# Chapter 4

# Linear maps

Before concentrating on linear maps, we provide a more general setting.

## 4.1 General maps

We start with the general definition of a map between two sets, and introduce some notations.

**Definition 4.1.1.** *Let $S, S'$ be two sets.* A map T from $S$ to $S'$ *is a rule which associates to each element of $S$ an element of $S'$. The notation*

$$\mathrm{T} : S \ni X \mapsto \mathrm{T}(X) \in S'$$

*will be used for such a map. If $X \in S$, then $\mathrm{T}(X) \in S'$ is called* the image of $X$ by T. *The set $S$ is often called* the domain of T *and is also denoted by* Dom(T)*, while*

$$\mathrm{T}(S) := \big\{\mathrm{T}(X) \mid X \in S\big\}$$

*is often called* the range of T *and is also denoted by* Ran(T)*.*

**Examples 4.1.2.** *(i) The function $f : \mathbb{R} \ni x \mapsto f(x) = x^2 - 3x + 2 \in \mathbb{R}$ is a map from $\mathbb{R}$ to $\mathbb{R}$,*

*(ii) Any $\mathcal{A} \in M_{mn}(\mathbb{R})$ defines a map $\mathrm{L}_\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^m$ by $\mathrm{L}_\mathcal{A}(X) := \mathcal{A}X$ for any $X \in \mathbb{R}^n$. More generally, for any field $\mathbb{F}$ and any $\mathcal{A} \in M_{mn}(\mathbb{F})$, one defines a map $\mathrm{L}_\mathcal{A} : \mathbb{F}^n \to \mathbb{F}^m$ by $\mathrm{L}_\mathcal{A}(X) := \mathcal{A}X$ for any $X \in \mathbb{F}^n$,*

*(iii) The rule $\mathrm{F} : \mathbb{R}^3 \ni \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \mathrm{F}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x^2+y \\ x+y+z+3 \end{pmatrix} \in \mathbb{R}^2$ is a map,*

*(iv) Let $C^1(\mathbb{R}) := \{$continuous functions $f$ on $\mathbb{R} \mid f'$ exists and is continuous $\}$ and let $C(\mathbb{R}) := \{$continuous functions $f$ on $\mathbb{R}\}$. Then the following rule defines a map:*

$$\mathrm{D} : C^1(\mathbb{R}) \ni f \mapsto \mathrm{D}f = f' \in C(\mathbb{R})$$

*(v)* *For any* $\mathcal{A}, \mathcal{B} \in M_n(\mathbb{R})$ *one can define a map by*

$$\mathrm{T}_{\mathcal{A},\mathcal{B}} : M_n(\mathbb{R}) \ni X \mapsto \mathrm{T}_{\mathcal{A},\mathcal{B}}(X) = \mathcal{A}X + \mathcal{B} \in M_n(\mathbb{R}),$$

*(vi)* *The function* $g : \mathbb{R}^* \ni x \mapsto g(x) = \frac{3x-2}{x} \in \mathbb{R}$ *is a map from* $\mathbb{R}^*$ *to* $\mathbb{R}$*, but is not a map from* $\mathbb{R}$ *to* $\mathbb{R}$ *because* $g(0)$ *is not defined,*

*(vii)* *For any fixed* $Y \in \mathbb{R}^n$*, a map is defined by* $\mathrm{T}_Y : \mathbb{R}^n \ni X \mapsto \mathrm{T}_Y(X) = X + Y \in \mathbb{R}^n$*, and is called* the translation by $Y$.

**Remark 4.1.3.** *For a map* $\mathrm{T} : S \to S'$*, the determination of* $\mathrm{Ran}(\mathrm{T})$ *is not always an easy task. For example if one considers* $f : \mathbb{R} \to \mathbb{R}$ *with* $f(x) = x^2 - 3x + 2$*, then one has to look for the minimum of* $f$*, which is* $-1/4$ *obtained for* $x = 3/2$*, and one can then set* $\mathrm{Ran}(f) = [-1/4, \infty)$*. Similarly, if* $\mathcal{A} \in M_{mn}(\mathbb{R})$*, then what is the range of* $\mathrm{L}_{\mathcal{A}}$*,* i.e. *the set of* $Y \in \mathbb{R}^m$ *such that* $Y = \mathcal{A}X$ *for some* $X \in \mathbb{R}^n$ *?*

We end this section with a natural definition.

**Definition 4.1.4.** *Let* $\mathrm{T} : S \to S'$ *be a map, let* $W \subset S$ *be a subset of* $S$ *and let* $Z$ *be a subset of* $S'$*. Then the set* $\mathrm{T}(W) := \{\mathrm{T}(X) \mid X \in W\}$ *is called* the image of $W$ by T, *while the set*

$$\mathrm{T}^{-1}(Z) := \big\{X \in S \mid \mathrm{T}(X) \in Z\big\}$$

*is called* the preimage of $Z$ by T.

## 4.2 Linear maps

From now on, we shall concentrate on the simplest maps, the linear ones. Note that in order to state the next definition, one has to deal with vector spaces instead of arbitrary sets, and in addition the two vector spaces have to be defined on the same field.

**Definition 4.2.1.** *Let* $V, W$ *be two vector spaces over the same field* $\mathbb{F}$*. A map* $\mathrm{T} : V \to W$ *is a linear map if the following two conditions are satisfied:*

*(i)* $\mathrm{T}(X + Y) = \mathrm{T}(X) + \mathrm{T}(Y)$ *for any* $X, Y \in V$*,*

*(ii)* $\mathrm{T}(\lambda X) = \lambda \mathrm{T}(X)$ *for any* $X \in V$ *and* $\lambda \in \mathbb{F}$*.*

Note that the examples (*ii*) and (*iv*) of Examples 4.1.2 were already linear maps. Let us still mention the map $\mathrm{Id} : V \to V$ (also denoted by $\mathbf{1}$) defined by $\mathrm{Id}(X) = X$ for any $X \in V$, which is clearly linear, and the map $\mathcal{O} : V \to W$ defined by $\mathcal{O}(X) = \mathbf{0}$ for any $X \in V$, which is also linear.

Let us now observe that linear maps are rather simple maps.

**Lemma 4.2.2.** *Let* $V, W$ *be vector spaces over the same field* $\mathbb{F}$*, and let* $\mathrm{T} : V \to W$ *be a linear map. Then,* *(i)* $\mathrm{T}(\mathbf{0}) = \mathbf{0}$,

*(ii)* $\mathrm{T}(-X) = -\mathrm{T}(X)$ *for any* $X \in V$.

*Proof.* (i) It is sufficient to observe that

$$\mathrm{T}(\mathbf{0}) = \mathrm{T}(\mathbf{0}+\mathbf{0}) = \mathrm{T}(\mathbf{0}) + \mathrm{T}(\mathbf{0}) = 2\mathrm{T}(\mathbf{0})$$

which implies the result.

(i) Observe that

$$\mathbf{0} = \mathrm{T}(\mathbf{0}) = \mathrm{T}(X - X) = \mathrm{T}(X) + \mathrm{T}(-X)$$

which directly leads to the result. □

Let go us step further in abstraction and consider families of linear maps. For that purpose, let us first define an addition of linear maps, and the multiplication of a linear map by a scalar. Namely, if $V, W$ are vector spaces over the same field $\mathbb{F}$ and if $\mathrm{T}_1, \mathrm{T}_2$ are linear maps from $V$ to $W$, one sets

$$\big(\mathrm{T}_1 + \mathrm{T}_2\big)(X) = \mathrm{T}_1(X) + \mathrm{T}_2(X) \qquad \text{for any } X \in V. \tag{4.2.1}$$

If $\lambda \in \mathbb{F}$ and if $\mathrm{T} : V \to W$ is linear, one also sets

$$\big(\lambda\mathrm{T}\big)(X) = \lambda\mathrm{T}(X) \qquad \text{for any } X \in V. \tag{4.2.2}$$

It is then easily observed that $\mathrm{T}_1 + \mathrm{T}_2$ is still a linear map, and that $\lambda\mathrm{T}$ is also a linear map. We can then even say more:

**Proposition 4.2.3.** *Let* $V, W$ *be vector spaces over the same field* $\mathbb{F}$. *Then*

$$\mathcal{L}(V, W) := \big\{\mathrm{T} : V \to W \mid \mathrm{T} \textit{ is linear } \big\},$$

*is a vector space over* $\mathbb{F}$, *once endowed with the addition defined by* (4.2.1) *and the multiplication by a scalar defined in* (4.2.2).

Before giving the proof, let us observe that if $V = \mathbb{R}^n$ and $W = \mathbb{R}^n$, then $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ corresponds to the set of all $\mathrm{L}_\mathcal{A}$ with $\mathcal{A} \in M_{mn}(\mathbb{R})$. Note that this statement also holds for arbitrary field $\mathbb{F}$, *i.e.*

$$\mathcal{L}(\mathbb{F}^n, \mathbb{F}^m) = \big\{\mathrm{L}_\mathcal{A} \mid \mathcal{A} \in M_{mn}(\mathbb{F})\big\}.$$

*Proof.* The proof consists in checking all conditions of Definition 3.1.3. For that purpose, consider $\mathrm{T}, \mathrm{T}_1, \mathrm{T}_2, \mathrm{T}_3$ be linear maps from $V$ to $W$, and let $\lambda, \mu \in \mathbb{F}$. Let also $X$ be an arbitrary element of $V$.

