambda_i} = e^{trace(X)}.$$

□

Definition E.8: *Given a Lie group G, we define the* **Lie algebra of** *G as*

$$Lie(G) := \{X \in \mathfrak{gl}(n, \mathbb{R}) \,|\, exp(tX) \in G, \ \forall t \in \mathbb{R}\}.$$

We will give an outline of the proof that $Lie(G)$ is not only a vector space but also a Lie algebra with Lie bracket the commutator.

Example E.9: *We will calculate $Lie(SL_n(\mathbb{R}))$, where*

$$SL_n(\mathbb{R}) := \{X \in \mathfrak{gl}(n, \mathbb{R}) \,|\, trace(X) = 0\}.$$

An application of Proposition E.7 yields that

$$\begin{aligned} Lie(SL_n(\mathbb{R})) &= \{M \in \mathfrak{gl}(n, \mathbb{R}) \,|\, det(exp(tM)) = 1, \ \forall t \in \mathbb{R}\} \\ &= \{M \in \mathfrak{gl}(n, \mathbb{R}) \,|\, e^{trace(M)t} = 1, \ \forall t \in \mathbb{R}\}. \end{aligned}$$

If we take the derivative at $t = 0$ we find

$$trace(M)e^0 = \frac{d}{dt} e^{trace(M)t} \Big|_{t=0} = \frac{d}{dt} 1 \Big|_{t=0} = 0.$$

Thus, $trace(M) = 0$. Conversely, suppose that $trace(M) = 0$. Then

$$det(exp(tM)) = e^{trace(M)t} = e^{0t} = 1,$$

and hence $M \in Lie(SL_n(\mathbb{R})$.

Lemma E.10: *If $U \in GL_n(\mathbb{R})$ and $X \in \mathfrak{gl}(n, \mathbb{R})$, then $e^{UXU^{-1}} = Ue^X U^{-1}$.*

Proof. We have that

$$\begin{aligned} exp(UXU^{-1}) &= \sum_{i=0}^{\infty} \frac{(UXU^{-1})^i}{i!} = \sum_{i=0}^{\infty} \frac{UX^iU^{-1}}{i!} = U \left(\sum_{i=0}^{\infty} \frac{X^i}{i!} \right) U^{-1} \\ &= Uexp(X)U^{-1}. \end{aligned}$$

□

The following theorem is important to prove that $Lie(G)$ is a Lie algebra but proving it involves defining a suitable matrix logarithm. We present the result without proof.

Theorem E.11 (Lie Product Formula)**:** *For* $X, Y \in \mathfrak{gl}(n, \mathbb{R})$, *we have*

$$exp(X+Y) = \lim_{m\to\infty} \left[exp\left(\frac{X}{m}\right) exp\left(\frac{Y}{m}\right)\right]^m.$$

We present the following useful result without proof.

Lemma E.12: *Let* G *be a matrix Lie group and* $X \in Lie(G)$. *Then,*

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

Proposition E.13: *Let* G *be a matrix Lie group. The following holds for all* $X, Y \in Lie(G)$.

(i) $AXA^{-1} \in Lie(G)$, *for all* $A \in G$.

(ii) $rX \in Lie(G)$, *for all* $r \in \mathbb{R}$.

(iii) $X + Y \in Lie(G)$.

(iv) $XY - YX \in Lie(G)$.

Proof.

(i): An application of Lemma E.10 yields that

$$exp(tAXA^{-1}) = exp(AtXA^{-1}) = Aexp(tX)A^{-1} \in G,$$

for all $t \in \mathbb{R}$. Hence, $AXA^{-1} \in Lie(G)$.

(ii): Note that $exp(t(rX)) = exp((tr)X) \in G$, for all $t \in \mathbb{R}$.

(iii): Applying Theorem E.11 we may write

$$exp(t(X+Y)) = exp(tX+tY) = \lim_{m\to\infty} \left[exp\left(\frac{tX}{m}\right) exp\left(\frac{tY}{m}\right)\right]^m.$$

For each m we have that

$$\left[exp\left(\frac{tX}{m}\right) exp\left(\frac{tY}{m}\right)\right]^m = \left[exp\left(\frac{t}{m}X\right) exp\left(\frac{t}{m}Y\right)\right]^m \in G,$$

since each $exp\left(\frac{t}{m}X\right) \in G$. By Corollary E.6 the limit is invertible and therefore, since G is closed, the limit is in G. Thus, $X + Y \in Lie(G)$.

