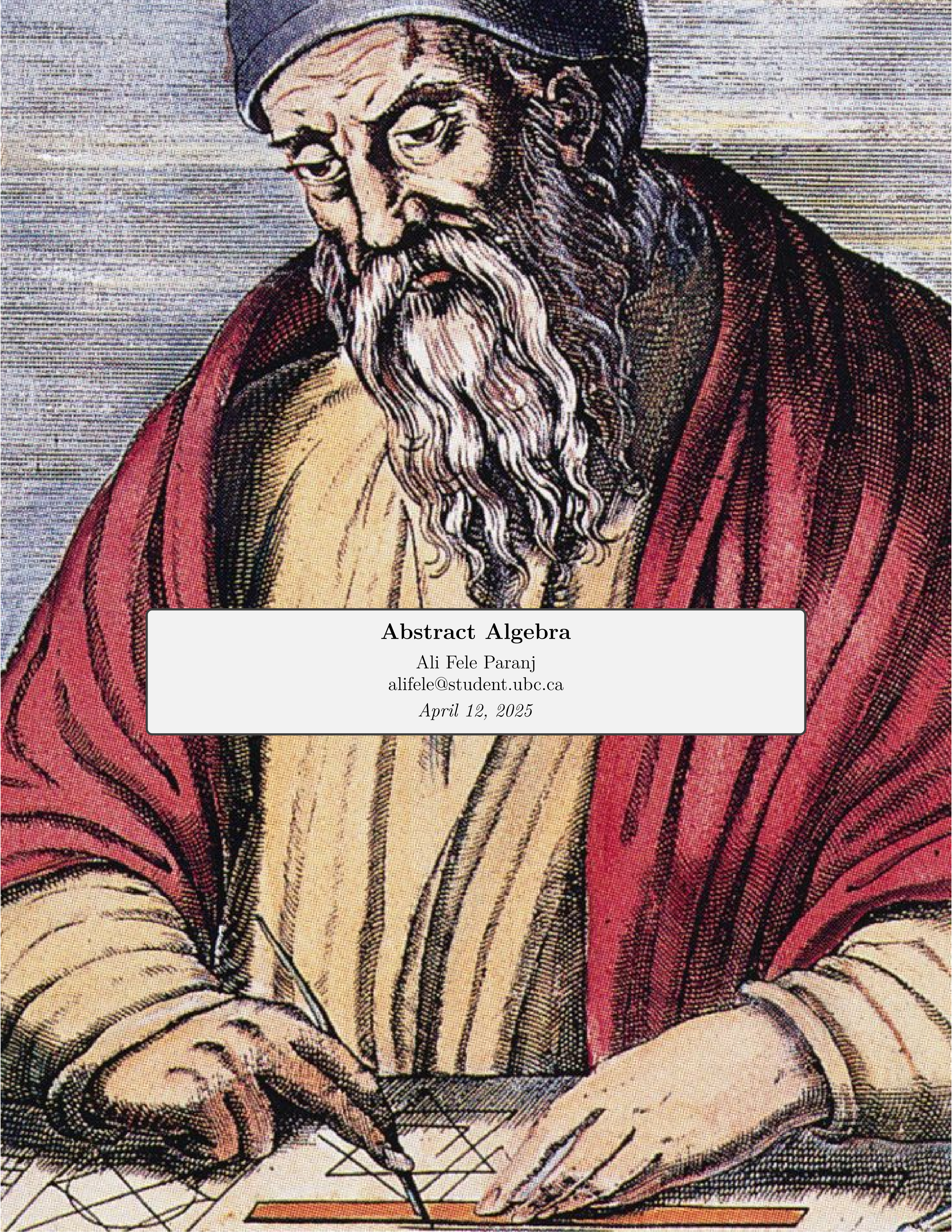
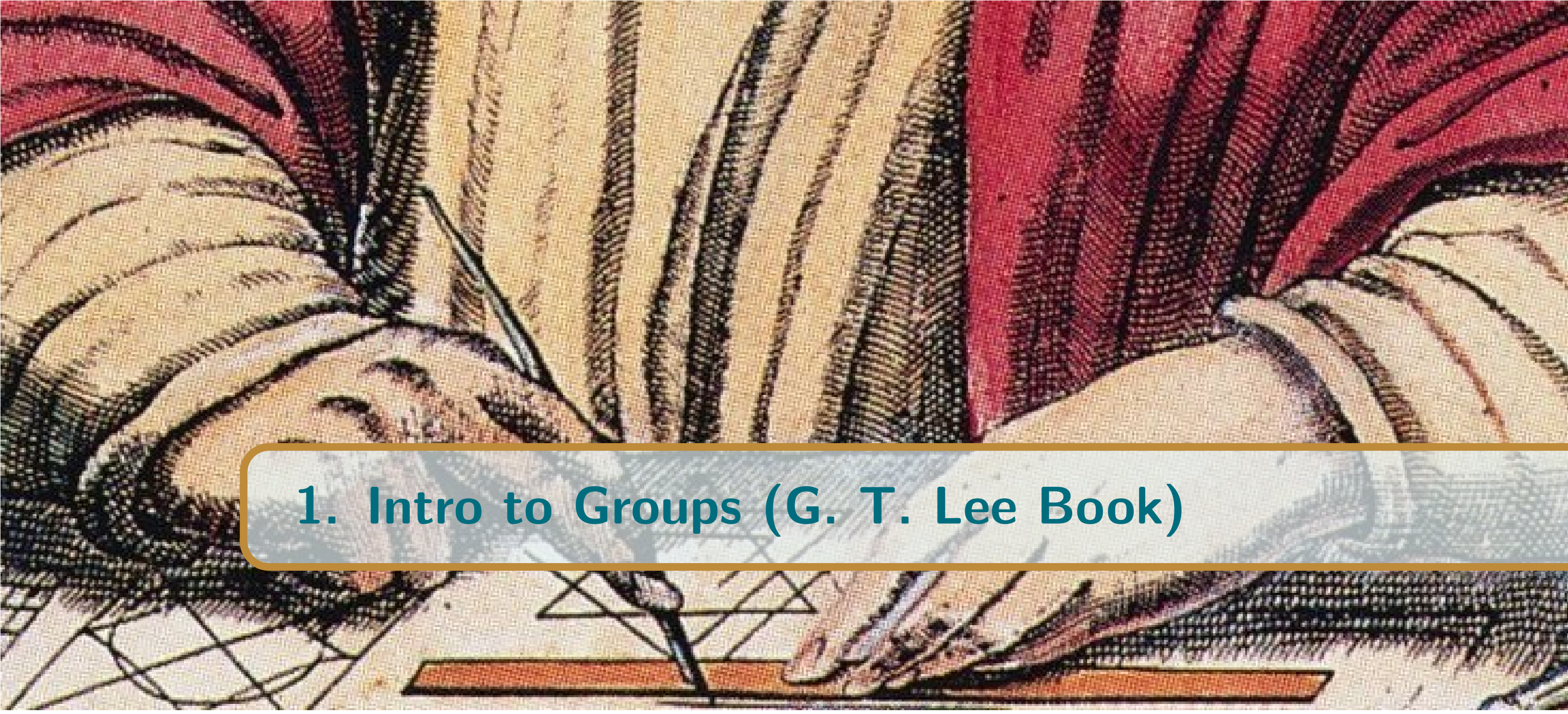


Abstract Algebra

Ali Fele Paranj
alifele@student.ubc.ca

April 12, 2025

0. Contents



1. Intro to Groups (G. T. Lee Book)

Theorem 1.1 Let G be a group and let $a \in G$. Suppose $i, j \in \mathbb{Z}$. Then

(i) If a has infinite order, then $a^i = a^j$ if and only if $i = j$.

(ii) If $|a| = n < \infty$, then $a^i = a^j$ if and only if $i \equiv j \pmod{n}$.

Proof. Proof for (i) and (ii) is as follows.

(i) $\boxed{\Rightarrow}$: Assume $a^i = a^j$ for some $i, j \in \mathbb{Z}$. Thus $a^{i-j} = e$. However, since a has infinite order, it implies that $i - j = 0$, hence $i = j$.

$\boxed{\Leftarrow}$: The converse direction follows immediately from the definition of group.

(ii) $\boxed{\Rightarrow}$: Assume $a^i = a^j$ for some $i, j \in \mathbb{Z}$. We can write $a^{i-j} = e$. Using division algorithm we can write $i - j = nq + r$ where $q, r \in \mathbb{Z}$ and $0 \leq r < n$. So

$$a^{i-j} = (a^n)^q a^r = e.$$

Since n is the order of a, it implies that $a^n = e$. Thus the equality above implies that $a^r = e$. By definition n was the smallest number with this property, and by division algorithm we have $0 \leq r < n$. This implies that $r = 0$. So $i - j = nq$ or equivalently $i \equiv j \pmod{n}$.

$\boxed{\Leftarrow}$: Assume $i \equiv j \pmod{n}$. This implies $i - j = qn$ for some $q \in \mathbb{Z}$. Thus $a^{i-j} = (a^n)^q = e$. This implies $a^i = a^j$.

$\square$

1.1 Solved Problems

The following problems are from Gregory T. Lee abstract algebra book in SUMS.

■ **Problem 1.1.1** In S_4 let $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$ and $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$. Calculate the followings.

(a) $\sigma\tau$

(b) $\tau\sigma$

(c) the inverse of σ

Solution (a)

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}.$$

(b)

$$\tau\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix}.$$

(c)

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}.$$

■ **Problem 1.1.2** In S_5, let $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 2 & 1 & 4 \end{pmatrix}$ and $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 3 & 5 \end{pmatrix}$. Calculate the following.

(a) $\sigma\tau\sigma$

(b) $\sigma\sigma\tau$

(c) the inverse of σ

Solution (a)

$$\sigma\tau\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 1 & 3 & 2 \end{pmatrix}.$$

(b)

$$\sigma\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix}.$$

(c)

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 2 & 5 & 1 \end{pmatrix}.$$

■ **Problem 1.1.3** How many permutations are there in S_n? How many of those permutation satisfy $\alpha(2) = 2$?

Solution There are n choices for $\alpha(1)$, $n-1$ choices for $\alpha(2)$, and so on. So there are in total $n!$ elements in S_n. Fixing the value of $\alpha(2) = 2$ will leave 4 possible values for $\alpha(1)$, 3 possible values for $\alpha(3)$, and so on. Thus there will be $4! = 24$ permutations satisfying $\alpha(2) = 2$.

■ **Problem 1.1.4** Let H be the set of all permutations $\alpha \in S_5$ satisfying $\alpha(2) = 2$. Which of the properties, closure, associativity, identity, and inverse does H enjoy under composition of functions?

Solution Closure is satisfied: Let $\alpha, \beta \in H$. Then $\alpha(\beta(2)) = \alpha(2) = 2$ and also $\beta(\alpha(2)) = \beta(2) = 2$. Associativity is satisfied which follows from the axioms of the group. The identity of the group is in H, which is given by

$$e = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}.$$

Every element in H also has an inverse. Let $\alpha \in H$. Let $\tau \in S_5$ be its inverse. We have

$$\tau(2) = \tau(\alpha(2)) = e(2) = 2.$$

Thus $\tau \in H$.

■ **Problem 1.1.5** Consider the set of all functions from $\{1, 2, 3, 4, 5\}$ to $\{1, 2, 3, 4, 5\}$. Which of the properties, i.e. closure, associativity, identity, and inverse does this set enjoy under the composition of functions.

Solution The composition of any two functions is a function, thus the set is closed under composition. The associativity follows from the properties of the function composition. The identity function is the function that maps every element to itself hence is in the set. But not every function necessarily has an inverse (injectivity, and surjectivity is needed to guarantee the inverse).

■ **Problem 1.1.6** Give group tables for the following additive groups

(a) $U(12)$,

(b) S_3.

Solution (a)

$*$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

(b)

$*$	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$	$(2,0)$	$(2,1)$
$(0,0)$	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$	$(2,0)$	$(2,1)$
$(0,1)$	$(0,1)$	$(0,0)$	$(1,1)$	$(1,0)$	$(2,1)$	$(2,0)$
$(1,0)$	$(1,0)$	$(1,1)$	$(2,0)$	$(2,1)$	$(0,0)$	$(0,1)$
$(1,1)$	$(1,1)$	$(1,0)$	$(2,1)$	$(2,0)$	$(0,1)$	$(0,0)$
$(2,0)$	$(2,0)$	$(2,1)$	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$(2,1)$	$(2,1)$	$(2,0)$	$(0,1)$	$(0,0)$	$(1,1)$	$(1,0)$

■ **Problem 1.1.7** Give group tables from the following groups.

(a) $U(12)$.

(b) S_3.

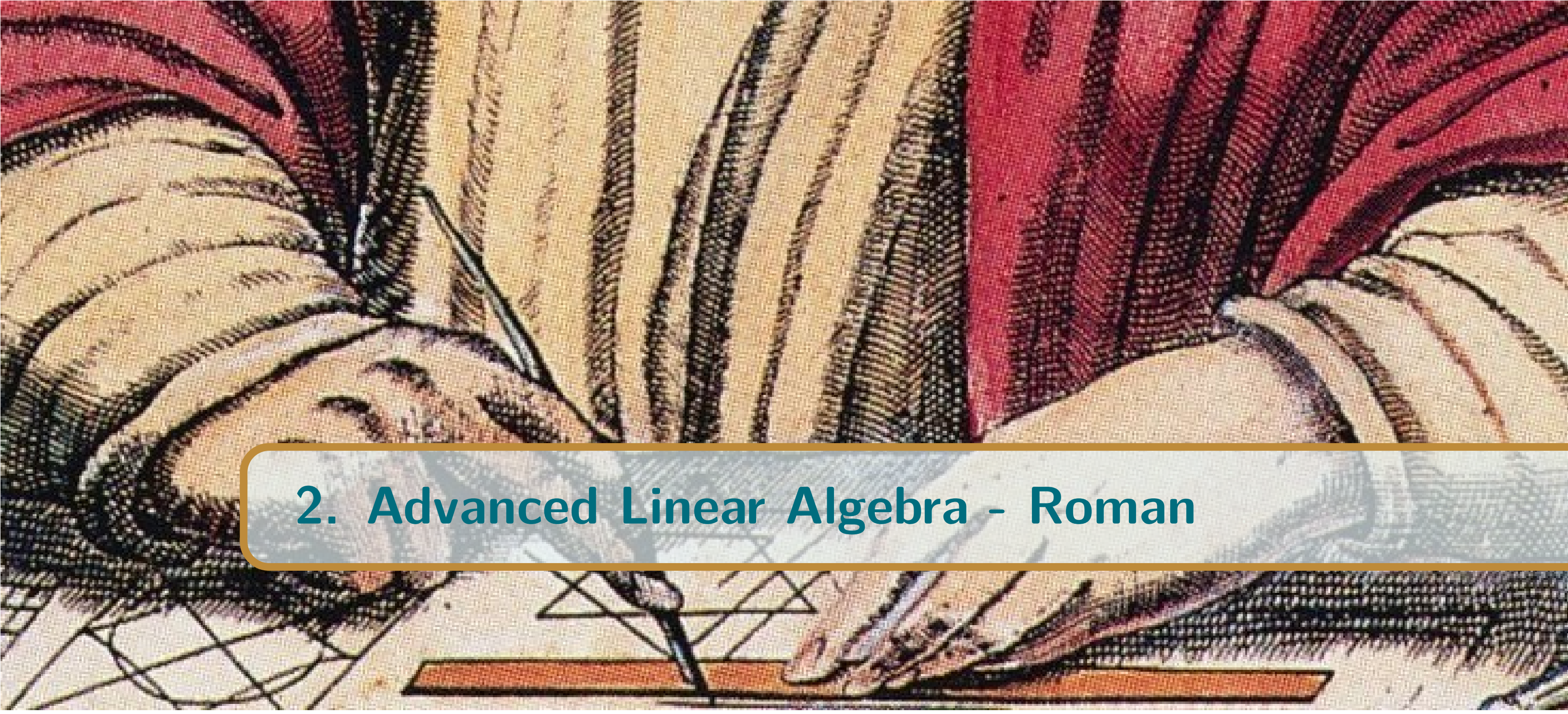
Solution (a) First observe that $U(12) = \{1, 5, 7, 11\}$. So

$*$	1	5	7	11
1	1	5	7	11
5	5	1	11	7
7	7	11	1	5
11	11	7	5	1

(b) Call the following permutations as $\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5$, and σ_6 respectively

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$

$*$	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6
σ_1	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6
σ_2	σ_2	σ_1	σ_5	σ_6	σ_3	σ_4
σ_3	σ_3	σ_4	σ_1	σ_2	σ_6	σ_5
σ_4	σ_4	σ_3	σ_6	σ_5	σ_1	σ_2
σ_5	σ_5	σ_6	σ_2	σ_1	σ_4	σ_3



2. Advanced Linear Algebra - Roman

2.1 Tensor Product

We start with the following proposition

Proposition 2.1 Let U, V be vector spaces. Then exists a unique linear map

$$\theta : U^* \otimes V^* \to (U \otimes V)^*,$$

defined by $f \otimes v = f \odot g$ where

$$(f \odot g)(u \otimes v) = f(u)g(v).$$

Moreover, θ is an embedding and is an isomorphism if U and V are finite dimensional. Thus the tensor product of linear functionals, i.e. $f \otimes g$ is a linear functions, i.e. $f \odot g$ on the tensor product.

