MATH1049 Linear Algebra II lecture notes

Ján Špakula

Spring 2021

Contents

Preface **2**
Acknowledgement . . . 2

1 Groups **3**

2 Fields and Vector Spaces **8**
2.1 Fields . . . 8
2.2 Vector spaces . . . 9
2.3 Subspaces . . . 13

3 Bases **16**
3.1 Spanning . . . 16
3.2 Linear independence . . . 17
3.3 Bases . . . 19
3.4 Dimension . . . 21

4 Linear Transformations **25**
4.1 Matrix representation I . . . 26
4.2 Kernel and image . . . 27
4.3 Isomorphism . . . 29
4.4 Dimension Theorem . . . 31
4.5 Matrix representation II . . . 33

5 Determinants **36**

6 Diagonalisability **40**
6.1 Eigen-things . . . 40
6.2 Diagonalisability . . . 42
6.3 Cayley–Hamilton Theorem . . . 46

7 Coursework Sheets **48**
7.1 Coursework Sheet 0 — not marked . . . 48
7.2 Coursework Sheet 1 . . . 49
7.3 Coursework Sheet 2 . . . 49
7.4 Coursework Sheet 3 . . . 51

7.5 Coursework Sheet 4 52
7.6 Coursework Sheet 5 54
7.7 Coursework Sheet 6 55
7.8 Coursework Sheet 7 56
7.9 Coursework Sheet 8 57

Preface

These lecture notes contain more–less verbatim what appears on the (black– or white–)board during the lectures.

The colour coding signifies the following: what is red is being *defined*. Blue are *named theorems and statements* (these could be referred to by these names). Finally green is used to reference lecture notes of other modules (mostly MATH1048 Linear Algebra I).

Acknowledgement

For the most part, these notes were designed and written by Dr Bernhard Koeck, and originally typed into LaTeX by the undergraduate student Thomas Blundell–Hunter in summer 2011. They have been regularly updated to reflect the changes in the syllabus. The current version is a complete overhaul of the technical side of the notes, so that they are available also in `html`.

1 Groups

The cartesian product $G \times H$ of two sets G and H is defined as the set

$$G \times H := \{(g,h) \mid g \in G, h \in H\},$$

consisting of all (ordered) pairs (g,h) where $g \in G$ and $h \in H$. For example, $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$.

Definition 1.1. (a) A binary operation on a set G is a function $G \times G \to G$ (from the cartesian product $G \times G = \{(a,b) \mid a \in G, b \in G\}$ to G). We write $a * b$ (or ab, or $a \circ b$, or $a + b$) for the image of the pair $(a,b) \in G \times G$ in G under this function.

(b) A group is a set G together with a binary operation $G \times G \to G$, $(a,b) \mapsto a * b$ (which we usually refer to as "group operation", or "group multiplication", or just "multiplication") such that the following axioms are satisfied:

(i) *Associativity*: For all $a,b,c \in G$, we have $(a * b) * c = a * (b * c)$ in G.

(ii) *Existence of the identity (or neutral) element*: There exists $e \in G$, such that for all $a \in G$ we have $e * a = a = a * e$.

(iii) *Existence of the right inverse*: For every $a \in G$ there exists $b \in G$ such that $a * b = e$.

(c) A group is called abelian (or commutative), if for all $a,b \in G$ we have $a * b = b * a$.

(d) A group is called finite if G has only finitely many elements; in this case its cardinality $|G|$ is called the order of G.

Example 1.2. (a) The usual addition $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$, $(m,n) \mapsto m + n$, defines a binary operation on $\mathbb{N}$ (and so does multiplication), but subtraction does not, because for instance $2 - 3 = -1 \notin \mathbb{N}$.

(b) The sets $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ together with addition as the binary operation are abelian groups. The sets $\mathbb{Q}^* := \mathbb{Q} \setminus \{0\}$, $\mathbb{R}^* := \mathbb{R} \setminus \{0\}$ and $\mathbb{C}^* := \mathbb{C} \setminus \{0\}$ together with multiplication as the binary operation are abelian groups.

(c) The set $\mathbb{N}$ with addition is not a group, because, for example, there doesn't exist a (right) inverse to $1 \in \mathbb{N}$. (In other words, the equation $1 + ? = 0$ has no solution in $\mathbb{N}$.)

(d) For any $n \in \mathbb{N}$, the set $\mathbb{R}^n$ together with vector addition is a group. For any $m,n \in \mathbb{N}$ the set $M_{m \times n}(\mathbb{R})$ of real m–by–n matrices together with matrix addition is a group. For every $n \in \mathbb{N}$, the set $\mathrm{GL}_n(\mathbb{R})$ of invertible real $(n \times n)$–matrices together with matrix multiplication is a group, called the *general linear group*. If $n > 1$, $\mathrm{GL}_n(\mathbb{R})$ is not abelian: e.g.

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad \text{but} \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Proposition 1.3. *Let G be a group. Then:*

(a) There exists precisely one identity element in G.

(b) For every $a \in G$ there exists precisely one right inverse (call it b) to a, and this b is also a left inverse of a (meaning that we also have $b * a = e$). We write a^{-1} for this inverse of a (or $-a$ if the group operation is "$+$").

Proof. (a) Suppose e and e' are two identity elements in G.
$\Longrightarrow$ $e * e' = e$ (because e' is an identity element)
and $e * e' = e'$ (because e is an identity element)
$\Longrightarrow$ $e = e'$.

(b) Let b be a right inverse of a, and let c be a right inverse of b.
$\Longrightarrow$ $a*(b*c) = a*e$ (because c is a right inverse of b)
$= a$ (because e is the identity element)
and $a*(b*c) = (a*b)*c$ (by associativity)
$= e*c$ (because b is a right inverse of a)
$= c$ (because e is the identity element)
$\Longrightarrow$ $a = c$
$\Longrightarrow$ $b*a = b*c = e$. (because c is a right inverse of b)
In other words, b is also a left inverse of a.
Suppose that b' is another right inverse of a.
$\Longrightarrow$ $b = b*e = b*(a*b') = (b*a)*b' = e*b' = b'$.
(These equalities hold by: definition of e, definition of b', associativity, the fact we just proved, and by definition of e, respectively.) □

Proposition 1.4. *Let G be a group.*

(a) (Cancellation) *If $a,b,z \in G$ and $a*z = b*z$ (or $z*a = z*b$), then $a = b$.*

(b) For all $a \in G$ both the equation $a*x = b$ and $y*a = b$ have a unique solution in G. (Another way to say this: for every $a \in G$, both the map $G \to G$, $x \mapsto a*x$ (called left translation by a) and $G \to G$, $y \mapsto y*a$ (called right translation by a) are bijective.)

(c) For every $a \in G$ we have $(a^{-1})^{-1} = a$.

(d) For all $a,b \in G$ we have $(a*b)^{-1} = b^{-1}*a^{-1}$.

(e) (Exponential laws) *For any $m \in \mathbb{Z}$ and $a \in G$, we define:*

$$a^m := \begin{cases} \underbrace{a*a*\cdots*a}_{m \text{ times}} & \text{if } m > 0; \\ e & \text{if } m = 0; \\ (a^{-1})^{|m|} & \text{if } m < 0. \end{cases}$$

(In additive notation, i.e. when the group operation is "+", we write ma instead of a^m.)
Then for all $m,n \in \mathbb{Z}$ and $a \in G$ we have $a^{m+n} = a^m * a^n$ and $a^{mn} = (a^m)^n$.
If $a,b \in G$ commute (i.e. $a*b = b*a$), then for all $m \in \mathbb{Z}$ we have $(a*b)^m = a^m * b^m$.

Proof. (a) Multiply both sides of the equation by z^{-1}.

(b) Proof of the "another way": *Injectivity*: use (a).
Surjectivity: If $b \in G$, then $b*a^{-1}$ is mapped to b under the right translation by a.

(c) Both $(a^{-1})^{-1}$ and a are solutions of the equation $a^{-1}*x = e$ (see Proposition 1.3 (b)). Now apply (b).

(d) We have $(a*b)*(b^{-1}*a^{-1}) = a*(b*(b^{-1}*a^{-1})) = a*((b*b^{-1})*a^{-1}) = a*(e*a^{-1})$
$= a*a^{-1} = e$. $\Longrightarrow$ $b^{-1}*a^{-1}$ is a right inverse to $a*b$.

(e) Left as an exercise. □

Example 1.5. (a) The group table of a finite group $G = \{a_1, a_2, \ldots, a_n\}$ is a table like this:

$*$	$\cdots$	a_j	$\cdots$
$\vdots$	$\ddots$	$\vdots$	
a_i	$\cdots$	$a_i * a_j$	$\cdots$
$\vdots$		$\vdots$	

The trivial group is the group with exactly one element, say e. Its group table must be

$*$	e
e	e

Any group with two elements, say e and a, is given by the group table:

$*$	e	a
e	e	a
a	a	e

Any group with three elements, say e, a and b, must have group table:

$*$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

(Note that $a * b = a$ would imply $b = e$ by 1.4(a).) There are two "essentially different" groups of order 4.
Note that Proposition 1.4(b) implies that group tables must satisfy "sudoku rules", i.e. that every group element must appear in each row and each column exactly once. However, not every table obeying this rule is a *group table* of a group; for example the table below does not. Why? (Hint: what is $a * a * b$?)

e	a	b	c	d
a	e	c	d	b
b	c	d	e	a
c	d	a	b	e
d	b	e	a	c

(b) Let $m \in \mathbb{N}$ and define $C_m := \{\overline{0}, \overline{1}, \ldots, \overline{m-1}\}$. Define a binary operation on the set C_m by:

$$\overline{x} \oplus \overline{y} := \begin{cases} \overline{x+y} & \text{if } x+y < m; \\ \overline{x+y-m} & \text{if } x+y \geq m. \end{cases}$$

Then C_m together with $\oplus$ is an abelian group called the cyclic group of order m. (Caveat: these notes will use $\oplus$ for the group operation on C_m, to distinguish it from "+" between numbers. However it is very common to just use "+" for the operation on C_m.)

Proof (that C_m is a group).
Associativity: Let $x, y, z \in \{0, 1, \ldots, m-1\}$. Want to show $(\overline{x} \oplus \overline{y}) \oplus \overline{z} = \overline{x} \oplus (\overline{y} \oplus \overline{z})$ in C_m.

First case: Suppose that $x+y+z<m$. Then also $x+y<m$ and $y+z<m$.
$\Longrightarrow$ LHS $=\overline{x+y}\oplus\overline{z}=\overline{(x+y)+z}=\overline{x+(y+z)}=\overline{x}\oplus\overline{y+z}=$ RHS (using associativity of addition of integers).

Second case: Suppose $m\le x+y+z<2m$.
Subcase (i): Suppose $x+y<m$.
$\Longrightarrow$ LHS $=\overline{x+y}\oplus\overline{z}=\overline{x+y+z-m}$.
Subcase (ii): Suppose $x+y\ge m$.
$\Longrightarrow$ LHS $=\overline{x+y-m}\oplus\overline{z}=\overline{x+y+z-m}$.
Hence in both subcases we have LHS $=\overline{x+y+z-m}$.
Similarly we obtain that also RHS $=\overline{x+y+z-m}=$ LHS.

Third case: Suppose $x+y+z\ge 2m \implies x+y\ge m$ and $y+z\ge m$.
$\Longrightarrow$ LHS $=\overline{x+y-m}\oplus\overline{z}=\overline{x+y+z-2m}$ (since $(x+y-m)+z\ge m$)
and RHS $=\overline{x}\oplus\overline{y+z-m}=\overline{x+y+z-2m}$ (since $x+(y+z-m)\ge m$).
$\Longrightarrow$ LHS = RHS.

Identity element: $\overline{0}$ is an identity element in C_m, because for any $\overline{x}\in C_m$ we have $\overline{x}\oplus\overline{0}=\overline{x+0}$ $=\overline{x}=\overline{0+x}=\overline{0}\oplus\overline{x}$.

Existence of right inverses: Let $\overline{x}\in C_m$. If $x\neq 0$, then the inverse to $\overline{x}$ is $\overline{m-x}\in C_m$. If $x=0$, then the inverse to $\overline{x}=\overline{0}$ is $\overline{0}$. □

(c) Let S be a set (such as $S=\{1,2,\dots,n\}$ for some $n\in\mathbb{N}$) and let Sym(S) denote the set of all bijective maps $\pi: S\to S$, also called permutations of S.
For any $\pi,\sigma\in\mathrm{Sym}(S)$, we denote $\sigma\circ\pi$ their composition (as functions, so $(\sigma\circ\pi)(s)=\sigma(\pi(s))$ for all $s\in S$). This defines a binary operation on Sym(S).
Then Sym(S) together with composition is a group, called the permutation group of S (or sometimes also the symmetric group of S).
The identity element in Sym(S) is the identity function (denoted id_S or just id), and the inverse of $\pi\in\mathrm{Sym}(S)$ is the inverse function π^{-1} (as in Calculus I).
If $S=\{1,\dots,n\}$ for some $n\in\mathbb{N}$, we write S_n for Sym(S) and use the "table notation" to describe permutations $\pi\in S_n$ as $\begin{pmatrix}1&2&\cdots&n\\ \pi(1)&\pi(2)&\cdots&\pi(n)\end{pmatrix}$. For example:

$$\begin{pmatrix}1&2&3&4\\3&1&2&4\end{pmatrix}^{-1}=\begin{pmatrix}1&2&3&4\\2&3&1&4\end{pmatrix}\quad\text{in } S_4,$$
$$\begin{pmatrix}1&2&3&4&5\\3&4&1&2&5\end{pmatrix}\circ\begin{pmatrix}1&2&3&4&5\\2&3&5&1&4\end{pmatrix}=\begin{pmatrix}1&2&3&4&5\\4&1&5&3&2\end{pmatrix}\quad\text{in } S_5.$$

Definition 1.6. Let $n\ge 1$ and $s\le n$. Let $a_1,\dots,a_s\in\{1,\dots,n\}$ be pairwise distinct. The permutation $\pi\in S_n$ such that

$$\pi(a_1)=a_2,\quad \pi(a_2)=a_3,\quad \dots,\quad \pi(a_{s-1})=a_s,\quad \pi(a_s)=a_1,$$
$$\text{and}\quad \pi(a)=a\quad \text{for } a\in\{1,\dots,n\}\setminus\{a_1,\dots,a_s\},$$

is denoted by $\langle a_1,\dots,a_s\rangle$. Any permutation of this form is called a cycle. If $s=2$, it is called a transposition. The number s is called the length (or order) of the cycle.

For example, $\langle 3,1,5,2\rangle=\begin{pmatrix}1&2&3&4&5&6\\5&3&1&4&2&6\end{pmatrix}$ in S_6; and $\langle 3,2\rangle=\begin{pmatrix}1&2&3&4\\1&3&2&4\end{pmatrix}$ in S_4.

Proposition 1.7. *Every permutation $\sigma\in S_n$ is a composition of cycles.*

Example 1.8. (a) Let $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 2 & 5 & 6 & 9 & 1 & 4 & 10 & 11 & 3 & 7 & 8 \end{pmatrix} \in S_{11}$. Then $\sigma = \langle 1,2,5\rangle \circ \langle 3,6,4,9\rangle \circ \langle 7,10\rangle \circ \langle 8,11\rangle$.

(b) Let $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix} \in S_4$. Then $\tau = \langle 1,4\rangle$.

General Recipe (which, with a bit of effort, can be turned into a proof of 1.7).
Denote by $\sigma \in S_n$ the permutation that we want to write as a composition of cycles.
Start with some $a \in \{1,\dots,n\}$ such that $\sigma(a) \neq a$ (e.g. $a := 1$). (If there is no such a then $\sigma = \mathrm{id}$ and we are done.)
Let $m \in \mathbb{N}$, $m > 1$, be the smallest number such that $\sigma^m(a) \in \{a, \sigma(a), \sigma^2(a), \dots, \sigma^{m-1}(a)\}$. (Actually then necessarily $\sigma^m(a) = a$.)
Let σ_1 be the cycle $\sigma_1 = \langle a, \sigma(a), \sigma^2(a), \dots, \sigma^{m-1}(a)\rangle$.
Now repeat the steps: take $b \in \{1,\dots,n\} \setminus \{a, \sigma(a), \dots, \sigma^{m-1}(a)\}$ such that $\sigma(b) \neq b$. Let $l \in \mathbb{N}$, $l > 1$, be the smallest number such that $\sigma^l(b) \in \{b, \sigma(b), \dots, \sigma^{l-1}(b)\}$. Let $\sigma_2 = \langle b, \sigma(b), \dots, \sigma^{l-1}(b)\rangle$.
Continuing in this way, we find a decomposition into cycles: $\sigma = \sigma_1 \circ \sigma_2 \circ \cdots$.

Definition 1.9. Let $n \geq 1$ and $\sigma \in S_n$. We write $\sigma = \sigma_1 \circ \sigma_2 \circ \cdots$ as a composition of cycles of lengths $s_1, s_2, \dots$. Then the number

$$\mathrm{sgn}(\sigma) := (-1)^{(s_1-1)+(s_2-1)+\cdots} \in \{\pm 1\}$$

is called the sign (or signum) of σ.

For example, with σ as in 1.8(a), we have $\mathrm{sgn}(\sigma) = (-1)^{2+3+1+1} = -1$.
We have $\mathrm{sgn}(\mathrm{id}) = 1$, and if τ is any transposition, then $\mathrm{sgn}(\tau) = -1$.

Theorem 1.10. *(a) The definition of* $\mathrm{sgn}(\sigma)$ *does not depend on the chosen cycle decomposition of* σ.

(b) For all $\sigma, \tau \in S_n$ *we have* $\mathrm{sgn}(\sigma \circ \tau) = \mathrm{sgn}(\sigma) \cdot \mathrm{sgn}(\tau)$.

Proof. In Group Theory in Year 2. (But for one possible proof for (a), see also an optional problem on one of the Courseworks.) □

2 Fields and Vector Spaces

2.1 Fields

Definition 2.1. A field is a set F together with binary operations on F, which we will refer to as *addition* and *multiplication*, such that:

- F together with addition is an abelian group (we use the notation $a+b$, 0 or 0_F, $-a$), and
- $F^\times := F \setminus \{0\}$ together with multiplication is an abelian group (we use the notation $a \cdot b$ or ab, 1, a^{-1}),

and such that the following axiom holds:
Distributivity: For all $a,b,c \in F$ we have $a(b+c) = ab+ac$ in F.

Example 2.2. (a) The sets $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ with the usual addition and multiplication are fields (see also 1.2(b)).

(b) The set $\mathbb{Z}$ with the usual addition and multiplication is *not* a field because for instance there is no multiplicative inverse of 2 in $\mathbb{Z}$.

(c) The set $\mathbb{F}_2 := \{0,1\}$ together with the following operations is a field.

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Note that $1+1=2$ in $\mathbb{Q}$ but $1+1=0$ in $\mathbb{F}_2$. $\mathbb{F}_2$ is the smallest field.

Proof (that $\mathbb{F}_2$ is a field): $\mathbb{F}_2$ with "+", and $\mathbb{F}_2 \setminus \{0\} = \{1\}$ with "$\cdot$", are abelian groups (see 1.5(a)).
Distributivity: We need to check $a(b+c) = ab+ac$ for all $a,b,c \in \mathbb{F}_2$.
First case: $a = 0 \implies$ LHS $= 0$, RHS $= 0+0 = 0$.
Second case: $a = 1 \implies$ LHS $= b+c =$ RHS. □

(d) (without proof) Let p be a prime. The set $\mathbb{F}_p := \{\overline{0}, \overline{1}, \ldots, \overline{p-1}\}$ together with the addition defined in Example 1.5(b) and the following multiplication is a field:

$$\overline{x} \cdot \overline{y} := \overline{\text{remainder left when } xy \text{ is divided by } p}.$$

(Why does this not work when p is not a prime, e.g. if $p = 4$?)

Proposition 2.3. *Let F be a field. Then:*

(a) For all $a \in F$ we have $0a = 0$.
(b) For all $a,b \in F$ we have $(-a)b = -(ab)$.

Proof. (a) We have $0+0a = 0a = (0+0)a = 0a+0a$ (0 is neutral for "+" and by distributivity)
$\implies 0 = 0a$. (cancel $0a$ on both sides using 1.4(a))

(b) We have $ab+(-a)b = (a+(-a))b$ (by distributivity)
$= 0b$ (by definition of the additive inverse)
$= 0$ (by part (a))
$\implies (-a)b$ is the additive inverse of ab, i.e. $(-a)b = -(ab)$. □

2.2 Vector spaces

Definition 2.4. Let F be a field. A vector space over F is an abelian group V (we will use "$+$" for the binary operation) together with a map $F \times V \to V$ (called scalar multiplication and written as $(a,x) \mapsto ax$), such that the following axioms are satisfied:

(i) 1^{st} *distributivity law*: For all $a,b \in F$ and $x \in V$ we have $(a+b)x = ax+bx$ in V.
(ii) 2^{nd} *distributivity law*: For all $a,b \in F$ and $x,y \in V$ we have $a(x+y) = ax+ay$ in V.
(iii) For all $a,b \in F$ and for all $x \in V$ we have $(ab)x = a(bx)$ in V.
(iv) For all $x \in V$ we have $1x = x$ in V.

The elements of V are called vectors. The elements of F will be referred to as scalars. We write 0_F and 0_V for the neutral elements of F and V, respectively, and often just 0 for both (when it is clear from the context if it is a scalar or a vector). Furthermore we use the notation $u-v$ for $u+(-v)$ when u,v are both vectors, or both scalars.

