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

1 Basic definitions and examples **3**
1.1 Lie algebras . . . 3
1.2 Examples . . . 5
1.2.1 Type A: the special linear algebra . . . 5
1.2.2 Type B: the odd-dimensional orthogonal algebra . . . 7
1.2.3 Type C: the symplectic algebra . . . 7
1.2.4 Type D: the even-dimensional orthogonal algebra . . . 7
1.3 Derivations, the adjoint representation . . . 8
1.4 Abstract Lie algebras . . . 9

2 Ideals and homomorphisms **10**
2.1 Ideals . . . 10
2.2 Homomorphisms . . . 11
2.3 Isomorphism theorems . . . 12

3 Automorphisms **14**

4 Solvable and nilpotent Lie algebras **15**
4.1 The derived series, solvability . . . 15
4.2 The descending central series, nilpotency . . . 16
4.3 Engel's theorem . . . 17

5 Solutions to exercises **18**

6 Appendix **21**
6.1 Definitions . . . 21
6.2 Some linear algebra . . . 22
6.2.1 Definitions . . . 22
6.2.2 Rank-nullity . . . 22
6.2.3 The matrix representation . . . 24
6.2.4 Trace . . . 28

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 $\mathbb{F}$ to denote generic vector spaces and fields respectively.

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— that 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 general linear algebra.

Definition 1.1.4 ($\mathfrak{gl}$, concretely)**.** Let $\mathbb{F}$ be some field and let n be a positive integer. The **general linear algebra** $\mathfrak{gl}_n(\mathbb{F})$ is defined to be the Lie algebra with underlying vector space the set of all $n \times n$ matrices over $\mathbb{F}$, with bracket given by

$$[xy] = xy - yx.$$

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

Proposition 1.1.5. Let $\{e_{pq}\}_{p,q=0}^{n}$ be the standard basis of $\mathfrak{gl}_n(\mathbb{F})$. Then

$$[e_{pq}e_{rs}] = \delta_{qr}e_{ps} - \delta_{sp}e_{rq},$$

where δ is the Kronecker delta.

Proof. Using the Iverson bracket,

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

and so

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

$$= \delta_{qr}(e_{ps})_{ij}.$$

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 $\mathfrak{gl}_n(\mathbb{F})$ 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 $\mathbb{F}$-vector space, and fix a basis $\{v_1, \ldots, v_n\}$ of V with a dual basis $\{v^1, \ldots, v^n\}$ of the dual space $V^\vee$. The **trace** $\operatorname{tr} x$ of an endomorphism $x \in \operatorname{End} V$ of V is defined to be the sum

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

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 6.2.13 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

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

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

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

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, $\mathfrak{sl}(V)$ is also a substructure of $\mathfrak{gl}(V)$.

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

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

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

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

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

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

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,

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

And, putting it together, we have proven:

Proposition 1.2.4.

$$\dim A_\ell = \dim \mathfrak{sl}(V) = \dim \mathfrak{sl}_n(\mathbb{F}) = \ell(\ell+2).$$

1.2.2 Type B: the odd-dimensional orthogonal algebra

Definition 1.2.5. The **orthogonal algebra** $\mathfrak{o}_{2\ell+1}(\mathbb{F})$ is defined to be

1.2.3 Type C: the symplectic algebra

Definition 1.2.6. A **symplectic form** on a vector space V is a bilinear form ω 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 1.2.7 (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** $\mathfrak{sp}(V)$, the set of all $x \in \operatorname{End} V$ such that

$$\mathsf{C}_\ell := \mathfrak{sp}(V) := \left\{x \in \mathfrak{gl}(V) : \omega\Big(x(\_), \_\Big) + \omega\Big(\_, x(\_)\Big) = 0\right\}$$

In matrix form, we define

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

where

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

is the standard symplectic form on $\mathbb{F}^{2\ell}$.

