Chapter 2

VECTOR SPACES AND SUBSPACES

§ 2.1 VECTOR SPACES, GENERATORS, BASES AND COORDINATES

The notion of ***vector space*** (synonim : ***linear space***) represents the fundamental mathematical structure for the **LINEAR ALGEBRA**. Its formal definition needs the reader to be acquainted with such algebraic structures as *groups* and *fields* (which are studied in the highschool). But, unlike the latter ones which are defined by means of internal operations (laws of composition), the notion of vector space involves both an internal operation – the vector *sum* (or vector *addition*) – and an external one – the *multiplication by scalars*.

The axiomatic definition of a vector space may be given in a condensed version (using the notion of group for the additive operation), but we start with the full definition (consisting of ten axioms):

Definition 1.1. Let V be a nonempty set (of *vectors*) and $\mathbb{F}$ a field (of *scalars*). Then V is a *vector* (or *linear*) *space over* $\mathbb{F}$ if the following properties (axioms) are satisfied :

$(\mathbf{L}_1)$ $(\forall\, x, y \in V)$ $\quad x + y \in V$;

$(\mathbf{L}_2)$ $(\forall\, x, y, z \in V)$ $\quad x + (y + z) = (x + y) + x$;

$(\mathbf{L}_3)$ $(\exists\, \mathbf{0} \in V)(\forall x \in V)$ $\quad x + \mathbf{0} = \mathbf{0} + x = x$;

$(\mathbf{L}_4)$ $(\forall\, x \in V)(\exists\, -x \in V)$ $\quad x + (-x) = (-x) + x = \mathbf{0}$;

$(\mathbf{L}_5)$ $(\forall\, x, y \in V)$ $\quad x + y = y + x$;

$(\mathbf{L}_6)$ $(\forall\, \lambda \in \mathbf{F})(\forall x \in V)$ $\quad \lambda x \in V$;

$(\mathbf{L}_7)$ $(\forall\, x \in V)$ $\quad 1x = x \quad (1 \in \mathbf{F})$;

$(\mathbf{L}_8)$ $(\forall \lambda, \mu \in \mathbf{F})(\forall\, x \in V)$ $\quad \lambda(\mu x) = (\lambda \mu) x$;

$(\mathbf{L}_9)$ $(\forall\, \lambda \in \mathbf{F})(\forall\, x, y \in V)$ $\quad \lambda(x + y) = \lambda x + \lambda y$;

$(\mathbf{L}_{10})$ $(\forall\, \lambda, \mu \in \mathbf{F})(\forall\, x \in V)$ $\quad (\lambda + \mu) x = \lambda x + \mu x$. ◊

Remarks 1.1. It follows from axioms $(\mathbf{L}_1, \ldots, \mathbf{L}_5)$ that $\langle V; +\rangle$ is an *Abelian group*. We recall that $(\mathbf{L}_1)$ means the closure of V under $+$, $(\mathbf{L}_2)$ is the *associativity* of the sum, $(\mathbf{L}_3)$ states the existence of the *zero* vector, $(\mathbf{L}_4)$ defines the *negative* of a vector x and $(\mathbf{L}_5)$ states the *commutativity* of the vector sum. The addition sign $+$ in $(\mathbf{L}_{10})$ is used with two different meanings: it stands for the scalar sum of the field $\mathbb{F}$ in the left hand side, while it denotes the sum of vectors in V in the right hand side, respectively. In the r.h.s. of $(\mathbf{L}_8)$ $\lambda\,\mu$ is the (commutative) multiplication of the scalars in the field $\mathbb{F}$, also denoted as $\mathbb{K}$.

Remark 1.2. (*Notational conventions*). As already used in *Def. 1.1*, the vectors in V are denoted by (italic) Latin letters $(x, y, u, v, \ldots)$, sometimes subscripted; the scalars in $\mathbb{F}$ are denoted by Greek letters $(\lambda, \mu, \alpha, \beta, \ldots)$, also subscripted when necessary. This latter convention will be eventually ignored for convenience. It is important to distinguish between the zero vector $\mathbf{0}$ and the zero scalar in $\mathbb{F}$: the latter one will be denoted by 0 or o. For this structure of vector space V over a field $\mathbb{F}$ we shall use the notation

$$\langle V, \mathbf{F}; +, \lambda V \rangle. \tag{1.1}$$

Remark 1.3. Not all the properties $(\mathbf{L_1})$ thru $(\mathbf{L_{10}})$ are independent or – in other words – the system of axioms $(\mathbf{L_1}, \ldots, \mathbf{L_{10}})$ *is not minimal*. For instance, the commutativity $(\mathbf{L_5})$ of the sum in V follows from other properties. Indeed, let us express $(1+1)(x+y)$ in two different ways :

$$(1+1)(x+y) = (1+1)x + (1+1)y = 1x + 1x + 1y + 1y = x + x + y + y\,; \tag{1.2}$$

$$(1+1)(x+y) = 1(x+y) + 1(x+y) = 1x + 1y + 1x + 1y = x + y + x + y. \tag{1.3}$$

Axioms $(\mathbf{L_9})$, $(\mathbf{L_{10}})$ and $(\mathbf{L_7})$ have been used in equations of (1.2), and $(\mathbf{L_{10}})$, $(\mathbf{L_9})$ and $(\mathbf{L_7})$ in (1.3), respectively (in the specified orders). It follows from (1.2) and (1.3) that

$$x + x + y + y = x + y + x + y. \tag{1.4}$$

By adding $-x$ to the left of both sides of Eq.(1.4) and $-y$ to the right of them, we derive – in view of $(\mathbf{L_2})$, $(\mathbf{L_4})$ and $(\mathbf{L_3})$ – that $x + y = y + x$, hence the addition in V is commutative. However, we keep this axiom for letting remain together the axioms of the Abelian group.

Remark 1.4. The field $\mathbb{F}$ underlying the vector space V may be, for instance, a numerical field like $\mathbb{R}$ (the real field) or $\mathbb{C}$ (the complex field). For $\mathbb{F} = \mathbb{R}$, V is said to be a *real* vector space.

Definition 1.1 of a vector (or linear) space has a couple of immediate consequences, stated together in

PROPOSITION 1.1. *The following properties of a linear space V hold* :

$(\mathbf{L}_{11})$ $(\forall \lambda \in \mathbf{F})(\forall x, y \in V)$ $\quad \lambda(x - y) = \lambda x - \lambda y\,;$

$(\mathbf{L}_{12})$ $(\forall \lambda, \mu \in \mathbf{F})(\forall x \in V)$ $\quad (\lambda - \mu)x = \lambda x - \mu x\,;$

$(\mathbf{L}_{13})$ $(\forall x \in V)$ $\quad 0x = \mathbf{0}$ $\;(0 \in \mathbf{F}$ and $\mathbf{0} \in V)\,;$

$(\mathbf{L}_{14})$ $(\forall \lambda \in \mathbf{F})$ $\quad \lambda \mathbf{0} = \mathbf{0}\,;$

($\mathbf{L}_{15}$) if $\lambda x = \mathbf{0}$, then $\lambda = 0$ or $x = \mathbf{0}$;

($\mathbf{L}_{16}$) $(\forall\, x \in V)$ $(-1)\, x = -x$.

Proofs. We shall prove properties ($\mathbf{L}_{11}$), ($\mathbf{L}_{13}$) and ($\mathbf{L}_{16}$) and leave the proofs of the remaining properties as exercises. We may write

$$\lambda x = \lambda(x + \mathbf{0}) = \lambda[x + (y - y)] = \lambda[(x - y) + y] = \lambda(x - y) + \lambda y \ ;$$

axioms ($\mathbf{L}_4$), ($\mathbf{L}_3$), ($\mathbf{L}_5$), ($\mathbf{L}_2$) and ($\mathbf{L}_{10}$) have been here applied. If we now add the negative $-(\lambda y)$ to both the leftmost and rightmost sides of this (multiple) equation we obtain ($\mathbf{L}_{11}$).

($\mathbf{L}_{13}$) readily follows from ($\mathbf{L}_{10}$) and ($\mathbf{L}_4$), with the zero scalar written as $0 = 1 - 1$; indeed, $0\,x = (1 - 1)\,x = 1x - 1x = x - x = \mathbf{0}$. Finally, to check ($\mathbf{L}_{16}$) we must prove that $x + (-1)x = \mathbf{0}$. To see this, let us remark that

$$x + (-1)x = 1x + (-1)\,x = [1 + (-1)]\,x = 0\,x = \mathbf{0}.$$

The last equation follows from the just proved property (L_{13}). ■

Examples of vector spaces

Example 1.1. The set of "geometrically defined" (free) vectors in the plane or in the space forms a linear space over the field $\mathbb{R}$ of real numbers. We do not insist here on this space since it will be extensively presented in the last chapter.

Example 1.2. For a given field $\mathbf{F}$ and any integer $n \geq 1$, denote by $\mathbf{F}^n$ the set of all ordered n-tuples

$$X = (x_1, x_2, \ldots, x_n) \quad \text{where} \quad (\forall\, i)\ x_i \in \mathbf{F}\,;$$

X and $Y = (y_1, y_2, \ldots, y_n)$ are distinct unless $x_1 = y_1, \ldots,\ x_n = y_n$. This set $\mathbf{F}^n$ forms a vector space with the operations defined by

$$X + Y \underset{\text{def}}{=} (x_1 + y_1, \ldots, x_n + y_n)\,, \tag{1.5}$$

$$\lambda X \underset{\text{def}}{=} (\lambda x_1, \ldots, \lambda x_n)\,. \tag{1.6}$$

This is an important example, which in many ways is typical (as it will follow from a theorem to be presented later on in this section). For $X = (x_1, x_2, \ldots, x_n)$, $x_1, \ldots, x_n$ are said to be the *components* of the vector X. In particular, it follows for $n = 1$ that any field $\mathbf{F}$ may be regarded as *a vector space over itself*. For $\mathbf{F} = \mathbb{R}$ we obtain the space $\mathbb{R}^n$ of the n-dimensional real vectors. Let us still mention that the vectors X in this space $\mathbf{F}^n$ may be regarded as one-row / one-column matrices :

$$X = [x_1\ x_2\ \ldots\ x_n] \quad \text{or} \quad X = [x_1\ x_2\ \ldots\ x_n]^{\mathrm{T}}. \tag{1.7}$$

Example 1.3. Let $\mathbf{F}$ be an arbitrary field, and let us denote by $\mathbf{F}^{\mathbb{N}}$ the set of *infinite sequences* over $\mathbf{F}$, $(x_1, x_2, \ldots, x_n, \ldots)$. Addition and multiplication by scalars can be defined as in (1.5) and (1.6), and it is obvious that the two operations endow this set with the structure of a vector space over $\mathbf{F}$. In particular, the set $\mathbb{R}^{\mathbb{N}}$ of the sequences of real numbers is a vector (or linear) space over $\mathbb{R}$.

Example 1.4. Let $\mathscr{M}_{m,n}$ denote the set of all matrices of size m – by – n (that is, with m rows and n columns) over a field $\mathbf{F}$ / over $\mathbb{R}$. The two operations involved in *Def.1.1* are defined on $\mathscr{M}_{m,n}$ as follows :

Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be in $\mathscr{M}_{m,n}$; then

$$A + B \underset{\text{def}}{=} [a_{ij} + b_{ij}] \quad \text{and} \quad \lambda A \underset{\text{def}}{=} [\lambda\, a_{ij}] \tag{1.8}$$

obviously are in $\mathscr{M}_{m,n}$ too, and the other eight properties of the addition and multiplication by scalars can be easily verified. Therefore, the set of m – by – n matrices over $\mathbf{F}$ / over $\mathbb{R}$ is a vector space over that field. In particular, the two sets $\mathscr{M}_{1,n}$ and $\mathscr{M}_{m,1}$ are practically identical to the space $\mathbf{F}^n / \mathbb{R}^n$, since they consist of (ordered) n-tuples of scalars in $\mathbf{F}$ / of real numbers ; the only difference regards the way these n-tuples are written, that is

$$X^{\mathrm{T}} = [x_1 \ x_2 \ \ldots \ x_n] \quad \text{or} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \ \text{respectively.} \tag{1.9}$$

Example 1.5. Let I denote an interval in the set $\mathbb{R}$ of the real numbers and let $\mathscr{F}_I$ be the set of all real functions defined on I. The two linear operations on $\mathscr{F}_I$ are introduced in a natural way by

$$(f+g)(x) \underset{\text{def}}{=} f(x) + g(x) \quad \text{and} \quad (\lambda f)(x) \underset{\text{def}}{=} \lambda f(x) \tag{1.10}$$

In particular, the set $\mathbf{C}_I$ of the *continuous* functions on interval I is a vector space over $\mathbb{R}$. The same property holds for the *differentiable* functions on I.

Example 1.6. We are closing this set of examples by a trivial one, namely the space consisting of a single vector $\mathbf{0}$ with the two linear operations defined by $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $\lambda\,\mathbf{0} = \mathbf{0}$ for any scalar $\lambda \in \mathbf{F}$. This space $V_0 = \{\mathbf{0}\}$ is called the *trivial space* (or the ***null space***); the reader can check that all the ten axioms in *Def. 1.1* are satisfied. □

Let now V be an arbitrary vector space over a field $\mathbb{F}$. A new (compound) operation involving **several** vectors (and scalars) may be introduced by

Definition 1.2. An expression of the form

$$\boxed{\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n = \sum_{i=1}^{n} \lambda_i x_i ,} \tag{1.11}$$

in which $\lambda_1, \lambda_2, \ldots, \lambda_n$ are scalars in $\mathbf{F}$ and $x_1, x_2, \ldots, x_n$ are vectors in V, is called a *linear combination* of the vectors $x_1, x_2, \ldots, x_n$ (with scalar coefficients $\lambda_1, \lambda_2, \ldots, \lambda_n$). If a vector x is equal to some linear combination of $x_1, x_2, \ldots, x_n$ it is said to be *expressible linearly* in terms of these vectors.

◊

It is clear that an expression of the form (1.11) is also a vector in V since its terms, that is $\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_n x_n$ are in V according to ($\mathbf{L}_6$) - *Def. 1.1*, and a multiple sum of terms is well defined in any additive structure with associative addition – see axiom ($\mathbf{L}_2$).

It will be convenient (in what follows) to use a special kind of a so-called "matrix notation" for writing linear combinations of the form (1.11). Let us denote

$$\mathcal{X} = [x_1 \; x_2 \ldots x_n] \quad \text{and} \quad \Lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} ; \tag{1.12}$$

then the linear combination (1.11) may be written as

$$\sum_{i=1}^{n} \lambda_i x_i = \Lambda^{\mathrm{T}} \cdot \mathcal{X}^{\mathrm{T}} = \mathcal{X} \cdot \Lambda = \sum_{i=1}^{n} x_i \lambda_i . \tag{1.13}$$

Note that the components x_i's of $\mathcal{X}$ in (1.12) are *vectors* and not scalars as the ones of X in (1.7). On another hand, the rightmost side in Eq. (1.13) is a sum of terms of the form $x_i \lambda_i$ with the scalars written *after* the vectors ; this is a matter of convention, since both $\lambda_i x_i$ and $x_i \lambda_i$ may stand for the vector x_i multiplied by scalar λ_i.

