Lie algebras

Jasper Ty

What is this?

These are notes based on my reading of Humphreys's "Introduction to Lie Algebras and Representation Theory".

Contents

Table of notation

$-$	A wildcard variable
$[n]$	The set $\{1, \ldots, n\}$
$\mathbb{Z}_{\geq 0}$	The set of nonnegative integers
$\mathbb{Z}_{>0}$	The set of positive integers
V	A generic vector space
$\mathbf{k}$	A generic field
$\mathrm{Mat}_n(\mathbf{k})$	The ring of $n \times n$ matrices over the field $\mathbf{k}$
e_{ij}	The standard basis of Mat_n
$x.v$	The action of x on v.
δ_{ij}	The Kronecker delta
$[-]^?$	The Iverson bracket
$\mathfrak{gl}$	The general linear Lie algebra
$\mathfrak{sl}$	The special linear Lie algebra
$\mathfrak{o}$	The orthogonal Lie algebra
$\mathfrak{sp}$	The symplectic Lie algebra
$\mathfrak{t}$	The Lie algebra of upper triangular matrices
$\mathfrak{n}$	The Lie algebra of strictly upper triangular matrices

1 Basic definitions and examples

Convention 1.0.1. All vector spaces considered are finite dimensional and no assumptions are made yet about underlying fields. We use V and $\mathbf{k}$ to denote generic vector spaces and fields respectively.

We will often use . to denote action in general, so if $v \in V$ and $x \in \operatorname{End} V$, we will define

$$x.v := x(v).$$

1.1 Lie algebras

Definition 1.1.1. A **Lie algebra** $\mathfrak{g}$ is a vector space equipped with a product

$$\begin{aligned} [-,-] : \mathfrak{g} \times \mathfrak{g} &\to \mathfrak{g}, \\ (x, y) &\mapsto [xy], \end{aligned}$$

such that

(L1) $[-,-]$ is bilinear,

(L2) $[xx] = 0$ for all $x \in \mathfrak{g}$, and

(L3) $[x[yz]] + [y[zx]] + [z[xy]] = 0$.

We refer to $[xy]$ as the **bracket** or the **commutator** of x and y.

(L3) is referred to as the *Jacobi identity*.

As an exercise in using this definition, we show the following:

Proposition 1.1.2. Brackets are anticommutative, i.e

$$[xy] = -[yx]. \tag{L2'}$$

is a relation in any Lie algebra.

Proof. By (L2), we have that

$$[x + y, x + y] = 0,$$

and by (L1),

$$[xx] + [xy] + [yx] + [yy] = 0.$$

By (L2) again,

$$[xy] + [yx] = 0,$$

which completes the proof. □

We will look at our first example of a Lie algebra, closely associated with the **general linear group** $\mathrm{GL}(V)$ of invertible endomorphisms of a vector space V.

Definition 1.1.3 ($\mathfrak{gl}$, abstractly)**.** Let V be a vector space. The **general linear algebra** $\mathfrak{gl}(V)$ is defined to be the Lie algebra with underlying vector space $\operatorname{End} V$ and bracket given by

$$[xy] = xy - yx$$

defined with $\operatorname{End} V$'s natural ring structure.

$\operatorname{End} V$'s aforementioned ring structure is exactly that of $n \times n$ matrices, where $n = \dim V$. Then, the following definition gives us a more concrete avatar of $\mathfrak{gl}$, and is in a sense "the only" finite dimensional $\mathfrak{gl}$.

Definition 1.1.4 ($\mathfrak{gl}$, concretely)**.** Let $\mathbf{k}$ be some field and let n be a positive integer. The **general linear algebra** $\mathfrak{gl}_n(\mathbf{k})$ is defined

$$\mathfrak{gl}_n(\mathbf{k}) := \mathfrak{gl}\big(\operatorname{Mat}_n(\mathbf{k})\big).$$

In this setting, we can easily compute the bracket of $\mathfrak{gl}$ relative to its standard basis:

Proposition 1.1.5. Let $\{e_{ij}\}_{i,j=0}^{n}$ be the standard basis of $\mathfrak{gl}_n(\mathbf{k})$. Then

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

where δ is the Kronecker delta.

Proof. Using the Iverson bracket (see Definition 8.1.1),

$$(e_{pq})_{ij} = [p = i \wedge q = j]^{?} = [p = i]^{?}\,[q = j]^{?}$$

and so

$$(e_{pq}e_{rs})_{ij} = \sum_{k=1}^{n} (e_{pq})_{ik}(e_{rs})_{kj}$$

$$
\begin{aligned}
&= \sum_{k=1}^{n} [p = i \wedge q = k]^? \, [r = k \wedge s = j]^? \\
&= \sum_{k=1}^{n} \left([q = k]^? \, [r = k]^? \right) [p = i]^? \, [s = j]^? \\
&= \left(\sum_{k=1}^{n} [q = r = k]^? \right) [p = i \wedge s = j]^? \\
&= \delta_{qr} (e_{ps})_{ij}.
\end{aligned}
$$

So $e_{pq} e_{rs} = \delta_{qr} e_{ps}$. Similarly, $e_{rs} e_{pq} = \delta_{sp} e_{rq}$. □

Importantly, many Lie algebras, and in fact all the Lie algebras we are concerned with, occur as subalgebras of the general linear algebra— a **subalgebra** of a Lie algebra $\mathfrak{g}$ is a subspace of $\mathfrak{g}$ that is closed under $\mathfrak{g}$'s bracket.

Definition 1.1.6. A **linear Lie algebra** is a subalgebra of $\mathfrak{gl}_n(\mathbf{k})$ for some n.

All finite dimensional Lie algebras are linear, in the sense that they are isomorphic to some linear Lie algebra.

1.2 Examples

We have four distinguished families of Lie algebras:

$$\mathsf{A}_\ell, \qquad \mathsf{B}_\ell, \qquad \mathsf{C}_\ell, \qquad \mathsf{D}_\ell.$$

These are parameterized by a positive integer ℓ, and they classify all but five of the so-called **semisimple Lie algebras**.

1.2.1 Type A: the special linear algebra

Definition 1.2.1. Let V be a vector space with basis $\mathbf{v} = (v_1, \ldots, v_n)$ and dual basis $\mathbf{v}^* = (v^1, \ldots, v^n)$ The **trace** $\operatorname{tr} x$ of an endomorphism $x \in \operatorname{End} V$ of V is defined to be the sum

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

In other words, it is the sum of the diagonal entries of the matrix representation of x. The trace is independent of the basis used to compute it (see Theorem 8.2.19 in the Appendix), hence it is a well defined quantity.

Definition 1.2.2 (The type A_ℓ Lie algebra)**.** Let V have dimension $n = \ell + 1$. We define A_ℓ to be the **special linear algebra** $\mathfrak{sl}(V)$, the set of all **traceless** endomorphisms of V, which means

$$A_\ell := \mathfrak{sl}(V) := \left\{x \in \mathfrak{gl}(V) : \operatorname{tr} x = 0\right\}.$$

As is the case with $\mathfrak{gl}(V)$ and $\mathfrak{gl}_n(\mathbf{k})$, we also define

$$A_\ell := \mathfrak{sl}_{\ell+1}(\mathbf{k}) := \left\{x \in \mathfrak{gl}_{\ell+1}(\mathbf{k}) : \operatorname{tr} x = 0\right\}$$

and will refer to them interchangeably.

This algebra is so named because of its connection with the **special linear group** $\mathrm{SL}(V)$, a distinguished subgroup of $\mathrm{GL}(V)$. Unsurprisingly, $\mathfrak{sl}(V)$ shares a similar relationship to $\mathfrak{gl}(V)$.

Proposition 1.2.3. $\mathfrak{sl}(V)$ is a subalgebra of $\mathfrak{gl}(V)$.

Proof. The trace is a linear operator $\operatorname{tr} : \mathfrak{gl}_n(\mathbf{k}) \to \mathbf{k}$. Since the kernel of a linear operator is a subspace of its domain, we conclude that $\mathfrak{sl}_n(\mathbf{k}) = \ker \operatorname{tr}$ is a subspace of $\mathfrak{gl}$.

Finally, the fact that $\operatorname{tr}(xy - yx) = \operatorname{tr}(xy) - \operatorname{tr}(yx) = 0$ for *all* $x, y \in \mathfrak{gl}_n(\mathbf{k})$ means that $\mathfrak{gl}_n(\mathbf{k})$'s Lie bracket is closed in $\mathfrak{sl}_n(\mathbf{k})$. □

Lastly, we will compute the dimension of $\mathfrak{sl}(V)$. Firstly, it has to be strictly less than that of $\mathfrak{gl}(V)$'s, as it is a proper subalgebra of $\mathfrak{gl}(V)$. Hence

$$\dim \mathfrak{sl}(V) < \dim \mathfrak{gl}(V) = (\ell+1)^2.$$

So

$$\dim \mathfrak{sl}(V) \leq (\ell+1)^2 - 1 = \ell(\ell+2)$$

However, we can explicitly name $\ell(\ell+2)$ linearly independent elements of $\mathfrak{sl}_n(\mathbf{k})$:

1. All the off-diagonal entries e_{ij} where $i \neq j$— there are $(\ell+1)^2 - (\ell+1) = \ell^2 + \ell$ of these.

2. All of the elements $e_{ii} - e_{i+1,i+1}$, of which there are $(\ell+1) - 1 = \ell$.

So,

$$\dim \mathfrak{sl}(V) \geq \ell + 2 + \ell + \ell = \ell(\ell+2).$$

And, putting it together, we have proven:

Proposition 1.2.4.

$$\dim \mathsf{A}_\ell = \dim \mathfrak{sl}(V) = \dim \mathfrak{sl}_n(\mathbf{k}) = \ell(\ell+2).$$

1.2.2 The rest; bilinear forms

Types B, C, and D are all defined with regards to certain bilinear forms.

Definition 1.2.5. Let V be a vector space over the field $\mathbf{k}$.
A **bilinear form** is a function $\omega : V \times V \to \mathbf{k}$ that is bilinear, i.e linear in each argument separately.

Definition 1.2.6. Let V be a vector space with a bilinear form ω.
If x is an endomorphism of V, we say that x is **ω-skew** if

$$\omega\Big(x.u, v\Big) + \omega\Big(u, x.v\Big) = 0$$

for all $u, v \in V$.
We denote the set of all ω-skew endomorphisms of V by $\mathfrak{o}_\omega(V)$.

Theorem 1.2.7. Let $\mathfrak{o}_\omega(V)$ is a Lie subalgebra of $\mathfrak{gl}(V)$.

