Chapter 7
Linear Transformations

Together with the theory of linear equations and matrices, the notion of linear transformations is crucial in both classical and quantum physics. In this chapter we introduce them and study their main properties.

7.1 Linear Transformations and Matrices

We have already seen that differently looking sets may have the same vector space structure. In this chapter we study mappings between vector spaces which are, in a proper sense, compatible with the vector space structure. The action of such maps will be represented by matrices.

Example 7.1.1 Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2,2}$. Let us define the map $f : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$f(X) = AX$$

where $X = (x, y)$ is a (column) vector representing a generic element in $\mathbb{R}^2$ and AX denotes the usual row by column product, that is

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

With $X = (x_1, x_2)$ and $Y = (y_1, y_2)$ two elements in $\mathbb{R}^2$, using the properties of the matrix calculus it is easy to show that

$$f(X + Y) = A(X + Y) = AX + AY = f(X) + f(Y)$$

© Springer International Publishing AG, part of Springer Nature 2018
G. Landi and A. Zampini, *Linear Algebra and Analytic Geometry for Physical Sciences*, Undergraduate Lecture Notes in Physics,
https://doi.org/10.1007/978-3-319-78361-1_7

as well as, with $\lambda \in \mathbb{R}$, that

$$f(\lambda X) = A(\lambda X) = \lambda A(X) = \lambda f(X).$$

This example is easily generalised to matrices of arbitrary dimensions.

Exercise 7.1.2 Given $A = (a_{ij}) \in \mathbb{R}^{m,n}$ one considers the map $f : \mathbb{R}^n \to \mathbb{R}^m$

$$f(X) = AX,$$

with $X = {}^t(x_1, \dots, x_n)$ and AX the usual row by column product. The above properties are easily generalised so this map satisfies the identities $f(X+Y) = f(X) + f(Y)$ for any $X, Y \in \mathbb{R}^n$ and $f(\lambda X) = \lambda f(X)$ for any $X \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$.

Example 7.1.3 Let $A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & -1 & 0 \end{pmatrix} \in \mathbb{R}^{2,3}$. The associated map $f : \mathbb{R}^3 \to \mathbb{R}^2$ is given by

$$f((x, y, z)) = \begin{pmatrix} 1 & 2 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + z \\ x - y \end{pmatrix}.$$

The above lines motivate the following.

Definition 7.1.4 Let V and W be two vector spaces over $\mathbb{R}$. A map $f : V \to W$ is called *linear* if the following properties hold:

(L1) $f(X+Y) = f(X) + f(Y)$ for all $X, Y \in V$,
(L2) $f(\lambda X) = \lambda f(X)$ for all $X \in V$, $\lambda \in \mathbb{R}$.

The proof of the following identities is immediate.

Proposition 7.1.5 *If $f : V \to W$ is a linear map then,*

(a) $f(0_V) = 0_W$,
(b) $f(-v) = -f(v)$ *for any* $v \in V$,
(c) $f(a_1 v_1 + \cdots + a_p v_p) = a_1 f(v_1) + \cdots + a_p f(v_p)$, *for any* $v_1, \dots, v_p \in V$ *and* $a_1, \dots, a_p \in \mathbb{R}$.

Proof (a) Since $0_V = 0_{\mathbb{R}} 0_V$ the (L2) defining property gives

$$f(0_V) = f(0_{\mathbb{R}} 0_V) = 0_{\mathbb{R}} f(0_V) = 0_W.$$

(b) Since $-v = (-1)v$, again from (L2) we have

$$f(-v) = f((-1)v) = (-1) f(v) = -f(v).$$

(c) This is proved by induction on p. If $p = 2$ the claim follows directly from (L1) and (L2) with

$$f(a_1v_1 + a_2v_2) = f(a_1v_1) + f(a_2v_2) = a_1 f(v_1) + a_2 f(v_2).$$

Let us assume it to be true for $p-1$. By setting $w = a_1v_1 + \cdots + a_{p-1}v_{p-1}$, we have

$$f(a_1v_1 + \cdots + a_pv_p) = f(w + a_pv_p) = f(w) + f(a_pv_p) = f(w) + a_p f(v_p)$$

(the first equality follows from (L1), the second from (L2)). From the induction hypothesis, we have $f(w) = a_1 f(v_1) + \cdots + a_{p-1} f(v_{p-1})$, so

$$f(a_1v_1 + \cdots + a_pv_p) = f(w) + a_p f(v_p) = a_1 f(v_1) + \cdots + a_{p-1} f(v_{p-1}) + a_p f(v_p),$$

which is the statement for p.

□

Example 7.1.6 The Example 7.1.1 and the Exercise 7.1.2 show how one associates a linear map between $\mathbb{R}^n$ and $\mathbb{R}^m$ to a matrix $A \in \mathbb{R}^{m,n}$. This construction can be generalised by using bases for vector spaces V and W.

Let us consider a basis $\mathcal{B} = (v_1, \ldots, v_n)$ for V and a basis $\mathcal{C} = (w_1, \ldots, w_m)$ for W. Given the matrix $A = (a_{ij}) \in \mathbb{R}^{m,n}$ we define $f : V \to W$ as follows. For any $v \in V$ we have uniquely $v = x_1v_1 + \cdots + x_nv_n$, that is $v = (x_1, \ldots, x_n)_{\mathcal{B}}$. With $X = {}^t(x_1, \ldots, x_n)$, we consider the vector $AX \in \mathbb{R}^m$ with $AX = {}^t(y_1, \ldots, y_m)_{\mathcal{C}}$. We write then

$$f(v) = y_1w_1 + \cdots + y_mw_m$$

which can be written as

$$f((x_1, \ldots, x_n)_{\mathcal{B}}) = \left(A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right)_{\mathcal{C}}.$$

Exercise 7.1.7 Let us consider the matrix $A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & -1 & 0 \end{pmatrix} \in \mathbb{R}^{2,3}$, with $V = \mathbb{R}[X]_2$ and $W = \mathbb{R}[X]_1$. With respect to the bases $\mathcal{B} = (1, X, X^2)$ for V and $\mathcal{C} = (1, X)$ for W the map corresponding to A as in the previous example is

$$f(a + bX + cX^2) = (A\, {}^t(a, b, c))_{\mathcal{C}}$$

that is

$$f(a + bX + cX^2) = (a + 2b + c, a - b)_{\mathcal{C}} = a + 2b + c + (a - b)X.$$

Proposition 7.1.8 *The map $f : V \to W$ defined in the Example 7.1.6 is linear.*

Proof Let $v, v' \in V$ with $v = (x_1, \ldots, x_n)_{\mathcal{B}}$ and $v' = (x'_1, \ldots, x'_n)_{\mathcal{B}}$. From the Remark 2.4.16 we have

$$v + v' = (x_1 + x'_1, \ldots, x_n + x'_n)_{\mathcal{B}}$$

so we get

$$\begin{aligned} f(v + v') &= (A^t(x_1 + x'_1, \ldots, x_n + x'_n))_{\mathcal{C}} \\ &= (A^t(x_1, \ldots, x_n))_{\mathcal{C}} + (A^t(x'_1, \ldots, x'_n))_{\mathcal{C}} \\ &= f(v) + f(v') \end{aligned}$$

(notice that the second equality follows from the Proposition 4.1.10). Along the same line one shows easily that for any $\lambda \in \mathbb{R}$ one has $f(\lambda v) = \lambda f(v)$. □

The following definition (a rephrasing of Example 7.1.6) plays a central role in the theory of linear transformations.

Definition 7.1.9 With V and W two vector spaces over $\mathbb{R}$ and bases $\mathcal{B} = (v_1, \ldots, v_n)$ for V and $\mathcal{C} = (w_1, \ldots, w_m)$ for W, consider a matrix $A = (a_{ij}) \in \mathbb{R}^{m,n}$. The linear map

$$f_A^{\mathcal{C},\mathcal{B}} \; : \; V \;\to\; W$$

defined by

$$V \;\ni\; v = x_1 v_1 + \cdots + x_n v_n \;\mapsto\; f_A^{\mathcal{B},\mathcal{C}}(v) = y_1 w_1 + \cdots + y_m w_m \;\in\; W$$

with

$${}^t(y_1, \ldots, y_m) = A\,{}^t(x_1, \ldots, x_n),$$

is the *linear map corresponding to the matrix A with respect to the basis $\mathcal{B}$ e $\mathcal{C}$.*

Remark 7.1.10 Denoting $f_A^{\mathcal{C},\mathcal{B}} = f$, one immediately sees that the n columns in A provide the components with respect to $\mathcal{C}$ in W of the vectors $f(v_1), \ldots, f(v_n)$, with $(v_1, \ldots, v_n)$ the basis $\mathcal{B}$ for V. One has

$$v_1 = 1v_1 + 0v_2 + \cdots + 0v_n = (1, 0, \ldots, 0)_{\mathcal{B}},$$

thus giving

$$\begin{aligned} f(v_1) = (A\,{}^t(1, 0, \ldots, 0))_{\mathcal{C}} &= {}^t(a_{11}, \ldots, a_{m1})_{\mathcal{C}} \\ &= f(v_1) = a_{11} w_1 + \cdots + a_{m1} w_m. \end{aligned}$$

It is straightforward now to show that $f(v_j) = (a_{1j}, \ldots, a_{mj})_{\mathcal{C}}$ for any index j.