1.2.4 Type D: the even-dimensional orthogonal algebra

Definition 1.2.8 (Type D Lie algebra)**.** Let $\dim V = 2\ell$. We define D to be the **orthogonal algebra** $\mathfrak{o}(V)$, the set of all compatible bilinear transformations.

$$\mathsf{D}_\ell := \mathfrak{o}(V) := \left\{x \in \mathfrak{gl}(V) : x + \right\}$$

1.3 Derivations, the adjoint representation

Definition 1.3.1. Let $\mathscr{A}$ be a $\mathbb{F}$-algebra. A **derivation** of $\mathscr{A}$ is a linear map $\mathrm{d} : \mathscr{A} \to \mathscr{A}$ 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 1.3.2. Let $\mathscr{A}$ be a $\mathbb{F}$-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,

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

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

Definition 1.3.3. 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

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

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

Proposition 1.3.4. $\operatorname{ad} x$ 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

$$(\operatorname{ad} x)\Big([yz]\Big) = [y, (\operatorname{ad} x)(z)] + [(\operatorname{ad} x)(y), z]$$

which is exactly the defining identity for derivations. □

1.4 Abstract Lie algebras

Definition 1.4.1. 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 1.4.2. An **abelian** Lie algebra $\mathfrak{g}$ is a Lie algebra with trivial bracket—$\left[x\,y\right] = 0$ for all $x, y \in \mathfrak{g}$.

Proposition 1.4.3. 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 1.4.4. 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 $\left[x\,y\right] = ax + by$, where at least one of a, b is nonzero. Without loss of generality, let a be nonzero. Then

$$\left[\left[x\,y\right]\,y\right] = \left[ax + by, y\right] = a\left[x\,y\right].$$

Now put $u = \left[x\,y\right]$ and $v = a^{-1}y$. Then

$$[uv] = \left[\left[x\,y\right], (a^{-1}y)\right] = \left[x\,y\right] = u.$$

□

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 $y \in \mathfrak{i}$.

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

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

Definition 2.1.2. 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

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

Proposition 2.1.3. $\mathfrak{g}/\mathfrak{i}$ 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:

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

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

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

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

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

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

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

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,

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

□

Theorem 2.3.2. 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:

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

□

Corollary 2.3.3. Any simple Lie algebra is isomorphic to a linear Lie algebra.

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

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

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

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

□

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

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

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.

Proposition 4.1.3. The Lie algebra of upper triangular matrices $\mathfrak{t}_n(\mathbb{F})$ is solvable.

Proof. TODO: boring proof— the diagonal keeps receding every time you do a commutator □

Theorem 4.1.4. 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 L/I and I's solvability— the former being solvable means that $L^{(i)} \subseteq I$ eventually, but that means that $L^{(i)}$ is a subalgebra of a solvable Lie algebra I, so it is itself solvable and we can go further.

Specifically, if $L^{(n)} \subseteq I$, and $I^{(m)} = 0$, then

$$L^{(n+m)} = 0.$$

□

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.

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 equivalent theorem:

Theorem 4.3.2. Let $\mathfrak{g}$ be a subalgebra of $\mathfrak{gl}(V)$, where V has positive dimension. If x is nilpotent for all $x \in \mathfrak{g}$, 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 so that $xv = 0$, and so $\mathfrak{g}v = (\mathbb{F}x)v = 0$.

Now supppose $\dim \mathfrak{g} > 1$. Let $\mathfrak{h}$ be a proper subalgebra of $\mathfrak{g}$ of positive dimension. Then the set

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

is a Lie algebra— a subalgebra of $\mathfrak{gl}(\mathfrak{g})$. Then, $\operatorname{ad}_{\mathfrak{g}/\mathfrak{h}}(\mathfrak{h})$ is also a Lie algebra. By the inductive hypothesis, we may find a nonzero vector $x+\mathfrak{h} \in \mathfrak{g}/\mathfrak{h}$ such that $\operatorname{ad}_{\mathfrak{g}/\mathfrak{h}}(\mathfrak{h})(x+$

$\mathfrak{h}) = 0$. This means that $(\mathrm{ad}_{\mathfrak{g}/\mathfrak{h}} h)(x + \mathfrak{h}) = \mathfrak{h}$ for all $h \in \mathfrak{h}$, so $\mathfrak{h} \subsetneq N_{\mathfrak{g}}(\mathfrak{h})$. Hence $[hx] \in \mathfrak{h}$ for all $h \in \mathfrak{h}$, but $x \notin \mathfrak{h}$.