Example 2.5. (a) For every $n \in \mathbb{N}$ the set $\mathbb{R}^n$ together with the usual addition and scalar multiplication (as seen in Linear Algebra I) is a vector space over $\mathbb{R}$. Similarly, for any field F, the set

$$F^n := \{(a_1,\dots,a_n) : \ a_1,\dots,a_n \in F\}$$

together with component-wise addition and the obvious scalar multiplication is a vector space over F. For example $\mathbb{F}_2^2 = \{(0,0),(0,1),(1,0),(1,1)\}$ is a vector space over $\mathbb{F}_2$; $F = F^1$ is a vector space over F, and finally $F^0 := \{0\}$ is a vector space over F.

(b) Let V be the additive group of $\mathbb{C}$. We view the usual multiplication $\mathbb{R} \times V \to V$, $(a,x) \mapsto ax$, as scalar multiplication of $\mathbb{R}$ on V. Then V is a vector space over $\mathbb{R}$. Similarly, we can think of $\mathbb{C}$ or $\mathbb{R}$ as vector spaces over $\mathbb{Q}$.

(c) Let V denote the abelian group $\mathbb{R}$ (with the usual addition). For $a \in \mathbb{R}$ and $x \in V$ we put $a \otimes x := a^2 x \in V$; this defines a scalar multiplication

$$\mathbb{R} \times V \to V, \quad (a,x) \mapsto a \otimes x,$$

of the field $\mathbb{R}$ on V. Which of the vector space axioms (see 2.4) hold for V with this scalar multiplication?

Solution:

(i) We need to check whether $(a+b) \otimes x = a \otimes x + b \otimes x$ for all $a,b \in \mathbb{R}$ and $x \in V$.
LHS $= (a+b)^2 x$; RHS $= a^2x + b^2x = (a^2+b^2)x$
$\Longrightarrow$ For $a=1, b=1$ and $x=1$ we have LHS $\neq$ RHS.
$\Longrightarrow$ First distributivity law does not hold.

(ii) We need to check whether $a \otimes (x+y) = a \otimes x + a \otimes y$ for all $a \in \mathbb{R}$ and $x,y \in V$.
$$\left.\begin{aligned} &\text{LHS} = a^2(x+y) \\ &\text{RHS} = a^2x + a^2y = a^2(x+y) \end{aligned}\right\} \Longrightarrow \text{LHS} = \text{RHS}$$
$\Longrightarrow$ Second distributivity law does hold.

(iii) We need to check whether $a \otimes (b \otimes x) = (ab) \otimes x$ for all $a,b \in \mathbb{R}$ and $x \in V$.
$$\left.\begin{aligned} &\text{LHS} = a \otimes (b^2 x) = a^2(b^2 x) \\ &\text{RHS} = (ab)^2 x = (a^2b^2)x = a^2(b^2x) \end{aligned}\right\} \Longrightarrow \text{LHS} = \text{RHS}$$
$\Longrightarrow$ Axiom (iii) does hold.

(iv) We have $1 \otimes x = 1^2 x = x$ for all $x \in V$.
$\Longrightarrow$ Axiom (iv) does hold.

Proposition 2.6. *Let V be a vector space over a field F and let $a, b \in F$ and $x, y \in V$. Then we have:*

(a) $(a-b)x = ax - bx$
(b) $a(x-y) = ax - ay$
(c) $ax = 0_V \iff a = 0_F$ *or* $x = 0_V$
(d) $(-1)x = -x$

Proof: (a) $(a-b)x + bx = ((a-b)+b)x$ (by first distributivity law)
$= (a+((-b)+b))x = (a+0_F)x = ax$ (using field axioms)
$\Longrightarrow (a-b)x = ax - bx.$ (add $-bx$ to both sides)

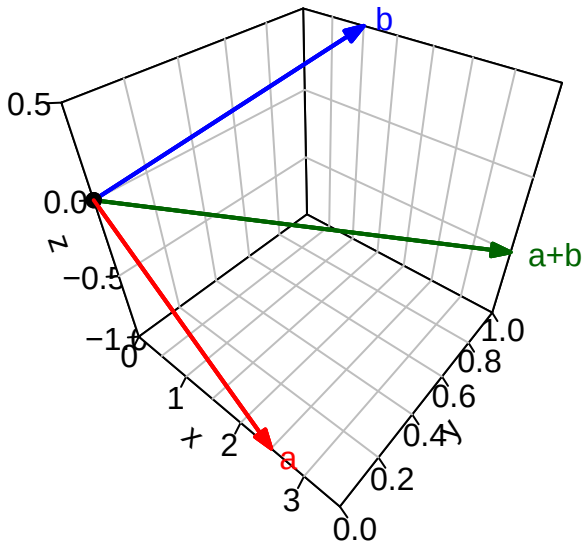
(b) On Coursework.

(c) "$\Longrightarrow$": On Coursework.
"$\Longleftarrow$": Put $a = b$ and $x = y$ in (a) and (b), respectively.

(d) Put $a = 0$ and $b = 1$ in (a) and use (c). □

The next example is the "mother" of almost all vector spaces. It vastly generalises the fourth of the following five ways of representing vectors and vector addition in $\mathbb{R}^3$.

(I)



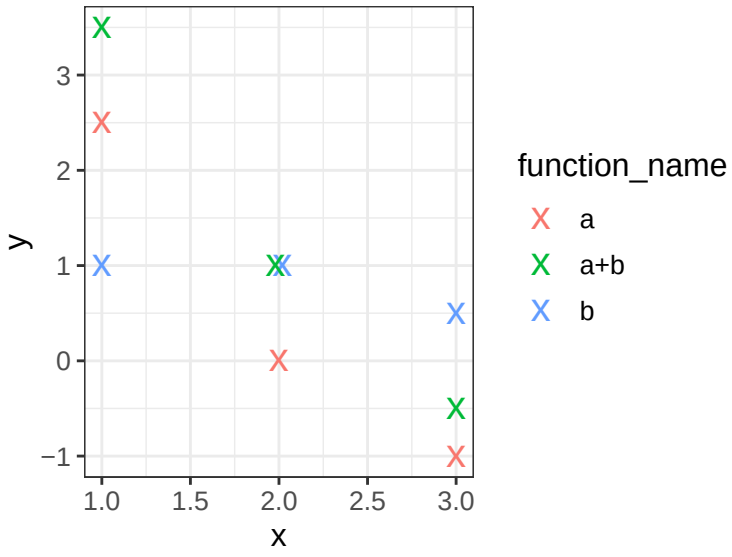
(II) $\underline{a} = (2.5, 0, -1) \qquad \underline{b} = (1, 1, 0.5) \qquad \underline{a} + \underline{b} = (3.5, 1, -0.5)$

(III) $\underline{a} = \begin{pmatrix} 1 & 2 & 3 \\ 2.5 & 0 & -1 \end{pmatrix} \qquad \underline{b} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 0.5 \end{pmatrix} \qquad \underline{a} + \underline{b} = \begin{pmatrix} 1 & 2 & 3 \\ 3.5 & 1 & -0.5 \end{pmatrix}$

(IV)

$$\begin{array}{lll} \underline{a} : \{1,2,3\} \to \mathbb{R} & \underline{b} : \{1,2,3\} \to \mathbb{R} & \underline{a} + \underline{b} : \{1,2,3\} \to \mathbb{R} \\ 1 \mapsto 2.5 & 1 \mapsto 1 & 1 \mapsto 3.5 \\ 2 \mapsto 0 & 2 \mapsto 1 & 2 \mapsto 1 \\ 3 \mapsto -1 & 3 \mapsto 0.5 & 3 \mapsto -0.5 \end{array}$$

(V)



Example 2.7. Let S be any set and let F be a field. Let

$$F^S := \{f : S \to F\}$$

denote the set of all maps from S to F. We define an addition on F^S and a scalar multiplication

of F on F^S as follows: When $f,g \in F^S$ and $a \in F$ we set:

$$(f+g)(s) := f(s)+g(s) \qquad \text{for any } s \in S$$
$$(af)(s) := af(s) \qquad \text{for any } s \in S.$$

Then F^S is a vector space over F (see below for the proof).

Special Cases:

(a) Let $S = \{1,\dots,n\}$. Identifying any map $f : \{1,\dots,n\} \to F$ with the corresponding tuple $(f(1),\dots,f(n))$, we see that F^S can be identified with the set F^n of all n-tuples $(a_1,\dots,a_n)$ considered in Example 2.5(a).

(b) Let $S = \{1,\dots,n\} \times \{1,\dots,m\}$. Identifying any map $f : \{1,\dots,n\} \times \{1,\dots,m\} \to F$ with the corresponding matrix:

$$\begin{pmatrix} f((1,1)) & \dots & f((1,m)) \\ \vdots & & \vdots \\ f((n,1)) & \dots & f((n,m)) \end{pmatrix}$$

we see that F^S can be identified with the set $M_{n\times m}(F)$ of $(n \times m)$-matrices

$$\begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix}$$

with entries in F. In particular $M_{n\times m}(F)$ is a vector space over F.

(c) Let $S = \mathbb{N}$. Identifying any map $f : \mathbb{N} \to F$ with the sequence $(f(1),f(2),f(3),\dots)$ we see that $F^{\mathbb{N}}$ can be identified with the set of all infinite sequences $(a_1,a_2,a_3,\dots)$ in F.

(d) Let $F = \mathbb{R}$ and let S be an interval I in $\mathbb{R}$. Then $F^S = \mathbb{R}^I$ is the set of all functions $f : I \to \mathbb{R}$. (We can visualise these functions via their graph, similarly as in (V) above.)

Proof (that F^S is a vector space over F): First, F^S with the above defined "+" is an abelian group:

Associativity: Let $f,g,h \in F^S$.
We need to show: $(f+g)+h = f+(g+h)$ in F^S
$\iff ((f+g)+h)(s) = (f+(g+h))(s)$ for all $s \in S$.
LHS $= (f+g)(s)+h(s) = (f(s)+g(s))+h(s)$
RHS $= f(s)+(g+h)(s) = f(s)+(g(s)+h(s))$ (by definition of addition in F^S)
$\implies$ LHS = RHS (by associativity in F)

Identity element: Let $\underline{0}$ denote the constant function $S \to F, s \mapsto 0_F$.
For any $f \in F^S$ and $s \in S$ we have $(f+\underline{0})(s) = f(s)+\underline{0}(s) = f(s)+0_F = f(s)$, hence $f+\underline{0} = f$.
Similarly we have $\underline{0}+f = f$. (using definitions of $\underline{0}$ and "+", and field axioms)
$\implies \underline{0}$ is the identity element.

Inverses: Let $f \in F^S$. Define $(-f)(s) := -f(s)$.
For any $s \in S$ we have $(f+(-f))(s) = f(s)+(-f)(s) = f(s)+(-f(s)) = 0_F = \underline{0}(s)$.
$\implies f+(-f) = \underline{0}$ in F^S, so $-f$ is the inverse to f. ($\uparrow$ defns of "+", "$-f$", $\underline{0}$, and field axioms)

Commutativity: Let $f, g \in F^S$.
For any $s \in S$ we have $(f+g)(s) = f(s)+g(s) = g(s)+f(s) = (g+f)(s)$.
$\Longrightarrow$ $f+g = g+f$. ($\uparrow$ by the definition of "+", and commutativity of $+$ in F)

Now the four axioms from Definition 2.4 (only (i) and (iii) spelled out here, the others are similar):

First distributivity law: Let $a, b \in F$ and $f \in F^S$. We want to check that $(a+b)f = af+bf$:
For all $s \in S$ we have

$((a+b)f)(s) = (a+b)(f(s))$ (by definition of the scalar multiplication)
$= a(f(s))+b(f(s))$ (by distributivity in F)
$= (af)(s)+(bf)(s)$ (by definition of the scalar multiplication)
$= (af+bf)(s)$ (by definition of addition in F^S)
$\Longrightarrow$ $(a+b)f = af+bf$.

Axiom (iii): Let $a, b \in F$ and $f \in F^S$. We want to check that $(ab)f = a(bf)$.
For all $s \in S$ we have

$((ab)f)(s) = (ab)(f(s))$ (by definition of scalar multiplication in F^S)
$= a(b(f(s)))$ (by associativity of multiplication in F)
$= a((bf)(s))$ (by definition of scalar multiplication in F^S)
$= (a(bf))(s)$ (by definition of scalar multiplication in F^S)
$\Longrightarrow$ $(ab)f = a(bf)$. □

2.3 Subspaces

Definition 2.8. Let V be a vector space over a field F. A subset W of V is called a subspace of V if the following conditions hold:

(a) $0_V \in W$.
(b) *"W is closed under addition"*: for all $x, y \in W$ we also have $x+y \in W$.
(c) *"W is closed under scalar multiplication"*: for all $a \in F$ and $x \in W$ we have $ax \in W$.

Note that condition (b) states that the restriction of the addition in V to W gives a binary operation $W \times W \to W$ on W (addition in W). Similarly, condition (c) states that the scalar multiplication of F on V yields a map $F \times W \to W$ which we view as a scalar multiplication of F on W.

Proposition 2.9. *Let V be a vector space over a field F and let W be a subspace of V. Then W together with the above mentioned addition and scalar multiplication is a vector space over F.*

Proof: The following axioms hold for W because they already hold for V:

- associativity of addition;
- commutativity of addition;
- all the four axioms in Definition 2.4.

There exists an additive identity element in W by condition 2.8(a) (i.e. $0_W := 0_V \in W$).
It remains to show that additive inverses exist: Let $x \in W$. Then $-x = (-1)x$ (see 2.6(d)) is in W by condition 2.8(c); and $-x$ satisfies $x+(-x) = 0_W = 0_V$ because it does so in V. □

Example 2.10. (a) Examples of subspaces of $\mathbb{R}^n$ as seen in Linear Algebra I, such as the nullspace of any real $(n \times m)$-matrix, or the column space of any real $(m \times n)$-matrix.

(b) The set of convergent sequences is a subspace of the vector space $\mathbb{R}^{\mathbb{N}}$ of all sequences $(a_1, a_2, a_3, \dots)$ in $\mathbb{R}$. A subspace of this subspace (and hence of $\mathbb{R}^{\mathbb{N}}$) is the set of all

sequences in $\mathbb{R}$ that converge to 0. (See Calculus I for proofs).

(c) Let $A \in M_{l\times m}(\mathbb{R})$. Then $W := \{B \in M_{m\times n}(\mathbb{R}) \mid AB = \underline{0}\}$ is a subspace of $M_{m\times n}(\mathbb{R})$.
Proof:

(i) We have $A \cdot \underline{0} = \underline{0} \implies \underline{0} \in W$.
(ii) Let $B_1, B_2 \in W$
$\implies A(B_1 + B_2) = AB_1 + AB_2 = \underline{0} + \underline{0} = \underline{0}$
$\implies B_1 + B_2 \in W$.
(iii) Let $a \in \mathbb{R}$ and $B \in W$
$\implies A(aB) = a(AB) = a\underline{0} = \underline{0}$
$\implies aB \in W$. □

(d) Let I be a non-empty interval in $\mathbb{R}$. The following subsets of the vector space $\mathbb{R}^I$ consisting of all functions from I to $\mathbb{R}$ are subspaces:

(i) For any $s_0 \in I$ the subset $W := \{f \in \mathbb{R}^I : f(s_0) = 0\}$ of $\mathbb{R}^I$.
Proof:
(1) The zero function $\underline{0}$ vanishes at $s_0 \implies \underline{0} \in W$.
(2) Let $f, g \in W$
$\implies (f+g)(s_0) = f(s_0) + g(s_0) = 0 + 0 = 0$
$\implies f + g \in W$.
(3) Let $a \in \mathbb{R}$ and $f \in W$
$\implies (af)(s_0) = a \cdot f(s_0) = a \cdot 0 = 0$
$\implies af \in W$. □
(ii) The set of all continuous functions $f : I \to \mathbb{R}$ (see Calculus I).
(iii) The set of all differentiable functions $f : I \to \mathbb{R}$ (see Calculus I).
(iv) For any $n \in \mathbb{N}$, the set $\mathbb{P}_n$ of polynomial functions $f : I \to \mathbb{R}$ of degree at most n, is a subspace by 3.2(c) and 3.3. A function $f : I \to \mathbb{R}$ is a polynomial function of degree at most n if there exists $a_0, \ldots, a_n \in \mathbb{R}$ such that:

$$f(s) = a_0 + a_1 s + \cdots + a_n s^n \text{ for all } s \in I.$$

Denoting the function $I \to \mathbb{R}, s \mapsto s^m$, by t^m, this means that $f = a_0 t^0 + a_1 t^1 + \cdots + a_n t^n$ as elements of the vector space $\mathbb{R}^I$. (We will also use the more natural notation 1 for t^0, and t for t^1.)
(v) The space of solutions of a homogeneous linear differential equation (without further explanation); e.g.:

$$\mathbb{P}_n = \{f \in \mathbb{R}^I : f \text{ is differentiable } (n+1) \text{ times and } f^{(n+1)} = \underline{0}\}$$

(e) The subset $\mathbb{Z}^n$ of the vector space $\mathbb{R}^n$ over $\mathbb{R}$ is closed under addition but not closed under scalar multiplication: For instance, $(1,0,\ldots,0) \in \mathbb{Z}^n$ and $\frac{1}{2} \in \mathbb{R}$, but $\frac{1}{2}(1,0,\ldots,0) \notin \mathbb{Z}^n$.

(f) The subsets $W_1 := \{(a,0) : a \in \mathbb{R}\}$ and $W_2 := \{(0,b) : b \in \mathbb{R}\}$ are subspaces of $\mathbb{R}^2$. The subset $W := W_1 \cup W_2$ of the vector space $\mathbb{R}^2$ is closed under scalar multiplication but not under addition because, for instance, $(1,0)$ and $(0,1)$ are in W but $(1,0) + (0,1) = (1,1) \notin W$.

Proposition 2.11. *Let W_1, W_2 be subspaces of a vector space V over a field F. Then the intersection $W_1 \cap W_2$ and the sum of subspaces*

$$W_1 + W_2 := \{x_1 + x_2 \in V \mid x_1 \in W_1,\ x_2 \in W_2\}$$

are subspaces of V as well.

Proof: *For $W_1 \cap W_2$:*

(a) We have $0_V \in W_1$ and $0_V \in W_2$ (because W_1 and W_2 are subspaces)
$\implies 0_V \in W_1 \cap W_2$. (by definition of intersection)

(b) Let $x, y \in W_1 \cap W_2$
$\implies x, y \in W_1$ and $x, y \in W_2$ (by definition of intersection)
$\implies x+y \in W_1$ and $x+y \in W_2$ (because W_1 and W_2 are subspaces)
$\implies x+y \in W_1 \cap W_2$. (by definition of intersection)

(c) Let $a \in F$ and $x \in W_1 \cap W_2$
$\implies x \in W_1$ and $x \in W_2$ (by definition of intersection)
$\implies ax \in W_1$ and $ax \in W_2$ (because W_1 and W_2 are subspaces)
$\implies ax \in W_1 \cap W_2$. (by definition of intersection)

For $W_1 + W_2$:

(a) We have $0_V = 0_V + 0_V \in W_1 + W_2$.

(b) Let $x, y \in W_1 + W_2$
$\implies \exists x_1, y_1 \in W_1$ and $\exists x_2, y_2 \in W_2$ with $x = x_1 + x_2$, $y = y_1 + y_2$ (by definition of $W_1 + W_2$)
$\implies x+y = (x_1+x_2)+(y_1+y_2) =$
$= (x_1+y_1)+(x_2+y_2) \in W_1 + W_2$. (because W_1 and W_2 are subspaces)

(c) Let $a \in F$ and $x \in W_1 + W_2$
$\implies \exists x_1 \in W_1$, $x_2 \in W_2$ such that $x = x_1 + x_2$ (by definition of $W_1 + W_2$)
$\implies ax = a(x_1+x_2) = ax_1 + ax_2 \in W_1 + W_2$. (because W_1 and W_2 are subspaces) $\square$

Example 2.12. Let W_1 and W_2 be as in 2.10(f). Then $W_1 + W_2 = \mathbb{R}^2$.

3 Bases

3.1 Spanning

Definition 3.1. Let V be a vector space over a field F. Let $x_1,\dots,x_n \in V$.

(a) An element $x \in V$ is called a linear combination of $x_1,\dots,x_n$ if there are $a_1,\dots,a_n \in F$ such that $x = a_1x_1 + \cdots + a_nx_n$.

(b) The subset of V consisting of all linear combinations of $x_1,\dots,x_n$ is called the span of $x_1,\dots,x_n$ and is denoted by $\mathrm{Span}(x_1,\dots,x_n)$ (or $\mathrm{Span}_F(x_1,\dots,x_n)$); i.e.

$$\mathrm{Span}(x_1,\dots,x_n) = \{a_1x_1 + \cdots + a_nx_n \mid a_1,\dots,a_n \in F\}.$$

(c) We say that $x_1,\dots,x_n$ span V, or that $x_1,\dots x_n$ form a spanning set of V, if $V = \mathrm{Span}(x_1,\dots,x_n)$, i.e. every $x \in V$ is a linear combination of $x_1,\dots,x_n$.

(See also Section 6.3 of L.A.I.)

Example 3.2. (a) Let $V := M_{n\times m}(F)$ be the vector space of $(n\times m)$-matrices with entries in F. For $i \in \{1,\dots,n\}$ and $j \in \{1,\dots,m\}$, let E_{ij} denote the $(n\times m)$-matrix with zeroes everywhere except at (ij) where it has the entry 1. Then the matrices $E_{ij}; i = 1,\dots,n; j = 1,\dots,m$ form a spanning set of V.
Proof: Let $A = (a_{ij}) \in M_{n\times m}(F)$ be an arbitrary matrix. Then $A = \sum_{i=1}^n \sum_{j=1}^m a_{ij}E_{ij}$.
For example: $\left(\begin{smallmatrix} 2 & 3 \\ -1 & 5 \end{smallmatrix}\right) = 2\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right) + 3\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) + (-1)\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right) + 5\left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right) = 2E_{11} + 3E_{12} + (-1)E_{21} + 5E_{22}$.