If $A = (a_{ij}) \in \mathbb{R}^{m,n}$ and $f : \mathbb{R}^n \to \mathbb{R}^m$ is the linear map defined by $f(X) = AX$, then the columns of A give the images under f of the vectors $(e_1, \ldots e_n)$ of the canonical basis $\mathcal{E}_n$ in $\mathbb{R}^n$. This can be written as

$$A \;=\; \big(f(e_1)\;f(e_2)\;\cdots\;f(e_n)\big)\,.$$

Exercise 7.1.11 Let us consider the matrix

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

with $\mathcal{B} = \mathcal{C} = \mathcal{E}_3$ the canonical basis in $\mathbb{R}^3$, and the corresponding linear map $f = f_A^{\mathcal{E}_3,\mathcal{E}_3} : \mathbb{R}^3 \to \mathbb{R}^3$. If $(x, y, z) \in \mathbb{R}^3$ then $f((x, y, z)) = A\,{}^t(x, y, z)$. The action of f is then given by

$$f((x, y, z)) = (x + y - z, y + 2z, x + y).$$

Being $\mathcal{B}$ the canonical basis, it is also

$$f(e_1) = (1, 0, 1), \quad f(e_2) = (1, 1, 1), \quad f(e_3) = (-1, 2, 0).$$

We see that $f(e_1)$, $f(e_2)$, $f(e_3)$ are the columns of A. This is not an accident: as mentioned the columns of A are, in the general situation, the components of $f(e_1)$, $f(e_2)$, $f(e_3)$ with respect to a basis $\mathcal{C}$—in this case the canonical one.

The Proposition 7.1.8 shows that, given a matrix A, the map $f_A^{\mathcal{C},\mathcal{B}}$ is linear. Our aim is now to prove that for *any* linear map $f : V \to W$ there exists a matrix A such that $f = f_A^{\mathcal{B},\mathcal{C}}$, with respect to two given bases $\mathcal{B}$ and $\mathcal{C}$ for V and W respectively.

In order to determine such a matrix we use the Remark 7.1.10: given a matrix A the images under $f_A^{\mathcal{C},\mathcal{B}}$ of the elements in the basis $\mathcal{B}$ of V are given by the column elements in A. This suggests the following definition.

Definition 7.1.12 Let $\mathcal{B} = (v_1, \dots, v_n)$ be a basis for the real vector space V and $\mathcal{C} = (w_1, \dots, w_m)$ a basis for the real vector space W. Let $f : V \to W$ be a linear map. The matrix *associated* to f with respect to the basis $\mathcal{B}$ and $\mathcal{C}$, that we denote by $M_f^{\mathcal{C},\mathcal{B}}$, is the element in $\mathbb{R}^{m,n}$ whose columns are given by the components with respect to $\mathcal{C}$ of the images under f of the basis elements in $\mathcal{B}$. That is, the matrix $M_f^{\mathcal{C},\mathcal{B}} = A = (a_{ij})$ is given by

$$\begin{aligned} f(v_1) &= a_{11}w_1 + \cdots + a_{m1}w_m \\ &\vdots \\ f(v_n) &= a_{1n}w_1 + \cdots + a_{mn}w_m, \end{aligned}$$

which can be equivalently written as

$$M_f^{\mathcal{C},\mathcal{B}} = (f(v_1), \quad \dots \quad , f(v_n)).$$

Such a definition inverts the one given in the Definition 7.1.9. This is the content of the following proposition, whose proof we omit.

Proposition 7.1.13 *Let V be a real vector space with basis $\mathcal{B} = (v_1, \ldots, v_n)$ and W a real vector space with basis $\mathcal{C} = (w_1, \ldots, w_m)$. The following results hold.*

(i) If $f : V \to W$ is a linear map, by setting $A = M_f^{\mathcal{C},\mathcal{B}}$ it holds that

$$f_A^{\mathcal{C},\mathcal{B}} = f.$$

(ii) If $A \in \mathbb{R}^{m,n}$, by setting $f = f_A^{\mathcal{C},\mathcal{B}}$ it holds that

$$M_f^{\mathcal{C},\mathcal{B}} = A.$$

Proposition 7.1.14 *Let V and W be two real vector spaces with $(v_1, \ldots, v_n)$ a basis for V. For any choice of $\{u_1, \ldots, u_n\}$ of n elements in W there exists a unique linear map $f : V \to W$ such that $f(v_j) = u_j$ for any $j = 1, \ldots, n$.*

Proof To define such a map one uses that any vector $v \in V$ can be written uniquely as

$$v = a_1 v_1 + \cdots + a_n v_n$$

with respect to the basis $(v_1, \ldots, v_n)$. By setting

$$f(v) = a_1 f(v_1) + \cdots + a_n f(v_n) = a_1 u_1 + \cdots + a_n u_n$$

we have a linear (by construction) map f that satisfies the required condition $f(v_j) = u_j$ for any $j \in 1, \ldots, n$.

Let us now suppose this map is not unique and that there exists a second linear map $g : V \to W$ with $g(v_j) = u_j$. From the Proposition 7.1.5 we could then write

$$g(v) = a_1 g(v_1) + \cdots + a_n g(v_n) = a_1 u_1 + \cdots + a_n u_n = f(v),$$

thus getting $g = f$. □

What we have discussed so far gives two equivalent ways to define a linear map between two vector spaces V and W.

I. Once a basis $\mathcal{B}$ for V, a basis $\mathcal{C}$ for W and a matrix $A = (a_{ij}) \in \mathbb{R}^{m,n}$ are fixed, from the Proposition 7.1.13 we know that the linear map $f_A^{\mathcal{C},\mathcal{B}}$ is uniquely determined.
II. Once a basis $\mathcal{B} = (v_1, \ldots, v_n)$ for V and n vectors $\{u_1, \ldots, u_n\}$ in W are fixed, we know from the Proposition 7.1.14 that there exists a unique linear map $f : V \to W$ with $f(v_j) = u_j$ for any $j = 1, \ldots, n$.

From now on, if $V = \mathbb{R}^n$ and $\mathcal{B} = \mathcal{E}$ is its canonical basis we shall denote by $f((x_1, \ldots, x_n))$ what we have previously denoted as $f((x_1, \ldots, x_n)_{\mathcal{B}})$. Analogously, with $\mathcal{C} = \mathcal{E}$ the canonical basis for $W = \mathbb{R}^m$ we shall write $(y_1, \ldots, y_m)$ instead of $(y_1, \ldots, y_m)_{\mathcal{C}}$.

With such a notation, if $f : \mathbb{R}^n \to \mathbb{R}^m$ is the linear map which, with respect to the canonical basis for both vector spaces corresponds to the matrix A, its action is written as

$$f((x_1, \dots, x_n)) = A^{\,t}(x_1, \dots, x_n)$$

or equivalently

$$f((x_1, \dots, x_n)) = (a_{11}x_1 + \cdots + a_{1n}x_n, \dots, a_{m1}x_1 + \cdots + a_{mn}x_n).$$

Exercise 7.1.15 Let $f_0 : V \to W$ be the *null (zero) map*, that is $f_0(v) = 0_W$ for any $v \in V$. With $\mathcal{B}$ and $\mathcal{C}$ arbitrary bases for V and W respectively, it is clearly

$$M_{f_0}^{\mathcal{C},\mathcal{B}} = 0_{\mathbb{R}^{m,n}},$$

that is the null matrix.

Exercise 7.1.16 If $\mathrm{id}_V(v) = v$ is the identity map on V then, using any basis $\mathcal{B} = (v_1, \dots, v_n)$ for V, one has the following expression

$$\mathrm{id}_V(v_j) = v_j = (0, \dots, 0, \underbrace{1}_{j}, 0, \dots, 0)_{\mathcal{B}}$$

for any $j = 1, \dots, n$. That is $M_{\mathrm{id}_V}^{\mathcal{B},\mathcal{B}}$ is the identity matrix I_n. Notice that $M_{\mathrm{id}_V}^{\mathcal{C},\mathcal{B}} \neq I_n$ if $\mathcal{B} \neq \mathcal{C}$.

Exercise 7.1.17 Let us consider for $\mathbb{R}^3$ both the canonical basis $\mathcal{E}_3 = (e_1, e_2, e_3)$ and the basis $\mathcal{B} = (v_1, v_2, v_3)$ with

$$v_1 = (0, 1, 1),\ \ v_2 = (1, 0, 1),\ \ v_3 = (1, 1, 0).$$

A direct computation gives

$$M_{\mathrm{id}}^{\mathcal{E}_3,\mathcal{B}} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \qquad M_{\mathrm{id}}^{\mathcal{E}_3,\mathcal{B}} = \frac{1}{2}\begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$$

and each of these matrices turns out to be the inverse of the other, that is $M_{\mathrm{id}}^{\mathcal{E}_3,\mathcal{B}}\, M_{\mathrm{id}}^{\mathcal{B},\mathcal{E}_3} = I_n$.