(i) One has

$$\begin{aligned}\big[(\mathrm{T}_1 + \mathrm{T}_2) + \mathrm{T}_3\big](X) &= (\mathrm{T}_1 + \mathrm{T}_2)(X) + \mathrm{T}_3(X) = \mathrm{T}_1(X) + \mathrm{T}_2(X) + \mathrm{T}_3(X)\\ &= \mathrm{T}_1(X) + (\mathrm{T}_2 + \mathrm{T}_3)(X) = \big[\mathrm{T}_1 + (\mathrm{T}_2 + \mathrm{T}_3)\big](X).\end{aligned}$$

Since $X$ is arbitrary, it follows that $(\mathrm{T}_1 + \mathrm{T}_2) + \mathrm{T}_3 = \mathrm{T}_1 + (\mathrm{T}_2 + \mathrm{T}_3)$.

(ii) One has

$$(\mathrm{T}_1 + \mathrm{T}_2)(X) = \mathrm{T}_1(X) + \mathrm{T}_2(X) = \mathrm{T}_2(X) + \mathrm{T}_1(X) = (\mathrm{T}_2 + \mathrm{T}_1)(X)$$

which implies that $\mathrm{T}_1 + \mathrm{T}_2 = \mathrm{T}_2 + \mathrm{T}_1$.

(iii) We already know that $\mathcal{O} : V \to W$ is linear, which means that $\mathcal{O} \in \mathcal{L}(V, W)$. In addition, one clearly has $\mathrm{T} + \mathcal{O} = \mathcal{O} + \mathrm{T} = \mathrm{T}$.

(iv) By setting $[-\mathrm{T}](X) := -\mathrm{T}(X)$, one readily observes that $-\mathrm{T} \in \mathcal{L}(V, W)$ and by using the addition (4.2.1) one infers that $\mathrm{T} + (-\mathrm{T}) = \mathcal{O}$.

(v) Similarly, $\lambda\mathrm{T} \in \mathcal{L}(V, W)$ and $1\mathrm{T} = \mathrm{T}$.

The remaining three properties are easily checked by using the definition 4.2.2 and the basic properties of vector spaces. □

**Question** : If $\dim(V) = n$ and if $\dim(W) = m$, what is the dimension of $\mathcal{L}(V, W)$ ?

Let us now consider a linear map $\mathrm{T} : V \to \mathbb{R}^n$ with $V$ a real vector space. Since for each $X \in V$ one has $\mathrm{T}(X) \in \mathbb{R}^n$, one often sets

$$\mathrm{T}(X) = \begin{pmatrix} \mathrm{T}_1(X) \\ \vdots \\ \mathrm{T}_n(X) \end{pmatrix} \tag{4.2.3}$$

with $\mathrm{T}_j(X) := \mathrm{T}(X)_j$ the $j^{th}$ component of $\mathrm{T}$ evaluated at $X$. Thus, $\mathrm{T}$ defines a family of maps $\mathrm{T}_j : V \to \mathbb{R}$, and reciprocally, any family $\{\mathrm{T}_j\}_{j=1}^n$ with $\mathrm{T}_j : V \to \mathbb{R}$ defines a map $\mathrm{T} : V \to \mathbb{R}^n$ by (4.2.3). Sometimes, the maps $\mathrm{T}_1, \dots, \mathrm{T}_n$ are called *the components of* $\mathrm{T}$.

**Example 4.2.4.** *If* $\mathrm{T} : \mathbb{R}^2 \to \mathbb{R}^3$ *is defined by*

$$\mathrm{T}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x - y \\ 3x + 4y \\ x - 5y \end{pmatrix},$$

*then* $\mathrm{T} = {}^t(\mathrm{T}_1, \mathrm{T}_2, \mathrm{T}_3)$ *with* $\mathrm{T}_1\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = 2x - y$, $\mathrm{T}_2\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = 3x + 4y$ *and* $\mathrm{T}_3\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = x - 5y$.

More generally:

**Lemma 4.2.5.** *Let* $V$ *be a vector space over a field* $\mathbb{F}$, *and let* $\mathrm{T} : V \to \mathbb{F}^n$ *with* $\mathrm{T} = {}^t(\mathrm{T}_1, \dots, \mathrm{T}_n)$ *the components of* $\mathrm{T}$. *Then* $\mathrm{T}$ *is a linear map if and only if each* $\mathrm{T}_j$ *is a linear map.*

*Proof.* One has $\mathrm{T}(X + Y) = {}^t\big(\mathrm{T}_1(X + Y), \dots, \mathrm{T}_n(X + Y)\big)$ and $\mathrm{T}(X) + \mathrm{T}(Y) = {}^t\big(\mathrm{T}_1(X) + \mathrm{T}_1(Y), \dots, \mathrm{T}_n(X) + \mathrm{T}_n(Y)\big)$. It then follows that

$$\mathrm{T}(X + Y) = \mathrm{T}(X) + \mathrm{T}(Y) \Longleftrightarrow \begin{pmatrix} \mathrm{T}_1(X + Y) \\ \vdots \\ \mathrm{T}_n(X + Y) \end{pmatrix} = \begin{pmatrix} \mathrm{T}_1(X) + \mathrm{T}_1(Y) \\ \vdots \\ \mathrm{T}_n(X) + \mathrm{T}_n(Y) \end{pmatrix},$$

which corresponds to half of the statement. A similar argument holds for the multiplication by a scalar. □

## 4.3 Kernel and range of a linear map

Let $V, W$ be two vector spaces over the same field $\mathbb{F}$ and let $\mathrm{T} : V \to W$ be a linear map. Recall that

$$\mathrm{Ran}(\mathrm{T}) := \{Y \in W \mid Y = \mathrm{T}(X) \text{ for some } X \in V\}$$

and

$$\mathrm{Ker}(\mathrm{T}) := \{X \in V \mid \mathrm{T}(X) = \mathbf{0}\}.$$

**Lemma 4.3.1.** *In the previous setting,* $\mathrm{Ker}(\mathrm{T})$ *is a subspace of* $V$ *while* $\mathrm{Ran}(\mathrm{T})$ *is a subspace of* $W$*.*

*Proof.* The first part of the statement is proved in Exercise 4.4. For the second part of the statement, consider $Y_1, Y_2 \in \mathrm{Ran}(\mathrm{T})$, *i.e.* there exist $X_1, X_2 \in V$ such that $Y_1 = \mathrm{T}(X_1)$ and $Y_2 = \mathrm{T}(X_2)$. Then one has

$$Y_1 + Y_2 = \mathrm{T}(X_1) + \mathrm{T}(X_2) = \mathrm{T}(X_1 + X_2)$$

with $X_1 + X_2 \in V$. In other words, $Y_1 + Y_2$ belongs to $\mathrm{Ran}(\mathrm{T})$. Similarly, for $\lambda \in \mathbb{F}$ and any $Y = \mathrm{T}(X)$ with $X \in V$ one has

$$\lambda Y = \lambda \mathrm{T}(X) = \mathrm{T}(\lambda X)$$

with $\lambda X \in V$. Again, it follows that $\lambda Y \in \mathrm{Ran}(\mathrm{T})$, from which one concludes that $\mathrm{Ran}(\mathrm{T})$ is a subspace of $W$. □

**Examples 4.3.2.** *(i) Let* $N \in \mathbb{R}^n$ *with* $N \neq \mathbf{0}$*, and let us set* $\mathrm{T}_N : \mathbb{R}^n \to \mathbb{R}$ *by* $\mathrm{T}_N(X) = N \cdot X$*. In this case,* $\mathrm{T}_N$ *is a linear map. Indeed, one has*

$$\mathrm{T}_N(X + Y) = N \cdot (X + Y) = N \cdot X + N \cdot Y = \mathrm{T}_N(X) + \mathrm{T}_N(Y),$$

*and similarly* $\mathrm{T}_N(\lambda X) = N \cdot (\lambda X) = \lambda(N \cdot X) = \lambda \mathrm{T}_N(X)$*. Then one observes that*

$$\mathrm{Ker}(\mathrm{T}_N) = \{X \in \mathbb{R}^n \mid N \cdot X = 0\} = \{X \in \mathbb{R}^n \mid X \cdot N = \mathbf{0} \cdot N\} = H_{N,\mathbf{0}}.$$

*On the other hand,* $\mathrm{Ran}(\mathrm{T}_N) = \mathbb{R}$*, as it can easily be checked by considering elements* $X$ *of the form* $\lambda N$*, for any* $\lambda \in \mathbb{R}$*.*

*(ii) Let* $\mathcal{A} \in M_{mn}(\mathbb{R})$ *and let us set* $\mathrm{L}_{\mathcal{A}} : \mathbb{R}^n \to \mathbb{R}^m$ *defined by* $\mathrm{L}_{\mathcal{A}}(X) = \mathcal{A}X$ *for any* $X \in \mathbb{R}^n$*. As already mentioned, this map is linear, and one has* $\mathrm{Ker}(\mathrm{L}_{\mathcal{A}}) = \{X \in \mathbb{R}^n \mid \mathcal{A}X = \mathbf{0}\}$, i.e. $\mathrm{Ker}(\mathrm{L}_{\mathcal{A}})$ *are the solutions of the linear system* $\mathcal{A}X = \mathbf{0}$*.*