(iv): Using the product rule and Lemma E.12 we can see that

$$\frac{d}{dt}(exp(tX)Y exp(-tX))\bigg|_{t=0} = XY - YX.$$

From (i) and Corollary E.6 we have that

$$exp(tX)Y exp(-tX) \in Lie(G),$$

for all $t \in \mathbb{R}$. Properties (ii) and (iii) show that $Lie(G)$ is a real subspace of $\mathfrak{gl}(n, \mathbb{R})$. Considering also that $exp(X)$ is continuous for any $X \in \mathfrak{gl}(n, \mathbb{R})$ yields that $Lie(G)$ is a closed subset of $\mathfrak{gl}(n, \mathbb{R})$. Therefore,

$$XY - YX = \lim_{h \to 0} \frac{exp(hX)Y exp(-hX) - Y}{h} \in Lie(G).$$

This proves (iv).

□

Corollary E.14: *Let G be a matrix Lie group. Then $\mathfrak{g} := Lie(G)$ equipped with Lie bracket $[X, Y] := XY - YX$ is a Lie algebra.*

Proof. Properties (ii) and (iii) show that $\mathfrak{g}$ is a real subspace of $\mathfrak{gl}(n, \mathbb{R})$. The bilinearity and skew-symmetry of the bracket both follow from basic properties of matrix multiplication. To see that the Jacobi identity holds, we note that for $X, Y, Z \in \mathfrak{g}$

$$\begin{aligned}
[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] &= X(YZ - ZY) - (YZ - ZY)X \\
&\quad + Y(ZX - XZ) - (ZX - XZ)Y \\
&\quad + Z(XY - YX) - (XY - YX)Z \\
&= 0.
\end{aligned}$$

The result then follows due to property (iv). □

Remark E.15: *By definition $Lie(GL_n(\mathbb{R})) = \mathfrak{gl}(n, \mathbb{R})$. Moreover, the map $exp\colon \mathfrak{gl}(n, \mathbb{R}) \to GL_n(\mathbb{R})$ can be restricted to $exp\colon Lie(G) \to G$ for any matrix Lie group G.*

Remark E.16: *Let us consider the matrix Lie group $\mathbb{C}^*$. By Remark E.2 have an isomorphism $\phi\colon \mathbb{C}^* \to G_{\mathbb{C}}$. From the definition of ϕ we have that*

$$\begin{aligned}
&Lie(\phi(\mathbb{C}^*)) = \{X \in \mathfrak{gl}(3, \mathbb{R}) \,|\, exp(tX) \in \phi(\mathbb{C}^*),\ \forall t \in \mathbb{R}\} \\
&= \left\{ X \in \mathfrak{gl}(3, \mathbb{R}) \,\middle|\, tX = \begin{pmatrix} ta & tb & 0 \\ -tb & ta & 0 \\ 0 & 0 & tz \end{pmatrix},\ t^3 z(a^2 + b^2) = 1\ \forall t \in \mathbb{R} \right\} \\
&= \left\{ X \in \mathfrak{gl}(3, \mathbb{R}) \,\middle|\, X = \begin{pmatrix} a & b & 0 \\ -b & a & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\} \cong \mathbb{C}.
\end{aligned}$$

Therefore, $Lie(\mathbb{C}^*) = \{z \in \mathbb{C} \,|\, e^z \in \mathbb{C}^*\} = \mathbb{C}$.

Definition E.17: *A homomorphism of matrix Lie groups is a map $\varphi\colon G \to H$, which is a group homomorphism which is continuous (for the induced topologies).*

Definition E.18: *A function $A\colon \mathbb{R} \to GL_n(\mathbb{C})$ is called a* **one-parameter subgroup** *of $GL_n(\mathbb{C})$ if*

1. *A is continuous,*
2. *$A(0) = I_n$, and*
3. *$A(t+s) = A(t)A(s)$, for all $t, s \in \mathbb{R}$.*

We present the following Lemma without proof.