Proof. Let $x, y \in \mathfrak{o}_\omega(V)$, and let $u, v \in V$.

$$\begin{aligned}
&\omega\Big([xy].u, v\Big) + \omega\Big(u, [xy].v\Big) \\
&= \omega\Big((xy - yx).u, v\Big) + \omega\Big(u, (xy - yx).v\Big) \\
&= \Big(\omega\Big(xy.u, v\Big) + \omega\Big(u, xy.v\Big)\Big) - \Big(\omega\Big(yx.u, v\Big) + \omega\Big(u, yx.v\Big)\Big) \\
&= \Big(\omega\Big(xy.u, v\Big) + \omega\Big(u, xy.v\Big)\Big) - \Big(\omega\Big(u, xy.v\Big) + \omega\Big(xy.u, v\Big)\Big) \\
&= 0.
\end{aligned}$$

Hence $[xy] \in \mathfrak{o}_\omega(V)$. $\square$

1.2.3 Type B: the odd-dimensional orthogonal algebra

Definition 1.2.8. A **symmetric nondegenerate form** on a vector space V is a bilinear form $\omega : V \times V \to \mathbf{k}$ such that

(a) $\omega(v,u) = \omega(u,v)$, and

(b) $\omega(v,u) = 0$ for all $v \in V$ implies that $u = 0$.

Definition 1.2.9 (The type B_ℓ Lie algebra)**.** Let $\dim V = 2\ell + 1$, and let V be endowed with a symmetric nondegenerate form ω.

We define B_ℓ to be the **orthogonal algebra** $\mathfrak{o}(V)$:

$$\mathsf{B}_\ell := \mathfrak{o}(V) := \mathfrak{o}_\omega(V).$$

1.2.4 Type C: the symplectic algebra

Definition 1.2.10. A **symplectic form** on a vector space V is a function form $\omega : V \times V \to \mathbf{k}$ such that

(a) ω is bilinear,

(b) $\omega(v,u) = -\omega(u,v)$, and

(c) $\omega(v,u) = 0$ for all $v \in V$ implies that $u = 0$.

Definition 1.2.11 (The type C_ℓ Lie algebra)**.** Let $\dim V = 2\ell$, and let V be endowed with a symplectic form ω.

We define C_ℓ to be the **symplectic algebra** $\mathfrak{sp}(V)$:

$$\mathsf{C}_\ell := \mathfrak{sp}(V) := \mathfrak{o}_\omega(V).$$

In matrix form, we define

$$\mathsf{C}_\ell := \mathfrak{sp}_{2\ell}(\mathbf{k}) := \left\{x \in \mathfrak{gl}_{2\ell}(\mathbf{k}) : Jx + x^\top J = 0\right\}$$

where

$$J = \begin{pmatrix} 0 & I_\ell \\ -I_\ell & 0 \end{pmatrix}$$

is the standard symplectic form on $\mathbf{k}^{2\ell}$.

1.2.5 Type D: the even-dimensional orthogonal algebra

Definition 1.2.12 (The type D_ℓ Lie algebra)**.** Let $\dim V = 2\ell + 1$, and let V be endowed with a symmetric nondegenerate form ω.

We define D_ℓ to be the **orthogonal algebra** $\mathfrak{o}(V)$:

$$\mathsf{D}_\ell := \mathfrak{o}(V) := \mathfrak{o}_\omega(V).$$

1.3 Lie algebras from algebras

Definition 1.3.1 (Algebras over a field)**.** Let $\mathbf{k}$ be a field. An **algebra over k**, or a **k-algebra** is a $\mathbf{k}$-vector space equipped with a bilinear product.

We will use qualifiers like *associative* and *unital* to indicate that this product is associative and has unit respectively.

Put another way, a unital associative algebra over a field is

- a vector space with a compatible ring structure, (vector space + bilinear product)
- or a ring with a compatible vector space structure. (ring + bilinear scaling map)

For example, $\mathrm{Mat}_n(\mathbf{k})$ is a unital associative algebra over $\mathbf{k}$.

However, we don't in general expect algebras to have unit or to be associative— $\mathbb{R}^3$ with the cross product is neither unital nor associative. Hence, the following is clear:

Proposition 1.3.2. Lie algebras are algebras, with the product given by the Lie bracket.

To go along with this definition, we have notion of a homomorphism of algebras.

Definition 1.3.3. An **algebra homomorphism** $f : \mathscr{A} \to \mathscr{B}$ between two algebras $\mathscr{A}$ and $\mathscr{B}$ is a vector space homomorphism that respects the product, i.e

$$f(xy) = f(x)f(y)$$

for all $x, y \in \mathscr{A}$.

We say that an algebra homomorphism is an **algebra isomorphism** if it is also a vector space isomorphism.

For example, the determinant is an algebra homomorphism from $\mathrm{Mat}_n(\mathbf{k})$ to $\mathbf{k}$.

$\mathbf{k}$-algebras can be turned into Lie algebras by defining the bracket $[x\,y] := xy - yx$.

Definition 1.3.4. Let $\mathscr{A}$ be a $\mathbf{k}$-algebra. Then $\mathrm{Lie}[\mathscr{A}]$ is defined to be the Lie algebra whose underlying vector space is $\mathscr{A}$ and whose bracket is given by

$$[xy] \coloneqq xy - yx$$

for all $x, y \in \mathscr{A}$.

We can check the following nice fact:

Proposition 1.3.5. Let $\mathscr{A}$ and $\mathscr{B}$ be two $\mathbf{k}$-algebras, and let $\phi : \mathscr{A} \to \mathscr{B}$ be an algebra homomorphism.

Then ϕ is also a *Lie algebra homomorphism* (see Definition 2.2.1) between $\mathrm{Lie}[\mathscr{A}]$ and $\mathrm{Lie}[\mathscr{B}]$.

Proof.

$$\begin{aligned}
\phi\Big([xy]\Big) &= \phi(xy - yx) \\
&= \phi(xy) - \phi(yx) \\
&= \phi(x)\phi(y) - \phi(y)\phi(x) \\
&= [\phi(x)\phi(y)].
\end{aligned}$$

□

Hence $\mathrm{Lie}[-]$ is actually *functorial*, with mapping of arrows given by the identity map.

What happens when we consider $\mathrm{Lie}[\mathfrak{g}]$, where $\mathfrak{g}$ is *already* a Lie algebra?

Let the new bracket of $\mathrm{Lie}[\mathfrak{g}]$ be denoted by $[\![-,-]\!]$. Then

$$[\![xy]\!] = [xy] - [yx] = [xy] + [xy] = 2[xy]$$

for all $x, y \in \mathfrak{g}$.

Then $[\![-,-]\!] = 2[-,-]$. This fact actually characterizes Lie algebras.

Proposition 1.3.6. Let $\mathscr{A}$ be a $\mathbf{k}$-algebra with product $*$. If $\mathrm{Lie}[\mathscr{A}]$ has product $2*$, then $\mathscr{A}$ is a Lie algebra with bracket given by $[xy] = x * y$.

Proof. The product is bilinear by definition, so we have (L1).

Next, we check (L2):

$$x * x = \frac{2(x * x)}{2} = \frac{[xx]}{2} = 0.$$

And finally, in the exact same way, we check the Jacobi identity, (L3):

$$x * (y * z) + y * (z * x) + z * (x * y) = \frac{[x[yz]] + [y[xz]] + [z[xy]]}{4} = 0.$$

□

1.4 Derivations, the adjoint representation

Definition 1.4.1. Let $\mathscr{A}$ be a $\mathbf{k}$-algebra. A **derivation** of $\mathscr{A}$ is a linear map $\mathrm{d} : \mathscr{A} \longrightarrow \mathscr{A}$ which satisfies the *Leibniz rule*:

$$\mathrm{d}(xy) = x(\mathrm{d}y) + (\mathrm{d}x)y.$$

The collection of all derivations of $\mathscr{A}$ is denoted $\operatorname{Der}\mathscr{A}$.

Derivations play nicely with the vector space structure of $\operatorname{End}\mathscr{A}$ as well as with the bracket inherited from $\mathfrak{gl}(\mathscr{A})$.

Proposition 1.4.2. Let $\mathscr{A}$ be a $\mathbf{k}$-algebra. Then $\operatorname{Der}\mathscr{A}$ is a subspace of $\operatorname{End}\mathscr{A}$. Moreover, it is a subalgebra of $\mathfrak{gl}(\mathscr{A})$

Proof. If d and d′ are two derivations, then

$$\begin{aligned}(a\mathrm{d} + b\mathrm{d}')(xy) &= (a\mathrm{d})(xy) + (b\mathrm{d}')(xy)\\ &= x(a\mathrm{d}y) + (a\mathrm{d}x)y + x(b\mathrm{d}'y) + (b\mathrm{d}'x)y\\ &= x\Big(a\mathrm{d}y + b\mathrm{d}'y\Big) + \Big(a\mathrm{d}x + b\mathrm{d}'x\Big)y\\ &= x(a\mathrm{d} + b\mathrm{d}')(y) + (a\mathrm{d} + b\mathrm{d}')(x)y.\end{aligned}$$

Hence $a\mathrm{d} + b\mathrm{d}' \in \operatorname{Der}\mathscr{A}$, so $\operatorname{Der}\mathscr{A}$ is a subspace of $\operatorname{End}\mathscr{A}$.

Moreover,

$$\begin{aligned}&[\mathrm{d}\mathrm{d}'](xy)\\ &= (\mathrm{d}\mathrm{d}' - \mathrm{d}'\mathrm{d})(xy)\\ &= (\mathrm{d}\mathrm{d}')(xy) - (\mathrm{d}'\mathrm{d})(xy)\\ &= \mathrm{d}\Big(x(\mathrm{d}'y) + (\mathrm{d}'x)y\Big) - \mathrm{d}'\Big(x(\mathrm{d}y) + (\mathrm{d}x)y\Big)\\ &= \mathrm{d}\Big(x(\mathrm{d}'y)\Big) + \mathrm{d}\Big((\mathrm{d}'x)y\Big) - \mathrm{d}'\Big(x(\mathrm{d}y)\Big) - \mathrm{d}'\Big((\mathrm{d}x)y\Big)\\ &= x\mathrm{d}\mathrm{d}'y + \mathrm{d}x\mathrm{d}'y + \mathrm{d}'x\mathrm{d}y + \mathrm{d}\mathrm{d}'xy - x\mathrm{d}'\mathrm{d}y - \mathrm{d}'x\mathrm{d}y - \mathrm{d}x\mathrm{d}'y - \mathrm{d}'\mathrm{d}xy\end{aligned}$$

$$
\begin{aligned}
&= x\mathrm{d}\mathrm{d}' y + \mathrm{d}\mathrm{d}' x y - x\mathrm{d}'\mathrm{d} y - \mathrm{d}'\mathrm{d} x y \\
&= x\Big(\mathrm{d}\mathrm{d}' y - \mathrm{d}'\mathrm{d} y\Big) + \Big(\mathrm{d}\mathrm{d}' x - \mathrm{d}'\mathrm{d} x\Big) y \\
&= x\Big((\mathrm{d}\mathrm{d}' - \mathrm{d}'\mathrm{d}) y\Big) + \Big((\mathrm{d}\mathrm{d}' - \mathrm{d}'\mathrm{d}) x\Big) y \\
&= x\Big([\mathrm{d}\mathrm{d}'] y\Big) + \Big([\mathrm{d}\mathrm{d}'] x\Big) y.
\end{aligned}
$$

So $\operatorname{Der} \mathscr{A}$ is a subalgebra of $\mathfrak{gl}(\mathscr{A})$. □

We have a special representation of *any* Lie algebra, which is given by its action on itself.