The linear combinations are essentially involved in defining a pair of new notions, more precisely ***two complementary types of relations*** among several vectors in a vector space V. They are introduced by

Definition 1.3. Let V be a vector space over a field $\mathbf{F}$ and let us consider the equation

$$\boxed{\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_m x_m = \mathbf{0}\ ;} \tag{1.14}$$

the vectors $x_1, x_2, \ldots, x_m \in V$ are said to be

(*i*) *linearly independent* if Eq. (1.14) $\Rightarrow \lambda_1 = \lambda_2 = \ldots = \lambda_m = 0\ ;$ (1.15)

(*ii*) *linearly dependent* if (1.14) also holds for at least one $\lambda_i \neq 0$, that is,

$$(\exists\, \lambda_i \in \mathbf{F})\ \ \lambda_i \neq 0 \quad \text{and Eq. (1.14) still holds.} \tag{1.16}$$

A family $\mathcal{Q} = \{u_1, u_2, \ldots, u_m\}$ of vectors in V is said to be *linearly independent* / *dependent* when the vectors it consists of satisfy (1.15) / (1.16), respectively. ◊

Another notion has also to be introduced as a preliminary to the the definition of the important notion of basis :

Definition 1.4. Let V be a vector space over a field $\mathbf{F}$ and $W \subseteq V$. The subset W is said to be *spanned* (or *generated*) by a family $\mathcal{Q} = \{u_1, u_2, \ldots, u_m\}$ of vectors in V if

$$(\forall\, w \in W)(\exists\, \lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbf{F})\ \ w = \lambda_1 u_1 + \lambda_2 u_2 + \ldots + \lambda_m u_m = \sum_{i=1}^{m} \lambda_i u_i . \tag{1.17}$$

The vectors $u_1, u_2, \ldots, u_m$ are said to be the *generators* of the subset *W*. ◊

Under the conditions in *Definition 1.4* we will use the notation $W = \mathcal{L}(\mathcal{Q})$.

Example 1.7. Let us consider three vectors in the space $\mathbb{R}^3$, namely

$$X_1 = [2\ \ 1\ \ 3]^{\mathrm{T}},\ X_2 = [1\ \ -1\ \ 3]^{\mathrm{T}},\ X_3 = [0\ \ 3\ \ -3]^{\mathrm{T}}.$$

It can be seen that $X_1 - 2X_2 - X_3 = \mathbf{0}$, therefore the three vectors are linearly dependent ; since $X_1 = 2X_2 + X_3$, it follows that

$$X_1 \in \mathcal{L}(\{X_2, X_3\}).$$ □

Remark 1.5. Given a family $\mathcal{Q}$ of m vectors and another vector x in *V*, it is possible that x may be written as a linear combination of the vectors of $\mathcal{Q}$, that is, as in Eq. (1.17). But it is also possible that *no such expression exists*, or – equivalently – that x is not linearly expressible in terms of the vectors in $\mathcal{Q}$. And, when it is expressible, its linear expression (1.17) may be *not unique*. Just this happens in the case when $\mathcal{Q}$ is a *dependent* family of vectors. To illustrate this latest remark, let us consider the vector $Y = [3\ \ 3\ \ 3]^{\mathrm{T}}$ and the three vectors of Example 1.7. It can be easily verified that

$$Y = X_1 + X_2 + X_3 = 2X_1 - X_2 .$$

Remark 1.6. *Any family of vectors including the zero vector is linearly dependent.*

Indeed, if the family $\mathcal{A} = \{u_1, u_2, \dots, u_m\}$ with $u_j = \mathbf{0}\ (1 \le j \le m)$ then we can take a single scalar

$$\lambda_j \ne 0 \text{ while } \lambda_1 = \dots = \lambda_{j-1} = \lambda_{j+1} = \dots = \lambda_m = 0\,; \tag{1.18}$$

$$(1.18)\ \Rightarrow\ \sum_{i=1}^{m} \lambda_i u_i = \mathbf{0}$$

and this latter equality obviously follows from properties ($\mathbf{L}_{13}$) & ($\mathbf{L}_{14}$) in PROPOSITION 1.1. Hence the vectors in are linearly dependent, in view of (1.16) in *Def. 1.3*.

Definition 1.5. Let V be a vector space over $\mathbf{F}$ and $\mathcal{A} = \{a_1, a_2, \dots, a_n\}$ a (finite) family of vectors in V. $\mathcal{A}$ is said to be a ***basis*** of V if

(i) $\mathcal{A}$ is linearly independent, and

(ii) $\mathcal{A}$ spans V, that is (see *Def. 1.4* – Eq.(1.17)), $V = \mathcal{L}(\mathcal{A})$. ◊

Remark 1.7. A basis $\mathcal{A}$ of a space V has been considered as a (finite) spanning family, therefore as a subset of V. The fact that $\mathcal{A}$ is written as a finite set in *Def. 1.5* is not essential. Moreover, there are vector spaces which do not admit any finite basis. But another problem appears concerning the nature of a basis: as it will be argued a little later, any basis should be considered – in fact – as an *ordered* family of vectors. Hence, a basis is an ordered n-tuple of vectors that may be written (for instance) as *a row whose entries are vectors :*

$$A = [a_1\ a_2\ \dots\ a_n] \in V^n . \tag{1.19}$$

For any vector $x \in V$, it follows from *Def. 1.5*, (ii) that x *can be linearly expressed* in terms of the vectors $a_1, a_2, \dots, a_n$:

$$\boxed{(\exists\ \xi_1, \xi_2, \dots, \xi_n \in \mathbf{F})\ \ x = \sum_{i=1}^{n} \xi_i a_i .} \tag{1.20}$$

Using the "matrix notation" (1.12) - (1.13) for linear combinations, an expression like (1.20) may be written as

$$x = A \cdot X_A \quad \text{with} \quad A \text{ of (1.19) and } X_A = [\xi_1\ \xi_2\ \dots\ \xi_n]^{\mathrm{T}} . \tag{1.21}$$

The scalar components of the column vector X_A, that is $\xi_1, \xi_2, \dots, \xi_n$, are said to be *coordinates* of x in the basis A. We do not yet say the coordinates (...) before stating and proving the next result :

PROPOSITION 1.2. *Let A be a basis of the vector space V and $x \in V$. Then the linear expression* (1.20) - (1.21) *of x in basis A is unique.*

Proof. We have to show that – given a basis A of V and a vector $x \in V$ – the vector of coordinates X_A of x in basis A is unique. Let us assume that x admits (at least) two linear expressions of the form (1.21) in basis A, that is

$$(\exists\, X_A\,,\, X_A' \in \mathbf{F}^n) \quad x = A \cdot X_A = A \cdot X_A' \tag{1.22}$$

or (using the explicit expression (1.20))

$$\left.\begin{aligned} x &= \sum_{i=1}^{n} \xi_i\, a_i\,, \\ x &= \sum_{i=1}^{n} \xi_i'\, a_i\,. \end{aligned}\right\} \tag{1.23}$$

Subtracting (side-by-side) the two equations in (1.23) and applying axiom ($\mathbf{L}_{10}$) we get

$$\sum_{i=1}^{n} (\xi_i - \xi_i')\, a_i = \mathbf{0}. \tag{1.24}$$

But Eq.(1.24) gives a linear combination of the vectors of basis A *equal to* $\mathbf{0}$. Taking into account condition (*ii*) of *Def. 1.5* (the linear independence of a_i's) and Eqs. (1.15) of *Def. 1.3,* we get

$$\xi_1 - \xi_1' = \xi_2 - \xi_2' = \ldots = \xi_n - \xi_n' = 0 \iff X_A = X_A'. \tag{1.25}$$

Eq. (1.25) shows the unicity of the coordinates of x in basis A. ■

Example 1.8. Let us find a basis in the space $\mathbb{R}^n$ presented in **Example 1.4**. The most convenient spanning family for this space consists of the vectors

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},\ e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots,\ e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}. \tag{1.26}$$

If we write together (in the natural order of their subscripts) the column vectors $e_i \in \mathbb{R}^n$ we obviously get the identity matrix of order n :

$$E = [e_1\, e_2 \ldots e_n] = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix} = I_n\,. \tag{1.27}$$

It is clear that E is an independent family (since rank $E = n$) and it spans the whole space $\mathbb{R}^n$. Indeed, if $X = [x_1\ x_2\ \ldots\ x_n] \in \mathbb{R}^n$ then it can be readily verified that

$$X = \sum_{i=1}^{n} x_i e_i . \tag{1.28}$$

It follows from expression (1.28) that the coordinates of a vector $X \in \mathbb{R}^n$ *are just its components*. The basis E of (1.26) - (1.27) is the only basis of space $\mathbb{R}^n$ with this property. It is called the *standard basis* (of the real *n*-dimensional space).

Example 1.9. A general *polynomial of order n*, with its coefficients in a field **F**, in particular in the real field $\mathbb{R}$, can be written as

$$p = a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n . \tag{1.29}$$

Let us denote the set of polynomials of order n over $\mathbf{F}$ / $\mathbb{R}$ by

$$\mathrm{POL}_n(\mathbf{F}) / \mathrm{POL}_n(\mathbb{R}). \tag{1.30}$$

If another polynomial of the general form (1.29) is considered, for instance

$$q = b_0 + b_1 t + b_2 t^2 + \ldots + b_n t^n ,$$

the two linear operations with polynomials are naturally defined by

$$p + q \underset{\text{def}}{=} (a_0 + b_0) + (a_1 + b_1)t + (a_2 + b_2)t^2 + \ldots + (a_n + b_n)t^n ; \tag{1.31}$$

$$\lambda p \underset{\text{def}}{=} (\lambda a_0) + (\lambda a_1)t + (\lambda a_2)t^2 + \ldots + (\lambda a_n)t^n . \tag{1.32}$$

It can be readily seen that these operations defined by (1.31) & (1.32) satisfy the ten axioms of a vector space in *Definition 1.1*. The proof is left as an exercise to the reader. Let us only specify the "special" elements: the *zero polynomial* is $\mathrm{O} = o + ot + ot^2 + \ldots + ot^n$; the *negative* of a polynomial p of the form (1.29) is $-p = (-a_0) + (-a_1)t + (-a_2)t^2 + \ldots + (-a_n)t^n$. It follows that the set(s) in (1.30), endowed with the linear operations of (1.31) & (1.32), is (are) vector space(s).

□

The next PROPOSITION presents two relevant properties of linearly dependent / independent sets (or families) of vectors.

PROPOSITION 1.3. *Let V be a vector space and*

$$\mathcal{Q} = \{u_1, u_2, \ldots, u_m (,\ldots)\} \subset V \tag{1.33}$$

a family of vectors.

(*i*) *If $\mathcal{Q}$ is linearly dependent and $\mathcal{Q} \subseteq \mathcal{Q}'$ then $\mathcal{Q}'$ is dependent, too ;*

(*ii*) *If $\mathcal{Q}$ is linearly independent and $\mathcal{Q}''\subseteq\mathcal{Q}$ then $\mathcal{Q}''$ is also independent.*

Proof. The way we have written the "contents" of the family $\mathcal{Q}$ in (1.33) suggests that it can be not necessarily finite: it may contain infinitely many vectors.

(*i*) The linear dependence of the family in (1.33) means, according to *Def. 1.3* - (1.16), that an equation like (1.14) holds even when at least one scalar is $\neq o$. Hence

$$(\exists\,\lambda_j\in\mathbf{F})\quad \lambda_j\neq 0\ \ \&\ \sum_{i=1}^{m}\lambda_i u_i=\mathbf{0}. \tag{1.34}$$

The inclusion relation in the statement means that the other family of vectors looks like

$$\mathcal{Q}'=\{u_1,u_2,\ldots,u_m(,\ldots),v_1,v_2,\ldots\}. \tag{1.35}$$

It follows that there exists a zero linear combination of the form

$$(\exists\,\lambda_j\in\mathbf{F})\ \lambda_j\neq o\ \ \&\ \sum_{i=1}^{m(\cdots)}\lambda_i u_i+\sum_{k\geq 1}\mu_k v_k=\mathbf{0} \tag{1.36}$$

with at least one nontrivial term in the first sum of (1.36), $\lambda_j u_j$ while we may take trivial (zero) scalar coefficients on the vectors in the complementary set $\mathcal{Q}'\setminus\mathcal{Q}$, that is

$$\mu_1=\mu_2=\ldots=0\ \Rightarrow\ \sum_{k\geq 1}\mu_k v_k=\mathbf{0}+\mathbf{0}+\ldots=\mathbf{0}. \tag{1.37}$$

Obviously, property ($\mathbf{L}_{13}$) in PROPOSITION 1.1 has been here involved. It follows from (1.34) with (1.37) that Eq. (1.36) holds for (at least) one non-zero scalar and the family $\mathcal{Q}'$ is thus linearly dependent.

(*ii*) The second implication in the statement immediately follows from former, by *reductio ad absurdum*. If the inclusion $\mathcal{Q}''\subseteq\mathcal{Q}$ holds with independent $\mathcal{Q}$ and we would assume that $\mathcal{Q}''$ could be linearly dependent, then, in view of (*i*), the larger family $\mathcal{Q}$ would be dependent ! This closes the proof. ■

Before stating (and proving) a couple of consequences of this simple result, let us reformulate it as follows:

(*i*) *Any subfamily of an independent family is independent, too* ;
(*ii*) *any superfamily of a dependent family is also dependent.*

COROLLARY 1.1. *Given a basis A of a vector space V, each element (vector) of V can be linearly expressed in basis A in one and only one way, and this is no longer true if any vector whatsoever is appended to or deleted from the basis A.*

Proof. Let A be a basis of space V. This implies (by *Def. 1.5* - (*ii*)) that

$$V = \mathcal{L}(A) \Rightarrow$$
$$\Rightarrow \left[x \in V \Rightarrow (\exists\, \xi_1, \xi_2, \ldots, \xi_n \in \mathbf{F}) \;\; x = \sum_{i=1}^{n} \xi_i a_i = A \cdot X_A \right]. \tag{1.38}$$

According to PROPOSITION 1.2 - (1.25), the linear expression of x in basis *A is unique*. If (at least one) vector a_j is deleted from the basis A it is obtained the smaller family $A' = A \setminus \{a_j\}$. We have just used the notation for the set subtraction although basis A is considered as an ordered family, as in (1.19). If $A' = A \setminus \{a_j\}$ would remain a basis of space V, the vector a_j would be linearly expressible in A':

$$a_j = \sum_{i=1}^{j-1} \mu_{ji} a_i + \sum_{i=j+1}^{n} \mu_{ji} a_i. \tag{1.39}$$

But this Eq. (1.39) implies the *linear dependence of the basis A*! Similarly, if a "new" vector *b* is adjoined to *A*, resulting in the larger (spanning) family

$$A'' = A \cup \{b\}, \tag{1.40}$$

this vector is also linearly expressible in basis A; hence, the set in (1.40) may be a spanning family for *V* but – in no case – a basis since *it is not independent*. ■

Comments. The statement of the above COROLLARY occurs in the textbook [S. LANG, 1988]. In the first edition of our textbook of LINEAR ALGEBRA [A. Carausu, 1999], we included in the statement of this result (COROLLARY 1.1, page 11) other two characterizations of the notion of basis. However, we are going to include them, more properly, in a theorem that follows and brings together the main characterizations of a basis.

Before presenting another result, let us state a remark, in fact a characterization of dependent vectors / families of vectors.

A family $\mathcal{Q}$ of vectors is linearly dependent iff (if and only if) *at least one vector in $\mathcal{Q}$ is linearly expressible in terms of the other vectors*.

The proof of this rather obvious property will be proposed as an exercise in the next section of Applications (exercises), **2-A.1** .

