Contents

Chapter 1

Algebraic structures

1.1 Group

1.1.1 Definitions and examples

Definition 1.1 (*Binary operation*)

*A **binary operation** (or binary law) on a nonempty set S is a map from $S \times S$ to S. Such a binary operation is usually denoted*

$$\begin{array}{rccc} \star : & S \times S & \longrightarrow & S \\ & (a,b) & \longmapsto & a \star b \end{array}$$

Example : The set $\mathbb{N}$ of all natural numbers comes with the binary operation "addition of two numbers".

$$\begin{array}{rccc} + : & \mathbb{N} \times \mathbb{N} & \longrightarrow & \mathbb{N} \\ & (a,b) & \longmapsto & a+b \end{array}$$

Example : Similarly, the addition provides a binary operation on the set $\mathbb{Z}$ of all integers.

$$\begin{array}{rccc} + : & \mathbb{Z} \times \mathbb{N} & \longrightarrow & \mathbb{Z} \\ & (a,b) & \longmapsto & a+b \end{array}$$

Moreover this binary operation satisfies the following properties

i) $\forall (a,b,c) \in \mathbb{Z}^3,\ (a+b)+c = a+(b+c) = a+b+c$

ii) $0 \in \mathbb{Z}$ is such that $\forall a \in \mathbb{Z},\ a+0 = 0+a = a$

iii) $\forall a \in \mathbb{Z},\ \exists b \in \mathbb{Z} /\ a+b = b+a = 0$ (actually $b = -a$, the opposite number of a)

Example : Let $\mathcal{B}$ be the set of all bijective functions from the segment $[0,1]$ to itself.

$$\mathcal{B} = \{f : [0,1] \to [0,1] \,/\, f \text{ bijective}\}$$

Notice this set is nonempty (for instance $\mathrm{Id}_{[0,1]} = (x \mapsto x) \in \mathcal{B}$). The composition of functions provides a binary operation on $\mathcal{B}$.

$$\begin{array}{rccc} \circ : & \mathcal{B} \times \mathcal{B} & \longrightarrow & \mathcal{B} \\ & (f,g) & \longmapsto & f \circ g \end{array}$$

Indeed $f \circ g$ is well in $\mathcal{B}$ since $f \circ g : [0,1] \to [0,1]$ is bijective as composition of bijective functions. Moreover this binary operation satisfies the following properties

i) $\forall (f,g,h) \in \mathcal{B}^3,\ (f \circ g) \circ h = f \circ (g \circ h) = f \circ g \circ h$

ii) $\mathrm{Id}_{[0,1]} \in \mathcal{B}$ is such that $\forall f \in \mathcal{B},\ f \circ \mathrm{Id}_{[0,1]} = \mathrm{Id}_{[0,1]} \circ f = f$

iii) $\forall f \in \mathcal{B},\ \exists g \in \mathcal{B} /\ f \circ g = g \circ f = \mathrm{Id}_{[0,1]}$ (actually $g = f^{-1}$, the bijective inverse function of f)

Definition 1.2 (*Group*)

Let $(G, \star)$ be a nonempty set with a binary operation. $(G, \star)$ is said to be a ***group*** *if it satisfies each of the following group axioms*

i) Associativity: $\forall (a,b,c) \in G^3,\ (a \star b) \star c = a \star (b \star c) = a \star b \star c$

ii) Identity element: $\exists e \in G /\ \forall a \in G,\ a \star e = e \star a = a$ *(e is called identity element)*

iii) Inverse element: $\forall a \in G,\ \exists a' \in G /\ a \star a' = a' \star a = e$ *(a' is called inverse element of a)*

Examples : $(\mathbb{Z}, +)$ and $(\mathcal{B}, \circ)$ are groups whose identity elements are respectively $0 \in \mathbb{Z}$ and $\mathrm{Id}_{[0,1]} \in \mathcal{B}$.

Definition 1.3 (*Abelian group*)

A binary operation $\star$ on a nonempty set S is said ***commutative*** *if*

$$\forall (a,b) \in S^2,\ a \star b = b \star a$$

An ***abelian group*** *(or commutative group) is a group $(G, \star)$ for which its binary operation $\star$ is commutative.*

Example : $(\mathbb{Z}, +)$ is an abelian group since the addition of numbers is commutative.

Example : $(\mathcal{B}, \circ)$ is not abelian since $\exists (f,g) \in \mathcal{B}^2 /\ f \circ g \neq g \circ f$. For instance, consider $f = (x \mapsto x^2) \in \mathcal{B}$ (its associated inverse element is $f^{-1} = (y \mapsto \sqrt{y}) \in \mathcal{B}$) and $g = (x \mapsto 1 - x) \in \mathcal{B}$ (its associated inverse element is $g^{-1} = (x \mapsto 1 - x) = g \in \mathcal{B}$). Then

$$\forall x \in [0,1],\ \begin{cases} f \circ g(x) = f(1-x) = (1-x)^2 = 1 - 2x + x^2 \\ g \circ f(x) = g(x^2) = 1 - x^2 \end{cases}$$

So $f \circ g \neq g \circ f$ (for instance at $x = \frac{1}{2}$).

Further examples :

a) Additive groups. $(\mathbb{Z}, +) \subset (\mathbb{Q}, +) \subset (\mathbb{R}, +) \subset (\mathbb{C}, +)$ are abelian groups.

b) Multiplicative groups. $(\mathbb{Q}^*, \times) \subset (\mathbb{R}^*, \times) \subset (\mathbb{C}^* \times)$ are abelian groups. The same goes for $(\mathbb{R}_+^*, \times)$ (but not for $\mathbb{R}_-^*$ because the multiplication is not a binary operation on this set).

Proof : For instance for $(\mathbb{R}^*, \times)$:

1. $\times$ is a binary operation on $\mathbb{R}^*$ since the multiplication of two real numbers not equal to 0 remains a real number not equal to 0.
2. $\times$ is associative: $\forall (a,b,c) \in (\mathbb{R}^*)^3,\ (ab)c = a(bc) = abc$
3. $1 \in \mathbb{R}^*$ is identity element: $\forall a \in \mathbb{R}^*,\ a \times 1 = 1 \times a = a$
4. Every real number a not equal to 0 has a multiplicative inverse $\frac{1}{a} \in \mathbb{R}^*$ which is inverse element of a (since $a \times \frac{1}{a} = \frac{1}{a} \times a = 1$).
5. Finally the group is abelian since the multiplication of numbers is commutative.

c) Counterexamples. $(\mathbb{N},+)$ is not a group since $1 \in \mathbb{N}$ but there is no positive integer $a \in \mathbb{N}$ such that $1+a=a+1=0$ (actually $-1 \notin \mathbb{N}$). The same goes for $(\mathbb{Z}-\{0\},\times)$ since $2 \in \mathbb{Z}$ but $\frac{1}{2} \notin \mathbb{Z}$.

Proposition 1.4

Let $(G,\star)$ be a group. Then the following properties hold

1. *The identity element $e \in G$ is unique.*
2. *For any element $a \in G$, the inverse element of a in $(G,\star)$ is unique, denoted a^{-1}.*
3. *Simplifications are possible:* $\forall (a,b,c) \in G^3,\ \begin{cases} a\star b = a\star c \Longrightarrow b=c \\ b\star a = c\star a \Longrightarrow b=c \end{cases}$
4. $\forall (a,b) \in G^2,\ (a\star b)^{-1} = b^{-1}\star a^{-1}$

Proof : 1. By contradiction, assume e_1 and e_2 in G are identity elements for $\star$. Then

$$\forall a \in G,\ \begin{cases} a\star e_1 = e_1 \star a = a \\ a\star e_2 = e_2 \star a = a \end{cases}$$

$a=e_2$ in the first line gives $e_2\star e_1 = e_1\star e_2 = e_2$.
$a=e_1$ in the second line gives $e_1\star e_2 = e_2\star e_1 = e_1$.
Consequently $e_1 = e_1\star e_2 = e_2$, that is all identity elements must be equal. In other words there is only one identity element.

2. Similarly, given any element $a \in G$, assume a' and a'' in G are inverse elements of a. Then

$$\begin{cases} a\star a' = a'\star a = e \\ a\star a'' = a''\star a = e \end{cases}$$

Using the identity element e and the associativity of binary operation $\star$, we get:

$$a' = a'\star e = a'\star(a\star a'') = (a'\star a)\star a'' = e\star a'' = a''$$

Thus, there is only one inverse element of a.

3. Using the inverse element of a and the associativity, we get:

$$\begin{aligned} a\star b = a\star c &\Longrightarrow a^{-1}\star(a\star b) = a^{-1}\star(a\star c) \\ &\Longrightarrow (a^{-1}\star a)\star b = (a^{-1}\star a)\star c \\ &\Longrightarrow e\star b = e\star c \\ &\Longrightarrow b = c \end{aligned}$$

The same goes for the second simplification.

4. We need to check the axiom of inverse element in Definition 1.2.

$$\begin{cases} (a\star b)\star(b^{-1}\star a^{-1}) = a\star(b\star b^{-1})\star a^{-1} = a\star e\star a^{-1} = a\star a^{-1} = e \\ (b^{-1}\star a^{-1})\star(a\star b) = b^{-1}\star(a^{-1}\star a)\star b = b^{-1}\star e\star b = b^{-1}\star b = e \end{cases}$$

So $b^{-1}\star a^{-1}$ is the inverse element of the element $a\star b$. ■

Example : If f and g are two bijective functions from the segment $[0,1]$ to itself then $(f\circ g)^{-1} = g^{-1}\circ f^{-1}$ (which is not necessary equal to $f^{-1}\circ g^{-1}$ since composition of functions is not a commutative binary operation).

1.1.2 Subgroup

Example : Consider the abelian group $(\mathbb{Z}, +)$ and the following subsets

$$\mathcal{E} = \{\text{even numbers}\} = \{2k,\ k \in \mathbb{Z}\}$$
$$\mathcal{O} = \{\text{odd numbers}\} = \{2k+1,\ k \in \mathbb{Z}\}$$

We have $\forall (a,b) \in \mathcal{E},\ a+b \in \mathcal{E}$ (if $a = 2k$ and $b = 2k'$ then $a+b = 2k''$ with $k'' = k+k' \in \mathbb{Z}$) but it is no longer true in $\mathcal{O}$. In other words, $+$ is a binary operation on $\mathcal{E}$ but not on $\mathcal{O}$.

Definition 1.5 (*Subgroup*)

Let $(G, \star)$ be a group and $H \subset G$ be a nonempty subset. $(H, \star)$ is a ***subgroup*** *of $(G, \star)$ if it satisfies each of the following subgroup axioms*

i) *$\star$ is a binary operation on H*

ii) *$(H, \star)$ is a group*

Example : For any group $(G, \star)$ with identity element e, $(\{e\}, \star)$ is a subgroup of $(G, \star)$ called the **trivial subgroup**.

Proposition 1.6

Let $(G, \star)$ be a group and $H \subset G$ be a nonempty subset. $(H, \star)$ is a subgroup of $(G, \star)$ if and only if it satisfies the following condition

$$\forall (a,b) \in H^2,\ a \star b^{-1} \in H$$

Proof : **Necessary.** Let a and b be two elements in H. The axiom of inverse element in Definition 1.2 gives $b^{-1} \in H$. And because $\star$ is a binary operation on H, we get $a \star b^{-1} \in H$.

Sufficient. We need to check that each axiom from Definition 1.2 is satisfied for $(H, \star)$ and $\star$ is a binary operation on H.

1. The axiom of associativity for $(H, \star)$ comes directly from that one satisfied by $(G, \star)$.
2. Denote $e \in G$ the identity element of $(G, \star)$. Since H is nonempty, there exists an element $a \in H$. It follows $e = a \star a^{-1} \in H$. Consequently there exists an identity element in $(H, \star)$ or equivalently, the axiom of identity element is checked for $(H, \star)$.
3. Now the inverse element in $(G, \star)$ of any $a \in H$ satisfy $a^{-1} = e \star a^{-1} \in H$ (since $e \in H$). In particular any element $a \in H$ has an inverse element in $(H, \star)$ or equivalently, the axiom of inverse element is checked for $(H, \star)$.
4. Finally for every elements a and b in H, we get $a \star b = a \star (b^{-1})^{-1} \in H$ (since $b^{-1} \in H$). In other words, $\star$ is well a binary operation on H.

■

Remark : In practice, it is more convenient to use Proposition 1.6 to show that $(H, \star)$ is a subgroup of $(G, \star)$ than Definition 1.5

Examples : $(\mathbb{Z}, +)$ is a subgroup of $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ and $(\mathbb{C}, +)$. $(\mathbb{Q}, +)$ is a subgroup of $(\mathbb{R}, +)$ and $(\mathbb{C}, +)$. $(\mathbb{R}_+^*, \times)$ is a subgroup of $(\mathbb{R}^*, \times)$ and $(\mathbb{C}^*, \times)$.

Proposition 1.7

For every positive integer $n \in \mathbb{N}^*$, consider the following subset of $\mathbb{Z}$

$$n\mathbb{Z} = \{\text{multiples of } n\} = \{nk,\ k \in \mathbb{Z}\}$$

Then $(n\mathbb{Z}, +)$ is a subgroup of $(\mathbb{Z}, +)$.

Proof: $n\mathbb{Z}$ is well a nonempty subset of $\mathbb{Z}$ (for instance $0 \in n\mathbb{Z}$). Let $a = nk$ and $b = nk'$ be two elements in $n\mathbb{Z}$. Then

$$a + (-b) = (nk) + \big(-(nk')\big) = n(k - k') \in n\mathbb{Z}$$

Hence, the result follows from Proposition 1.6. ■

Example: $\mathcal{E} = \{\text{even numbers}\} = 2\mathbb{Z}$ is a subgroup of $\mathbb{Z}$.

Further examples:

a) Consider the following set of all complex numbers with absolute value 1 (that is the unital circle in the complex plane)

$$\mathbb{T} = \{z \in \mathbb{C} \,/\, |z| = 1\}$$

Then $(\mathbb{T}, \times)$ is a subgroup of $(\mathbb{C}^*, \times)$ called the **circle group**.

Proof: If $|z| = 1$ for a complex number $z \in \mathbb{C}$ then $z \neq 0$. Moreover $1 \in \mathbb{T}$. So $\mathbb{T}$ is a nonempty subset of $\mathbb{C}^*$. Now using the claims that the absolute value of the multiplicative inverse of a complex number is the multiplicative inverse of its absolute value and the absolute value of the product of two complex numbers is equal to the product of their absolute values, we get

$$\forall (z, w) \in \mathbb{T}^2,\ |z \times w^{-1}| = \frac{|z|}{|w|} = 1$$

The conclusion follows with Proposition 1.6. ■

b) For every positive integer $n \in \mathbb{N}^*$, consider the following subset of $\mathbb{C}$

$$\mathcal{R}_n = \{n^{\text{th}} \text{ root of unity}\} = \{z \in \mathbb{C} \,/\, z^n = 1\}$$

Then $(\mathcal{R}_n, \times)$ is a subgroup of $(\mathbb{T}, \times)$ (and of $(\mathbb{C}^*, \times)$ as well). It is finite with cardinality n.

Proof: If $z^n = 1$ for a complex number $z \in \mathbb{C}$ then $|z|^n = |z^n| = 1$ and consequently $|z| = 1$ (because $|z| \in \mathbb{R}_+$). Moreover 1 is a root of unitaly for every power $n \in \mathbb{N}^*$. So $\mathcal{R}_n$ is a nonempty subset of $\mathbb{T}$. Actually $\mathcal{R}_n$ may be written as follows

$$\mathcal{R}_n = \left\{z = e^{\imath \frac{2k\pi}{n}},\ k \in \{0, 1, \ldots, n-1\}\right\} \text{ where } \forall \theta \in \mathbb{R},\ e^{\imath\theta} = \cos(\theta) + \imath \sin(\theta)$$

In particular, $\mathcal{R}_n$ is a finite set with cardinality n. We conclude as in the previous proof for $(\mathbb{T}, \times)$ since the exponentiation function $z \mapsto z^n$ satisfies the same required properties as the absolute value for complex numbers. ■

1.1.3 Group homomorphism

Definition 1.8 (*Group homomorphism*)

Let $(G, \star)$ and $(H, \cdot)$ be two groups. A ***group homomorphism*** *from $(G, \star)$ to $(H, \cdot)$ is a function $\varphi : G \to H$ such that*

$$\forall (a, b) \in G^2,\ \varphi(a \star b) = \varphi(a) \cdot \varphi(b)$$

Moreover

- *If a group homomorphism $\varphi : G \to H$ is a bijection, then it is called a* ***group isomorphism****.*
- *If $\psi : G \to G$ is a group homomorphism from $(G, \star)$ to itself, then it is called a* ***group endomorphism****. If furthermore ψ is bijective and hence a group isomorphism, it is called a* ***group automorphism****.*

Remark : If $\varphi : G \to H$ is a group isomorphism, then its bijective inverse $\varphi^{-1} : H \to G$ is also a group homomorphism.

Proof : Let a and b be two elements in H. Since φ is a surjective function, there exist two elements a' and b' in G such that $a = \varphi(a')$ and $b = \varphi(b')$. Then

$$\varphi^{-1}(a \cdot b) = \varphi^{-1}(\varphi(a') \cdot \varphi(b')) = \varphi^{-1}(\varphi(a' \star b')) = a' \star b' = \varphi^{-1}(a) \star \varphi^{-1}(b)$$

■

Some examples :

a) The constant function $\varphi : a \mapsto \varphi(a) = e_H$ equal to the identity element of $(H, \cdot)$ is a group homomorphism from $(G, \star)$ to $(H, \cdot)$ (even onto the trivial subgroup $(\{e_H\}, \cdot)$).

b) The identity function $\varphi = \mathrm{Id}_G = (a \mapsto a)$ is a group automorphism of $(G, \star)$.

c) Given $n \in \mathbb{Z}$, $\varphi_n : \mathbb{Z} \to \mathbb{Z},\ k \mapsto nk$ is a group endomorphism from $(\mathbb{Z}, +)$ to itself and a group isomorphism from $(\mathbb{Z}, +)$ onto its subgroup $(n\mathbb{Z}, +)$.

d) Given $\lambda \in \mathbb{R}$, $\varphi_\lambda : \mathbb{R} \to \mathbb{R},\ x \mapsto \lambda x$ is a group endomorphism from $(\mathbb{R}, +)$ to itself. Moreover it is a group automorphism if and only if $\lambda \neq 0$.

Proof :

$$\forall (x, y) \in \mathbb{R}^2,\ \varphi_\lambda(x + y) = \lambda(x + y) = \lambda x + \lambda y = \varphi_\lambda(x) + \varphi_\lambda(y)$$

And φ_λ is a bijective function over $\mathbb{R}$ if and only if $\lambda \neq 0$ (in this case, the associated bijective inverse is $(x \mapsto x/\lambda) = \varphi_{1/\lambda}$). ■

e) $\exp : \mathbb{R} \to \mathbb{R}_+^*,\ x \mapsto \exp(x)$ is a group isomorphism from $(\mathbb{R}, +)$ to $(\mathbb{R}_+^*, \times)$.

Proof :

$$\forall (x, y) \in \mathbb{R}^2,\ \exp(x + y) = \exp(x) \exp(y)$$

And $\exp : \mathbb{R} \to \mathbb{R}_+^*$ is a bijective function (whose bijective inverse is $\ln : \mathbb{R}_+^* \to \mathbb{R}$). ■

1.2 Ring

1.2.1 Definitions and examples

Definition 1.9 (*Ring*)

Let $(R,+,\times)$ be a nonempty set with two binary operations denoted $+$ and $\times$. $(R,+,\times)$ is said to be a ***ring*** *if it satisfies each of the following ring axioms*

i) *$(R,+)$ is an abelian group*

ii) *the binary operation $\times$ is associative:*

$$\forall (a,b,c) \in R^3,\ (a\times b)\times c = a\times(b\times c) = a\times b\times c$$

iii) *the binary operation $\times$ is* ***distributive*** *over the binary operation $+$:*

$$\forall (a,b,c) \in R^3,\ \begin{cases} a\times(b+c) = (a\times b)+(a\times c) \\ (b+c)\times a = (b\times a)+(c\times a) \end{cases}$$

Moreover a ring $(R,+,\times)$ is said

- ***unital*** *if there exists an identity element for the binary operation $\times$:*

$$\exists e \in R/\ \forall a \in R,\ a\times e = e\times a = a$$

- ***commutative*** *if the binary operation $\times$ is commutative:*

$$\forall (a,b) \in R^2,\ a\times b = b\times a$$

We denote

- *0_R the identity element for the binary operation $+$, called the* ***additive identity***
- *1_R the identity element for the binary operation $\times$ in case $(R,+,\times)$ is unital, called the* ***multiplicative identity***
- *$-a$ the inverse element of an element $a \in R$ for the binary operation $+$, called the* ***additive inverse*** *(or* ***opposite****)*

Example : $(\mathbb{Z},+,\times) \subset (\mathbb{Q},+,\times) \subset (\mathbb{R},+,\times) \subset (\mathbb{C},+,\times)$ are commutative unital rings.

Proof : For instance for $(\mathbb{Q},+,\times)$:

1. $+$ and $\times$ are binary operations on $\mathbb{Q}$: $\forall (a,b) \in \mathbb{Q}^2,\ a+b \in \mathbb{Q}$ and $ab \in \mathbb{Q}$
2. $(\mathbb{Q},+)$ is an abelian group.
3. $\times$ is associative on $\mathbb{Q}$: $\forall (a,b,c) \in \mathbb{Q}^3,\ (ab)c = a(bc) = abc$
4. $\times$ is distributive over $+$: $\forall (a,b,c) \in \mathbb{Q}^3,\ a(b+c) = (ab)+(ac)$
5. $\times$ has a multiplicative identity in $\mathbb{Q}$: $1_{\mathbb{Q}} = 1$ since $\forall a \in \mathbb{Q},\ a\times 1 = 1\times a = a$
6. $\times$ is commutative on $\mathbb{Q}$: $\forall (a,b) \in \mathbb{Q}^2,\ ab = ba$

■

Example : Let $\mathcal{F}$ be the set of all functions from $\mathbb{R}$ to itself.

$$\mathcal{F} = \{f : \mathbb{R} \to \mathbb{R}\}$$

For every functions f and g in $\mathcal{F}$, the function $f+g \in \mathcal{F}$ is defined by $\forall x \in \mathbb{R},\ (f+g)(x) = f(x)+g(x)$ (using addition of real numbers) and the function $f \times g \in \mathcal{F}$ by $\forall x \in \mathbb{R},\ (f \times g)(x) = f(x)g(x)$ (using multiplication of real numbers). Then $(\mathcal{F}, +, \times)$ is a commutative unital ring.

Proof : 1. $+$ and $\times$ are binary operations on $\mathcal{F}$ by definition.

2. $(\mathcal{F}, +)$ is an abelian group: the associativity and the commutativity come from those ones of $(\mathbb{R}, +)$, the additive identity is the constant function $0_{\mathcal{F}} : x \mapsto 0$ and the additive inverse of a function $f \in \mathcal{F}$ is the function $-f$ defined by $-f : x \mapsto -x$.
3. $\times$ is associative on $\mathbb{R}$, so the same does on $\mathcal{F}$ as well.
4. $\times$ is distributive over $+$ on $\mathbb{R}$, so the same does on $\mathcal{F}$ as well.
5. The constant function $1_{\mathcal{F}} : x \mapsto 1$ is a multiplicative identity for $\times$ in $\mathcal{F}$.
6. $\times$ is commutative on $\mathbb{R}$, so the same does on $\mathcal{F}$ as well. ■

Proposition 1.10

Let $(R, +, \times)$ be a ring. Then the following properties hold

1. *If $(R, +, \times)$ is unital, then the multiplicative identity $1_R \in R$ is unique.*
2. $\forall a \in R,\ a \times 0_R = 0_R \times a = 0_R$
3. *If $(R, +, \times)$ is unital, then* $\forall a \in R,\ (-1_R) \times a = a \times (-1_R) = -a$
4. $\forall (a, b) \in R^2,\ (-a) \times b = a \times (-b) = -(a \times b)$

Proof : 1. Actually the same proof as for the first property from Proposition 1.4 still holds.

2. Using the distributivity of $+$ over $\times$ and the additive identity 0_R, we get:

$$0_R \times a + 0_R \times a = (0_R + 0_R) \times a = 0_R \times a = 0_R + 0_R \times a$$

Now a simplification on each side by $0_R \times a$ (see Proposition 1.4) gives $0_R = 0_R \times a$ as needed. The same goes for $a \times 0_R$.

3. We have

$$a + (-1_R) \times a = 1_R \times a + (-1_R) \times a = (1_R + (-1_R)) \times a = 0_R \times a = 0_R$$

Consequently $(-1_R) \times a$ is the additive inverse of the element a (since $(R, +)$ is an abelian group whose its additive identity is 0_R). The same goes for $a \times (-1_R)$.

4. We have

$$a \times b + (-a) \times b = (a + (-a)) \times b = 0_R \times b = 0_R$$

The conclusion follows. ■

Remark : In particular, the second point of the previous proposition shows that the additive identity 0_R of a ring $(R, +, \times)$ has no multiplicative inverse: there is no element $a \in R$ such that $0_R \times a = a \times 0_R = 1_R$.

Definition 1.11 (*Subring*)

Let $(R,+,\times)$ be a ring and $T \subset R$ be a nonempty subset. $(T,+,\times)$ is a ***subring*** *of $(R,+,\times)$ if it satisfies each of the following subring axioms*

i) *$(T,+)$ is a subgroup of $(R,+)$*

ii) *$\times$ is a binary operation on T*

In this case, $(T,+,\times)$ is a ring.

Trivial example : For any ring $(R,+,\times)$, $(\{0_R\},+,\times)$ is a subring called the **trivial subring**.

Examples : $(\mathbb{Z},+,\times)$ is a subring of $(\mathbb{Q},+,\times)$, $(\mathbb{R},+,\times)$ and $(\mathbb{C},+,\times)$. For every positive integer $n \in \mathbb{N}^*$, $(n\mathbb{Z},+,\times)$ is a subring of $(\mathbb{Z},+,\times)$ which is not unital as soon as $n \geqslant 2$ (since $1 \notin n\mathbb{Z}$).

Proof : For instance for $(n\mathbb{Z},+,\times)$:

1. $(n\mathbb{Z},+)$ is a subgroup of $(\mathbb{Z},+)$ (see Proposition 1.7)
2. If $a = nk$ and $b = nk'$ are two elements in $n\mathbb{Z}$ then $a \times b = (nk) \times (nk') = n(knk') \in n\mathbb{Z}$. Hence, $\times$ is a binary operation on $n\mathbb{Z}$.

■

Definition 1.12 (*Ring homomorphism*)

Let $(R,+,\times)$ and $(T,\oplus,\otimes)$ be two rings. A ***ring homomorphism*** *from $(R,+,\times)$ to $(T,\oplus,\otimes)$ is a function $\varphi : R \to T$ such that*

$$\forall (a,b) \in R^2,\ \begin{cases} \varphi(a+b) = \varphi(a) \oplus \varphi(b) \\ \varphi(a \times b) = \varphi(a) \otimes \varphi(b) \end{cases}$$

Moreover

- *If a ring homomorphism $\varphi : R \to T$ is a bijection, then it is called a* ***ring isomorphism****.*
- *If $\psi : R \to R$ is a ring homomorphism from $(R,+,\times)$ to itself, then it is called a* ***ring endomorphism****. If furthermore ψ is bijective and hence a ring isomorphism, it is called a* ***ring automorphism****.*

Remark : The same remark as for group isomorphism holds: the bijective inverse of any ring isomorphism is also a ring homomorphism.

1.2.2 The rings $\mathbb{Z}/n\mathbb{Z}$

Remind the following result:

Theorem 1.13 (*Division algorithm*)

For any given two integers a and d with $d \neq 0$, there exist unique integers q and r such that

$$\begin{cases} a = qd + r \\ 0 \leqslant r < |d| \end{cases}$$

The integer q is called the ***quotient****, r the* ***remainder****, d the* ***divisor*** *and a the* ***dividend****.*

Fix a positive integer $n \in \mathbb{N}^*$.

Definition 1.14 (*Congruence modulo n*)

Two integers a and b are said ***congruent modulo n*** *if n divides their difference $a-b$, that is if there exists an integer q such that $a-b=qn$ (equivalently if the remainder of the division algorithm with dividend $a-b$ and divisor n is equal to 0). In this case, we write $\boldsymbol{a \equiv b[n]}$.*

Remark : In particular, any integer a is congruent modulo n to its associated remainder r of the division algorithm with dividend a and divisor n (since $a-r=qn$). Moreover if two integers a and b have the same remainder r from the division algorithm with divisor n, then they are congruent modulo n (since $a=qn+r$ and $b=q'n+r'$ with $r=r'$ imply $a-b=(q-q')n$).

Definition 1.15 (*Congruence class*)

For any given integer $a \in \mathbb{Z}$, the ***congruence class modulo n*** *of a is the following set*

$$\begin{aligned}\overline{a} &= \{\text{integers congruent modulo } n \text{ to } a\}\\ &= \{b \in \mathbb{Z} \,/\, b \equiv a[n]\}\\ &= \{\ldots, a-2n, a-n, a, a+n, a+2n, a+3n, \ldots\}\\ &= a+n\mathbb{Z}\end{aligned}$$

Proposition 1.16

Let $\overline{a}$ and $\overline{b}$ be two congruence classes modulo n associated to two integers a and b. Then $\overline{a}$ and $\overline{b}$ are equal if and only if a and b are congruent modulo n.

$$\forall (a,b) \in \mathbb{Z}^2,\ \overline{a}=\overline{b} \Leftrightarrow a \equiv b[n]$$

Proof : Necessary. If $\overline{a}=\overline{b}$ then in particular $b \in \overline{a}$ that is b is congruent modulo n to a.