Definition 1.4.3. The **adjoint representation** of a Lie algebra $\mathfrak{g}$ is the mapping

$$
\begin{aligned}
\operatorname{ad}_{\mathfrak{g}} : \mathfrak{g} &\to \operatorname{Der} \mathfrak{g} \\
x &\mapsto \operatorname{ad}_{\mathfrak{g}} x
\end{aligned}
$$

where $\operatorname{ad}_{\mathfrak{g}} x$ is defined to be the linear map

$$
\begin{aligned}
\operatorname{ad}_{\mathfrak{g}} x : \mathfrak{g} &\to \mathfrak{g} \\
y &\mapsto [xy].
\end{aligned}
$$

We will write $\operatorname{ad} x$ for $\operatorname{ad}_{\mathfrak{g}} x$ unless there is any ambiguity.

As a set, we define $\operatorname{ad} \mathfrak{g} \coloneqq \operatorname{ad}_{\mathfrak{g}}(\mathfrak{g}) \subseteq \mathfrak{gl}(\mathfrak{g})$.

Proposition 1.4.4. $\operatorname{ad} x$ is a derivation.

Proof. We start with the Jacobi identity (L3)

$$[x[yz]] + [y[zx]] + [z[xy]] = 0,$$

which, using the anticommutation relations $[y[zx]] = -[y[xz]]$ and $[z[xy]] = -[[xy]z]$, is equivalent to

$$[x[yz]] = [y[xz]] + [[xy]z].$$

But this is saying that

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

which is exactly the defining identity for derivations. □

1.5 Abstract Lie algebras

Definition 1.5.1. Let $\mathfrak{g}$ be a Lie algebra, and fix some basis $\{x_1, \ldots, x_n\}$ of $\mathfrak{g}$. We define $\mathfrak{g}$'s **structure constants** a_{ij}^k relative to this basis to be the basis coefficients of the Lie brackets of basis elements— the numbers such that

$$[x_i x_j] = \sum_{k=1}^{n} a_{ij}^k x_k.$$

Definition 1.5.2. An **abelian** Lie algebra $\mathfrak{g}$ is a Lie algebra with trivial bracket— $[xy] = 0$ for all $x, y \in \mathfrak{g}$.

Proposition 1.5.3. Let V be a vector space with basis $x_1, \ldots, x_n$, and let a_{ij}^k be an array of structure coefficients. Then, the bracket defined by a_{ij}^k gives V a Lie algebra structure if and only if

$$\begin{cases} a_{ii}^k = 0 \\ a_{ij}^k + a_{ji}^k = 0 \\ \sum_k a_{ij}^k a_{kl}^m + a_{jl}^k a_{ki}^m + a_{l}^k a_{kij}^m = 0 \end{cases}$$

for any values of i, j, k, l, m.

We will classify all the Lie algebras of dimensions 1 and 2.

Proposition 1.5.4. There are only two Lie algebras of dimension two up to isomorphism:

(a) The abelian two-dimensonal Lie algebra,

(b) and the Lie algebra with basis (x, y) and product $[x, y] = x$.

Proof. If $\mathfrak{g}$ is nonabelian, then $[xy] = ax + by$, where at least one of a, b is nonzero. Without loss of generality, let a be nonzero. Then

$$[[xy]y] = [ax + by, y] = a[xy].$$

Now put $u = [xy]$ and $v = a^{-1}y$. Then

$$[uv] = [[xy], (a^{-1}y)] = [xy] = u.$$

□

2 Ideals and homomorphisms

2.1 Ideals

Definition 2.1.1. A subspace $\mathfrak{i}$ of a Lie algebra $\mathfrak{g}$ is called an **ideal** of $\mathfrak{g}$ if $[xy] \in \mathfrak{i}$ for all $x \in \mathfrak{g}$ and $y \in \mathfrak{i}$.

Convention 2.1.2. In accordance with group theoretic notation, we write $\mathfrak{h} \leq \mathfrak{g}$ whenever $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$, and $\mathfrak{h} \trianglelefteq \mathfrak{g}$ whenever $\mathfrak{h}$ is an ideal of $\mathfrak{g}$.

The **sum** and the **bracket** of the ideals $\mathfrak{i}, \mathfrak{j}$ are defined in the obvious way:

$$\mathfrak{i} + \mathfrak{j} := \left\{ x + y : x \in \mathfrak{i}, y \in \mathfrak{j} \right\}, \quad [\mathfrak{i}\mathfrak{j}] := \left\{ \sum_{i=0}^{r} c_i [x_i y_i] : c_i \in \mathbf{k}, x_i \in \mathfrak{i}, y_i \in \mathfrak{j} \right\}.$$

Theorem 2.1.3. If $\mathfrak{a}$ and $\mathfrak{b}$ are ideals of a Lie algebra $\mathfrak{g}$, then so are $\mathfrak{a} + \mathfrak{b}$, $\mathfrak{a} \cap \mathfrak{b}$ and $[\mathfrak{a}\mathfrak{b}]$.

Proof. These are all easy to show.

($\mathfrak{a} + \mathfrak{b}$) Let $a + b \in \mathfrak{a} + \mathfrak{b}$ and $g \in \mathfrak{g}$. Then

$$[g, a + b] = \underbrace{[ga]}_{\in \mathfrak{a}} + \underbrace{[gb]}_{\in \mathfrak{b}}.$$

So $[g, a + b] \in \mathfrak{a} + \mathfrak{b}$.

($\mathfrak{a} \cap \mathfrak{b}$) Let $x \in \mathfrak{a} \cap \mathfrak{b}$ and $g \in \mathfrak{g}$. We have that $[gx] \in \mathfrak{a}$ and $[gx] \in \mathfrak{b}$ since $x \in \mathfrak{a}$ and $x \in \mathfrak{b}$ respectively. So $[gx] \in \mathfrak{a} \cap \mathfrak{b}$.

($[\mathfrak{a}\mathfrak{b}]$) Let $a \in \mathfrak{a}$, $b \in \mathfrak{b}$, and $g \in \mathfrak{g}$. We have that $[ab] \in [\mathfrak{a}\mathfrak{b}]$, and by the Jacobi identity,

$$[g[ab]] = [a[gb]] + [[ga]b],$$

hence $[g[ab]] \in [\mathfrak{a}\mathfrak{b}]$. Linearity extends this to the general case.

□

As a nice consequence, we have effectively shown the following:

Proposition 2.1.4. Ideals of a Lie algebra form a lattice, with order given by containment and whose join and meet correspond to sums and intersections of ideals respectively.

Proof. Ideals of $\mathfrak{g}$ are subspaces of $\mathfrak{g}$. By the previous theorem, it's clear that the set of ideals of $\mathfrak{g}$ are a *sublattice* of the set of subspaces of $\mathfrak{g}$. □

Definition 2.1.5. The **quotient of a Lie algebra** $\mathfrak{g}$ by an ideal $\mathfrak{i}$, denoted $\mathfrak{g}/\mathfrak{i}$, is defined to be the quotient of $\mathfrak{g}$ as a vector space by $\mathfrak{i}$ as a subspace, equipped with the product

$$[x+\mathfrak{i}, y+\mathfrak{i}] \coloneqq [xy]+\mathfrak{i}.$$

Proposition 2.1.6. $\mathfrak{g}/\mathfrak{i}$ is a Lie algebra.

Proof. These are all easy to check.

$$\begin{aligned}
[ax+by+\mathfrak{i}, z+\mathfrak{i}] &= \Big([ax+by, z]\Big)+\mathfrak{i} \\
&= \Big(a[x,z]+b[y,z]\Big)+\mathfrak{i} \\
&= \Big(a[x,z]+\mathfrak{i}\Big)+\Big(b[y,z]+\mathfrak{i}\Big) \\
&= a[x+\mathfrak{i}, z+\mathfrak{i}]+b[y+\mathfrak{i}, z+\mathfrak{i}].
\end{aligned}$$

$$[x+\mathfrak{i}, x+\mathfrak{i}] = [xx]+\mathfrak{i} = 0+\mathfrak{i}$$

□

2.2 Homomorphisms

There is a natural definition of a Lie algebra homomorphism— it's a map that respects brackets.

Definition 2.2.1. Let $\mathfrak{g}$ and $\mathfrak{h}$ be two Lie algebras. We say that a map $\phi : \mathfrak{g} \to \mathfrak{h}$ is a **Lie algebra homomorphism** if it is a linear map for which

$$\phi\Big([xy]\Big) = [\phi(x)\phi(y)]$$

for all $x, y \in \mathfrak{g}$. A **Lie algebra isomorphism** is a Lie algebra homomorphism that is also an isomorphism of vector spaces.

Definition 2.2.2. A **representation** of a Lie algebra $\mathfrak{g}$ is a Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{gl}(V)$ where V is some vector space.

2.3 Isomorphism theorems

Theorem 2.3.1 (Lie algebra isomorphism theorems)**.** Let $\mathfrak{g}$ and $\mathfrak{h}$ be Lie algerbas.

(a) If $\phi : \mathfrak{g} \to \mathfrak{h}$ is a homomorphism, then $\mathfrak{g}/\ker\phi \simeq \operatorname{im}\phi$. If $\mathfrak{i} \subseteq \ker\phi$ is an ideal of $\mathfrak{g}$, there exists a unique homomorphism $\overline{\phi} : \mathfrak{g}/\mathfrak{i} \to \mathfrak{h}$ that makes the following diagram commute:

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\phi"] \arrow[d, "\pi"'] & \mathfrak{h} \\ \mathfrak{g}/\mathfrak{i} \arrow[ur, "\overline{\phi}"'] & \end{tikzcd}$$

(b) If $\mathfrak{a}$ and $\mathfrak{b}$ are ideals of $\mathfrak{g}$ such that $\mathfrak{b} \subseteq \mathfrak{a}$, then $\mathfrak{a}/\mathfrak{b}$ is an ideal of $\mathfrak{g}/\mathfrak{b}$ and there is a natural isomorphism

$$(\mathfrak{g}/\mathfrak{b})/(\mathfrak{a}/\mathfrak{b}) \simeq \mathfrak{g}/\mathfrak{a}.$$

(c) If $\mathfrak{a}, \mathfrak{b}$ are ideals of $\mathfrak{g}$, there is a natural isomorphism

$$(\mathfrak{a}+\mathfrak{b})/\mathfrak{b} \simeq \mathfrak{a}/(\mathfrak{a}\cap\mathfrak{b}).$$

Proof. (a) The map

$$\begin{aligned} \overline{\phi} : \mathfrak{g}/\ker\phi &\to \operatorname{im}\phi \\ x + \ker\phi &\mapsto \phi(x) \end{aligned}$$

is the desired isomorphism $\mathfrak{g}/\ker\phi \simeq \operatorname{im}\phi$. We verify that it is well defined: let $x+\ker\phi = x'+\ker\phi$. Then there exists $k, k' \in \ker\phi$ such that $x+k = x'+k'$, and we have that

$$\phi(x) = \phi(x+k) = \phi(x+k') = \phi(x'),$$

so $\overline{\phi}$ is a well-defined function on the cosets in $\mathfrak{g}/\ker\phi$.

