Willoughby Seago

Notes from

Lie Theory

September 26th, 2024
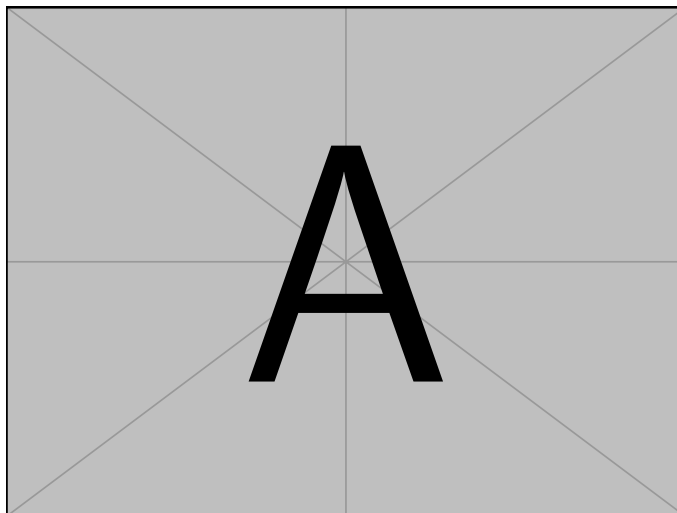
UNIVERSITY OF GLASGOW

Lie Theory

Willoughby Seago

September 26th, 2024

These are my notes from the master's course *Lie Theory* taught at the University of Glasgow by Dr Dinakar Muthiah. I also found this video series[1] from Michael Penn useful. These notes were last updated at 17:13 on November 18, 2024.

Chapters

$$\text{Page}$$

Contents

Page

One

Introduction

In this course we will study Lie algebras. The motivation for the definition of a Lie algebra is as a tool for studying Lie groups. Despite this we will rarely mention Lie groups, since the study of Lie groups requires a knowledge of differential geometry. We will stick to the study of Lie algebras, which can be defined purely algebraically.

Roughly speaking, the storey is thus: Lie groups are both smooth manifolds and groups in a compatible way. By this, we mean that if G is a Lie group then the maps

- $m \colon G \times G \to G$ given by $m(x, y) = xy$
- $i \colon G \to G$ given by $i(g) = g^{-1}$

are smooth and these maps satisfy the obvious requirements to make G a group. In fancy terms, a Lie group is a group object in the category of smooth manifolds.

Lie groups were first used by Sophus Lie, who lends his name to these objects. Lie is pronounced Lee by the way, not lye. He was looking at the continuous symmetries possessed by the solutions to differential equations, and found that these naturally formed what we now know as a Lie group. The problem with studying Lie groups, apart from all of the topology, is that they are generally very non-linear. We can think of Lie algebras as being a linear approximation of Lie groups, given by expanding about the identity and discarding non-linear terms.

To get a Lie algebra out of a Lie group we take the tangent space at the identity, T_eG, which is a vector space. It turns out that the group commutator, $[g, h] = ghg^{-1}h^{-1}$, induces a natural operation on the tangent space, called the Lie bracket. A Lie algebra is then this tangent space equipped with this Lie bracket.

It is possible to abstract the properties of this Lie bracket to define a Lie algebra without reference to a Lie group. It is also possible to go in reverse, given a Lie algebra there is always a corresponding Lie group. It should be noted that this assignment is not unique, in general many different Lie groups can have the same Lie algebra. The assignment does become invertible if we restrict ourselves to simply connected Lie groups, but then we're really getting into topology in a way that this course hopes to avoid. This assignment of a Lie algebra to a Lie group is functorial.

1.1 Notation

Throughout $\Bbbk$ will denote a field. Often we will place further restrictions on $\Bbbk$, such as being algebraically closed. Most of the time we'll be interested in the case

$\Bbbk = \mathbb{C}$, with occasional use of $\Bbbk = \mathbb{R}$. Unless stated otherwise vector spaces are assumed to be vector spaces over $\Bbbk$.

We will denote by 0 the zero vector space (for a given field), {0}, which consists of only the zero vector.

If we have a category, C, we will denote by $\mathsf{C}(A, B) = \mathrm{Hom}_{\mathsf{C}}(A, B)$ the set of all morphisms $A \to B$, or simply $\mathrm{Hom}(A, B)$ if C is clear from context. This is just notation, we won't make much use of categories apart from borrowing some of the language. The main category we'll make use of is $\Bbbk$-Vect, the category of vector spaces over $\Bbbk$ and linear maps.

We will mostly denote Lie algebras with lowercase fraktur letters, or so called \mathfrak letters. For reference, here's the alphabet in \mathfrak:

$$\mathfrak{a}\,\mathfrak{b}\,\mathfrak{c}\,\mathfrak{d}\,\mathfrak{e}\,\mathfrak{f}\,\mathfrak{g}\,\mathfrak{h}\,\mathfrak{i}\,\mathfrak{j}\,\mathfrak{k}\,\mathfrak{l}\,\mathfrak{m}\,\mathfrak{n}\,\mathfrak{o}\,\mathfrak{p}\,\mathfrak{q}\,\mathfrak{r}\,\mathfrak{s}\,\mathfrak{t}\,\mathfrak{u}\,\mathfrak{v}\,\mathfrak{w}\,\mathfrak{x}\,\mathfrak{y}\,\mathfrak{z}$$

Two

Direct Sums and Diagonalisation

Our goal in this section is to state a definition of a diagonalisable linear operator in such a basis free way. We will then generalise this to get an "almost diagonal form" for arbitrary linear operators when $\Bbbk$ is algebraically closed. To this end we will recap some linear algebra in this first section.

2.1 Direct Sums

2.1.1 Binary Direct Sums

> **Definition 2.1.1 — Binary Direct Sum** Let V be a vector space over $\Bbbk$, and let $U_1, U_2 \subseteq V$ be subspaces. Then we say that $V = U_1 \oplus U_2$, that is V is the (internal) **direct sum** of U_1 and U_2, if
>
> - $V = U_1 + U_2 := \{u_1 + u_2 \mid u_1 \in U_1 \text{ and } u_2 \in U_2\}$; and
>
> - $U_1 \cap U_2 = 0$.
>
> If U_1 and U_2 are vector spaces over $\Bbbk$ then we can construct a vector space $V = U_1 \oplus U_2$, also over $\Bbbk$, by defining
>
> $$V = \{(u_1, u_2) \mid u_1 \in U_1 \text{ and } u_2 \in U_2\} \tag{2.1.2}$$
>
> and defining addition and scalar multiplication by
>
> $$(u_1, u_2) + (u'_1, u'_2) = (u_1 + u'_1, u_2 + u'_2) \qquad \text{and} \qquad \lambda(u_1, u_2) = (\lambda u_1, \lambda u_2) \tag{2.1.3}$$
>
> for all $u_1, u'_1 \in U_1$, $u_2, u'_2 \in U_2$, and $\lambda \in \Bbbk$.

After constructing the external direct sum we may identify U_1 with the subspace consisting of elements of the form $(u_1, 0)$ with $u_1 \in U_1$, and U_2 with the subspace of elements of the form $(0, u_2)$ with $u_2 \in U_2$. Then the external direct sum coincides with the internal direct sum. For this reason we won't distinguish internal and external direct sums, making such identifications as necessary.

The following lemma gives an alternative characterisation of the direct sum.

3

Lemma 2.1.4 If $V = U_1 \oplus U_2$ then every vector $v \in V$ can be written *uniquely* as a sum $v = u_1 + u_2$ with $u_1 \in U_1$ and $u_2 \in U_2$. Conversely, if every v has a unique decomposition as $v = u_1 + u_2$ with $u_1 \in U_1$ and $u_2 \in U_2$ then $V = U_1 \oplus U_2$.

Proof. Suppose that $V = U_1 \oplus U_2$. From the definition of a direct sum we know that $v \in V = U_1 + U_2$ and as such $v = u_1 + u_2$ for some $u_1 \in U_1$ and $u_2 \in U_2$ because this is how elements of $U_1 + U_2$ are defined. We need only prove uniqueness. Suppose that $v = u_1 + u_2$ and $v = u_1' + u_2'$ with $u_1, u_1' \in U_1$ and $u_2, u_2' \in U_2$ are two decompositions of v. Then we have $u_1 + u_2 = u_1' + u_2'$, which we can rearrange to get $(u_1 - u_1') + (u_2 - u_2') = 0$. This means that $u_1 - u_1' = u_2' - u_2 =: w$. Now, $u_1 - u_1' \in U_1$, since it's a linear combination of elements of U_1, and similarly $u_2' - u_2 \in U_2$. Thus, $w \in U_1 \cap U_2 = 0$ and so $w = 0$.

Suppose instead that every $v \in V$ has a unique decomposition as $v = u_1 + u_2$. Then clearly v corresponds to $(u_1, u_2) \in U_1 \oplus U_2$ and if $v = u_1 + u_2$ and $v' = u_1' + u_2'$ and $\lambda \in \Bbbk$ then $v + \lambda v' = (u_1 + u_2) + \lambda(u_1' + u_2')$ corresponds to $(u_1, u_2) + \lambda(u_1', u_2')$, but also $v + \lambda v' = (u_1 + \lambda u_1') + (u_2 + \lambda u_2')$ corresponds to $(u_1 + \lambda u_1', u_2 + \lambda u_2')$. This shows that this correspondence defines a linear map. Clearly this correspondence is invertible, and thus we have an isomorphism $V \cong U_1 \oplus U_2$. $\qquad\square$

Thus, the direct sum may be characterised as giving a unique decomposition of each vector into a pair of vectors from two subspaces with only the zero vector in common.

2.1.2 Finite Direct Sums

Direct sums of two vector spaces generalise to direct sums of a finite number of vector spaces in an obvious way.

Definition 2.1.5 — Finite Direct Sum Let V be a vector space over $\Bbbk$, and let $U_1, \dots, U_r \subseteq V$ be subspaces. Then we say that

$$V = U_1 \oplus \cdots \oplus U_r = \bigoplus_{i=1}^{r} U_i, \qquad (2.1.6)$$

that is V is the (internal) **direct sum** of the U_i, if

- $V = U_1 + \cdots + U_r := \{u_1 + \cdots + u_r \mid u_i \in U_i \text{for all } i = 1, \dots, r\}$; and

- for all $i = 1, \dots, r$ we have $U_i \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_r) = 0$ where the sum is over all subspaces apart from U_i.

Note that we can define the external direct sum, but care has to be taken as if we define $U_1 \oplus U_2 \oplus U_3$ to consist of elements of the form (u_1, u_2, u_3) then this is not the same as defining $U_1 \oplus (U_2 \oplus U_3)$ to consist of elements of the form $(u_1, (u_2, u_3))$ and $(U_1 \oplus U_2) \oplus U_3$ to consist of elements of the form $((u_1, u_2), u_3)$. However these spaces are naturally (in the technical sense) isomorphic, and as such we will identify them with each other. A complicated way to put this is that the direct sum

is associative up to natural isomorphism. An even more complicated way to put this, along with the fact that $V \oplus 0 \cong V \cong 0 \oplus V$, is that $(\Bbbk\text{-Vect}, \oplus, 0)$ is a monoidal category.

The same characterisation of the direct sum giving a unique decomposition of $v \in V$ carries over to finite direct sums.

> **Lemma 2.1.7** If $V = U_1 \oplus \cdots \oplus U_r$ then for all $v \in V$ there exist unique $u_i \in U_i$ such that $v = u_1 + \cdots + u_r$.
>
> *Proof.* We proceed by induction on r. The case $r = 2$ is Lemma 2.1.4. Suppose that the result holds some $k \geq 2$ and that $V = U_1 \oplus \cdots \oplus U_k \oplus U_{k+1}$. Take $v \in V$. Writing $V = (U_1 \oplus \cdots \oplus U_k) \oplus U_{k+1}$ we see that there are unique vectors $u \in U_1 \oplus \cdots \oplus U_k$ and $u_{k+1} \in U_{k+1}$ such that $v = u + u_{k+1}$. By the induction hypothesis since $U_1 \oplus \cdots \oplus U_k$ is a k-fold direct sum we can uniquely decompose u as $u = u_1 + \cdots + u_k$. Then $v = u_1 + \cdots + u_k + u_{k+1}$ gives a unique decomposition of v. Thus, by induction the result holds for any finite direct sum. $\square$

2.1.3 Arbitrary Direct Sums

We can further generalise direct sums to arbitrary collections of spaces. One thing we have to be cautious about is that infinite sums of vectors are not defined, and we get around this by considering sums over infinite sets, but requiring that only finitely many of the vectors in the sum are not zero. More formally, we can view a sequence $(u_i)_{i \in I}$ of vectors as a function $u : I \to V$ with $u(i) = u_i$ and then we require that u has finite support.

> **Definition 2.1.8 — Arbitrary Direct Sums** Let V be a vector space over $\Bbbk$, and let $\{U_i\}_{i \in I}$ be an indexed family of subspaces, that is $U_i \subseteq V$ is a subspace for all $i \in I$. Then we say that
>
> $$V = \bigoplus_{i \in I} U_i, \tag{2.1.9}$$
>
> that is V is the (internal) **direct sum** of the U_i, if
>
> - we have
>
> $$V = \sum_{i \in I} U_i := \left\{ \sum_{i \in I} u_i \,\middle|\, u_i \in U_i \text{ and } u_i \neq 0 \text{ for only finitely many terms} \right\};$$
>
> - for all $i \in I$ we have $U_i \cap \sum_{j \in I \setminus \{i\}} U_j = 0$.

We can define the external direct sum by taking a family of vector spaces, $\{U_i\}_{i \in I}$, all over $\Bbbk$, and defining

$$V = \bigoplus_{i \in I} U_i := \{(u_i)_{i \in I} \mid u_i \text{ nonzero for only finitely many } i\} \subseteq \prod_{i \in I} U_i \tag{2.1.10}$$

where

$$\prod_{i \in I} U_i = \{(u_i)_{i \in I}\} \tag{2.1.11}$$

is the direct product of vector spaces (that is, the product in the category $\Bbbk$-Vect). Note that when I is finite the product and direct sum coincide.

Lemma 2.1.12 If $V = \bigoplus_{i \in I} U_i$ then for all $v \in V$ there exist unique $u_i \in U_i$ such that $v = \sum_{i \in I} u_i$ with only finitely many of the u_i being nonzero.

Proof. The proof is essentially the same as before, but now we work with formally infinite sums in which most terms are zero, and all we have to do is check that at each stage this is still the case, which it is, since adding two such sums cannot result in infinitely many nonzero terms. □

Example 2.1.13 Let $V = \Bbbk[x]$ be the vector space of polynomials in x with coefficients in $\Bbbk$. That is

$$V = \left\{ \sum_{i=0}^{N} a_i x^i \,\middle|\, N \in \mathbb{N} \text{ and } a_i \in \Bbbk \right\}. \tag{2.1.14}$$

Addition is just addition of polynomials, and scaling is simply scaling each coefficient, which we can phrase as multiplying by the corresponding constant polynomial. We have the subspaces

$$U_i = \text{span}\{x^i\} := \{a x^i \mid a \in \Bbbk\} \tag{2.1.15}$$

and it's easy to see that

$$V = \bigoplus_{i \in \mathbb{N}} U_i. \tag{2.1.16}$$

Specifically, if we have a polynomial, $f(x) \in \Bbbk[x]$, then

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_N x^N \tag{2.1.17}$$

and $a_0 \in U_0$, $a_1 x \in U_1$, $a_2 x^2 \in U_2$, and so on up to $a_N x^N \in U_N$, so we get a finite decomposition of $f(x)$, and clearly this is unique, showing that $\Bbbk[x]$ is indeed this direct sum as claimed.

(R) There's actually more structure here. First, we can multiply polynomials, so we have an algebra (see Definition 3.3.3), not just a vector space. Define $\deg f(t)$ to be the highest power of t appearing in f with a nonzero coefficient. Then U_i consists of all homogenous polynomials of degree i (homogenous meaning each term has the same degree). Further, since $x^n x^m = x^{n+m}$ we have $\deg(x^n x^m) = n + m$, and so $U_i U_j := \{u_i u_j \mid u_i \in U_i \text{ and } u_j \in U_j\} = U_{i+j}$. When we have a decomposition like this and $U_i U_j \subseteq U_{i+j}$ we say that V is an $\mathbb{N}$-graded algebra.

Remark 2.1.18 The most general definition of the direct sum is as the coproduct in the category $\Bbbk$-Vect. Recall (if you know anything about cat-

egories) that the coproduct $V = \bigoplus_{i \in I} U_i$ comes equipped with inclusion maps $\iota_i \colon U_i \hookrightarrow V$ which are such that if we have a family of maps $f_i \colon U_i \to W$ for some other vector space W then there is a unique linear map $f \colon V \to W$ such that $f\iota_i = f_i$. This is the universal property of the coproduct, and can be summarised as the following diagram commuting for all $j \in I$:

$$
\begin{array}{ccc}
U_j & \xrightarrow{\iota_j} & \bigoplus_{i \in I} U_i \\
& {\scriptstyle f_j} \searrow & \downarrow {\scriptstyle \exists! \, f} \\
& & W.
\end{array}
\tag{2.1.19}
$$

The inclusion map $\iota_i \colon U_i \hookrightarrow V$ is what allows us to identify U_i in the external direct sum with the subspace $\iota_i(U_i) \subseteq V$, and since ι_i is injective we have $U_i \cong \iota_i(U_i)$.

2.2 Diagonalisable Operators and the Direct Sum Decomposition

In this section we'll give the standard definition of a diagonalisable operator, which should be familiar. We'll then give an equivalent definition which makes no mention of a basis.

Let V be a finite dimensional vector space over $\Bbbk$ with basis $\mathcal{B} = \{v_1, \ldots, v_n\}$. Recall that if $T \colon V \to V$ is a linear operator then its matrix in basis $\mathcal{B}$ is the matrix

$$
[T]_{\mathcal{B}} := \begin{bmatrix} \uparrow & & \uparrow \\ Tv_1 & \cdots & Tv_n \\ \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}
\tag{2.2.1}
$$

where c_{ij} is the coefficient of v_i in Tv_j when expressed in the basis $\mathcal{B}$. That is, the ith column of $[T]_{\mathcal{B}}$ is Tv_i as a column vector in basis $\mathcal{B}$.

Definition 2.2.2 — Diagonalisable Operator Let V be a finite dimensional vector space over $\Bbbk$. Then a linear operator, $T \colon V \to V$, is **diagonalisable** if there exists some basis, $\mathcal{B}$, in which the matrix $[T]_{\mathcal{B}}$ is diagonal.

Notice that if $T \colon V \to V$ is a linear operator diagonalised by the basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ then

$$
[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}
\tag{2.2.3}
$$

for some $\lambda_i \in \Bbbk$. We can also express the basis vectors v_i in this basis, they are simply the standard basis, in which $[v_i]_{\mathcal{B}} = e_i$ is the column vector with 1 in the

ith position and 0 everywhere else. Then, for example, we have

$$[Tv_1]_{\mathcal{B}} = [T]_{\mathcal{B}}[v_1]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}\begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \vdots \\ 0 \end{bmatrix} = \lambda_1\begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix} = \lambda_1[v_1]_{\mathcal{B}}. \quad (2.2.4)$$

Thus, we have $Tv_1 = \lambda v_1$, that is, v_1 is an eigenvector of T with eigenvalue λ_1. In general, v_i will be an eigenvector of T with eigenvalue λ_i.

Now make the following observations. We may define the subspaces $U_i = \mathrm{span}\{v_i\}$. The fact that $V = \bigoplus_{i=1}^{n} U_i$ follows immediately from (and is equivalent to) the fact that $\mathcal{B}$ is a basis. Each subspace U_i is T-**invariant**, meaning that $T(U_i) \subseteq U_i$, which simply means that if $u \in U_i$ then $Tu \in U_i$ also, so it's not possible to leave U_i just by the action of T. This is clear in this case because T acts on each subspace, U_i, by scalar multiplication, specifically, by multiplication by λ_i. Because of this it makes sense to consider the restricted linear map $T|_{U_i}: U_i \to U_i$, defined by $T|_{U_i}(u) = T(u)$ for $u \in U_i$. By definition $T(u) = \lambda_i u$ for $u \in U_i$, and thus we have $T = \lambda_i \mathrm{id}_{U_i}$ where $\mathrm{id}_W: W \to W$ is the identity linear map.

2.2.1 Basis Independent Definition

We can now move towards making a basis independent definition of diagonalisability. To do so we need the notion of an eigenspace.

Definition 2.2.5 — Eigenspace Let $\Bbbk$ be a field and let V be a vector space over V. Let $T: V \to V$ be a linear operator. Then for $\alpha \in \Bbbk$ define the **eigenspace** corresponding to α to be

$$W_\alpha := \{w \in V \mid T(w) = \alpha w\}. \quad (2.2.6)$$

Note that W_α can alternatively be characterised as

$$W_\alpha = \{w \in V \mid (T - \alpha)(w) = 0\} = \ker(T - \alpha) \quad (2.2.7)$$

where we perform the common abuse of notation writing $T - \alpha$ when we mean $T - \alpha \mathrm{id}_V$.

For example, in the case of a diagonalisable operator where the eigenvalues are all distinct, as discussed at the end of the previous section, the subspaces U_i are exactly the eigenspaces W_{λ_i}. Also, $W_\alpha = 0$ if $\alpha \neq \lambda_i$ for any i. This is a general fact, if α is not an eigenvalue then there will be no $w \in V$ satisfying $T(w) = \alpha w$.

The problem is that eigenvalues needn't be distinct, in fact, eigenvalues are distinct if and only if the nonzero eigenspaces have dimension 1. Regardless of this problem, we can still perform the direct sum decomposition from the last part, replacing the U_i with the eigenspaces W_α. In fact, this gives us an equivalent definition of diagonalisability without reference to a basis.

Definition 2.2.8 — Diagonalisable Operator Let V be a finite dimensional vector space over $\Bbbk$. Then a linear operator, $T: V \to V$, is **diagonalisable** if

$$V = \bigoplus_{\alpha \in \Bbbk} W_\alpha. \quad (2.2.9)$$

2.2.2 Decomposition Theorem

We now ask if there is a generalisation of this decomposition to not-necessarily-diagonalisable operators. We'll start with an example.

Example 2.2.10 — Non-diagonalisable Operator Consider the case $\Bbbk = \mathbb{C}$, $V = \mathbb{C}^2$, and

$$T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad (2.2.11)$$

in the standard basis

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \text{and} \qquad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \qquad (2.2.12)$$

To find the eigenvalues we solve the characteristic polynomial, which is

$$\chi_T(t) = \det(\lambda - T) = \det \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \lambda^2. \qquad (2.2.13)$$

The only solution to $\lambda^2 = 0$ is $\lambda = 0$, and thus we have one repeated eigenvalue. This means that T is not diagonalisable. For $\alpha \neq 0$ we have $W_\alpha = 0$, and we have

$$W_0 = \{w \in V \mid Tw = 0\} = \ker(T - 0) = \ker T = \operatorname{span}\{e_1\}. \qquad (2.2.14)$$

This follows simply by considering the action of T on the basis vectors, $Te_1 = 0$ and $Te_2 = e_1 \neq 0$. This means that

$$\bigoplus_{\alpha \in \mathbb{C}} W_\alpha = W_0 = \operatorname{span}\{e_1\} \neq \mathbb{C}^2. \qquad (2.2.15)$$

From this calculation we see that $T^2 e_2 = 0$, and so if we instead consider the operator T^2 then the eigenspaces of this cover all of $\mathbb{C}^2$, so this suggests that we should look not just at T, but also powers of T in our decomposition.

Remark 2.2.16 The matrix

$$T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = E \qquad (2.2.17)$$

is important for the Lie algebra $\mathfrak{sl}(2, \mathbb{C})$, which is one of the most important Lie algebras. The fact that this matrix is not diagonalisable is important in that it effects whether E commutes with other matrices. Recall that if two matrices can be simultaneously diagonalised then they commute, an important fact in quantum mechanics.

Following on from the idea that we should look at powers of T we replace the eigenspace with the following definition.

Definition 2.2.18 — Generalised Eigenspace Let $\Bbbk$ be a field and let T be a vector space over $\Bbbk$. Let $T \colon V \to V$ be a linear operator. Then for $\alpha \in \Bbbk$ define the **generalised eigenspace** corresponding to α to be

$$V_\alpha := \{w \in V \mid (T - \alpha)^N(w) = 0 \text{ for some } N \in \mathbb{N}\}. \tag{2.2.19}$$

We can then often make a decomposition into a direct sum of generalised eigenspaces. There's just one hitch, this only works if $\Bbbk$ is algebraically closed.

Definition 2.2.20 — Algebraically Closed Field A field, $\Bbbk$, is algebraically closed if every nonconstant polynomial $f(t) \in \Bbbk[t]$ has a root, $\alpha \in \Bbbk$, such that $f(\alpha) = 0$.

The complex numbers, $\mathbb{C}$, are algebraically closed, this is the fundamental theorem of algebra. The real numbers, $\mathbb{R}$, are not algebraically closed, for example $t^2 + 1$ has no (real) roots.

The following proof relies on three standard results, which we state without proof.

Theorem 2.2.21 — Cayley–Hamilton. If $\chi_T(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$ is the characteristic polynomial of the linear operator $T \colon V \to V$ then the linear operator

$$T^n + a_{n-1}T^{n-1} + \cdots + a_1 T + a_0 \mathrm{id}_V \colon V \to V \tag{2.2.22}$$

is the zero map. That is, T satisfies it's own characteristic polynomial, $\chi_T(T) = 0$.

Lemma 2.2.23 — Bézout's Lemma If R is a principal ideal domain and $x, y \in R$ have greatest common divisor d then there exist $a, b \in R$ such that $ax + by = d$.

Lemma 2.2.24 If $\Bbbk$ is a field then $\Bbbk[x]$ is a PID, and in particular Bézout's lemma applies to polynomials.

Theorem 2.2.25 — Decomposition Theorem. Let $\Bbbk$ be an algebraically closed field, V a finite dimensional vector space, and $T \colon V \to V$ a linear operator. Then

$$V = \bigoplus_{\alpha \in \Bbbk} V_\alpha \tag{2.2.26}$$

where V_α is the generalised eigenspace associated with α and only finitely many of the V_α are nonzero.

Proof. We start by showing that $V_\alpha \cap \sum_{\beta \in \Bbbk \setminus \{\alpha\}} V_\beta = 0$. Take $v V_\alpha \cap \sum_{\beta \in \Bbbk \setminus \alpha} V_\beta$, then we have $v \in V_\alpha$ and $v \in V_{\beta_1} + \cdots + V_{\beta_r}$ for some $\beta_1, \dots, \beta_r \in \Bbbk \setminus \{\alpha\}$. This means that there exist some $N, N_1, \dots, N_r \in \mathbb{N}$ such that

$$(T - \alpha)^N v = (T - \beta_1)^{N_1} \cdots (T - \beta_r)^{N_r} v = 0. \tag{2.2.27}$$

Since $\alpha \neq \beta_i$ we know that the greatest common factor of $(t - \alpha)^N$ and $(t - \beta_1)^{N_1} \cdots (t - \beta_r)^{N_r}$ is 1. Thus, by Bézout's lemma there exist polynomials $f(t), g(t) \in \Bbbk[t]$ such that

$$f(t)(t - \alpha)^N + g(t)(t - \beta_1)^{N_1} \cdots (t - \beta_r)^{N_r} = 1. \tag{2.2.28}$$

Evaluating at T and applying this to v we then have

$$f(T)(T - \alpha)^N v + g(T)(T - \beta_1)^{N_1} v \cdots (T - \beta_r)^{N_r} v = v, \tag{2.2.29}$$

and we know that these operators acting on v both give zero, so the left hand side vanishes, and thus $v = 0$, and so $V_\alpha \cap \sum_{\beta \in \Bbbk \setminus \{\alpha\}} V_\beta = 0$.

We now show that $\sum_{\alpha \in \Bbbk} V_\alpha = V$. Factorise the characteristic polynomial, $\chi_T(t)$, as follows

$$\chi_T(t) = (t - \alpha_1)^{N_1} \cdots (t - \alpha_s)^{N_s}, \tag{2.2.30}$$

with $\alpha_i \neq \alpha_j$ for $i \neq j$. Note that the existence of such a factorisation relies on $\Bbbk$ being algebraically closed. We claim that $v \in V_{\alpha_1} + \cdots + V_{\alpha_s}$. This can be proven by induction on s. The basis case, $s = 1$, has the characteristic polynomial factorise as

$$\chi_T(t) = (t - \alpha_1)^{N_1}. \tag{2.2.31}$$

Then by the Cayley–Hamilton theorem we have

$$(T - \alpha_1)^{N_1} v = 0 \tag{2.2.32}$$

for all $v \in V$, and thus $V_{\alpha_1} = V$. For the inductive step suppose that $s > 1$ and the statement is true for $s - 1$. The highest common factor of $(t - \alpha_s)^{N_s}$ and $(t - \alpha_1)^{N_1} \cdots (t - \alpha_{s-1})^{N_{s-1}}$ is 1. Thus, by Bézout's lemma there exist $f(t), g(t) \in \Bbbk[t]$ such that

$$f(t)(t - \alpha_s)^{N_s} + g(t)(t - \alpha_1)^{N_1} \cdots (t - \alpha_{s-1})^{N_{s-1}} = 1. \tag{2.2.33}$$

Evaluating at T and applying the map to $v \in V$ we have

$$f(T)(T - \alpha_s)^{N_s} v + g(T)(T - \alpha_1)^{N_1} \cdots (T - \alpha_{s-1})^{N_{s-1}} v = v. \tag{2.2.34}$$

Define

$$v' = f(T)(T - \alpha_s)^{N_s} v, \qquad \text{and} \qquad v_s = g(T)(T - \alpha_1)^{N_1} \cdots (T - \alpha_{s-1})^{N_{s-1}} v, \tag{2.2.35}$$

so we have $v = v' + v_s$. Note that we have $(T - \alpha_s)^{N_s} v_s = 0$ and $(T - \alpha_1)^{N_1} \cdots (T - \alpha_{s-1})^{N_{s-1}} v' = 0$. This is simply the Cayley–Hamilton theorem applied to the linear map T restricted to the subspaces corresponding to v_s and v' respectively. Define

$$V' = \{u \in V \mid (T - \alpha_1)^{N_1} \cdots (T - \alpha_{s-1})^{N_{s-1}} u = 0\}. \qquad (2.2.36)$$

Then $v_s \in V_{\alpha_s}$ and $v' \in V'$. The characteristic polynomial of the restricted map, $T|_{V'}$, has all of its roots in $\{\alpha_1, \ldots, \alpha_{s-1}\}$. By induction, we therefore have $v' = v_1 + \cdots + v_{s-1}$ with $v_i \in V'_{\alpha_i} = V' \cap V_{\alpha_i}$. Thus, $v = v_1 + \cdots + v_{s-1} + v_s in V_{\alpha_1} + \cdots + V_{\alpha_s}$. Since $v \in V$ was arbitrary this shows that $V = V_{\alpha_1} + \cdots + V_{\alpha_s}$, proving that $V = V_{\alpha_1} \oplus \cdots \oplus V_{\alpha_s}$. This gives the final result by realising that if $\beta \neq \alpha_i$ then $V_\beta = 0$ since $V_\beta \cap (V_{\alpha_1} + \cdots + V_{\alpha_s}) = 0$ so $V_\beta \cap V = 0$, so $V_\beta = 0$. Thus, $V = \bigoplus_{\alpha \in \Bbbk} V_\alpha$ and V_α is only nonzero for the finite index set $\{\alpha_1, \ldots, \alpha_s\}$. $\qquad\square$

Note that both conditions, the finite dimension of V and the algebraic closure of $\Bbbk$, are required here. Dropping either condition will result in cases where this decomposition doesn't exist.

Example 2.2.37 Consider again the example of $\Bbbk = \mathbb{C}$, $V = \mathbb{C}^2$, and

$$T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \qquad (2.2.38)$$

As we saw before (Example 2.2.10) T is not diagonalisable, and $\bigoplus_{\alpha \in \mathbb{C}} W_\alpha = W_0 = \text{span}\{e_1\} \neq \mathbb{C}^2$. We also saw there that $T^2 e_2 = 0$, meaning that when considering generalised eigenspaces we also have $e_2 \in V_0$, and thus

$$\bigoplus_{\alpha \in \mathbb{C}} V_\alpha = V_0 = \mathbb{C}^2. \qquad (2.2.39)$$

Definition 2.2.40 — Jordan Block and Jordan Normal Form A **Jordan block** is a matrix of the form

$$\begin{bmatrix} \alpha & 1 & & & \\ & \alpha & 1 & & \\ & & \ddots & \ddots & \\ & & & \alpha & 1 \\ & & & & \alpha \end{bmatrix}. \qquad (2.2.41)$$

That is, we have some $\alpha \in \Bbbk$ on the diagonal and 1s above the diagonal, and 0 everywhere else.

A matrix, X, is a sum of Jordan blocks, or is in **Jordan normal form** if it

is given by the block diagonal matrix

$$X = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{bmatrix} \tag{2.2.42}$$

with the J_i Jordan blocks.

Corollary 2.2.43 — Existence of Jordan Normal Form Let $\Bbbk$ be an algebraically closed field, let V be a finite dimensional vector space, and let $T : V \to V$ be a linear map. There exists a basis of V such that $[T]_{\mathcal{B}}$ is in Jordan normal form.