7.2 Basic Notions on Maps

Before we proceed we recall in a compact and direct way some of the basic notions concerning injectivity, surjectivity and bijectivity of mappings between sets.

Definition 7.2.1 Let X and Y be two non empty sets and $f : X \to Y$ a map between them. The element $f(x)$ in Y is called the *image* under f of the element $x \in X$. The set

$$\mathrm{Im}(f) = \{y \in Y \mid \exists\, x \in X \ : \ y = f(x)\}$$

is called the *image* (or *range*) of f in Y. The set (that might be empty)

$$f^{-1}(y) = \{x \in X \ : \ f(x) = y\}.$$

defines the *pre-image* of the element $y \in Y$.

Definition 7.2.2 Let X and Y be two non empty sets, with a map $f : X \to Y$. One says that:

(i) f is *injective* if, for any pair $x_1,\ x_2 \in X$ with $x_1 \neq x_2$, it is $f(x_1) \neq f(x_2)$,
(ii) f is *surjective* if $\mathrm{Im}(f) = Y$,
(iii) f is *bijective* if f is both injective and surjective.

Definition 7.2.3 Let $f : X \to Y$ and $g : Y \to Z$ be two maps. The composition of g with f is the map

$$g \circ f : X \to Z$$

defined as $(g \circ f)(x) = g(f(x))$ for any $x \in X$.

Definition 7.2.4 A map $f : X \to Y$ is *invertible* if there exists a map $g : Y \to X$ such that $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$. In such a case the map g is called the *inverse* of f and denoted by f^{-1}. It is possible to prove that, if f is invertible, then f^{-1} is unique.

Proposition 7.2.5 *A map $f : X \to Y$ is invertible if and only if it is bijective. In such a case the map f^{-1} is invertible as well, with $(f^{-1})^{-1} = f$.*

7.3 Kernel and Image of a Linear Map

Injectivity and surjectivity of a linear map are measured by two vector subspaces that we now introduce and study.

Definition 7.3.1 Consider a linear map $f : V \to W$. The set

$$V \supseteq \ker(f) = \{v \in V \ : \ f(v) = 0_W\}$$

is called the *kernel* of f, while the set

$$W \supseteq \operatorname{Im}(f) = \{w \in W : \exists\, v \in V : w = f(v)\}$$

is called the *image* of f.

Theorem 7.3.2 *Given a linear map $f : V \to W$, the set* $\ker(f)$ *is a vector subspace in V and* $\operatorname{Im}(f)$ *is a vector subspace in W.*

Proof We recall the Proposition 2.2.2. Given $v, v' \in \ker(f)$ and $\lambda, \lambda' \in \mathbb{R}$ we need to compute $f(\lambda v + \lambda' v')$. Since $f(v) = 0_W = f(v')$ by hypothesis, from the Proposition 7.1.5 we have $f(\lambda v + \lambda' v') = \lambda f(v) + \lambda' f(v') = 0_W$. This shows that $\ker(f)$ is a vector subspace in V.

Analogously, let $w, w' \in \operatorname{Im}(f)$ and $\lambda, \lambda' \in \mathbb{R}$. From the hypothesis there exist $v, v' \in V$ such that $w = f(v)$ and $w' = f(v')$; thus we can write $\lambda w + \lambda' w' = \lambda f(v) + \lambda' f(v') = f(\lambda v + \lambda' v') \in \operatorname{Im}(f)$ again from he Proposition 7.1.5. This shows that $\operatorname{Im}(f)$ is a vector subspace in W. □

Having proved that $\operatorname{Im}(f)$ and $\ker(f)$ are vector subspaces we look for a system of generators for them. Such a task is easier for the image of f as the following lemma shows.

Lemma 7.3.3 *With $f : V \to W$ a linear map, one has that* $\operatorname{Im}(f) = \mathcal{L}(f(v_1), \dots, f(v_n))$, *where $\mathcal{B} = (v_1, \dots, v_n)$ is an arbitrary basis for V. The map f is indeed surjective if and only if $f(v_1), \dots, f(v_n)$ generate W.*

Proof Let $w \in \operatorname{Im}(f)$, that is $w = f(v)$ for some $v \in V$. Being $\mathcal{B}$ a basis for V, one has $v = a_1 v_1 + \cdots + a_n v_n$ and since f is linear, one has $w = a_1 f(v_1) + \cdots + a_n f(v_n)$, thus giving $w \in \mathcal{L}(f(v_1), \dots, f(v_n))$. We have then $\operatorname{Im}(f) \subseteq \mathcal{L}(f(v_1), \dots, f(v_n))$. The opposite inclusion is obvious since $\operatorname{Im}(f)$ is a vector subspace in W and contains the vectors $f(v_1), \dots, f(v_n)$.

The last statement is the fact that f is surjective (Definition 7.2.2) if and only if $\operatorname{Im}(f) = W$. □

Exercise 7.3.4 Let us consider the linear map $f : \mathbb{R}^3 \to \mathbb{R}^2$ given by

$$f((x, y, z)) = (x + y - z, x - y + z).$$

From the lemma above, the vector subspace $\operatorname{Im}(f)$ is generated by the images under f of an arbitrary basis in $\mathbb{R}^3$. With the canonical basis $\mathcal{E} = (e_1, e_2, e_3)$ we have $\operatorname{Im}(f) = \mathcal{L}(f(e_1), f(e_2), f(e_3))$, with

$$f(e_1) = (1, 1), \qquad f(e_2) = (1, -1), \qquad f(e_3) = (-1, 1).$$

It is immediate to see that $\operatorname{Im}(f) = \mathbb{R}^2$, that is f is surjective.

Lemma 7.3.5 *Let $f : V \to W$ be a linear map between two real vector spaces. Then,*

(i) f is injective if and only if $\ker(f) = \{0_V\}$,
(ii) if f is injective and $(v_1, \dots, v_n)$ is a basis for V, the vectors $f(v_1), \dots, f(v_n)$ are linearly independent.

Proof (i) Let us assume that f is injective and $v \in \ker(f)$, that is $f(v) = 0_W$. From the Proposition 7.1.5 we know that $f(0_V) = 0_W$. Since f is injective it must be $v = 0_V$, that is $\ker(f) = \{0_V\}$.
Viceversa, let us assume that $\ker(f) = 0_V$ and let us consider two vectors $v_1,\ v_2$ such that $f(v_1) = f(v_2)$. Since f is linear this reads $0_W = f(v_1) - f(v_2) = f(v_1 - v_2)$, that is $v_1 - v_2 \in \ker(f)$ which, being the latter the null vector subspace, thus gives $v_1 = v_2$.

(ii) In order to study the linear independence of the system of vectors $\{f(v_1), \dots, f(v_n)\}$ let us take scalars $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ such that $\lambda_1 f(v_1) + \cdots + \lambda_n f(v_n) = 0_W$. Being f linear, this gives $f(\lambda_1 v_1 + \cdots + \lambda_n v_n) = 0_W$ and then $\lambda_1 v_1 + \cdots + \lambda_n v_n \in \ker(f)$. Since f is injective, from (i) we have $\ker(f) = \{0_V\}$ so it is $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0_V$. Being $(v_1, \dots, v_n)$ a basis for V, we have that $\lambda_1 = \cdots = \lambda_n = 0_{\mathbb{R}}$ thus proving that also $f(v_1), \dots, f(v_n)$ are linearly independent.

□

Exercise 7.3.6 Let us consider the linear map $f : \mathbb{R}^2 \to \mathbb{R}^3$ given by

$$f((x, y)) = (x + y, x - y, 2x + 3y).$$

The kernel of f is given by

$$\ker(f) = \{(x, y) \in \mathbb{R}^2 \mid f((x, y)) = (x + y, x - y, 2x + 3y) = (0, 0, 0)\}$$

so we have to solve the linear system

$$\begin{cases} x + y = 0 \\ x - y = 0 \\ 2x + 3y = 0 \end{cases} .$$

Its unique solution is $(0, 0)$ so $\ker(f) = \{0_{\mathbb{R}^2}\}$ and we can conclude, from the lemma above, that f is injective. From the same lemma we also know that the images under f of a basis for $\mathbb{R}^2$ make a linearly independent set of vectors. If we take the canonical basis for $\mathbb{R}^2$ with $e_1 = (1, 0)$ and $e_2 = (0, 1)$, we have

$$f(e_1) = (1, 1, 2), \quad f(e_2) = (1, -1, 3).$$

7.4 Isomorphisms

Definition 7.4.1 Let V and W be two real vector spaces. A bijective linear map $f : V \to W$ is called an *isomorphism*. Two vector spaces are said to be *isomorphic* if there exists an isomorphism between them. If $f : V \to W$ is an isomorphism we write $V \cong W$.