Sufficient. If b is congruent modulo n to a then there exists an integer q such that $a=b+qn$. Consequently

$$\overline{a}=a+n\mathbb{Z}=b+qn+n\mathbb{Z}=b+n(q+\mathbb{Z})=b+n\mathbb{Z}=\overline{b}$$

■

Remark : In particular, for any given integer $a \in \mathbb{Z}$, if r denotes the remainder of the division algorithm with dividend a and divisor n then $\overline{a}=\overline{r}$.

Definition 1.17 ($\mathbb{Z}/n\mathbb{Z}$)

The set of all congruence classes modulo n is denoted by $\boldsymbol{\mathbb{Z}/n\mathbb{Z}}$ (read "$\mathbb{Z}$ over $n\mathbb{Z}$").

$$\begin{aligned}\mathbb{Z}/n\mathbb{Z} &= \{\text{congruence classes modulo } n\}\\ &= \{\overline{a},\ a \in \mathbb{Z}\}\\ &= \{\text{congruence classes modulo } n \text{ of the remainders}\\ &\qquad\qquad \text{from the division algorithm with divisor } n\}\\ &= \{\overline{r},\ r \in \mathbb{Z} \text{ and } 0 \leqslant r < n\}\\ &= \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}\end{aligned}$$

It is a finite set with cardinality n.

Example : For $n = 2$, $\mathbb{Z}/2\mathbb{Z} = \{\overline{0}, \overline{1}\}$. For instance $\overline{2} = \overline{0}$ and $\overline{5} = \overline{1}$ (equivalently we may write $2 \equiv 0[2]$ and $5 \equiv 1[2]$). Actually

$$\begin{cases} a \text{ even} & \Longrightarrow & \overline{a} = \overline{0} \\ a \text{ odd} & \Longrightarrow & \overline{a} = \overline{1} \end{cases}$$

Example : For $n = 3$, $\mathbb{Z}/3\mathbb{Z} = \{\overline{0}, \overline{1}, \overline{2}\}$. For instance $\overline{6} = \overline{0}$ (or $6 \equiv 0[3]$) and $\overline{13} = \overline{1}$ (or $13 \equiv 1[3]$) because $13 = 4 \times 3 + 1$.

Definition 1.18 (*Addition and multiplication in $\mathbb{Z}/n\mathbb{Z}$*)

Two binary operations on $\mathbb{Z}/n\mathbb{Z}$ denoted by $+$ and $\times$ are defined as follows (using addition and multiplication of integers)

$$\forall (r, r') \in \{0, 1, \ldots, n-1\}^2, \ \begin{cases} \overline{r} + \overline{r'} &= \overline{r + r'} \\ \overline{r} \times \overline{r'} &= \overline{rr'} \end{cases}$$

Proposition 1.19

$$\forall (a, b) \in \mathbb{Z}^2, \ \begin{cases} \overline{a} + \overline{b} &= \overline{a + b} \\ \overline{a} \times \overline{b} &= \overline{ab} \end{cases}$$

Proof : From Theorem 1.13, $a = qn + r$ and $b = q'n + r'$ for some integers q, q', r, r' with $0 \leqslant r, r' < n$.

- We have $(a + b) - (r + r') = (q + q')n$. So $(a + b)$ and $(r + r')$ are congruent modulo n. It follows from Proposition 1.16 that

$$\overline{a + b} = \overline{r + r'} = \overline{r} + \overline{r'} = \overline{a} + \overline{b}$$

- We have

$$ab = (qn + r)(q'n + r') = qq'n^2 + (qr' + q'r)n + rr'$$

Thus $ab - rr' = (qq'n + qr' + q'r)n$ that is ab and rr' are congruent modulo n. It follows from Proposition 1.16 that

$$\overline{ab} = \overline{rr'} = \overline{r} \times \overline{r'} = \overline{a} \times \overline{b}$$

■

Corollary 1.20

$(\mathbb{Z}/n\mathbb{Z}, +, \times)$ is a commutative unital ring whose identity elements are respectively $\overline{0}$ for the binary operation $+$ and $\overline{1}$ for the binary operation $\times$.

Proof : Proposition 1.19 gives everything we need from the fact that $(\mathbb{Z}, +, \times)$ is a commutative unital ring whose identity elements are respectively 0 and 1. ■

Example : Some computations in $\mathbb{Z}/6\mathbb{Z} = \{\overline{0}, \overline{1}, \overline{2}, \overline{3}, \overline{4}, \overline{5}\}$:

a) $\overline{15} = \overline{3}$ since $15 = 2 \times 6 + 3$

b) $\overline{4} + \overline{5} = \overline{4 + 5} = \overline{9} = \overline{3 + 6} = \overline{3}$

c) $\overline{4} + \overline{2} = \overline{6} = \overline{0}$ thus $-\overline{4} = \overline{2}$

d) $\overline{5} \times \overline{4} = \overline{5 \times 4} = \overline{20} = \overline{3 \times 6 + 2} = \overline{2}$

Remark : In $\mathbb{Z}/6\mathbb{Z}$, we have $\overline{3} \times \overline{2} = \overline{6} = \overline{0}$ but $\overline{3} \neq \overline{0}$ and $\overline{2} \neq \overline{0}$.

1.3 Field

Definition 1.21

*Let $(F,+,\times)$ be a nonempty set with two binary operations. $(F,+,\times)$ is said to be a **field** if it satisfies each of the following field axioms*

i) *$(F,+,\times)$ is an unital ring*

ii) *every element except 0_F has an inverse element in F for the binary operation $\times$:*

$$\forall a \in F - \{0_F\},\ \exists a^{-1} \in F/\ a \times a^{-1} = a^{-1} \times a = 1_F$$

*Moreover a field $(F,+,\times)$ is said **commutative** if the binary operation $\times$ is commutative:*

$$\forall (a,b) \in F^2,\ a \times b = b \times a$$

*The same notations as for ring (0_F, 1_F and $-a$) are used and we denote a^{-1} the inverse element of an element $a \in F - \{0_F\}$ for the binary operation $\times$ (called the **multiplicative inverse**).*

Remark : Equivalently $(F,+,\times)$ is a field if and only if it satisfies each of the following conditions

i) $(F,+)$ is an abelian group (whose its additive identity is denoted by 0_F)

ii) $(F - \{0_F\},\times)$ is a group

iii) the binary operation $\times$ is distributive over the binary operation $+$

Examples : $(\mathbb{Q},+,\times) \subset (\mathbb{R},+,\times) \subset (\mathbb{C},+,\times)$ are commutative fields.

Counterexample : $(\mathbb{Z},+,\times)$ is not a field since $2 \in \mathbb{Z} - \{0\}$ has no multiplicative inverse in $\mathbb{Z}$: there is no integer $a \in \mathbb{Z}$ such that $2a = 1$ (equivalently $(\mathbb{Z} - \{0\},\times)$ is not a group).

Proposition 1.22

Let $(F,+,\times)$ be a field. Then

$$\forall (a,b) \in F^2,\ a \times b = 0_F \Longrightarrow \text{either } a = 0_F \text{ or } b = 0_F$$

Proof : Assume $a \neq 0_F$. Consequently, there exists a multiplicative inverse $a^{-1} \in F$ and we get:

$$b = 1_F \times b = (a^{-1} \times a) \times b = a^{-1} \times (a \times b) = a^{-1} \times 0_F = 0_F$$

■

Remark : In particular, if there exist two elements a and b in a ring $(R,+,\times)$ such that $a \times b = 0_R$ but $a \neq 0_R$ and $b \neq 0_R$ (such elements are called **zero divisors**) then $(R,+,\times)$ is not a field. More precisely, any zero divisor does not have multiplicative inverse.

Example : $(\mathbb{Z}/6\mathbb{Z},+,\times)$ is not a field (since $\overline{3} \times \overline{2} = \overline{0}$ but $\overline{3} \neq \overline{0}$ and $\overline{2} \neq \overline{0}$).

Theorem 1.23

Let $n \in \mathbb{N}^*$ be a positive integer. Then $(\mathbb{Z}/n\mathbb{Z},+,\times)$ is a commutative field if and only if n is a prime number.

Proof : **Necessary.** The proof is the same as the previous example. By contradiction, assume that $n \in \mathbb{N}^*$ is not a prime number. In other words, there exist two integers p and q such that $1 < p, q < n$ and $n = pq$. Then $\overline{p} \times \overline{q} = \overline{pq} = \overline{n} = \overline{0}$ but $\overline{p} \neq \overline{0}$ and $\overline{q} \neq \overline{0}$. That is a contradiction with Proposition 1.22.

Sufficient. Assume n is a prime number and let $\overline{r} \in \mathbb{Z}/n\mathbb{Z}$ be a congruence class not equal to the congruence class $\overline{0}$. We may assume that $0 < r < n$. In particular, r and n are relatively prime. Remind the following result:

Theorem 1.24 (*Bézout's identity*)

If two integers a and b are relatively prime then there exist integers x and y such that

$$ax + by = 1$$

Here we get two integers x and y such that $rx + ny = 1$. Consequently

$$\overline{r} \times \overline{x} = \overline{rx} = \overline{1 - ny} = \overline{1}$$

In other words, $\overline{x}$ is the multiplicative inverse of $\overline{r}$. So, any congruence class in $\mathbb{Z}/n\mathbb{Z}$ not equal to $\overline{0}$ has a multiplicative inverse. ■

Example : In $\mathbb{Z}/7\mathbb{Z}$, the multiplicative inverse of $\overline{2}$ is $\overline{4}$ since $\overline{2} \times \overline{4} = \overline{2 \times 4} = \overline{8} = \overline{1}$ (and then the multiplicative inverse of $\overline{4}$ is $\overline{2}$). Moreover, we have $\overline{3}^{-1} = \overline{5}$ (since $\overline{3} \times \overline{5} = \overline{15} = \overline{1 + 2 \times 7} = \overline{1}$) and $\overline{6}^{-1} = \overline{6}$ (since $\overline{6} \times \overline{6} = \overline{36} = \overline{1 + 5 \times 7} = \overline{1}$).

Chapter 2

Polynomials

In this chapter, fix a commutative field $(\mathbb{K}, +, \times)$ (for instance $\mathbb{K} = \mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$). We denote

- 0 the additive identity
- $-a$ the additive inverse of an element $a \in \mathbb{K}$
- 1 the multiplicative identity
- a^{-1} the multiplicative inverse of an element $a \in \mathbb{K}^* = \mathbb{K} - \{0\}$

2.1 The ring $\mathbb{K}[X]$

2.1.1 Definition and operations

Definition 2.1 (*Polynomial*)

Let $(a_n)_{n\in\mathbb{N}} = (a_0, a_1, a_2, \dots, a_n, \dots)$ be an infinite sequence of elements in $\mathbb{K}$ which are eventually equal to zero, that is

$$\exists d \in \mathbb{N} / \ \forall n > d, \ a_n = 0$$

*The **polynomial** with **coefficients** $(a_n)_{n\in\mathbb{N}}$ is the following formal expression*

$$P(X) = a_0 + a_1 X + a_2 X^2 + \cdots + a_d X^d = \sum_{n=0}^{d} a_n X^n = \sum_{n=0}^{+\infty} a_n X^n$$

Moreover

- *the formal symbol X is called the **variable***
- *the formal symbols $X^0 = 1, X^1 = X, X^2, \dots, X^n, \dots$ are called the **powers of X***
- *for any $n \in \mathbb{N}$, a_n is called the **coefficient of the term with degree n***
- *a_0 is called the **constant term***

We denote $\mathbb{K}[X]$ the set of all polynomials with coefficients in $\mathbb{K}$.

Examples : $P(X) = X + X^3 + X^5 = 0 + 1X + 0X^2 + 1X^3 + 0X^4 + 1X^5$ is a polynomial in $\mathbb{K}[X]$ but also $Q(X) = X^2$ or $R(X) = 0$. $S(X) = \frac{1}{2} - \sqrt{2}X^2 + 5X^4$ is a polynomial in $\mathbb{R}[X]$ or $\mathbb{C}[X]$ but not in $\mathbb{Q}[X]$.

Definition 2.2 (*Addition in* $\mathbb{K}[X]$)

Let $P(X)$ and $Q(X)$ be two polynomials in $\mathbb{K}[X]$ with coefficients respectively $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$. We define the sum of $P(X)$ and $Q(X)$, denoted $(P+Q)(X)$, to be the polynomial in $\mathbb{K}[X]$ with coefficients $(a_n+b_n)_{n\in\mathbb{N}}$. That provides a binary operation $+$ on $\mathbb{K}[X]$.

$$P(X)+Q(X)=(P+Q)(X)=\sum_{n=0}^{+\infty}(a_n+b_n)X^n$$

Example : For instance, we have in $\mathbb{R}[X]$:

$$\begin{aligned}(1+2X+3X^3)+(4-X+5X^4) &= (1+4)+(2-1)X+(0+0)X^2+(3+0)X^3+(0+5)X^4\\ &= 5+X+3X^3+5X^4\end{aligned}$$

Definition 2.3 (*Multiplication in* $\mathbb{K}[X]$)

Let $P(X)$ and $Q(X)$ be two polynomials in $\mathbb{K}[X]$ with coefficients respectively $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$. We define the product of $P(X)$ and $Q(X)$, denoted $(PQ)(X)$, to be the polynomial in $\mathbb{K}[X]$ with coefficients $(\sum_{k=0}^{n}a_kb_{n-k})_{n\in\mathbb{N}}$. That provides a binary operation $\times$ on $\mathbb{K}[X]$.

$$P(X)Q(X)=(PQ)(X)=\sum_{n=0}^{+\infty}\left(\sum_{k=0}^{n}a_kb_{n-k}\right)X^n=\sum_{n=0}^{+\infty}\left(\sum_{\substack{k,\ell\in\mathbb{N}\\ k+\ell=n}}a_kb_\ell\right)X^n$$

Remark : The previous definition is natural with respect to the following property

$$\forall(k,\ell)\in\mathbb{N}^2,\ X^kX^\ell=X^{k+\ell}$$

Example : For instance, we have in $\mathbb{R}[X]$:

$$\begin{aligned}(3-X-2X^2)(2+6X+4X^2) &= 6+(18-2)X+(12-6-4)X^2+(-4-12)X^3-8X^4\\ &= 6+16X+2X^2-16X^3-8X^4\end{aligned}$$

Proposition 2.4

$(\mathbb{K}[X],+,\times)$ is a commutative unital ring whose identity elements are respectively the constant polynomials 0 for addition and 1 for multiplication.

Proof : At first, $\mathbb{K}[X]$ is nonempty (for instance the constant polynomial 0 is in $\mathbb{K}[X]$).

1. $+$ and $\times$ are binary operations on $\mathbb{K}[X]$ by definition.
2. We have:
   (a) $+$ is associative on $\mathbb{K}$, so the same does on $\mathbb{K}[X]$ as well.
   (b) The constant polynomial $0\in\mathbb{K}[X]$ is the additive identity for $+$ in $\mathbb{K}[X]$.
   (c) For any polynomial in $\mathbb{K}[X]$ with coefficients $(a_n)_{n\in\mathbb{N}}$, the polynomial in $\mathbb{K}[X]$ with coefficients $(-a_n)_{n\in\mathbb{N}}$ is its additive inverse.
   (d) $+$ is commutative on $\mathbb{K}$, so the same does on $\mathbb{K}[X]$ as well.

   Consequently $(\mathbb{K}[X],+)$ is an abelian group.

3. Let $P(X)$, $Q(X)$ and $R(X)$ be three polynomials in $\mathbb{K}[X]$ with coefficients respectively $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ and $(c_n)_{n\in\mathbb{N}}$. We have:

$$\begin{aligned}
[P(X)Q(X)]\,R(X) &= \left[\sum_{n=0}^{+\infty}\left(\sum_{\substack{k,\ell\in\mathbb{N}\\ k+\ell=n}} a_k b_\ell\right) X^n\right]\sum_{j=0}^{+\infty} c_j X^j \\
&= \sum_{N=0}^{+\infty}\left(\sum_{\substack{n,j\in\mathbb{N}\\ n+j=N}}\left(\sum_{\substack{k,\ell\in\mathbb{N}\\ k+\ell=n}} a_k b_\ell\right) c_j\right) X^N \\
&= \sum_{N=0}^{+\infty}\left(\sum_{\substack{k,\ell,j\in\mathbb{N}\\ k+\ell+j=N}} a_k b_\ell c_j\right) X^N \\
&= \sum_{N=0}^{+\infty}\left(\sum_{\substack{k,m\in\mathbb{N}\\ k+m=N}} a_k\left(\sum_{\substack{\ell,j\in\mathbb{N}\\ \ell+j=m}} b_\ell c_j\right)\right) X^N \\
&= \sum_{k=0}^{+\infty} a_k X^k\left[\sum_{m=0}^{+\infty}\left(\sum_{\substack{\ell,j\in\mathbb{N}\\ \ell+j=m}} b_\ell c_j\right) X^m\right] \;=\; P(X)\,[Q(X)R(X)]
\end{aligned}$$

   Thus, $\times$ is associative on $\mathbb{K}[X]$.

4. Let $P(X)$, $Q(X)$ and $R(X)$ be three polynomials in $\mathbb{K}[X]$ with coefficients respectively $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ and $(c_n)_{n\in\mathbb{N}}$. We have:

$$\sum_{n=0}^{+\infty}\left(\sum_{\substack{k,\ell\in\mathbb{N}\\ k+\ell=n}} a_k(b_\ell+c_\ell)\right) X^n = \sum_{n=0}^{+\infty}\left(\sum_{\substack{k,\ell\in\mathbb{N}\\ k+\ell=n}} a_k b_\ell\right) X^n + \sum_{n=0}^{+\infty}\left(\sum_{\substack{k,\ell\in\mathbb{N}\\ k+\ell=n}} a_k c_\ell\right) X^n$$

   Equivalently, $P(X)\,(Q(X)+R(X)) = P(X)Q(X)+P(X)R(X)$. And the same goes for $(Q(X)+R(X))\,P(X) = Q(X)P(X)+R(X)P(X)$. Thus, $\times$ is distributive over $+$ on $\mathbb{K}[X]$.

5. Let $P(X)$ be a polynomial in $\mathbb{K}[X]$ with coefficients $(a_n)_{n\in\mathbb{N}}$. Since every coefficient of the constant polynomial $1\in\mathbb{K}[X]$ is equal to zero except the constant term equal to 1, we have:

$$P(X)1 = \sum_{n=0}^{+\infty}\left(a_0 0 + a_1 0 + \cdots + a_{n-1}0 + a_n 1\right) X^n = \sum_{n=0}^{+\infty} a_n X^n = P(X)$$

   And the same goes for $1P(X) = P(X)$. Thus, the constant polynomial $1\in\mathbb{K}[X]$ is the multiplicative identity for $\times$ in $\mathbb{K}[X]$.

6. $+$ and $\times$ are commutative on $\mathbb{K}$, so the same goes for $\times$ on $\mathbb{K}[X]$ as well.

Finally, $\mathbb{K}[X]$ satisfies all conditions to be a commutative unital ring. ■

Remark: But $(\mathbb{K}[X],+,\times)$ is not a field. For instance, the polynomial $P(X)=X\in\mathbb{K}[X]-\{0\}$ has no multiplicative inverse in $\mathbb{K}[X]$: there is no polynomial $Q(X)\in\mathbb{K}[X]$ such that $XQ(X)=1$.

Proof: For any polynomial $Q(X)=a_0+a_1X+a_2X^2+\cdots+a_dX^d\in\mathbb{K}[X]$, the constant term of $XQ(X)=a_0X+a_1X^2+a_2X^3+\cdots+a_dX^d\in\mathbb{K}[X]$ is 0 but that one of the constant polynomial $1\in\mathbb{K}[X]$ is 1. So the equality can not hold. ■

2.1.2 Degree

Definition 2.5 (*Degree*)

The ***degree*** *of a polynomial $P(X) \in \mathbb{K}[X]$, denoted* $\mathbf{deg(P(X))}$ *or shortly* $\mathbf{deg(P)}$*, is the highest exponent for terms with non zero coefficient. More precisely if $(a_n)_{n\in\mathbb{N}}$ are the coefficients of $P(X)$ then $\deg(P)$ is an element of $\mathbb{N}\cup\{-\infty\}$ defined by*

$$\deg(P) = \begin{cases} -\infty & \text{if } P(X) = 0 \\ \max\{n \in \mathbb{N} \,/\, a_n \neq 0\} & \text{otherwise} \end{cases}$$

Moreover, the coefficient $a_{\deg(P)} \in \mathbb{K}^*$ (in case $P(X) \neq 0$) is called the ***leading coefficient****.*

Some examples :

a) The (non zero) **constant polynomials** $P(X) = a_0 \in \mathbb{K}[X]$ where $a_0 \neq 0$ are of degree 0.

b) The **linear polynomials** $P(X) = a_0 + a_1X \in \mathbb{K}[X]$ where $a_1 \neq 0$ are of degree 1.

c) The **quadratic polynomials** $P(X) = a_0 + a_1X + a_2X^2 \in \mathbb{K}[X]$ where $a_2 \neq 0$ are of degree 2.

d) The **cubic polynomials** $P(X) = a_0 + a_1X + a_2X^2 + a_3X^3 \in \mathbb{K}[X]$ where $a_3 \neq 0$ are of degree 3.

etc.

Remark : It is usefull to define the degree of the zero constant polynomial to be $-\infty$ (furthermore it is convenient to take $\max\emptyset = -\infty$ as a convention). In the following, we introduce the rules:

$$\forall k \in \mathbb{N}\cup\{-\infty\},\ \begin{cases} -\infty \leqslant k \\ \max\{k, -\infty\} = k \\ k + (-\infty) = (-\infty) + k = -\infty \end{cases}$$

Proposition 2.6

Let $P(X)$ and $Q(X)$ be two polynomials in $\mathbb{K}[X]$. Then the following properties hold

1. $\deg(P+Q) \leqslant \max\{\deg(P), \deg(Q)\}$ *with equality if* $\deg(P) \neq \deg(Q)$
2. $\deg(PQ) = \deg(P) + \deg(Q)$

Proof : We write $P(X) = \sum_{n=0}^{\deg(P)} a_nX^n$ and $Q(X) = \sum_{n=0}^{\deg(Q)} b_nX^n$.

1. In case $\deg(P) < \deg(Q)$ we get

$$(P+Q)(X) = \sum_{n=0}^{\deg(P)} (a_n + b_n)X^n + \sum_{n=\deg(P)+1}^{\deg(Q)} b_nX^n$$

And $b_{\deg(Q)} \neq 0$ implies that $\deg(P+Q) = \deg(Q) = \max\{\deg(P), \deg(Q)\}$. The same goes when $\deg(P) > \deg(Q)$. Now if $\deg(P) = \deg(Q) = d$ then

$$(P+Q)(X) = \sum_{n=0}^{d} (a_n + b_n)X^n$$

Consequently, we have $\deg(P+Q) \leqslant d = \max\{\deg(P), \deg(Q)\}$.

2. We have:

$$(PQ)(X) = \sum_{n=0}^{+\infty} \left(\sum_{\substack{k,\ell \in \mathbb{N} \\ k+\ell=n}} a_k b_\ell \right) X^n = \sum_{n=0}^{+\infty} \left(\sum_{\substack{k \leqslant \deg(P) \\ \ell \leqslant \deg(Q) \\ k+\ell=n}} a_k b_\ell \right) X^n = \sum_{n=0}^{\deg(P)+\deg(Q)} \left(\sum_{\substack{k \leqslant \deg(P) \\ \ell \leqslant \deg(Q) \\ k+\ell=n}} a_k b_\ell \right) X^n$$

Moreover the coefficient of the term with degree $n = \deg(P) + \deg(Q)$ is

$$\sum_{\substack{k \leqslant \deg(P) \\ \ell \leqslant \deg(Q) \\ k+\ell=\deg(P)+\deg(Q)}} a_k b_\ell = a_{\deg(P)} b_{\deg(Q)} \neq 0$$

The conclusion follows. ■

Remarks :

- The results remain true if $P(X)$ or $Q(X)$ (or both) is the zero constant polynomial.
- In the first property, the sufficient condition for equality is not necessary. For instance

$$\deg(X + X) = \deg(2X) = 1 = \deg(X)$$

Actually if $P(X)$ and $Q(X)$ are two polynomials of same degree $d \in \mathbb{N}$ whose their leading coefficients are respectively a_d and b_d then

$$\deg(P + Q) = d \Leftrightarrow a_d \neq -b_d$$

2.1.3 Polynomial arithmetic

Theorem 2.7 (*Polynomial division algorithm*)

For any given two polynomials $P(X)$ and $D(X)$ with $D(X) \neq 0$, there exist unique polynomials $Q(X)$ and $R(X)$ such that

$$\begin{cases} P(X) = Q(X)D(X) + R(X) \\ \deg(R) < \deg(D) \end{cases}$$

*The polynomial $Q(X)$ is called the **quotient**, $R(X)$ the **remainder**, $D(X)$ the **divisor** and $P(X)$ the **dividend**.*

Proof : Existence. Fix a polynomial $D(X) \in \mathbb{K}[X]$ with $D(X) \neq 0$ and call $d \geqslant 0$ its degree. Remark that if $d = 0$, that is $D(X) = b_0 \neq 0$, then $Q(X) = b_0^{-1}P(X)$ and $R(X) = 0$ are suitable. So we may assume that $d \geqslant 1$.

We will prove by induction the following property for every $k \geqslant 0$

$$\mathcal{P}_k = \text{“the existence part is true for every } P(X) \in \mathbb{K}[X] \text{ with } \deg(P) \leqslant k\text{”}$$

At first, if $\deg(P) \leqslant d - 1$ then $Q(X) = 0$ and $R(X) = P(X)$ are suitable. Hence, $\mathcal{P}_k$ is satisfied for every integer k such that $0 \leqslant k \leqslant d - 1$ (and at least for $k = 0$ since $d \geqslant 1$).

Now assume $\mathcal{P}_k$ is satisfied for a given integer $k \geqslant d - 1$. Let $P(X)$ be a polynomial in $\mathbb{K}[X]$ of degree $\deg(P) = k + 1 \geqslant d$. We write:

$$\begin{array}{rcll} P(X) & = & a_0 + a_1X + a_2X^2 + \cdots + a_dX^d + \cdots + a_{k+1}X^{k+1} & \text{with } a_{k+1} \neq 0 \\ D(X) & = & b_0 + b_1X + b_2X^2 + \cdots + b_dX^d & \text{with } b_d \neq 0 \end{array}$$

Consider the polynomial $P_1(X) = P(X) - a_{k+1}b_d^{-1}X^{k+1-d}D(X) \in \mathbb{K}[X]$. From Proposition 2.6, we have:

$$\deg(a_{k+1}b_d^{-1}X^{k+1-d}D(X)) = \deg(a_{k+1}b_d^{-1}X^{k+1-d}) + \deg(D) = (k+1-d) + d = k+1$$

and

$$\deg(P_1) \leqslant \max\Big\{\deg(P), \deg(a_{k+1}b_d^{-1}X^{k+1-d}D(X))\Big\} = \max\{k+1, k+1\} = k+1$$

Moreover the coefficient of the term in P_1 with degree $k+1$ is $a_{k+1} - a_{k+1}b_d^{-1}b_d = 0$. Then we have $\deg(P_1) \leqslant k$. By inductive hypothesis $\mathcal{P}_k$ applied to $P_1(X)$, there exist two polynomials $Q_1(X)$ and $R_1(X)$ such that

$$\begin{cases} P_1(X) = Q_1(X)D(X) + R_1(X) \\ \deg(R_1) < d \end{cases}$$

Now take $Q(X) = a_{k+1}b_d^{-1}X^{k+1-d} + Q_1(X)$ and $R(X) = R_1(X)$ then we get

$$\begin{cases} P(X) = a_{k+1}b_d^{-1}X^{k+1-d}D(X) + P_1(X) = Q(X)D(X) + R(X) \\ \deg(R) < d \end{cases}$$

Consequently $\mathcal{P}_{k+1}$ is satisfied and the conclusion follows by induction.