PROPOSITION 1.4. *Let V be a vector space and*

$$\mathcal{Q} = \{u_1, u_2, \ldots, u_p\} \subset V \tag{1.41}$$

a family of linearly independent vectors. Then any $p + 1$ *vectors in* $\mathcal{L}(\mathcal{Q})$ *are linearly dependent.*

Proof. By the hypothesis, the independence of $\mathcal{Q}$ means that

$$\sum_{j=1}^{p} \lambda_j u_j = \mathbf{0} \;\Rightarrow\; \lambda_1 = \lambda_2 = \ldots = \lambda_p = 0. \tag{1.42}$$

Consider now the $p + 1$ vectors in the statement,

$$x_1, x_2, \ldots, x_{p+1} \in \mathcal{L}(\mathcal{Q}). \tag{1.43}$$

By *Def. 1.4*, it follows that each of them admits a linear expression in terms of the vectors of (1.41). But we must write the scalars in the expressions of the form in Eq. (1.17) with doubly indexed coefficients: for each vector $x_i (1 \le i \le p + 1)$ of (1.43), p scalars exist such that

$$x_i = \sum_{j=1}^{p} \alpha_{ij} u_j. \tag{1.44}$$

Let us now write the equation of the form (1.14), involved in the definition of both linear dependence and independence, for the vectors in (1.44):

$$\sum_{i=1}^{p+1} \beta_i x_i = \mathbf{0} \underset{(1.44)}{\Rightarrow} \sum_{i=1}^{p+1} \beta_i \sum_{j=1}^{p} \alpha_{ij} u_j = \mathbf{0}. \tag{1.45}$$

In the double (or iterated) sum of Eq. (1.45), the summation order can be inverted and it thus follows that

$$\sum_{j=1}^{p} \Big(\sum_{i=1}^{p+1} \alpha_{ij} \beta_i \Big) u_j = \mathbf{0}. \tag{1.46}$$

The sums between the big parentheses of Eq. (1.46) play the role of the scalars λ_j of (1.42) and they must therefore vanish, due to the independence of the set $\mathcal{Q}$:

$$\sum_{i=1}^{p+1} \alpha_{ij} \beta_i = 0, \;\; j = \overline{1, p}. \tag{1.47}$$

The p equations in (1.47) form a homogeneous system in the unknowns β_i $(i = \overline{1, p+1})$.

Since the matrix $[\alpha_{ij}]$ of its coefficients is of size p-by-$(p+1)$, its rank is at most $= p < p + 1 =$ the number of unknowns. It is known (from the highschool algebra) that such a homogeneous system admits nontrivial solutions. Hence, at least one β_i in the first equation of (1.45) is $\neq 0$ and the conclusion in the statement thus holds. ■

It follows from this PROPOSITION 1.3 the next property.

COROLLARY 1.2. *If a subset $W \subseteq V$ is generated by p linearly independent vectors then any m vectors of W are linearly dependent if $m > p \Longleftrightarrow$ $\Longleftrightarrow m \geq p+1$.*

This property could be equivalently stated as follows.

If a subset $W \subseteq V$ is generated by p linearly independent vectors then at most p vectors in W can be linearly independent.

The properties so far presented make possible to state and prove an important result concerning the number of vectors in the bases of the same vector space.

THEOREM 1.1. *Let V be a (finitely generated) vector space over the field $\mathbf{F}$. If V is spanned by two bases A and B then the number of vectors in A and B is the same.*

Proof. Let the two bases in the statement be

$$A = [a_1 \ a_2 \ldots a_n] \in V^n \ \ \& \ \ B = [b_1 \ b_2 \ldots b_m] \in V^m. \tag{1.48}$$

Regarding the number(s) of vectors in the two bases, let us assume that

$$n \neq m \Longleftrightarrow [m < n \text{ or } m > n]. \tag{1.49}$$

In the first alternative of (1.49), all the vectors of basis A are in $\mathscr{L}(A)$; hence, expressions of the form

$$a_i = \sum_{j=1}^{m} \alpha_{ij} b_j \tag{1.50}$$

hold. But any such equation implies the linear dependence of the vectors in A and this contradicts condition (*i*) in the definition of a basis – *Def. 1.5*. Similarly, the other inequality between m & n is also impossible and we thus have $m = n$ and the proof is over. ■

COROLLARY 1.3. *The number of vectors in all the bases of a* (finitely generated) *vector space is the same*.

Hence this (natural) number *is an intrinsic feature of a vector space*, in the sense that it does not depend on a particular basis that spans the space. It is therefore natural to state the following

Definition 1.6. The common number of vectors in every basis of a (finitely generated) vector space V is called the *dimension of* V and it is denoted as

$\dim V$. ◊

Remarks 1.8. In the previous definition, only finitely generated have been considered. If $V \neq \{\mathbf{0}\}$ does not admit a finite spanning family it is not finitely generated. We may write $\dim V = \infty$. The earlier mentioned functional spaces (see Example 1.5 at page 4) are infinite-dimensional. In order to determine the dimension of a finitely generated vector space it suffices to find a basis which spans it. The following examples are relevant in this sense.

Examples 1.10. We presented the space $\mathbb{R}^n$ of the ordered n-tuples of real numbers, in Example 1.8 - page 8, with its standard basis E consisting of n (column) vectors. Therefore

$$\boxed{\dim \mathbb{R}^n = n.}$$

The space $\mathcal{M}_{m,n}$ of m - by - n matrices over a field $\mathbf{F}$ / over $\mathbb{R}$ also admits a finite (standard) basis. It suffices to consider the mn "elementary" matrices

$$e_{ij} = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}. \tag{1.51}$$

The only nonzero (unit) entry of this matrix appears in its i-th row and j-th column. It is very easy to see that any matrix $A = [a_{ij}] \in \mathcal{M}_{m,n}$ can be written as a linear combination of such matrices, the scalars being just its entries:

$$A = [a_{ij}] = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}\, e_{ij}. \tag{1.52}$$

The linear independence of the matrices in (1.51) is obvious: if a linear combination thereof (written as a double sum like in (1.52)) is equated to the zero matrix $O = [0]_{m,n}$ it obviously follows that all the scalars should be $= 0 : \lambda_{ij} = 0 \;\; (\forall\, i \in \overline{1,m},\ \forall\, j \in \overline{1,n})$. Therefore $\dim \mathcal{M}_{m,n} = m\,n$.

The space of *polynomials* (of order n) was presented in Example 1.9 at page 10. For this space $\mathrm{POL}_n(\mathbf{F}) / \mathrm{POL}_n(\mathbb{R})$, a standard (most convenient) basis can be obviously considered, taking into account the general expression of a polynomial - Eq. (1.29):

$$p = a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n. \tag{1.29}$$

This basis is

$$B = [1, t, t^2, \ldots, t^n] \tag{1.53}$$

and it is easy to see that a polynomial of the form (1.29) can be written, with our "matrix notations" (1.19) - (1.21) at page 7, as

$$p = [1, t, t^2, \ldots, t^n] \cdot [a_0, a_1, a_2, \ldots, a_n]^{\mathrm{T}}. \tag{1.54}$$

Expression (1.54) implies that $\mathrm{POL}_n(\mathbf{F}) / \mathrm{POL}_n(\mathbb{R}) = \mathscr{L}(B)$.

The independence of the "elementary" polynomials of (1.53) – the components of B – immediately follows from the definition of the identically zero polynomial O – the polynomial with $a_0 = a_1 = a_2 = \ldots = a_n = 0$. Hence $\dim \mathrm{POL}_n(\mathbf{F}) = n+1 = \dim \mathrm{POL}_{\mathrm{n}}(\mathbb{R})$. In fact, this space of n-order polynomials is in 1-to-1 correspondence with the space $\mathbf{F}^{n+1} / \mathbb{R}^{n+1}$ since any polynomial of the form (1.29) is uniquely determined by its coefficients that appear as the components of the column vector – the second factor in the "formal"product of (1.54). □

PROPOSITION 1.5. *Let V be a vector space over the field* $\mathbf{F}$. *Any independent family of vectors $\mathscr{A} = \{a_1, a_2, \ldots, a_m\} \subset V$ can be extended up to a basis of the space V.*

Proof. The first condition in the definition of a basis (*Def. 1.5*) is met by the family $\mathscr{A}$, but it could not satisfy the second condition, (*ii*): $V = \mathscr{L}(\mathscr{A})$. Here the dimension of the space V should be considered. Let us assume that $\dim V = n$. If $m > n$, it follows from COROLLARY 1.2 that $\mathscr{A}$ is dependent and it cannot form a basis. Moreover, it cannot be extended up to basis since any superfamily of $\mathscr{A}$ will be also dependent, according to PROPOSITION 1.3 (pages 9-10). If $m = n$ then $\mathscr{A}$ is already a basis. Indeed, it is independent as stated and any vector $x \in V$ admits a linear representation in terms of (the vectors) of $\mathscr{A}$: in the particular (or even trivial) case when $x \in \mathscr{A}$ it follows that

$$(\exists\, j \in \overline{1,m})\ \ x = a_j = 0\, a_1 + \ldots + 0\, a_{j-1} + 1\, a_j + 0\, a_{j-1} + \ldots + 0\, a_n.$$

It follows that the relevant case is the one when $x \notin \mathscr{A}$, what implies that

$$\mathscr{A}' = \{a_1, a_2, \ldots, a_n, x\} \tag{1.55}$$

cannot be a basis of V: it spans the space but it consists of $n+1 > n$ = the dimension of V: the family in (1.55) is necessarily dependent. Therefore, the family in the statement could be effectively extended up to a basis only if $m < n$. This extension proceeds step by step, in fact vector by vector: the family $\mathscr{A} = \{a_1, a_2, \ldots, a_m\}$ is extended by adjoining one vector at a time, as it follows.

$$\mathscr{A}'' = \{a_1, a_2, \ldots, a_m, a_{m+1}\} \ \text{ with } \ a_{m+1} \neq 0, a_1, a_2, \ldots, a_m. \tag{1.56}$$

This adjoined vector can be selected in such a way that the family $\mathscr{A}''$ remains independent. Indeed, if no such vector would exist then it would follow that the greatest number of independent vectors in V would be $m < n =$

$= \mathbf{dim}\, V$, what contradicts the assumption on the dimension of the space. This extension process can be continued by adjoining another vector a_{m+2} to $\mathcal{A}''$ and so on, until an independent family consisting of n vectors is obtained. The property that it spans the whole space V follows by the previous argument - see (1.55). This completes the proof. ∎

Another – and somehow dual – property regards the possibility to reduce a (larger) spanning family of a vector space to a basis of its.

PROPOSITION 1.5. *Let V be a vector space over the field* $\mathbf{F}$. *Any spanning family of vectors $\mathcal{A} = \{a_1, a_2, \ldots, a_m\} \subset V$ can be reduced down to a basis of the space V.*

We do not give a (detailed) proof of this result: it may remain as an exercise to the reader.

The earlier presented properties and characterizations of the bases of a vector space are stated together in the result that follows, including *Def. 1.5*.

THEOREM 1.2. *Let V be a (finitely generated) vector space over the field* $\mathbf{F}$.

① *A family of vectors $\mathcal{A} = \{a_1, a_2, \ldots, a_n(,\ldots)\}$ in V is a **basis** of V if*
 (*i*) *$\mathcal{A}$ is linearly independent, and*
 (*ii*) *$\mathcal{A}$ spans V,* that is $V = \mathcal{L}(\mathcal{A})$.

② *If $\mathcal{A} = \{a_1, a_2, \ldots, a_n\}$ is a set (or family) of n linearly independent vectors in an n-dimensional space V then it is a basis for V.*

③ *If $\mathcal{A} = \{a_1, a_2, \ldots, a_n\}$ is a set (or family) of n vectors that span the n–dimensional space V then it is a basis for V.*

④ *Given a basis A of a vector space V, each element (vector) of V can be linearly expressed in basis A in one and only one way, and this is no longer true if any vector whatsoever is appended to or deleted from the basis A.*

⑤ *If $\mathcal{A} = \{a_1, a_2, \ldots, a_m\}$ is an independent family in the n-dimensional space V and $m < n$, it can be extended up to a basis of the space V.*

⑥ *If $\mathcal{A} = \{a_1, a_2, \ldots, a_p\}$ is a spanning family of the n-dimensional space V and $p > n$, it can be reduced down to a basis of the space V.*

⑦ *An independent family $\mathcal{A}$ in V is a basis $\Longleftrightarrow$ no set which properly contains $\mathcal{A}$ is independent.*

⑧ *A spanning family $\mathcal{A}$ of V is a basis $\Longleftrightarrow$ no proper subset of $\mathcal{A}$ still spans*

Comments. The most part of the properties just stated in THEOREM 1.2 were earlier proved. Other ones will be proposed for being proved in the next section of APPLICATIONS - EXERCISES. The notion of basis in a vector space was synthetically characterized in the excellent textbook of LINEAR ALGEBRA [G. STRANG, 1988 - page 86] as follows :

A basis is a ***maximal independent set***. A basis is also a ***minimal spanning set***.

Examples 1.11. Let $A = [a_{ij}]$ be an m – by – n matrix. Obviously, its columns are vectors in the vector space $\mathbb{R}^m$. They could form a basis for this space it the conditions in *Def. 1.5* would be satisfied.

As it was be presented in section **§ 1.2**, mainly devoted to matrices, we used the following notations for the columns / rows of a matrix :

$$A^j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} (1 \le j \le n) \;\Rightarrow\; A = [A^1 \dots A^j \dots A^n] \;; \tag{1.57}$$

$$A_i = [a_{i1}\ a_{i2} \dots\ a_{in}]\ (1 \le i \le m) \;\Rightarrow\; A = \begin{bmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_m \end{bmatrix}. \tag{1.58}$$

Coming back to the columns of matrix $A = [a_{ij}]$ that occur in (1.57), any vector $X \in \mathbb{R}^m$ will admit a linear expression in terms of $A^1, \dots, A^j, \dots, A^n$ if an equation of the form

$$X = \lambda_1 A^1 + \dots + \lambda_j A^j + \dots + \lambda_n A^n \tag{1.59}$$

will be satisfied. This equation (1.59) is equivalent to a non-homogeneous linear system whose augmented matrix is

$$\tilde{A} = [A^1 \dots A^j \dots A^n \mid X]. \tag{1.60}$$

The system corresponding to Eq. (1.59) admits (at least) a solution, i.e. it is consistent, if and only if

$$(\forall\, X \in \mathbb{R}^m)\ \mathrm{rank}\,[A^1 \dots A^j \dots A^n] = \mathrm{rank}\,[A^1 \dots A^j \dots A^n \mid X], \tag{1.61}$$

as it was stated and proved in **§ 1.2** (LINEAR SYSTEMS). Hence, condition

$$(1.61) \;\Rightarrow\; \mathbb{R}^m = \mathscr{L}([A^1 \dots A^j \dots A^n]). \tag{1.62}$$

But the first condition for $A = [A^1 \dots A^j \dots A^n]$ to be a basis for the space it spans consists in its linear independence. The columns of matrix $A = [A^1 \dots A^j \dots A^n]$ are independent $\Longleftrightarrow$ $\mathrm{rank}\,A = n$.

Taking into account the former condition (1.61) we can now write that

$$\boxed{[A^1 \dots A^j \dots A^n] = \text{ a basis for } \mathbb{R}^m}$$

$$\Updownarrow$$

$$\text{rank}\,[A^1 \dots A^j \dots A^n] = n + \text{Condition (1.61)}.$$

But the condition for the rank implies $n \leq m$ while condition (1.61) can be satisfied for *any* $X \in \mathbb{R}^m$ if and only if $\text{rank}\,[A^1 \dots A^j \dots A^n] = m$ since the nonhomogeneous linear system that is equivalent to Eq. (1.59) could be inconsistent for some vectors $X \in \mathbb{R}^m$ when $\text{rank}\,[A^1 \dots A^j \dots A^n] < m$. To conclude this discussion, it follows that all the columns of an m-by-n matrix can form a basis for space $\mathbb{R}^m \Longleftrightarrow$

$$\Longleftrightarrow \text{rank}\,[A^1 \dots A^j \dots A^n] = m = n.$$

Hence the matrix should be square. If we relax the condition in the statement of this example asking if only *some* of the columns of the matrix can form a basis for then it will suffice that $\text{rank}\,[A^1 \dots A^j \dots A^n] = m \leq n$.