Proof. Consider the decomposition of V into a sum of generalised eigenspaces of T. Each generalised eigenspace is invariant under T, and thus if $v_\alpha \in V_\alpha$ then $Tv_\alpha = c_1 v_1 + \cdots c_r v_r$ where $\{v_1, \dots, v_r\}$ is a basis for V_α. We can further choose this basis by scaling and taking linear combinations untile $Tv_i = \alpha v_i + v_{i+1}$, which can be shown by induction on the size of the dimension of V_α. Then T acts on V_α in this basis as a Jordan block, and so T acts on V, in the basis formed by the union of these bases, as a matrix in Jordan normal form. $\square$

Three

Lie Algebras

3.1 Definition and Remarks

Definition 3.1.1 — Lie Algebra Let $\Bbbk$ be a field. A **Lie algebra** over $\Bbbk$ is a vector space, $\mathfrak{g}$, over $\Bbbk$ equipped with a bilinear map,

$$[-,-] \colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} \tag{3.1.2}$$
$$(X,Y) \mapsto [X,Y], \tag{3.1.3}$$

called the Lie bracket, or simply the bracket. This must satisfy two properties:

1. **alternating**: for all $X \in \mathfrak{g}$ we have $[X,X] = 0$;

2. **Jacobi identity**: for all $X, Y, Z \in \mathfrak{g}$ we have

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0. \tag{3.1.4}$$

Bilinearity means that for all $\lambda \in \Bbbk$ and $X, Y, Z \in \mathfrak{g}$ we have

$$[X + \lambda Y, Z] = [X,Z] + \lambda[Y,Z], \tag{3.1.5}$$
$$[X, Y + \lambda Z] = [X,Y] + \lambda[X,Z]. \tag{3.1.6}$$

We say that $[-,-]$ is antisymmetric if $[X,Y] = -[Y,X]$. It turns out that this property is almost the same as, but not quite, the alternating property, and as such the definition of a Lie algebra is often given with antisymmetry in place of alternativity, the catch being that over fields of characteristic 2 these aren't equivalent. Alternativity is the "correct" condition, but if we're only looking at $\Bbbk = \mathbb{C}, \mathbb{R}$, as is often the case, then there's little harm in taking antisymmetry as the defining condition.

Lemma 3.1.7 The Lie bracket is antisymmetric.

Proof. Let $\mathfrak{g}$ be a Lie algebra and take $X, Y \in \mathfrak{g}$. Consider the bracket of $X+Y$ with itself. On the one hand, alternativity tells us that $[X+Y, X+Y] =$

0, and on the other hand using bilinearity and alternativity we have

$$[X + Y, X + Y] = [X, X] + [X, Y] + [Y, X] + [Y, Y] \tag{3.1.8}$$
$$= [X, Y] + [Y, X] \tag{3.1.9}$$

Now, if this is to vanish we must have

$$[X, Y] = -[Y, X]. \tag{3.1.10}$$

Lemma 3.1.11 Over a field of characteristic other than 2 antisymmetry of a bilinear bracket implies alternativity.

Proof. Let $\mathfrak{g}$ be a vector space over $\Bbbk$ with char $\Bbbk \neq 2$, equipped with an *antisymmetric* bracket, $[-, -] \colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$. Antisymmetry means that $[X, Y] = -[Y, X]$ for all $X, Y \in \mathfrak{g}$. In particular, we are free to take $X = Y$, then we have $[X, X] = -[X, X]$. Rearranging this gives

$$[X, X] + [X, X] = 2[X, X] = 0. \tag{3.1.12}$$

Now, in a field with $2 \neq 0$ (that is, in a field of characteristic other than 2) this immediately implies $[X, X] = 0$. $\qquad\square$

This distinction is subtle, and ultimately not that important since we'll mostly concern ourselves with $\Bbbk = \mathbb{C}$, which has characteristic 0 (and char $\mathbb{R} = 0$ also).

The Jacobi identity is a bit weird the first time you see it. There are a couple of ways to think about it that help to remember the definition. The first is to notice that the second and third term are cyclic permutations of the first, so it's often shortened to

$$[X, [Y, Z]] + \text{cycilc permutations} = 0. \tag{3.1.13}$$

The second is to write it like

$$[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]], \tag{3.1.14}$$

which is an equivalent form. This doesn't look much simpler, but fix some $X \in \mathfrak{g}$ and define a linear map $D \colon \mathfrak{g} \to \mathfrak{g}$ by $D(Y) = [X, Y]$. Then this becomes

$$D([Y, Z]) = [D(Y), Z] + [Y, D(Z)]. \tag{3.1.15}$$

To make this even simpler, write $A \cdot B$ in place of $[A, B]$, and we have

$$D(Y \cdot Z) = D(Y) \cdot Z + Y \cdot D(Z) \tag{3.1.16}$$

and we see that this is a version of the product rule, or Leibniz rule. So D acts a bit like a derivative. The fancy way to say this is that the adjoint representation of $\mathfrak{g}$ acts on $\mathfrak{g}$ by derivations. A **derivation** is just any linear map on an algebra satisfying the Leibniz rule with respect to the product of that algebra, which for a Lie group is the Lie bracket. The "adjoint representation" (Definition 4.1.3) is simply $\mathfrak{g}$ acting on itself where X acts on Y by $X \cdot Y \mapsto [X, Y]$.

[1]I've brought this forwards in the notes, it feels wrong to define objects and not move on to morphisms immediately.

Now[1] that we've defined Lie algebras, an algebraic object, we should define maps between them. The appropriate maps will be "structure preserving", in the same way that a group homomorphism is structure preserving (it preserves the multiplicative structure) and a linear map is structure preserving (it preserves addition and scalar multiplication). The structure that we have to preserve here is the Lie bracket, as well as the underlying vector space structure of the Lie algebra.

> **Definition 3.1.17 — Homomorphism** Let $\mathfrak{g}$ be a Lie algebra with bracket $[-,-]_\mathfrak{g}$, and let $\mathfrak{h}$ be a Lie algebra with bracket $[-,-]_\mathfrak{h}$. Then a **Lie algebra homomorphism**, $\varphi : \mathfrak{g} \to \mathfrak{h}$, is a linear map which preserves the bracket, meaning that for all $X, Y \in \mathfrak{g}$ we have
>
> $$\varphi([X, Y]_\mathfrak{g}) = [\varphi(X), \varphi(Y)]_\mathfrak{h}. \tag{3.1.18}$$
>
> An invertible Lie algebra homomorphism is a **Lie algebra isomorphism**.

We'll usually just say "homomorphism", or simply "morphism", rather than "Lie algebra homomorphism", and likewise we'll just speak of an "isomorphism" rather than a "Lie algebra isomorphism". We will writhe $\mathfrak{g} \cong \mathfrak{h}$ if there is a Lie algebra isomorphism $\mathfrak{g} \to \mathfrak{h}$. For notational simplicity we'll usually write $[-,-]$ for the bracket of both $\mathfrak{g}$ and $\mathfrak{h}$, allowing context (i.e., what elements we put into it) to make clear which we mean.

We now prove the standard results about morphisms.

> **Lemma 3.1.19** If $\varphi : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra isomorphism then $\varphi^{-1} : \mathfrak{h} \to \mathfrak{g}$ is a Lie algebra isomorphism.
>
> *Proof.* We need to show that for all $X', Y' \in \mathfrak{h}$ we have
>
> $$\varphi^{-1}([X', Y']_\mathfrak{h}) = [\varphi^{-1}(X'), \varphi^{-1}(Y')]_\mathfrak{g}. \tag{3.1.20}$$
>
> To do so note that φ^{-1} is an invertible linear map, and as such is a bijection. This means that there exist unique $X, Y \in \mathfrak{g}$ such that $\varphi(X) = X'$ and $\varphi(Y) = Y'$. Thus, we may write the left hand side as
>
> $$\varphi^{-1}([X', Y']_\mathfrak{h}) = \varphi^{-1}([\varphi(X), \varphi(Y)]_\mathfrak{h}). \tag{3.1.21}$$
>
> Using the fact that φ is an isomorphism, and hence is a homomorphism, we can pull the φ out of the bracket on the right, to give
>
> $$\varphi^{-1}([X', Y']_\mathfrak{h}) = \varphi^{-1}(\varphi([X, Y]_\mathfrak{g})). \tag{3.1.22}$$
>
> Then, since by definition $\varphi^{-1} \circ \varphi = \mathrm{id}_\mathfrak{g}$ we have
>
> $$\varphi^{-1}([X', Y']_\mathfrak{h}) = [X, Y]_\mathfrak{g}. \tag{3.1.23}$$
>
> We can then invert $X' = \varphi(X)$ and $Y' = \varphi(Y)$ to write $X = \varphi^{-1}(X')$ and $Y = \varphi^{-1}(Y')$. This lets us write this as
>
> $$\varphi^{-1}([X', Y']_\mathfrak{h}) = [\varphi^{-1}(X'), \varphi^{-1}(Y')]_\mathfrak{g}, \tag{3.1.24}$$
>
> which is the result we wanted. $\square$

Lemma 3.1.25 The composite of Lie algebra homomorphisms is a Lie algebra homomorphism.

Proof. Let $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{l}$ be Lie algebras with brackets $[-,-]_\mathfrak{g}$, $[-,-]_\mathfrak{h}$, and $[-,-]_\mathfrak{l}$ respectively. Let $\varphi \colon \mathfrak{g} \to \mathfrak{h}$ and $\psi \colon \mathfrak{h} \to \mathfrak{l}$ be Lie algebra homomorphisms. Then we may consider $\psi \circ \varphi \colon \mathfrak{g} \to \mathfrak{l}$. We wish to show that this is a Lie algebra homomorphism.

First, note that ψ and φ are linear, so their composite is too. We then need only show that for all $X, Y \in \mathfrak{g}$ we have

$$(\psi \circ \varphi)([X,Y]_\mathfrak{g}) = [(\psi \circ \varphi)(X), (\psi \circ \varphi)(Y)]_\mathfrak{l}. \tag{3.1.26}$$

Starting with the left-hand side we can use the definition of composition to write

$$(\psi \circ \varphi)([X,Y]_\mathfrak{g}) = \psi(\varphi([X,Y]_\mathfrak{g})). \tag{3.1.27}$$

Using the fact that φ is a Lie algebra homomorphism into $\mathfrak{h}$ we have $\varphi([X,Y]_\mathfrak{g}) = [\varphi(X), \varphi(Y)]_\mathfrak{h}$, so we have

$$(\psi \circ \varphi)([X,Y]_\mathfrak{g}) = \psi([\varphi(X), \varphi(Y)]_\mathfrak{h}). \tag{3.1.28}$$

Now, using the fact that ψ is a Lie algebra homomorphism into $\mathfrak{l}$ we have $\psi([X',Y']_\mathfrak{h}) = [\psi(X'), \psi(Y')]_\mathfrak{l}$ for any $X', Y' \in \mathfrak{h}$, and in particular this is true when $X' = \varphi(X)$ and $Y' = \varphi(Y)$, giving

$$(\psi \circ \varphi)([X,Y]_\mathfrak{g}) = [\psi(\varphi(X)), \psi(\varphi(Y))]_\mathfrak{l}. \tag{3.1.29}$$

Finally, using the definition of composition again, we get our result,

$$(\psi \circ \varphi)([X,Y]_\mathfrak{g}) = [(\psi \circ \varphi)(X), (\psi \circ \varphi)(Y)]_\mathfrak{l}. \tag{3.1.30}$$

Noting that function composition is associative we have the following result.

Corollary 3.1.31 For a field $\Bbbk$ there is a category $\Bbbk$-LieAlg whose objects are Lie algebras over $\Bbbk$ and whose morphisms are Lie algebra homomorphisms.

The objects of this category, Lie algebras, are vector spaces, and the morphisms are linear maps. This means that $\Bbbk$-LieAlg is a subcategory of $\Bbbk$-Vect, and as such inherits a lot of its structure from $\Bbbk$-Vect. For example, the coproduct (direct sum) of Lie algebras is just the coproduct (direct sum) of the underlying vector spaces, equipped with an appropriate bracket. That is, given Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ we can equip the vector space $\mathfrak{g} \oplus \mathfrak{h}$ with a Lie algebra structure by defining

$$[X_1 + Y_1, X_2 + Y_2]_{\mathfrak{g} \oplus \mathfrak{h}} = [X_1, X_2]_\mathfrak{g} + [Y_1, Y_2]_\mathfrak{h} \tag{3.1.32}$$

where $X_1 + Y_1$ and $X_2 + Y_2$ are two arbitrary elements of $\mathfrak{g} \oplus \mathfrak{h}$ decomposed into $X_1, X_2 \in \mathfrak{g}$ and $Y_1, Y_2 \in \mathfrak{h}$. The alternativity and Jacobi identity for $\mathfrak{g}$ and $\mathfrak{h}$ apply

directly when brackets in $\mathfrak{g} \oplus \mathfrak{h}$ are written as on the right-hand side here, and so these properties also hold for the bracket on $\mathfrak{g} \oplus \mathfrak{h}$. This gives the external direct sum. We can similarly give an internal direct sum by saying that if $\mathfrak{g}$ is a Lie algebra with subalgebras $\mathfrak{h}_1$ and $\mathfrak{h}_2$ then $\mathfrak{g} = \mathfrak{h}_1 \oplus \mathfrak{h}_2$ if $\mathfrak{g} = \mathfrak{h}_1 \oplus \mathfrak{h}_2$ as vector spaces and for all $X \in \mathfrak{h}_1$ and $Y \in \mathfrak{h}_2$ we have $[X, Y] = 0$. The idea here is that the two parts of the direct sum should not "interact". For vector spaces this simply means that they should be disjoint except for 0, and for Lie algebras this means that elements from the two parts should commute (which is what we say when the bracket vanishes, since often the bracket is literally a commutator).

3.2 Subalgebras, Ideals, and Quotients

To state these definitions it's useful to abuse the notation slightly as follows. Let $U, V \subseteq \mathfrak{g}$ be subspaces of a Lie algebra. Define the subspace

$$[U, V] = \text{span}\{[u, v] \mid u \in U \text{ and } v \in V\}. \tag{3.2.1}$$

> **Definition 3.2.2 — Subalgebra** Let $\mathfrak{g}$ be a Lie algebra. A **Lie subalgebra** (or just **subalgebra**), $\mathfrak{h}$, is a subspace which is closed under the Lie bracket. That is, $[\mathfrak{h}, \mathfrak{h}] \subseteq \mathfrak{h}$, or equivalently for all $h, h' \in \mathfrak{h}$ we have $[h, h'] \in \mathfrak{h}$.

The notion of a Lie subalgebra is the Lie algebra analogue of a subset of a set, subspace of a vector space, subgroup of a group, or subring of a ring. We can't form a quotient by an arbitrary subset, subgroup or subring, we need a set generated by an equivalence relation, a normal subgroup, or an ideal. This is exactly the case with Lie algebras as well.

Note that a Lie subalgebra, $\mathfrak{h}$, is necessarily a Lie algebra in its own right after restricting the bracket to $\mathfrak{h}$. This follows immediately because the requirements for a Lie algebra are universally quantified so remain true after restricting to a subspace.

> **Definition 3.2.3 — Ideal** Let $\mathfrak{g}$ be a Lie algebra. An **ideal** of $\mathfrak{g}$ is a Lie subalgebra, $\mathfrak{i}$, such that $[\mathfrak{i}, \mathfrak{g}] \subseteq \mathfrak{i}$, or equivalently for all $X \in \mathfrak{g}$ and $I \in \mathfrak{i}$ we have $[X, I] \in \mathfrak{i}$.

Note that there's no notion of left- or right-ideals, any ideal is two sided since $[\mathfrak{h}, \mathfrak{h}'] = [\mathfrak{h}, \mathfrak{h}']$, as elements of the two differ only by a sign and the fact that these are subspaces means that all the elements that differ only by a sign are also included.

With ideals we can define quotients.

> **Definition 3.2.4 — Quotient** Let $\mathfrak{g}$ be a Lie algebra, and let $\mathfrak{i} \subseteq \mathfrak{g}$ be an ideal. Then the **quotient** vector space $\mathfrak{g}/\mathfrak{i}$ is a Lie algebra when we define the bracket by
>
> $$[X + \mathfrak{i}, Y + \mathfrak{i}] = [X, Y] + \mathfrak{i}. \tag{3.2.5}$$

Note that in the definition of the bracket of $\mathfrak{g}/\mathfrak{i}$ the bracket $[X + \mathfrak{i}, Y + \mathfrak{i}]$ is computed in $\mathfrak{g}/\mathfrak{i}$, and the bracket $[X, Y]$ is computed in $\mathfrak{g}$.

Lemma 3.2.6 The quotient algebra as defined above is a Lie algebra.

Proof. Let $\mathfrak{g}$ be a Lie algebra and let $\mathfrak{i}$ be an ideal. We have four things to show:

1. The bracket on $\mathfrak{g}/\mathfrak{i}$ is well-defined;

2. The bracket on $\mathfrak{g}/\mathfrak{i}$ is bilinear;

3. The bracket on $\mathfrak{g}/\mathfrak{i}$ is alternating;

4. The bracket on $\mathfrak{g}/\mathfrak{i}$ satisfies the Jacobi identity.

The first relies on $\mathfrak{i}$ being an ideal, and the others are simply an exercise in algebra.

STEP 1: WELL-DEFINED

Let $X, X' \in \mathfrak{g}$ be such that $X + \mathfrak{i} = X' + \mathfrak{i}$. Recall that this means $X - X' \in \mathfrak{i}$. Then for all $Y \in \mathfrak{g}$ we have

$$[X + \mathfrak{i}, Y + \mathfrak{i}] = [X, Y] + \mathfrak{i} \tag{3.2.7}$$

and

$$[X' + \mathfrak{i}, Y + \mathfrak{i}] = [X', Y] + \mathfrak{i}. \tag{3.2.8}$$

We need to show that these are equal. This means we need to show that

$$[X, Y] - [X', Y] \in \mathfrak{i}. \tag{3.2.9}$$

This is necessarily true however, since bilinearity gives us

$$[X, Y] - [X', Y] = [X - X', Y] \tag{3.2.10}$$

which is in $\mathfrak{i}$ since $X - X' \in \mathfrak{i}$ and $\mathfrak{i}$ is an ideal.

STEP 2: BILINEAR

We will demonstrate linearity in the first argument. Linearity in the second argument then follows from a similar argument, or in a field of characteristic other than 2 it follows from linearity in the first argument and alternativity, which implies anticommutativity. Let $X, X', Y \in \mathfrak{g}$, then

$$[(X + \mathfrak{i}) + (X' + \mathfrak{i}), Y + \mathfrak{i}] = [(X + X') + \mathfrak{i}, Y + \mathfrak{i}] \tag{3.2.11}$$
$$= [X + X', Y] + \mathfrak{i} \tag{3.2.12}$$
$$= ([X, Y] + [X', Y]) + \mathfrak{i} \tag{3.2.13}$$
$$= ([X, Y] + \mathfrak{i}) + ([X', Y] + \mathfrak{i}). \tag{3.2.14}$$

Further, take $\lambda \in \Bbbk$ and we have

$$[(\lambda X) + \mathfrak{i}, Y + \mathfrak{i}] = [\lambda X, Y] + \mathfrak{i} \tag{3.2.15}$$
$$= (\lambda[X, Y]) + \mathfrak{i} \tag{3.2.16}$$
$$= \lambda([X, Y] + \mathfrak{i}). \tag{3.2.17}$$

Thus, the bracket is linear in the first argument.

STEP 3: ALTERNATING
Let $X \in \mathfrak{g}$, then

$$[X + \mathfrak{i}, X + \mathfrak{i}] = [X, X] + \mathfrak{i} = 0 + \mathfrak{i} = \mathfrak{i} = 0 \qquad (3.2.18)$$

since $\mathfrak{i} = 0 + \mathfrak{i}$ is the zero vector in the quotient space $\mathfrak{g}/\mathfrak{i}$.

STEP 4: JACOBI IDENTITY
Let $X, Y, Z \in \mathfrak{g}$,

$$[X + \mathfrak{i}, [Y + \mathfrak{i}, Z + \mathfrak{i}]] = [X + \mathfrak{i}, [Y, Z] + \mathfrak{i}] = [X, [Y, Z]] + \mathfrak{i}. \quad (3.2.19)$$

Doing this with the other terms in the Jacobi identity we see that we end up with the left hand side of the Jacobi identity in $\mathfrak{g}/\mathfrak{i}$ being

$$([X, [Y, Z]] + \mathfrak{i}) + ([Y, [Z, X]] + \mathfrak{i}) + ([Z, [X, Y]] + \mathfrak{i}). \qquad (3.2.20)$$

Recall that addition in the quotient space is defined by $(X + \mathfrak{i}) + (Y + \mathfrak{i}) = (X + Y) + \mathfrak{i}$, and so this becomes

$$([X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]]) + \mathfrak{i} = 0 + \mathfrak{i} = 0 \qquad (3.2.21)$$

where we've used the Jacobi identity in $\mathfrak{g}$, and then identified $0 + \mathfrak{i}$ as the zero vector in the quotient $\mathfrak{g}/\mathfrak{i}$. $\qquad\square$

3.3 Examples of Lie Algebras

We've done quite a lot now without ever actually looking at any examples of a Lie algebra, so let's change that.

3.3.1 Abelian Lie Algebra

We start with the simplest example, where the bracket always vanishes.

Definition 3.3.1 — Abelian Lie Algebra A Lie algebra, $\mathfrak{g}$, is **abelian** if $[\mathfrak{g}, \mathfrak{g}] = 0$, that is, if $[X, Y] = 0$ for all $X, Y \in \mathfrak{g}$.

The terminology here, abelian, should not be confused with this terminology referring to groups. An abelian Lie algebra is not one in which the bracket is "commutative" (although it is, since every bracket is the 0 we do technically have $[X, Y] = [Y, X]$). Instead, this terminology is inherited from the Lie group, an abelian Lie group will give us an abelian Lie algebra.

Note that, up to isomorphism, there is one abelian Lie algebra of each dimension for each field. Any vector space can be made into an abelian Lie algebra by equipping it with the trivial bracket, and any two such Lie algebras are isomorphic so long as the underlying vector spaces are. For this reason abelian Lie algebras on their own are "boring", we only look at them when they arise naturally as subalgebras of nonabelian Lie algebras.

3.3.2 Low Dimension

In an attempt to classify Lie algebras (something we'll return to later) one might start with dimension. Let's do that. For dimension 0 there is one Lie algebra, the zero vector space with the trivial bracket.

For dimension 1 (and a fixed field) there is one Lie algebra (up to isomorphism), the one-dimensional space equipped with the trivial bracket. Note that there cannot be a nonabelian one-dimensional Lie algebra. Suppose that $\mathfrak{g}$ was such a Lie algebra, then $\mathfrak{g} = \Bbbk X = \mathrm{span}\{X\}$ for some X. Then if $Y, Z \in \mathfrak{g}$ we know that $Y = \lambda X$ and $Z = \mu X$, and then

$$[Y,Z] = [\lambda X, \mu X] = \lambda\mu[X,X] = \lambda\mu \cdot 0 = 0. \tag{3.3.2}$$

Thus, all one-dimensional Lie algebras are necessarily abelian.

For dimension 2 (and a fixed field) we of course have the abelian Lie algebra, but there is another. Suppose that $\mathfrak{g}$ is two-dimensional with basis $\{X, Y\}$. Assuming that $\mathfrak{g}$ is nonabelian the bracket $[X, Y] = -[Y, X]$ must be nonzero. This means that $[\mathfrak{g}, \mathfrak{g}]$ is a one-dimensional subspace of $\mathfrak{g}$, since every element of $[\mathfrak{g}, \mathfrak{g}]$ is a linear combination of brackets of elements of $\mathfrak{g}$, but all nonzero such brackets are multiples of $[X, Y]$. Further, we can rescale X so that $[X, Y] = Y$, and this is the only 2-dimensional non-abelian Lie algebra up to isomorphism (for a fixed field). So, there are two Lie algebras of dimension 2 up to isomorphism.

In three dimensions one can make similar arguments, and it turns out that there are infinitely many isomorphism classes of three-dimensional Lie algebra, but we can still classify them. Here are the isomorphism classes over $\mathbb{C}$ (or any algebraically closed field of characteristic 0):

- The abelian Lie algebra of dimension 3.

- $\mathfrak{g} = \mathrm{span}\{X, Y, Z\}$ with $[X, Y] = Z$ and $[X, Z] = [Y, Z] = 0$[2].

- $\mathfrak{g} = \mathrm{span}\{X, Y, Z\}$ with $[X, Y] = Y$ and $[X, Z] = [X, Z] = 0$.

- $\mathfrak{g}_b = \mathrm{span}\{X, Y, Z\}$ with $[X, Y] = Y$, $[Y, Z] = bZ$, and $[Y, Z] = 0$ for $b \in \mathbb{C}^\times$. Note that $\mathfrak{g}_b \cong \mathfrak{g}_{b'}$ if and only if $b = b'$ or $b = 1/b'$.

- $\mathfrak{g} = \mathrm{span}\{X, Y, Z\}$ with $[X, Y] = Y$, $[X, Z] = Y + Z$, and $[Y, Z] = 0$.

- $\mathfrak{g} = \mathrm{span}\{X, Y, Z\}$ with $[X, Y] = 2X$, $[X, Z] = -2Z$, and $[Y, Z] = X$.

[2]This is the three-dimensional **Heisenberg algebra**.

Of these, by far the most important is the last one, which we'll later see is the Lie algebra known as $\mathfrak{sl}(2, \mathbb{C})$.

If we classify over $\mathbb{R}$ instead then we get another familiar example of a Lie algebra, $\mathbb{R}^3$ equipped with the bracket given by the cross-product.

Hopefully, these examples are enough to convince you that a full classification for Lie algebras of dimension 4 is, while maybe possible, probably quite challenging.

All hope is not lost, we do have, as we'll see later, a classification of a particularly nice type of Lie algebra, called simple Lie algebras, over $\mathbb{C}$.

3.3.3 Associative Algebras

> **Definition 3.3.3 — Algebra** An **algebra** is a vector space, A, equipped with a bilinear product, $\cdot : A \times A \to A$. An **associative algebra** is an algebra for which the product is associative.

Note that Lie algebras are algebras in the above sense, but they are generally not associative. Examples of associative algebras include:

- $n \times n$ matrices with matrix multiplication.

- $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$ (the quaternions) with their usual multiplication are associative algebras over $\mathbb{R}$. In fact, these are *division* algebras: an algebra, D, is a division algebra if for any $a, b \in D$ with $b \neq 0$ there is exactly one $x \in D$ with $a = bx$ and exactly one $y \in D$ with $a = yb$. It turns out that these are the only finite-dimensional associative division algebras over $\mathbb{R}$. If we drop the associativity condition then we also get $\mathbb{O}$, the octonions.

Our interest in associative algebras is mostly the following.

> **Lemma 3.3.4** If A is an associative algebra then defining the bracket by $[X, Y] = XY - YX$ defines a Lie algebra.
>
> *Proof.* First note that we clearly have
>
> $$[X, X] = XX - XX = 0. \tag{3.3.5}$$
>
> The commutator is bilinear, here we show linearity in the second argument, taking $X, Y, Z \in A$ and $\lambda \in \mathbb{k}$
>
> $$[X, Y + \lambda Z] = X(Y + \lambda Z) - (Y + \lambda Z)X \tag{3.3.6}$$
> $$= XY + \lambda XZ - YX - \lambda ZX \tag{3.3.7}$$
> $$= XY - YX + \lambda(XZ - ZX) \tag{3.3.8}$$
> $$= [X, Y] + \lambda[X, Z]. \tag{3.3.9}$$
>
> The Jacobi identity follows by some algebra. First note that
>
> $$[X, [Y, Z]] = [X, YZ - ZY] \tag{3.3.10}$$
> $$= X(YZ - ZY) - (YZ - ZY)X \tag{3.3.11}$$
> $$= XYZ - XZY - YZX + ZYX. \tag{3.3.12}$$
>
> Then, using the fact that the other terms are simply linear combinations of these, we have that the Jacobi relation is reduced to
>
> $$XYZ - XZY - YZX + ZYX + YZX - YXZ - ZXY + XZY$$
> $$+ ZXY - ZYX - XYZ + YXZ \tag{3.3.13}$$
>
> and these terms all cancel to give zero. $\square$

When the Lie bracket can be written as $[X, Y] = XY - YX$ we call it the **commutator**, because it is a measure of the failure of A to be a commutative algebra. This terminology is also often used for the Lie bracket in general, but we'll try to avoid it in the general case.

3.3.4 Classical Lie Algebras

The complex classical Lie algebras are some classes of Lie algebras defined with similar definitions. These were first of interest because they correspond to particularly common Lie groups. First, some notation.

Notation 3.3.14 Let $\mathrm{Mat}(n, \Bbbk)$ denote the set of $n \times n$ matrices with entries in $\Bbbk$.

For $A \in \mathrm{Mat}(n, \Bbbk)$ denote the transpose by A^T.

For $A \in \mathrm{Mat}(n, \Bbbk)$ with $\Bbbk = \mathbb{R}, \mathbb{C}$ denote the complex conjugate by $\bar{A}$.

For $A \in \mathrm{Mat}(n, \Bbbk)$ with $\Bbbk = \mathbb{R}, \mathbb{C}$ denote the Hermitian conjugate by A^*, recall that $A^* = \bar{A}^\mathsf{T} = \overline{A^\mathsf{T}}$

> (!) Physicists will denote the complex conjugate by A^*, and the Hermitian conjugate by $A^\dagger$.

Definition 3.3.15 — General Linear Lie Algebra Let V be a vector space, and define

$$\mathfrak{gl}(V) = \mathrm{End}\, V := \{T : V \to V \mid T \text{ is linear}\} = \Bbbk\text{-Vect}(V, V). \quad (3.3.16)$$

This is an associative algebra under composition of linear maps, and as such $\mathfrak{gl}(V)$ is a Lie algebra under the commutator.

Typically we write $\mathfrak{gl}(V)$ when thinking of the Lie algebra, and $\mathrm{End}\, V$ when thinking of the associative algebra.

When V is finite dimensional, say $\dim V = n$, we can fix a basis and then each linear map corresponds to an $n \times n$ matrix with entries in $\Bbbk$. This leads us to make the following definition.

$$\mathfrak{gl}(n, \Bbbk) := \mathrm{Mat}(n, \Bbbk) \cong \mathfrak{gl}(V). \quad (3.3.17)$$

Note that if $\dim V = n$ then $V \cong \Bbbk^n$ and $\mathfrak{gl}(n, \Bbbk) = \mathfrak{gl}(\Bbbk^n)$.

The following are some of the subalgebras of $\mathfrak{gl}(n, \mathbb{C})$ viewed as a *real* vector space. This is important, not all of these are closed under multiplication by arbitrary complex numbers (in particular, if A is skew-Hermitian then iA is Hermitian).

- $\mathfrak{gl}(n, \mathbb{R}) = \mathrm{Mat}(n, \mathbb{R})$, the general linear Lie algebra.

- $\mathfrak{sl}(n, \mathbb{R}) := \{A \in \mathfrak{gl}(n, \mathbb{R}) \mid \mathrm{tr}\, A = 0\}$, the **special linear Lie algebra**;

- $\mathfrak{o}(n, \mathbb{R}) = \mathfrak{so}(n, \mathbb{R}) := \{A \in \mathfrak{gl}(n, \mathbb{R}) \mid A^\mathsf{T} + A = 0\}$, the **orthogonal Lie algebra** and **special orthogonal Lie algebra**. Note that these two are equal but given different names because they both come from important Lie groups which are not equal. Note that $A^\mathsf{T} + A = 0$ means A is antisymmetric.

- $\mathfrak{u}(n) := \{A \in \mathfrak{gl}(n, \mathbb{C}) \mid A^* + A = 0\}$, the **unitary Lie algebra**. Note that $A^* + A = 0$ means A is skew-Hermitian.

- $\mathfrak{su}(n) := \{A \in \mathfrak{gl}(n, \mathbb{C}) \mid \operatorname{tr} A = 0 \text{ and } A^* + A = 0\}$, the **special unitary Lie algebra**;

- $\mathfrak{sp}(n, \mathbb{R}) := \{A \in \mathfrak{gl}(2n, \mathbb{R}) \mid A^\top J_n + J_n A = 0\}$, the **symplectic Lie algebra**, where $J_n \in \operatorname{Mat}(2n, \mathbb{R})$ is the block matrix

$$\begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \tag{3.3.18}$$

where $I_n \in \operatorname{Mat}(n, \mathbb{R})$ is the $n \times n$ identity matrix. The matrix J_n (or rather, the symmetric bilinear form it represents) is called a **symplectic form**.

> (!) Physicists like operators to be Hermitian, for quantum mechanics, so they will define things with a few factors of i different so that $\mathfrak{u}(n)$ and $\mathfrak{su}(n)$ consist of Hermitian matrices, as opposed to *skew*-Hermitian matrices. This distinction unfortunately causes these factors of i to propagate through many formulae, so be careful when looking things up that the source uses the correct convention.

> (!) Notation differs for the symplectic Lie algebra (and group), some authors denote it $\mathfrak{sp}(2n)$, it's the D_n vs D_{2n} debate all over again.

The names here are all derived from the names of the associated Lie groups, $\mathrm{GL}(n, \mathbb{R})$ the general linear group, $\mathrm{SL}(n, \mathbb{R})$ the special linear group, $\mathrm{O}(n, \mathbb{R})$ the orthogonal group, $\mathrm{SO}(n, \mathbb{R})$ the special orthogonal group, $\mathrm{U}(n)$ the unitary group, $\mathrm{SU}(n)$ the special unitary group, and various different but related symplectic Lie groups.

These are defined as follows:

- $\mathrm{GL}(n, \mathbb{R}) := \{A \in \operatorname{Mat}(n, \mathbb{R}) \mid \det A \neq 0\}$;

- $\mathfrak{sl}(n, \mathbb{R}) := \{A \in \mathrm{GL}(n, \mathbb{R}) \mid \det A = 1\}$;

- $\mathrm{O}(n, \mathbb{R}) := \{A \in \mathrm{GL}(n, \mathbb{R}) \mid A^\top A = I_n\}$;

- $\mathrm{SO}(n, \mathbb{R}) := \{A \in \mathrm{GL}(n, \mathbb{R}) \mid A^\top A = I_n \text{ and } \det A = 1\}$.

- $\mathrm{U}(n) := \{A \in \mathrm{GL}(n, \mathbb{C}) \mid U^* U = I_n\}$;

- $\mathrm{SU}(n) := \{A \in \mathrm{GL}(n, \mathbb{C}) \mid U^* U = I_n \text{ and } \det U = 1\}$;

- $\mathrm{Sp}(n, \mathbb{R}) := \{A \in \mathrm{GL}(2n, \mathbb{R}) \mid A^\top J_n A = J_n\}$.

These definitions are all about preserving some structure on $\mathbb{R}^n$ (or $\mathbb{R}^{2n}$ for the symplectic group).

- The general linear group preserves the vector space structure, including dimension.

- The special linear group preserves volues.