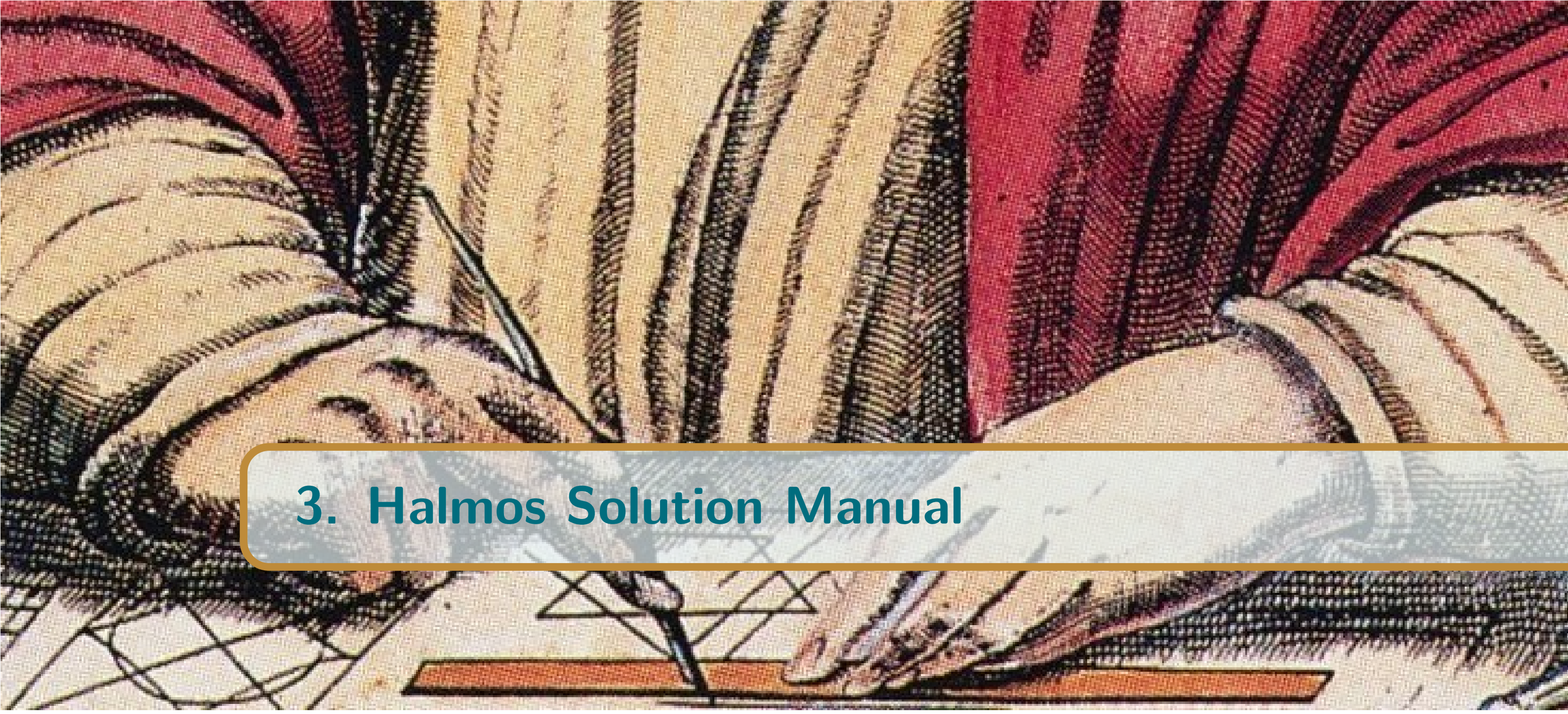
Proof. Fix some $f \in U^*$ and $g \in V^*$. Consider the bilinear map

$$G : U^* \times V^* \to (U \otimes V)^*$$

given by $G(f, g) = f \odot g$ where $(f \odot g)(u \otimes v) = f(u)g(v)$. This map $f \odot g$ exists, since the map $F_{f,g}(u, v) : U \times V \to F$ given by $F_{f,g}(u, v) = f(u)g(v)$ is bilinear, and by the universal property of the tensor product, there exist some linear map from $U \otimes V$ to F whose values matches $f(u)g(v)$, and we call this map $f \odot g$. The bilinear map G induces a linear map $\theta : U^* \otimes V^* \to (U \otimes V)^*$ given by

$$\theta(f \otimes g) = f \odot g.$$

For the rest of proof see Roman 14.7. □



3. Halmos Solution Manual

3.1 Fields

Observation 3.1.1 In a group we can only add one element an integer number of times with itself. For instance we can only have $\alpha + \alpha + \cdot + \alpha = m\alpha$ for some $m \in \mathbb{N}$. However, field is a generalization of group in the sense that we can have more general many times addition with itself for every element. For instance we can have $q\alpha$ where q is not necessarily an integer. The same is true in vector spaces. The fact that we can multiply a vector by some element of the underlying field (i.e. the scalar) shows this.

■ **Problem 3.1.1**

Solution (a) Holds because $(F, 0, +)$ is an abelean group.

(b) Since $(\mathcal{F}, 0, +)$ is an abelian group, then α has an inverse (i.e. $-\alpha$). Add this to both sides of the equation.

(c) We can write

$$\begin{aligned} \alpha + (\beta - \alpha) &= \alpha + (\beta + (-\alpha)) \\ &= \alpha + \beta + (-\alpha) && \text{(distributivity of multiplication)} \\ &= \alpha + (-\alpha) + \beta = \beta && ((F, 0, +) \text{ is abelian group}) \end{aligned}$$

(d) We can write

$$\alpha \cdot 0 = \alpha \cdot (0 + 0) = \alpha \cdot 0 + \alpha \cdot 0 = 2\alpha \cdot 0.$$

Adding the inverse of $\alpha \cdot 0$ to both sides we will get

$$\alpha \cdot 0 = 0.$$

(e) We can use the distributivity of multiplication and write

$$(-1)\alpha + \alpha = (-1 + 1)\alpha = 0 \cdot \alpha = 0.$$

Adding the inverse of α to both sides we will get

$$(-1)\alpha = -\alpha.$$

(f) We can write

$$\begin{aligned}
(-\alpha)(-\beta) &= (-1(\alpha))(-\beta) && \text{(property proved above)}\\
&= ((-1)(\alpha))((-1)(\beta)) && \text{(property proved above)}\\
&= (-1)(-1)\alpha\beta && \text{(associativity of product)}\\
&= -(-1)\alpha\beta = \alpha\beta
\end{aligned}$$

(g) If both α and β are zero, then it follows that $\alpha\beta = 0$. If one of them is not zero, WLOG we can assume $\beta \neq 0$, then β^{-1} exists, and multiplying it on the both sides we will get

$$\alpha = 0.$$

In a second run, I observed that my proof above might be wrong. I am somehow using the conclusion to prove the hypothesis.

■ **Problem 3.1.2**

Solution (a) No. $(F, 0, +)$ is not a group. $(F\backslash\{0\}, 1, \cdot)$ is not a group.

(b) No. $(F\backslash\{0\}, 1, \cdot)$ is not a group. The set of all integers is a Ring with identity.

(c) Yes. For instance, consider the set of all integers $\mathbb{Z}$. Let $\phi : \mathbb{Z} \to \mathbb{Q}$ be a bijection. Define the addition and multiplication as

$$m \oplus n = \phi^{-1}(\phi(m) + \phi(n)), \qquad m \odot n = \phi^{-1}(\phi(m) \cdot \phi(n)).$$

Let $z_1 = \phi^{-1}(1)$ and $z_0 = \phi^{-1}(0)$. Then we claim that $(\mathbb{Z}, z_0, z_1, \oplus, \odot)$ is a field. It is straightforward to check that $(\mathbb{Z}, z_0, \oplus)$, and $(\mathbb{Z}\backslash\{z_0\}, z_1, \odot)$ are groups, and the distributitivty law holds. For instance, to check for the associativity of addition let $m, n, l \in \mathbb{Z}$. Then we have

$$\begin{aligned}
(m \oplus n) \oplus l &= \phi^{-1}(\phi(m \oplus n) + \phi(l))\\
&= \phi^{-1}(\phi(m) + \phi(n) + \phi(l))\\
&= \phi^{-1}(\phi(m) + \phi(n \oplus l))\\
&= m \oplus (n \oplus l),
\end{aligned}$$

where we have used the fact that $\phi(m \oplus n) = \phi(m) + \phi(n)$.

■ **Problem 3.1.3**

Solution (a) We can solve this part with different levels of abstraction. But we want to use the fact that the multiplicative group $U(n)$ (i.e. the set of numbers in $\mathbb{Z}_n$ that are prime relative to n) is a group under multiplication (see example 3.3 Lee Abstract Algebra). When n is prime, then $U(n)$ contains all numbers $1, \cdots, n-1$. Thus when n is prime, $(\mathbb{Z}_n\backslash\{0\}, \cdot)$ and $(\mathbb{Z}_n, +)$ are both groups, and since the distribution law holds, it follows that $\mathbb{Z}_n$ is a field if and only if n is a prime number.

(b) $-1 = 4$ in $\mathbb{Z}_5$.

(c) $\frac{1}{3} = 5$ in $\mathbb{Z}_7$.

Observation 3.1.2 The field $\mathbb{Z}_p$ has characteristic p.

■ Problem 3.1.4

Solution First, observe that if the cardinality of the underlying set of the field is infinite, then we have $\underbrace{1+\cdots+1}_{m} = m \cdot 1 \neq 0$ for all $m \in \mathbb{N}$. However, when F is finite, then

$$F \text{ is finite } \implies m \cdot 1 = 0 \text{ for some } m \in \mathbb{N}.$$

We can see this by contrapositive. If $m \cdot 1 \neq 0$ for all $m \in \mathbb{N}$ then F has at least $\mathbb{N}$ many elements. We want to show that the smallest such m is prime. Assume otherwise. Then $m = pq$ for some $p, q \neq 0$. Then

$$0 = m \cdot 1 = pq \cdot 1 = pq.$$

Since every field is an integral domain (has no zero divisors) (see Theorem 8.9 Lee Abstract Algebra), then it implies that $p = 0$ or $q = 0$, this is a contradiction.

■ Problem 3.1.5

Solution (a) Yes. It is easy to check that $(\mathbb{Q}(\sqrt{2}), 0, +)$ and $(\mathbb{Q}(\sqrt{2})\backslash\{0\}, 1, \cdot)$ are groups. The associativity and closedness of the operators can be shown directly. For instance

$$(\alpha + \beta\sqrt{2})(\eta + \gamma\sqrt{2}) = (\alpha\eta + 2\beta\gamma) + \sqrt{2}(\alpha\gamma + \beta\eta),$$

hence the multiplication is closed. Also, $0, 1 \in \mathbb{Q}$ are the same as $0, 1 \in \mathbb{Q}(\sqrt{2})$. Also, it is easy to check that the additive inverse of $\alpha + \sqrt{2}\beta$ is $-\alpha - \sqrt{2}\beta$. And the multiplicative inverse is easy to calculate and follows from the observation that

$$(\alpha + \beta\sqrt{2}) \cdot (\frac{\alpha - \beta\sqrt{2}}{\alpha^2 - 2\beta^2}) = 1.$$

So the multiplicative inverse of $\alpha + \beta\sqrt{2}$ is

$$\frac{\alpha - \beta\sqrt{2}}{\alpha^2 - 2\beta^2}.$$

(b) No. Because $1 + \sqrt{2}$ has no inverse of the form $\alpha + \beta\sqrt{2}$ where $\alpha, \beta \in \mathbb{Z}$.

■ Problem 3.1.6

Solution (a) No. Not every polynomial has an inverse with integer coefficients. For instance, $p = 2x^2 - 1$ should be multiplied by

$$-1 + 2x^2 - 4x^4 + 8x^6 - \cdots$$

to get the 1 polynomial. But the expression above is not a polynomial.

(b) No. The same problem above. The set of all polynomials with integer or real coefficients forms a commutative Ring.

■ Problem 3.1.7

Solution (a) The addition part of OK! I.e. $(F, (0,0), +)$ forms an abelian group. However, $(F \backslash \{(0,0)\}, \mathbb{1}, \cdot)$ does not form a group as defined above. Because by the provided definition of multiplication we need to have $(\alpha, \beta)\mathbb{1} = (\alpha, \beta)$ that implies that the only choice for $\mathbb{1}$ is

$$\mathbb{1} = (1, 1).$$

But then the elements $(0,1)$ and $(1,0)$ have no multiplicative inverses. This is not the only obstacle though.

(b) Yes. This multiplication resolves the obstacles above and $(F \backslash \{(0,0)\}, \mathbb{1}, \cdot)$ is an abelian group. It is easy to check that the multiplicative identity should be

$$\mathbb{1} = (1, 0).$$

I.e. this is the only choice that satisfies $(\alpha, \beta) \cdot \mathbb{1} = (\alpha, \beta)$. It is also easy to check that the inverse for a non-zero element (α, β) is

$$(\frac{\alpha}{\alpha^2 + \beta^2}, \frac{-\beta}{\alpha^2 + \beta^2}).$$

(c) It will lead to the same kind of structure.

3.2 Vector Spaces

■ Problem 3.2.1

Solution (a) This follows from $(V, 0, +)$ being an abelian group.

(b) The additive inverse of the zero element in an additive group is itself. So this follows from $(V, 0, +)$ being an abelian group.

(c) We can write

$$\alpha \cdot 0 = \alpha \cdot (0 + 0) = \alpha \cdot 0 + \alpha \cdot 0$$

Since the set of vectors is an additive abelian group, we can add the inverse of $\alpha \cdot 0$ to both sides and get

$$\alpha \cdot 0 = 0.$$

(d) We can write

$$0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x.$$

Since the set of vectors is an additive abelian group, then adding the inverse of $0 \cdot x$ to both sides we will get

$$0 \cdot x = 0.$$

■ Remark Note that in the expression above, the zero on the LHS is the zero element of the field, and the zero on the RHS is the zero element of the vector field.

(e) Still thinking. I was trying to do a similar proof as for problem 1 part (g), but I realized that my proof for that part is also not correct.

(f) We can add $x = 1 \cdot x$ to $(-1)x$. Then using the distributivity law we can write

$$x + (-1)x = 0.$$

Adding the additive inverse of x to both sides we will get

$$(-1)x = -x.$$

(g) We can write

$$y + (x - y) = 1 \cdot y + 1 \cdot (x - y) = 1 \cdot (y + x + (-y)) = 1 \cdot x = x.$$

■ **Problem 3.2.2** The elements of $\mathbb{Z}_p^n$ are the n-tuples, or equivalently the set of all functions $f : [p] \to \mathbb{Z}_p$ where $[n] = \{1, 2, \cdots, n\}$. There are p^n such functions.

■ **Problem 3.2.3**

Solution No. One of the immediate problems that I can see is that the scalar 1 does not interact nicely with the vectors. I.e. in the vector space axioms we have $1 \cdot x = x$ for all $x \in V$. However, in the definition above we have $1 \cdot (\xi_1, \xi_2) = (1\xi_1, 0) = (\xi_1, 0) \neq (\xi_1, \xi_2)$.

■ **Problem 3.2.4**

Solution We assume that the vector space $\mathbb{C}^3$ is defined on $\mathbb{C}$ rather than $\mathbb{R}$.

(a) No. While the vector space $(V, 0, +)$ forms a group, but the scalars does not behave nicely. For instance $i \cdot (r_1, \xi_2, \xi_3) = (ir_1, \xi_2, \xi_3)$ and the first argument is not longer a real number.

(b) Yes.

(c) No. because $(\xi_1, 0, \xi_2) + (0, \tilde{\xi}_2, \tilde{\xi}_3) = (\xi_1, \tilde{\xi}_2, \xi_2 + \tilde{\xi}_3)$ and neither its first or second argument is zero.

(d) The vector space $(V, 0, +)$ forms a group. The addition is closed: Let (ξ_1, ξ_2, ξ_3) and $(\tilde{\xi}_1, \tilde{\xi}_2, \tilde{\xi}_3)$ be in the subspace. Then

$$\alpha(\xi_1, \xi_2, \xi_3) + \beta(\tilde{\xi}_1, \tilde{\xi}_2, \tilde{\xi}_3) = (\alpha\xi_1 + \beta\tilde{\xi}_1, \alpha\xi_2 + \beta\tilde{\xi}_2, \xi_3 + \tilde{\xi}_3).$$

Since

$$\alpha(\xi_1 + \xi_2) + \beta(\tilde{\xi}_1 + \tilde{\xi}_2) = 0,$$

then the sum is also in the subspace, and the addition is closed. Also, in additive inverse of (ξ_1, ξ_2, ξ_3) is

$$(-\xi_1, -\xi_2, -\xi_3)$$

where since $-(\xi_1 + \xi_2) = 0$ it implies that the inverse is also in the subspace. It is also easy to check that the scalars behave nicely with the vector space.

(e) No. This subspace does not contain the origin $(0, 0, 0)$.

■ **Problem 3.2.5**

Solution The answers to this question depends on the set where the coefficients of the polynomial belongs, as well as the scalars on which the vector space is defined. We assume that the coefficients are complex numbers and the scalars are also complex numbers.

(a) Yes.