Next, we check that it respects brackets:

$$\begin{aligned}
\overline{\phi}\Big([x + \ker\phi, y + \ker\phi]\Big) &= \overline{\phi}\Big([x\,y] + \ker\phi\Big) \\
&= \phi\Big([x\,y]\Big) \\
&= [\phi(x)\phi(y)] \\
&= [\overline{\phi}\Big(x + \ker\phi\Big), \overline{\phi}\Big(y + \ker\phi\Big)].
\end{aligned}$$

Then, it is a homomorphism. To show that it is an isomorphism, we note that it has a trival kernel, trivially:

$$\ker\overline{\phi} = \{x + \ker\phi : x + \ker\phi = \ker\phi\} = \{0 + \ker\phi\}.$$

Now, let $\mathfrak{i}$ be an ideal of $\mathfrak{g}$ contained in $\ker\phi$. We define in a similar way

$$\begin{aligned}
\overline{\phi} : \mathfrak{g}/\mathfrak{i} &\longrightarrow \operatorname{im}\phi \\
x + \mathfrak{i} &\longmapsto \phi(x),
\end{aligned}$$

and via a similar argument as above, this map is well-defined. It is moreover clear that $\overline{\phi} \circ \pi = \phi$ and that it is the only such homomorphism that has these properties.

(b) Let $\mathfrak{a}$ and $\mathfrak{b}$ be ideals of $\mathfrak{g}$ such that $\mathfrak{b} \subseteq \mathfrak{a}$. We define the map

$$\begin{aligned}
\phi : \mathfrak{g}/\mathfrak{b} &\longrightarrow \mathfrak{g}/\mathfrak{a} \\
x + \mathfrak{b} &\longmapsto x + \mathfrak{a}.
\end{aligned}$$

This map is surjective. The kernel of this map is all the cosets $a + \mathfrak{b}$, namely the ideal $\mathfrak{a}/\mathfrak{b}$. Then, by (a),

$$(\mathfrak{g}/\mathfrak{b})(\mathfrak{a}/\mathfrak{b}) = (\mathfrak{g}/\mathfrak{b})/\ker\phi \simeq \operatorname{im}\phi = \mathfrak{g}/\mathfrak{a}.$$

(c) Let $\mathfrak{a}$ and $\mathfrak{b}$ be ideals of $\mathfrak{g}$. Define the map

$$\begin{aligned}
\phi : \mathfrak{a} &\longrightarrow (\mathfrak{a} + \mathfrak{b})/(\mathfrak{b}) \\
a &\longmapsto a + \mathfrak{b}.
\end{aligned}$$

This map is surjective, as, if $(a + b) + \mathfrak{b} \in (\mathfrak{a} + \mathfrak{b})/(\mathfrak{b})$, then

$$\phi(a) = a + \mathfrak{b} = a + (b + \mathfrak{b}) = (a + b) + \mathfrak{b}.$$

Moreover, since

$$\ker \phi = \mathfrak{a} \cap \mathfrak{b}$$

we have that, by (a) again,

$$(\mathfrak{a} + \mathfrak{b})/\mathfrak{b} = \operatorname{im} \phi \simeq \mathfrak{a}/\ker \phi = \mathfrak{a}/(\mathfrak{a} \cap \mathfrak{b}).$$

□

We have a useful theorem, usually considered a consequence of the third isomorphism theorem, which is important enough to state on its own:

Theorem 2.3.2 (Correspondence theorem)**.** Let $\mathfrak{i} \trianglelefteq \mathfrak{g}$. Then there is an order isomorphism

$$\begin{aligned} \text{subalgebras of } \mathfrak{g} \text{ containing } \mathfrak{i} &\leftrightarrow \text{subalgebras of } \mathfrak{g}/\mathfrak{i} \\ \mathfrak{h} &\leftrightarrow \mathfrak{h}/\mathfrak{i}. \end{aligned}$$

Proof. Similarly to Proposition 2.3.2, this is true on the level of a Lie algebra's vector space structure.

The maps

$$\mathfrak{h} \mapsto \mathfrak{h}/\mathfrak{i}$$

and

$$\mathfrak{h} \mapsto \bigcup_{h \in \mathfrak{h}} (h + \mathfrak{i})$$

are obviously mutual inverses, and this is an order is

Hence, it proves the bijection. TODO: details! □

Theorem 2.3.3. The adjoint representation $\operatorname{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is a representation of $\mathfrak{g}$.

Proof. ad is evidently linear. Next, we just check that it is a homomorphism:

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

$$
\begin{aligned}
&= [x[yz]] + [y[zx]] \\
&= [[xy]z] \\
&= \mathrm{ad}[xy].z.
\end{aligned}
$$

□

Corollary 2.3.4. Any simple Lie algebra is isomorphic to a linear Lie algebra.

Proof. Let $\mathfrak{g}$ be a Lie algebra. We have that

$$\ker \mathrm{ad} = \left\{x \in \mathfrak{g} : \mathrm{ad}\, x = 0\right\} = \left\{x \in \mathfrak{g} : [xy] = 0 \text{ for all } y \in \mathfrak{g}\right\} = Z(\mathfrak{g}).$$

Hence, if $\mathfrak{g}$ is simple, i.e if $Z(\mathfrak{g}) = 0$, then ad has a trivial kernel, so it is an isomorphism. □

3 Automorphisms

Definition 3.0.1. A **automorphism** of a Lie algebra $\mathfrak{g}$ is an isomorphism $\mathfrak{g} \to \mathfrak{g}$.

Proposition 3.0.2. Let V be a vector space and let $g \in \mathrm{GL}(V)$. Then the map

$$x \mapsto gxg^{-1}$$

is an automorphism of $\mathfrak{gl}(V)$.

Proof. The aforementioned map is a vector space isomorphism, with explicit inverse

$$x \mapsto g^{-1}xg$$

and it is a homomorphism, as

$$
\begin{aligned}
g[xy]g^{-1} &= g\left(xy - yx\right)g^{-1} \\
&= \left(gxyg^{-1}\right) - \left(gyxg^{-1}\right) \\
&= \left(gxg^{-1}gyg^{-1}\right) - \left(gyg^{-1}gxg^{-1}\right) \\
&= [gxg^{-1}, gyg^{-1}].
\end{aligned}
$$

□

4 Solvable and nilpotent Lie algebras

4.1 The derived series, solvability

Definition 4.1.1. The **derived series** of a Lie algebra $\mathfrak{g}$ is a sequence of ideals $\mathfrak{g}^{(0)}, \mathfrak{g}^{(1)}, \ldots$ defined

$$\begin{cases} \mathfrak{g}^{(0)} := \mathfrak{g} \\ \mathfrak{g}^{(i)} := [\mathfrak{g}^{(i-1)}\mathfrak{g}^{(i-1)}] \end{cases}.$$

In other words, $\mathfrak{g}^{(i)}$ is all those elements of $\mathfrak{g}$ which can be written as linear combinations of i "full binary trees" of brackets in $\mathfrak{g}$.

Definition 4.1.2. A Lie algebra $\mathfrak{g}$ is said to be **solvable** if $\mathfrak{g}^{(n)} = 0$ for some n.

For example, abelian Lie algebras are solvable, whereas simple Lie algebras are never solvable.

In group theory, solvable groups are precisely those which can be constructed with abelian extensions— solvable Lie algebras are analogous.

Proposition 4.1.3. A Lie algebra $\mathfrak{g}$ is solvable if and only if there exists a filtration of ideals

$$\mathfrak{g} = \mathfrak{g}_0 \triangleright \mathfrak{g}_1, \triangleright \cdots \triangleright \mathfrak{g}_{k-1} \triangleright \mathfrak{g}_k = \{0\}$$

such that $\mathfrak{g}_i/\mathfrak{g}_{i+1}$ is abelian.

Proof. Since $\mathfrak{h}/[\mathfrak{h}\mathfrak{h}]$ is *always* abelian for *any* Lie algebra $\mathfrak{h}$, it's clear that if $\mathfrak{g}$ is solvable it suffices to take its derived series as the filtration, as

$$\mathfrak{g}^{(i)}/\mathfrak{g}^{(i+1)} = \mathfrak{g}^{(i)} \Big/ [\mathfrak{g}^{(i)}\mathfrak{g}^{(i)}].$$

On the flip side, if we have such a descending sequence of ideals $\mathfrak{g}_0, \ldots, \mathfrak{g}_k$, it must be that $[\mathfrak{g}_i\mathfrak{g}_i] \subseteq \mathfrak{g}_{i+1}$. Let $[xy] \in [\mathfrak{g}_i\mathfrak{g}_i]$. Then

$$[xy] + \mathfrak{g}_{i+1} = [x + \mathfrak{g}_i, y + \mathfrak{g}_i] = \mathfrak{g}_{i+1}.$$

Then, by an easy induction $\mathfrak{g}^{(i)} \subseteq \mathfrak{g}_i$, which proves that the derived series terminates, since $\mathfrak{g}_i$ does. $\square$

Proposition 4.1.4. The Lie algebra of upper triangular matrices $\mathfrak{t}_n(\mathbf{k})$ is solvable.

Proof. We use the following definition of an upper triangular matrix:

$$(a_{ij}) \text{ is upper triangular} \iff a_{ij} = 0 \text{ if } j - i < 0.$$

Let (a_{ij}) and (b_{ij}) be two upper triangular matrices, and let $j - i < 1$, then

$$\begin{aligned}
(ab - ba)_{ij} &= (ab)_{ij} - (ba)_{ij} \\
&= \sum_{k=1}^{n} a_{ik} b_{kj} - \sum_{k=1}^{n} b_{ik} a_{kj} \\
&= \left(\sum_{k=1}^{i-1} a_{ik} b_{kj} + \sum_{k=i}^{j} a_{ik} b_{kj} + \sum_{k=j+1}^{n} a_{ik} b_{kj} \right) - \sum_{k=1}^{n} b_{ik} a_{kj} \\
&= \left(\sum_{k=1}^{i-1} 0 \cdot b_{kj} + \sum_{k=i}^{j} a_{ik} b_{kj} + \sum_{k=j+1}^{n} a_{ik} \cdot 0 \right) - \sum_{k=1}^{n} b_{ik} a_{kj} \\
&= \sum_{k=i}^{j} a_{ik} b_{kj} - \sum_{k=1}^{n} b_{ik} a_{kj} \\
&= \sum_{k=i}^{j} a_{ik} b_{kj} - \sum_{k=i}^{j} b_{ik} a_{kj} \\
&= \sum_{k=i}^{j} (a_{ik} b_{kj} - b_{ik} a_{kj}) \\
&= \begin{cases} 0 & \text{if } j < i \\ a_{jj} b_{jj} - b_{jj} a_{jj} & \text{if } j = i \end{cases} \\
&= 0.
\end{aligned}$$

Hence, $(ab - ba)$ is *strictly* upper triangular, so $[ab] \in \mathfrak{n}$. Then $\mathfrak{t}^{(1)} = [\mathfrak{t}\mathfrak{t}] \subseteq \mathfrak{n}$.

