Chapter 4

MORPHISMS OF VECTOR SPACES

§ 4.1 LINEAR TRANSFORMATIONS

The linear transformations are mappings between (two) vector spaces. They are similar to the *homomorphisms* met in the theory of algebraic structures : homomorphisms of groups, rings and fields. That is why such mappings are also called *linear morphisms* or *morphisms of vector spaces*. However, the terminology for such mappings differs from one textbook (or monograph) to another, from one author to another. Let us mention that, in the excellent monograph [G. Strang, 1988], page 116, the term of *transformation* is used. Such mappings are presented in close connection with matrices and they are initially defined on the most usual vector space, that is $\mathbb{R}^n$. For transformations defined on space $\mathbb{R}^2$, geometric interpretations are also given. In the textbook [E. Sernesi, 1993], the term of *linear maps* is preferred : Chapter 11 (from page 145) bears just this title. We prefer the term of *linear transformations* for mappings defined on a vector space and taking values into / onto another vector space. But the more general term of *linear morphisms* covers all the cases.

Definition 1.1. Let U & V be two vector spaces over the same field $\mathbf{K}$ (= $\mathbb{R}$ / = $\mathbb{C}$). A mapping $f : U \longrightarrow V$ is said to be a *linear transformation* (or *linear morphism*) if it satisfies both of the following properties (or axioms):

$$(\mathbf{LM}_1) \quad (\forall\, x_1, x_2 \in U)\ f(x_1 + x_2) = f(x_1) + f(x_2)\ ;$$

$$(\mathbf{LM}_2) \quad (\forall\, \lambda \in \mathbf{K})(\forall\, x \in U)\ f(\lambda x) = \lambda f(x)\,.$$

◊

Remark 1.1. Property $(\mathbf{LM}_1)$ states that a linear transformation or morphism is *additive* with respect to the vector sum in both spaces U & V, respectively. Property $(\mathbf{LM}_2)$ states that a linear morphism is *homogeneous* with respect to the external operation of multiplication by scalars (also defined in both spaces). Let us recall that – formally speaking – these two properties were also satisfied by a linear form : *Definition 1.1* in **§ 3.1** – page 49. Certainly, the two operations should be differently understood in the two sides of each equation : in the case of linear forms, the two linear operations were acting in the vector space V for the left-hand sides, while they were the field operations of addition and multiplication in the field of scalars, for the right-hand sides : see $(\mathbf{LF}_1)$ &

$(\mathbf{LF}_2)$ at page 49. In the case of linear morphisms, the addition and multiplication by scalars occur in both sides but they act on (possibly) different vector spaces, U & V.

Other properties of the linear forms (LFs) are formally retrieved for the linear transformations / morphisms. For instance, properties (axioms) $(\mathbf{LM}_1)$ and $(\mathbf{LM}_2)$ in *Def. 1.1* may be replaced by a single property / axiom ensuring that a linear mapping $f : U \longrightarrow V$ is a linear transformation.

Definition 1.1'. Let U & V be two vector spaces over the same field $\mathbf{K}$ (= $\mathbb{R}$ / = $\mathbb{C}$). A mapping

$$f : U \longrightarrow V \tag{1.1}$$

is a *linear transformation* (or *linear morphism*) if it satisfies

$$(\forall\, \lambda_1, \lambda_2 \in \mathbf{K})\,(\forall\, x_1, x_2 \in U)$$

$$\text{(LIN)}$$

$$\boxed{f(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 f(x_1) + \lambda_2 f(x_2).} \tag{1.2}$$

$\diamondsuit$

The equivalence between the two definitions is rather obvious. It can be checked as the similar equivalence for the LFs. Indeed,

$$(\forall\, \lambda_1, \lambda_2 \in \mathbf{K})\,(\forall\, x_1, x_2 \in U)$$
$$f(\lambda_1 x_1 + \lambda_2 x_2) \underset{(\mathbf{LM}_1)}{=} f(\lambda_1 x_1) + f(\lambda_2 x_2) \underset{(\mathbf{LM}_2)}{=} \lambda_1 f(x_1) + \lambda_2 f(x_2).$$

Conversely,

$$\begin{cases} \lambda_1 = \lambda_2 = 1 \ \text{ in (LIN)} \Rightarrow (\mathbf{LM}_1), \\ \lambda_1 = \lambda,\ \lambda_2 = 0 \ \&\ x_1 = x \ \text{ in (LIN)} \Rightarrow (\mathbf{LM}_2). \end{cases}$$

Property (LIN) is the ***linearity*** and it just gives the terms of *linear* morphism or *linear* transformation.

Let us also see that this property (LIN) or (1.2) could be replaced by an even simpler one :

$$(\forall\, \alpha \in \mathbf{K})\,(\forall\, x_1, x_2 \in U)\ \ f(x_1 + \alpha x_2) = f(x_1) + \alpha f(x_2), \tag{1.3}$$

but we prefer (1.2) since it admits a generalization to arbitrary linear combinations of (several) vectors under f :

PROPOSITION 1.1. *If* $f : U \longrightarrow V$ *is a linear transformation / morphism* *then*

$$(\forall\, \lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbf{K})\,(\forall\, x_1, x_2, \ldots, x_m \in U)$$

$$\boxed{f\Big(\sum_{i=1}^{m} \lambda_i\, x_i\Big) = \sum_{i=1}^{m} \lambda_i\, f(x_i).} \tag{1.4}$$

Proof. Formally, the proof is just the same as for the linear forms, in **§ 3.1** – PROPOSITION 1.1 at page 50 : it goes by induction on m. ■

This property (1.4) may be called the *extended linearity*. It can be written in a simpler way if we use the so-called 'matrix notations' introduced in **§ 1.2** for linear combinations of several vectors with several scalars. Let us recall those notations :

$$\mathcal{X} = [\, x_1\; x_2 \ldots\, x_m\,] \quad \text{and} \quad \Lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix}; \tag{1.5}$$

With notations of (1.5), a linear combination may be written as

$$\sum_{i=1}^{m} \lambda_i\, x_i = \Lambda^{\mathrm{T}} \cdot \mathcal{X}^{\mathrm{T}} = \mathcal{X} \cdot \Lambda = \sum_{i=1}^{m} x_i\, \lambda_i\,. \tag{1.6}$$

It follows from (1.4) with (1.6) that

$$f\left(\sum_{i=1}^{m} \lambda_i x_i\right) = f(\Lambda^{\mathrm{T}} \cdot \mathcal{X}^{\mathrm{T}}) = \Lambda^{\mathrm{T}} \cdot f(\mathcal{X}^{\mathrm{T}})\,. \tag{1.4'}$$

In (1.4'), $f(\mathcal{X}^{\mathrm{T}})$ represents the column vector of the values

$$f(x_i)\,,\; i = \overline{1, m}\,.$$

Using the alternative way to write a linear combination with the matrix notation (1.4'), the property of extended linearity can be written as

$$\boxed{f(\mathcal{X} \cdot \Lambda) = f(\mathcal{X}) \cdot \Lambda.} \tag{1.7}$$

In this formula (1.7), the linear form's values of (1.5) appear as the components of a row vector :

$$f(\mathcal{X}) = [\, f(x_1)\;\; f(x_2)\, \ldots\, f(x_m)\,]. \tag{1.8}$$

In what follows, we will prefer the notational alternative (1.7).

PROPOSITION 1.2. *If* $f : U \longrightarrow V$, $\dim U = m$ & $\dim V = n$, *is a linear morphism,* U *is spanned by basis* $A = [a_1\ a_2\ \dots\ a_m]$ *and* V *is spanned by* $B = [b_1\ b_2\ \dots\ b_n]$ *then the morphism* f *uniquely determines an m-by-n matrix* $F_{A,B}$ *defined by*

$$\boxed{f(A^{\mathrm{T}}) = F_{A,B}\, B^{\mathrm{T}}.} \tag{1.9}$$

Proof. For any $i \in \{1,2,\dots,m\}$, $f(a_i) \in V = \mathcal{L}(B)$. Therefore $f(a_i)$ admits a unique linear expression in the basis B of V :

$$f(a_i) = \sum_{i=1}^{n} \varphi_{ij} b_j. \tag{1.10}$$

More explicitly, the scalars φ_{ij} $(j = 1,2,\dots,n)$ are the coordinates of $f(a_i)$ in basis B. The linear expression (1.10) can be equivalently written using a matrix notation :

$$f(a_i) = F_i\, B^{\mathrm{T}} \quad \text{with} \quad F_i = [\varphi_{i1}\ \varphi_{i2}\ \dots\ \varphi_{in}],\ i = \overline{1,m}. \tag{1.11}$$

The m linear expressions of the form (1.11) can be written one under the other resulting a system of equations (equalities) which is equivalent to the matrix equation

$$f(A^{\mathrm{T}}) = F_{A,B}\, B^{\mathrm{T}} \quad \text{with} \quad F_{A,B} = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{bmatrix},\ A^{\mathrm{T}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix},\ B^{\mathrm{T}} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}. \tag{1.12}$$

Therefore the expression (1.9) is proved and it uniquely defines the matrix of the morphism in the (given) pair of bases. ■

The matrix $F_{A,B}$ is the matrix of the linear morphism / transformation in the pair of bases (A,B). As we shall see a little later, it essentially depends on the two bases selected in the two spaces. Let us also mention that this PROPOSITION 1.2 with formula (1.9) holds in the case when the two spaces are finite-dimensional, only.

The property (1.4) / (1.7) is involved in formulating the analytic expression of a morphism in a pair of bases (A,B) of the spaces U & V, respectively.

PROPOSITION 1.3. *If the vector space U is spanned by the basis* $A = [a_1\ a_2\ \dots\ a_m]$, *space* V *is spanned by basis* $B = [b_1\ b_2\ \dots\ b_n]$ *and the* *matrix of the linear morphism* $f : U \longrightarrow V$ *in the pair of bases* (A, B) *is* $F_{A,B}$ *then the image* $f(x)$ *of a vector* $x = A X_A \in U$ is

$$\boxed{f(x) = X_A^{\mathrm{T}} F_{A,B} B^{\mathrm{T}}.} \tag{1.13}$$

Proof. Formula (1.13) immediately follows from PROPOSITIONS 1.1 & 1.2 and – more precisely – from the property of extended linearity (1.4') by replacing

$\mathfrak{X} \to A$ and $\Lambda \to X_A$, with $x = A X_A$. ∎

Remarks 1.2. The proof is over, but we can give a more explicit (expanded) version of this formula (1.13), recalling – from **§ 1.1** – that

$$X_A^{\mathrm{T}} = [\,\xi_1\ \xi_2\ \ldots\ \xi_m\,] \;\Rightarrow\; x = \sum_{i=1}^{n} \xi_i a_i \underset{(1.4)}{\Rightarrow} f(x) = \sum_{i=1}^{m} \xi_i f(a_i) =$$

$$= \sum_{i=1}^{m} \xi_i \sum_{j=1}^{n} \varphi_{ij} b_j = \sum_{i=1}^{m} \sum_{j=1}^{n} \xi_i \varphi_{ij} b_j = \sum_{j=1}^{n} \Big(\sum_{i=1}^{m} \xi_i \varphi_{ij} \Big) b_j \,. \tag{1.14}$$

Formulas (1.13) and (1.14) effectively give the analytical expression of the image $f(x)$ of a vector x through the morphism f, in basis B of space V.

On another hand, PROPOSITION 1.3 may be considered as giving the converse result to PROPOSITION 1.2 : indeed, if an *m*-by-*n* matrix $[\varphi_{ij}]_{m,n}$ is given and the linear expression of a vector x of the form

$$x = A \cdot X_A = X_A^{\mathrm{T}} \cdot A^{\mathrm{T}} \in U$$

is known, then $X_A^{\mathrm{T}} \cdot [\varphi_{ij}] \cdot B^{\mathrm{T}}$ is just the image of vector x through f if $f(A^{\mathrm{T}}) = [\varphi_{ij}]_{m,n} \cdot B^{\mathrm{T}}$ gives the connection between basis A of U and basis B of V through f (see Eq. (1.9)). In other words, a linear transformation from U to V uniquely determines a matrix in the pair of bases (A, B) of the two spaces and – conversely – an m-by-n matrix $[\varphi_{ij}]_{m,n}$ plus a pair of bases uniquely determine a morphism $f : U \longrightarrow V$. However, this "equivalence" should not be formally understood ; moreover, *it is dependent on the bases in the two spaces.*

Example 1.1. If U is a vector space of dimension 4 over the field $\mathbb{R}$ and V is another real vector space with $\dim V = 3$ over the field $\mathbb{R}$, $A = [a_1\ a_2\ a_3\ a_4]$ is a basis spanning U and $B = [b_1\ b_2\ b_3]$ spans V, then the linear morphism $f : U \longrightarrow V$ with its matrix in the bases (A, B)

$$F_{A,B} = \begin{bmatrix} 2 & -4 & 0 \\ -1 & 0 & 3 \\ 0 & 0 & 1 \\ 2 & -2 & 2 \end{bmatrix},$$

then the image of the vector $x = -a_1 + 3\,a_2 + a_3$ in basis B is

$$f(x) = X_A^{\mathrm{T}} \cdot F_{A,B} \cdot B^{\mathrm{T}} = [-1 \;\; 3 \;\; 1 \;\; 0] \cdot \begin{bmatrix} 2 & -4 & 0 \\ -1 & 0 & 3 \\ 0 & 0 & 1 \\ 2 & -2 & 2 \end{bmatrix} \cdot B^{\mathrm{T}} =$$

$$= [-5 \;\; 4 \;\; 10] \cdot \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = -5\,b_1 + 4\,b_2 + 10\,b_3. \qquad \square$$