(b) Do the vectors $\begin{pmatrix} 1 \\ i \end{pmatrix}, \begin{pmatrix} i \\ 2 \end{pmatrix} \in \mathbb{C}^2$ span the vector space $\mathbb{C}^2$ over $\mathbb{C}$?
Solution: Let $\begin{pmatrix} w \\ z \end{pmatrix} \in \mathbb{C}^2$ be an arbitrary vector.
We want to know if we can find $a_1, a_2 \in \mathbb{C}$ such that $a_1\begin{pmatrix} 1 \\ i \end{pmatrix} + a_2\begin{pmatrix} i \\ 2 \end{pmatrix} = \begin{pmatrix} w \\ z \end{pmatrix}$.
Hence $\left(\begin{array}{cc|c} 1 & i & w \\ i & 2 & z \end{array}\right) \xrightarrow{R2\mapsto R2-iR1} \left(\begin{array}{cc|c} 1 & i & w \\ 0 & 3 & z-iw \end{array}\right)$.
As in Linear Algebra I we conclude that this system is solvable. (Theorem 3.15 of L.A.I.)
Thus $\begin{pmatrix} 1 \\ i \end{pmatrix}, \begin{pmatrix} i \\ 2 \end{pmatrix}$ span $\mathbb{C}^2$.

(c) $\mathbb{P}_n = \mathrm{Span}(t^0, t^1, t^2, \dots, t^n) = \mathrm{Span}(1, t, t^2, \dots, t^n)$ (c.f. 2.10(d)(iv)).

Proposition 3.3. *Let V be a vector space over a field F. Let $x_1,\dots,x_n \in V$. Then* $\mathrm{Span}(x_1,\dots,x_n)$ *is the smallest subspace of V that contains $x_1,\dots,x_n$. Furthermore:*

(a) If $x \in \mathrm{Span}(x_1,\dots,x_n)$ then $\mathrm{Span}(x_1,\dots,x_n,x) = \mathrm{Span}(x_1,\dots,x_n)$.

(b) For any $a_2,\dots,a_n \in F$ we have $\mathrm{Span}(x_1,\dots,x_n) = \mathrm{Span}(x_1, x_2 - a_2x_1, \dots, x_n - a_nx_1)$.

Notes:

- The first statement means:
  1. $\mathrm{Span}(x_1,\dots,x_n)$ is a subspace of V that contains $x_1,\dots,x_n$, *and*
  2. among all the subspaces of V with this property it is the smallest;
     in other words if W is a subspace of V that contains $x_1,\dots,x_n$, then $\mathrm{Span}(x_1,\dots x_n) \subseteq W$.

- The statement (b) implies that the span of the column vectors of any matrix does not change when performing (standard) column operations.

Proof: $\mathrm{Span}(x_1,\dots,x_n)$ *is a subspace of* V:

(i) We have $0_V = 0_F x_1 + \cdots + 0_F x_n \in \mathrm{Span}(x_1,\dots,x_n)$.

(ii) Let $x, y \in \mathrm{Span}(x_1,\dots,x_n)$.
$\implies \exists a_1,\dots,a_n \in F,\ \exists b_1,\dots b_n \in F$ such that $x = a_1x_1 + \cdots + a_nx_n$ and $y = b_1x_1 + \cdots + b_nx_n$;
$\implies x + y = (a_1x_1 + \cdots + a_nx_n) + (b_1x_1 + \cdots + b_nx_n)$
$= (a_1x_1 + b_1x_1) + \cdots + (a_nx_n + b_nx_n)$ (using commutativity and associativity)
$= (a_1 + b_1)x_1 + \cdots + (a_n + b_n)x_n$ (using first distributivity law)
$\in \mathrm{Span}(x_1,\dots,x_n)$ (by definition of Span)

(iii) Let $x \in \mathrm{Span}(x_1,\dots,x_n)$ and $a \in F$.
Write $x = a_1x_1 + \dots a_nx_n$ with $a_1,\dots,a_n \in F$ as above.
$\implies ax = a(a_1x_1 + \cdots + a_nx_n) \quad = a(a_1x_1) + \cdots + a(a_nx_n)$ (using distributivity)
$= (aa_1)x_1 + \cdots + (aa_n)x_n$ (by axiom 2.4(iii))
$\in \mathrm{Span}(x_1,\dots,x_n)$ (by definition of Span)

$\mathrm{Span}(x_1,\dots,x_n)$ *contains* $x_1,\dots,x_n$:
$\dots$ because $x_i = 0_F \cdot x_1 + \dots 0_F \cdot x_{i-1} + 1 \cdot x_i + 0_F \cdot x_{i+1} + \cdots + 0_F \cdot x_n \in \mathrm{Span}(x_1,\dots,x_n)$.

$\mathrm{Span}(x_1,\dots,x_n)$ *is the smallest*:
Let W be a subspace of V such that $x_1,\dots,x_n \in W$.
Let $x \in \mathrm{Span}(x_1,\dots,x_n)$. Write $x = a_1x_1 + \cdots + a_nx_n$ with $a_1,\dots a_n \in F$.
$\implies a_1x_1,\dots,a_nx_n \in W$ (by condition 2.8(c))
$\implies x = a_1x_1 + \cdots + a_nx_n \in W$. (by condition 2.8(b))
$\implies \mathrm{Span}(x_1,\dots,x_n) \subseteq W$.

Part (a): $\mathrm{Span}(x_1,\dots,x_n) \subseteq \mathrm{Span}(x_1,\dots,x_n,x) =: W$
(because W is a subspace of V and $x_1,\dots,x_n \in W$)
$\mathrm{Span}(x_1,\dots,x_n,x) \subseteq \mathrm{Span}(x_1,\dots,x_n) =: \widetilde{W}$
(because $\widetilde{W}$ is a subspace of V and $x_1,\dots,x_n,x \in \widetilde{W}$)

Part (b): $\mathrm{Span}(x_1,\dots,x_n) \subseteq \mathrm{Span}(x_1, x_2 - a_2x_1,\dots,x_n - a_nx_1) =: W$
(because W is a subspace of V and $x_1,\dots,x_n \in W$)
$\mathrm{Span}(x_1, x_2 - a_2x_1,\dots,x_n - a_nx_1) \subseteq \mathrm{Span}(x_1,\dots,x_n) =: \widetilde{W}$
(because $\widetilde{W}$ is a subspace of V and $x_1 \in \widetilde{W}$ and for $i = 2,\dots,n$ also $x_i - a_ix_1 \in \widetilde{W}$)$\square$

3.2 Linear independence

Definition 3.4. Let V be a vector space over a field F. Let $x_1,\dots,x_n \in V$. We say that $x_1,\dots,x_n$ are linearly independent (over F) if the following condition holds:

if $a_1,\dots,a_n \in F$ and $a_1x_1 + \cdots + a_nx_n = 0_V$ then $a_1 = \cdots = a_n = 0_F$.

Otherwise we say that $x_1,\dots,x_n$ are linearly dependent. A linear combination $a_1x_1 + \cdots + a_nx_n$ is called trivial if $a_1 = \cdots = a_n = 0$, otherwise it is called non-trivial. (See also Section 6.5 of L.A.I.)

Note: $x_1,\dots,x_n$ are linearly dependent $\iff \exists a_1,\dots,a_n \in F$, not all zero, such that $a_1x_1 + \cdots + a_nx_n = 0_V$. In other words, $\iff$ there exists a non-trivial linear combination of $x_1,\dots,x_n$ which equals 0_V.

Example 3.5. (a) Examples as seen in Linear Algebra I. (Section 6.5 of L.A.I.)

(b) The three vectors $\underline{x}_1 = \begin{pmatrix}1\\0\\1\end{pmatrix}, \underline{x}_2 = \begin{pmatrix}1\\1\\0\end{pmatrix}, \underline{x}_3 = \begin{pmatrix}0\\-1\\1\end{pmatrix} \in F^3$ are *not* linearly independent because $\underline{x}_1 - \underline{x}_2 - \underline{x}_3 = \underline{0}$.

(c) Determine all vectors $\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix} \in \mathbb{C}^3$ such that $\underline{x}_1 := \begin{pmatrix}1\\i\\1\end{pmatrix}, \underline{x}_2 := \begin{pmatrix}0\\1\\0\end{pmatrix}, \underline{x}_3 := \begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix} \in \mathbb{C}^3$ are linearly dependent.
Solution: We apply Gaussian elimination:

$$\begin{pmatrix}1 & 0 & c_1\\ i & 1 & c_2\\ 1 & 0 & c_3\end{pmatrix} \xrightarrow[R2\mapsto R2-iR1]{R3\mapsto R3-R1} \begin{pmatrix}1 & 0 & c_1\\ 0 & 1 & c_2-ic_1\\ 0 & 0 & c_3-c_1\end{pmatrix}$$

$\Longrightarrow$ The equation $a_1\underline{x}_1 + a_2\underline{x}_2 + a_3\underline{x}_3 = \underline{0}$ has a non-trivial solution (a_1,a_2,a_3) if and only if $c_3 - c_1 = 0$.
$\Longrightarrow$ $\underline{x}_1, \underline{x}_2, \underline{x}_3$ are linearly dependent if and only if $c_1 = c_3$.

(d) The two functions $\sin, \cos \in \mathbb{R}^{\mathbb{R}}$ are linearly independent.
Proof: Let $a,b \in \mathbb{R}$ such that $a\sin + b\cos = \underline{0}$ in $\mathbb{R}^{\mathbb{R}}$.
$\Longrightarrow$ For all $s \in \mathbb{R}$ we have $a\cdot\sin(s) + b\cdot\cos(s) = \underline{0}(s) = 0$.
$\Longrightarrow$ $\begin{cases} a\cdot\sin(0) + b\cdot\cos(0) = 0 \Longrightarrow a\cdot 0 + b\cdot 1 = 0 \Longrightarrow b = 0 \\ a\cdot\sin(\frac{\pi}{2}) + b\cdot\cos(\frac{\pi}{2}) = 0 \Longrightarrow a\cdot 1 + b\cdot 0 = 0 \Longrightarrow a = 0 \end{cases}$ □

(e) Let $I \subseteq \mathbb{R}$ be a non-empty open interval. Recall from 2.10(d)(iv) that for any $i \in \mathbb{N}_0$, we denote t^i the polynomial function $I \to \mathbb{R}, s \mapsto s^i$.
The vectors $t^0, t^1, t^2, \ldots, t^n$ are linearly independent in $\mathbb{R}^I$.
Proof: Let $a_0, \ldots, a_n \in \mathbb{R}$ such that $a_0t^0 + a_1t^1 + \cdots + a_nt^n = \underline{0}$
$\Longrightarrow$ $a_0 + a_1s + \cdots + a_ns^n = 0$ for all $s \in I$
$\Longrightarrow$ $a_0 = \cdots = a_n = 0$, because any non-zero real polynomial of degree n has at most n real roots. (This follows from the Fundamental Theorem of Algebra. Alternatively, it can be proved by induction and using long division.)

Proposition 3.6. *Let V be a vector space over a field F.*

(a) A single vector $x \in V$ is linearly independent if and only if $x \neq 0_V$.

(b) Every subset of any set of linearly independent vectors is linearly independent again. (This is equivalent to: If a subset of a set of vectors is linearly dependent, then the set itself is linearly dependent.)

(c) Let $x_1, \ldots, x_n \in V$ and suppose that $x_i = 0_V$ for some $i \in \{1, \ldots, n\}$, or that $x_i = x_j$ for some $i \neq j$. Then $x_1, \ldots, x_n$ are linearly dependent.

(d) If $x_1, \ldots, x_n \in V$ are linearly dependent then at least one vector x_i among $x_1, \ldots, x_n$ is a linear combination of the other ones.

(e) Let $x_1, \ldots, x_n \in V$ and $x \in \mathrm{Span}(x_1, \ldots, x_n)$. Then $x_1, \ldots, x_n, x$ are linearly dependent.

Proof:

(a) "⇒": If $x = 0_V$ then $1 \cdot x = 0_V$ is a non-trivial linear combination of x.
"⇐": Let $x \neq 0_V$ and let $a \in F$ such that $ax = 0_V$. Then $a = 0$ by 2.6(c).

(b) Let $x_1, \ldots, x_n \in V$ be linearly independent and let $y_1, \ldots, y_m$ be a subset of $x_1, \ldots, x_n$ for some $m \leq n$.
Proceed by contradiction: Suppose that $\exists b_1, \ldots, b_m \in F$, not all zero, such that $b_1 y_1 + \cdots + b_m y_m = 0_V$.
Extending the list $b_1, \ldots, b_m$ by zeroes we get $a_1, \ldots, a_n \in F$, not all zero, such that $a_1 x_1 + \cdots + a_n x_n = 0_V$.
So $x_1, \ldots, x_n$ are linearly dependent, a contradiction.

(c) If $x_i = 0_V$ then $x_1, \ldots, x_n$ are linearly dependent by (a) and (b).
If $x_i = x_j$ for some $i \neq j$, then x_i, x_j are linearly dependent, because $1x_i + (-1)x_j = 0_V$.
Now apply (b).

(d) $\exists a_1, \ldots, a_n \in F$, not all zero, such that $a_1 x_1 + \cdots + a_n x_n = 0_V$.
After reordering we may assume $a_n \neq 0$.
$\implies x_n = -a_n^{-1}(a_1 x_1 + \cdots + a_n x_n) = (-a_n^{-1} a_1)x_1 + \cdots + (-a_n^{-1} a_{n-1})x_{n-1}$
$\implies x_n$ is a linear combination of $x_1, \ldots, x_{n-1}$.

(e) Let $x \in \mathrm{Span}(x_1, \ldots, x_n)$.
$\implies \exists a_1, \ldots, a_n \in F$, such that $x = a_1 x_1 + \cdots + a_n x_n$.
$\implies 0_V = a_1 x_1 + \cdots + a_n x_n + (-1)x$
$\implies$ We have a a non-trivial linear combination of $x_1, \ldots, x_n, x$ which equals 0_V.
(Because $-1 \neq 0$ in any field.) □

3.3 Bases

Definition 3.7. Let V be a vector space over a field F. Let $x_1, \ldots, x_n \in V$. We say that $x_1, \ldots, x_n$ form a basis of V if $x_1, \ldots, x_n$ both span V and are linearly independent. (Compare Definition 6.40 of L.A.I.)

Example 3.8. (a) Let F be a field. The vectors $\underline{e}_1 := (1, 0, \ldots, 0)$; … ; $\underline{e}_n := (0, \ldots, 0, 1)$ form a basis F^n, called the standard basis of F^n (as in Linear Algebra I (Ex 6.41(a)).

(b) The polynomials $1, t, \ldots, t^n$ form a basis of $\mathbb{P}_n$. (see 2.10(d)(iv), 3.2(c)).

(c) $1, i$ form a basis of the vector space $\mathbb{C}$ over $\mathbb{R}$ (cf 2.5(b)).

(d) Determine a basis of the nullspace $N(A) \subseteq \mathbb{R}^4$ of the matrix

$$A := \begin{pmatrix} 1 & -1 & 3 & 2 \\ 2 & -1 & 6 & 7 \\ 3 & -2 & 9 & 9 \\ -2 & 0 & -6 & -10 \end{pmatrix} \in M_{4\times 4}(\mathbb{R}).$$

Solution: We perform Gaussian elimination until the reduced lower echelon form (row

operations):

$$A = \begin{pmatrix} 1 & -1 & 3 & 2 \\ 2 & -1 & 6 & 7 \\ 3 & -2 & 9 & 9 \\ -2 & 0 & -6 & -10 \end{pmatrix} \xrightarrow[\substack{R3\mapsto R3-3R1 \\ R4\mapsto R4+2R1}]{R2\mapsto R2-2R1} \begin{pmatrix} 1 & -1 & 3 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 1 & 0 & 3 \\ 0 & -2 & 0 & -6 \end{pmatrix}$$

$$\xrightarrow[\substack{R3\mapsto R3-R2 \\ R4\mapsto R4+2R2}]{R1\mapsto R1+R2} \begin{pmatrix} 1 & 0 & 3 & 5 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} =: \widetilde{A}$$

Because performing row operations does not change the nullspace of a matrix (see Note below), we have:

$$\begin{aligned} N(A) = N(\widetilde{A}) &= \{\underline{x} \in \mathbb{R}^4 : \widetilde{A}\underline{x} = \underline{0}\} \\ &= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{R}^4 : x_1 = -3x_3 - 5x_4;\ x_2 = -3x_4 \right\} \\ &= \left\{ \begin{pmatrix} -3x_3 - 5x_4 \\ -3x_4 \\ x_3 \\ x_4 \end{pmatrix} : x_3, x_4 \in \mathbb{R} \right\} \\ &= \left\{ x_3 \begin{pmatrix} -3 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -5 \\ -3 \\ 0 \\ 1 \end{pmatrix} : x_3, x_4 \in \mathbb{R} \right\} \\ &= \text{Span} \left(\begin{pmatrix} -3 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -5 \\ -3 \\ 0 \\ 1 \end{pmatrix} \right). \end{aligned}$$

(Denote $\underline{u}_1 = (-3,0,1,0)^T$, $\underline{u}_2 = (-5,-3,0,1)^T$).
• $\underline{u}_1$ and $\underline{u}_2$ are also linearly independent (because they are not multiples of each other).
$\Longrightarrow$ $\underline{u}_1$ and $\underline{u}_2$ form a basis of $N(A)$.

(e) Let $A \in M_{n\times n}(\mathbb{R})$. Then:
A invertible $\iff$ The columns of A form a basis of $\mathbb{R}^n$ (for a proof see Theorem 5.6).

Note: Row operations do not change the nullspace of a matrix (say A). This is because vectors $\underline{x}$ are in $N(A)$ exactly if they are solutions to $A\underline{x} = \underline{0}$, i.e. solutions to the homogeneous system of linear equations described by the matrix A. Row operations on A correspond to the "allowed" operations on the linear system of equations.

Proposition 3.9. *Let V be a vector space over a field F. Let $x_1,\ldots,x_n \in V$. The following statements are equivalent:*

(a) $x_1,\ldots,x_n$ form a basis of V.
(b) $x_1,\ldots,x_n$ form a minimal spanning set of V (i.e. $x_1,\ldots,x_n$ span V and after removing any vector from $x_1,\ldots,x_n$ the remaining ones don't span V anymore). (Compare Def 6.23 of L.A.I.)

(c) $x_1,\dots,x_n$ *form a maximal linearly independent subset of* V *(i.e.* $x_1,\dots,x_n$ *are linearly independent and for any* $x \in V$ *the* $n+1$ *vectors* $x_1,\dots,x_n,x$ *are linearly dependent).*

(d) *Every vector* $x \in V$ *can be written in the form*

$$x = a_1x_1 + \cdots + a_nx_n$$

with coefficients $a_1,\dots,a_n \in F$ *uniquely determined by* x.

Proof: "*(a)* $\Longrightarrow$ *(b)*": 1/*Spanning*: $x_1,\dots,x_n$ span V by definition of a basis.
2/*Minimality*: Suppose that the spanning set $x_1,\dots,x_n$ is not minimal.
$\Longrightarrow$ After reordering we may assume that $x_1,\dots,x_{n-1}$ span V.
$\Longrightarrow$ $x_n \in V = \mathrm{Span}(x_1,\dots,x_{n-1})$.
$\Longrightarrow$ $x_1,\dots,x_{n-1},x_n$ are linearly dependent (by 3.6(e)). Contradiction.

(b) $\Longrightarrow$ *(c)*: 1/*Independence*: Suppose that $x_1,\dots,x_n$ are linearly dependent.
$\Longrightarrow$ After reordering we have $x_n \in \mathrm{Span}(x_1,\dots,x_{n-1})$ (by 3.6(d))
$\Longrightarrow$ $V = \mathrm{Span}(x_1,\dots,x_n) = \mathrm{Span}(x_1,\dots,x_{n-1})$ (by 3.3(a))
This contradicts the minimality assumed in (b).
2/*Maximality*: Let $x \in V = \mathrm{Span}(x_1,\dots,x_n)$
$\Longrightarrow$ $x_1,\dots,x_n,x$ are linearly dependent (by 3.6(c)).

(c) $\Longrightarrow$ *(d)*: 1/*Existence*: Let $x \in V$.
$\Longrightarrow$ $\exists b_1,\dots,b_n,b \in F$, not all zero, with $b_1x_1 + \cdots + b_nx_n + bx = 0$.
(because $x_1,\dots,x_n,x$ are linearly dependent)
$\Longrightarrow$ $b \neq 0$ (because $x_1,\dots,x_n$ are linearly independent)
$\Longrightarrow$ $x = a_1x_1 + \cdots + a_nx_n$, where $a_i := -b^{-1}b_i$.
2/*Uniqueness*: Suppose $x = a_1x_1 + \cdots + a_nx_n = b_1x_1 + \cdots + b_nx_n$ for some $a_1,\dots,a_n,b_1,\dots,b_n \in F$.
$\Longrightarrow$ $0_V = x - x = (a_1 - b_1)x_1 + \cdots + (a_n - b_n)x_n$
$\Longrightarrow$ $a_1 = b_1,\dots,a_n = b_n$. (because $x_1,\dots,x_n$ are linearly independent)

(d) $\Longrightarrow$ *(a)*: 1/*Spanning*: Directly from (d).
2/*Independence*: Let $a_1,\dots,a_n \in F$ such that $a_1x_1 + \dots a_nx_n = 0_V$.
$\Longrightarrow$ $a_1 = \cdots = a_n = 0$. (from uniqueness, because also $0x_1 + \cdots + 0x_n = 0_V$) □

Corollary 3.10. *Let* V *be a vector space over a field* F. *Suppose* $V = \mathrm{Span}(x_1,\dots,x_n)$ *for some* $x_1,\dots x_n \in V$. *Then a subset of* $x_1,\dots,x_n$ *forms a basis of* V. *In particular* V *has a basis.*

Proof: By successively removing vectors from $x_1,\dots,x_n$ we arrive at a minimal spanning set $y_1,\dots,y_m$ for some $m \leq n$. Then $y_1,\dots,y_m$ form a basis of V (by 3.9 (b) $\Longrightarrow$ (a)). □

3.4 Dimension

Theorem 3.11. *(This Theorem allows us to define the dimension of a vector space.)*
Let V *be a vector space over a field* F. *Suppose* $x_1,\dots,x_n$ *and* $y_1,\dots,y_m$ *both form a basis of* V. *Then* $m = n$. *(Compare Thm 6.44 from L.A.I.)*

Definition 3.12. Let V be a vector space over a field F. If the vectors $x_1,\dots,x_n \in V$ form a basis of V, we say that V is of finite dimension, and call n the dimension of V. We write $\dim_F(V)$ or just $\dim(V)$ for n. Note that n does not depend on the chosen basis $x_1,\dots,x_n$ by 3.11.