**Remark 4.3.3.** *The kernel of a linear map is never empty, indeed it always contains the element* $\mathbf{0}$*.*

**Lemma 4.3.4.** *Let* $\mathrm{T} : V \to W$ *be a linear map between vector spaces over the same field* $\mathbb{F}$*, and assume that* $\mathrm{Ker}(\mathrm{T}) = \{\mathbf{0}\}$*. If* $\{X_1, \dots, X_r\}$ *are linearly independent elements of* $V$*, then* $\{\mathrm{T}(X_1), \dots, \mathrm{T}(X_n)\}$ *are linearly independent elements of* $W$*.*

*Proof.* Let $\lambda_1, \dots, \lambda_n$ such that

$$\lambda_1 \mathrm{T}(X_1) + \lambda_2 \mathrm{T}(X_2) + \cdots + \lambda_n \mathrm{T}(X_n) = \mathbf{0}.$$

By linearity, this is equivalent to $\mathrm{T}\big(\lambda_1 X_1 + \cdots + \lambda_n X_n\big) = \mathbf{0}$, but since the kernel of T is reduced to $\mathbf{0}$ it means that $\lambda_1 X_1 + \cdots + \lambda_n X_n = \mathbf{0}$. Finally, by the linear independence of $X_1, \dots, X_n$ it follows that $\lambda_j = 0$ for any $j \in \{1, \dots, n\}$. As a consequence, the elements $\mathrm{T}(X_1), \dots, \mathrm{T}(X_n)$ of $W$ are linearly independent. □

Let us now come to an important result of this section. For this, we just recall that for a vector space, its dimension corresponds to the number of elements of any of its bases. It also corresponds to the maximal number of linearly independent elements of this vector space.

**Theorem 4.3.5.** *Let* $\mathrm{T} : V \to W$ *be a linear map between two vector spaces over the same field* $\mathbb{F}$*, and assume that* $V$ *is of finite dimension. Then*

$$\dim\big(\mathrm{Ker}(\mathrm{T})\big) + \dim\big(\mathrm{Ran}(\mathrm{T})\big) = \dim(V).$$

*Proof.* Let $\{Y_1, \dots, Y_n\}$ be a basis for $\mathrm{Ran}(\mathrm{T})$, and let $X_1, \dots, X_n \in V$ such that $\mathrm{T}(X_j) = Y_j$ for any $j \in \{1, \dots, n\}$. Let also $\{K_1, \dots, K_m\}$ be a basis for $\mathrm{Ker}(\mathrm{T})$. Note that if one shows that $\{X_1, \dots, X_n, K_1, \dots, K_m\}$ is a basis for $V$, then the statement is proved (with $\dim(V) = m + n$).

So, let $X$ be an arbitrary element of $V$. Then there exist $\lambda_1, \dots, \lambda_n \in \mathbb{F}$ such that $\mathrm{T}(X) = \lambda_1 Y_1 + \cdots + \lambda_n X_n$, since $\{Y_1, \dots, Y_n\}$ is a basis for $\mathrm{Ran}(\mathrm{T})$. It follows that

$$\begin{aligned}\mathbf{0} &= \mathrm{T}(X) - \lambda_1 X_1 - \cdots - \lambda_n Y_n \\ &= \mathrm{T}(X) - \lambda_1 \mathrm{T}(X_1) - \cdots - \lambda_n \mathrm{T}(X_n) \\ &= \mathrm{T}\big(X - \lambda_1 X_1 - \cdots - \lambda_n X_n\big),\end{aligned}$$

which means that $X - \lambda_1 X_1 - \cdots - \lambda_n X_n$ belongs to $\mathrm{Ker}(\mathrm{T})$. As a consequence, there exist $\lambda'_1, \dots, \lambda'_m \in \mathbb{F}$ such that

$$X - \lambda_1 X_1 - \cdots - \lambda_n X_n = \lambda'_1 K_1 + \cdots + \lambda'_m K_m,$$

since $\{K_1, \dots, K_m\}$ is a basis for $\mathrm{Ker}(\mathrm{T})$. Consequently, one gets

$$X = \lambda_1 X_1 + \cdots + \lambda_n X_n + \lambda'_1 K_1 + \cdots + \lambda'_m K_m,$$

or in other words

$$\mathrm{Vect}\big(X_1, \dots, X_n, K_1, \dots, K_m\big) = V.$$

Let us now show that these vectors are linearly independent. By contraposition, assume that

$$\lambda_1 X_1 + \cdots + \lambda_n X_n + \lambda'_1 K_1 + \cdots + \lambda'_m K_m = \mathbf{0} \tag{4.3.1}$$

for some $\lambda_1, \ldots, \lambda_n, \lambda'_1, \ldots, \lambda'_m$. Then one infers from (4.3.1) that

$$\begin{aligned}\mathbf{0} &= \mathrm{T}(\mathbf{0}) \\ &= \mathrm{T}\big(\lambda_1 X_1 + \cdots + \lambda_n X_n + \lambda'_1 K_1 + \cdots + \lambda'_m K_m\big) \\ &= \mathrm{T}\big(\lambda_1 X_1 + \cdots + \lambda_n X_n\big) + \mathbf{0} \\ &= \lambda_1 \mathrm{T}(X_1) + \cdots + \lambda_n \mathrm{T}(X_n) \\ &= \lambda_1 Y_1 + \cdots + \lambda_n Y_n.\end{aligned}$$

Since $Y_1, \ldots, Y_n$ are linearly independent, one already concludes that $\lambda_j = 0$ for any $j \in \{1, \ldots, n\}$. It then follows from (4.3.1) that $\lambda'_1 K_1 + \cdots + \lambda'_m K_m = \mathbf{0}$, which implies that $\lambda'_i = 0$ for any $i \in \{1, \ldots, m\}$ since the vectors $K_i$ are linearly independent.

In summary, one has shown that $V$ is generated by the family of linearly independent elements $X_1, \ldots, X_n, K_1, \ldots, K_m$ of $V$. Thus, these elements define a basis, as expected. □

## 4.4 Rank and linear maps

Let us come back to matrices over $\mathbb{F}$. For any $\mathcal{A} \in M_{mn}(\mathbb{F})$, recall that we denote by $\mathcal{A}^j$ the $j^{th}$ column of $\mathcal{A}$ and by $\mathcal{A}_k$ the $k^{th}$ row of $\mathcal{A}$. We also denote by $\mathrm{L}_\mathcal{A} : \mathbb{F}^n \to \mathbb{F}^m$ the linear map defined by $\mathrm{L}_\mathcal{A}(X) = \mathcal{A}X$. Observe finally that $\{E_j\}_{j=1}^n$ is a basis of $\mathbb{F}^n$ (note that the 1 at the entry $j$ of $E_j$ is the 1 of the field $\mathbb{F}$). Thus, for any $X \in \mathbb{F}^n$ one has

$$X = {}^t(x_1, \ldots, x_n) = x_1 E_1 + x_2 E_2 + \cdots + x_n E_n$$

and in addition

$$\begin{aligned}\mathrm{L}_\mathcal{A}(X) &= \mathcal{A}\big(x_1 E_1 + x_2 E_2 + \cdots + x_n E_n\big) \\ &= x_1 \mathcal{A}E_1 + x_2 \mathcal{A}E_2 + \cdots + x_n \mathcal{A}E_n \\ &= x_1 \mathcal{A}^1 + x_2 \mathcal{A}^2 + \ldots x_n \mathcal{A}^n.\end{aligned}$$

With such equalities, one directly infers the following statement:

**Lemma 4.4.1.** *The range of* $\mathrm{L}_\mathcal{A}$ *corresponds to the subspace generated by the columns of* $\mathcal{A}$.

*Proof.* It is enough to remember the following equality

$$\mathrm{Ran}(\mathrm{L}_\mathcal{A}) = \big\{\mathrm{L}_\mathcal{A}(X) \mid X \in \mathbb{F}^n\big\}$$

and to take into account the computation performed before the statement. □

Considering the dimensions of these spaces one directly gets:

**Corollary 4.4.2.** *The dimension of the range of the linear map* $\mathrm{L}_{\mathcal{A}}$ *is equal to the rank of* $\mathcal{A}$*,* i.e.

$$\dim\big(\mathrm{Ran}(\mathrm{L}_{\mathcal{A}})\big) = \mathrm{rank}(\mathcal{A}).$$

**Theorem 4.4.3.** *Let* $\mathcal{A} \in M_{mn}(\mathbb{F})$ *with* $\mathrm{rank}(\mathcal{A}) = r$*. Then one has* $\dim\big(\mathrm{Ker}(\mathrm{L}_{\mathcal{A}})\big) = n - r$*.*

*Proof.* Since $\mathrm{L}_{\mathcal{A}} : \mathbb{F}^n \to \mathbb{F}^m$ is a linear map, one has from Theorem 4.3.5

$$\dim\big(\mathrm{Ker}(\mathrm{L}_{\mathcal{A}})\big) + \underbrace{\dim\big(\mathrm{Ran}(\mathrm{L}_{\mathcal{A}})\big)}_{=r} = n,$$

from which the statement follows. □

**Example 4.4.4.** *What is the dimension of the space of solutions of the system*

$$\begin{cases} 2x_1 - x_2 + x_3 + 2x_4 = 0 \\ x_1 + x_2 - 2x_3 - x_4 = 0 \end{cases} \qquad ?$$