Now if $\mathfrak{h}$ is maximal, then this means 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 an ideal of $\mathfrak{g}$. We will show that it has codimension one. Suppose it has codimension at least two. Then, we can pull back a one-dimensional subalgebra of the quotient $\mathfrak{g}/\mathfrak{h}$ along the projection map and obtain a proper subalgebra of $\mathfrak{g}$ that properly contains $\mathfrak{h}$, which is impossible.

Now, consider the subspace $W = \{v \in V : \mathfrak{h}v = 0\}$. 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

$$hgw = \Big(gh - \big[gh\big]\Big)w = g(hw) + \big[hg\big]\, w = 0 + 0 = 0.$$

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, $(\mathfrak{h} + \mathbb{F}g)v = 0$, completing the theorem. □

Proof of Engel's theorem. □

Corollary 4.3.3.

5 Solutions to exercises

> **Exercise 5.1** (Humphreys 1.1). Verify that $\mathbb{R}^3$ with the bracket given by the *cross product*
> $$[xy] \coloneqq x \times y$$
> is a Lie algebra, and write down its structure constants relative to the usual basis of $\mathbb{R}^3$.

Let

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \qquad y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, \qquad z = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}, \qquad w = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}.$$

The cross product is defined

$$x \times y = \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}$$

Then we directly verify the Lie algebra axioms.

For (L1),

$$\begin{aligned}
\left(ax+by\right)\times z &= \begin{pmatrix} (ax_2+by_2)z_3-(ax_3+by_3)z_2 \\ (ax_3+by_3)z_1-(ax_1+by_1)z_3 \\ (ax_1+by_1)z_2-(ax_2+by_2)z_1 \end{pmatrix} \\
&= \begin{pmatrix} \left(ax_2z_3+by_2z_3\right)-\left(ax_3z_2+by_3z_2\right) \\ \left(ax_3z_1+by_3z_1\right)-\left(ax_1z_3+by_1z_3\right) \\ \left(ax_1z_2+by_1z_2\right)-\left(ax_2z_1+by_2z_1\right) \end{pmatrix} \\
&= \begin{pmatrix} a\left(x_2z_3-x_3z_2\right)+b\left(y_2z_3+y_3z_2\right) \\ a\left(x_3z_1-x_1z_3\right)+b\left(y_3z_1+y_1z_3\right) \\ a\left(x_1z_2-x_2z_1\right)+b\left(y_1z_2+y_2z_1\right) \end{pmatrix} \\
&= a(x\times z)+b(y\times z).
\end{aligned}$$

And, via an almost identical calculation,

$$x\times\left(ay\times bz\right)=a(x\times y)+b(x\times z).$$

Next, we verify (L2)

$$x\times x=\begin{pmatrix} x_2x_3-x_3x_2 \\ x_3x_1-x_1x_3 \\ x_1x_2-x_2x_1 \end{pmatrix}=0.$$

And finally, we verify the Jacobi identity (L3).

$$\begin{aligned}
&x\times(y\times z)+y\times(z\times x)+z\times(x\times y) \\
&=\varepsilon_{ijk}x_j(y\times z)_k+\varepsilon_{ijk}y_j(z\times x)_k+\varepsilon_{ijk}z_j(x\times y)_k \\
&=\varepsilon_{ijk}\left(x_j(y\times z)_k+y_j(z\times x)_k+z_j(x\times y)_k\right) \\
&=\varepsilon_{ijk}\left(x_j(\varepsilon_{klm}y_lz_m)+y_j(\varepsilon_{klm}z_lx_m)+z_j(\varepsilon_{klm}x_ly_m)\right) \\
&=\varepsilon_{ijk}\varepsilon_{klm}\left(x_jy_lz_m+y_jz_lx_m+z_jx_ly_m\right) \\
&=(\delta_{im}\delta_{lj}-\delta_{il}\delta_{jm})\left(x_jy_lz_m+y_jz_lx_m+z_jx_ly_m\right) \\
&=\delta_{im}\delta_{lj}\left(x_jy_lz_m+y_jz_lx_m+z_jx_ly_m\right)-\delta_{il}\delta_{jm}\left(x_jy_lz_m+y_jz_lx_m+z_jx_ly_m\right)
\end{aligned}$$

$$
\begin{aligned}
&= \left(x_l y_l z_i + y_l z_l x_i + z_l x_l y_i\right) - \left(x_m y_i z_m + y_m z_i x_m + z_m x_i y_m\right) \\
&= (y_l z_l - z_m y_m) x_i + (x_l y_l - y_m x_m) z_i + (z_l x_l - x_m z_m) y_i \\
&= 0.
\end{aligned}
$$