Example 3.13. (a) $\dim_F(F^n) = n$ (by Example 3.8(a)).
(b) $\dim_{\mathbb{R}}(\mathbb{P}_n) = n+1$ (by Example 3.8(b)).

(c) $\dim_{\mathbb{R}}(\mathbb{C}) = 2$ (by Example 3.8(c)).
(d) $\dim_{\mathbb{C}}(\mathbb{C}^3) = 3$, $\dim_{\mathbb{R}}(\mathbb{C}^3) = 6$. In general $\dim_{\mathbb{C}}(\mathbb{C}^n) = n$, $\dim_{\mathbb{R}}(\mathbb{C}^n) = 2n$.
(e) $\mathbb{R}$ as a vector space over $\mathbb{Q}$ is *not* finite dimensional (see 2.5(b)).
(f) $\dim_{\mathbb{R}}(M_{n\times m}(\mathbb{R})) = nm$.
(g) $\dim_{\mathbb{Q}}(\mathbb{Q}(\sqrt{-3})) = 2$ (see Coursework 2).
(h) About $\dim_F(\mathrm{Span}(x_1,\dots,x_n))$. We determine its dimension by finding a basis, i.e. subset of $x_1,\dots,x_n$ which still spans $\mathrm{Span}(x_1,\dots,x_n)$, and it is linearly independent. For example:
Let $x_1 \in V$, $x_1 \neq 0 \implies \dim_F(\mathrm{Span}(x_1)) = 1$.
Let x_2 be another vector in V.
$$\implies \dim_F(\mathrm{Span}(x_1,x_2)) = \begin{cases} 2 & \text{if } x_1,x_2 \text{ are linearly independent} \\ 1 & \text{if } x_1,x_2 \text{ are linearly dependent} \end{cases}$$

Proof of Theorem 3.11: It follows from the following Proposition. □

Proposition 3.14. *Let V be a vector space over a field F. Let $x_1,\dots,x_n \in V$ and $y_1,\dots,y_m \in V$. Suppose $x_1,\dots,x_n$ span V. If $y_1,\dots,y_m$ are linearly independent then $m \leq n$.*

Proof: We will show the contrapositive:

- If $m > n$ then there exist $c_1,\dots,c_m \in F$, not all zero, such that $c_1y_1 + \cdots + c_my_m = 0$.

For every $i \in \{1,\dots,m\}$ we can write $y_i = a_{i1}x_1 + \cdots + a_{in}x_n$ for some $a_{i1},\dots a_{in} \in F$.
(because $x_1,\dots,x_n$ span V)
$\implies$ For all $c_1,\dots,c_m \in F$ we have: (using the axioms of a vector space)

$$\begin{aligned} c_1y_1 + \cdots + c_my_m &= c_1(a_{11}x_1 + \cdots + a_{1n}x_n) + \cdots + c_m(a_{m1}x_1 + \cdots + a_{mn}x_n) \\ &= (a_{11}c_1 + \cdots + a_{m1}c_m)x_1 + \cdots + (a_{1n}c_1 + \cdots + a_{mn}c_m)x_n \end{aligned}$$

$\implies$ It suffices to show that the system of linear equations

$$\begin{aligned} a_{11}c_1 + a_{21}c_2 + \cdots + a_{m1}c_m &= 0 \\ &\vdots \\ a_{1n}c_1 + a_{2n}c_2 + \cdots + a_{mn}c_m &= 0 \end{aligned}$$

has a solution $(c_1,\dots,c_m) \in F^m$ different from $(0,\dots,0)$.
This follows from Gaussian elimination (as seen in Linear Algebra I for $F = \mathbb{R}$ (Thm 3.15(b) from L.A.I.)): since $m > n$, we have more unknowns than equations. □

Corollary 3.15. *(Two-out-of-three basis criterion.)*
Let V be a vector space over a field F. Let $x_1,\dots x_n \in V$. Suppose two of the following three statements hold. Then $x_1,\dots,x_n$ form a basis:

(a) $x_1,\dots,x_n$ are linearly independent.
(b) $x_1,\dots,x_n$ span V.
(c) $n = \dim_F(V)$.

Proof: If (a) and (b) hold, then $x_1,\dots,x_n$ form a basis by definition.
Suppose (a) and (c) hold. If $x_1,\dots,x_n$ would not form a basis we could find an $x \in V$ such that $x_1,\dots,x_n,x$ are still linearly independent (by 3.9 (a) $\iff$ (c)).
$\implies n+1 \leq n$ by Proposition 3.14. Contradiction.
Suppose (b) and (c) hold. After reordering we may assume that $x_1,\dots,x_m$ form a minimal spanning set, for some $m \leq n$.

$\Longrightarrow$ $x_1,\ldots,x_m$ form a basis (by 3.9 (a) $\Longleftrightarrow$ (b)).
$\Longrightarrow$ $m=n$ (by 3.11); i.e. $x_1,\ldots,x_n$ form a basis. $\square$

Example 3.16. (a) The vectors $\underline{x}_1 := \begin{pmatrix}1\\2\\3\end{pmatrix}$, $\underline{x}_2 := \begin{pmatrix}-2\\1\\0\end{pmatrix}$, $\underline{x}_3 := \begin{pmatrix}1\\0\\1\end{pmatrix}$ form a basis of the vector space $\mathbb{Q}^3$ over $\mathbb{Q}$, of the vector space $\mathbb{R}^3$ over $\mathbb{R}$ and of the vector space $\mathbb{C}^3$ over $\mathbb{C}$.

Proof: We first show that $\underline{x}_1,\underline{x}_2,\underline{x}_3$ are linearly independent over $\mathbb{C}$:
Let $c_1,c_2,c_3\in\mathbb{C}$ such that $c_1\underline{x}_1+c_2\underline{x}_2+c_3\underline{x}_3=0$.

$$\Longrightarrow \quad \begin{pmatrix}1&-2&1\\2&1&0\\3&0&1\end{pmatrix}\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}.$$

Gaussian elimination yields:

$$A:=\begin{pmatrix}1&-2&1\\2&1&0\\3&0&1\end{pmatrix}\xrightarrow[R3\mapsto R3-3R1]{R2\mapsto R2-2R1}\begin{pmatrix}1&-2&1\\0&5&-2\\0&6&-2\end{pmatrix}\xrightarrow{R3\mapsto R3-\frac{6}{5}R2}\begin{pmatrix}1&-2&1\\0&5&-2\\0&0&2/5\end{pmatrix}$$

$\Longrightarrow$ $\underline{c}=(c_1,c_2,c_3)=(0,0,0)$ is the only solution of $A\underline{c}=0$.
$\Longrightarrow$ $\underline{x}_1,\underline{x}_2,\underline{x}_3$ are linearly independent over $\mathbb{C}$ and then also over $\mathbb{R}$ and $\mathbb{Q}$.
$\Longrightarrow$ $\underline{x}_1,\underline{x}_2,\underline{x}_3$ form a basis of $\mathbb{C}^3$, $\mathbb{R}^3$ and $\mathbb{Q}^3$ over the respective fields.
(by 3.15 since $3=\dim_{\mathbb{C}}(\mathbb{C}^3)=\dim_{\mathbb{R}}(\mathbb{R}^3)=\dim_{\mathbb{Q}}(\mathbb{Q}^3)$) $\square$

(b) We view $\mathbb{C}$ as a vector space over $\mathbb{C}$, $\mathbb{R}$ and $\mathbb{Q}$ (see 2.5(b)).
Let $x_1:=1,x_2:=2,x_3:=\sqrt{2},x_4:=i,x_5:=i\sqrt{3}\in\mathbb{C}$.
Determine $\dim_F(\mathrm{Span}_F(x_1,x_2,x_3,x_4,x_5))$ for $F=\mathbb{C}$, $\mathbb{R}$ and $\mathbb{Q}$.

Solution: *For $F=\mathbb{C}$*:
We have $\mathbb{C}=\mathrm{Span}_{\mathbb{C}}(x_1)\subseteq\mathrm{Span}_{\mathbb{C}}(x_1,\ldots,x_5)\subseteq\mathbb{C}$.
$\Longrightarrow$ $\dim_{\mathbb{C}}(\mathrm{Span}_{\mathbb{C}}(x_1,\ldots,x_5))=\dim_{\mathbb{C}}(\mathbb{C})=1$.

For $F=\mathbb{R}$:
x_1 and x_4 span $\mathbb{C}$ as a vector space over $\mathbb{R}$.
$\Longrightarrow$ $\mathrm{Span}_{\mathbb{R}}(x_1,\ldots,x_5)=\mathbb{C}$.
$\Longrightarrow$ $\dim_{\mathbb{R}}(\mathrm{Span}(x_1,\ldots,x_5))=\dim_{\mathbb{R}}(\mathbb{C})=2$.

For $F=\mathbb{Q}$: Observations:
x_1,x_2 are LD over $\mathbb{Q}$ $\Longrightarrow$ $\mathrm{Span}_{\mathbb{Q}}(x_1,\ldots,x_5)=\mathrm{Span}_{\mathbb{Q}}(x_1,x_3,x_4,x_5)$.
Also x_1,x_3 are LI over $\mathbb{Q}$; x_1,x_4 are LI over $\mathbb{R}$.
$\rightsquigarrow$ Let us try to prove that x_1,x_3,x_4,x_5 are linearly independent over $\mathbb{Q}$.
Let $a_1,a_3,a_4,a_5\in\mathbb{Q}$ be such that $a_1x_1+a_3x_3+a_4x_4+a_5x_5=0$.
$\Longrightarrow$ $(a_1+a_3\sqrt{2})+i(a_4+a_5\sqrt{3})=0$.
$\Longrightarrow$ $a_1+a_3\sqrt{2}=0$ and $a_4+a_5\sqrt{3}=0$ (because 1 and i are linearly independent over $\mathbb{Q}$)
$\Longrightarrow$ $a_1=a_3=0$ (If $a_3\neq 0$ $\Longrightarrow$ $\sqrt{2}=-\frac{a_1}{a_3}\in\mathbb{Q}$. Contradiction.)
and $a_4=a_5=0$ (similarly).
$\Longrightarrow$ x_1,x_3,x_4,x_5 are linearly independent, so form a basis of $\mathrm{Span}_{\mathbb{Q}}(x_1,x_3,x_4,x_5)$.
$\Longrightarrow$ $\dim_{\mathbb{Q}}(\mathrm{Span}_{\mathbb{Q}}(x_1,x_2,x_3,x_4,x_5))=\dim_{\mathbb{Q}}(\mathrm{Span}_{\mathbb{Q}}(x_1,x_3,x_4,x_5))=4$.

(c) Let V be a vector space of finite dimension over a field F and let W be a subspace of V. Then $\dim_F(W) \leq \dim_F(V)$.

Proof: If vectors are L.I. in W, they are also L.I. in V. (by def. of L.I.)
$\Longrightarrow$ Any L.I. subset of W has at most $\dim(V)$ elements (use 3.14 and 3.9(a) $\Longleftrightarrow$ (c))
$\Longrightarrow$ $\dim(W) \leq \dim(V)$ (by 3.9(a) $\Longleftrightarrow$ (c))

4 Linear Transformations

Let F be a field (e.g. $F = \mathbb{R}, \mathbb{Q}, \mathbb{C}$ or $\mathbb{F}_2$).

Definition 4.1. An $m \times n$-matrix A over a field F is an array

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

with entries a_{ij} in F. We use the notation $M_{m\times n}(F)$ for the set of all $(m \times n)$-matrices over F (see also 2.7(b)). We define addition and multiplication of matrices (and other notions) in the same way as in the case $F = \mathbb{R}$ (as seen in Linear Algebra I).

For example:

- $\begin{pmatrix} 1 & 1+i \\ 2 & 1-i \end{pmatrix} \begin{pmatrix} 1-i \\ 3 \end{pmatrix} = \begin{pmatrix} 1-i+3+3i \\ 2-2i+3-3i \end{pmatrix} = \begin{pmatrix} 4+2i \\ 5-5i \end{pmatrix}$ (matrices over $\mathbb{C}$)
- $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ (as matrices over $\mathbb{F}_2$)
- but $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$ (as matrices over $\mathbb{R}$)

Definition 4.2. Let V, W be vector spaces over a field F. A map $L : V \to W$ is called a linear transformation if the following two conditions hold:

(a) For all $x, y \in V$ we have $L(x+y) = L(x) + L(y)$ in W.
(b) For all $a \in F$ and $x \in V$ we have $L(ax) = a(L(x))$ in W.

Note: Then we also have $L(0_V) = 0_W$ and $L(x-y) = L(x) - L(y)$ for all $x, y \in V$.

Proof: $L(0_V) = L(0_V + 0_V) = L(0_V) + L(0_V)$
$\implies L(0_V) = 0_W$. (by cancelling $L(0_V)$)
$L(x-y) = L(x+(-1)y) = L(x) + L((-1)y) = L(x) + (-1)L(y) = L(x) - L(y)$. (using 2.6(d))
□

Example 4.3. (a) Let $A \in M_{m\times n}(F)$. Then the map

$$L_A : F^n \to F^m$$
$$\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto A\underline{x} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix}$$

is a linear transformation. (Compare with Lemma 5.3 in L.A.I.)
For example, if $A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$ for some $a \in \mathbb{R}$ then $L_A : \mathbb{R}^2 \to \mathbb{R}^2$ is given by $\underline{x} \mapsto a\underline{x}$, i.e. it is a stretch of the plane by a factor of a.
If $A = \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix}$ for some $0 \le \phi < 2\pi$ then $L_A : \mathbb{R}^2 \to \mathbb{R}^2$ is the clockwise rotation by the angle ϕ.

Proof that L_A is a linear transformation:

1/ Let $\underline{x}, \underline{y} \in F^n$.
$\Longrightarrow L_A(\underline{x}+\underline{y}) = A(\underline{x}+\underline{y}) = A\underline{x}+A\underline{y} = L_A(\underline{x})+L_A(\underline{y})$.
2/ Let $a \in F$ and $\underline{x} \in F^n$.
$\Longrightarrow L_A(a\underline{x}) = A(a\underline{x}) = a(A\underline{x}) = a(L_A(\underline{x}))$.
(The middle equality of both chains of equalities has been proved in Linear Algebra I for $F = \mathbb{R}$, see Thm 2.13(i) and (ii), the same proof works for any field F.) □

(b) Let V be a vector space over a field F. Then the following maps are linear transformations (cf. Example 5.4(c),(d) in L.A.I.):

- id: $V \to V$, $x \mapsto x$ (identity)
- $\underline{0}$: $V \to V$, $x \mapsto 0_V$ (zero map)
- the map $V \to V$, given by $x \mapsto ax$, for any given $a \in F$ fixed (stretch)

(c) Let $L : V \to W$ and $M : W \to Z$ be linear transformation between vector spaces over a field F. Then their composition $M \circ L : V \to Z$ is again a linear transformation. (See also Section 5.3 of L.A.I.)

Proof that $M \circ L$ is a linear transformation:
1/ Let $x, y \in V$.
$\Longrightarrow (M \circ L)(x+y) = M(L(x+y)) = M(L(x)+L(y))$
$= M(L(x))+M(L(y)) = (M \circ L)(x)+(M \circ L)(y)$.
2/ Let $a \in F$ and $x \in V$.
$\Longrightarrow (M \circ L)(ax) = M(L(ax)) = M(a(L(x))) = a(M(L(x))) = a(M \circ L)(x)$. □

(d) Let V be the subspace of $\mathbb{R}^{\mathbb{R}}$ consisting of all differentiable functions. Then differentiation $D : V \to \mathbb{R}^{\mathbb{R}}$, $f \mapsto f'$, is a linear transformation.

Proof:
1/ Let $f, g \in V \Longrightarrow D(f+g) = (f+g)' = f'+g' = D(f)+D(g)$.
2/ Let $a \in \mathbb{R}$ and $f \in V \Longrightarrow D(af) = (af)' = af' = a(D(f))$.
(The middle equality in both chains of equalities has been proved in Calculus.) □

(e) The map $L : \mathbb{R}^2 \to \mathbb{R}$, $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1 x_2$, is not a linear transformation.

Proof: Let $a = 2$ and $\underline{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in \mathbb{R}^2$. Then:

$L(a\underline{x}) = L\left(\begin{pmatrix} 2 \\ 2 \end{pmatrix}\right) = 4$, but $aL(\underline{x}) = 2 \cdot 1 = 2$. □

4.1 Matrix representation I

Proposition 4.4. *(Matrix representation I)*
Let F be a field. Let $L : F^n \to F^m$ be a linear transformation. Then there exists a unique matrix $A \in M_{m \times n}(F)$ such that $L = L_A$ (as defined in 4.3(a)). In this case we say that A represents L (with respect to the standard bases of F^n and F^m).

(See also Theorem 5.6 of L.A.I.)

For example, the map $\mathbb{R}^3 \to \mathbb{R}^2$, $\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \mapsto \begin{pmatrix} 2c_1 + c_3 - 4c_2 \\ c_2 \end{pmatrix}$, is represented by $A = \begin{pmatrix} 2 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix} \in M_{2\times 3}(\mathbb{R})$.

Proof: Let $\underline{e}_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \underline{e}_n := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ denote the standard basis of F^n.

Uniqueness: Suppose $A \in M_{m\times n}(F)$ satisfies $L = L_A$.
$\implies$ The j^{th} column of A is $A\underline{e}_j = L_A(\underline{e}_j) = L(\underline{e}_j)$ (for $j = 1, \ldots, n$)
$\implies$ $\boxed{A \text{ is the } (m \times n)\text{-matrix with the column vector } L(\underline{e}_1), \ldots, L(\underline{e}_n).}$

Existence: Let A be defined this way. We want to show $L = L_A$.
Let $\underline{c} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in F^n \implies \underline{c} = c_1\underline{e}_1 + \cdots + c_n\underline{e}_n$;
$\implies L(\underline{c}) = L(c_1\underline{e}_1) + \cdots + L(c_n\underline{e}_n) = c_1L(\underline{e}_1) + \cdots + c_nL(\underline{e}_n)$
and $L_A(\underline{c}) = \quad \ldots \quad = c_1L_A(\underline{e}_1) + \cdots + c_nL_A(\underline{e}_n)$
(because L and L_A are linear transformations)
$\implies L(\underline{c}) = L_A(\underline{c})$ because $L(\underline{e}_j) = L_A(\underline{e}_j)$ for all $j = 1, \ldots, n$). □

4.2 Kernel and image

Definition 4.5. Let $L : V \to W$ be a linear transformation between vector spaces V, W over a field F. Then

$$\ker(L) := \{x \in V : \ L(x) = 0_W\}$$

is called the kernel of L, and

$$\text{im}(L) := \{y \in W : \ \exists x \in V : \ y = L(x)\}$$

is called the image of L.

Remark 4.6. Let F be a field and $A \in M_{m\times n}(F)$. Then

$$\ker(L_A) = \text{N}(A)$$

where $\text{N}(A) = \{\underline{c} \in F^n : A\underline{c} = \underline{0}\}$ denotes the nullspace of A (see also Section 6.2 of L.A.I.) and

$$\text{im}(L_A) = \text{Col}(A)$$

where $\text{Col}(A)$ denotes the column space of A; i.e. $\text{Col}(A) = \text{Span}(\underline{a}_1, \ldots, \underline{a}_n)$, where $\underline{a}_1, \ldots, \underline{a}_n$ denote the n columns of A. (See also Section 6.4 of L.A.I.)

Proof:
First assertion: by definition.
Second assertion: follows from 4.9(a) applied to the standard basis of F^n. □

Proposition 4.7. *Let V and W be vector spaces over a field F and let $L : V \to W$ be a linear transformation. Then:*

(a) $\ker(L)$ *is a subspace of* V.
(b) $\mathrm{im}(L)$ *is a subspace of* W.

Proof: (a) We verify the three subspace axioms.

(i) We have $0_V \in \ker(L)$ (see Note after Defn 4.2.)
(ii) Let $x, y \in \ker(L)$;
$\implies L(x+y) = L(x)+L(y) = 0_W + 0_W = 0_W$;
$\implies x+y \in \ker(L)$.
(iii) Let $a \in F$ and $x \in \ker(L)$;
$\implies L(ax) = a(L(x)) = a0_W = 0_W$;
$\implies ax \in \ker(L)$.

(b) We verify the three subspace axioms.

(i) We have $0_W = L(0_V) \in \mathrm{im}(L)$.
(ii) Let $x, y \in \mathrm{im}(L)$;
$\implies \exists v, w \in V$ such that $x = L(v)$ and $y = L(w)$;
$\implies x+y = L(v)+L(w) = L(v+w) \in \mathrm{im}(L)$.
(iii) Let $y \in \mathrm{im}(L)$ and $a \in F$;
$\implies \exists x \in V$ such that $y = L(x)$;
$\implies ay = a(L(x)) = L(ax) \in \mathrm{im}(L)$. □

Example 4.8. Let $A \in M_{4\times 4}(\mathbb{R})$ be as in 3.8(d). Find a basis of the image, $\mathrm{im}(L_A)$, of $L_A : \mathbb{R}^4 \to \mathbb{R}^4, \underline{c} \mapsto A\underline{c}$.