*Since this system is equivalent to* $\mathrm{L}_{\mathcal{A}}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ *with* $\mathcal{A} = \begin{pmatrix} 2 & -1 & 1 & 2 \\ 1 & 1 & -2 & -1 \end{pmatrix}$ *and since* $\mathrm{rank}(\mathcal{A}) = 2$*, one directly infers from the previous result that* $\dim\big(\mathrm{Ker}(\mathrm{L}_{\mathcal{A}})\big) = 4-2 = 2$*. This corresponds to the dimension of the space of solutions of the homogeneous equation.*

One ends up this section with an important result:

**Theorem 4.4.5.** *Let* $\mathcal{A} \in M_{mn}(\mathbb{F})$ *and* $B \in \mathbb{F}^m$*, and consider the equation* $\mathcal{A}X = B$ *for some* $X \in \mathbb{F}^n$*. If this equation has a solution, then its set of all solutions is of dimension equal to* $\dim\big(\mathrm{Ker}(\mathrm{L}_{\mathcal{A}})\big)$*.*

*Proof.* Assume that $Y_0 \in \mathbb{F}^n$ satisfies $\mathcal{A}Y_0 = B$. Then if $Y \in \mathbb{F}^n$ satisfies $\mathcal{A}Y = \mathbf{0}$, one infers that $\mathcal{A}(Y_0 + Y) = B$, which means that $Y_0 + Y$ is a solution of the original problem, for any $Y \in \mathrm{Ker}(\mathrm{L}_{\mathcal{A}})$. Now, if one can show that all solutions of $\mathcal{A}X = B$ are of the form $X = Y_0 + Y$ for some $Y \in \mathrm{Ker}(\mathrm{L}_{\mathcal{A}})$, then the statement is proved. For that purpose, it is sufficient to observe that if $X \in \mathbb{F}^n$ satisfies $\mathcal{A}X = B$, then one has $\mathcal{A}(X - Y_0) = B - B = \mathbf{0}$, or in other words $X - Y_0 =: Y$ for some $Y \in \mathrm{Ker}(\mathrm{L}_{\mathcal{A}})$. As a consequence, one infers that $X = Y_0 + Y$ with $Y \in \mathrm{Ker}(\mathrm{L}_{\mathcal{A}})$, as expected. □

## 4.5 Matrix associated with a linear map

Let us start with a question: If $V, W$ are vector spaces over a field $\mathbb{F}$ and if $\mathrm{T} : V \to W$ is a linear map, how can one associate with this linear map a matrix ?

In fact, this can be done only once a choice of bases for $V$ and $W$ has been done, and the resulting matrix will depend on the choice of bases, as we shall see. So, let us introduce a new notation: a basis for a vector space $V$ over $\mathbb{F}$ will be denoted by $\mathcal{V} := \{V_1, \dots, V_n\}$ with $\{V_1, \dots, V_n\}$ a family of linearly independent elements of $V$ which generate $V$. In addition, let us denote by $\mathcal{X}$ an arbitrary element of $V$ (which was simply denoted by $X$ up to now). Then, since $\mathcal{V}$ is a basis for $V$ there exists $X := {}^t(x_1, \dots, x_n) \in \mathbb{F}^n$ such that

$$\mathcal{X} = x_1V_1 + x_2V_2 + \dots + x_nV_n.$$

The vector $X \in \mathbb{F}^n$ is called *the coordinate vector of $\mathcal{X}$ with respect to the basis $\mathcal{V}$ of $V$*, and we shall use the notation

$$(\mathcal{X})_{\mathcal{V}} = X$$

meaning precisely that the coordinates of $\mathcal{X}$ with respect to the basis $\mathcal{V}$ are $X$.

**Remark 4.5.1.** *Clearly, if $V = \mathbb{R}^n$ and if $V_j = E_j$, one just says that $X$ are the coordinates of $\mathcal{X}$ and one uses to identify $\mathcal{X}$ and $X$. This is what we have done until now since we have only considered the usual basis $\{E_j\}_{j=1}^n$ on $\mathbb{R}^n$. However, if one needs to consider different bases on $\mathbb{R}^n$, the above notations are necessary. Note for example that $\mathcal{X}$ exists without any choice of a particular basis, while $X$ depends on such a choice.*

Now, if $\mathcal{Y}$ is another element of $V$ with $(\mathcal{Y})_{\mathcal{V}} = Y = {}^t(y_1, \dots, y_n)$, let us observe that

$$(\mathcal{X} + \mathcal{Y})_{\mathcal{V}} = X + Y \qquad \text{and} \qquad (\lambda\mathcal{X})_{\mathcal{V}} = \lambda X \tag{4.5.1}$$

for any $\lambda \in \mathbb{F}$. Indeed, this follows from the equalities

$$\begin{aligned}\mathcal{X} + \mathcal{Y} &= x_1V_1 + \dots + x_nV_n + y_1V_1 + \dots + y_nV_n \\ &= (x_1 + y_1)V_1 + \dots + (x_n + y_n)V_n\end{aligned}$$

and

$$\lambda\mathcal{X} = \lambda(x_1V_1 + \dots + x_nV_n) = (\lambda x_1)V_1 + \dots + (\lambda x_n)V_n.$$

Thus, choosing a basis $\mathcal{V}$ for $V$ allows one to identity any point of $V$ with an element of $\mathbb{F}^n$ via its coordinate vector. By taking (4.5.1) into account, one also observes that $\mathcal{V}$ allows one to define a linear map $(\cdot)_{\mathcal{V}} : V \to \mathbb{F}^n$.

We also consider a vector space $W$ over $\mathbb{F}$ endowed with a basis $\mathcal{W} := \{W_1, \dots, W_m\}$. In this case, for any $\mathcal{Z} \in \mathcal{W}$ we set $(\mathcal{Z})_{\mathcal{W}} = Z = {}^t(z_1, \dots, z_m) \in \mathbb{F}^m$ for the coordinate vector of $\mathcal{Z}$ with respect to the basis $\mathcal{W}$ of $W$. Thus, if $\mathrm{T} : V \to W$ is a linear map, there exists $\mathcal{T} := (t_{ij}) \in M_{mn}(\mathbb{F})$, called *the matrix associated with* T *with respect to the basis $\mathcal{V}$ of $V$ and $\mathcal{W}$ of $W$* defined by

$$\mathrm{T}(V_j) = \sum_{i=1}^m t_{ij}W_i = \sum_{i=1}^m {}^tt_{ji}W_i \tag{4.5.2}$$

for any $j \in \{1, \dots, n\}$. On the other hand, we shall show just below that the following equality also holds

$$\big(\mathrm{T}(\mathcal{X})\big)_{\mathcal{W}} = \mathcal{T}(\mathcal{X})_{\mathcal{V}}. \tag{4.5.3}$$

In other words, the action of T on a basis of $V$ is given in terms of ${}^t\mathcal{T}$ by relation (4.5.2), while the action of T on the coordinate vectors is given in terms of $\mathcal{T}$ by relation (4.5.3). Note that this is related to the more general notion of *covariant* or *contravariant* transformations.

For the proof of (4.5.3) it is enough to observe that one has

$$\begin{aligned}\mathrm{T}(\mathcal{X}) = \mathrm{T}\Big(\sum_{j=1}^n x_j V_j\Big) &= \sum_{j=1}^n x_j \mathrm{T}(V_j)\\ &= \sum_{j=1}^n x_j \sum_{i=1}^m t_{ij} W_i = \sum_{i=1}^m \Big(\sum_{j=1}^n t_{ij}\, x_j\Big) W_i = \sum_{i=1}^m (\mathcal{T}X)_i W_i,\end{aligned}$$

which implies that

$$\big(\mathrm{T}(\mathcal{X})\big)_{\mathcal{W}} = \mathcal{T}X = \mathcal{T}(\mathcal{X})_{\mathcal{V}}. \tag{4.5.4}$$

**Example 4.5.2.** *If $V = \mathbb{R}^n$, $W = \mathbb{R}^m$ and $V_j = E_j$ while $W_i = E_i$ for any $j \in \{1,\dots,n\}$ and $j \in \{1,\dots,m\}$, and if* T *is a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$ then one deduces from* (4.5.2) *that*

$$\mathrm{T}(E_j) = \sum_{i=1}^m t_{ij} E_i = \begin{pmatrix} t_{1j} \\ t_{2j} \\ \vdots \\ t_{mj} \end{pmatrix} = \mathcal{T}^j$$

*where $\mathcal{T}^j$ corresponds to the $j^{th}$ column of the matrix $\mathcal{T}$. In other words one has*

$$\mathcal{T} = \big(\mathrm{T}(E_1)\ \mathrm{T}(E_2)\ \dots\ \mathrm{T}(E_n)\big)\,.$$

**Example 4.5.3.** *If $V$ is a real vector space with basis $\mathcal{V} = \{V_1, V_2, V_3\}$ and if $\mathrm{T} : V \to V$ is the linear map such that*

$$\mathrm{T}(V_1) = 2V_1 - V_2, \quad \mathrm{T}(V_2) = V_1 + V_2 - 4V_3, \quad \mathrm{T}(V_3) = 5V_1 + 4V_2 + 2V_3,$$