(b) Yes. It is easier to see this if we identify each such polynomial with a 4-tuple (a_0, a_1, a_2, a_3), where these coordinates record the coefficients of the monomials $1, x^1, x^2$, and x^3 respectively. Then the subset of interest is

$$\tilde{V} = \{(a_0, a_1, a_2, a_3) : a_1 = 2a_0\}.$$

This is subspace. Because the structure $(\tilde{V}, +, 0)$ is a group (it is easy to check that under addition of two such tuples the resulting tuple still satisfies $\xi_1 = 2\xi_0$). Also, it is straightforward to see that $(0, 0, 0, 0) \in \tilde{V}$. And the inverse of $(a_0, 2a_0, a_2, a_3)$ is $(-a_0, -2a_0, -a_2, -a_3)$.

(c) No. let $x \in V$. So $x(t) \geq 0$ for $t \in [0,1]$. Let $\alpha = 2$ be a scalar. Then $\alpha x(t) \leq 0$ for $t \in [0,1]$. So $\alpha x(t) \notin V$.

(d) Yes. The specified condition leads to

$$a_0 + a_1 t + a_2 t^2 + a_3 t^3 = a_0 + a_1(1-t) + a_2(1-t)^2 + a_3(1-t)^3.$$

This simplifies to

$$(2t-1)a_1 + (t^2 - (1-t)^2)a_2 + (t^3 - (1-t)^3)a_3 = 0.$$

Observe that

$$t^2 - (1-t)^2 = 2t - 1.$$

Also

$$t^3 - (1-t)^3 = 2t^3 - 3t^2 + 3t - 1.$$

So we will have

$$(2t-1)a_1 + (2t-1)a_2 + (2t^3 - 3t^2 + 3t - 1)a_3 = 0.$$

This is a hyperplane passing through the origin in $\mathbb{R}^4$ (if we identify each polynomial with a 4-tuple). This defines a subspace. Note: We could have said this without simplifying the coefficients, and I did that for no good reason!

3.3 Bases

Observation 3.3.1 — Dimension depends on basis. Consider the vector space $\mathbb{C}^1$. Let $z_1, z_2 \in \mathbb{C}^1$ be any non-zero vectors. Then z_1, z_2 are linearly dependent vectors. The reason that I am highlighting this is that I was putting too much emphasis on the looking at $\mathbb{C}^1$ as $\mathbb{R}^2$ that I was not expecting to see that any two non-zero vectors in $\mathbb{C}^1$ is linearly dependent (which is definitely false, in $\mathbb{R}^2$ as not any two non-zero vectors are linearly dependent). Then I realized that it is the magic of scalars that makes the difference. If the vector space of consideration is $(\mathbb{C}^1, \mathbb{R}^1)$, i.e. the scalars are real numbers, then not every two vectors in $\mathbb{C}^1$ is linearly dependent. For instance the vectors 1 and i are linearly independent (as there are no ways to multiply 1 at a real number and get i). But in the case of $(\mathbb{C}^1, \mathbb{C})$ any two vectors are linearly dependent. Because in the example above we can multiply i be $-i$ and get 1. I.e. in the second case the scalars are not only for enlarging the vectors, but also to rotate them.

Observation 3.3.2 — $\mathbb{R}$ can be infinite dimensional. We often omit the underlying field when talking about the vector space $\mathbb{R}$. But it turns out to be crucial when considering the dimension of the space. For instance, it is easy to see that $(\mathbb{R}, \mathbb{R})$ (i.e. the underlying field is $\mathbb{R}$) is one dimensional. However, $(\mathbb{R}, \mathbb{Q})$ is infinite dimensional.

■ **Problem 3.3.1**

Solution To show that x, y, z are linearly independent, consider the linear combination

$$\alpha x + \beta y + \gamma z = 0.$$

Considering the components of x, y, z we will have

$$\begin{cases} \alpha = 0, \\ \beta = 0, \\ \gamma = 0. \end{cases}.$$

This $\{x, y, z\}$ is linearly independent. For other combinations of vectors, one can easily mimic a similar approach.

■ **Problem 3.3.2**

Solution $\boxed{\Rightarrow}$ We want to show that $1, \xi$ being linearly independent implies $\xi \notin \mathbb{Q}$. We use prove by contrapositive. Let $\xi \in \mathbb{Q}$. Then we will have

$$\alpha + \beta\xi = 0$$

when $\beta = -\alpha/\xi \in \mathbb{Q}$. So $1, \xi$ is not linearly independent.
$\boxed{\Leftarrow}$ We want to show that $\xi \notin \mathbb{Q}$ implies $1, \xi$ are linearly independent. We again use prove by contrapositive. Assume $1, \xi$ be linearly dependent. Then

$$\alpha + \beta\xi = 0$$

holds while not both of $\alpha, \beta = 0$. However note that either α or β being zero forces the other one to be zero as well. So $\xi = -\alpha/\beta$ thus $\xi \in \mathbb{Q}$. This completes the proof.

■ **Problem 3.3.3**

Solution Consider the linear combination

$$\alpha(x + y) + \beta(y + z) + \gamma(z + x) = 0.$$

By rearranging the terms we will have

$$(\alpha + \gamma)x + (\alpha + \beta)y + (\beta + \gamma)z = 0.$$

Invoking the linear independence of x, y, z we will have

$$\begin{cases} \alpha = -\gamma, \\ \alpha = -\beta, \\ \beta = -\gamma. \end{cases}$$

This implies $\alpha = \gamma = \beta = 0$. So the specified vectors are linearly independent.

■ **Problem 3.3.4**

Solution (a) Let $A = (1 + \xi)$ and $B = (1 - \xi)$. Consider the equation

$$\alpha \begin{bmatrix} A \\ B \end{bmatrix} + \beta \begin{bmatrix} B \\ A \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

We can write the system of equations as

$$\begin{pmatrix} A & B \\ B & A \end{pmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

If the determinant of the matrix above is non-zero, then the only solution is $\alpha = 0, \beta = 0$ (multiply both sides at the inverse of the matrix). Thus the vectors are linearly dependent if we have

$$A^2 - B^2 = 0.$$

This implies

$$\xi = 0.$$

(b) Similar to the solution above we need to have

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

This happens when

$$\xi^3 = 2\xi,$$

The solutions for the equation above is

$$\xi = 0, \qquad \xi = \pm\sqrt{2}.$$

(c) Then the vectors are linearly independent only when $\xi = 0$ (note that ξ can not take the values $\pm\sqrt{2}$.)

■ **Problem 3.3.5**

Solution (a) In order for (ξ_1, ξ_2) and (η_1, η_2) be linearly dependent, we need to have

$$\det\begin{pmatrix}\xi_1 & \eta_1\\ \xi_2 & \eta_2\end{pmatrix} = 0.$$

This holds if $\xi_1\eta_2 = \xi_2\eta_1$.

(b) For two vectors $x, y \in \mathbb{C}^3$ to be *linearly independent*, the following matrix need to have rank 2.

$$\begin{pmatrix}x_1 & y_1\\ x_2 & y_2\\ x_3 & y_3\end{pmatrix}.$$

This does not hold true if

$$x_1y_2 = x_2y_1, \quad \text{and} \quad x_2y_3 = x_3y_2.$$

And for three vectors $x, y, z \in \mathbb{C}^3$ to be linearly dependent, the matrix below should have zero determinant:

$$\det\begin{pmatrix}x_1 & y_1 & z_1\\ x_2 & y_2 & z_2\\ x_3 & y_3 & z_3\end{pmatrix} = 0.$$

(c) No.

■ **Problem 3.3.6** Skipped. Will be added later.

■ **Problem 3.3.7** The following basis will do the job

$$\mathcal{B}_1 = \left\{ \begin{bmatrix}0\\0\\1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix} \right\},$$

and

$$\mathcal{B}_1 = \left\{ \begin{bmatrix}0\\0\\1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix} \right\}.$$

■ **Problem 3.3.8** Skipped. Will be added later.

■ **Problem 3.3.9** Skipped. Will be added later.

■ **Problem 3.3.10** Skipped. Will be added later.

■ **Problem 3.3.11** Skipped. Will be added later.

3.4 Isomorphism

Observation 3.4.1 — A better proof of Theorem in §9 . Here we give a more formal prove of the theorem below.

Theorem 3.1 Every n-dimensional vector space (defined over the field F) is isomorphic to F^n.

Proof. Let $\{b_i\}$ be a basis for V and let $\{e_i\}$ be the standard basis for F^n. Define the map $\phi : V \to F^n$ by

$$\phi(b_i) = e_i$$

and extend to the whole space by linearity. We claim that ϕ is an isomorphism. First, observe that ϕ is a function. Indeed $x \in V$ can be expanded $x = \sum_i \alpha_i b_i$, and

$$T(x) = T(\sum_i \alpha_i b_i) = \sum_i \alpha_i T(b_i) = \sum_i e_i.$$

Also, ϕ is linear by construction. Furthermore, we claim that ϕ is injective. To show this we prove that $\ker T = \{0\}$. Let $x \in \ker T$. We expand $x = \sum_i \alpha_i b_i$ and by applying T we have

$$0 = T(x) = T(\sum_i \alpha_i b_i) = \sum_i \alpha_i T(b_i) = \sum_i \alpha_i e_i.$$

Since $\{e_i\}$ is linearly independent, it follows that $\alpha_i = 0$ for all i, thus $x = 0$. Also, we claim that T is surjective. To see this let $y \in F^n$ with expansion $y = \sum_i \alpha_i e_i$. Let $x = \sum_i \alpha_i b_i$. We claim that $T(x) = y$. Indeed

$$T(x) = T(\sum_i \alpha_i b_i) = \sum_i \alpha_i T(b_i) = \sum_i \alpha_i e_i = y.$$

□

■ **Problem 3.4.1**

Solution (a) For $V = (\mathbb{C}, \mathbb{R})$ the dimension is 2. Let $\{1, i\} \subset \mathbb{C}$. Construct the map

$$\phi : \mathbb{C} \to \mathbb{R}^2,$$

given by $\phi(1) = e_1, \phi(i) = e_2$, extended to the whole space by linearity. It is straight forward to see that ϕ is an isomorphism of vector spaces. Thus V is two dimensional.

(b) We can choose the basis $\{(1, 0 \cdots, 0), (i, 0, \cdots, 0), (0, 1, 0, \cdots, 0), (0, i, 0, \cdots, 0), \cdots (0, \cdots, 0, i)\}$, and we can construct an isomorphism to $\{e_1, e_2, \cdots, e_{2n}\}$ similar to above. Thus V^- is $2n$ dimensional.

■ **Problem 3.4.2**

Solution No! $(\mathbb{R}, \mathbb{Q})$ is in fact infinite dimensional vector space with $\dim_F V = \mathfrak{c}$. A basis for this space is $\mathbb{R}/\mathbb{Q}$.

■ **Problem 3.4.3**

Solution For an n-dimensional vector space over the field $\mathbb{Z}_p$ there is an isomorphism to $(\mathbb{Z}_p)^n$. There are p^n vectors in this space. To see this note that the product space $(\mathbb{Z}_p)^n$ is the space of all functions $f : [n] \to \mathbb{Z}_p$ where $f(i) \in \mathbb{Z}_p$ (using the notation $[n] = \{1, 2, \cdots, n\}$). There are p^n such maps.

■ **Remark** The functions in the product space is the same as the tuples if the dimension is finite.

■ **Problem 3.4.4**

Solution We assume that the vector spaces of interest of finite dimensional. Then the assertion is not true. Because $\mathbb{Q}^n$ is countable for all $n \in \mathbb{N}$, thus its cardinality is $\aleph_0$. However, dimension is an isomorphism invariant. So the assertion is not true. In particular, $\mathbb{Q}^2$ and $\mathbb{Q}^3$ both as the same cardinality $\aleph_0$, but they are not isomorphic (because they have different dimension).

3.5 Dimension of a subspace

Observation 3.5.1 Denote by $\mathcal{S}$, the set of all subspaces of a vector space V. Observe that $V, \varnothing \in \mathcal{S}$, and $\mathcal{S}$ is closed under intersection (it is not closed under union).

Observation 3.5.2 — Definition of spanning. In this definition box, I am highlighting the definition of the span of a set.

Definition 3.1 — Span of a set. Let $S \subset V$ where V is some finite dimensional vector space. The span of S, denoted by $\operatorname{Span} S$, is the smallest linear subspace containing all S.

Proposition 3.1 Let S, V be as in the definition above. Then $\operatorname{Span} S$ is the same as the all linear combinations of vectors in S.

Proof. The proof will have two parts. First, we prove that

$$\operatorname{Span} S \subset \{\sum_i \alpha_i s_i | s_i \in S\}.$$

For easier notation we denote $W = \{\sum_i \alpha_i s_i | s_i \in S\}$. To see the set inclusion above observe that W is a subspace. Indeed, any linear combination of linear combinations of vectors in S can be written as a linear combination of vectors in S. Also, note that W contains S. However, by definition $\operatorname{Span} S$ is the smallest subspace containing S. So W must also contain S, i.e. $S \subset W$. Now we prove the reverse inclusion

$$W \subset \operatorname{Span} S.$$

To see this note that $\operatorname{Span} S$ is a subspace, and it contains S. So it also contains all linear combination of vectors in S as well. So it contains W. I.e. $W \subset S$. □

The proof of the following proposition is left to the reader in §11 Halmos.

Proposition 3.2 Let K, H be two linear subspaces of V and let $S = \text{Span}(K \cup H)$. Then S is the same as the set of all vectors of the form $x + y$ where $x \in K$ and $y \in H$.

Proof. From the proposition above we know that S contains all of linear combination of of elements of $K \cup H$. Such a linear combination can be written as $x+y$ where $x \in K$ and $y \in H$, and any linear combination of the form $x+y$ is a linear combination of elements of $K \cup H$. So S contains all such vectors. ☐

■ **Problem 3.5.1**

Solution Let $\{b_i\}_{i=1}^n$ be a basis for M. Since $M \subset N$, and both of the are subspaces, then we can extend this basis to a basis on N. Since N is also n-dimensional, $\{b_i\}_{i=1}^n$ is already a basis for N (adding any new vector will make the collection linearly dependent). Let $y \in N$. We can expand $y = \sum_i \alpha_i b_i$. This implies $y \in M$. So $N \subset M$. Then it follows that $M = N$.

■ **Problem 3.5.2**

Solution We state the problem in a better language: If M, N are subspaces of a vector space V, and if $V \subset M \cup N$, then $V = M$ or $V = N$. We prove by contrapositive. Assuming $M \subsetneq V$ and $N \subsetneq V$. Then $\exists x \in V$ such that $x \notin M$ and $x \notin N$. I.e. $x \in M^c \cap N^c = (M \cup N)^c$. Thus $x \notin M \cup C$. This completes the proof.

■ **Problem 3.5.3**

Solution We want to show that

$$\text{Span}\{x, y\} = \text{Span}\{y, z\}.$$

Let $v \in \text{Span}\{x, y\}$. Then $v = \alpha x + \beta y$ for some α, β in the underlying field. From $x + y + z = 0$ we can write $x = -y - z$. Substituting this in the expansion we will get

$$v = \alpha(-y - z) + \beta y = (\beta - \alpha)y - \alpha z \in \text{Span}\{y, z\}.$$

Showing the reverse inclusion is very similar.

■ **Problem 3.5.4**

Solution We have

$$x, y \in V, \qquad M \subset_L V,$$

and

$$H = \text{Span}\{x, M\}, \qquad K = \text{Span}\{y, M\}.$$

We assume $y \in H \cap M^c$. If $x \in M$, then it immediately follows that $x \in K$. Assuming $x \notin M$ we will have $y \in \text{Span}\{x\}$. So $y = \alpha x$ for some scalar α. On the other hand we can write $x = y/\alpha$, thus $x \in \text{Span}\{y\}$, hence $x \in K$.

■ **Problem 3.5.5**

Solution (a) This fails to hold if $L \subset M + N$ but $L \cap M = L \cap N = \{0\}$. For instance, in $\mathbb{R}^2$, let $M = \text{Span}\{e_1\}$, and $N = \text{Span}\{e_2\}$, and $L = \text{Span}\{e_1 + e_2\}$. Then

$$L \cap (M + N) = L, \qquad (L \cap M) + (L \cap N) = \{0\}.$$

(b) First, we prove the following inclusion

$$L \cap (M + (L \cap N)) \subset (L \cap M) + (L \cap N).$$

Let $x \in L \cap (M + (L \cap N))$. This implies that $x = m + n$ for $m \in M$ and $n \in L \cap N$, and also $x \in L$. So we cam write $m = x - n$. Observe that $x, n \in L$, this implies $m \in L$. And since $m \in M$ as well, we will have $m \in L \cap M$.

$$x = (x - n) + n,$$

where $x - n \in L \cap M$ and $n \in L \cap N$. Thus $x \in (L \cap M) + (L \cap N)$. Now we prove the reverse inclusion

$$L \cap (M + (L \cap N)) \supset (L \cap M) + (L \cap N).$$

let $y \in (L \cap N) + (L \cap N)$. Then we can write $y = \ell_1 + \ell_2$ where $\ell_1 \in L \cap M$ and $\ell_2 \in L \cap N$. Notice that it follows $x \in L$. Also, since $\ell_1 \in L \cap M$, it means that $\ell_1 \in L$ and $\ell_1 \in M$. So for $x = \ell_1 + \ell_2$ we can state that $\ell_1 \in L$, $\ell_2 \in L \cap N$, thus $x \in L + (L \cap M)$. Note that what we are trying to say is the trivial fact that $(L \cap M) + (L \cap N) \subset L + (L \cap N)$.,m;n',mn';

■ **Remark** Part (a) above looks like the following identity regarding the closure of sets in topology:

$$\overline{(A \cap B)} \neq \overline{A} \cap \overline{B}.$$

■ **Problem 3.5.6**

Solution (a) No! Let M be a non-trivial subspace of V, and let $\{b_i\}_{i=1}^m$ be a basis for M. We can extend this basis to $\{b_i\}_{i=1}^n \cup \{b_i\}_{i=n+1}^{\dim V}$. Any choice of $\{b_i\}_{i=n+1}^{\dim V}$ will lead to a complement subspace to M. So it is not unique.