Lemma E.19: *If A is a one-parameter subgroup of $GL_n(\mathbb{C})$, then there exists a unique $X \in M_n(\mathbb{C})$ such that $A(t) = exp(tX)$.*

Theorem E.20: *Let G and H be matrix Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively. For a Lie group homomorphism $\phi\colon G \to H$ there exists a unique real-linear map $d\phi(1)\colon \mathfrak{g} \to \mathfrak{h}$ such that $d\phi(1)$ has the following properties*

(i) $\phi(e^X) = e^{d\phi(1)(X)}$, *for all* $X \in \mathfrak{g}$.

(ii) $d\phi(1)(AXA^{-1}) = \phi(A)d\phi(1)(X)\phi(A)^{-1}$, *for all* $X \in \mathfrak{g},\ A \in G$.

(iii) $d\phi(1)([X,Y]) = [d\phi(1)(X), d\phi(1)(Y)]$, *for all* $X, Y \in \mathfrak{g}$.

(iv) $d\phi(1)(X) = \frac{d}{dt}\phi(e^{tX})\Big|_{t=0}$, *for all* $X \in \mathfrak{g}$.

Proof. As ϕ is continuous, $\phi(e^{tX})$ will be a one-parameter subgroup of H for all $X \in \mathfrak{g}$. An application of Lemma E.19 yields that there is a unique $Z \in M_n(\mathbb{C})$ such that $\phi(e^{tX}) = e^{tZ}$ for all $t \in \mathbb{R}$. We define $d\phi(1)(X) := Z$. Property (i) follows by setting $t = 1$. For $s \in \mathbb{R}$ we have that $\phi(e^{tsX}) = e^{tsZ}$, and so $d\phi(1)(sX) = sd\phi(1)(X)$. Since ϕ is continuous, by an application of the Lie Product Formula E.11 we have that

$$\begin{aligned} e^{td\phi(1)(X+Y)} &= \phi\left(\lim_{m\to\infty}\left[exp\left(\frac{tX}{m}\right)exp\left(\frac{tY}{m}\right)\right]^m\right) \\ &= \lim_{m\to\infty}\left[\phi\left(exp\left(\frac{tX}{m}\right)\right)\phi\left(exp\left(\frac{tY}{m}\right)\right)\right]^m \\ &= \lim_{m\to\infty}\left[exp\left(\frac{td\phi(1)X}{m}\right)exp\left(\frac{td\phi(1)(Y)}{m}\right)\right]^m \\ &= e^{td\phi(1)(X)+td\phi(1)(Y)}. \end{aligned}$$

Using Lemma E.12 we differentiate at $t = 0$ and find

$$d\phi(1)(X+Y) = d\phi(1)(X) + d\phi(1)(Y).$$

Thus, $d\phi(1)$ is real-linear. Suppose τ is another real-linear map which satisfies (i). Then

$$e^{t\tau(X)} = \phi(e^{tX}) = e^{td\phi(1)(X)}.$$

The uniqueness of $d\phi(1)$ follows by differentiating at $t = 0$. We move onto demonstrating that properties (ii)-(iv) also hold.

$$\begin{aligned} e^{td\phi(1)(AXA^{-1})} &= \phi(e^{tAXA^{-1}}) = \phi(Ae^{tX}A^{-1}) = \phi(A)\phi(e^{tX})\phi(A)^{-1} \\ &= \phi(A)e^{td\phi(1)(X)}\phi(A)^{-1}. \end{aligned}$$

Differentiating at $t = 0$ gives property (ii). Now, using the product rule and Lemma E.12 we remark that

$$\frac{d}{dt}e^{tX}Ye^{-tX}\bigg|_{t=0} = XY - YX = [X, Y].$$

Using the property that taking a derivative commutes with a linear transformation and property (ii), we can then see that

$$\begin{aligned} d\phi(1)([X,Y]) = d\phi(1)\left(\frac{d}{dt}e^{tX}Ye^{-tX}\bigg|_{t=0}\right) &= \frac{d}{dt}d\phi(1)(e^{tX}Ye^{-tX})\bigg|_{t=0} \\ &= \frac{d}{dt}\phi(e^{tX})d\phi(1)(Y)\phi(e^{-tX})\bigg|_{t=0} \\ &= \frac{d}{dt}e^{td\phi(1)(X)}d\phi(1)(Y)e^{-td\phi(1)(X)}\bigg|_{t=0} \\ &= d\phi(1)(X)d\phi(1)(Y) - d\phi(1)(Y)d\phi(1)(X) \\ &= [d\phi(1)(X), d\phi(1)(Y)], \end{aligned}$$