*then the matrix associated with* T *with respect to the basis $\mathcal{V}$ is given by*

$$\mathcal{T} = \begin{pmatrix} 2 & 1 & 5 \\ -1 & 1 & 4 \\ 0 & -4 & 2 \end{pmatrix}.$$

Let us still consider the notion of a change of basis. Indeed, given the matrix associated to a linear map in a prescribed basis, it is natural to wonder about the matrix associated to the same linear map but with respect to another basis. So, let $\mathcal{V} = \{V_1, \dots, V_n\}$ and $\mathcal{V}' = \{V_1', \dots, V_n'\}$ be two basis of the same vector space $V$. Let $\mathcal{B} = (b_{ij}) \in M_n(\mathbb{F})$ be the matrix defined by

$$V_j' = \sum_{i=1}^n b_{ij} V_i \equiv \sum_{i=1}^n {}^t b_{ji} V_i.$$

It is easily observed that the matrix $\mathcal{B}$ is invertible. Then, for any $\mathcal{X} \in V$ with $X = (\mathcal{X})_{\mathcal{V}}$ and $X' = (\mathcal{X})_{\mathcal{V}'}$, one has

$$\sum_{i=1}^{n} x_i V_i = \mathcal{X} = \sum_{j=1}^{n} x'_j V'_j = \sum_{j=1}^{n} x'_j \sum_{i=1}^{n} b_{ij} V_i = \sum_{i=1}^{n} \Big( \sum_{j=1}^{n} b_{ij} x'_j \Big) V_i .$$

Since the vectors $V_1, \dots, V_n$ are linearly independent, this implies that

$$X = \mathcal{B} X' \quad \text{or equivalently} \quad (\mathcal{X})_{\mathcal{V}} = \mathcal{B}(\mathcal{X})_{\mathcal{V}'}. \tag{4.5.5}$$

Let us now consider a linear map $\mathrm{T} : V \to V$, and let $\mathcal{T}$ be the matrix associated with T with respect to the basis $\mathcal{V}$, and let $\mathcal{T}'$ be the matrix associated to T with respect to the basis $\mathcal{V}'$. The original question corresponds then to the link between $\mathcal{T}$ and $\mathcal{T}'$ ? In order to answer this question, observe that for any $\mathcal{X} \in V$ one gets by equations (4.5.4) and (4.5.5) that

$$\mathcal{T}\mathcal{B}X' = \mathcal{T}X = \big(\mathrm{T}(\mathcal{X})\big)_{\mathcal{V}} = \mathcal{B}\big(\mathrm{T}(\mathcal{X})\big)_{\mathcal{V}'} = \mathcal{B}\mathcal{T}'X'.$$

Since $X'$ is arbitrary, one infers that $\mathcal{T}\mathcal{B} = \mathcal{B}\mathcal{T}'$, or equivalently

$$\mathcal{T}' = \mathcal{B}^{-1}\mathcal{T}\mathcal{B}. \tag{4.5.6}$$

One deduces in particular that the matrix $\mathcal{T}$ and $\mathcal{T}'$ are similar, see Definition 2.1.16.

Note that a similar (but slightly more complicated) computation can be realized for a linear map between two vector spaces $V$ and $W$ over the same field $\mathbb{F}$ endowed with two different bases $\mathcal{V}, \mathcal{V}'$ and $\mathcal{W}, \mathcal{W}'$.

## 4.6 Composition of linear maps

Let us now consider three sets $U, V, W$ and let $\mathrm{F} : U \to V$ and $\mathrm{G} : V \to W$ be maps. Then the map

$$\mathrm{G} \circ \mathrm{F} : U \to W,$$

defined by $\big(\mathrm{G} \circ \mathrm{F}\big)(X) = \mathrm{G}\big(\mathrm{F}(X)\big)$ for any $X \in U$, is called *the composition map of* F *with* G. Notice that if $W \not\subset U$ the composition map $\mathrm{F} \circ \mathrm{G}$ has simply no meaning.

**Examples 4.6.1.** *(i) Let $U = V = W = \mathbb{R}$ and $\mathrm{F}, \mathrm{G}$ be two real functions defined on $\mathbb{R}$. Then $\mathrm{G} \circ \mathrm{F}$ just corresponds to the composition of functions.*

*(ii) If $U = \mathbb{R}^n$, $V = \mathbb{R}^m$, $W = \mathbb{R}^p$, $\mathcal{A} \in M_{mn}(\mathbb{R})$ and $\mathcal{B} \in M_{pm}(\mathbb{R})$, then for any $X \in \mathbb{R}^n$ one has*

$$\big(\mathrm{L}_{\mathcal{B}} \circ \mathrm{L}_{\mathcal{A}}\big)(X) = \mathrm{L}_{\mathcal{B}}\big(\mathrm{L}_{\mathcal{A}}(X)\big) = \mathcal{B}\mathcal{A}X = (\mathcal{B}\mathcal{A})X = \mathrm{L}_{\mathcal{B}\mathcal{A}}(X). \tag{4.6.1}$$

Let us now observe an important property of the composition of maps, namely the associativity. Indeed, If $U, V, W, S$ are sets and $\mathrm{F} : U \to V$, $\mathrm{G} : V \to W$ and $\mathrm{H} : W \to S$ are maps, one has

$$(\mathrm{H} \circ \mathrm{G}) \circ \mathrm{F} = \mathrm{H} \circ (\mathrm{G} \circ \mathrm{F}).$$

Indeed, for any $X \in U$ one has

$$\big[(\mathrm{H} \circ \mathrm{G}) \circ \mathrm{F}\big](X) = \big(\mathrm{H} \circ \mathrm{G}\big)(\mathrm{F}(X)) = \mathrm{H}\big(\mathrm{G}(\mathrm{F}(X))\big)$$

and

$$\big[\mathrm{H} \circ (\mathrm{G} \circ \mathrm{F})\big](X) = \mathrm{H}\big((\mathrm{G} \circ \mathrm{F})(X)\big) = \mathrm{H}\big(\mathrm{G}(\mathrm{F}(X))\big),$$

and the equality of the two right hand sides implies the statement.

**Lemma 4.6.2.** *Let $U, V, W$ be vector spaces over a field $\mathbb{F}$, and let $\mathrm{G} : U \to V$, $\mathrm{G}' : U \to V$, $\mathrm{H} : V \to W$ and $\mathrm{H}' : V \to W$ be linear maps. Then*

*(i)* $\mathrm{H} \circ \mathrm{G} : U \to W$ *is a linear map,*

*(ii)* $(\mathrm{H} + \mathrm{H}') \circ \mathrm{G} = \mathrm{H} \circ \mathrm{G} + \mathrm{H}' \circ \mathrm{G}$*,*

*(iii)* $\mathrm{H} \circ (\mathrm{G} + \mathrm{G}') = \mathrm{H} \circ \mathrm{G} + \mathrm{H} \circ \mathrm{G}'$*,*

*(iv)* $(\lambda \mathrm{H}) \circ \mathrm{G} = \mathrm{H} \circ (\lambda \mathrm{G}) = \lambda(\mathrm{H} \circ \mathrm{G})$, *for all* $\lambda \in \mathbb{F}$.

The proof will be provided in Exercise 4.17.

**Remark 4.6.3.** *If $V$ is a vector space and if $\mathrm{T} : V \to V$ is a linear map, then $\mathrm{T}^n = \underbrace{\mathrm{T} \circ \mathrm{T} \cdots \circ \mathrm{T}}_{n \text{ terms}}$ is a linear map from $V$ to $V$. By convention, one sets $\mathrm{T}^0 = \mathbf{1}$, and observes that $\mathrm{T}^{r+s} = \mathrm{T}^r \circ \mathrm{T}^s = \mathrm{T}^s \circ \mathrm{T}^r$.*

## 4.7 Inverse of a linear map

**Definition 4.7.1.** *For a map $\mathrm{F} : V \to W$ between two sets $V$ and $W$, one says that $\mathrm{F}$ has* an inverse *if there exists $\mathrm{G} : W \to V$ such that $\mathrm{G} \circ \mathrm{F} = \mathbf{1}_V$ and $\mathrm{F} \circ \mathrm{G} = \mathbf{1}_W$. In this case, one also says that $\mathrm{F}$ is* invertible *and write $\mathrm{F}^{-1}$ for this inverse.*

**Example 4.7.2.** *If $\mathcal{A} \in M_n(\mathbb{R})$ is invertible, then the linear map $\mathrm{L}_\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^n$ is invertible, with inverse $\mathrm{L}_{\mathcal{A}^{-1}}$. This follows from equation* (4.6.1)*, or more precisely*

$$\mathrm{L}_\mathcal{A} \circ \mathrm{L}_{\mathcal{A}^{-1}} = \mathrm{L}_{\mathcal{A}\mathcal{A}^{-1}} = \mathbf{1} = \mathrm{L}_{\mathcal{A}^{-1}\mathcal{A}} = \mathrm{L}_{\mathcal{A}^{-1}} \circ \mathrm{L}_\mathcal{A}.$$

Due to the following lemma, there is no ambiguity in speaking about the inverse (and not only about an inverse) of a invertible map.