Uniqueness. By contradiction, assume $Q_1(X)$, $R_1(X)$ and $Q_2(X)$, $R_2(X)$ are such that

$$\begin{cases} P(X) = Q_1(X)D(X) + R_1(X) \\ \deg(R_1) < \deg(D) \end{cases} \quad \text{and} \quad \begin{cases} P(X) = Q_2(X)D(X) + R_2(X) \\ \deg(R_2) < \deg(D) \end{cases}$$

Then $Q_1(X)D(X) + R_1(X) = Q_2(X)D(X) + R_2(X)$ or equivalently

$$(Q_1(X) - Q_2(X))\, D(X) = R_2(X) - R_1(X)$$

From Proposition 2.6, we have:

$$\begin{cases} \deg\left((Q_1 - Q_2)\, D\right) &= \deg\left(Q_1 - Q_2\right) + \deg(D) \\ \deg(R_2 - R_1) &\leqslant \max\left\{\deg(R_1), \deg(R_2)\right\} < \deg(D) \end{cases}$$

So we get $\deg(Q_1 - Q_2) < 0$ (since $D(X) \neq 0$ implies $\deg(D) \geqslant 0$) that is $\deg(Q_1 - Q_2) = -\infty$ and hence $Q_1(X) - Q_2(X) = 0$. It follows $Q_1(X) = Q_2(X)$ and hence $R_1(X) = R_2(X)$. Finally the quotient and the remainder of the polynomial division algorithm are unique. ■

Examples : In order to compute the quotient and the remainder of a polynomial division algorithm, one may use a long division algorithm as follows

a) For $P(X) = X^3 - 12X^2 - 42$ and $D(X) = X - 3$

$$\begin{array}{rrcrll} X^3 & & = & X^2 & (X-3) & +3X^2 \\ -12X^2 & +3X^2 & = & -9X & (X-3) & -27X \\ 0 & -27X & = & -27 & (X-3) & -81 \\ -42 & -81 & = & 0 & (X-3) & -123 \end{array}$$

and the sum of all these equalities gives after simplifications

$$X^3 - 12X^2 - 42 = (X^2 - 9X - 27)(X - 3) - 123$$

that is $Q(X) = X^2 - 9X - 27$ and $R(X) = -123$

b) For $P(X) = X^4 + 7X^3 - 3X^2 - 11X + 5$ and $D(X) = X^2 - 2X - 3$

$$\begin{array}{rrcrl} X^4 & & = & X^2 & (X^2-2X-3) \quad +2X^3+3X^2 \\ 7X^3 & +2X^3 & = & 9X & (X^2-2X-3) \quad +18X^2+27X \\ -3X^2 & +3X^2+18X^2 & = & 18 & (X^2-2X-3) \quad +36X+54 \\ -11X & +27X+36X & = & 0 & (X^2-2X-3) \quad +52X \\ 5 & +54 & = & 0 & (X^2-2X-3) \quad +59 \end{array}$$

and the sum of all these equalities gives after simplifications

$$X^4 + 7X^3 - 3X^2 - 11X + 5 = (X^2 + 9X + 18)(X^2 - 2X - 3) + 52X + 59$$

that is $Q(X) = X^2 + 9X + 18$ and $R(X) = 52X + 59$

c) For $P(X) = X^5 + X^4 - X - 1$ and $D(X) = X^2 - 1$

$$\begin{array}{rrcrl} X^5 & & = & X^3 & (X^2-1) \quad +X^3 \\ X^4 & & = & X^2 & (X^2-1) \quad +X^2 \\ 0 & +X^3 & = & X & (X^2-1) \quad +X \\ 0 & +X^2 & = & 1 & (X^2-1) \quad +1 \\ -X & +X & = & 0 & (X^2-1) \\ -1 & +1 & = & 0 & (X^2-1) \end{array}$$

and the sum of all these equalities gives after simplifications

$$X^5 + X^4 - X - 1 = (X^3 + X^2 + X + 1)(X^2 - 1)$$

that is $Q(X) = X^3 + X^2 + X + 1$ and $R(X) = 0$

Definition 2.8 (*Multiple and divisor*)

*Let $A(X)$ and $B(X)$ be two polynomials in $\mathbb{K}[X]$. We say $B(X)$ **divides** $A(X)$ or equivalently $A(X)$ is a **multiple** of $B(X)$ if*

$$\exists Q(X) \in \mathbb{K}[X] / \ A(X) = Q(X)B(X)$$

In this case, we write $\boldsymbol{B(X)|A(X)}$ or shortly $\boldsymbol{B|A}$.

Remark : A polynomial $B(X) \in \mathbb{K}[X]$ divides a polynomial $A(X) \in \mathbb{K}[X]$ if and only if

- either $B(X) = 0$ and $A(X) = 0$
- or $B(X) \neq 0$ and the remainder from the polynomial division algorithm with dividend $A(X)$ and divisor $B(X)$ is $R(X) = 0$

Proposition 2.9

Let $A(X)$ and $B(X)$ be two polynomials in $\mathbb{K}[X]$. If $B|A$ with $A(X) \neq 0$ then

$$\deg(B) \leqslant \deg(A)$$

Proof : There exists a polynomial $Q(X) \in \mathbb{K}[X]$ such that $A(X) = Q(X)B(X)$. Moreover $Q(X) \neq 0$ since $A(X) \neq 0$. In particular $\deg(Q) \geqslant 0$ and from Proposition 2.6

$$\deg(A) = \deg(QB) = \deg(Q) + \deg(B) \geqslant \deg(B)$$

■

Proposition 2.10

The ring $(\mathbb{K}[X], +, \times)$ has no zero divisor. That is

$$\forall (A(X), B(X)) \in (\mathbb{K}[X])^2,\ A(X)B(X) = 0 \Longrightarrow \textit{either } A(X) = 0 \textit{ or } B(X) = 0$$

Proof : Assume $A(X)B(X) = 0$ with $B(X) \neq 0$. The polynomial division algorithm with dividend 0 and divisor $B(X)$ is

$$0 = 0 \times B(X) + 0$$

But we have

$$0 = A(X) \times B(X) + 0$$

Consenquently the uniqueness part of Theorem 2.7 gives $A(X) = 0$. ■

2.2 Polynomial maps

2.2.1 Definition

Definition 2.11 (*Polynomial map*)

*Let $P(X)$ be a polynomial in $\mathbb{K}[X]$ with coefficients $(a_n)_{n\in\mathbb{N}}$. The **polynomial map** associated to $P(X)$ is the following map*

$$\begin{array}{rccl} P: & \mathbb{K} & \to & \mathbb{K} \\ & x & \mapsto & P(x) = \displaystyle\sum_{n=0}^{+\infty} a_n x^n \end{array}$$

Some examples :

a) The polynomial map associated to $P(X) = 0$ is the constant map $x \mapsto 0$.

b) The polynomial map associated to $P(X) = X$ is the identity map $x \mapsto x$.

c) The polynomial map associated to $P(X) = -1 + 2X + X^3 \in \mathbb{R}[X]$ is the cubic map

$$\begin{array}{rccl} P: & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & P(x) = x^3 + 2x - 1 \end{array}$$

d) Recall that $\mathbb{Z}/2\mathbb{Z} = \{\overline{0}, \overline{1}\}$ is a commutative field since 2 is a prime number. Moreover we have

$$\overline{0} + \overline{0} \times \overline{0} = \overline{0} \quad \text{and} \quad \overline{1} + \overline{1} \times \overline{1} = \overline{2} = \overline{0}$$

Consequently the polynomial map associated to $P(X) = X + X^2 \in \mathbb{Z}/2\mathbb{Z}[X]$ is the constant map

$$\begin{array}{rccl} P: & \mathbb{Z}/2\mathbb{Z} & \to & \mathbb{Z}/2\mathbb{Z} \\ & x & \mapsto & P(x) = \overline{0} \end{array}$$

But notice $P(X) \neq 0$.

Proposition 2.12

Denote $\mathcal{F}(\mathbb{K})$ the set of all functions from $\mathbb{K}$ to itself. $\mathcal{F}(\mathbb{K})$ has a natural ring structure coming from that one of $\mathbb{K}$. Then the following map

$$\begin{array}{rcl} \mathbb{K}[X] & \to & \mathcal{F}(\mathbb{K}) \\ P(X) & \mapsto & P \end{array}$$

is a ring homomorphism. In particular for any given $\alpha \in \mathbb{K}$, the following map

$$\begin{array}{rcl} \mathbb{K}[X] & \to & \mathbb{K} \\ P(X) & \mapsto & P(\alpha) \end{array}$$

is a ring homomorphism as well.

Proof: Everything comes from the definitions of addition and multiplication of two polynomials and from the ring homomorphism $\mathcal{F}(\mathbb{K}) \to \mathbb{K},\ f \mapsto f(\alpha)$. ■

2.2.2 Derivative polynomial

Definition 2.13 (*Derivative polynomial*)

*Let $P(X)$ be a polynomial in $\mathbb{K}[X]$ with coefficients $(a_n)_{n\in\mathbb{N}}$. The **derivative polynomial** of $P(X)$ is the following polynomial*

$$P'(X) = a_1 + 2a_2X + 3a_3X^2 + \cdots = \sum_{n=0}^{+\infty}(n+1)a_{n+1}X^n = \sum_{n=1}^{+\infty} na_nX^{n-1}$$

*By induction over $k \geqslant 1$, we define the $\boldsymbol{k^{th}}$ **order polynomial derivative**, denoted $\boldsymbol{P^{(k)}(X)}$, to be the polynomial derivative of $P^{(k-1)}(X)$ with the notation $P^{(0)} = P$ (and then $P^{(1)} = P'$).*

Remarks:

- Notice that any integer $n \in \mathbb{N}$ may be considered in $\mathbb{K}$ if we write it as follows
$$n = \underbrace{1+1+1+\cdots+1}_{n \text{ times}} \in \mathbb{K} \quad \text{with } 1 \in \mathbb{K}$$
In particular the derivative polynomial of any polynomial with coefficients in $\mathbb{K}$ is well in $\mathbb{K}[X]$.
- The polynomial map P' associated to the derivative polynomial $P'(X)$ of a polynomial $P(X)$ is the derivative of the map P as expected. But limit, differentiation or any calculus tool are not needed here to define the derivative of a polynomial map.

Example: Consider $P(X) = -7 + 8X - 5X^2 + 2X^3 - X^5 \in \mathbb{R}[X]$. Then

$$\begin{array}{rcl} P'(X) = P^{(1)}(X) & = & 8 - 10X + 6X^2 - 5X^4 \\ P^{(2)}(X) & = & -10 + 12X - 20X^3 \\ P^{(3)}(X) & = & 12 - 60X^2 \\ P^{(4)}(X) & = & -120X \\ P^{(5)}(X) & = & -120 \\ P^{(6)}(X) & = & 0 \\ \text{etc.} & & \end{array}$$

Proposition 2.14

The following properties hold

1. $\forall (P(X), Q(X)) \in (\mathbb{K}[X])^2 \ \left\{ \begin{array}{l} (P+Q)'(X) = P'(X) + Q'(X) \\ (PQ)'(X) = P'(X)Q(X) + P(X)Q'(X) \end{array} \right.$
2. *Consider the polynomial* $P(X) = X^n$ *with* $n \in \mathbb{N}$. *Then*
$$\forall k \geqslant 1, \ (P)^{(k)}(X) = \left\{ \begin{array}{cl} \frac{n!}{(n-k)!} X^{n-k} & \text{if } 1 \leqslant k \leqslant n \\ 0 & \text{if } k \geqslant n+1 \end{array} \right.$$
where $\frac{n!}{(n-k)!} = n(n-1)(n-2)\dots(n-k+1)$
3. $P(X) \in \mathbb{K}[X]$ *is a constant polynomial if and only if* $P'(X) = 0$
4. *If* $P(X) \in \mathbb{K}[X]$ *is a non constant polynomial then* $\deg(P') = \deg(P) - 1$
More generally, if $P^{(k)}(X) \neq 0$ *for some* $k \geqslant 1$ *then* $\deg(P^{(k)}) = \deg(P) - k$

Proof : 1. Denote respectively $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ the coefficients of $P(X)$ and $Q(X)$. Then
$$(P+Q)'(X) = \sum_{n=1}^{+\infty} n(a_n + b_n)X^{n-1} = \sum_{n=1}^{+\infty} n a_n X^{n-1} + \sum_{n=1}^{+\infty} n b_n X^{n-1} = P'(X) + Q'(X)$$
And, using new indices $n' = n-1$, $k' = k-1$ and $\ell' = \ell - 1$, we get
$$\begin{aligned}
(PQ)'(X) &= \sum_{n=1}^{+\infty} n \left(\sum_{\substack{k,\ell\in\mathbb{N} \\ k+\ell=n}} a_k b_\ell \right) X^{n-1} \\
&= \sum_{n=1}^{+\infty} \left(\sum_{\substack{k,\ell\in\mathbb{N} \\ k+\ell=n}} (k+\ell) a_k b_\ell \right) X^{n-1} \\
&= \sum_{n'=0}^{+\infty} \left(\sum_{\substack{k',\ell\in\mathbb{N} \\ k'+\ell=n'}} (k'+1) a_{k'+1} b_\ell \right) X^{n'} + \sum_{n'=0}^{+\infty} \left(\sum_{\substack{k,\ell'\in\mathbb{N} \\ k+\ell'=n'}} a_k (\ell'+1) b_{\ell'+1} \right) X^{n'} \\
&= \left(\sum_{k'=0}^{+\infty} (k'+1) a_{k'+1} X^{k'} \right) \left(\sum_{\ell=0}^{+\infty} b_\ell X^\ell \right) + \left(\sum_{k=0}^{+\infty} a_k X^k \right) \left(\sum_{k'=0}^{+\infty} (\ell'+1) b_{\ell'+1} X^{\ell'} \right) \\
&= P'(X)Q(X) + P(X)Q'(X)
\end{aligned}$$

2. An induction over the order $k \geqslant 1$ gives the result.
3. If $P(X) = a_0 \in \mathbb{K}$ then $P'(X) = 0$ by definition of the polynomial derivative. Conversely, $P'(X) = \sum_{n=1}^{+\infty} n a_n X^{n-1} = 0$ implies that $a_n = 0$ for every $n \geqslant 1$. The result follows.
4. Let $P(X)$ be a non constant polynomial in $\mathbb{K}[X]$ and denote by $a_{\deg(P)}$ its leading coefficient. By definition of the polynomial derivative, we have $\deg(P') \leqslant \deg(P) - 1$. Moreover the coefficient of the term with degree $n = \deg(P) - 1$ is $\deg(P) a_{\deg(P)}$ which is not zero since $\deg(P) \geqslant 1$ (P is non constant) and $a_{\deg(P)} \neq 0$ (as leading coefficient of $P(X)$). Thus, we have $\deg(P') = \deg(P) - 1$. The remain follows by induction over the order $k \geqslant 1$. ■

Corollary 2.15

Let $P(X)$ be a polynomial in $\mathbb{K}[X]$ of degree $d = \deg(P)$. Then

$$\forall k \geqslant d+1,\ P^{(k)}(X) = 0$$

Theorem 2.16 (*Exact Taylor's formula*)

Let $P(X)$ be a polynomial in $\mathbb{K}[X]$ of degree $d = \deg(P)$. Then

$$P(X) = P(0) + P'(0)X + \frac{P^{(2)}(0)}{2}X^2 + \cdots + \frac{P^{(d)}(0)}{d!}X^d = \sum_{n=0}^{d} \frac{P^{(n)}(0)}{n!}X^n$$

where $\frac{1}{n!} = (1 \times 2 \times 3 \times \cdots \times n)^{-1}$

More generally, for any $a \in \mathbb{K}$ we have

$$P(X) = \sum_{n=0}^{d} \frac{P^{(n)}(a)}{n!}(X-a)^n$$

Proof: Call $R(X)$ the following polynomial

$$R(X) = P(X) - \sum_{n=0}^{d} \frac{P^{(n)}(a)}{n!}(X-a)^n$$

We will prove by induction for every integer k with $0 \leqslant k \leqslant d$ that $R^{(d-k)}(X) = 0$. In particular, $k = d$ will give the result. At first, using Proposition 2.6, we have:

$$\det(R) \leqslant \max\left\{\deg(P), \deg\left(\sum_{n=0}^{d} \frac{P^{(n)}(a)}{n!}(X-a)^n\right)\right\} \leqslant d$$

Then from Corollary 2.15 and Proposition 2.14, we get $(R^{(d)})'(X) = R^{(d+1)}(X) = 0$ that is $R^{(d)}(X)$ is a constant polynomial. Consequently Proposition 2.14 gives

$$R^{(d)}(X) = R^{(d)}(a) = P^{(d)}(a) - \sum_{n=0}^{d-1} \frac{P^{(n)}(a)}{n!}0 - \frac{P^{(d)}(a)}{d!}d! = P^{(d)}(a) - P^{(d)}(a) = 0$$

So the inductive hypothesis is true for $k = 0$. Now assume the inductive hypothesis is satisfied for a given integer k with $0 \leqslant k \leqslant d-1$. Then $(R^{(d-k-1)})'(X) = R^{(d-k)}(X) = 0$ that is $R^{(d-k-1)}(X)$ is a constant polynomial. Consequently Proposition 2.14 gives

$$\begin{aligned}
R^{(d-k-1)}(X) &= R^{(d-k-1)}(a) \\
&= P^{(d-k-1)}(a) - \sum_{n=0}^{d-k-2} \frac{P^{(n)}(a)}{n!}0 - \frac{P^{(d-k-1)}(a)}{(d-k-1)!}(d-k-1)! \\
&\quad - \sum_{n=d-k}^{d} \frac{P^{(n)}(a)}{n!}\frac{n!}{(n-(d-k-1))!}(a-a)^n \\
&= P^{(d-k-1)}(a) - 0 - P^{(d-k-1)}(a) - 0 \\
&= 0
\end{aligned}$$

Finally the inductive hypothesis is still true for $k+1$. The result follows by induction. ■

2.2.3 Root

Definition 2.17 (*Root*)

*$\alpha \in \mathbb{K}$ is said to be a **root** of a polynomial $P(X) \in \mathbb{K}[X]$ if $P(\alpha) = 0$.*

Example : 1 and 3 are roots of $P(X) = 3-4X+X^2$ since $P(1) = 3-4+1 = 0$ and $P(3) = 3-12+9 = 0$. Actually $P(X) = (X-1)(X-3)$ in order that $(X-1)|P(X)$ and $(X-3)|P(X)$.

Example : 1 and -1 are roots of $P(X) = 1 - 2X^2 + X^4 = (X^2-1)^2 = (X-1)^2(X+1)^2$.

Proposition 2.18

$\alpha \in \mathbb{K}$ is a root of $P(X) \in \mathbb{K}[X]$ if and only if $(X-\alpha)|P(X)$.

Proof : Sufficient. If $(X-\alpha)|P(X)$ then there exists a polynomial $Q(X) \in \mathbb{K}[X]$ such that $P(X) = (X-\alpha)Q(X)$ and then $P(\alpha) = (\alpha-\alpha)Q(\alpha) = 0$.

Necessary. From Theorem 2.7, we get two polynomials $Q(X)$ and $R(X)$ such that

$$\begin{cases} P(X) = (X-\alpha)Q(X) + R(X) \\ \deg(R) < \deg(X-\alpha) = 1 \end{cases}$$

In particular, $R(X)$ is a constant polynomial that is $R(X) = r \in \mathbb{K}$. But α is a root of $P(X)$ implies

$$0 = P(\alpha) = (\alpha-\alpha)Q(\alpha) + r = r$$

Consequently $R(X) = 0$ and $P(X) = (X-\alpha)Q(X)$ as needed. ■

Further example :

a) The polynomial $P(X) = 1 + X^2 \in \mathbb{R}[X]$ has no root since $\forall x \in \mathbb{R},\ P(x) = 1 + x^2 \geqslant 1 > 0$. In particular, $P(X)$ can not be written as a product of two linear polynomials in $\mathbb{R}[X]$.

b) But $\imath$ and $-\imath$ are roots of $Q(X) = 1 + X^2 \in \mathbb{C}[X]$ since $Q(X) = (X-\imath)(X+\imath)$.

Definition 2.19 (*Root of higher multiplicity*)

*Let $k \geqslant 1$ be a positive integer. $\alpha \in \mathbb{K}$ is said to be a **root of multiplicity k** of a polynomial $P(X) \in \mathbb{K}[X]$ if $(X-\alpha)^k|P(X)$ and $(X-\alpha)^{k+1} \nmid P(X)$, or equivalently if*

$$\exists Q(X) \in \mathbb{K}[X] /\ P(X) = (X-\alpha)^k Q(X) \text{ and } Q(\alpha) \neq 0$$

*Furthermore, the **multiplicity** of a root $\alpha \in \mathbb{K}$ of a polynomial $P(X) \neq 0$ is the following positive integer*

$$k_\alpha = \max\left\{k \geqslant 1 \ /\ (X-\alpha)^k | P(X)\right\}$$

Proposition 2.20

$\alpha \in \mathbb{K}$ is a root of multiplicity $k \geqslant 1$ of the polynomial $P(X) \neq 0$ if and only if

$$P(\alpha) = P'(\alpha) = P^{(2)}(\alpha) = \cdots = P^{(k-1)}(\alpha) = 0 \quad \text{and} \quad P^{(k)}(\alpha) \neq 0$$

Proof : Sufficient. The Taylor's formula (see Theorem 2.16) gives

$$\begin{aligned} P(X) &= \sum_{n=0}^{+\infty} \frac{P^{(n)}(\alpha)}{n!}(X-\alpha)^n \\ &= \sum_{n=0}^{k-1} \frac{P^{(n)}(\alpha)}{n!}(X-\alpha)^n + \sum_{n=k}^{+\infty} \frac{P^{(n)}(\alpha)}{n!}(X-\alpha)^n \\ &= 0 + (X-\alpha)^k \sum_{n=0}^{+\infty} \frac{P^{(n+k)}(\alpha)}{(n+k)!}(X-\alpha)^n \\ &= (X-\alpha)^k Q(X) \end{aligned}$$

with $Q(\alpha) = \frac{P^{(k)}(\alpha)}{k!} + \sum_{n=1}^{+\infty} \frac{P^{(n+k)}(\alpha)}{(n+k)!}(\alpha-\alpha)^n = \frac{P^{(k)}(\alpha)}{k!} + 0 \neq 0$ since $P^{(k)}(\alpha) \neq 0$

Necessary. If $P(X) = (X-\alpha)^k Q(X)$ with $Q(\alpha) \neq 0$ then Proposition 2.14 gives

$$P(\alpha) = P'(\alpha) = P^{(2)}(\alpha) = \cdots = P^{(k-1)}(\alpha) = 0 \quad \text{and} \quad P^{(k)}(\alpha) = k!Q(\alpha) \neq 0$$

■

Proposition 2.21

Let $P(X) \neq 0$ be a polynomial in $\mathbb{K}[X]$. Denote by $k_1, k_2, \ldots, k_n$ the multiplicities of the roots of $P(X)$. Then

$$k_1 + k_2 + \cdots + k_n \leqslant \deg(P)$$

In particular $P(X)$ has at most $\deg(P)$ roots.

Proof : Denote by $\alpha_1, \alpha_2, \ldots, \alpha_n$ the roots of $P(X)$ associated to the multiplicities $k_1, k_2, \ldots, k_n$. Then the polynomial $(X-\alpha_1)^{k_1}(X-\alpha_2)^{k_2}\ldots(X-\alpha_n)^{k_n}$ divides $P(X)$. But from Proposition 2.6 we have:

$$\deg\left((X-\alpha_1)^{k_1}(X-\alpha_2)^{k_2}\ldots(X-\alpha_n)^{k_n}\right) = k_1 + k_2 + \cdots + k_n$$

Hence, the conclusion follows from Proposition 2.9. ■

To conclude, just state the following important and powerful result without proof.

Theorem 2.22 (*Fundamental theorem of algebra*)

Every non constant polynomial in $\mathbb{C}[X]$ has at least one root.

Remark : In particular, the inequality of Proposition 2.21 becomes an equality in $\mathbb{C}[X]$. More precisely, any non constant polynomial $P(X) \in \mathbb{C}[X]$ may be written as a product of linear polynomials:

$$P(X) = C(X-\alpha_1)^{k_1}(X-\alpha_2)^{k_2}\ldots(X-\alpha_n)^{k_n}$$

where

- $C \in \mathbb{C}^*$ is the leading coefficient of $P(X)$
- $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the roots of $P(X)$
- $k_1, k_2, \ldots, k_n$ are their associated multiplicities

Chapter 3

Vector spaces

In this chapter, fix a commutative field $(\mathbb{K}, +, \times)$ (for instance $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$).

3.1 The structure of vector space

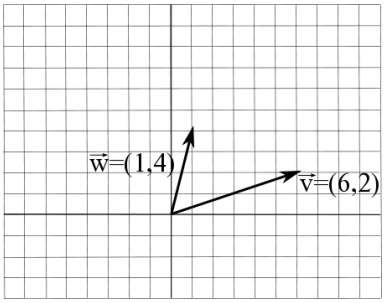
3.1.1 First examples

Example of the real plane $\mathbb{R}^2$:

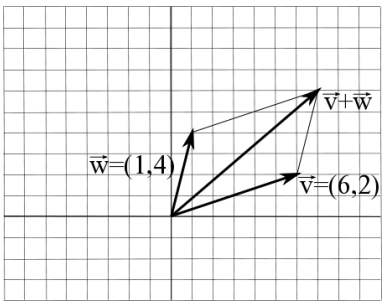
The set of all vectors in the plane consists of the set of all arrows starting at one fixed point in the plane . We may write it as follows

$$\mathbb{R}^2 = \{\overrightarrow{v} = (x, y) \ / \ x \in \mathbb{R} \text{ and } y \in \mathbb{R}\}$$

The real numbers x and y are called the **coordinates** of the vector $\overrightarrow{v}$.



Given two arrows $\overrightarrow{v}$ and $\overrightarrow{w}$ starting at one fixed point in a plane, the parallelogram spanned by these two arrows contains one diagonal arrow which starts at the same fixed point. This new arrow defines the sum of $\overrightarrow{v}$ and $\overrightarrow{w}$.

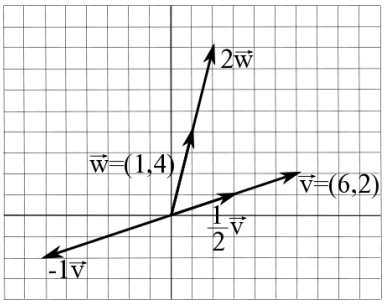


More precisely, the addition of two vectors in $\mathbb{R}^2$ is defined by the following binary operation

$$\begin{array}{rccl} + : & \mathbb{R}^2 \times \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & (\overrightarrow{v}, \overrightarrow{w}) & \longmapsto & \overrightarrow{v} + \overrightarrow{w} = (x + x', y + y') \quad \text{where} \left\{ \begin{array}{l} \overrightarrow{v} = (x, y) \\ \overrightarrow{w} = (x', y') \end{array} \right. \end{array}$$

Notice that this binary operation provides an abelian group structure on $\mathbb{R}^2$.

Moreover, any arrow $\overrightarrow{v}$ starting at one fixed point in a plane may be scaled: given any positive real number λ, the scaling of $\overrightarrow{v}$ by λ is the arrow whose diection is the same as $\overrightarrow{v}$ but is dilated or shrunk by multiplying its length by λ. When λ is negative, the scaling of $\overrightarrow{v}$ by λ is defined as the arrow pointing in the opposite direction, instead.



More precisely, the multiplication of a vector in $\mathbb{R}^2$ by a scalar is defined as follows

$$\begin{array}{rccl} . : & \mathbb{R} \times \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & (\lambda, \overrightarrow{v}) & \longmapsto & \lambda.\overrightarrow{v} = (\lambda x, \lambda y) \quad \text{where } \overrightarrow{v} = (x, y) \end{array}$$

But notice that . is not a binary operation.

Example of the real space $\mathbb{R}^3$:

Similarly to the plane, it is the same for the space. Indeed the set of all vectors in the space consists of the set of all arrows starting at one fixed point in the space.

$$\mathbb{R}^3 = \{\overrightarrow{v} = (x, y, z) \ / \ x \in \mathbb{R},\ y \in \mathbb{R} \text{ and } z \in \mathbb{R}\}$$

Each vector in the space $\mathbb{R}^3$ has three coordinates (instead of two for a vector in the plane $\mathbb{R}^2$).

The addition of two vectors in $\mathbb{R}^3$ is defined by the following binary operation

$$\begin{array}{rccl} + : & \mathbb{R}^3 \times \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & (\overrightarrow{v}, \overrightarrow{w}) & \longmapsto & \overrightarrow{v} + \overrightarrow{w} = (x + x', y + y', z + z') \quad \text{where} \left\{ \begin{array}{l} \overrightarrow{v} = (x, y, z) \\ \overrightarrow{w} = (x', y', z') \end{array} \right. \end{array}$$

The multiplication of a vector in $\mathbb{R}^3$ by a scalar is defined as follows

$$\begin{array}{rccl} . : & \mathbb{R} \times \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & (\lambda, \overrightarrow{v}) & \longmapsto & \lambda.\overrightarrow{v} = (\lambda x, \lambda y, \lambda z) \quad \text{where } \overrightarrow{v} = (x, y, z) \end{array}$$

We get a similar algebraic structure as for the real plane $\mathbb{R}^2$: a binary operation $+$ such that $(\mathbb{R}^3, +)$ is an abelian group and a multiplication by a scalar which is not a binary operation.