Such a basis will consist of only m independent columns of A:

$$B = [A^{j_1}\, A^{j_2} \dots A^{j_m}] \text{ with } \det[A^{j_1}\, A^{j_2} \dots A^{j_m}] \neq 0.$$

Numerical example. Let us consider the matrix 3-by-4 matrix

$$A = \begin{bmatrix} 1 & 5 & 3 & -1 \\ 2 & 9 & 5 & -1 \\ -1 & -2 & 0 & 1 \end{bmatrix}. \tag{1.63}$$

It can be easily seen that $2A^1 + A^3 = A^2$; hence, the first three columns of A are linearly dependent and they cannot form a basis for $\mathbb{R}^3$. But columns A^1, A^3, A^4 are independent since

$$\det[A^1\, A^3\, A^4] = \begin{vmatrix} 1 & 3 & -1 \\ 2 & 5 & -1 \\ -1 & 0 & 1 \end{vmatrix} = -3 \neq 0.$$

Therefore $B = [A^1\ A^3\ A^4]$ is a basis for the space $\mathbb{R}^3$, but not the only possible one. The reader is invited to check that $C = [A^2\ A^3\ A^4]$ is also a basis. A similar discussion can regard the rows of matrix A of (1.63): can the three rows form a basis for the space $\mathbb{R}^4$? Anyway, the reader can check that the three rows are linearly independent ; but the condition

$$\mathbb{R}^4 = \mathscr{L}\left(\begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix}\right)$$

is or is not satisfied ? □

A final problem concerning the bases of a vector space regards the

Changes of bases and coordinates

As stated in the previous results (THEOREM 1.1, COROLLARY 1.3, ...), a vector space V *admits several bases*; in fact, it admits *infinitely many bases*. Indeed, if a basis $\mathscr{A} = \{a_1, a_2, \ldots, a_n\}$ exists and each vector is scalar-multiplied by an arbitrary non-zero scalar α_i $(1 \le i \le n)$ then the new family of vectors $\mathscr{A}' =$ $= \{\alpha_1 a_1, \alpha_2 a_2, \ldots, \alpha_n a_n\}$ is still a basis of the space. And there exist not less than ∞^n possibilities.

Let $A = [a_1, a_2, \ldots, a_n]$ and $B = [b_1, b_2, \ldots, b_n]$ be two bases of the space V. Any vector b_i $(1 \le i \le n)$ is a vector in $V = \mathscr{L}(A)$. Therefore a unique n-vector of coordinates exists such that each b_i $(1 \le i \le n)$ can be linearly expressed in terms of basis A ; formally,

$$(\forall\, i \in \{1, 2, \ldots, n\})\,(\exists^!\ [\tau_{i1}\ \tau_{i2} \ldots \tau_{in}] \in \mathbf{F}^n)\ b_i = \sum_{j=1}^{n} \tau_{ij}\, a_j \ \ (1 \le i \le n). \tag{1.64}$$

The n equations in (1.64), in fact n *unique* linear expressions of the vectors of the "new" basis B in terms of the initial (or "old") basis A can be written together, one under the other and also using the "matrix notations" for linear combinations and linear expressions: see Eqs. (1.13) at page 5 and (1.21) at page 7, respectively. To this end, let us write the scalar coefficients that occur in (1.64) as a row vector (or a row of a matrix) :

$$(\forall\, i \in \{1, 2, \ldots, n\})\ \ T_i = [\tau_{i1}\ \tau_{i2} \ldots \tau_{in}]\,. \tag{1.65}$$

Equations (1.64) can be written one under the other, for $i = 1, 2, \ldots, n$ resulting in

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_n \end{bmatrix} A^{\mathrm{T}} = \begin{bmatrix} \tau_{11}\ \tau_{12} \ldots \tau_{1n} \\ \tau_{21}\ \tau_{22}\ \ldots \tau_{2n} \\ \vdots \quad\ \vdots \qquad \vdots \\ \tau_{n1}\ \tau_{n2} \ldots \tau_{nn} \end{bmatrix} A^{\mathrm{T}}. \tag{1.66}$$

Obviously, the matrix equation (1.66) is equivalent to the set of (explicit) vector equations or equalities

$$\begin{cases} b_1 = \tau_{11} a_1 + \tau_{12} a_2 + \ldots + \tau_{1n} a_n \\ b_2 = \tau_{21} a_1 + \tau_{22} a_2 + \ldots + \tau_{2n} a_n \\ \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\ b_n = \tau_{n1} a_1 + \tau_{n2} a_2 + \ldots + \tau_{nn} a_n \end{cases}, \tag{1.67}$$

or also to its (more synthetic) matrix expression

$$\boxed{B^{\mathrm{T}} = T A^{\mathrm{T}} \text{ or } B = A\, T^{\mathrm{T}}.} \tag{1.68}$$

Obviously, the two equations in (1.68) can be obtained from each other by transposition, and the matrix

$$T = \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_n \end{bmatrix} = \begin{bmatrix} \tau_{11} & \tau_{12} & \ldots & \tau_{1n} \\ \tau_{21} & \tau_{22} & \ldots & \tau_{2n} \\ \vdots & \vdots & & \vdots \\ \tau_{n1} & \tau_{n2} & \ldots & \tau_{nn} \end{bmatrix} \tag{1.69}$$

is said to be the *transformation matrix from basis A to basis B.*

The next remark needs a couple of notions regarding matrices to be recalled. In fact, they were presented in **§ 1.1**.

The *rank* of a matrix $A \in \mathcal{M}_{m,n}$ is the maximum number of its linearly independent rows / columns. A square matrix A is *invertible* if there exists another square matrix A^{-1} such that $A A^{-1} = A^{-1} A = I_n$ = the identity matrix of order n. A square matrix A (of order n) is invertible if and only if it is nonsingular, what is equivalent to $\det A \neq 0 \iff \operatorname{rank} A = n$.

Remark 1.9. The transformation matrix T which is involved in Eqs. (1.66) and (1.68-69) is *nonsingular*, hence *invertible*. Indeed, let us assume that

$$\operatorname{rank} T < n \iff \det T = 0. \tag{1.70}$$

But both properties in (1.70) are equivalent to the linear dependence of the rows (or columns) of matrix T. This implies the existence of n scalars

$$\lambda_1, \lambda_2, \ldots, \lambda_n : \lambda_1 T_1 + \lambda_2 T_2 + \ldots + \lambda_n T_n = \mathbf{0}. \tag{1.71}$$

Obviously, $\mathbf{0}$ in (1.71) is the row zero vector in $\mathbb{R}^n$. Assume that there exists (at least one scalar) $\lambda_i \neq 0$ in this equation. Then it follows from Eq. (1.71) that the i - th row of the transformation matrix can be linearly expressed in terms of the other rows :

$$T_i = -\sum_{k \neq i} \frac{\lambda_k}{\lambda_i} T_k = \sum_{k \neq i} \mu_k T_k. \tag{1.72}$$

Let us now consider a null linear combination of the n vectors b_i $(1 \le i \le n)$ of basis B :

$$\alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_n b_n = \mathbf{0} \iff \iff \boldsymbol{\alpha} B^{\mathrm{T}} = \mathbf{0} \text{ with } \boldsymbol{\alpha} = [\alpha_1 \ \alpha_2 \ \ldots \ \alpha_n]. \tag{1.73}$$

The first expression of (1.68), taken to (1.73), gives the equation

$$\boldsymbol{\alpha}(TA^{\mathrm{T}}) = \mathbf{0} \iff (\boldsymbol{\alpha}T)A^{\mathrm{T}} = \mathbf{0} \in V. \tag{1.74}$$

But A is a basis and its independence implies that the row of scalars that premultiplies it in (1.74) should be the zero row vector : $\boldsymbol{\alpha} = \mathbf{0}^{\mathrm{T}} \in \mathbf{F}^n$. This matrix-form equation may be rewritten using the row-by-row structure of T (see (1.69)) as

$$\sum_{i=1}^{n} \alpha_i T_i = \mathbf{0}^{\mathrm{T}} = [0 \ 0 \ \ldots \ 0]. \tag{1.75}$$

This Eq. (1.75) can be rewritten by transposing it :

$$\sum_{i=1}^{n} \alpha_i T_i^{\mathrm{T}} = \sum_{i=1}^{n} T_i^{\mathrm{T}} \alpha_i = \mathbf{0} \iff T^{\mathrm{T}} \boldsymbol{\alpha}^{\mathrm{T}} = \mathbf{0}. \tag{1.76}$$

The last matrix equation in (1.76) can be written in terms of the column vectors using expression (1.72) of row T_i :

$$T_1^{\mathrm{T}} \alpha_1 + \ldots + \left(\sum_{k \ne i} \mu_k T_k^{\mathrm{T}} \right) \alpha_i + \ldots + T_n^{\mathrm{T}} \alpha_n = [0 \ 0 \ \ldots \ 0]^{\mathrm{T}}. \tag{1.77}$$

The system (1.77) is a homogeneous linear system of n equations in the n unknowns $\alpha_1, \alpha_2, \ldots, \alpha_n$. But its rank is at most $n - 1$ since the i-th column of its matrix is a linear combination of the other columns. Therefore it admits non-zero solutions, and Eq. (1.73) thus holds for some $\alpha_j \ne 0$, what contradicts the independence of (the vectors of) basis B. Hence the properties in (1.70) are false and the transformation matrix T is nonsingular :

$$\operatorname{rank} T = n \iff \det T \ne 0.$$

T is thus invertible. The same conclusion can be derived using the columns of T instead of its rows. ■

This conclusion gives the ground for proving the next result which gives a formula for changing the coordinates of a vector when the basis is changed.

PROPOSITION 1.6. *Let V be a vector space over the field $\mathbf{F}$ and A, B two bases spanning V. If basis A is changed for B with the transformation matrix T then T is invertible and the connection between the coordinates of a vector $x \in V$ in the two bases is given by*

$$\boxed{X_B = (T^{\mathrm{T}})^{-1} X_A \underset{\text{not}}{=} T^{-\mathrm{T}} X_A .} \tag{1.78}$$

Proof. The connection between the two bases is expressed by (one of the) Eqs. (1.68), with matrix-type notations. The non-singularity of the transformation matrix T has been stated and proved in the earlier *Remark*. The connection (1.78) between the coordinates of vector x now easily follows if we use the matrix form (1.21) of the expression of a vector in a basis, together with connection (1.68) between the two bases.

$$x = A X_A = B X_B \iff x = X_A^{\mathrm{T}} A^{\mathrm{T}} = X_B^{\mathrm{T}} B^{\mathrm{T}} \underset{(1.68-1)}{=} \tag{1.79}$$

$$= X_B^{\mathrm{T}} (T A^{\mathrm{T}}) = (X_B^{\mathrm{T}} T) A^{\mathrm{T}}. \tag{1.80}$$

The equality in (1.80) follows from the associativity of the matrix product; this property also holds when the third factor is a column of vectors – the column form of the basis A. But the first equation in (1.79), after $\iff$, and the last expression in (1.80) represent two linear expressions of the vector x in the same basis A. By the unicity of the coordinates, it follows that

$$X_A^{\mathrm{T}} = X_B^{\mathrm{T}} T \iff X_B = (T^{\mathrm{T}})^{-1} X_A. \tag{1.81}$$

The equivalence in (1.81) holds by post-multiplying the first equation with the inverse of the transformation matrix and – then – by transposing the two sides of the equation thus obtained.

$$X_A^{\mathrm{T}} = X_B^{\mathrm{T}} T \,|_{\cdot T^{-1}} \Rightarrow X_A^{\mathrm{T}} T^{-1} = X_B^{\mathrm{T}} (T\, T^{-1}) = X_B^{\mathrm{T}}. \tag{1.82}$$

The equations in (1.82), next to $\Rightarrow$, have followed by the associativity of the matrix product, with the definition of the inverse of a matrix: in this case,

$$T\, T^{-1} = I_n \;\;\&\;\; X_B^{\mathrm{T}} I_n = X_B^{\mathrm{T}}.$$

The equality of the rightmost to the leftmost side of (1.82) is just the transpose form of the formula (1.78) in the statement. This completes the proof. ■

Remarks 1.10. The proof of this rather important result offers an example of the usefulness of our matrix notations. The equivalent form of the connection formula (1.78) is

$$\boxed{X_B^{\mathrm{T}} = X_A^{\mathrm{T}} T^{-1}.} \tag{1.78'}$$

As regards the (somehow strange) notation that occurs in (1.78), that is

$$(T^{\mathrm{T}})^{-1} \underset{\text{not}}{=} T^{-\mathrm{T}},$$

it could be formally accepted in view of the rule of multiplication of powers.

As an application of the formulas (1.78) / (11.78'), let us see how the coordinates of a vector $X \in \mathbf{F}^n$ or $X \in \mathbb{R}^n$ in a given basis B can be effectively found. Let this basis be

$$B = [b_1, b_2, \dots, b_n].$$

The vectors b_i $(1 \le i \le n)$ are, in fact, column vectors in $\mathbf{F}^n$ or in $\mathbb{R}^n$, and therefore B is a square matrix of order n. Moreover, it is a nonsingular matrix since its columns are linearly independent (see the definition of the rank given earlier – at page 20). Let us also recall that the coordinates of a vector $X \in \mathbf{F}^n$ or $X \in \mathbb{R}^n$ in the standard basis E of this space are *identical to its components*. Hence such a vector may be written as

$$X = EX = BX_B. \tag{1.83}$$

But the standard basis $E = [e_1\ e_2 \dots e_n]$ as a matrix is just the identity matrix I_n: see the structure of this basis in **Example 1.8**, Eqs. (1.26-27) – page 8. Hence we may write

$$B = I_n B = EB \iff B^{\mathrm{T}} = B^{\mathrm{T}} I_n = B^{\mathrm{T}} E^{\mathrm{T}}. \tag{1.84}$$

The property of the transposing operator on the product of two matrices has been here applied : $(AB)^{\mathrm{T}} = B^{\mathrm{T}} A^{\mathrm{T}}$, as well as the property of the identity matrix (or of any diagonal matrix) – the transposing operator leaves it unchanged. If we compare equation (1.83) with (1.68), it is clear that the transformation matrix from the standard basis E to the basis B is just $T = B^{\mathrm{T}} \iff T^{\mathrm{T}} = B$. It now follows from (1.83) with formula (1.78) that

$$\boxed{X_B = B^{-1} X.} \tag{1.85}$$

But this formula (1.85) shows that the (column vector of) the coordinates of vector X in basis B are obtained from the solution to the matrix equation

$$BX_B = X. \tag{1.86}$$

In its turn, this equation is equivalent to a non-homogeneous linear system.