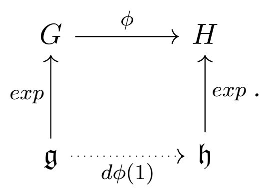
which proves (iii).

Lastly, we have that

$$\phi(e^{tX}) = e^{td\phi(1)(X)},$$

and so (iv) follows using Lemma E.12 and differentiating at $t = 0$. □

Remark E.21: *Property (i) of the above theorem says that the following diagram commutes*



E.3 Automorphism Group

In this section we see that the automorphism group of an algebra form a matrix Lie group. We see the connection between the Lie algebra of this group is this the derivations of the algebra. We explore some properties of this connection.

In this section we denote by U an n-dimensional algebra over $\mathbb{R}$.

Proposition E.22: *The automorphisms of U, denoted $Aut(U)$, form a matrix Lie group.*

Proof. We may identify the vector space U with $\mathbb{R}$. Under this identification $Aut(U)$ is a subgroup of $GL_n(\mathbb{R})$. Set a basis $(u_i)_{i=1}^n$ for U. Then for each $1 \leq i, j \leq n$ we then find that $u_i u_j = \sum_{k=1}^{n} m_{ijk} u_k$, for some $m_{ijk} \in \mathbb{R}$. An automorphism of U is an element $f \in GL_n(\mathbb{R})$, such that for $1 \leq i, j \leq n$

$$f(u_i u_j) = f(u_i) f(u_j).$$

For each $1 \leq i \leq n$ we may write $f(u_i) = \sum_{r=1}^{n} a_{ir} u_r$. Then the following must be satisfied

$$\begin{aligned}
f(u_i u_j) &= f(u_i) f(u_j) \\
\sum_{k=1}^{n} m_{ijk} f(u_k) &= \left(\sum_{s=1}^{n} a_{is} u_s \right) \left(\sum_{t=1}^{n} a_{jt} u_t \right) \\
\sum_{k=1}^{n} m_{ijk} \left(\sum_{r=1}^{n} a_{kr} u_r \right) &= \sum_{s,t=1}^{n} a_{is} a_{jt} u_s u_t \\
\sum_{k,r=1}^{n} m_{ijk} a_{kr} u_r &= \sum_{s,t=1}^{n} a_{is} a_{jt} \left(\sum_{r=1}^{n} m_{str} u_r \right) \\
\sum_{k,r=1}^{n} m_{ijk} a_{kr} u_r &= \sum_{s,t,r=1}^{n} a_{is} a_{jt} m_{str} u_r,
\end{aligned}$$

and so f is an automorphism of U if and only if the following is satisfied

$$\sum_{k,r=1}^{n} m_{ijk} a_{kr} = \sum_{s,t,r=1}^{n} a_{is} a_{jt} m_{str}. \tag{E.1}$$

This is a system of second degree polynomial equations. To see that $Aut(U)$ is closed we may use the same trick as we did in Remark E.2. We may see $Aut(U)$ as the inverse image of a closed set under a continuous mapping (continuous because (E.1) is a system of second degree polynomial equations). □

Lemma E.23: *Let $f \in Der(U)$. Then for $x, y \in U$*

$$f^n(xy) = \sum_{k=0}^{n} \binom{n}{k} f^k(x) f^{n-k}(y),$$

for all $n \in \mathbb{N}$.

Proof. We proceed by induction on n. The case $n = 1$ is trivial.