In the case when the two spaces are U & V are the most usual linear spaces (in examples and applications), that is – for instance – $U = \mathbf{K}^m$ & $V = \mathbf{K}^n$ or $U = \mathbb{R}^m$ & $V = \mathbb{R}^n$, the matrix of such a morphism and the linear expression of an image $f(X)$ have to be adapted, from (1.9) and (1.13). Hence it comes to linear transformations / morphisms of the form

$$f : \mathbf{K}^m \longrightarrow \mathbf{K}^n \quad \text{or} \quad f : \mathbb{R}^m \longrightarrow \mathbb{R}^n. \tag{1.15}$$

It is natural and convenient to see how the formulas (1.9) and (1.13) look if the two general bases (A, B) are replaced by the standard bases in the two spaces:

$$(A, B) \rightarrow (E_m, E_n),\; \mathbb{R}^m = \mathcal{L}(E_m) \;\&\; \mathbb{R}^n = \mathcal{L}(E_n) \tag{1.16}$$

where

$$E_m = [e_1 \; e_2 \; \ldots \; e_m] \;\&\; E_n = [e_1' \; e_2' \; \ldots \; e_n']. \tag{1.17}$$

With the standard bases in (1.16), formula (1.9) becomes

$$\boxed{f(E_m^{\mathrm{T}}) = F_{E_m, E_n} E_n^{\mathrm{T}}.} \tag{1.18}$$

As regards the analytical expression of $f(X)$, $X_A \rightarrow X_{E_m} = X$ and Eq. (1.13) turns to

$$\boxed{f(X) = X^{\mathrm{T}} \cdot F_{E_m, E_n} E_n^{\mathrm{T}} = X^{\mathrm{T}} \cdot F_{E_m, E_n} \cdot I_n.} \tag{1.19}$$

If we denote $f(X) = Y \in \mathbf{K}^n \,/\, \in \mathbb{R}^n$, it follows from (1.19) that

$$\boxed{f(X) = [x_1 \; x_2 \; \ldots \; x_m] \cdot [\varepsilon_{i,j}]_{m,n} \cdot I_n = Y_{E_n} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = Y.} \tag{1.20}$$

A simple example illustrates this formula (1.20).

Example 1.2. Let $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ be a linear morphism with its matrix in the pairs of bases (E_3, E_2) given as

$$F_{E_3,E_2} =_{\text{not}} [\varepsilon] = \begin{bmatrix} 2 & -1 \\ 0 & 3 \\ 1 & 4 \end{bmatrix} \;\&\; X = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}. \tag{1.21}$$

From Eq. (1.20) and the data in (1.21) we get

$$f(X) = [-1 \;\; 2 \;\; 1] \cdot \begin{bmatrix} 2 & -1 \\ 0 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} e_1' \\ e_2' \end{bmatrix} = [-1 \;\; 11] \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 11 \end{bmatrix} = Y. \qquad \square$$

Two subsets associated to a linear morphism are defined next.

***Definition* 1.2.** Let U & V be two vector spaces over the same field $\mathbf{K}$ (= $\mathbb{R}$ / = $\mathbb{C}$) and $f : U \longrightarrow V$ a *linear transformation* (or *linear morphism*). Then the *kernel* and the *image* (or *range*) of are defined by

$$\boxed{\operatorname{Ker} f \underset{\text{def}}{=} \{x \in U : f(x) = \mathbf{0}' \in V\};} \tag{1.22}$$

$$\boxed{\operatorname{Im} f \underset{\text{def}}{=} \{y \in V : (\exists\, x \in U)\;\; y = f(x)\}.} \tag{1.23}$$

$\Diamond$

Before stating (and proving) a result concerning these two subsets of spaces U and respectively V, let us remark that $\mathbf{0}'$ that occurs in (1.22) is the zero vector of the space V, which is, in general, different from $\mathbf{0} \in U$. Next, let us notice that the two subsets of (1.22) & (1.23) may be equivalently defined (or written) as

$$\boxed{\operatorname{Ker} f = f_{-1}(\mathbf{0}') \subseteq U;} \tag{1.22'}$$

$$\boxed{\operatorname{Im} f = f(U) = \{ f(x) : x \in U \} \subseteq V.} \qquad (1.23')$$

The notation f_{-1} used in (1.22') should be read as "the counterimage of" ; we do not use -1 as a superscript since f should not necessarily be invertible, and the notation f^{-1} is reserved for the *inverse* of the (bijective) mapping f. As regards notation $f(U)$, it deserves almost no explanation : in equation (1.23'), it denotes the image of a morphism defined on space U and taking values in V, hence $f(U)$ is simply the set of the images of all vectors in U.

The next result shows that these two subsets are more than simple subsets of the respective spaces.

PROPOSITION 1.4. *Let* U & V *be two vector spaces over the same field* $\mathbf{K}$ $(= \mathbb{R} \;/ = \mathbb{C})$ *and* $f : U \longrightarrow V$ *a linear transformation* (*or linear morphism*). *Then* $\operatorname{Ker} f$ *and* $\operatorname{Im} f$ *are subspaces of* U *and* V, *respectively.*

Proof. To prove that $\operatorname{Ker} f$ is a subspace of U it suffices to show that the subset in (1.22) / (1.22') is closed under arbitrary linear combinations of (two) vectors in it (see **§ 1.2**). Indeed, by property (LIN) in *Definition 1.1'*,

$$(\forall\, \lambda_1, \lambda_2 \in \mathbf{K})(\forall\, x_1, x_2 \in \operatorname{Ker} f)$$

$$f(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 f(x_1) + \lambda_2 f(x_2) \underset{(1.22)}{=} \lambda_1 \mathbf{0'} + \lambda_2 \mathbf{0'} = \mathbf{0'} \Rightarrow$$

$$\Rightarrow \lambda_1 x_1 + \lambda_2 x_2 \in \operatorname{Ker} f \Rightarrow \operatorname{Ker} f \subseteq_{\text{subsp}} U.$$

As regards $\operatorname{Im} f$, it is also closed under arbitrary linear combinations of (two) vectors :

$$(\forall\, \lambda_1, \lambda_2 \in \mathbf{K})(\forall\, y_1, y_2 \in \operatorname{Im} f)(\exists\, x_1, x_2 \in U) \;\; y_i = f(x_i) \;\; (i = \overline{1,2})$$

$$\text{and} \;\; \lambda_1 y_1 + \lambda_2 y_2 = \lambda_1 f(x_1) + \lambda_2 f(x_2) \underset{(\text{LIN})}{=} f(\lambda_1 x_1 + \lambda_2 x_2) \Rightarrow$$

$$\Rightarrow (\exists\, x \in U) \; x = \lambda_1 x_1 + \lambda_2 x_2 \;\&\; \lambda_1 y_1 + \lambda_2 y_2 = f(x) \Rightarrow \qquad (1.23)$$

$$\Rightarrow \lambda_1 y_1 + \lambda_2 y_2 \in \operatorname{Im} f \Rightarrow \operatorname{Im} f \subseteq_{\text{subsp}} V.$$

Although this is quite clear, let us notice that – on line (1.23) – we have made use of the property that a linear combinations of (two) vectors in U is also a vector in U, namely $x = \lambda_1 x_1 + \lambda_2 x_2$. This concludes the proof.

■

Before introducing the notion of rank of a morphism, before seeing its connections with the two subspaces of PROPOSITION 1.2 and before presenting some special types of morphisms, let us see how the kernel and the image of a morphism can be practically found. We start with an example of a morphism from an Euclidean space to another.

Example 1.3. Let $f: \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ be a linear morphism with its matrix in the pairs of bases (E_3, E_4) given as

$$F_{E_3,E_4} =_{\text{not}} [\varepsilon] = \begin{bmatrix} 1 & -2 & 0 & 1 \\ 3 & 0 & 6 & -3 \\ 1 & 4 & 6 & -5 \end{bmatrix}. \tag{1.24}$$

It is required to determine $\operatorname{Ker} f$ and $\operatorname{Im} f$ for this morphism.

By *Definition 1.2* - Eq. (1.22), adapted to our particular case,

$$\operatorname{Ker} f = \{X \in \mathbb{R}^3 : f(X) = \mathbf{0}' \in \mathbb{R}^4\}. \tag{1.25}$$

$$(1.20)\ \&\ (1.24) \Rightarrow f(X) = X^{\mathrm{T}} [\varepsilon]\, E_4^{\mathrm{T}}; \tag{1.26}$$

$$(1.24)\ \&\ (1.26) \Rightarrow f(X) = [x_1\ x_2\ x_3] \begin{bmatrix} 1 & -2 & 0 & 1 \\ 3 & 0 & 6 & -3 \\ 1 & 4 & 6 & -5 \end{bmatrix} E_4^{\mathrm{T}}; \tag{1.27}$$

For an easier way to obtain the general solution of the equation in (1.25) it is convenient to transpose the product in the right-hand side of (1.27) :

$$(1.27)\ \&\ (1.25) \Rightarrow E_4 \begin{bmatrix} 1 & -2 & 0 & 1 \\ 3 & 0 & 6 & -3 \\ 1 & 4 & 6 & -5 \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}' \in \mathbb{R}^4. \tag{1.28}$$

But (as a matrix) E_4 is just the identity matrix of order 4, I_4; hence it may be omitted from (1.28) and the resulting matrix equation *is equivalent to a homogeneous system* that can be solved by the Gaussian elimination method, as in § **1.1** :

$$F_{E_3,E_4}^{\mathrm{T}} = \begin{bmatrix} 1 & 3 & 1 \\ -2 & 0 & 4 \\ 0 & 6 & 6 \\ 1 & -3 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 1 \\ 0 & 6 & 6 \\ 0 & 6 & 6 \\ 0 & -6 & -6 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix} \sim$$

$$\sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \end{bmatrix} \Rightarrow X(\alpha) = \begin{bmatrix} 2\alpha \\ -\alpha \\ \alpha \end{bmatrix} = \alpha \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}. \tag{1.29}$$

Hence the vector in (1.29) is a general vector in $\operatorname{Ker} f$.

The image is easier to be found. It follows from (1.26) - (1.27) that an arbitrary image vector $Y = f(X)$ is given by the right-hand sides of these equation. But the respective products yield a row vector. In order to obtain a column vector in $\mathbb{R}^4$ the product in the r.h.s. (right side) of (1.27) has to be transposed, as we have did it for determining the kernel as the solution set of system (1.28). Hence we arrive at the image $Y = f(X)$ given by

$$Y = F_{E_3,E_4}^{\mathrm{T}} X = \begin{bmatrix} 1 & 3 & 1 \\ -2 & 0 & 4 \\ 0 & 6 & 6 \\ 1 & -3 & -5 \end{bmatrix} X = x_1 \begin{bmatrix} 1 \\ -2 \\ 0 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 0 \\ 6 \\ -3 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 4 \\ 6 \\ -5 \end{bmatrix}. \tag{1.30}$$

It follows from (1.30) that the subspace $\mathrm{Im}\, f$ is generated by the three columns of the matrix in (1.30). Let us recall from **§ 2.1** that to find a subspace practically means *to determine a basis spanning it.* It is therefore necessary to check whether these three vector are linearly independent and – if not – to select a basis as a subfamily thereof. As we proceeded in **§ 1.1** , we can easily obtain this basis by a couple of transformations on that matrix. It follows to obtain a *quasi-triangular* equivalent matrix, as we did it for finding the rank of a matrix. We may see that the second vector can be replaced by a simpler (or "shorter") vector by taking one third of its.

$$F_{E_3,E_4}^{\mathrm{T}} \sim \begin{bmatrix} 1 & 1 & 1 \\ -2 & 0 & 4 \\ 0 & 2 & 6 \\ 1 & -1 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 6 \\ 0 & 2 & 6 \\ 0 & -2 & -6 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{1.31}$$

It follows from (1.31) that the first two vectors in (1.30), or the first two columns of the first matrix in the chain of (1.31). To conclude with this example, the two required subspaces are spanned by the following bases :

$$\mathrm{Ker}\, f = \mathscr{L}(A),\ A = [a], \qquad a = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}; \tag{1.32}$$

$$\mathrm{Im}\, f = \mathscr{L}(B),\ B = [b_1\ \ b_2]:\ b_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \\ 1 \end{bmatrix}, \quad b_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \\ -1 \end{bmatrix}. \tag{1.33}$$

□

The example just closed involved a morphism between two Euclidean spaces. But the problem of finding the kernel and the image of a morphism can be approached and solved in a more general setting, when $f : U \longrightarrow V$ and the two spaces are respectively spanned by (let us say) abstract bases,

(A, B). We are going to reformulate an exam subject from our web page, http://math.etc.tuiasi.ro/ac/, namely subject $\boxed{\text{C.4}}$ from section **AG.1**. In the statement of that subject, the morphism was of the form $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ but we replace the two spaces by more general ones, keeping their respective dimensions.

Example 1.4. Le $f : U \longrightarrow V$ be a linear morphism between the spaces U with $\dim U = 3$ and V with $\dim V = 2$. they are assumed to be spanned by the respective bases $A = [a_1\ a_2\ a_3]$ and $B = [b_1\ b_2]$. The morphism is defined by its matrix in the pair of bases (A, B), according to formula (1.9) :

$$F_{A,B} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 0 & -1 \end{bmatrix}. \tag{1.34}$$

In is required to find the coordinates X_A of a vector in $\operatorname{Ker} f$, respectively the coordinates Y_B of a vector in $\operatorname{Im} f$.