Example 1.12. Let us consider, in the space $\mathbb{R}^3$, a vector X and a basis B :

$$X = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix};\ B:\ b_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix},\ b_2 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix},\ b_3 = \begin{bmatrix} 0 \\ 4 \\ -2 \end{bmatrix}. \tag{1.87}$$

The matrix equation of (1.86) with the four vectors in (1.87) becomes

$$\begin{bmatrix} 1 & 3 & 0 \\ 2 & 0 & 4 \\ -1 & 1 & -2 \end{bmatrix} X_B = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}. \tag{1.88}$$

The matrix B has $\det B = -4 \neq 0$. Hence Eq. (1.88) will have a unique solution. This unicity theoretically follows from the unicity of the coordinates of a vector in a (given) basis - PROPOSITION 1.2. But it also follows from the properties of (non-homogeneous) linear systems, with the previous remark following after Eq. (1.86). A square n-by-n nonhomogeneous system with non-singular coefficient matrix has a unique solutions (as it is known from the highschool). Such systems are sometimes said to be *determined*, and we all call them *Cramer-type systems*. The linear system (equivalent to) Eq. (1.88) can be solved using transformations on the rows of its augmented matrix $\tilde{B}$ (as it will be explained, in more detail, in the next section), thus avoiding the effective determination of the inverse B^{-1}:

$$\tilde{B} = \left[\begin{array}{rrr|r} 1 & 3 & 0 & 3 \\ 2 & 0 & 4 & -1 \\ -1 & 1 & -2 & 2 \end{array}\right] \sim \left[\begin{array}{rrr|r} 4 & 0 & 6 & -3 \\ -1 & 1 & -2 & 2 \\ 2 & 0 & 4 & -1 \end{array}\right] \sim$$

$$\sim \left[\begin{array}{rrr|r} 1 & 0 & 3/2 & -3/4 \\ 0 & 1 & -1/2 & 5/4 \\ 0 & 0 & 1 & 1/2 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & 0 & 0 & -3/2 \\ 0 & 1 & 0 & 3/2 \\ 0 & 0 & 1 & 1/2 \end{array}\right] \Rightarrow X_B = \begin{bmatrix} -3/2 \\ 3/2 \\ 1/2 \end{bmatrix}. \tag{1.88}$$

The column vector of the coordinates of vector X in basis B means that we can write that

$$X = -\frac{3}{2} b_1 + \frac{3}{2} b_2 + \frac{1}{2} b_3. \tag{1.89}$$

This linear expression of (1.89) can be checked by the reader, with the data in (1.87). □

The next example illustrates the change of bases and coordinates, in the real Euclidean space.

Example 1.13. Check that the following two families of vectors are bases in space $\mathbb{R}^3$ and find the transformation matrix T from A to B.

$$A: a_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \ a_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ a_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}; \tag{1.90}$$

$$B: \ b_1 = \begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix}, \ b_2 = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}, \ b_3 = \begin{bmatrix} 6 \\ 7 \\ 9 \end{bmatrix}. \tag{1.91}$$

The two families A and B are bases since both of them consist of three vectors (= the dimension of $\mathbb{R}^3$) and their determinants are non-zero:

$$\det A = -3 \ \ \& \ \ \det B = 3.$$

The matrix equation connecting the three matrices A, B and T is either of the two equations in (1.68) - page 20. If we take the first of them,

$$B^{\mathrm{T}} = T A^{\mathrm{T}}, \tag{1.92}$$

the matrices A and B are known from (1.90) & (1.91) while T is the unknown matrix, to be found. With the above data, Eq. (1.92) becomes

$$T \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 3 \\ 2 & 2 & 3 \\ 6 & 7 & 9 \end{bmatrix}. \tag{1.93}$$

In order to get the transformation matrix , the two sides of Eq. (1.93) should be post-multiplied - that is multiplied at right - by the inverse of A^{T}. This inverse can be determined by a formula known from the highschool (involving the "adjugate" or cofactor matrix of A and its determinant $|A|$), but a more efficient method consists in using transformations on the rows of a block matrix of size 3-by-6, similar to those applied in the former example. But an even easier way to obtain T avoids the prior determination of $A^{-\mathrm{T}}$. To this end it is easier to skip to the second equation in (1.68), that is

$$B = A\,T^{\mathrm{T}} \ \Rightarrow \ T^{\mathrm{T}} = A^{-1} B. \tag{1.94}$$

The matrix product in the r.h.s. of Eq. (1.94) can be calculated by a transformation method applied on the rows of a block matrix of size 3-by-6, as earlier mentioned. This method is called *Gauss-Jordan* or *Gaussian elimination*, and it was presented and explained in the former section, **§ 1.2.** However, we are going to employ it to this example, after a brief recalling. Starting from the block matrix $[A \,|\, B]$, the rows of this 3-by-6 matrix are transformed in the way - known from the highschool - used to simplify the calculation of determinants: a row may be multiplied or divided by a non-zero number, two rows may be interchanged, or a row multiplied by a (non-zero) number may be added to another row. The aim is to obtain the identity matrix, in our case I_3, instead of the left block A. When this aim is reached, the desired matrix of Eq. (1.94) will appear as the right block, instead of matrix B. Hence, the scheme of these transformations is

$$[A|B] \rightarrow \ldots \rightarrow [I_3|A^{-1}B = T^{\mathrm{T}}]. \tag{1.95}$$

With the data in (1.90-91) we have

$$[A|B] = \begin{bmatrix} 1 & 1 & 1 & | & 1 & 2 & 6 \\ 1 & 0 & 2 & | & 3 & 2 & 7 \\ 0 & 0 & 3 & | & 3 & 3 & 9 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 2 & | & 3 & 2 & 7 \\ 0 & 1 & -1 & | & -2 & 0 & -1 \\ 0 & 0 & 3 & | & 3 & 3 & 9 \end{bmatrix} \rightarrow$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 2 & | & 3 & 2 & 7 \\ 0 & 1 & -1 & | & -2 & 0 & -1 \\ 0 & 0 & 1 & | & 1 & 1 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & | & 1 & 0 & 1 \\ 0 & 1 & 0 & | & -1 & 1 & 2 \\ 0 & 0 & 1 & | & 1 & 1 & 3 \end{bmatrix}. \tag{1.96}$$

It follows from the last matrix in (1.96) that the transpose of the transformation matrix is

$$T^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} \Rightarrow T = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}. \tag{1.97}$$

For a better understanding of this result, let us explicitly write the connection between the vectors of basis and the ones of basis *B*, corresponding to the transformation matrix of (1.97):

$$\begin{cases} b_1 = a_1 - a_2 + a_3, \\ b_2 = \phantom{a_1 - {}} a_2 + a_3, \\ b_3 = a_1 + 2a_2 + 3a_3. \end{cases} \tag{1.98}$$

We are going to extend this example by considering a vector X expressed in basis A and looking for its coordinates in the "new" basis B of (1.91). Let us consider the vector

$$X = 2\,a_1 - 3\,a_2 + 5\,a_3. \tag{1.99}$$

$$(1.99) \Rightarrow \quad X_A = \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}. \tag{1.100}$$

The coordinates of X in the "new" basis B can be obtained by applying formula (1.78). The expression of X_B means that the vector of the "new" coordinates is the solution to the matrix equation in (1.81),

$$X_A^{\mathrm{T}} = X_B^{\mathrm{T}}\,T \iff T^{\mathrm{T}} X_B = X_A. \tag{1.101}$$

The last matrix equation in Eq. (1.101) is equivalent to a non-homogeneous linear system which can be solved by the Gaussian elimination, as in **Example 1.12**. The transformation scheme is

$$[T^{\mathrm{T}} \,|\, X_A] \rightarrow \ldots \rightarrow [I_3 \,|\, T^{-\mathrm{T}} X_A = X_B]. \tag{1.102}$$

From (1.97) and (1.100) – the first matrix – we have

$$[T^{\mathrm{T}} \,|\, X_A] = \left[\begin{array}{rrr|r} 1 & 0 & 1 & 2 \\ -1 & 1 & 2 & -3 \\ 1 & 1 & 3 & 5 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 0 & 1 & 2 \\ 0 & 1 & 3 & -1 \\ 0 & 1 & 2 & 3 \end{array}\right] \rightarrow$$

$$\rightarrow \left[\begin{array}{rrr|r} 1 & 0 & 1 & 2 \\ 0 & 1 & 3 & -1 \\ 0 & 0 & -1 & 4 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 0 & 0 & 6 \\ 0 & 1 & 0 & 11 \\ 0 & 0 & 1 & -4 \end{array}\right]. \tag{1.103}$$

Hence

$$X_B = \begin{bmatrix} 6 \\ 11 \\ -4 \end{bmatrix} \Rightarrow X = 6\,b_1 + 11\,b_2 - 4\,b_3. \tag{1.104}$$

The result in (1.104) can be checked as follows: the linear expression (1.99) of X in basis A leads to the vector X as an element in $\mathbb{R}^3$:

$$X = 2\,a_1 - 3\,a_2 + 5\,a_3 = 2\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - 3\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 5\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 12 \\ 15 \end{bmatrix}. \tag{1.105}$$

If we pass to the "new" basis B, the expression (1.104) with the vectors in (1.91) leads to

$$X = 6b_1 + 11b_2 - 4b_3 = 6\begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix} + 11\begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} - 4\begin{bmatrix} 6 \\ 7 \\ 9 \end{bmatrix} = \begin{bmatrix} 4 \\ 12 \\ 15 \end{bmatrix}. \tag{1.106}$$

Hence, the linear expressions in (1.105) & (1.106) represent the same vector $X \in \mathbb{R}^3$. □

The example is over. However, let us see another way to check that the coordinates and the expression of X in basis B, that occur in (1.104) are / is correct. This method can be employed even for such problems formulated in a general vector space and not necessarily in $\mathbb{R}^3$. It simply consists in taking the vectors b_i, expressed in basis A as in (1.98), what should give back the expression (1.99) of X in the "initial" basis A:

$$
\begin{aligned}
X = 6b_1 + 11b_2 - 4b_3 = 6(a_1 - a_2 + a_3) &+ \\
+ 11(a_2 + a_3) &- \\
-4(a_1 + 2a_2 + 3a_3) &= \\
= 2a_1 - 3a_2 + 5a_3 &.
\end{aligned}
$$

Hence, expression (1.99) in basis A of X has been retrieved.

§ 1.1-A VECTOR SPACES, BASES AND COORDINATES - APPLICATIONS

1-A.1 Study the linear dependence / independence of the four vectors in $\mathbb{R}^4$ given below, for the real values of parameter λ:

$$U_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \\ \lambda \end{bmatrix}, \; U_2 = \begin{bmatrix} 1 \\ \lambda \\ 0 \\ 1 \end{bmatrix}, \; U_3 = \begin{bmatrix} 3 \\ 2 \\ 2 \\ 3 \end{bmatrix}, \; U_4 = \begin{bmatrix} \lambda \\ 1 \\ 1 \\ \lambda \end{bmatrix}.$$

1-A.2 Find the coordinates X_A of $X = [\,0\ 0\ 0\ 1\,]^{\mathrm{T}}$ in the basis

$A:\ a_1 = [1\ 1\ 0\ 1]^{\mathrm{T}},\ a_2 = [1\ 1\ 0\ 0]^{\mathrm{T}},\ a_3 = [2\ 1\ 3\ 1]^{\mathrm{T}},\ a_4 = [0\ 1\ -1\ -1]^{\mathrm{T}},$

and then in basis $B:\ B^{\mathrm{T}} = T A^{\mathrm{T}}$ with $T = \begin{bmatrix} 3 & 5 & 3 & 0 \\ -1 & 4 & 4 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$

1-A.3 Find a linear expression of the vector $X = \begin{bmatrix} 5 \\ -2 \\ 12 \end{bmatrix}$ in terms of

$$a_1 = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}, \; a_2 = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}, \; a_3 = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}.$$ Is it unique ?

1-A.4 Show that the space $\mathrm{POL}_2(\mathbb{R})$ of the polynomials over the field $\mathbb{R}$ can be spanned by the set of "polynomial vectors" consisting of

$$p_1 = 2t, \; p_2 = t^2 + t, \; p_3 = t - 1$$

and find the coordinates of $p = t^2 - 5t$ and $q = 2t - 1$ in this basis.

1-A.5 Check that the column vectors of matrix

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

are linearly dependent, find a linear dependence relation among them and determine the dimension of and a basis in the vector space spanned by the four columns

$$A^1,\ A^2,\ A^3,\ A^4.$$

1-A.6 Determine the real values of parameter m such that vector $X = [4\ \ 3\ \ 4]^{\mathrm{T}}$ can be linearly expressed in terms of

$$U_1 = [m\ \ 1\ \ 1]^{\mathrm{T}},\ U_2 = [1\ \ 1\ \ 2]^{\mathrm{T}},\ \text{and } U_3 = [1\ \ 1\ \ 1]^{\mathrm{T}}.$$

Then find the coordinates of X in basis $A = [U_1\ U_2\ U_3]$.

1-A.7 Find a basis in the matrix space $\mathscr{M}_{2\times 3}(\mathbb{R})$, so that the coordinates of any matrix $A = [a_{ij}]_{2\times 3}$ coincide with its entries.

1-A.8 Is it true that if $v_1, v_2, v_3 \in V$ are linearly dependent, then the vectors $w_1 = v_1 + v_2,\ w_2 = v_1 + v_3,\ w_3 = v_2 + v_2$ are linearly dependent, too ?

(*Hint*: Assume some combination $\lambda_1 w_1 + \lambda_2 w_2 + \lambda_3 w_3 = 0$ and find which λ_i are possible).

1-A.9 Decide the dependence or independence of

(*a*) $(1,1,2),\ (1,2,1),\ (3,1,1) \in \mathbb{R}^3$;

(*b*) $v_1 - v_2,\ v_2 - v_3,\ v_3 - v_4,\ v_4 - v_1$ for any vectors $v_1, v_2, v_3, v_4 \in V$.

(*c*) $(1,1,0),\ (1,0,0),\ (0,1,1)$ and $(x,y,z) \in \mathbb{R}^3$ for any real numbers x, y, z.

1-A.10 In the space of 2-by-2 matrices, find a basis for the subset of matrices whose row sums and column sums are equal. Find five linearly independent 3-by-3 matrices with this property.

1-A.11 Decide whether it is *True* or *False* :

(*a*) If the columns of A are linearly independent, then equation $AX = b$ has exactly one solution for any $b \in \mathbb{R}^m$;

(*b*) A 5-by-7 matrix never has linearly independent columns.

1-A.12 Check which one(s) of the polynomials t^2 and $t-1$ are (is) in the space spanned by

$$\{t^3 - t + 1,\ 3t^2 + 2t,\ t^3\}.$$

1-A.13 A vector space is spanned by a basis A, that is $V = \mathcal{L}(A)$; find the coordinates of $x = -a_1 + 5a_2 + a_3$

in another basis $B = AT^{\mathrm{T}}$ where $T = \begin{bmatrix} 1 & 3 & 1 \\ 2 & 0 & -1 \\ -1 & 2 & 0 \end{bmatrix}$

and then check the result.

Hint: The method employed in the latest example (Example 1.13 at page 25) can be here applied.

1-A.14 The space of polynomials of order 3, that is $\mathrm{POL}_3(\mathbb{R})$, can be spanned by its standard basis

$$E = \{1, t, t^2, t^3\}.$$

Check that the family

$$B = \{1, (t-1), (t-1)^2, (t-1)^3\}$$

is also a basis for this space. Find the transformation matrix from E to B and find the coordinates of the polynomial $p = 3 - 2t + 4t^3$ in both these bases. Check the coordinates in basis B .

§ 2.2 SUBSPACES OF VECTOR SPACES

The notion of *algebraic structure* was met in the highschool, together with the one of *substructure*. For instance, a group may admit one or several *subgroups*. Moreover, a rather general structure may admit substructures with *richer properties*. As a typical example, the ring (with identity) of the square matrices $\mathscr{M}_n(\mathbf{F})$ admits the subring of *nonsingular* matrices, but the latter one (augmented with the zero matrix **0**) is a non-commutative *field*, since any nonsingular matrix is invertible (as we saw in **§ 1.2**). Since the basic structure of the LINEAR ALGEBRA is the *vector* (or *linear*) *space*, it is therefore natural to see what would mean a *substructure of a vector space*.