Assume the result holds for $n - 1 \geq 1$. Then,

$$f^n(xy) = f^{n-1}(f(xy)) = f^{n-1}(xf(y)) + f^{n-1}(f(x)y).$$

Applying the inductive hypothesis twice we find

$$\begin{aligned}
f^n(xy) &= \sum_{k=0}^{n-1} \binom{n-1}{k} f^k(x) f^{n-1-k+1}(y) + \sum_{k=0}^{n-1} \binom{n-1}{k} f^{k+1}(x) f^{n-1-k}(y) \\
&= \sum_{k=0}^{n-1} \binom{n-1}{k} f^k(x) f^{n-k}(y) + \sum_{k=0}^{n-1} \binom{n-1}{k} f^{k+1}(x) f^{n-(k+1)}(y) \\
&= \binom{n-1}{0} x f^n(y) + \left[\binom{n-1}{1} + \binom{n-1}{0}\right] f(x) f^{n-1}(y) \\
&+ \ldots + \left[\binom{n-1}{n-1} + \binom{n-1}{n-2}\right] f^{n-1}(x) f(y) + \binom{n-1}{n-1} f^n(x) y \\
&= \binom{n}{0} x f^n(y) + \binom{n}{1} f(x) f^{n-1}(y) \\
&+ \ldots + \binom{n}{n-1} f^{n-1}(x) f(y) + \binom{n}{n} f^n(x) y \\
&= \sum_{k=0}^{n} \binom{n}{k} f^k(x) f^{n-k}(y).
\end{aligned}$$

□

Proposition E.24: *$Lie(Aut(U)) = Der(U)$.*

Proof. We have that $f \in Lie(Aut(U))$ if and only if $exp(tf) \in Aut(U)$ for any $t \in \mathbb{R}$. Equivalently the following must hold

$$\begin{aligned}
&exp(tf)(xy) = exp(tf)(x) exp(tf)(y) \\
&= (x + tf(x) + \frac{t^2}{2!} f^2(x) + \ldots)(y + tf(y) + \frac{t^2}{2!} f^2(y) + \ldots) \\
&= xy + t[xf(y) + yf(x)] + t^2[\frac{xf^2(y)}{2!} + \frac{yf^2(x)}{2!} + f(x)f(y)] + \ldots,
\end{aligned}$$

where the last equality follows because we may multiply the series term by term ($exp(tf)$ converges absolutely). Since

$$exp(tf)(xy) = xy + tf(xy) + \frac{t^2 f^2(xy)}{2!} + \dots$$

and

$$exp(tf)(xy) = exp(tf)(x)exp(tf)(y),$$

we can subtract xy and divide by t to find that the following must hold:

$$f(xy) + \frac{tf^2(xy)}{2!} + \dots = xf(y) + yf(x) + t[\frac{xf^2(y)}{2!} + \frac{yf^2(x)}{2!} + f(x)f(y)] + \dots.$$

Letting $t \to 0$ we find that

$$f(xy) = xf(y) + f(x)y,$$

and so $f \in Der(U)$.

Take $f \in Der(U)$ and $x, y \in U$, then by Lemma E.23 we have that

$$\begin{aligned}
exp(tf)(xy) &= \sum_{n=0}^{\infty} \frac{t^n}{n!} f^n(xy) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \left(\sum_{k=0}^{n} \binom{n}{k} f^k(x) f^{n-k}(y) \right) \\
&= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \frac{t^n}{k!(n-k)!} f^k(x) f^{n-k}(y) \right) \\
&= xy + t[xf(y) + yf(x)] + t^2 \left[\frac{xf^2(y)}{2!} + \frac{yf^2(x)}{2!} + f(x)f(y) \right] + \dots \\
&= \left(x + tf(x) + \frac{t^2}{2!} f^2(x) + \dots \right) \left(y + tf(y) + \frac{t^2}{2!} f^2(y) + \dots \right) \\
&= exp(tf)(xy) = exp(tf)(x)exp(tf)(y),
\end{aligned}$$

which shows that $exp(tf) \in Aut(U)$ and so $f \in Lie(Aut(U))$. □

Remark E.25: *If $\psi \in Aut(U)$ and $d \in Der(U)$. Then $\psi d \psi^{-1} \in Der(U)$.*

Proof. For $x, y \in U$ we have

$$\begin{aligned}
\psi d \psi^{-1}(xy) &= \psi d(\psi^{-1}(x)\psi^{-1}(y)) = \psi(\psi^{-1}(x) d\psi^{-1}(y) + d\psi^{-1}(x)\psi^{-1}(y)) \\
&= x\psi d\psi^{-1}(y) + \psi d \psi^{-1}(x) y.
\end{aligned}$$