Exercise 5.2 (Humphreys 1.2)**.** Verify that the following equations and those implied by (L2) (L1) define a Lie algebra structure on a three dimensional vector space with basis (x, y, z): $[xy] = z$, $[xz] = y$, $[yz] = 0$.

The only $[\_\ [\_, \_]]$ terms consisting of basis elements that are nonzero are $[x\ [xy]]$ and $[x\ [xz]]$, so if

$$a = a_x x + a_y y + a_z z,$$

$$
\begin{aligned}
&\left[a_x x + a_y y + a_z z\ \left[b_x x + b_y y + b_z z, c_x x + c_y y + c_z z\right]\right] \\
&= (a_x b_x c_z - a_x b_z c_x) z + (a_x b_x c_y - a_x b_x c_x) y
\end{aligned}
$$

so, permuting indices,

$$(a_x b_x c_z - a_x b_z c_x) z + (a_x b_x c_y - a_x b_x c_x) y + (b_x c_x a_z - b_x c_z a_x) z + (b_x c_x a_y - b_x c_x a_x) y + (c_x a_x b_z - c$$

So the Jacobi identity is satisfied.

Exercise 5.3 (Humphreys 1.3)**.** Let

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

be an ordered basis for $\mathfrak{sl}(2, \mathbb{F})$. Compute the matrices of $\operatorname{ad} e$, $\operatorname{ad} f$, $\operatorname{ad} h$ relative to this basis.

We will compute the structure constants relative to e, f, h— there are $3(3-1)/2 = 3$ brackets to check:

$$[ef] = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = h.$$

$$[he] = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} = 2e.$$

$$[hf] = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = -2f.$$

Now, if we order this basis as, (e, f, h), the matrix representing $\operatorname{ad} e$ is

$$\begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

And similarly,

$$\operatorname{ad} f = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ -1 & 0 & 0 \end{pmatrix}, \qquad \operatorname{ad} h = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Exercise 5.4. Find a linear Lie algebra isomorphic to the nonabelian two dimensional algebra constructed in (1.4). [Hint: Look at the adjoint representation.]

We look at the adjoint representation at the Lie algebra given by x, y

$$x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$y = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$$

and we verify that $[x\,y] = x$:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Exercise 5.5. Verify the assertions made in (1.2) about $\mathfrak{t}(n, \mathbb{F})$, $\mathfrak{d}(n, \mathbb{F})$, $\mathfrak{n}(n, \mathbb{F})$ and compute the dimension of each algebra, by exhibiting bases.

A basis for $\mathfrak{t}(n, \mathbb{F})$ is all e_{ij} where $1 \le i \le j \le n$.
A basis for $\mathfrak{d}(n, \mathbb{F})$ is all e_{ii} where $1 \le i \le n$.
A basis for $\mathfrak{n}(n, \mathbb{F})$ is all e_{ij} where $1 \le i < j \le n$.

6 Appendix

6.1 Definitions

Definition 6.1.1. 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.

6.2 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.

6.2.1 Definitions

Definition 6.2.1. The **endomorphism ring** $\operatorname{End} V$ of the vector space V is the collection of all linear maps from V to itself.

Definition 6.2.2. Let $\mathbb{K}$ be a field. The **$n \times n$ matrix ring** $M_n(\mathbb{K})$ is defined to be the ring whose underlying set is $\mathbb{K}^{n\times n}$ with pointwise scaling and addition, and with product given by matrix multiplication.

Definition 6.2.3. Let V be a vector space over the field $\mathbb{K}$. The **dual space** $V^\vee$ of V is the collection of all linear maps $V \to \mathbb{K}$.