***Definition* 2.1.** Let V be a vector space over a field $\mathbf{F}$. A subset $W \subseteq V$ is a *subspace* of V if W itself is a vector space over the same field (under the vector sum and the multiplication of vectors by scalars in $\mathbf{F}$). ♢

It follows that a simple subset $W \subseteq V$ is not necessarily a subspace. The ten axioms of *Def. 1.1* $(\mathbf{L}_{1,\ldots,10})$ should be satisfied on W, too. However, let us remark that *it suffices only two* of the ten axioms to be satisfied, namely $(\mathbf{L}_1)$ and $(\mathbf{L}_6)$, , as stated in

PROPOSITION 2.1. $\langle W, \mathbf{F}\,;\, +, \lambda\, w \rangle$ *with* $W \subseteq V$ *is a subspace of* V *iff*

$(\mathbf{L}_1)$ $(\forall\, u, w \in W)$ $u + w \in W$, *and*

$(\mathbf{L}_6)$ $(\forall\, \lambda \in \mathbf{F})(\forall\, w \in W)$ $\lambda\, w \in W$.

Proof. We have to show that the other eight axioms of *Def. 1.1* are also satisfied on W. But some of them are "inherited" from the corresponding properties of the sum and multiplication by scalars satisfied on the entire space $\langle V, \mathbf{F}\,;\, +, \lambda\, x \rangle$: this is the case with the associativity $(\mathbf{L}_2)$ and commutativity $(\mathbf{L}_5)$ of the vector sum, as well as with axioms $(\mathbf{L}_7)$, $(\mathbf{L}_8)$, $(\mathbf{L}_9)$ and $(\mathbf{L}_{10})$. The only axioms which have to be effectively checked are $(\mathbf{L}_3)$ and $(\mathbf{L}_4)$, since it would be (theoretically) possible that the zero vector $\mathbf{0} \in V$ be no more in the subset W; similarly, it would also be possible that the negative $-w$ be no more in W for some $w \in W$. But, taking $\lambda = -1 \in \mathbf{F}$ for $\forall\, w \in W$ and applying the consequence $(\mathbf{L}_{16})$ of PROPOSITION 1.1 (which holds for $\forall\, x \in V$, hence also for $\forall\, w \in W \subseteq V \Rightarrow w \in V)$, we derive from $(\mathbf{L}_6)$ that $(-1)\, w = -w \in W$ for $\forall\, w \in W$. As regards the defining property of the negative $-w:\ w + (-w) = (-w) + w = \mathbf{0}$, it is "inherited" from V. As regards the membership $\mathbf{0} \in W$, it readily follows from $-w \in W$ for $\forall\, w \in W$ and from $(\mathbf{L}_1)$ with u replaced by $w:\ w + (-w) \in W$, but $w + (-w) = 0$, hence $\mathbf{0} \in W$. This completes the proof. ■

In other words, a subset W of a vector space V is a subspace of V if and only if W *is closed* (or *stable*) under the two linear operations: the sum of vectors and the multiplication of vectors by scalars. Therefore, *Definition 3.1* may be replaced by a "minimal" and equivalent one :

***Definition 2.1'.* (Subspaces)** Let V be a vector space over a field $\mathbf{F}$. A subset $W \subseteq V$ is a *subspace* of V if axioms $(\mathbf{L}_1)$ & $(\mathbf{L}_6)$ are satisfied on W. ♢

If *Definition 2.1'* (hence *Definition 2.1*) is satisfied by a subset $W \subseteq V$, we use the special notation

$$\boxed{W \subseteq_{\text{subsp}} V.} \tag{2.1}$$

This notation should be read "W *is included as a subspace in the vector space* V ".

Remark 2.1. Conditions $(\mathbf{L}_1)$ & $(\mathbf{L}_6)$ for checking whether a subset $W \subseteq V$ is a *subspace* of V can be replaced by a single condition :

$$\boxed{(\forall\, \lambda_1, \lambda_2 \in \mathbf{F})\, (\forall\, w_1, w_2 \in W)\;\; \lambda_1 w_1 + \lambda_2 w_2 \in W.} \tag{2.2}$$

Indeed, (2.2) $\Rightarrow$ $(\mathbf{L}_1)$ for $\lambda_1 = \lambda_2 = 1$ & $w_1 = u$, $w_2 = w$ and applying $(\mathbf{L}_7)$, too. (2.2) $\Rightarrow$ $(\mathbf{L}_6)$ for $\lambda_1 = 0, \lambda_2 = \lambda$ and w_2 replaced by w (also applying consequence $(\mathbf{L}_{13})$ in PROPOSITION 1.1).

Conversely, if *Definition 2.1 (which is equivalent to* P. 2.1*)* is satisfied on W, then

$$(\forall\, \lambda, \mu \in \mathbf{F})(\forall\, u, w \in W)\; \lambda u \in W \wedge \mu w \in W$$

by $(\mathbf{L}_6)$ and $\lambda u + \mu w \in W$ **by** $(\mathbf{L}_1)$; the equivalence is thus proved. Hence, the preceding two definitions may be replaced by

Definition 2.1". Let V be a vector space over a field $\mathbf{F}$. A subset $W \subseteq V$ is a *subspace* of V iff (if and only if) the membership in (2.2) is satisfied. ♢

In fact, the property in (2.2) implies the axioms $(\mathbf{L}_1)$ & $(\mathbf{L}_6)$, but the scalars and the vectors are subscripted. The property in (2.2) means that

> ***A subset of a vector space is a subspace iff it is closed under arbitrary linear combinations*** (of two vectors in that subset).

This latter characterization of a vector subspace admits a generalization, given by

PROPOSITION 2.2. *If* $W \subseteq_{\text{subsp}} V$ *then*

$$\boxed{(\forall\, \lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbf{F})(\forall\, w_1, w_2, \ldots, w_m \in W) \sum_{i=1}^{m} \lambda_i w_i \in W.} \qquad (2.3)$$

and conversely.

Proof (by induction $/m$). Let us denote property (2.3) by $(\mathbf{P}_m)$. $(\mathbf{P}_2)$ is just (2.2). Let us assume that $(\mathbf{P}_m)$ is satisfied on W. Then

$$(\forall\, \lambda_1, \lambda_2 \ldots, \lambda_m, \lambda_{m+1} \in \mathbf{F})(\forall\, w_1, w_2, \ldots, w_m, w_{m+1} \in W)$$

$$\sum_{i=1}^{m+1} \lambda_i w_i = \sum_{i=1}^{m} \lambda_i w_i + \lambda_{m+1} w_{m+1} \underset{\text{not}}{=} u + v \in W \text{ by } (\mathbf{P}_2).$$

$$u = \sum_{i=1}^{m} \lambda_i w_i \in W \text{ by } (\mathbf{P}_m), \text{ and } v = \lambda_{m+1} w_{m+1} \in W \text{ by } (\mathbf{L}_6).$$

Hence $(\mathbf{P}_{m+1})$ is satisfied, and $(\mathbf{P}_m)$ thus holds for $\forall\, m \in \mathbb{N}$. The converse implication is quite obvious, for $m = 2$ with *Definition 2.1"*. ■

Before continuing with other definitions and results involving the subspaces, let us see a couple of examples of subspaces. In general, a subspace can be defined by specifying one or more properties of the elements (vectors) in the vector space it is a part of. But a subspace $W \subseteq_{\text{subsp}} V$ can also be characterized by identifying a *basis* that spans it : $W = \mathcal{L}(B)$. Implicitly, its dimension can also be determined.

Example 2.1. Let $W \subseteq \mathbb{R}^n$ be the subset of the vectors X with the first component $= 0$, that is

$$W = \{ X = [0 \;\; x_2 \; \ldots \; x_n]^{\mathrm{T}} : x_i \in \mathbb{R} \text{ for } i = 2, 3, \ldots, n \}. \qquad (2.4)$$

It is clear that $W \subset \mathbb{R}^n$, and it easy to check condition (2.2) for any two vectors $X, Y \in W$ since the first component of $\lambda X + \mu Y$ will also be $= 0$, hence $\lambda X + \mu Y \in W$. A basis spanning W consists of $e_2, e_3, \ldots, e_n$ – the vectors of the standard basis E of $\mathbb{R}^n$ except the first one. Hence, the dimension of W is $= n - 1$. This example may be obviously extended by defining W as the subset of vectors X with the components at certain (but fixed) positions $= 0$. □

Example 2.2. Let us consider a subset of the space of square matrices of order 2, $\mathcal{M}_2(\mathbb{R})$ defined as

$$W = \left\{ \begin{bmatrix} a & b \\ c & a \end{bmatrix} :\ a, b, c \in \mathbb{R} \right\}.$$

It is easy to see that W is spanned by the basis consisting of three matrices, namely

$$B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ B_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\ B_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

since $(\forall\, M \in W)\ M = aB_1 + bB_2 + cB_3$. Hence, $\dim W = 3$. □

Example 2.3. The set of polynomials over $\mathbb{R}$, $\mathrm{POL}_{n,\,\text{even}}(\mathbb{R})$, whose terms with odd powers are $= 0$, is a subspace of $\mathrm{POL}_n(\mathbb{R})$. The reader is invited to prove this assertion and show that the dimension of this subspace equals $[\frac{n}{2}] + 1$, where $[x]$ denotes the integer part of the real number x. □

A rather general and important example of subspace is given in

PROPOSITION 2.3. *If $\mathcal{A} = \{u_1, u_2, \ldots, u_m\}$ is a family of vectors in the vector space V, then the set spanned by A, $\mathcal{L}(A)$ is a subspace of V.*

Proof. Let us consider two arbitrary scalars $\lambda, \mu \in \mathbf{F}$ and two vectors in the set $\mathcal{L}(\mathcal{A})$:

$$v = \sum_{i=1}^{m} \alpha_i u_i,\ w = \sum_{i=1}^{m} \beta_i u_i. \tag{2.5}$$

A linear combination with the two vectors in (2.5) gives

$$\lambda v + \mu w = \lambda \sum_{i=1}^{m} \alpha_i u_i + \mu \sum_{i=1}^{m} \beta_i u_i = \sum_{i=1}^{m} \lambda(\alpha_i u_i) + \sum_{i=1}^{m} \mu(\beta_i u_i) =$$

$$= \sum_{i=1}^{m} (\lambda\alpha_i) u_i + \sum_{i=1}^{m} (\mu\beta_i) u_i = \sum_{i=1}^{m} (\lambda\alpha_i + \mu\beta_i) u_i = \sum_{i=1}^{m} \gamma_i u_i; \tag{2.6}$$

it follows from (2.6) that any linear combination of two vectors in $\mathcal{L}(\mathcal{A})$ is also in this set, hence $\mathcal{L}(\mathcal{A})$ is a subspace of V. ■

Remark 2.2. It follows from the proof of PROPOSITION 3.1 that *any subspace of a vector space V includes the zero vector* $\mathbf{0}$. The set consisting of this zero vector, that is $\{\mathbf{0}\}$, is just the "least" subspace of any vector space. Hence, any subspace W satisfies the double inclusion

$$\boxed{\{\mathbf{0}\} \subseteq W \subseteq V.} \tag{2.7}$$

Thus, any vector space V admits two improper subspaces : the zero subspace $\{\mathbf{0}\}$ and the space V itself. We denote by SUBSP_V the set of all subspaces of V. The fact that the trivial subspace $\{\mathbf{0}\}$ is actually a subspace can be easily verified according to *Def.3.1'* : $\mathbf{0}+\mathbf{0}=\mathbf{0}\in\{\mathbf{0}\}$, according to $(\mathbf{L}_3)$; $(\mathbf{L}_6)$: $(\forall\,\lambda\in\mathbf{F})\ \lambda\,\mathbf{0}=\mathbf{0}\in\{\mathbf{0}\}$ according to $(\mathbf{L}_{14})$ in PROPOSITION 2.1.

As regards the dimension of a subspace, let us state

PROPOSITION 2.3. *If* $W\subseteq_{\text{subsp}}V$ *then* $\dim W\le\dim V.$ (2.8)

Proof. In the trivial case when $W=V$, the inequality in (2.8) obviously becomes an equality: $\dim W=\dim V$. If $W\subset_{\text{subsp}}V$ (W is a proper subspace of V), let us denote $\dim W=m$, $\dim V=n$. Assume that $m=n$ and let $B=\{b_1,b_2,\ldots,b_m\}$ be a basis of W. But it follows from the hypothesis that B is also a basis in V; hence, any vector $x\in V$ can be linearly expressed in terms of B:

$$x=\sum_{i=1}^{m}\xi_i b_i\ \Rightarrow\ x\in W.$$

Thus we get the inclusion $V\subseteq W$, what contradicts $W\subset_{\text{subsp}}V$; therefore $m=n$ is impossible and it follows that $\dim W<\dim V$ if $W\subset_{\text{subsp}}V$. The proof is over. ■

Several operations can be defined on the subspaces of a vector space V. Since the subspaces are subsets (parts) of V, the set-theoretic operations as the set union and set intersection are naturally possible in $\mathbf{SUBSP}_V$. But another (specific) operation can also be defined :

***Definition* 2.2.** Let $W_1,W_2\subseteq_{\text{subsp}}V$. Then

(i) $W_1+W_2\underset{\text{def}}{=}\{w_1+w_2:\ w_1\in W_1\wedge w_2\in W_2\}$;

(ii) $W_1\cup W_2\underset{\text{def}}{=}\{u:\ u\in W_1\vee u\in W_2\}$;

(iii) $W_1\cap W_2=_{\text{def}}\{u:\ u\in W_1\wedge u\in W_2\}$. ♢

Operations (ii) and (iii) should not be explained. As regards (i), the *sum of two subspaces* is simply the sets of sums of two vectors, each of them in one of the subspaces, respectively. The results of these three operations with subspaces could be a subspace or merely a subset of V. The problem is established in

PROPOSITION 2.4. *If* $W_1,W_2\subseteq_{\text{subsp}}V$ *then* :

(i) $W_1+W_2\subseteq_{\text{subsp}}V$;

(ii) $W_1\cap W_2\subseteq_{\text{subsp}}V$;

(iii) $W_1\cup W_2\subseteq_{\text{not-subsp}}V$.

Proofs. (*i*) Let us take two arbitrary scalars λ, μ in the field $\mathbf{F}$ and two vectors in $W_1 + W_2$: $u_1 + u_2$ and $w_1 + w_2$. The corresponding linear combination thereof is

$$\lambda(u_1 + u_2) + \mu(w_1 + w_2) = \lambda u_1 + \lambda u_2 + \mu w_1 + \mu w_2 =$$
$$= (\lambda u_1 + \mu w_1) + (\lambda u_2 + \mu w_2) \in W_1 + W_2$$

since $W_1, W_2 \subset_{\text{subsp}} V$ and *Definition 2.1*" gives the membership to W_1, W_2 of the vectors between parentheses in the rightmost side of the above equation. *Def. 2.1* of a vector subspace has also been involved.

(*ii*) Let us take two arbitrary scalars $\lambda, \mu \in \mathbf{F}$ and two vectors in the intersection : $u, v \in W_1 \cap W_2$. Since

$$W_1 \cap W_2 \subseteq W_1, W_2 \subseteq_{\text{subsp}} V,$$

it follows that

$$\lambda u + \mu v \in W_1 \ \& \ \lambda u + \mu v \in W_2 \Rightarrow \lambda u + \mu v \in W_1 \cap W_2.$$

According to *Definition 2.1"*, property (*ii*) holds, too. The proof is over.