□

Definition E.26: *Let $G := Aut(U)$ and $\mathfrak{g} := Lie(Aut(U)) = Der(U)$. We define*

$$\begin{aligned}
Ad \colon G &\to aut(\mathfrak{g}) \\
\psi &\mapsto Ad_\psi \colon \mathfrak{g} \to \mathfrak{g} \\
&\qquad\qquad d \mapsto \psi d \psi^{-1},
\end{aligned}$$

that is, $Ad_\psi(d) = \psi d \psi^{-1}$.

Proposition E.27: *Let $G := Aut(U)$ and $\mathfrak{g} := Der(U)$. The map $Ad\colon G \to Aut(\mathfrak{g})$ is a matrix Lie group homomorphism.*

Proof. For $\alpha, \beta \in G$ and $f \in \mathfrak{g}$ we have

$$Ad_{\alpha\beta}(f) = \alpha\beta f \beta^{-1}\alpha^{-1} = \alpha Ad_{\beta}(f)\alpha^{-1} = Ad_{\alpha}Ad_{\beta}(f).$$

It remains to show that Ad is continuous. Let F be an open set in $Aut(\mathfrak{g}$. We claim that $C := G \backslash Ad^{-1}(F)$ is closed. Take $(\psi_m) \subseteq C$ such that $\lim_{m\to\infty} \psi_m = \psi$. Suppose $Ad_{\psi} \in F$. Since F is open there is an open ball B around Ad_{ψ} such that $B \subseteq F$. This means that there must be an M such that $M \leq m$ implies that $Ad_{\psi_m} \in F$ but this contradicts $\psi_m \in C$. Thus, C is closed and so Ad is continuous. □

Corollary E.28: *There is a unique linear map*

$$\begin{aligned} ad\colon Der(U) &\to Der(\mathfrak{g}) \\ x \mapsto ad_x\colon \quad & \mathfrak{g} \to \mathfrak{g} \\ & y \mapsto xy - yx, \end{aligned}$$

making the following diagram commute

$$\begin{array}{ccc} Aut(U) & \xrightarrow{Ad} & Aut(\mathfrak{g}) \\ {\scriptstyle exp}\uparrow & & \uparrow{\scriptstyle exp} \\ Der(U) & \underset{ad}{\dashrightarrow} & Der(\mathfrak{g}). \end{array}$$

Note that $Der(U) = \mathfrak{g}$ and $Aut(U) = G$. We will denote $[x, y] := ad_x(y)$.

Proof. From the above proposition we may apply Theorem E.20 and set $ad := dAd(1)$. Proposition E.24 makes it clear that $\mathfrak{g} = Der(U)$, while Remark E.21 shows that the diagram commutes. Property (iv) of Theorem E.20 yields that

$$ad_x = dAd(1)(x) = \frac{d}{dt}Ad(e^{tx})\bigg|_{t=0},$$

and so

$$ad_x(y) = \frac{d}{dt}Ad_{e^{tx}}(y)\bigg|_{t=0} = \frac{d}{dt}e^{tx}ye^{-tx}\bigg|_{t=0} = [x, y].$$

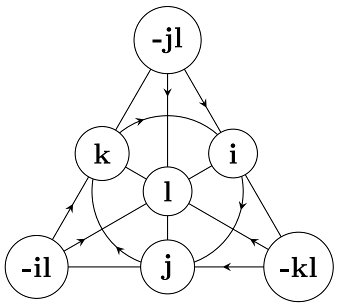
□

Appendix F

Octonions

This short section holds some important information about the octonions which is necessary for the work in Chapter 4. We introduce some maps and results which will also prove useful. All vector spaces in this chapter are finite-dimensional and defined over the field of real numbers.