3.1.2 Definitions

Definition 3.1 (*External binary operation*)

*An **external binary operation** (or external binary law) over $\mathbb{K}$ on a nonempty set S is a map from $\mathbb{K} \times S$ to S. Such a binary operation is usually denoted*

$$\begin{array}{rccc} . : & \mathbb{K} \times S & \longrightarrow & S \\ & (\lambda, x) & \longmapsto & \lambda.x \end{array}$$

Definition 3.2 (*Vector space*)

*Let $(V, +, .)$ be a nonempty set with a binary operation denoted $+$ and an external binary operation over $\mathbb{K}$ denoted $.$ and called the **scalar multiplication**. $(V, +, .)$ (or simply V) is said to be a **vector space over** $\mathbb{K}$ (or a $\mathbb{K}$**-vector space**) if it satisfies each of the following vector space axioms*

i) *$(V, +)$ is an abelian group*

ii) *the scalar multiplication $.$ is **distributive** over the binary operation $+$ on V:*

$$\forall \lambda \in \mathbb{K},\ \forall (v, w) \in V^2,\ \lambda.(v + w) = \lambda.v + \lambda.w$$

iii) *the scalar multiplication $.$ is **distributive** over the binary operation $+$ on $\mathbb{K}$:*

$$\forall (\lambda, \mu) \in \mathbb{K}^2,\ \forall v \in V,\ (\lambda + \mu).v = \lambda.v + \mu.v$$

iv) *the scalar multiplication $.$ is **compatible** with the binary operation $\times$ on $\mathbb{K}$:*

$$\forall (\lambda, \mu) \in \mathbb{K}^2,\ \forall v \in V,\ (\lambda \times \mu).v = \lambda.(\mu.v)$$

v) *$1_{\mathbb{K}} \in \mathbb{K}$ is the **identity element** for the scalar multiplication $.$:*

$$\forall v \in V,\ 1_{\mathbb{K}}.v = v$$

*In this case, the elements of $\mathbb{K}$ are called **scalars** and those of V are called **vectors**. Furthermore if $\lambda_1, \lambda_2, \ldots \lambda_n$ are scalars and $v_1, v_2, \ldots, v_n$ are vectors, then the **linear combination** of those vectors with those scalars is given by*

$$\lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_n.v_n = \sum_{k=1}^{n} \lambda_k.v_k \in V$$

Examples : The real plane $\mathbb{R}^2$ and the real space $\mathbb{R}^3$ are vector spaces over $\mathbb{R}$.

Proof : Notice that the scalar multiplication $.$ on $\mathbb{R}^2$ or $\mathbb{R}^3$ corresponds for every coordinate to the multiplication $\times$ on $\mathbb{R}$. Consequently, the distributivity of the scalar multiplication $.$ over the addition $+$ on V or $\mathbb{R}$ follows from the distributivity of the multiplication $\times$ over the addition $+$ on $\mathbb{R}$, the compatibility of the scalar multiplication $.$ with the multiplication $\times$ on $\mathbb{R}$ follows from the associativity of the multiplication $\times$ on $\mathbb{R}$, and $1_{\mathbb{R}} \in \mathbb{R}$ is the identity element for the scalar multiplication $.$ since it is the identity element for the multiplication $\times$ on $\mathbb{R}$.

Proposition 3.3

Let $(V, +, .)$ be a vector space over $\mathbb{K}$. Then the following properties hold

1. $\forall \lambda \in \mathbb{K},\ \lambda.0_V = 0_V$
2. $\forall v \in V,\ 0_{\mathbb{K}}.v = 0_V$
3. $\forall \lambda \in \mathbb{K},\ \forall v \in V,\ \lambda.v = 0_V \Longrightarrow$ *either* $\lambda = 0_{\mathbb{K}}$ *or* $v = 0_V$
4. $\forall \lambda \in \mathbb{K},\ \forall v \in V,\ (-\lambda).v = -(\lambda.v)$

Proof: 1. From the distributivity of the scalar multiplication over $+$ on V we get

$$\lambda.0_V + \lambda.0_V = \lambda.(0_V + 0_V) = \lambda.0_V = 0_V + \lambda.0_V$$

Now, a simplification on each side by $\lambda.0_V$ gives $\lambda.0_V = 0_V$ as needed.

2. From the distributivity of the scalar multiplication over $+$ on $\mathbb{K}$ we get

$$0_{\mathbb{K}}.v + 0_{\mathbb{K}}.v = (0_{\mathbb{K}} + 0_{\mathbb{K}}).v = 0_{\mathbb{K}}.v = 0_V + 0_{\mathbb{K}}.v$$

Now, a simplification on each side by $0_{\mathbb{K}}.v$ gives $0_{\mathbb{K}}.v = 0_V$ as needed.

3. Assume $\lambda \neq 0_{\mathbb{K}}$. Consequently, there exists a multiplicative inverse $\lambda^{-1} \in \mathbb{K}$ and using the compatibility of the scalar multiplication with $\times$ on $\mathbb{K}$ and the first point, we get

$$v = 1_{\mathbb{K}}.v = (\lambda^{-1} \times \lambda).v = \lambda^{-1}.(\lambda.v) = \lambda^{-1}.0_V = 0_V$$

Finally either $\lambda = 0_{\mathbb{K}}$ or $v = 0_V$.

4. From the distributivity of the scalar multiplication over $+$ on $\mathbb{K}$ and the second point we get

$$\lambda.v + (-\lambda.v) = (\lambda + (-\lambda)).v = 0_{\mathbb{K}}.v = 0_V$$

■

3.1.3 More examples

Definition 3.4 (*Coordinate space*)

Fix a positive integer $n \in \mathbb{N}^*$. The ***coordinate space*** *of dimension n is the set of all n-tuples of elements of $\mathbb{K}$ which is denoted*

$$\mathbb{K}^n = \{v = (x_1, x_2, \ldots, x_n) \ / \ x_1 \in \mathbb{K}, x_2 \in \mathbb{K}, \ldots, \text{ and } x_n \in \mathbb{K}\}$$

It is a $\mathbb{K}$-vector space for the following operations

$$\begin{array}{rccl} + : & \mathbb{K}^n \times \mathbb{K}^n & \longrightarrow & \mathbb{K}^n \\ & (v, w) & \longmapsto & v + w = (x_1 + x'_1, x_2 + x'_2, \ldots, x_n + x'_n) \quad \text{where} \left\{ \begin{array}{l} v = (x_1, x_2, \ldots, x_n) \\ w = (x'_1, x'_2, \ldots, x'_n) \end{array} \right. \end{array}$$

$$\begin{array}{rccl} . : & \mathbb{K} \times \mathbb{K}^n & \longrightarrow & \mathbb{K}^n \\ & (\lambda, v) & \longmapsto & \lambda.v = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n) \quad \text{where } v = (x_1, x_2, \ldots, x_n) \end{array}$$

The elements $x_1, x_2, \ldots, x_n$ in $\mathbb{K}$ are called the ***coordinates*** *of the vector $v = (x_1, x_2, \ldots, x_n)$.*

Examples :

a) For $\mathbb{K} = \mathbb{R}$, we get the $\mathbb{R}$-vector space $(\mathbb{R}^n, +, .)$ called the **real coordinate space** (or simply the **real plane** in case $n = 2$ and the **real space** in case $n = 3$). And for $\mathbb{K} = \mathbb{C}$, we get the $\mathbb{C}$-vector space $(\mathbb{C}^n, +, .)$ called the **complex coordinate space**.

b) In particular $n = 1$ and $\mathbb{K} = \mathbb{R}$ gives that $\mathbb{R}$ is a $\mathbb{R}$-vector space. In this case, the scalar multiplication . corresponds to the multiplication $\times$ on $\mathbb{R}$ (and then the external binary operation is actually a binary operation).

c) Similarly, $\mathbb{C}$ is a $\mathbb{C}$-vector space.

d) But $\mathbb{C}$ is also a $\mathbb{R}$-vector space. To show this, we may identify $\mathbb{C}$ and its operations as follows

$$\mathbb{C} = \{z = a + \imath b \ / \ a \in \mathbb{R} \text{ and } b \in \mathbb{R}\}$$

$$\begin{array}{rccl} + : & \mathbb{C} \times \mathbb{C} & \longrightarrow & \mathbb{C} \\ & (z, z') & \longmapsto & z + z' = (a + a') + \imath(b + b') \quad \text{where } \left\{ \begin{array}{l} z = a + \imath b \\ z' = a' + \imath b' \end{array} \right. \end{array}$$

$$\begin{array}{rccl} . : & \mathbb{R} \times \mathbb{C} & \longrightarrow & \mathbb{C} \\ & (\lambda, z) & \longmapsto & \lambda.z = (\lambda a) + \imath(\lambda b) \quad \text{where } z = a + \imath b \end{array}$$

Actually, $\mathbb{C}$ as a $\mathbb{R}$-vector space is very "similar" to the real plane $\mathbb{R}^2$ (the coordinates of a vector $z = a + \imath b \in \mathbb{C}$ are $a \in \mathbb{R}$ and $b \in \mathbb{R}$).

Further examples :

a) Polynomial space. The ring $\mathbb{K}[X]$ of all polynomials with coefficients in $\mathbb{K}$ is a $\mathbb{K}$-vector space. The addition comes from the ring structure and the scalar multiplication is defined as follows

$$\begin{array}{rccl} . : & \mathbb{K} \times \mathbb{K}[X] & \longrightarrow & \mathbb{K}[X] \\ & (\lambda, P(X)) & \longmapsto & (\lambda.P)(X) = \lambda P(X) = \displaystyle\sum_{n=0}^{+\infty} \lambda a_n X^n \quad \text{where } P(X) = \sum_{n=0}^{+\infty} a_n X^n \end{array}$$

b) Function space. The set $\mathcal{F}(S, V) = \{f : S \to V, x \mapsto f(x)\}$ of all functions from a set S to a $\mathbb{K}$-vector space $(V, +, .)$ is a $\mathbb{K}$-vector space with the following operations

$$\begin{array}{rccl} + : & \mathcal{F}(S, V) \times \mathcal{F}(S, V) & \longrightarrow & \mathcal{F}(S, V) \\ & (f, g) & \longmapsto & (f + g) : x \mapsto f(x) + g(x) \end{array}$$

$$\begin{array}{rccl} . : & \mathbb{K} \times \mathcal{F}(S, V) & \longrightarrow & \mathcal{F}(S, V) \\ & (\lambda, f) & \longmapsto & (\lambda.f) : x \mapsto \lambda.f(x) \end{array}$$

In particular, the set $\mathcal{F}(\mathbb{K}) = \mathcal{F}(\mathbb{K}, \mathbb{K}) = \{f : \mathbb{K} \to \mathbb{K}, x \mapsto f(x)\}$ of all functions from $\mathbb{K}$ to itself is a $\mathbb{K}$-vector space.

c) Sequence space. Choosing S in the previous example to be the set of all natural numbers $\mathbb{N}$, we get the $\mathbb{K}$-vector space $\mathcal{F}(\mathbb{N}, V) = \{u : \mathbb{N} \to V, n \mapsto u_n\}$ of all infinite sequences of elements of V.

3.2 Subspace

3.2.1 Definition and characterization

Definition 3.5 (*Subspace*)

Let $(V,+,.)$ be a $\mathbb{K}$-vector space and $W \subset V$ be a nonempty subset. $(W,+,.)$ (or simply W) is a ***subspace*** *of V over $\mathbb{K}$ if it satisfies each of the following subspace axioms*

i) *$(W,+)$ is a subgroup of $(V,+)$*

ii) *. is an external binary operation over $\mathbb{K}$ on W*

In this case, $(W,+,.)$ is a $\mathbb{K}$-vector space.

Trivial example : For any $\mathbb{K}$-vector space $(V,+,.)$, $(\{0_V\},+,.)$ is a subspace called the **trivial subspace**.

Proposition 3.6

Let $(V,+,.)$ be a $\mathbb{K}$-vector space and $W \subset V$ be a nonempty subset. W is a subspace of V over $\mathbb{K}$ if and only if it satisfies the following condition

$$\forall \lambda \in \mathbb{K},\ \forall (v,w) \in W^2,\ v + \lambda.w \in W$$

Proof : **Necessary.** If $(W,+,.)$ is a $\mathbb{K}$-vector space then $v + \lambda.w = 1_{\mathbb{K}}.v + \lambda.w \in W$ as a linear combination of vectors in W.

Sufficient. $\lambda = -1_{\mathbb{K}}$ gives $v + (-w) \in W$ for every vectors $v \in W$ and $w \in W$. Consequently $(W,+)$ is a subroup of $(V,+)$. In particular 0_V is in W and $v = 0_V$ gives $\lambda.w \in W$ for every $\lambda \in \mathbb{K}$ and $w \in W$ that is . is an external binary operation over $\mathbb{K}$ on W. The conclusion follows. ■

Remark : Equivalently, a nonempty subset W of a vector space $(V,+,.)$ is a subspace if and only if any linear combination of vectors in W belongs to W.

Some examples :

a) Let v be a non zero vector in a $\mathbb{K}$-vector space $(V,+,.)$, that is $v \in V - \{0_V\}$. Consider the following set of all vectors **colinear** to v

$$\mathbb{K}.v = \{\lambda.v \ / \ \lambda \in \mathbb{K}\}$$

Then $\mathbb{K}.v$ is a subspace of V called the **linear line spanned by v**.

Proof : For every scalar $\mu \in \mathbb{K}$ and every vectors $w = \lambda.v$ and $w' = \lambda'.v$ in $\mathbb{K}.v$, we have:

$$w + \mu.w' = (\lambda.v) + \mu.(\lambda'.v) = (\lambda + \mu\lambda').v \in \mathbb{K}.v$$

The conclusion follows from Proposition 3.6. ■

b) Let a, b and c be three real numbers. Consider the following subset of the real space $\mathbb{R}^3$

$$\mathcal{P} = \left\{v = (x,y,z) \in \mathbb{R}^3 \ / \ ax + by + cz = 0\right\}$$

Then $\mathcal{P}$ is a subspace of $\mathbb{R}^3$.

Proof : For every real $\lambda \in \mathbb{R}$ and every vectors $v = (x, y, z)$ and $w = (x, y, z)$ in $\mathcal{P}$ we have $v + \lambda.w = (x + \lambda x', y + \lambda y', z + \lambda z')$ and

$$\begin{aligned} a(x + \lambda x') + b(y + \lambda y') + c(z + \lambda z') &= (ax + by + cz) + \lambda(ax' + by' + cz') \\ &= 0 + \lambda \times 0 \\ &= 0 \end{aligned}$$

Then $v + \lambda.w$ is in $\mathcal{P}$ and the conclusion follows from Proposition 3.6. ■

c) Counterexample. Consider the following subset of the real plane $\mathbb{R}^2$

$$\mathcal{C} = \left\{ v = (x, y) \in \mathbb{R}^2 \ / \ x^2 + y^2 = 1 \right\}$$

Then $\mathcal{C}$ is not a subspace of $\mathbb{R}^2$. Indeed notice that $v = (1, 0)$ and $w = (0, 1)$ are in $\mathcal{C}$ but their linear combination $v + w = (1, 1)$ is not. Furthermore, the zero vector $0_{\mathbb{R}^2} = (0, 0)$ is not in $\mathcal{C}$.

Further examples :

a) Polynomial subspaces. Fix $d \in \mathbb{N}$. The set $\mathbb{K}_d[X]$ of all polynomials with degree at most d is a subspace of $\mathbb{K}[X]$ since for every scalar $\lambda \in \mathbb{K}$ and every polynomials $P(X)$ and $Q(X)$ in $\mathbb{K}_d[X]$

$$\deg(P + \lambda.Q) \leqslant \max\{\deg(P), \deg(\lambda.Q)\} \leqslant \max\{\deg(P), \deg(Q)\} \leqslant d$$

But the set of all polynomials with degree exactly d is not (for instance the linear combination $X^d + (-1_{\mathbb{K}}).X^d = 0_{\mathbb{K}}$ of polynomials of degree d is not of degree d).

b) Function subspaces. Let S be a nonempty set, a be an element in S and $(V, +, .)$ be a $\mathbb{K}$-vector space. Then the set of all functions $f : S \to V$ such that $f(a) = 0_V$ is a subspace of $\mathcal{F}(S, V)$. But the set of all functions $f : S \to V$ such that $f(a)$ is a given non zero vector in V is not.

Furthermore, the set of all continuous (respectively differentiable) functions from $\mathbb{R}$ to itself is a subspace of $\mathcal{F}(\mathbb{R})$ since the sum and the product of any continuous (respectively differentiable) functions is continuous (respectively differentiable).

c) Sequence subspaces. The set of all infinite sequences of elements in a $\mathbb{K}$-vector space $(V, +, .)$ which are eventually equal to zero (that is all elements except a finite number are equal to 0_V) is a subspace of $\mathcal{F}(\mathbb{N}, V)$. But the set of all infinite sequences which are eventually equal to a given non zero vector is not.

Furthermore, the set of all infinite sequences of elements in $\mathbb{K}$ which are convergent is a subspace of $\mathcal{F}(\mathbb{N}, \mathbb{K})$. The same goes for the set of all infinite sequences of elements in $\mathbb{K}$ which converge to $0_{\mathbb{K}}$. But the set of all infinite sequences of elements in $\mathbb{K}$ which converge to a given non zero element is not.

Proposition 3.7

Let W and W' be two subspaces of a $\mathbb{K}$-vector space $(V, +, .)$. Then $W \cap W'$ is also a subspace of V.

Proof : The result follows from Proposition 3.6 ■

Remark : The same does not hold for $W \cup W'$ as soon as $W \not\subset W'$ and $W' \not\subset W$.

Proof : Let w and w' be two vectors in $W \cup W'$ such that $w \in W - W'$ and $w' \in W' - W$. Then the vector $v = w + w'$ is neither in W (otherwise $w' = v - w$ would be in W as a linear combination of vectors in W that is a contradiction since $w' \notin W$) nor in W (by a similar contradiction for $w = v - w'$). So $W \cup W'$ does not contain the linear combination $w + w'$ of vectors in $W \cup W'$ and consequently $W \cup W'$ is not a subspace (from Proposition 3.6). ■

3.2.2 Linear span

Definition 3.8

*Let $(V,+,.)$ be a $\mathbb{K}$-vector space and $S \subset V$ be a subset of vectors. The **linear span** of S, denoted* $\mathbf{Span(S)}$*, is the smallest subspace of V over $\mathbb{K}$ containing S.*

Remark : More precisely, $\text{Span}(S)$ is a subspace containing S and such that $\text{Span}(S) \subset W$ for every subspace W containing S. Hence, the linear span of S is given by the following intersection

$$\text{Span}(S) = \bigcap_{\substack{W \text{ subspace of } (V,+,.) \\ \text{such that } S \subset W}} W$$

Proof : Denote by $I(S)$ the intersection above. It follows from Proposition 3.7 that $I(S)$ is a subspace. Moreover $I(S)$ contains S since S is included in every set of the intersection. Then $\text{Span}(S) \subset I(S)$ by definition of the linear span of S.

Conversely, $\text{Span}(S)$ is a subspace containing S. It follows that $\text{Span}(S)$ is one of the sets of the intersection $I(S)$ and then $I(S) \subset \text{Span}(S)$. Finally, $\text{Span}(S) = I(S)$ as needed. ■

Proposition 3.9

Let $(V,+,.)$ be a $\mathbb{K}$-vector space and $S \subset V$ be a subset of vectors. Then the linear span of S is the set of all linear combinations of vectors in S.

$$\text{Span}(S) = \left\{ \lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_n.v_n \ / \ n \in \mathbb{N} \text{ and } \begin{cases} (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{K}^n \\ (v_1, v_2, \ldots, v_n) \in S^n \end{cases} \right\}$$

Proof : Denote by $L(S)$ the set of all linear combinations of vectors in S. $L(S)$ is a subspace containing S (with linear combination of the form $n = 1$ and $\lambda_1 = 1_{\mathbb{K}}$). Consequently $\text{Span}(S) \subset L(S)$ by definition of the linear span of S.

Conversely, any linear combination of vectors in S is a linear combination of vectors in the subspace $\text{Span}(S)$ (since $S \subset \text{Span}(S)$) and then is in $\text{Span}(S)$ (from Proposition 3.6). It follows that $L(S) \subset \text{Span}(S)$ and finally $\text{Span}(S) = L(S)$ as needed. ■

Examples :

a) $\text{Span}(\emptyset) = \{0_V\}$

b) For every subspace W, we have $\text{Span}(W) = W$.

c) The linear line spanned by a non zero vector v is defined to be

$$\mathbb{K}.v = \text{Span}(\{v\}) = \{\lambda.v \ / \ \lambda \in \mathbb{K}\}$$

d) Consider the following subspace of the real space $\mathbb{R}^3$

$$\mathcal{P} = \left\{ v = (x,y,z) \in \mathbb{R}^3 \ / \ x - 2y + z = 0 \right\}$$

Then $\mathcal{P} = \text{Span}(S)$ where $S = \{(2,1,0), (-1,0,1)\}$.

Proof : At first $(2,1,0)$ and $(-1,0,1)$ are in $\mathcal{P}$ since $2 - 2 \times 1 + 0 = 0$ and $-1 - 2 \times 0 + 1 = 0$. Consequently, $\text{Span}(S) \subset \mathcal{P}$. Conversely, let $v = (x,y,z)$ be a vector in $\mathcal{P}$. Then we have

$$x = 2y - z \text{ and } v = (x,y,z) = (2y - z, y, z) = y.(2,1,0) + z.(-1,0,1)$$

It follows that $v \in \text{Span}(S)$ as a linear combination of $(2,1,0)$ and $(-1,0,1)$, then $\mathcal{P} \subset \text{Span}(S)$. ■

Further examples :

a) Polynomial space. For any element $a \in \mathbb{K}$, the exact Taylor's formula implies that the vector space $\mathbb{K}[X]$ is spanned by the infinite set $\big\{(X-a)^k \ / \ k \in \mathbb{N}\big\}$. Similarly given $d \in \mathbb{N}$, the subspace $\mathbb{K}_d[X]$ of all polynomials with degree at most d is spanned by the finite set $\big\{(X-a)^k \ / \ k \in \{0,1,\ldots,d\}\big\}$, for any element $a \in \mathbb{K}$.

b) Function space. The linear span of $\{f_k : \mathbb{K} \to \mathbb{K}, x \mapsto x^k \ / \ k \in \mathbb{N}\}$ is the subspace of $\mathcal{F}(\mathbb{K})$ of all polynomial maps.

c) Sequence subspaces. For every integer $k \in \mathbb{N}$, define an infinite sequence $u^k \in \mathcal{F}(\mathbb{N}, \mathbb{K})$ as follows

$$\forall i \in \mathbb{N},\ u_i^k = \begin{cases} 1_{\mathbb{K}} & \text{if } i = k \\ 0_{\mathbb{K}} & \text{otherwise} \end{cases}$$

Then the linear span of $S = \{u^k \ / \ k \in \mathbb{N}\}$ is not the whole vector space $\mathcal{F}(\mathbb{N}, \mathbb{K})$. For instance, the infinite sequence $v \in \mathcal{F}(\mathbb{N}, \mathbb{K})$ whose all elements are equal to $1_{\mathbb{K}}$ can not be written as a linear combination of vectors in S.

Proof : By contradiction, assume that v can be written as a linear combination of vectors in S. Then there exist some integers $k_1, k_2, \ldots, k_n$ such that

$$v = \lambda_1.u^{k_1} + \lambda_2.u^{k_2} + \cdots + \lambda_n.u^{k_n}$$

Choose an integer i distinct from k_1, k_2, ..., and k_n. Then

$$u_i^{k_1} = u_i^{k_2} = \cdots = u_i^{k_n} = 0_{\mathbb{K}}$$
$$\text{and} \quad (\lambda_1.u^{k_1} + \lambda_2.u^{k_2} + \cdots + \lambda_n.u^{k_n})_i = \lambda_1 u_i^{k_1} + \lambda_2 u_i^{k_2} + \cdots + \lambda_n u_i^{k_n} = 0_{\mathbb{K}}$$

That is a contradiction with $v_i = 1_{\mathbb{K}}$. ■

Actually, $\mathrm{Span}(S)$ is the subspace of all infinite sequences of elements in $\mathbb{K}$ which are eventually equal to zero.

3.2.3 Sum and direct sum

Definition 3.10 (*Sum of subspaces*)

Let W and W' be two subspaces of a $\mathbb{K}$-vector space $(V, +, .)$. The ***sum*** *of W and W' is defined to be the following set*

$$W + W' = \{w + w' \ / \ w \in W \text{ and } w' \in W'\}$$

Proposition 3.11

Let W and W' be two subspaces of a $\mathbb{K}$-vector space $(V, +, .)$. Then

$$W + W' = \mathrm{Span}(W \cup W')$$

In particular, $W + W'$ is a subspace of V.

Proof : Let v be a vector in $W + W'$. Then there exist $w \in W$ and $w' \in W'$ such that $v = w + w'$. In other words, v is a linear combination of vectors in $W \cup W'$. Consequently $v \in \mathrm{Span}(W \cup W')$ and then $W + W' \subset \mathrm{Span}(W \cup W')$.

Conversely, let v be a vector in $\mathrm{Span}(W \cup W')$. Then v is a linear combination of vectors in $W \cup W'$ that is

$$\exists n \in \mathbb{N} / \begin{cases} \exists (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{K}^n \\ \exists (w_1, w_2, \ldots, w_n) \in (W \cup W')^n \end{cases} / \; v = \lambda_1.w_1 + \lambda_2.w_2 + \cdots + \lambda_n.w_n$$

Without loss of generality, we may reorder the vectors w_1, w_2, ..., and w_n such that the first ones are in W and the next ones are in W'. More precisely, we may assume there exists $k \in \{0, 1, \ldots, n\}$ such that

$$(w_1, w_2, \ldots, w_k) \in W^k \quad \text{and} \quad (w_{k+1}, w_{k+2}, \ldots, w_n) \in (W')^k$$

Then we may write v as follows

$$v = \underbrace{\lambda_1.w_1 + \lambda_2.w_2 + \cdots + \lambda_k.w_k}_{w} + \underbrace{\lambda_{k+1}.w_{k+1} + \lambda_{k+2}.w_{k+2} + \cdots + \lambda_n.w_n}_{w'} = w + w'$$

with $\begin{cases} w = \lambda_1.w_1 + \lambda_2.w_2 + \cdots + \lambda_k.w_k \in W \text{ as a linear combination of vectors in W} \\ w' = \lambda_{k+1}.w_{k+1} + \lambda_{k+2}.w_{k+2} + \cdots + \lambda_n.w_n \in W' \text{ as a linear combination of vectors in W'} \end{cases}$
Consequently $v = w + w' \in W + W'$ and $\mathrm{Span}(W \cup W') \subset W + W'$. ■

Example : Let $\mathbb{K}.v$ and $\mathbb{K}.v'$ be two linear lines spanned by two non zero vectors v and v'. Then

$$\mathbb{K}.v + \mathbb{K}.v' = \left\{\lambda.v + \lambda'.v' \;/\; (\lambda, \lambda') \in \mathbb{K}^2\right\} = \mathrm{Span}(\{v, v'\})$$

In particular, the sum of the linear lines $\mathbb{K}.v$ and $\mathbb{K}.v'$ is a linear line if and only if v and v' are colinear (that is there exists $\mu \in \mathbb{K}^*$ such that $v = \mu.v'$). In this case $\mathbb{K}.v = \mathbb{K}.v' = \mathbb{K}.v + \mathbb{K}.v'$.

Proof : At first, assume $\mathbb{K}.v + \mathbb{K}.v' = \mathbb{K}.w$ for some non zero vector w. We have:

$$\begin{cases} v = 1_{\mathbb{K}}.v + 0_{\mathbb{K}}.v' \in \mathbb{K}.v + \mathbb{K}.v' = \mathbb{K}.w \\ v' = 0_{\mathbb{K}}.v + 1_{\mathbb{K}}.v' \in \mathbb{K}.v + \mathbb{K}.v' = \mathbb{K}.w \end{cases} \quad \text{then} \quad \begin{cases} \exists \lambda \in \mathbb{K} / \; v = \lambda.w \\ \exists \lambda' \in \mathbb{K} / \; v' = \lambda'.w \end{cases}$$

Moreover λ and λ' are not equal to $0_{\mathbb{K}}$ since v and v' are not equal to 0_V. It follows that $w = \lambda'^{-1}.v'$ and $v = \lambda.w = (\lambda\lambda'^{-1}).v' = \mu.v'$ with $\mu = \lambda\lambda'^{-1} \in \mathbb{K}^*$.
Conversly, assume $v = \mu.v'$ for some $\mu \in \mathbb{K}^*$. Then

$$\mathbb{K}.v + \mathbb{K}.v' = \left\{\lambda.v + \lambda'.(\mu.v) \;/\; (\lambda, \lambda') \in \mathbb{K}^2\right\} = \left\{(\lambda + \lambda'\mu).v \;/\; (\lambda, \lambda') \in \mathbb{K}^2\right\} = \{\lambda.v \;/\; \lambda \in \mathbb{K}\} = \mathbb{K}.v$$

In particular, $\mathbb{K}.v + \mathbb{K}.v'$ is a linear line. Furthermore, a similar reasoning shows $\mathbb{K}.v + \mathbb{K}.v' = \mathbb{K}.v'$ (with $v' = \mu^{-1}.v$). The conclusion follows. ■

Remarks :

- More generally, if S and S' are two subsets of vectors then

$$\mathrm{Span}(S) + \mathrm{Span}(S') = \mathrm{Span}(S \cup S')$$

- That provides a binary operation $+$ on the set $\mathcal{V}$ of all subspaces of a $\mathbb{K}$-vector space $(V, +, .)$. This binary operation is associative and commutative. Moreover the trivial subspace $\{0_V\}$ is an identity element. But $(\mathcal{V}, +)$ is not an abelian group since any non trivial subspace has no inverse element. Actually, the same holds for the binary operation $\cap$ on $\mathcal{V}$ (with the whole subspace V as identity element).

- The sum $W + W'$ of two subspaces of a $\mathbb{K}$-vector space $(V, +, .)$ satisfies the following property

$$\forall v \in W + W',\ \exists (w, w') \in W \times W' / \ v = w + w'$$

But the vectors w and w' are not necessarily unique. For instance if $W = W' \neq \emptyset$ then $W + W = W$ and the zero vector $0_V \in W$ may be written as $0_V = w + (-w)$ for every vector $w \in W$.