(*iii*) The union of two subspaces W_1, W_2 of a vector space V is clearly a subset of its, but *it is not necessarily a subspace*. Let us take two vectors in $W_1 \cup W_2$, namely $u \in W_1 \setminus W_2$ and $v \in W_2 \setminus W_1$ (provided these two set differences are not empty). Obviously, $u, v \in W_1 \cup W_2 \Rightarrow u + v \in W_1 \cup W_2$ if $W_1 \cup W_2$ were a subspace. But $u \notin W_2$ and $v \notin W_1$; hence the sum of the two vectors could be not in the union. A simple example would better prove this possibility. Let us take two subspaces of $\mathbb{R}^2$, namely

$$W_1 = \left\{\begin{bmatrix} 0 \\ b \end{bmatrix} : b \in \mathbb{R}\right\}, \ W_2 = \left\{\begin{bmatrix} a \\ 0 \end{bmatrix} : a \in \mathbb{R}\right\};$$

it follows that their union $W_1 \cup W_2$ consists of vectors *with at least one component* $= 0$. But let us now consider

$$u = \begin{bmatrix} 0 \\ b \end{bmatrix} \text{ with } b \neq 0 \ \& \ v = \begin{bmatrix} a \\ 0 \end{bmatrix} \text{ with } a \neq 0;$$

We clearly have $w = u + v = \begin{bmatrix} 0 \\ b \end{bmatrix} + \begin{bmatrix} a \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$ with $a, b \neq 0$.

Hence the sum $w = u + v \notin W_1 \cup W_2$ as it should be if $W_1 \cup W_2$ were a subspace. ■

Remark 2.2. Certainly, the three operations introduced by *Definition 2.2* can be extended to sums, intersections and unions of several subspaces, and the properties of PROPOSITION 3.4 still hold :

PROPOSITION 2.5. *If* $(W_i)_{i \in I} \subseteq \text{SUBSP}_V$ *then* :

(*i*) $\sum_{i \in I} W_i \in \text{SUBSP}_V$;

(*ii*) $\bigcap_{i \in I} W_i \in \text{SUBSP}_V$.

(*iii*) $\bigcup_{i \in I} W_i \notin \text{SUBSP}_V$;

The proofs of these properties follow by induction on $m = \text{card}\, I$ (if I is a finite family of indices) and from PROPOSITION 2.4 ; we do not give them here (leaving them as exercises to the reader). As regards property (*iii*) above, it has be understood as in the proof of PROPOSITION 2.4 : the union of several subspaces is <u>not necessaril</u>y a subspace of V, but only a subset. It is, however, possible that the union of certain p<u>articular</u> subspaces gives a subspace.

The sum of two subspaces introduced in *Def. 2.2* - (*i*) gives rise to an interesting problem: a (fixed) vector $x \in W_1 + W_2$ may be obviously written, by definition, as

$$x = x_1 + x_2 \text{ with } x_2 \in W_2 \text{ ;} \tag{2.9}$$

but <u>when</u> is a decomposition like (3.9) unique ? The answer is given by

PROPOSITION 2.6. *If* $W_1, W_2 \subseteq_{\text{subsp}} V$ *then the decomposition* $x = x_1 + x_2 \in W_1 + W_2$ *with* $x_1 \in W_1 \wedge x_2 \in W_2$ *is unique if and only if*

$$W_1 \cap W_2 = \{\mathbf{0}\}. \tag{2.10}$$

Proof. $(\Rightarrow)$: Assume that decomposition (2.9) is unique but – however – $W_1 \cap W_2 \neq \{\mathbf{0}\}$. Let us take a vector $u \in W_1 \cap W_2$, $u \neq \mathbf{0}$. Then we obtain <u>two different</u> decompositions of x, namely

$$x = x_1 + x_2 = x_1 + \mathbf{0} + x_2 = x_1 - u + u + x_2 \in W_1 + W_2 \text{;}$$

the latter membership follows from $u \in W_1 \cap W_2 \subseteq W_1, W_2 \subseteq_{\text{subsp}} V$. But $u \neq \mathbf{0}$, hence $x_1 - u \neq x_1$ and $x_2 + u \neq x_2$. Thus, decomposition (2.9) would be not unique !

$(\Leftarrow)$: Let us now assume that condition (2.10) holds, but the decomposition (2.9) would be not unique :

$$x = x_1 + x_2 = y_1 + y_2 \text{ with } x_1, y_1 \in W_1 \text{ and } y_2, x_2 \in W_2. \tag{2.11}$$

It follows from the double decomposition of x in (2.11) that

$$x_1 - y_1 = y_2 - x_2 \underset{\text{not}}{=} u \in W_1 \cap W_2 \tag{2.11'}$$

since both W_1 and W_2 are subspaces of V. But the membership in (2.11') and

(2.10) imply that $x_1 - y_1 = y_2 - x_2 = \mathbf{0} \Rightarrow x_1 = y_1 \wedge x_2 = y_2$. Hence the decomposition of x in (3.10) is – in fact – unique. ■

***Definition* 2.3.** Let $W_1, W_2 \subseteq_{\text{subsp}} V$. If condition (2.10) satisfied, then the sum of the two subspaces is said to be *direct* and it is denoted by $\oplus$:

$$W_1 + W_2 \underset{\text{not}}{=} W_1 \oplus W_2 \Longleftrightarrow W_1 \cap W_2 = \{\mathbf{0}\}. \tag{2.12}$$

The direct sum of several subspaces may also be defined :

$$W_1 \oplus W_2 \oplus \ldots \oplus W_m \underset{\text{not}}{=} \bigoplus_{i=1}^{m} W_i \Longleftrightarrow [i \neq j \Rightarrow W_i \cap W_j = \{\mathbf{0}\}]. \tag{2.13}$$

If $W_1, W_2 \subseteq_{\text{subsp}} V$ and $W_1 \oplus W_2 = V$ then the subspaces W_1, W_2 are said to be *supplementary* (with respect to V).

An interesting relation between the dimensions of $W_1, W_2, W_1 \cap W_2$ and $W_1 + W_2$ is given by

THEOREM 2.1. (Grassmann) *If* $W_1, W_2 \subseteq_{\text{subsp}} V$ *then*

$$\boxed{\dim W_1 + \dim W_2 = \dim(W_1 \cap W_2) + \dim(W_1 + W_2).} \tag{2.14}$$

Proof. Let

$$\dim W_1 = m,\ \dim W_2 = n,\ \dim(W_1 \cap W_2) = k, \tag{2.15}$$

and let $B = \{b_1, b_2, \ldots, b_k\} \subset W_1 \cap W_2$ be a basis. We complete B up to a basis of W_1 : $B_1 = \{b_1, b_2, \ldots, b_k, c_{k+1}, \ldots, c_m\}$ of W_1. Since $B \subset W_2$ too, it can also be completed up to a basis $B_2 = \{b_1, b_2, \ldots, b_k, d_{k+1}, \ldots, d_m\}$ of W_2. Let us show that the family

$$B_+ = \{b_1, b_2, \ldots, b_k, c_{k+1}, \ldots, c_m, d_{k+1}, \ldots, d_n\} \tag{2.16}$$

is a basis of $W_1 + W_2$. For any $x = x_1 + x_2 \in W_1 + W_2$ we have the linear expressions of x_1, x_2 in the two bases of W_1, W_2 :

$$x_1 = \lambda_1 b_1 + \lambda_2 b_2 + \ldots + \lambda_k b_k + \lambda_{k+1} c_{k+1} + \ldots + \lambda_m c_m, \tag{2.17}$$

$$x_2 = \mu_1 b_1 + \mu_2 b_2 + \ldots + \mu_k b_k + \mu_{k+1} d_{k+1} + \ldots + \mu_n d_n. \tag{2.18}$$

It follows from (2.17) & (2.18) that

$$x = x_1 + x_2 = (\lambda_1 + \mu_1) b_1 + (\lambda_2 + \mu_2) b_2 + \ldots + (\lambda_k + \mu_k) b_k + \\ + \lambda_{k+1} c_{k+1} + \ldots + \lambda_m c_m + \mu_{k+1} d_{k+1} + \ldots + \mu_n d_n.$$

Hence $x \in \mathcal{L}(B_+)$ and B_+ of (2.16) is thus a spanning family for $W_1 + W_2$.

Let us show that B_+ is linearly independent. The defining equation is

$$\beta_1 b_1 + \ldots + \beta_k b_k + \gamma_{k+1} c_{k+1} + \ldots + \gamma_m c_m + \delta_{k+1} d_{k+1} + \ldots + \delta_n d_n = \mathbf{0}. \quad (2.19)$$

If

$$\beta_1 b_1 + \ldots + \beta_k b_k + \gamma_{k+1} c_{k+1} + \ldots + \gamma_m c_m = \mathbf{0}$$

then $\beta_1 = \ldots = \beta_k = \gamma_{k+1} = \ldots = \gamma_m = 0$ since B_1 is independent, and (2.19) with the next equations imply

$$\delta_{k+1} d_{k+1} + \ldots + \delta_n d_n = \mathbf{0} \Rightarrow \delta_{k+1} = \ldots = \delta_n = 0$$

because $\{d_{k+1}, \ldots, d_n\}$ is also independent as a subfamily of basis B_2. Let us now assume that $\beta_1 b_1 + \beta_2 b_2 + \ldots + \beta_k b_k + \gamma_{k+1} c_{k+1} + \ldots + \gamma_m c_m = x \neq \mathbf{0}$. Since $x \in W_1$ and Eq. (2.19) $\Rightarrow$

$$\Rightarrow -x = \delta_{k+1} d_{k+1} + \ldots + \delta_n d_n \text{ or}$$

$$x = 0\, b_1 + \ldots + 0\, b_k + (-\delta_{k+1})\, d_{k+1} + \ldots (-\delta_n)\, d_n,$$

it follows that $x \in W_2$, too. Hence

$$x \in W_1 \cap W_2 \Rightarrow x = \alpha_1 b_1 + \ldots + \alpha_k b_k.$$

From this expression of x and from the former expression of $-x$ we get

$$\beta_1 b_1 + \ldots + \beta_k b_k = -(\delta_{k+1} d_{k+1} + \ldots + \delta_n d_n) \Rightarrow \text{(in view of (2.15))}$$

$$\beta_1 b_1 + \ldots + \beta_k b_k + \delta_{k+1} d_{k+1} + \ldots + \delta_n d_n = \mathbf{0} \Rightarrow \text{(since } B_2 \text{ is independent)}$$

$$\beta_1 = \ldots = \beta_k = \delta_{k+1} = \ldots = \delta_n = 0 \Rightarrow x = \mathbf{0},$$

what contradicts the assumption that $x \neq \mathbf{0}$!

Therefore B_+ is a basis in $W_1 + W_2 \Rightarrow$

$$\Rightarrow \dim(W_1 + W_2) = k + (m - k) + (n - k) = m + n - k \Rightarrow (2.14).$$

■

***Remark* 2.3.** If the sum $W_1 + W_2$ is direct, then – according to (2.12) and to the obvious property $\dim \{\mathbf{0}\} = 0$ – we get the equation

$$\dim (W_1 \oplus W_2) = \dim W_1 + \dim W_2. \quad (2.20)$$

We close this section with some more examples of subspaces. The first of the next examples is very often met in applications :

Example 2.4. If $AX = \mathbf{0}$ is a homogeneous system with nontrivial solutions, that is (see **§ 1.2**), $\operatorname{rank} A = r < n$, then its set of solutions S is a subspace of $\mathbb{R}^n$. This readily follows from (2.32) in *Remark 2.4* of **§ 1.2**, from definition (2.31) of the set S and from *Def. 3.1"* of a subspace :

$$X_1, X_2 \in S \Rightarrow A(\lambda_1 X_1 + \lambda_2 X_2) = \lambda_1 A X_1 + \lambda_2 A X_2 =$$
$$= \lambda_1 \mathbf{0} + \lambda_2 \mathbf{0} = \mathbf{0} + \mathbf{0} \Rightarrow \lambda_1 X_1 + \lambda_2 X_2 \in S \Rightarrow S \subseteq_{\text{subsp}} \mathbb{R}^n.$$

Moreover, its dimension equals $n - r$, since – according to (2.43) in **§ 1.2** – any solution X of system (2.31) depends on $n - r$ parameters and therefore S is linearly spanned by the $n - r$ columns of the block $\overline{A}_{12}$. □

Example 2.5. If $A \in \mathcal{M}_{m,n}(\mathbb{R})$ then the *row* (*sub*)*space* of A is the set spanned by its rows $A_1, A_2, \ldots, A_m$ while the *column* (*sub*)*space* of A is the set spanned by its columns $A^1, A^2, \ldots, A^n$. Let us denote these sets by ROWSP_A & COLSP_A, respectively. The fact that

$$\text{ROWSP}_A \subseteq_{\text{subsp}} \mathbb{R}^n \ \& \ \text{COLSP}_A \subseteq_{\text{subsp}} \mathbb{R}^m \tag{2.21}$$

immediately follows from PROPOSITION 3.3. The dimension of both these subspaces is obviously equal to the rank of the matrix A; formally

$$\text{ROWSP}_A = \text{L}(A_1, A_2, \ldots, A_m), \ \text{COLSP}_A = \text{L}(A^1, A^2, \ldots, A^n) \ \&$$
$$\dim \text{ROWSP}_A = \dim \text{COLSP}_A = r = \text{rank}\, A. \tag{2.22}$$

□

With reference to the preceding example, the set of solutions of a homogeneous system may be written as

$$S = \text{COLSP}_{B_{12}} \subseteq_{\text{subsp}} \mathbb{R}^n \text{ where } B_{12} = \begin{bmatrix} \overline{A}_{12} \\ I_{n-r} \end{bmatrix}. \tag{2.23}$$

Let us also illustrate the row and column (sub)spaces of a matrix A, and also the two operations on subspaces giving a subspace (the sum and the intersection) by a numerical example :

Example 2.6. Given the matrix

$$A = \begin{bmatrix} 1 & -1 & 2 & 3 & 4 \\ 2 & 1 & -1 & 2 & 0 \\ -1 & 2 & 1 & 1 & 3 \end{bmatrix}, \tag{2.24}$$

find a basis spanning each of the subspaces ROWSP_A and COLSP_A, respectively. Then write the general form of a vector in each of them.