The significance of the matrix in (1.34) is the following :

$$\begin{cases} f(a_1) = \ b_1, \\ f(a_2) = 2b_1 + b_2, \\ f(a_3) = \ -b_2. \end{cases} \tag{1.35}$$

The matrix equation (and then the corresponding homogeneous system) that should be satisfied by the coordinates X_A follow from definition (1.22) of the kernel and from formula (1.13) – at page 126 – for the image $f(x)$ when the morphism is given by its matrix $F_{A,B}$.

$$f(x) = \mathbf{0}' \in \mathbb{R}^2 \underset{(1.13)}{\Rightarrow} X_A^{\mathrm{T}} F_{A,B} B^{\mathrm{T}} = \mathbf{0}' = [0\ \ 0] \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \tag{1.36}$$

In the rightmost side of (1.36) we have taken into account the obvious property of the *zero vector* of having zero coordinates in any basis. It follows from this Eq. (1.36), due to the uniqueness of the coordinates of a vector in any given basis, the vector equation which we write and transpose :

$$X_A^{\mathrm{T}} F_{A,B} B^{\mathrm{T}} = [0\ \ 0] \Rightarrow F_{A,B}^{\mathrm{T}} X_A = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{1.37}$$

The homogeneous system (1.37) is solved on its matrix, that is the transpose of the matrix in (1.34) :

$$F_{A,B}^{\mathrm{T}} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \end{bmatrix} \Rightarrow X_A = \begin{bmatrix} -2\alpha \\ \alpha \\ \alpha \end{bmatrix} = \alpha \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}. \quad (1.38)$$

The explicit coordinates of a vector in $\operatorname{Ker} f$, with the notations of **§ 1.1**, are

$$X_A = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} = \alpha \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \Rightarrow \begin{cases} \xi_1 = -2\alpha, \\ \xi_2 = \alpha, \\ \xi_3 = \alpha \end{cases} \Rightarrow x = \alpha(-2a_1 + a_2 + a_3). \quad (1.39)$$

The result in (1.38-39) can be checked by determining the image of this vector, from the data (matrix) in (1.34) and by formula (1.13) :

$$f(x) \underset{(1.13)}{=} X_A^{\mathrm{T}} F_{A,B} B^{\mathrm{T}} \underset{(1.34,39)}{=} [-2\alpha \ \ \alpha \ \ \alpha] \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 0 & -1 \end{bmatrix} B^{\mathrm{T}} = [0 \ \ 0] B^{\mathrm{T}} = \mathbf{0}'.$$

The image of the morphism with matrix (1.34) can be linearly expressed in basis B using the explicit expressions of the vectors of A, earlier given in (1.35) :

$$x = \xi_1 a_1 + \xi_2 a_2 + \xi_3 a_3 \Rightarrow f(x) = \xi_1 f(a_1) + \xi_2 f(a_2) + \xi_3 f(a_3) \underset{(1.35)}{=}$$

$$= \xi_1 b_1 + \xi_2(2b_1 + b_2) - \xi_3 b_2 = (\xi_1 + 2\xi_2) b_1 + (\xi_2 - \xi_3) b_2 =$$

$$= \beta_1 b_1 + \beta_2 b_2. \quad (1.40)$$

The notation we have used for writing expression (1.40) is obvious. When the coordinates ξ_1, ξ_2, ξ_3 of x independently vary over $\mathbb{R}$, its coordinates β_1, β_2 can take every real value. If this assertion is not so evident, we can consider the non-homogeneous system

$$\begin{cases} \xi_1 + 2\xi_2 = \beta_1, \\ \xi_2 - \xi_3 = \beta_2. \end{cases} \quad (1.41)$$

This system (1.41) is not determined, in the sense that its solution depends on a parameter (for instance $\xi_2 = \lambda$). But it is essential that it has solutions.

Hence, any would be a vector $y \in V$ with its real coordinates $Y_B = [\beta_1 \ \ \beta_2]^{\mathrm{T}}$, a vector x with $X_A = [\xi_1 \ \ \xi_2 \ \ \xi_3]^{\mathrm{T}}$ exists such that

$$y = f(x) \Rightarrow \operatorname{Im} f = V.$$

□

Definition 1.3. **(Rank of a morphism).** Let U & V be two vector spaces over the same field $\mathbf{K}$ (= $\mathbb{R}$ / = $\mathbb{C}$) and $f : U \longrightarrow V$ a linear transformation (or linear morphism). Then the *rank* of f is defined as the rank of its matrix $F_{A,B}$ in any pair of bases A and B of space U, respectively V.

Hence, if the matrix $F_{A,B}$ is defined as in PROPOSITION 1.3 - Eq. (1.9), that is $f(A^{\mathrm{T}}) = F_{A,B}\, B^{\mathrm{T}}$, then

$$\boxed{\operatorname{rank} f = \operatorname{rank} F_{A,B}\,.} \tag{1.42}$$

It would follow, from the defining Eq. (1.42), that this notion of rank would be dependent on the two bases A & B. However, we shall see – a little later – how the change of bases affect the matrix of a morphism, but not its rank.

There exists a connection between the rank of a morphism, its kernel and its image. But let us firstly notice that

$$\begin{aligned}\dim U = m \;\;\&\;\; \dim V = n \;\Rightarrow\; F_{A,B} \in \mathscr{M}_{m,n}(\mathbb{R}) \;\Rightarrow\\ \Rightarrow\; \operatorname{rank} f \underset{\text{not}}{=} r \le \min\{m, n\}\,.\end{aligned} \tag{1.43}$$

Certainly, this inequality also holds when the two (finite-dimensional) vector spaces and the morphism are considered on a more general field, $\mathbf{K}$ instead of $\mathbb{R}$.

PROPOSITION 1.5. *Let U & V be two vector spaces over the same field $\mathbf{K}$ (= $\mathbb{R}$ / = $\mathbb{C}$) and $f : U \longrightarrow V$ a linear transformation (or linear morphism). If $\dim U = m$, $\dim V = n$ and $\operatorname{rank} f = r$ then*

$$\boxed{\dim \operatorname{Ker} f = m - r} \quad \& \quad \boxed{\dim \operatorname{Im} f = r\,.} \tag{1.44}$$

Proof. For the dimension of the *kernel*, let us see that the coordinates of a vector in the kernel should satisfy the homogeneous system of the form (1.37), that is

$$X_A^{\mathrm{T}} F_{A,B}\, B^{\mathrm{T}} = [0 \,\ldots\, 0]\, B^{\mathrm{T}} \;\Rightarrow\; F_{A,B}^{\mathrm{T}} X_A = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^n. \tag{1.45}$$

The matrix of this system (1.45) is of size (n, m). It is known from **§ 1.2** (and from the highschool Algebra as well) that the solution set S of a homogeneous system of this size depends of $m - r$ parameters ; in terms of subspaces and their dimensions, this means that

$$\operatorname{Ker} f = S \;\Rightarrow\; \dim \operatorname{Ker} f = m - r\,.$$

As regards the *image*, we should also go back to an earlier formula, Eq. (1.13) at page 149 :

$$f(x) = X_A^{\mathrm{T}} F_{A,B} B^{\mathrm{T}}. \tag{1.46}$$

Therefore $y = B\,Y_B = Y_B^{\mathrm{T}} B^{\mathrm{T}} \in \operatorname{Im} f$ if and only

$$(\exists\, x \in U)\; x = X_A^{\mathrm{T}} A^{\mathrm{T}} :\, y = f(x) = X_A^{\mathrm{T}} F_{A,B} B^{\mathrm{T}}. \tag{1.47}$$

But $y = B\,Y_B = Y_B^{\mathrm{T}} B^{\mathrm{T}}$ and thus we get, with Eq. (1.47), the equation

$$X_A^{\mathrm{T}} F_{A,B} B^{\mathrm{T}} = Y_B^{\mathrm{T}} B^{\mathrm{T}} \iff X_A^{\mathrm{T}} F_{A,B} = Y_B^{\mathrm{T}} \iff F_{A,B}^{\mathrm{T}} X_A = Y_B. \tag{1.48}$$

The last matrix - vector equation in (1.48) represents a non-homogeneous system. Assuming that a vector $y = B\,Y_B \in V$ is given, we have to establish whether it is contained in $\operatorname{Im} f$. By *Definition 1.2* - Eqs. (1.23) / (1.23') at pages 128 / 129, we have to check the existence of (at last) an x satisfying (1.47) - (1.48). We have already noticed that the size of $F_{A,B}^{\mathrm{T}}$ is (n, m). The system (1.48) has solutions $\iff$ a condition of consistency is satisfied. By Rouché's Theorem, (see **§ 1.2**) we should have

$$\operatorname{rank} F_{A,B}^{\mathrm{T}} = \operatorname{rank}\,[F_{A,B}^{\mathrm{T}} \mid Y_B]. \tag{1.49}$$

If (1.49) holds, this means that $\operatorname{rank} F_{A,B}^{\mathrm{T}} = \operatorname{rank}\,[F_{A,B}^{\mathrm{T}} \mid Y_B] = r$ and $n - r$ equations of the system (1.48) may be removed / deleted. The system is consistent and (the vectors in) its solution set S will depend on $m - r$ parameters (the secondary unknowns). The vectors in S will be counter-images of

$$Y_B :\, X_A(\lambda_1, \ldots, \lambda_{m-r}) \in f_{-1}(Y_B).$$

But the consistency of the system (1.48) is equivalent to the condition that its vector of free terms belongs to the (sub)space generated by the columns of its coefficient matrix. See the definition of this subspace for a general matrix A, denoted COLSP_A, in **§ 1.3** - Eq. (3.21) at page 66, and the just mentioned condition for system's consistency as PROPOSITION 2.10 - Eq. (2.35) at page 46 in **§ 1.2** of ⟦𝒜. 𝒞., 2014⟧ . We slightly change this notation to $\mathrm{COLSP}\,(A)$ and we arrive to the equivalence

$$y = B\,Y_B \in \operatorname{Im} f \iff Y_B \in \mathrm{COLSP}\,(F_{A,B}^{\mathrm{T}}). \tag{1.50}$$

But the rank of the matrix in (1.50) is just r and therefore (a minimum number of) only r columns of $F_{A,B}^{\mathrm{T}}$ generate the vectors in $\operatorname{Im} f$ and thus the second equality in (1.44) is also proved. ■

Remarks 1.2. Our proof for Eq. (1.44-2) essentially consists of Eqs. (1.47) + (1.48) + (1.49). However, we have offered more details and explanations. Other proofs, involving the bases spanning U and $\operatorname{Ker} f$, can be found in textbooks of LINEAR ALGEBRA like [E. Sernesi, 1993] and [C. Radu, 1996]. In fact, the proofs in these two references are essentially the same. We presented them, with appropriate changes of notations, in ⟦$\mathcal{A}$. $\mathcal{C}$., 2014⟧ (that is, the extended version of our textbook of 1999).

We saw, in **Examples 1.3** & **1.4**, how the kernel and the image of a linear morphism can be found, starting from the matrix of such a mapping. But, in very many applications (exercises) where these two subspaces are required to be found, the morphisms of the form $f : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ are given by the image of the vector $X \in \mathbb{R}^m$ written as the (column) vector $Y = f(X) \in \mathbb{R}^n$ whose components are *linear forms* in the components of $X : x_1, x_2, \ldots, x_m$. If the morphism is given this way, its kernel and image can be very easily found. The kernel is just the solution subspace of the homogeneous system $f(X) = \mathbf{0}' \in \mathbb{R}^n$ while a basis spanning $\operatorname{Im} f$ is described below.

This informal description becomes more explicit if we denote by M the matrix of the homogeneous system, just mentioned. Thus

$$f(X) = \mathbf{0}' \in \mathbb{R}^n \iff MX = [0\ 0\ \ldots\ 0]^{\mathrm{T}} \in \mathbb{R}^n. \tag{1.51}$$

If we compare this equation with Eqs. (1.19 - 20) at pages 127 / 128, it clearly follows that our matrix M of (1.51) is just the transpose of the matrix F_{E_m, E_n} there involved :

$$M = F_{E_m, E_n}^{\mathrm{T}}. \tag{1.52}$$

Hence $\operatorname{Ker} f$ is the solution set S of the homogeneous system (1.51) whose matrix is

$$M \in \mathcal{M}_{n,m}(\mathbb{R}),\ M = [\,M^1\ \ldots\ M^j\ \ldots\ M^m\,]. \tag{1.53}$$

With this structure in (1.53), the homogeneous system (1.51) can be – more explicitly – written as

$$M^1 x_1 + \ldots + M^j x_j + \ldots + M^m x_m = [0\ 0\ \ldots\ 0]^{\mathrm{T}} \in \mathbb{R}^n. \tag{1.51'}$$

As regards $\operatorname{Im} f$, a basis B spanning it can be rather easily found by selecting r linearly independent columns among the columns of (1.53). The method we presented in **§ 1.2** can be conveniently used : M must be brought to a quasi-triangular form for identifying these r independent columns, while the same Gaussian elimination technique can be applied until a quasi-diagonal form of the matrix M is obtained for getting $S = \operatorname{Ker} f$. The next example illustrates this descriptive presentation.

Example 1.5. Let f and g be the two linear morphisms given below :

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2,\ f(X) = \begin{bmatrix} x_1 - x_3 \\ x_2 + x_3 \end{bmatrix}; \tag{1.54}$$

$$g : \mathbb{R}^3 \longrightarrow \mathbb{R}^4,\ g(X) = \begin{bmatrix} x_3 \\ x_1 + x_2 \\ x_1 \\ x_1 - x_2 \end{bmatrix}. \tag{1.55}$$

It is required to find their kernels and images.