**Lemma 4.7.3.** *Let $\mathrm{F} : V \to W$ be an invertible map between two sets $V$ et $W$. Then this inverse is unique.*

*Proof.* Let us assume that there exists $\mathrm{G} : W \to V$ and $\mathrm{G}' : W \to V$ such that $\mathrm{G} \circ \mathrm{F} = \mathbf{1}_V$, $\mathrm{F} \circ \mathrm{G} = \mathbf{1}_W$, $\mathrm{G}' \circ \mathrm{F} = \mathbf{1}_V$, and $\mathrm{F} \circ \mathrm{G}' = \mathbf{1}_W$. Then one gets

$$\mathrm{G} = \mathbf{1}_V \circ \mathrm{G} = (\mathrm{G}' \circ \mathrm{F}) \circ \mathrm{G} = \mathrm{G}' \circ (\mathrm{F} \circ \mathrm{G}) = \mathrm{G}' \circ \mathbf{1}_W = \mathrm{G}'$$

from which the result follows. □

Let us now come to two more refined notions related to a maps, linear or not.

**Definition 4.7.4.** *A map* $\mathrm{F} : V \to W$ *between two sets is* injective *or* one-to-one *if* $\mathrm{F}(X_1) \neq \mathrm{F}(X_2)$ *whenever* $X_1, X_2 \in V$ *with* $X_1 \neq X_2$. *The map* $\mathrm{F}$ *is called* surjective *if for any* $Y \in W$ *there exists at least one* $X \in V$ *such that* $\mathrm{F}(X) = Y$. *The map* $\mathrm{F}$ *is* bijective *if it is both injective and surjective.*

The following result links the notions of invertibility and of bijectivity.

**Theorem 4.7.5.** *A map* $\mathrm{F} : V \to W$ *between two sets is invertible if and only if* $\mathrm{F}$ *is bijective.*

*Proof.* (i) Assume first that F is bijective. In particular, since F is surjective, for any $Y \in W$, there exists $X \in V$ such that $\mathrm{F}(X) = Y$. Note that $X$ is unique because F is also injective. Thus if one sets $\mathrm{F}^{-1}(Y) := X$ then one has

$$(\mathrm{F}^{-1} \circ \mathrm{F})(X) = \mathrm{F}^{-1}\big(\mathrm{F}(X)\big) = \mathrm{F}^{-1}(Y) = X$$

which implies that $\mathrm{F}^{-1} \circ \mathrm{F} = \mathbf{1}_V$, and similarly

$$\big(\mathrm{F} \circ \mathrm{F}^{-1}\big)(Y) = \mathrm{F}\big(\mathrm{F}^{-1}(Y)\big) = \mathrm{F}(X) = Y$$

which implies that $\mathrm{F} \circ \mathrm{F}^{-1} = \mathbf{1}_W$. One has thus define an inverse for F.

(ii) Let us now assume that F is invertible, with inverse denoted by $\mathrm{F}^{-1}$. Let first $X_1, X_2 \in V$ with $\mathrm{F}(X_1) = \mathrm{F}(X_2)$. One then deduces that

$$X_1 = \mathbf{1}_V X_1 = \big(\mathrm{F}^{-1} \circ \mathrm{F}\big)X_1 = \mathrm{F}^{-1}\big(\mathrm{F}(X_1)\big) = \mathrm{F}^{-1}\big(\mathrm{F}(X_2)\big) = \big(\mathrm{F}^{-1} \circ \mathrm{F}\big)(X_2) = X_2,$$

and thus F is injective. Secondly, let $Y \in W$, and observe that

$$Y = \mathbf{1}_W Y = \big(\mathrm{F} \circ \mathrm{F}^{-1}\big)(Y) = \mathrm{F}\big(\mathrm{F}^{-1}(Y)\big)$$

which implies that $Y = \mathrm{F}(X)$ for $X$ given by $\mathrm{F}^{-1}(X)$. Thus F is surjective. Since F is both injective and surjective, F is bijective. □

For linear maps the general theory simplifies a lot, as we shall see now.

**Theorem 4.7.6.** *Let* $V, W$ *be two vector spaces over the same field* $\mathbb{F}$, *and let* $\mathrm{T} : V \to W$ *be an invertible linear map. Then its inverse* $\mathrm{T}^{-1} : W \to V$ *is also a linear map.*

*Proof.* Let $Y_1, Y_2 \in W$ and set $X_1 := \mathrm{T}^{-1}(Y_1)$ and $X_2 := \mathrm{T}^{-1}(X_2)$. Since $\mathrm{T} \circ \mathrm{T}^{-1} = \mathbf{1}_W$ one has for $j \in \{1, 2\}$

$$Y_j = \big(\mathrm{T} \circ \mathrm{T}^{-1}\big)(Y_j) = \mathrm{T}\big(\mathrm{T}^{-1}(Y_j)\big) = \mathrm{T}(X_j).$$

Then, one infers that

$$\begin{aligned} \mathrm{T}^{-1}\big(Y_1 + Y_2\big) &= \mathrm{T}^{-1}\big(\mathrm{T}(X_1) + \mathrm{T}(X_2)\big) \underbrace{=}_{linearity} \mathrm{T}^{-1}\big(\mathrm{T}(X_1 + X_2)\big) \\ &= \mathrm{T}^{-1} \circ \mathrm{T}(X_1 + X_2) = X_1 + X_2 = \mathrm{T}^{-1}(Y_1) + \mathrm{T}^{-1}(Y_2). \end{aligned} \tag{4.7.1}$$

Similarly one has for any $\lambda \in \mathbb{F}$ and $Y \in W$ (with $Y := \mathrm{T}(X)$)

$$\begin{aligned} \mathrm{T}^{-1}(\lambda Y) &= \mathrm{T}^{-1}\big(\lambda \mathrm{T}(X)\big) \underbrace{=}_{linearity} \mathrm{T}^{-1}\big(\mathrm{T}(\lambda X)\big) \\ &= \big(\mathrm{T}^{-1} \circ \mathrm{T}\big)(\lambda X) = \lambda X = \lambda \mathrm{T}^{-1}(Y). \end{aligned} \tag{4.7.2}$$

It is then sufficient to observe that (4.7.1) and (4.7.2) correspond to the linearity conditions for $\mathrm{T}^{-1}$. □

In the next statement we give an equivalent property for the injectivity of a linear map.

**Lemma 4.7.7.** *A linear map* $\mathrm{T} : V \to W$ *between two vector spaces over the same field is injective if and only if* $\mathrm{Ker}(\mathrm{T}) = \{\mathbf{0}\}$*.*

*Proof.* (i) The first part of the proof is a contraposition argument: instead of proving $A \Rightarrow B$ we show equivalently that $\bar{B} \Rightarrow \bar{A}$. Thus, let us assume first that $\mathrm{Ker}(\mathrm{T}) \neq \{\mathbf{0}\}$, then there exists $X_0 \neq \mathbf{0}$ such that $\mathrm{T}(X_0) = \mathbf{0}$. In addition, for any $X \in V$ one has

$$\mathrm{T}(X + X_0) = \mathrm{T}(X) + \mathrm{T}(X_0) = \mathrm{T}(X) + \mathbf{0} = \mathrm{T}(X).$$

Since $X \neq X + X_0$ but $\mathrm{T}(X) = \mathrm{T}(X + X_0)$, one concludes that T is not injective. By contraposition, one has shown that T injective implies that $\mathrm{Ker}(\mathrm{T}) = \{\mathbf{0}\}$.

(ii) Assume now that $\mathrm{Ker}(\mathrm{T}) = \{\mathbf{0}\}$, and consider $X_1, X_2 \in V$ with $X_1 \neq X_2$. Then one has

$$\mathrm{T}(X_1) - \mathrm{T}(X_2) = \mathrm{T}(X_1 - X_2) \neq \mathbf{0}$$

since $X_1 - X_2 \neq \mathbf{0}$. As a consequence, $\mathrm{T}(X_1) \neq \mathrm{T}(X_2)$. □

Let us provide a final theorem for this section, which is useful in many situations.

**Theorem 4.7.8.** *Let* $\mathrm{T} : V \to W$ *be a linear map between the vector spaces* $V$ *and* $W$*, and assume that* $\dim(V) = \dim(W) < \infty$*. Then the following assertions are equivalent:*

*(i)* $\mathrm{Ker}(\mathrm{T}) = \{\mathbf{0}\}$,

*(ii)* T *is invertible,*

*(iii)* T *is surjective.*

*Proof.* The implication $(ii) \Rightarrow (i)$ and $(ii) \Rightarrow (iii)$ are direct consequences of Theorem 4.7.5 and Lemma 4.7.7.

Assume now $(i)$, and recall from Lemma 4.7.7 that this condition corresponds to T injective. Then from Theorem 4.3.5 and more precisely from the equality

$$\underbrace{\dim\big(\mathrm{Ker}(\mathrm{T})\big)}_{0} + \dim\big(\mathrm{Ran}(\mathrm{T})\big) = \dim(V)$$

one deduces that $\dim\big(\mathrm{Ran}(\mathrm{T})\big) = \dim(V) = \dim(W)$, where the assumption about the dimension has been taken into account. It is enough then to observe that

$$\dim\big(\mathrm{Ran}(\mathrm{T})\big) = \dim(W)$$

means that T is surjective. Since T is also injective, it follows that T is bijective. Since bijectivity corresponds to invertibility by Theorem 4.7.5, one infers that $(ii)$ holds.