- The orthogonal group preserves angles.

- The special orthogonal group preserves the inner product on $\mathbb{R}^n$.

- The (special) unitary group is a complex analogue of the (special) orthogonal group.

- The symplectic group preserves J_n.

Note that "special" means that we impose the condition $\det A = 1$. In the case of the Lie algebras this becomes the condition that $\operatorname{tr} A = 0$. This is the reason that $\mathfrak{o}(n, \mathbb{R}) = \mathfrak{so}(n, \mathbb{R})$, the requirement that A is antisymmetric means that the diagonal of A is zero, so it automatically has trace zero. In terms of Lie groups, the reason is that both $\mathrm{O}(n, \mathbb{R})$ and $\mathrm{SO}(n, \mathbb{R})$ are very similar, in particular, $\mathrm{SO}(n, \mathbb{R}) \cong \mathrm{O}(n, \mathbb{R})/(\mathbb{Z}/2\mathbb{Z})$ where $\mathbb{Z}/2\mathbb{Z}$ here is $\{\pm I_n\}$, so $\mathrm{SO}(n, \mathbb{R})$ preserves everything that $\mathrm{O}(n, \mathbb{R})$ preserves, and also preserves orientation. This similarity means that $\mathrm{O}(n, \mathbb{R})$ has two connected components, one preserving orientation, corresponding to the subgroup $\mathrm{SO}(n, \mathbb{R})$, and one reversing orientation. Then the Lie algebra corresponds only to the component of the Lie group connected to the identity, which is $\mathrm{SO}(n, \mathbb{R})$ in both cases, so the Lie algebras are the same.

In some of these cases, but not all, it is possible to extend the field to $\mathbb{C}$, giving

- $\mathfrak{gl}(n, \mathbb{C}) := \mathrm{Mat}(n, \mathbb{C})$;

- $\mathfrak{o}(n, \mathbb{C}) := \mathfrak{so}(n, \mathbb{C}) = \{A \in \mathfrak{gl}(n, \mathbb{C}) \mid A^\mathsf{T} + A\} = 0$.

- $\mathfrak{sp}(n, \mathbb{C}) := \{A \in \mathfrak{gl}(2n, \mathbb{C}) \mid A^\mathsf{T} J_n + J_n A = 0\}$.

Showing that the classical Lie algebras are indeed Lie algebras, specifically subalgebras of $\mathfrak{gl}(N, \Bbbk)$, requires a little bit of work. Mostly we have to show that they are closed under the bracket, which is the commutator in all cases. There are a few tricks though. The first is that $\mathfrak{sl}(n, \Bbbk) = \ker \operatorname{tr}$, when we view $\operatorname{tr} : \mathfrak{gl}(n, \Bbbk) \to \mathfrak{gl}(n, \Bbbk)$ as a linear operator, and the kernel of a linear operator is always a subspace. We also have to show that the commutator of two traceless matrices is again traceless for $\mathfrak{sl}(n, \Bbbk)$ to be closed under the bracket:

$$\operatorname{tr}([X, Y]) = \operatorname{tr}(XY - YX) = \operatorname{tr}(XY) - \operatorname{tr}(YX) = \operatorname{tr}(XY) - \operatorname{tr}(XY) = 0 \quad (3.3.19)$$

where we've used the fact that the trace of a product is invariant under cyclic permutations of that product (henceforth, "the trace is cyclic"). Notice that this doesn't actually use that X and Y are traceless, the trace of a commutator is always zero. This means that if $X \in \mathfrak{gl}(n, \Bbbk)$ and $Y \in \mathfrak{sl}(n, \Bbbk)$ we still have $\operatorname{tr}([X, Y]) = 0$, and thus $[X, Y] \in \mathfrak{sl}(n, \Bbbk)$, meaning that $\mathfrak{sl}(n, \Bbbk)$ is an ideal of $\mathfrak{gl}(n, \Bbbk)$

3.3.4.1 Dimensions

An important exercise is to compute the dimensions of the classical Lie algebras. We'll do it for the real ones.

- $\dim \mathfrak{gl}(n, \mathbb{R}) = n^2$, since an arbitrary element of $\mathfrak{gl}(n, \mathbb{R})$ is an $n \times n$ matrix, which is parametrised by n^2 entries. A basis for $\mathfrak{gl}(n, \mathbb{R})$ is

$$\{E_{ij} \mid i = 1, \dots, n \text{ and } j = 1, \dots, n\} \quad (3.3.20)$$

where E_{ij} has 1 in the ith row and jth column and 0 everywhere else. Note that

$$[E_{ij}, E_{kl}] = E_{ij}E_{kl} - E_{kl}E_{ij} = \delta_{ij}E_{il} - \delta_{il}E_{kj}. \quad (3.3.21)$$

- $\dim \mathfrak{sl}(n, \mathbb{R}) = n^2 - 1$, since setting $\operatorname{tr} A = 0$ fixes one element on the diagonal, say the last element, by the requirement that if $A = (a_{ij})$ then $a_{11} + \cdots + a_{n-1,n-1} = -a_{nn}$ to get $\operatorname{tr} A = a_{11} + \cdots + a_{n-1,n-1} + a_{nn} = 0$. Alternatively, note that $\mathfrak{sl}(n, \mathbb{R}) = \ker \operatorname{tr}$, and $\operatorname{im} \operatorname{tr} = \mathbb{R}$ and so by the rank-nullity theorem we have

$$\dim \mathfrak{gl}(n, \mathbb{R}) = \dim(\ker \operatorname{tr}) + \dim(\operatorname{im} \operatorname{tr}) = \dim \mathfrak{sl}(n, \mathbb{R}) + 1. \quad (3.3.22)$$

  The result then follows from this and our calculation of $\dim \mathfrak{gl}(n, \mathbb{R}) = n^2$.

- $\dim \mathfrak{so}(n, \mathbb{R}) = n(n-1)/2$, requiring that A be antisymmetric means that A is zero on the diagonal, and below the diagonal is fixed by above the diagonal. The above diagonal elements form a triangle with a base of $n - 1$ elements, and thus the number of entries above the diagonal is the $(n - 1)$st triangle number, $T_{n-1} = n(n - 1)/2$ ($T_n = n(n + 1)/2$). We can check this by identifying that SO(3) corresponds to rotations in three dimensions, and this is a three dimensional group, since any rotation is specified by either a) three Euler angles, or b) an angle and an axis of rotation (requiring two numbers to fix the direction, the third component fixed by requiring it to be a unit vector). Note that $\dim \mathfrak{so}(3) = 3$ is a coincidence, we have $\dim \mathfrak{so}(2) = 1$ (rotations in two dimensions are specified by just an angle) and $\dim \mathfrak{so}(4) = 6$. Of course, here we're using the fact that a Lie group and its associated Lie algebra have the same dimension, this is because the tangent space has the same dimension as the manifold (assuming the manifold is connected, which is always the case for at least one Lie group corresponding to a given Lie algebra).

- $\dim \mathfrak{u}(n) = n^2$, note that we are talking about the dimension as a real vector space, despite the entries being complex numbers. This is the dimension because an arbitrary $n \times n$ matrix with entries in $\mathbb{C}$ has $2n^2$ real parameters (one for the real part and one for the imaginary part of each entry). Requiring that the matrix is skew-Hermitian means that each entry on the diagonal must be equal to its conjugate, so it must have zero imaginary part (fixing n parameters), then the entries below the diagonal are fixed by the entries above the diagonal (fixing $2n(n - 1)/2 = n^2 - n$ parameters). Thus, the dimension is $2n^2 - n - (n^2 - n) = n^2$.

- $\dim \mathfrak{su}(n) = n^2 - 1$, since any matrix in $\mathfrak{su}(n)$ is in $\mathfrak{u}(n)$, and setting its trace to zero fixes the final entry on the diagonal (which is purely real).

- $\dim \mathfrak{sp}(n) = 2n^2 + n$, a general $2n \times 2n$ matrix with entries in $\mathbb{R}$ has $(2n)^2 = 4n^2$ parameters. Write this as a block matrix,

$$A = \begin{bmatrix} X & Y \\ Z & W \end{bmatrix}, \qquad A^\mathsf{T} = \begin{bmatrix} X^\mathsf{T} & Z^\mathsf{T} \\ Y^\mathsf{T} & W^\mathsf{T} \end{bmatrix} \qquad (3.3.23)$$

  with $X, Y, Z, W \in \operatorname{Mat}(n, \mathbb{R})$. Then computing the defining relationship, $A^\mathsf{T} J_n + J_n A = 0$, we have

$$\begin{bmatrix} Z^\mathsf{T} - Z & -W - X^\mathsf{T} \\ W^\mathsf{T} + X & Y - Y^\mathsf{T} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \qquad (3.3.24)$$

  We see that this fixes $Z = Z^\mathsf{T}$, $Y = Y^\mathsf{T}$, and $W = -X^\mathsf{T}$. Requiring that Z be symmetric means that the elements below the diagonal are fixed, removing $n(n - 1)/2$ free parameters. Similarly, requiring that Y is symmetric

removes $n(n-1)/2$ parameters. Since X determines W we also fix another n^2 parameters. Thus, the number of remaining free parameters is

$$4n^2 - \left(\frac{1}{2}n(n-1) + \frac{1}{2}n(n-1) + n^2\right) = 2n^2 + n. \qquad (3.3.25)$$

3.3.4.2 $\mathfrak{sl}(2, \mathbb{C})$

As we've said a couple of times already the most important Lie algebra is $\mathfrak{sl}(2, \mathbb{C})$, which is a Lie algebra over $\mathbb{C}$. A general element of this algebra is a 2×2 matrix with complex entries such that the trace vanishes, we can write such an element as

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix} \qquad (3.3.26)$$

for $a, b, c \in \mathbb{C}$. This shows that $\dim \mathfrak{sl}(2, \mathbb{C}) = 3$.

An explicit basis for $\mathfrak{sl}(2, \mathbb{C})$ is

$$H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad E = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \text{and} \qquad F = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}. \qquad (3.3.27)$$

Then

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix} = aH + bE + cF. \qquad (3.3.28)$$

A simple calculation of commutators shows that we have the relations

$$[H, E] = 2E, \qquad [H, F] = -2F, \qquad \text{and} \qquad [E, F] = H. \qquad (3.3.29)$$

Later we will work more abstractly, rather than defining $\mathfrak{sl}(2, \mathbb{C})$ to be 2×2 complex traceless matrices we'll define $\mathfrak{sl}(2, \mathbb{C})$ to be a three dimensional Lie algebra with basis $\{H, E, F\}$ satisfying the above bracket relations. Note that this doesn't require the bracket to be the commutator, in fact, the commutator isn't even defined in this abstract setting since we don't have a notion of multiplication[3] with which to form, say, HE. Once we do this we see that the matrices above form a two-dimensional representation (Definition 4.1.1) of $\mathfrak{sl}(2, \mathbb{C})$, called the defining representation, which has an obvious action on $\mathbb{C}^2$.

An important observation here is that if we define the map[4] $\mathrm{ad}_H : \mathfrak{sl}(2, \mathbb{C}) \to \mathfrak{sl}(2, \mathbb{C})$ by $\mathrm{ad}_H(X) = [H, X]$ then we see that this map has E and F as eigenvectors, with eigenvalues 2 and -2 respectively. The other eigenvector of ad_H is H itself, since $\mathrm{ad}_H(H) = [H, H] = 0 = 0H$. Then we have the eigenspaces $V_0 = \mathbb{C}H = \mathfrak{h}$, $V_2 = \mathbb{C}E$, and $V_{-2} = \mathbb{C}F$, giving the decomposition

$$\mathfrak{sl}(2, \mathbb{C}) = \mathbb{C}H \oplus \mathbb{C}E \oplus \mathbb{C}F = \mathfrak{h} \oplus V_2 \oplus V_{-2} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \mathbb{C}\backslash\{0\}} V_\alpha. \qquad (3.3.30)$$

The important thing here is that we get $\mathfrak{h}$, which is a special subalgebra called the Cartan subalgebra which we'll see a lot later, and we get the generalised eigenspaces V_α. We'll see later[5] that replacing the generalised eigenspaces with something called root spaces makes this construction generalise to finite dimensional semisimple (Definition 5.3.20) Lie algebras over $\mathbb{C}$, and further that the subalgebras appearing in this decomposition are all simple Definition 5.2.1, so a classification of simple Lie algebras extends to a classification of semisimple Lie algebras.

---

[3]It turns out that every Lie algebra, $\mathfrak{g}$, can be embedded into an associative algebra, $\mathcal{U}(\mathfrak{g})$ in which the Lie bracket on $\mathfrak{g}$ coincides with the commutator in $\mathcal{U}(\mathfrak{g})$. This larger algebra is called the **universal enveloping algebra**. Formally, this algebra can be constructed by taking the tensor algebra, $T(\mathfrak{g}) = \mathbb{k} \oplus \mathfrak{g} \oplus (\mathfrak{g} \otimes \mathfrak{g}) \oplus (\mathfrak{g} \otimes \mathfrak{g} \otimes \mathfrak{g}) \oplus \cdots$, and defining the quotient $\mathcal{U}(\mathfrak{g}) = T(\mathfrak{g})/I$ where I is the two sided ideal generated by elements of the form $a \otimes b - b \otimes a - [a, b] \in \mathfrak{g} \oplus (\mathfrak{g} \otimes \mathfrak{g}) \subseteq T(\mathfrak{g})$. A general element of this ideal is thus of the form $c \otimes d \otimes \cdots \otimes (a \otimes b - b \otimes a - [a, b]) \otimes e \otimes f \otimes \cdots$ with $a, b, c, d, e, f \in \mathfrak{g}$. Note that the product of two elements of $T(\mathfrak{g})$ is their tensor product, although we usually write $a \otimes b = ab$.

[4]This is the so called adjoint representation (Definition 4.1.3).

[5]so ignore any terminology here that is unfamiliar

This requires replacing the indexing set $\mathbb{C} \setminus \{0\}$ with the root space $\mathfrak{h}^* \setminus \{0\}$, or rather we've already done that, it's just that $\dim \mathfrak{h} = 1$ so $\mathfrak{h}^* \cong \mathfrak{h}$ and $\dim \mathfrak{h}^* = 1$ so $\mathfrak{h}^* \cong \mathbb{C}$.

While this last paragraph is beyond what we're ready for yet it's important to start seeing the recurring pattern of this type of decomposition now. The fact that $\mathfrak{sl}(2, \mathbb{C})$ is the simplest example of such a decomposition is why it's so important. In particular, every finite dimension semisimple Lie algebra over $\mathbb{C}$ of dimension at least 3 contains (many) copies of $\mathfrak{sl}(2, \mathbb{C})$.

3.3.5   Other Subalgebras of $\mathfrak{gl}(n, \mathbb{k})$

There are two further subalgebras of $\mathfrak{gl}(n, \mathbb{k})$ which are worth mentioning.

> **Definition 3.3.31 — Triangular Matrices** Denote by $\mathfrak{b}(n, \mathbb{k}) \subseteq \mathfrak{gl}(n, \mathbb{k})$ the subalgebra of upper triangular matrices, and by $\mathfrak{n}(n, \mathbb{k}) \subseteq \mathfrak{gl}(n, \mathbb{k})$ the subalgebra of *strictly* upper triangular matrices (meaning the diagonal is all zeros).

> **Remark 3.3.32** Here $\mathfrak{b}$ stands for "Borel" because $\mathfrak{b}(n, \mathbb{k})$ is a **Borel subalgebra**, meaning it's a maximal solvable subalgebra (Definition 5.3.6). The $\mathfrak{n}$ stands for "nilpotent" because $\mathfrak{n}(n, \mathbb{k})$ is a (maximal) nilpotent algebra (Definition 5.3.1)
> These properties ultimately come from the fact that taking nested commutators of (strictly) upper triangular matrices will eventually result in zero. The difference in how the commutators are nested is the distinction between solvable (e.g., $[[-, -], [-, -]]$) and nilpotent (e.g., $[-, [-, [-, -]]]$).

It turns out that $\mathfrak{n}(n, \mathbb{k})$ is an ideal of $\mathfrak{b}(n, \mathbb{k})$, since the product of an upper triangular matrix and a strictly upper triangular matrix is strictly upper triangular, and so the commutator is too. This means we can consider the quotient $\mathfrak{g} = \mathfrak{b}(n, \mathbb{k})/\mathfrak{n}(n, \mathbb{k})$. Two upper triangular matrices in $\mathfrak{b}(n, \mathbb{k})$ are identified in $\mathfrak{g}$ if their difference is a strictly upper triangular matrix in $\mathfrak{n}(n, \mathbb{k})$. This means that they must have the same diagonal. Suppose then that $X, Y \in \mathfrak{b}(n, \mathbb{k})$ have the same diagonal, so $X_{ii} = Y_{ii}$ for all $i = 1, \dots, n$. By definition of upper triangular we also know that $X_{ij} = Y_{ij} = 0$ if $i > j$. Basic matrix multiplication tells us an element of the diagonal of XY is

$$(XY)_{ii} = \sum_{j=1}^{n} X_{ij} Y_{ji}. \tag{3.3.33}$$

If $i > j$ then $X_{ij} = 0$, and if $i < j$ then $Y_{ji} = 0$, so the only nonzero diagonal term comes from $X_{ii} Y_{ii}$. The same analysis can be applied to YX and we see that the only nonzero diagonal term is $(YX)_{ii} = Y_{ii} X_{ii}$. In the commutator these terms cancel out, and as such $[X, Y] \in \mathfrak{n}(n, \mathbb{k})$ for all $X, Y \in \mathfrak{b}(n, \mathbb{k})$. This means that in the quotient $\mathfrak{g}$ we identify $[X, Y]$ with $[X, Y] + \mathfrak{n}(n, \mathbb{k}) = 0 + \mathfrak{n}(n, \mathbb{k})$, which is zero, and so this quotient is abelian.

Note that as $\mathbb{k}$-vector spaces $\dim \mathfrak{b}(n, \mathbb{k}) = n(n + 1)/2 = T_n$ and $\dim \mathfrak{b}(n, \mathbb{k}) = n(n - 1)/2 = T_{n-1}$. Thus, the quotient $\mathfrak{g}$ has dimension $T_n - T_{n-1} = n$.

If A is an algebra with then another important Lie subalgebra of $\mathfrak{gl}(A)$ is Der A, the algebra of derivations of A. Recall that a derivation is a linear map $D : A \to A$ such that

$$D(ab) = aD(b) + D(a)b \tag{3.3.34}$$

for all $a, b \in A$. To show that this is a subalgebra of $\mathfrak{gl}(A)$ we need to show that Der A is closed under the commutator, that is, that the commutator of two derivations is a derivation. This follows from the following calculation with $D_1, D_2 \in$ Der A and $a, b \in A$:

$$
\begin{aligned}
[D_1, D_2](ab) &= (D_1 D_2 - D_2 D_1)(ab) \tag{3.3.35}\\
&= D_1(D_2(ab)) - D_2(D_1(ab))\\
&= D_1(aD_2(b) + D_2(a)b) - D_2(aD_1(b) - D_1(a)b)\\
&= D_1(aD_2(b)) + D_1(D_2(a)b) - D_2(aD_1(b)) - D_2(D_1(a)b)\\
&= aD_1(D_2(b)) + D_1(a)D_2(b) + D_2(a)D_1(b) + D_1(D_2(a))b\\
&\quad - aD_2(D_1(b)) - D_2(a)D_1(b) - D_1(a)D_2(b) - D_2(D_1(a))b\\
&= aD_1(D_2(b)) + D_1(D_2(a))b - aD_2(D_1(b)) - D_2(D_1(a))b\\
&= a\{D_1(D_2(b)) - D_2(D_1(b))\} + \{D_1(D_2(a)) - D_2(D_1(a))\}b\\
&= a(D_1 D_2 - D_2 D_1)(b) + (D_1 D_2 - D_2 D_1)(a)b\\
&= a[D_1, D_2](b) + [D_1, D_2](a)b.
\end{aligned}
$$

This shows that $[D_1, D_2]$ is a derivation, so Der A is closed under the commutator.

Four

Representation Theory

4.1 Definition

> **Definition 4.1.1 — Representation** Let $\mathfrak{g}$ be a Lie algebra over a field, $\Bbbk$. A **representation** of $\mathfrak{g}$ is a pair, (V, ρ), where V is a vector space over $\Bbbk$ and $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is a Lie algebra homomorphism.

The idea behind this definition is that $\mathfrak{g}$ is acting on V by specifying an endomorphism $\varphi_X = \rho(X) \in \operatorname{End} V = \mathfrak{gl}(V)$ for each $X \in \mathfrak{g}$. Then X acts on $v \in V$ by $x \mapsto \varphi_X(V)$.

This should be compared to the idea of a group action, which can similarly be specified by a map $\rho : G \to \operatorname{Perm}(X)$ where $\operatorname{Perm}(X)$ is the group of permutations of elements of X. If you're familiar with group representations then the idea here is the same as a group representation (V, ρ) with $\rho : G \to \operatorname{GL}(V)$ a group homomorphism.

> **Example 4.1.2 — Trivial Representation** Let $\mathfrak{g}$ be a Lie algebra over $\Bbbk$ and V any vector space over $\Bbbk$. Then the homomorphism $\rho_{\text{triv}} : \mathfrak{g} \to \mathfrak{gl}(V)$ defined by $\rho_{\text{triv}}(X) = 0$ is the **trivial representation** of $\mathfrak{g}$ on V.

[1]We say "the" trivial representation, because V really isn't important in this case. When V is not specified the trivial representation is typically taken to be carried by a one-dimensional vector space.

Note that the[1] trivial Lie algebra representation acts as zero, whereas a trivial group representation acts as the identity. This is consistent since $e^0 = 1$, so we recover the trivial action of the corresponding Lie group from the trivial action of the Lie algebra.

One of the most important examples of a representation is when the Lie algebra acts on itself in the obvious way, through the bracket.

> **Definition 4.1.3 — Adjoint Representation** Let $\mathfrak{g}$ be a Lie algebra. There is a Lie algebra representation called the **adjoint representation**, $(\mathfrak{g}, \operatorname{ad})$ where $\operatorname{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is defined by $\operatorname{ad}(X) = \operatorname{ad}_X$ where $\operatorname{ad}_X \in \mathfrak{gl}(\mathfrak{g})$ is the linear map $\operatorname{ad}_X : \mathfrak{g} \to \mathfrak{g}$ defined by $\operatorname{ad}_X(Y) = [X, Y]$.

This is related to the adjoint action of a the corresponding Lie group on itself, which is by conjugation.

Definition 4.1.4 — Modules Let $\mathfrak{g}$ be a Lie algebra over a field, $\Bbbk$. A (left) $\mathfrak{g}$-**module**, V, is a vector space over $\Bbbk$ equipped with a bilinear map $- \cdot - : \mathfrak{g} \times V$ such that for all $X, Y \in \mathfrak{g}$ and $v \in V$ we have

$$[X, Y] \cdot v = X \cdot (Y \cdot v) - Y \cdot (X \cdot v). \tag{4.1.5}$$

If you're familiar with the idea of an R-module for R a ring then the idea is the same here. An R-module is an abelian group equipped with an R-action, and a $\mathfrak{g}$-module is a vector space equipped with a $\mathfrak{g}$-action.

Proposition 4.1.6 A Lie algebra representation of $\mathfrak{g}$ carries the same information as a $\mathfrak{g}$-module.

Proof. Let (V, ρ) be a representation of $\mathfrak{g}$. Then we may define a $\mathfrak{g}$-action on V by $X \cdot v := \rho(X)(v)$. This is linear in the first argument, X, because ρ is linear, and linear in the second argument, v, because $\rho(X) \in \mathfrak{gl}(V)$ is linear. Thus, this is a bilinear map. We further have

$$[X, Y] \cdot v = \rho([X, Y])(v) \tag{4.1.7}$$
$$= [\rho(X), \rho(Y)](v) \tag{4.1.8}$$
$$= \rho(X)(\rho(Y)(v)) - \rho(Y)(\rho(X)(v)) \tag{4.1.9}$$
$$= X \cdot (Y \cdot v) - Y \cdot (X \cdot v) \tag{4.1.10}$$

where we've used the fact that ρ is a homomorphism and in $\mathfrak{gl}(V)$ the Lie bracket is just the commutator.

Conversely, if V is a $\mathfrak{g}$-module then we may define a homomorphism $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ by $X \mapsto X \cdot -$ which acts by $v \mapsto X \cdot v$. This is a linear map $V \to V$ because the $\mathfrak{g}$-action is bilinear in the second argument, and it's a Lie algebra homomorphism because by definition the Lie bracket acts as a commutator, which is the Lie bracket in $\mathfrak{gl}(V)$.

Thus, every representation of $\mathfrak{g}$ uniquely specifies a $\mathfrak{g}$-module and vice versa. $\qquad \square$

The "modern" treatment of representations is mostly via the notion of modules, and we shall prefer this method.

4.1.1 Morphism of Representations

Definition 4.1.11 Let V_1 and V_2 be $\mathfrak{g}$-modules for some Lie algebra $\mathfrak{g}$. A **morphism** of $\mathfrak{g}$-modules is then a linear map $\varphi : V_1 \to V_2$ such that for all $X \in \mathfrak{g}$ and $v \in V_1$ we have

$$\varphi(X \cdot v) = X \cdot v. \tag{4.1.12}$$

We also call φ an **equivariant map**.
An invertible $\mathfrak{g}$-module homomorphism is a $\mathfrak{g}$-module isomorhpism.

Writing $X \cdot -$ for the map $v \mapsto X \cdot v$ the definition above is equivalent to the commutativity of the following square for all $X \in \mathfrak{g}$:

$$
\begin{array}{ccc}
V_1 & \xrightarrow{\;\varphi\;} & V_2 \\
{\scriptstyle X\cdot-}\Big\downarrow & & \Big\downarrow{\scriptstyle X\cdot-} \\
V_1 & \xrightarrow[\;\varphi\;]{} & V_2.
\end{array}
\tag{4.1.13}
$$

Note that $X \cdot -$ on the left is the action of X on V_1, whereas on the right it's the action of X on V_2. We get away with the same notation by looking at the domain of the action to specify which we mean.

Now that we have a notion of morphisms between representations we can form a category. Specifically, $\mathfrak{g}$-Rep is the category whose objects are $\mathfrak{g}$-modules (or equivalently representations of $\mathfrak{g}$) and whose morhpisms are morphisms of representations. Note that formally the category with representations and the category with $\mathfrak{g}$-modules are different, but they are isomorphic, which is why we don't need to distinguish between them. The isomorphism is given by the functor sending (V, ρ) to the $\mathfrak{g}$-module V with the action given by $X \cdot v = \rho(X)(v)$.

Lemma 4.1.14 If $\varphi \colon V_1 \to V_2$ is an isomorphism of $\mathfrak{g}$-modules then $\varphi^{-1} \colon V_1 \to V_2$ is also.

Proof. We need to show that for all $X \in \mathfrak{g}$ and $v \in V_2$ we have

$$
\varphi^{-1}(X \cdot v) = X \cdot \varphi^{-1}(v).
\tag{4.1.15}
$$

To do so note that φ^{-1} is an invertible linear map, and as such is a bijection. This means that there exists unique $v' \in V_1$ such that $\varphi(v') = v$ and $v' = \varphi^{-1}(v)$. Thus, we may write the left hand side as

$$
\varphi^{-1}(X \cdot v) = \varphi^{-1}(X \cdot \varphi(v')).
\tag{4.1.16}
$$

Using the fact that φ is a morphism of $\mathfrak{g}$-modules we have

$$
\varphi^{-1}(X \cdot v) = \varphi^{-1}(\varphi(X \cdot v')) = X \cdot v' = X \cdot \varphi^{-1}(v).
\tag{4.1.17}
$$

Notation 4.1.18 If M and N are $\mathfrak{g}$-modules over $\Bbbk$ we write

$$
\operatorname{Hom}_{\Bbbk}(M, N) = \operatorname{Hom}(M, N) = \{T \colon M \to N \mid T \text{ is linear}\}
\tag{4.1.19}
$$
$$
= \Bbbk\text{-Vect}(M, N)
\tag{4.1.20}
$$

and

$$
\operatorname{Hom}_{\mathfrak{g}}(M, N) = \{T \colon M \to N \mid T \text{ is a } \mathfrak{g}\text{-module homomorphism}\}
$$
$$
= \mathfrak{g}\text{-Rep}(M, N).
\tag{4.1.21}
$$

Similarly, we write $\operatorname{End}_{\Bbbk}(M) = \operatorname{End}(M) = \operatorname{Hom}(M, M)$ and $\operatorname{End}_{\mathfrak{g}}(M) = \operatorname{Hom}_{\mathfrak{g}}(M, M)$.

4.2 Simple Modules

Definition 4.2.1 — Submodule Let $\mathfrak{g}$ be a Lie algebra over $\Bbbk$ and let V be a $\mathfrak{g}$-module. A subspace, $W \subseteq V$, is a **submodule** if it is a $\mathfrak{g}$-module under the restricted action of $\mathfrak{g}$. Equivalently, W is a submodule if $X \cdot w \in W$ for all $X \in \mathfrak{g}$ and for all $w \in W$.

A **proper submodule** of V is a submodule that is neither 0 nor V.

Lemma 4.2.2 Let $\varphi \colon M \to N$ be a $\mathfrak{g}$-module homomorphism. Then $\ker \varphi$ is a submodule of M and $\operatorname{im} \varphi$ is a submodule of N.

Proof. We already know that $\ker \varphi$ and $\operatorname{im} \varphi$ are subspaces of M and N respectively, so it remains only to show that they are submodules, that they are closed under the action of $\mathfrak{g}$. Let $v \in \ker \varphi$, then $\varphi(v) = 0$. From this we know that

$$\varphi(X \cdot v) = X \cdot \varphi(v) = X \cdot 0 = 0 \tag{4.2.3}$$

and so $X \cdot v \in \ker \varphi$. Now, let $w \in \operatorname{im} \varphi$, then there exists some $w' \in M$ with $\varphi(w') = w$. We then have

$$X \cdot w = X \cdot \varphi(w') = \varphi(X \cdot w') \tag{4.2.4}$$

which means $X \cdot w \in \operatorname{im} \varphi$. $\qquad\square$

Definition 4.2.5 — Simple Module A **simple $\mathfrak{g}$-module** is a $\mathfrak{g}$-module with no proper submodules. The corresponding representation is called an **irreducible representation**, often shortened to **irrep**.

The simple $\mathfrak{g}$-modules turn out to be the most important. Often, but not always (but always over $\mathbb{C}$) it is possible to write any module, V, as a direct sum of simple $\mathfrak{g}$-modules. Because of this one of the first goals when studying representations of a given Lie algebra is to classify all simple $\mathfrak{g}$-modules. This is often possible, but not always easy.

Lemma 4.2.6 — Schur's Lemma Let V be a finite-dimensional $\mathfrak{g}$-module over an algebraically closed field, $\Bbbk$. Then eery $\mathfrak{g}$-module endomorphism of V is a multiple of the identity and $\operatorname{End}_{\mathfrak{g}}(V) \cong \Bbbk$.

Proof. Let $\varphi \colon V \to V$ be a $\mathfrak{g}$-module endomorphism. Then φ is a linear map. Since $\Bbbk$ is assumed to be algebraically closed we know that this linear map has an eigenvalue, call it λ and let $v \in V$ be the associated eigenvector. We know that $v \in \ker \varphi'$ by construction. The kernel is a submodule, and since it's not zero (it contains v) and V is simple the kernel must be all of V. Thus, $\varphi'(w) = 0$ for all $w \in V$, and so $\varphi - \lambda\text{id}$ is the zero map, and so $\varphi = \lambda\text{id}$ is just a scalar multiple of the identity. $\qquad\square$

Note that from this we can immediately conclude that the only $\mathfrak{g}$-module endomorphisms are invertible (if $\lambda \neq 0$) or the zero map ($\lambda = 0$).

Lemma 4.2.7 If V and W are non-isomorphic simple $\mathfrak{g}$-modules then $\mathrm{Hom}_{\mathfrak{g}}(V, W) = 0$.

Proof. Suppose that $\varphi \in \mathrm{Hom}_{\mathfrak{g}}(V, W)$. That is, $\varphi : V \to W$ is a $\mathfrak{g}$-module homomorphism. Then we may consider the submodule $\ker \varphi \subseteq V$. Supposing that $\varphi \neq 0$ there must be some $v \in V$ such that $\varphi(v) \neq 0$, thus $\ker \varphi$ is not the whole of V, and as a submodule of the simple module V it must be that $\ker \varphi = 0$, so φ is injective. Now consider the submodule $\mathrm{im}\, \varphi \subseteq W$. We know that $\mathrm{im}\, \varphi \neq 0$ since $0 \neq \varphi(v) \in \mathrm{im}\, \varphi$. Thus, since W is simple, $\mathrm{im}\, \varphi$ must be all of W, so φ is surjective. This leaves us with two cases, either φ is a injective and surjective, and hence an isomorphism, or φ is the zero map. By assumption there are no isomorphisms $V \to W$, and further $\varphi \neq 0$. Thus, there are no non-trivial morphisms $V \to W$ as claimed. $\square$

4.2.1 New Modules From Old

Lemma 4.2.8 Let $\mathfrak{m}$ and $\mathfrak{l}$ be Lie algebras, and let V and W be simple $\mathfrak{m}$- and $\mathfrak{l}$-modules, respectively. Then $V \otimes W$ is made into an $(\mathfrak{m} \oplus \mathfrak{l})$-module by extending the action

$$(X, Y) \cdot (v \otimes w) = (X \cdot v) \otimes w + v \otimes (Y \cdot w) \tag{4.2.9}$$

by linearity to non-simple tensors. Note that $X \in \mathfrak{m}$, $Y \in \mathfrak{l}$, $v \in V$ and $w \in W$.