It follows from (1.54) that

$$X \in \operatorname{Ker} f \iff \begin{bmatrix} x_1 - x_3 \\ x_2 + x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow X(\alpha) = \begin{bmatrix} \alpha \\ -\alpha \\ \alpha \end{bmatrix}. \tag{1.56}$$

The matrix M also comes from (1.54), but we give it a subscript :

$$M_f = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}. \tag{1.57}$$

In fact, the homogeneous system (1.56) was too simple for needing this matrix for its solution ; but we have written it for illustrating (1.52). It is also clear that the first two columns of M_f are linearly independent and they form the basis for $\operatorname{Im} f$; in fact, this basis is just the standard basis $E = [e'_1\ e'_2]$ of $\mathbb{R}^2$. Hence $\operatorname{Im} f = \mathbb{R}^2$. Equations (1.44) are trivially satisfied since

$$(1.56) \Rightarrow \operatorname{rank} F = \dim \operatorname{Im} f = 2 \ \&\ \operatorname{null} f = 1.$$

The matrix of this simple morphism in the pair of standard bases (E_3, E_2) is

$$F_{E_3, E_2} = M_f^{\mathrm{T}} \underset{(1.57)}{=} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}. \tag{1.58}$$

For the second morphism in (1.55) we proceed analogously.

$$(1.65) \Rightarrow \quad M_g = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}. \tag{1.59}$$

The kernel is obtained from (1.55). The H-system is

$$X \in \operatorname{Ker} f \iff \begin{cases} \qquad x_3 = 0, \\ x_1 + x_2 = 0, \\ x_1 = 0, \\ x_1 - x_2 = 0 \end{cases} \Rightarrow X(\beta) = \begin{bmatrix} 0 \\ \beta \\ 0 \\ -\beta \end{bmatrix}. \tag{1.60}$$

The three columns of the matrix in (1.59) are linearly independent since the determinant formed with its first 3 rows is $= -1 \neq 0$. $\operatorname{Im} g$ is spanned by these three column vectors, hence a vector in the image is of the form (taking $x_1 = \beta,\ x_2 = \delta,\ x_3 = \varepsilon$)

$$Y(\gamma, \delta, \varepsilon) = \gamma \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \delta \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix} + \varepsilon \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{1.61}$$

The three column vectors in (1.61) form a basis for the image of g. But the same result can be obtained, in a simpler way, by expanding the vector in the r.h.s. of (1.55) :

$$Y = g(X) = \begin{bmatrix} x_3 \\ x_1 + x_2 \\ x_1 \\ x_1 - x_2 \end{bmatrix} = x_1 \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

$$(1.60)\ \&\ (1.61) \Rightarrow \quad \operatorname{null} g = 1,\ \dim \operatorname{Im} g = \operatorname{rank} g = 3.$$

Equations (1.44) are thus satisfied. Let us close this example with the remark that the images of f cover the space where it takes values while the images of g do not cover $\mathbb{R}^4$: $\operatorname{Im} g \subset \mathbb{R}^4$. □

Properties & Classification of Linear Morphisms

A series of properties of the linear morphisms are going to be defined and studied (characterized). In fact, they are not specific to the mappings between vector spaces. They are met and studied in the highschool ALGEBRA, in connection with algebraic structures and mappings between them like the homomorphisms, isomorphisms, etc. They are also met in MATHEMATICAL ANALYSIS - CALCULUS.

Definition 1.4. **(Properties of morphisms).** Let U & V be two vector spaces over the same field $\mathbf{K}$ (= $\mathbb{R}$ / = $\mathbb{C}$) and $f : U \longrightarrow V$ a linear transformation (or linear morphism). f is said to be *injective* if ,

$$(\forall\, x_1, x_2 \in U)\ \ x_1 \neq x_2 \ \Rightarrow\ f(x_1) \neq f(x_2) \ \text{, or} \tag{1.62}$$

$$(\forall\, x_1, x_2 \in U)\ \ f(x_1) = f(x_2) \ \Rightarrow\ x_1 = x_2 . \tag{1.63}$$

The morphism f is said to be *surjective* (or *onto*) if

$$\mathrm{Im}\, f = f(U) = V \iff (\forall\, y \in V)\ \ f_{-1}(y) \neq \varnothing . \tag{1.64}$$

A linear transformation / morphism $f : U \longrightarrow V$ is *bijective* if it is both injective and surjective.

Remarks 1.3. **1°** It is clear that definitions (1.62) & (1.63) for the injectivity of a morphism are equivalent. For instance, if we take some $x_1 \neq x_2$ but assume that $f(x_1) = f(x_2)$ it would follow by (1.63) that $x_1 = x_2$! Hence (1.63) $\Rightarrow$ (1.62) and the converse implication holds, too.

2° Property (1.62) tells that an injective morphism takes different values for distinct arguments, while (1.63) states that an injective morphism takes the same value on two vectors only if the two arguments coincide.

3° A surjective morphism has an image (or range) that covers the whole space where it takes values : any vector $y \in V$ has a nonempty counterimage. This means that $(\forall\, y \in V)(\exists\, x \in U)\ \ y = f(x)$.

4° As a matter of terminology, a bijective morphism is called an *isomorphism*, and two vector spaces U & V such that there exists an isomorphism from one to the other are said to be *isomorphic*. In such a case it can be used the notation

$$U \overset{\text{iso}}{\longleftrightarrow} V \quad \text{or} \quad U \overset{\text{iso}}{\sim} V.$$

These two properties (that can be verified or not by a certain morphism) are connected with the other two notions, the kernel and the image. These connections are stated in

PROPOSITION 1.6. *Let U & V be two vector spaces over the same field* $\mathbf{K}$ (= $\mathbb{R}$ / = $\mathbb{C}$) *and $f : U \longrightarrow V$ a linear transformation (or linear morphism).*

① f *is injective* $\iff \mathrm{Ker}\, f = \{\mathbf{0}\}$; (1.65)

② f *is an isomorphism* $\iff [\mathrm{Ker}\, f = \{\mathbf{0}\}\ \&\ \mathrm{Im}\, f = V]$. (1.66)

Proof. ① ($\Rightarrow$) Let us assume that f is an injective morphism, that is both properties (1.61) & (1.62) are satisfied ; in fact, it suffices to assume that one of them holds since they are equivalent. Let us also suppose that

$$\mathrm{Ker}\, f \neq \{\mathbf{0}\} \;\Rightarrow\; (\exists\, u \in \mathrm{Ker}\, f)\; u \neq \mathbf{0}. \tag{1.67}$$

Then (1.67) $\Rightarrow$ $x + u \neq x$. For some fixed $x \in U$ we have

$$f(x+u) = f(x) + f(u) = f(x) + \mathbf{0}' = f(x). \tag{1.68}$$

But the equation (1.68) contradicts the injectivity of f by (1.62) : different vectors would have the same image through f. The converse implication ($\Leftarrow$) can be proved as follows. Let us consider two vectors $x_1, x_2 \in U$ such that $f(x_1) = f(x_2) \Rightarrow f(x_1) - f(x_2) = \mathbf{0}' \in V$. But this latter equation plus the defining property (LIN) of any morphism leads to

$$f(x_1 - x_2) = \mathbf{0}' \underset{(1.22)}{\Rightarrow} x_1 - x_2 \in \mathrm{Ker}\, f \underset{(1.65)}{=} \{\mathbf{0}\} \Rightarrow$$
$$\Rightarrow x_1 - x_2 = \mathbf{0} \Rightarrow x_1 = x_2 \underset{(1.72)}{\Rightarrow} f \text{ is injective}.$$

② Immediately follows from *Definition 1.4* and ①. We have introduced this characterization in the statement taking into account the importance of this type of morphisms and their characterization in terms of kernel and image.

■

A couple of other properties of morphisms to be introduced next need another preliminary definition.

Definition 1.5. **(Special morphisms, composite morphisms).** Let U & V be two vector spaces over the same field $\mathbf{K}$ (= $\mathbb{R}$ / = $\mathbb{C}$). The *identical* morphisms on each of the two spaces are respectively defined by :

$$\mathrm{id}_U : U \longrightarrow U,\ (\forall\, x \in U)\ \mathrm{id}_U(x) = x\,; \tag{1.69}$$

$$\mathrm{id}_V : V \longrightarrow V,\ (\forall\, y \in V)\ \mathrm{id}_V(y) = y. \tag{1.70}$$

The *zero morphism* is defined by

$$\mathrm{O}_U : U \longrightarrow U,\ (\forall\, x \in U)\ \mathrm{O}_U(x) = \mathbf{0}' \in V. \tag{1.71}$$

Given two morphisms $f : U \longrightarrow V$ & $g : V \longrightarrow W$, the *composite* morphism of f with g is defined by

$$g \circ f : U \longrightarrow W\ (\forall\, x \in U)\ (g \circ f)(x) \underset{\text{def}}{=} g[f(x)]. \tag{1.72}$$

◊

Remarks 1.5. (i) As a matter of notation, the identical morphisms of (1.79) and (1.70) are denoted, in many textbooks, as $\mathbf{1}_U$, respectively $\mathbf{1}_V$. In fact, the definitions in (1.69) and (1.70) are the same, but the spaces on which each of them is defined differ. Obviously, $\mathrm{Ker}\,\mathbf{1}_U = \{\mathbf{0}\}$ & $\mathrm{Im}\,\mathbf{1}_U = U$. The same equations hold for $\mathbf{1}_V$.

(ii) The zero morphism of (1.71) is a constant mapping : it takes a unique value on any argument :

$$x \in U:\ \mathrm{O}_U(U) = \{\mathbf{0}'\} \subset V \ \Rightarrow\ \mathrm{Ker}\,\mathrm{O}_U = U \ \&\ \mathrm{Im}\,\mathrm{O}_U = \{\mathbf{0}'\}.$$

It can be easily checked that this degenerate (or trivial) morphism satisfies the definition of a linear morphism, *Def. 1.1* at page 122.

(iii) The operation of composing two morphisms, by (1.72), is the usual operation of composing two mappings or functions. The set of all linear morphisms between two spaces U, V is denoted as $\mathrm{Hom}(U, V)$. With this notation, (1.72) can be rewritten as follows : for any

$$f \in \mathrm{Hom}(U, V) \ \&\ g \in \mathrm{Hom}(V, W),$$

$$g \circ f : U \longrightarrow W \ (\forall\, x \in U)\ (g \circ f)(x) \underset{\text{def}}{=} g[f(x)].$$

Let us check that the composite mapping thus defined is a linear morphism, too. Denote

$$\boxed{g \circ f = h} \tag{1.73}$$

and let us check that $h \in \mathrm{Hom}(U, V)$.

$$(\forall\, x \in U)\ h(x) \underset{(1.73)}{=} (g \circ f)(\lambda_1 x_1 + \lambda_2 x_2) \underset{\text{def}}{=} g[f(\lambda_1 x_1 + \lambda_2 x_2) \underset{(\text{LIN})}{=}$$

$$= g[f(\lambda_1 x_1 + \lambda_2 x_2)] \underset{(\text{LIN})}{=} g[\lambda_1 f(x_1) + \lambda_2 f(x_2)] =$$

$$= g(\lambda_1 y_1 + \lambda_2 y_2) \underset{(\text{LIN})}{=} \lambda_1 g(y_1) + \lambda_2 g(y_2) \underset{(1.83)}{=}$$

$$= \lambda_1 h(x_1) + \lambda_2 h(x_2) \ \Rightarrow\ h \in \mathrm{Hom}(U, V).$$

The matrix of a composite morphism, in a triple of bases A , B , C with

$$U = \mathcal{L}(A),\ V = \mathcal{L}(B),\ W = \mathcal{L}(C),$$

is obtained from the two matrices

$$F_{A,B} = [\varphi_{i,j}]_{m,n},\ G_{B,C} = [\gamma_{j,k}]_{n,p}. \tag{1.74}$$

It follows from (1.74) that

$$f(A^{\mathrm{T}}) = F_{A,B}\, B^{\mathrm{T}} \ \&\ g(B^{\mathrm{T}}) = G_{B,C}\, B^{\mathrm{T}}. \tag{1.75}$$

$$(1.73)\ \&\ (1.75) \Rightarrow \quad h(A^{\mathrm{T}}) \underset{(1.73)}{=} (g \circ f)(A^{\mathrm{T}}) \underset{(1.75-1)}{=} g(F_{A,B} B^{\mathrm{T}}) =$$
$$= F_{A,B}\, g(B^{\mathrm{T}}) \underset{(1.75-2)}{=} F_{A,B}\, G_{B,C}\, C^{\mathrm{T}}. \tag{1.76}$$

$$(1.76) \quad \Rightarrow \quad h(A^{\mathrm{T}}) = \ldots = F_{A,B}\, G_{B,C}\, C^{\mathrm{T}} \ \Rightarrow \ H_{A,C} = F_{A,B}\, G_{B,C}.$$

Therefore, we have effectively proved the next

PROPOSITION 1.7. *Let* U, V, W *be three finitely generated vector spaces over the same field* $\mathbf{K}\,(=\mathbb{R}\,/=\mathbb{C})$ *and* $f: U \longrightarrow V$, $g: V \longrightarrow W$ *two linear transformations* (*or linear morphisms*). If A, B, C *are three bases respectively spanning the spaces U, V, W and the two matrices of the morphisms* f & g *are the ones in* (1.74), *then the composite mapping* $g \circ f = h$ *defined by* (1.72) *is also a linear morphism from* U *to* W *and its matrix in the pair of bases* (A,C) *is*

$$\boxed{H_{A,C} = F_{A,B}\, G_{B,C}.} \tag{1.77}$$

Proof. As we have just mentioned, the proof was already presented in equations (1.72), . . . , (1.77). Let us only see that

$$\dim U = m,\ \dim V = n,\ \dim W = p \ \Rightarrow$$
$$\Rightarrow \ F_{A,B} = [\varphi_{i,j}]_{m,n}\ \&\ G_{B,C} = [\gamma_{j,k}]_{n,p} \ \Rightarrow \ H_{A,C} = [\chi_{i,k}]_{m,p}.$$

■

Example 1.6. Let f and g be the two linear morphisms

$$f: U \longrightarrow V \ \&\ g: V \longrightarrow W$$

with $\dim U = 3$, $\dim V = 2$, $\dim W = 4$. The two morphisms are given, in the pairs of bases (A,B) and (B,C), by their respective matrices

$$F_{A,B} = \begin{bmatrix} 2 & -1 \\ 5 & 5 \\ 3 & 0 \end{bmatrix},\ G_{B,C} = \begin{bmatrix} 2 & 1 & 0 & 4 \\ -2 & 0 & 0 & 1 \end{bmatrix}. \tag{1.78}$$

It is required to calculate the matrix $H_{A,C}$ of the composite morphism $g \circ f = h$ and to determine the linear expression $h(x)$ of the vector

$$x = 3a_1 - a_2 + 4a_2 \tag{1.79}$$

in the basis C of W using this matrix and – also – the intermediate expression of $y = f(x)$ in basis B of V.