(b) Let $\{b_i\}_{i=1}^m$ be a basis for M. This can be extended to $\{b_i\}_{i=1}^m \cup \{b_i\}_{i=m+1}^n$, a basis for the whole space, where $\{b_i\}_{i=m+1}^n$ is a basis for the complement of M. It is easy to see that the complement has dimension $n - m$.

■ **Problem 3.5.7**

Solution (a) Assume otherwise. Let $\{b_1, b_2, b_3\} \subset M$ and $\{a_1, a_2, a_3\} \subset N$ be a basis for the corresponding spaces. Since M, N are assumed to be disjoint, then $M \cap N = \{0\}$, thus $\{b_1, b_2, b_3, a_1, a_2, a_3\}$ is a basis for $M + N$. $M + N$ is itself a subspace of V. But it dimension (i.e. 6) is larger than the dimension of V (i.e. 5). So this is a contradiction.

(b) Start with a basis for $M \cap N$ given as $\{b_i\}_{i=1}^l$. Since $M \cap N \subset N$ and $M \cap N \subset M$ we can extend this basis to a basis to M, as well as N. I.e.

$$\mathcal{B}_1 = \{b_i\}_{i=1}^l \cup \{a_i\}_{i=1}^{d_1}, \qquad \mathcal{B}_2 = \{b_i\}_{i=1}^l \cup \{c_i\}_{i=1}^{d_2},$$

where $\mathcal{B}_1$ is a basis for M and $\mathcal{B}_2$ is a basis for N. Note that $d_1 = \dim M - l$ and $d_2 = \dim N - l$. We claim that the collection

$$\mathcal{B} = \{b_i\}_{i=1}^l \cup \{a_i\}_{i=1}^{d_1}, \cup \{c_i\}_{i=1}^{d_2}$$

is a basis for $N + M$. Indeed, each $\{a_i\}_{i=1}^{d_1}$ and $\{c_i\}_{i=1}^{d_2}$ is linearly independent set. However, since $\text{Span}\{a_i\}_{i=1}^{d_1}$ is a complement subspace to $M \cap N$ with respect to M, and $\text{Span}\{c_i\}_{i=1}^{d_2}$ is a complement subspace to $M \cap N$ with respect to N, then $\{a_i\}_{i=1}^{d_1} \cup \{c_i\}_{i=1}^{d_2}$ is a linearly independent collection (otherwise, if a_i can be written as a linear combination of $\{c_i\}_{i=1}^{d_2}$, then $a_i \in N$, and since $a_i \in M$, it follows that $a_i \in M \cap N$, which is a contradiction). So

$\mathcal{B}$ is a collection of linearly independent vectors. We also claim that $\mathcal{B}$ is spanning. Because $v \in M + N$ can be written as $v = m + n$ for $m \in M$ and $n \in N$. So we can write m as a linear combination of vectors in $\{b_i\}_{i=1}^{l} \cup \{a_i\}_{i=1}^{d_1}$ and n as a linear combination of vectors in $\mathcal{B}_2 = \{b_i\}_{i=1}^{l} \cup \{c_i\}_{i=1}^{d_2}$. There are $\dim M + \dim N - l$ vectors in $\mathcal{B}$. So

$$\dim(M + N) = \dim M + \dim N - \dim M \cap N.$$

■ **Problem 3.5.8**

Solution (a) Let $\mathcal{E}$ be the set of all polynomials that their evaluation at t is the same as their evaluation at $-t$ for all $t \in \mathbb{R}$. let $p_1, p_2 \in \mathcal{E}$. Then evaluating $\alpha p_1 + \beta p_2$ at $t \in \mathbb{R}$ we will get

$$E_t(\alpha p_1+\beta p_2) = (\alpha p_1+\beta p_2)(t) = \alpha p_1(t)+\beta p_2(t) = \alpha p_1(-t)+\beta p_2(-t) = (\alpha p_1+\beta p_2)(-t) = E_{-t}(\alpha p_1+\beta p_2),$$

where E_t is the evaluation functional at t. From above we see that $\alpha p_1 + \beta p_2 \in \mathcal{E}$. Thus $\mathcal{E}$ is a subspace. The proof for the odd subspace is very similar.

(b) It only remains that the intersection of these subspaces is $\{0\}$. Let $p \in \mathcal{P}$ be a polynomial that is both even and odd. Then

$$E_t(p) = p(t) = p(-t) = E_{-t}(p), \qquad E_t(p) = p(t) = -p(-t) = -E_{-t}(p).$$

It follows that $p(-t) = -p(-t)$. Since the codomain of the evaluation functional is $\mathbb{R}$, ir follows that $p(t) = 0$ for all t. Since the only vector that vanishes on any linear functional is the origin, it follows that $p = 0$.

■ **Remark** Geometric argument

3.6 Dual Spaces

■ **Problem 3.6.1**

Solution (a) Linear functional.

(b) Linear functional.

(c) Not a linear functional

(d) Not a linear functional. Because the codomain of a linear functional is the underlying field, that is $\mathbb{R}$ in this case. But the codomain of f is $\mathbb{C}$.

(e) Not a linear functional.

■ **Problem 3.6.2**

Solution (a) Linear functional.

(b) Not a linear functional.

(c) Not a linear functional.

(d) Linear functional.

■ **Problem 3.6.3**

Solution (a) Linear functional.

(b) Not a linear functional.

(c) Linear functional.

(d) Linear functional.

(e) Linear functional.

(f) Linear functional.

■ Problem 3.6.4

Solution (a) Let $x_1, x_2 \in \mathcal{P}$ with degree n_1, n_2 respectively. WLOG we can assume $n_2 \geq n_1$. Then we can write

$$x_1(t) = \sum_{i=1}^{n_2} \xi_i t^i, \qquad x_2(t) = \sum_{i=1}^{n_2} \eta_i t^i,$$

where $\xi_i = 0$ for all $i > n_1$. Applying y we will get

$$y(x_1 + x_2) = \sum_{i=1}^{n_2} (\xi_i + \eta_i)\alpha_i = \sum_{i=1}^{n_2} \xi_i \alpha_+ \sum_{i=1}^{n_2} \eta_i \alpha_i = \sum_{i=1}^{n_1} \xi_i \alpha_+ \sum_{i=1}^{n_2} \eta_i \alpha_i = y(x_1) + y(x_2).$$

where we used the fact that $\xi_i = 0$ for all $i > n_1$.

To show that any linear functional on $\mathcal{P}$ is determined this way, let y be any linear functional. Define $\alpha_i = y(t^i)$. Let $x(t) = \sum_{i=1}^{n} \xi_i t^i$ be any polynomial. Then

$$y(\sum_{i=1}^{n} \xi_i t^i) = \sum_{i=1}^{n} \xi_i y(t^i) = \sum_{i=1}^{n} \xi_i \alpha_i.$$

■ Problem 3.6.5

Solution Yes. Let $\{b_i\}$ be a basis for the space. There is at least one basis vector for which $y(b_i) \neq 0$, otherwise y is a zero functional. Let $\beta = y(b_i)$. Then for $x = \frac{\alpha}{\beta} b_i$ we have $y(x) = \alpha/\beta \cdot \beta = \alpha$.

■ Problem 3.6.6

Solution To answer this question we will be using some facts that will be covered later in the book. Since $y(x) = 0$ whenever $z(x) = 0$, this implies that $\ker z \subset \ker y$. Choose $x_0 \notin \ker z$. Define $\alpha = y(x_0)/z(x_0)$. Since $\ker z$ and $\mathrm{Span}\{x_0\}$ are complement subspaces, for any $x \in V$ we can write

$$x = w + t x_0,$$

where $w \in \ker z$ and $t x_0 \in \mathrm{Span}\{x_0\}$. Applying y, z we will get

$$z(x) = t z(x_0), \qquad y(x) = t y(x_0).$$

Using the definition of α, we can see that $y(x) = \alpha z(x)$.

Proposition 3.3 It is easy from the problem above, that for two linear functionals if $\ker z \subset \ker y$, then it follows that $\ker y \subset \ker z$ (using the fact that $y = \alpha z$). Thus $\ker y = \ker z$.

3.7 Annihilators

■ **Problem 3.7.1**

Solution Let $f = \alpha_1 e^1 + \alpha_2 e^2 + \alpha_3 e^3$. The following values for the coefficients is one choice to satisfy the condition $y(x_1) = y(x_2)$.

$$\alpha_1 = 1, \quad \alpha_2 = -1, \quad \alpha_3 = 0.$$

■ **Problem 3.7.2**

Solution Let

$$y_j = \sum_{i=1}^{3} \alpha_{ji} x^i,$$

where x^i is the dual basis. We need to have $y_j(x_i) = \delta_{ij}$. I.e. we will have the following matrix equation

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

It is easy to see

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}.$$

So the result of $y_1(x), y_2(x)$, and $y_3(x)$ will be the columns of the following matrix on the RHS

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

■ **Problem 3.7.3**

Solution Denote this set as $\ker y$. Let $v, u \in \ker f$ and α, β scalars. Then

$$y(\alpha u + \beta v) = \alpha y(u) + \beta y(v) = 0.$$

So $\alpha u + \beta v \in \ker y$.

■ **Problem 3.7.4**

Solution Let $x \in \ker y$. Assume $x = \sum \xi_i e_i$. Then we will have

$$y(x) = \xi_1 + \xi_2 + \xi_3 = 0.$$

Writing $\xi_3 = -\xi_1 - \xi_2$. Thus the components of x is $x = (\xi_1, \xi_2, -\xi_1 - \xi_2)$. Thus we can parameterize any $x \in \ker y$ by ξ_1, ξ_2 only. We can write the components of x as

$$(\xi_1, \xi_2, -\xi_1 - \xi_2) = \xi_1(1, 0, -1) + \xi_2(0, 1, -1).$$

So any $x \in \ker y$ can be written as the linear combination above. So

$$\mathcal{B} = \{(1, 0, -1), (0, 1, -1)\},$$

is a basis for $\ker y$.

■ **Problem 3.7.5** Skipped. Will be added later.

■ **Problem 3.7.6** Skipped. Will be added later.

■ **Problem 3.7.7**

Solution We will have a combinatorical proof. First, pick one non-zero vector from the space. Recall that a vector space over a finite filed has q^p vectors (where p is the characteristic of the field). There are $q^p - 1$ possible choices. Now pick another vector independent of this vector. We can not chose q vectors that are in the span of the first chosen vector. So there are $q^p - p$ possible vectors. Choose the third vector, linearly independent from the first two, i.e. it should not belong to the span of the first two vectors (where there are q^2 of them). So there are $q^p - q^2$ choices. So to choose m linearly independent vectors (ordered basis) we have

$$(q^n - 1)(q^n - q^2)(q^n - q^3)\cdots(q^n - q^{n-m+1}).$$

To count the number of distinct subspaces, we need to divide the number above to the number of all ordered basis of a m-dimensional vector space, where we use a similar method of counting, and we will have

$$(q^m - 1)(q^m - q^2)(q^m - q^3)\cdots(q^n - q^m).$$

So the total number of subspaces will be

$$\frac{(q^n - 1)(q^n - q^2)(q^n - q^3)\cdots(q^n - q^{n-m+1})}{(q^m - 1)(q^m - q^2)(q^m - q^3)\cdots(q^n - q^m)}.$$

This is the same as the Guassian binomial coefficient

$$\binom{n}{m}_q,$$

that turns out to be symmetric in the sense

$$\binom{n}{m}_q = \binom{n}{n-m}_q$$

■ **Problem 3.7.8**

Solution (a) First, note the Lemma below.

Lemma 3.1 Let $S \subset V$ a subset of V. Then we have

$$(\operatorname{Span} S)^0 = S^0.$$

In particular $\dim S^0 = \dim(\operatorname{Span} S)^0$.

Proof. Let $f \in (\operatorname{Span} S)^0$. Then $f(x) = 0$ for all $x \in \operatorname{Span} S$, in particular $x \in S$. So $f \in S^0$. For the converse, let $f \in S^0$, and let $x \in \operatorname{Span} S$. We can write $x = \sum_i \alpha_i s_i$ for $s_i \in S$. Then it implies that

$$f(x) = f(\sum \alpha_i s_i) = \sum_i \alpha_i f(s_i) = 0.$$

So $f \in (\operatorname{Span} S)^0$. □

Also, observe that $S \subset S^{00}$ (from theorem 2) §17, this implies that $\operatorname{Span} S \subset S^{00}$. To see this let $x \in \operatorname{Span} S$. From lemma above, $\forall y \in S^0$ we have $y(x) = 0$. Furthermore, from the natural correspondence, it follows that for all $f \in S^0$ $z(f) = 0$ for $x \mapsto z$ under natural correspondence. So $z \in S^{00}$. This proves that $\operatorname{Span} S \subset S^{00}$. Now we do a similar dimension argument as in Theorem 2 §17. If $\operatorname{Span} S$ is m dimensional, then S^0 is $n - m$ dimensional, and the dimension of S^{00} is $n - (n - m) = m$. So $\operatorname{Span} S = S^{00}$.

(b) Let $f \in J^0$. Then it implies that for all $x \in J$ we have $f(x) = 0$. This holds true in particular for $x \in S$ as $S \subset J$.

(c) First we will prove $(M+N)^0 = M^0 \cap N^0$. Let $f \in (M+N)^0$. Using the fact that $M \subset M+N$, and $N \subset M + N$, it follows that $f(x) = 0$ for all $x \in M$ and $x \in N$. Thus $f \in M^0 \cap N^0$. For the converse, let $f \in M^0 \cap N^0$. Let $x \in M+N$. Then $x = m+n$ for $m \in M$ and $n \in N$. Then it follows that $f(x) = f(m + n) = f(m) + f(n) = 0$ for all $x \in M + N$. So $f \in (M + N)^0$. To prove $(M \cap N)^0 = M^0 + N^0$, replace M, N with M^0, N^0 in the identity that we proved above, and using the fact that $M^{00} = M$ (true in finite dimension) and applying $(\cdot)^0$ to both sides, we will prove the second statement.

(d) TODO: Not sure about my answer for this section.

Theorem 3.2 — Summary of the problem above. Let $S \subset V$ be a subset of a finite dimensional vector space. Then

$$S^{00} = \operatorname{Span} S.$$

■ **Problem 3.7.9** Skipped. Will be added later.

3.8 Dual of a direct sum

■ **Problem 3.8.1**

Solution (a) As in the characterization in Roman, if $\mathbb{C}^4 = M \oplus N$, then 0 should have a unique representation. I.e. if $m \in M$ and $n \in N$, both of which non-zero, and $m + n = 0$, then the sum is not direct. First observe that the sets $\{x, y\}$ and $\{u, v\}$ are both linearly independent. Let $0 = a_1 x + a_2 y + a_3 u + a_4 v$. This defines

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

It is easy to see that the determinant of the matrix above is zero. So there is a non-zero vector (a_1, a_2, a_3, a_4) that satisfies the system above. With Gaussian elimination, it turns out that the solution space is $\{w(1, -1, -1, 1), w \in \mathbb{C}\}$.

(b) With a similar argument as above, since

$$\det \begin{pmatrix} -1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} = -2 \neq 0,$$

and using the fact that each $\{x, y\}$ and $\{u, v\}$ are linearly independent, we conclude that the sum is direct.

(c) We can see that

$$\det \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} = 0,$$

and since $\{x, y\}$ and $\{u, v\}$ are both linearly independent, it follows that the sum is not unique, and there is a non-trivial representation of the origin. The solution space is $\{w(-1, -1, 1, 1) | w \in \mathbb{C}\}$.

■ **Problem 3.8.2**

Solution Let $x \in M \cap N$. Then because $x \in M$ it follows that $x_1 = \cdots = x_n = 0$. And because $x \in N$ if follows that $x_j = x_{n+j}$ for all $j = 1, \cdots, n$. This implies $x_1 = x_2 = \cdots = x_n = x_{n+1} = \cdots = x_{2n} = 0$. Thus $x = 0$. So $M \cap N = \{0\}$. So the sum is direct.

■ **Problem 3.8.3**

Solution Let $V = \mathbb{R}^2$, $M = \text{Span}\{e_1\}$, $N_1 = \text{Span}\{e_2\}$, and $N_2 = \text{Span}\{e_1 + e_2\}$.

■ **Problem 3.8.4**

Solution (a) We can do a universal property argument as we can find in Roman textbook. However, here we adopt a more straight forward approach. We claim that the spaces $U \oplus (V \oplus W)$ is isomorphic to $(U \oplus V) \oplus W$. Let $x \in U \oplus (V \oplus W)$. So we can write $x = x_U + x_{V \oplus W}$. And also $x_{V \oplus W} = x_V + x_W$. So $x = x_U + x_V + x_W$. Observe that $x_U + x_V \in U \oplus V$. Denote this as $x_{U \oplus V} = x_U \oplus x_V$. So $x = x_{U \oplus V} + x_W$. We claim that the mapping $\phi : U \oplus (V \oplus W) \to (U \oplus V) \oplus W$ defined by $x = x_U + x_{V \oplus W} \mapsto x_{U \oplus V} + x_W$ as in our discussion above. It is straightforward to check that this map linear, surjective, and injective. Thus an isomorphism.