Proposition 7.4.2 *If the map $f : V \to W$ is an isomorphism, such is its inverse $f^{-1} : W \to V$.*

Proof From the Proposition 7.2.5 we have that f is invertible, with an invertible inverse map f^{-1}. We need to prove that f^{-1} is linear. Let us consider two arbitrary vectors $w_1, w_2 \in W$ with $v_1 = f^{-1}(w_1)$ and $v_2 = f^{-1}(w_2)$ in V; this is equivalent to $w_1 = f(v_1)$ and $w_2 = f(v_2)$. Let us consider also $\lambda_1, \lambda_2 \in \mathbb{R}$. Since f is linear we can write

$$\lambda_1 w_1 + \lambda_2 w_2 = f(\lambda_1 v_1 + \lambda_2 v_2).$$

For the action of f^{-1} is then

$$f^{-1}(\lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 v_1 + \lambda_2 v_2 = \lambda_1 f^{-1}(w_1) + \lambda_2 f^{-1}(w_2),$$

which amounts to say that f^{-1} is a linear map. □

In order to characterise isomorphisms we first prove a preliminary result.

Lemma 7.4.3 *Let $f : V \to W$ be a linear map with $(v_1, \dots, v_n)$ a basis for V. The map f is an isomorphism if and only if $(f(v_1), \dots, f(v_n))$ is a basis for W.*

Proof If f is an isomorphism, it is both injective and surjective. From the Lemma 7.3.3 the system $f(v_1), \dots, f(v_n)$ generates W, while from the Lemma 7.3.5 such a system is linearly independent. This means that $(f(v_1), \dots, f(v_n))$ is a basis for W.

Let us now assume that the vectors $(f(v_1), \dots, f(v_n))$ are a basis for W. From the Proposition 7.1.14 there exists a linear map $g : W \to V$ such that $g(f(v_j)) = v_j$ for any $j = 1, \dots, n$. This means that the linear maps $g \circ f$ and id_V coincide on the basis $(v_1, \dots, v_n)$ in V and then (again from Proposition 7.1.14) they coincide, that is $g \circ f = \mathrm{id}_V$. Along the same lines it is easy to show that $f \circ g = \mathrm{id}_W$, so we have $g = f^{-1}$; the map f is then invertible so it is an isomorphism. □

Theorem 7.4.4 *Let V and W be two real vector spaces. They are isomorphic if and only if* $\dim(V) = \dim(W)$.

Proof Let us assume V and W to be isomorphic, that is there exists an isomorphism $f : V \to W$. From the previous lemma, if $(v_1, \dots, v_n)$ is a basis for V, then $(f(v_1), \dots, f(v_n))$ is a basis for W and this gives $\dim(V) = n = \dim(W)$.

Let us now assume $n = \dim(V) = \dim(W)$ and try to define an isomorphism $f : V \to W$. By fixing a basis $\mathcal{B} = (v_1, \dots, v_n)$ for V and a basis $\mathcal{C} = (w_1, \dots, w_n)$

for W, we define the linear map $f(v_j) = w_j$ for any j. Such a linear map exists and it is unique from the Proposition 7.1.14. From the lemma above, f is an isomorphism since it maps the basis $\mathcal{B}$ to the basis $\mathcal{C}$ for W. □

Corollary 7.4.5 *If V is a real vector space with* $\dim(V) = n$*, then $V \cong \mathbb{R}^n$. Any choice of a basis $\mathcal{B}$ for V induces the natural isomorphism*

$$\alpha : V \xrightarrow{\cong} \mathbb{R}^n \quad \textit{given by} \quad (x_1, \ldots, x_n)_{\mathcal{B}} \mapsto (x_1, \ldots, x_n).$$

Proof The first claim follows directly from the Theorem 7.4.4 above. Once the basis $\mathcal{B} = (v_1, \ldots, v_n)$ is chosen the map α is defined as the linear map such that $\alpha(v_j) = e_j$ for any $j = 1, \ldots, n$. From the Lemma 7.4.3 such a map α is an isomorphism. It is indeed immediate to check that the action of α on any vector in V is given by $\alpha : (x_1, \ldots, x_n)_{\mathcal{B}} \mapsto (x_1, \ldots, x_n)$. □

Exercise 7.4.6 Let $V = \mathbb{R}[X]_2$ be the space of the polynomials whose degree is not higher than 2. As we know, V has dimension 3 and a basis for it is given by $\mathcal{B} = (1, X, X^2)$. The isomorphism $\alpha : \mathbb{R}[X]_2 \xrightarrow{\cong} \mathbb{R}^3$ corresponding to such a basis reads

$$a + bX + cX^2 \quad \mapsto \quad (a, b, c).$$

It is simple to check whether a given system of polynomials is a basis for $\mathbb{R}[X]_2$. As an example we consider

$$p_1(X) = 3X - X^2, \quad p_2(X) = 1 + X, \quad p_3(X) = 2 + 3X^2.$$

By setting $v_1 = \alpha(p_1) = (0, 3, -1)$, $v_2 = \alpha(p_2) = (1, 1, 0)$ and $v_3 = \alpha(p_3) = (2, 0, 3)$, it is clear that the rank of the matrix whose columns are the vectors v_1, v_2, v_3 is 3, thus proving that (v_1, v_2, v_3) is a basis for $\mathbb{R}^3$. Since α is an isomorphism, the inverse $\alpha^{-1} : \mathbb{R}^3 \to \mathbb{R}[X]_2$ is an isomorphism as well: the vectors $\alpha^{-1}(v_1)$, $\alpha^{-1}(v_2)$, $\alpha^{-1}(v_2)$ provide a basis for $\mathbb{R}[X]_2$ and coincide with the given polynomials $p_1(X)$, $p_2(X)$, $p_3(X)$.

Theorem 7.4.4 shows that a linear isomorphism exists only if its domain has the same dimension of its image. A condition that characterises isomorphism can then be introduced only for vector spaces with the same dimensions. This is done in the following sections.

7.5 Computing the Kernel of a Linear Map

We have seen that isomorphisms can be defined only between spaces with the same dimension. Being not an isomorphism indeed means for a linear map to fail to be injective or surjective. In this section and the following one we characterise injectivity and surjectivity of a linear map via the study of its kernel and its image. In particular, we shall describe procedures to exhibit bases for such spaces.

Proposition 7.5.1 *Let $f : V \to W$ be a linear map between real vector spaces, and $\dim(V) = n$. Fix a basis $\mathcal{B}$ for V and a basis $\mathcal{C}$ for W, with associated matrix $A = M_f^{\mathcal{C},\mathcal{B}}$. By denoting $\Sigma : AX = 0$ the linear system associated to A, the following hold:*

(i) $S_\Sigma \cong \ker(f)$ *via the isomorphism* $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_n)_\mathcal{B}$*,*
(ii) $\dim(\ker(f)) = n - \mathrm{rk}(A)$*,*
(iii) if $(v_1, \dots, v_p)$ *is a basis for* S_Σ*, the vectors* $\big((v_1)_\mathcal{B}, \dots, (v_p)_\mathcal{B}\big)$ *are a basis for* $\ker(f)$.

Proof (i) With the given hypothesis, from the definition of the kernel of a linear map we can write

$$\begin{aligned}
\ker(f) &= \{v \in V \ : \ f(v) = 0_W\} \\
&= \left\{ v = (x_1, \dots, x_n)_\mathcal{B} \in V \ : \ \left(A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right)_\mathcal{C} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}_\mathcal{C} \right\} \\
&= \{(x_1, \dots, x_n)_\mathcal{B} \in V \ : \ (x_1, \dots, x_n) \in S_\Sigma\}
\end{aligned}$$

with S_Σ denoting the space of solutions for Σ. As in Corollary 7.4.5 we can then write down the isomorphism $S_\Sigma \to \ker(f)$ given by

$$(x_1, \dots, x_n) \quad \mapsto \quad (x_1, \dots, x_n)_\mathcal{B}.$$

(ii) From the isomorphism of the previous point we then have

$$\dim(\ker(f)) = \dim(S_\Sigma) = n - \mathrm{rk}(A)$$

where the last equality follows from the Theorem 6.4.3.

(iii) From the Lemma 7.4.3 we know that, under the isomorphism $S_\Sigma \to \ker(f)$, a basis for S_Σ is mapped into a basis for $\ker(f)$.

□

Exercise 7.5.2 Consider the linear map $f : \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$f((x, y, z)_\mathcal{B}) = (x + y - z, x - y + z, 2x)_\mathcal{E}$$

where $\mathcal{B} = \big((1, 1, 0), (0, 1, 1), (1, 0, 1)\big)$ and $\mathcal{E}$ is the canonical basis for $\mathbb{R}^3$. We determine $\ker(f)$ and compute a basis for it with respect to both $\mathcal{B}$ and $\mathcal{E}$. Start by considering the matrix associated to the linear map f with the given basis,

$$A = M_f^{\mathcal{E},\mathcal{B}} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ 2 & 0 & 0 \end{pmatrix}.$$

To solve the linear system $\Sigma : AX = 0$ we reduce the matrix A by rows:

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

and the space S_Σ of the solutions of Σ is then given by

$$S_\Sigma = \{(0, a, a) : a \in \mathbb{R}\} = \mathcal{L}((0, 1, 1)).$$

This reads

$$\ker(f) = \{(0, a, a)_{\mathcal{B}} : a \in \mathbb{R}\},$$

with a basis given by the vector $(0, 1, 1)_{\mathcal{B}}$. With the explicit expression of the elements of $\mathcal{B}$,

$$(0, 1, 1)_{\mathcal{B}} = (0, 1, 1) + (1, 0, 1) = (1, 1, 2).$$

This shows that the basis vector for $\ker(f)$ given by $(0, 1, 1)$ on the basis $\mathcal{B}$ is the same as the basis vector $(1, 1, 2)$ with respect to the canonical basis $\mathcal{E}$ for $\mathbb{R}^3$.