Assume now that $(iii)$ holds. By taking again Theorem 4.3.5 into account, one deduces that from the equality

$$\dim\big(\mathrm{Ran}(\mathrm{T})\big) = \dim(W) = \dim(V)$$

that $\dim\big(\mathrm{Ker}(\mathrm{T})\big) = 0$, meaning that T is injective. Again, it implies that T is bijective, and thus invertible, and thus that $(ii)$ holds. □

**Corollary 4.7.9.** *For any $\mathcal{A} \in M_n(\mathbb{F})$, the following statements are equivalent:*

*(i) There exists $\mathcal{B} \in M_n(\mathbb{F})$ such that $\mathcal{B}\mathcal{A} = \mathbf{1}_n$,*

*(ii) There exists $\mathcal{C} \in M_n(\mathbb{F})$ such that $\mathcal{A}\mathcal{C} = \mathbf{1}_n$.*

*In addition, whenever $(i)$ or $(ii)$ holds, then $\mathcal{B} = \mathcal{C}$, and $\mathcal{A}$ is invertible with $\mathcal{A}^{-1} = \mathcal{B} = \mathcal{C}$.*

## 4.8 Exercises

**Exercise 4.1.** *Let* $\mathrm{F}:\mathbb{R}^2\to\mathbb{R}^2$ *be the map defined by* $\mathrm{F}\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)=\left(\begin{smallmatrix}2x\\3y\end{smallmatrix}\right)$ *for any* $\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)\in\mathbb{R}^2$. *Describe the image by* F *of the points lying on the unit circle centered at* $\mathbf{0}$*, i.e.* $\left\{\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)\in\mathbb{R}^2\mid x^2+y^2=1\right\}$.

**Exercise 4.2.** *Let* $\mathrm{F}:\mathbb{R}^2\to\mathbb{R}^2$ *be the map defined by* $\mathrm{F}\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)=\left(\begin{smallmatrix}xy\\y\end{smallmatrix}\right)$ *for any* $\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)\in\mathbb{R}^2$. *Describe the image by* F *of the line* $\{\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)\in\mathbb{R}^2\mid x=2\}$.

**Exercise 4.3.** *Let* $V$ *be a vector space of dimension* $n$*, and let* $\{X_1,\dots,X_n\}$ *be a basis for* $V$*. Let* F *be a linear map from* $V$ *into itself. Show that* F *is uniquely defined if one knows* $\mathrm{F}(X_j)$ *for* $j\in\{1,\dots,n\}$*. Is it also true if* F *is an arbitrary map from* $V$ *into itself ?*

**Exercise 4.4.** *Let* $V,W$ *be vector spaces over the same field, and let* $\mathrm{T}:V\to W$ *be a linear map. Show that the following set is a subspace of* $V$*:*

$$\{X\in V\mid \mathrm{T}(X)=\mathbf{0}\}.$$

*This subspace is called the kernel of* T.

**Exercise 4.5.** *Show that the image of a convex set under a linear map is a convex set.*

**Exercise 4.6.** *Determine which of the following maps are linear:*

*a)* $\mathrm{F}:\mathbb{R}^3\to\mathbb{R}^2$ *defined by* $\mathrm{F}\begin{pmatrix}x\\y\\z\end{pmatrix}=\left(\begin{smallmatrix}x\\z\end{smallmatrix}\right)$,

*b)* $\mathrm{F}:\mathbb{R}^4\to\mathbb{R}^4$ *defined by* $\mathrm{F}(X)=-X$ *for all* $X\in\mathbb{R}^4$,

*c)* $\mathrm{F}:\mathbb{R}^3\to\mathbb{R}^3$ *defined by* $\mathrm{F}(X)=X+\begin{pmatrix}0\\-1\\0\end{pmatrix}$ *for all* $X\in\mathbb{R}^3$,

*d)* $\mathrm{F}:\mathbb{R}^2\to\mathbb{R}^2$ *defined by* $\mathrm{F}\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)=\left(\begin{smallmatrix}2x\\y-x\end{smallmatrix}\right)$,

*e)* $\mathrm{F}:\mathbb{R}^2\to\mathbb{R}^2$ *defined by* $\mathrm{F}\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)=\left(\begin{smallmatrix}y\\x\end{smallmatrix}\right)$,

*f)* $\mathrm{F}:\mathbb{R}^2\to\mathbb{R}$ *defined by* $\mathrm{F}\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)=xy$.

**Exercise 4.7.** *Determine the kernel and the range of the maps defined in the previous exercise.*

**Exercise 4.8.** *Consider the subset of* $\mathbb{R}^n$ *consisting of all vectors* ${}^t(x_1,\dots,x_n)$ *such that* $x_1+x_2+\dots+x_n=0$*. Is it a subspace of* $\mathbb{R}^n$ *? If so, what is its dimension ?*

**Exercise 4.9.** *Let* $\mathrm{P}:M_n(\mathbb{R})\to M_n(\mathbb{R})$ *be the map defined for any* $\mathcal{A}\in M_n(\mathbb{R})$ *by*

$$\mathrm{P}(\mathcal{A})=\frac{1}{2}\big(\mathcal{A}+{}^t\!\mathcal{A}\big).$$

1. *Show that* P *is a linear map.*

2. *Show that the kernel of* $\mathrm{P}$ *consists in the vector space of all skew-symmetric matrices.*

3. *Show that the range of* $\mathrm{P}$ *consists in the vector space of all symmetric matrices.*

4. *What is the dimension of the vector space of all symmetric matrices, and the dimension of the vector space of all skew-symmetric matrices ?*

**Exercise 4.10.** *Let* $C^\infty(\mathbb{R})$ *be the vector space of all real functions on* $\mathbb{R}$ *which admit derivatives of all orders. Let* $\mathrm{D} : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ *be the map which associates to any* $f \in C^\infty(\mathbb{R})$ *its derivative, i.e.* $\mathrm{D}f = f'$.

1. *Is* $\mathrm{D}$ *a linear map ?*

2. *What is the kernel of* $\mathrm{D}$ *?*

3. *What is the kernel of* $\mathrm{D}^n$, *for any* $n \in \mathbb{N}$, *and what is the dimension of this vector space ?*

**Exercise 4.11.** *Consider the map* $\mathrm{F} : \mathbb{R}^3 \to \mathbb{R}^4$ *defined by*

$$\mathrm{F}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ x-y \\ x-z \\ x-y-z \end{pmatrix}.$$

1. *Is* $\mathrm{F}$ *a linear map ? (Justify your answer)*

2. *Determine the kernel of* $\mathrm{F}$.

3. *Determine the range of* $\mathrm{F}$.

**Exercise 4.12.** *What is the dimension of the space of solutions of the following systems of linear equations ? In each case, find a basis for the space of solutions.*

$$a) \begin{cases} 2x+y-z &= 0 \\ 2x+y+z &= 0 \end{cases} \qquad b) \begin{cases} x-y+z &= 0 \end{cases} \qquad c) \begin{cases} 4x+7y-\pi z &= 0 \\ 2x-y+z &= 0 \end{cases}$$

*and*

$$d) \begin{cases} x+y+z &= 0 \\ x-y &= 0 \\ y+z &= 0 \end{cases}$$

**Exercise 4.13.** *Let* $\mathcal{A}$ *be the matrix given by* $\mathcal{A} = \begin{pmatrix} 0 & 1 & 3 & -2 \\ 2 & 1 & -4 & 3 \\ 2 & 3 & 2 & -1 \end{pmatrix}$ *and consider the linear map* $\mathrm{L}_\mathcal{A} : \mathbb{R}^4 \to \mathbb{R}^3$ *defined by* $\mathrm{L}_\mathcal{A} X = \mathcal{A}X$ *for all* $X \in \mathbb{R}^4$.

1. *Determine the rank of* $\mathcal{A}$ *and the dimension of the range of* $\mathrm{L}_\mathcal{A}$.

2. *Deduce the dimension of the kernel of* $L_{\mathcal{A}}$*, and exhibit a basis for the kernel of* $L_{\mathcal{A}}$.

3. *Find the set of all solutions of the equation* $\mathcal{A}X = \left(\begin{smallmatrix} 0 \\ 2 \\ 2 \end{smallmatrix}\right)$.

**Exercise 4.14.** *Let* $F : \mathbb{R}^3 \to \mathbb{R}^2$ *be the map indicated below. What is the matrix associated with* F *in the canonical bases of* $\mathbb{R}^3$ *and* $\mathbb{R}^2$ *?*

$$a) \quad F(E_1) = \begin{pmatrix} 1 \\ -3 \end{pmatrix}, \ F(E_2) = \begin{pmatrix} -4 \\ 2 \end{pmatrix}, \ F(E_3) = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

*and*

$$b) \quad F\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3x_1 - 2x_2 + x_3 \\ 4x_1 - x_2 + 5x_3 \end{pmatrix}.$$

**Exercise 4.15.** *Let* $L : \mathbb{R}^3 \to \mathbb{R}^3$ *be a linear map which associated matrix has the form* $\left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{smallmatrix}\right)$ *with respect to the canonical basis of* $\mathbb{R}^3$. *What is the matrix associated with* L *in the basis generated by the three vectors* $V_1 = \left(\begin{smallmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{smallmatrix}\right)$, $V_2 = \left(\begin{smallmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{smallmatrix}\right)$, $V_3 = \left(\begin{smallmatrix} 0 \\ 0 \\ -1 \end{smallmatrix}\right)$

**Exercise 4.16.** *For any* $\mathcal{A}, \mathcal{B} \in M_n(\mathbb{R})$*, one says that* $\mathcal{A}$ *and* $\mathcal{B}$ commute *if* $\mathcal{AB} = \mathcal{BA}$.

a) *Show that the set of all matrices which commute with* $\mathcal{A}$ *is a subspace of* $M_n(\mathbb{R})$,

b) *If* $\mathcal{A} = \left(\begin{smallmatrix} 0 & 1 \\ 2 & 3 \end{smallmatrix}\right)$*, exhibit a basis of the subspace of all matrices which commute with* $\mathcal{A}$.