Solution: We perform column operations:

$$A = \begin{pmatrix} 1 & -1 & 3 & 2 \\ 2 & -1 & 6 & 7 \\ 3 & -2 & 9 & 9 \\ -2 & 0 & -6 & -10 \end{pmatrix} \xrightarrow[C4\mapsto C4-2C1]{C4\mapsto C4+3C2 \\ C3\mapsto C3-3C1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 3 \\ 3 & 1 & 0 & 3 \\ -2 & -2 & 0 & -6 \end{pmatrix}$$

$$\xrightarrow{C2\mapsto C2+C1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ -2 & -2 & 0 & 0 \end{pmatrix} =: \widetilde{A}$$

$\implies$ The two vectors $\begin{pmatrix} 1 \\ 2 \\ 3 \\ -2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ -2 \end{pmatrix}$ span $\mathrm{im}(L_A)$

(because $\mathrm{im}(L_A) = \mathrm{Col}(A) = \mathrm{Col}(\tilde{A})$ by 4.6 and 3.3(b))

$\implies$ They form a basis on $\mathrm{im}(L_A)$.

(because they are also L.I., as they are not multiples of each other)□

Proposition 4.9. *Let* V *and* W *be vector spaces over a field* F *and let* $L : V \to W$ *be a linear transformation. Let* $x_1, \ldots, x_n \in V$. *Then:*

(i) If $x_1, \ldots, x_n \in V$ *span* V, *then* $L(x_1), \ldots, L(x_n)$ *span* $im(L)$.
(ii) If $L(x_1), \ldots, L(x_n)$ *are linearly independent, then* $x_1, \ldots, x_n$ *are linearly independent.*

Proof:

(a) First, $\text{Span}(L(x_1),\dots,L(x_n)) \subseteq \text{im}(L)$ (by 3.3 Note (i)).
For the other inclusion, let $y \in \text{im}(L)$;
$\implies \exists x \in V$ such that $y = L(x)$
and $\exists a_1,\dots,a_n \in F$ such that $x = a_1x_1 + \cdots + a_nx_n$ (since $V = \text{Span}(x_1,\dots,x_n)$)
$\implies y = L(x) = L(a_1x_1 + \cdots + a_nx_n) =$
$a_1L(x_1) + \cdots + a_nL(x_n) \in \text{Span}(L(x_1),\dots,L(x_n))$;
$\implies \text{im}(L) \subseteq \text{Span}(L(x_1),\dots,L(x_n))$;
$\implies \text{im}(L) = \text{Span}(L(x_1),\dots,L(x_n))$. (i.e. $L(x_1),\dots,L(x_n)$ span $\text{im}(L)$)

(b) Let $a_1,\dots,a_n \in F$ such that $a_1x_1 + \cdots + a_nx_n = 0_V$;
$\implies 0_W = L(0_V) = L(a_1x_1 + \cdots + a_nx_n) = a_1L(x_1) + \cdots + a_nL(x_n)$;
$\implies a_1 = \cdots = a_n = 0$ (since $L(x_1),\dots,L(x_n)$ are linearly independent)
$\implies x_1,\dots,x_n$ are linearly independent. □

Proposition 4.10. *(Kernel Criterion)*
Let V and W be vector spaces over a field F, and let $L : V \to W$ be a linear transformation. Then:

$$L \text{ is injective} \iff \ker(L) = \{0_V\}.$$

Proof: "$\implies$":
Let $x \in \ker(L) \implies L(x) = 0_W$.
We also have $L(0_V) = 0_W$.
$\implies x = 0_V$. (by injectivity)
"$\impliedby$":
Let $x, y \in V$ such that $L(x) = L(y)$;
$\implies L(x-y) = L(x) - L(y) = 0_W$;
$\implies x - y = 0_V$ (since $\ker(L) = \{0_V\}$)
$\implies x = y$. □

4.3 Isomorphism

Definition 4.11. Let V, W be vector spaces over a field F. A bijective linear transformation $L : V \to W$ is called an isomorphism. The vector spaces V and W are called isomorphic if there exists an isomorphism $L : V \to W$; we then write $V \cong W$.

Example 4.12. (a) For any vector space V over a field F, the identity $\text{id} : V \to V$ is an isomorphism.

(b) If $L : V \to W$ is an isomorphism then the inverse map $L^{-1} : W \to V$ is an isomorphism as well. (See also Def 5.21 from L.A.I.)

(c) If $A \in M_{n\times n}(\mathbb{R})$ is invertible then $L_A : \mathbb{R}^n \to \mathbb{R}^n$ is an isomorphism.

(d) The map $L : \mathbb{R}^2 \to \mathbb{C}, \begin{pmatrix} a \\ b \end{pmatrix} \mapsto a + bi$, is an isomorphism between the vector spaces $\mathbb{R}^2$ and $\mathbb{C}$ over $\mathbb{R}$.

(e) For any $n \in \mathbb{N}$, the map

$$L : \mathbb{R}^{n+1} \to \mathbb{P}_n, \quad \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix} \mapsto a_0 + a_1t + \cdots + a_nt^n$$

is an isomorphism between the vector spaces $\mathbb{R}^{n+1}$ and $\mathbb{P}_n$ over $\mathbb{R}$.

(f) For any $m,n \in \mathbb{N}$ we have $\mathbb{R}^{mn} \cong M_{m\times n}(\mathbb{R})$.

Proof: (b) and (c) see Coursework.
(d) and (e) follow from the following proposition and 3.8(c) and (d), respectively.
(f) (only in the case $m = n = 2$) The map

$$\mathbb{R}^4 \to M_{2\times 2}(\mathbb{R}), \qquad \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} \mapsto \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$$

is clearly an isomorphism.

Proposition 4.13. *Let V be a vector space over a field F with basis $x_1, \dots x_n$. Then the map*

$$L : F^n \to V, \ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mapsto a_1x_1 + \cdots + a_nx_n$$

is an isomorphism. (We will later use the notation $I_{x_1,\dots,x_n}$ for the map L.)

Proof:

(a) Let $\underline{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ and $\underline{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in F^n$;

$$\begin{aligned} \implies L(\underline{a}+\underline{b}) &= L\left(\begin{pmatrix} a_1+b_1 \\ \vdots \\ a_n+b_n \end{pmatrix}\right) \\ &= (a_1+b_1)x_1 + \cdots + (a_n+b_n)x_n && \text{(by definition of } L\text{)} \\ &= (a_1x_1 + \cdots + a_nx_n) + (b_1x_1 + \cdots + b_nx_n) \\ & && \text{(by distributivity, commutativity and associativity)} \\ &= L(\underline{a}) + L(\underline{b}). && \text{(by definition of } L\text{)} \end{aligned}$$

(b) Let $a \in F$ and $\underline{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in F^n$;

$$\begin{aligned} \implies L(a\underline{b}) &= L\left(\begin{pmatrix} ab_1 \\ \vdots \\ ab_n \end{pmatrix}\right) \\ &= (ab_1)x_1 + \cdots + (ab_n)x_n && \text{(by definition of } L\text{)} \\ &= a(b_1x_1 + \cdots + b_nx_n) && \text{(using the axioms of a vector space)} \\ &= a(L(\underline{b})). && \text{(by definition of } L\text{)} \end{aligned}$$

(c) $$\ker(L) = \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in F^n : a_1x_1 + \cdots + a_nx_n = 0_V \right\} = \left\{ \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \right\}$$
(because $x_1, \dots, x_n$ are linearly independent)

$\implies$ L is injective. (by 4.10)

(d) $\operatorname{im}(L) = \operatorname{Span}(L(\underline{e}_1), \ldots, L(\underline{e}_n))$ (by 4.9(a))
$= \operatorname{Span}(x_1, \ldots, x_n) = V$ (because $x_1, \ldots, x_n$ span V)
$\Longrightarrow$ L is surjective. □

Theorem 4.14. *Let V and W be vector spaces over a field F of dimension n and m, respectively. Then V and W are isomorphic if and only if $n = m$.*

Proof: "$\Longleftarrow$":
We assume that $n = m$.
$\Longrightarrow$ We have isomorphisms $L_V : F^n \to V$ and $L_W : F^n \to W$ (by 4.13)
$\Longrightarrow$ $L_W \circ L_V^{-1}$ is an isomorphism between V and W. (by 4.3(b) and 4.12(b))

"$\Longrightarrow$":
We assume that V and W are isomorphic.
Let $L : V \to W$ be an isomorphism and let $x_1, \ldots, x_n$ be a basis of V.
$\Longrightarrow$ $L(x_1), \ldots, L(x_n)$ span $\operatorname{im}(L) = W$ (by 4.9(a))
and are linearly independent (by 4.9(b) applied to L^{-1} and $L(x_1), \ldots, L(x_n)$)
$\Longrightarrow$ $L(x_1), \ldots, L(x_n)$ form a basis of W
$\Longrightarrow$ $n = \dim_F(W) = m$. □

4.4 Dimension Theorem

Theorem 4.15. *(Dimension Theorem) Let V be a vector space over a field F of finite dimension and let $L : V \to W$ be a linear transformation from V to another vector space W over F. Then:*

$$\boxed{\dim_F(\ker(L)) + \dim_F(\operatorname{im}(L)) = \dim_F(V).}$$

(In the textbooks this is sometimes called the *rank–nullity theorem*.)

Example 4.16.

$\mathbf{L : V \to W}$	$\dim(\ker(\mathbf{L}))$	$\dim(\operatorname{im}(\mathbf{L}))$	$\dim(\mathbf{V})$	**Verification**
L_A for $A \in M_{4\times 4}(\mathbb{R})$ as in 4.8	$= 2$ by 3.8(d)	$= 2$ by 4.8	$= 4$	$2+2=4$
L_A for $A \in M_{3\times 5}(\mathbb{R})$ below	$= 3$	$= 2$	$= 5$	$3+2=5$
Isomorphism	$= 0$ by 4.10	$= \dim(W)$	$= \dim(V)$	$0 + \dim(W) = \dim(V)$ by 4.14
Zero map	$= \dim(V)$	$= 0$	$= \dim(V)$	$\dim(V) + 0 = \dim(V)$

Let

$$A = \begin{pmatrix} 1 & -2 & 2 & 3 & -1 \\ -3 & 6 & -1 & 1 & -7 \\ 2 & -4 & 5 & 8 & -4 \end{pmatrix}.$$

We want to find $\dim_{\mathbb{R}}(\ker(L_A))$ and $\dim_{\mathbb{R}}(\operatorname{im}(L_A))$ – we do this by finding bases for both $\ker(L_A)$ and $\operatorname{im}(L_A)$.

We find a basis of the nullspace $N(A)$:

$$A = \begin{pmatrix} 1 & -2 & 2 & 3 & -1 \\ -3 & 6 & -1 & 1 & -7 \\ 2 & -4 & 5 & 8 & -4 \end{pmatrix} \xrightarrow{\text{Gaussian elimination}} \begin{pmatrix} 1 & -2 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & -2 \end{pmatrix} =: \tilde{A}.$$

Then

$$\begin{aligned} N(A) &= \{\underline{x} \in \mathbb{R}^5 : A\underline{x} = \underline{0}\} = \{\underline{x} \in \mathbb{R}^5 : \tilde{A}\underline{x} = \underline{0}\} = \\ &= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \in \mathbb{R}^5 : x_1 = 2x_2 + x_4 - 3x_5, x_3 = -2x_4 + 2x_5 \right\} = \\ &= \left\{ \begin{pmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{pmatrix} : x_2, x_4, x_5 \in \mathbb{R} \right\} = \\ &= \left\{ x_2 \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix} : x_2, x_4, x_5 \in \mathbb{R} \right\} = \\ &= \text{Span}_{\mathbb{R}} \left(\begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix} \right). \end{aligned}$$

The three vectors above are linearly independent, since if $a_1, a_2, a_3 \in \mathbb{R}$ and

$$a_1 \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} + a_3 \begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

then $a_1 = a_2 = a_3 = 0$ by looking at the second, fourth and fifth coordinates, respectively. Thus these three vectors are a basis of $N(A)$, so $\dim_{\mathbb{R}}(\ker(L_A)) = \dim_{\mathbb{R}}(N(A)) = 3$.

We now find a basis of the image $\text{im}(L_A)$:

$$A = \begin{pmatrix} 1 & -2 & 2 & 3 & -1 \\ -3 & 6 & -1 & 1 & -7 \\ 2 & -4 & 5 & 8 & -4 \end{pmatrix} \xrightarrow{\text{column operations}} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -3 & 0 & 5 & 0 & 0 \\ 2 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

So the vectors $\begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \\ 1 \end{pmatrix}$ span $\text{im}(L_A)$. Since they are obviously not multiples of each other, they are linearly independent, hence form a basis of $\text{im}(L_A)$. Consequently, $\dim_{\mathbb{R}}(\text{im}(L_A)) = 2$.

Proof of the Dimension Theorem:
Let $x_1,\dots,x_r$ be a basis of $\ker(L)$.
We extend $x_1,\dots,x_r$ to a basis $x_1,\dots,x_n$ of the whole V for some $n \geq r$ (by adding L.I. vectors in V until we obtain a maximal L.I. subset) and show below that $L(x_{r+1}),\dots,L(x_n)$ form a basis of $\text{im}(L)$. Then we have

$$\dim_F(\ker(L)) + \dim_F(\text{im}(L)) = r + (n-r) = n = \dim_F(V),$$

as we wanted to prove.

Proof that $L(x_{r+1}),\dots,L(x_n)$ form a basis of $\text{im}(L)$:

- $L(x_{r+1}),\dots,L(x_n)$ *span* $\text{im}(L)$:
  Let $y \in \text{im}(L)$.
  $\implies \exists x \in V$ such that $y = L(x)$ (by definition of $\text{im}(L)$)
  and $\exists a_1,\dots,a_n \in F$ such that $x = a_1x_1 + \cdots + a_nx_n$ (since $x_1,\dots,x_n$ span V)
  $\implies y = L(x) = L(a_1x_1 + \cdots + a_nx_n)$
  $= a_1L(x_1) + \cdots + a_nL(x_n)$ (because L is a linear transformation)
  $= a_{r+1}L(x_{r+1}) + \cdots + a_nL(x_n)$ (because $x_1,\dots,x_r \in \ker(L)$)
  $\in \text{Span}(L(x_{r+1}),\dots,L(x_n))$
  $\implies \text{im}(L) \subseteq \text{Span}(L(x_{r+1}),\dots,L(x_n))$.
  We also have $\text{Span}(L(x_{r+1}),\dots,L(x_n)) \subseteq \text{im}(L)$ (by 3.3/Note(i))
  $\implies \text{im}(L) = \text{Span}(L(x_{r+1}),\dots,L(x_n))$.

- $L(x_{r+1}),\dots,L(x_n)$ *are linearly independent*:
  Let $a_{r+1},\dots,a_n \in F$ such that $a_{r+1}L(x_{r+1}) + \cdots + a_nL(x_n) = 0_W$.
  $\implies L(a_{r+1}x_{r+1} + \cdots + a_nx_n) = 0_W$ (because L is a linear transformation)
  $\implies a_{r+1}x_{r+1} + \cdots + a_nx_n \in \ker(L)$ (by definition of kernel)
  $\implies \exists a_1,\dots a_r \in F$ such that $a_{r+1}x_{r+1} + \cdots + a_nx_n = a_1x_1 + \cdots + a_rx_r$
  (because $x_1,\dots,x_r$ span $\ker(L)$)
  $\implies a_1x_1 + \cdots + a_rx_r - a_{r+1}x_{r+1} - \cdots - a_nx_n = 0_V$
  $\implies a_1 = \cdots = a_r = -a_{r+1} = \cdots = -a_n = 0$ (because $x_1,\dots,x_n$ are linearly independent)
  $\implies a_{r+1} = \cdots = a_n = 0.$ □

4.5 Matrix representation II

Proposition 4.17. *(Matrix representation II) Let V and W be vector spaces over a field F with bases $x_1,\dots,x_n$ and $y_1,\dots,y_m$, respectively. Let $L : V \to W$ be a linear transformation. Then there exists a unique matrix $A \in M_{m\times n}(F)$ that represents L with respect to $x_1,\dots,x_n$ and $y_1,\dots,y_m$. Here we say that $A = (a_{ij}) \in M_{m\times n}(F)$ represents L with respect to $x_1,\dots,x_n$ and $y_1,\dots,y_m$ if for all $c_1,\dots,c_n,d_1,\dots,d_m \in F$ we have*

$$L(c_1x_1 + \cdots + c_nx_n) = d_1y_1 + \cdots + d_my_m \iff \begin{pmatrix} d_1 \\ \vdots \\ d_m \end{pmatrix} = A \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

Proof: Let $A \in M_{m\times n}(F)$. Then:
A represents L with respect to $x_1,\dots,x_n$ and $y_1,\dots,y_m$

$\overset{(\star)}{\iff}$ The diagram

$$\begin{array}{ccc} V & \xrightarrow{L} & W \\ I_{x_1,\dots,x_n} \uparrow & & \uparrow I_{y_1,\dots,y_m} \\ F^n & \xrightarrow{L_A} & F^m \end{array}$$

commutes, i.e. $L \circ I_{x_1,\dots,x_n} = I_{y_1,\dots,y_m} \circ L_A$ (see proof below)
$\iff L_A = I^{-1}_{y_1,\dots,y_m} \circ L \circ I_{x_1,\dots,x_n} =: M$
$\iff A$ represents $M : F^n \to F^m$ (with respect to the standard bases of F^n and F^m).
Hence 4.17 follows from 4.4.

Proof of ($\star$):

Let $c_1,\dots,c_n \in F$ and let $d_1,\dots,d_m \in F$ be given by $A \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} d_1 \\ \vdots \\ d_m \end{pmatrix}$

$$\implies (L \circ I_{x_1,\dots,x_n}) \left(\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \right) = L(c_1x_1 + \cdots + c_nx_n)$$

$$\text{and } (I_{y_1,\dots,y_m} \circ L_A) \left(\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \right) = I_{y_1,\dots,y_m} \left(\begin{pmatrix} d_1 \\ \vdots \\ d_m \end{pmatrix} \right) = d_1x_1 + \cdots + d_my_m.$$

Hence: $L(c_1x_1 + \cdots + c_nx_n) = d_1y_1 + \cdots + d_my_m$

$$\iff (L \circ I_{y_1,\dots,y_m}) \left(\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \right) = (I_{y_1,\dots,y_m} \circ L_A) \left(\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \right)$$

Therefore: A represents $L \iff L \circ I_{x_1,\dots,x_n} = I_{y_1,\dots,y_m} \circ L_A$. □

Note: Given $L, x_1,\dots,x_n$ and $y_1,\dots,y_m$ as in 4.17 we find the corresponding matrix A as follows: For each $i = 1,\dots,n$ we compute $L(x_i)$, represent $L(x_i)$ as a linear combination of $y_1,\dots,y_m$ and write the coefficients of this linear combination into the i^{th} column of A.

Example 4.18. Find the matrix $A \in M_{3\times 4}(\mathbb{R})$ representing differentiation $D : \mathbb{P}_3 \to \mathbb{P}_2, f \mapsto f'$, with respect to the bases $1,t,t^2,t^3$ and $1,t,t^2$ of $\mathbb{P}_3$ and $\mathbb{P}_2$, respectively.

Solution: We have

$$\begin{aligned} D(1) &= 0 = 0 + 0t + 0t^2 \\ D(t) &= 1 = 1 + 0t + 0t^2 \\ D(t^2) &= 2t = 0 + 2t + 0t^2 \\ D(t^3) &= 3t^2 = 0 + 0t + 3t^2 \end{aligned}$$

$$\implies A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Example 4.19. Let $B := \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$. Find the matrix $A \in M_{2\times 2}(\mathbb{R})$ representing the linear transformation $L_B : \mathbb{R}^2 \to \mathbb{R}^2, \underline{x} \mapsto B\underline{x}$, with respect to the basis $\begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ of $\mathbb{R}^2$ (used for both source and target space).

Solution:

$$L_B\left(\begin{pmatrix}1\\2\end{pmatrix}\right) = \begin{pmatrix}1 & -1\\2 & 4\end{pmatrix}\begin{pmatrix}1\\-2\end{pmatrix} = \begin{pmatrix}3\\-6\end{pmatrix} = 3\begin{pmatrix}1\\-2\end{pmatrix} + 0\begin{pmatrix}1\\-1\end{pmatrix}$$
$$L_B\left(\begin{pmatrix}1\\-1\end{pmatrix}\right) = \begin{pmatrix}1 & -1\\2 & 4\end{pmatrix}\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}2\\-2\end{pmatrix} = 0\begin{pmatrix}1\\-2\end{pmatrix} + 2\begin{pmatrix}1\\-1\end{pmatrix}$$

$$\implies A = \begin{pmatrix}3 & 0\\0 & 2\end{pmatrix}.$$

5 Determinants

In Linear Algebra I the determinant of a square matrix has been defined axiomatically (cf. Theorem 5.3 here). Here we begin with the following closed formula.

Definition 5.1. (Leibniz' definition of determinant) Let F be a field. Let $n \geq 1$ and $A = (a_{ij})_{i,j=1,\dots,n} \in M_{n\times n}(F)$. Then

$$\det(A) := \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} \in F$$

is called the determinant of A. We also write $|A|$ for $\det(A)$.