Proof. We first show that this action is bilinear. Linearity in the first argument follows from the calculation

$$
\begin{align}
(\lambda(X, Y)) \cdot (v \otimes w) &= (\lambda X, \lambda Y) \cdot (v \otimes w) \tag{4.2.10} \\
&= (\lambda X \cdot w) \otimes v + w \otimes (\lambda Y \cdot w) \tag{4.2.11} \\
&= \lambda((X \cdot w) \otimes v) + \lambda(w \otimes (Y \cdot w)) \tag{4.2.12} \\
&= \lambda((X \cdot w) \otimes v + w \otimes (Y \cdot w)) \tag{4.2.13} \\
&= \lambda((X, Y) \cdot (v \otimes w)) \tag{4.2.14}
\end{align}
$$

and

$$
\begin{align}
((X, Y) + (X', Y')) \cdot (v \otimes w) &= (X + X', Y + X') \cdot (v \otimes w) \tag{4.2.15} \\
= ((X + X') \cdot v) \otimes w &+ v \otimes ((Y + Y') \cdot w) \tag{4.2.16} \\
= (X \cdot v + X' \cdot v) \otimes w &+ v \otimes (Y \cdot w + Y' \cdot w) \tag{4.2.17} \\
= (X \cdot v) \otimes w + (X' \cdot v) \otimes w &+ v \otimes (Y \cdot w) + v \otimes (Y' \cdot w) \tag{4.2.18} \\
= (X \cdot v) \otimes w + v \otimes (Y \cdot w) &+ (X' \cdot v) \otimes w + v \otimes (Y' \cdot w) \tag{4.2.19} \\
= (X, Y) \cdot (v \otimes w) &+ (X', Y') \cdot (v \otimes w). \tag{4.2.20}
\end{align}
$$

Linearity in the second argument may be demonstrated by a similar argument. We can also demonstrate that brackets act as they must:

$$[(X,Y),(X',Y')] \cdot (v \otimes w) = ([X,X'],[Y,Y']) \cdot (v \otimes w) \qquad (4.2.21)$$
$$= ([X,X'] \cdot v) \otimes w + v \otimes ([Y,Y'] \cdot w) \qquad (4.2.22)$$
$$= (X \cdot (X' \cdot v) - X' \cdot (X \cdot v)) \otimes w + v \otimes (Y \cdot (Y' \cdot w) + Y' \cdot (Y \cdot w))$$
$$= (X \cdot (X' \cdot v)) \otimes w - (X' \cdot (X \cdot v)) \otimes w \qquad (4.2.23)$$
$$\qquad + v \otimes (Y \cdot (Y' \cdot w)) - v \otimes (Y' \cdot (Y \cdot w))$$
$$= (X \cdot (X' \cdot v)) \otimes w + v \otimes (Y \cdot (Y' \cdot w)) \qquad (4.2.24)$$
$$\qquad - (X' \cdot (X \cdot v)) \otimes w - v \otimes (Y' \cdot (Y \cdot w))$$
$$= (X,Y) \cdot ((X',Y') \cdot (v \otimes w)) - (X',Y') \cdot ((X,Y) \cdot (v \otimes w)).$$

Thus, this is indeed an $(\mathfrak{m} \oplus \mathfrak{l})$-module. □

The goal of the next couple of lemmas is to prove Proposition 4.2.28, that the tensor product of simple modules is again simple.

Lemma 4.2.25 Let V be a simple $\mathfrak{m}$-module and $v_1, v_2 \in V$ such that v_1 is not proportional to v_2. Then the smallest submodule of $V \oplus V$ containing (v_1, v_2) is $V \oplus V$.

Proof. Let $U \subseteq V \oplus V$ be the minimal submodule (with respect to inclusion) containing (v_1, v_2). Consider the inclusion map $i \colon U \hookrightarrow V \oplus V$ and the projection maps $p_1, p_2 \colon V \oplus V \to V$. These are all module homomorphisms, and so the composites $p_1 \circ i, p_2 \circ i \colon U \to V$ are also homomorphisms. Since V is simple and U is nonzero and $p_1 \circ i$ is a nonzero map this map must be surjective. The kernel of $p_1 \circ i$ is contained in the kernel of p_1, which is $0 \oplus V$. Clearly $\{0\} \oplus \cong V$ is simple, and as such the kernel of $p_1 \circ i$ is either 0, in which case $U \cong V$, or it's $0 \oplus V$, in which case $U = V \oplus V$. Suppose that the kernel of $p_1 \circ i$ is zero, then $p_1 \circ i \colon U \to V$ is an isomorphism, and we can denote its inverse by $\varphi \colon V \to U$. The map $p_1 \circ i \colon U \to V$ will also be an isomorphism in this case, and thus we can consider the composite $p_2 \circ i \circ \varphi$, which is an isomorphism $V \to V$. Since V is simple we know by Schur's lemma (Lemma 4.2.6) that $p_2 \circ i \circ \varphi$ is a multiple of the identity, so $p_2 \circ i \circ \varphi(v_1) = \lambda v_1$ for some $\lambda \in \Bbbk$. However, we have $p_2 \circ i \circ \varphi(v_1) = p_2 \circ i(v_1, v_2) = p_2(v_1, v_2) = v_2$, contradicting the assumption that v_2 is not proportional to v_1. Thus, we are left with the case where $U = V \oplus V$. □

The following lemma is then a special case of the result of Proposition 4.2.28.

Lemma 4.2.26 Let V be a simple $\mathfrak{m}$-module and W a simple $\mathfrak{l}$-module. The smallest $(\mathfrak{m} \oplus \mathfrak{l})$-submodule of $V \otimes W$ containing a pure tensor $v \otimes w \neq 0$ is $V \otimes W$.

Proof. The module $V \otimes W$ is spanned by vectors of the form $v' \otimes w'$ for $v' \in V$ and $w' \in W$. Therefore, it suffices to show that $v' \otimes w'$ is contained in the smallest submodule, U, of $V \otimes W$ containing $v \otimes w$, and then necessarily U will contain all elements of $V \otimes W$. Linearity of the $\mathfrak{g}$-action tells us that $0 \cdot v = 0$ for all $v \in V$, and thus we have

$$(0, Y) \cdot (v \otimes w) = (0 \cdot v) \otimes w + v \otimes (Y \cdot w) = 0 \otimes w + v \otimes (Y \cdot w) = v \otimes (Y \cdot w). \tag{4.2.27}$$

This is true for all $Y \in \mathfrak{l} \subseteq \mathfrak{m} \oplus \mathfrak{l}$. Now, we can generate a submodule in this way starting with $v \otimes w$ and acting with $Y \in \mathfrak{l}$ as above. Doing so we generate a submodule of W which is not zero, and thus must be isomorphic to all of W. This submodule is thus $\Bbbk v \otimes W$ where $\Bbbk v = \text{span}\{v\}$. Thus, $v \otimes w' \in U$ for any $w' \in W$. Similarly, one can show $v' \otimes w \in U$ for any $v' \in V$. Thus, $v' \otimes w' \in U$. $\square$

We are now ready to prove our result.

Proposition 4.2.28 Let V be a simple $\mathfrak{m}$-module and W a simple $\mathfrak{l}$-module. Then $V \otimes W$ is a simple $(\mathfrak{m} \oplus \mathfrak{l})$-module.

Proof. Let $u \in V \otimes W$ be such that $u \neq 0$. Let U be the smallest $(\mathfrak{m} \oplus \mathfrak{l})$-submodule of $V \otimes W$ containing u. We want to show that $U = V \otimes W$. If we know that $0 \neq v \otimes w \in U$ then the result follows from Lemma 4.2.26. As an element of $V \otimes W$ we can express u as a sum of simple tensors,

$$u = \sum_{i=1}^{k} v_i \otimes w_i. \tag{4.2.29}$$

Note that we don't need scalar factors since we can just absorb them into the definition of the v_i. We may further assume that the v_i are linearly independent, and that no pair of w_is is proportional, since if $w_i = \alpha w_j$ then $v_i \otimes w_i + v_j \otimes w_j = (v_i + \alpha v_j) \otimes w_i$. It is also possible to find some element of U in place of u such that k is minimal in making the above sum true. If $k = 1$ then u is a simple tensor and we're done. Suppose then that $k > 1$. Since the v_i are linearly independent we can define an injective $\mathfrak{l}$-module homomorphism $\psi \colon W \oplus \cdots \oplus W \to V \otimes W$ by

$$(w'_1, \dots, w'_k) \mapsto \sum_{i=1}^{k} v_i \otimes w'_i. \tag{4.2.30}$$

Then u is the image of some $\boldsymbol{w} = (w_1, \dots, w_k)$. If $\boldsymbol{w}'$ is some element of the smallest $\mathfrak{l}$-submodule W' of $W \oplus \cdots \oplus W$ containing $\boldsymbol{w}$ then $\psi(\boldsymbol{w}') \in U$. Consider the projection of W' onto the first two terms, $W \oplus W \oplus 0 \oplus \cdots$. This is an $\mathfrak{l}$-submodule of $W \oplus W$ and contains (w_1, w_2). We have assumed that w_1 and w_2 are not proportional, and thus by Lemma 4.2.25 this projection must be the whole of $W \oplus W$. In particular, $(w_1, 0)$ is in the image of the

projection. Let $\boldsymbol{w}' \in W'$ be some element projecting to $(w_1, 0)$. Then $\boldsymbol{w}' = (w_1, 0, w_2', w_3', \ldots, w_k')$. This then gives us an element in U that can be formed by summing $k - 1$ vectors of the form $v_i \otimes w_i'$, which contradicts the minimality assumption of k, and thus we must have $k = 1$ and we are done. $\qquad\square$

Lemma 4.2.31 Let V and W be finite-dimensional $\mathfrak{g}$-modules and let $\mathrm{Hom}_{\mathbb{k}}(V, W)$ be the space of linear maps $V \to W$. Equipping this space with the $\mathfrak{g}$-action

$$(X \cdot f)(v) = X \cdot f(v) - f(X \cdot v) \qquad X \in \mathfrak{g}, v \in V, \text{ and } f \in \mathrm{Hom}_{\mathbb{k}}(V, W)$$
$$(4.2.32)$$

defines a $\mathfrak{g}$-module.

Proof. Linearity is immediate because f is linear so we may pull out any scalar factors and pull apart any sums. We need only show that the bracket acts as required:

$$
\begin{aligned}
([X, Y] \cdot f)(v) &= [X, Y] \cdot f(v) - f([X, Y] \cdot f) & (4.2.33) \\
&= X \cdot (Y \cdot f(v)) - Y \cdot (X \cdot f(v)) - f(X \cdot (Y \cdot v) - Y \cdot (X \cdot v)) \\
&= X \cdot (Y \cdot f(v)) - Y \cdot (X \cdot f(v)) - f(X \cdot (Y \cdot v)) + f(Y \cdot (X \cdot v)) \\
&= X \cdot (Y \cdot f(v)) - f(X \cdot (Y \cdot v)) - (Y \cdot (X \cdot f(v)) - f(Y \cdot (X \cdot v))) \\
&= (X \cdot (Y \cdot f))(v) - (Y \cdot (X \cdot f))(v). & (4.2.34)
\end{aligned}
$$

This holds for all $v \in V$, and thus

$$[X, Y] \cdot f = X \cdot (Y \cdot f) - Y \cdot (X \cdot f) \qquad\qquad (4.2.35)$$

as required. $\qquad\square$

Since the sum of two $\mathfrak{g}$-module homomorphisms is again a $\mathfrak{g}$-module homomorphism this shows that $\mathrm{Hom}_{\mathfrak{g}}(V, W)$ is a $\mathfrak{g}$-submodule of $\mathrm{Hom}_{\mathbb{k}}(V, W)$, and hence $\mathrm{Hom}_{\mathfrak{g}}(V, W)$ is a $\mathfrak{g}$-module in its own right. A fancy way of putting this is that $\mathfrak{g}$-Rep is enriched over itself.

4.3   Representation Theory of $\mathfrak{sl}(2, \mathbb{C})$

Recall that the complex Lie algebra $\mathfrak{sl}(2, \mathbb{C})$ can be defined as the span of $\{E, F, H\}$ with the bracket defined by

$$[H, E] = 2E, \quad [H, F] = -2F, \quad \text{and} \quad [E, F] = H. \qquad\qquad (4.3.1)$$

These relations make H a **semisimple element**, meaning that the linear map $X \mapsto [H, X]$ is diagonalisable. We can see this is the case because the bracket's definition above shows that E and F are eigenvectors of the map $[H, -]$, with eigenvalues $+2$ and -2 respectively, and alternativity tells us that $[H, H] = 0$, so H is an eigenvector of $[H, -]$ with eigenvalue 0.

Let V be a finite-dimensional $\mathfrak{sl}(2,\mathbb{C})$-module. Then we can decompose V into generalised eigenspaces of H:

$$V = \bigoplus_{\alpha \in \mathbb{C}} V_\alpha \tag{4.3.2}$$

where

$$V_\alpha = \{v \in V \mid (H - \alpha)^N \cdot v = 0 \text{ for some } N \in \mathbb{N}\}. \tag{4.3.3}$$

We call the V_α the **α-weight spaces** for H. The **weight vectors** of V are those $v \in V$ such that $v \in V_\alpha$ for some $\alpha \in \mathbb{C}$. We call α the **weight** of such a v. Note that not all vectors are weight vectors, any $v \in V$ can be written as $\sum_{\alpha \in \mathbb{C}} v_\alpha$ with $v_\alpha \in V_\alpha$, and so v is a weight vector only if $v_\alpha = 0$ for all but one value of α.

Lemma 4.3.4 If v is an eigenvector of H with eigenvalue α then $E \cdot v \, (F \cdot v)$ is an eigenvector of H with eigenvalue $\alpha + 2 \, (\alpha - 2)$.

Proof. We know that $H \cdot v = \alpha v$. Using the action of a Lie bracket we have

$$[H, E] \cdot v = H \cdot (E \cdot v) - E \cdot (H \cdot v) \tag{4.3.5}$$
$$= H \cdot (E \cdot v) - E \cdot (\alpha v) \tag{4.3.6}$$
$$= H \cdot (E \cdot v) - \alpha E \cdot v. \tag{4.3.7}$$

Using the defining relations of $\mathfrak{sl}(2,\mathbb{C})$ we have

$$[H, E] = 2E \implies 2E \cdot v = H \cdot (E \cdot v) - \alpha E \cdot v. \tag{4.3.8}$$

Rearranging this we have

$$H \cdot (E \cdot v) = 2E \cdot v + \alpha E \cdot v = (\alpha + 2)(E \cdot v), \tag{4.3.9}$$

and thus $E \cdot v$ is an eigenvector of H with eigenvalue $\alpha + 2$. The proof for $F \cdot v$ is analogous. $\qquad\square$

This extends to the following result about generalised eigenspaces.

Lemma 4.3.10 If $v \in V_\alpha$ is a weight vector of weight α then $E \cdot v \, (F \cdot v)$ is a weight vector of weight $\alpha + 2 \, (\alpha - 2)$.

If $V_\alpha \neq 0$ but $V_{\alpha+2} = 0 \, (V_{\alpha-2} = 0)$ then this implies that $E \cdot v = 0 \, (F \cdot v)$ for all $v \in V_\alpha$. We call $v \in V$ a **highest weight** (**lowest weight**) if $E \cdot v = 0 \, (F \cdot v = 0)$ and $H \cdot v = \alpha v$.

Lemma 4.3.11 Let V be a finite-dimensional $\mathfrak{sl}(2,\mathbb{C})$-module. Then V contains a highest weight vector, v_0. Setting $v_{-1} = 0$ and $v_i = \frac{1}{i!} F^i \cdot v_0$ for $i \geq 0$ if v_0 has weight α then

$$H \cdot v_i = (\alpha - 2i)v_i, \quad E \cdot v_i (\alpha - i + 1)v_{i-1}, \quad \text{and} \quad F \cdot v_i = (i+1)v_{i+1}. \tag{4.3.12}$$

Proof. STEP 1: $H \cdot v_i$

First note that $H \cdot v_0 = \alpha v_0$ so the $i = 0$ case holds. We proceed by induction. For the base case, $i = 1$, we have $v_1 = F \cdot v_0$. We then have $H \cdot (F \cdot v_0) = (\alpha - 2)(F \cdot v_0)$ by Lemma 4.3.4, and so $H \cdot v_1 = (\alpha - 2 \cdot 1)v_1$. Now suppose that the result holds for some $k \in \mathbb{N}$, that is, we have $H \cdot v_k = (\alpha - 2k)v_k$. Consider $H \cdot v_{k+1}$:

$$H \cdot v_{k+1} = H \cdot \left(\frac{1}{k+1} F \cdot v_k \right) \tag{4.3.13}$$

$$= \frac{1}{k+1} H \cdot (F \cdot v_k) \tag{4.3.14}$$

$$= \frac{1}{k+1} ([H, F] \cdot v_k + F \cdot (H \cdot v_k)) \tag{4.3.15}$$

$$= \frac{1}{k+1} (-2F \cdot v_k + F \cdot (\alpha - 2k)v_k) \tag{4.3.16}$$

$$= \frac{1}{k+1} (-2 + (\alpha - 2k))(F \cdot v_k) \tag{4.3.17}$$

$$= (\alpha - 2(k+1))v_{k+1}. \tag{4.3.18}$$

Thus, by induction, the result holds for all $i \in \mathbb{N}$.

STEP 2: $E \cdot v_i$

First note that $E \cdot v_0 = 0 = (\alpha - 0 + 1)v_{0-1}$ as $v_{-1} = 0$, so the $i = 0$ case holds. We again proceed by induction. for the base case, $i = 1$, we have $v_1 = F \cdot v_0$. We then have

$$E \cdot v_1 = E \cdot (F \cdot v_0) \tag{4.3.19}$$

$$= \underbrace{[E, F]}_{=H} \cdot v_0 + F \cdot \underbrace{(E \cdot v_0)}_{=0} \tag{4.3.20}$$

$$= H \cdot v_0 \tag{4.3.21}$$

$$= \alpha v_0 \tag{4.3.22}$$

$$= (\alpha - 1 + 1)v_{1-1}. \tag{4.3.23}$$

Now suppose that the result holds for some $k \in \mathbb{N}$, that is, we have $E \cdot v_k = (\alpha - k + 1)v_{k-1}$. Consider $E \cdot v_{k+1}$:

$$E \cdot v_{k+1} = \frac{1}{k+1} E \cdot (F \cdot v_k) \tag{4.3.24}$$

$$= \frac{1}{k+1} ([E, F] \cdot v_k + F \cdot (E \cdot v_k)) \tag{4.3.25}$$

$$= \frac{1}{k+1} (H \cdot v_k + (\alpha - k + 1)F \cdot v_{k-1}) \tag{4.3.26}$$

$$= \frac{1}{k+1} ((\alpha - 2k)v_k + (\alpha - k + 1)(k+1)v_k) \tag{4.3.27}$$

$$= \text{Something isn't working here} \tag{4.3.28}$$

$$= (\alpha - k)v_k \tag{4.3.29}$$

$$= (\alpha - (k+1) + 1)v_k \tag{4.3.30}$$

STEP 3: $F \cdot v_i = (i + 1)v_{i+1}$
This is evident from the definition of v_i.   □

Since we have $H \cdot v_i = (\alpha - 2i)v_i$ the vectors v_i are all in different eigen-spaces of H, and thus are linearly independent. Since, by assumption, V is finite-dimensional there must therefore be some $N \in \mathbb{N}$ such that $v_N \neq 0$ but $v_{N+1} = 0$. Considering the action of E according to the above result we have $0 = E \cdot 0 = E \cdot v_{N+1} = (\alpha - N)v_N$ and since $v_N \neq 0$ we must therefore have $\alpha - N = 0$, so $\alpha = N$, and thus the only possible eigenvalues of H are non-negative integers.

Further assume that V is a simple module, and that $V \neq 0$. Then by simplicity the submodule $\text{span}\{v_0, \dots, v_N\}$ must be all of V, that is, $\{v_0, \dots, v_N\}$ provides a basis for V, and $\dim V = N + 1$. Further, the weight spaces are V_{N-2i} for $i = 0, \dots, N$, and each is one-dimensional, being spanned by v_i, that is $V_{N-2i} = \mathbb{C}v_i$.

Combining these results we have proven the following.

> **Theorem 4.3.31 — Classification of Finite Dimensional Simple $\mathfrak{sl}(2, \mathbb{C})$-Modules.** Let V be a finite-dimensional simple $\mathfrak{sl}(2, \mathbb{C})$-module with a highest weight vector of weight α. Then
>
> 1. α is a non-negative integer, N;
>
> 2. $\dim V = N + 1$ and the nonzero weight spaces of V are $V_N, V_{N-2}, \dots, V_{-N}$, each of which is one-dimensional.
>
> Conversely, for any positive integer, N, there exists a unique simple $\mathfrak{sl}(2, \mathbb{C})$-module $V(N)$ with a highest weight vector of weight N.

This leads to the following picture of $\mathfrak{sl}(2, \mathbb{C})$ representations:

$$
\begin{array}{c}
0 \\
E \uparrow \\
V_N \circlearrowleft H \\
\vdots \\
V_{\alpha+2} \circlearrowleft H \\
E \uparrow \downarrow F \\
V_\alpha \circlearrowleft H \\
E \uparrow \downarrow F \\
V_{\alpha-2} \circlearrowleft H \\
\vdots \\
V_{-N} \circlearrowleft H \\
\downarrow F \\
0
\end{array}
\qquad (4.3.32)
$$

Consider the defining $\mathfrak{sl}(2,\mathbb{C})$ module $\mathbb{C}^2$ where H, E, and F act as the matrices

$$H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad E = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \text{and} \quad F = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}. \tag{4.3.33}$$

The eigenvectors of H are

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{4.3.34}$$

These have eigenvalues $+1$ and -1 respectively. Of these only the first vanishes when we act with E, giving

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{4.3.35}$$

This means that the first eigenvector is a highest weight vector of $\mathbb{C}^2$. The weight of this vector, its H eigenvalue, is $+1$ Further, this representation is simple, because if we start with some arbitrary vector then acting with E gives a vector along e_1, and acting with F gives a vector along e_2, spanning the whole space. So, this representation is isomorphic to $V(1)$.

Now consider the adjoint action of $\mathfrak{sl}(2,\mathbb{C})$. Recall that this is the action of $\mathfrak{sl}(2,\mathbb{C})$ on itself by the bracket. Then we may identify the representation space with $\mathbb{C}^3$ where

$$E = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad H = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad F = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{4.3.36}$$

From the defining bracket relations we can see that we have

$$\mathrm{ad}_E = \begin{bmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathrm{ad}_H = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \quad \text{and} \quad \mathrm{ad}_F = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}.$$

These matrices are such that $\mathrm{ad}_X(Y) = [X,Y]$. For example, we have

$$\mathrm{ad}_H(E) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} = 2E = [H,E]. \tag{4.3.37}$$

The eigenvectors of H are E, H, and F, with eigenvalues $+2$, 0, and -2 respectively. Of these, only E has the property that $E \cdot E = \mathrm{ad}_E(E) = [E,E]$ vanishes, and thus E is a highest weight vector, and it has weight 2. Acting with F we have $F \cdot E = \mathrm{ad}_F(E) = [F,E] = -H$, and acting with F again we have $F \cdot H = \mathrm{ad}_F(H) = [F,H] = 2F$. Thus, E generates the whole space, showing that any submodule containing E must be all of $\mathfrak{sl}(2,\mathbb{C})$, and a similar argument holds for H or F. This shows that the module is simple. Then this module is isomorphic to $V(2)$.

<div align="center">

Five

</div>

<div align="center">

The Structure of Lie Algebras

</div>

In this chapter we will analyse the structure of general Lie algebras. We do this by looking at some special types of Lie algebras which we define here.

5.1 Centre

> **Definition 5.1.1** Let $\mathfrak{g}$ be a Lie algebra. Its **centre**, $\mathfrak{z}(\mathfrak{g})$, is the ideal
>
> $$\mathfrak{z}(\mathfrak{g}) := \{X \in \mathfrak{g} \mid [X, Y] = 0 \forall Y \in \mathfrak{g}\}. \tag{5.1.2}$$

When we interpret the bracket as a commutator the centre is exactly the collection of all elements which commute with all other elements of the Lie algebra when viewed as an associative algebra. In general we say that two elements commute if their bracket vanishes. This definition is then analogous to the centre of a group, which is all elements that commute with all other elements of the group.

> **Example 5.1.3** We have that $\mathfrak{z}(\mathfrak{gl}(n, \mathbb{C})) = \mathrm{span}_{\mathbb{C}}\{I_n\}$, this is just Schur's lemma (Lemma 4.2.6).

> **Lemma 5.1.4** The centre of a Lie algebra is an ideal.
>
> *Proof.* Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{z}(\mathfrak{g})$ its centre. If $X, Y \in \mathfrak{z}(\mathfrak{g})$ then we have
>
> $$[X + Y, Z] = [X, Z] + [Y, Z] = 0 \tag{5.1.5}$$
>
> for all $Z \in \mathfrak{g}$, and so $\mathfrak{z}(\mathfrak{g})$ is closed under addition. Similarly, if $\lambda \in \Bbbk$ and $X \in \mathfrak{z}(\mathfrak{g})$ then we have
>
> $$[\lambda X, Y] = \lambda[X, Y] = 0 \tag{5.1.6}$$
>
> for all $Y \in \mathfrak{g}$, so $\mathfrak{z}(\mathfrak{g})$ is closed under scalar multiplication. Thus, $\mathfrak{z}(\mathfrak{g})$ is a subspace of $\mathfrak{g}$. Now notice that if $X \in \mathfrak{z}(\mathfrak{g})$ and $Y \in \mathfrak{g}$ then for all $Z \in \mathfrak{g}$ we have
>
> $$[[X, Y], Z] = [0, Z] = 0, \tag{5.1.7}$$
>
> and so $[X, Y] \in \mathfrak{z}(\mathfrak{g})$ so $\mathfrak{z}(\mathfrak{g})$ is an ideal. $\qquad \square$

Notice that a Lie algebra, $\mathfrak{g}$, is abelian if and only if we have $\mathfrak{g} = \mathfrak{z}(\mathfrak{g})$.

5.2 Simple Lie Algebra

Definition 5.2.1 A Lie algebra, $\mathfrak{g}$, is **simple** if it is nonabelian has no proper ideals.

We exclude abelian Lie algebras in part because this excludes the one-dimensional Lie algebra, which doesn't poses all of the properties that we want from simple Lie algebras. These properties will become clear later.

Example 5.2.2 The Lie algebra $\mathfrak{sl}(n, \mathbb{C})$ is simple.

5.3 Solvable and Nilpotent Lie Algebras

Definition 5.3.1 — Nilpotent Lie Algebra The **lower central series** of a Lie algebra, $\mathfrak{g}$, is the series of subalgebras

$$\mathfrak{g} = \mathfrak{g}_0 \subseteq \mathfrak{g}_1 \subseteq \mathfrak{g}_2 \subseteq \cdots . \tag{5.3.2}$$

These are defined recursively by

$$\mathfrak{g}_0 = \mathfrak{g}, \quad \mathfrak{g}_1 = [\mathfrak{g}, \mathfrak{g}], \quad \text{and} \quad \mathfrak{g}_{k+1} = [\mathfrak{g}, \mathfrak{g}_k]. \tag{5.3.3}$$

A Lie algebra is **nilpotent** if its lower central series eventually terminates with 0. That is, there is some $n \in \mathbb{N}$ such that $\mathfrak{g}_n = 0$.

Example 5.3.4 The Lie algebra $\mathfrak{n}(n, \mathbb{C})$ consisting of strictly upper triangular matrices is nilpotent. To see this note that the product of any two strictly upper triangular matrices has zero on both the diagonal and the superdiagonal. Then the product of any such matrix with a strictly upper triangular matrix has zero on the leading diagonal and the two diagonals above it. Continuing on like this we must eventually have all zeros, since the matrices are finite, and thus the lower central series terminates.

Lemma 5.3.5 Each term in the lower central series is an ideal of $\mathfrak{g}$.

Proof. Since $[\mathfrak{g}, \mathfrak{g}]$ is defined to be the span of all elements of the form $[X, Y]$ with $X, Y \in \mathfrak{g}$ it is clearly a subspace of $\mathfrak{g}$. Further, since the bracket of two brackets is just a bracket of elements of $\mathfrak{g}$ it is again found in $[\mathfrak{g}, \mathfrak{g}]$. Thus, $[\mathfrak{g}, \mathfrak{g}]$ is a subalgebra of $\mathfrak{g}$. Now suppose that $X, Y, Z \in \mathfrak{g}$, then $[X, [Y, Z]]$ is clearly an element of $[\mathfrak{g}, \mathfrak{g}]$, and by linearity this is also true if we replace $[Y, Z]$ with a linear combination of brackets, thus $[\mathfrak{g}, \mathfrak{g}]$ is an ideal of $\mathfrak{g}$. The same proof shows that $\mathfrak{g}_2$ is an ideal of $\mathfrak{g}_1$ and so on,

and hence all of these are ideals of $\mathfrak{g}$ since an ideal of an ideal is again an ideal.   □

Definition 5.3.6 — Solvable Lie Algebra The **derived series** or **upper central series** of a Lie algebra, $\mathfrak{g}$, is the series of subalgebras

$$\mathfrak{g} = \mathfrak{g}^0 \subseteq \mathfrak{g}^1 \subseteq \mathfrak{g}^2 \subseteq \cdots . \tag{5.3.7}$$

These are defined recursively by

$$\mathfrak{g}_0 = \mathfrak{g}, \quad \mathfrak{g}^1 = [\mathfrak{g}, \mathfrak{g}], \quad \text{and} \quad \mathfrak{g}^{k+1} = [\mathfrak{g}^k, \mathfrak{g}^k]. \tag{5.3.8}$$

The Lie subalgebra $\mathfrak{g}^1 = \mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$ is sometimes called the **derived subalgebra** of $\mathfrak{g}$.
A Lie algebra is **solvable** if its derived series eventually terminates with 0. That is, there is some $n \in \mathbb{N}$ such that $\mathfrak{g}^n = 0$.

Note the difference between this definition and the lower central series, we have

$$\mathfrak{g}_2 = [\mathfrak{g}, \mathfrak{g}_1] = [\mathfrak{g}, [\mathfrak{g}, \mathfrak{g}]] \tag{5.3.9}$$

for the lower central series, and

$$\mathfrak{g}^2 = [\mathfrak{g}^1, \mathfrak{g}^1] = [[\mathfrak{g}, \mathfrak{g}], [\mathfrak{g}, \mathfrak{g}]] \tag{5.3.10}$$

for the derived series.

Example 5.3.11 The Lie algebra $\mathfrak{b}(n, \mathbb{C})$ consisting of upper triangular matrices is nilpotent. To see this note that the commutator of two upper triangular matrices is strictly upper triangular, and the commutator of two strictly upper triangular matrices has zero on the leading diagonal and super diagonal, and the commutator of any two such matrices has three diagonals of zero, and so on. Thus, the derived series eventually terminates with 0.

Lemma 5.3.12 Every nilpotent algebra is solvable.

Proof. Suppose that $\mathfrak{g}$ is a nilpotent Lie algebra. Then its upper central series terminates. This means that we eventually have

$$[\mathfrak{g}, [\mathfrak{g}, [\cdots, \mathfrak{g}]]] = 0. \tag{5.3.13}$$

This then necessitates that if the rightmost copy of $\mathfrak{g}$ is replaced with any subalgebra of $\mathfrak{g}$ we still have 0, and in particular we may replace it with some subalgebra given by iterated commutators to achieve an element of the derived series, and thus the derived series terminates.   □

Lemma 5.3.14 If $\mathfrak{g}$ is a Lie algebra with an ideal $\mathfrak{i}$ such that $\mathfrak{i}$ is solvable and $\mathfrak{g}/\mathfrak{i}$ is solvable then $\mathfrak{g}$ is solvable.

Proof. STEP 1: $(\mathfrak{g}/\mathfrak{i})^n \cong (\mathfrak{g}^n + \mathfrak{i})/\mathfrak{i}$

We prove this by induction on n. The base case, $n = 0$, is that $\mathfrak{l}/\mathfrak{i} \cong (\mathfrak{g}+\mathfrak{i})/\mathfrak{i}$. This is true because the second isomorphism theorem tells us that $(\mathfrak{g} + \mathfrak{i})/\mathfrak{i} \cong \mathfrak{l}/(\mathfrak{g} \cap \mathfrak{i})$, and since $\mathfrak{i} \subseteq \mathfrak{g}$ we have $\mathfrak{g} \cap \mathfrak{i} = \mathfrak{i}$, so we have $(\mathfrak{g} + \mathfrak{i})/\mathfrak{i} \cong \mathfrak{l}/\mathfrak{i}$. For the inductive step suppose the result holds for some $k \in \mathbb{N}$, that is, $(\mathfrak{g}/\mathfrak{i})^k \cong (\mathfrak{g}^k + \mathfrak{i})/\mathfrak{i}$. Then we have

$$
\begin{aligned}
(\mathfrak{g}/\mathfrak{i})^{k+1} &= ((\mathfrak{g}/\mathfrak{i})^k)' && \text{definition of derived subalgebra} \\
&\cong ((\mathfrak{g}^k + \mathfrak{i})/\mathfrak{i})' && \text{induction hypothesis} && (5.3.15) \\
&\cong ((\mathfrak{g}^k + \mathfrak{i})' + \mathfrak{i})/\mathfrak{i} && \text{base case} && (5.3.16) \\
&\cong (\mathfrak{g}^{k+1} + \mathfrak{i}' + \mathfrak{i})/\mathfrak{i} && \mathfrak{i}' \subseteq \mathfrak{i} && (5.3.17) \\
&= (\mathfrak{g}^{k+1} + \mathfrak{i})/\mathfrak{i} && && (5.3.18)
\end{aligned}
$$

and the result holds for all $n \in \mathbb{N}$.

STEP 2: $(\mathfrak{g}^n)^m = \mathfrak{g}^{n+m}$

We prove this by induction on m. For the base case we take $m = 1$, then we have $(\mathfrak{g}^n)^m = (\mathfrak{g}^n)^1 = (\mathfrak{g}^n)' = \mathfrak{g}^{n+1}$ by definition. Now suppose that for some $k \in \mathbb{N}$ we have $(\mathfrak{g}^n)^k = \mathfrak{g}^{n+k}$. Then consider $k + 1$, we have $(\mathfrak{g}^n)^{k+1} = ((\mathfrak{g}^n)^k)' = (\mathfrak{g}^{n+k})' = \mathfrak{g}^{n+k+1}$, and so the result holds for all $m \in \mathbb{N}$.