Definition 3.12 (*Direct sum of subspaces*)

Let W and W' be two subspaces of a $\mathbb{K}$-vector space $(V, +, .)$. The sum $W + W'$ is said to be ***direct*** *if every vector in $W + W'$ may be written uniquely as a sum $w + w'$ with $w \in W$ and $w' \in W'$. Equivalently, $W + W'$ is a direct sum if it satisfies the following condition*

$$\forall v \in W + W',\ \exists ! (w, w') \in W \times W' / \ v = w + w'$$

In this case we write the sum $W \oplus W'$.

Proposition 3.13

Let W and W' be two subspaces of a $\mathbb{K}$-vector space $(V, +, .)$. Then the sum $\Sigma = W + W'$ is a direct sum if and only if W and W' have no intersection except the zero vector. Equivalently

$$\Sigma = W \oplus W' \Longleftrightarrow \left\{ \begin{array}{l} \Sigma = W + W' \\ W \cap W' = \{0_V\} \end{array} \right.$$

Proof : **Necessary.** Let v be a vector in $W \cap W'$. Then we may write

$$\begin{array}{rcll} v & = & v + 0_V & \text{with } v \in W \text{ and } 0_V \in W' \\ & = & 0_V + v & \text{with } 0_V \in W \text{ and } v \in W' \end{array}$$

In particular we get that v is in $W + W' = \Sigma = W \oplus W'$ and then v may be written uniquely as a sum of a vector in W and a vector in W'. It follows that $v = 0_V$ and finally $W \cap W' = \{0_V\}$.

Sufficient. Let v be a vector in $\Sigma = W + W'$. By contradiction, assume there exist two vectors w_1 and w_2 in W and two vectors w'_1 and w'_2 in W' such that $v = w_1 + w'_1 = w_2 + w'_2$. Then the vector $w = w_1 - w_2 = w'_2 - w'_1$ is in W as a linear combination of vectors in W and in W' as a linear combination of vectors in W'. It follows $w \in W \cap W' = \{0_V\}$ that is $w = 0_V$. Consequently, $w_1 = w_2$ and $w'_1 = w'_2$. In other words, every vector $v \in \Sigma$ may be written uniquely as a sum of a vector in W and a vector in W', that is $\Sigma = W \oplus W'$ ■

Example : In $\mathbb{C}$ seen as a $\mathbb{R}$-vector space, the subspaces $\mathbb{R}$ and $\imath\mathbb{R} = \{\imath b \ / \ b \in \mathbb{R}\}$ are such that $\mathbb{R} \cap \imath\mathbb{R} = \{0\}$. In particular, $\mathbb{C} = \mathbb{R} \oplus \imath\mathbb{R}$, and every complex number $z \in \mathbb{C}$ may be written uniquely as $z = a + \imath b$, where $a \in \mathbb{R}$ is called the **real part** of z and $b \in \mathbb{R}$ is called the **imaginary part** of z.

Definition 3.14 (*Supplementary subspaces*)

Two subspaces W and W' of a $\mathbb{K}$-vector space $(V, +, .)$ are said ***supplementary subspaces*** *in V if $V = W \oplus W'$ or equivalently if*

$$\forall v \in V,\ \exists ! (w, w') \in W \times W' / \ v = w + w'$$

In this case, we also say that W' is a ***supplementary subspace*** *of W in V.*

Example : Consider the sets of all even or odd functions from $\mathbb{R}$ to itself:

$$\mathcal{E}(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \text{ even}\} = \{f : \mathbb{R} \to \mathbb{R} \ / \ \forall x \in \mathbb{R},\ f(-x) = f(x)\}$$
$$\mathcal{O}(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \text{ odd}\} = \{f : \mathbb{R} \to \mathbb{R} \ / \ \forall x \in \mathbb{R},\ f(-x) = -f(x)\}$$

It follows from Proposition 3.6 that $\mathcal{E}(\mathbb{R})$ and $\mathcal{O}(\mathbb{R})$ are subspaces of $\mathcal{F}(\mathbb{R})$. Furthermore they are supplementary subspaces in $\mathcal{F}(\mathbb{R})$.

Proof : Let f be a function from $\mathbb{R}$ to itself. Define the two following maps

$$\begin{array}{rcl} f_{\mathcal{E}} : \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{1}{2}\Big(f(x) + f(-x)\Big) \end{array} \quad \text{and} \quad \begin{array}{rcl} f_{\mathcal{O}} : \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{1}{2}\Big(f(x) - f(-x)\Big) \end{array}$$

For every $x \in \mathbb{K}$, we have:

$$\left\{ \begin{array}{l} f_{\mathcal{E}}(-x) = \frac{1}{2}\Big(f(-x) + f(x)\Big) = \frac{1}{2}\Big(f(x) + f(-x)\Big) = f_{\mathcal{E}}(x) \\ f_{\mathcal{O}}(-x) = \frac{1}{2}\Big(f(-x) - f(x)\Big) = -\frac{1}{2}\Big(f(x) - f(-x)\Big) = -f_{\mathcal{O}}(x) \\ f_{\mathcal{E}}(x) + f_{\mathcal{O}}(x) = \frac{1}{2}\Big(f(x) + f(-x) + f(x) - f(-x)\Big) = \frac{1}{2}\Big(2f(x)\Big) = f(x) \end{array} \right.$$

Then $f = f_{\mathcal{E}} + f_{\mathcal{O}}$ with $f_{\mathcal{E}}$ even and $f_{\mathcal{O}}$ odd. Consequently $\mathcal{F}(\mathbb{R}) = \mathcal{E}(\mathbb{R}) + \mathcal{O}(\mathbb{R})$.
Now let f be a function from $\mathbb{R}$ to itself which is odd and even. For every $x \in \mathbb{R}$, we have:

$$f(x) = f(-x) = -f(x)$$

Then $2f(x) = 0$ that is $f(x) = 0$. Consequently $\mathcal{E}(\mathbb{R}) \cap \mathcal{O}(\mathbb{R}) = \{0_{\mathcal{F}(\mathbb{R})}\}$. The conclusion follows from Proposition 3.13 ■

Remark : A supplementary subspace of a given subspace is not necessarily unique (if there exists). For instance consider the real line $\mathbb{R}.v$ in the real plane $\mathbb{R}^2$ with $v = (1, 0) \in \mathbb{R}^2$. Then $v' = (0, 1)$ and $v'' = (1, 1)$ give two distinct supplementary subspaces $\mathbb{R}.v'$ and $\mathbb{R}.v''$ of $\mathbb{R}.v$. Actually, we have $\mathbb{R}^2 = \mathbb{R}.v \oplus \mathbb{R}.w$ for every non zero vector $w \in \mathbb{R}^2$ which is non colinear with v (that is $w \notin \mathbb{R}.v$).

3.3 Linear map

3.3.1 Definition and examples

Definition 3.15 (*Linear map*)

Let $(V, +, .)$ and $(W, +, .)$ be two $\mathbb{K}$-vector spaces. A ***linear map*** *from $(V, +, .)$ to $(W, +, .)$ is a function $f : V \to W$ such that*

$$\left\{ \begin{array}{rl} \forall (v, w) \in V^2, & f(v + w) = f(v) + f(w) \\ \forall \lambda \in \mathbb{K},\ \forall v \in V, & f(\lambda.v) = \lambda.f(v) \end{array} \right.$$

In particular, the image of a linear combination of vectors in V under a linear map $f : V \to W$ is a linear combination of vectors in W, that is

$$\forall n \in \mathbb{N},\ \forall (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{K}^n,\ \forall (v_1, v_2, \ldots, v_n) \in V^n,$$
$$f\left(\sum_{k=1}^{n} \lambda_k.v_k\right) = \sum_{k=1}^{n} \lambda_k.f(v_k)$$

Remark : In particular, $n = 0$ gives $f(0_V) = 0_W$.

Proof : We have:

$$f(0_V) + f(0_V) = f(0_V + 0_V) = f(0_V)$$

Now, a simplification on each side by $f(0_V)$ gives $f(0_V) = 0_W$ as needed. ■

Proposition 3.16

Let $(V, +, .)$ and $(W, +, .)$ be two $\mathbb{K}$-vector spaces. A function $f : V \to W$ is a linear map if and only if it satisfies the following condition

$$\forall \lambda \in \mathbb{K},\ \forall (v, w) \in V^2,\ f(v + \lambda.w) = f(v) + \lambda.f(w)$$

Proof : Necessary. If f is a linear map then

$$f(v + \lambda.w) = f(v) + f(\lambda.w) = f(v) + \lambda.f(w)$$

Suficient. $\lambda = 1_{\mathbb{K}}$ gives $f(v + w) = f(v) + f(w)$ for every vectors v and w in V. And $v = 0_V$ gives $f(\lambda.w) = f(0_V) + \lambda.f(w) = 0_W + \lambda.f(w) = \lambda.f(w)$ for every $\lambda \in \mathbb{K}$ and $w \in V$. ■

Examples :

a) The constant function $f : v \mapsto f(v) = 0_V$ equal to the zero vector is a linear map from $(V, +, .)$ to itself (even onto the trivial subspace $\{0_V\}$).

b) The idendity function $f = \mathrm{Id}_V = (v \mapsto v)$ is a linear map from $(V, +, .)$ onto itself.

c) The homothety $f_\mu : v \mapsto \mu.v$ of ratio $\mu \in \mathbb{K}$ is a linear map from $(V, +, .)$ onto itself.

Proof : For every $\lambda \in \mathbb{K}$ and every $(v, w) \in V^2$ we have:

$$\mu.(v + \lambda.w) = \mu.v + \mu.(\lambda.w) = \mu.v + (\mu\lambda).w = \mu.v + (\lambda\mu).w = \mu.v + \lambda.(\mu.w)$$

Consequently $f_\mu(v + \lambda.w) = f_\mu(v) + \lambda.f_\mu(w)$ and the conclusion follows from Proposition 3.16. ■

d) Consider the following map

$$f :\ v = (x, y, z) \longmapsto f(v) = (2x - y, 3z)$$

Then f is a linear map from the real space $\mathbb{R}^3$ to the real plane $\mathbb{R}^2$.

e) Counterexample. The function $f : x \mapsto x^2$ from the vector space $\mathbb{R}$ to itself is not a linear map. Indeed notice that $f(1) = 1$ but $f(1 + 1) = f(2) = 4$ is not equal to $1 + 1 = 2$.

f) Counterexample. The function $f : x \mapsto x + 1$ from the vector space $\mathbb{R}$ to itself is not a linear map. Actually $f(x) + f(0) = f(x) + 1$ is not equal to $f(x) = f(x + 0)$.

Further examples :

a) Polynomial space. The differentiation map $P \mapsto P'$ is a linear map from $\mathbb{K}[X]$ to itself. Given $\alpha \in \mathbb{K}$, the map $P \mapsto P(\alpha)$ is a linear map from $\mathbb{K}[X]$ to $\mathbb{K}$ (seen as the coordinate space of dimension 1). But in general, a polynomial map $P : \mathbb{K} \to \mathbb{K}$ is not a linear map.

b) Function space. The differentiation map $f \mapsto f'$ is a linear map from the subspace of all differentiable functions to $\mathcal{F}(\mathbb{R})$. Given $(a, b) \in \mathbb{R}^2$, the integral $f \mapsto \int_a^b f(t)dt$ is a linear map from the subspace of all continuous functions to $\mathbb{R}$.

c) Sequence space. Consider the following maps from $\mathcal{F}(\mathbb{N}, \mathbb{K})$ to itself

$$\begin{aligned} S_L : u = (u_0, u_1, u_2, \dots) &\longmapsto S_L(u) = (u_1, u_2, u_3, \dots) \\ S_R : u = (u_0, u_1, u_2, \dots) &\longmapsto S_R(u) = (0_{\mathbb{K}}, u_0, u_1, \dots) \end{aligned}$$

Then S_L and S_R are linear maps called respectively the **left hand side shift** and the **right hand side shift**.

3.3.2 Linear map and subspaces

Proposition 3.17

Let $f : V \to W$ be a linear map between two $\mathbb{K}$-vector spaces. Then the following properties hold

1. *If V' is a subspace of V then the image of V' under f defined as follows*

$$f(V') = \{f(v') \in W \ / \ v' \in V'\}$$

is a subspace of W.

2. *if W' is a subspace of W then the inverse image of W' under f defined as follows*

$$f^{-1}(W') = \{v \in V \ / \ f(v) \in W'\}$$

is a subspace of V.

Proof: 1. Let λ be a scalar in $\mathbb{K}$ and w_1, w_2 be two vectors in $f(V')$. Then there exist two vectors v_1', v_2' in V' such that $w_1 = f(v_1')$ and $w_2 = f(v_2')$. Consequently we get

$$w_1 + \lambda.w_2 = f(v_1') + \lambda.f(v_2') = f(v_1' + \lambda.v_2') \in f(V')$$

since $v_1' + \lambda.v_2' \in V'$ as a linear combination of vector in V'. The result follows from Proposition 3.6.

2. Let λ be a scalar in $\mathbb{K}$ and v_1, v_2 be two vectors in $f^{-1}(W')$. In particular $f(v_1), f(v_2)$ are two vectors in W'. Consequently, we get

$$f(v_1 + \lambda.v_2) = f(v_1) + \lambda.f(v_2) \in W'$$

as a linear combination of vectors in W'. Then $v_1 + \lambda.v_2 \in f^{-1}(W')$ and the result follows from Proposition 3.6. ■

Definition 3.18 (*Image and Kernel*)

Let $f : V \to W$ be a linear map between two $\mathbb{K}$-vector spaces.

- *The **image** of f is defined to be the following subspace of W*

$$\mathrm{Im}(f) = f(V) = \{f(v) \in W \ / \ v \in V\}$$

- *The **kernel** of f is defined to be the following subspace of V*

$$\mathrm{Ker}(f) = f^{-1}(\{0_W\}) = \{v \in V \ / \ f(v) = 0_W\}$$

Proposition 3.19

Let $f : V \to W$ be a linear map between two $\mathbb{K}$-vector spaces. Then the following equivalence hold

1. *f is a surjective map if and only if* $\mathrm{Im}(f) = W$
2. *f is an injective map if and only if* $\mathrm{Ker}(f) = \{0_V\}$

Proof : 1. **Necessary.** If f is surjective then every vector $w \in W$ may be written as $w = f(v)$ with $v \in V$, that is $W = \mathrm{Im}(f)$.

Sufficient. if $\mathrm{Im}(f) = W$ then for every vector $w \in W$ there exists $v \in V$ such that $f(v) = w$, that is f is surjective.

2. **Necessary.** If f is injective then $f(v) = 0_W = f(0_V)$ implies $v = 0_V$, that is $\mathrm{Ker}(f) = \{0_V\}$.

Sufficient. Assume $\mathrm{Ker}(f) = \{0_V\}$ and let v_1, v_2 be two vectors in V such that $f(v_1) = f(v_2)$.

$$f(v_1 - v_2) = f\left(v_1 + (-1_{\mathbb{K}}).v_2\right) = f(v_1) + (-1_{\mathbb{K}}).f(v_2) = f(v_1) - f(v_2) = 0_W$$

Then $v_1 - v_2 \in \mathrm{Ker}(f) = \{0_V\}$, that is $v_1 - v_2 = 0_V$ or equivalently $v_1 = v_2$. It follows that f is injective. ■

Example : The left hand side shift map S_L is surjective whereas the right hand side shift map S_R is not. But S_R is injective whereas S_L is not (indeed $\mathrm{Ker}(S_L) = \{u = (u_0, 0_{\mathbb{K}}, 0_{\mathbb{K}}, 0_{\mathbb{K}}, \dots) \ / \ u_0 \in \mathbb{K}\}$).

Definition 3.20 (*Projection*)

Let W and W' be two supplementary subspaces of a $\mathbb{K}$-vector space V, that is $V = W \oplus W'$. Since every vector $v \in V$ may be written uniquely as $v = w + w'$ with $w \in W$ and $w' \in W'$, the following map is well defined.

$$\begin{array}{rccc} p : & V & \longrightarrow & V \\ & v = w + w' & \longmapsto & p(v) = w \end{array}$$

*The linear map p is called the **projection along W' onto W**.*

Example : In $\mathbb{C}$ seen as a $\mathbb{R}$-vector space, we have $\mathbb{C} = \mathbb{R} \oplus \imath\mathbb{R}$. So, we may define

$$\begin{array}{rccc} \Re : & \mathbb{C} & \longrightarrow & \mathbb{R} \\ & z = a + \imath b & \longmapsto & \Re(z) = a \end{array} \quad \text{and} \quad \begin{array}{rccc} \Im : & \mathbb{C} & \longrightarrow & \mathbb{R} \\ & z = a + \imath b & \longmapsto & \Im(z) = b \end{array}$$

The real part function $\Re$ is the projection along $\imath\mathbb{R}$ onto $\mathbb{R}$. And the projection along $\mathbb{R}$ onto $\imath\mathbb{R}$ is given by $\imath\Im : z = a + \imath b \mapsto \imath b$ where $\Im$ is the imaginary part function.

Proposition 3.21

Let W and W' be two supplementary subspaces of a $\mathbb{K}$-vector space V, that is $V = W \oplus W'$. If p is the projection along W' onto W then

$$\mathrm{Im}(p) = W \quad \textit{and} \quad \mathrm{Ker}(p) = W'$$

In particular, if p is a projection then $V = \mathrm{Im}(p) \oplus \mathrm{Ker}(p)$.

Proof : It follows from the definition of a projection that $\mathrm{Im}(p) \subset W$ and $W' \subset \mathrm{Ker}(p)$. If w is a vector in W then $p(w) = p(w+0_V) = w$, that is $w = p(w) \in \mathrm{Im}(p)$. Consequently $W \subset \mathrm{Im}(p)$. Now if v is a vector in $\mathrm{Ker}(p)$, let $w \in W$ and $w' \in W'$ be such that $v = w + w'$, then $0_V = p(v) = p(w + w') = w$. It follows that $v = 0_V + w' = w' \in W'$ and consequently $\mathrm{Ker}(p) \subset W'$. ■

3.3.3 Isomorphic vector spaces

Definition 3.22 (*Isomorphic vector spaces*)

*Two $\mathbb{K}$-vector spaces V and W are said **isomorphic** if there exists a bijective linear map f from V to W. In this case, we write $\boldsymbol{V \approx W}$ and f is called a **vector space isomorphism**.*

Example : Consider the following function

$$\begin{array}{rccc} \varphi: & \mathbb{C} & \longrightarrow & \mathbb{R}^2 \\ & z = a + \imath b & \longmapsto & \varphi(z) = (a, b) \end{array}$$

Then φ is a linear map since $\varphi : z \mapsto (\Re(z), \Im(z))$ where $\Re$ and $\Im$ are linear maps. Moreover φ is bijective. Consequently, the $\mathbb{R}$-vector spaces $\mathbb{C}$ and $\mathbb{R}^2$ are isomorphic.

Further example : The $\mathbb{K}$-vector space $\mathbb{K}[X]$ of all polynomials with coefficients in $\mathbb{K}$ is isomorphic to the subspace of $\mathcal{F}(\mathbb{N}, V)$ of all infinite sequences of elements in $\mathbb{K}$ which are eventually equal to zero.

3.3.4 Sets of linear maps

Definition 3.23 (*Sets of linear maps*)

Let V and W be two $\mathbb{K}$-vector spaces.

- *We denote by $\boldsymbol{L(V, W)}$ (or $\boldsymbol{Hom(V, W)}$) the set of all linear maps from V to W, that is the set of all **vector space homomorphisms** from $(V, +, .)$ to $(W, +, .)$.*
- *We denote by $\boldsymbol{L(V)}$ (or $\boldsymbol{End(V)}$) the set of all linear maps from V to itself, that is the set of all **vector space endomorphisms** of $(V, +, .)$.*
- *We denote by $\boldsymbol{GL(V)}$ (or $\boldsymbol{Aut(V)}$) the set of all bijective linear maps from V to itself, that is the set of all **vector space automorphisms** of $(V, +, .)$.*

Remarks :

- We also denote by $\boldsymbol{Isom(V, W)}$ the set of all bijective linear maps from V to W, that is the set of all vector space isomorphisms from $(V, +, .)$ to $(W, +, .)$. In particular, V and W are isomorphic if and only if $\mathrm{Isom}(V, W) \neq \emptyset$.
- If $f \in GL(V)$ then its bijective inverse $f^{-1} : V \to V$ is also a linear map, that is $f^{-1} \in GL(V)$.

  Proof : Let λ be a scalar in $\mathbb{K}$ and v, w be two vectors in V. Since f is a surjective function, there exist two vectors v', w' in V such that $f(v') = v$ and $f(w') = w$. Then

  $$f^{-1}(v + \lambda.w) = f^{-1}(f(v') + \lambda.f(w')) = f^{-1}(f(v' + \lambda.w')) = v' + \lambda.w' = f^{-1}(v) + \lambda.f^{-1}(w)$$

  The conclusion follows from Proposition 3.16. ■

Proposition 3.24

Let V and W be two $\mathbb{K}$-vector spaces. For every scalar $\lambda \in \mathbb{K}$ and every linear maps f, g from V to W, the maps $f+g$ and $\lambda.f$ defined as follows

$$\begin{array}{rcl} f+g : V & \longrightarrow & W \\ v & \longmapsto & (f+g)(v) = f(v)+g(v) \end{array} \quad \text{and} \quad \begin{array}{rcl} \lambda.f : V & \longrightarrow & W \\ v & \longmapsto & (\lambda.f)(v) = \lambda.f(v) \end{array}$$

are linear maps from V to W. That provides an addition and a scalar multiplication on $L(V,W)$, and the following properties hold

1. *$(L(V,W),+,.)$ is a $\mathbb{K}$-vector space.*
2. *$(L(V),+,.)$ is a $\mathbb{K}$-vector space.*

Proof : The second point follows from the first one with $V = W$. Now, let f and g be two linear maps from V to W and λ be a scalar in $\mathbb{K}$. Then for every vector v, w in V and every scalar μ in $\mathbb{K}$, we have:

$$\begin{array}{rcl} (f+\lambda.g)(v+\mu.w) & = & f(v+\mu.w)+\lambda.g(v+\mu.w) \\ & = & f(v)+\mu.f(w)+\lambda.(g(v)+\mu.g(w)) \\ & = & f(v)+\mu.f(w)+\lambda.g(v)+\lambda.(\mu.g(w)) \\ & = & f(v)+\lambda.g(v)+\mu.f(w)+(\lambda\mu).g(w) \\ & = & (f+\lambda.g)(v)+\mu.f(w)+(\mu\lambda).g(w) \\ & = & (f+\lambda.g)(v)+\mu.f(w)+\mu.(\lambda.g(w)) \\ & = & (f+\lambda.g)(v)+\mu.(f(w)+\lambda.g(w)) \\ & = & (f+\lambda.g)(v)+\mu.(f+\lambda.g)(w) \end{array}$$

It follows from Proposition 3.16 that $f+\lambda.g$ is a linear map and from Proposition 3.6 that $L(V,W)$ is a subspace of the $\mathbb{K}$-vector space $\mathcal{F}(V,W)$ of all functions from V to W. In particular, $L(V,W)$ is a $\mathbb{K}$-vector space. ■

Remark : But a linear combination of bijective linear maps is not always bijective.

Proposition 3.25

Let U, V and W be three $\mathbb{K}$-vector spaces. For every linear map f from V to W and every linear map g from U to V, the composition map $f \circ g$ defined as follows

$$\begin{array}{rcl} f \circ g : U & \longrightarrow & W \\ u & \longmapsto & (f \circ g)(u) = f(g(u)) \end{array}$$

is a linear map from U to W. In case $U = V = W$, that provides a binary operation on $L(V)$, and the following properties hold

1. *$(L(V),+,\circ)$ is an unital ring.*
2. *$(GL(V),\circ)$ is a group.*

The identity elements are respectively the constant linear map equal to 0_V for addition and the identity function Id_V for composition.

Proof : The second point follows from the first one and from the claim that the bijective inverse of a linear map is also a linear map. For the first point, we have:

1. $(L(V), +)$ is an abelian group.
2. $\circ$ is associative.
3. $\circ$ is distributive over $+$. Indeed for every linear maps f, g, h from V to V, we have:

$$\forall v \in V, \begin{cases} f((g+h)(v)) = f(g(v)+h(v)) = f(g(v)) + f(h(v)) \\ (g+h)(f(v)) = g(f(v)) + h(f(v)) \end{cases}$$

that is

$$\begin{cases} f \circ (g+h) = f \circ g + f \circ h \\ (g+h) \circ f = g \circ f + h \circ f \end{cases}$$

4. The identity map $\mathrm{Id}_V : v \mapsto v$ is the identity element for $\circ$.

The result follows. ■

Remark : But the composition of linear maps is not commutative.

Chapter 4

Finite-dimensional vector spaces

In this chapter, fix a commutative field $(\mathbb{K}, +, \times)$ (for instance $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$).

4.1 Family of vectors

In this section, fix a $\mathbb{K}$-vector space V.

4.1.1 Linearly independent family

Definition 4.1 (*Linearly independent family*)

Let $(v_1, v_2, \ldots, v_n)$ be a family of vectors in V. This family is said to be ***linearly independent*** *if it satisfies the following condition*

$$\forall (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{K}^n,\ \lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_n.v_n = 0_V \Longrightarrow \lambda_1 = \lambda_2 = \cdots = \lambda_n = 0_{\mathbb{K}}$$

Conversely this family is said ***linearly dependent*** *if there exist some scalars $\lambda_1, \lambda_2, \ldots, \lambda_n$ not all zero such that the linear combination $\lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_n.v_n$ is equal to the zero vector.*

Examples :

a) Any family of only one non zero vector $v \in V - \{0_V\}$ is linearly independent since

$$\forall \lambda \in \mathbb{K},\ \lambda.v = 0_V \Longleftrightarrow \lambda = 0_{\mathbb{K}}$$

b) Any family of vectors containing the zero vector is linearly dependent.

Proof : Let $v_1, v_2, \ldots, v_n$ be such a family of vectors. Without loss of generality, we may reorder the family such that $v_1 = 0_V$. Now choosing any scalar $\lambda_1 \neq 0_{\mathbb{K}}$ and $\lambda_2 = \cdots = \lambda_n = 0_{\mathbb{K}}$, we get

$$\lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_n.v_n = \lambda_1.0_V + 0_{\mathbb{K}}.(v_2 + \cdots + v_n) = 0_V$$

■

c) Let v and v' be two non zero vectors in $V - \{0_V\}$. Then the family (v, v') is linearly dependent if and only if v and v' are colinear (that is there exists $\mu \in \mathbb{K}$ such that $v = \mu.v'$).

Proof : If v and v' are colinear then $\lambda = 1_{\mathbb{K}} \neq 0_{\mathbb{K}}$ and $\lambda' = -\mu$ give $\lambda.v + \lambda'.v' = v - \mu.v' = 0_V$. Conversely, if there exist λ and λ' not both zero in $\mathbb{K}$ such that $\lambda.v + \lambda'.v' = 0_V$ then $\lambda \neq 0_{\mathbb{K}}$ (otherwise $\lambda' \neq 0_{\mathbb{K}}$ and $\lambda'.v' = 0_V$ are a contradiction with $v' \neq 0_V$) and $\mu = -\lambda'\lambda^{-1}$ implies $v = \mu.v'$ as needed. ■

d) Consider the following family of the real space $\mathbb{R}^3$

$$\Big(v_1 = (0,0,1), v_2 = (1,0,-1), v_3 = (1,2,3)\Big)$$

It is a linearly independent family of vectors in $\mathbb{R}^3$.

Proof: For every $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$, we have:

$$\begin{aligned}
\lambda_1.v_1 + \lambda_2.v_2 + \lambda_3.v_3 = 0_{\mathbb{R}^3} &\iff \lambda_1.(0,0,1) + \lambda_2.(1,0,-1) + \lambda_3.(1,2,3) = (0,0,0) \\
&\iff (\lambda_2 + \lambda_3, 2\lambda_3, \lambda_1 - \lambda_2 + 3\lambda_3) = (0,0,0) \\
&\iff \begin{cases} 2\lambda_3 &= 0 \\ \lambda_2 + \lambda_3 &= 0 \\ \lambda_1 - \lambda_2 + 3\lambda_3 &= 0 \end{cases} \\
&\iff \begin{cases} \lambda_3 = 0 \\ \lambda_2 = -\lambda_3 = 0 \\ \lambda_1 = \lambda_2 - 3\lambda_3 = 0 \end{cases}
\end{aligned}$$

The result follows. ■

Proposition 4.2

Every subfamily of a linearly independent family of vectors in V is also linearly independent.

Proof: Let $(v_1, v_2, \ldots, v_n)$ be a linearly independent family of vectors in V. Consider a subfamily of this family. Without loss of generality, we may reorder the family such that the subfamily is $(v_1, v_2, \ldots, v_k)$ with $k \leqslant n$. Now let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be scalars in $\mathbb{K}$ such that

$$\lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_k.v_k = 0_V$$

Then $\lambda_{k+1} = \lambda_{k+2} = \cdots = \lambda_n = 0_{\mathbb{K}}$ give

$$\lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_n.v_n = \underbrace{\lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_k.v_k}_{=0_V} + \underbrace{\lambda_{k+1}.v_{k+1} + \lambda_{k+2}.v_{k+2} + \cdots + \lambda_n.v_n}_{=0_{\mathbb{K}}.(v_{k+1}+v_{k+2}+\cdots+v_n)=0_V} = 0_V$$

Since $(v_1, v_2, \ldots, v_n)$ is linearly independent, it follows $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0_{\mathbb{K}}$, and it particular $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0_{\mathbb{K}}$. Consequently the subfamily $(v_1, v_2, \ldots, v_k)$ is also linearly independent. ■

4.1.2 Spanning family and basis

Definition 4.3 (*Spanning family*)

*Let $(v_1, v_2, \ldots, v_n)$ be a family of vectors in V. This family is said to be a **spanning family** if the linear span of $\{v_1, v_2, \ldots, v_n\}$ is the whole vector space V. Equivalently, $(v_1, v_2, \ldots, v_n)$ is a spanning family if it satisfies the following condition*

$$\forall v \in V,\ \exists (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{K}^n \ / \ v = \lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_n.v_n$$

Proposition 4.4

Every superfamily of a spanning family of vectors in V is also a spanning family.