Exercise 7.5.3 With canonical bases $\mathcal{E}$, consider the linear map $f : \mathbb{R}^3 \to \mathbb{R}^3$ given by

$$f((x, y, z)) = (x + y - z, x - y + z, 2x).$$

To determine the space $\ker(f)$, we observe that the matrix associated to f is the same matrix of the previous exercise, so the linear system $\Sigma : AX = 0$ has solutions $S_\Sigma = \mathcal{L}((0, 1, 1)) = \ker(f)$, since $\mathcal{E}$ is the canonical basis.

Since the kernel of a linear map is the preimage of the null vector in the image space, we can generalise the above procedure to compute the preimage of any element $w \in W$. We denote it as $f^{-1}(w) = \{v \in V : f(v) = w\}$, with $\ker(f) = f^{-1}(0_W)$. Notice that we denote the preimage of a set under f by writing f^{-1} also when f is not invertible.

Proposition 7.5.4 *Consider a real vector space V with basis $\mathcal{B}$ and a real vector space W with basis $\mathcal{C}$. Let $f : V \to W$ be a linear map with $A = M_f^{\mathcal{C},\mathcal{B}}$ its corresponding matrix. Given any $w = (y_1, \ldots, y_m)_{\mathcal{C}} \in W$, it is*

$$f^{-1}(w) = \{(x_1, \ldots, x_n)_{\mathcal{B}} \in V : A\,{}^t(x_1, \ldots, x_n) = {}^t(y_1, \ldots, y_m)\}.$$

Proof It is indeed true that, with $v = (x_1, \ldots, x_n)_{\mathcal{B}}$, one has

$$f(v) \;=\; f((x_1,\ldots,x_n)_{\mathcal{B}}) \;=\; (A\,{}^t(x_1,\ldots,x_n))_{\mathcal{C}}.$$

The equality $f(v) = w$ is the equality of components, given by $A\,{}^t(x_1,\ldots,x_n) = {}^t(y_1,\ldots,y_m)$, on the basis $\mathcal{C}$. □

Remark 7.5.5 This fact can be expressed via linear systems. Given $w \in W$, its preimage $f^{-1}(w)$ is made of vectors in V whose components with respect to $\mathcal{B}$ solve the linear system $AX = B$, where B is the column of the components of w with respect to $\mathcal{C}$.

Exercise 7.5.6 Consider the linear map $f : \mathbb{R}^3 \to \mathbb{R}^3$ given in the Exercise 7.5.2. We compute $f^{-1}(w)$ for $w = (1, 1, 1)$. We have then to solve the system $\Sigma : AX = B$, with $B = {}^t(1, 1, 1)$. We reduce the matrix (A, B) as follows

$$(A, B) \;=\; \begin{pmatrix} 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 2 & 0 & 0 & 1 \end{pmatrix} \;\mapsto\; \begin{pmatrix} 1 & 1 & -1 & 1 \\ 2 & 0 & 0 & 2 \\ 2 & 0 & 0 & 1 \end{pmatrix} \;\mapsto\; \begin{pmatrix} 1 & 1 & -1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

This shows that the system Σ has no solution, that is $w \notin \mathrm{Im}(f)$.

Next, let us compute $f^{-1}(u)$ for $u = (2, 0, 2)$, so we have the linear system $\Sigma : AX = B$ with $B = {}^t(2, 0, 2)$. Reducing by row, we have

$$(A, B) \;=\; \begin{pmatrix} 1 & 1 & -1 & 2 \\ 1 & -1 & 1 & 0 \\ 2 & 0 & 0 & 2 \end{pmatrix} \;\mapsto\; \begin{pmatrix} 1 & 1 & -1 & 2 \\ 2 & 0 & 0 & 2 \\ 2 & 0 & 0 & 2 \end{pmatrix} \;\mapsto\; \begin{pmatrix} 1 & 1 & -1 & 2 \\ 2 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The system Σ is then equivalent to

$$\begin{cases} x = 1 \\ y = z + 1 \end{cases}$$

whose space of solutions is $S_\Sigma = \{(1, a+1, a) \; : \; a \in \mathbb{R}\}$. We can then write

$$f^{-1}(2,0,2) \;=\; \{(1, a+1, a)_{\mathcal{B}} \; : \; a \in \mathbb{R}\} \;=\; \{(a+1, a+2, 2a+1) \; : \; a \in \mathbb{R}\}.$$

7.6 Computing the Image of a Linear Map

We next turn to the study of the image of a linear map.

Proposition 7.6.1 *Let $f : V \to W$ be a linear map between real vector spaces, with $\dim(V) = n$ and $\dim(W) = m$. Fix a basis $\mathcal{B}$ for V and a basis $\mathcal{C}$ for W, with associated matrix $A = M_f^{\mathcal{C},\mathcal{B}}$ and with $C(A)$ its space of columns. The following results hold:*

(i) $\mathrm{Im}(f) \cong C(A)$ *via the isomorphism* $(y_1,\ldots,y_m)_{\mathcal{C}} \mapsto (y_1,\ldots,y_m)$,

(ii) $\dim(\mathrm{Im}(f)) = \mathrm{rk}(A)$,

(iii) *if* $(w_1, \ldots, w_r)$ *is a basis for* $C(A)$*, then* $((w_1)_{\mathcal{C}}, \ldots, (w_r)_{\mathcal{C}})$ *is a basis for* $\mathrm{Im}(f)$.

Proof (i) With the given hypothesis, from the definition of the image of a linear map we can write

$$\begin{aligned}\mathrm{Im}(f) &= \{w \in W \,:\, \exists\, v \in V \,:\, w = f(v)\}\\ &= \left\{ w = (y_1, \ldots, y_m)_{\mathcal{C}} \in W \,:\, \exists\,(x_1, \ldots, x_n)_{\mathcal{B}} \in V \,:\, \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}_{\mathcal{C}} = \left(A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right)_{\mathcal{C}} \right\}\\ &= \left\{ (y_1, \ldots, y_m)_{\mathcal{C}} \in W \,:\, \exists\,(x_1, \ldots, x_n) \in \mathbb{R}^n \,:\, \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\}.\end{aligned}$$

Representing the matrix A by its columns, that is $A = (C_1 \; \cdots \; C_n)$, we have

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 C_1 + \cdots + x_n C_n.$$

We can therefore write

$$\begin{aligned}\mathrm{Im}(f) &= \left\{ (y_1, \ldots, y_m)_{\mathcal{C}} \in W \,:\, \exists\,(x_1, \ldots, x_n) \in \mathbb{R}^n \,:\, \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = x_1 C_1 + \cdots x_n C_n \right\}\\ &= \left\{ (y_1, \ldots, y_m)_{\mathcal{C}} \in W \,:\, \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \in C(A) \right\}.\end{aligned}$$

We have then the isomorphism $C(A) \to \mathrm{Im}(f)$ defined by

$$(y_1, \ldots, y_m) \quad \mapsto \quad (y_1, \ldots, y_m)_{\mathcal{C}}$$

(compare this with the one in the Corollary 7.4.5).

(ii) Being $\mathrm{Im}(f) \cong C(A)$, it is $\dim(\mathrm{Im}(f)) = \dim(C(A)) = \mathrm{rk}(A)$.

(iii) The claim follows from (i) and the Lemma 7.4.3.

□

Remark 7.6.2 To determine a basis for $C(A)$ as in (iii) above, one can proceed as follows.

(a) If the rank of A is known, one has to select n linearly independent columns: they will give a basis for $C(A)$.

(b) If the rank of A is not known, by denoting A' the matrix obtained from A by reduction by columns, a basis for $C(A)$ is given by the r non zero columns of A'.

Exercise 7.6.3 Let $f : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear map with associated matrix

$$A = M_f^{\mathcal{C},\mathcal{E}} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -3 \\ 2 & -1 & 1 \end{pmatrix}$$

for the canonical basis $\mathcal{E}$ and basis $\mathcal{C} = (w_1, w_2, w_3)$, with $w_1 = (1, 1, 0)$, $w_2 = (0, 1, 1)$, $w_3 = (1, 0, 1)$. We reduce A by columns

$$A \xrightarrow[C_3 \mapsto C_3 - 2C_1]{C_2 \mapsto C_2 + C_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 2 & 1 & -3 \end{pmatrix} \xrightarrow{C_3 \mapsto C_3 + 3C_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 1 & 0 \end{pmatrix} = A'.$$

Being A' reduced by columns, its non zero columns yield a basis for the space $C(A)$. Thus, $C(A) = C(A') = \mathcal{L}((1, 0, 2), (0, 1, 1))$. From the Proposition 7.6.1 a basis for $\mathrm{Im}(f)$ is given by the pair (u_1, u_2),

$$\begin{aligned} u_1 &= (1, 0, 2)_{\mathcal{C}} = w_1 + 2w_3 = (3, 1, 2) \\ u_2 &= (0, 1, 1)_{\mathcal{C}} = w_2 + w_3 = (1, 1, 2). \end{aligned}$$

Clearly, $\dim(\mathrm{Im}(f)) = 2 = \mathrm{rk}(A)$.