STEP 3: $\mathfrak{g}$ IS SOLVABLE

By assumption $\mathfrak{i}$ and $\mathfrak{g}/\mathfrak{i}$ are solvable, so there exist $M, N \in \mathbb{N}$ such that $\mathfrak{i}^M = 0$ and $(\mathfrak{g}/\mathfrak{i})^N = 0$. We therefore have $0 = (\mathfrak{g}/\mathfrak{i})^N \cong (\mathfrak{g}^N + \mathfrak{i})/\mathfrak{i}$. From this it follows that $\mathfrak{g}^N + \mathfrak{i} \subseteq \mathfrak{i}$, because this is the condition for a quotient to be zero. Thus, we must have $\mathfrak{g}^N \subseteq \mathfrak{i}$. Taking the Mth derived subalgebra of each side we have $(\mathfrak{g}^N)^M \subseteq \mathfrak{i}^M = 0$, and thus $\mathfrak{g}^{N+M} = (\mathfrak{g}^N)^M = 0$, so $\mathfrak{g}$ is solvable. $\square$

Lemma 5.3.19 If $\mathfrak{g}$ is a Lie algebra with solvable ideals $\mathfrak{i}$ and $\mathfrak{j}$ then $\mathfrak{i} + \mathfrak{j}$ is a solvable ideal of $\mathfrak{g}$.

Proof. Since $\mathfrak{i}$ and $\mathfrak{j}$ are solvable there exist $M, N \in \mathbb{N}$ such that $\mathfrak{i}^N = 0$ and $\mathfrak{j}^M = 0$. By the second isomorphism theorem we have $(\mathfrak{i} + \mathfrak{j})/\mathfrak{j} \cong \mathfrak{i}/(\mathfrak{i} \cap \mathfrak{j})$, and taking the Mth derived subalgebra of each side we get $((\mathfrak{i} + \mathfrak{j})/\mathfrak{j})^M = (\mathfrak{i}/(\mathfrak{i} \cap \mathfrak{j}))^M$. By the first step in the proof of Lemma 5.3.14 we know that $(\mathfrak{i}/(\mathfrak{i} \cap \mathfrak{j}))^M \cong (\mathfrak{i}^M + (\mathfrak{i} \cap \mathfrak{j}))/(\mathfrak{i} \cap \mathfrak{j}) = (0 + (\mathfrak{i} \cap \mathfrak{j}))/(\mathfrak{i} \cap \mathfrak{j}) = (\mathfrak{i} \cap \mathfrak{j})/(\mathfrak{i} \cap \mathfrak{j}) = 0$. Thus, $((\mathfrak{i} + \mathfrak{j})/\mathfrak{j})^M = 0$, so $(\mathfrak{i} + \mathfrak{j})/\mathfrak{j}$ is solvable. Hence, by Lemma 5.3.14 we know that $\mathfrak{i} + \mathfrak{j}$ is solvable. $\square$

Definition 5.3.20 — Semisimple Lie Algebra A Lie algebra, $\mathfrak{g}$, is semisimple if it contains no proper solvable ideals.

We immediately see that every simple Lie algebra is semisimple, since a simple Lie algebra contains no nonzero proper ideals, so it certainly doesn't contain any solvable ideals. Note that this is not true if we relax the definition of simple Lie algebras to allow abelian Lie algebras.

Lemma 5.3.21 Let $\mathfrak{g}$ be a finite dimensional Lie algebra. Then there exists a maximal (with respect to inclusion) solvable ideal of $\mathfrak{g}$ called the **radical** of $\mathfrak{g}$, denoted rad $\mathfrak{g}$.

Proof. Let $\{\mathfrak{i}_\alpha\}_{\alpha \in A}$ be the collection of all solvable ideals. Define

$$\mathfrak{i} = \sum_{\alpha \in A} \mathfrak{i}_\alpha = \left\{ \sum_{\alpha \in A} i_\alpha \mid i_\alpha \in \mathfrak{i}_\alpha \text{ and } i_\alpha = 0 \text{ for all but finitely many } \alpha \in A \right\}. \tag{5.3.22}$$

This is a subspace of $\mathfrak{g}$ since it is a sum of subspaces.
We claim that we actually have

$$\mathfrak{i} = \mathfrak{i}_{\alpha_1} + \cdots + \mathfrak{i}_{\alpha_n} \tag{5.3.23}$$

for some finite indexing subset $\{\alpha_1, \ldots, \alpha_n\} \subseteq A$. This is the case because $\mathfrak{g}$ is finite dimensional, so $\mathfrak{i}$ must also be finite dimensional. If $\mathfrak{i} = \mathfrak{i}_{\alpha_1}$ for some $\alpha_1 \in A$ then we are done. If this is not the case then we can find some $\alpha_2 \in A$ such that $\mathfrak{i}_{\alpha_1} + \mathfrak{i}_{\alpha_2} \neq \mathfrak{i}_{\alpha_1}$. If $\mathfrak{i} = \mathfrak{i}_{\alpha_1} + \mathfrak{i}_{\alpha_2}$ we are done. Iterating this process we see that since we get a new subspace each time we add an $\mathfrak{i}_{\alpha_k}$ and we clearly have $\mathfrak{i}_{\alpha_1} + \cdots \mathfrak{i}_{\alpha_{k-1}} \subseteq \mathfrak{i}_{\alpha_1} + \cdots + \mathfrak{i}_{\alpha_{k-1}} + \mathfrak{i}_{\alpha_k}$ the dimension must increase at each step, and thus this process must terminate for $\mathfrak{i}$ to be finite dimensional.
By construction any solvable ideal, $\mathfrak{j}$, can be found in $\mathfrak{i}$, since we have $\mathfrak{j} = \mathfrak{i}_\alpha$ for some $\alpha \in A$ and we therefore have terms of the form $j = 0 + \cdots + 0 + j + 0 + \cdots + 0 \in \mathfrak{i}$ for each $j \in \mathfrak{j}$. Thus, $\mathfrak{i}$ is a maximal subspace of $\mathfrak{g}$ containing all solvable ideals.
We now need only show that $\mathfrak{i}$ is an ideal. That it is a subalgebra follows by linearity of the bracket and the fact that $\mathfrak{i}$ is defined as a sum of subalgebras. That it is an ideal follows because each ideal in the sum defining $\mathfrak{i}$ has the absorbing property, and so the whole sum has the absorbing property. $\square$

This result tells us that $\mathfrak{g}$ is semisimple if and only if rad $\mathfrak{g} = 0$. One thing we can say is that $\mathfrak{g}/\text{rad } \mathfrak{g}$ is semisimple. This is because an ideal of the quotient $\mathfrak{g}/\text{rad } \mathfrak{g}$ is of the form $\mathfrak{i}/\text{rad } \mathfrak{g}$ where $\mathfrak{i} \supseteq \text{rad } \mathfrak{g}$ is an ideal of $\mathfrak{i}$. Supposing that $\mathfrak{i}/\text{rad } \mathfrak{g}$ is solvable, and knowing that rad $\mathfrak{g}$ is solvable we must have that $\mathfrak{i}$ is solvable, meaning that $\mathfrak{i} \subseteq \text{rad } \mathfrak{g}$. Thus, we have $\mathfrak{i} = \text{rad } \mathfrak{g}$, and so $\mathfrak{i}/\text{rad } \mathfrak{g} = \mathfrak{i}/\mathfrak{i} = 0$. Thus, there are no nontrivial proper solvable ideals of $\mathfrak{g}/\text{rad } \mathfrak{g}$, so $\mathfrak{g}/\text{rad } \mathfrak{g}$ is semisimple.

5.4 Engel's Theorem

Let V be a vector space. A **nilpotent** of V is a linear map $\varphi : V \to V$ such that there exists some $N \in \mathbb{N}$ such that φ^N is the zero map. That is, $\varphi^N(v) = 0$ for all $v \in V$.

Let $\mathfrak{g}$ be a Lie algebra. An element, X, of $\mathfrak{g}$ is called ad-**nilpotent** if ad_X is a nilpotent endomorphism of $\mathfrak{g}$.

Before we can prove the big result of this section, Engel's theorem, we need a couple of lemmas. The proof of the first is quite a lot of algebra, manipulating various sums, the details of the proof aren't important, so I'd suggest skipping it.

Lemma 5.4.1 Let A be an associative algebra with the Lie algebra structure given by the commutator. Then for $n \in \mathbb{N}$ and $X, Y \in A$ we have

$$\mathrm{ad}_X^n(Y) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} x^{n-k} y x^k. \tag{5.4.2}$$

Proof. The proof is by induction on n. For $n = 1$ we have

$$\sum_{k=0}^{1} (-1)^k \binom{2}{k} X^{1-k} Y X^k = (-1)^0 \binom{2}{0} XY + (-1)^1 \binom{2}{1} YX \tag{5.4.3}$$

$$= XY - YX \tag{5.4.4}$$

$$= [X, Y] \tag{5.4.5}$$

$$= \mathrm{ad}_X^1(Y). \tag{5.4.6}$$

Now suppose that the result holds for some $m \in \mathbb{N}$. Consider the following calculation

$$\mathrm{ad}_X^{m+1}(Y) = [x, \mathrm{ad}_X^m(Y)] \tag{5.4.7}$$

$$= \left[X, \sum_{k=0}^{m} (-1)^k \binom{m}{k} X^{m-k} Y X^k \right] \tag{5.4.8}$$

$$= \sum_{k=0}^{m} (-1)^k \binom{m}{k} [X, X^{m-k} Y X^k] \tag{5.4.9}$$

$$= \sum_{k=0}^{m} (-1)^k \binom{m}{k} (XX^{m-k} Y X^k - X^{m-k} Y X^k X) \tag{5.4.10}$$

$$= \sum_{k=0}^{m} (-1)^k \binom{m}{k} (X^{m-k+1} Y X^k - X^{m-k} Y X^{k+1}) \tag{5.4.11}$$

$$= \sum_{k=0}^{m} (-1)^k \binom{m}{k} X^{m-k+1} Y X^k - \sum_{k=0}^{m} (-1)^k \binom{m}{k} X^{m-k} Y X^{k+1}.$$

We can reindex the second sum so that k runs from 1 to $m + 1$. To do so we replace each k in the second sum with $k - 1$:

$$\sum_{k=0}^{m} (-1)^k \binom{m}{k} X^{m-k} Y X^{k+1} = \sum_{k=1}^{m+1} (-1)^{k-1} \binom{m}{k-1} X^{m-k+1} Y X^k. \tag{5.4.12}$$

Thus, we have

$$\text{ad}_X^{m+1}(Y) = \sum_{k=0}^{m} (-1)^k \binom{m}{k} X^{m-k+1} Y X^k \tag{5.4.13}$$

$$- \sum_{k=1}^{m+1} (-1)^{k-1} \binom{m}{k-1} X^{m-k+1} Y X^k$$

$$= (-1)^0 \binom{m}{0} X^{m+1} Y X^0 + \sum_{k=1}^{m} (-1)^k \binom{m}{k} X^{m-k+1} Y X^k$$

$$- \sum_{k-1}^{m} (-1)^{k-1} \binom{m}{k-1} X^{m-k-1} Y X^k - (-1)^m \binom{m}{m} X^0 Y X^{m+1}$$

$$= X^{m+1} Y + \sum_{k=1}^{m} (-1)^k \binom{m}{k} X^{m-k+1} Y X^k \tag{5.4.14}$$

$$- \sum_{k-1}^{m} (-1)^{k-1} \binom{m}{k-1} X^{m-k-1} Y X^k - (-1)^m Y X^{m+1}$$

$$= \sum_{k=1}^{m} \left((-1)^k \binom{m}{k} + (-1)^k \binom{m}{k-1} \right) X^{m-k+1} Y X^k$$

$$+ X^{m+1} Y - (-1)^m Y X^{m+1} \tag{5.4.15}$$

$$= \sum_{k=1}^{m} (-1)^k \binom{m+1}{k} X^{m-k+1} Y X^k + X^{m+1} Y - (-1)^m Y X^{m+1}$$

where we've used the result

$$\binom{m}{k} + \binom{m}{k-1} = \frac{m+1}{k}, \tag{5.4.16}$$

which can be proved by manipulating some factorials. We can then recognise that the two terms not in the sum are just the $k = 0$ term and the $k = m + 1$ term, so that we have

$$\text{ad}_X^{m+1}(Y) = \sum_{k=0}^{m+1} (-1)^k \binom{m+1}{k} X^{m+1-k} Y X^k. \tag{5.4.17}$$

So, the result holds for all $n \in \mathbb{N}$. $\square$

Lemma 5.4.18 Let V be a finite dimensional vector space and let $X \in \mathfrak{gl}(V)$ be a nilpotent endomorphism. Then X is ad-nilpotent. Further, if $n \in \mathbb{N}$ is the smallest positive integer making $X^n = 0$ then $2n + 1$ is such that $\text{ad}_X^{2n+1} = 0$.

Proof. Let $X \in \mathfrak{gl}(V)$ be a nilpotent endomorphism. That is, there is some $n \in \mathbb{N}$ such that $X^n = 0$. Since $\mathfrak{gl}(V)$ is a Lie algebra under the commutator

we may apply the results of Lemma 5.4.1. For all $X, Y \in \mathfrak{gl}(V)$ we have

$$\mathrm{ad}_X^{2n+1}(Y) = \sum_{k=0}^{2n+1} (-1)^k \binom{2n+1}{k} X^{2n+1-k} Y X^k. \qquad (5.4.19)$$

This always vanishes, and we can show this by showing that every term must vanish:

- if $k \geq n$ then the term contains $X^k = 0$; or

- if $k < n$ then $2n + 1 - k > n$ then the term contains $X^{2n+1-k} = 0$.

So, every term is zero, and the whole expression vanishes. $\qquad \square$

Theorem 5.4.20. Let V be a finite dimensional vector space over $\Bbbk$ with $V \neq 0$. Let $\mathfrak{n}$ be a subalgebra of $\mathfrak{gl}(V)$ such that all elements of $\mathfrak{n}$ are nilpotent endomorphisms of V. Then there exists some vector $v \in V$ with $v \neq 0$ such that $\mathfrak{n} \cdot v = 0$, that is, $v \in \ker \varphi$ for all $\varphi \in \mathfrak{n}$.

Proof. The proof is by induction on $\dim \mathfrak{n}$.

STEP 1: BASE CASE
The base case is $\dim \mathfrak{n} = 1$. In this case we know that $\mathfrak{n} = \Bbbk X$ for some (in fact, any) $X \in \mathfrak{n} \setminus \{0\}$. We know that X is nilpotent by assumption. So, there exists some minimal $n \in \mathbb{N}$ such that $X^n = 0$, and since $X \neq 0$ we know that $n \geq 2$. By the assumption that n is minimal we know that $X^{n-1} \neq 0$. Recalling that X^{n-1} is an element of $\mathfrak{gl}(V)$ this means that there exists some $w \in V$ such that $X^{n-1}(w) \neq 0$. However, we have that

$$X(X^{n-1}(w)) = X^n(W) = 0(w) = 0, \qquad (5.4.21)$$

and so $X^{n-1}(w) \in \ker X$. Thus, $v = X^{n-1}(w)$ is as needed, and we are done with the base case.

STEP 2: INDUCTION
Suppose that the result holds when $\dim \mathfrak{n} = k$ for some $k \in \mathbb{N}$. We first show that there is an ideal of $\mathfrak{n}$ of codimension 1. That is, there is an ideal, $\mathfrak{l} \subseteq \mathfrak{n}$ such that $\dim \mathfrak{l} = k - 1$.
To do this we first suppose that $\mathfrak{l} \subseteq \mathfrak{n}$ is a maximal proper subalgebra of $\mathfrak{n}$. This makes sense because

- $\mathfrak{n}$ definitely has subalgebras, any one-dimensional subspace is a subalgebra; and

- $\dim \mathfrak{n}$ is finite, and so the dimension of $\mathfrak{l}$ is bounded, and as such we can simply take all subalgebras and pick one with maximal dimension.

Since $\mathfrak{n}$ is a vector space and $\mathfrak{l} \subseteq \mathfrak{n}$ is a subspace we can form the quotient *vector space* $\mathfrak{n}/\mathfrak{l}$. We do now know if this is a quotient *Lie algebra*, since

we only know that $\mathfrak{l}$ is a subalgebra, not that it's an ideal. Consider the adjoint action of $\mathfrak{n}$ on itself. We can restrict this to consider the action of $\mathfrak{l}$ on $\mathfrak{n}$, given by $X \cdot Y = [X, Y]$ for $X \in \mathfrak{l}$ and $Y \in \mathfrak{n}$. This makes $\mathfrak{n}$ an $\mathfrak{l}$-module. Since $\mathfrak{l}$ is a subalgebra we have $[\mathfrak{l}, \mathfrak{l}] \subseteq \mathfrak{l}$, which means that this action naturally gives the quotient vector space, $\mathfrak{n}/\mathfrak{l}$, an $\mathfrak{l}$-module structure given by

$$X \cdot (Y + \mathfrak{l}) = [X, Y] + \mathfrak{l} \tag{5.4.22}$$

for $X\mathfrak{l}$ and $Y \in \mathfrak{n}$.

By assumption elements of $\mathfrak{n}$ are nilpotent, and thus by Lemma 5.4.18 they are also ad-nilpotent. Hence, $\mathrm{ad}(\mathfrak{l}) = \{\mathrm{ad}_X \mid X \in \mathfrak{l}\}$ consists of nilpotent endomorphisms in $\mathfrak{gl}(\mathfrak{n})$, and $\mathrm{ad}(\mathfrak{l})$ is a Lie algebra under the commutator. Since $\mathfrak{l}$ is a proper subalgebra of $\mathfrak{n}$ we know that $\dim \mathfrak{l} < \dim \mathfrak{n}$, and hence $\dim \mathrm{ad}(\mathfrak{l}) < \dim \mathfrak{n} = k$. Hence, we can apply the inductive hypothesis to $\mathrm{ad}(\mathfrak{l})$, meaning that there is some $\bar{Y} \in \mathfrak{n}/\mathfrak{l}$ such that $\bar{Y} \neq 0$ and $\mathfrak{l} \cdot \bar{Y} = 0$.

Let $Y \in \mathfrak{n}$ be a representative of $\bar{Y}$. That is, $\bar{Y} = Y + \mathfrak{n}$. Then we have $[\mathfrak{l}, Y] \subseteq \mathfrak{l}$ and . This implies that $[\mathfrak{l} \oplus \Bbbk Y, \mathfrak{l} \oplus \Bbbk Y] \subseteq \mathfrak{l} \oplus \Bbbk Y$. This means that $\mathfrak{l} \oplus \Bbbk\mathfrak{l}$ is a subalgebra of $\mathfrak{n}$. Now, $\dim(\mathfrak{l} \oplus \Bbbk Y) = \dim(\mathfrak{l}) + 1$, and so by the assumption that $\mathfrak{l}$ is a *maximal* proper subalgebra we must have that $\mathfrak{l} \oplus \Bbbk Y = \mathfrak{n}$. This proves that $\dim \mathfrak{l} = \dim(\mathfrak{n}) - 1$. It also implies that $\mathfrak{l}$ is an ideal, since all elements in $\mathfrak{n}$ which are not in $\mathfrak{l}$ are just scalar multiples of Y, and we've seen that $[\mathfrak{l}, Y] \in \mathfrak{l}$, so $\mathfrak{l}$ has the required absorbing property. Since $\dim \mathfrak{l} < \dim \mathfrak{n}$ we can apply the induction hypothesis, and we know that there is some $v \in \mathfrak{l}$ with $v \neq 0$ such that $\mathfrak{l} \cdot v = 0$. Let $W \subseteq V$ be the subspace of all vectors with this property. Taking the decomposition $\mathfrak{n} = \mathfrak{l} \oplus \Bbbk Y$ it is sufficient to show that there is some $w \in W$ with $w \neq 0$ such that $Y(w) = 0$, since then we will certainly have $(X + Y)(w) = X(w) + Y(w) = 0 + 0 = 0$ for $X \in \mathfrak{l}$. To show that such a w exists take $X \in \mathfrak{l}$, then since the bracket in $\mathfrak{gl}(V)$ is just the commutator we can evaluate at some $w \in W$ and we have

$$XY(w) = YX(w) + [X, Y](w) = 0, \tag{5.4.23}$$

where we've used the fact that $X, [X, Y] \in \mathfrak{l}$ (since $\mathfrak{l}$ is an ideal) and thus $X(w) = [X, Y](w) = 0$ by definition of W. Therefore, $Y(W) \subseteq W$. Since Y is a nilpotent endomorphism of V and preserves W it restricts to a nilpotent endomorphism of W. Thus, there exists some $w \in W$ with $w \neq 0$ such that $Y(w) = 0$. $\qquad\square$

5.4.1   Corollaries of Engel's Theorem

Before we can prove some corollaries of Engel's theorem we need a definition and a couple of lemmas.

Definition 5.4.24 — Flag Let V be a vector space. A **flag**, $V_\bullet$, in V is an ascending sequence of vector spaces

$$0 = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n = V. \tag{5.4.25}$$

A flag is **complete** if $\dim V_{i+1} = \dim V_i + 1$, or equivalently $\dim(V_{i+1}/V_i) = 1$.

For $\Bbbk^n$ with basis $\{e_1, \dots, e_n\}$ the **standard flag** is the complete flag given by taking $V_0 = 0$ and $V_i = \mathrm{span}_{\Bbbk}\{e_1, \dots, e_i\}$.

Each choice of a basis gives rise to a complete flag, although not uniquely, since $\{e_1, e_2, e_3, \dots, e_n\}$ and $\{e_1, 2e_2, e_3, \dots, e_n\}$ both give rise to the same standard flag.

Lemma 5.4.26 Take $V_{\bullet}$ to be the standard flag in $\mathbb{C}^n$ with basis $\{e_1, \dots, e_n\}$. Then for $X \in \mathrm{GL}(n, \mathbb{C})$ we have

- $X \in \mathfrak{b}(n, \mathbb{C})$ if and only if $X(V_i) \subseteq V_i$; and

- $X \in \mathfrak{n}(n, \mathbb{C})$ if and only if $X(V_i) \subseteq V_{i-1}$.

Proof. This is just restating the definition of the (strictly) upper triangular matrices using the language of flags. We can identify V_i as the space of all column vectors which are zero after the ith row. Then if $v \in V_i$ we know that $X(v) \in V_i$ if and only if $X(v)$ has zeros after the ith row, which is true exactly when X is upper triangular. If we replace X with a strictly upper triangular matrix then $X(v)$ will have zeros after the $(i-1)$st row, and so will be in V_{i-1}. $\square$

We say that $X \in \mathfrak{b}(n, \mathbb{C})$ preserves the flag and $X \in \mathfrak{n}(n, \mathbb{C})$ strictly preserves the flag.

Note that each basis choice gives us a (not necessarily unique) algebra $\mathfrak{b}(n, \mathbb{C})$ (called a Borel subalgebra), of matrices which are upper triangular when expressed in that basis. While this pairing of bases to Borel subalgebras is not a bijection the pairing of standard flags to Borel subalgebras is, since two bases give the same Borel subalgebra if and only if they are related by rescaling the basis elements, if and only if they correspond to the same standard flag.

Corollary 5.4.27 Let V be a finite dimensional vector space with $V \neq 0$. Let $\mathfrak{n}$ be a subalgebra of $\mathfrak{gl}(V)$ such that all elements of $\mathfrak{n}$ are nilpotent endomorphisms of V. Then there exists a basis of V such that all elements of $\mathfrak{n}$ correspond to strictly upper triangular matrices in this basis.

Proof. By Lemma 5.4.26 this result is equivalent to the statement that there exists a complete flag, $V_{\bullet}$, such that $\mathfrak{n}(V_i) \subseteq V_{i-1}$, that is $X(V_i) \subseteq V_{i-1}$ for all $X \in \mathfrak{n}$.

We can prove this by induction on $\dim V$. The base case is $\dim V = 1$. If $X \colon V \to V$ is nilpotent, so that $X^n = 0$ for some $n \in \mathbb{N}$, then in any basis of V we can express X as $[X]_{\mathcal{B}} = (b)$ for some $b \in \mathbb{C}$. Then the nilpotent condition is that $(b)^n = (b^n) = 0$, which is true if and only if $b = 0$, and hence $[X]_{\mathcal{B}} = (0)$, which is strictly upper triangular.

For the inductive step suppose that for any vector space of dimension less than $\dim V$ there is a complete flag which is completely preserved by el-

ements of $\mathfrak{n}$. By Engel's theorem (Theorem 5.4.20) we know that there is
some $v_1 \in V$ such that $v_1 \neq 0$ and $\mathfrak{n}v_1 = 0$ (that is, $X(v_1) = 0$ for all $X \in \mathfrak{n}$).
Consider the flag

$$0 = V_0 \subsetneq V_1 = \mathbb{C}v_1 \subsetneq V. \tag{5.4.28}$$

Note that we're assuming that $\dim V > \dim V_1 = 1$, else we would already
be done by the base case.

We now consider the quotient V/V_1. This has $\dim(V/V_1) = \dim V - \dim V_1 = \dim V - 1 \leq \dim V$, and so the induction hypothesis applies.
This tells us that there is a flag

$$0 = \bar{V}_1 \subsetneq \bar{V}_2 \subsetneq \cdots \subsetneq \bar{V}_{n-1} = V/V_1 \tag{5.4.29}$$

which is strictly preserved by $\mathfrak{n}$, so $\mathfrak{n}\bar{V}_i \subseteq \bar{V}_{i-1}$. We use the preimage of the
$\bar{V}_i$ under the quotient map $V \twoheadrightarrow V/V_1$ to extend our initial flag by defining

$$V_i = \{v \in V \mid v + V_1 \in \bar{V}_i\}. \tag{5.4.30}$$

Note that this agrees with our earlier definition of V_1 since $v + V_1 \in \bar{V}_1 = 0$
if and only if $v \in V_1$. So, we have a complete flag

$$0 = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_{n-1} = V/V_1 \subsetneq V_n = V. \tag{5.4.31}$$

We just need to show that this is strictly preserved by $\mathfrak{n}$. To see this note
that if $X \in \mathfrak{n}$ then $X(\bar{V}_i) \subseteq \bar{V}_{i-1}$. This means that $X(V_i) \subseteq V_{i-1} + V_1 = V_{i-1}$.
Thus, this flag is as required. $\square$

We now have another lemma before we can prove another corollary of Engel's
theorem, which is really a corollary of this corollary.

Lemma 5.4.32 Let $\mathfrak{n}$ be a Lie algebra such that $\mathfrak{n}/\mathfrak{z}(\mathfrak{n})$ is nilpotent. Then $\mathfrak{n}$
is nilpotent.

Proof. For $\mathfrak{g}$ a Lie algebra and $\mathfrak{i} \subseteq \mathfrak{g}$ an ideal we have $(\mathfrak{g}/\mathfrak{i})_n = \mathfrak{g}_n + \mathfrak{i}/\mathfrak{i}$. Taking $\mathfrak{g} = \mathfrak{n}$ and $\mathfrak{i} = \mathfrak{z}(\mathfrak{n})$ we have $(\mathfrak{n}_n + \mathfrak{z}(\mathfrak{n}))/\mathfrak{z}(\mathfrak{n}) = (\mathfrak{n}/\mathfrak{z}(\mathfrak{n}))_n$. By assumption
$\mathfrak{n}/\mathfrak{z}(\mathfrak{n})$ is nilpotent, so there is some $N \in \mathbb{N}$ such that $(\mathfrak{n}/\mathfrak{z}(\mathfrak{n}))_N = 0$. We
therefore have $(\mathfrak{n}_N + \mathfrak{z}(\mathfrak{n}))/\mathfrak{z}(\mathfrak{n}) = 0$, which is only possible if $\mathfrak{n}_N = 0$, which
means that $\mathfrak{n}$ is nilpotent. $\square$

This result is useful because we always have the adjoint action, $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$,
and by definition $\ker \mathrm{ad} = \mathfrak{z}(\mathfrak{g})$. Then using the first isomorphism theorem we
have $\mathrm{ad}(\mathfrak{g}) \cong \mathfrak{g}/\mathfrak{z}(\mathfrak{g})$. This means that we always have an embedding of $\mathfrak{g}/\mathfrak{z}(\mathfrak{g})$ into
$\mathfrak{gl}(\mathfrak{g})$ for any Lie algebra, nilpotent or not. This lets us study $\mathfrak{g}$ by studying $\mathfrak{gl}(\mathfrak{g})$, so
long as we're prepared to set all central elements to zero.

Corollary 5.4.33 Let $\mathfrak{n}$ be a finite dimensional Lie algebra. If every element
of $\mathfrak{n}$ is ad-nilpotent then $\mathfrak{n}$ is nilpotent.

Proof. By Lemma 5.4.32 it is sufficient to prove that $\mathfrak{n}/\mathfrak{z}(\mathfrak{n})$. By the remark following Lemma 5.4.32 we have that $\mathfrak{n}/\mathfrak{z}(\mathfrak{n}) \cong \text{ad}(\mathfrak{n}) \subseteq \mathfrak{gl}(\mathfrak{n})$. By Corollary 5.4.27 there exists a basis for $\mathfrak{n}$ such that elements of $\text{ad}(\mathfrak{n})$ are strictly upper triangular matrices in this basis. This implies that, in this basis, $\text{ad}(\mathfrak{n}) \subseteq \mathfrak{n}(n, \mathbb{C})$, and this implies that $\mathfrak{n}/\mathfrak{z}(\mathfrak{n})$ is nilpotent, so $\mathfrak{n}$ is nilpotent by Lemma 5.4.32. $\qquad \square$

5.5 Lie's Theorem

Given the notation setup in the statement of Engel's theorem (Theorem 5.4.20) we may state the conclusion as "there is a common eigenvector, $v \in V$, for all endomorphisms $X \in \mathfrak{n}$ such that the eigenvalue of v for X is zero". Lie's theorem is the analogue for solvable Lie algebras. It states that there is a common eigenvector, but does not force the eigenvalue to be zero. Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ be some Lie algebra. Then $v \in V$ is an eigenvector of all $X \in \mathfrak{g}$ if there exist scalars $\alpha_X \in \Bbbk$ for each $X \in \mathfrak{g}$ such that $Xv = \alpha_X v$. We can then view this as a map $\alpha : \mathfrak{g} \to \Bbbk, X \mapsto \alpha_X$. Further, this map is linear since α_{X+Y} is the eigenvalue of v under the operation $X + Y$ and $(X + Y)v = Xv + Yv = \alpha_X v + \alpha_Y v$ so we can identify $\alpha_{X+Y} = \alpha_X + \alpha_Y$, and similarly, $(\lambda X)v = \lambda(Xv) = \lambda \alpha_X v$ so $\alpha_{\lambda X} = \lambda \alpha_X$. Thus, $\alpha \in \mathfrak{g}^*$. So, $v \in V$ with $v \neq 0$ is a common eigenvalue of all elements of $\mathfrak{g}$ if and only if there exists some $\alpha \in \mathfrak{g}^*$ such that $Xv = \alpha(X)v$ for all $X \in \mathfrak{g}$.

Theorem 5.5.1 — Lie. Let V be a finite-dimensional vector space over an algebraically closed field of characteristic zero (so basically, $\Bbbk = \mathbb{C}$), $\Bbbk$, with $V \neq 0$, and let $\mathfrak{s}$ a solvable subalgebra of $\mathfrak{gl}(V)$. Then there exists a common eigenvector for all endomorphisms in $\mathfrak{s}$.

Proof. The steps taken in this proof are the same as in Enge's theorem (Theorem 5.4.20), but the justifications for why each step holds differ slightly. The proof is by induction on $\dim \mathfrak{s}$. For the base case, $\dim \mathfrak{s} = 1$, the result is trivial as $\mathfrak{s} = \mathbb{C}X$ for some nonzero endomorphism $X \in \mathfrak{gl}(V)$, and by the assumption that we work over an algebraically closed field X has an eigenvector, $v \in V$ with $v \neq 0$, so $Xv = \alpha v$ for some $\alpha \in \Bbbk$, and this is an eigenvector of all $Y \in \mathfrak{s}$ since $Y = yX$ for some $y \in \Bbbk$, since $Yv = yXv = y\alpha v$, so v is an eigenvector of Y with eigenvalue $y\alpha$.

Now suppose that $\dim \mathfrak{s} > 1$ and that the hypothesis holds for all solvable Lie algebras of dimension less than $\dim \mathfrak{s}$. Since $\mathfrak{s}$ is solvable the derived subalgebra, $[\mathfrak{s}, \mathfrak{s}]$ is either zero or a proper ideal of $\mathfrak{s}$. Importantly, $[\mathfrak{s}, \mathfrak{s}] \neq \mathfrak{s}$, since then the derived series would never terminate. If $[\mathfrak{s}, \mathfrak{s}] = 0$ then $\mathfrak{s}$ is abelian and $\mathfrak{s}/[\mathfrak{s}, \mathfrak{s}] = \mathfrak{s}/0 = \mathfrak{s}$ is abelian. If $[\mathfrak{s}, \mathfrak{s}]$ is a proper ideal of $\mathfrak{s}$ then $\mathfrak{s}/[\mathfrak{s}, \mathfrak{s}]$ is abelian since for all $X, Y \in \mathfrak{s}$ we have

$$[X + \mathfrak{s}', Y + \mathfrak{s}'] = [X, Y] + \mathfrak{s}' = \mathfrak{s}' \qquad (5.5.2)$$

where $\mathfrak{s}' = [\mathfrak{s}, \mathfrak{s}]$ and $[X, Y] \in \mathfrak{s}'$ gives us the final equality. This means that regardless of whether $\mathfrak{s}'$ is zero or a proper ideal $\mathfrak{s}/\mathfrak{s}'$ is abelian and hence every subspace of $\mathfrak{s}/\mathfrak{s}'$ is an ideal.

Take some subspace of codimension 1, that is a subspace with dimension $\dim(\mathfrak{s}/\mathfrak{s}') - 1$. Call this subspace $\bar{\mathfrak{n}}$, and note that this is an ideal of $\mathfrak{s}/\mathfrak{s}'$. Denote by $\mathfrak{n}$ the preimage of $\bar{\mathfrak{n}}$ under the quotient map. This is an ideal of $\mathfrak{s}$ of codimension 1. This holds because the preimage of an ideal under a homomorphism is always an ideal and the projection map is surjective, so the dimensions match up as needed. By induction we may posit that there is some common eigenvector for all elements of $\mathfrak{n}$, since it has dimension strictly less than $\dim \mathfrak{s}$ and is solvable as a subalgebra of a solvable algebra. Then there exists some $\alpha \in \mathfrak{n}^*$ such that $Xv = \alpha(X)v$ for all $X \in \mathfrak{n}$.

