Willoughby Seago

Notes from

Lie Theory

September 26th, 2024

UNIVERSITY OF GLASGOW

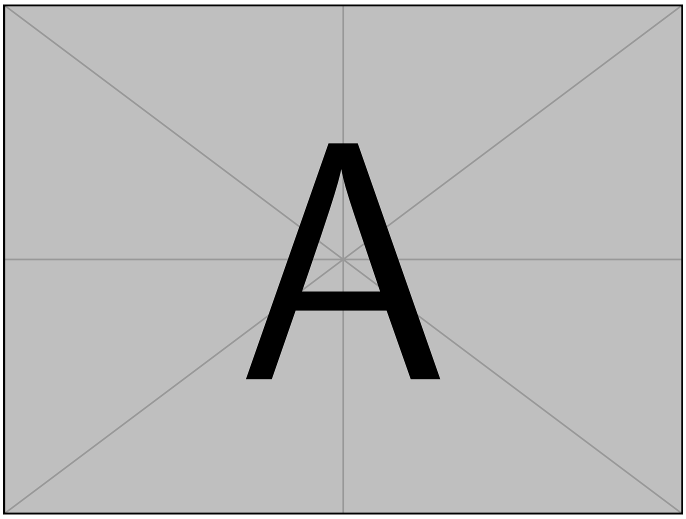
Lie Theory

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 10:55 on September 27, 2024.

[1] https://www.youtube.com/playlist?list=PLVMgvCDIRy1zJ6iJHLC_lWgIwY4AhviZ-



Chapters

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.

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.

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 $V \cong U_1 \oplus U_2$. □

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

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

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. □

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 $u \colon I \to V$ 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

$$U_i = \operatorname{span}\{x^i\} \coloneqq \{ax^i \mid a \in \Bbbk\} \tag{2.1.15}$$

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 $U_i U_j \coloneqq \{u_i u_j \mid u_i \in U_i \text{ and } u_j \in U_j\} = U_{i+j}$. 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 categories) 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 $j \in I$:

$$\begin{tikzcd} U_j \arrow[r, hook, "\iota_j"] \arrow[dr, "f_j"'] & \bigoplus_{i\in I} U_i \arrow[d, dashed, "\exists!\, f"] \\ & W. \end{tikzcd} \tag{2.1.19}$$

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 $\mathcal{B} = \{v_1, \dots, v_n\}$. Recall that if $T\colon V \to V$ is a linear operator then its matrix in basis $\mathcal{B}$ is the matrix

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

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 $\mathcal{B} = \{v_1, \dots, v_n\}$ 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 $U_i = \operatorname{span}\{v_i\}$. 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 $T|_{U_i} \colon U_i \to U_i$, 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 $\mathrm{id}_W \colon W \to W$ 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 $T \colon V \to V$ 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, $T \colon V \to V$, 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

$$T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{2.2.11}$$

in the standard basis

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

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

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

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

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

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

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

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

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

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

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

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 $\{\alpha_1, \dots, \alpha_{s-1}\}$. 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 $\{\alpha_1, \dots, \alpha_s\}$. □

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

$$T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{2.2.38}$$

As we saw before (Example 2.2.10) T is not diagonalisable, and $\bigoplus_{\alpha \in \mathbb{C}} W_\alpha = W_0 = \operatorname{span}\{e_1\} \neq \mathbb{C}^2$. 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

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

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

$$\begin{bmatrix} \alpha & 1 & & & \\ & \alpha & 1 & & \\ & & \ddots & \ddots & \\ & & & \alpha & 1 \\ & & & & \alpha \end{bmatrix}. \tag{2.2.41}$$

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 $T\colon V \to V$ 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,

$$[-,-]: \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} \tag{3.1.2}$$

$$(X, Y) \mapsto [X, Y], \tag{3.1.3}$$

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 $\operatorname{char}\Bbbk \neq 2$, equipped with an *antisymmetric* bracket, $[-,-]: \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$. 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 $[X,X] = 0$. □

This distinction is subtle, and ultimately not that important since we'll mostly concern ourselves with $\Bbbk = \mathbb{C}$, which has characteristic 0 (and $\operatorname{char}\mathbb{R} = 0$ 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 $D: \mathfrak{g} \to \mathfrak{g}$ 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 satisfying the Leibniz rule, and the "adjoint representation" is simply $\mathfrak{g}$ acting on itself where X acts on Y by $X \cdot Y \mapsto [X,Y]$.

Now[1] that we've defined Lie algebras, an algebraic object, we should define

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

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. □

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.