Definition F.1: *The* **octonions** *are an alternative non-associative non-commutative normed division algebra over the real numbers, denoted by* $\mathbb{O}$. *They have a well-known vector space basis* $B_{\mathbb{O}} := \{1, \mathbf{i}, \mathbf{j}, \mathbf{k}, \mathbf{l}, \mathbf{il}, \mathbf{jl}, \mathbf{kl}\}$ *and their products are fully described by the following diagram.*



To find the product of two elements we follow along the line on which they appear, their product is the next element in the line. The direction of the arrows indicates the product being the positive or negative of the third element along a line. As examples we have $\mathbf{ij} = \mathbf{k}$ *and* $(-\mathbf{il})\mathbf{j} = \mathbf{kl}$. *Using the Fano plane, we can see that* $[\mathbb{O}, \mathbb{O}] = \mathbb{O}_0$.

We can also obtain the octonions by applying the Cayley-Dickson doubling process to the quaternions. Then $\mathbb{O} = \mathbb{H} \oplus \mathbb{H}\mathbf{l}$ *with the product defined as follows for* $q_1, q_2, q_3, q_4 \in \mathbb{H}$.

$$(q_1 + q_2\mathbf{l})(q_3 + q_4\mathbf{l}) = (q_1q_3 - \overline{q_4}q_2) + (q_4q_1 + q_2\overline{q_3})\mathbf{l}.$$

Let $\sigma := \sum_{q \in B_{\mathbb{O}}} a_q q \in \mathbb{O}$. *We call* $\overline{\sigma} := a_1 1 - \sum_{q \in B_{\mathbb{O}} - \{1\}} a_q q$ *the* **conjugate** *of* σ. *The conjugate map is an involution. The* **trace** *of* σ *is defined* *as* $tr(\sigma) = \sigma + \overline{\sigma} \in \mathbb{R}1$. *We set* $\mathbb{O}_0 := \{\sigma \in \mathbb{O} \,|\, tr(\sigma) = 0\}$. *We define an inner product* $n\colon \mathbb{O} \times \mathbb{O} \to \mathbb{R}$ *by* $n(\sigma_1, \sigma_2) := tr(\sigma_1 \overline{\sigma_2})$. *The basis* $B_{\mathbb{O}}$ *is an orthonormal basis relative to* n. *There is a norm* $n\colon \mathbb{O} \to \mathbb{R}_{\geq 0}$ *related to the inner product which is defined by* $n(\sigma) := \frac{1}{2} n(\sigma, \sigma) = \sigma\overline{\sigma}$.

Remark F.2: *To see that* $\mathbb{O}$ *is non-associative, we note that*

$$((\mathbf{ij})\mathbf{l} = \mathbf{kl} \neq -\mathbf{kl} = \mathbf{i}(\mathbf{jl}).$$

Moreover, the only elements which are associative (and commutative) with all of $\mathbb{O}$ *are multiples of* 1.

Proposition F.3: *Suppose* $x, y, z, t \in \mathbb{O}$. *Then*

1. $n(xy) = n(x)n(y)$,

2. $n(x)n(y, z) = n(xy, xz)$,

3. $n(x, t)n(y, z) = n(xy, tz) + n(ty, xz)$, *and*

4. $n(L_x(y), z) = n(y, L_{\overline{x}}(z))$, *where* $L_u(v) := uv$, *for all* $u, v \in \mathbb{O}$.

Proof.

(1): Using alternativity and the fact that the norm of any octonion is real, we find

$$n(xy) = (xy)(\overline{xy}) = (xy)(\overline{y}\,\overline{x}) = xn(y)\overline{x} = x\overline{x}n(y) = n(x)n(y).$$

(2): Using part (1), we note that

$$n(x(y+z)) = n(x)n(y+z) = n(x)\frac{1}{2}[n(y,y) + n(z,z) + 2n(y,z)]$$
$$= n(x)[n(y) + n(z) + n(y,z)] = n(xy) + n(xz) + n(x)n(y,z).$$

Moreover, we also have

$$n(x(y+z)) = n(xy + xz) = \frac{1}{2}[n(xy,xy) + n(xz,xz) + 2n(xy,xz)]$$
$$= n(xy) + n(xz) + n(xy,xz).$$

Therefore, $n(xy) + n(xz) + n(x)n(y,z) = n(xy) + n(xz) + n(xy,xz)$. The result follows.