6.2.2 Rank-nullity

Theorem 6.2.4 (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 $\{p_1, p_2, \dots, p_\ell\}$ be a basis for $\ker x$.

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

Then, we claim that the set $\{x(q_{\ell+1}), \dots, x(q_n)\}$ is a basis of $\operatorname{im} x$.

Evidently, it spans im x, as

$$\begin{aligned}
&\operatorname{im} x \\
&= \left\{x(v) : v \in V\right\} \\
&= \left\{x(a_1 p_1 + \cdots + a_\ell p_\ell + a_{\ell+1} q_{\ell+1} + \cdots + a_n q_n) : a_1, \ldots, a_n \in \mathbb{F}\right\} \\
&= \left\{\underbrace{x(a_1 p_1 + \cdots + a_\ell p_\ell)}_{=0} + x(a_{\ell+1} q_{\ell+1} + \cdots + a_n q_n) : a_1, \ldots, a_n \in \mathbb{F}\right\} \\
&= \left\{x(a_{\ell+1} q_{\ell+1} + \cdots + a_n q_n) : a_{\ell+1}, \ldots, a_n \in \mathbb{F}\right\} \\
&= \left\{a_{\ell+1} x(q_{\ell+1}) + \cdots + a_n x(q_n) : a_{\ell+1}, \ldots, a_n \in \mathbb{F}\right\} \\
&= \operatorname{span}\left\{x(q_{\ell+1}), \ldots, x(q_n)\right\}.
\end{aligned}$$

Moreover, it is linearly independent— suppose that there existed $a_{\ell+1}, \ldots, a_n$ such that

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

But this means that

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

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 span of $\{p_1, \ldots, p_\ell\}$, which contradicts the fact that $p_1, \ldots, p_\ell$ is a basis for ker x.

Hence $\{x(q_{\ell+1}), \ldots, x(q_n)\}$ is linearly independent, and thus we have proved that it is a basis of im x.

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

$$r + \ell = n,$$

which proves the theorem. □

Corollary 6.2.5. 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. □

6.2.3 The matrix representation

Definition 6.2.6. Let V be a vector space 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 function

$$\begin{aligned} \mathbf{v} : \mathbb{K}^n &\to V \\ (a_1, \ldots, a_n) &\mapsto a_1 v_1 + \cdots + a_n v_n, \end{aligned}$$

and the function

$$\begin{aligned} \mathbf{v}^* : V &\to \mathbb{K}^n \\ u &\mapsto \Big(v^1(u), \cdots, v^n(u)\Big). \end{aligned}$$

$$\begin{aligned} \mathbf{v}^* : \mathbb{K}^n &\to V^* \\ (a_1, \ldots, a_n) &\mapsto a_1 v^1 + \cdots + a_n v^n, \end{aligned}$$

Evidently $\mathbf{v}^*\mathbf{v} = \mathrm{id}_V$ and $\mathbf{v}\mathbf{v}^* = \mathrm{id}_{\mathbb{K}^n}$.

It's also clear that both maps have trivial kernel, so by rank-nullity they are both vector isomorphisms.

Let e_{ij} be the standard basis for $M_n(\mathbb{K})$. Let $v_{i\to j}$ denote the map $x \mapsto v^i(x)v_j$. Then

$$\mathbf{v} e_{ij} \mathbf{v}^* = v_{i\to j}$$

and

$$\mathbf{v}^* v_{i\to j} \mathbf{v} = e_{ij}$$

Moreover

Proposition 6.2.7.
$$(\mathbf{v}^* T \mathbf{v})(\mathbf{v}^* x) = \mathbf{v}^* T x$$

We also have the map
$$\mathrm{id} \otimes \mathbf{v}$$
which embeds $M_n(\mathbb{K})$ in $V^* \otimes V$.

$$\begin{aligned}
&\Big(v^1 \otimes (c_{11}, \cdots, c_{1n}) + \cdots + v^n \otimes (c_{n1}, \ldots, c_{nn})\Big) \\
\mapsto &\Big(v^1 \otimes (c_{11} v_1 + \cdots + c_{1n} v_n) + \cdots + v^n \otimes (c_{n1} v_1 + \ldots + c_{nn} v_n)\Big) \\
= &\sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} (v^i \otimes v_j).
\end{aligned}$$

And the map
$$\mathbf{v}^* \otimes \mathrm{id}$$
goes backwards.