Now consider the subspace $W = \{w \in V \mid Xw = \alpha(X)w \forall X \in \mathfrak{n}\}$. This subspace can be thought of as the space of all candidates for a common eigenvector of $\mathfrak{s}$, we just need to find some element of this subspace that is also an eigenvector of some vector spanning the missing dimension that $\mathfrak{s}$ has but $\mathfrak{n}$ doesn't. We know by the previous paragraph that W is nonempty, and thus by the following lemma (Lemma 5.5.3) we know that W is an $\mathfrak{s}$-submodule of V. Choose some $Y \in \mathfrak{s} \setminus \mathfrak{n}$, so $\mathfrak{s} = \mathfrak{n} \oplus \mathbb{C} Y$. Then $Y(W) \subseteq W$, as otherwise W would not be a $\mathfrak{s}$-submodule. Thus, there exists some $w \in W$ with $w \neq 0$ such that w is an eigenvector of Y, and so w is an eigenvalue of all elements of $\mathfrak{s}$. $\square$

Lemma 5.5.3 Work over an algebraically closed field of characteristic zero. Let $\mathfrak{n}$ be an ideal of a Lie algebra $\mathfrak{s}$ and let V be an $\mathfrak{s}$-submodule and $\alpha \in \mathfrak{s}^*$. Let

$$W = \{v \in V \mid Xv = \alpha(X)v \forall X \in \mathfrak{n}\}. \tag{5.5.4}$$

Then W is an $\mathfrak{s}$-submodule of V.

Proof. We need to show that $Y(W) \subseteq W$ for all $Y \in \mathfrak{s}$. If $Y \in \mathfrak{n}$ then this is clearly true since Y just acts on W by scaling by $\alpha(Y)$. We need to show in general that $Yw \in W$ for all $w \in W$. This means we need to show that $X(Yw) = \alpha(X)Yw$ for all $X \in \mathfrak{n}$. We can do this using the definition of how the bracket acts and the fact that w is an X eigenvector:

$$X(Yw) = Y(Xw) + [X, Y]w \tag{5.5.5}$$
$$= Y(\alpha(X)w) + \alpha([X, Y])w \tag{5.5.6}$$
$$= \alpha(X)Yw + \alpha([X, Y])w. \tag{5.5.7}$$

Here we've used the fact that $[X, Y] \in \mathfrak{n}$ since $\mathfrak{n}$ is an ideal of $\mathfrak{s}$, and so the action of the bracket is determined by α. We are then left to show that $\alpha([X, Y]) = 0$.

Let U be the subspace of V spanned by $\{w, Yw, Y^2w, \dots\}$. Then $Y(U)$ is the subspace spanned by $\{Yw, Y^2w, Y^3w, \dots\}$, which is clearly a subspace of U, so $Y(U) \subseteq U$. We claim that U is also an $\mathfrak{n}$-submodule of V. That is, we claim that $X(U) \subseteq U$ for all $X \in \mathfrak{n}$. Clearly we have $X(w) = \alpha(X)w \in U$, and Equation (5.5.7) shows that $X(Yw)$ is a linear combination of Yw and w, so is an element of U. We take this as the base case for an induction proof where we wish now to show that $X(Y^nw) \in U$ for some $n \in \mathbb{N}$ and

we assume that $X(Y^k w) \in U$ for all $k < n$. Then we have

$$XY^n w = YXY^{n-1}w + [X,Y]Y^{n-1}w \tag{5.5.8}$$

and since $\mathfrak{n}$ is an ideal and $X \in \mathfrak{n}$ we have $[X,Y] \in \mathfrak{n}$, so by the induction hypothesis $[X,Y]Y^{n-1}w \in U$. Thus, $XY^n w \in U$ as it's a linear combination of elements of U. This shows that $XY^n w$ is given by $\alpha(X)Y^n w$ plus terms involving $Y^k w$ with $k < n$. Thus, there is a basis for U in which X is upper triangular, with the diagonals given by $\alpha(X)$. In particular, we have that $\operatorname{tr}(X|_U) = \alpha(X)\dim U$. The same is true fo $[X,Y] \in \mathfrak{n}$, $\operatorname{tr}([X,Y]|_U) = \alpha([X,Y])\dim U$. However, the trace of a commutator must vanish by the cyclic property of the trace, and thus we have $\alpha([X,Y]) = 0$. $\qquad\square$

Note that we require characteristic zero to conclude from $0 = \alpha([X,Y])\dim U$ that $\alpha([X,Y]) = 0$, since in a field of characteristic dividing $\dim U$ this will hold with nonzero $\alpha([X,Y])$. We need algebraic closure to guarantee the existence of an eigenvector in the one-dimensional base case. So, Lie's theorem pretty much holds only over $\mathbb{C}$, and a few other more "constructed" fields, such as $\bar{\mathbb{Q}}$, the algebraic closure of $\mathbb{Q}$.

Corollary 5.5.9 There exists a basis of V such that any solvable subalgebra $\mathfrak{s} \subseteq \mathfrak{gl}(V)$ is a subalgebra of $\mathfrak{b}(n, \mathbb{C})$.

Proof. This is the same as Corollary 5.4.27 but with solvable and upper triangular in place of nilpotent and strictly upper triangular. $\qquad\square$

What we have shown with Engel and Lie's theorems is the following:

- A nilpotent subalgebra of $\mathfrak{gl}(V)$ is (up to isomorphism) a subalgebra of $\mathfrak{n}(n, \mathbb{C})$.

- A solvable subalgebra of $\mathfrak{gl}(V)$ is (up to isomorphism) a subalgebra of $\mathfrak{b}(n, \mathbb{C})$.

This is the sense in which $\mathfrak{n}(n, \mathbb{C})$ and $\mathfrak{b}(n, \mathbb{C})$ are the maximal nilpotent and solvable subalgebras of $\mathfrak{gl}(V)$.

Further clarifying nilpotent and solvable subalgebras is hard to impossible, and so we shan't try. We move our focus to the classifiable semisimple Lie algebras.

Six

Killing Form

In this chapter we develop one of the key tools in working with semisimple Lie algebras, the Killing form. Before we define the Killing form we'll have a short introduction to bilinear forms.

6.1 Bilinear Forms

Definition 6.1.1 Let V be a $\Bbbk$-vector space. A **bilinear form**, η, is a map $\eta\colon V{\times}V \to \Bbbk$ which is linear in both arguments. That is, for all $u, v, w \in V$ and $\lambda \in \Bbbk$ we have

$$\eta(u{+}\lambda v, w) = \eta(u,w){+}\lambda\eta(v,w), \qquad \text{and} \qquad \eta(u, v{+}\lambda w) = \eta(u,v){+}\lambda\eta(u,w). \tag{6.1.2}$$

A bilinear form, η, is **symmetric** if for all $u, v \in V$ we have $\eta(u,v) = \eta(v,u)$.

A symmetric bilinear form is **non-degenerate** if for all $u \in V$ with $u \neq 0$ there exists some $v \in V$ such that $\eta(u,v) \neq 0$.

Note that for $\Bbbk = \mathbb{R}$ non-degeneracy is equivalent to requiring that $\eta(u,u) \neq 0$ for all $u \neq 0$. The notion of a non-degenerate bilinear form can also be extended to non-symmetric bilinear forms, but we do not have need of this. We just have to repeat the definition with u and v swapped.

Suppose that $\eta\colon V{\times}V \to \Bbbk$ is a symmetric bilinear form. Then we can partially evaluate at some $u \in V$ and consider a map $\eta(u,-)\colon V \to \Bbbk$ given by $v \mapsto \eta(u,v)$. This is linear since η is linear in the second argument, so $\eta(u,-) \in V^*$. Likewise, $\eta(-,v)\colon V \to \Bbbk$ defined by $u \mapsto \eta(u,v)$ is linear and so $\eta(-,v) \in V^*$. Of course, since η is symmetric these two maps are actually the same.

So, each symmetric[1] bilinear form, η, gives rise to a canonical linear map, $\iota_\eta\colon V \to V^*$, given by $\iota_\eta(u) = \eta(u,-)$. We can use this to restate the non-degeneracy condition as saying that $\ker \iota_\eta = 0$ if and only if η is non-degenerate. To see this first let $v \in \ker \iota_\eta$ so $\iota_\eta(v) = 0$, which means $\eta(v,-)$ is the zero map, so for all $u \in V$ we have $\eta(v,u) = 0$, so if η is non-degenerate we must have $v = 0$ and vice-versa if $\eta(v,-)$ is not the zero map then there exists some $u \in V$ such that $\eta(v,u) \neq 0$, and so η is non-degenerate.

Further, in the finite dimensional case if η is non-degenerate then ι_η is an isomorphism, since we showed above that ι_η is an injection $V \hookrightarrow V^*$, and it is a well-

[1] a non-symmetric bilinear form gives rise to two such maps, and we have to choose one, so we lose the "canonical" part of this map.

known fact that $\dim V = \dim V^*$ for finite-dimensional V, and so an injective linear map between spaces of the same dimension is automatically an isomorphism.

So, a non-degenerate bilinear form gives a canonical isomorphism $V \cong V^*$. Note that we always have *an* isomorphism $V \cong V^*$, it's just that in general there are many such isomorphisms and no reason to prefer one over another. This particular canonical isomorphism is important when our space is a Hilbert space, in which case this result is called the riesz representation theorem.

> **Definition 6.1.3 — Radical** Let η be a bilinear form on V. The **radical** is $\operatorname{rad}\eta = \ker \iota_\kappa \subseteq V$.

6.2 Killing Form

> **Definition 6.2.1** Let $\mathfrak{g}$ be a Lie algebra. Then the **Killing form** on $\mathfrak{g}$ is the symmetric bilinear form
>
> $$\kappa \colon \mathfrak{g} \times \mathfrak{g} \to \Bbbk \tag{6.2.2}$$
> $$(X, Y) \mapsto \kappa(X, Y) = \operatorname{tr}(\operatorname{ad}_X \circ \operatorname{ad}_Y). \tag{6.2.3}$$

Note that symmetry of the Killing form follows immediately from the cyclic property of the trace, namely that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, and linearity follows immediately from the linearity of the trace and composition of endomorphisms.

The most important property of the Killing form is arguably that it gives us a way to show that a finite-dimensional Lie algebra is semisimple. Recall that a Lie algebra is semisimple if it contains no proper solvable ideals.

> **Lemma 6.2.4 — Associativity of the Killing Form** Let $\mathfrak{g}$ be a Lie algebra and κ the Killing form on $\mathfrak{g}$, then
>
> $$\kappa(X, [Y, Z]) = \kappa([X, Y], Z) \tag{6.2.5}$$
>
> for all $X, Y, Z \in \mathfrak{g}$.
>
> *Proof.* This follows by a quick calculation using the fact that ad is a representation and some properties of the trace:
>
> $$\begin{aligned} \kappa(X, [Y, Z]) &= \operatorname{tr}(\operatorname{ad}_X \circ \operatorname{ad}_{[Y,Z]}) & (6.2.6)\\ &= \operatorname{tr}(\operatorname{ad}_X \circ [\operatorname{ad}_Y, \operatorname{ad}_Z]) & (6.2.7)\\ &= \operatorname{tr}(\operatorname{ad}_X \circ (\operatorname{ad}_Y \circ \operatorname{ad}_Z - \operatorname{ad}_Z \circ \operatorname{ad}_Y)) & (6.2.8)\\ &= \operatorname{tr}(\operatorname{ad}_X \circ \operatorname{ad}_Y \circ \operatorname{ad}_Z - \operatorname{ad}_X \circ \operatorname{ad}_Z \circ \operatorname{ad}_Y) & (6.2.9)\\ &= \operatorname{tr}(\operatorname{ad}_X \circ \operatorname{ad}_Y \circ \operatorname{ad}_Z - \operatorname{ad}_Y \circ \operatorname{ad}_X \circ \operatorname{ad}_Z) & (6.2.10)\\ &= \operatorname{tr}((\operatorname{ad}_X \circ \operatorname{ad}_Y - \operatorname{ad}_Y \circ \operatorname{ad}_X) \circ \operatorname{ad}_Z) & (6.2.11)\\ &= \operatorname{tr}([\operatorname{ad}_X, \operatorname{ad}_Y] \circ \operatorname{ad}_Z) & (6.2.12)\\ &= \operatorname{tr}(\operatorname{ad}_{[X,Y]} \circ \operatorname{ad}_Z) & (6.2.13)\\ &= \kappa([X, Y], Z). & \square \end{aligned}$$

> **Theorem 6.2.14 — Cartan's First Criterion.** A finite-dimensional Lie alge-
> bra, $\mathfrak{g}$, over $\mathbb{C}$ is solvable if and only if $\kappa(X, Y) = 0$ for all $X \in \mathfrak{g}'$ and $Y \in \mathfrak{g}$.

Proof. Suppose that $\mathfrak{g}$ is solvable. Then Lie's theorem tells us that there is a
basis, $\mathcal{B}$, for $\mathfrak{g}$ such that if $W \in \mathfrak{g}$ then $[\mathrm{ad}_W]_{\mathcal{B}}$ is upper triangular. This im-
plies that $[[\mathrm{ad}_X, \mathrm{ad}_Y]]_{\mathcal{B}} = [[\mathrm{ad}_X]_{\mathcal{B}}, [\mathrm{ad}_Y]_{\mathcal{B}}]$ is *strictly* upper triangular for
$X, Y \in \mathfrak{g}$, since the commutator of upper triangular matrices is strictly up-
per triangular. This then implies that $\kappa([X, Y], Z) = \mathrm{tr}([\mathrm{ad}_X, \mathrm{ad}_Y] \circ \mathrm{ad}_Z) =$
0 using similar calculations to Lemma 6.2.4. Note then that $[X, Y]$ is exactly
an element of $\mathfrak{g}'$ and Z an element of $\mathfrak{g}$.

Now suppose that $X, Y, Z \in \mathfrak{g}$ are such that $\kappa([X, Y], Z) = 0$. If $L' = [\mathfrak{g}, \mathfrak{g}]$
is nilpotent then L is solvable by Engel's theorem. Notice that if $W =$
$[X, Y] \in \mathfrak{g}'$ then $\kappa(W, Z) = 0$. We can now use the Jordan decomposition
to write

$$\mathrm{ad}_W = (\mathrm{ad}_W)_{\mathrm{d}} + (\mathrm{ad}_W)_{\mathrm{n}} =: W_{\mathrm{d}} + W_{\mathrm{n}} \tag{6.2.15}$$

where W_{d} is diagonalisable and W_{n} is nilpotent, meaning that in some basis
W_{n} is upper triangular, and so has vanishing trace. Then we have

$$\mathrm{tr}(W\bar{W}_{\mathrm{d}}) = \mathrm{tr}((W_{\mathrm{d}} + W_{\mathrm{n}})\bar{W}_{\mathrm{d}}) = \mathrm{tr}(W_{\mathrm{d}}\bar{W}_{\mathrm{d}}) + \mathrm{tr}(W_{\mathrm{n}}\bar{W}_{\mathrm{d}}) = \mathrm{tr}(W_{\mathrm{d}}\bar{W}_{\mathrm{d}}) \tag{6.2.16}$$

where we've used the fact that the product $W_{\mathrm{n}}\bar{W}_{\mathrm{d}}$ will be strictly upper tri-
angular in whichever basis W_{n} is strictly upper triangular in, and as such
has trace zero. Writing λ_i with $i = 1, \dots, m$ for the eigenvalues of ad_W we
have

$$\mathrm{tr}(W\bar{W}_{\mathrm{d}}) = \mathrm{tr}(W_{\mathrm{d}}\bar{W}_{\mathrm{d}}) = |\lambda_1|^2 + \cdots + |\lambda_m|^2 \tag{6.2.17}$$

since W_{d} and $\bar{W}_{\mathrm{d}}$ are diagonal with entries λ_i and $\bar{\lambda}_i$ as the ith diagonal entry
respectively. We also have

$$\mathrm{tr}(W\bar{W}_{\mathrm{d}}) = \mathrm{tr}(\mathrm{ad}_W \circ (\mathrm{ad}_{\bar{W}})_{\mathrm{d}}) \tag{6.2.18}$$
$$= \mathrm{tr}(\mathrm{ad}_{[X,Y]}(\mathrm{ad}_{\bar{W}})_{\mathrm{d}}) \tag{6.2.19}$$
$$= \mathrm{tr}(\mathrm{ad}_X \circ (\mathrm{ad}_{[Y,\bar{W}]})_{\mathrm{d}}) \tag{6.2.20}$$
$$= 0 \tag{6.2.21}$$

where we've used similar calculations to the proof of Lemma 6.2.4 in the
last step.

So, we must have that $|\lambda_1|^2 + \cdots + |\lambda_m|^2 = 0$, which implies that $\lambda_i = 0$
for all $i = 1, \dots, m$. Thus, $W_{\mathrm{d}} = 0$ so $\mathrm{ad}_W = W_{\mathrm{n}}$, and hence W is nilpotent,
because this shows it is ad-nilpotent so we can apply Corollary 5.4.33. Since
W was chosen as an arbitrary element of $\mathfrak{g}'$ this shows that $\mathfrak{g}'$ is nilpotent,
and hence that $\mathfrak{g}$ is solvable by Engel's theorem. $\square$

Lemma 6.2.22 Every nonzero solvable ideal of a finite dimensional Lie algebra, $\mathfrak{g}$, contains a nonzero abelian ideal of $\mathfrak{g}$.

Proof. Let $\mathfrak{l}$ be a solvable ideal of $\mathfrak{g}$ such that $\mathfrak{l}^n = 0$. Consider $\mathfrak{l}^{n-1}$. This is such that $[\mathfrak{l}^{n-1}, \mathfrak{l}^{n-1}] = \mathfrak{l}^n = 0$, and so $\mathfrak{l}^{n-1}$ is abelian, and it's also an ideal. $\qquad\square$

Lemma 6.2.23 Let $\mathfrak{s}$ be a Lie algebra. Then $\mathfrak{s}$ is solvable if and only if $\mathrm{ad}(\mathfrak{s}) \subseteq \mathfrak{gl}(\mathfrak{s})$ is solvable.

Proof. Suppose $\mathrm{ad}(\mathfrak{s})$ is solvable. We have seen that $\mathrm{ad}(\mathfrak{s}) \cong \mathfrak{s}/\mathfrak{z}(\mathfrak{s})$. The centre, $\mathfrak{z}(\mathfrak{s})$ is abelian, and thus solvable, and this isomorphism tells us that $\mathfrak{s}/\mathfrak{z}(\mathfrak{s})$ is solvable, and so it follows by Lemma 5.3.14 that $\mathfrak{s}$ is solvable. $\quad\square$

Lemma 6.2.24 Let $\mathfrak{g}$ be a Lie algebra and let $\mathfrak{s} = \mathrm{rad}\,\kappa$ be the radical of the Killing form. Then $\mathfrak{s}$ is a solveable ideal of $\mathfrak{g}$.

Proof. Recall that $\mathfrak{s} = \mathrm{rad}\,\kappa = \ker \iota_\kappa$ where $\iota_\kappa : \mathfrak{g} \to \mathfrak{g}^*$ is the map $X \mapsto \kappa(X, -)$. Suppose that $X \in \mathfrak{s}$ and $Y \in \mathfrak{g}$. Then we want to show that $[Y, X] \in \mathfrak{s}$. That is, we want to show that $\kappa([Y, X], -)$ is the zero map, which means showing that $\kappa([Y, X], Z) = 0$ for all $Z \in \mathfrak{g}$. This is true by Lemma 6.2.4 because we have

$$\kappa([Y, X], Z) = \kappa(Y, [X, Z]) = 0 \tag{6.2.25}$$

with the last equality following because $\mathfrak{s}$ is solvable, which follows because $\kappa(X, X') = 0$ for $X, X' \in \mathfrak{s}$ proving solvability by Cartan's first criterion (Theorem 6.2.14). $\qquad\square$

Theorem 6.2.26 — Cartan's Second Criterion. A Lie algebra, $\mathfrak{g}$, over $\mathbb{C}$ is semisimple if and only if the Killing form is non-degenerate.

Proof. Suppose that $\mathfrak{g}$ is semisimple, then we want to show that κ is non-degenerate. Since $\mathfrak{g}$ is semisimple it has no solveable ideals, and by Lemma 6.2.24 is a solvable ideal, which means that we must have $\mathfrak{s} = 0$, and thus κ is non-degenerate as this shows that ι_κ is an injection so the discussion following Definition 6.1.1 applies.

Now suppose that κ is non-degenerate. Then $\mathfrak{s} = 0$. Let $I \subseteq \mathfrak{g}$ be a solveable ideal. By ... we know that I contains an abelian ideal, $\mathfrak{l}$. We will show that $\mathfrak{l} \subseteq \mathfrak{s}$ and thus $\mathfrak{l} = 0$ and so $I = 0$. Take elements $X \in \mathfrak{l}$ and $Y \in \mathfrak{g}$. We want to show that $\kappa(X, Y) = 0$, since this implies that $\mathfrak{l} \subseteq \mathfrak{s}$. To do this we consider the map $\mathrm{ad}_X \circ \mathrm{ad}_Y : \mathfrak{g} \to \mathfrak{g}$, the trace of which is $\kappa(X, Y)$. For any $Z \in \mathfrak{g}$ we always have that $[X, [Y, Z]] \in \mathfrak{l}$ because $X \in \mathfrak{l}$ and $\mathfrak{l}$ is an ideal.

Further, since $\mathfrak{l}$ is abelian we have

$$[X, [Y, [X, [Y, Z]]]] = 0 \tag{6.2.27}$$

since $X \in \mathfrak{l}$ and $[Y, [X, [Y, Z]]] \in \mathfrak{l}$ (since $[X, [Y, Z]] \in \mathfrak{l}$ and $\mathfrak{l}$ is an ideal), so this is indeed the bracket of two elements of the abelian ideal $\mathfrak{l}$ and so vanishes. We can then identify the operator $\mathrm{ad}_X \circ \mathrm{ad}_Y$ in the above, as what we have here is

$$(\mathrm{ad}_X \circ \mathrm{ad}_Y)^2(Z) = 0. \tag{6.2.28}$$

This holds for all $Z \in \mathfrak{g}$ and so we can conclude that $(\mathrm{ad}_X \circ \mathrm{ad}_Y)^2 = 0$. So, $\mathrm{ad}_X \circ \mathrm{ad}_Y$ is a nilpotent operator. This means that $\mathrm{ad}_X \circ \mathrm{ad}_Y$ has trace zero, since this is true of all nilpotent operators, and so we are done. This general fact follows because nilpotent operators are strictly upper triangular in some basis, and computing the trace in this basis must give zero, and trace is basis independent. An alternative proof of this general fact is that the minimal polynomial of $T \colon V \to V$ if $T^n = 0$ (and n is minimal in this respect) is $m(t) = t^n$, and the roots of this are zero, and the roots of the minimal polynomial are the eigenvalues (without multiplicity) of the operator, and the trace is the sum of the eigenvalues (with multiplicity), and so here all the eigenvalues are zero and so their sum vanishes. $\square$

Theorem 6.2.29. Let $\mathfrak{g}$ be a finite dimensional semisimple Lie algebra. Then there is a unique (up to isomorphisms of summands and the order of terms in the sum) decomposition

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_k \tag{6.2.30}$$

with $\mathfrak{g}_i$ simple Lie algebras.

Proof. STEP 1: EXISTENCE
Suppose $\mathfrak{g}$ is a semisimple Lie algebra. If such a decomposition exists then the $\mathfrak{g}_i$ will be ideals of $\mathfrak{g}$ which pairwise commute, that is $[\mathfrak{g}_i, \mathfrak{g}_j] = 0$ for $i \neq j$. Another way to put this is that if $i \neq j$ then $\mathrm{ad}(\mathfrak{g}_i)$ vanishes on $\mathfrak{g}_j$ and vice versa. Very concretely, ad_X, for $X \in \mathfrak{g}_i$, will admit a basis such that it is block diagonal and the blocks corresponding to $\mathfrak{g}_j$ $(j \neq i)$ will be zero. In other words, the subspace $\mathfrak{g}_i$ will be perpendicular to the subspace $\mathfrak{g}_j$ with respect to the Killing form, κ.
This gives us a way to start to construct such a decomposition. Let $\mathfrak{l}$ be some proper ideal, that is $\mathfrak{l} \neq \mathfrak{g}, 0$. Define

$$\mathfrak{l}^\perp := \{X \in \mathfrak{g} \mid \kappa(X, Y) = 0 \, \forall Y \in \mathfrak{l}\}. \tag{6.2.31}$$

This is the space of all X which are perpendicular to all elements of $\mathfrak{l}$. This is an ideal, since if $X \in \mathfrak{l}^\perp$, $Y \in \mathfrak{l}$, and $Z \in \mathfrak{g}$ then

$$\kappa([X, Z], Y) = \kappa(X, [Z, Y]) = 0 \tag{6.2.32}$$

having used the associativity of the Killing form (Lemma 6.2.4) and then the fact that $\mathfrak{l}$ is an ideal and $Y \in \mathfrak{l}$ so $[Z, Y] \in \mathfrak{l}$ and since $X \in \mathfrak{l}^{\perp}$ it must be that $\kappa(X, [Z, Y])$ vanishes.

The intersection of ideals is again an ideal, and so $\mathfrak{l} \cap \mathfrak{l}^{\perp}$ is an ideal of $\mathfrak{g}$. Further, κ vanishes on $\mathfrak{l} \cap \mathfrak{l}^{\perp}$ since if $X, Y \in \mathfrak{l} \cap \mathfrak{l}^{\perp}$ then, in particular, $X \in \mathfrak{l}^{\perp}$ and $Y \in \mathfrak{l}$ so by the definition of $\mathfrak{l}^{\perp}$ we must have that $\kappa(X, Y) = 0$. This implies that $\mathfrak{l} \cap \mathfrak{l}^{\perp}$ is solveable by Cartan's criterion (Theorem 6.2.26). ∗∗∗ The semisimplicity of $\mathfrak{g}$ therefore requires that $\mathfrak{l} \cap \mathfrak{l}^{\perp} = 0$, note that we can't have $\mathfrak{l} \cap \mathfrak{l}^{\perp} = \mathfrak{g}$ because $\mathfrak{l} \subsetneq \mathfrak{g}$. From this we can see that $\mathfrak{l} + \mathfrak{l}^{\perp} = \mathfrak{l} \oplus \mathfrak{l}^{\perp}$.

We claim that $\mathfrak{l} \oplus \mathfrak{l}^{\perp} = \mathfrak{g}$. This follows by an argument on dimensions. Clearly, $\mathfrak{l} \oplus \mathfrak{l}^{\perp} \subseteq \mathfrak{g}$, and so we get the result if we can show that $\dim \mathfrak{l} + \dim \mathfrak{l}^{\perp} = \dim \mathfrak{g}$. We get this in turn by noting that $\mathfrak{l}^{\perp} = \ker(\iota_{\kappa}|_{\mathfrak{l}})$ and that $\iota_{\kappa}|_{\mathfrak{l}} \in \mathfrak{l}^*$. Since all of our vector spaces are finite dimensional we have $\dim \mathfrak{l}^* = \dim \mathfrak{l}$, thus we have that $\dim \mathfrak{l} + \dim \mathfrak{l}^{\perp} = \dim \mathfrak{l}^* + \dim \mathfrak{l}^{\perp} = \dim \mathfrak{g}^*$, which follows since any linear map $\mathfrak{g} \to \mathbb{C}$ may be written as a linear map on $\mathfrak{l}$ plus a linear map on $\mathfrak{g}$ which vanishes on $\mathfrak{l}$, which is exactly what $\mathfrak{l}^* + \mathfrak{l}^{\perp}$ gives us. Finally, we have $\dim \mathfrak{g}^* = \dim \mathfrak{g}$, and so $\mathfrak{l} \oplus \mathfrak{l}^{\perp} = \mathfrak{g}$.

If $\mathfrak{l}$ and $\mathfrak{l}^{\perp}$ are simple we are done. If this is not the case then we can recurse on any non-simple term. Note that $\mathfrak{l}$ and $\mathfrak{l}^{\perp}$ are ideals of a semisimple Lie algebra and so are themselves semisimple.

The dimensions of $\mathfrak{l}$ and $\mathfrak{l}^{\perp}$ are strictly smaller than the dimension of $\mathfrak{g}$. Because the dimensions are finite decrease with each recursion this process must terminate eventually. It will terminate only when all terms appearing in the sums are simple, and then $\mathfrak{g}$ will be a sum of all of these simple Lie algebras.

STEP 2: UNIQUENESS

Suppose that $\mathfrak{h}$ is a simple ideal of $\mathfrak{g}$. We claim that $\mathfrak{h} = \mathfrak{g}_i$ for some i. The space $[\mathfrak{g}, \mathfrak{h}]$ is an ideal of $\mathfrak{h}$. It is nonzero since the centre of $\mathfrak{h}$ is trivial. Since $\mathfrak{h}$ is assumed to be simple it must then be that $[\mathfrak{g}, \mathfrak{h}] = \mathfrak{h}$. Now consider the decomposition

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_k. \tag{6.2.33}$$

Taking the bracket with $\mathfrak{h}$ we have

$$[\mathfrak{g}, \mathfrak{h}] = [\mathfrak{g}_1, \mathfrak{h}] \oplus [\mathfrak{g}_2, \mathfrak{h}] \oplus \cdots \oplus [\mathfrak{g}_k, \mathfrak{h}] = \mathfrak{h}. \tag{6.2.34}$$

Since each bracket is either 0 or $\mathfrak{h}$ by the same logic as above there must be a unique i such that $[\mathfrak{g}_i, \mathfrak{h}] = \mathfrak{h}$. Since each $\mathfrak{g}_i$ is simple and $[\mathfrak{g}_i, \mathfrak{h}] = \mathfrak{h} \neq 0$ we must have that $\mathfrak{h} = [\mathfrak{g}_i, \mathfrak{h}] = \mathfrak{g}_i$. Thus, every ideal of $\mathfrak{g}$ corresponds uniquely to some $\mathfrak{g}_i$, and clearly the $\mathfrak{g}_i$ are ideals of $\mathfrak{g}$, and so the decomposition is uniquely determined by the simple ideals of $\mathfrak{g}$, again, up to order in the sum and isomorphism of each $\mathfrak{g}_i$. □

The idea of this result is that a semisimple Lie algebra is determined by its simple ideals. Further, we've shown that if $\mathfrak{g}$ is semisimple then $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$. In particular, this means that semisimple Lie algebras are neither solveable or nilpotent.

Seven

Complete Reducibility

7.1 Definition and Statement of Weyl's Theorem

Definition 7.1.1 — Complete Reducibility Let $\mathfrak{g}$ be a Lie algebra over $\Bbbk$. A $\mathfrak{g}$-module, V, is **completely reducible** if there exists simple $\mathfrak{g}$-modules V_i such that

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_n \tag{7.1.2}$$

for some $n \in \mathbb{N}$.

Note that in such a decomposition $X \in \mathfrak{g}$ acts on $\bigoplus_i V_i$ by

$$X(v_1, v_2, \ldots, v_n) = (Xv_1, \ldots, Xv_n) \tag{7.1.3}$$

where the action Xv_i is that of the $\mathfrak{g}$-module structure of V_i.

There is an obvious inclusion

$$V_i \hookrightarrow \bigoplus_{i=1}^{n} V_i \tag{7.1.4}$$

$$v \mapsto (0, \ldots, 0, \underset{i\text{th place}}{\underline{v}}, 0, \ldots, 0). \tag{7.1.5}$$

We write $v_1 + \cdots + v_n$ for $(v_1, \ldots, v_n)$. With this the above inclusion is then even more obvious, it's just $v \mapsto 0 + \cdots + 0 + v + 0 + \cdots + 0 = v$, where the zeros are implicit in the final expression.

Note that, unlike the decomposition of a semisimple Lie algebra, the decomposition of a completely reducible $\mathfrak{g}$-module need not be unique. For example, if we take $\mathfrak{g} = 0$ then any vector space is trivially a $\mathfrak{g}$-module, and all subspaces of dimension 1 are simple submodules. Then any decomposition of a vector space is a decomposition into simple $\mathfrak{g}$-modules, and such decompositions are not unique (for dimension at least 2).

Our goal of this chapter is to prove the following theorem, but before we can do that we'll spend some time developing a few lemmas and some extra theory.

Theorem 7.1.6 — Weyl's Theorem. Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{C}$. Then every finite-dimensional representation of $\mathfrak{g}$ is completely reducible.

This result is as strong as possible. If $\mathfrak{g}$ is not semisimple then this result doesn't hold. If the representation is not finite dimensional then this result doesn't hold. For this second case simply take any infinite dimensional vector space as a 0-module, such a space does not have a decomposition as a *finite* sum of one-dimensional spaces, so is not completely reducible.

Note that since we can always decompose a semisimple Lie algebra, $\mathfrak{g}$, as a direct sum of simple Lie algebras it is sufficient to prove Weyl's theorem for a simple Lie algebra.

7.2 Complements

Lemma 7.2.1 A finite dimensional $\mathfrak{g}$-module, V, is completely reducible if and only if every submodule, $W \subseteq V$, admits a complement, W', which is a submodule $W' \subseteq V$ such that $V = W \oplus W'$.

Proof. STEP 1: COMPLETE REDUCIBILITY IMPLIES COMPLEMENTS
Suppose that V is a completely reducible $\mathfrak{g}$-module. Then we can write $V = V_1 \oplus \cdots \oplus V_n$. Let $W \subseteq V$ be a submodule of V. Then W has the decomposition

$$W = (W \cap V_1) \oplus \cdots \oplus (W \cap V_n). \tag{7.2.2}$$

Since each V_i is a simple module and $W \cap V_i \subseteq V_i$ is a submodule of V_i we have that either $W \cap V_i = 0$ or $W \cap V_i = V_i$. Let $\{i_1, \ldots, i_k\}$ be the set of indices for which $W \cap V_i = V_i$, then we can write the decomposition of W as

$$W = V_{i_1} \oplus \cdots \oplus V_{i_k}. \tag{7.2.3}$$

Let $\{j_1, \ldots, j_\ell\}$ be the set of indices for which $W \cap V_j = 0$, and let

$$W' = V_{j_1} \oplus \cdots \oplus V_{j_\ell}. \tag{7.2.4}$$

Then we have

$$W \oplus W' = V_{i_1} \oplus \cdots \oplus V_{i_k} \oplus V_{j_1} \oplus \cdots \oplus V_{j_\ell} = V \tag{7.2.5}$$

and so W' is the complement of W.