(b) Similar to our discussion above, $U \oplus V$ and $V \oplus U$ are isomorphic.

Theorem 3.3 The external direct sum is associative and commutative in the sense of isomorphicity. I.e. Let U, V, W be vectors spaces, then $U \oplus (V \oplus W)$ is isomorphic to $(U \oplus V) \oplus W$. Also $U \oplus V$ is isomorphic to $V \oplus U$.

■ **Problem 3.8.5**

Solution (a) We want to prove if L, M, N are independent and $V = L+M+N$, then $L \oplus (M \oplus N)$ is direct. First observe that $M \cap (L+N) = \{0\}$ and $N \cap (L+M) = \{0\}$ implies that $M \cap N = \{0\}$, otherwise, it contradicts the assumptions. So $M \oplus N$ is direct. Furthermore, since $L \cap (M+N) = \{0\}$, replacing $M + N = M \oplus N$ we will have $L \cap (M \oplus N) = \{0\}$. This and the fact that $V = M + N + L$ together imply that the sum $L \oplus (M \oplus N)$ is direct. For the converse, we prove by contrapositive. I.e. we assume that V is not the join of the subspaces, or the subspaces are not independent, and conclude that the sum is not direct. If V is not the join of the subspaces, then the sum $L \oplus (M \oplus N)$ can not be direct. Also since if $M \cap N$ is not trivial implies that $M \oplus N$ is not direct, hence $L \oplus (M \oplus N)$ is not direct, in the following cases we will assume that $M \cap N$ is trivial. If $L \cap (M+N)$ is non-trivial, (assuming $M \cap N = \{0\}$) it implies that $L \cap (M \oplus N) \neq \{0\}$ so the sum $L \oplus (M \oplus N)$ is not direct. Also, if $M \cap (L+N)$ is non-trivial, (assuming $M \cap N = \{0\}$) it follows that $M \cap L \neq \{0\}$, thus $L \cap (M \oplus N) \neq \{0\}$, hence the sum $L \oplus (M \oplus N)$ is not direct. The reasoning for the case $N \cap (L+M) \neq \{0\}$ is also similar to above.

(b) Let $V = \mathbb{R}^2$, $L = \text{Span}\{e_1\}$, $M = \text{Span}\{e_2\}$, and $N = \text{Span}\{e_1 + e_2\}$.

(c) First we prove L, M, N being independent implies x, y, z are linearly independent. We prove by contrapositive. Assume $\{x, y, z\}$ be linearly dependent. If $\{x, y\}$ is linearly dependent, then $M \oplus N$ is not direct, hence $L \oplus (M \oplus N)$ is not direct. If $\{x, y\}$ is linearly independent but $\{x, y, z\}$ is linearly dependent, then $z = \alpha x + \beta y$ for some α, β in the field. This implies $L \cap (M \oplus N) \neq \{0\}$. So $L \oplus (M \oplus N)$ is not direct (note that we used the fact that forming a direct sum is equivalent to being independent, as we proved in part (a))

Now we prove that x, y, z being linearly independent implies $L \oplus (M \oplus N)$ is direct. Since y, z is linearly independent, $M \cap N = \{0\}$. This implies that $M \oplus N$ is direct. Since x, y, z is linearly independent, $L \cap (M \oplus N) = \{0\}$, thus $L \oplus (M \oplus N)$ is direct.

(d) We use the principle of inclusion-exclusion generalized for three subspaces. We have

$$\dim(L+M+N) = \dim(L)+\dim(M)+\dim(N)-\dim(M\cap N)-\dim(M\cap L)-\dim(N\cap L)+\dim(M\cap N\cap L).$$

So if the dimension of the sum of the subspaces is sum of their dimension, it follows that

$$\dim(L \cap M) = \dim(L \cap N) = \dim(M \cap N) = \dim(M \cap N \cap L) = 0.$$

This implies that L, M, N are independent. So $L \oplus (M \oplus N)$ is direct. It is also easy to have a similar argument for the converse.

(e)

3.9 Dimension of Quotient Space

We start with a different proof of the Theorem 1 in §22.

Theorem 3.4 If M and N are complementary subspaces of a vector space V, then the correspondence that assigns to each vector $y \in N$ the vector $y + M \in V/M$ is an isomorphism.

Proof. Consider the mapping $N \ni y \mapsto y + M \in V/M$. We claim that this mapping is an isomorphism. To see this first observe that the mapping is linear. Because

$$(\alpha y_1 + y_2) \mapsto (\alpha y_1 + y_2 + M = (\alpha y_1 + M) + (y_2 + M)).$$

Now we check to see if the mapping is injective. Let $y \in N$ that maps to the zero vector in V/M by the mapping above. So $y \mapsto 0 + M$. This implies that $y \in M$. Since $M \cap N = \{0\}$, we conclude that $y = 0$. Furthermore, we want to show that this mapping is surjective. To see this let $y + M$ be a coset in V/M. Since $V = M \oplus N$, then we can writ $y = y_m + y_n$ where $y_m \in M$ and $y_n \in N$. We claim that $y_n \mapsto y + M$. That is because $y + M = y_n + y_m + M = y_n + M$. So this mapping is surjective. □

■ **Problem 3.9.1**

Solution (a) If $M = P_n$, then P/M is not finite dimensional. Because $x \in P$ can be written as

$$x = P_n + \alpha_1 t^{n+1} + \alpha_2 t^{n+2} + \cdots \in P_n + W,$$

and W is infinite dimensional.

(b) If M is the set of all even polynomials, then P/M is not finite dimensional. Because $x \in O$ can be written as

$$x = M + \alpha_0 + \alpha_1 t^3 + \alpha_2 t^5 + \cdots \in M + Y,$$

where Y is infinite dimensional.

(c) If M is the set of all polynomials that are divisible by t^n, the P/M is finite dimensional. We can write M as

$$M = \mathrm{Span}(\bigcup_{i=0}^{\infty} t^n P_i),$$

where $t^n P_i$ is the set of all polynomials of the form $t^n x$, where $x \in P_i$. So elements od P/M will be of the form

$$x = \alpha_0 + \alpha_1 t + \cdots + \alpha_{n-1} t^{n-1} + M \in P_n + M,$$

■ **Problem 3.9.2**

Solution Both empty set and the set $\{0\}$ acts as the origin of the vector space. Any set that contains more than one element will not have an additive inverse. This is not an exhaustive list of possible problems.

■ **Problem 3.9.3**

Solution (a) The relation is reflexive because $x - x = 0 \in M$, as M as a subspace should contain the origin. The relation is symmetric because if $x - y \in M$, its additive inverse will also be in M, and we have $M \ni -(x - y) = y - x$. The relation if transitive, because $x \sim y$ implies $x - y \in M$, and $y \sim z$ implies $y - z \in M$. So we can write

$$M \ni x - y = x - y \pm z = (x - z) + (z - y),$$

Using the fact that $y - z \in M$ it follows that $x - z \in M$. Thus $x \sim z$.

(b) Observe that $x_1 - y_1 \in M$, and $x_2 - y_2 \in M$. Since M is a subspace, it follows that $\alpha(x_1 - y_1) + \beta(x_2 - y_2) \in M$. Rearranging the terms we will have

$$(\alpha x_1 + \beta x_2) - (\alpha y_1 + \beta y_2) \in M.$$

So it follows that

$$\alpha x_1 + \beta x_2 \sim \alpha y_1 + \beta y_2.$$

(c) Let $A \subset V$ be an equivalence class. Pick $y \in A$. We claim that $A = y + M$, i.e. it is a coset of M. To see this let $x \in A$. Then we can write $x = (x - y) + y$. Since M is an equivalence class, it follows that $x - y \in M$. So $x \in y + M$. So $A \subset y + M$. Now let $x \in y + M$. Then $x - y \in M$. Thus $y + M \subset A$. For the reverse inclusion, we assume that $A = y + M$ for some $y \in V$. Let $x_1, x_2 \in y + M$. So $\exists m_1, m_2 \in M$ such that $x_1 = y + m_1$ and $x_2 = y + m_2$. Thus $x_1 - x_2 = m_1 - m_2 \in M$. So A is an equivalence class.

3.10 §33 Transformation as vectors

■ **Problem 3.10.1**

Solution (a) Linear map. This follows from the property of complex conjugation, $\overline{(\alpha z_1 + \beta z_0)} = \alpha \overline{z}_1 + \beta \overline{z}_0$.

Observation 3.10.1 Notice that the complex conjugation is a linear operator on $\mathbb{C}$ only when considered on the field $\mathbb{R}$. Otherwise, (when we consider the vector space over the field $\mathbb{C}$) then the complex conjugation is no longer a linear map. There are several ways to see this. First, we can see by letting $\{1\}$ be a basis for the space $(\mathbb{C}, \mathbb{C})$, and observe that the complex conjugation map maps $1 \mapsto 1$. So it should map every element to itself, which is obviously not correct. A more straight forward way to show this is to observe that if $\mathbb{C}$ is defined over the field $\mathbb{C}$, then the scalars are complex variables, thus

$$\overline{(\alpha z_1 + \beta z_0)} = \overline{\alpha}\ \overline{z}_1 + \overline{\beta}\overline{z}_0,$$

which is not the desired value $\alpha \overline{z}_1 + \beta \overline{z}_0$.

Linear map. let $x, x' \in \mathcal{P}$ and α be any scalar. Then

$$\begin{aligned}(A(\alpha x + x'))(t) &= (\alpha x + x')(t+1) - (\alpha x + x')(t) \\ &= (\alpha x(t+1) - x'(t)) + (x'(t+1) - x'(t)) = \alpha(Ax)(t) + (Ax')(t).\end{aligned}$$

(b) Yes. Since A permutes the indices but preserves the scalar coefficients.

(c) Yes.

(d) For the question above, as well as this question, I have drafted a more complete solution. I feel it is correct, but tat the same time I feel that it is not correct! However, I am writing it here for my future reference in order to improve it. Let w be a k-linear map. Then we can write it as

$$w = \sum_{i_1, \cdots, i_k = 1}^{n} \alpha_{i_1, \cdots, i_k} e^{i_1} \otimes \cdots \otimes e^{i_k}.$$

Note that the sum above is the same as running the sum over all the functions $f : \{1, \cdots, n\} \to \{1, \cdots, n\}$. Denote this set as $[n]^{[n]}$. So we can write the sum above as

$$w = \sum_{f \in [n]^{[n]}} \alpha_{f(1), \cdots, f(k)} e^{f(1)} \otimes \cdots \otimes e^{f(k)}.$$

Now in a similar fashion for Aw given as in the problem statement we have

$$w = \sum_{i_{\pi(1)},\cdots,i_{\pi(k)}=1}^{n} \alpha_{i_{\pi(1)},\cdots,i_{\pi(k)}} e^{i_{\pi(1)}} \otimes \cdots \otimes e^{i_{\pi(k)}}.$$

TODO: I don't know how to continue from here.w

(e) TODO

(f) TODO

■ **Problem 3.10.2**

Solution There are different ways to approach this question. But I will adopt the following approach. Let $A \in \mathrm{End}_F(V)$, and let $\{e_i\}$ be a basis for the space V. Then we can write

$$Ae_i = \sum \alpha_{ji} e_j.$$

The coefficients α_{ji} for $j = 1, \cdots, n$ and $i = 1, \cdots, n$ determine A. Because for any $x \in V$ we can write $x = \sum_i \beta_i e_i$ and Ax can we expressed by the coefficients α_{ji}. Because

$$A(x) = A(\sum_i \beta_i e_i) = \sum_i \beta_i A(e_i) = \sum_i \beta_i \sum_{i,j} \alpha_{j,i} e_j.$$

So it is easy to see that the dimension of the space is n^2.

■ **Problem 3.10.3**

Solution (a) This is the dual space of U, which is a vector space.

(b) Let $U = V \oplus W$. Let $U \ni u = v + w$ where $v \in V$ and $W \in w$. By the definition of A we have $Au = v$. To show that this mapping is linear, let $u' = v' + w'$ be any other vector in U, where $v' \in V$ and $w' \in W$. Let α be any scalar. Then $\alpha u + u' = (\alpha v + v') + (\alpha w + w')$. By the definition of A we have

$$A(\alpha u + u') = \alpha v + v' = \alpha A(u) + A(v').$$

So A is a linear map from U to V.

(c) Let $V = U/W$ where W is some subspace. The the map A is in fact the quotient map. There are many different ways to show that this map is in fact a linear map (for instance using the universal property, etc). Or we can do manually. I will invoke the universal property to prove the linearity of the universal property.

4. Random Notes

4.1 Interesting Observations from Roman

Observation 4.1.1 — Geometric Interpretation of Dual Vectors. The notion of the dual space of a vector space is somewhat abstract and one usually struggles to have a geometric realization of the functionals and dual spaces. Here, I provide a very interesting point of view. Let V be a finite dimensional vectors space. Then every $f \in V^*$ is characterized by a hyperplane H such that $H = \ker f$.
With this point of view, $f(x) = 0$ corresponds to the fact that $x \in H$. Also, it is very straight forward to see the following properties of functionals with this geometric point of view.

Proposition 4.1

(a) If $f(x) \neq 0$ then

$$V = \langle x \rangle \oplus \ker f.$$

(b) For every $x \in V$ there exists $f \in V^*$ such that $f(x) \neq 0$.

(c) For $x \in V$, $f(x) = 0$ for all $f \in V^*$ implies $x = 0$.

(d)

Proof. (Geometric interpretation)

(a) If $x \notin H$ for some hyperplane H, then

$$V = \langle x \rangle \oplus H.$$

(b) Given any point of the space, there is some hyperplane that misses that particular point.s

(c) The only point that belongs to all hyperplanes is the origin.

□

Observation 4.1.2 — More Geometric Interpretation of Dual Vectors. The characterization above, i.e. identifying the linear functionals with their kernel, i.e. hyperplanes, work surprisingly well in characterizing very interesting facts. For instance, we can have the following definition of the annihilators of a set.

Definition 4.1 — Annihilators. Let $M \subset V$ (no necessarily a linear subspace). Then the annihilators of M, denoted by M^0 is the set of all linear functionals that kills M. I.e.

$$M^0 = \{f \in V^* | f(M) = 0\}.$$

With the geometric point of view above, the annihilators of M is the set of all hyperplanes that contain M.

For instance, let L be a one dimensional linear subspace of $\mathbb{R}^3$. Then L^0 will be the set of all hyperplanes containing L. Each such hyperplane can be represented by a normal vectors. So the set of all hyperplanes containing L is isomorphic to a plane perpendicular to L and going through the origin (more generally, any 2-dimensional linear subspace of $\mathbb{R}^3$ that does not contain L). It is now very straightforward to see the result of Theorem 3.14 part (2). The set M^{00} is the set of all hyperplanes containing M^0. There is just one such hyperplane, and since it can be parameterized using one normal vector (along L), we have

$$M^{00} \simeq \operatorname{span} L.$$

Observation 4.1.3 — Double Dual Map. We start with the following definition.

Definition 4.2 Let $\tau \in \mathcal{L}(U, V)$. The dual map $\tau^\times \in \mathcal{L}(V^*, U^*)$ and, the double dual map $\tau^{\times\times} = \mathcal{L}(U^{**}, V^{**})$ is defined as

$$(\tau^\times f)(u) = f(\tau u), \qquad \text{for } u \in U,\ f \in V^*,$$

and

$$(\tau^{\times\times} E)(f) = E(\tau^\times f), \qquad \text{for } E \in V^{**},\ f \in W^*.$$

In finite dimension, the following is a very useful characterization of $\tau^{\times\times}$. Let $u \in U$ and using the canonical map $u \mapsto E_u \in V^{**}$, where E_u is the evaluation map at u.Also let $f \in V^*$. Then we can write

$$\begin{aligned}(\tau^{\times\times} E_u)(f) &= E_u(\tau^\times f)\\ &= (\tau^\times f)(u)\\ &= f(\tau u)\\ &= E_{\tau u}(f).\end{aligned}$$

Thus we have

$$\tau^{\times\times} E_u = E_{\tau u}.$$

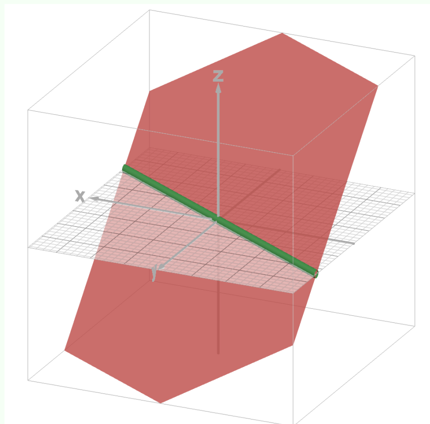
Observation 4.1.4 — Geometric Interpretation of Dual Map. For $\tau \in \mathcal{L}(V, W)$, the dual map $\tau^\times \in \mathcal{L}(W^*, V^*)$ is given by

$$(\tau^\times f)(v) = f(\tau v),$$

where $f \in W^*$ and $v \in V$. Using out geometric point of view of the functionals (as hyperplanes) we can have a geometric interpretation of what is the dual of a map. The following is a high level summary:

> Let $f \in W^*$ be a functional, i.e. a hyperplane. Then $\tau^\times$ returns a hyperplane in V that is the pre-image of restriction of f to $\mathrm{im}(\tau)$.