From the previous results we have the following theorem.

Theorem 7.6.4 *Let $f : V \to W$ be a linear map. It holds that*

$$\dim(\ker(f)) + \dim(\mathrm{Im}(f)) = \dim(V).$$

Proof Let A be any matrix associated to f (that is irrespective of the bases chosen in V and W). From the Proposition 7.5.1 one has $\dim(\ker(f)) = \dim(V) - \mathrm{rk}(A)$, while from the Proposition 7.6.1 one has $\dim(\mathrm{Im}(f)) = \mathrm{rk}(A)$. The claim follows. □

From this theorem, the next corollary follows easily.

Corollary 7.6.5 *Let $f : V \to W$ be a linear map, with* $\dim(V) = \dim(W)$. *The following statements are equivalent.*

(i) *f is injective,*
(ii) *f is surjective,*
(iii) *f is an isomorphism.*

Proof Clearly it is sufficient to prove the equivalence (i) $\Leftrightarrow$ (ii). From the Lemma 7.3.5 we know that f is injective if and only if $\dim(\ker(f)) = 0$. We also known that f is surjective if and only if $\dim(\mathrm{Im}(f)) = \dim(W)$. Since $\dim(V) = \dim(W)$ by hypothesis, the statement thus follows from the Theorem 7.6.4. □

7.7 Injectivity and Surjectivity Criteria

In this section we study conditions for injectivity and surjectivity of a linear map through properties of its associated matrix.

Proposition 7.7.1 (Injectivity criterion) *Let $f : V \to W$ be a linear map. Then f is injective if and only if* $\mathrm{rk}(A) = \dim(V)$ *for any matrix A associated to f (that is, irrespective of the bases with respect to which the matrix A is given).*

Proof From (i) in the Lemma 7.3.5 we know that f is injective if and only if $\ker(f) = \{0_V\}$, which means $\dim(\ker(f)) = 0$. From the Proposition 7.5.1 we have that $\dim(\ker(f)) = \dim(V) - \mathrm{rk}(A)$ for any matrix A associated to f. We then have that f is injective if and only if $\dim(V) - \mathrm{rk}(A) = 0$. □

Exercise 7.7.2 Let $f : \mathbb{R}[X]_2 \to \mathbb{R}^{2,2}$ be the linear map associated to the matrix

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

with respect to two given basis. Since$A$ is already reduced by column, $\mathrm{rk}(A) = 3$, the number of its non zero columns. Being $\dim(\mathbb{R}[X]_2) = 3$ we have, from the Proposition 7.7.1, that f is injective.

Proposition 7.7.3 (Surjectivity criterion) *Let $f : V \to W$ be a linear map. The map f is surjective if and only if* $\mathrm{rk}(A) = \dim(W)$ *for any matrix associated to f (again irrespective of the bases with respect to which the matrix A is given).*

Proof This follows directly from the Proposition 7.6.1. □

Exercise 7.7.4 Let $f : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear map given by

$$f(x, y, z) = (x + y - z, 2x - y + 2z).$$

With $\mathcal{E}$ the canonical basis in $\mathbb{R}^3$ and $\mathcal{C}$ the canonical basis in $\mathbb{R}^2$, we have

$$A = M_f^{\mathcal{C},\mathcal{E}} = \begin{pmatrix} 1 & 1 & -1 \\ 2 & -1 & 2 \end{pmatrix} :$$

by reducing by rows,

$$A \mapsto \begin{pmatrix} 1 & 1 & -1 \\ 3 & 0 & 1 \end{pmatrix} = A'.$$

We know that $\mathrm{rk}(A) = \mathrm{rk}(A') = 2$, the number of non zero rows in A'. Being $\dim(\mathbb{R}^2) = 2$, the map f is surjective from the Proposition 7.7.3.

We have seen in the Proposition 7.4.2 that if a linear map f is an isomorphism, then its domain and image have the same dimension. Injectivity and surjectivity of a linear map provide necessary conditions on the relative dimensions of the domain and the image of the map.

Remark 7.7.5 Let $f : V \to W$ be a linear map. One has:

(a) If f is injective, then $\dim(V) \leq \dim(W)$. This claim easily follows from the Lemma 7.3.5, since the images under f of a basis for V gives linearly independent vectors in W.
(b) If f is surjective, then $\dim(V) \geq \dim(W)$. This claim follows from the Lemma 7.3.3, since the images under f of a basis for V generate (that is they linearly span) W.

Remark 7.7.6 Let $f : V \to W$ be a linear map, with A its corresponding matrix with respect to any basis. One has:

(a) With $\dim(V) < \dim(W)$, f is injective if and only if $\mathrm{rk}(A)$ is maximal;
(b) With $\dim(V) > \dim(W)$, f is surjective if and only if $\mathrm{rk}(A)$ is maximal;
(c) With $\dim(V) = \dim(W)$, f is an isomorphism if and only if $\mathrm{rk}(A)$ is maximal.

Exercise 7.7.7 The following linear maps are represented with respect to canonical bases.

(1) Let the map $f : \mathbb{R}^3 \to \mathbb{R}^4$ be defined by

$$(x, y, z) \quad \mapsto \quad (x - y + 2z, y + z, -x + z, 2x + y).$$

To compute the rank of the corresponding matrix A with respect to the canonical basis, as usual we reduce it by rows. We have

$$A = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

and the rank of A is maximal, $\mathrm{rk}(A) = 3$. Since $\dim(V) < \dim(W)$ we have that f is injective.

(2) Let the map $f : \mathbb{R}^4 \to \mathbb{R}^3$ be defined by

$$(x, y, z, t) \quad \mapsto \quad (x - y + 2z + t, y + z + 3t, x - y + 2z + 2t).$$

We proceed as above and compute, via the following reduction,

$$A = \begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & 1 & 3 \\ 1 & -1 & 2 & 2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

that $\mathrm{rk}(A) = 3$. Since $\mathrm{rk}(A)$ is maximal, with $\dim(V) > \dim(W)$, f turns out to be surjective.

(3) Let $f : \mathbb{R}^3 \to \mathbb{R}^3$ be represented as before by the matrix

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & -1 & 2 \end{pmatrix},$$

which, by reduction, becomes

$$A \mapsto \begin{pmatrix} 1 & 2 & 1 \\ 0 & -3 & -1 \\ 0 & 0 & 2 \end{pmatrix},$$

whose rank is clearly maximal. Thus f is an isomorphism since $\dim(V) = \dim(W)$.

7.8 Composition of Linear Maps

We rephrase the general Definition 7.2.3 of composing maps.

Definition 7.8.1 Let $f : V \to W$ and $g : W \to Z$ be two linear maps between real vector spaces. The composition between g and f is the map

$$g \circ f : X \to Z$$

defined as $(g \circ f)(v) = g(f(v))$, for any $v \in X$.

Proposition 7.8.2 *If $f : V \to W$ and $g : W \to Z$ are two linear maps, the composition map $g \circ f : V \to Z$ is linear as well.*

Proof For any $v, v' \in V$ and $\lambda, \lambda' \in \mathbb{R}$, the linearity of both f and g allows one to write:

$$\begin{aligned}(g \circ f)(\lambda v + \lambda' v') &= g(f(\lambda v + \lambda' v')) \\ &= g(\lambda f(v) + \lambda' f(v')) \\ &= \lambda g(f(v)) + \lambda' g(f(v')) \\ &= \lambda (g \circ f)(v) + \lambda' (g \circ f)(v'),\end{aligned}$$

showing the linearity of the composition map. □

The following proposition, whose proof we omit, characterises the matrix corresponding to the composition of two linear maps.

Proposition 7.8.3 *Let V, W, Z be real vector spaces with basis $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ respectively. Given linear maps $f : V \to W$ and $g : W \to Z$, the corresponding matrices with respect to the given bases are related by*

$$M_{g\circ f}^{\mathcal{D},\mathcal{B}} = M_g^{\mathcal{D},\mathcal{C}} \cdot M_f^{\mathcal{C},\mathcal{B}}.$$

The following theorem characterises an isomorphism in terms of its corresponding matrix.