Eq. (1.77) with the matrices in (1.78) leads to

$$H_{A,C} = \begin{bmatrix} 2 & -1 \\ 5 & 5 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 & 4 \\ -2 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 2 & 0 & 7 \\ 0 & 5 & 0 & 25 \\ 6 & 3 & 0 & 12 \end{bmatrix}. \tag{1.80}$$

If we denote $h(x) = u$ & U_C = the coordinates of u in the basis C of W then, with the coordinates X_A of (1.79) and the matrix of (1.80) we get

$$u = h(x) = h(X_A^{\mathrm{T}} A^{\mathrm{T}}) = X_A^{\mathrm{T}} h(A^{\mathrm{T}}) = X_A^{\mathrm{T}} H_{A,C} C^{\mathrm{T}} = \tag{1.81}$$

$$= [\,3 \;\; -1 \;\; 4\,] \begin{bmatrix} 6 & 2 & 0 & 7 \\ 0 & 5 & 0 & 25 \\ 6 & 3 & 0 & 12 \end{bmatrix} C^{\mathrm{T}} = [42 \;\; 13 \;\; 0 \;\; 44]\, C^{\mathrm{T}} =$$

$$= U_C C^{\mathrm{T}} = 42\, c_1 + 13\, c_2 + 44\, c_4. \tag{1.82}$$

The image $y = f(x)$ in basis B can be similarly obtained :

$$y = f(x) = f(X_A^{\mathrm{T}} A^{\mathrm{T}}) = X_A^{\mathrm{T}} f(A^{\mathrm{T}}) = X_A^{\mathrm{T}} F_{A,B} B^{\mathrm{T}} =$$

$$= [\,3 \;\; -1 \;\; 4\,] \begin{bmatrix} 2 & -1 \\ 5 & 5 \\ 3 & 0 \end{bmatrix} B^{\mathrm{T}} = [13 \;\; -8]\, B^{\mathrm{T}} = 13\, b_1 - 8\, b_2. \tag{1.83}$$

Next, the image $u = g(y)$ in basis C is

$$u = g(y) = f(Y_B^{\mathrm{T}} B^{\mathrm{T}}) = Y_B^{\mathrm{T}} g(B^{\mathrm{T}}) = Y_B^{\mathrm{T}} G_{B,C} C^{\mathrm{T}} =$$

$$= [13 \;\; -8] \begin{bmatrix} 2 & 1 & 0 & 4 \\ -2 & 0 & 0 & 1 \end{bmatrix} C^{\mathrm{T}} = [42 \;\; 13 \;\; 0 \;\; 4]\, C^{\mathrm{T}} =$$

$$= 42\, c_1 + 13\, c_2 + 44\, c_4. \tag{1.84}$$

Therefore, the expressions in (1.82) and (1.84) coincide and the formulas in PROPOSITION 1.7 have been illustrated / verified on this example. □

PROPOSITION 1.8. *If*

$$\mathcal{Q} = \{u_1, u_2, \ldots, u_m\} \subset U \tag{1.95}$$

is a linearly independent / dependent family of vectors in U and $f(\mathcal{Q})$ is its image through the bijective (or only injective) morphism f then it also *independent / dependent.*

Proof. The image through f of the family in (1.95) is

$$f(\mathcal{A}) = \{f(u_1), f(u_2), \dots, f(u_m)\} = \{v_1, v_2, \dots, v_m\} \subset V. \tag{1.86}$$

Let us assume that the family in (1.95) is linearly independent. Hence

$$\sum_{i=1}^{m} \lambda_i a_i = \mathbf{0} \;\Rightarrow\; \lambda_1 = \lambda_2 = \dots = \lambda_m = 0. \tag{1.87}$$

If the linear morphism f is applied to the first equation in (1.87), that is to the zero linear combination, it follows by the extended linearity of f – see PROPOSITION 1.1 in this section – that

$$f\left(\sum_{i=1}^{m} \lambda_i u_i\right) = \sum_{i=1}^{m} \lambda_i f(u_i) = \sum_{i=1}^{m} \lambda_i v_i = \mathbf{0}. \tag{1.88}$$

Let us assume that the last equation in (1.98) also holds for some $\lambda_j \neq 0$, $(1 \leq j \leq m)$ this would imply that the vector $v_j = f(u_j) \in \mathcal{B} = f(\mathcal{A})$ would be linearly expressible in terms of the other vectors of $\mathcal{B}$:

$$v_j = \sum_{i \neq j} \mu_i v_i. \tag{1.89}$$

The injectivity of f (and so more its bijectivity) ensure that every $v_i, v_j \in \mathcal{B}$ have unique counter-images $u_i = f^{-1}(v_i)$, $u_j = f^{-1}(v_j) \in \mathcal{A}$, respectively. Thus, expression (1.99) + (1.96) imply

$$u_j = \sum_{i \neq j} \mu_i u_i \iff \sum_{i<j} \mu_i u_i - u_j + \sum_{i>j} \mu_i u_i = \mathbf{0}. \tag{1.90}$$

But (1.90) represents a linear dependence relation among the vectors of the family (since, in Eq. (1.90), $\mu_j = -1 \neq 0)$, what contradicts its assumed independence, formally characterized by (1.87). This proof can be conversely restated for showing that the independence / dependence of the family $\mathcal{B} \subset V$ implies the same relation for the family $\mathcal{A} \subset U$, $\mathcal{A} = f^{-1}(\mathcal{B})$. ■

As a matter of terminology, the following equivalent (synonim) terms are used for naming the types of linear morphisms we have just discussed :

f is injective :	f is a *monomorphism* ;
f is surjective :	f is an *epimorphism* ;
f is bijective :	f is an *isomorphism* .

Definition 1.6. **(The inverse of a morphism).** Let U & V be two vector spaces over the same field $\mathbf{K}$ and let $f : U \longrightarrow V$, $g : V \longrightarrow U$ be two linear

morphisms. The morphism g is the *inverse* of f iff

$$(\forall\, x \in U)\ f(x) = y \in V \Rightarrow\ g(y) = x. \tag{1.91}$$

The usual notation for the inverse of the morphism f is $g =_{\text{not}} f^{-1}$. Hence $f^{-1} : V \longrightarrow U$ and the definition of (1.101) becomes

$$(\forall\, x \in U)\ f(x) = y \in V \Rightarrow\ f^{-1}(y) = x. \tag{1.91'}$$

◊

THEOREM 1.1. *The following properties of linear morphisms and isomorphisms hold :*

① $\mathrm{Hom}(U, V)$ *has the natural structure of a vector space.*

② *If* f *is an isomorphism then* $g = f^{-1}$ *is an isomorphism, too.*

③ *If* f *is an isomorphism and* $g = f^{-1}$ *is its inverse then*

$$f^{-1} \circ f = \mathbf{1}_U\ \ \&\ \ f \circ f^{-1} = \mathbf{1}_V. \tag{1.92}$$

④ *If* U, V *are finite dimensional spaces over* $\mathbf{K}$ *then they are isomorphic* $\Longleftrightarrow$ $\dim U = \dim V$.

⑤ *If* $\dim V = n$ *then* V *is isomorphic to* $\mathbf{K}^n$.

⑥ *If* $f : U \longrightarrow V$, $g : V \longrightarrow U$ *are inverse isomorphisms, that is* $g = f^{-1}$, *and* if A, B *are two bases of the spaces* U, V *with* $\dim U = \dim V = n$ *then*

$$G_{B,A} = (F_{A,B})^{-1}. \tag{1.93}$$

Note. The rather long and technical proof of this Theorem, with its 6 points, is not given here. It can be found in the extended textbook in ⟦𝒜. 𝒞., 2014⟧, **§ 3.1** (pages 171 - 174).

Comments. A couple of remarks on this Theorem could be appropriate. *Two vector spaces that are isomorphic are - for algebraic purposes - the same,* as stated in [G. Strang, 1988 - page 200], even when they are practically different. *They match completely : linearly independent sets correspond to linearly independent sets, and a basis in one corresponds to a basis in the other. Their dimensions coincide.* As regards part ⑤ in the previous Theorem, the isomorphism of any *n*-dimensional space V to $\mathbf{K}^n$ or to the "standard" real space $\mathbb{R}^n$ implies that any definition or result, stated (and proved) in $\mathbb{R}^n$ can pe transferred to the space V, with appropriate formulations and notations.

Example 1.7. We gave in **§ 1.1** , as an example of a vector space, the set (space) of *polynomials of order* n with real coefficients, $\mathrm{POL}_n(\mathbb{R})$. We there remarked that such a polynomial is completely and uniquely determined by its $n+1$ coefficients $a_0, a_1, \ldots, a_n$. This entails the one-to-one correspondence

$$\begin{gathered} p = a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n \in \mathrm{POL}_n(\mathbb{R}) \\ \updownarrow \\ [a_0 \ a_1 \ a_2 \ \ldots \ a_n] \in \mathbb{R}^{n+1}. \end{gathered} \tag{1.94}$$

This is a typical case of isomorphism. The correspondence in (1.94) is clearly bijective, and it as also linear. The linear operations with polynomials were presented in **Example 1.9** of **§ 1.1**, Eqs. (1.31) & (1.32). A standard or canonical basis spanning the space $\mathrm{POL}_n(\mathbb{R})$ is offered by the $n+1$ elementary polynomials $\{1, t, t^2, \ldots, t^n\}$. □

The Matrix of a Linear Morphism after a Change of Bases

Let $f: U \longrightarrow V$ be a linear linear morphism. If $\dim U = m$, $\dim V = n$ and the two vector spaces are respectively spanned by the bases

$$A = [a_1 \ a_2 \ \ldots \ a_m] \quad \text{and} \quad B = [b_1 \ b_2 \ \ldots \ b_n],$$

the matrix $F_{A,B}$ in this pair is uniquely determined by the formula (1.9) in PROPOSITION 1.2 (page 125). We recall that formula (giving it a new number):

$$\boxed{f(A^{\mathrm{T}}) = F_{A,B} B^{\mathrm{T}}.} \tag{1.95}$$

If the bases A & B are changed for a pair of "new" bases $\bar{A}$ & $\bar{B}$, f's matrix in this pair of bases will naturally change. We met this situation in the cases of linear forms (LFs), bilinear forms (BLFs), in **Chapter 3**, **§§ 3.1** & **3.2**. The way the matrix $F_{A,B}$ changes when $(A, B) \to (\bar{A}, \bar{B})$ is presented in

PROPOSITION 1.12. (Changing the bases and matrix of an LM). *If*

$$f: U \longrightarrow V, \ \dim U = m \ \& \ \dim V = n,$$

is a linear morphism with its matrix $F_{A,B}$ *defined by Eq.* (1.95) *in the pair of bases* A, B *of the spaces* U & V *(respectively) and if*

$$A \to \overline{A} \;\; by \;\; \overline{A}^{\mathrm{T}} = S \cdot A^{\mathrm{T}} \;\; and \;\; B \to \overline{B} \;\; by \;\; \overline{B}^{\mathrm{T}} = T \cdot B^{\mathrm{T}} \tag{1.96}$$

then the coefficient matrix of f *in the new bases* $\overline{A}, \overline{B}$ *is given by*

$$\boxed{F_{\overline{A}, \overline{B}} = S \cdot F_{A,B} \cdot T^{-1}.} \tag{1.97}$$

Proof. As in the case of the BLFs, the proof of the formula (1.97) follows from the transformation equations of the two bases, that is (1.20) in **§ 1.1**, and from the property of extended linearity of morphism f applied (simultaneously) to several linear combinations of vectors, in this case the linear expressions of the "new" vectors of bases $\overline{A}, \overline{B}$ in terms of the vectors of initial bases A, B. Our "matrix notations" are very useful in presenting this proof.