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

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

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 three things to show:

1. The bracket on $\mathfrak{g}/\mathfrak{i}$ is well-defined;
2. The bracket on $\mathfrak{g}/\mathfrak{i}$ is alternating;
3. 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: ALTERNATING
Let $X \in \mathfrak{g}$, then

$$[X + \mathfrak{i}, X + \mathfrak{i}] = [X, X] + \mathfrak{i} = 0 + \mathfrak{i} = \mathfrak{i} = 0 \tag{3.2.11}$$

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

STEP 3: JACOBI IDENTITY
Let $X, Y, Z \in \mathfrak{g}$,

$$[X + \mathfrak{i}, [Y + \mathfrak{i}, Z + \mathfrak{i}]] = [X + \mathfrak{i}, [Y, Z] + \mathfrak{i}] = [X, [Y, Z]] + \mathfrak{i}. \tag{3.2.12}$$

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

$$([X, [Y, Z]] + \mathfrak{i}) + ([Y, [Z, X]] + \mathfrak{i}) + ([Z, [X, Y]] + \mathfrak{i}). \tag{3.2.13}$$

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

$$([X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]]) + \mathfrak{i} = 0 + \mathfrak{i} = 0 \tag{3.2.14}$$

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

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 $\mathfrak{g} = \Bbbk X = \operatorname{span}\{X\}$ 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.
- $\mathfrak{g} = \operatorname{span}\{X, Y, Z\}$ with $[X, Y] = Z$ and $\langle X|Z\rangle = [Y, Z] = 0$.
- $\mathfrak{g} = \operatorname{span}\{X, Y, Z\}$ with $[X, Y] = Y$ and $[X, Z] = [X, Z] = 0$.

- $\mathfrak{g}_b = \operatorname{span}\{X, Y, Z\}$ 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 $b = 1/b'$.
- $\mathfrak{g} = \operatorname{span}\{X, Y, Z\}$ with $[X, Y] = Y$, $[X, Z] = Y + Z$, and $[Y, Z] = 0$.
- $\mathfrak{g} = \operatorname{span}\{X, Y, Z\}$ with $[X, Y] = 2X$, $[X, Z] = -2Z$, and $[Y, Z] = X$.

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 $\lambda \in \Bbbk$

$$\begin{aligned}
[X, Y + \lambda Z] &= X(Y + \lambda Z) - (Y + \lambda Z)X && (3.3.6)\\
&= XY + \lambda XZ - YX - \lambda ZX && (3.3.7)\\
&= XY - YX + \lambda(XZ - ZX) && (3.3.8)\\
&= [X, Y] + \lambda[X, Z]. && (3.3.9)
\end{aligned}$$

The Jacobi identity follows by some algebra. First note that

$$\begin{aligned}[X,[Y,Z]] &= [X, YZ - ZY] && (3.3.10)\\ &= X(YZ - ZY) - (YZ - ZY)X && (3.3.11)\\ &= XYZ - XZY - YZX + ZYX. && (3.3.12)\end{aligned}$$

Then, using the fact that the other terms are simply linear combinations of these, we have that the Jacobi relation is reduced to

$$\begin{aligned}&XYZ - XZY - YZX + ZYX + YZX - YXZ - ZXY + XZY\\ &\qquad + ZXY - ZYX - XYZ + YXZ \quad (3.3.13)\end{aligned}$$

and these terms all cancel to give zero. □

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

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

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 $\operatorname{End} V$ 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.

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

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

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

  where $I_n \in \mathrm{Mat}(n, \mathbb{R})$ 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:

- $\mathrm{GL}(n, \mathbb{R}) := \{A \in \mathrm{Mat}(n, \mathbb{R}) \mid \det A \neq 0\}$;
- $\mathfrak{sl}(n, \mathbb{R}) := \{A \in \mathrm{GL}(n, \mathbb{R}) \mid \det A = 1\}$;
- $\mathrm{O}(n, \mathbb{R}) := \{A \in \mathrm{GL}(n, \mathbb{R}) \mid A^{\mathsf{T}} A = I_n\}$;

- $\mathrm{SO}(n, \mathbb{R}) \coloneqq \{A \in \mathrm{GL}(n, \mathbb{R}) \mid A^{\mathsf{T}} A = I_n \text{ and } \det A = 1\}$.
- $\mathrm{U}(n) \coloneqq \{A \in \mathrm{GL}(n, \mathbb{C}) \mid U^* U = I_n\}$;
- $\mathrm{SU}(n) \coloneqq \{A \in \mathrm{GL}(n, \mathbb{C}) \mid U^* U = I_n \text{ and } \det U = 1\}$;
- $\mathrm{Sp}(n, \mathbb{R}) \coloneqq \{A \in \mathrm{GL}(2n, \mathbb{R}) \mid A^{\mathsf{T}} J_n A = J_n\}$.