Theorem 7.8.4 *Let $f : V \to W$ be a linear map. The map f is an isomorphism if and only if, for any choice of the bases $\mathcal{B}$ for V and $\mathcal{C}$ for W, the corresponding matrix $M_f^{\mathcal{C},\mathcal{B}}$ with respect to the given bases is invertible, with*

$$M_{f^{-1}}^{\mathcal{B},\mathcal{C}} = \left(M_f^{\mathcal{C},\mathcal{B}}\right)^{-1}.$$

Proof Let us assume that f is an isomorphism: we can then write $\dim(V) = \dim(W)$, so $M_f^{\mathcal{C},\mathcal{B}}$ is a square matrix whose size is $n \times n$ (say). From the Proposition 7.4.2 we know that f^{-1} exists as a linear map whose corresponding matrix, with the given bases, will be $M_{f^{-1}}^{\mathcal{B},\mathcal{C}}$. From the Proposition 7.8.3 we can write

$$M_{f^{-1}}^{\mathcal{B},\mathcal{C}} \cdot M_f^{\mathcal{C},\mathcal{B}} = M_{f^{-1}\circ f}^{\mathcal{B},\mathcal{B}} = M_{\mathrm{id}_V}^{\mathcal{B},\mathcal{B}} = I_n \quad \Rightarrow \quad M_{f^{-1}}^{\mathcal{B},\mathcal{C}} = \left(M_f^{\mathcal{C},\mathcal{B}}\right)^{-1}.$$

We set now $A = M_f^{\mathcal{C},\mathcal{B}}$. By hypothesis A is a square invertible matrix, with inverse A^{-1}, so we can consider the linear map

$$g = f_{A^{-1}}^{\mathcal{B},\mathcal{C}} : W \to V.$$

In order to show that g is the inverse of f, consider the matrix corresponding to $g \circ f$ with respect to the basis $\mathcal{B}$. From the Proposition 7.8.3,

$$M_{g\circ f}^{\mathcal{B},\mathcal{B}} = M_g^{\mathcal{B},\mathcal{C}} \cdot M_f^{\mathcal{C},\mathcal{B}} = A^{-1} \cdot A = I_n.$$

Since linear maps are in bijection with matrices, we have that $g \circ f = \mathrm{id}_V$. Along the same lines we can show that $f \circ g = \mathrm{id}_W$, thus proving $g = f^{-1}$. □

Exercise 7.8.5 Consider the linear map $f : \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$f((x, y, z)) = (x - y + z, 2y + z, z).$$

With the canonical basis $\mathcal{E}$ for $\mathbb{R}^3$ the corresponding matrix is

$$A = M_f^{\mathcal{E},\mathcal{E}} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Since $\mathrm{rk}(A) = 3$, f is an isomorphism, with f^{-1} the linear map corresponding to A^{-1}. From the Proposition 5.3.3, we have

$$A^{-1} = \frac{1}{\det(A)} \big(\alpha_{ji}\big) = \begin{pmatrix} 1 & 1/2 & -3/2 \\ 0 & 1/2 & -1/2 \\ 0 & 0 & 1 \end{pmatrix} = M^{\mathcal{E},\mathcal{E}}_{f^{-1}}.$$

7.9 Change of Basis in a Vector Space

In this section we study how to relate the components of the *same* vector in a vector space with respect to *different* bases. This problem has a natural counterpart in physics, where different bases for the same vector space represent different reference systems. Thus different observers measuring observables of the same physical system in a compatible way.

Example 7.9.1 We start by considering the vector space $\mathbb{R}^2$ with two bases given by

$$\mathcal{E} = \big(e_1 = (1,0),\ e_2 = (0,1)\big), \qquad \mathcal{B} = \big(b_1 = (1,2),\ b_2 = (3,4)\big).$$

Any vector $v \in \mathbb{R}^2$ will then be written as

$$v = (x_1, x_2)_{\mathcal{B}} = (y_1, y_2)_{\mathcal{E}},$$

or, more explicitly,

$$v = x_1 b_1 + x_2 b_2 = y_1 e_1 + y_2 e_2.$$

By writing the components of the elements in $\mathcal{B}$ in the basis $\mathcal{E}$, that is

$$b_1 = e_1 + 2e_2, \qquad b_2 = 3e_1 + 4e_2,$$

we have

$$\begin{aligned} y_1 e_1 + y_2 e_2 &= x_1(e_1 + 2e_2) + x_2(3e_1 + 4e_2) \\ &= (x_1 + 3x_2)e_1 + (2x_1 + 4x_2)e_2. \end{aligned}$$

We have then obtained

$$y_1 = x_1 + 3x_2, \qquad y_2 = 2x_1 + 4x_2.$$

These expression can be written in matrix form

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + 3x_2 \\ 2x_1 + 4x_2 \end{pmatrix} \quad \Leftrightarrow \quad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Such a relation can be written as

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

where

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

Notice that the columns of A above are given by the components of the vectors in $\mathcal{B}$ with respect to the basis $\mathcal{E}$. We have the following general result.

Proposition 7.9.2 *Let V be a real vector space with $\dim(V) = n$. Let $\mathcal{B}$ and $\mathcal{C}$ be two bases for V and denote by $(x_1, \ldots, x_n)_{\mathcal{B}}$ and $(y_1, \ldots, y_n)_{\mathcal{C}}$ the component of the same vector v with respect to them. It is*

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = M^{\mathcal{C},\mathcal{B}}_{\mathrm{id}_V} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Such an expression will also be written as

$${}^t(y_1, \ldots, y_n) = M^{\mathcal{C},\mathcal{B}}_{\mathrm{id}_V} \cdot {}^t(x_1, \ldots, x_n).$$

Proof This is clear, by recalling the Definition 7.1.12 and the Proposition 7.1.13. □

Definition 7.9.3 The matrix $M^{\mathcal{C},\mathcal{B}} = M^{\mathcal{C},\mathcal{B}}_{\mathrm{id}_V}$ is called the *matrix of the change of basis* from $\mathcal{B}$ to $\mathcal{C}$. The columns of this matrix are given by the components with respect to $\mathcal{C}$ of the vectors in $\mathcal{B}$.

Exercise 7.9.4 Let $\mathcal{B} = (v_1, v_2, v_3)$ and $\mathcal{C} = (w_1, w_2, w_3)$ two different bases for $\mathbb{R}^3$, with

$$\begin{aligned} v_1 &= (0, 1, -1), \quad v_2 = (1, 0, -1), \quad v_3 = (2, -2, 2), \\ w_1 &= (0, 1, 1), \quad w_2 = (1, 0, 1), \quad w_3 = (1, 1, 0). \end{aligned}$$

We consider the vector $v = (1, -1, 1)_{\mathcal{B}}$ and we wish to determine its components with respect to $\mathcal{C}$. The solution to the linear system

$$\begin{aligned} v_1 &= a_{11} w_1 + a_{21} w_2 + a_{31} w_3 \\ v_2 &= a_{12} w_1 + a_{22} w_2 + a_{32} w_3 \\ v_3 &= a_{13} w_1 + a_{23} w_2 + a_{33} w_3 \end{aligned}$$

give the entries for the matrix of the change of basis, which is found to be

$$M^{\mathcal{C},\mathcal{B}} = \begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & 3 \\ 1 & 1 & -1 \end{pmatrix}.$$

We can then write

$$v = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}_{\mathcal{C}} = \left(M^{\mathcal{C},\mathcal{B}} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right)_{\mathcal{C}} = \left(\begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & 3 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right)_{\mathcal{C}} = \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix}_{\mathcal{C}}.$$

Theorem 7.9.5 *Let $\mathcal{B}$ and $\mathcal{C}$ be two bases for the vector space V over $\mathbb{R}$. The matrix $M^{\mathcal{C},\mathcal{B}}$ is invertible, with*

$$\left(M^{\mathcal{C},\mathcal{B}}\right)^{-1} = M^{\mathcal{B},\mathcal{C}}.$$

Proof This easily follows by applying the Theorem 7.8.4 to $M^{\mathcal{C},\mathcal{B}} = M^{\mathcal{C},\mathcal{B}}_{\mathrm{id}_V}$, since $\mathrm{id}_V = \mathrm{id}_V^{-1}$. □

Theorem 7.9.6 *Let $A \in \mathbb{R}^{n,n}$ be an invertible matrix. Denoting by $v_1, \ldots, v_n$ the column vectors in A and setting $\mathcal{B} = (v_1, \ldots, v_n)$, it holds that:*

(i) $\mathcal{B}$ is a basis for $\mathbb{R}^n$,
(ii) $A = M^{\mathcal{B},\mathcal{E}}$ with $\mathcal{E}$ the canonical basis in $\mathbb{R}^n$.

Proof (i) From the Remark 7.7.6, we know that A has maximal rank, that is $\mathrm{rk}(A) = n$. Being the column vectors in A, the system $v_1, \ldots, v_n$ is then free. A system of n linearly independent vectors in $\mathbb{R}^n$ is indeed a basis for $\mathbb{R}^n$ (see the Corollary 2.5.5 in Chap. 2).
(ii) It directly follows from the Definition 7.9.3.

□

Remark 7.9.7 From the Theorems 7.9.5 and 7.9.6 we have that the group $\mathrm{GL}(n, \mathbb{R})$ of invertible matrices of order n, is the same as (the group of) matrices providing change of basis in $\mathbb{R}^n$.