Proof : Let $(v_1, v_2, \ldots, v_n)$ be a superfamily of a spanning family of vectors in V. Without loss of generality, we may reorder the family such that $(v_1, v_2, \ldots, v_k)$ with $k \leqslant n$ is a spanning family. Now let v be a vector in V. Since $(v_1, v_2, \ldots, v_k)$ is a spanning family, it follows that there exist some $\lambda_1, \lambda_2, \ldots, \lambda_k$ scalars in $\mathbb{K}$ such that

$$v = \lambda_1.v_1 + \lambda_2 + v_2 + \cdots + \lambda_k.v_k$$

Then $\lambda_{k+1} = \lambda_{k+2} = \cdots = \lambda_n = 0_{\mathbb{K}}$ give

$$v = \underbrace{\lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_k.v_k}_{=v} + \underbrace{\lambda_{k+1}.v_{k+1} + \lambda_{k+2}.v_{k+2} + \cdots + \lambda_n.v_n}_{=0_{\mathbb{K}}.(v_{k+1}+v_{k+2}+\cdots+v_n)=0_V} = \lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_n.v_n$$

Consequently the superfamily $(v_1, v_2, \ldots, v_n)$ is also a spanning family. ■

Remark : If $(v_1, v_2, \ldots, v_n)$ is a spanning family then every vector in V may be written as a linear combination of vectors in this family with scalars $(\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{K}^n$. But these scalars are not necessarily unique.

Definition 4.5 (*Basis*)

A ***basis*** *of the vector space V is a linearly independent spanning family $\mathcal{B} = (v_1, v_2, \ldots, v_n)$.*

Proposition 4.6

Let $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ be a family of vectors in V. $\mathcal{B}$ is a basis of V if and only if every vector in V may be written uniquely as a linear combination of $v_1, v_2, \ldots, v_n$. Equivalently, $\mathcal{B}$ is a basis if it satisfies the following condition

$$\forall v \in V,\ \exists!(\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{K}^n \ / \ v = \lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_n.v_n$$

Proof : **Sufficient.** If every vector may be written uniquely as a linear combination of vectors in $\mathcal{B}$, then in particular $\mathcal{B}$ is a spanning family. Now assume the linear combination $\lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_n.v_n$ is equal to the zero vector for some scalars $\lambda_1, \lambda_2, \ldots, \lambda_n$ in $\mathbb{K}$. Then the unicity of the writing $0_V = 0_{\mathbb{K}}.v_1 + 0_{\mathbb{K}}.v_2 + \cdots + 0_{\mathbb{K}}.v_n$ implies $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0_{\mathbb{K}}$. Consequently $\mathcal{B}$ is linearly independent and thus a basis of V.

Necessary. Let v be a vector in V. Since $\mathcal{B}$ is a spanning family, there exist some scalars $\lambda_1, \lambda_2, \ldots, \lambda_n$ in $\mathbb{K}$ such that $v = \lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_n.v_n$. Now assume there exist some others scalars $\lambda'_1, \lambda'_2, \ldots, \lambda'_n$ in $\mathbb{K}$ such that $v = \lambda'_1.v_1 + \lambda'_2.v_2 + \cdots + \lambda'_n.v_n$. Then

$$\begin{aligned} 0_V &= v - v \\ &= (\lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_n.v_n) - (\lambda'_1.v_1 + \lambda'_2.v_2 + \cdots + \lambda'_n.v_n) \\ &= (\lambda_1 - \lambda'_1).v_1 + (\lambda_2 - \lambda'_2).v_2 + \cdots + (\lambda_n - \lambda'_n).v_n \end{aligned}$$

Since $\mathcal{B}$ is linearly independent, it follows that $\lambda_1 - \lambda'_1 = \lambda_2 - \lambda'_2 = \cdots = \lambda_n - \lambda'_n = 0_{\mathbb{K}}$ or equivalently $\lambda_1 = \lambda'_1$, $\lambda_2 = \lambda'_2$, …, and $\lambda_n = \lambda'_n$. Consequently the writing $v = \lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_n.v_n$ is unique. ■

Example : Consider the following family of the real space $\mathbb{R}^3$

$$\mathcal{B} = \Big(v_1 = (0, 0, 1), v_2 = (1, 0, -1), v_3 = (1, 2, 3)\Big)$$

It is a basis of $\mathbb{R}^3$.

Proof: We have already shown that $\mathcal{B}$ is a linearly independent family of vectors in $\mathbb{R}^3$. Now let $v = (x, y, z)$ be a vector in $\mathbb{R}^3$. For every $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{K}^3$, we have:

$$\begin{aligned} v = \lambda_1.v_1 + \lambda_2.v_2 + \lambda_3.v_3 &\iff (x, y, z) = \lambda_1.(0, 0, 1) + \lambda_2.(1, 0, -1) + \lambda_3.(1, 2, 3) \\ &\iff (x, y, z) = (\lambda_2 + \lambda_3, 2\lambda_3, \lambda_1 - \lambda_2 + 3\lambda_3) \\ &\iff \left\{ \begin{array}{rcl} 2\lambda_3 & = & y \\ \lambda_2 + \lambda_3 & = & x \\ \lambda_1 - \lambda_2 + 3\lambda_3 & = & z \end{array} \right. \\ &\iff \left\{ \begin{array}{l} \lambda_3 = \frac{y}{2} \\ \lambda_2 = x - \lambda_3 = x - \frac{y}{2} \\ \lambda_1 = z + \lambda_2 - 3\lambda_3 = z + x - \frac{y}{2} - \frac{3y}{2} = x - 2y + z \end{array} \right. \end{aligned}$$

Consequently $v = (x - 2y + z).v_1 + (x - \frac{y}{2}).v_2 + \frac{y}{2}.v_3$. It follows that $\mathcal{B}$ is a spanning family and thus a basis of $\mathbb{R}^3$. ■

4.2 Finite dimension

4.2.1 Finite-dimensional vector space

Definition 4.7 (*Finite dimensional vector space*)

Let V be a $\mathbb{K}$-vector space. V is said to be ***finite-dimensional*** *if there exist a spanning family of vectors in V.*

Examples:

a) For every positive integer $n \in \mathbb{N}^*$, the coordinate space $\mathbb{K}^n$ is finite-dimensional. Indeed the following family of vectors in $\mathbb{K}^n$ is a basis

$$\mathcal{B} = \Big(e_1, e_2, \ldots, e_n\Big) \quad \text{where } \forall k \in \{1, 2, \ldots, n\},\ e_k = (0_{\mathbb{K}}, 0_{\mathbb{K}}, \ldots, 0_{\mathbb{K}}, \underbrace{1_{\mathbb{K}}}_{k^{\text{th}} \text{ position}}, 0_{\mathbb{K}}, \ldots, 0_{\mathbb{K}})$$

This family is called the **canonical basis** of $\mathbb{K}^n$.

Proof: Linearly independent. For every $(\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{K}^n$, we have:

$$\begin{aligned} \lambda_1.e_1 + \lambda_2.e_2 + \cdots + \lambda_n.e_n = 0_{\mathbb{K}^n} &\iff (\lambda_1, \lambda_2, \ldots, \lambda_n) = (0_{\mathbb{K}}, 0_{\mathbb{K}}, \ldots, 0_{\mathbb{K}}) \\ &\iff \lambda_1 = \lambda_2 = \cdots = \lambda_n = 0_{\mathbb{K}} \end{aligned}$$

Spanning family. Let $v = (x_1, x_2, \ldots, x_n)$ be a vector in $\mathbb{K}^n$. Then

$$v = (x_1, x_2, \ldots, x_n) = x_1.e_1 + x_2.e_2 + \cdots + x_n.e_n$$

■

b) Counterexample. The $\mathbb{K}$-vector space $\mathbb{K}[X]$ of all polynomials with coefficients in $\mathbb{K}$ is not finite-dimensional.

Proof: Assume there exists a spanning family, say $(P_1(X), P_2(X), \ldots, P_n(X))$. Then every polynomial $P(X) \in \mathbb{K}[X]$ may be written as $P(X) = \lambda_1.P_1(X) + \lambda_2.P_2(X) + \cdots + \lambda_n.P_n(X)$ where $(\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{K}^n$. Consequently

$$\begin{aligned} \deg(P) &= \deg(\lambda_1.P_1 + \lambda_2.P_2 + \cdots + \lambda_n.P_n) \\ &\leqslant \max\{\lambda_1.P_1, \lambda_2.P_2, \ldots, \lambda_n.P_n\} \\ &\leqslant \max\{P_1, P_2, \ldots, P_n\} \end{aligned}$$

But $P(X) = X^{d+1}$ with $d = \max\{P_1, P_2, \ldots, P_n\} \geqslant 0$ gives $d+1 \leqslant d$ which is a contradiction. ■

c) But the subspace $\mathbb{K}_d[X]$ of all polynomials with degree at most a given integer d is finite-dimensional. Indeed the family $(1, X, X^2, \ldots, X^d)$ of successive powers of X is a basis (it is linearly independent from the definition of polynomials and a spanning family from the exact Taylor's formula).

Proposition 4.8

Let V be a finite-dimensional $\mathbb{K}$-vector space. Then the following properties hold

1. *Every linearly independent family of vectors in V is a subfamily of some basis of V.*
2. *Every spanning family of vectors in V contains some basis of V as subfamily.*

Proof: Let $\mathcal{F}$ be a spanning family of vectors in V and n be the number of vectors in this family.

1. Let $\mathcal{F}_0 = (v_1, v_2, \ldots, v_m)$ be a linearly independent family of vectors in V. Now denote by $(v_{m+1}, v_{m+2}, \ldots, v_{m+n})$ the vectors in $\mathcal{F}$ and by $\mathcal{F}_n$ the family $(v_1, v_2, \ldots, v_{m+n})$ which is a spanning family from Proposition 4.4.

$$\mathcal{F}_0 = (\underbrace{v_1, v_2, \ldots, v_m}_{\text{linearly independent}}) \subset \mathcal{F}_n = (\underbrace{\underbrace{v_1, v_2, \ldots, v_m}_{\text{linearly independent}}, \underbrace{v_{m+1}, v_{m+2}, \ldots, v_{m+n}}_{\text{spanning family}}}_{\text{spanning family}})$$

If $\mathcal{F}_0$ is a spanning family then it is a basis and the conclusion holds. Thus assume that $\mathcal{F}_0$ is not a spanning family. Consequently there exists at least one vector in $(v_{m+1}, v_{m+2}, \ldots, v_{m+n})$ which may not be written as a linear combination of vectors in $\mathcal{F}_0$ (otherwise the linear span $\text{Span}(v_{m+1}, v_{m+2}, \ldots, v_{m+n}) = V$ would be included in $\text{Span}(v_1, v_2, \ldots, v_m) \subsetneq V$ which is a contradiction). Without loss of generality, we may reorder the family $(v_{m+1}, v_{m+2}, \ldots, v_{m+n})$ such that v_{m+1} may not be written as a linear combination of vectors in $\mathcal{F}_0$. Now denote by $\mathcal{F}_1$ the family $(v_1, v_2, \ldots, v_m, v_{m+1})$. We are going to prove that $\mathcal{F}_1$ is linearly independent. Assume $\lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_{m+1}.v_{m+1} = 0_V$ for some scalars $(\lambda_1, \lambda_2, \ldots, \lambda_{m+1}) \in \mathbb{K}^{m+1}$.

First case. $\lambda_{m+1} \neq 0_{\mathbb{K}}$
Then we get $v_{m+1} = (-\lambda_{m+1}^{-1}\lambda_1).v_1 + (-\lambda_{m+1}^{-1}\lambda_2).v_2 + \cdots + (-\lambda_{m+1}^{-1}\lambda_m).v_m$ that is v_{m+1} is a linear combination of vectors in $\mathcal{F}_0$ which is a contradiction.

Second case. $\lambda_{m+1} = 0_{\mathbb{K}}$
Then $\lambda_1.v_1 + \lambda_2.v_2 + \cdots + \lambda_m.v_m = 0_V$ implies $\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0_{\mathbb{K}}$ since $\mathcal{F}_0$ is linearly independent.

Finally $\lambda_1 = \lambda_2 = \cdots = \lambda_m = \lambda_{m+1} = 0_{\mathbb{K}}$ and thus $\mathcal{F}_1$ is linearly independent.

$$\mathcal{F}_1 = (\underbrace{v_1, v_2, \ldots, v_{m+1}}_{\text{linearly independent}}) \subset \mathcal{F}_n = (\underbrace{v_1, v_2, \ldots, v_m, v_{m+1}, v_{m+2}, \ldots, v_{m+n}}_{\text{spanning family}})$$

If $\mathcal{F}_1$ is a spanning family then it is a basis which contains the starting linearly independent family $\mathcal{F}_0$ and the conclusion holds. Otherwise we may iterate the previous process to construct an increasing sequence of linearly independent families

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}_n$$

This algorithm stops as soon as one constructed family $\mathcal{F}_k$ for $k \in \{0, 1, 2, \ldots, n\}$ is a spanning family (at worst for $k = n$). Then $\mathcal{F}_k$ is a basis which contains the starting linearly independent family $\mathcal{F}_0$ and the conclusion holds.

2. Let $\mathcal{F}_n = \mathcal{F} = (v_1, v_2, \ldots, v_n)$ be the starting spanning family. Without loss of generality, we may reorder this family such that $v_1 \neq 0_V$. Now denote by $\mathcal{F}_1$ the family (v_1) of only one vector which is linearly independent.

$$\mathcal{F}_1 = (\underbrace{v_1}_{\text{linearly independent}}) \subset \mathcal{F}_n = (\underbrace{v_1, v_2, \ldots, v_n}_{\text{spanning family}})$$

Then we may use the same algorithm as the previous point to construct an increasing sequence of linearly independent families

$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}_n$$

When the algorithm stops, we get a basis $\mathcal{F}_k$ with $k \in \{0, 1, 2, \ldots, n\}$ which is contained in the starting spanning family $\mathcal{F}_n$ and the conclusion holds. ■

Remark : In particular, every finite-dimensional $\mathbb{K}$-vector space has a basis. Indeed, any family of only one non zero vector is linearly independent thus is contained in some basis.

4.2.2 Dimension

Theorem 4.9 (*Dimension theorem*)

Let V be a finite-dimensional $\mathbb{K}$-vector space. Then all bases of V have the same cardinality.

Proof : At first, we are going to prove the following result:

Lemma 4.10

If $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ is a basis of V and $\mathcal{F} = (w_1, w_2, \ldots, w_m)$ is a spanning family of vectors in V then $n \leqslant m$.

Proof of the lemma : Since $\mathcal{F}$ is a spanning family, the vector v_1 may be written as a linear combination of vectors in $\mathcal{F}$.

$$\exists (\lambda_1, \lambda_2, \ldots, \lambda_m) \in \mathbb{K}^m \ / \ v_1 = \lambda_1.w_1 + \lambda_2.w_2 + \cdots + \lambda_m.w_m$$

Since $\mathcal{B}$ is linearly independent, v_1 is not equal to the zero vector and the scalars $\lambda_1, \lambda_2, \ldots, \lambda_m$ are not all zero. Without loss of generality, we may reorder $\mathcal{F}$ such that $\lambda_1 \neq 0_{\mathbb{K}}$. Then

$$w_1 = \lambda_1^{-1}.v_1 + (-\lambda_1^{-1}\lambda_2).w_2 + \cdots + (-\lambda_1^{-1}\lambda_m).w_m$$

It follows that the linear span $\text{Span}(w_1, w_2, \ldots, w_m) = V$ is included in $\text{Span}(v_1, w_2, \ldots, w_m)$ and then $\mathcal{F}_1 = (v_1, w_2, \ldots, w_m)$ is a spanning family.

By induction, we may show as well that $\mathcal{F}_k = (v_1, \ldots, v_k, w_{k+1}, w_{k+2}, \ldots, w_m)$ is a spanning family for any $k \in \{1, 2, \ldots, n\}$ (after some possible reordering of $\mathcal{F}$ if necessary). Indeed assume $\mathcal{F}_k$ is a spanning family for some $k \in \{1, 2, \ldots, n-1\}$. Then the vector v_{k+1} may be written as a linear combination of vectors in $\mathcal{F}_k$.

$$\begin{aligned} \exists (\lambda_1, \lambda_2, \ldots, \lambda_m) \in \mathbb{K}^m \ / \quad & v_{k+1} = \lambda_1.v_1 + \cdots + \lambda_k.v_k + \lambda_{k+1}.w_{k+1} + \cdots + \lambda_m.w_m \\ & \iff (-\lambda_1).v_1 + \cdots + (-\lambda_k).v_k + v_{k+1} = \lambda_{k+1}.w_{k+1} + \cdots + \lambda_m.w_m \end{aligned}$$

Since $\mathcal{B}$ is linearly independent, $(-\lambda_1).v_1 + \cdots + (-\lambda_k).v_k + v_{k+1}$ is not equal to the zero vector and the scalars $\lambda_{k+1}, \lambda_{k+2}, \ldots, \lambda_m$ are not all zero. Without loss of generality, we may reorder $(w_{k+1}, w_{k+2}, \ldots, w_m)$ such that $\lambda_{k+1} \neq 0_{\mathbb{K}}$. Then

$$w_{k+1} = (-\lambda_{k+1}^{-1}\lambda_1).v_1 + \cdots + (-\lambda_{k+1}^{-1}\lambda_k).v_k + \lambda_{k+1}^{-1}.v_{k+1} + (-\lambda_{k+1}^{-1}\lambda_{k+2}).w_{k+2} + \cdots + (-\lambda_{k+1}^{-1}\lambda_m).w_m$$

It follows that the linear span $\text{Span}(v_1, \ldots, v_k, w_{k+1}, w_{k+2}, \ldots, w_m) = V$ (by the inductive hypothesis) is included in $\text{Span}(v_1, \ldots, v_k, v_{k+1}, w_{k+2}, \ldots, w_m)$ and then $\mathcal{F}_{k+1}$ is a spanning family. That concludes the induction.

In particular, we get that $\mathcal{F}_n = (v_1, \ldots, v_n, w_{n+1}, w_{n+2}, \ldots, w_m)$ is a spanning family and that $n \leqslant m$. ■

Now let $\mathcal{B}$ and $\mathcal{B}'$ be two bases of V. From Lemma 4.10, we get $\text{Card}(\mathcal{B}) \leqslant \text{Card}(\mathcal{B}')$ since $\mathcal{B}'$ is a spanning family and $\text{Card}(\mathcal{B}') \leqslant \text{Card}(\mathcal{B})$ since $\mathcal{B}$ is a spanning family. Consequently $\mathcal{B}$ and $\mathcal{B}'$ have the same cardinality and the conclusion follows. ■

Definition 4.11 (*Dimension*)

Let V be a finite-dimensional $\mathbb{K}$-vector space. The ***dimension*** *of V, denoted* $\mathbf{dim(V)}$*, is the number of vectors in any basis of V.*

Examples :

a) For every $n \in \mathbb{N}^*$, $\dim(\mathbb{K}^n) = n$ since the canonical basis of $\mathbb{K}^n$ contains exactly n vectors.

b) $\dim(\{0_V\}) = 0$ since $\emptyset$ is a basis of the trivial vector space.

c) For any given integer $d \in \mathbb{N}$, $\dim(\mathbb{K}_d[X]) = d+1$ since the basis $(1, X, X^2, \ldots, X^d)$ contains exactly $d+1$ polynomials.

Proposition 4.12

Let V be a finite-dimensional $\mathbb{K}$-vector space. The the following properties hold

1. *Every linearly independent family contains at most $\dim(V)$ vectors and is a basis if and only if it contains exactly $\dim(V)$ vectors.*
2. *Every spanning family contains at least $\dim(V)$ vectors and is a basis if and only if it contains exactly $\dim(V)$ vectors.*

Proof : That follows from Proposition 4.8 and Dimension theorem 4.9. ■

Example : For any given integer $d \in \mathbb{N}$ and any element $a \in \mathbb{K}$, the family

$$\mathcal{B} = \Big(1, (X-a), (X-a)^2, \ldots, (X-a)^d\Big)$$

is a basis of $\mathbb{K}_d[X]$. Indeed it is a spanning family from the exact Taylor's formula and it contains exactly $d+1 = \dim(\mathbb{K}_d[X])$ polynomials.

4.2.3 Finite-dimensional subspace

Proposition 4.13

Let V be a finite-dimensional $\mathbb{K}$-vector space and W be a subspace of V. Then W is also finite-dimensional and

$$\dim(W) \leqslant \dim(V)$$

Furthermore

$$\dim(W) = \dim(V) \iff W = V$$

Proof: At first, we prove that W is finite-dimensional. Let $\mathcal{F}_1 = (w_1)$ be a family of only one non zero vector in W. In particular, $\mathcal{F}_1$ is linearly independent. If $\mathcal{F}_1$ is a spanning family of W then it is a basis of W and the conclusion holds. Thus assume that $\mathcal{F}_1$ is not a spanning family of W. Consequently there exists at least one vector in W, say w_2, which may not be written as a linear combination of vectors in $\mathcal{F}_1$ (that is w_2 is not colinear to w_1). As in the proof of Proposition 4.8, we deduce that the family $\mathcal{F}_2 = (w_1, w_2)$ is linearly independent. If $\mathcal{F}_2$ is a spanning family of W then it is a basis and the conclusion holds. Otherwise we may iterate the previous process to construct an increasing sequence of linearly independent families

$$\mathcal{F}_1 = (w_1) \subset \mathcal{F}_2 = (w_1, w_2) \subset \cdots \subset \mathcal{F}_k = (w_1, w_2, \ldots, w_k) \subset \ldots$$

This algorithm stops as soon as one constructed family $\mathcal{F}_k$ for $k \geqslant 1$ is a spanning family. It happens at worst for $k = \dim(V)$, since we would get a basis of V (from Proposition 4.12) and thus a spanning family of $W \subset V$. Consequently, W is finite-dimensional.

The remaining follows from Proposition 4.12. ■

Proposition 4.14

Let W and W' be two subspaces of a finite-dimensional $\mathbb{K}$-vector space V. Then

$$\dim(W + W') = \dim(W) + \dim(W') - \dim(W \cap W')$$

In particular

$$\begin{cases} \dim(V) = \dim(W) + \dim(W') \\ W \cap W' = \{0_V\} \end{cases} \iff V = W \oplus W'$$

Proof: From Proposition 4.13, $W \cap W'$ is finite-dimensional as subspace of V. Let $\mathcal{B}_\cap = (v_1, \ldots, v_k)$ be a basis of $W \cap W'$. From the first point of Proposition 4.8, $\mathcal{B}_\cap$ is a subfamily of

- some basis of W, say $\mathcal{B} = (v_1, \ldots, v_k, w_1, \ldots, w_n)$
- some basis of W', say $\mathcal{B}' = (v_1, \ldots, v_k, w'_1, \ldots, w'_{n'})$

We are going to prove that $\mathcal{B}_+ = (v_1, \ldots, v_k, w_1, \ldots, w_n, w'_1, \ldots, w'_{n'})$ is a basis of $W + W'$.

Linearly independent. Assume $\lambda_1.v_1 + \cdots + \lambda_k.v_k + \mu_1.w_1 + \cdots + \mu_n.w_n + \mu'_1.w'_1 + \cdots + \mu'_{n'}.w'_{n'} = 0_V$ for some scalars $(\lambda_1, \ldots, \lambda_k, \mu_1, \ldots, \mu_n, \mu'_1, \ldots, \mu'_{n'}) \in \mathbb{K}^{k+n+n'}$. Then

$$w' = \mu'_1.w'_1 + \cdots + \mu'_{n'}.w'_{n'} = -(\lambda_1.v_1 + \cdots + \lambda_k.v_k + \mu_1.w_1 + \cdots + \mu_n.w_n) \in W \cap W'$$

Consequently $w' = \lambda'_1.v_1 + \cdots + \lambda'_k.v_k$ for some scalars $(\lambda'_1, \ldots, \lambda'_k) \in \mathbb{K}^k$. It follows

$$w' - w' = \lambda'_1.v_1 + \cdots + \lambda'_k.v_k + (-\mu'_1).w'_1 + \cdots + (-\mu'_{n'}).w'_{n'} = 0_V$$

Since $\mathcal{B}'$ is linearly independent, we get $\lambda'_1 = \cdots = \lambda'_k = \mu'_1 = \cdots = \mu'_{n'} = 0_\mathbb{K}$. Now we have $\lambda_1.v_1 + \cdots + \lambda_k.v_k + \mu_1.w_1 + \cdots + \mu_n.w_n = 0_V$ and thus $\lambda_1 = \cdots = \lambda_k = \mu_1 = \cdots = \mu_{n'} = 0_\mathbb{K}$ since $\mathcal{B}$ is linearly independent. Finally $\mathcal{B}_+$ is linearly independent.

Spanning family. Let v be a vector in $W + W'$. Then there exist $w \in W$ and $w' \in W'$ such that $v = w + w'$. Moreover

$$\begin{cases} \exists(\lambda_1, \ldots, \lambda_k, \mu_1, \ldots, \mu_n) \in \mathbb{K}^{k+n} \ / & w = \lambda_1.v_1 + \cdots + \lambda_k.v_k + \mu_1.w_1 + \cdots + \mu_n.w_n \\ \exists(\lambda'_1, \ldots, \lambda'_k, \mu'_1, \ldots, \mu'_{n'}) \in \mathbb{K}^{k+n'} \ / & w' = \lambda'_1.v_1 + \cdots + \lambda'_k.v_k + \mu'_1.w'_1 + \cdots + \mu'_{n'}.w'_{n'} \end{cases}$$

It follows that

$$\begin{aligned} v &= w + w' \\ &= (\lambda_1.v_1 + \cdots + \lambda_k.v_k + \mu_1.w_1 + \cdots + \mu_n.w_n) + (\lambda'_1.v_1 + \cdots + \lambda'_k.v_k + \mu'_1.w'_1 + \cdots + \mu'_{n'}.w'_{n'}) \\ &= (\lambda_1 + \lambda'_1).v_1 + \cdots + (\lambda_k + \lambda'_k).v_k + \mu_1.w_1 + \cdots + \mu_n.w_n + \mu'_1.w'_1 + \cdots + \mu'_{n'}.w'_{n'} \end{aligned}$$

Consequently $\mathcal{B}_+$ is a spanning family of $W + W'$.

Finally $\mathcal{B}_+$ is a basis of $W + W'$. It follows that

$$\begin{aligned} \dim(W + W') &= \mathrm{Card}(\mathcal{B}_+) \\ &= k + n + n' \\ &= (k + n) + (k + n') - k \\ &= \mathrm{Card}(\mathcal{B}) + \mathrm{Card}(\mathcal{B}') - \mathrm{Card}(\mathcal{B}_\cap) \\ &= \dim(W) + \dim(W') - \dim(W \cap W') \end{aligned}$$

■

Remark : It follows from Proposition 4.8 that every subspace of a finite-dimensional $\mathbb{K}$-vector space has a supplementary subspace.

4.2.4 Linear map on finite-dimensional vector space

Proposition 4.15

Let $f : V \to W$ be a linear map between two $\mathbb{K}$-vector spaces with V finite-dimensional. Then $\mathrm{Im}(f)$ is finite-dimensional.

Proof : Because if S is a spanning set of V then $f(S)$ is a spanning set of $\mathrm{Im}(f)$. ■

Definition 4.16 (*Rank*)

*Let $f : V \to W$ be a linear map between two $\mathbb{K}$-vector spaces with V finite-dimensional. The **rank** of f, denoted* $\mathbf{rank}(f)$*, is the dimension of the image of f.*

$$\mathrm{rank}(f) = \dim(\mathrm{Im}(f))$$

Theorem 4.17 (*Rank theorem*)

Let $f : V \to W$ be a linear map between two $\mathbb{K}$-vector spaces with V finite-dimensional. Then

$$\dim(V) = \dim(\mathrm{Ker}(f)) + \mathrm{rank}(f)$$

Proof : Let $\mathcal{B}_{\mathrm{Ker}} = (v_1, \ldots, v_k)$ be a basis of $\mathrm{Ker}(f)$. From the first point of Proposition 4.8, $\mathcal{B}_{\mathrm{Ker}}$ is a subfamily of some basis of V, say $\mathcal{B} = (v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+m})$. We are going to prove that $\mathcal{B}_{\mathrm{Im}} = (f(v_{k+1}), \ldots, f(v_{k+m}))$ is a basis of $\mathrm{Im}(f)$.