Example 5.2. (a) Let $n = 1$ and $A = (a_{11}) \in M_{1\times 1}(F)$.
We have $S_1 = \{\mathrm{id}\}$ and

$$\det(A) = \operatorname{sgn}(\mathrm{id})a_{11} = a_{11}.$$

(b) Let $n = 2$ and $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in M_{2\times 2}(F)$.
We have $S_2 = \{\mathrm{id}, \langle 1,2\rangle\}$ and

$$\det(A) = \operatorname{sgn}(\mathrm{id})a_{11}a_{22} + \operatorname{sgn}(\langle 1,2\rangle)a_{12}a_{21} = a_{11}a_{22} - a_{12}a_{21}.$$

For example: if $A = \begin{pmatrix} 1+2i & 3+4i \\ 1-2i & 2-i \end{pmatrix} \in M_{2\times 2}(\mathbb{C})$ then
$\det(A) = (1+2i)(2-i) - (3+4i)(1-2i) = (2+2+4i-i) - (3+8+4i-6i) = -7+5i.$

(c) Let $n = 3$ and $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in M_{3\times 3}(F)$.
We have $S_3 = \{\mathrm{id}, \langle 1,2,3\rangle, \langle 1,3,2\rangle, \langle 1,3\rangle, \langle 2,3\rangle, \langle 1,2\rangle\}$ and

$$\begin{aligned}\det(A) = {} & \underbrace{\operatorname{sgn}(\mathrm{id})}_{=+1} a_{11}a_{22}a_{33} + \underbrace{\operatorname{sgn}(\langle 1,2,3\rangle)}_{=+1} a_{12}a_{23}a_{31} + \underbrace{\operatorname{sgn}(\langle 1,3,2\rangle)}_{=+1} a_{13}a_{21}a_{32} \\ & + \underbrace{\operatorname{sgn}(\langle 1,3\rangle)}_{=-1} a_{13}a_{22}a_{31} + \underbrace{\operatorname{sgn}(\langle 2,3\rangle)}_{=-1} a_{11}a_{23}a_{32} + \underbrace{\operatorname{sgn}(\langle 1,2\rangle)}_{=-1} a_{12}a_{21}a_{33}.\end{aligned}$$

Trick to memorise:

$$\begin{array}{ccccccccc} a_{11} & & a_{12} & & a_{13} & & a_{11} & & a_{12} \\ & \searrow & & \times & & \times & & \swarrow & \\ a_{21} & & a_{22} & & a_{23} & & a_{21} & & a_{22} \\ & \swarrow & & \times & & \times & & \searrow & \\ a_{31} & & a_{32} & & a_{33} & & a_{31} & & a_{32} \end{array}$$

(Rule of Sarrus)

For example, let $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \in M_{3\times 3}(\mathbb{F}_2)$.
$\Longrightarrow \det(A) = (0+0+0) - (1+0+0) = 1$ (because $-1 = 1$ in $\mathbb{F}_2$).

(d) Let $A = (a_{ij})$ be an upper (or lower) triangular matrix. So A is of the form $\begin{pmatrix} a_{11} & \dots & \dots & a_{1n} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}$. In other words, we have $a_{ij} = 0$ if $i > j$.

Then $\det(A) = a_{11}a_{22}\cdots a_{nn}$, i.e. $\det(A)$ is the product of the entries on the main diagonal. For example, $\det(I_n) = 1$.

Proof:
Let $\sigma \in S_n, \sigma \neq \mathrm{id} \implies \exists i_0 \in \{1,\ldots,n\}$ such that $i_0 > \sigma(i_0)$;
$\implies a_{i_0,\sigma(i_0)} = 0 \implies \prod_{i=1}^n a_{i,\sigma(i)} = 0.$
$\implies \det(A) = \mathrm{sgn}(\mathrm{id}) \prod_{i=1}^n a_{i,\mathrm{id}(i)} = a_{11}a_{22}\cdots a_{nn}.$ □

Theorem 5.3. *(Weierstrass' axiomatic description of the determinant map) See Defn 4.1 of L.A.I. Let F be a field. Let $n \geq 1$. The map*

$$\det : M_{n\times n}(F) \to F, A \mapsto \det(A),$$

has the following properties and is uniquely determined by these properties:

(a) det *is linear in each column:*

$$\det\begin{pmatrix} & a_{1s}+b_{1s} & \\ C & \vdots & D \\ & a_{ns}+b_{ns} & \end{pmatrix} = \det\begin{pmatrix} & a_{1s} & \\ C & \vdots & D \\ & a_{ns} & \end{pmatrix} + \det\begin{pmatrix} & b_{1s} & \\ C & \vdots & D \\ & b_{ns} & \end{pmatrix}.$$

(b) *Multiplying any column of a matrix $A \in M_{n\times n}(F)$ with a scalar $\lambda \in F$ changes* $\det(A)$ *by the factor λ:*

$$\det\begin{pmatrix} & \lambda a_{1s} & \\ C & \vdots & D \\ & \lambda a_{ns} & \end{pmatrix} = \lambda \cdot \det\begin{pmatrix} & a_{1s} & \\ C & \vdots & D \\ & a_{ns} & \end{pmatrix}.$$

(c) *If two columns of $A \in M_{n\times n}(F)$ are equal, then* $\det(A) = 0$.

(d) $\det(I_n) = 1$.

Proof: Omitted here (please see the extended notes).

Remark 5.4. Theorem 5.3 and the following Corollary 5.5 also hold when "columns" are replaced with "rows" (similar proofs).

Corollary 5.5. *Let F be a field. Let $A \in M_{n\times n}(F)$. Then:*

(a) *For all $\lambda \in F$ we have* $\det(\lambda A) = \lambda^n \det(A)$.
(b) *If a column of A is the zero column then* $\det(A) = 0$.
(c) *Let B be obtained from A by swapping two columns of A. Then* $\det(B) = -\det(A)$.
(d) *Let $\lambda \in F$ and let B be obtained from A by adding the λ-multiple of the j^{th} column of A to the i^{th} column of A $(i \neq j)$. Then* $\det(B) = \det(A)$.

Proof:

(a) Apply Theorem 5.3(b) n times.

(b) Apply Theorem 5.3(b) with $\lambda = 0$.

(c) Let $\underline{a}$, $\underline{b}$ denote the two columns of A to be swapped.

$$\begin{aligned} \implies \det(A) + \det(B) &= \det(\ldots\underline{a}\ldots\underline{b}\ldots) + \det(\ldots\underline{b}\ldots\underline{a}\ldots) \\ &= \det(\ldots\underline{a}\ldots\underline{b}\ldots) + \det(\ldots\underline{b}\ldots\underline{a}\ldots) \\ &\quad + \underbrace{\det(\ldots\underline{a}\ldots\underline{a}\ldots)}_{=0} + \underbrace{\det(\ldots\underline{b}\ldots\underline{b}\ldots)}_{=0} && \text{(using 5.3(c))} \\ &= \det(\ldots\underline{a}\ldots\underline{a}+\underline{b}\ldots) + \det(\ldots\underline{b}\ldots\underline{a}+\underline{b}\ldots) && \text{(by 5.3(a))} \end{aligned}$$

$$= \det(\ldots \underline{a}+\underline{b} \ldots \underline{a}+\underline{b} \ldots) \qquad \text{(by 5.3(a))}$$
$$= 0 \qquad \text{(by 5.3(c))}$$

(d) $\det(B) = \det(\ldots \underline{a}+\lambda\underline{b} \ldots \underline{b} \ldots)$
$= \det(\ldots \underline{a} \ldots \underline{b} \ldots) + \lambda \underbrace{\det(\ldots \underline{b} \ldots \underline{b} \ldots)}_{=0}$ (by 5.3(a),(b))
$= \det(A)$. (by 5.3(c))□

Theorem 5.6. *Let F be a field. Let $A \in M_{n\times n}(F)$. Then the following are equivalent:*

(a) A is invertible.
(b) The columns of A span F^n.
(c) $N(A) = \{\underline{0}\}$.
(d) $\det(A) \neq 0$.

For (a) $\iff$ (d), compare Thm 4.14 of L.A.I.

Before the proof, recall that we denote the standard basis vectors of F^n as $\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \underline{e}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$. So, in particular, $I_n = (\underline{e}_1, \ldots, \underline{e}_n)$.

Proof:

"(a) $\Longrightarrow$ (b)": We'll prove the following more precise statement:
$(\star)$ $\exists B \in M_{n\times n}(F)$ such that $AB = I_n \iff$ The columns of A span F^n.
Proof of $(\star)$: Let $\underline{a}_1 \ \ldots, \underline{a}_n$ denote the columns of A. Then:
LHS $\iff \exists \underline{b}_1, \ldots, \underline{b}_n \in F^n$ such that $A\underline{b}_i = \underline{e}_i$ for all $i = 1, \ldots, n$
(we can use the columns of $B = (\underline{b}_1, \ldots, \underline{b}_n)$)
$\iff \exists \underline{b}_1, \ldots, \underline{b}_n \in F^n$ such that $\underline{a}_1 b_{i1} + \cdots + \underline{a}_n b_{in} = \underline{e}_i$ for all $i = 1, \ldots, n$
$\iff \underline{e}_1, \ldots, \underline{e}_n \in \text{Span}(\underline{a}_1, \ldots, \underline{a}_n)$
$\iff$ RHS.

"(b) $\iff$ (c)": The columns of A span F^n
$\iff$ The columns of A form a basis of F^n (by 3.15)
$\iff$ The columns of A are linearly independent (by 3.15)
$\iff N(A) = \{\underline{0}\}$ (by definition of $N(A)$)

"(b) $\Longrightarrow$ (a)": $\exists B \in M_{n\times n}(F)$ such that $AB = I_n$ (by $(\star)$)
$\Longrightarrow A(BA) = (AB)A = I_n A = A = AI_n$
$\Longrightarrow A(BA - I_n) = \underline{0}$
$\Longrightarrow$ Every column of $BA - I_n$ belongs to $N(A)$
$\Longrightarrow BA - I_n = \underline{0}$ (because $N(A) = \{\underline{0}\}$ by (b) $\iff$ (c))
$\Longrightarrow BA = I_n$
$\Longrightarrow A$ is invertible (as both $AB = I_n$ and $BA = I_n$)

"(b) $\iff$ (d)": We apply column operations to the matrix A until we arrive at a lower triangular matrix C. Then:
The columns of A span F^n
$\iff$ the columns of C span F^n (by 3.3(b))
$\iff$ all the diagonal elements of C are non-zero (because C is triangular)
$\iff \det(C) \neq 0$ (by 5.2(d))
$\iff \det(A) \neq 0$ (because $\det(C) = \lambda \det(A)$ for some non-zero $\lambda \in F$ by 5.5)□

Theorem 5.7. *(See also Thm 4.21 of L.A.I.) Let F be a field. Let $A, B \in M_{n\times n}(F)$. Then:*

$$\det(AB) = \det(A) \cdot \det(B)$$

Proof: Omitted. (See the proof of Thm 4.21 in L.A.I., it works for any field F.) □

Example 5.8. For each $m \in \mathbb{N}$ compute $\det(A^m) \in \mathbb{C}$, here $A = \begin{pmatrix} 1+4i & 1 \\ 5+i & 1-i \end{pmatrix} \in M_{2\times 2}(\mathbb{C})$.

Solution: $\det(A) = (1+4i)(1-i) - (5+i) = (1+4+4i-i) - (5+i) = 2i$.

$$\Longrightarrow \det(A^m) = \det(A)^m = (2i)^m = \begin{cases} 2^m & \text{if } m \text{ is of the form } 4k; \\ 2^m i & \text{if } m \text{ is of the form } 4k+1; \\ -(2^m) & \text{if } m \text{ is of the form } 4k+2; \\ -(2^m)i & \text{if } m \text{ is of the form } 4k+3. \end{cases} \quad \text{(by 5.7)}$$

6 Diagonalisability

6.1 Eigen-things

Definition 6.1. Let V be a vector space over a field F and let $L : V \to V$ be a linear transformation from V to itself.

(a) For any $\lambda \in F$ the set
$$E_\lambda(L) := \{x \in V : L(x) = \lambda x\}$$
is called the eigenspace of L corresponding to λ.

(b) An element $\lambda \in F$ is called an eigenvalue of L if $E_\lambda(L)$ is not the zero space. In this case any vector x in $E_\lambda(L)$ different from the zero vector is called an eigenvector of L with eigenvalue λ.

(c) Let $A \in M_{n\times n}(F)$. The eigenspaces, eigenvalues and eigenvectors of A are, by definition, those of $L_A : F^n \to F^n, \underline{x} \mapsto A\underline{x}$.

(Compare Definition 7.1 of L.A.I.)

Proposition 6.2. *Let F, V and L be as in Defn 6.1. Then $E_\lambda(L)$ is a subspace of V for every $\lambda \in F$.*

Proof:

(a) We have $0_V \in E_\lambda(L)$ because $L(0_V) = 0_V = \lambda \cdot 0_V$.
(b) Let $x, y \in E_\lambda(L)$
$\implies L(x+y) = L(x) + L(y) = \lambda x + \lambda y = \lambda(x+y)$
$\implies x+y \in E_\lambda(L)$.
(c) Let $a \in F$ and $x \in E_\lambda(L)$
$\implies L(ax) = aL(x) = a(\lambda x) = (a\lambda)x = (\lambda a)x = \lambda(ax)$
$\implies ax \in E_\lambda(L)$. □

Proposition 6.3. *Let F be a field. Let $A \in M_{n\times n}(F)$ and $\lambda \in F$. Then:*

$$\lambda \text{ is an eigenvalue of } A \iff \det(\lambda I_n - A) = 0.$$

($p_A(\lambda) := \det(\lambda I_n - A)$ is called the characteristic polynomial of A) (See also Proposition 7.5 in L.A.I.)

Proof: λ is an eigenvalue of A
$\iff \exists \underline{x} \in F^n,\ \underline{x} \neq \underline{0},$ such that $A\underline{x} = \lambda\underline{x}$
$\iff \exists \underline{x} \in F^n,\ \underline{x} \neq \underline{0},$ such that $(\lambda I_n - A)\underline{x} = 0$
$\iff N(\lambda I_n - A) \neq \{\underline{0}\}$
$\iff \det(\lambda I_n - A) = 0.$ (by 5.6 (c) $\iff$ (d)) □

Example 6.4. Determine the (complex!) eigenvalues of the matrix
$$A := \begin{pmatrix} 5i & 3 \\ 2 & -2i \end{pmatrix} \in M_{2x2}(\mathbb{C})$$
and a basis of the eigenspace of A for each eigenvalue of A.

Solution:

$$\begin{aligned} p_A(\lambda) &= \det(\lambda I_2 - A) = \det\begin{pmatrix} \lambda - 5i & -3 \\ -2 & \lambda + 2i \end{pmatrix} \\ &= (\lambda - 5i)(\lambda + 2i) - 6 = \lambda^2 - (3i)\lambda + 4 \end{aligned}$$

The two roots of this polynomial are $\lambda_{1,2} = \frac{3i \pm \sqrt{-9-16}}{2} = \frac{3i \pm 5i}{2} = 4i$ or $-i$;
$\implies$ eigenvalues of A are $4i$ and $-i$.

Basis of $E_{4i}(A)$: Apply Gaussian elimination to

$$4iI_2 - A = \begin{pmatrix} -i & -3 \\ -2 & 6i \end{pmatrix} \xrightarrow{R1 \mapsto iR1} \begin{pmatrix} 1 & -3i \\ -2 & 6i \end{pmatrix} \xrightarrow{R2 \mapsto R2 + 2R1} \begin{pmatrix} 1 & -3i \\ 0 & 0 \end{pmatrix}$$

$$\begin{aligned} \implies E_{4i}(A) &= \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{C}^2 : x_1 = (3i)x_2 \right\} \\ &= \left\{ \begin{pmatrix} (3i)x_2 \\ x_2 \end{pmatrix} : x_2 \in \mathbb{C} \right\} = \mathrm{Span}\left(\begin{pmatrix} 3i \\ 1 \end{pmatrix} \right) \end{aligned}$$

$\implies$ a basis of $E_{4i}(A)$ is $\begin{pmatrix} 3i \\ 1 \end{pmatrix}$ (as it is L.I.).

Basis of $E_{-i}(A)$: Apply Gaussian elimination to

$$-iI_2 - A = \begin{pmatrix} -6i & -3 \\ -2 & i \end{pmatrix} \xrightarrow{R1 \leftrightarrow R2} \begin{pmatrix} -2 & i \\ -6i & -3 \end{pmatrix} \xrightarrow{R2 \mapsto R2 - (3i)R1} \begin{pmatrix} -2 & i \\ 0 & 0 \end{pmatrix}$$

$$\begin{aligned} \implies E_{-i}(A) &= \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{C}^2 : x_1 = \frac{i}{2}x_2 \right\} \\ &= \left\{ \begin{pmatrix} \frac{i}{2}x_2 \\ x_2 \end{pmatrix} : x_2 \in \mathbb{C} \right\} = \mathrm{Span}\left(\begin{pmatrix} \frac{i}{2} \\ 1 \end{pmatrix} \right) \end{aligned}$$

$\implies$ a basis of $E_{-i}(A)$ is $\begin{pmatrix} i \\ 2 \end{pmatrix}$ (as it is L.I.).

Example 6.5. Let V be the real vector space of infinitely often differentiable functions from $\mathbb{R}$ to $\mathbb{R}$ and let $D : V \to V,\ f \mapsto f'$, denote differentiation (cf. 4.3(d)). Then for every $\lambda \in \mathbb{R}$ the eigenspace of D with eigenvalue λ is of dimension 1 with basis given by the function $\exp_\lambda : \mathbb{R} \to \mathbb{R},\ t \mapsto e^{\lambda t}$.

Proof: $(e^{\lambda t})' = \lambda(e^{\lambda t})$ (by the chain rule)
$\implies \exp_\lambda \in E_\lambda(D)$.
Conversely, suppose $f \in E_\lambda(D)$
$\implies (f(t)e^{-\lambda t})' = f'(t)e^{-\lambda t} + f(t)(e^{-\lambda t})'$ (by the product rule)
$= \lambda f(t)e^{-\lambda t} - \lambda f(t)e^{-\lambda t}$ (because $f \in E_\lambda(D)$ and by the chain rule)
$= 0$
$\implies f(t)e^{-\lambda t}$ is a constant, say $a \in \mathbb{R}$ (by Calculus)
$\implies f(t) = ae^{\lambda t}$, i.e. $f = a\exp_\lambda$.
Hence $E_\lambda(D) = \mathrm{Span}(\exp_\lambda)$. □

6.2 Diagonalisability

Definition 6.6. (a) Let F, V and $L : V \to V$ be as in Defn 6.1. We say that L is diagonalisable if there exists a basis $x_1, \ldots, x_n$ of V such that the matrix D representing L with respect to this basis is a diagonal matrix.

(b) Let F be a field. We say that a square matrix $A \in M_{n\times n}(F)$ is diagonalisable if the linear transformation $L_A : F^n \to F^n$, $\underline{x} \mapsto A\underline{x}$, is diagonalisable.

6.2.1 Diagonalisability (version 1)

Proposition 6.7. *Let F, V and $L : V \to V$ be as in Defn 6.1. Then L is diagonalisable if and only if V has a basis $x_1, \ldots, x_n$ consisting of eigenvectors of L.*

Proof: "$\Longrightarrow$":
Suppose $\exists$ a basis $x_1, \ldots, x_n$ of V such that the matrix D representing L is diagonal, with some $\lambda_1, \ldots, \lambda_n \in F$ on the main diagonal.
$\Longrightarrow$ for any $c_1, \ldots, c_n \in F$ we have

$$L(c_1x_1 + \cdots + c_nx_n) = (\lambda_1c_1)x_1 + \cdots + (\lambda_nc_n)x_n,$$

$$\text{because } \begin{pmatrix} \lambda_1c_1 \\ \vdots \\ \lambda_nc_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

$\Longrightarrow$ in particular when $c_1 = 0, \ldots, c_{i-1} = 0, c_i = 1, c_{i+1} = 0, \ldots, c_n = 0$ for some $i \in \{1, \ldots, n\}$, we get

$$L(x_i) = \lambda_i x_i$$

$\Longrightarrow$ $x_1, \ldots, x_n$ are eigenvectors of L with eigenvalues $\lambda_1, \ldots, \lambda_n$, respectively.

"$\Longleftarrow$":
Let $x_1, \ldots, x_n$ be a basis of V consisting of eigenvectors of L and let $\lambda_i \in F$ denote the eigenvalue corresponding to x_i.

Define a diagonal matrix D by $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$.

$\Longrightarrow$ D represents L with respect to $x_1, \ldots, x_n$ because

$$L(c_1x_1 + \cdots + c_nx_n) = c_1L(x_1) + \cdots + c_nL(x_n) = \lambda_1c_1x_1 + \cdots + \lambda_nc_nx_n$$

$$\text{and } \begin{pmatrix} \lambda_1c_1 \\ \vdots \\ \lambda_nc_n \end{pmatrix} = D \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \quad \forall c_1, \ldots, c_n \in F.$$

□

6.2.2 Diagonalisability (version 2)

Proposition 6.8. *Let F be a field. Let $A \in M_{n\times n}(F)$. Then A is diagonalisable if and only if there exists an* invertible *matrix $M \in M_{n\times n}(F)$ such that $M^{-1}AM$ is a diagonal matrix. (In this case we say that M diagonalises A.)*

Proof preparation: Let $M \in M_{n\times n}(F)$ with column vectors $\underline{x}_1, \ldots, \underline{x}_n$. Suppose M is invertible. Then:

$$\begin{aligned}
&\underline{x}_i \text{ is an eigenvector of } A \text{ with eigenvalue } \lambda_i \\
&\iff A\underline{x}_i = \lambda_i \underline{x}_i \\
&\iff AM\underline{e}_i = \lambda_i (Me_i) && \text{(because } x_i = M\underline{e}_i \text{ is the } i^{\text{th}} \text{ column of } M\text{)} \\
&\iff AM\underline{e}_i = M(\lambda_i \underline{e}_i) \\
&\iff M^{-1}AM\underline{e}_i = \lambda_i \underline{e}_i && \text{(multiply with } M^{-1}\text{)}
\end{aligned}$$

$$\iff i^{\text{th}} \text{ column of } M^{-1}AM \text{ is } \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda_i \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i^{\text{th}} \text{ place.}$$

Proof of "$\Longrightarrow$":
A is diagonalisable
$\Longrightarrow$ $\exists$ a basis $\underline{x}_1, \ldots, \underline{x}_n$ of F^n consisting of eigenvectors of A (by 6.7)
$\Longrightarrow$ The matrix $M \in M_{n\times n}(F)$ with columns $\underline{x}_1, \ldots, \underline{x}_n$ is invertible (by 5.6(b) $\Longrightarrow$ (a))
and $M^{-1}AM$ is diagonal. (by "preparation" above)

Proof of "$\Longleftarrow$":
There exists an invertible $M \in M_{n\times n}(F)$ such that $M^{-1}AM$ is diagonal
$\Longrightarrow$ the columns of M are eigenvectors of A (by "preparation" above)
and they form a basis of F^n. (by 5.6(a) $\Longrightarrow$ (b) and 3.15)
$\Longrightarrow$ A is diagonalisable. (by 6.7) $\square$

Example 6.9. Show that the matrix

$$A := \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} \in M_{3\times 3}(\mathbb{R})$$

is diagonalisable and find an invertible matrix $M \in M_{3\times 3}(\mathbb{R})$ that diagonalises it.