Exercise 7.9.8 The matrix

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

is invertible since $\mathrm{rk}(A) = 3$ (the matrix A is reduced by rows, so its rank is the number of non zero columns). The column vectors in A, that is

$$v_1 = (1, 0, 0), \quad v_2 = (1, 1, 0), \quad v_3 = (-1, 2, 1),$$

form a basis for $\mathbb{R}^3$. It is also clear that $A = M^{\mathcal{E},\mathcal{B}} = M^{\mathcal{E},\mathcal{B}}_{\mathrm{id}_{\mathbb{R}^3}}$, with $\mathcal{B} = (v_1, v_2, v_3)$.

We next turn to study how $M_f^{\mathcal{C},\mathcal{B}}$, the matrix associated to a linear map $f : V \to W$ with respect to the bases $\mathcal{B}$ for V and $\mathcal{C}$ in W, is transformed under a change of basis in V and W. In the following pages we shall denote by $V_{\mathcal{B}}$ the vector space V referred to its basis $\mathcal{B}$.

Theorem 7.9.9 *Let $\mathcal{B}$ and $\mathcal{B}'$ be two bases for the real vector space V and $\mathcal{C}$ and $\mathcal{C}'$ two bases for the real vector space W. With $f : V \to W$ a linear map, one has that*

$$M_f^{\mathcal{C}',\mathcal{B}'} = M^{\mathcal{C}',\mathcal{C}} \cdot M_f^{\mathcal{C},\mathcal{B}} \cdot M^{\mathcal{B},\mathcal{B}'}.$$

Proof The *commutative diagram*

$$\begin{tikzcd} V_{\mathcal{B}} \arrow[r, "f"] & W_{\mathcal{C}} \arrow[d, "\mathrm{id}_W"] \\ V'_{\mathcal{B}} \arrow[u, "\mathrm{id}_V"] \arrow[r, "f"'] & W'_{\mathcal{C}} \end{tikzcd}$$

shows the claim: going from $V'_{\mathcal{B}}$ to $W'_{\mathcal{C}}$ along the bottom line is equivalent to going around the diagram, that is

$$f \;=\; \mathrm{id}_W \circ f \circ \mathrm{id}_V .$$

Such a relation can be translated in a matrix form,

$$M_f^{\mathcal{C}',\mathcal{B}'} \;=\; M_{\mathrm{id}_W \circ f \circ \mathrm{id}_V}^{\mathcal{C}',\mathcal{B}'}$$

and, by recalling the Proposition 7.8.3, we have

$$M_{\mathrm{id}_W \circ f \circ \mathrm{id}_V}^{\mathcal{C}',\mathcal{B}'} \;=\; M_{\mathrm{id}_W}^{\mathcal{C}',\mathcal{C}} \cdot M_f^{\mathcal{C},\mathcal{B}} \cdot M_{\mathrm{id}_V}^{\mathcal{B},\mathcal{B}'},$$

which proves the claim. □

Exercise 7.9.10 Consider the linear map $f : \mathbb{R}^2_{\mathcal{B}} \to \mathbb{R}^3_{\mathcal{C}}$ whose corresponding matrix is

$$A \;=\; M_f^{\mathcal{C},\mathcal{B}} \;=\; \begin{pmatrix} 1 & 2 \\ -1 & 0 \\ 1 & 1 \end{pmatrix}$$

with respect to $\mathcal{B} = \big((1, 1), (0, 1)\big)$ and $\mathcal{C} = \big((1, 1, 0), (1, 0, 1), (0, 1, 1)\big)$. We determine the matrix $B \;=\; M_f^{\mathcal{E}_3,\mathcal{E}_2}$, with $\mathcal{E}_2$ the canonical basis for $\mathbb{R}^2$ and $\mathcal{E}_3$ the canonical basis for $\mathbb{R}^3$. The commutative diagram above turns out to be

$$
\begin{tikzcd}
\mathbb{R}^2{}_{\mathcal{B}} \arrow[r, "f_A^{\mathcal{C},\mathcal{B}}"] & \mathbb{R}^3{}_{\mathcal{C}} \arrow[d, "\mathrm{id}_{\mathbb{R}^3}"] \\
\mathbb{R}^2{}_{\mathcal{E}_2} \arrow[u, "\mathrm{id}_{\mathbb{R}^2}"] \arrow[r, "f_B^{\mathcal{E}_3,\mathcal{E}_2}"'] & \mathbb{R}^3{}_{\mathcal{E}_3}
\end{tikzcd} ,
$$

which reads

$$B = M_f^{\mathcal{E}_3,\mathcal{E}_2} = M^{\mathcal{E}_3,\mathcal{C}} \, A \, M^{\mathcal{B},\mathcal{E}_2}.$$

We have to compute the matrices $M^{\mathcal{E}_3,\mathcal{C}}$ and $M^{\mathcal{B},\mathcal{E}_2}$. Clearly,

$$M^{\mathcal{E}_3,\mathcal{C}} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

and, from the Theorem 7.9.5, it is

$$M^{\mathcal{B},\mathcal{E}_2} = \left(M^{\mathcal{E}_2,\mathcal{B}}\right)^{-1} = \left(\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\right)^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

(the last equality follows from the Proposition 5.3.3). We have then

$$B = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ -1 & 3 \\ -1 & 1 \end{pmatrix}.$$

We close this section by studying how to construct linear maps with specific properties.

Exercise 7.9.11 We ask whether there is a linear map $f : \mathbb{R}^3 \to \mathbb{R}^3$ which fulfils the conditions $f(1, 0, 2) = 0$ and $\mathrm{Im}(f) = \mathcal{L}((1, 1, 0), (2, -1, 0))$. Also, if such a map exists, is it unique?

We start by setting $v_1 = (1, 0, 2)$, $v_2 = (1, 1, 0)$, $v_3 = (2, -1, 0)$. Since a linear map is characterised by its action on the elements of a basis and v_1 is required to be in the kernel of f, we complete v_1 to a basis for $\mathbb{R}^3$. By using the elements of the canonical basis $\mathcal{E}_3$, we may take the set (v_1, e_1, e_2), which is indeed a basis: the matrix

$$\begin{pmatrix} 1 & 0 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

whose rows are given by (v_1, e_1, e_2) has rank 3 (the matrix is already reduced by rows). So we can take the basis $\mathcal{B} = (v_1, e_1, e_2)$ and define

$$\begin{aligned} f(v_1) &= 0_{\mathbb{R}^3}, \\ f(e_1) &= v_2, \\ f(e_2) &= v_3. \end{aligned}$$

Such a linear map satisfies the required conditions, since $\ker(f) = \{v_1\}$ and

$$\mathrm{Im}(f) = \mathcal{L}(f(v_1), f(e_1), f(e_2)) = \mathcal{L}(v_2, v_3).$$

With respect to the bases $\mathcal{E}$ and $\mathcal{B}$ we have

$$M_f^{\mathcal{E},\mathcal{B}} = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

In order to understand whether the required conditions can be satisfied by a different linear map f, we start by analysing whether the set (v_1, v_2, v_3) itself provides a basis for $\mathbb{R}^3$. As usual, we reduce by rows the matrix associated to the vectors,

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

Such a reduction gives $\mathrm{rk}(A) = 3$, that is $\mathcal{C} = (v_1, v_2, v_3)$ is a basis for $\mathbb{R}^3$. Then, let $g : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by

$$\begin{aligned} g(v_1) &= 0_{\mathbb{R}^3}, \\ g(v_2) &= v_2, \\ g(v_3) &= v_3. \end{aligned}$$

Also the linear map g satisfies the conditions we set at the beginning and the matrix

$$M_g^{\mathcal{C},\mathcal{C}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

represents its action by the basis $\mathcal{C}$. It seems clear that f and g are different. In order to prove this claim, we shall see that their corresponding matrices with respect to the same pair of bases differ. We need then to find $M_g^{\mathcal{E},\mathcal{B}}$. We know that $M_g^{\mathcal{E},\mathcal{B}} = M_g^{\mathcal{E},\mathcal{C}}\, M^{\mathcal{C},\mathcal{B}}$, with

$$M_g^{\mathcal{E},\mathcal{C}} = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix},$$

since the column vectors in $M_g^{\mathcal{E},\mathcal{C}}$ are given by $g(v_1)$, $g(v_2)$, $g(v_3)$. The columns of the matrix $M_{\mathcal{C},\mathcal{B}}$ are indeed the components with respect to $\mathcal{C}$ of the vectors in $\mathcal{B}$, that is

$$\begin{aligned} v_1 &= v_1, \\ e_1 &= \frac{1}{3}\, v_2 + \frac{1}{3}\, v_3, \\ e_2 &= \frac{2}{3}\, v_2 - \frac{1}{3}\, v_3. \end{aligned}$$

Thus we have

$$M^{\mathcal{C},\mathcal{B}} = \tfrac{1}{3} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 1 & -1 \end{pmatrix},$$

and in turn,

$$M_g^{\mathcal{E},\mathcal{B}} = M_g^{\mathcal{E},\mathcal{C}}\, M^{\mathcal{C},\mathcal{B}} = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \tfrac{1}{3} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

This shows that $M_g^{\mathcal{E},\mathcal{B}} \neq M_f^{\mathcal{E},\mathcal{B}}$, so that $g \neq f$.