Linearly independent. Assume that $\lambda_{k+1}.f(v_{k+1}) + \cdots + \lambda_{k+m}.f(v_{k+m}) = 0_W$ for some scalars $(\lambda_{k+1}, \ldots, \lambda_{k+m}) \in \mathbb{K}^m$. Then $v = \lambda_{k+1}.v_{k+1} + \cdots + \lambda_{k+m}.v_{k+m}$ is in the kernel of f since

$$f(v) = \lambda_{k+1}.f(v_{k+1}) + \cdots + \lambda_{k+m}.f(v_{k+m}) = 0_W$$

Consequently $v = \lambda_1.v_1 + \cdots + \lambda_k.v_k$ for some scalars $(\lambda_1, \ldots, \lambda_k) \in \mathbb{K}^k$. It follows

$$v - v = \lambda_1.v_1 + \cdots + \lambda_k.v_k + (-\lambda_{k+1}).v_{k+1} + \cdots + (-\lambda_{k+m}).v_{k+m} = 0_V$$

Since $\mathcal{B}$ is linearly independent, we get $\lambda_1 = \cdots = \lambda_k = \lambda_{k+1} = \cdots = \lambda_{k+m} = 0_{\mathbb{K}}$. In particular $\mathcal{B}_{\mathrm{Im}}$ is linearly independent.

Spanning family. Let w be a vector in the image of f. Then there exists $v \in V$ such that $w = f(v)$. Moreover

$$\exists(\lambda_1, \ldots, \lambda_k, \lambda_{k+1}, \ldots, \lambda_{k+m}) \in \mathbb{K}^{k+m} \ / \ v = \lambda_1.v_1 + \cdots + \lambda_k.v_k + \lambda_{k+1}.v_{k+1} + \cdots + \lambda_{k+m}.v_{k+m}$$

It follows that

$$\begin{aligned} w &= f(v) \\ &= f(\lambda_1.v_1 + \cdots + \lambda_k.v_k + \lambda_{k+1}.v_{k+1} + \cdots + \lambda_{k+m}.v_{k+m}) \\ &= f(\underbrace{\lambda_1.v_1 + \cdots + \lambda_k.v_k}_{\in \mathrm{Ker}(f)}) + f(\lambda_{k+1}.v_{k+1} + \cdots + \lambda_{k+m}.v_{k+m}) \\ &= 0_W + f(\lambda_{k+1}.v_{k+1} + \cdots + \lambda_{k+m}.v_{k+m}) \\ &= \lambda_{k+1}.f(v_{k+1}) + \cdots + \lambda_{k+m}.f(v_{k+m}) \end{aligned}$$

Consequently $\mathcal{B}_{\mathrm{Im}}$ is a spanning family of $\mathrm{Im}(f)$.

Finally $\mathcal{B}_{\mathrm{Im}}$ is a basis of $\mathrm{Im}(f)$. It follows that

$$\begin{aligned} \dim(V) &= \mathrm{Card}(\mathcal{B}) \\ &= k + m \\ &= \mathrm{Card}(\mathcal{B}_{\mathrm{Ker}}) + \mathrm{Card}(\mathcal{B}_{\mathrm{Im}}) \\ &= \dim(\mathrm{Ker}(f)) + \mathrm{rank}(f) \end{aligned}$$

■

Corollary 4.18

Let $f \in L(V)$ be a linear map from a finite-dimensional $\mathbb{K}$-vector space V to itself. Then the following properties are equivalent

1. *f is bijective (that is $f \in GL(V)$)*
2. *f is injective*
3. *f is surjective*
4. $\mathrm{Ker}(f) = \{0_V\}$
5. $\mathrm{rank}(f) = \dim(V)$

Proposition 4.19

Let $f : V \to W$ be a linear map between two $\mathbb{K}$-vector spaces with V finite-dimensional. Then f is uniquely determined by the image of a basis of V. More precisely, if $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ is a basis of V then for every family $(w_1, w_2, \ldots, w_n)$ of vectors in W, there exists an unique linear map $f : V \to W$ such that $\forall k \in \{1, 2, \ldots, n\},\ f(v_k) = w_k$.

Proof : Since every vector $v \in V$ may be written uniquely as a linear combination of vectors in $\mathcal{B}$, the linear map f must be defined by

$$f(v) = \lambda_1.w_1 + \lambda_1.w_2 + \cdots + \lambda_1.w_n \quad \text{where } v = \lambda_1.v_1 + \lambda_1.v_2 + \cdots + \lambda_k.v_n$$

■

Proposition 4.20

Let V and W be two finite-dimensional $\mathbb{K}$-vector spaces. Then V and W are isomorphic if and only if they have the same dimension.

$$V \approx W \iff \dim(V) = \dim(W)$$

In particular, every finite-dimensional $\mathbb{K}$-vector space V is isomorphic to $\mathbb{K}^{\dim(V)}$.

Proof : **Necessary.** Assume there exists a bijective linear map $f : V \to W$. Then $\mathrm{Ker}(f) = \{0_V\}$ since f is injective and $\dim(W) = \dim(\mathrm{Im}(V)) = \mathrm{rank}(f)$ since f is surjective. Rank theorem 4.17 gives $\dim(V) = 0 + \dim(W) = \dim(W)$.

Sufficient. Let $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ be a basis of V and $\mathcal{B}' = (w_1, w_2, \ldots, w_n)$ be a basis of W. Now consider the linear map from V to W defined by

$$\forall k \in \{1, 2, \ldots, n\},\ f(v_k) = w_k$$

(f is well defined by Proposition 4.19.)
Then f is surjective since the image of f contains every vector of the spanning family $\mathcal{B}'$. In particular $\mathrm{rank}(f) = \dim(\mathrm{Im}(f)) = \dim(W) = \dim(V)$. Moreover Rank theorem 4.17 gives $\dim(\mathrm{Ker}(f)) = \dim(V) - \mathrm{rank}(f) = 0$ that is $\mathrm{Ker}(f) = \{0_V\}$. It follows that f is injective and thus an isomorphism. ■

Examples :

a) The dimension of $\mathbb{C}$ is 1 as a $\mathbb{C}$-vector space but 2 as a $\mathbb{R}$-vector space. More generally for any given integer $n \in \mathbb{N}$, the dimension of $\mathbb{C}^n$ is n as a $\mathbb{C}$-vector space but $2n$ as a $\mathbb{R}$-vector space.

b) For any given integer $d \in \mathbb{N}$, the space $\mathbb{K}_d[X]$ of all polynomials with degree at most d is isomorphic to the coordinate space $\mathbb{K}^{d+1}$ of dimension $d + 1$ (for instance the map which associates to a polynomial the family of its first $d + 1$ coefficients is an isomorphism).

Chapter 5

Matrices

In this chapter, fix a commutative field $(\mathbb{K}, +, \times)$ (for instance $\mathbb{K} = \mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$). We denote

- 0 the additive identity
- $-a$ the additive inverse of an element $a \in \mathbb{K}$
- 1 the multiplicative identity
- a^{-1} the multiplicative inverse of an element $a \in \mathbb{K}^* = \mathbb{K} - \{0\}$

5.1 Definition and operations

In this section, fix two positive integers $n \in \mathbb{N}^*$ and $p \in \mathbb{N}^*$.

5.1.1 The vector space $\mathcal{M}_{p,n}(\mathbb{K})$

Definition 5.1 (*Matrix*)

*Let $(m_{i,j})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}}$ be a family of elements in $\mathbb{K}$. The **matrix** with **entries** $(m_{i,j})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}}$ is the following rectangular arrangement of these elements*

$$M = (m_{i,j})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} = \begin{pmatrix} m_{1,1} & m_{1,2} & \dots & m_{1,n} \\ m_{2,1} & m_{2,2} & \dots & m_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{p,1} & m_{p,2} & \dots & m_{p,n} \end{pmatrix}$$

Moreover

- *the horizontal and vertical lines are respectively called **rows** and **columns***
- *the size $\boldsymbol{p \times n}$ of the rectangular arrangement, where p and n are respectively the numbers of rows and columns, is called the **dimension***
- *for any $i \in \{1, 2, \dots, p\}$ and any $j \in \{1, 2, \dots, n\}$, $m_{i,j}$ is called the **entry of the $\boldsymbol{i}^{th}$ row and the $\boldsymbol{j}^{th}$ column** or shortly the **$\boldsymbol{(i,j)}^{th}$ entry***

We denote $\boldsymbol{\mathcal{M}_{p,n}(\mathbb{K})}$ the set of all matrices with dimension $p \times n$ and entries in $\mathbb{K}$.

Examples :

a)
$$A = \begin{pmatrix} 1 & -1 & 0 & \sqrt{2} \\ \frac{3}{5} & 0 & 17 & 3 \\ 5\pi & 1+\sqrt{5} & 0 & -\frac{7}{2} \end{pmatrix} \in \mathcal{M}_{3,4}(\mathbb{R}) \quad \text{or} \quad B = \begin{pmatrix} 0 & -1+\imath \\ -1-\imath & 0 \end{pmatrix} \in \mathcal{M}_{2,2}(\mathbb{C})$$

The $(2,3)^{\text{th}}$ entry of A is 17 and its dimension is 3×4.

b) A matrix of dimension $p \times 1$ is called a **column vector** and a matrix of dimension $1 \times n$ is called a **row vector.**

$$C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{K}) \quad \text{and} \quad R = \begin{pmatrix} r_1 & r_2 & \dots & r_n \end{pmatrix} \in \mathcal{M}_{1,n}(\mathbb{K})$$

c) A 1×1 matrix is no more than an element in $\mathbb{K}$.

Definition 5.2 (*Addition and scalar multiplication in $\mathcal{M}_{n,p}(\mathbb{K})$*)

- *Let A and B be two matrices in $\mathcal{M}_{p,n}(\mathbb{K})$ with entries respectively $(a_{i,j})_{\substack{1\leqslant i\leqslant p \\ 1\leqslant j\leqslant n}}$ and $(b_{i,j})_{\substack{1\leqslant i\leqslant p \\ 1\leqslant j\leqslant n}}$. We define the sum of A and B, denoted $A+B$, to be the matrix in $\mathcal{M}_{p,n}(\mathbb{K})$ with entries $(a_{i,j}+b_{i,j})_{\substack{1\leqslant i\leqslant p \\ 1\leqslant j\leqslant n}}$. That provides a binary operation $+$ on $\mathcal{M}_{p,n}(\mathbb{K})$.*

$$\begin{aligned} A+B &= \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p,1} & a_{p,2} & \dots & a_{p,n} \end{pmatrix} + \begin{pmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,n} \\ b_{2,1} & b_{2,2} & \dots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p,1} & b_{p,2} & \dots & b_{p,n} \end{pmatrix} \\ &= \begin{pmatrix} a_{1,1}+b_{1,1} & a_{1,2}+b_{1,2} & \dots & a_{1,n}+b_{1,n} \\ a_{2,1}+b_{2,1} & a_{2,2}+b_{2,2} & \dots & a_{2,n}+b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p,1}+b_{p,1} & a_{p,2}+b_{p,2} & \dots & a_{p,n}+b_{p,n} \end{pmatrix} \end{aligned}$$

- *Let λ be a scalar in $\mathbb{K}$ and A be a matrix in $\mathcal{M}_{p,n}(\mathbb{K})$ with entries $(a_{i,j})_{\substack{1\leqslant i\leqslant p \\ 1\leqslant j\leqslant n}}$. We define the scalar multiplication of λ with A, denoted $\lambda.A$, to be the matrix in $\mathcal{M}_{p,n}(\mathbb{K})$ with entries $(\lambda a_{i,j})_{\substack{1\leqslant i\leqslant p \\ 1\leqslant j\leqslant n}}$. That provides an external binary operation . over $\mathbb{K}$ on $\mathcal{M}_{p,n}(\mathbb{K})$.*

$$\lambda.A = \lambda. \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p,1} & a_{p,2} & \dots & a_{p,n} \end{pmatrix} = \begin{pmatrix} \lambda a_{1,1} & \lambda a_{1,2} & \dots & \lambda a_{1,n} \\ \lambda a_{2,1} & \lambda a_{2,2} & \dots & \lambda a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{p,1} & \lambda a_{p,2} & \dots & \lambda a_{p,n} \end{pmatrix}$$

Proposition 5.3

$(\mathcal{M}_{p,n}(\mathbb{K}),+,.)$ is a finite-dimensional $\mathbb{K}$-vector space with $\dim(\mathcal{M}_{p,n}(\mathbb{K}))=pn$. Moreover the family of matrices $\mathcal{B}=(E_{1,1},E_{1,2},\ldots,E_{1,n},E_{2,1},E_{2,2},\ldots,E_{2,n},\ldots,E_{p,1},E_{p,2},\ldots,E_{p,n})$ where

$$\forall k\in\{1,2,\ldots,p\},\ \forall \ell\in\{1,2,\ldots,n\},\ E_{k,\ell}=\left(\delta_{i,j}^{k,\ell}=\left\{\begin{array}{l}1 \text{ if } (i,j)=(k,\ell)\\ 0 \text{ otherwise}\end{array}\right.\right)_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant n}}$$

*is a basis of $\mathcal{M}_{p,n}(\mathbb{K})$ called the **canonical basis**.*

Proof: Everything does as for $\mathbb{K}^{pn}$. Actually $\mathcal{M}_{p,n}(\mathbb{K})$ is isomorphic to $\mathbb{K}^{pn}$ (as $\mathbb{K}$-vector spaces). ■

Definition 5.4 (*Transpose*)

*Let A be a matrix in $\mathcal{M}_{p,n}(\mathbb{K})$ with entries $(a_{i,j})_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant n}}$. The **transpose** of A is the matrix ${}^t\boldsymbol{A}$ in $\mathcal{M}_{n,p}(\mathbb{K})$ with entries $(a_{j,i})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}}$.*

Example: The transpose of $A=\begin{pmatrix}1&2&3\\0&1&-1\end{pmatrix}\in\mathcal{M}_{2,3}(\mathbb{R})$ is ${}^tA=\begin{pmatrix}1&0\\2&1\\3&-1\end{pmatrix}\in\mathcal{M}_{3,2}(\mathbb{R})$.

Remark: The transpose map ${}^t:\mathcal{M}_{p,n}(\mathbb{K})\to\mathcal{M}_{n,p}(\mathbb{K}), A\mapsto {}^tA$ is a vector space isomorphism.

5.1.2 Multiplication of matrices

Definition 5.5 (*Multiplication of matrices*)

Let A be a matrix in $\mathcal{M}_{p,n}(\mathbb{K})$ with entries $(a_{i,j})_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant n}}$ and B be a matrix in $\mathcal{M}_{n,q}(\mathbb{K})$ (where $q\in\mathbb{N}^*$) with entries $(b_{i,j})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant q}}$. We define the product of A and B, denoted $A\times B$ or shortly AB, to be the matrix in $\mathcal{M}_{p,q}(\mathbb{K})$ with entries $(\sum_{k=1}^n a_{i,k}b_{k,j})_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant q}}$.

$$\begin{aligned}A\times B=AB &= \underbrace{\begin{pmatrix}a_{1,1}&a_{1,2}&\ldots&a_{1,n}\\a_{2,1}&a_{2,2}&\ldots&a_{2,n}\\\vdots&\vdots&\ddots&\vdots\\a_{p,1}&a_{p,2}&\ldots&a_{p,n}\end{pmatrix}}_{p\times n}\times\underbrace{\begin{pmatrix}b_{1,1}&b_{1,2}&\ldots&b_{1,q}\\b_{2,1}&b_{2,2}&\ldots&b_{2,q}\\\vdots&\vdots&\ddots&\vdots\\b_{n,1}&b_{n,2}&\ldots&b_{n,q}\end{pmatrix}}_{n\times q}\\ &= \underbrace{\begin{pmatrix}\sum_{k=1}^n a_{1,k}b_{k,1}&\sum_{k=1}^n a_{1,k}b_{k,2}&\ldots&\sum_{k=1}^n a_{1,k}b_{k,q}\\\sum_{k=1}^n a_{2,k}b_{k,1}&\sum_{k=1}^n a_{2,k}b_{k,2}&\ldots&\sum_{k=1}^n a_{2,k}b_{k,q}\\\vdots&\vdots&\ddots&\vdots\\\sum_{k=1}^n a_{p,k}b_{k,1}&\sum_{k=1}^n a_{p,k}b_{k,2}&\ldots&\sum_{k=1}^n a_{p,k}b_{k,q}\end{pmatrix}}_{p\times q}\end{aligned}$$

Remark: For a row vector $R\in\mathcal{M}_{1,n}(\mathbb{K})$ and a column vector $C\in\mathcal{M}_{n,1}(\mathbb{K})$, we get

$$RC=\begin{pmatrix}r_1&r_2&\ldots&r_n\end{pmatrix}\times\begin{pmatrix}c_1\\c_2\\\vdots\\c_n\end{pmatrix}=\sum_{k=1}^n r_kc_k=r_1c_1+r_2c_2+\cdots+r_nc_n\in\mathbb{K}$$

Consequently if we write $A \in \mathcal{M}_{p,n}(\mathbb{K})$ as a list of its rows and $B \in \mathcal{M}_{n,q}(\mathbb{K})$ as a list of its columns, we get

$$\begin{cases} A = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_p \end{pmatrix} \quad \text{where } \forall i \in \{1, 2, \dots p\},\ R_i = \begin{pmatrix} a_{i,1} & a_{i,2} & \dots & a_{i,n} \end{pmatrix} \in \mathcal{M}_{1,n}(\mathbb{K}) \\ B = \begin{pmatrix} C_1 & C_2 & \dots & C_q \end{pmatrix} \quad \text{where } \forall j \in \{1, 2, \dots q\},\ C_j = \begin{pmatrix} b_{1,j} \\ b_{2,j} \\ \vdots \\ b_{n,j} \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{K}) \end{cases}$$

and thus

$$AB = \begin{pmatrix} R_1C_1 & R_1C_2 & \dots & R_1C_q \\ R_2C_1 & R_2C_2 & \dots & R_2C_q \\ \vdots & \vdots & \ddots & \vdots \\ R_pC_1 & R_pC_2 & \dots & R_pC_q \end{pmatrix} \in \mathcal{M}_{p,q}(\mathbb{K})$$

Example : For instance, consider the following matrices:

$$A = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 1 & 0 \end{pmatrix} \in \mathcal{M}_{2,3}(\mathbb{R}) \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 6 & 2 & 3 \\ 1 & 3 & 4 & 1 \end{pmatrix} \in \mathcal{M}_{3,4}(\mathbb{R})$$

Then we get

$$\begin{aligned} AB &= \begin{pmatrix} 1\times1+3\times0+5\times1 & 1\times0+3\times6+5\times3 & 1\times2+3\times2+5\times4 & 1\times1+3\times3+5\times1 \\ 2\times1+1\times0+0\times1 & 2\times0+1\times6+0\times3 & 2\times2+1\times2+0\times4 & 2\times1+1\times3+0\times1 \end{pmatrix} \\ &= \begin{pmatrix} 6 & 33 & 28 & 15 \\ 2 & 6 & 6 & 5 \end{pmatrix} \in \mathcal{M}_{2,4}(\mathbb{R}) \end{aligned}$$

Proposition 5.6

Let A be a matrix in $\mathcal{M}_{p,n}(\mathbb{K})$ and B be a matrix in $\mathcal{M}_{n,q}(\mathbb{K})$ (where $q \in \mathbb{N}^*$). Then

$${}^t(\underbrace{A \times B}_{p\times q}) = \underbrace{{}^tB}_{q\times n} \times \underbrace{{}^tA}_{n\times p} \in \mathcal{M}_{q,p}(\mathbb{K})$$

Proof : Denote respectively $(a_{i,j})_{\substack{1\leqslant i\leqslant p \\ 1\leqslant j\leqslant n}}$ and $(b_{i,j})_{\substack{1\leqslant i\leqslant p \\ 1\leqslant j\leqslant q}}$ the entries of A and B. For every $i \in \{1, 2, \dots, q\}$ and $j \in \{1, 2, \dots, p\}$ the $(i,j)^{\text{th}}$ entry of the matrix ${}^t(AB)$ is

$$\sum_{k=1}^{n} a_{j,k}b_{k,i}$$

But the $(i,j)^{\text{th}}$ entry of the matrix ${}^tB{}^tA$ is also

$$\sum_{k=1}^{n} b_{k,i}a_{j,k} = \sum_{k=1}^{n} a_{j,k}b_{k,i}$$

The conclusion follows. ■

5.1.3 The ring $\mathcal{M}_n(\mathbb{K})$

Definition 5.7 (*Square matrix*)

*A **square matrix** is a matrix with the same number of rows and columns. We denote $\boldsymbol{\mathcal{M}_n(\mathbb{K})}$ the set of all square matrices with dimension $n \times n$ and entries in $\mathbb{K}$.*

Remark : In case $n = p = q$ in Definition 5.5, $\times$ provides a binary operation on $\mathcal{M}_n(\mathbb{K})$.

Proposition 5.8

$(\mathcal{M}_n(\mathbb{K}), +, \times)$ is an unital ring whose identity elements are respectively the zero matrix 0_n for addition and I_n for multiplication, where

$$0_n = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{K}) \quad \textit{and} \quad I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$$

Proof : 1. $(\mathcal{M}_n(\mathbb{K}), +)$ is an abelian group.

2. Let A, B and C be three matrices in $\mathcal{M}_n(\mathbb{K})$ with entries respectively $(a_{i,j})_{1\leqslant i,j\leqslant n}$, $(b_{i,j})_{1\leqslant i,j\leqslant n}$ and $(c_{i,j})_{1\leqslant i,j\leqslant n}$. For every indices i and j in $\{1, 2, \dots, n\}$ the $(i,j)^{\text{th}}$ entry of the matrix $(AB)C$ is
$$\sum_{\ell=1}^{n} \left(\sum_{k=1}^{n} a_{i,k} b_{k,\ell} \right) c_{\ell,j} = \sum_{1\leqslant k,\ell\leqslant n} a_{i,k} b_{k,\ell} c_{\ell,j}$$
And the $(i,j)^{\text{th}}$ entry of the matrix $A(BC)$ is
$$\sum_{k=1}^{n} a_{i,k} \left(\sum_{\ell=1}^{n} b_{k,\ell} c_{\ell,j} \right) = \sum_{1\leqslant k,\ell\leqslant n} a_{i,k} b_{k,\ell} c_{\ell,j}$$
Consequently $(AB)C = A(BC)$ and $\times$ is associative.

3. Let A, B and C be three matrices in $\mathcal{M}_n(\mathbb{K})$ with entries respectively $(a_{i,j})_{1\leqslant i,j\leqslant n}$, $(b_{i,j})_{1\leqslant i,j\leqslant n}$ and $(c_{i,j})_{1\leqslant i,j\leqslant n}$. For every indices i and j in $\{1, 2, \dots, n\}$ the $(i,j)^{\text{th}}$ entry of the matrix $A(B+C)$ is
$$\sum_{k=1}^{n} a_{i,k}(b_{k,j} + c_{k,j}) = \sum_{k=1}^{n} a_{i,k} b_{k,j} + \sum_{k=1}^{n} a_{i,k} c_{k,j}$$
Consequently $A(B+C) = AB + AC$ and the same goes for $(B+C)A = BA + CA$.Thus, $\times$ is distributive over $+$ on $\mathcal{M}_n(\mathbb{K})$.

4. Let A be a matrix in $\mathcal{M}_n(\mathbb{K})$ with entries $(a_{i,j})_{1\leqslant i,j\leqslant n}$. For every indices i and j in $\{1, 2, \dots, n\}$ the $(i,j)^{\text{th}}$ entry of the matrix AI_n is
$$\begin{aligned} & \sum_{k=1}^{n} a_{i,k} \delta_{k,j} \quad \text{where } \delta_{k,j} = \begin{cases} 1 \text{ if } k = j \\ 0 \text{ otherwise} \end{cases} \\ = \; & a_{i,1} \times 0 + \dots + a_{i,j-1} \times 0 + a_{i,j} \times 1 + a_{i,j+1} \times 0 + \dots + a_{i,n} \times 0 \\ = \; & a_{i,j} \end{aligned}$$
Consequently $AI_n = A$ and the same goes for $I_n A = A$. Thus, the matrix I_n is the multiplicative identity for $\times$ in $\mathcal{M}_n(\mathbb{K})$.

Finally, $\mathcal{M}_n(\mathbb{K})$ satisfies all conditions to be an unital ring. ■

Remarks :

- But $\times$ is not commutative on $\mathcal{M}_n(\mathbb{K})$ (as soon as $n \geqslant 2$). For instance, we have:

$$E_{1,n} \times E_{n,1} = \begin{pmatrix} 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} = E_{1,1}$$

  but

$$E_{n,1} \times E_{1,n} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} = 0_n$$

- Moreover $(\mathcal{M}_n(\mathbb{K}), +, \times)$ is not a field (as soon as $n \geqslant 2$) since there exist some zero divisors (for instance $E_{n,1}$ and $E_{1,n}$) which can not have multiplicative inverse.
- Actually we have:

$$\forall (i,j,k,\ell) \in \{1,2,\dots,n\}^4,\ E_{i,j} \times E_{k,\ell} = \delta_{j,k}.E_{i,\ell} \quad \text{where } \delta_{j,k} = \begin{cases} 1 \text{ if } j = k \\ 0 \text{ otherwise} \end{cases}$$

Definition 5.9 (*General linear group*)

We denote $\boldsymbol{\mathcal{GL}_n(\mathbb{K})}$ *the set of all square matrices in* $\mathcal{M}_n(\mathbb{K})$ *which have a multiplicative inverse.*

$$\mathcal{GL}_n(\mathbb{K}) = \{A \in \mathcal{M}_n(\mathbb{K}) \ / \ \exists B \in \mathcal{M}_n(\mathbb{K}),\ AB = BA = I_n\}$$

Example : $\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \in \mathcal{GL}_2(\mathbb{R})$ with $\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$.

Remark : $(\mathcal{GL}_n(\mathbb{K}), \times)$ is a group but a sum of matrices in $\mathcal{GL}_n(\mathbb{K})$ is not always in $\mathcal{GL}_n(\mathbb{K})$.

5.2 Matrices associated to vectors and linear maps

In this section, fix two finite-dimensional $\mathbb{K}$-vector spaces V and W.

5.2.1 Coordinate vector

Example for the coordinate space $V = \mathbb{K}^n$:

Let $v = (x_1, x_2, \dots, x_n)$ be a vector in $\mathbb{K}^n$. Since $\mathbb{K}^n$ is isomorphic to $\mathcal{M}_{n,1}(\mathbb{K})$ (both vector spaces are of dimension n), we may associate to v the following column vector

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{K})$$

Equivalently if $\mathcal{B} = (e_1, e_2, \dots, e_n)$ is the canonical basis of $\mathbb{K}^n$ then

$$\begin{cases} v &= x_1.e_1 &+ x_2.e_2 &+ \dots &+ x_n.e_n \\ X &= x_1.E_{1,1} &+ x_2.E_{2,1} &+ \dots &+ x_n.X_{n,1} \end{cases}$$

Example : For instance we may associate to the vector $v = (2, 4, 3) \in \mathbb{R}^3$ the following column vector

$$X = \begin{pmatrix} 2 \\ 4 \\ 3 \end{pmatrix} \in \mathcal{M}_{3,1}(\mathbb{R})$$

with respect to the canonical basis $\mathcal{B} = \Big(e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)\Big)$ of $\mathbb{R}^3$.

But notice that $\mathcal{B}' = \Big(v_1 = (0, 0, 1), v_2 = (1, 0, -1), v_3 = (1, 2, 3)\Big)$ is also a basis of $\mathbb{R}^3$ and we have $v = -3.v_1 + 0.v_2 + 2.v_3$. So we may also associate to the vector v the following column vector

$$Y = \begin{pmatrix} -3 \\ 0 \\ 2 \end{pmatrix} \in \mathcal{M}_{3,1}(\mathbb{R})$$

with respect to the basis $\mathcal{B}'$.

Definition 5.10 (*Coordinate vector*)

Let $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ be a basis of V and v be a vector in V. Then v may be written uniquely as a linear combination of vectors in $\mathcal{B}$:

$$\exists!(x_1, x_2, \ldots, x_n) \in \mathbb{K}^n \ / \ v = x_1.v_1 + x_2.v_2 + \cdots + x_n.v_n$$

*The **coordinate vector of v in $\mathcal{B}$**, denoted $\mathbf{Mat_{\mathcal{B}}(v)}$, is the following column vector*

$$\mathrm{Mat}_{\mathcal{B}}(v) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{K}) \quad \text{where } n = \dim(V)$$

Remark : For every basis $\mathcal{B}$ of V, the map $\mathrm{Mat}_{\mathcal{B}} : V \to \mathcal{M}_{n,1}(\mathbb{K})$ is a vector space isomorphism.

5.2.2 Matrix associated to a linear map

Recall that a linear map $f : V \to W$ is uniquely determined by the image of a basis of V.