STEP 2: COMPLEMENTS IMPLY COMPLETE REDUCIBILITY
Suppose that V is a $\mathfrak{g}$-module such that every submodule, $W \subseteq V$, admits a complement, W'. If V is simple then it is trivially a direct sum of simple modules, so we're done. If V is not simple then we induct on $\dim V$. For $\dim V = 1$ then V is simple and so we're done. Now suppose that having complements implies complete reducibility for all $\mathfrak{g}$-modules of dimension less than $\dim V$. By assumption V is not simple, so we may find some proper submodule $W \subsetneq V$. Take $W' \subseteq V$ such that $W \oplus W' = V$, such a complement exists by assumption. Since W was chosen to be proper we

> know that $\dim W, \dim W' < \dim V$, and thus W and W' both have decompositions into sums of simple $\mathfrak{g}$-modules,
>
> $$W = W_1 \oplus \cdots \oplus W_n, \quad \text{and} \quad W' = W_1' \oplus \cdots \oplus W_m'. \qquad (7.2.6)$$
>
> Note that W_i' need not be the complement to W_1. Then we have
>
> $$V = W \oplus W' = W_1 \oplus \cdots \oplus W_n \oplus W_1' \oplus \cdots \oplus W_m', \qquad (7.2.7)$$
>
> and so V is a direct sum of simple $\mathfrak{g}$-modules as required. $\qquad \square$

This result fails when $\mathfrak{g}$ is not semisimple. For example, taking $\mathfrak{g}$ to be the one-dimensional Lie algebra spanned by X we can define a two-dimensional $\mathfrak{g}$-module $V = \text{span}\{e_1, e_2\}$ by $Xe_1 = 0$ and $Xe_1 = 0$. Then $W = \mathbb{C}e_1$ is a $\mathfrak{g}$-submodule of V, since it's invariant under the action of X (just mapping everything in W to $0 \in W$). Any complementary subspace to W as a vector space must be of the form $W' = \mathbb{C}(\alpha e_1 + e_2)$ for some fixed value of $\alpha \in \mathbb{C}$ (we are taken the span of $\alpha e_1 + e_2$, so we're free to choose the overall scaling to have the coefficient of e_2 be 1). This is the only way that $W + W'$ can contain e_2. However, $X(\alpha e_1 + e_2) = \alpha Xe_1 + Xe_2 = 0 + e_1 = e_1$, and so W' is not a $\mathfrak{g}$-submodule.

Another example of this failing is when we take $\mathfrak{g} = \mathfrak{n}(3, \mathbb{C})$, which is not semisimple, and we take $V = \text{span}\{e_1, e_2, e_3\}$ to be the standard representation of a subalgebra of $\mathfrak{gl}(3, \mathbb{C})$, so e_i are just the standard column vector basis vectors. Elements of $\mathfrak{n}(3, \mathbb{C})$ are of the form

$$\begin{bmatrix} 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{bmatrix}. \qquad (7.2.8)$$

The fact that the first column of any nilpotent matrix must be zero means that any $X \in \mathfrak{n}(3, \mathbb{C})$ acts trivially on e_1, we have $Xe_1 = 0$. This means that $W = \mathbb{C}e_1$ is a submodule of V, since every element of W just maps to $0 \in W$. However, there is no $\mathfrak{g}$-submodule, W', such that $V = W \oplus W'$, since if such a submodule existed we could take $u = \alpha e_1 + \beta e_2 + \gamma e_3 \in W'$ and then since $W \cap W' = 0$ we must have that either β or γ is non-zero in any such vector. For non-zero β we have

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} \beta \\ 0 \\ 0 \end{bmatrix} = \beta e_1 \qquad (7.2.9)$$

and for non-zero γ we have

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} \gamma \\ 0 \\ 0 \end{bmatrix} = \gamma e_1. \qquad (7.2.10)$$

For either case we have left W' and ended up with an element of W, and as such it cannot be that $W \oplus W' = V$.

7.3 Faithful Modules

Definition 7.3.1 Let (V, ρ) be a representation of a Lie algebra, $\mathfrak{g}$. We say that this is a **faithful representation** or that the corresponding $\mathfrak{g}$-module, V, is a **faithful module** if $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is injective.

The idea of a faithful representation is that in mapping $\mathfrak{g}$ to a Lie-algebra of operators on V we don't lose any information by mapping two elements of $\mathfrak{g}$ to the same operator on V.

Lemma 7.3.2 Let $\mathfrak{g}$ be a simple Lie algebra and V a $\mathfrak{g}$-module. Then V is either a faithful module, or a trivial module.

Proof. Let (V, ρ) be an irreducible representation (that is, V is a simple $\mathfrak{g}$-module with the action given by ρ). Then $\ker \rho \subseteq \mathfrak{g}$ is an ideal of $\mathfrak{g}$ since $\ker \rho$ is a subspace of $\mathfrak{g}$ by basic linear algebra, and if $X \in \ker \rho$ and $Y \in \mathfrak{g}$ we have

$$\rho([X, Y])(v) = [\rho(X), \rho(Y)](v) = [0, \rho(Y)](v) = 0(v) = 0 \qquad (7.3.3)$$

for all $v \in V$ and so $[X, Y] \in \ker \rho$.

Since $\mathfrak{g}$ is assumed to be simple we must therefore have either $\ker \rho = 0$, in which case ρ is injective, or $\ker \rho = \mathfrak{g}$, in which case ρ is the zero map. This corresponds to V being a faithful or trivial $\mathfrak{g}$-module respectively. $\quad\square$

Corollary 7.3.4 If $\mathfrak{g}$ is a simple Lie algebra and V is a non-trivial $\mathfrak{g}$-module then $\dim V \geq \sqrt{\dim \mathfrak{g}}$.

Proof. Let (ρ, V) be the corresponding representation, which must be faithful, since we assume it is not trivial. Then $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is injective, and so $\mathfrak{g}$ may be realised as a subalgebra of $\mathfrak{gl}(V)$, and $\dim \mathfrak{gl}(V) = (\dim V)^2$ and since a subalgebra must have at most the dimension of the full algebra we find that $\dim \mathfrak{g} \leq (\dim V)^2$. $\quad\square$

For example, any non-trivial $\mathfrak{sl}(n, \mathbb{C})$ representation must have dimension at least $\sqrt{n^2 - 1}$.

The intuition here is that a faithful module (which any non-trivial module of a simple Lie algebra is) must have "enough room" to fit all of the elements of $\mathfrak{g}$ in, and it turns out that enough room is $\sqrt{\dim \mathfrak{g}}$ dimensions. My first assumption would be that enough room would be $\dim \mathfrak{g}$ dimensions, but the bracket relations between elements of $\mathfrak{g}$ reduces the degrees of freedom required, and so we don't quite need one dimension for each dimension of $\mathfrak{g}$.

7.4 Casimir Operators

In this section I diverge slightly from the material of the lectures. This course never introduces the notion of the universal enveloping algebra, which I believe is a mistake, and so I will take a slightly more general approach here of introducing this universal enveloping algebra and then a general Casimir element, of which the quadratic

Casimir is just one example, before restricting to representations of Casimir elements, which brings us back in line with the content of the course.

> **Definition 7.4.1 — Enveloping Algebra** Let $\mathfrak{g}$ be a Lie algebra over $\Bbbk$. An **enveloping algebra** of $\mathfrak{g}$ is a pair (φ, U) where U is a unital associative algebra over $\Bbbk$ and $\varphi : \mathfrak{g} \to U$ is a Lie algebra homomorphism, where the Lie algebra structure of U is that of the commutator, $[X, Y] = XY - YX$ for $X, Y \in U$.

Note the similarity to the definition of a representation, (ρ, V), where $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is a Lie algebra homomorphism, where $\mathfrak{gl}(V)$ is the Lie algebra of $\mathrm{End}\, V$ equipped with the commutator as a bracket. What this means is that representations, (ρ, V), correspond to enveloping algebras $(\rho, \mathrm{End}\, V)$.

There is a distinguished enveloping algebra for each Lie algebra, called the universal enveloping algebra. As the name suggests it may be defined via a universal property. It is, in a sense, the most universal enveloping algebra because any other enveloping algebra factors through it.

> **Definition 7.4.2 — Universal Enveloping Algebra** Let $\mathfrak{g}$ be a Lie algebra. The **universal enveloping algebra** of $\mathfrak{g}$ is the enveloping algebra $(\Phi, \mathcal{U}(\mathfrak{g}))$ satisfying the universal property that for any enveloping algebra, (φ, U), of $\mathfrak{g}$ there is a unique homomorphism of associative algebras $f : \mathcal{U}(\mathfrak{g}) \to U$ such that $\varphi = f \circ \Phi$. That is, the following diagram commutes
>
> $$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\ \varphi\ } & U \\ {\scriptstyle \Phi}\downarrow & \nearrow{\scriptstyle \exists! f} & \\ \mathcal{U}(\mathfrak{g}) & & \end{array} \qquad (7.4.3)$$
>
> commutes.

Note that the above definition doesn't really tell us what $\mathcal{U}(\mathfrak{g})$ is, just what it does in relation to other enveloping algebras. In fact, this definition doesn't even guarantee the existence of such an object. This is usually the case with objects defined via a universal property. We shall now prove that the universal enveloping algebra always exists and is unique (up to isomorphism) when it does.

> **Lemma 7.4.4** The universal enveloping algebra exists for any Lie algebra.
>
> *Proof.* Let $T(\mathfrak{g})$ be the tensor algebra of $\mathfrak{g}$. This may variously be defined as the free unital associative algebra on some basis, $\{e_1, e_2, \dots\}$ of $\mathfrak{g}$, or by
>
> $$T(\mathfrak{g}) = \Bbbk \oplus \mathfrak{g} \oplus (\mathfrak{g} \otimes \mathfrak{g}) \oplus (\mathfrak{g} \otimes \mathfrak{g} \otimes \mathfrak{g}) \oplus \cdots . \qquad (7.4.5)$$
>
> Let I be the two-sided ideal of $T(\mathfrak{g})$ generated by the elements $e_i e_j - e_j e_i - [e_i, e_j]$. Then $\mathcal{U}(\mathfrak{g}) = T(\mathfrak{g})/I$ has the desired properties of the universal enveloping algebra if we define $\Phi : \mathfrak{g} \to T(\mathfrak{g})/I$ by taking $\Phi(e_i) = e_i + I$ and extending linearly to all of $\mathfrak{g}$.

We first show that $(\Phi, T(\mathfrak{g})/I)$ is an enveloping algebra. To do so we just need to check that Φ is a Lie algebra homomorphism. Since Φ is defined by linearly extending a mapping of a basis it is sufficient to check on a basis that Φ is a Lie algebra homomorphism. That is, we need to check that $\Phi([e_i, e_j]) = [\Phi(e_i), \Phi(e_j)]$. This is true because

$$\Phi([e_i, e_j]) - [\Phi(e_i), \Phi(e_j)] = [e_i, e_j] - e_i e_j - e_j e_i + I = 0 + I. \quad (7.4.6)$$

Since $0 + I$ is zero in $T(\mathfrak{g})/I$ this shows that we have the desired equality. We now show that $(\Phi, T(\mathfrak{g})/I)$ is universal. Let (φ, U) be some other enveloping algebra. An element of $T(\mathfrak{g})$ is a finite sum of a finite product of e_is. Define the map $\tilde{f} : T(\mathfrak{g}) \to U$ by defining

$$f\left(\prod_{i=1}^{k} e_{i_j}\right) = \prod_{i=1}^{k} \varphi(e_{i_j}) \quad (7.4.7)$$

and extending linearly over sums. This is, by construction, a homomorphism of unital associative algebras. We further have

$$f([e_i, e_j] - e_i e_j - e_j e_i) = \varphi([e_i, e_j]) - \varphi(e_i)\varphi(e_j) - \varphi(e_j)\varphi(e_i) = 0 \quad (7.4.8)$$

where the last equality follows because $\varphi : \mathfrak{g} \to U$ is a Lie algebra homomorphism. Then $I \subseteq \ker \tilde{f}$, and so the map $f : \mathcal{U}(\mathfrak{g}) \to U$ given by $f(X + I) = \tilde{f}(X)$ is well-defined by definition of the quotient and is a homomorphism of associative unital algebras.

Uniqueness of f is immediate when we realise that at no point in this construction were we free to make any assumptions. $\qquad\square$

Lemma 7.4.9 The universal enveloping algebra is unique up to isomorphism.

Proof. Let (Φ_1, U_1) and (Φ_2, U_2) be universal enveloping algebras of $\mathfrak{g}$. Then by the universal properties of these two objects we have unique unital associative algebra homomorphisms $f_{ij} : U_i \to U_j$ such that $\Phi_i = f_{ij} \circ \Phi_j$ for $i, j \in \{1, 2\}$. Since $\Phi_i = \mathrm{id}_{U_i} \circ \Phi_i$ uniqueness imposes that $f_{ii} = \mathrm{id}_{U_i}$. We also have $\Phi_i = f_{ij} \circ \Phi_j = f_{ij} \circ f_{ji} \circ \Phi_i$, since $\Phi_j = f_{ji} \circ \Phi_i$, and thus by uniqueness $f_{ij} \circ f_{ji} = \mathrm{id}_{U_i}$. This means that f_{12} is invertible, and so is an isomorphism (of unital associative algebras) giving us the required uniqueness of universal enveloping algebras up to isomorphism. $\qquad\square$

One can also set up the category of enveloping algebras, the objects of which are enveloping algebras and a morphism $(\varphi, U) \to (\psi, V)$ is a homomorphism of unital associative algebras, $f : U \to V$, such that $\psi = f \circ \varphi$. Then the universal enveloping algebra is the initial object in this category and is unique by abstract nonsense.

The idea of the universal enveloping algebra is that it is the freest unital associative algebra which, when equipped with the commutator as a bracket, contains $\mathfrak{g}$ as a Lie-subalgebra. This is pretty much all we'll actually need.

Definition 7.4.10 — Casimir Element Let $\mathfrak{g}$ be a A **Casimir element**, Ω, is a distinguished element of the centre of $\mathcal{U}(\mathfrak{g})$.

That is, Ω is just some element of the universal enveloping algebra of $\mathfrak{g}$ which is such that it commutes with all elements of $\mathcal{U}(\mathfrak{g})$ (commuting with respect to the associative product in the normal way, so it's commutator with any element of $\mathcal{U}(\mathfrak{g})$ is always zero).

It turns out that there's a way to construct such an element that works for any finite-dimensional Lie algebra. Before we can do this we need the notion of a dual basis.

Let V be a finite-dimensional vector space with basis $\mathcal{B} = \{v_1, \dots, v_n\}$. Then any $v \in V$ may uniquely be expressed as a sum

$$v = \sum_{i=1}^{n} a_i w_i \tag{7.4.11}$$

with $a_i \in \mathbb{C}$. This defines n linear maps $\lambda_i : V \to \mathbb{C}$ given by $\lambda_i(v) = a_i$. These maps form a basis, $\{\lambda_1, \dots, \lambda_n\}$, for V^*. We say that this is the dual basis in V^* to $\mathcal{B}$, by which we mean that

$$\lambda_i(v_j) = \delta_{ij}. \tag{7.4.12}$$

Now suppose that $\eta : V \times V \to \mathbb{C}$ is a non-degenerate symmetric bilinear form. Then we can use this to define a second basis, $\mathcal{C} = \{u_1, \dots, u_n\}$ of V, which is such that $\eta(v_i, u_j) = \delta_{ij}$. Confusingly, we also call $\mathcal{C}$ the dual basis to $\mathcal{B}$, but now $\mathcal{C}$ is a basis for V, not V^*.

Lemma 7.4.13 Given a finite-dimensional vector space, V, with basis $\mathcal{B}$, and a non-degenerate symmetric bilinear form, $\eta : V \times V \to \mathbb{C}$, there is a unique dual basis, $\mathcal{C}$, such that $\eta(v_i, u_j) = \delta_{ij}$ for all $v_i \in \mathcal{B}$ and $u_j \in \mathcal{C}$.

Proof. Non-degeneracy of η means that we have an isomorphism $\iota_\eta : V \to V^*$. We may then define $u_i = \iota_\eta^{-1}(\lambda_i)$ where $\{\lambda_i\}$ is the dual basis to $\mathcal{B}$ in V^*. Then

$$\eta(v_i, u_j) = \eta(u_j, v_i) = \iota_\eta(u_j)(v_i) = \iota_\eta(\iota_\eta^{-1}(\lambda_j))(v_i)$$
$$= \mathrm{id}_{V^*}(\lambda_j)(v_i) = \lambda_j(v_i) = \delta_{ji} = \delta_{ij}. \tag{7.4.14}$$

This proves that $\mathcal{C} = \{u_1, \dots, u_n\}$ has the desired property. Uniqueness follows by noting that the above chain of equalities only holds for this choice of u_j. $\square$

One way of stating the above construction of $\mathcal{C}$ is that $\mathcal{C}$ is the pullback of the dual basis to $\mathcal{B}$ along ι_η. This is just fancy talk for saying that we're using the inverse of the map ι_η to take something defined in V^* to something defined in V.

Definition 7.4.15 — Quadratic Casimir Let $\mathfrak{g}$ be a finite-dimensional Lie algebra, and let $\mathcal{B} = \{X_1, \dots, X_n\}$ be a basis for $\mathfrak{g}$. Let $\mathcal{C} = \{Y_1, \dots, Y_n\}$ be the

dual basis for $\mathfrak{g}$ corresponding to some non-degenerate symmetric bilinear form, η. If $\mathfrak{g}$ is semisimple then we will assume $\eta = \kappa$ is the Killing form, but it is always possible to find such an η in any case. The **quadratic Casimir element**, Ω, is the Casimir element defined by

$$\Omega_2 = \sum_{i=1}^{n} X_i Y_i. \tag{7.4.16}$$

Note that this sum, and in particular the product $X_i Y_i$, is all done in $\mathcal{U}(\mathfrak{g})$. That Ω_2 is a Casimir element can be shown by an explicit calculation, or it can be seen by the following argument chaining together linear maps. This argument also proves that Ω_2 is independent of the choice of basis.

- There is a linear map $\Bbbk \to \operatorname{End}\mathfrak{g}$ given by sending $\lambda \in \Bbbk$ to the map $X \mapsto \lambda X$, that is the map which is just scaling by λ.

- For $\mathfrak{g}$ finite dimensional there is a canonical isomorphism of vector spaces[1] $\mathfrak{g} \otimes \mathfrak{g} \to \operatorname{End}_{\Bbbk} \mathfrak{g}$ given by $(X, \lambda) \mapsto (Y \mapsto (\lambda(Y)X))$, that is (X, λ) maps to the linear map given by scaling X by $\lambda(Y) \in \mathbb{C}$.

- For $\mathfrak{g}$ finite dimensional there is a canonical isomorphism of vector spaces $\iota_\eta : \mathfrak{g}^* \to \mathfrak{g}$

- There is a linear map $\mathfrak{g} \otimes \mathfrak{g} \to \mathcal{U}(\mathfrak{g})$ given by $X \otimes Y \mapsto XY$.

[1] We write $\operatorname{End}_{\Bbbk} \mathfrak{g}$ to mean the vector space of linear maps $\mathfrak{g} \to \mathfrak{g}$, rather than the set of Lie algebra homomorphisms $\mathfrak{g} \to \mathfrak{g}$, which we denote $\operatorname{End}\mathfrak{g}$.

Combining these we have

$$\Bbbk \hookrightarrow \operatorname{End}_{\Bbbk} \mathfrak{g} \xrightarrow{\sim} \mathfrak{g} \otimes \mathfrak{g}^* \xrightarrow{\sim} \mathfrak{g} \otimes \mathfrak{g} \to \mathcal{U}(\mathfrak{g}). \tag{7.4.17}$$

Now consider where $1 \in \Bbbk$ maps under this chain of linear maps.

- $1 \mapsto \operatorname{id}_{\mathfrak{g}}$ since scaling by 1 is the identity.

- $\operatorname{id}_{\mathfrak{g}} \mapsto \sum_{i=1}^{n} X_i \otimes \lambda_i$, this is the completeness relation and works for any choice of basis, $\{X_i\}$, and its dual, $\{\lambda_i\}$.

- $\lambda_i \mapsto Y_i$ so $X_i \otimes \lambda_i \mapsto X_i \otimes Y_i$ and $\sum_{i=1}^{n} X_i \otimes \lambda_i \mapsto \sum_{i=1}^{n} X_i \otimes Y_i$.

- $X_i \otimes Y_i$ maps to $X_i Y_i$ so $\sum_{i=1}^{n} X_i \otimes Y_i \mapsto \sum_{i=1}^{n} X_i Y_i = \Omega_2$.

Since $1 \in \Bbbk$ is $\mathfrak{g}$-invariant and commutes with all elements the same is true of its image under this composite algebra morphism, and the image is exactly the quadratic Casimir, Ω_2.

7.4.1 Casimir According to the Course

We now give the definition of the representation of Ω, which is what the course calls *the* Casimir operator.

Let (ρ, V) be a representation of a simple Lie algebra, $\mathfrak{g}$. Then V is either faithful or trivial. Trivial modules aren't interesting, so we'll assume V is faithful. There is an obvious generalisation of the Killing form, namely

$$\beta_V : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C} \tag{7.4.18}$$
$$(X, Y) \mapsto \beta_V(X, Y) = \operatorname{tr}_V(\rho(X)\rho(Y)) \tag{7.4.19}$$

where tr_V is the trace of an operator on V. Note that the Killing form is then simply $\kappa = \beta_{\mathfrak{g}}$ where we're considering the representation $(\mathrm{ad}, \mathfrak{g})$. Many of the properties of the Killing form, such as bilinearity and associativity, hold for β_V without much modification to the proofs. The fact that $\mathfrak{g}$ is simple means that by similar proof to Cartan's criterion we also know that β_V is non-degenerate.

We now define "the Casimir" as the course defines it.

> **Definition 7.4.20** Let V be a faithful $\mathfrak{g}$-module, and let (ρ, V) be the corresponding representation. We define the **Casimir** of V to be
>
> $$\Omega_V = \sum_{i=1}^{n} \rho(X_i)\rho(Y_i) \tag{7.4.21}$$
>
> where $\{X_1, \ldots, X_n\}$ is some basis of $\mathfrak{g}$ and $\{Y_1, \ldots, Y_n\}$ is the dual basis in $\mathfrak{g}$ corresponding to the symmetric non-degenerate associative bilinear form β_V.

That is, $\Omega_V = \tilde{\rho}(\Omega_2)$ with our more general definition of the quadratic Casimir, where $\tilde{\rho} \colon \mathcal{U}(\mathfrak{g}) \to \mathfrak{gl}(V)$ is the unital associative algebra homomorphism corresponding to the Lie algebra homomorhpism ρ under the universal property of $\mathcal{U}(\mathfrak{g})$ with $(\rho, \mathrm{End}\, V)$ viewed as an enveloping algebra.

That Ω_V commutes with all elements of $\mathrm{End}\, V$ follows immediately from the fact that Ω_2 commutes with all elements of $\mathcal{U}(\mathfrak{g})$. It can also be proven by a calculation from the definition. First take $X \in \mathfrak{g}$ and let $\{X_i\}$ be a basis for the n-dimensional $\mathfrak{g}$. Write

$$[X, X_i] = \sum_{j=1}^{n} a_{ij} X_j \tag{7.4.22}$$

for some $a_{ij} \in \mathbb{C}$. This is just expressing $[X, X_i] \in \mathfrak{g}$ as a linear combination of basis vectors, X_i. Similarly, write

$$[X, Y_i] = \sum_{j=1}^{n} b_{ij} Y_j \tag{7.4.23}$$

where $\{Y_i\}$ is the dual basis to $\{X_i\}$ corresponding to some symmetric non-degenerate

associative bilinear form, β_V. Then we have

$$a_{ij} = \sum_{k=1}^{n} a_{ik}\delta_{kj} \tag{7.4.24}$$

$$= \sum_{k=1}^{n} a_{ik}\beta_V(X_k, Y_j) \tag{7.4.25}$$

$$= \beta_V\left(\sum_{k=1}^{n} a_{ik}X_k, Y_j\right) \tag{7.4.26}$$

$$= \beta_V([X, X_i], Y_j) \tag{7.4.27}$$

$$= \beta_V(-[X_i, X], Y_j) \tag{7.4.28}$$

$$= \beta_V(-X_i, [X, Y_j]) \tag{7.4.29}$$

$$= \beta_V\left(-X_i, \sum_{k=1}^{n} b_{jk}Y_k\right) \tag{7.4.30}$$

$$= \sum_{k=1}^{n} -b_{jk}\beta_V(X_i, Y_k) \tag{7.4.31}$$

$$= \sum_{k=1}^{n} -b_{jk}\delta_{ik} \tag{7.4.32}$$

$$= -b_{ji}. \tag{7.4.33}$$

Now we can use the fact that in End V (and more generally in any algebra for which the bracket is a commutator) we have

$$[XY, Z] = [X, Z]Y + X[Y, Z], \tag{7.4.34}$$

which can be shown by expanding these brackets as commutators. This gives us

$$[\Omega_V, \rho(X)] = \left[\sum_{i=1}^{n} \rho(X_i)\rho(Y_i), \rho(X)\right] \tag{7.4.35}$$

$$= \sum_{i=1}^{n} ([\rho(X_i), \rho(X)]\rho(Y_i) + \rho(X_i)[\rho(Y_i), \rho(X)]) \tag{7.4.36}$$

$$= \sum_{i=1}^{n} (\rho([X_i, X])\rho(Y_i) + \rho(X_i)\rho([Y_i, X])) \tag{7.4.37}$$

$$= \sum_{i=1}^{n} (\rho(-[X, X_i])\rho(Y_i) + \rho(X_i)\rho(-[X, Y_i])) \tag{7.4.38}$$

$$= -\sum_{i=1}^{n} \left(\rho\left(\sum_{j=1}^{n} a_{ij}X_j\right)\rho(Y_i) + \rho(X_i)\rho\left(\sum_{j=1}^{n} b_{ij}Y_j\right)\right) \tag{7.4.39}$$

$$= -\sum_{i=1}^{n}\sum_{j=1}^{n} (a_{ij}\rho(X_j)\rho(Y_i) + b_{ij}\rho(X_i)\rho(Y_j)) \tag{7.4.40}$$

$$= 0 \tag{7.4.41}$$

where the last equality follows since $a_{ij} = -b_{ji}$ so the two terms in the sum cancel.

Note that we have

$$\text{tr}(\Omega_V) = \text{tr}\left(\sum_{i=1}^{n} \rho(X_i)\rho(Y_i)\right) \tag{7.4.42}$$

$$= \sum_{i=1}^{n} \text{tr}(\rho(X_i)\rho(Y_i)) \tag{7.4.43}$$

$$= \sum_{i=1}^{n} \beta_V(X_i, Y_i) \tag{7.4.44}$$

$$= \sum_{i=1}^{n} \delta_{ii} \tag{7.4.45}$$

$$= \dim \mathfrak{g}. \tag{7.4.46}$$

7.5 Proving Weyl's Theorem

Lemma 7.2.1 implies that if V is a $\mathfrak{g}$-module then it is sufficient to show that each $\mathfrak{g}$-submodule, W, admits a complementary $\mathfrak{g}$-submodule, W'. To do this we'll make use of the decomposition theorem (Theorem 2.2.25). Recall that this states that for a finite dimensional vector space, V, and some endomorphism $X \in \text{End}_{\mathbb{C}}(V)$ we have the decomposition

$$V = \bigoplus_{\alpha \in \mathbb{C}} V_\alpha \tag{7.5.1}$$

where

$$V_\alpha = \{v \in V \mid (X - \alpha)^N v = 0 \text{ for some } N \in \mathbb{N}\} \tag{7.5.2}$$

is the generalised eigenspace of V associated with X.

If in addition V is a $\mathfrak{g}$-module and $X \in \text{End}_{\mathfrak{g}}(V)$ then each V_α is a $\mathfrak{g}$-submodule since the action of $\mathfrak{g}$ on V commutes with the $\mathfrak{g}$-module endomorphism, X (and also the identity), and thus leaves each V_α invariant. That is, $Y \cdot (X - \alpha)^N v = 0$ if and only if $(X - \alpha)^N (Y \cdot v) = 0$.

Lemma 7.5.3 If $\mathfrak{g}$ is a semisimple Lie algebra and V is a one-dimensional $\mathfrak{g}$-module then V is trivial.

Proof. Let $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ be the corresponding Lie algebra representation. For V one-dimensional, so $V \cong \mathbb{C}$, then $\mathfrak{gl}(V) \cong \mathfrak{gl}(\mathbb{C}) = \mathbb{C}$, the only linear maps $\mathbb{C} \to \mathbb{C}$ are scalar multiplication. The image $\rho(\mathfrak{g})$ is a Lie subalgebra of $\mathfrak{gl}(V)$, and so in this case $\rho(\mathfrak{g})$ is a Lie subalgebra of $\mathbb{C}$, and thus is abelian. Then $X, Y \in \mathfrak{g}$ act on V by scalar multiplication, say $X \cdot v = xv$ and $Y \cdot v = yv$ for some $x, y \in \mathbb{C}$ and $v \in V$. Then we have $X \cdot (Y \cdot v) = xyv = yxv = Y \cdot (X \cdot v)$ so $X \cdot (Y \cdot v) - Y \cdot (X \cdot v) = 0$ which means that we must have $[X, Y] \cdot v = 0$. Thus, we have that $[\mathfrak{g}, \mathfrak{g}]$ must act trivially on V, but $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ for a semisimple Lie algebra, so $\mathfrak{g}$ acts trivially on V. This last fact follows because $[\mathfrak{g}, \mathfrak{g}]$ is a solvable subalgebra of $\mathfrak{g}$ and therefore for semisimple $\mathfrak{g}$ must be either 0 or $\mathfrak{g}$, and it's not 0 unless $\mathfrak{g}$ is abelian, which is assumed not to be the case for a semisimple

Lie algebra. □

Before we prove Weyl's theorem for all simple Lie algebras we prove a special case, and then the proof of the full theorem is just reduction to this special case.

Proposition 7.5.4 Let V be a $\mathfrak{g}$-module and W a submodule of codimension one. Then, there exists a $\mathfrak{g}$-submodule, W', of dimension one with $W \oplus W' = V$.

Proof. The proof is by induction on $\dim V$. The base case, $\dim V = 1$, is trivial since any submodule of codimension one is of dimension $1 - 1 = 0$ so the only codimension one submodule is 0, and we trivially have $\dim V = 1$ and $0 \oplus V = V$.

For the inductive step assume that any $\mathfrak{g}$-module of dimension less than $\dim V$ admits complementary submodules for all submodules of codimension 1. Let W be a $\mathfrak{g}$-submodule of V of codimension 1. We know that V is either a faithful representation or trivial (Lemma 7.3.2). If V is trivial then any dimension 1 $\mathfrak{g}$-submodule (which is any dimension 1 subspace) will work for W'. Suppose V is faithful. We can use the Casimir, $\Omega_V \in \operatorname{End}_{\mathfrak{g}} V$, to generate W'. To do this we consider the decomposition

$$V = \bigoplus_{\alpha \in \mathbb{C}} V_\alpha \tag{7.5.5}$$

where $V_\alpha = \{v \in V \mid (\Omega_V - \alpha)^N(v) = 0 \text{ for some } N \in \mathbb{N}\}$ is the generalised eigenspace associated with Ω_V. We know that Ω_V preserves W, that is $\Omega_V(W) \subseteq W$, since $\mathfrak{g}$ preserves W, and Ω_V acts as a sum of (products) of elements of $\mathfrak{g}$. We can therefore consider the restriction of Ω_V to W, $\Omega_V|_W \colon W \to W$ where $\Omega_V|_W(w) = \Omega_V(w)$, and $\Omega_V|_W \in \operatorname{End}_{\mathfrak{g}} W$. We then get a decomposition

$$W = \bigoplus_{\alpha \in \mathbb{C}} W_\alpha \tag{7.5.6}$$

where

$$W_\alpha = \{w \in W \mid (\Omega_V|_W - \alpha)^N(w) = 0 \text{ for some } N \in \mathbb{N}\} \tag{7.5.7}$$

is the generalised eigenspace associated with $\Omega_V|_W$. From these definitions we can immediately see that W_α consists of all elements of W for which $(\Omega_V - \alpha)^N(w)$ is eventually 0 for some $N \in \mathbb{N}$. That is, $W_\alpha = W \cap V_\alpha$.

Claim 7.5.8 With the notation as above we have $V_\alpha = W_\alpha$ for all $\alpha \neq 0$ and $\dim(V_0/W_0) = 1$.