Let us recall that the transformation formulas in (1.96) can be more explicitly written with the transformation matrices S, T written as stacks of rows. Taking into account the dimensions of the spaces U, V we have

$$S \in \mathcal{M}_m(\mathbf{K}) \text{ and } T \in \mathcal{M}_n(\mathbf{K}). \tag{1.98}$$

The matrices in (1.96) are nonsingular. We can write them as column vectors whose compunents are rows in $\mathbf{K}^m$ & $\mathbf{K}^n$, respectively :

$$S = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_m \end{bmatrix}, \quad T = \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_n \end{bmatrix}.$$

$$(1.96\text{-}1) \Rightarrow \overline{A}^{\mathrm{T}} = S \cdot A^{\mathrm{T}} \iff \begin{bmatrix} \overline{a}_1 \\ \overline{a}_2 \\ \vdots \\ \overline{a}_m \end{bmatrix} = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_m \end{bmatrix} \cdot A^{\mathrm{T}}, \tag{1.99}$$

$$(1.96\text{-}2) \Rightarrow \overline{B}^{\mathrm{T}} = T \cdot B^{\mathrm{T}} \iff \begin{bmatrix} \overline{b}_1 \\ \overline{b}_2 \\ \vdots \\ \overline{b}_n \end{bmatrix} = \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_n \end{bmatrix} \cdot B^{\mathrm{T}}. \tag{1.100}$$

For instance, the i- th vector of basis $\overline{A}$ is linearly expressed in basis A as

$$\bar{a}_i = S_i \cdot A^{\mathrm{T}} = \sum_{i=1}^{m} \sigma_{ij}\, a_j. \tag{1.101}$$

If the linear morphism f is applied to Eq. (1.99), Eq. (1.7) with $\mathcal{X}^{\mathrm{T}} \rightarrow A^{\mathrm{T}}$ and $\Lambda^{\mathrm{T}} \rightarrow S_i \quad \Rightarrow$

$$\Rightarrow f(\bar{a}_i) = f(S_i \cdot A^{\mathrm{T}}) = S_i \cdot f(A^{\mathrm{T}}). \tag{1.102}$$

The m equations of the form (1.102) can be written one under the other resulting in

$$f(\bar{A}^{\mathrm{T}}) = f(S \cdot A^{\mathrm{T}}) = S \cdot f(A^{\mathrm{T}}). \tag{1.103}$$

Next, we have to use expression (1.95) of $f(A^{\mathrm{T}})$:

$$f(\bar{A}^{\mathrm{T}}) \underset{(1.103)}{=} S \cdot f(A^{\mathrm{T}}) = S \cdot F_{A,B}\, B^{\mathrm{T}}. \tag{1.104}$$

But it follows from the second equation in (1.96) that

$$B = T^{-1}\bar{B} \underset{(1.104)}{\Rightarrow} f(\bar{A}^{\mathrm{T}}) = S \cdot F_{A,B}\, B^{\mathrm{T}} = (S \cdot F_{A,B}\, T^{-1})\, \bar{B}. \tag{1.105}$$

The formula connecting the new bases $(\bar{A}, \bar{B})$ with the matrix of f comes from (1.95) with bars on the bases :

$$f(\bar{A}^{\mathrm{T}}) = F_{\bar{A},\bar{B}}\, \bar{B}^{\mathrm{T}}. \tag{1.106}$$

Equations (1.104) & (1.105), under their matrix forms, offer the m images through f of the vectors of $\bar{A}$ in the basis $\bar{B}$. According to the uniqueness of the coordinates (of one or several vectors) in a basis,

$$(1.143)\ \&\ (1.144) \Rightarrow F_{\bar{A},\bar{B}} = S \cdot F_{A,B} \cdot T^{-1} : (1.197)$$

The proof is thus complete. ■

Example 1.7. Let us consider two vector spaces U, V (over the same field $\mathbb{R}$) with $\dim U = 3$, $\dim V = 4$, respectively spanned by their bases A, B –

$$A = [a_1\ a_2\ a_3] \ \ \&\ \ B = [b_1\ b_2\ b_3\ b_4]$$

– and a morphism $f: U \longrightarrow V$ with its matrix in bases (A, B)

$$F_{A,B} = \begin{bmatrix} 1 & 3 & 0 & -2 \\ 0 & 2 & 1 & 4 \\ -1 & 1 & 1 & -1 \end{bmatrix}. \tag{1.107}$$

These two bases are changed for $(\bar{A}, \bar{B})$ with the transformation matrices

$$S = \begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 5 \\ 3 & 2 & 0 \end{bmatrix}, \quad T = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 0 & 2 & 1 & -1 \\ 1 & -1 & -1 & 2 \\ -1 & 0 & 0 & -1 \end{bmatrix}. \tag{1.108}$$

It is required to write the matrix of f in the pair of the new bases $(\bar{A}, \bar{B})$. Next, it is required to find the image through f of the vector

$$x = 2\,a_1 - a_2 + 3\,a_3 \tag{1.109}$$

using both matrices $F_{A,B}$ & $F_{\bar{A},\bar{B}}$.

In order to apply the formula of matrix change (1.197) to the data in (1.107) - (1.108), the *inverse of the matrix* T is needed. The most convenient way to obtain it is the one based upon transformations (Gaussian elimination), presented in **§ 1.2**. We recall it, with T instead of A :

$$[\,T \mid I_4\,] \longrightarrow \ldots \longrightarrow [\,I_4 \mid T^{-1}\,]. \tag{1.110}$$

$$[\,T \mid I_4\,] = \left[\begin{array}{cccc|cccc} 2 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & -1 & 0 & 1 & 0 & 0 \\ 1 & -1 & -1 & 2 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & -1 & 0 & 0 & 0 & 1 \end{array}\right] \longrightarrow$$

$$\longrightarrow \left[\begin{array}{cccc|cccc} 1 & -1 & -1 & 2 & 0 & 0 & 1 & 0 \\ 0 & 2 & 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 3 & 3 & -4 & 1 & 0 & -2 & 0 \\ 0 & -1 & -1 & 1 & 0 & 0 & 1 & 1 \end{array}\right] \longrightarrow$$

$$\longrightarrow \left[\begin{array}{cccc|cccc} 1 & 1 & 1 & 0 & 0 & 0 & -1 & -2 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & -1 & -1 & 0 & 1 & 0 & 2 & 4 \\ 0 & -1 & -1 & 1 & 0 & 0 & 1 & 1 \end{array}\right] \longrightarrow$$

$$\longrightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 0 & -1 & -2 & -3 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & -1 & 0 & 1 & 1 & 3 & 5 \\ 0 & 0 & -1 & 1 & 1 & 0 & 2 & 2 \end{array}\right] \longrightarrow$$

$$\longrightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & -1 & -1 & -3 & -5 \\ 0 & 0 & 0 & 1 & 0 & -1 & -1 & -3 \end{array}\right] \Rightarrow$$

$$\Rightarrow\ T^{-1} = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ -1 & -1 & -3 & -5 \\ 0 & -1 & -1 & -3 \end{bmatrix}. \tag{1.111}$$

$$(1.97)\,,\ (1.107)\ \&\ (1.101)\ \Rightarrow\quad F_{\bar{A},\bar{B}} = S \cdot F_{A,B} \cdot T^{-1} =$$

$$= \begin{bmatrix} 3 & 2 & -1 \\ 1 & 1 & 5 \\ 3 & 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 3 & 0 & -2 \\ 0 & 2 & 1 & 4 \\ -1 & 1 & 1 & -1 \end{bmatrix} \cdot T^{-1} = \begin{bmatrix} 4 & 12 & 1 & 3 \\ -4 & 10 & 6 & -3 \\ 3 & 13 & 2 & 2 \end{bmatrix} \cdot T^{-1} =$$

$$\underset{(3.111)}{=} \begin{bmatrix} 4 & 12 & 1 & 3 \\ -4 & 10 & 6 & -3 \\ 3 & 13 & 2 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ -1 & -1 & -3 & -5 \\ -1 & 0 & -1 & -3 \end{bmatrix} = \begin{bmatrix} 0 & 11 & 10 & 6 \\ -7 & 4 & -9 & -19 \\ -1 & 11 & 8 & 3 \end{bmatrix}.$$

The image of the vector x in (1.109) follows from the formula (1.13) at page 126. With the matrix in (1.107) and the coordinates resulting from expression (1.147), that is $X_A^{\mathrm{T}} = [2\ \ -1\ \ 3]$, we obtain

$$f(x) = X_A^{\mathrm{T}} \cdot F_{A,B} \cdot B^{\mathrm{T}} = [2\ \ -1\ \ 3] \begin{bmatrix} 1 & 3 & 0 & -2 \\ 0 & 2 & 1 & 4 \\ -1 & 1 & 1 & -1 \end{bmatrix} B^{\mathrm{T}} =$$

$$= [-1\ \ 7\ \ 2\ \ -11]\, B^{\mathrm{T}} = -b_1 + 7b_2 + 2b_3 - 11b_4. \tag{1.112}$$

The coordinates of x in the transformed basis $\bar{A}$ of space U can be determined by formula (1.78) in **§ 2.1**, with $B \rightarrow \bar{A}\ \&\ T \rightarrow S$:

$$X_{\bar{A}} = S^{-\mathrm{T}} X_A = (S^{\mathrm{T}})^{-1} X_A. \tag{1.113}$$

The column vector of the "new" coordinates in (1.113) can be effectively obtained as the solution to a nonhomogeneous system of augmented matrix

$$\left[\begin{array}{ccc|c} 3 & 1 & 3 & 2 \\ 2 & 1 & 2 & -1 \\ -1 & 5 & 0 & 3 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 2 & 1 & 2 & -1 \\ -11 & 0 & -10 & 8 \end{array}\right] \sim$$

$$\sim \left[\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 2 & 1 & 2 & -1 \\ -11 & 0 & -10 & 8 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 1 & 0 & -7 \\ 0 & 0 & 1 & 41 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & -38 \\ 0 & 1 & 0 & -7 \\ 0 & 0 & 1 & 41 \end{array}\right].$$

Hence we get

$$X_{\bar{A}}^{\mathrm{T}} = [-38 \quad -7 \quad 41]^{\mathrm{T}} \Rightarrow x = -38\,\bar{a}_1 - 7\bar{a}_2 + 41\bar{a}_3. \tag{1.114}$$

The image of x, with the transformed bases, is

$$f(x) = X_{\bar{A}}^{\mathrm{T}} \cdot F_{\bar{A},\bar{B}} \cdot \bar{B}^{\mathrm{T}} = [-38 \quad -7 \quad 41] \cdot \begin{bmatrix} 0 & 11 & 10 & 6 \\ -7 & 4 & -9 & -19 \\ -1 & 11 & 8 & 3 \end{bmatrix} \cdot \bar{B}^{\mathrm{T}} =$$

$$= [\,8 \quad 5 \quad 11 \quad 28\,] \cdot \bar{B}^{\mathrm{T}}. \tag{1.115}$$

Checking the coordinates $Y_{\bar{B}} = [8 \quad 5 \quad 11 \quad 28]$ is possible if we replace the vectors of $\bar{B}^{\mathrm{T}} = [\bar{b}_1 \quad \bar{b}_2 \quad \bar{b}_3 \quad \bar{b}_4]$ by their expressions in the initial basis B of space V, in expression (1.115), using the transformation matrix T of (1.108):

$$\begin{aligned} f(x) = y &= 8\,\bar{b}_1 + 5\,\bar{b}_2 + 11\,\bar{b}_3 + 28\,\bar{b}_4 = \\ &= 8\,(2b_1 + b_2 + b_3) - \\ &\quad 5\,(\quad 2b_2 + b_3 - b_4) + \\ &+ 11\,(\,b_1 - b_2 - b_3 + 2b_4) + \\ &+ 28\,(-b_1 \qquad\qquad - b_4) = -b_1 + 7b_2 + 2\,b_3 - 11b_4. \end{aligned}$$

Hence, the linear expression of the image $f(x) = y$ in the initial basis B, (1.109), has been retrieved. □

* * * *

Before continuing with other definitions and results on the linear transformations, let us remark that the morphisms from $U = \mathbb{R}^m$ to $V = \mathbb{R}^n$ (or from $\mathbf{K}^m$ to $\mathbf{K}^n$) are usually given, mainly for applications, by the image

$$Y = f(X) = [\,f_1(X) \;\ldots\; f_n(X)]^{\mathrm{T}}, \tag{1.116}$$

where the components in the rightmost side of (1.116) are linear forms. The equation (1.18) at page 127 gave the definition of the matrix of a morphism $f: \mathbb{R}^m \longrightarrow \mathbb{R}^n$ in the pair of the standard bases of these spaces. Let us recall that formula, also valid for a possibly more general morphism of the form $f: \mathbf{K}^m \longrightarrow \mathbf{K}^n$:

$$f(E_m^{\mathrm{T}}) = F_{E_m,E_n}\, E_n^{\mathrm{T}}.$$

It is the (matrix of the) morphism that has to be brought to a simpler form – this will be the case with diagonalization of endomorphisms, to be presented in

Section 4.3 – the standard bases should be changed for other bases. We approached this issue in connection with the BLFs, in **§ 3.2** . If the standard bases are changed to other (more general) bases, $(E_m, E_n) \to (A, B)$, the respective transformation matrices are

$$A = S E_m \;\;\&\;\; B = T E_n \;\;\text{with}\;\; S = A^{\mathrm{T}} \;\;\&\;\; T = B^{\mathrm{T}}. \tag{1.118}$$

These transformation matrices can be taken to the previous formula (1.97) at page 147 giving the matrix of a morphism after a change of bases. The necessary replacings are

$$(A, B) \to (E_m, E_n),\; (\bar{A}, \bar{B}) \to (A, B),\; S \to A^{\mathrm{T}} \;\&\; T \to B^{\mathrm{T}}. \tag{1.119}$$

From (1.97) with (1.118) we get

$$\boxed{F_{A,B} = A^{\mathrm{T}} \cdot F_{E_m, E_n} \cdot B^{-\mathrm{T}}.} \tag{1.120}$$

Example 1.8. The morphism $f: \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ is given by

$$Y = f(X) = \begin{bmatrix} f_1(X) \\ f_2(X) \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 \\ -x_1 + 5x_2 + x_3 \end{bmatrix}. \tag{1.121}$$

It is required to write the matrix of this morphism in the pair of bases

$$(A, B) \in \mathbb{R}^3 \times \mathbb{R}^2,\; A = [a_1 \;\; a_2 \;\; a_3] \;\&\; B = [b_1 \;\; b_2]$$

with

$$a_1 = \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix},\; a_2 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix},\; a_3 = \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix};\; b_1 = \begin{bmatrix} 3 \\ 5 \end{bmatrix},\; b_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}. \tag{1.122}$$

It is also required to find the image of $X = [7 \;\; 1 \;\; -1]^{\mathrm{T}}$ as a (column) vector in $\mathbb{R}^2$ and by its linear expression in basis B of (1.122).

$$(1.121) \Rightarrow \quad M = \begin{bmatrix} 2 & 3 & 0 \\ -1 & 5 & 1 \end{bmatrix} \Rightarrow F_{E_3, E_2} = M^{\mathrm{T}} = \begin{bmatrix} 2 & -1 \\ 3 & 5 \\ 0 & 1 \end{bmatrix}. \tag{1.123}$$

$$(1.122) \Rightarrow \quad A^{\mathrm{T}} = \begin{bmatrix} 2 & -1 & 4 \\ 3 & 0 & 1 \\ 1 & 5 & 5 \end{bmatrix},\; B^{\mathrm{T}} = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}. \tag{1.124}$$

In order to use formula (1.120), the matrix B^{T} has to be inverted. Since $\det B = 1$, it is convenient to apply the formula $B^{-1} = (1/\det B)\,B^{\star}$.