Now suppose that, for some $l \geq 0$,

$$(a_{ij}) \in \mathfrak{n}^{(l)} \implies a_{ij} = 0 \text{ if } j - i < m.$$

Then, we can do a similar, in fact easier calculation to show that if $(a_{ij}), (b_{ij}) \in \mathfrak{t}^{(m)}$

and $j - i < 2m$.

$$(ab - ba)_{ij} = \sum_{k=i+m}^{j-m} (a_{ik}b_{kj} - b_{ik}a_{kj}) = 0.$$

Hence, we have shown that

$$(a_{ij}) \in \mathfrak{t}^{(l+1)} \implies a_{ij} = 0 \text{ if } j - i < 2m.$$

Combined with our initial conditions, we have shown in general that

$$(a_{ij}) \in \mathfrak{t}^{(l)} \implies a_{ij} = 0 \text{ if } j - i < 2^l.$$

Clearly, if l is large enough, (a_{ij}) is forced to be the zero matrix. Hence $\mathfrak{n}$ is solvable, as $\mathfrak{n}^{(l)} = 0$ for some positive integer l. Then $\mathfrak{t}$ is also solvable, as $\mathfrak{t}^{(l+1)} \subseteq \mathfrak{n}^{(l)} = 0$. □

Theorem 4.1.5. Let $\mathfrak{g}$ be a Lie algebra.

(a) If $\mathfrak{g}$ is solvable, then so are all subalgebras and homomorphic images of $\mathfrak{g}$.

(b) If $\mathfrak{i}$ is a solvable ideal of $\mathfrak{g}$ such that $\mathfrak{g}/\mathfrak{i}$ is also solvable, then $\mathfrak{g}$ is solvable.

(c) If $\mathfrak{i}, \mathfrak{j}$ are solvable ideals of $\mathfrak{g}$, then so is $\mathfrak{i} + \mathfrak{j}$.

Proof. The first statement of (a) follows if we show that

$$\mathfrak{h}^{(i)} \subseteq \mathfrak{g}^{(i)}$$

for any subalgebra $\mathfrak{h}$ of $\mathfrak{g}$— this is an easy induction. Similarly, the second statement of (a) follows from

$$(\phi\mathfrak{g})^{(i)} = \phi\left(\mathfrak{g}^{(i)}\right)$$

for any homomorphism ϕ. This is another easy induction.

For (b), we stack together $\mathfrak{g}/\mathfrak{i}$ and $\mathfrak{i}$'s solvability— the former being solvable means that $\mathfrak{g}^{(n)} \subseteq \mathfrak{i}$ for large enough n, but that means that $\mathfrak{g}^{(i)}$ is a subalgebra of $\mathfrak{i}$, for which $\mathfrak{i}^{(m)} = 0$ for large enough m, so we can "push in" $\mathfrak{g}$ further, namely

$$\mathfrak{g}^{(n+m)} = \left(\mathfrak{g}^{(n)}\right)^{(m)} \subseteq \mathfrak{i}^{(m)} = 0.$$

□

The solvability of a Lie algebra measures how "structured" its nonabelianness is.

Definition 4.1.6. The **radical** $\operatorname{rad}\mathfrak{g}$ of a Lie algebra $\mathfrak{g}$ is defined to be the maximal solvable ideal of $\mathfrak{g}$.

Definition 4.1.7. A Lie algebra $\mathfrak{g}$ is said to be **semisimple** if $\operatorname{rad}\mathfrak{g} = 0$.

4.2 The descending central series, nilpotency

Definition 4.2.1. The **descending central series** of a Lie algebra $\mathfrak{g}$ is a sequence of ideals $\mathfrak{g}^0, \mathfrak{g}^1, \ldots$ defined to be

$$\begin{cases} \mathfrak{g}^0 := \mathfrak{g} \\ \mathfrak{g}^i := [\mathfrak{g}\mathfrak{g}^{i-1}] \end{cases}.$$

Definition 4.2.2. A Lie algebra $\mathfrak{g}$ is said to be **nilpotent** if $\mathfrak{g}^n = 0$ for some n.

Proposition 4.2.3. All nilpotent Lie algebras are solvable.

Definition 4.2.4. Let $\mathfrak{g}$ be a Lie algebra. We say that $x \in \mathfrak{g}$ is **ad-nilpotent** if $(\operatorname{ad} x)^n = 0$ for some n.

Theorem 4.2.5. Let $\mathfrak{g}$ be a Lie algebra.

(a) If $\mathfrak{g}$ is nilpotent, then so are all subalgebras and homomorphic images of $\mathfrak{g}$.

(b) If $\mathfrak{g}/Z(\mathfrak{g})$ is nilpotent, then so is $\mathfrak{g}$.

(c) If $\mathfrak{g}$ is nilpotent and nonzero, then $Z(\mathfrak{g})$ is nonzero.

4.3 Engel's theorem

We will prove **Engel's theorem**.

Theorem 4.3.1 (Engel)**.** Let $\mathfrak{g}$ be a Lie algebra. Then the following are equivalent:

(i) $\mathfrak{g}$ is nilpotent.

(ii) All the elments of $\mathfrak{g}$ are ad-nilpotent.

We will prove the following equivalent theorem:

Theorem 4.3.2. Let $\mathfrak{g}$ be a subalgebra of $\mathfrak{gl}(V)$, where V has positive dimension. If $\mathfrak{g}$ consists only of nilpotent transformations, then there exists a nonzero vector $v \in V$ so that $\mathfrak{g}.v = 0$.

Proof. We induct on $\dim \mathfrak{g}$.

The $\dim \mathfrak{g} = 0$ case is trivial— $\mathfrak{g}$ will only contain the zero transformation.

The $\dim \mathfrak{g} = 1$ case is also easy. Let $x \in \mathfrak{g}$ be nonzero and nilpotent. Then we can find a nonzero vector $v \in V$ so that $x.v = 0$, and so $\mathfrak{g}.v = \mathbf{k}x.v = 0$.

Now supppose $\dim \mathfrak{g} > 1$. The induction step is very algebraic, so we will break it down

Step 1 LOCATE AN IDEAL $\mathfrak{h}$ OF CODIMENSION ONE.

We will do this by demonstrating that subalgebras of $\mathfrak{g}$ are *not* self-normalizing, which will allow us to produce a maximal ideal $\mathfrak{h}$ of codimension one.

Let $\mathfrak{h}$ be a proper subalgebra of $\mathfrak{g}$ of positive dimension. Then,

$$\operatorname{ad} \mathfrak{g}/\mathfrak{h} := \left\{ \operatorname{ad}_{\mathfrak{g}/\mathfrak{h}}(x + \mathfrak{h}) : x \in \mathfrak{g} \right\}$$

is a Lie algebra— it is the homomorphic image of $\mathfrak{g}$ under the composition

$$\mathfrak{g} \xrightarrow{\pi} \mathfrak{g}/\mathfrak{h} \xrightarrow{\operatorname{ad}} \operatorname{ad} \mathfrak{g}/\mathfrak{h}.$$

Moreover,

$$\dim \mathfrak{g} > \dim \mathfrak{g}/\mathfrak{h} \geq \dim \operatorname{ad} \mathfrak{g}/\mathfrak{h},$$

as $\mathfrak{h}$ has positive dimension. By the inductive hypothesis, we may find a nonzero vector $x + \mathfrak{h} \in \mathfrak{g}/\mathfrak{h}$ such that

$$\operatorname{ad} \mathfrak{g}/\mathfrak{h}.(x + \mathfrak{h}) = 0 + \mathfrak{h} = \mathfrak{h}.$$

This means that

$$\begin{aligned} [hx] + \mathfrak{h} &= [h + \mathfrak{h}, x + \mathfrak{h}] \\ &= \operatorname{ad}_{\mathfrak{g}/\mathfrak{h}}(h + \mathfrak{h}).(x + \mathfrak{h}) \\ &= \mathfrak{h} \end{aligned}$$

for all $h \in \mathfrak{h}$, so $x \in N_{\mathfrak{g}}(\mathfrak{h})$.

But $x + \mathfrak{h}$ being nonzero in $\mathfrak{g}/\mathfrak{h}$ means exactly that $x \notin \mathfrak{h}$, so $\mathfrak{h} \subsetneq N_{\mathfrak{g}}(\mathfrak{h})$. We will use this fact to produce a nontrivial maximal ideal of $\mathfrak{g}$.

We are always able to find a proper subalgebra of positive dimension— choose the span of any single element in $\mathfrak{g}$. Then, there must exist maximal proper subalgebras. Let $\mathfrak{h}$ be maximal now. Then we have that $N_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{g}$, as otherwise $N_{\mathfrak{g}}(\mathfrak{h})$ is a larger proper subalgebra of $\mathfrak{g}$.

Hence, $\mathfrak{h}$ is a proper ideal of $\mathfrak{g}$, so $\mathfrak{g}/\mathfrak{h}$ must contain a one-dimensional subalgebra. By Theorem 2.3.2, this one-dimensional subalgebra has the form $\mathfrak{a}/\mathfrak{h}$, where $\mathfrak{h} \lhd \mathfrak{a} \leq \mathfrak{g}$. Now, it must be that $\mathfrak{a} = \mathfrak{g}$, as otherwise $\mathfrak{a}$ is a proper subalgebra of $\mathfrak{g}$ containing $\mathfrak{h}$. Then $\mathfrak{a}/\mathfrak{h} = \mathfrak{g}/\mathfrak{h}$, so $\mathfrak{g}/\mathfrak{h}$ is one-dimensional. This shows that $\mathfrak{h}$ has codimension one in $\mathfrak{g}$.

Now, consider the subspace $W = \{v \in V : \mathfrak{h}.v = 0\}$ of V. Since $\mathfrak{h}$ is an ideal of $\mathfrak{g}$, $\mathfrak{g}$ stabilizes W— for all $g \in \mathfrak{g}$, $h \in \mathfrak{h}$, and $w \in W$, we have that

$$\begin{aligned} h.g.w &= hg.w \\ &= \Big(gh - [gh]\Big).w \\ &= \Big(g.\underbrace{h.w}_{=0}\Big) + \Big(\underbrace{[hg]}_{\in\mathfrak{h}}.w\Big) \\ &= \Big(g.0\Big) + 0 \\ &= 0, \end{aligned}$$

hence $\mathfrak{g}.W \subseteq W$.