Proof. Each V_α is a $\mathfrak{g}$-module and W_α is a $\mathfrak{g}$-submodule of V_α. Thus, the quotient V_α/W_α is a $\mathfrak{g}$-module. Further, we have

$$V/W = \bigoplus_{\alpha \in \mathbb{C}} V_\alpha/W_\alpha. \tag{7.5.9}$$

Since $\dim V/W = 1$ (because W is of codimension 1 by assumption) there is therefore exactly one $\alpha \in \mathbb{C}$ for which $V_\alpha/W_\alpha \neq 0$ and for that value of α we have $V_\alpha/W_\alpha = 1$. This holds because the dimension of the sum is the sum of the dimensions. We then only need to show that $\alpha = 0$ is this distinguished value. This follows because the quotients $V_\alpha/W_\alpha = V_\alpha/(V_\alpha \cap W)$ are such that $\mathfrak{g}$ acts trivially on these, since these quotients have dimension at most one. Then Ω_V also acts trivially on the quotients since Ω_V acts as a sum of products of elements of $\mathfrak{g}$. However, by definition Ω_V also acts on V_α by scalar multiplication by α and this carries across into the quotient, so it must be that $\alpha = 0$ for our non-zero quotient. $\square$

To complete the proof from here it is sufficient to show that $\dim V_0 < \dim V$, since $W_0 \subseteq V_0$ is a submodule of codimension one and this will allow us to apply the inductive hypothesis. We know that $\dim V_0 < \dim V$ because $\operatorname{tr}(\Omega_V) = \dim \mathfrak{g} \neq 0$ so Ω_V is not the zero operator, so there is at least one $\alpha \in \mathbb{C} \setminus \{0\}$ for which $V_\alpha \neq 0$ and thus V_0 cannot be all of V. $\square$

We're now almost ready to prove Weyl's theorem for any simple $\mathfrak{g}$-module. In order to do this we just need a few more details. If V and W is a $\mathfrak{g}$-module then so is $\operatorname{Hom}_{\mathbb{C}}(V, W)$, which has its vector space structure defined pointwise and the $\mathfrak{g}$-module structure defined by taking $X \in \mathfrak{g}$ to act on $f \in \operatorname{Hom}_{\mathbb{C}}(V, W)$ by defining $X \cdot f$ to be the map

$$(X \cdot f)(v) = X \cdot f(v) - f(X \cdot v) \tag{7.5.10}$$

for all $v \in V$.

Definition 7.5.11 — Invariants If U is a $\mathfrak{g}$-module denote by $U^{\mathfrak{g}}$ the $\mathfrak{g}$-submodule consisting of $\mathfrak{g}$-invariant elements, that is $u \in U^{\mathfrak{g}}$ if and only if $X \cdot u = 0$ for all $X \in \mathfrak{g}$.

Lemma 7.5.12 Let V and W be $\mathfrak{g}$-modules. Then $\operatorname{Hom}_{\mathbb{C}}(V, W)^{\mathfrak{g}} = \operatorname{Hom}_{\mathfrak{g}}(V, W)$. That is, the $\mathfrak{g}$-equivaraint maps are precisely those which are fixed under the action of $\mathfrak{g}$.

Proof. Suppose that $f \in \operatorname{Hom}_{\mathfrak{g}}(V, W)$, that is $f : V \to W$ is linear and $X \cdot f(v) = f(X \cdot v)$ for all $X \in \mathfrak{g}$ and $v \in V$. Then by definition X acts

trivially on f since the right hand side of Equation (7.5.10) vanishes. Suppose instead that $f \in \mathrm{Hom}_{\mathbb{C}}(V, W)^{\mathfrak{g}}$, so that $X \cdot f = 0$ for all $X \in \mathfrak{g}$. Then the left hand side of Equation (7.5.10) vanishes and it follows that $X \cdot f(v) = f(X \cdot v)$, so f is $\mathfrak{g}$-equivariant, so $f \in \mathrm{Hom}_{\mathfrak{g}}(V, W)$. $\qquad\square$

We are now ready to prove Weyl's Theorem for simple Lie algebras.

Theorem 7.5.13 — Weyl. Let $\mathfrak{g}$ be a simple Lie algebra over $\mathbb{C}$. Then every finite-dimensional representation of $\mathfrak{g}$ is completely reducible.

Proof. Let V be a $\mathfrak{g}$-module with $W \subseteq V$ a $\mathfrak{g}$-submodule. Then $\mathrm{Hom}_{\mathbb{C}}(V, W)$ is a $\mathfrak{g}$-module. Define

$$U = \{f \in \mathrm{Hom}_{\mathbb{C}}(V, W) \mid f|_W = \lambda \mathrm{id}_W \text{ for some } \lambda \in \mathbb{C}\}. \qquad (7.5.14)$$

That is, U is the subspace of $\mathrm{Hom}_{\mathbb{C}}(V, W)$ consisting of linear maps which restrict to scalar multiplication on W. Let

$$U_0 = \{f \in \mathrm{Hom}_{\mathbb{C}}(V, W) \mid f|_W = 0\} \qquad (7.5.15)$$

be the subspace of U of linear maps which $V \to W$ which restrict to the zero map on W.

Claim 7.5.16 With notation as above U and U_0 are $\mathfrak{g}$-submodules of $\mathrm{Hom}_{\mathbb{C}}(V, W)$.

Proof. We need to show that U and U_0 are closed under the action of $\mathfrak{g}$. Let $X \in \mathfrak{g}$ and $f \in U$. Then X acts on f according to Equation (7.5.10). That is

$$(X \cdot f)(v) = X \cdot f(v) - f(X \cdot v) \qquad (7.5.17)$$

for all $v \in V$. We need to show that $X \cdot f \in U$, that is, we want to show that $(X \cdot f)|_W = \lambda \mathrm{id}_W$ for some $\lambda \in \mathbb{C}$. This is simple because if $v = w \in W$ in the above we have

$$(X \cdot f)|_W(w) = (X \cdot f)(w) = X \cdot f(w) - f(X \cdot w) \qquad (7.5.18)$$

and since W is a submodule we know that $X \cdot w \in W$ so both applications of f above are to some element of W, and so both are just scalar multiplication by some $\lambda \in \mathbb{C}$, and thus

$$(X \cdot f)|_W(w) = X \cdot (\lambda w) - \lambda(X \cdot w) = 0 \qquad (7.5.19)$$

using linearity of the action of X. This shows that $X \cdot f \in U$ as $X \cdot f$ simply acts as $0\mathrm{id}_W$. This also proves that U_0 is a submodule, it's just the special case where $\lambda = 0$ in the above equations. $\qquad\square$

Claim 7.5.20 With notation as above dim $U/U_0 = 1$.

Proof. Consider the restriction map

$$\text{res}: U \to \text{Hom}_{\mathbb{C}}(W, V) \qquad\qquad (7.5.21)$$
$$f \mapsto f|_W. \qquad\qquad (7.5.22)$$

By definition of U we know that res $f = \lambda\text{id}_W$ where $\lambda \in \mathbb{C}$, and certainly we hit every $\lambda \in \mathbb{C}$ in this way, so $\text{res}(U) \cong \mathbb{C}$. We also have that ker(res) consists of those maps which map to 0id_W, which is exactly ker(res) $= U_0$. Thus, by the rank-nullity theorem we have that

$$\dim U = \dim(\text{res}(U)) + \dim(\text{ker}(\text{res})) = 1 + \dim U_0 \quad (7.5.23)$$

and so dim $U/U_0 = 1$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

By the special case of Weyl's theorem (Proposition 7.5.4) we know that U_0 admits a complement U_0' such that $U_0 \oplus U_0' = U$. Now take some $\varphi \in U_0'$, then we know that $\text{res}\,\varphi = \varphi|_W = c\text{id}_W$ for some $c \in \mathbb{C}$. Since $\varphi \in U_0'$ we know that $\varphi \in U_0$ only if $\varphi = 0$. Assuming that $\varphi \neq 0$ we therefore have that $c \neq 0$, and so we can rescale φ to $\frac{1}{c}\varphi$ which has $\text{res}(\frac{1}{c}\varphi) = \text{id}_W$. Therefore without loss of generality we may assume that any nonzero $\varphi \in U_0'$ is such that $\text{res}(\varphi) = \text{id}_W$.

Claim 7.5.24 With the notation as above we have $\varphi \in \text{Hom}_{\mathfrak{g}}$.

Proof. We know that dim $U_0' = 1$, and thus the action of $\mathfrak{g}$ on U_0' is trivial (Lemma 7.5.3). Thus, for all $X \in \mathfrak{g}$ we have $X \cdot \varphi(v) = \varphi(X \cdot v)$, which is Equation (7.5.10) defining the action of X on $\text{Hom}_{\mathbb{C}}(V, W)$ when the action is trivial. This result is exactly what it means for φ to be a $\mathfrak{g}$-module homomorphism. $\qquad\qquad \square$

This allows us to construct the following sequence of $\mathfrak{g}$-modules linked by $\mathfrak{g}$-module homomorphisms:

$$W \hookrightarrow V \xrightarrow{\varphi} W. \qquad\qquad (7.5.25)$$

Further, by definition we know that embedding W in V then applying φ is the same sa applying the identity map since $\text{res}(\varphi) = \text{id}_W$. The existence of such a chain here tells us that $V = W \oplus \ker\varphi$ and thus we have found a complement for W and so V is completely reducible by Lemma 7.2.1. $\quad \square$

Here we've used the fact that a short exact sequence

$$0 \to A \xrightarrow{f} B \xrightarrow{g} C \to 0 \qquad\qquad (7.5.26)$$

splits (meaning that the map $B \to C$ has a right inverse, $h: C \to B$, which we've

shown is the case for φ here) if and only if $B = C/A$. This is a general result that holds for vector spaces and more generally for R-modules. It follows then that $B \cong A \oplus C$ because we can explicitly construct an isomorphism $A \oplus C \to B$ by $(a, c) \mapsto f(a) + h(c)$. Note that a sequence

$$\cdots \to X \xrightarrow{f} Y \xrightarrow{g} Z \to \cdots \tag{7.5.27}$$

is exact at Y if $f(X) = \ker g$, which for a short exact sequence, such as Equation (7.5.26), implies that f is injective and g is surjective.

Note that the choice of U_0' is not unique, since in the proof of Proposition 7.5.4 any one-dimensional subspace we choose will work.

It remains only to show that this result generalises to semisimple Lie algebras. To do this first note that if V is an $(\mathfrak{m} \oplus \mathfrak{l})$-module then defining $X \cdot f = (X, 0) \cdot f$ for all $f \in \mathrm{Hom}_\ell(W, V)$ makes $\mathrm{Hom}_\ell(W, V)$ into a $\mathfrak{m}$-module.

Corollary 7.5.28 — Weyl for Semisimple Lie Algebras Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{C}$. Then every finite-dimensional representation of $\mathfrak{g}$ is completely reducible.

Proof. Recall that any semisimple Lie algebra, $\mathfrak{g}$, can be decomposed as a sum of simple Lie algebras, $\mathfrak{g}_i$:

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_k. \tag{7.5.29}$$

We prove this result by induction on k. The case of $k = 1$ simply has $\mathfrak{g} = \mathfrak{g}_1$, which means that $\mathfrak{g}$ is simple, so this is just Weyl's theorem (Theorem 7.5.13). Suppose then that every semisimple Lie algebra which decomposes into fewer than k simple Lie algebras has only completely reducible finite-dimensional representations over $\mathbb{C}$.

Let $\mathfrak{g}$ be a semisimple Lie group decomposing as $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{m}$ where $\mathfrak{g}_1$ is simple and $\mathfrak{m} = \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_k$ where the $\mathfrak{g}_i$ are also simple. Let V be a finite-dimensional $\mathfrak{g}$-module. Using the simplicity of $\mathfrak{g}_1$ we can decompose V as a direct sum $V = V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}$ where the V_i are simple, pairwise non-isomorphic $\mathfrak{g}_1$-modules. Define a map

$$\mathrm{Hom}_{\mathfrak{g}_1}(V_i, V) \otimes V_i \to V \tag{7.5.30}$$

$$(\varphi, v) \mapsto \varphi(v). \tag{7.5.31}$$

We know by the discussion preceding this corollary that $\mathrm{Hom}_{\mathfrak{g}_1}(V_i, V)$ is an $\mathfrak{m}$-module. Thus, $\mathrm{Hom}_{\mathfrak{g}_1}(V_i, V) \otimes V_i$ is a $\mathfrak{m}$ module, and V is a $\mathfrak{g}_1$-module, because it is a $\mathfrak{g}$-module and $\mathfrak{g}_1 \subseteq \mathfrak{g}$. Then $\mathrm{Hom}_{\mathfrak{g}_1}(V_i, V) \otimes V_1$ is an $(\mathfrak{m} \oplus \mathfrak{g}_1)$-module, and $\mathfrak{m} \oplus \mathfrak{g}_1 \cong \mathfrak{g}_1 \oplus \mathfrak{m} = \mathfrak{g}$, so $\mathrm{Hom}_{\mathfrak{g}_1}(V_i, V) \otimes V$ is a $\mathfrak{g}$-module. The map defined above then extends to an isomorphism of $\mathfrak{g}$-modules

$$(\mathrm{Hom}_{\mathfrak{g}_1}(V_1, V) \otimes V_1) \oplus \cdots \oplus (\mathrm{Hom}_{\mathfrak{g}_1}(V_r, V) \otimes V_r) \to V. \tag{7.5.32}$$

Since each $\mathrm{Hom}_{\mathfrak{g}_1}(V_i, V)$ is a completely reducible $\mathfrak{m}$-module by the induction hypothesis we can decompose each $\mathrm{Hom}_{\mathfrak{g}_1}(V_i, V) \otimes V_i$ as a direct sum of simple $\mathfrak{g}$-modules. $\square$

Eight

Cartan Subalgebras

8.1 Representation Theory of $\mathfrak{sl}_2$

We start this section with a recap of $\mathfrak{sl}_2 = \mathfrak{sl}(2, \mathbb{C})$ representation theory. We saw in Theorem 4.3.31 that any simple $\mathfrak{sl}_2$-module is isomorphic to an $\mathfrak{sl}_2$-module, $V(n)$, for some $n \in \mathbb{Z}$ such that there exists some highest weight vector $v_0 \in V(n)_n$ where

$$V(n)_k = \{v \in V(n) \mid Hv = kv\} \tag{8.1.1}$$

for integer k with $-n \leq k \leq n$. By "highest weight" we mean that $Hv_0 = nv_0$ and $Ev_0 = 0$. The dimension of $V(n)$ is $\dim V(n) = n + 1$, with a basis given by $\{v_i = F^i v_0 / i! \mid i = 0, 1, \ldots, n\}$. This leads to the picture in Equation (4.3.32) where the $\mathfrak{sl}_2$ generators act as $E \colon V(n)_k \to V(n)_{k+2}$, $H \colon V(n)_k \to V(n)_k$, and $F \colon V(n)_k \to V(n)_{k-2}$, where we define $V(n)_{n+2} = 0 = V(n)_{-n-2}$.

Since $\mathfrak{sl}_2$ is simple the work of the previous chapter tells us that any finite-dimensional $\mathfrak{sl}_2$-module, V, admits a decomposition as

$$V \cong \bigoplus_{i \geq 0} V(i)^{\oplus n_i} \tag{8.1.2}$$

where the **multiplicity** is $n_i \in \mathbb{Z}_{\geq 0}$ with $n_i = 0$ for $i \gg 0$. These n_i are uniquely determined by V, and can be computed as $n_i = \dim(V_i \cap \ker E)$. That is, we get one $V(i)$ submodule for each basis vector in V_i (that is, an eigenvector of H with eigenvalue i) that also vanishes under the action of E, so the highest weight vectors of V determine n_i.

We now list some special properties of H and which we will seek to generalise to other simple Lie algebras.

- If V is a finite dimensional $\mathfrak{sl}_2$-module then H acts diagonalisably on V, this is an immediate consequence of the direct sum decomposition of V being into eigenspaces of H. In fact, one can take "diagonalisable" to mean that there is an eigenbasis of the operator corresponding to such a decomposition.

- All of H's eigenvalues are integers.

- For any integer $\alpha \in \mathbb{Z}$ we have $\dim V_\alpha = \dim V_{-\alpha}$, since the generalised H-eigenspace of α and $-\alpha$ coincide.

These last two facts are particularly nice because $\mathfrak{sl}_2$-modules tend to arise in lots of areas of maths, such as Hodge theory and the cohomology of Kähler manifolds.

8.1.1   Representation Theory of $\mathfrak{sl}_3$

Now consider $\mathfrak{sl}_3 = \mathfrak{sl}(3, \mathbb{C})$, which is also a simple Lie algebra, consisting of trace-less 3×3 matrices over $\mathbb{C}$. There is no one equivalent "H" for $\mathfrak{sl}_3$. Instead, we can consider

$$H_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \text{and} \quad H_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \tag{8.1.3}$$

More specifically, we want to consider $\mathfrak{h} = \operatorname{span}_{\mathbb{C}}\{H_1, H_2\}$ which consists of all traceless diagonal 3×3 matrices over $\mathbb{C}$. Compare this to the case of $\mathfrak{sl}_2$ where considering the same set but for 2×2 matrices would mean considering $\mathfrak{h} = \operatorname{span}_{\mathbb{C}}\{H\}$. We can generalise the definition of a weight space from $\mathfrak{sl}_2$ to $\mathfrak{sl}_3$ by defining a weight space to be a the common eigenspace of all $H \in \mathfrak{h}$. Then any finite-dimensional $\mathfrak{sl}_3$-module, V, has a decomposition into weight-spaces with multiplicity, just as in the $\mathfrak{sl}_2$ case.

    We will follow this generalisation through for arbitrary simple finite-dimensional Lie algebras over $\mathbb{C}$, and we shall see that we get the same result each time. We'll also see that there are generalisations of E and F, as well as the hidden integrality of the eigenvalues of H which lead us to define something called a root system, which then leads to a full classification of all finite-dimensional simple Lie algebras over $\mathbb{C}$.

8.1.2   Preliminary Definitions

> **Definition 8.1.4** Let V be a finite-dimensional vector space over $\mathbb{C}$. We say that $A \in \operatorname{End}_{\mathbb{C}} V$ is **nilpotent** if there exists $n \in \mathbb{N}$ such that $A^n = 0$. We say that A is **semisimple** if it is diagonalisable.

    For example, for any finite-dimensional $\mathfrak{sl}_2$-module E and F are nilpotent, since applying them repeatedly we always end up with the zero module, and H is semisimple since we've shown that H is diagonalisable. The benefit of the "semisimple" wording over "diagonalisable" is that it reduces the emphasis of matrices, which is generally a good thing for more advanced linear algebra where we aim to talk about linear maps rather than matrices.

> **Proposition 8.1.5** Let V be a finite-dimensional vector space over $\mathbb{C}$ and let $A \in \operatorname{End}_{\mathbb{C}} V$. Then there exists a unique decomposition $A = A_{\mathrm{s}} + A_{\mathrm{n}}$ with A_{s} semisimple, A_{n} nilpotent, and $[A_{\mathrm{s}}, A_{\mathrm{n}}] = 0$.
>
> ---
>
> *Proof.* Since we work over $\mathbb{C}$ we know that V decomposes into generalised-eigenspaces of A as
>
> $$V = \bigoplus_{\alpha \in \mathbb{C}} V_\alpha \tag{8.1.6}$$
>
> where
>
> $$V_\alpha = \{v \in V \mid (A - \alpha)^N v = 0 \text{ for some } N \in \mathbb{N}\}. \tag{8.1.7}$$
>
> We can then define A_{s} and A_{n} by their action on each generalised eigen-

space. In particular, take

$$A_{\mathrm{s}} = \sum_{\alpha \in \mathbb{C}} \alpha \mathrm{id}_{V_\alpha} \qquad (8.1.8)$$

so A_{s} acts on V_α by multiplication by α. Define A_{n} to be what is left over,

$$A_{\mathrm{n}} = \sum_{\alpha \in \mathbb{C}} (A|_{V_\alpha} - \alpha \mathrm{id}_{V_\alpha}). \qquad (8.1.9)$$

Then by construction $A = A_{\mathrm{s}} + A_{\mathrm{n}}$ and A_{s} acts on each generalised eigenspace by scalar multiplication so it's action commutes with that of A_{n}. We do not prove uniqueness. $\qquad \square$

Note that this works over $\mathbb{C}$, since it's algebraically closed, guaranteeing the generalised eigenspace decomposition exists. It does not work in general over non-algebraically closed fields, it only works for the special cases where the field contains the eigenvalues.

8.1.3 Cartan Subalgebra

Definition 8.1.10 Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{C}$. We say that $X \in \mathfrak{g}$ is nilpotent (semisimple), or if we need to distinguish ad-nilpotent (ad-semisimple) if ad_X is a nilpotent (semisimple) element of $\mathrm{End}_{\mathbb{C}} \mathfrak{g}$.

Definition 8.1.11 — Cartan Subalgebra Let $\mathfrak{g}$ be a Lie algebra. A **Cartan subalgebra** is a subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ containing only semisimple elements and which is maximal with respect to this property. That is, if $\mathfrak{h}' \subseteq \mathfrak{g}$ is also a subalgebra consisting of semisimple elements and $\mathfrak{h} \subseteq \mathfrak{h}'$ then $\mathfrak{h} = \mathfrak{h}'$.

Lemma 8.1.12 Every finite-dimensional Lie algebra has a Cartan subalgebra.

Proof. The zero space is always a subalgebra, and the zero element is semisimple. This shows that every Lie algebra has a subalgebra consisting of semisimple elements. Since we assume a finite number of dimensions it is always possible to take some maximal such subalgebra and this will then be a Cartan subalgebra. $\qquad \square$

Note that we say *a* Cartan subalgebra, Cartan subalgebras are not necessarily unique. However, when there are multiple Cartan subalgebras they are all conjugate, and as such it's common to talk of "the" Cartan subalgebra.

One property of Cartan subalgebras is that they are always abelian. This is simply a special case of the following.

Lemma 8.1.13 Let $\mathfrak{g}$ be a Lie algebra. Any subalgebra, $\mathfrak{h} \subseteq \mathfrak{g}$, consisting of semisimple elements is abelian.

Proof. Consider some $X \in \mathfrak{h}$. Since $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$ it is closed under the bracket, and thus we can restrict the adjoint action of X on $\mathfrak{g}$ to a map $\mathrm{ad}_X : \mathfrak{h} \to \mathfrak{h}$. Our goal is to show that $\mathfrak{h}$ is abelian, which we can do by showing that $\mathrm{ad}_X = 0$ when restricted to $\mathfrak{h}$. By assumption $\mathfrak{h}$ consists of semisimple elements so ad_X is semisimple. Therefore it suffices to show that all eigenvalues of ad_X are zero, since then in the basis in which it is diagonal it has only zeros on the diagonal, and thus is just the zero matrix. More generally, if all eigenvalues vanish then the operator is also nilpotent, and semisimplicity plus nilpotency imply that the operator is zero.

Let $Y \in \mathfrak{h}$ be an eigenvector of ad_X. Then there exists some $a \in \mathbb{C}$ such that $\mathrm{ad}_X(Y) = aY$. Note that Y is also an eigenvector of ad_Y since $\mathrm{ad}_Y(Y) = [Y, Y] = 0$, so the eigenvalue is 0. We may extend this to some basis of eigenvectors of Y, $\{X_1, \ldots, X_n\}$, with $X_1 = Y$. Then by definition there exist some $\alpha_i \in \mathbb{C}$ such that $\mathrm{ad}_Y(X_i) = \alpha_i X_i$, and in particular we must have that $\alpha_1 = 0$. Now expand X in this eigenbasis, which lets us write

$$X = \sum_{i=1}^{n} u_i X_i \tag{8.1.14}$$

for some $u_i \in \mathbb{C}$. A direct computation in this basis then shows that

$$\mathrm{ad}_X(Y) = [X, Y] \tag{8.1.15}$$
$$= -[Y, X] \tag{8.1.16}$$
$$= -\left[Y, \sum_{i=1}^{n} u_i X_i \right] \tag{8.1.17}$$
$$= -\sum_{i=1}^{n} u_i [Y, X_i] \tag{8.1.18}$$
$$= -\sum_{i=1}^{n} u_i \mathrm{ad}_Y(X_i) \tag{8.1.19}$$
$$= -\sum_{i=2}^{n} u_i \alpha_i X_i \tag{8.1.20}$$

where we've used the fact that $\mathrm{ad}_Y(X_1) = 0$ to get rid of the first term of the sum. We also have

$$\mathrm{ad}_X(Y) = aY = aX_1. \tag{8.1.21}$$

Equating these we have that

$$aX_1 = -\sum_{i=1}^{n} \alpha_i u_i X_i \implies aX_1 + \sum_{i=2}^{n} \alpha_i u_i X_i = 0. \tag{8.1.22}$$

Since the X_i form a basis this implies that $a = \alpha_i u_i = 0$ for all i by linear independence. Thus, $a = 0$ and ad_X has only zero eigenvalues and is semisimple, so it is zero. Hence, $\mathfrak{h}$ is abelian. $\qquad \square$

Corollary 8.1.23 Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{h} \subseteq \mathfrak{g}$ a subalgebra consisting of only semisimple elements. Then if any subalgebra strictly containing $\mathfrak{h}$ is nonabelian $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$.

Proof. By the contrapositive of Lemma 8.1.13 the subalgebras containing $\mathfrak{h}$ cannot contain only semisimple elements, and therefore $\mathfrak{h}$ is a maximal subalgebra containing only semisimple elements, and thus is a Cartan subalgebra of $\mathfrak{g}$. □

This corollary gives us a method of checking if a given subalgebra is Cartan that is often easier to check than directly checking that there is no larger subalgebra.

Definition 8.1.24 — Normaliser Let $\mathfrak{g}$ be a Lie algebra and let $\mathfrak{m} \subseteq$ be a subalgebra. The **normaliser**, $N_{\mathfrak{g}}(\mathfrak{m})$, of $\mathfrak{m}$ in $\mathfrak{g}$ is the largest subalgebra of $\mathfrak{g}$ such that $[N_{\mathfrak{g}}(\mathfrak{m}), \mathfrak{m}] \subseteq \mathfrak{m}$. That is,

$$N_{\mathfrak{g}}(\mathfrak{m}) = \{X \in \mathfrak{g} \mid [X, \mathfrak{m}] \subseteq \mathfrak{m}\}. \tag{8.1.25}$$

Lemma 8.1.26 The normaliser of a subalgebra is again a subalgebra. Further the normaliser of a subalgebra in $\mathfrak{g}$ is equal to $\mathfrak{g}$ if and only if the subalgebra is an ideal.

Proof. Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{m} \subseteq \mathfrak{g}$ a subalgebra. To show that $N_{\mathfrak{g}}(\mathfrak{m})$ is a subalgebra we need to show that it is a subspace which is closed under the bracket. First note that if $X \in N_{\mathfrak{g}}(\mathfrak{m})$ then $[X, \mathfrak{m}] \subseteq \mathfrak{m}$, that is $[X, Z] \in \mathfrak{m}$ for all $Z \in \mathfrak{m}$. Then it follows by linearity that $[aX, Z] = a[X, Z] \in \mathfrak{m}$ for any $a \in \mathbb{C}$. If $Y \in N_{\mathfrak{g}}(\mathfrak{m})$ also then $[Y, Z] \in \mathfrak{m}$ for all $Z \in \mathfrak{m}$ and again by linearity we have

$$[X + Y, Z] = [X, Z] + [Y, Z] \in \mathfrak{m}. \tag{8.1.27}$$

So, $N_{\mathfrak{g}}(\mathfrak{m})$ is a subspace.
Now suppose that $X, Y \in N_{\mathfrak{g}}(\mathfrak{m})$. Then for $Z \in \mathfrak{m}$ using the Jacobi identity we hvae

$$[[X, Y], Z] = -[Z, [X, Y]] = -[[Z, X], Y] - [X, [Z, Y]] \tag{8.1.28}$$

and each of these brackets is in $N_{\mathfrak{g}}(\mathfrak{m})$ since $[Z, X]$ and $[Z, Y]$ are in $\mathfrak{m}$ and so the bracket of these with Y and X respectively is again in $\mathfrak{m}$. Thus, $N_{\mathfrak{g}}(\mathfrak{m})$ is closed under the bracket so it is a subalgebra.
Now suppose that $\mathfrak{m}$ is an ideal. Then $[X, Z] \in \mathfrak{m}$ for any $X \in \mathfrak{g}$ and $Z \in \mathfrak{m}$, so $N_{\mathfrak{g}}(\mathfrak{m}) = \mathfrak{g}$. Conversely, suppose that $N_{\mathfrak{g}}(\mathfrak{m}) = \mathfrak{g}$. Then $[X, Z] \in \mathfrak{m}$ for any $X \in \mathfrak{g}$ and $Z \in \mathfrak{m}$, so $N_{\mathfrak{g}}(\mathfrak{m})$ is all of $\mathfrak{g}$. □

From the last part of this proof we see that $N_{\mathfrak{g}}(\mathfrak{m}) = \mathfrak{g}$ is just a slightly more complicated way of writing the absorbing property of an ideal.

An important property of Cartan subalgebras is that they are equal to their own normalisers. That is, if $X \in \mathfrak{g}$ such that $[X, \mathfrak{h}] \subset \mathfrak{h}$ then $X \in \mathfrak{h}$. This follows immediately from the maximality of Cartan subalgebras, since the normaliser of a subalgebra of semisimple elements consists only of semisimple elements.

8.1.3.1 Examples of Cartan Subalgebras

Cartan Subalgebras of $\mathfrak{sl}_2$. The simplest example of a Cartan subalgebra is the[1] Cartan subalgebra of $\mathfrak{sl}_2$, which is simply $\mathfrak{h} = \mathbb{C}H$. To prove this is a Cartan subalgebra we need to show that

[1] a slight abuse of language, but one that is common

a) every element of $\mathfrak{h}$ is semisimple; and

b) $\mathfrak{h}$ is maximal.

First note that we need to show that all elements of $\mathfrak{h}$ are ad-semisimple, that is that ad_X is diagonalisable, not that X is diagonalisable. An arbitrary element of $\mathfrak{h}$ is of the form cH for some $c \in \mathbb{C}$. We then have

$$\text{ad}_{cH}(E) = [cH, E] = 2cE, \tag{8.1.29}$$
$$\text{ad}_{cH}(H) = [cH, H] = 0, \tag{8.1.30}$$
$$\text{ad}_{cH}(F) = [cH, F] = -2cF. \tag{8.1.31}$$
$$\tag{8.1.32}$$

This shows that in the basis $\{E, H, F\}$ we have

$$\text{ad}_{cH} = \begin{bmatrix} 2c & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2c \end{bmatrix} \tag{8.1.33}$$

which shows that ad_{cH} is diagonalisable, and thus cH is semisimple.

Now suppose that $\tilde{\mathfrak{h}}$ is some other subalgebra of $\mathfrak{sl}_2$ such that $\mathfrak{h} \subseteq \tilde{\mathfrak{h}}$ and $\tilde{\mathfrak{h}}$ is abelian. Then for all $Y \in \tilde{\mathfrak{h}}$ we have $\text{ad}_X(Y) = 0$, which implies that all $Y \in \tilde{\mathfrak{h}}$ lie in the zero eigenspace of ad_X for all $X \in \mathfrak{h}$. However, from our calculations above we see that the zero eigenspace of ad_X is simply the span of H, which is just $\mathfrak{h}$. Thus, $\tilde{\mathfrak{h}} \subseteq \mathfrak{h}$, and so we have $\tilde{\mathfrak{h}} = \mathfrak{h}$. This proves maximality.

It turns out that this is really the *only* example of a Cartan subalgebra of $\mathfrak{sl}_2$ up to a very specific type of isomorhpism. The Lie group SL_2 consists of 2×2 matrices over $\mathbb{C}$ with unit determinant. There is an action of SL_2 on $\mathfrak{sl}_2$ given by conjugation. For $A \in \text{SL}_2$ denote this action by c_A, that is $c_A(X) = AXA^{-1}$ for $X \in \mathfrak{sl}_2$. Since the bracket in $\mathfrak{sl}_2$ is just the commutator it's easy to show that c_A preserves the bracket:

$$[c_A(X), c_A(Y)] = [AXA^{-1}, AYA^{-1}] \tag{8.1.34}$$
$$= AXA^{-1}AYA^{-1} - AYA^{-1}AXA^{-1} \tag{8.1.35}$$
$$= AXYA^{-1} - AYXA^{-1} \tag{8.1.36}$$
$$= A[X, Y]A^{-1} \tag{8.1.37}$$
$$= c_A([X, Y]). \tag{8.1.38}$$

Clearly conjugation is linear. Thus, c_A is an automorphism of $\mathfrak{sl}_2$.

We claim that $\text{span}_{\mathbb{C}}\{AHA^{-1}\}$ for some fixed $A \in \text{SL}_2$ is also a Cartan subalgebra of $\mathfrak{sl}_2$. This follows in exactly the same way as the proof that $\text{span}_{\mathbb{C}}\{H\}$ is a Cartan subalgebra. The conjugate As just factor out at each step. In fact, one can show that all Cartan subalgebras of $\mathfrak{sl}_2$ arise in this way.

Cartan Subalgebras of $\mathfrak{sl}_n$.   The next simplest example is the generalisation of this to $\mathfrak{sl}_n$. Note that $\mathfrak{sl}_n$ consists of all $n \times n$ traceless matrices over $\mathbb{C}$. The Cartan subalgebra, $\mathfrak{h}$, consists of all such matrices which are also diagonal. A basis for this is $H_i = E_{ii} - E_{i+1,i+1}$ for $i = 1, \dots, n-1$ where E_{ij} is the elementary matrix with a 1 in the ith row and jth column and zero everywhere else.

To show that this is indeed the Cartan subalgebra we need to show that each H_i is ad-semisimple. Rather than computing the action of H_i on $\mathfrak{sl}_n$ it's simpler to find a simultaneous eigenbasis of the H_i for $\mathfrak{sl}_n$. This can be done because the adjoint action of the H_is all commute with each other. The H_i give us the diagonal matrices in $\mathfrak{sl}_n$. To get the other matrices we can simply take E_{ij} with $i \neq j$. Then a basis for $\mathfrak{sl}_n$ is

$$\{H_i \mid i = 1, \dots, n-1\} \cup \{E_{ij} \mid i, j \in \{1, \dots, n\} \text{ and } i \neq j\}. \tag{8.1.39}$$

This is an eigenbasis for the adjoint action of the H_is which can be shown using the relation

$$E_{ij}E_{k\ell} = \delta_{jk}E_{i\ell}. \tag{8.1.40}$$

We then have

$$\begin{aligned}
\mathrm{ad}_{H_i}(H_j) &= (E_{ii} - E_{i+1,i+1})(E_{jj} - E_{j+1,j+1}) & \text{(8.1.41)}\\
&= E_{ii}E_{jj} - E_{i+1,i+1}E_{jj} - E_{ii}E_{j+1,j+1} + E_{i+1,i+1}E_{j+1,j+1} & \text{(8.1.42)}\\
&= \delta_{ij}E_{ij} - \delta_{i+1,j}E_{i+1,j} - \delta_{i,j+1}E_{i,j+1} + \delta_{i+1,j+1}E_{i+1,j+1} & \text{(8.1.43)}
\end{aligned}$$