Definition 5.11 (*Matrix associated to a linear map*)

Let $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ and $\mathcal{C} = (w_1, w_2, \ldots, w_p)$ be two bases of respectively V and W and $f : V \to W$ be a linear map. Then for every $j \in \{1, 2, \ldots, n\}$, $f(v_j)$ may be written as a linear combination of vectors in $\mathcal{C}$:

$$\forall j \in \{1, 2, \ldots, n\}, \ \exists!(\lambda_{1,j}, \lambda_{2,j}, \ldots, \lambda_{p,j}) \in \mathbb{K}^p \ / \ f(v_j) = \lambda_{1,j}.w_1 + \lambda_{2,j}.w_2 + \cdots + \lambda_{p,j}.w_p$$

*The **matrix associated to f in $\mathcal{B}$ and $\mathcal{C}$**, denoted $\mathbf{Mat_{\mathcal{C},\mathcal{B}}(f)}$, is the following matrix*

$$\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f) = (\lambda_{i,j})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} = \begin{pmatrix} \lambda_{1,1} & \lambda_{1,2} & \ldots & \lambda_{1,n} \\ \lambda_{2,1} & \lambda_{2,2} & \ldots & \lambda_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{p,1} & \lambda_{p,2} & \ldots & \lambda_{p,n} \end{pmatrix} \in \mathcal{M}_{p,n}(\mathbb{K}) \quad \text{where } \begin{cases} n = \dim(V) \\ p = \dim(W) \end{cases}$$

Remark : If we write $\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f)$ as a list of its columns, we get

$$\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f) = \Big(\begin{array}{cccc} \mathrm{Mat}_{\mathcal{C}}(f(v_1)) & \mathrm{Mat}_{\mathcal{C}}(f(v_2)) & \dots & \mathrm{Mat}_{\mathcal{C}}(f(v_n)) \end{array} \Big) \in \mathcal{M}_{p,n}(\mathbb{K})$$

where

$$\forall j \in \{1, 2, \dots, n\},\ \mathrm{Mat}_{\mathcal{C}}(f(v_j)) \in \mathcal{M}_{p,1}(\mathbb{K})$$

Examples :

a) Consider the following linear map from the real space $V = \mathbb{R}^3$ provided with its canonical basis $\mathcal{B}$ to the real plane $W = \mathbb{R}^2$ provided with its canonical basis $\mathcal{C}$

$$f : v = (x, y, z) \longmapsto f(v) = (2x - y, 3z)$$

Then

$$\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f) = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix} \in \mathcal{M}_{2,3}(\mathbb{R})$$

b) Let $r_{\pi/2}$ be the rotation by the angle $\frac{\pi}{2}$ counterclockwise about the origin in the real plane $V = \mathbb{R}^2$ provided with its canonical basis $\mathcal{B} = (e_1, e_2)$. Then $r_{\pi/2}(e_1) = e_2$ and $r_{\pi/2}(e_2) = -e_1$, that is

$$\begin{array}{rccc} r_{\pi/2} : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & v = (x, y) & \longmapsto & f(v) = (-y, x) \end{array}$$

And

$$\mathrm{Mat}_{\mathcal{B},\mathcal{B}}(r_{\pi/2}) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathcal{M}_{2,2}(\mathbb{R})$$

More generally, if r_θ is the rotation by the angle $\theta \in \mathbb{R}$ counterclockwise about the origin then

$$R_\theta = \mathrm{Mat}_{\mathcal{B},\mathcal{B}}(r_\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \in \mathcal{M}_{2,2}(\mathbb{R})$$

c) Let ∂ be the differentiation map $P \mapsto P'$ from $V = \mathbb{R}_4[X]$ to itself. In the basis $\mathcal{B} = (1, X, X^2, X^3, X^4)$ we have:

$$\mathrm{Mat}_{\mathcal{B},\mathcal{B}}(\partial) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathcal{M}_{5,5}(\mathbb{R})$$

Remark : The matrix associated to the identity function Id_V from V to itself in any basis $\mathcal{B}$ of V is

$$\mathrm{Mat}_{\mathcal{B},\mathcal{B}}(\mathrm{Id}_V) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} = I_n \in \mathcal{M}_n(\mathbb{K}) \quad \text{where } n = \dim(V)$$

Proposition 5.12

Let $\mathcal{B}$ and $\mathcal{C}$ be two bases of respectively V and W. Let $f : V \to W$ be a linear map and v be a vector in V. Then

$$\underbrace{\mathrm{Mat}_{\mathcal{C}}(f(v))}_{\dim(W)\times 1} = \underbrace{\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f)}_{\dim(W)\times\dim(V)} \times \underbrace{\mathrm{Mat}_{\mathcal{B}}(v)}_{\dim(V)\times 1}$$

Proof : Denote by $(v_1, v_2, \ldots, v_n)$ the vectors in $\mathcal{B}$, by $(w_1, w_2, \ldots, w_p)$ the vectors in $\mathcal{C}$, by $(\lambda_{i,j})_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant n}}$ the entries of $\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f)$ and by $(x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$ the coordinates of the vector v written in the basis $\mathcal{B}$. Then

$$\begin{aligned}
f(v) &= f(x_1.v_1 + x_2.v_2 + \cdots + x_n.v_n)\\
&= x_1.f(v_1) + x_2.f(v_2) + \cdots + x_n.f(v_n)\\
&= \sum_{k=1}^{n} x_k.f(v_k)\\
&= \sum_{k=1}^{n} x_k.(\lambda_{1,k}.w_1 + \lambda_{2,k}.w_2 + \cdots + \lambda_{p,k}.w_p)\\
&= \sum_{k=1}^{n} \Big((x_k\lambda_{1,k}).w_1 + (x_k\lambda_{2,k}).w_2 + \cdots + (x_k\lambda_{p,k}).w_p\Big)\\
&= \left(\sum_{k=1}^{n} \lambda_{1,k}x_k\right).w_1 + \left(\sum_{k=1}^{n} \lambda_{2,k}x_k\right).w_2 + \cdots + \left(\sum_{k=1}^{n} \lambda_{p,k}x_k\right).w_p
\end{aligned}$$

The result follows. ■

Examples :

a) The image of $v = (x, y) \in \mathbb{R}^2$ under rotation by the angle $\theta \in \mathbb{R}$ counterclockwise about the origin is $r_\theta(v) = (\cos(\theta)x - \sin(\theta)y, \sin(\theta)x + \cos(\theta)y)$ since

$$\begin{aligned}
\mathrm{Mat}_{\mathcal{B}}(r_\theta(v)) &= \mathrm{Mat}_{\mathcal{B},\mathcal{B}}(r_\theta)\,\mathrm{Mat}_{\mathcal{B}}(v)\\
&= R_\theta \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(\theta)x - \sin(\theta)y \\ \sin(\theta)x + \cos(\theta)y \end{pmatrix}
\end{aligned}$$

b) The derivative polynomial of $P(X) = -7 + 8X - 5X^2 + 2X^4 \in \mathbb{R}_4[X]$ may be computed as follows

$$\mathrm{Mat}_{\mathcal{B}}(P') = \mathrm{Mat}_{\mathcal{B},\mathcal{B}}(\partial)\,\mathrm{Mat}_{\mathcal{B}}(P) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -7 \\ 8 \\ -5 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 8 \\ -10 \\ 0 \\ 8 \\ 0 \end{pmatrix}$$

Consequently, $P'(X) = 8 - 10X + 8X^3$.

Proposition 5.13

Let $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ be three bases of respectively U, V and W (where U is a finite-dimensional $\mathbb{K}$-vector space). Let $f : V \to W$ and $g : U \to V$ be two linear maps. Then $f \circ g : U \to W$ is a linear map and

$$\underbrace{\mathrm{Mat}_{\mathcal{C},\mathcal{A}}(f \circ g)}_{\dim(W)\times\dim(U)} = \underbrace{\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f)}_{\dim(W)\times\dim(V)} \times \underbrace{\mathrm{Mat}_{\mathcal{B},\mathcal{A}}(g)}_{\dim(V)\times\dim(U)}$$

Proof : Denote by $(u_1, u_2, \dots, u_n)$ the vectors in $\mathcal{A}$. Writing $\mathrm{Mat}_{\mathcal{B},\mathcal{A}}(g)$ as a list of its columns, we get

$$\begin{aligned}
&\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f) \times \mathrm{Mat}_{\mathcal{B},\mathcal{A}}(g)\\
=\ &\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f) \times \begin{pmatrix} \mathrm{Mat}_{\mathcal{B}}(g(u_1)) & \mathrm{Mat}_{\mathcal{B}}(g(u_2)) & \dots & \mathrm{Mat}_{\mathcal{B}}(g(u_n)) \end{pmatrix}\\
=\ &\begin{pmatrix} \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f) \times \mathrm{Mat}_{\mathcal{B}}(g(u_1)) & \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f) \times \mathrm{Mat}_{\mathcal{B}}(g(u_2)) & \dots & \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f) \times \mathrm{Mat}_{\mathcal{B}}(g(u_n)) \end{pmatrix}\\
=\ &\begin{pmatrix} \mathrm{Mat}_{\mathcal{C}}(f(g(u_1))) & \mathrm{Mat}_{\mathcal{C}}(f(g(u_2))) & \dots & \mathrm{Mat}_{\mathcal{C}}(f(g(u_n))) \end{pmatrix}\\
=\ &\begin{pmatrix} \mathrm{Mat}_{\mathcal{C}}((f\circ g)(u_1)) & \mathrm{Mat}_{\mathcal{C}}((f\circ g)(u_2)) & \dots & \mathrm{Mat}_{\mathcal{C}}((f\circ g)(u_n)) \end{pmatrix}\\
=\ &\mathrm{Mat}_{\mathcal{C},\mathcal{A}}(f\circ g)
\end{aligned}$$

where the second equality comes from Definition 5.5 and the third equality from Proposition 5.12. ■

Examples :

a) For every angles $\theta \in \mathbb{R}$ and $\varphi \in \mathbb{R}$ we have

$$\begin{aligned}
R_\theta R_\varphi &= \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}\\
&= \begin{pmatrix} \cos(\theta)\cos(\varphi) - \sin(\theta)\sin(\varphi) & -\cos(\theta)\sin(\varphi) - \sin(\theta)\cos(\varphi) \\ \sin(\theta)\cos(\varphi) + \cos(\theta)\sin(\varphi) & -\sin(\theta)\sin(\varphi) + \cos(\theta)\sin(\varphi) \end{pmatrix}\\
&= \begin{pmatrix} \cos(\theta+\varphi) & -\sin(\theta+\varphi) \\ \sin(\theta+\varphi) & \cos(\theta+\varphi) \end{pmatrix}\\
&= R_{\theta+\varphi}
\end{aligned}$$

Consequently $r_\theta \circ r_\varphi = r_{\theta+\varphi}$ as expected.

b) We have after some computation:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}^5 = \underbrace{\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \times \cdots \times \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}}_{5 \text{ times}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Consequently for every polynomial $P(X)$ in $\mathbb{R}_4[X]$:

$$\mathrm{Mat}_{\mathcal{B}}(P^{(5)}) = \mathrm{Mat}_{\mathcal{B},\mathcal{B}}(\partial^5) \times \mathrm{Mat}_{\mathcal{B}}(P) = (\mathrm{Mat}_{\mathcal{B},\mathcal{B}}(\partial))^5 \times \mathrm{Mat}_{\mathcal{B}}(P) = 0_5 \times \mathrm{Mat}_{\mathcal{B}}(P) = \mathrm{Mat}_{\mathcal{B}}(0)$$

Equivalently the 5^{th} order polynomial derivative of any polynomial in $\mathbb{R}_4[X]$ is the zero constant polynomial.

Corollary 5.14

Let $\mathcal{B}$ and $\mathcal{C}$ be two bases of respectively V and W with dimension respectively n and p. The following properties hold

1. $\mathrm{Mat}_{\mathcal{C},\mathcal{B}} : L(V,W) \to \mathcal{M}_{p,n}(\mathbb{K})$ *is a vector space isomorphism.*
2. $\mathrm{Mat}_{\mathcal{B},\mathcal{B}} : L(V) \to \mathcal{M}_n(\mathbb{K})$ *is a vector space isomorphism and a ring isormorphism.*
3. $\mathrm{Mat}_{\mathcal{B},\mathcal{B}} : GL(V) \to \mathcal{G}\ell_n(\mathbb{K})$ *is a group isomorphism.*

Remark : Matrices allow arbitrary linear maps on finite-dimensional vector spaces to be represented in a convenient form, suitable for computation. Conversely, every matrix in $\mathcal{M}_{p,n}(\mathbb{K})$ may be considered as a matrix associated to a linear map from $\mathbb{K}^n$ to $\mathbb{K}^p$ in the canonical bases.

5.2.3 Change of basis

Definition 5.15 (*Transition matrix*)

Let $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ and $\mathcal{B}' = (v'_1, v'_2, \ldots, v'_n)$ be two bases of V. For every $j \in \{1, 2, \ldots, n\}$, v'_j may be written as a linear combination of vectors in $\mathcal{B}$:

$$\forall j \in \{1, 2, \ldots, n\},\ \exists!(\lambda_{1,j}, \lambda_{2,j}, \ldots, \lambda_{n,j}) \in \mathbb{K}^n \ / \ v'_j = \lambda_{1,j}.v_1 + \lambda_{2,j}.v_2 + \cdots + \lambda_{n,j}.v_n$$

*The **transition matrix from $\mathcal{B}$ to $\mathcal{B}'$**, denoted $\boldsymbol{T_{\mathcal{B}\to\mathcal{B}'}}$, is the following matrix*

$$T_{\mathcal{B}\to\mathcal{B}'} = (\lambda_{i,j})_{1\leqslant i,j\leqslant n} = \begin{pmatrix} \lambda_{1,1} & \lambda_{1,2} & \ldots & \lambda_{1,n} \\ \lambda_{2,1} & \lambda_{2,2} & \ldots & \lambda_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{n,1} & \lambda_{n,2} & \ldots & \lambda_{n,n} \end{pmatrix} \in \mathcal{M}_n(\mathbb{K}) \quad \text{where } n = \dim(V)$$

Remark : If we write $T_{\mathcal{B}\to\mathcal{B}'}$ as a list of its columns, we get

$$T_{\mathcal{B}\to\mathcal{B}'} = \begin{pmatrix} \mathrm{Mat}_{\mathcal{B}}(v'_1) & \mathrm{Mat}_{\mathcal{B}}(v'_2) & \ldots & \mathrm{Mat}_{\mathcal{B}}(v'_n) \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$$

Example : Consider the canonical basis $\mathcal{B} = \Big(e_1 = (1,0,0), e_2 = (0,1,0),\ e_3 = (0,0,1)\Big)$ and the basis $\mathcal{B}' = \Big(v_1 = (0,0,1), v_2 = (1,0,-1), v_3 = (1,2,3)\Big)$ of the real space $V = \mathbb{R}^3$. Then

$$T_{\mathcal{B}\to\mathcal{B}'} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 1 & -1 & 3 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$$

Moreover we have:

$$\begin{cases} v_1 &=& 0.e_1 + 0.e_2 + 1.e_3 \\ v_2 &=& 1.e_1 + 0.e_2 - 1.e_3 \\ v_3 &=& 1.e_1 + 2.e_2 + 3.e_3 \end{cases} \iff \begin{cases} v_1 &=& e_3 \\ v_2 &=& e_1 - e_3 \\ v_3 &=& e_1 + 2.e_2 + 3.e_3 \end{cases}$$

$$\iff \begin{cases} e_3 &=& v_1 \\ e_1 &=& v_2 + e_3 = v_2 + v_1 \\ e_2 &=& \frac{1}{2}.(v_3 - e_1 - 3.e_3) = \frac{1}{2}.(v_3 - v_2 - v_1 - 3.v_1) \end{cases}$$

$$\iff \begin{cases} e_1 &=& 1.v_1 + 1.v_2 + 0.v_3 \\ e_2 &=& -2.v_1 - \frac{1}{2}.v_2 + \frac{1}{2}.v_3 \\ e_2 &=& 1.v_1 + 0.v_2 + 0.v_3 \end{cases}$$

Consequently

$$T_{\mathcal{B}'\to\mathcal{B}} = \begin{pmatrix} 1 & -2 & 1 \\ 1 & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$$

Remark that

$$T_{\mathcal{B}\to\mathcal{B}'} \times T_{\mathcal{B}'\to\mathcal{B}} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 1 & -1 & 3 \end{pmatrix} \begin{pmatrix} 1 & -2 & 1 \\ 1 & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3$$

$$T_{\mathcal{B}\to\mathcal{B}'} \times T_{\mathcal{B}'\to\mathcal{B}} = \begin{pmatrix} 1 & -2 & 1 \\ 1 & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 1 & -1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3$$

Proposition 5.16

Let $\mathcal{B}$ and $\mathcal{B}'$ be two bases of V with dimension n. The following properties hold

1. $T_{\mathcal{B}\to\mathcal{B}'} = \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(\mathrm{Id}_V)$
2. $T_{\mathcal{B}\to\mathcal{B}'} \in \mathcal{G}\ell_n(\mathbb{K})$ *and* $(T_{\mathcal{B}\to\mathcal{B}'})^{-1} = T_{\mathcal{B}'\to\mathcal{B}}$

Proof : 1. That follows from Definition 5.11.

2. From Proposition 5.13, we get

$$T_{\mathcal{B}\to\mathcal{B}'} \times T_{\mathcal{B}'\to\mathcal{B}} = \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(\mathrm{Id}_V) \times \mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{Id}_V) = \mathrm{Mat}_{\mathcal{B},\mathcal{B}}(\mathrm{Id}_V) = I_n$$

And the same goes for $T_{\mathcal{B}'\to\mathcal{B}} \times T_{\mathcal{B}\to\mathcal{B}'} = I_n$. The conclusion follows. ■

Corollary 5.17 (*Change of basis*)

Let $\mathcal{B}$ and $\mathcal{B}'$ be two bases of V. Let $\mathcal{C}$ and $\mathcal{C}'$ be two bases of W. Let $f : V \to W$ be a linear map and v be a vector in V. Then the following properties hold

1. $\mathrm{Mat}_{\mathcal{B}'}(v) = T_{\mathcal{B}'\to\mathcal{B}} \times \mathrm{Mat}_{\mathcal{B}}(v) = (T_{\mathcal{B}\to\mathcal{B}'})^{-1} \times \mathrm{Mat}_{\mathcal{B}}(v)$
2. $\mathrm{Mat}_{\mathcal{C}',\mathcal{B}'}(f) = T_{\mathcal{C}'\to\mathcal{C}} \times \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f) \times T_{\mathcal{B}\to\mathcal{B}'} = (T_{\mathcal{C}\to\mathcal{C}'})^{-1} \times \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f) \times T_{\mathcal{B}\to\mathcal{B}'}$

Remark : In case $f : V \to V$ is a vector space endomorphism:

$$\mathrm{Mat}_{\mathcal{B}',\mathcal{B}'}(f) = (T_{\mathcal{B}\to\mathcal{B}'})^{-1} \times \mathrm{Mat}_{\mathcal{B},\mathcal{B}}(f) \times T_{\mathcal{B}\to\mathcal{B}'}$$

Example : Consider the linear map f from the real space $V = \mathbb{R}^3$ to itself whose associated matrix in the canonical basis $\mathcal{B}$ is

$$\mathrm{Mat}_{\mathcal{B},\mathcal{B}}(f) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ -1 & 2 & 1 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$$

Equivalently, f may be defined as follows

$$\begin{array}{rccc} f : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & v = (x, y, z) & \longmapsto & f(v) = (2x, 2y, -x + 2y + z) \end{array}$$

Now in the basis $\mathcal{B}' = \Big(v_1 = (0, 0, 1), v_2 = (1, 0, -1), v_3 = (1, 2, 3)\Big)$, we have:

$$\begin{aligned} \mathrm{Mat}_{\mathcal{B}',\mathcal{B}'}(f) &= T_{\mathcal{B}'\to\mathcal{B}} \times \mathrm{Mat}_{\mathcal{B},\mathcal{B}}(f) \times T_{\mathcal{B}\to\mathcal{B}'} \\ &= \begin{pmatrix} 1 & -2 & 1 \\ 1 & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 1 & -1 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -2 & 1 \\ 1 & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 0 & 2 & 2 \\ 0 & 0 & 4 \\ 1 & -2 & 6 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \end{aligned}$$

It follows that f may be simply described as follows

$$f(v) = x_1.v_1 + 2x_2.v_2 + 2x_3.v_3 \quad \text{where } v = x_1.v_1 + x_2.v_2 + x_3.v_3 \in \mathbb{R}^3$$

For instance, consider the vector $v = (2, 4, 3) \in \mathbb{R}^3$. The coordinate vector of v in $\mathcal{B}$ is

$$X = \mathrm{Mat}_{\mathcal{B}}(v) = \begin{pmatrix} 2 \\ 4 \\ 3 \end{pmatrix} \in \mathcal{M}_{3,1}(\mathbb{R})$$

And the coordinate vector of v in $\mathcal{B}'$ is

$$Y = \mathrm{Mat}_{\mathcal{B}'}(v) = T_{\mathcal{B}'\to\mathcal{B}} \times \mathrm{Mat}_{\mathcal{B}}(v) = \begin{pmatrix} 1 & -2 & 1 \\ 1 & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 4 \\ 3 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \\ 2 \end{pmatrix} \in \mathcal{M}_{3,1}(\mathbb{R})$$

Consequently $v = -3.v_1 + 2.v_3$ and $f(v) = -3.v_1 + 4.v_3$.

5.2.4 Rank

Definition 5.18 (*Rank*)

Fix two positive integers $n \in \mathbb{N}^*$ and $p \in \mathbb{N}^*$ and let M be a matrix in $\mathcal{M}_{p,n}(\mathbb{K})$ with entries $(m_{i,j})_{\substack{1\leqslant i\leqslant p \\ 1\leqslant j\leqslant n}}$. M may be written as a list of its columns

$$M = \begin{pmatrix} C_1 & C_2 & \dots & C_n \end{pmatrix} \in \mathcal{M}_{p,n}(\mathbb{K}) \quad \text{where } \forall j \in \{1, 2, \dots, n\},\ C_j = \begin{pmatrix} m_{1,j} \\ m_{2,j} \\ \vdots \\ m_{p,j} \end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{K})$$

*The **rank** of M, denoted **rank(M)**, is the dimension of the subspace of $\mathcal{M}_{p,1}(\mathbb{K})$ spanned by the columns of M.*

$$\mathrm{rank}(M) = \dim\Big(\mathrm{Span}(\{C_1, C_2, \dots, C_n\})\Big)$$

Remark : The rank of $M \in \mathcal{M}_{p,n}(\mathbb{K})$ is less than $\dim(\mathcal{M}_{p,1}(\mathbb{K})) = p$ by definition and than n since $(C_1, C_2, \dots, C_n)$ is a spanning family of n vectors (in the subspace spanned by the columns of M). It follows that $\mathrm{rank}(M) \in \{0, 1, \dots, \min\{n, p\}\}$.

Proposition 5.19

The rank of a matrix associated to a linear map $f : V \to W$ in any bases $\mathcal{B}$ and $\mathcal{C}$ of respectively V and W is equal to the rank of f.

$$\mathrm{rank}\Big(\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f)\Big) = \mathrm{rank}(f)$$

Proof : Denote by $(v_1, v_2, \dots, v_n)$ the vectors in $\mathcal{B}$. If we write $\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f)$ as a list of its columns, we get

$$\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f) = \begin{pmatrix} \mathrm{Mat}_{\mathcal{C}}(f(v_1)) & \mathrm{Mat}_{\mathcal{C}}(f(v_2)) & \dots & \mathrm{Mat}_{\mathcal{C}}(f(v_n)) \end{pmatrix} \in \mathcal{M}_{p,n}(\mathbb{K})$$

Then

$$\begin{aligned}\text{rank}\Big(\text{Mat}_{\mathcal{C},\mathcal{B}}(f)\Big) &= \dim\Big(\text{Span}(\{\text{Mat}_{\mathcal{C}}(f(v_1)),\text{Mat}_{\mathcal{C}}(f(v_2)),\dots,\text{Mat}_{\mathcal{C}}(f(v_n))\})\Big)\\ &= \dim\Big(\text{Mat}_{\mathcal{C}}(f(\underbrace{\text{Span}(\{v_1,v_2,\dots,v_n\})}_{V}))\Big)\\ &= \dim\Big(\text{Mat}_{\mathcal{C}}(f(V))\Big)\end{aligned}$$

Since $\text{Mat}_{\mathcal{C}} : W \to \mathcal{M}_{p,1}(\mathbb{K})$ is a vector space isomorphism, it follows that

$$\text{rank}\Big(\text{Mat}_{\mathcal{C},\mathcal{B}}(f)\Big) = \dim\Big(\text{Mat}_{\mathcal{C}}(f(V))\Big) = \dim\Big(f(V)\Big) = \dim\Big(\text{Im}(f)\Big) = \text{rank}(f)$$

■

Proposition 5.20

Fix two positive integers $n \in \mathbb{N}^*$ and $p \in \mathbb{N}^*$ and let M be a matrix in $\mathcal{M}_{p,n}(\mathbb{K})$. The rank of M is equal to $r \in \{0,1,\dots,\min\{n,p\}\}$ if and only if there exist $P \in \mathcal{G}\ell_p(\mathbb{K})$ and $Q \in \mathcal{G}\ell_n(\mathbb{K})$ such that

$$PMQ = \underbrace{\left(\begin{array}{cccc|ccc} 1 & 0 & \dots & 0 & 0 & \dots & 0\\ 0 & 1 & \dots & 0 & 0 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & 1 & 0 & \dots & 0\\ \hline 0 & 0 & \dots & 0 & 0 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{array}\right)}_{r \text{ columns}\quad (n-r) \text{ columns}} \begin{array}{l}\left.\begin{array}{l}\\\\\\\end{array}\right\} r \text{ rows}\\ \left.\begin{array}{l}\\\\\\\end{array}\right\} (p-r) \text{ rows}\end{array} \in \mathcal{M}_{p,n}(\mathbb{K})$$

Proof : Denote by J_r the matrix above and by $f : \mathbb{K}^n \to \mathbb{K}^p$ the linear map such that $M = \text{Mat}_{\mathcal{C},\mathcal{B}}(f)$ where $\mathcal{B}$ and $\mathcal{C}$ are the canonical bases of respectively $\mathbb{K}^n$ and $\mathbb{K}^p$.

Necessary. If $r = \text{rank}(M) = \text{rank}(f) = \dim(\text{Im}(f))$ then $\dim(\text{Ker}(f)) = \dim(\mathbb{K}^n) - \dim(\text{Im}(f)) = n - r$ from the rank theorem. Denote by $(v_{r+1}, v_{r+2}, \dots, v_n)$ a basis of $\text{Ker}(f)$. Since this family is linearly independent, it is a subfamily of some basis of $\mathbb{K}^n$, say $\mathcal{B}' = (v_1, v_2, \dots, v_n)$. Now, consider the family $(w_1 = f(v_1), w_2 = f(v_2), \dots, w_r = f(v_r))$. By construction, it is a basis of $\text{Im}(f)$, thus a subfamily of some basis of $\mathbb{K}^p$, say $\mathcal{C}' = (w_1, w_2, \dots, w_p)$. Finally we have $\text{Mat}_{\mathcal{C}',\mathcal{B}'}(f) = J_r$. Consequently if $P = T_{\mathcal{C}'\to\mathcal{C}} \in \mathcal{G}\ell_p(\mathbb{K})$ and $Q = T_{\mathcal{B}\to\mathcal{B}'} \in \mathcal{G}\ell_n(\mathbb{K})$ then

$$PMQ = T_{\mathcal{C}'\to\mathcal{C}} \times \text{Mat}_{\mathcal{C},\mathcal{B}}(f) \times T_{\mathcal{B}\to\mathcal{B}'} = \text{Mat}_{\mathcal{C}',\mathcal{B}'}(f) = J_r$$

Sufficient. Denote by $g : \mathbb{K}^p \to \mathbb{K}^p$ and $h : \mathbb{K}^n \to \mathbb{K}^n$ the linear maps such that $P = \text{Mat}_{\mathcal{C},\mathcal{C}}(g)$ and $Q = \text{Mat}_{\mathcal{B},\mathcal{B}}(h)$. Then

$$\begin{aligned} r \;=\; \text{rank}\Big(J_r\Big) &= \text{rank}\Big(PMQ\Big)\\ &= \text{rank}\Big(\text{Mat}_{\mathcal{C},\mathcal{C}}(g) \times \text{Mat}_{\mathcal{C},\mathcal{B}}(f) \times \text{Mat}_{\mathcal{B},\mathcal{B}}(h)\Big)\\ &= \text{rank}\Big(\text{Mat}_{\mathcal{C},\mathcal{B}}(g \circ f \circ h)\Big)\\ &= \text{rank}\Big(g \circ f \circ h\Big)\\ &= \dim\Big(\text{Im}(g \circ f \circ h)\Big)\end{aligned}$$

Moreover g and h are vector space isomorphisms since their associated matrices P and Q have multiplicative inverse. It follows that

$$\begin{aligned} r \;=\; \mathrm{rank}\Big(J_r\Big) &= \dim\Big(\mathrm{Im}(g\circ f\circ h)\Big) \\ &= \dim\Big(g(f(h(\mathbb{K}^n)))\Big) \\ &= \dim\Big(f(\mathbb{K}^n)\Big) \\ &= \dim\Big(\mathrm{Im}(f)\Big) \\ &= \mathrm{rank}\Big(f\Big) \\ &= \mathrm{rank}\Big(\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(f)\Big) \;=\; \mathrm{rank}(M) \end{aligned}$$

■

Corollary 5.21

Fix a positive integer $n \in \mathbb{N}^*$ and let M be a matrix in $\mathcal{M}_n(\mathbb{K})$. Then $M \in \mathcal{G}\ell_n(\mathbb{K})$ if and only if $\mathrm{rank}(M) = n$.

Remark : Furthermore, every matrix in $\mathcal{G}\ell_n(\mathbb{K})$ may be considered as a transition matrix from the canonical basis of $\mathbb{K}^n$ to the basis formed with its column vectors (by identifying $\mathbb{K}^n$ and $\mathcal{M}_{n,1}(\mathbb{K})$).

Corollary 5.22

Fix two positive integers $n \in \mathbb{N}^*$ and $p \in \mathbb{N}^*$ and let M be a matrix in $\mathcal{M}_{p,n}(\mathbb{K})$. Then

$$\mathrm{rank}({}^tM) = \mathrm{rank}(M)$$

Remark : In particular, the rank of a matrix is the maximal number of its linearly independent columns or rows.