For instance, if $\tau : \mathbb{R}^2 \to \mathbb{R}^3$ the inclusion map that sends $\mathbb{R}^2$ to the xy plane in $\mathbb{R}^3$, the $\tau^\times$ map maps the following red hyperplane (as a functional in $\mathbb{R}^3$) to the green hyperplane (as a functional in $\mathbb{R}^2$).



Using the interpretation above we can have the following "geometric" proof of the following facts in Roman (presented in Theorem 3.19).

Proposition 4.2 Let $\tau \in \mathcal{L}(V, W)$. Then

(a) $\ker(\tau^\times) = \mathrm{im}(\tau)^0$.

(b) $\mathrm{im}(\tau^\times) = \ker(\tau)^0$.

Geometric proof. (a) We want to show $\ker(\tau^\times) \subset \mathrm{im}(\tau)^0$. Let $f \in \ker(\tau^\times)$ be a hyperplane (i.e. functional). This means that if we restrict f to $\mathrm{im}(\tau)$ and then consider its pre-image, it should be the whole space (i.e. the zero functional). Thus f should contain $\mathrm{im}(\tau)$. So $f \in \mathrm{im}(\tau)^0$ (remember that $\mathrm{im}(\tau)^0$) is the set of all hyperplanes containing $\mathrm{im}(\tau)$. For the converse, we want to show $\mathrm{im}(\tau)^0 \subset \ker(\tau^\times)$. Let $f \in \mathrm{im}(\tau)^0$. I.e. f is a hyperplane that contains $\mathrm{im}(\tau)$. So restricting f to $\mathrm{im}(\tau)$ will be whole $\mathrm{im}(\tau)$. So the pre-image of the restriction of f to $\mathrm{im}(\tau)$ will be the whole space V (thus the zero functional). So $f \in \ker(\tau^\times)$. *Note: We have used the fact that for any linear map τ we have $\mathrm{im}(\tau) \simeq \mathrm{dom}(\tau)$.*

(b)

□

Observation 4.1.5 — Coordinate maps. Let (V, F) be a vector space (defined on the field F) with finite dimension n. Once we choose an ordered basis for V, like $\mathcal{B} = (v_1, \cdots, v_n)$, we can define the coordinate map

$$\phi_{\mathcal{B}} : V \to F^n,$$

that

$$v = \sum_i \alpha_i v_n \mapsto \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

In particular, for the basis vectors we have $\phi(v_i) = e_i$, where e_i is a column vector whose entries are all zero, but the i^{th} row. This coordinate map ϕ justifies the name "vector space" for this algebraic structure. The elements of any finite dimensional vector space defined on F can be "coordinated" by the elements of F^n.

Observation 4.1.6 As a continuiation of the note above, lets now focus on the linear maps $\mathcal{L}(F^n, F^m)$. We know that every matrix in $A \in \mathcal{M}_{n,m}$ induces a linear map $\tau_A \in \mathcal{L}(F^n, F^m)$, given by

$$\tau_A(v) = Av.$$

The converse is also true. Every linear map $\tau \in \mathcal{L}(F^n, F^m)$ has a matrix representation $A \in \mathcal{M}_{n,m}$ given by

$$A = (\tau e_1 | \cdots | \tau e_n),$$

i.e. apply τ on the basis vectors, write the coordinates of the resulting vector in the columns of a matrix to get the matrix representation of the linear transformation.

Observation 4.1.7 I have started to notice a very interesting interaction between the following objects, and each pair of these notions induces a similar feeling. I have not yet been able to quantify this feeling. But I am sure there is some connection there.

surjective	ker	spanning	exists
injective	img	linearly independent	for all

Observation 4.1.8 — Change of basis. Consider the following diagram.

$$\begin{tikzcd}
F^n \arrow[ddd, bend right, "{[\tau]_{\tilde{\mathcal{B}}\tilde{\mathcal{C}}}}"'] & & F^n \arrow[ll, "M_{\mathcal{B}\tilde{\mathcal{B}}}=\phi_{\mathcal{B}\tilde{\mathcal{B}}}"'] \arrow[ddd, bend left, "{[\tau]_{\mathcal{B}\mathcal{C}}}"] \\
& V \arrow[ul, "\phi_{\tilde{\mathcal{B}}}"] \arrow[ur, "\phi_{\mathcal{B}}"'] \arrow[d, "\tau"] & \\
& W \arrow[dl, "\phi_{\tilde{\mathcal{C}}}"'] \arrow[dr, "\phi_{\mathcal{C}}"] & \\
F^m & & F^m \arrow[ll, "M_{\mathcal{C}\tilde{\mathcal{C}}}=\phi_{\mathcal{C}\tilde{\mathcal{C}}}"]
\end{tikzcd}$$

where V, W are vector spaces with dimension n and m respectively. Furthermore, $\mathcal{B}$ and $\tilde{\mathcal{B}}$ are two ordered basis for V with the coordinate maps $\phi_{\mathcal{B}}$ and $\phi_{\tilde{\mathcal{B}}}$ respectively. Similarly $\mathcal{C}$ and $\tilde{\mathcal{C}}$ are two ordered basis for W with $\phi_{\mathcal{C}}$ and $\phi_{\tilde{\mathcal{C}}}$ as the corresponding coordinate maps. The change of basis matrices are given in Theorem 2.12 Roman. This diagram summarizes the relation between the vectors in the abstract vector spaces V and W with their representation with we change the basis in either of the spaces. Also it capture the transformation that happens for the representation of linear maps when we change basis. For instance it is very easy to see

$$[\tau]_{\tilde{\mathcal{B}}\tilde{\mathcal{C}}} = M_{\mathcal{C}\tilde{\mathcal{C}}}[\tau]_{BC} M^{-1}_{\mathcal{B}\tilde{\mathcal{B}}}.$$

It is also easier to memorize the following relation instead

$$[\tau]_{\tilde{\mathcal{B}}\tilde{\mathcal{C}}} M_{\mathcal{B}\tilde{\mathcal{B}}} = M_{\mathcal{C}\tilde{\mathcal{C}}}[\tau]_{\mathcal{B}\mathcal{C}}.$$

Observation 4.1.9 — Characteristic of a filed and the alternating, v.s. skew-symmetric forms. It is only when $\mathrm{Char}(F) \neq 2$ we have the equivalence between the alternating forms and the skew-symmetric forms

$$\text{alternating} \qquad \Longleftrightarrow \qquad \text{skew-symmetric.}$$

For the forward direction we don't need any restriction on the characteristic of the field. To see this let f be an alternating bi-linear form.

$$f(u+v, u+v) = f(u,u) + f(v,v) + f(u,v) + f(v,u).$$

Since f is alternating, we have $f(u+v, u+v) = 0$ as well as $f(u,u) = f(v,v) = 0$. So we can conclude that

$$f(u,v) = -f(v,u).$$

However, for the converse, we need $\mathrm{Char}(F)$ $neq 2$. Because if f is skew-symmetric, then $f(u,u) = -f(u,u)$, which implies $2f(u,u) = 0$. We can only conclude $f(u,u) = 0$ when $\mathrm{Char}(F) \neq 2$, i.e. when 2 is invertible in the field.

Observation 4.1.10 — Symmetric and Antisymmetric tensor products with Roman's notation. In Roman text book, he introduces the notion of the symmetric and anti-symmetric tensor products with a notation that is not easy to understand, unless there is a running example. Here in this box, I will give an explicit example. Let V has dimension $n = 3$ and let $\{e_1, e_2, e_3\}$ be a basis for V. We want to explicitly construct the basis vectors for the $\mathrm{ST}^p(V)$ and $\mathrm{AT}^p(V)$. We will have the following cases

(i) $p = 2$. Then the basis elements of $\mathrm{ST}^2(V)$ will be

M	$\sum_{t \in G_M} \mathrm{t}$	equiv in $F_2[e_1, e_2, e_3]$
$\{1,1\}$	$e_1 \otimes e_1$	$e_1 \vee e_1$
$\{2,2\}$	$e_2 \otimes e_2$	$e_2 \vee e_2$
$\{3,3\}$	$e_3 \otimes e_3$	$e_3 \vee e_3$
$\{1,2\}$	$e_1 \otimes e_2 + e_2 \otimes e_1$	$e_1 \vee e_2$
$\{1,3\}$	$e_1 \otimes e_3 + e_3 \otimes e_1$	$e_1 \vee e_3$
$\{2,3\}$	$e_1 \otimes e_2 + e_2 \otimes e_1$	$e_1 \vee e_2$

And the basis elements of $\mathrm{AT}^2(V)$ will be

M	$\sum_{t\in G_M} \mathrm{t}$	equiv in $F_2[e_1,e_2,e_3]$
$\{1,2\}$	$e_1\otimes e_2 - e_2\otimes e_1$	$e_1\wedge e_2$
$\{1,3\}$	$e_1\otimes e_3 - e_3\otimes e_1$	$e_1\wedge e_3$
$\{2,3\}$	$e_2\otimes e_3 - e_3\otimes e_2$	$e_2\wedge e_3$

(ii) $p=3$. Then the basis elements of $\mathrm{ST}^3(V)$ will be

M	$\sum_{t\in G_M} \mathrm{t}$	equiv in $F_2[e_1,e_2,e_3]$
$\{1,1,1\}$	$e_1\otimes e_1\otimes e_1$	$e_1\vee e_1\vee e_1$
$\{2,2,2\}$	$e_2\otimes e_2\otimes e_2$	$e_2\vee e_2\vee e_2$
$\{3,3,3\}$	$e_3\otimes e_3\otimes e_3$	$e_3\vee e_3\vee e_3$
$\{1,2,2\}$	$e_1\otimes e_2\otimes e_2+e_2\otimes e_1\otimes e_2+e_2\otimes e_2\otimes e_1$	$e_1\vee e_2\vee e_2$
$\{1,3,3\}$	$e_1\otimes e_3\otimes e_3+e_3\otimes e_1\otimes e_3+e_3\otimes e_3\otimes e_1$	$e_1\vee e_3\vee e_3$
$\{2,1,1\}$	$e_2\otimes e_1\otimes e_1+e_1\otimes e_2\otimes e_1+e_2\otimes e_2\otimes e_1$	$e_1\vee e_1\vee e_2$
$\{2,3,3\}$	$e_2\otimes e_3\otimes e_3+e_3\otimes e_2\otimes e_3+e_3\otimes e_3\otimes e_2$	$e_2\vee e_3\vee e_3$
$\{3,1,1\}$	$e_3\otimes e_1\otimes e_1+e_1\otimes e_3\otimes e_1+e_1\otimes e_1\otimes e_3$	$e_1\vee e_1\vee e_3$
$\{3,2,2\}$	$e_3\otimes e_2\otimes e_2+e_2\otimes e_3\otimes e_2+e_2\otimes e_2\otimes e_3$	$e_2\vee e_2\vee e_3$
$\{1,2,3\}$	$e_1\otimes e_2\otimes e_3+e_1\otimes e_3\otimes e_2+e_2\otimes e_1\otimes e_3$ $+e_2\otimes e_3\otimes e_1+e_3\otimes e_1\otimes e_2+e_3\otimes e_2\otimes e_1$	$e_1\vee e_2\vee e_3$

And for $\mathrm{AT}^3(V)$ we have

M	$\sum_{t\in G_M} \mathrm{t}$	equiv in $F_2[e_1,e_2,e_3]$
$\{1,2,3\}$	$e_1\otimes e_2\otimes e_3-e_1\otimes e_3\otimes e_2+e_2\otimes e_1\otimes e_3$ $-e_2\otimes e_3\otimes e_1+e_3\otimes e_1\otimes e_2-e_3\otimes e_2\otimes e_1$	$e_1\wedge e_2\wedge e_3$

This we have following two theorems.

Theorem 4.1 Let V be a finite dimensional vector space, and let $\mathcal{B}=\{e_1,\cdots,e_n\}$ be a basis. Then

$$\mathrm{ST}^p(V)\simeq F_p[e_1,\cdots,e_n],$$

and

$$\mathrm{AT}^p(V)\simeq F_p^-[e_1,\cdots,e_n].$$

In words,

> The symmetric tensor space $\mathrm{ST}^p(V)$ is isomorphic to the algebra $F_p[e_1,\cdots,e_n]$ of homogeneous polynomials of degree p.

And similarly

> The anti-symmetric tensor space $\mathrm{AT}^p(V)$ is isomorphic to the algebra $F_p^-[e_1,\cdots,e_n]$ of anti-commutative homogeneous polynomials of degree p.

It is easy to see (Roman Theorem 14.18) that

$$\dim(\mathrm{AT}^p(V))=\binom{n}{p},\qquad \dim(\mathrm{ST}^p(V))=\binom{n+p-1}{p}.$$

The formula for the dimension of $\mathrm{ST}^p(V)$ resembles the formula for all possible distribution of p units on energy in n containers (see Schroeder, Equation 2.9).

Observation 4.1.11 — Suitable basis for nilpotent maps. This note is meant to accompany the sections 8.4 to 8.8 of Leohen's advanced linear algebra. I demonstrate a concrete example to show how we can find a suitable basis for a nilpotent map such that its matrix representation has Jordan blocks. Consider the following nilpotent matrix.

$$A = \begin{pmatrix} -16 & -20 & 34 \\ 2 & 52 & -29 \\ 0 & 72 & -36 \end{pmatrix}.$$

It is easy tot check that $A^3 = 0$. We want to find a suitable basis in which the matrix above has Jordan blocks in its structure. To do so, we first consider image A. By writing the matrix in reduced row echelon form it is easy to see that the column space of A is

$$W = \mathrm{im}(A) = \mathrm{Span}\{ \begin{bmatrix} -16 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -20 \\ 52 \\ 72 \end{bmatrix} \}.$$

The subspace W is A-invariant. So we can restrict A to this subspace. Fix the vectors above as the basis of W. In this basis, let the matrix representation of $A|_W$ is $[A|_W]$. Then this needs to satisfy

$$AM = M[A|_W],$$

where

$$M = \begin{pmatrix} -16 & 20 \\ 2 & 52 \\ 0 & 72 \end{pmatrix}.$$

One possible solution for B will be

$$[A|_W] = \begin{pmatrix} -16 & 128 \\ 2 & 16 \end{pmatrix}.$$

Again, we need to consider the image of this matrix. Writing the matrix $A|_W$ in reduced row echelon form we can see that

$$W' = \mathrm{im}(A|_W) = \mathrm{Span}\{ \begin{bmatrix} -16 \\ 2 \end{bmatrix} \}.$$

Now we need to restrict $A|_W$ to W' above. However since W' is one dimensional and $A|_W$ in nilpotent (since A is nilpotent), $(A|_W)|_{W'}$ will be the zero map. So we can choose

$$e_1 = \begin{bmatrix} -16 \\ 2 \end{bmatrix},$$

as the first vector in the desired basis. Since $\ker(A|_W)$ is the same as $\mathrm{im}(A|_W)$, then $\mathrm{im}(A|_W)+\ker(A|_W) = \mathrm{im}(A|_W)$ and the basis can not be extended further (stage to of finding suitable basis for the nilpotent map $A|_W$). We need to perform stage 3 and find a vector in W that maps to W'. One possible choice is

$$e_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Note that in both of the representations above we are assuming that the basis vectors for the space W is as fixed above. Lastly, we need to perform stage 2 for the map A, and since $\ker(A) \subset \mathrm{im}(A)$, this has no interesting result. Finally, for stage 3 for A, we need to find some $e_3 \in V$ such that $Ae_3 = e_2$. By using the pseudo inverse for A we can find

$$e_3 = \frac{1}{14} \begin{bmatrix} 5 \\ -3 \\ -6 \end{bmatrix}.$$

So the suitable basis is

$$e_1 = -16 \begin{bmatrix} -6 \\ 2 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 20 \\ 52 \\ 72 \end{bmatrix}, \quad e_2 = \begin{bmatrix} -16 \\ 2 \\ 0 \end{bmatrix}, \quad e_3 = \frac{1}{14} \begin{bmatrix} 5 \\ -3 \\ -6 \end{bmatrix}.$$

It is easy to see

$$Ae_1 = 0, \quad A^2 e_2 = 0, \quad A^3 e_3 = 0.$$

You can experiment with the code written in python to perform the computations above, where the code is included in the jupyter notebook file in the Codes folder of this latex project.

```
from sympy import Matrix
import sympy as sp

# Define the matrix

# Perform row reduction (RREF - Reduced Row Echelon Form)
# rref_matrix, pivot_columns = A_sym.rref()


def imageBasis(matrix):
# List of basis vectors
colSpace = matrix.columnspace()
# Stack the first two vectors into a matrix
return Matrix.hstack(*colSpace[:])



def kernelBasis(matrix):
"""Computes a basis for the null space (kernel) of a matrix."""
nullSpace = matrix.nullspace()
# Returns a list of basis vectors for the null space
return Matrix.hstack(*nullSpace[:])


def restrictTo(A,M):
```

```
        """
        Note that colSpace(B) should be image of A (thus A invariant)
        """

        return (M.T * M).pinv() * (M.T * A * M)
```