Then, if we pick $g \in \mathfrak{g}$ and restrict it to W, we have a nilpotent endomorphism of W, hence g has an eigenvector v in W.

Then, $(\mathfrak{h} + \mathbf{k}g).v = 0$, completing the theorem. □

Now, we can prove Engel's theorem:

Proof of Engel's theorem. As before, the $\dim \mathfrak{g} = 0$ and $\dim \mathfrak{g} = 1$ cases are trivial. So, we induct on $\dim \mathfrak{g}$.

Let $\mathfrak{g}$ be a Lie algebra whose elements are all ad-nilpotent.

Then $\operatorname{ad} \mathfrak{g}$ is a subalgebra of $\mathfrak{gl}(\mathfrak{g})$ consisting of nilpotent transformations, hence there exists a nonzero vector $x \in \mathfrak{g}$ such that $\operatorname{ad} \mathfrak{g}.x = 0$.

But, from the definition of ad, this means that $[\mathfrak{g}x] = 0$, hence $x \in Z(\mathfrak{g})$, so $Z(\mathfrak{g})$ has positive dimension, and $\dim \mathfrak{g}/Z(\mathfrak{g}) < \dim \mathfrak{g}$.

Now, we want to show that $\mathfrak{g}/Z(\mathfrak{g})$ consists of ad-nilpotent elements. This follows from the observation that

$$\operatorname{ad}\Big(x + Z(\mathfrak{g})\Big).\Big(y + Z(\mathfrak{g})\Big) = [x + Z(\mathfrak{g}), y + Z(\mathfrak{g})]$$

$$= [xy] + Z(\mathfrak{g})$$
$$= (\operatorname{ad} x.y) + Z(\mathfrak{g}),$$

hence it easily follows that $\operatorname{ad}\left(x + Z(\mathfrak{g})\right)$ is nilpotent given that $\operatorname{ad} x$ is nilpotent.

Then, by the induction hypothesis, $\mathfrak{g}/Z(\mathfrak{g})$ is a nilpotent Lie algebra.

By Theorem, $\mathfrak{g}$ is a nilpotent Lie algebra, completing the proof. □

Corollary 4.3.3. If $\mathfrak{g}$ is a nilpotent subalgebra of $\mathfrak{gl}(V)$, then there exists a flag in V such that $\mathfrak{g}.V_i \subseteq V_{i-1}$ for all i.

Namely, there exists a basis of V for which all the matrices of $\mathfrak{g}$ are strictly upper triangular.

Proof. □

5 Semisimple Lie algebras

Convention 5.0.1. Let $\mathbf{k}$ denote an algebraically closed field of characteristic zero.

5.1 Lie's theorem

Similar to Engel's theorem, which concerned *nilpotent* Lie algebras, we have **Lie's theorem**, which concerns *solvable* Lie algebras.

Theorem 5.1.1 (Lie's theorem)**.** Let $\mathfrak{g}$ be a solvable subalgebra of $\mathfrak{gl}(V)$. Then $\mathfrak{g}$ stabilizes some flag in V.

In other words, relative to some basis of V, the matrix representation of all elements of $\mathfrak{g}$ are upper triangular.

Again, we will prove it by proving an equivalent formulation in terms of the existence of a common eigenvector.

Theorem 5.1.2. Let $\mathfrak{g}$ be a solvable subalgebra of $\mathfrak{gl}(V)$. Then there exists $v \in V$ that is an eigenvector for all $x \in \mathfrak{g}$.

In other words, there exists a linear functional $\lambda : \mathfrak{g} \to \mathbf{k}$ such that

$$x.v = \lambda(x)v$$

for all $x \in \mathfrak{g}$.

Proof. We will use a similar strategy as with the proof of Engel's theorem.

Step 1 LOCATE AN IDEAL $\mathfrak{h}$ OF CODIMENSION ONE.

Since $\mathfrak{g}$ is solvable, $[\mathfrak{g}\mathfrak{g}] < \mathfrak{g}$, and so $\mathfrak{g}/[\mathfrak{g}\mathfrak{g}]$ has positive dimension.

Combined with the fact it is abelian, it then has an *ideal* $\mathfrak{h}/[\mathfrak{g}\mathfrak{g}] \lhd \mathfrak{g}/[\mathfrak{g}\mathfrak{g}]$ of codimension one, which, by the correspondence theorem (2.3.2), gives us an ideal $\mathfrak{h} \lhd \mathfrak{g}$ of codimension one.

Step 2 USE INDUCTION TO FIND A COMMON EIGENVECTOR FOR $\mathfrak{h}$.

Suppose that the theorem were true for all $\mathfrak{h} \leq \mathfrak{gl}(V)$ such that $\dim \mathfrak{h} < \dim \mathfrak{g}$.

Then there exists a linear functional $\lambda : \mathfrak{h} \to \mathbf{k}$ such that

$$x.v = \lambda(x)v.$$

Now, define the amalgamation of eigenspaces

$$W = \Big\{ w \in V : x.w = \lambda(x)w \Big\}.$$

Step 3 PROVE THAT $\mathfrak{g}$ LEAVES W INVARIANT.

Let $x \in \mathfrak{g}$ and $w \in W$. Then if $x.w \in W$, that means that for all $y \in \mathfrak{h}$

$$y.x.w = \lambda(y)(x.w) = \lambda(y)x.w.$$

But also,

$$\begin{aligned} y.x.w &= yx.w \\ &= \Big(xy - [xy]\Big).w \\ &= \Big(xy.w\Big) - \Big([xy].w\Big) \\ &= \Big(x.\lambda(y)w\Big) - \lambda\Big([xy]\Big)w \\ &= \Big(\lambda(y)x.w\Big) - \lambda\Big([xy]\Big)w \end{aligned}$$

Hence

$$\lambda(y)x.w = \Big(\lambda(y)x.w\Big) - \lambda\Big([xy]\Big)w,$$

so it must be that $\lambda\Big([xy]\Big) = 0$ if $x.w \in W$. We will show this directly.

Let $z \in \mathfrak{h}$. Define

$$W_i := \operatorname{span}\{w, x.w, \ldots, x^{i-1}.w\},$$

and let n be the smallest integer for which $W_n = W_{n+1}$. We would like to show the following

$$zw^i.x \equiv_{W_i} \lambda(z)(w^i.x),$$

which allows us to immediately conclude that the matrix representation of z acting on W_n is upper triangular, with diagonal entries $\lambda(z)$.

Hence, $\mathrm{tr}_{W_n}(z) = n\lambda(z)$.

Now, put $z = [xy]$, we immediately see that

$$\mathrm{tr}_{W_n}\Big([xy]\Big) = n\lambda\Big([xy]\Big).$$

However,

$$[xy]\Big|_{W_n} = [x|_{W_n}, y|_{W_n}],$$

as x and y both stabilize W_n, hence

$$\mathrm{tr}_{W_n}\Big([xy]\Big) = \mathrm{tr}\Big([x|_{W_n}, y|_{W_n}]\Big) = 0,$$

being the commutator of two elements of $\mathfrak{gl}(W_n)$.

Hence

$$n\lambda\Big([xy]\Big) = 0,$$

which, because $\mathrm{char}\,\mathbf{k} = 0$, implies that $\lambda\Big([xy]\Big) = 0$.

Hence y stabilizes W.

Step 4 Find an eigenvector in W for an endomorphism in $\mathfrak{g} - \mathfrak{h}$

Now, write $\mathfrak{g} = \mathfrak{h} + \mathbf{k}z$ for some $z \in \mathfrak{g}$.

Since z stabilizes W, and since $\mathbf{k}$ is algebraically closed, it has an eigenvector v_0 in W.

But, by definition of W, v_0 is also an eigenvector for all endomorphisms in $\mathfrak{h}$, hence we conclude that v_0 is a common eigenvector for all endomorphisms in $\mathfrak{g}$.

This completes the proof of the theorem.

□

5.2 Jordan-Chevalley decomposition

We say that an endomorphism $x \in \operatorname{End} V$ where V is a vector space over an algebraically closed field is **semisimple** if all of its eigenvalues are distinct.

Theorem 5.2.1 (Jordan-Chevalley decomposition). Let $x \in \operatorname{End} V$. Then there exist unique $x_s, x_n \in \operatorname{End} V$ such that

(a) $x = x_s + x_n$,

(b) x_s is semismple and x_n is nilpotent.

(c) x_s and x_n are polynomials in x.

Proof. We will explicitly construct x_s as a certain polynomial in x.

Step 1 Break up V into x's generalized eigenspaces.

Let

$$p_x(t) = \prod_{i=1}^{k} (t - a_i)^{m_i}$$

be the characteristic polynomial of x.

Define the generalized eigenspace

$$V_i := \ker(x - a_i 1)^{m_i}$$

for each a_i. Then

$$V = V_1 \oplus \cdots \oplus V_k.$$

Since x stabilizes each V_i, we can consider the endomorphism $x_i \in \operatorname{End} V_i$ defined by restricting x to V_i. Then, clearly the characteristic polynomial of x_i is

$$p_i(t) := p_{x_i}(t) = (t - a_i)^{m_i}.$$

Step 2 Use the Chinese Remainder Theorem to create an operator that is diagonal on each eigenspace.

The ideals generated by each $p_i(t)$ is coprime— consider $i \neq j$. The maximal ideals $\langle t - a_i \rangle$ and $\langle t - a_j \rangle$ are coprime, hence $\langle (t - a_i)^{m_i} \rangle$ and $\langle (t - a_j)^{m_j} \rangle$ are coprime.

Now, use the Chinese remainder theorem to locate a polynomial $p(t) \in \mathbf{k}[t]$ that satisfies the congruences

$$p(t) \equiv a_i \mod p_i(t), \qquad p(t) \equiv 0 \mod t.$$

We will show that $p(x)$ is semisimple, so put $x_s := p(x)$.

For all i, $t - a_i$ annihilates x_s restricted to V_i:

$$(x_s - a_i)\big|_{V_i} = p(x_i) - a_i = \left[a_i + h(x_i)p_i(x_i)\right] - a_i = a_i - a_i = 0.$$

where we have used the fact that $p(t) \equiv a_i \mod p_i(t)$ to write $p(x_i) = a_i + h(x_i)p_i(x_i)$ and that $p_i(x_i) = 0$.

□

5.3 Cartan's criterion

Theorem 5.3.1 (Cartan's criterion)**.** Let $\mathfrak{g}$ be a subalgebra of $\mathfrak{gl}(V)$. The following are equivalent

(a) $\mathfrak{g}$ is solvable.