**Exercise 4.17.** *Let* $U, V, W$ *be vector spaces over a field* $\mathbb{F}$*, and let* $G : U \to V$, $G' : U \to V$, $H : V \to W$ *and* $H' : V \to W$ *be linear maps. Show that*

(i) $H \circ G : U \to W$ *is a linear map,*

(ii) $(H + H') \circ G = H \circ G + H' \circ G$,

(iii) $H \circ (G + G') = H \circ G + H \circ G'$,

(iv) $(\lambda H) \circ G = H \circ (\lambda G) = \lambda(H \circ G)$, *for all* $\lambda \in \mathbb{F}$.

**Exercise 4.18.** *Let* $V$ *be a real vector space, and let* $P : V \to V$ *be a linear map satisfying* $P^2 = P$. *Such a linear map is called* a projection.

(i) *Show that* $\mathbf{1} - P$ *is also a projection, and that* $(\mathbf{1} - P)P = P(\mathbf{1} - P) = \mathbf{0}$,

(ii) *Show that* $V = \mathrm{Ker}(P) + \mathrm{Ran}(P)$,

(iii) *Show that the intersection of* $\mathrm{Ker}(P)$ *and* $\mathrm{Ran}(P)$ *is* $\{\mathbf{0}\}$.

**Exercise 4.19.** *Let* $\mathrm{L} : \mathbb{R}^2 \to \mathbb{R}^2$ *be the linear map defined by* $\mathrm{L}\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \left(\begin{smallmatrix} x+y \\ x-y \end{smallmatrix}\right)$. *Show that* L *is invertible and find its inverse. Same question with the map* $\mathrm{L} : \mathbb{R}^3 \to \mathbb{R}^3$ *defined by* $\mathrm{L}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x-y \\ x+z \\ x+y+3z \end{pmatrix}$.

**Exercise 4.20.** *Let* $\mathrm{F}, \mathrm{G}$ *be invertible linear maps from a vector space into itself. Show that* $(\mathrm{G} \circ \mathrm{F})^{-1} = \mathrm{F}^{-1} \circ \mathrm{G}^{-1}$.

**Exercise 4.21.** *Show that the matrix* $\mathcal{B} : \mathbb{R}^n \to \mathbb{R}^n$ *defining a change of basis in* $\mathbb{R}^n$ *is always invertible.*

**Exercise 4.22.** *Let* $V$ *be the set of all infinite sequences of real numbers* $(x_1, x_2, x_3, \dots)$. *We endow* $V$ *with the pointwise addition and multiplication, i.e.*

$$(x_1, x_2, x_3, \dots) + (x_1', x_2', x_3', \dots) = (x_1 + x_1', x_2 + x_2', x_3 + x_3', \dots)$$

*and* $\lambda(x_1, x_2, x_3, \dots) = (\lambda x_1, \lambda x_2, \lambda x_3, \dots)$, *which make* $V$ *an infinite dimensional vector space.*

*Define the map* $\mathrm{F} : V \to V$, *called* shift operator, *by*

$$\mathrm{F}(x_1, x_2, x_3, \dots) = (0, x_1, x_2, x_3, \dots).$$

*(i) Is* F *a linear map ?*

*(ii) Is* F *injective, and what is the kernel of* F *?*

*(iii) Is* F *surjective ?*

*(iv) Show that there is a linear map* $\mathrm{G} : V \to V$ *such that* $\mathrm{G} \circ \mathrm{F} = \mathbf{1}$.

*(v) Does the map* G *have the property that* $\mathrm{F} \circ \mathrm{G} = \mathbf{1}$ *?*

*(vi) What is different from the finite dimensional case, i.e. when* $V$ *is of finite dimension ?*

**Exercise 4.23.** *Consider the matrices*

$$\mathcal{A} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad \mathcal{B} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathcal{C} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathcal{D} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$
$$\mathcal{E} = \begin{pmatrix} 1 & 0.2 \\ 0 & 1 \end{pmatrix}, \quad \mathcal{F} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

*and show their effect on the letter* $L$ *defined by the three points* $\left(\begin{smallmatrix} 0 \\ 2 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ *of* $\mathbb{R}^2$.

**Exercise 4.24.** *Let* $N = \left(\begin{smallmatrix} n_1 \\ n_2 \end{smallmatrix}\right)$ *be a vector in* $\mathbb{R}^2$ *with* $\|N\| = 1$, *and let* $\ell$ *be the line in* $\mathbb{R}^2$ *passing trough* $\mathbf{0} \in \mathbb{R}^2$ *and parallel to* $N$. *Then any vector* $X \in \mathbb{R}^2$ *can be written uniquely as* $X = X_\parallel + X_\perp$, *where* $X_\parallel$ *is a vector parallel to* $\ell$ *and* $X_\perp$ *is a vector perpendicular to* $\ell$. *Show that there exists a projection* $\mathrm{P} \in M_2(\mathbb{R})$ *such that* $X_\parallel = \mathrm{P}X$, *and express* P *in terms of* $n_1$ *and* $n_2$.

**Exercise 4.25.** *1) Do the same exercise in* $\mathbb{R}^3$ *with* $N$ *given by* $\begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}$.

*2) Show that there also exists a projection* $\mathrm{Q}$ *such that* $X_\perp = \mathrm{Q}\, X$. *If* $H_{\mathbf{0},N}$ *is the plane passing through* $\mathbf{0} \in \mathbb{R}^3$ *and perpendicular to* $N$, *show that* $X_\perp \in H_{\mathbf{0},N}$.

**Exercise 4.26.** *In the framework of the previous exercise, a reflection of* $X$ *about* $H_{\mathbf{0},N}$ *is defined by the vector* $X_{\mathrm{ref}} := X_\perp - X_\parallel$. *Show that* $\|X_{\mathrm{ref}}\| = \|X\|$, *and provide the expression for the linear map transforming* $X$ *into* $X_{\mathrm{ref}}$.

**Exercise 4.27.** *Prove Corollary 4.7.9.*

**Exercise 4.28.** *Block matrices are matrices which are partitioned into rectangular sub-matrices called* blocks. *For example, let* $\mathcal{A} \in M_{n+m}(\mathbb{R})$ *be the block matrix*

$$\mathcal{A} = \begin{pmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{pmatrix}$$

*with* $\mathcal{A}_{11} \in M_n(\mathbb{R})$, $\mathcal{A}_{22}(\mathbb{R}) \in M_m(\mathbb{R})$, $\mathcal{A}_{12} \in M_{n\times m}(\mathbb{R})$, *and* $\mathcal{A}_{21} \in M_{m\times n}(\mathbb{R})$. *Such matrices can be multiplied as if every blocks where scalars (with the usual multiplication of matrices), as long as the products are well defined. For example, check this statement by computing the product* $\mathcal{AB}$ *in two different ways with the following matrices:* $\mathcal{A} = \begin{pmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \end{pmatrix}$ *with* $\mathcal{A}_{11} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ *and* $\mathcal{A}_{12} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, *and* $\mathcal{B} = \begin{pmatrix} \mathcal{B}_{11} & \mathcal{B}_{12} \\ \mathcal{B}_{21} & \mathcal{B}_{22} \end{pmatrix}$ *with* $\mathcal{B}_{11} = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}$, $\mathcal{B}_{12} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$, $\mathcal{B}_{21} = \begin{pmatrix} 7 & 8 \end{pmatrix}$, *and* $\mathcal{B}_{22} = \begin{pmatrix} 9 \end{pmatrix}$.

**Exercise 4.29.** *Let* $\mathcal{A} \in M_{n+m}(\mathbb{R})$ *be the block matrix*

$$\mathcal{A} = \begin{pmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{O} & \mathcal{A}_{22} \end{pmatrix}$$

*with* $\mathcal{A}_{11} \in M_n(\mathbb{R})$, $\mathcal{A}_{22}(\mathbb{R}) \in M_m(\mathbb{R})$ *and* $\mathcal{A}_{12} \in M_{n\times m}(\mathbb{R})$.

*(i) For which choice of* $\mathcal{A}_{11}$, $\mathcal{A}_{12}$ *and* $\mathcal{A}_{22}$ *is* $\mathcal{A}$ *invertible ?*

*(ii) If* $\mathcal{A}$ *is invertible, what is* $\mathcal{A}^{-1}$, *in terms of* $\mathcal{A}_{11}$, $\mathcal{A}_{12}$ *and* $\mathcal{A}_{22}$ *?*