Observation 4.1.12 — Step by step walk through for calculating the generalized eigenspaces. Consider the following matrix

$$A = \begin{bmatrix} 7 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 7 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -4 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -4 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & i & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & i & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & i \end{bmatrix}.$$

We want to construct the generalized eigenspaces for this matrix. First, observe that the spectrum of this matrix is $\text{Spect}_A = \{7, -4, i\}$. This can be calculated by

```
A.eigenvals()
```

From the form of the matrix, it is immediate that for $\lambda = 7$, there are 3 associated generalized eigenspaces. To find these three, we first calculate the eigenvalues associated to $\lambda = 7$. We will have

$$\ker(A - 7I) = \text{Span}\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \}$$

These computations can be done by

```
I = sp.eye(J.shape[0])
v1,v2,v3 = ((A - (7)*I)).nullspace()
```

Now for each of the vectors above, we calculate their Jordan chain. I.e. we solve

$$(A - \lambda I)w = v,$$

where we start by v as one of the vector above to get w (second vector in the chain), and again, we replace v with w calculated above and solve the equation again to find a new w, which will be the third element in the chain, and so on. This process will stop eventually (the solution of the equation will be the zero vector). Starting with the each of the vectors above, we will get

$$V_1^1 = \mathrm{Span}\{ \begin{bmatrix}1\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\end{bmatrix} \}, \qquad V_1^2 = \mathrm{Span}\{ \begin{bmatrix}0\\0\\1\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\end{bmatrix} \}, \qquad V_1^3 = \mathrm{Span}\{ \begin{bmatrix}0\\0\\0\\0\\1\\0\\0\\0\\0\\0\\0\\0\\0\\0\end{bmatrix} \}.$$

Now for $\lambda = -4$ do a similar process. Observe that for this eigenvalue we have only one Jordan block matrix. First, we calculate the eigenvectors, and starting from each of them we compute the generalized eigenvectors. It turns out that there is just one eigenvalue for $\lambda = -4$

$$\ker(A - (-4)I) = \mathrm{Span}\{ \begin{bmatrix}0\\0\\0\\0\\0\\1\\0\\0\\0\\0\\0\\0\\0\\0\end{bmatrix} \}.$$

Similar to above, we can calculate the Jordan chain for starting with this vector. We will see

$$V_2^1 = \mathrm{Span}\{\begin{bmatrix}0\\0\\0\\0\\0\\1\\0\\0\\0\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\0\\0\\1\\0\\0\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\0\\0\\0\\1\\0\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\0\\0\\0\\0\\1\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\0\\0\\0\\0\\0\\1\\0\\0\\0\\0\end{bmatrix}\}.$$

Finally, for $\lambda = i$, it turns out (as also evident from the matrix), there are two eigenvalues (translates to two Jordan blocks) as

$$\ker(A - iI) = \mathrm{Span}\{\begin{bmatrix}0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\1\end{bmatrix}\}.$$

And with the calculations similar to above we will get

$$V_3^1 = \text{Span}\{\begin{bmatrix}0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\1\\0\\0\end{bmatrix}\}, \qquad V_3^2 = \text{Span}\{\begin{bmatrix}0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\1\end{bmatrix}\}.$$

The calculation above can be found under Example 3 in the JordanForm.ipynb file inside the Codes directory.

Observation 4.1.13 — Step by step calculation of Jordan canonical form. See StepByStepJordanCanonicalForm.ipynb in the Codes directory of this latex project for a complete example of calculating the Jordan canonical form.

Observation 4.1.14 — Yet another chance to understand the universal property of tensor product. The universal property for the tensor product states that if U, V are two vector spaces, and W is any other vector space, where $f : U \times V \to W$ is a bilinear map, then there exists a universal pair $(\tau, U \otimes V)$ such that f can be factored through τ, i.e. $f = t \circ \tau$. In other words the following diagram commutes

$$\begin{tikzcd}
& U \times V \arrow[dl, "f"'] \arrow[d, "\tau"] \\
W & U \otimes V \arrow[l, "t"]
\end{tikzcd}$$

where $f = t \circ \tau$. Here is an intuitive construction for this concept. Let $V = \mathbb{R}^3$ and $U = \mathbb{R}^2$ that $V = \text{Span}\{e_1', e_2', e_3'\}$ and $U = \text{Span}\{e_1, e_2\}$. Let $f : U \times V \to W$ be a bilinear map where W is any vector space $\mathbb{R}^n$. For $(a, b) \in U \times V$ we can write

$$\begin{aligned}
f(a,b) &= f(a_1e_1 + a_2e_2, b_1e_1' + b_2e_2' + b_3e_3') \\
&= a_1b_1f(e_1,e_1') + a_1b_2f(e_1,e_2') + a_1b_3f(e_1,e_3') + a_2b_1f(e_2,e_1') + a_2b_2f(e_2,e_2') + a_2b_3f(e_2,e_3').
\end{aligned}$$

So f is fully specified by knowing its effect on (e_i, e_j') tuple.s Assume the map f sends

$$(e_1,e_1') \mapsto w_1, \quad (e_1,e_2') \mapsto w_2, \quad (e_1,e_3') \mapsto w_3, \quad (e_2,e_1') \mapsto w_4, \quad (e_2,e_2') \mapsto w_5, \quad (e_2,e_3') \mapsto w_6,$$

where $w_1, \cdots, w_6 \in W$. So the effect of f can be written as

$$\begin{pmatrix} | & | & \cdots & | \\ w_1 & w_2 & \cdots & w_6 \\ | & | & \cdots & | \end{pmatrix} \begin{bmatrix} a_1 \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \\ a_2 \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \end{bmatrix} = \begin{pmatrix} | & | & \cdots & | \\ w_1 & w_2 & \cdots & w_6 \\ | & | & \cdots & | \end{pmatrix} \begin{bmatrix} a_1 b_1 \\ a_1 b_2 \\ a_1 b_3 \\ a_2 b_1 \\ a_2 b_2 \\ a_2 b_3 \end{bmatrix}$$

Intuitively, we form a basis that is the Cartesian product of the basis of the initial spaces to be able to write the bilinear form as a matrix multiplication.

Observation 4.1.15 — Yet another chance to understand the universal property of direct sum. Let $U = \mathbb{R}^2 = \operatorname{Span}\{e_1, e_2\}$ and $V = \mathbb{R}^3 = \operatorname{Span}\{e_1', e_2', e_3'\}$. Let (f, g) be two linear maps $f : U \to W$ and $g : V \to W$ where $W = \mathbb{R}^n$ is any vector space. Let $a \in U$ and $b \in V$ and observe that any such pair of linear maps (f, g) can be characterized by action of f and g as

$$f(a) = f(a_1 e_1 + a_2 e_2) = a_1 f(e_1) + a_2 f(e_2),$$

and

$$g(b) = g(b_1 e_1' + b_2 e_2' + b_3 e_3') = b_1 g(e_1') + b_2 g(e_2') + b_3 (e_3').$$

Assume

$$f : \quad e_1 \mapsto w_1, \quad e_2 \mapsto w_2,$$

and

$$g : \quad e_1' \mapsto w_1', \quad e_2' \mapsto w_2', \quad e_3' \mapsto w_3'$$

where $w_1, w_2, w_1', w_2', w_3' \in W$. So this follows that we can write the action of f, g together as

$$\begin{pmatrix} \begin{pmatrix} | & | \\ w_1 & w_2 \\ | & | \end{pmatrix} & 0 \\ 0 & \begin{pmatrix} | & | & | \\ w_1' & w_2' & w_3' \\ | & | & | \end{pmatrix} \end{pmatrix} \begin{bmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

So in fact, we concatenate the basis vectors of the older spaces to get a new basis for the new space so that we can express any pair of linear maps (f, g) as a single linear map.

Observation 4.1.16 — How to quickly define a bilinear form?. Let $U = \mathbb{R}^2$ and $V = \mathbb{R}^3$ and assume we want to define a bilinear form. It is enough to specify

$$(e_1, e_1) \mapsto \alpha_1, \quad (e_1, e_2) \mapsto \alpha_2, \quad (e_1, e_3) \mapsto \alpha_3, \quad (e_2, e_1) \mapsto \alpha_4, \quad (e_2, e_2) \mapsto \alpha_5, \quad (e_2, e_3) \mapsto \alpha_6.$$

Let $u \in U$ and $v \in V$. Then we can write $f(u, v)$ as

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 & \alpha_6 \end{pmatrix} \begin{bmatrix} u_1 \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \\ u_2 \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \end{bmatrix}.$$

Observation 4.1.17 — A simple example for Theorem 14.6 Roman. In this observation box I will demonstrate some example for the representation of tensors. I will use the vector notation to represent the tensors (instead of the matrix notation) as it makes more sense to work with this notation in this example. Let $U = \mathbb{R}^2$ and $V = \mathbb{R}^3$, and we are using the convention that

$$u \otimes v = \begin{bmatrix} u_1 \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \\ u_2 \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} u_1 v_1 \\ u_1 v_2 \\ u_1 v_3 \\ u_2 v_1 \\ u_2 v_2 \\ u_2 v_3 \end{bmatrix}.$$

Consider the following tensor

$$z = r_{11} e_1 \otimes e_1 + r_{12} e_1 \otimes e_2 + r_{13} e_1 \otimes e_3 + r_{21} e_2 \otimes e_1 + r_{22} e_2 \otimes e_2 + r_{23} e_2 \otimes e_3.$$

One representation for this tensor is to write (representation (1) in theorem 14.6 Roman)

$$z = \begin{bmatrix} r_{11} \\ r_{12} \\ r_{13} \\ r_{21} \\ r_{22} \\ r_{23} \end{bmatrix}$$

Another possible way to write this is (representation (2) in theorem 14.6)

$$\begin{bmatrix} r_{11} \\ r_{12} \\ r_{13} \\ r_{21} \\ r_{22} \\ r_{23} \end{bmatrix} = \begin{bmatrix} r_{11} \\ r_{12} \\ r_{13} \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ r_{21} \\ r_{22} \\ r_{23} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} r_{11} \\ r_{12} \\ r_{13} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} r_{21} \\ r_{22} \\ r_{23} \end{bmatrix}$$

Also, we can write (representation (3) in theorem 14.6 Roman)

$$\begin{bmatrix} r_{11} \\ r_{12} \\ r_{13} \\ r_{21} \\ r_{22} \\ r_{23} \end{bmatrix} = \begin{bmatrix} r_{11} \\ 0 \\ 0 \\ r_{21} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ r_{12} \\ 0 \\ 0 \\ r_{22} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ r_{13} \\ 0 \\ 0 \\ r_{23} \end{bmatrix} = \begin{bmatrix} r_{11} \\ r_{21} \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} r_{12} \\ r_{22} \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} r_{13} \\ r_{23} \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Observation 4.1.18 — Theorem 14.5 Roman. Let $\{u_1 \cdots, u_n\} \subset U$ be linearly independent, and $\{v_1 \cdots, v_n\} \subset V$ be arbitrary vectors. Then

$$\sum_{i=1}^{n} u_i \otimes v_i = 0$$

implies that $v_i = 0$ for all i.

Proof. Let $f = \tau \circ t$ be any bilinear form, where τ is its mediating morphism. Then

$$\tau(\sum_{i=1}^{n} u_i \otimes v_i) = \sum_{i=1}^{n} \tau(u_i \otimes v_i) = \sum_{i=1}^{n} f(u_i, v_i) = 0.$$

This holds true for any bilinear form f. Choose f to be given as

$$f(u, v) = u^k(u)\beta(v),$$

where u^k is the dual basis for $\{u_k\}$ and $\beta \in V^*$. Using above, we can see that

$$\tau(\sum_{i=1}^{n} u_i \otimes v_i) = \sum_{i=1}^{n} u^k(u)\beta(v_i) = \beta(v_k) = 0.$$

This holds true for all k and all β. So $v_k = 0$ for all k. □

Observation 4.1.19 — Injectivity of inclusion map in Theorem 14.7 Roman. In this theorem, we define a linear map $\theta : U^* \times V^* \to (U \times V)^*$ by using the universal property of tensor products. I.e. we exhibit a bilinear map $F : U^* \times V^* \to (U \otimes V)^*$ defined by

$$(F(f, g))(u \otimes v) = f(u)g(v),$$

for $f \in U^*$ and $g \in V^*$. So this induces a unique linear map $\theta : U^* \otimes V^* \to (U \otimes V)^*$. We now want to show that this linear map is indeed an injection. Let $h \in U^* \otimes V^*$ be in the kernel of θ. I.e. $\theta(h)$ will be the zero functional on $U \otimes V$. I.e. for all $z \in U \otimes V$ we can write

$$(\theta(h))(z) = 0.$$

However, we can represent h as

$$h = \sum_{i=1}^{m} f_i \otimes g_i$$

where $\{f_i\} \subset U^*$ and $\{g_i\} \subset V^*$ are linearly independent. So

$$\theta(h) = \sum_{i=1}^{m} \theta(f_i \otimes g_i) = 0.$$

Let $\{e_k \otimes e'_l\}_{k,l}$ be basis vectors of $U \otimes V$. Since $\theta(h)$ is the zero functional, for all k, l we need to have

$$(\theta(h))(e_k \otimes e'_l) = \sum_{i=1}^{m} f_i(e_k) g_i(e'_l) = 0 \quad \forall k, l.$$

We can interpret this in two ways:

$$\sum_{i=1}^{m} g_i(e'_l) f_i = 0, \quad \text{and} \quad \sum_{i=1}^{m} f_i(e_k) g_i = 0,$$

Since $\{f_i\}$ and $\{g_i\}$ are linearly independent, it follows that $g_i(e'_l) = 0$ for all i and l. So $\{g_i\}$ are all zero functionals. We can similarly conclude that $\{f_i\}$ are all zero functionals.

Observation 4.1.20 — Useful fact!. When we are seeking a linear map whose domain is a tensor product, we need to find a bilinear form whose domain is a Cartesian product of the spaces, and then to use to universal property of the tensor products get the linear map from the tensor product space.

Observation 4.1.21 — Infinitesimals Vs. Nilpotents. In the context of linear algebra, the nilpotent maps are the maps that if raised to a large enough power, will be a zero map. This is a very interesting concept because it has very interesting similarities to the infinitesimals. When working with Floating Point representation of numbers in computers, then we are also dealing with nilpotents. For instance, in a floating point system, we can have $(0.1)**5 = 0$. So in this system 0.1 acts as a nilpotent of degree less than five.

Presence of nilpotents in linear algebra, terminates some power series involving matrices. With a similar idea, all of the power series in a floating point representation is a finite sum. That is because if small values are raised to high enough powers, then at some point they become negligible compared to the machine precision, and this leads to the termination of the power series.

Observation 4.1.22 — A convenient way to represent any bilinear form . Any bilinear form $(\cdot, \cdot) : U \times V \to F$ can be represented in a very convenient way. Let $u \in U$ and $v \in V$ with $u = \sum_{i=1}^{n} u_i e_i$ and $v = \sum_{j=1}^{m} v_j e'_j$. So we can write the bilinear form as

$$(u, v) = (\sum_{i=1}^{n} u_i e_i, \sum_{j=1}^{m} v_j e'_j) = \sum_{i,j} u_i v_j (e_i, e'_j) = \sum_i u_i \sum_j (e_i, e'_j) v_j = \sum_i u_i (Fv)_i = a^t F b,$$

where

$$(F)_{i,j} = (e_i, e_j).$$

Observation 4.1.23 — Some thoughts on the tensor product of weird stuff with other weird stuff. It is quote common in the topic of tensor products to see tensor product of spaces that look very bizarre! For instance, $U \otimes \operatorname{End} V$, $U^* \otimes U$, $U^* \otimes \operatorname{Hom}(V, W)$, and etc are few examples. Trying to interpret these objects can be quite messy and there is a danger of overthinking leading to a very detailed and non functional intuition about their nature.

However, there is a simple intuitive way to look at these objects. And in fact, everybody has been interacting with these kind of objects from their high school. One of the prime examples is working with functions that are parameterized. For instance, consider $f(x) = x^2 - \alpha$ that arises in studying the bifurcations in dynamical systems. Assume we want to consider the behaviour of this function for different values of its parameters. One way is to define a function with a same name but $f : \mathbb{R} \times I \to \mathbb{R}$, where $\alpha \in I$. I.e. we can assume that the function is defined on the Cartesian product of $\mathbb{R}$ (where x lives) and I (where α lives). However, this is not continent in different ways, and instead, we might choose to talk about a family of functions $\{f_\alpha\}_{\alpha \in I}$ and $f_\alpha : \mathbb{R} \to \mathbb{R}$ given by $f_\alpha(x) = x^2 - \alpha$. In this way, the emphasis is still on the the values of the function defined on the original space $\mathbb{R}$, but we also consider the parameter α in an elegant way.