(1.122-1) , (1.123) , (1.122-2) $\Rightarrow$

$$\Rightarrow F_{A,B} = \begin{bmatrix} 2 & -1 & 4 \\ 3 & 0 & 1 \\ 1 & 5 & 5 \end{bmatrix} \cdot \begin{bmatrix} 2 & -1 \\ 3 & 5 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -3 \\ 6 & -2 \\ 17 & 29 \end{bmatrix} \cdot \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} =$$

$$= \begin{bmatrix} 5 & -14 \\ 14 & -36 \\ 5 & 2 \end{bmatrix}. \tag{1.125}$$

From (1.123) we get the image of $X = [7 \;\; 1 \;\; -1]^{\mathrm{T}}$:

$$Y = f(X) = M \cdot X = \begin{bmatrix} 2 & 3 & 0 \\ -1 & 5 & 1 \end{bmatrix} \cdot \begin{bmatrix} 7 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 17 \\ -3 \end{bmatrix}. \tag{1.126}$$

In order to find this image using the matrix of (1.125), the coordinates X_A of X in basis A have to be found. They are easily get from a nonhomogeneous system (under its matrix form), as presented in **§ 1.1** and **§ 1.2** :

$$[A\,|\,X] = \left[\begin{array}{ccc|c} 2 & 3 & 1 & 7 \\ -1 & 0 & 5 & 1 \\ 4 & 1 & 5 & -1 \end{array}\right] \sim \left[\begin{array}{ccc|c} -1 & 0 & 5 & 1 \\ 4 & 1 & 5 & -1 \\ -10 & 0 & -14 & 10 \end{array}\right] \sim$$

$$\sim \left[\begin{array}{ccc|c} 1 & 0 & -5 & -1 \\ 0 & 1 & 25 & 3 \\ 0 & 0 & -64 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 0 \end{array}\right] \Rightarrow X_A = \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix}. \tag{1.127}$$

The image $Y = f(X)$, with bases (A, B) and formula (1.13) at page 126, is obtained, with the matrix in (1.125), as

$$f(X) = X_A^{\mathrm{T}} F_{A,B} B^{\mathrm{T}} = [-1 \;\; 3 \;\; 0] \begin{bmatrix} 5 & -14 \\ 14 & -36 \\ 5 & 2 \end{bmatrix} B^{\mathrm{T}} =$$

$$= [37 \;\; -94]\, B^{\mathrm{T}} = 37 b_1 - 94 b_2. \tag{1.128}$$

This expression in (1.128) can be checked by replacing b_1 & b_2 from (1.122) :

$$37\,b_1 - 94\,b_2 = 37\begin{bmatrix}3\\5\end{bmatrix} - 94\begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}17\\-3\end{bmatrix}.$$

Thus, the value of (1.128) is retrieved and the example is complete. □

Example 1.9. A mapping $\mathbf{diag} : \mathscr{M}_n(\mathbb{R}) \longrightarrow \mathbb{R}^n$ is defined by

$$(\forall\, A = [a_{ij}]_{n,n} \in \mathscr{M}_n(\mathbb{R}))\ \ \mathbf{diag}(A) = [a_{11}\ a_{22}\ \dots\ a_{nn}]. \tag{1.129}$$

Check that this is a linear morphism but it is not injective. Is it surjective ?

We saw in **§ 2.1** that the set of matrices of any size $m \times n$ over an arbitrary field $\mathbf{F}$ form a vector space over that field. This property clearly holds for the particular case of the square matrices, too. The mapping $\mathbf{diag}$ applied to a linear combination of two matrices $\lambda A + \mu B = [\lambda a_{ij} + \mu b_{ij}]_{n,n}$ gives

$$\begin{aligned}
\mathbf{diag}(\lambda A + \mu B) &= \mathbf{diag}([\lambda a_{ij} + \mu b_{ij}]_{n,n}) = \\
&= [\lambda a_{11} + \mu b_{11}\ \ \lambda a_{22} + \mu b_{22}\ \dots\ \lambda a_{nn} + \mu b_{nn}] = \\
&= \lambda [a_{11}\ a_{22}\ \dots\ a_{nn}] + \mu [b_{11}\ b_{22}\ \dots\ b_{nn}] = \\
&= \lambda\,\mathbf{diag}(A) + \mu\,\mathbf{diag}(B) \Rightarrow (\mathrm{LIN})\ \text{ for }\ \mathbf{diag}.
\end{aligned} \tag{1.130}$$

The mapping is obviously surjective. For any $X = [x_1\ x_2\ \dots\ x_n] \in \mathbb{R}^n$ there exists at least one square matrix of order n whose image through $\mathbf{diag}$ is just this X; for instance, the simplest matrix $A : \mathbf{diag}(A) = X$ is just the diagonal matrix $A_X = \lceil x_1\ x_2\ \dots\ x_n \rfloor$.

The mapping is not injective since two distinct matrices

$$A, B \in \mathscr{M}_n(\mathbb{R}) \ \text{ with }\ \mathbf{diag}(A) = \mathbf{diag}(B),$$

having just the same entries on their main diagonals, may have different entries in their lower and upper triangles : $(\exists\, i, j \in \overline{1,n})\ \ i \neq j\ \ \&\ \ a_{ij} \neq b_{ij}$. □

The next result regards the *composite morphisms*. Therefore, the definition of $h = g \circ f$ (*Def 1.5* at page 140, Eq. (1.72)) is going to be involved and we recall it (under a new number) :

$$g \circ f : U \longrightarrow W\ \ (\forall\, x \in U)\ \ (g \circ f)(x) \underset{\mathrm{def}}{=} g[f(x)]. \tag{1.131}$$

THEOREM 1.2. (Properties of composite morphisms) *Let the composite of the two linear morphisms* $f : U \longrightarrow V$ & $g : V \longrightarrow W$ *be* $h = g \circ f$ *defined by* (1.131). *Then* :

$(i)\quad h = g \circ f \;\; surjective\; (epimorphism) \quad \Rightarrow \;\; g \;\; surjective\,;$

$(ii)\quad h = g \circ f \;\; injective \;\; (monomorphism) \Rightarrow \;\; f \;\; injective\,;$

$(iii)\quad h = g \circ f \;\; bijective \;\; (isomorphism) \quad \Rightarrow \;\; f \;\&\; g \;\; bijective\,;$

$(iv)\quad W = U \;\&\; h = g \circ f = \mathbf{1}_U \;\; \Rightarrow \;\; f \;\&\; g \;\; bijective\; and \;\; g = f^{-1}.$

Proofs. These four assertions offer properties of the "factor" morphisms $f \;\&\; g$ resulting from a specific property of the composite morphism $h = g \circ f$. Obviously, the possible properties of a linear morphism have to be reviewed from *Def. 1.4* at page 139, Eqs. (1.62) - (1.63). Before starting the proof, we offer a (possibly useful) diagram representing the three mappings (morphisms).

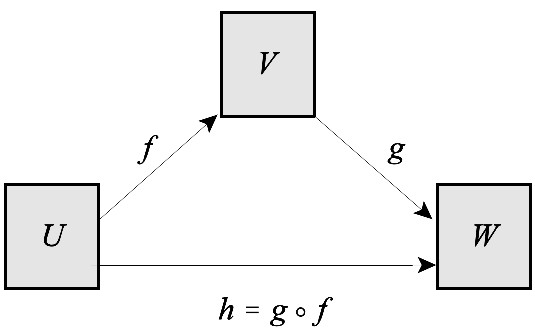

Fig. 1.1 The composite of two morphisms

(i) If h is surjective, this means (by *Def. 1.4*, Eq. (1.73)) that

$$\operatorname{Im} h = h(U) = W \iff (\forall\; w \in W)\;\; h_{-1}(w) \neq \varnothing. \qquad (1.132)$$

Since $h = g \circ f$, (1.132) means that

$$(\forall\; w \in W)\,(\exists\; u \in U)\;\; h(u) \underset{\text{def}}{=} g[f(u)] = w \;\; \Rightarrow$$

$$\Rightarrow \;\; (\forall\; w \in W)\,(\exists\; v \in V)\;\; v = f(u) \;\&\; g(v) = w \;\; \Rightarrow \;\; g \text{ is surjective.}$$

(ii) If h is injective, it follows by *Def. 1.4*, Eq. (1.72) that

$$(\forall\; x_1, x_2 \in U)\;\; x_1 \neq x_2 \;\; \Rightarrow \;\; h(x_1) \neq h(x_2), \text{ or} \qquad (1.133)$$

$$(\forall\; x_1, x_2 \in U)\;\; h(x_1) = h(x_2) \;\; \Rightarrow \;\; x_1 = x_2. \qquad (1.134)$$

We can use these equivalent definitions, but is seems more convenient to use the characterization in terms of the kernel(s) - PROPOSITION 1.6 at page 139, Eq. (1.75). Let us denote the three zero vectors of the spaces U, V, W as

$$\mathbf{0} \in U,\ \mathbf{0}' \in V,\ \mathbf{0}'' \in W.$$

We must prove that $\operatorname{Ker} f = \{\mathbf{0}\}$ if $\operatorname{Ker} h = \{\mathbf{0}\}$.

$$\operatorname{Ker} h = \{\mathbf{0}\} \Rightarrow \{u \in U:\ h(u) = \mathbf{0}'' \in W\} = \{\mathbf{0}\} \Rightarrow$$

$$\Rightarrow \{u \in U:\ g[f(u)] = \mathbf{0}'' \in W\} = \{\mathbf{0}\} \Rightarrow$$

$$\Rightarrow \{u \in U:\ f(u) = v\ \&\ g(v) = \mathbf{0}'' \in W\} = \{\mathbf{0}\} \Rightarrow \tag{1.135}$$

$$\Rightarrow \{u \in U:\ f(u) = \mathbf{0}'\} = \operatorname{Ker} f = \{\mathbf{0}\}. \tag{1.136}$$

Let us assume that the implication from (1.135) to (1.136) would not hold. This would mean that

$$(\exists\ u \in U)\ u \neq \mathbf{0}\ \&\ f(u) = \mathbf{0}'. \tag{1.137}$$

But any morphism maps the zero vector onto the zero vector (of the "next" space), hence (1.137) $\Rightarrow$

$$\Rightarrow (\exists\ u \in U)\ u \neq \mathbf{0}\ \&\ f(u) = \mathbf{0}' \Rightarrow g[f(u)] = g(\mathbf{0}') = \mathbf{0}'' \Rightarrow$$

$$\Rightarrow (\exists\ u \in U)\ u \neq \mathbf{0}\ \&\ h(u) = \mathbf{0}'',$$

what contradicts the injectivity of h ; thus (1.136) holds.

Let us also see the proof of (ii) by using (1.134). From h injective it follows that

$$h(x_1) = h(x_2) \Rightarrow x_1 = x_2. \tag{1.138}$$

If we take $y_1 = y_2 \Rightarrow f(x_1) = f(x_2)$ but suppose that $x_1 \neq x_2$ then, by the other definition of injectivity for h, (1.62), it would follow that

$$h(x_1) \neq h(x_2)$$

what contradicts the hypothesis in (1.138).

(iii) If $h = g \circ f$ is bijective, it is both surjective and injective. In view of (i) & (ii) it follows that g is surjective and f is injective. Let us show that g is injective, too. If $\operatorname{Ker} g \neq \{\mathbf{0}'\}$ then

$$(\exists\ v \in V)\ v \neq \mathbf{0}'\ \&\ g(v) = \mathbf{0}'' \Rightarrow g[f(u)] = \mathbf{0}'' \tag{1.139}$$

for some $u \in U$. But (1.139) $\Rightarrow$ $h(u) = \mathbf{0}'' \Rightarrow u = \mathbf{0}$ since h is bijective. Hence $v = f(u)$ of (1.175) cannot be $\neq \mathbf{0}'$ and $\operatorname{Ker} g = \{\mathbf{0}'\} \Rightarrow g$ is injective. It remains to verify that and f is surjective. Let us assume that

$$f(U) \subset V \Rightarrow (\exists\, v \in V \setminus f(U)(\nexists\, u \in U)\ v = f(u) \Rightarrow$$
$$\Rightarrow (\exists\, w \in W)(\nexists\, u \in U)\ w = h(u)$$

since the two morphisms f & g cannot be composed through this v. Thus, v could not be surjective ! Thus both morphisms f & g are surjective and injective, hence bijective.