Solution: First compute the characteristic polynomial of A:

$$\begin{aligned}
p_A(\lambda) = \det(\lambda I_3 - A) &= \det \begin{pmatrix} \lambda & 1 & -1 \\ 3 & \lambda+2 & -3 \\ 2 & 2 & \lambda-3 \end{pmatrix} \\
&= \lambda(\lambda+2)(\lambda-3) + 1(-3)2 + (-1)3\cdot 2 - (-1)(\lambda+2)2 - \lambda(-3)(2) - 1\cdot 3(\lambda-3) \\
&= \lambda(\lambda^2 - \lambda - 6) - 6 - 6 + 2\lambda + 4 + 6\lambda - 3\lambda + 9 \\
&= \lambda^3 - \lambda^2 - \lambda + 1 = \lambda^2(\lambda-1) - (\lambda-1) = (\lambda^2-1)(\lambda-1) = (\lambda-1)^2(\lambda+1).
\end{aligned}$$

$\Longrightarrow$ Eigenvalues of A are 1 and -1.

Basis of $E_1(A)$: We apply Gaussian elimination to $1\cdot I_3 - A$:

$$1\cdot I_3 - A = \begin{pmatrix} 1 & 1 & -1 \\ 3 & 3 & -3 \\ 2 & 2 & -2 \end{pmatrix} \xrightarrow[R3\mapsto R3-2R1]{R2\mapsto R2-3R1} \begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} =: \tilde{A}$$

$$\implies \quad E_1(A) = N(1\cdot I_3 - A) = N(\tilde{A}) = \left\{ \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \in \mathbb{R}^3 : b_1 + b_2 - b_3 = 0 \right\}$$
$$= \left\{ \begin{pmatrix} -b_2 + b_3 \\ b_2 \\ b_3 \end{pmatrix} : b_2, b_3 \in \mathbb{R} \right\}$$
$$= \left\{ b_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + b_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} : b_2, b_3 \in \mathbb{R} \right\} = \text{Span}\left(\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right)$$

Also $\underline{x}_1 := \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \underline{x}_2 := \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ are L.I. (as they are not multiples of each other)

$\implies$ A basis of $E_1(A)$ is $\underline{x}_1, \underline{x}_2$.

Basis of $E_{-1}(A)$: We apply Gaussian elimination to $(-1)I_3 - A$:

$$-I_3 - A = \begin{pmatrix} -1 & 1 & -1 \\ 3 & 1 & -3 \\ 2 & 2 & -4 \end{pmatrix} \xrightarrow[R3 \mapsto R3+2R1]{R2 \mapsto R2+3R1} \begin{pmatrix} -1 & 1 & -1 \\ 0 & 4 & -6 \\ 0 & 4 & -6 \end{pmatrix}$$
$$\xrightarrow[R3 \mapsto R3-R2]{R \mapsto R1 - \frac{1}{4}R2} \begin{pmatrix} -1 & 0 & \frac{1}{2} \\ 0 & 4 & -6 \\ 0 & 0 & 0 \end{pmatrix} =: \hat{A}$$

$$\implies \quad E_{-1}(A) = N((-1)\cdot I_3 - A) = N(\hat{A}) =$$
$$= \left\{ \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \in \mathbb{R}^3 : -b_1 + \frac{1}{2}b_3 = 0 \text{ and } 4b_2 - 6b_3 = 0 \right\}$$
$$= \left\{ \begin{pmatrix} \frac{1}{2}b_3 \\ \frac{3}{2}b_3 \\ b_3 \end{pmatrix} : b_3 \in \mathbb{R} \right\} = \text{Span}\left(\begin{pmatrix} \frac{1}{2} \\ \frac{3}{2} \\ 1 \end{pmatrix} \right)$$

Also $\underline{x}_3 := \begin{pmatrix} \frac{1}{2} \\ \frac{3}{2} \\ 1 \end{pmatrix}$ is linearly independent (as it is not $\underline{0}$)

$\implies$ A basis of $E_{-1}(A)$ is $\underline{x}_3$.

For $M := (\underline{x}_1, \underline{x}_2, \underline{x}_3) = \begin{pmatrix} -1 & 1 & \frac{1}{2} \\ 1 & 0 & \frac{3}{2} \\ 0 & 1 & 1 \end{pmatrix}$ we have $\det(M) = \frac{1}{2} + \frac{3}{2} - 1 = 1 \neq 0$

$\implies$ $\underline{x}_1, \underline{x}_2, \underline{x}_3$ form a basis of $\mathbb{R}^3$ consisting of eigenvectors of A (by 5.6 and 3.13)
$\implies$ A is diagonalisable and M diagonalises A. (by the proof of 6.8)

6.2.3 Diagonalisability (version 3)

Definition 6.10. Let F be a field. Let $A \in M_{n\times n}(F)$ and $\lambda \in F$ be an eigenvalue of A.

(a) The algebraic multiplicity $a_\lambda(A)$ of λ is its multiplicity as a root of the characteristic polynomial of A.

(b) The geometric multiplicity $g_\lambda(A)$ of λ is the dimension of the eigenspace $E_\lambda(A)$.

Example 6.11. (a) In Example 6.9 we had $p_A(\lambda) = (\lambda-1)^2(\lambda+1)$, so $a_1(A) = 2$ and $a_{-1}(A) = 1$. Looking at the eigenspaces, we had $g_1(A) = 2$ and $g_{-1}(A) = 1$.

(b) Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in M_{2\times 2}(F)$ (for any field F)

$\implies p_A(\lambda) = (\lambda-1)^2$ and a basis of $E_1(A)$ is $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$\implies a_1(A) = 2$ but $g_1(A) = 1$.

Theorem 6.12. *Let F be a field. Let $A \in M_{n\times n}(F)$. Then A is diagonalisable if and only if the characteristic polynomial of A splits into linear factors and the algebraic multiplicity equals the geometric multiplicity for each eigenvalue of A.*

Proof: Omitted.

Example 6.13. Determine whether the matrix $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is diagonalisable when viewed as an element of $M_{2\times 2}(\mathbb{R})$, of $M_{2\times 2}(\mathbb{C})$ and of $M_{2\times 2}(\mathbb{F}_2)$. If A is diagonalisable then determine an invertible matrix M that diagonalises A.

Solution: $p_A(\lambda) = \det \begin{pmatrix} \lambda & -1 \\ 1 & \lambda \end{pmatrix} = \lambda^2 + 1$.

For $\mathbb{R}$:
λ^2+1 does not split into linear factors
$\implies$ as an element of $M_{2\times 2}(\mathbb{R})$ the matrix A is not diagonalisable. (by 6.12)
(Actually A is a rotation by 90° about the origin.)

For $\mathbb{C}$:
$p_A(\lambda) = \lambda^2+1 = (\lambda+i)(\lambda-i)$
$\implies a_{+i}(A) = 1$ and $a_{-i}(A) = 1$.

Basis of $E_i(A)$: We apply Gaussian elimination to $iI_2 - A$:

$$iI_2 - A = \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} \xrightarrow{R1\mapsto(-i)R1} \begin{pmatrix} 1 & i \\ 1 & i \end{pmatrix} \xrightarrow{R2\mapsto R2-R1} \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix} =: \tilde{A}$$

$$\implies \quad E_i(A) = N(iI_2 - A) = N(\tilde{A}) = \left\{ \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{C}^2 : b_1 + ib_2 = 0 \right\}$$
$$= \left\{ \begin{pmatrix} -ib_2 \\ b_2 \end{pmatrix} : b_2 \in \mathbb{C} \right\} = \text{Span}\left(\begin{pmatrix} -i \\ 1 \end{pmatrix} \right)$$

Also $\begin{pmatrix} -i \\ 1 \end{pmatrix}$ is linearly independent (as it is not $\underline{0}$)

$\implies \begin{pmatrix} -i \\ 1 \end{pmatrix}$ is a basis of $E_i(A)$
$\implies g_i(A) = 1$.

Basis of $E_{-i}(A)$: We apply Gaussian elimination to $(-i)I_2 - A$:

$$-iI_2 - A = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \xrightarrow{R1\mapsto iR1} \begin{pmatrix} 1 & -i \\ 1 & -i \end{pmatrix} \xrightarrow{R2\mapsto R2-R1} \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix} =: \hat{A}$$

$$\implies \quad E_{-i}(A) = N((-i)I_2 - A) = N(\hat{A}) = \left\{ \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{C}^2 : b_1 - ib_2 = 0 \right\}$$
$$= \left\{ \begin{pmatrix} ib_2 \\ b_2 \end{pmatrix} : b_2 \in \mathbb{C} \right\} = \text{Span}\left(\begin{pmatrix} i \\ 1 \end{pmatrix} \right)$$

Also $\begin{pmatrix} i \\ 1 \end{pmatrix}$ is linearly independent (as it is not $\underline{0}$)

$\implies \begin{pmatrix} i \\ 1 \end{pmatrix}$ is a basis of $E_{-i}(A)$

$\implies g_{-i}(A) = 1.$

$\implies$ A is diagonalisable when viewed as an element of $M_{2\times 2}(\mathbb{C})$ (by 6.12)
and $M = \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix}$ diagonalises A.

For $\mathbb{F}_2$: $p_A(\lambda) = \lambda^2 + 1 = (\lambda + 1)^2$ (since $1 + 1 = 0$ in $\mathbb{F}_2$)

$\implies$ A has a single eigenvalue $1 = -1$ and $a_1(A) = 2$.

Basis of $E_1(A)$: We apply Gaussian elimination to $1 \cdot I_2 - A$:

$$I_2 - A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \xrightarrow{R2 \mapsto R2 - R1} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \widehat{A}$$

$$\implies \quad E_1(A) = N(I_2 - A) = N(\widehat{A}) = \left\{ \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{F}_2^2 : b_1 + b_2 = 0 \right\}$$
$$= \left\{ \begin{pmatrix} b_2 \\ b_2 \end{pmatrix} : b_2 \in \mathbb{F}_2 \right\} = \text{Span}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right)$$

Also $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is linearly independent (as it is not $\underline{0}$)

$\implies \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is a basis of $E_1(A)$

$\implies g_1(A) = 1.$

$\implies$ A is not diagonalisable. (by 6.12, since $a_1(A) = 2 \neq 1 = g_1(A)$)

6.3 Cayley–Hamilton Theorem

Theorem 6.14. *(Cayley–Hamilton Theorem) Let F be a field, let $A \in M_{n\times n}(F)$ and let p_A be the characteristic polynomial of A. Then $p_A(A)$ is the zero matrix.*

Example 6.15. Let $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in M_2(F)$

$\implies p_A(\lambda) = \lambda^2 + 1$ (see Example 6.13)

$\implies p_A(A) = A^2 + 1 \cdot I_2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$

Proof of the Cayley–Hamilton Theorem 6.14:

First Case: When $A = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$ is a diagonal matrix.

$$\begin{aligned}
\implies \quad & p_A(\lambda) = \det(\lambda \cdot I_n - A) = (\lambda - a_1)\dots(\lambda - a_n) \\
\implies \quad & p_A(A) = (A - a_1 I_n)\cdots(A - a_n I_n) \\
& = \begin{pmatrix} 0 & & 0 \\ & a_2 - a_1 & \\ & & \ddots \\ 0 & & & a_n - a_1 \end{pmatrix} \begin{pmatrix} a_1 - a_2 & & & 0 \\ & 0 & & \\ & & \ddots & \\ 0 & & & a_n - a_2 \end{pmatrix} \cdots \begin{pmatrix} a_1 - a_n & & & 0 \\ & \ddots & & \\ & & a_{n-1} - a_n & \\ 0 & & & 0 \end{pmatrix} \\
& = \underline{0},
\end{aligned}$$

because the product of any two diagonal matrices with diagonal entries $b_1, \dots, b_n$ and $c_1, \dots, c_n$ respectively, is the diagonal matrix with diagonal entries $b_1 c_1, \dots, b_n c_n$.

Preparatory Step: If $A, M, D \in M_{n\times n}(F)$ are such that M is invertible and $D = M^{-1}AM$, then:

$$\begin{aligned}
p_D(\lambda) &= \det(\lambda \cdot I_n - D) = \det(\lambda \cdot I_n - M^{-1}AM) \\
&= \det(M^{-1}(\lambda I_n)M - M^{-1}AM) = \det(M^{-1}(\lambda I_n - A)M) \\
&= \det(M)^{-1}\det(\lambda I_n - A)\det(M) = \det(\lambda I_n - A) \\
&= p_A(\lambda).
\end{aligned}$$

In other words, the characteric polynomials of A and D are the same.

Another Preparatory Computation: If $M, D \in M_{n\times n}(F)$, M invertible, and $k \geq 0$, then:

$$\begin{aligned}
(MDM^{-1})^k &= (MDM^{-1})(MDM^{-1})\cdots(MDM^{-1})) \\
&= MD(M^{-1}M)D(M^{-1}M)\cdots(M^{-1}M)DM^{-1} \\
&= MD^kM^{-1}.
\end{aligned}$$

Second Case: When A is a diagonalisable matrix.
$\implies \exists M \in GL_n(F)$ such that $M^{-1}AM = D$ where D is a diagonal matrix (by 6.8)
Denote $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$ the characteristic polynomial of A

$$\begin{aligned}
\implies \; p_A(A) &= A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_n \\
&= (MDM^{-1})^n + a_{n-1}(MDM^{-1})^{n-1} + \cdots + a_1(MDM^{-1}) + a_0 I_n \\
&\qquad \text{(by Preparatory Computation above)} \\
&= MD^nM^{-1} + a_{n-1}MD^{n-1}M^{-1} + \cdots + a_1 MDM^{-1} + a_0 MM^{-1} \\
&= M(D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0 I_n)M^{-1} \\
&= Mp_A(D)M^{-1} \\
&= Mp_D(D)M^{-1} \qquad \text{(by Preparatory Step above)} \\
&= M\underline{0}M^{-1} = \underline{0}. \qquad \text{(by the First Case)}
\end{aligned}$$

General Case: Omitted. □

7 Coursework Sheets

7.1 Coursework Sheet 0 — not marked

Do NOT submit, not marked.

Exercise 1

Consider the matrix $A = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 1 & 0 \\ 2 & 0 & -1 \end{pmatrix}$.

(a) Show that the equation $A\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ has no solution in $\mathbb{R}^3$.

(b) Find two linearly independent vectors $\mathbf{y}$ in $\mathbb{R}^3$ such that the equation $A\mathbf{x} = \mathbf{y}$ has a solution in $\mathbb{R}^3$.

Exercise 2

(a) Let w, z be complex numbers. Solve the linear equation $wx = z$; in other words, find all $x \in \mathbb{C}$ such that $wx = z$. (*Hint: You need to distinguish three cases.*)

(b) Solve the following system of linear equations:

$$\begin{aligned} (5i)x_1 + 3x_2 &= 12 + i \\ x_1 - (2i)x_2 &= 3 - i. \end{aligned}$$

Exercise 3

Let A be a real $m \times n$ matrix. Recall that its *nullspace* is the following set of vectors: $N(A) = \{\mathbf{u} \in \mathbb{R}^n \mid A\mathbf{u} = \mathbf{0} \in \mathbb{R}^m\} \subseteq \mathbb{R}^n$. Prove that $N(A)$ is a subspace of $\mathbb{R}^n$.

Extra question (not assessed)

Multiplication of complex numbers defines a binary operation on $\mathbb{C}^\times := \mathbb{C} \setminus \{0\}$. Show that $\mathbb{C}^\times$ together with this operation is an abelian group. (*Here we consider the multiplication of complex numbers defined like so: if $a+bi$ and $c+di$ ($a,b,c,d \in \mathbb{R}$) are complex numbers, their product is* declared *to be the complex number $(ac-bd)+(ad+bc)i$. In your arguments, you may without further discussion use the usual laws of algebra for $\mathbb{R}$. such as associativity for addition and multiplication of real numbers.*)

Challenge question (not assessed)

Let G be a group such that for all $a \in G$ we have $a * a = e$. Show that G is abelian.

7.2 Coursework Sheet 1

Submit a single pdf with scans of your work to Blackboard by Monday, 15 February 2021, 17:00.

Exercise 1

Let G and H be groups with binary operations $\boxplus$ and $\odot$, respectively. We define a binary operation $*$ on the cartesian product $G \times H$ by

$$(a,b) * (a',b') := (a \boxplus a', b \odot b') \quad (\text{for } a,a' \in G \text{ and } b,b' \in H).$$

Show that $G \times H$ together with this operation is a group.

Exercise 2

For $a,b \in \mathbb{R}$ we define $a * b := a + b + ab \in \mathbb{R}$. Furthermore let $G := \mathbb{R} \setminus \{-1\}$.

(a) Show that $a * b \in G$ for all $a,b \in G$.

(b) Show that G together with the binary operation $G \times G \to G$, $(a,b) \mapsto a * b$, is a group.

Exercise 3

Let $G = \{s,t,u,v\}$ be a group with $s * u = u$ and $t * t = v$. Determine the group table of G. (*There is only one way of completing the group table for G. Give a reason for each step.*)

Exercise 4

Write down the group tables for the groups C_4 and $C_2 \times C_2$ (cf. Exercise 1). For every element a in C_4 and $C_2 \times C_2$ determine the smallest positive integer m such that ma equals the identity element.

Extra question (not assessed — no need to submit)

Let G be a group whose binary operation is written additively, i.e. $G \times G \to G$, $(a,b) \mapsto a + b$. Show that $m(na) = (mn)a$ for all $a \in G$ and $m,n \in \mathbb{Z}$. (*Hint: You need to distinguish up to 9 cases.*) Write down the other two exponential laws in additive notation as well. (*Formulate these laws as complete mathematical statements including all quantifiers. No proofs are required.*)

7.3 Coursework Sheet 2

Submit a single pdf with scans of your work to Blackboard by Monday, 22 February 2021, 17:00.

Exercise 1

Write down the group table for the permutation group S_3 and show that S_3 is not abelian. (*You may find it more convenient to write all elements of S_3 in cycle notation.*)

Exercise 2

Let $\sigma := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 7 & 6 & 1 & 8 & 9 & 4 & 2 & 5 \end{pmatrix} \in S_9$, $\tau := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 2 & 5 & 6 & 3 & 1 \end{pmatrix} \in S_6$

and $\eta := \begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ n & n-1 & \dots & 2 & 1 \end{pmatrix} \in S_n$ (for any even $n \in \mathbb{N}$).

(a) Determine the sign of σ, τ and η.

(b) Write σ^2, σ^{-1}, τ^2, τ^{-1}, η^2 and η^{-1} as a composition of cycles.

(c) Determine the sign of σ^2, τ^2 and η^2 in two ways, firstly using (b) and secondly using (a) and Theorem 1.10 (b).

Exercise 3

Let $n \geq 1$. Let $\langle a_1, \dots, a_s \rangle \in S_n$ be a cycle and let $\sigma \in S_n$ be arbitrary. Show that

$$\sigma \circ \langle a_1, \dots, a_s \rangle \circ \sigma^{-1} = \langle \sigma(a_1), \dots, \sigma(a_s) \rangle \text{ in } S_n.$$

(*Note this is an equality between maps. Hence, in order to show this equality you need to show that both sides are equal after applying them to an arbitrary element b of* $\{1, 2, \dots, n\}$. *To do so you will need to distinguish whether b belongs to* $\{\sigma(a_1), \dots, \sigma(a_s)\}$ *or not.*)

Exercise 4

Let $\mathbb{Q}(\sqrt{5})$ denote the set of real numbers z of the form $z = a + b\sqrt{5}$ where $a, b \in \mathbb{Q}$. Show that $\mathbb{Q}(\sqrt{5})$ together with the usual addition and multiplication of real numbers is a field. (*Hint: You need to show that for any* $w, z \in \mathbb{Q}(\sqrt{5})$ *also* $w+z, wz, -z$ *and* z^{-1} *(if* $z \neq 0$*) are in* $\mathbb{Q}(\sqrt{5})$ *and that* 0 *and* 1 *are in* $\mathbb{Q}(\sqrt{5})$. *Distributivity, commutativity and associativity for addition and multiplication hold in* $\mathbb{Q}(\sqrt{5})$ *because they hold in* $\mathbb{R}$.)

Exercise 5

Let F be a field. For any $a, b \in F$, $b \neq 0$, we write $\frac{a}{b}$ for ab^{-1}. Prove the following statements for any $a, a' \in F$ and $b, b' \in F \setminus \{0\}$:

(i) $\dfrac{a}{b} + \dfrac{a'}{b'} = \dfrac{ab' + a'b}{bb'}$; (ii) $\dfrac{a}{b}\dfrac{a'}{b'} = \dfrac{aa'}{bb'}$.