(b) $\operatorname{tr}(xy) = 0$ for all $x \in \mathfrak{g}$ and $y \in [\mathfrak{g}\mathfrak{g}]$.

Proof. Put $x = s + n$, and let s have □

5.4 Killing form

Definition 5.4.1. The **Killing form** of a Lie algebra $\mathfrak{g}$ is the bilinear form defined by

$$(x, y) \mapsto \operatorname{tr}(\operatorname{ad} x, \operatorname{ad} y).$$

Proposition 5.4.2. A finite dimensional Lie algebra is semisimple if and only if its Killing form is nondegenerate.

Proof. □

Theorem 5.4.3. The Killing form on $\mathfrak{gl}_n(\mathbb{F})$ is given by

$$(x, y) \mapsto 2n \cdot \operatorname{tr}(xy) - 2\operatorname{tr}(x)\operatorname{tr}(y).$$

Proof. Let $x, y \in \mathfrak{gl}_n(\mathbb{F})$, and put $x = (x_{ij})$, $y = (y_{ij})$.

Then, by expanding the definition of matrix multiplication, we can see that

$$[xy]_{ij} = x_{ik}y_{kj} - y_{i\ell}x_{\ell j},$$

where here we are using the Einstein summation convention.

We can manipulate the right hand side as follows

$$\begin{aligned} &x_{ik} \underbrace{y_{kj}}_{=\delta_{\ell j} y_{k\ell}} - \underbrace{y_{i\ell}}_{=\delta_{ik} y_{k\ell}} x_{\ell j} \\ &= x_{ik}(\delta_{\ell j} y_{k\ell}) - (\delta_{ik} y_{k\ell}) x_{\ell j} \\ &= y_{k\ell}\Big(x_{ik}\delta_{\ell j} - \delta_{ik}x_{\ell j}\Big). \end{aligned}$$

Now define $\hat{x}_{ij}^{k\ell} := x_{ik}\delta_{\ell j} - \delta_{ik}x_{\ell j}$. Then we have shown that $\hat{x}_{ij}^{k\ell} y_{k\ell} = [xy]_{ij}$, which is namely the fact $\hat{x}_{ij}^{k\ell}$ is the matrix representation of $\operatorname{ad} x$ relative to the standard basis e_{ij} of $\mathfrak{gl}_n(\mathbb{F})$.

We now wish to know the value of $\operatorname{tr}(\hat{x}\hat{y})$. This is given by the contraction

$$\hat{x}_{ij}^{k\ell}\hat{y}_{k\ell}^{ij},$$

which we easily compute:

$$\begin{aligned} &\hat{x}_{ij}^{k\ell}\hat{y}_{k\ell}^{ij} \\ &= \Big(x_{ik}\delta_{\ell j} - \delta_{ik}x_{\ell j}\Big)\Big(y_{ki}\delta_{j\ell} - \delta_{ki}y_{j\ell}\Big) \\ &= (x_{ik}\delta_{\ell j})(y_{ki}\delta_{j\ell}) + (\delta_{ik}x_{\ell j})(\delta_{ki}y_{j\ell}) - (x_{ik}\delta_{\ell j})(\delta_{ki}y_{j\ell}) - (\delta_{ik}x_{\ell j})(y_{ki}\delta_{j\ell}) \\ &= \underbrace{(\delta_{j\ell}\delta_{\ell j})}_{=\delta_{jj}=n} \underbrace{x_{ik}y_{ki}}_{=\operatorname{tr}(xy)} + \underbrace{(\delta_{ik}\delta_{ki})}_{=\delta_{jj}=n} \underbrace{x_{j\ell}y_{\ell j}}_{=\operatorname{tr}(xy)} - \underbrace{(x_{ik}\delta_{ki})}_{=x_{ii}=\operatorname{tr} x} \underbrace{(\delta_{\ell j}y_{j\ell})}_{=y_{jj}=\operatorname{tr} y} - \underbrace{(x_{\ell j}\delta_{j\ell})}_{=x_{ii}=\operatorname{tr} x} \underbrace{(y_{ki}\delta_{ik})}_{=y_{ii}=\operatorname{tr} y} \\ &= 2n\operatorname{tr}(xy) + 2\operatorname{tr}(x)\operatorname{tr}(y). \end{aligned}$$

□

5.5 $\mathfrak{g}$-modules

Definition 5.5.1. Let $\mathfrak{g}$ be a Lie algebra. A **$\mathfrak{g}$-module** is a vector space V equipped with a *scaling map*

$$\begin{aligned} -.- : \mathfrak{g} \times V &\to V \\ (x, v) &\mapsto x.v \end{aligned}$$

which satisfies the following axioms:

(M1) $(ax + by).v = ax.v + by.v$,

(M2) $x.(av + bw) = ax.v + bx.w$,

(M3) $[xy].v = x.y.v - y.x.v$.

Proposition 5.5.2. $\mathfrak{g}$-modules are in one-to-one correspondence with representations of $\mathfrak{g}$.

Proof. Let V be a vector space, and let $\mathfrak{g}$ be a Lie algebra. We will demonstrate a correspondence between $\mathfrak{g}$-module structures on V and representations of $\mathfrak{g}$ in $\mathfrak{gl}(V)$.

Let $\phi : \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation of $\mathfrak{g}$.

Define a $\mathfrak{g}$-module structure on V by

$$x.v := \phi(x).v.$$

Then, (M1) and (M2) follow easily from the fact that $\phi(x) \in \mathfrak{gl}(V)$.

Then, the fact that ϕ is a Lie algebra homomorphism shows (M3), as

$$\begin{aligned} [xy].v &= \phi\Big([xy]\Big).v \\ &= [\phi(x)\phi(y)].v \\ &= \Big(\phi(x)\phi(y) - \phi(y)\phi(x)\Big).v \\ &= \Big(\phi(x).\phi(y).v\Big) - \Big(\phi(y).\phi(x).v\Big) \\ &= x.y.v - y.x.v. \end{aligned}$$

Conversely, suppose that V has a $\mathfrak{g}$-module structure. Then for all $x \in \mathfrak{g}$ we can define $\phi(x) \in \operatorname{End} V$ by

$$\phi(x).v := x.v.$$

$\square$

Theorem 5.5.3 (Schur's lemma)**.** Let

Proof. □

5.6 Weyl's theorem

Theorem 5.6.1 (Weyl's theorem)**.** If $\mathfrak{g}$ is a semisimple Lie algebra, then any representation of $\mathfrak{g}$ is completely reducible.

6 The root space decomposition

6.1 Maximal toral subalgebras

Definition 6.1.1. A **maximal toral subalgebra** $\mathfrak{h}$ of $\mathfrak{g}$ is an algebra for $\operatorname{ad} x = \operatorname{ad} x_s$ for all $x \in \mathfrak{h}$.

Proposition 6.1.2. Let $\mathfrak{h}$ be a maximal toral subalgebra of $\mathfrak{g}$.

We have an isomorphism $\mathfrak{h} \simeq \mathfrak{h}^*$ induced by the Killing form of $\mathfrak{g}$

7 Root systems

Definition 7.0.1.

8 Appendix

8.1 Definitions

Definition 8.1.1. Let ψ be some statement that can be evaluated to be true or false. The **Iverson bracket** of ψ is

$$[\psi]^? := \begin{cases} 1, & \text{if } \psi \text{ is true} \\ 0, & \text{otherwise.} \end{cases}$$

a function of the free variables of ψ.

8.2 Some linear algebra

I never really got a chance to learn much foundational *abstract* linear algebra. Learning this material was a great way for me to brush up on a lot of this stuff, so here's a short dump of some important results.

8.2.1 Definitions

Definition 8.2.1. The **endomorphism ring** $\operatorname{End} V$ of the vector space V is the collection of all linear maps from V to itself.

If $T \in \operatorname{End} V$ and $v \in V$, we will write $T.v$ to denote $T(v)$.

Definition 8.2.2. Let $\mathbf{k}$ be a field. The **$\mathbf{n} \times \mathbf{n}$ matrix ring** $\operatorname{Mat}_n(\mathbf{k})$ is defined to be the ring whose underlying set is $\mathbf{k}^{n \times n}$ with pointwise scaling and addition, and with product given by matrix multiplication.

Definition 8.2.3. Let V be a vector space over the field $\mathbf{k}$. The **dual space** V^* of V is the collection of all linear maps $V \to \mathbf{k}$.

8.2.2 Rank-nullity

Theorem 8.2.4 (Rank-nullity). Let $x \in \operatorname{End} V$. xhen

$$\operatorname{rank} x + \operatorname{nullity} x = \dim V,$$

where

$$\operatorname{rank} x := \dim \operatorname{im} x, \qquad \operatorname{nullity} x := \dim \ker x.$$

Proof. Let $n = \dim V$, $r = \operatorname{rank} x$ and let $\ell = \operatorname{nullity} x$.

Let $\mathbf{p} = (p_1, p_2, \dots, p_\ell)$ be a basis for $\ker x$.

We may extend this into a basis of V by adjoining more vectors $\mathbf{q} = (q_{\ell+1}, \dots, q_n)$, so that $(\mathbf{p}, \mathbf{q}) = (p_1, \dots, p_\ell, q_{\ell+1}, \dots, q_n)$ is a basis of V.

Then, we claim that

$$x.\mathbf{q} = \left(x.q_{\ell+1}, \cdots x.q_n,\right)$$

is a basis for $\operatorname{im} x$. We first show that it spans $\operatorname{im} x$: let $v \in V$, then $v = a_1 p_1 + \cdots + a_\ell p_\ell + a_{\ell+1} q_{\ell+1} + \cdots + a_n q_n$.

So

$$\begin{aligned} x.v &= x.\left(a_1 p_1 + \cdots + a_\ell p_\ell + a_{\ell+1} q_{\ell+1} + \cdots + a_n q_n\right) \\ &= \underbrace{\left(x.a_1 p_1 + \cdots + a_\ell p_\ell\right)}_{=0} + \left(x.a_{\ell+1} q_{\ell+1} + \cdots + a_n q_n\right) \\ &= x.\left(a_{\ell+1} q_{\ell+1} + \cdots + a_n q_n\right) \end{aligned}$$

$$= a_{\ell+1}\Big(x.q_{\ell+1}\Big) + \cdots + a_n\Big(x.q_n\Big).$$

Hence $x.v$ is in the span of $x.\mathbf{q}$. Next, we show that it is linearly independent— suppose that there existed $a_{\ell+1}, \ldots, a_n$ such that

$$a_{\ell+1}\Big(x.q_{\ell+1}\Big) + \cdots + a_n\Big(x.q_n\Big) \neq 0.$$

But this means that

$$x.\Big(a_{\ell+1}q_{\ell+1} + \cdots + a_n q_n\Big) \neq 0,$$

and so the vector $a_{\ell+1}q_{\ell+1} + \cdots + a_n q_n$ is in the kernel of x, however it is not in the kernel of x because it is not in the span of $\mathbf{p}$, a contradiction.