Then if $T \in \operatorname{End} V$,
$$\mathrm{id} \otimes T$$

$$\begin{aligned}
&\Big(v^1 \otimes (c_{11}, \cdots, c_{1n}) + \cdots + v^n \otimes (c_{n1}, \ldots, c_{nn})\Big) \\
\mapsto &\Big(v^1 \otimes (c_{11} v_1 + \cdots + c_{1n} v_n) + \cdots + v^n \otimes (c_{n1} v_1 + \ldots + c_{nn} v_n)\Big) \\
= &\sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} (v^i \otimes v_j).
\end{aligned}$$

Definition 6.2.8. Let V be a vector space 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^*.

The **matrix representation** of a linear map $T \in \operatorname{End} V$ is the matrix a_{ij} defined by
$$= v^i\Big(x(v_j)\Big)(e^i \otimes e_j)$$

$$\begin{aligned}
\mathbf{V} : V &\to V^* \otimes V \\
u &\mapsto \mathbf{v}^* \otimes u
\end{aligned}$$

$$\begin{aligned}\mathbf{V}^* : V^* \otimes V &\to V^* \otimes V \\ u &\mapsto \sum_i v^i \otimes u\end{aligned}$$

Theorem 6.2.9. Let V be a vector space over $\mathbb{K}$ of dimension n. Then

$$\operatorname{End} V \simeq V^* \otimes V \simeq M_n(\mathbb{K}).$$

Proof. Fix a basis $\mathbf{v}$ and dual basis $\mathbf{v}^*$ of V. The tensor product $V^* \otimes V$ has the basis

$$\left\{v^i \otimes v_j : v^i \in \mathbf{v}^*, v_j \in \mathbf{v}\right\}.$$

Now consider the spaces $V^* \otimes \mathbb{K}$ and $\mathbb{K} \otimes V$. By abuse of language, let $\mathbf{v}$ and $\mathbf{v}^*$ denote

$$\mathbf{v} := 1 \otimes v_1 + \cdots + 1 \otimes v_n, \qquad \mathbf{v}^* := v^1 \otimes 1 + \cdots + v^n \otimes 1$$

We can endow an action of V on $V^* \otimes V$:

$$x(v^i \otimes v_j) = v^i x \otimes v_j.$$

And similarly for V^*:

$$(v^i \otimes v_j)x = v^i \otimes x(v_j).$$

and we can also endow an action of $V^* \otimes V$ on V:

$$(v^i \otimes v_j)x = v^i(x)v_j.$$

Now, by abuse of language, let $\mathbf{v}$ and $\mathbf{v}^*$ denote

$$\mathbf{v} := v_1 + \cdots + v_n, \qquad \mathbf{v}^* := v^1 + \cdots + v^n$$

Then

$$\mathbf{v}^*\mathbf{v} = \sum_{i=1}^n v^i \otimes v_i.$$

We can turn $V^* \otimes V$ into an algebra by defining

$$(v^i \otimes v_j)(v^k \otimes v_l) = v^k v_j (v^i \otimes v_l).$$

Note that this means $\mathbf{v}^*\mathbf{v}$ is the identity element, making $V^* \otimes V$ unital. Now define more generally

$$\mathbf{v}^* T\mathbf{v} := \sum_{i=1}^{n} v^i \otimes Tv_i.$$

So

$$\mathbf{v}^* AB\mathbf{v} := \mathbf{v}^* \otimes AB\mathbf{v}.$$

Which we expand: Also, define

$$AB = \sum_{i,j} \sum_{k=1}^{n} a_{ik} b_{kj} v^i \otimes v_j$$

□

Proposition 6.2.10. Let $T, S \in \operatorname{End} V$. Then

$$(\mathbf{v}^* T\mathbf{v})(\mathbf{v}^* S\mathbf{v}) = \mathbf{v}^* TS\mathbf{v}.$$