(3): Using part (2), we note that

$$n(x+t)n(y,z) = n((x+t)y,(x+t)z) = n(xy+ty,xz+tz)$$
$$= n(xy,xz) + n(xy,tz) + n(ty,xz) + n(ty,tz).$$

Moreover,

$$n(x+t)n(y,z) = [n(x)+n(t)+n(x,t)]n(y,z)$$
$$= n(xy,xz) + n(ty,tz) + n(x,t)n(y,z).$$

The result follows since the equations above are all equal.

(4): Using part (3) we find

$$n(xy,z) + n(y,xz) = n(x,1)n(y,z) = tr(x)n(y,z) = n(y,tr(x)z).$$

Therefore,

$$n(L_x(y),z) = n(xy,z) = n(y,tr(x),z) - n(y,xz) = n(y,(tr(x)-x)z)$$
$$= n(y,\overline{x}z) = n(y,L_{\overline{x}}(z)).$$

□

Example F.4: *We consider some examples of Lie algebras arising from the octonions.*

1. *The algebra of all linear transformations of the octonions,* $\mathfrak{gl}(\mathbb{O})$*. Since* $\dim(\mathbb{O}) = 8$ *we can see* $\dim(\mathfrak{gl}(\mathbb{O})) = 64$*. The following examples are all subalgebras of* $\mathfrak{gl}(\mathbb{O})$*.*

2. *We know that* $\dim(\mathbb{O}_0) = 7$*, therefore* $\mathfrak{gl}(\mathbb{O}_0) = 49$*.*

3. *We consider the algebra of all derivations of the octonions*

   $$\mathfrak{g}_2 := Der(\mathbb{O}) := \{d \in \mathfrak{gl}(\mathbb{O}) \,|\, d(xy) = xd(y) + d(x)y, \ \forall x,y \in \mathbb{O}\}.$$

   We discuss this algebra further in Chapter 4.

4. *We consider the special orthogonal algebra relative to the inner product* n

   $$\mathfrak{so}(\mathbb{O}_0,n) := \{f \in \mathfrak{gl}(\mathbb{O}) \,|\, n(x,f(y)) = -n(y,f(x)), \ \forall x,y \in \mathbb{O}_0\}.$$

   This has the same dimension as $\mathfrak{so}(7)$ *(*$\dim(\mathbb{O}_0) = 7$*) which has dimension* 21*.*

Bibliography

[1] Keith Conrad, *Tensor Products*, https://kconrad.math.uconn.edu/blurbs/linmultialg/tensorprod.pdf.

[2] M. de Montigny and J. Patera, *Discrete and continuous graded contractions of Lie algebras and superalgebras*, Journal of Physics: Mathematical and General **24** (1991), 525–547.

[3] Cristina Draper, *Notes from g_2 sessions*, Hand-written notes sent to me in personal correspondence.

[4] Cristina Draper and Cándido Martín, *Gradings on g_2*, Linear Algebra and its Applications **418** (2006), 85–111.

[5] Alberto Elduque and Mikhail Kochetov, *Weyl groups of fine gradings on matrix algebras, octonions and the Albert algebra*, J. Algebra **366** (2012), 165–186.

[6] Karin Erdmann and Mark J. Wildon, *Introduction to Lie Algebras*, no. 978-1-84628-040-5, Springer-Verlag, U.S.A, 2006.

[7] Brian C. Hall, *Lie Groups, Lie Algebras, and Representations*, 2 ed., no. 978-3-319-13467-3, Springer, 2015.

[8] J. E. Humphreys, *Introduction to Lie Algebras and Representation Theory*, no. 0-387-90053-5, Springer-Verlag, 1972.

[9] Cándido Martín, *On the Differential Map*, Document sent during personal correspondence, July 2020.

[10] James S. Milne, *Basic Theory of Affine Group Schemes*, http://www.jmilne.org/math/, March 2012.

[11] Patrick J. Morandi, *The Smith Normal Form of a Matrix*, http://sierra.nmsu.edu/morandi/notes/SmithNormalForm.pdf, February 2005.

[12] William C. Waterhouse, *Introduction to Affine Group Schemes*, no. 978-1-4612-6217-6, Springer-Verlag, 1979.

[13] Steven H. Weintraub, *Jordan Canonical Form: Theory and Practice*, no. 9781608452514, Morgan and Claypool, 2009.