In order to find the required bases in the subspaces spanned by the rows and (respectively) the columns of matrix A, let us recall that any such a basis will consist of r rows / columns, with $r = \text{rank}\, A$. Moreover, the method of RANK PRESERVING TRANSFORMATIONS (see (2.36) in **§ 1.2**) gives the possibility to identify r independent rows and columns giving these bases: they will be the rows / columns passing through the triangular block $\overline{A}_{11}$. Therefore, we have to transform the matrix in (2.24) until reaching a quasi-triangular form :

$$A = \begin{bmatrix} 1 & -1 & 2 & 3 & 4 \\ 2 & 1 & -1 & 2 & 0 \\ -1 & 2 & 1 & 1 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -1 & 2 & 3 & 4 \\ 0 & 3 & -5 & -4 & -8 \\ 0 & 1 & 3 & 4 & 7 \end{bmatrix} \rightarrow$$

$$\rightarrow \begin{bmatrix} 1 & -1 & 2 & 3 & 4 \\ 0 & 1 & 3 & 4 & 7 \\ 0 & 0 & -14 & -16 & -29 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 5 & 7 & 11 \\ 0 & 1 & 3 & 4 & 7 \\ 0 & 0 & 1 & 8/7 & 29/14 \end{bmatrix} \rightarrow$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 5 & 7 & 11 \\ 0 & 1 & 0 & 4/7 & 11/14 \\ 0 & 0 & 1 & 8/7 & 29/14 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 9/7 & 9/14 \\ 0 & 1 & 0 & 4/7 & 11/14 \\ 0 & 0 & 1 & 8/7 & 29/14 \end{bmatrix}.$$

The quasi-triangular form is just the third matrix in the chain above, and it is a triangular matrix of order 3 ; it follows that the rank of A is $= 3$. Moreover, the rows and the columns corresponding to this block

$$\bar{A}_{11} = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & -14 \end{bmatrix} \tag{2.25}$$

give two possible bases :

$$B_{\text{ROWSP}} = [A_1 \; A_2 \; A_3]^{\text{T}}, \; B_{\text{COLSP}} = [A^1 \; A^2 \; A^3]. \tag{3.26}$$

Therefore, the general form of a vector in ROWSP_A is $\lambda_1 A_1 + \lambda_2 A_2 + \lambda_3 A_3 =$

$$= \lambda_1 [1 \;\; -1 \;\; 2 \;\; 3 \;\; 4] + \lambda_2 [2 \;\; 1 \;\; -1 \;\; 2 \;\; 0] + \lambda_3 [-1 \;\; 2 \;\; 1 \;\; 1 \;\; 3] =$$

$$= [\lambda_1 + 2\lambda_2 - \lambda_3 \quad -\lambda_1 + \lambda_2 - \lambda_3 \quad 2\lambda_1 - \lambda_2 + \lambda_3 \quad 3\lambda_1 + 2\lambda_2 + \lambda_3 \quad 4\lambda_1 + 3\lambda_3].$$

The general form of a vector in COLSP_A is

$$\mu_1 A^1 + \mu_2 A^2 + \mu_3 A^3 = \ldots \underset{(2.25)}{=} \begin{bmatrix} \mu_1 - \mu_2 + 2\mu_3 \\ 2\mu_1 + \mu_2 - \mu_3 \\ -\mu_1 + 2\mu_2 + \mu_3 \end{bmatrix}.$$

The last (quasi-diagonal) matrix in the above chain gives the expression of the last two columns in (2.25) in the corresponding basis of (2.26), since the transformations have been applied on the rows only :

$$A^4 = \frac{1}{7}(9A^1 + 4A^2 + 8A^3), \; A^5 = \frac{1}{14}(9A^1 + 11A^2 + 29A^3).$$

□

The preceding example illustrates a way to establish the membership of some vector in $\mathbb{R}^m$ to the subspace spanned by other vectors in the same space. In fact, in the spanning vectors are $U_1, U_2, \ldots, U_k \in \mathbb{R}^m$ and the candidate vectors for the membership to

$$\mathcal{L}(U_1, U_2, \dots, U_k) \ \text{ are } \ V_1, V_2, \dots, V_\ell \in \mathbb{R}^m \tag{2.27}$$

(all of them written as columns), then this problem is equivalent to the simultaneous solution of nonhomogeneous linear systems with the same coefficient matrix, $A = [U_1 \ U_2 \ \dots \ U_k]$ and with the matrix of the free terms $B = [V_1 \ V_2 \ \dots \ V_\ell]$. This latter problem was discussed (in some detail) in **§ 1.2** - pages 50-51. Hence, the problem of deciding whether

$$V_j \in \mathcal{L}(U_1, U_2, \dots, U_k) \underset{\text{not}}{=} W \ \text{ for } \ j = 1, 2, \dots, \ell \tag{2.28}$$

or not can be reduced to the study and solution of several systems with the augmented matrix $\tilde{A} = [A \,|\, B]$. If $\operatorname{rank} A = m$ (what is possible only for $k \geq m$) then the answer is always positive: the multiple solution of the system

$$A X = B \tag{2.29}$$

will give – on the columns of $X = [X^1 \dots X^j \dots X^\ell]$ – the (coefficients of the) linear expression of the vectors $V_1, V_2, \dots, V_\ell$ in terms of $U_1, U_2, \dots, U_k$. This expression will be unique iff (that is, $\Longleftrightarrow$) $k = m = \operatorname{rank} A$. But, in this case, A is just a basis for the space $\mathbb{R}^m$. The more interesting case is $\operatorname{rank} A = r < m$. In this situation (see **§ 1.2**) some of vectors $V_1, V_2, \dots, V_\ell$ may be in W of (2.28) while others may be outside, that is in $\mathbb{R}^m \setminus W$. But the expressions of the vectors inside W will be not unique: they will depend on $k - r$ parameters (for $k > r$), $\mu_1, \mu_2, \dots, \mu_{k-r}$. Similar problems can occur when it is required to find a basis spanning the sum or the intersection of two subspaces.

Example 2.7. Let us consider two subspaces $U \,\&\, W$ of $\mathbb{R}^3$ respectively spanned by

$$A: \, a_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \, a_2 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \ \& \ B: \, b_1 = \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix}, \, b_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \tag{2.30}$$

The general expression of a vector in the sum $U \,\&\, W$ is

$$X = \alpha_1 a_1 + \alpha_2 a_2 + \beta_1 b_1 + \beta_2 b_2 = \alpha_1 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \beta_1 \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix} + \beta_2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \tag{2.31}$$

It is obvious that not all of the four vector in (2.30-31) are independent since their number is $4 > \dim \mathbb{R}^3 = 3$ and it can also be seen that

$$b_1 = 2\,a_2 = \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix} \;\&\; b_2 = a_1 + a_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \tag{2.32}$$

Moreover, the dependence relations (2.32) $\Rightarrow$ $[a_1\ \ b_1]$ or $[a_1\ \ a_2]$ are bases for $U + W \Rightarrow$ $\dim(U + W) = 2$. It is also clear that the dimensions of the two subspaces are both = 2 since the generating families in (2.30) are independent.

As regards the *intersection* of the two subspaces, we have to look for a vector

$$Y \in U \cap W \Rightarrow Y = \alpha_1 a_1 + \alpha_2 a_2 = \beta_1 b_1 + \beta_2 b_2 . \tag{2.33}$$

This (latter) vector equation (2.33) is – in fact – a homogeneous system in the unknown vector $[\alpha_1\ \ \alpha_2\ \ \beta_1\ \ \beta_2]^T$. The (coefficient) matrix of this system is

$$[a_1\ a_2\ -b_1\ -b_2] = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 2 & -1 & 2 & -1 \\ 0 & 1 & -2 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -2 & -1 \\ 2 & 0 & 0 & -2 \end{bmatrix} \sim$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -2 & -1 \end{bmatrix} \Rightarrow$$

$$\Rightarrow \text{(for } \beta_1 \underset{\text{not}}{=} \beta \;\&\; \beta_2 \underset{\text{not}}{=} \gamma) \;\; Y = \gamma \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + (2\,\beta + \gamma) \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} \gamma \\ -2\,\beta + \gamma \\ 2\,\beta + \gamma \end{bmatrix} = \beta \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix} + \gamma \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \beta_1 b_1 + \beta_2 b_2 . \tag{2.34}$$

We have thus retrieved, in view of the rightmost side of (2.34), the second expression of a vector Y in the intersection written in (2.33), what means that

$$U \cap W = \mathcal{L}\,(B) = W. \tag{2.35}$$

But a possibility for checking the general expression of (2.34) of a vector in the intersection subspaces is available if we use the solution of the homogeneous system that has led to (3.34), with the vectors of basis B:

$$Y = \beta_1 b_1 + \beta_2 b_2 \underset{(2.30)}{=} \beta \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix} + \gamma \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \Rightarrow \ldots \Rightarrow (2.34).$$

Therefore, the intersection has been found to be equal to the second subspace W, that is

$$U \cap W = W \Rightarrow W \subseteq U. \tag{2.36}$$

But it can be easily checked that the converse inclusion to that of (2.36) also holds. This follows from the fact that the two bases can be replaced by each other. Indeed,

$$(2.32) \Rightarrow \mathscr{L}(B) \subseteq \mathscr{L}(A) = U \Rightarrow W \subseteq U.$$

But the same relationship in (2.32) allows to express the vectors of A in terms of (those in) B :

$$(2.32) \Rightarrow a_2 = \frac{1}{2} b_1 \ \& \ a_1 = b_2 - a_1 = b_2 - \frac{1}{2} b_1 .$$

Hence $U = W \Rightarrow U \cap W = U \cup W = U = W$. This is a rather strange case, but it follows, from the above discussion, that Grassmann's formula is satisfied by these two subspaces :

$$\dim U = 2, \ \dim W = 2, \ \dim (U + W) = 2, \ \dim(U \cap W) = 2.$$

□

§ 2.2-A SUBSPACES OF VECTOR SPACES - APPLICATIONS

2 – A.1 There are considered, in the space $\mathbb{R}^4$, the vectors

$$u_1 = \begin{bmatrix} 2 \\ 4 \\ 1 \\ 3 \end{bmatrix}, \ u_2 = \begin{bmatrix} 7 \\ 4 \\ -9 \\ 5 \end{bmatrix}, \ u_3 = \begin{bmatrix} 4 \\ 8 \\ -3 \\ 7 \end{bmatrix}, \ u_4 = \begin{bmatrix} 5 \\ 5 \\ -5 \\ 5 \end{bmatrix}, \ u_5 = \begin{bmatrix} 8 \\ 4 \\ -14 \\ 6 \end{bmatrix}.$$

Which is $\dim W$ with W spanned by u_1, u_2, u_3, u_4, u_5 ? Choose a basis in W and express the other vectors in it.

2 – A.2 Determine a basis of the subspace of $\mathbb{R}^5$ spanned by the vectors

$$u_1 = \begin{bmatrix} 1 \\ 2 \\ -4 \\ 3 \\ 1 \end{bmatrix}, \ u_2 = \begin{bmatrix} 2 \\ 5 \\ -3 \\ 4 \\ 8 \end{bmatrix}, \ u_3 = \begin{bmatrix} 6 \\ 17 \\ -7 \\ 10 \\ 22 \end{bmatrix}, \ u_4 = \begin{bmatrix} 1 \\ 3 \\ -3 \\ 2 \\ 0 \end{bmatrix}.$$

2 – A.3 Check which of the polynomials t^2 and $t-1$ belong to the subspace generated by

$$\{t^3 - t + 1, \ 3t^2 + 2t, \ t^3\} \subset \mathrm{POL}_3(\mathbb{R}).$$

Then find a linear expression for that one(s) which are in this subspace.

2 – A.4 Prove that the subsets (of the corresponding vector spaces) given in Examples **2.4** thru **2.6** (in **§ 2.2**) are actually subspaces, and find bases spanning them.

2 – A.5 Determine $\dim W_{(\ell)}$ and a basis for each of $W_{(\ell)}$ $(\ell = \overline{i, ii})$ – the solution (sub)spaces of the homogeneous systems

$$\begin{cases} x_1 - 4x_2 + 3x_3 - x_4 = 0, \\ 2x_1 - 8x_2 + 6x_3 - 2x_4 = 0; \end{cases} \qquad \begin{cases} x_1 - 3x_2 + x_3 = 0, \\ 2x_1 - 6x_2 + 2x_3 = 0, \\ 3x_1 - 9x_2 + 3x_3 = 0. \end{cases}$$

2 – A.6 The subspaces $W_1, W_2 \subseteq \mathbb{R}^3$ are respectively spanned by

$$\mathcal{A} : a_1 = \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix}, \ a_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \ a_3 = \begin{bmatrix} 4 \\ 3 \\ 1 \end{bmatrix}$$

and $\mathcal{B} : b_1 = \begin{bmatrix} -1 \\ 3 \\ -4 \end{bmatrix}, \ b_2 = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$.

Find a basis in each of the subspaces $W_1, W_2, W_1 \cap W_2, W_1 \cup W_2$.

2 – A.7 Find a basis in each of ROWSP_A & COLSP_A where

$$A = \begin{bmatrix} 1 & -13 & 2 & 2 & 1 \\ 0 & 3 & 6 & 0 & -2 \\ 2 & -3 & -2 & 4 & 4 \\ 3 & -3 & 6 & 6 & 3 \\ 5 & -3 & 10 & 10 & 5 \end{bmatrix}.$$

2 – A.8 Show that the set of functions

$$W = \{ f : f(x) = a\cos x + b\sin x \ \text{ with } \ a, b \in \mathbb{R} \}$$

is a subspace of the space of real functions $\mathcal{F}_{\mathbb{R}}$.

2 – A.9 Check whether the family of matrices

$$\mathcal{Q} : \ A = \begin{bmatrix} 2 & -2 \\ 4 & -6 \end{bmatrix}, \ B = \begin{bmatrix} 1 & -4 \\ 5 & 3 \end{bmatrix}, \ C = \begin{bmatrix} 0 & -4 \\ 4 & 8 \end{bmatrix}$$

is linearly dependent or independent, and find a basis in the subspace of spanned by this family.

2 – A.10 Let us consider a subset of $\mathbb{R}^n$ defined by

$$W = \{ X = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_1 + x_2 + \ldots + x_n = 0 \ \wedge \ x_1 + x_2 = 0 \}.$$

Show that W is a subspace of $\mathbb{R}^n$, find its dimension and a basis spanning it.

2 – A.11 Let W be the subset of $\mathcal{M}_3(\mathbb{R})$ of matrices of the form

$$\begin{bmatrix} a & b & c \\ c & a-c & 0 \\ b & 0 & 0 \end{bmatrix}.$$

Show that $W \subseteq_{\text{subsp}} \mathcal{M}_3(\mathbb{R})$ and identify a basis of W.

2 – A.12 Determine the dimensions of the sum and intersection spanned by

$$\mathcal{Q} : a_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \ a_2 = \begin{bmatrix} 3 \\ 4 \\ -2 \end{bmatrix}, \ a_3 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} \ \text{ and } \ \mathcal{B} : b_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \ b_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix},$$

respectively.

2 – A.13 Show that the family of functions

$$\{f_n\}_{n\in\mathbb{N}} \subset \mathcal{F}_{[-\pi,\pi]} : f_n(x) = \cos nx$$

is linearly independent. What about $\dim \{f_n\}_{n\in\mathbb{N}}$?

2 – A.14 Establish which of the following subsets of $\mathrm{POL}_{\leq n}(\mathbb{R})$ (the space of polynomials of degree $\leq n$) are vector subspaces:

(*a*) $S_1 = \{p(x) : p(0) = \alpha \in \mathbb{R}\}$;

(*b*) $S_2 = \{p(x) : p(x) = p'(x)\}$;

(*c*) $S_3 = \{p(x) : 2p(0) + p(2) = 0\}$;

(*d*) $S_4 = \{p(x) : p(1) + p(2) + \ldots + p(n) = 0\}$.

2 – A.15 Determine $\alpha, \beta \in (\mathbb{R})$ so that the dimension of the subspace spanned by the matrices

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & \alpha \\ 1 & 0 & 1 \end{bmatrix},\ A_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \beta & 1 \\ 1 & 0 & 1 \end{bmatrix},\ A_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

be minimum (the least possible).

2 – A.16 Two subspaces $W_1, W_2 \subseteq \mathbb{R}^3$ are respectively spanned by $A : a = [1\ \ 2\ \ 0]^{\mathrm{T}}$ and $B : b_1 = [2\ \ 1\ \ 0]^{\mathrm{T}}$ & $b_2 = [1\ \ {-1}\ \ \alpha]^{\mathrm{T}}$. Find α such that $W_1 \cap W_2 = \{\mathbf{0}\}$. Then find the sum decomposition of

$$u = [1\ \ 1\ \ 1]^{\mathrm{T}} = w_1 + w_2 : w_1 \in W_1\ \ \&\ \ w_2 \in W_2 .$$