Extra items for Exercise 5 (not assessed, do not submit):

(iii) $\dfrac{a}{b} = \dfrac{a'}{b'}$ if and only if $ab' = a'b$;

(iv) $\dfrac{\frac{a}{b}}{\frac{a'}{b'}} = \dfrac{ab'}{a'b}$ (if in addition $a' \neq 0$).

Extra problems to think about (do not submit)

The solutions for this will not be provided (but possible to find in a book or google). Not necessary for the rest of the module at all. Feel free to ignore.

Task 1. The aim is to prove Thm 1.10 from the notes, about the sign function on the symmetric groups S_n. Here's one possible path to a proof.

- Every cycle of length k can be written as a product of $k-1$ transpositions.
- Thus, every permutation can be written as a product of transpositions.
- Let σ be a permutation, and write it as a product of transpositions. Define the number $\mathrm{nsgn}(\sigma)$ (for "new sign' ') to be equal to 1 if the number of transpositions is even, and -1 if the number of transpositions is odd. Again, apriori nsgn depends on *how* do we write σ as a product of transpositions. However, by the first point above, $\mathrm{nsgn}(\sigma) = \mathrm{sgn}(\sigma)$, since every cycle decomposition of σ gives also a way to write σ as a product of transpositions. So the goal now is to prove that nsgn is well defined, and that it's multiplicative.
- A way to prove the above is to find a way to characterise nsgn to be something *intrinsic* to a permutation. Here's such a thing: Given a permutation $\sigma \in S_n$, we say that σ *reverses the pair* (i,j), if $i,j \in \{1,\dots,n\}$, $i<j$ and $\sigma(i) > \sigma(j)$. Let $\mathrm{isgn}(\sigma)$ be 1 if σ reverses even number of pairs, and -1 if σ reverses odd number of pairs.
- Prove that if σ is a permutation and τ is a transposition, then $\mathrm{isgn}(\sigma\circ\tau) = -\mathrm{isgn}(\sigma) = \mathrm{isgn}(\tau\circ\sigma)$.
- From the previous point, conclude that $\mathrm{isgn}(\sigma) = \mathrm{nsgn}(\sigma)$ (thus the sign is well defined).
- From the definition of nsgn, show that $\mathrm{nsgn}(\sigma\circ\tau) = \mathrm{nsgn}(\sigma)\mathrm{nsgn}(\tau)$ for any two permutations σ and τ.

Task 2. Prove that the number of elements of S_n (i.e. the order of the symmetric group S_n) is $n!$.

7.4 Coursework Sheet 3

Submit a single pdf with scans of your work to Blackboard by Monday, 1 March 2021, 17:00.

Exercise 1

The set $\mathbb{R}^2$ together with the usual vector addition forms an abelian group. For $a \in \mathbb{R}$ and $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$ we put $a \otimes \mathbf{x} := \begin{pmatrix} ax_1 \\ 0 \end{pmatrix} \in \mathbb{R}^2$; this defines a scalar multiplication

$$\mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2, \quad (a, \mathbf{x}) \mapsto a \otimes \mathbf{x},$$

of the field $\mathbb{R}$ on $\mathbb{R}^2$. Determine which of the axioms defining a vector space hold for the abelian group $\mathbb{R}^2$ with this scalar multiplication. (*Proofs or counterexamples are required.*)

Exercise 2

The set $\mathbb{R}_{>0}$ of positive real numbers together with multiplication forms an abelian group. Let $\mathbb{R}^n_{>0}$ denote the n-fold cartesian product of $\mathbb{R}_{>0}$ with itself (cf. Exercise 1 on Sheet 2). (*You may find it convenient to use the symbol $\oplus$ for the binary operation in the abelian group $\mathbb{R}^n_{>0}$, that is $(b_1,\dots,b_n)\oplus(c_1,\dots,c_n) = (b_1c_1,\dots,b_nc_n)$ for $b_1,\dots,b_n,c_1,\dots,c_n \in \mathbb{R}_{>0}$.*) Furthermore, for $a \in \mathbb{Q}$ and $\mathbf{b} = (b_1,\dots,b_n) \in \mathbb{R}^n_{>0}$ we put $a \otimes \mathbf{b} := (b_1^a,\dots,b_n^a)$. Show that the abelian group $\mathbb{R}^n_{>0}$ together with the scalar multiplication

$$\mathbb{Q} \times \mathbb{R}^n_{>0} \to \mathbb{R}^n_{>0}, \quad (a, \mathbf{b}) \mapsto a \otimes \mathbf{b},$$

is a vector space over $\mathbb{Q}$.

Exercise 3

Let V be a vector space over the field F and let $a \in F$ and $x, y \in V$.

(a) Show that $a(x-y) = ax - ay$ in V.

(b) If $ax = 0_V$ show that $a = 0_F$ or $x = 0_V$.

(*Remember to give a reason for each step.*)

Exercise 4

Let S be a set and let V be a vector space over a field F. Let V^S denote the set of all maps from S to V. We define an addition on V^S and a scalar multiplication of F on V^S as follows: let $f, g \in V^S$ and let $a \in F$; then

$$(f+g)(s) := f(s) + g(s) \text{ and } (af)(s) := a\,(f(s)) \text{ (for any } s \in S).$$

Show that V^S is a vector space over F. (*For a complete proof many axioms need to be checked. In order to save you some writing, your solution will be considered complete, if you check that there exists an additive identity element in V^S, that every element in V^S has an additive inverse and that the second distributivity law holds.*)

7.5 Coursework Sheet 4

Submit a single pdf with scans of your work to Blackboard by Monday, 15 March 2021, 17:00.

Exercise 1

Let $n \geq 2$. Which of the conditions defining a subspace are satisfied for the following subsets of the vector space $M_{n\times n}(\mathbb{R})$ of real $(n \times n)$-matrices? (*Proofs or counterexamples are required.*)

$$\begin{aligned} U &:= \{A \in M_{n\times n}(\mathbb{R}) \,|\, \text{rank}(A) \leq 1\} \\ V &:= \{A \in M_{n\times n}(\mathbb{R}) \,|\, \det(A) = 0\} \\ W &:= \{A \in M_{n\times n}(\mathbb{R}) \,|\, \text{trace}(A) = 0\} \end{aligned}$$

(*Recall that rank(A) denotes the number of non-zero rows in a row-echelon form of A and trace(A) denotes the sum $\sum_{i=1}^{n} a_{ii}$ of the diagonal elements of the matrix $A = (a_{ij})$.*)

Exercise 2

Which of the following subsets of the vector space $\mathbb{R}^{\mathbb{R}}$ of all functions from $\mathbb{R}$ to $\mathbb{R}$ are subspaces? (*Proofs or counterexamples are required.*)

$$\begin{aligned} U &:= \{f \in \mathbb{R}^{\mathbb{R}} \,|\, f \text{ is differentiable and } f'(-5) = 0\} \\ V &:= \{f \in \mathbb{R}^{\mathbb{R}} \,|\, f \text{ is polynomial of the form } f = at^2 \text{ for some } a \in [0,\infty)\} \\ &= \{f \in \mathbb{R}^{\mathbb{R}} \,|\, \exists a \in [0,\infty] : \forall s \in \mathbb{R} : f(s) = as^2\} \\ W &:= \{f \in \mathbb{R}^{\mathbb{R}} \,|\, f \text{ is polynomial of the form } f = at^3 \text{ or } f = at^5 \text{ for some } a \in \mathbb{R}\} \\ &= \{f \in \mathbb{R}^{\mathbb{R}} \,|\, \exists i \in \{3,5\}\ \exists a \in \mathbb{R} : \forall s \in \mathbb{R} : f(s) = as^i\} \\ X &:= \{f \in \mathbb{R}^{\mathbb{R}} \,|\, f \text{ is even}\} \end{aligned}$$

(*Recall that a function* $f : \mathbb{R} \to \mathbb{R}$ *is called* even *if* $f(-s) = f(s)$ *for all* $s \in \mathbb{R}$.)

Exercise 3

Let $\mathbb{F}_2 = \{0, 1\}$ denote the field with 2 elements.

(a) Let V be a vector space over $\mathbb{F}_2$. Show that every non-empty subset W of V which is closed under addition is a subspace of V.

(b) Show that $\{(0,0),(1,0)\}$ is a subspace of the vector space $\mathbb{F}_2^2$ over $\mathbb{F}_2$.

(c) Write down all subsets of $\mathbb{F}_2^2$ and underline those subsets which are subspaces. (*No explanations are required.*)

Exercise 4 (optional, not marked)

Let V be a vector space over a field F. Putting $S = V$ in Example 2.7 we obtain the vector space F^V consisting of *all* functions from V to F. Consider the subset

$$V^* := \{L : V \to F \mid L \text{ is a linear transformation}\},$$

consisting of all linear transformations from the vector space V to the (one-dimensional) vector space F. Show that V^* is a subspace of F^V. (*To get you started, at the end of this sheet you'll find a detailed proof of the first of the three conditions that need to be verified for a subspace.*)

Extra question (not marked, do not submit)

Let V be a vector space over a field F and let X, Y and Z be subspaces of V, such that $X \subseteq Y$. Show that $Y \cap (X + Z) = X + (Y \cap Z)$. (*Note: this is an equality of sets, so you need to show that every vector in the LHS also belongs to RHS, and vice versa.*)

Verification of the first condition of being a subspace, for V^* from Exercise 4

(You don't need to reproduce this in your solution, just say that the first condition is proved.)

The first condition for a subspace asserts that the zero vector of the "big" vector space F^V belongs to set V^* that we are showing to be a subspace.

The zero vector (= the additive identity element for vector addition) of F^V is the zero function $\underline{0} : V \to F$, defined by $\underline{0}(v) = 0_F$ for all $v \in V$, that is, it maps *every* vector v from V to the additive identity element 0_F in the field F.

We need to show that this function $\underline{0}$ belongs to the set V^*, in other words, that it is a linear transformation from V to F. This entails checking two conditions:

(a) $\underline{0}$ is compatible with addition: take arbitrary vectors $x, y \in V$. We need to check that $\underline{0}(x+y) = \underline{0}(x) + \underline{0}(y)$ in F:
LHS $= 0_F$ (by definition of $\underline{0}$)
RHS $= 0_F + 0_F = 0_F$ (by definition of $\underline{0}$ and the field axioms)
So LHS = RHS.

(b) $\underline{0}$ is compatible with scalar multiplication: take a vector $x \in V$ and a scalar $a \in F$. We need to check that $\underline{0}(ax) = a(\underline{0}(x))$ in F:
LHS $= 0_F$ (by definition of $\underline{0}$)

RHS $= a0_F = 0_F$ (by definition of $\underline{0}$ and Prop. 2.3(a))
So again LHS = RHS.

7.6 Coursework Sheet 5

Submit a single pdf with scans of your work to Blackboard by Monday, 12 April 2021, 17:00.

Exercise 1

Which of the following are spanning sets for the vector space $\mathbb{P}_2$ of polynomial functions of degree at most 2? (*Give reasons for your answers.*)

(a) $\frac{1}{2}, t^2+t, t^2-1$

(b) $1, 2t, t^2, 3t^2+5$

(c) $t+1, t^2+t$

Exercise 2

Determine whether the following are linearly independent sets of vectors in the vector space $\mathbb{R}^{\mathbb{R}}$ of all functions from $\mathbb{R}$ to $\mathbb{R}$. (*Give reasons for your answers.*)

(a) $1+t, 1+t+t^2, 1+t+t^2+t^3, 1+t+t^2+t^4$

(b) $\sin, \cos^2, \sin^3$

(c) $1, \sin^2, \cos^2$

(*Here for example* $\sin^2$ *denotes the function* $\mathbb{R} \to \mathbb{R}, s \mapsto (\sin(s))^2$.)

Exercise 3

Find a basis of the null space $N(A) \subset \mathbb{R}^5$ of the matrix

$$A = \begin{pmatrix} 1 & -3 & 3 & -1 & -1 \\ -2 & 6 & -1 & -3 & -8 \\ 3 & -9 & 10 & -4 & -5 \end{pmatrix} \in M_{3\times 5}(\mathbb{R})$$

and hence determine its dimension.

Exercise 4

(a) Determine whether the following (2×2)-matrices form a basis of the vector space $M_{2\times 2}(\mathbb{R})$ of all (2×2)-matrices over $\mathbb{R}$:

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 3 & 2 \\ 1 & 0 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 4 & 3 \\ 2 & 1 \end{pmatrix}.$$

(b) Find a basis of the subspace $W := \{A \in M_{2\times 2}(\mathbb{R}) \mid \operatorname{trace}(A) = 0\}$ of the vector space $M_{2\times 2}(\mathbb{R})$ and hence determine the dimension of W (see Exercise 1 on Coursework Sheet 4 for the definition of trace).

Extra exercise (not marked, do not submit)

We view $\mathbb{C}^2 = \left\{ \begin{pmatrix} w \\ z \end{pmatrix} : w,z \in \mathbb{C} \right\}$ as a vector space over $\mathbb{C}$, $\mathbb{R}$ and $\mathbb{Q}$ (cf. Example 3.16 (b)). Let $\mathbf{x}_1 := \begin{pmatrix} i \\ 0 \end{pmatrix}, \mathbf{x}_2 := \begin{pmatrix} \sqrt{2} \\ \sqrt{5} \end{pmatrix}, \mathbf{x}_3 := \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \mathbf{x}_4 := \begin{pmatrix} i\sqrt{3} \\ \sqrt{3} \end{pmatrix}, \mathbf{x}_5 := \begin{pmatrix} 1 \\ 3 \end{pmatrix} \in \mathbb{C}^2$. Determine $\dim_F(\mathrm{Span}_F(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4,\mathbf{x}_5))$ for $F = \mathbb{C}$, $\mathbb{R}$ and $\mathbb{Q}$.

7.7 Coursework Sheet 6

Submit a single pdf with scans of your work to Blackboard by Monday, 19 April 2021, 17:00.

Exercise 1

Determine whether the following maps are linear transformations. (For a matrix A, A^{T} denotes its transpose, see Section 2.3 in L.A.I.) (*Proofs or counterexamples are required.*)

(a) $L : \mathbb{R}^2 \to \mathbb{R}^3, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_2 \\ x_1 + 2x_2 \\ x_1 - x_2 \end{pmatrix}$

(b) $L : \mathbb{R}^2 \to \mathbb{R}, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1^2 + x_1 x_2$

(c) $L : M_{n\times n}(\mathbb{R}) \to M_{n\times n}(\mathbb{R}), \quad A \mapsto A^{\mathrm{T}} - A$

(d) $L : \mathbb{P}_3 \to \mathbb{P}_2, \quad f \mapsto 2f' + (f(3))t^2$

Exercise 2

Consider the linear transformation $\mathbb{R}^3 \to \mathbb{R}^4$ given by $L(\mathbf{x}) = A\mathbf{x}$ where A is the matrix

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & 2 \\ 1 & 2 & -5 \\ 2 & 1 & 1 \end{pmatrix} \in M_{4\times 3}(\mathbb{R}).$$

Find a basis of the image of L. Using the Dimension Theorem show that L is injective.

Exercise 3

Let F be a field.

(a) Let $A \in M_{n\times n}(F)$ be an invertible matrix. Show that the linear transformation

$$L_A : F^n \to F^n, \quad \mathbf{x} \mapsto A\mathbf{x},$$

(cf. Example 4.3(a)) is an isomorphism.

(b) Let $L : V \to W$ be an isomorphism between vector spaces over F. Show that the inverse map $L^{-1} : W \to V$ is a linear transformation (and hence an isomorphism as well).

Exercise 4

For $\mathbf{y} \in \mathbb{R}^n$ let $L_{\mathbf{y}} : \mathbb{R}^n \to \mathbb{R}$ denote the map given by $\mathbf{x} \mapsto L_{\mathbf{y}}(\mathbf{x}) = \mathbf{x} \cdot \mathbf{y}$ where $\mathbf{x} \cdot \mathbf{y}$ denotes the dot product of $\mathbf{x}$ and $\mathbf{y}$ introduced in Linear Algebra I.

(a) For each $\mathbf{y} \in \mathbb{R}^n$ show that $L_{\mathbf{y}}$ is a linear transformation and compute $\dim_{\mathbb{R}}(\ker(L_{\mathbf{y}}))$.

(b) (optional, not marked) Let $(\mathbb{R}^n)^*$ denote the vector space introduced in Coursework 4/Exercise 4. Show that the map $L : \mathbb{R}^n \to (\mathbb{R}^n)^*$, $\mathbf{y} \mapsto L_{\mathbf{y}}$, is an isomorphism. (*Hint: For surjectivity use Proposition 4.4.*)

7.8 Coursework Sheet 7

Submit a single pdf with scans of your work to Blackboard by Tuesday, 4 May 2021, 17:00.

Exercise 1

From Calculus we know that for any polynomial function $f : \mathbb{R} \to \mathbb{R}$ of degree at most n, the function $I(f) : \mathbb{R} \to \mathbb{R}$, $s \mapsto \int_0^s f(u)\,du$, is a polynomial function of degree at most $n+1$. Show that the map

$$I : \mathbb{P}_n \to \mathbb{P}_{n+1}, \quad f \mapsto I(f),$$

is an injective linear transformation, determine a basis of the image of I and find the matrix $M \in M_{(n+2)\times(n+1)}(\mathbb{R})$ that represents I with respect to the basis $1, t, \ldots, t^n$ of $\mathbb{P}_n$ and the basis $1, t, \ldots, t^{n+1}$ of $\mathbb{P}_{n+1}$.

Exercise 2

(a) Let $\alpha \in \mathbb{C}$ and $A := \begin{pmatrix} 1-i & \alpha & i \\ i-\alpha & 1-\alpha & \alpha-i \\ 1-\alpha & 1 & 2+\alpha \end{pmatrix} \in M_{3\times 3}(\mathbb{C})$. Compute $\det(A) \in \mathbb{C}$.

(b) Let F be a field, n be even and let $c_1, \ldots, c_n \in F$. Follow the blueprint of the proof of Example 5.2(d) and use Exercise 2(a) on Coursework Sheet 2 to compute the determinant of the matrix

$$B := \begin{pmatrix} 0 & \ldots & 0 & c_1 \\ \vdots & \iddots & \iddots & 0 \\ 0 & \iddots & \iddots & \vdots \\ c_n & 0 & \ldots & 0 \end{pmatrix} \in M_{n\times n}(F).$$

Exercise 3

Let F be a field, let $n \geq 1$ and let $a, b \in F$. Furthermore let

$$A := \begin{pmatrix} b & a & a & \ldots & a \\ a & b & a & \ldots & a \\ a & a & b & \ldots & a \\ \vdots & \vdots & \vdots & & \vdots \\ a & a & a & \ldots & b \end{pmatrix} \in M_{n\times n}(F).$$

Show that $\det(A) = (b + (n-1)a)(b-a)^{n-1}$. (*Hint: Begin with $R_1 \mapsto R_1 + R_2 + \ldots + R_n$, apply Theorem 5.3(b) and use further row operations to arrive at an upper triangular matrix.*)

Exercise 4

Let $A = \begin{pmatrix} 1+i & 1-i & 2 \\ 3 & i & -i \\ 1 & 2-i & 2+i \end{pmatrix} \in M_{3\times 3}(\mathbb{C})$ and $B = \begin{pmatrix} 2 & 3+i & 2i \\ 1 & 2-i & 2+2i \\ 1-i & i & 3 \end{pmatrix} \in M_{3\times 3}(\mathbb{C})$.

Compute $\det(A)$, $\det(B)$, $\det(AB)$ and $\det(A^3)$.

7.9 Coursework Sheet 8

Submit a single pdf with scans of your work to Blackboard by Monday, 10 May 2021, 17:00.

Exercise 1

Let F be a field and let $A \in M_{n\times n}(F)$.

(a) If $n = 2$ show that $p_A(\lambda) = \lambda^2 - \text{trace}(A)\lambda + \det(A)$. (*See Exercise 1 on Coursework Sheet 4 for the definition of trace*(A).)

(b) Let $k \geq 1$. Show that if λ is an eigenvalue of A then λ^k is an eigenvalue of A^k.

(c) Suppose that $F = \mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$ and that $A^2 = I_n$. Show that if λ is an eigenvalue of A then $\lambda = 1$ or $\lambda = -1$. Show that $\ker(L_{I_n+A}) = E_{-1}(A)$ and that $\text{im}(L_{I_n+A}) = E_1(A)$. (*Note: The notation "*L_{matrix}*" is from Example 4.3 (a).*)

Exercise 2

Let F be a field and let $A \in M_{n\times n}(F)$ be a diagonalizable matrix.

(a) Let $k \geq 1$. Show that A^k is diagonalizable.

(b) Show that the transpose A^T of A is diagonalizable.

(c) Show that if A is invertible then A^{-1} is diagonalizable.

Exercise 3

Find the eigenvalues of each of the following matrices and determine a basis of the eigenspace for each eigenvalue. Determine which of these matrices are diagonalizable; if so, write down a diagonalizing matrix.

$$A = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{pmatrix} \in M_{3\times 3}(\mathbb{R}), \quad B = \begin{pmatrix} 3 & -1 \\ 1 & 5 \end{pmatrix} \in M_{2\times 2}(\mathbb{R}),$$

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 2 \\ 1 & -1 & 1 \end{pmatrix} \text{ as element of } M_{3\times 3}(\mathbb{R}) \text{ and as element of } M_{3\times 3}(\mathbb{C}).$$

Compute C^{2020}.

Exercise 4

Let V be a vector space over a field F and let L, M be two linear transformations from V to itself.

(a) Suppose that $L \circ M = M \circ L$. Show that $L(E_\lambda(M)) \subseteq E_\lambda(M)$ for all $\lambda \in F$.

(b) Suppose that V is of finite dimension. Show that L is injective if and only if it is surjective.