Proof. Let $T = (t_{ij})$ and let $S = (s_{ij})$. Then

$$\mathbf{v}^* T\mathbf{v} = \sum_{i,j} t_{ij} v^i \otimes v_j, \qquad \mathbf{v}^* S\mathbf{v} = \sum_{i,j} s_{ij} v^i \otimes v_j.$$

so

$$\begin{aligned}(\mathbf{v}^* T\mathbf{v})(\mathbf{v}^* S\mathbf{v}) &= \left(\sum_{i,j} t_{ij} v^i \otimes v_j\right)\left(\sum_{i,j} s_{ij} v^i \otimes v_j\right) \\ &= \sum_{i,j,k,l} t_{ij} s_{kl} (v^k v_j)(v^i \otimes v_l) \\ &= \sum_{i,l} \left(\sum_k t_{ik} s_{kl}\right) (v^i \otimes v_l) \\ &= \sum_{i,l} v^i \otimes \left(\sum_k t_{ik} s_{kl}\right) v_l \\ &= \sum_{i,l} v^i \otimes TSv_l.\end{aligned}$$

$$\begin{aligned}
\sum_{i,j} v^i \otimes TSv_j &= \sum_{i,j} v^i \otimes T\left(\sum_k s_{jk} v_k\right) \\
&= \sum_{i,j,k} s_{jk} v^i \otimes T v_k \\
&= \sum_{i,j,k} s_{jk} v^i \otimes \left(\sum_l t_{kl} v_l\right) \\
&= \sum_{i,j,k,l} s_{jk} t_{kl} (v^i \otimes v_l)
\end{aligned}$$

□

6.2.4 Trace

Definition 6.2.11. Let V be a vector space and fix a basis $\{v_1, \ldots, v_n\}$ of V with a dual basis $\{v^1, \ldots, v^n\}$ of the dual space $V^\vee$. 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 6.2.12. The trace is a linear operator, i.e if $x, y \in \operatorname{End} V$ and $a, b \in \mathbb{F}$,

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

Proof.

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

$$= a \operatorname{tr} x + b \operatorname{tr} y.$$

□

Theorem 6.2.13. 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.

Proof. Let $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_n\}$ be two bases of V, and let $\{v^1, \ldots, v^n\}$ and $\{w^1, \ldots, w^n\}$ be the corresponding dual bases of $V^\vee$.

We write the transition coefficients S_{ij} and S^{ij}, which record the expansions of w_i and w^i in terms of v_j and v^j respectively.

$$w_i = \sum_{k=1}^n S_{ki} v_k, \qquad w^i = \sum_{k=1}^n S^{ik} v^k.$$

Importantly,

$$\begin{aligned}
\delta_{ij} &= w^i w_j \\
&= \left(\sum_{k=1}^n S^{ik} v^k\right)\left(\sum_{l=1}^n S_{jl} v_l\right) \\
&= \sum_{k=1}^n \sum_{l=1}^n S^{ik} S_{jl} v^k v_l \\
&= \sum_{k=1}^n \sum_{l=1}^n S^{ik} S_{jl} \delta_{kl} \\
&= \sum_{k=1}^n S_{jk} S^{ik},
\end{aligned}$$

Hence,

$$\begin{aligned}
\sum_{i=1}^n w^i\Big(x(w_i)\Big) &= \sum_{i=1}^n \left(\sum_{k=1}^n S^{ik} v^k\right)\left(x\left(\sum_{j=1}^n S_{ji} v_j\right)\right) \\
&= \sum_{i=1}^n \left(\sum_{k=1}^n S^{ik} v^k\right)\left(\sum_{j=1}^n S_{ji} x(v_j)\right) \\
&= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n S_{ji} S^{ik} v^i \Big(x(v_j)\Big)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=1}^{n} \sum_{k=1}^{n} \left(\sum_{i=1}^{n} S_{ji} S^{ik} \right) v^k \Big(x(v_j) \Big) \\
&= \sum_{j=1}^{n} \sum_{k=1}^{n} \delta_{jk} v^k \Big(x(v_j) \Big) \\
&= \sum_{j=1}^{n} v^j \Big(x(v_j) \Big).
\end{aligned}
$$

Hence the trace gives the same value regardless of basis. □