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

- $\mathfrak{gl}(n, \mathbb{C}) \coloneqq \mathrm{Mat}(n, \mathbb{C})$;
- $\mathfrak{o}(n, \mathbb{C}) \coloneqq \mathfrak{so}(n, \mathbb{C}) = \{A \in \mathfrak{gl}(n, \mathbb{C}) \mid A^{\mathsf{T}} + A\} = 0$.
- $\mathfrak{sp}(n, \mathbb{C}) \coloneqq \{A \in \mathfrak{gl}(2n, \mathbb{C}) \mid A^{\mathsf{T}} J_n + J_n A = 0\}$.

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 $\operatorname{tr} \colon \mathfrak{gl}(n, \Bbbk) \to \mathfrak{gl}(n, \Bbbk)$ 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 $X \in \mathfrak{gl}(n, \mathbb{k})$ and $Y \in \mathfrak{sl}(n, \mathbb{k})$ we still have $\operatorname{tr}([X, Y]) = 0$, and thus $[X, Y] \in \mathfrak{sl}(n, \mathbb{k})$, meaning that $\mathfrak{sl}(n, \mathbb{k})$ is an ideal of $\mathfrak{gl}(n, \mathbb{k})$

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

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

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

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

- $\dim \mathfrak{sl}(n, \mathbb{R}) = n^2 - 1$, 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 $a_{11} + \dots + a_{n-1,n-1} = -a_{nn}$ to get $\operatorname{tr} A = a_{11} + \dots + a_{n-1,n-1} + a_{nn} = 0$. 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

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

  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).
- $\dim \mathfrak{sp}(n) = 2n^2 + n$, but showing this is a bit tricky.

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

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix} \tag{3.3.23}$$

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

$$H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad E = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \text{and} \qquad F = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}. \tag{3.3.24}$$

Then

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix} = aH + bE + cF. \tag{3.3.25}$$

A simple calculation of commutators shows that we have the relations

$$[H, E] = 2E, \qquad [H, F] = -2F, \qquad \text{and} \qquad [E, F] = H. \tag{3.3.26}$$

An important observation here is that if we define the map[2] $\mathrm{ad}_H : \mathfrak{sl}(2, \mathbb{C}) \to \mathfrak{sl}(2, \mathbb{C})$ 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 eigensspaces $V_0 = \mathbb{C}H = \mathfrak{h}$, $V_2 = \mathbb{C}E$, and $V_{-2} = \mathbb{C}F$, giving the decomposition

$$\mathfrak{sl}(2, \mathbb{C}) = \mathbb{C}H \oplus \mathbb{C}E \oplus \mathbb{C}F = \mathfrak{h} \oplus V_2 \oplus V_{-2} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \mathbb{C} \setminus \{0\}} V_\alpha. \tag{3.3.27}$$

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[3] that replacing the generalised eigenspaces with something called root spaces makes this construction generalise to finite dimensional semisimple Lie algebras over $\mathbb{C}$, and further that the subalgebras appearing in this decomposition are all simple, so a classification of simple Lie algebras extends to a classification of semisimple Lie algebras. 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 $\mathfrak{sl}(2, \mathbb{C})$.

[2] This is the so called adjoint representation.

[3] so ignore any terminology here that is unfamiliar

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

Remark 3.3.29 Here $\mathfrak{b}$ stands for "Borel" because $\mathfrak{b}(n, \Bbbk)$ is a **Borel subalgebra**, meaning it's a maximal solvable subalgebra (we'll define solvable shortly). The $\mathfrak{n}$ stands for "nilpotent" because $\mathfrak{n}(n, \Bbbk)$ is a (maximal) nilpotent algebra.
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 $\mathfrak{n}(n, \Bbbk)$ is an ideal of $\mathfrak{b}(n, \Bbbk)$, 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 $\mathfrak{g} = \mathfrak{b}(n, \Bbbk)/\mathfrak{n}(n, \Bbbk)$. Two upper triangular matrices in $\mathfrak{b}(n, \Bbbk)$ are identified in $\mathfrak{g}$ if their difference is a strictly upper triangular matrix in $\mathfrak{n}(n, \Bbbk)$. This means that they must have the same diagonal. Suppose then that $X, Y \in \mathfrak{b}(n, \Bbbk)$ 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

$$(XY)_{ii} = \sum_{j=1}^{n} X_{ij} Y_{ji}. \tag{3.3.30}$$

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 $[X, Y] \in \mathfrak{n}(n, \Bbbk)$ for all $X, Y \in \mathfrak{b}(n, \Bbbk)$. This means that in the quotient $\mathfrak{g}$ we identify $[X, Y]$ with $[X, Y] + \mathfrak{n}(n, \Bbbk) = 0 + \mathfrak{n}(n, \Bbbk)$, which is zero, and so this quotient is abelian.

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