In the terminology at page 144,

(i) if $h = g \circ f$ is a monomorphism then f is a monomorphism ;

(ii) if $h = g \circ f$ is an epimorphism then g is an epimorphism ;

(iii) if $h = g \circ f$ is an isomorphism then both f & g are isomorphisms.

(iv) This property follows from the pervious one. If

$$W = U \;\&\; h = g \circ f = \mathbf{1}_U$$

then f & g are bijective since the identical morphism on any space is bijective (an isomorphism) – the most trivial isomorphism. Hence each of them has an inverse. It remains to show that $g = f^{-1}$. Let us denote by γ <u>another</u> possible inverse of f. By *Definition 1.6* at page 170, Eqs. (1.101) - (1.101'), and also by THEOREM 1.1 - ③ (Eq. (1.102) ,

$$\gamma \circ f = \mathbf{1}_U \Rightarrow (\forall\, x \in U)\ \gamma[f(x)] = x \in U. \tag{1.176}$$

If we consider an arbitrary vector (or point) $x \in U$ and its image $y = f(x)$, from the equation $h = g \circ f = \mathbf{1}_U$ in the statement plus (1.176) we have

$$\begin{cases} g(y) = g[f(x)] = (g \circ f)(x) = \mathbf{1}_U(x) = x, \\ \gamma(y) = \gamma[f(x)] = (\gamma \circ f)(x) = \mathbf{1}_U(x) = x \end{cases} \Rightarrow \gamma \equiv g = f^{-1}. \tag{1.177}$$

Therefore, two morphisms whose composite - or product - is the identity morphism $(h = g \circ f = \mathbf{1}_U)$ are both isomorphisms and thus invertible, the inverse of f is $g = f^{-1}$ and it is unique. ■

Comments. **1°** In the proofs of all the four points (i) thru (iv) of this THEOREM, nowhere was used the property of f & g to be *linear morphism*s between (pairs among) the three vector spaces. Thus, the four properties would hold for more general maps from a set to another set. However, the zero vectors and the kernels have been involved in some of the proofs, and we stated and proved this four fold result dedicated to linear morphisms.

2° The second chapter entitled LINEAR MAPS, in the monograph **Functional**

Analysis [P. Lax, 2002], includes two additional properties that can complete the points (i) thru (iv) in our **THEOREM 1.2**. We present them under the next numbers, with slight changes in Professor **Peter D. Lax**'s notations [***Functional Analysis***, John Wiley & Sons, Inc., 2002] at page 9.

(v) If the morphisms $f : U \longrightarrow V$ & $g : V \longrightarrow W$ are both invertible, so is their product, $h = g \circ f$ and

$$h^{-1} = f^{-1} \circ g^{-1}. \tag{1.178}$$

(vi) If $h = g \circ f$ is invertible, then

$$\operatorname{Ker} f = \{\mathbf{0}\} \ \ \& \ \ \operatorname{Im} g = W. \tag{1.179}$$

Let us see that these two equations were stated (and proved) in our **TH. 1.2**, (i) & (ii), taking into account the characterizations of injective and bijective linear morphisms, earlier presented in our PROPOSITION 1.6 - ① & ② at page 144. We defined the surjective morphisms in *Def. 1.4* at page 139.

3° The author **P. Lax** adds the following ***Remark*** : *When* $\square = V = W$ *are finite dimensional, then the invertibility of the product* $g\,f$ ($h = g \circ f$ in our notations) *implies that* g *and* f *separately are invertible. This is not so in the infinite-dimensional case ; take, for instance, the space of infinite sequences*

$$x = (a_1, a_2, \dots)$$

and define R *and* L *to be the right and left shifts* :

$$Rx = (0, a_1, a_2, \dots), \quad Lx = (a_2, a_3, \dots).$$

Clearly LR *is the identity map, but neither* R *nor* L *are invertible ; nor is the* RL *identity.*

4° Part (iv) in **THEOREM 1.2**. is the reciprocal (converse implication) to the immediate property stated as part ③ in **THEOREM 1.1** at page 145.

4-A APPLICATIONS TO LINEAR MORPHISMS

§ 4.1-A APPLICATIONS TO LINEAR TRANSFORMATIONS

1 – A.1 Show that the following mappings are linear morphisms and write their matrices in the standard bases of the spaces they are defined on.

(i) $\quad f:\mathbb{R}^3 \longrightarrow \mathbb{R}^2,\ f([x_1\ \ x_2\ \ x_3]^{\mathrm{T}}) = [3x_1 + x_2\ \ \ -4x_3]^{\mathrm{T}};$

(ii) $\quad f:\mathbb{R}^2 \longrightarrow \mathbb{R}^3,\ f([x_1\ \ x_2]^{\mathrm{T}} = [x_1\ \ 3x_1 - x_2\ \ -2x_1]^{\mathrm{T}};$

(iii) $\quad f:\mathbb{R}^2 \longrightarrow \mathbb{R}^2,\ f([x_1\ \ x_2]^{\mathrm{T}} = [0\ \ 0]^{\mathrm{T}};$

(iv) $\quad f:\mathrm{POL}_n(\mathbb{R}) \longrightarrow \mathrm{POL}_{n-1}(\mathbb{R}),$

$$f(a_0 + a_1 t + \ldots a_n t^n) = a_1 + 2a_2 t + \ldots n a_n t^{n-1}.$$

1 – A.2 Show that the following mappings are linear morphisms, find their kernels and images and establish which of them are isomorphisms.

$a)$ $\quad f:\mathbb{R}^2 \longrightarrow \mathbb{R}^2,\ f([x_1\ \ x_2]^{\mathrm{T}} = [x_1 - 2x_2\ \ x_2 - x_1]^{\mathrm{T}};$

$b)$ $\quad f:\mathbb{R}^3 \longrightarrow \mathrm{POL}_2(\mathbb{R}),$

$$f([x_1\ \ x_2\ \ x]^{\mathrm{T}} = x_1 + x_3 + 2x_2 t - (x_1 + x_2)t^2;$$

$c)$ $\quad f:\mathrm{POL}_1(\mathbb{R}) \longrightarrow \mathbb{R}^2,\ f(a_1 + a_2 t) = [2a_0 + a_1\ \ 3a_0 + a_1]^{\mathrm{T}};$

$d)$ $\quad f:\mathrm{POL}_1(\mathbb{R}) \longrightarrow \mathrm{POL}_2(\mathbb{R}),\ f(a_1 + a_2 t) = a_0 + 3a_0 t\ \ a_1 t^2;$

$e)$ $\quad f:\mathbb{R}^3 \longrightarrow \mathbb{R}^3,$

$$f([x_1\ \ x_2\ \ x_3]^{\mathrm{T}} = [x_1 - 2x_3\ \ 2x_1 + x_2\ \ x_2 + 3x_3]^{\mathrm{T}}.$$

1 – A.3 Determine the composite morphisms $g \circ f$ and (or) $f \circ g$ and check whether they are linear morphisms, where:

$$a)\quad \begin{cases} f:\mathbb{R}^3 \longrightarrow \mathbb{R}^2,\ f([x_1\ x_2\ x_3]^{\mathrm{T}}) = [x_1 - x_2 + x_3\ \ \ x_1 + 3x_3]^{\mathrm{T}} \\ g:\mathbb{R}^2 \longrightarrow \mathbb{R}^3,\ g([y_1\ \ y_2\ \ y_3]^{\mathrm{T}}) = [y_1\ \ y_1 - y_2\ \ y_2]^{\mathrm{T}}; \end{cases}$$

$$b)\quad \begin{cases} f:\mathbb{R}^2 \longrightarrow \mathbb{R}^4,\ f([x_1\ x_2]^{\mathrm{T}}) = [x_1\ \ 3x_2 + x_1\ \ 2x_1 - x_2\ \ x_2]^{\mathrm{T}} \\ g:\mathbb{R}^3 \longrightarrow \mathbb{R}^2,\ g([y_1\ \ y_2\ \ y_3]^{\mathrm{T}}) = [2y_1\ \ y_2]^{\mathrm{T}}, \end{cases}$$

Write the matrices of $f \circ g$ in the respective standard bases and calculate the matrices of the composite morphisms ; check the formula (1.77) in PROPOSITION 1.7 - page 142 for the matrices of the composite morphisms.

1 – A.4 Given the morphism

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3,\ f([x_1\ \ x_2]^{\mathrm{T}} = [3x_1 - 2x_2\ \ 2x_1 - x_2\ \ -x_1 + x_2]^{\mathrm{T}}$$

and the (linearly independent) vectors

$$U_1 = [1\ \ -2]^{\mathrm{T}}\ \ \&\ \ U_2 = [-1\ \ 1]^{\mathrm{T}},$$

check for independence / dependence their images $f(U_1)\ \&\ f(U_2)$. Find the counter-images (vectors or sets of vectors)

$$f_{-1}([1\ \ 2\ \ -1]^{\mathrm{T}})\ \ \&\ \ f_{-1}([10\ \ 0\ \ 0]^{\mathrm{T}}).$$

Can the subscript -1 be raised as a superscript ?

1 – A.5 The morphism $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ is given by its matrix in the pair of standard bases

$$F_{E_3,E_2} = \begin{bmatrix} -1 & 2 \\ 1 & 1 \\ 2 & 2 \end{bmatrix}.$$

Find its matrix $F_{A,B}$ in the pair of bases

$$A : a_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix},\ a_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix},\ a_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix};\quad B : b_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ b_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

Then find (a basis) spanning $\mathrm{Ker}\, f$ and $f_{-1}([2\ \ 4]^{\mathrm{T}})$.

1 – A.6 Check whether some linear morphism $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ can map the vectors

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix},\ \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ \begin{bmatrix} 1 \\ 1 \end{bmatrix}\quad \text{onto}\quad \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

(respectively).

3 – A.7 Find a basis spanning the subspace W of the solutions to the homogeneous system whose matrix is

$$A = \begin{bmatrix} 3 & 4 & 2 \\ 2 & 3 & 1 \\ 1 & 2 & 0 \end{bmatrix}.$$

Then find (a basis spanning) the image $f(W)$ through the morphism $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} -x_1 + 2x_2 + x_3 \\ 3x_2 - x_3 \end{bmatrix}.$$

1 – A.8 Let $f : U \longrightarrow V$ & $g : V \longrightarrow W$ be linear morphisms. Prove that

$$\operatorname{Ker} f \subseteq \operatorname{Ker}(g \circ f), \quad \operatorname{Im} g \supseteq \operatorname{Im}(g \circ f).$$

1 – A.9 Let $f : U \longrightarrow V$ & $g : V \longrightarrow W$ be two linear morphisms with the property $g \circ f = O$ (the zero morphism). Show that

(i) If f is surjective then $g = O$;

(ii) if g is injective then $f = O$;

(iii) $\operatorname{Im} f \subseteq \operatorname{Ker} g \iff g \circ f = O$.

Note **:** The same symbol O is used for the zero morphism in this statement although three different zero morphisms are here involved. Property (iii) is not directly connected with (i) & (ii) since $g \circ f = O$ should be checked to be a necessary and sufficient condition for the inclusion.

1 – A.10 Let $A = [a_1 \; a_2 \; a_3 \; a_4]$ be a basis in $\mathbb{R}^4$ and $B = [b_1 \; b_2]$ – a basis in $\mathbb{R}^2$. The linear morphism $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ is defined by

$$f(a_1) = b_1, \; f(a_2) = b_2, \; f(a_3) = b_1 + b_2, \; f(a_4) = b_1 - b_2.$$

(i) Write the matrix of this morphism in the pair of bases (A, B);

(ii) write the image of a vector $X = \sum_{i=1}^{4} \xi_i a_i$ as a linear expression in basis B;

(iii) find the counter-image $f_{-1}(-b_1 + 4b_2)$;

(iv) determine $\operatorname{Ker} f$, $\operatorname{Im} f$, $\operatorname{null} f$, $\operatorname{rank} f$.

1 – A.11 It is considered the mapping $T : \mathfrak{F}_{[-1,1]} \longrightarrow \mathfrak{F}_{[0,2\pi]}$ defined by

$$T(f)(x) = f(\sin x).$$

Show that T is a linear morphism and check whether it is an isomorphism.

1 – A.12 There are considered two linear morphisms,

$$f : U \longrightarrow V \ \& \ g : V \longrightarrow U$$

with the two spaces respectively spanned by the bases $A = [a_1 \ a_2 \ a_3]$ and $B = [b_1 \ b_2]$. The two morphisms are differently defined, as follows :

$$F_{A,B} = \begin{bmatrix} -2 & 3 \\ 3 & -2 \\ -1 & 5 \end{bmatrix}, \quad \begin{cases} g(b_1) = \ \ 3a_1 + 2a_2 - a_3, \\ g(b_2) = -2a_1 - 2a_2 + a_3. \end{cases}$$

It is required to :

1° Find the expression of $f(-2a_1 + 3a_2 - a_3)$ in the basis B ;

2° show that f is surjective ; **3°** show that g is injective.

4° Write the matrix H_A of the composite morphism

$$h = g \circ f : U \longrightarrow U.$$