Hence $x.\mathbf{q}$ is linearly independent, completing our assertion that it is a basis of $\operatorname{im} x$.

Then $r = \dim \operatorname{im} x = n - \ell$, and so

$$r + \ell = n,$$

which proves the theorem. □

Corollary 8.2.5. Let $x \in \operatorname{End} V$. The following are equivalent:

(a) x is injective.

(b) x is surjective.

(c) x is bijective.

Proof. We have the easily verifiable propositions:

$$\dim \ker x = 0 \iff x \text{ is injective}$$

$$\dim \operatorname{im} x = \dim V \iff x \text{ is surjective}$$

And, by rank nullity,

$$\dim \ker x = 0 \iff \dim \operatorname{im} x = \dim V,$$

hence x is injective if and only if it is surjective. □

8.2.3 The matrix representation

We recall the definition of a tensor product:

Definition 8.2.6. Let V and W be two $\mathbf{k}$-vector spaces with bases $\mathbf{v} = (v_1, \ldots, v_n)$ and $\mathbf{w} = (w_1, \ldots, w_m)$ respectively.

The **tensor product of vector spaces** $V \otimes W$, is the $\mathbf{k}$-vector space with basis

$$\Big\{v_i \otimes w_j : 1 \leq i \leq n, 1 \leq j \leq m\Big\}.$$

As a structure, there isn't really "anything happening" with this construction. The following definition makes

Definition 8.2.7. Let V, W be $\mathbf{k}$-vector spaces as before.

Let $v = a_1v_1 + \cdots + a_nv_n \in V$ and $w = b_1w_1 + \cdots + b_mw_m \in W$. The **tensor product of vectors** $v \otimes w$ is defined

$$v \otimes w = \left(\sum_{i=1}^n a_iv_i\right) \otimes \left(\sum_{j=1}^n b_jw_j\right) := \sum_{i=1}^n \sum_{j=1}^m a_ib_j(v_i \otimes w_j).$$

This defines a map $i : V \times W \to V \otimes W$ given by $(v, w) \mapsto v \otimes w$.

Together, this pair of constructions satisfies a *universal property*:

Theorem 8.2.8. If U and V are two $\mathbf{k}$-vector spaces, then any bilinear map $f : U \times V \to W$ factors through $\otimes : U \times V \to U \otimes V$ — there exists a unique linear map $\overline{f}$ that makes the following diagram commute:

$$\begin{tikzcd}
U \times V \arrow[d, "i"'] \arrow[dr, "f"] & \\
U \otimes V \arrow[r, "\overline{f}"'] & W
\end{tikzcd}$$

Proof. Fix bases $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots v_n)$ of U and V.

Let $u = a_1u_1 + \cdots + a_nu_n$ and $v = b_1v_1 + \cdots + b_mv_m$.

Then, by bilinearity,

$$\begin{aligned}
f(u, v) &= f\Big(a_1u_1 + \cdots + a_nu_n, b_1v_1 + \cdots + b_mv_m\Big) \\
&= \sum_{i=1}^n a_i \cdot f\Big(u_i, b_1v_1 + \cdots + b_mv_m\Big)
\end{aligned}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j \cdot f(u_i, v_j).$$

Hence, f is completely determined by its values $f(u_i, v_j)$ where $1 \leq i \leq n$, $1 \leq j \leq m$. Conversely, any array $w_{ij} \in V$ defines a bilinear map by putting $(u_i, v_j) \mapsto w_{ij}$.

Pick some i, j. If $f = \overline{f} \circ i$, it must be that

$$f(u_i, v_j) = \left(\overline{f} \circ i\right)(u_i, v_j) = \overline{f}(u_i \otimes v_j).$$

$\square$

Definition 8.2.9. Let U, V, X, Y be $\mathbf{k}$-vector spaces, and let $f : U \to X$ and $g : V \to Y$ be linear maps. We define the **tensor product of linear maps** $f \otimes g$ to be the map

$$\begin{aligned} f \otimes g : U \otimes V &\to X \otimes Y \\ \sum_i u_i \otimes v_i &\mapsto \sum_i f(u_i) \otimes g(v_i). \end{aligned}$$

Definition 8.2.10. Let U, V, W be vector spaces with bases $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, then any linear map $f : V \to W$ induces a linear map $U \otimes V \to U \otimes W$ given by $\mathrm{id} \otimes f$

Definition 8.2.11. Let V be a vector space over $\mathbf{k}$ and fix a basis $\mathbf{v} = (v_1, \ldots, v_n)$ of V with a dual basis $\mathbf{v}^* = (v^1, \ldots, v^n)$ of the dual space V^*.

By abuse of notation, we define the corresponding elements

$$\mathbf{v} := \sum_{i=1}^{n} e^i \otimes v_i \in (\mathbf{k}^n)^* \otimes V, \qquad \mathbf{v}^* := \sum_{i=1}^{n} v^i \otimes e_i \in V^* \otimes \mathbf{k}^n$$

for the basis $\mathbf{v}$ and dual basis $\mathbf{v}^*$.

Definition 8.2.12. Let U, V, W be vector spaces with bases $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, then we define a product

$$\begin{aligned} (U^* \otimes V) \times (V^* \otimes W) &\to U^* \otimes W \\ (u^i \otimes v_j)(v^k \otimes w_l) &\mapsto v^k v_j (u^i \otimes w_l). \end{aligned}$$

Theorem 8.2.13. Let $X : V \to W$ and $Y : U \to V$. Then

$$\mathbf{u}^*(X \circ Y)\mathbf{w} = (\mathbf{w}^* X \mathbf{v})(\mathbf{v}^* Y \mathbf{u}).$$

Theorem 8.2.14. Let V be a vector space over $\mathbf{k}$ of dimension n. Then

$$\operatorname{End} V \simeq V^* \otimes V \simeq M_n(\mathbf{k}).$$

Proof. Fix a basis $\mathbf{v}$ and dual basis $\mathbf{v}^*$ of V.

Now, if U is a vector space and $x \in \operatorname{End} V$, it has a linear action on $U^* \otimes V$ given by $u^i \otimes v_j \mapsto u^i \otimes (x.v_j)$.

The map $T \mapsto \mathbf{v}^* T \mathbf{v}$ is the desired isomorphism between $\operatorname{End} V$ and $V^* \otimes V$.

Then, the map $v^i \otimes v_j \mapsto e_{ij}$ provides the isomorphism between $V^* \otimes V$ and $M_n(\mathbf{k})$. □

8.2.4 Change of basis

Proposition 8.2.15. If $\mathbf{v}$ and $\mathbf{w}$ are two bases of V, then

$$\mathbf{v}\mathbf{w}^* \in (\mathbf{k}^n)^* \otimes \mathbf{k}^n$$

encodes the change of basis matrix expressing coordinates in $\mathbf{v}$ as coordinates in $\mathbf{w}$.

Similarly,

$$\mathbf{w}^*\mathbf{v} \in V^* \otimes V$$

encodes the linear map $v_i \mapsto w_i$.

Proof. Define the array S_{ij} to be the numbers for which

$$v_i = \sum_{j=1}^n S_{ij} w_j.$$

Then

$$\begin{aligned} \mathbf{v}\mathbf{w}^* &= \left(\sum_{i=1}^n e^i \otimes v_i\right)\left(\sum_{j=1}^n w^j \otimes e_j\right) \\ &= \sum_{i=1}^n \sum_{j=1}^n w^j(v_i)(e_i \otimes e_j) \end{aligned}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} S_{ij}(e_i \otimes e_j).$$

Now, consider the map $T \in \operatorname{End} V$ given by $w_i \mapsto v_i$. Then

$$\begin{aligned}
\mathbf{w}^* T \mathbf{w} &= \sum_{i=1}^{n} w^i \otimes (T.w_i) \\
&= \sum_{i=1}^{n} w^i \otimes v_i \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \delta_{ij} (w^i \otimes v_j) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} (e^i . e_j)(w^i \otimes v_j) \\
&= \left(\sum_{i=1}^{n} w^i \otimes e_i \right) \left(\sum_{j=1}^{n} e^j \otimes v_j \right) \\
&= \mathbf{w}^* \mathbf{v}.
\end{aligned}$$

□

8.2.5 Trace

Definition 8.2.16. Let V be a vector space with basis $\mathbf{v} = (v_1, \ldots, v_n)$ The **trace** $\operatorname{tr} x$ of an endomorphism $x \in \operatorname{End} V$ of V is defined to be the sum

$$\sum_{i=1}^{n} v^i \Big(x(v_i) \Big).$$

Theorem 8.2.17. The trace is a linear operator, i.e if $x, y \in \operatorname{End} V$ and $a, b \in \mathbf{k}$,

$$\operatorname{tr}(ax + by) = a \operatorname{tr} x + b \operatorname{tr} y.$$

Proof.

$$\operatorname{tr}(ax + by) = \sum_{i=1}^{n} v^i \Big((ax + by)(v_i) \Big)$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} v^i \Big(ax(v_i) + by(v_i) \Big) \\
&= \sum_{i=1}^{n} av^i \Big(x(v_i) \Big) + bv^i \Big(y(v_i) \Big) \\
&= a \sum_{i=1}^{n} v^i \Big(x(v_i) \Big) + b \sum_{i=1}^{n} v^i \Big(y(v_i) \Big) \\
&= a \operatorname{tr} x + b \operatorname{tr} y.
\end{aligned}
$$

□

Theorem 8.2.18. Let V be a vector space.
For all $x, y \in \operatorname{End} V$, $\operatorname{tr}(xy) = \operatorname{tr}(yx)$.

Proof. Fix a basis $\mathbf{v} = (v_1, \ldots, v_n)$ of V.

$$
\begin{aligned}
\operatorname{tr}(xy) &= \sum_{i=1}^{n} v^i.xy.v_i \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \Big(v^i.x.v_j \Big) \Big(v^j.y.v_i \Big) \\
&= \sum_{j=1}^{n} \sum_{i=1}^{n} \Big(v^j.y.v_i \Big) \Big(v^i.x.v_j \Big) \\
&= \sum_{i=1}^{n} v^i.yx.v_i \\
&= \operatorname{tr}(yx).
\end{aligned}
$$

□

Theorem 8.2.19. The trace of a linear operator $x \in \operatorname{End} V$ is basis invariant— its value is independent of the basis used to compute it.

8.3 Some commutative algebra

8.3.1 The Chinese Remainder Theorem

> **Theorem 8.3.1** (Chinese Remainder Theorem). Let R be a principal ideal domain, and let $I_1, \ldots, I_n$ be coprime ideals of R.
>
> Put $I = I_1 \cap \cdots \cap I_n$.
>
> The map $R/I \to R/I_1 \times \cdots \times R/I_n$ given by
>
> $$x + I \mapsto \left(x + I_1, \cdots x + I_n\right)$$
>
> is an isomorphism.

Proof. Let $I_{1,\ldots,n}$ □