The objects in the tensor spaces above has a very similar stories, but with even more information. For instance, consider $U \otimes \operatorname{End} V$. The objects in the space, are *endomorphisms* from V to V that are parameterized by U, and the parameterization is *linear*. So $\sigma \in U \otimes \operatorname{End} V$, in its input, should receive a package, that contains about $v \in V$ (on which the elements of $\operatorname{End} V$ acts). And in the output, we need to distribute these inputs among the appropriate consumers and combine their results in a suitable way (there might be several ways to do so). What we have described above, is simply saying that there is an injection

$$\theta : U \otimes \operatorname{End} V \to \operatorname{Hom}(V, U \otimes V),$$

where for instance for $u \otimes \sigma$ we can write

$$(\theta(u \otimes \sigma))(v) = u \otimes \sigma(v)$$

Or, as another example, consider $U^* \otimes \operatorname{Hom}(V, W)$. Objects in this space are linear operators from V to W that are parameterized by some linear functional on U. So for instance for $f \otimes \sigma$ where $f \in U^*$ and $\sigma \in \operatorname{Hom}(V, W)$ we can write

$$(f \otimes \sigma)(u, v) = f(u)\sigma(v).$$

Observation 4.1.24 — Some verbal explanation on a hard looking problem. Somewhere in one of the homework problems, one asks to find a natural isomorphism

$$\theta : \operatorname{End}(U \otimes V) \to \operatorname{Hom}(U, U \otimes \operatorname{End}(V)).$$

At first, this might seem like a hard problem, but there is nothing hard about it. The proof is writing down almost everything comes into your mind in your natural language. First, lets make a guess on what this isomorphism can be. I.e. given any element in $\operatorname{End}(U \otimes V)$, then $\theta(\sigma)$ should be interpreted by an element in $\operatorname{Hom}(U, U \otimes \operatorname{End}(V))$. Let's go through this one step at a time. If $\theta(\sigma)$ is an element in $\operatorname{Hom}(U, U \otimes \operatorname{End}(V))$, then it should accept $u \in U$, and $[\theta\sigma](u)$ should be an element in $U \otimes \operatorname{End} V$, i.e. an endomorphism on V parameterized by a vector u. So $[\theta\sigma](u)$, thought as a parameterized endomorphism, should receive $v \in V$ and do with it anything that a parameterized endomorphsim does (i.e. parameter $\otimes$ result of the endomorphism on v). So

$$([\theta\sigma](u))(v) = \sigma(u \otimes v).$$

As another example, we want to find a natural isomorphism

$$\theta : \mathrm{Hom}(U \otimes V_1, U \otimes V_2) \to \mathrm{Hom}(U, U \otimes \mathrm{Hom}(V_1, V_2)).$$

Let $\sigma \in \mathrm{Hom}(U \otimes V_1, U \otimes V_2)$. So we can write

$$((\theta\sigma)(u))(v_1) = \sigma(u \otimes v_1).$$

Observation 4.1.25 Let U, V be real vector spaces equipped with non-degenerate inner products. We can show that

$$\langle u_1 \otimes v_1, u_2 \otimes v_2 \rangle_{U \otimes V} := \langle u_1, u_2 \rangle_U \langle v_1, v_2 \rangle.$$

Let $A \in \mathrm{End}(V)$ and $B \in \mathrm{End}(V)$. Then we can show that

$$(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger.$$

To see this observe that

$$\begin{aligned}
\langle (A \otimes B)^\dagger u \otimes v, x \otimes y \rangle &= \langle u \otimes v, (A \otimes B)(x \otimes y) \rangle \\
&= \langle u \otimes v, Ax \otimes By \rangle \\
&= \langle u, Ax \rangle \langle v, By \rangle \\
&= \langle A^\dagger u, x \rangle \langle B^\dagger v, y \rangle \\
&= \langle A^\dagger u \otimes B^\dagger v, x \otimes y \rangle \\
&= \langle (A^\dagger \otimes B^\dagger)(u \otimes v), x \otimes y \rangle
\end{aligned}$$

Observation 4.1.26 Theorem 14.5 in Roman and P2 in Lior's homework problems seems to be quite uncompatible. But there is a very subtle difference. What Lior is saying is that a pure tensor is zero if and only if and only if one of their components is zero. But the statement in Theorem 14.5 is not saying this.

Observation 4.1.27 — Some notes and the dual basis. The dual basis might be quite confusing for people who encounter it for the first time. However, there is a very simple and intuitive semantics behind it, which is as described below. Let $f \in U^*$ be a linear functional $f : U \to F$, where F is the underlying field. f will be uniquely determined by its values on the basis vectors and extending it linearly to the whole space. So if $U = \mathrm{Span}\, u_1, \cdots, u_n$ then to fully specify f we need to determine

$$u_1 \mapsto \alpha_1, \quad u_2 \mapsto \alpha_2, \quad \cdots, \quad u_n \mapsto \alpha_n.$$

An alternative way to write the mappings above in a more compact and computationally useful way is to write

$$f = \alpha_1 u^1 + \alpha_2 u^2 + \cdots + \alpha_n u^n,$$

where u^i is a special kind of linear function that is determined by mapping u_i to one and all the rest basis vectors to zero. Now we can record all these information in a row matrix, and $f(x)$ for some vector x will by simply multiplying that row vector in x from left.

Also, using the machinery of the tensor product we can do a similar thing with the bilinear forms. Let $F : U \otimes U \to F$ be a bilinear form. By the universal property of the tensor products we can see that it induces a linear form $\tau : U \otimes V$. So its value can be determined by its values on the basis vectors of $U \otimes U$, that is $\{e_i \otimes e_j\}_{i,j}$. One way to write down this specification is to write

$$e_1 \otimes e_1 \mapsto \alpha_{11}, \quad e_1 \otimes e_2 \mapsto \alpha_{12}, \quad \cdots, e_i \otimes e_j \mapsto \alpha_{i,j}, \quad e_n \otimes e_n \mapsto \alpha_{n,n}.$$

Or since σ is a linear functional, similar to the previous example we can write it simply as

$$\sigma = \sum_{i,j} \alpha_{i,j} e^i \otimes e^j,$$

where $e^i \otimes e^j$ is a special kind of linear functional on $U \otimes U$ that is specified by sending $e_i \otimes e_j$ to zero and everything else to zero. In other words

$$(e^i \otimes e^j)(e_k \otimes e_l) = \delta_{i,k}\delta_{j,l}.$$

Now we can record all this information in a matrix as

$$M = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

And then $\sigma(u \otimes v) = F(u, v)$ will simply be

$$u^t M v.$$

This way, it is very easy to get a high level characterization of quite complicated bilinear forms. Consider the following observation box for more details on this.

Observation 4.1.28 — Characterizing anti-symmetric and symmetric forms. Using the detailed explained in the observation box above, we can arrive at a very high level and easy to understand characterization of anti-symmetric and symmetric forms. Let $U = \mathbb{R}^2$ and we want to characterize all the anti-symmetric bilinear forms on this space. Let $f : U \times U \to \mathbb{R}$ be such functional. By the universal property of tensor products this map induces a linear map $\sigma : U \otimes U \ \to \mathbb{R}$. To specify this linear map we need to tabulate the values that it takes on the basis vectors. The requirement for being anti-symmetric forces this map to be

$$e_1 \otimes e_1 \mapsto 0, \quad (e_1 \otimes e_2 - e_2 \otimes e_1) \mapsto b, \quad e_2 \otimes e_2 \mapsto 0.$$

This suggest that an anti-symmetric bilinear form should have a matrix representation as

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

■ **Remark** The form of the matrix above implies that any complex number ib when viewed as a matrix via the following inclusion map

$$ib \mapsto \begin{pmatrix} 0 & -b \\ b & 0 \end{pmatrix},$$

can be viewed as a bilinear form if we interpret the matrix above as the matrix representation of a bilinear form.

So on $\mathbb{R}^2$ every anti-symmetric bilinear form will have a matrix representation as

$$\begin{pmatrix} 0 & ♫ \\ -♫ & 0 \end{pmatrix}$$

Similarly, on $U = \mathbb{R}^4$, any anti-symmetric bilinear form, when viewed as a linear form on the tensor space $U \otimes U$ should have

$$\begin{aligned} (e_1 \otimes e_2 - e_2 \otimes e_1) &\mapsto a, & (e_1 \otimes e_3 - e_3 \otimes e_1) &\mapsto b, \\ (e_1 \otimes e_4 - e_4 \otimes e_1) &\mapsto c, & (e_2 \otimes e_3 - e_3 \otimes e_2) &\mapsto d, \\ (e_2 \otimes e_4 - e_4 \otimes e_2) &\mapsto e, & (e_3 \otimes e_4 - e_4 \otimes e_3) &\mapsto f. \end{aligned}$$

Instead of the tabulation above, this can be summarized in the following expression

$$\begin{aligned} f = a(e^1 \otimes e^2 - e^2 \otimes e^1) + b(e^1 \otimes e^3 - e^3 \otimes e^1) \\ +c(e^1 \otimes e^4 - e^4 \otimes e^1) + d(e^2 \otimes e^3 - e^3 \otimes e^2) \\ +e(e^2 \otimes e^4 - e^4 \otimes e^2) + f(e^3 \otimes e^4 - e^4 \otimes e^3). \end{aligned}$$

So the matrix representation will be of the form

$$M = \begin{pmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{pmatrix}.$$

It is very easy to see that the dimension of this space (the space of all alternating bilinear forms on $\mathbb{R}^4$ is $3 + 2 + 1 = 6 = \binom{4}{2}$). We can do a similar argument for symmetric bilinear forms on $\mathbb{R}^4$. Any such symmetric map should have

$$\begin{aligned} e_1 \otimes e_1 \mapsto a, \quad e_2 \otimes e_2 \mapsto b, \quad e_3 \otimes e_3 \mapsto c, \quad e_4 \otimes e_4 \mapsto d \\ (e_1 \otimes e_2 + e_2 \otimes e_1) \mapsto e, \quad (e_1 \otimes e_3 + e_3 \otimes e_1) \mapsto f, \quad (e_1 \otimes e_4 + e_4 \otimes e_1) \mapsto g, \\ (e_2 \otimes e_3 + e_3 \otimes e_2) \mapsto h, \quad (e_2 \otimes e_4 + e_4 \otimes e_2) \mapsto i, \quad (e_3 \otimes e_4 + e_4 \otimes e_3) \mapsto j. \end{aligned}$$

Its matrix representation will be

$$M = \begin{pmatrix} a & e & f & g \\ e & b & h & i \\ f & h & c & j \\ g & i & j & d \end{pmatrix}$$

It is also very clear that the dimension of this space is

$$4 + (3 + 2 + 1) = 4 + \frac{4(4-1)}{2} = 10.$$

Observation 4.1.29 One possible way to write the content of two observation boxes above, is to write

$$\{\text{symmetric bilinear forms}\} \cong (\operatorname{Sym}^2 U)' \cong \operatorname{Sym}^2(U').$$

It is very easy to see this if we have a simple running example in our mind.

To show the first isomorphism, let $U = \mathbb{R}^3$. Now let f be a symmetric bilinear form. By the universal property of tensor products this induces a linear form $\tau : U \otimes U \to F$ where $f = \tau \circ t$ where t is the tensor map. Now restrict τ to the subspace

$$\operatorname{Sym}^2 U = \operatorname{Span}\{e_1 \otimes e_1, e_2 \otimes e_2, e_3 \otimes e_3, e_1 \otimes e_2 + e_2 \otimes e_1, e_1 \otimes e_3 + e_3 \otimes e_1, e_2 \otimes e_3 + e_3 \otimes e_2\},$$

so $\tau|_{\operatorname{Sym}^2 U} \in (\operatorname{Sym}^2 U)'$. For the converse, let $f \in (\operatorname{Sym}^2 U)'$. So f is characterized by its values on the basis vectors:

$$\begin{aligned}
&e_1 \otimes e_1 \mapsto \alpha, \quad e_2 \otimes e_2 \mapsto \beta, \quad e_3 \otimes e_3 \mapsto \gamma, \\
&e_1 \otimes e_2 + e_2 \otimes e_1 \mapsto \omega, \quad e_1 \otimes e_3 + e_3 \otimes e_1 \mapsto \lambda, \quad e_2 \otimes e_3 + e_3 \otimes e_2 \mapsto \mu.
\end{aligned}$$

Define bilinear map $F : U \times U \to \mathbb{R}$ given by $(e_1, e_1) \mapsto \alpha$, $(e_2, e_2) \mapsto \beta$, and so on, and extend by bilinearity to the whole space.

For the second isomorphism, Let $f \in (\operatorname{Sym}^2 U)'$ (similar to above). Now let $h \in \operatorname{Sym}^2(U')$ given by

$$\begin{aligned}
g = {}&\alpha e_1 \otimes e_1 + \beta e_2 \otimes e_2 + \gamma e_3 \otimes e_3 \\
&+ \omega(e_1 \otimes e_2 + e_2 \otimes e_1) + \lambda(e_1 \otimes e_3 + e_3 \otimes e_1) + \mu(e_2 \otimes e_3 + e_3 \otimes e_2).
\end{aligned}$$

For the converse, we will do the a similar thing by in reverse.

Observation 4.1.30 — On the universal property of direct sum. The universal property of direct sum (demonstrated here for two spaces), simply states that when I have a pair of vector spaces (V_1, V_2) and $A_1 \in \operatorname{Hom}(V_1, W)$ and $A_2 \in \operatorname{Hom}(V_2, W)$ where W is any other vector space, then there is a vector space $V_1 \oplus V_2$ on which (A_1, A_2) can be thought of as a single linear map $T : V_1 \oplus V_2 \to W$, where

$$(A_1, A_2) = T \circ (\iota_1, \iota_2) = (T \circ \iota_1, T \circ \iota_2),$$

where $\iota_1 : V_1 \to V_1 \oplus V_2$ and $\iota_2 : V_2 \to V_1 \oplus V_2$ are the inclusion maps. Solving the problem below, will demonstrate an application.

■ **Problem 4.1.1** Let $\{V_i\}_{i \in I}$ be a family of vector spaces, and let $A_i \in \operatorname{End}(V_i)$.

(a) Show that there is a unique element $\oplus_i A_i \in \operatorname{End}(\oplus_i V_i)$ whose restriction to the image of V_i in the sum is A_i.

Solution (a) We are looking for a map whose domain is $\oplus_i V_i$. So we need to first find a linear map for each of the vector spaces in the direct sum above. Consider

$$\iota_i \circ A_i : V_i \mapsto \oplus_i V_i.$$

So by the universal property of the direct sum, there exists $\oplus_i A_i : \oplus_i V_i \to \oplus_i V_i$, such that

$$\iota_i \circ A_i = \oplus_i A_i \circ \iota_i.$$

4.2 Ongoing thoughts

Observation 4.2.1 — Ongoing thought on the relation of the space of linear operators and the tensor product. In many instances, I have noticed a similar structure between $\mathcal{L}(V, W)$ and $V \otimes W$. For instance, we know that while $u \otimes v$ is a tensor (a pure tensor), but not every tensor can be written like this, but rather it is a linear combination of pure tensors. This is very similar to the idea that for $A : V \to W$ and $B : U \to W$, we can construct a linear map $C : U \oplus V \to W$, that has a block diagonal representation. But, we can not write every matrix in a block diagonal representation.

Also, another hint is that $\dim(\mathcal{L}(V, W)) = n \times m$, and similarly, $\dim(U \otimes V) = n \times m$. Yet another hint is that every elements of $U \otimes V$ has a matrix coordinate. I need to make this connection more clear and easy to see / understand.

4.3 Questions

■ **Problem 4.3.1** (a) Let Z be any vector space, and suppose we have for each i a linear map $g_i : Z \to V_i$. Show that there is a unique $g : Z \to \prod_i V_i$ such that $\pi_i \circ g = g_i$ for all i.

Solution A quick reminder that $\prod_{i \in I} V_i$ is

$$\prod_{i \in I} V_i = \{f : I \to \bigcup_i V_i \ : \ f(i) \in V_i\}.$$

This a generalization of n-tuple. C

$$\begin{tikzcd} & Z \arrow[dl, "(g_i)_{i \in I}"'] \arrow[d, "g"] \\ (V_i)_{i \in I} & \prod_{i \in I} V_i \arrow[l, "\pi_i"] \end{tikzcd}$$

Defien $g : Z \to \prod_i V_i$ by

$$\pi_i \circ g = g_i.$$

Note that g is uniquely determined by this definition. Indeed, for $v \in Z$ we have

$$(g(v))(i) = \pi_i(g(v)) = (\pi_i \circ g)(v) = g_i(v).$$

We need to show that g is linear. Let $v, u \in Z$ and α, β scalars. Then

$$(g(\alpha v + \beta u))(i) = (\pi_i \circ g)(\alpha v + \beta u) = g_i(\alpha v + \beta u) = \alpha g_i(v) + \beta g_i(u) = \alpha(g(u))(i) + \beta(g(v))(i).$$

Since this is true for the g component-wise, then it follows that

$$g(\alpha v + \beta u) = \alpha g(v) + \beta g(u).$$

■ **Problem 4.3.2** Let U, V, W be vector spaces, and let $T : U \to V$ and $S : V \to W$.

(a) Suppose U, V are finite dimensional with bases $\{u_j\}_{j=1}^m \subset U$ and $\{v_i\}_{i=1}^n \subset V$, and let $A \in M_{n,m}(F)$ be the matrix of T in those bases. Show that the matrix of the dual map $T' : V' \to U'$ with respect to the dual bases $\{u^j\}, \{v^j\}$ is the transpose of A.

Solution (a) Let $\tilde{A}$ denote the matrix representation of T'. Then $(\tilde{A})_{i,j}$, is the i^{th} component (in the dual basis) of $T'v^j$. Note that $T'v^j$ is a functional on U and to get its i^{th} component we need to apply it on u_i. So

$$(\tilde{A})_{ij} = (T'(v^j))(u_i) = v^j(Tu_i) = (A)_{ji}.$$