8

Espaces vectoriels réels ou complexes

On notera $\mathbb{K}$ le corps de réels ou des complexes, en précisant quand cela sera nécessaire s'il s'agit de $\mathbb{R}$ ou $\mathbb{C}$. Par scalaire on entend réel ou complexe.

8.1 L'espace vectoriel $\mathbb{K}^n$

On peut utiliser les nombres réels $x \in \mathbb{R}$ pour représenter tous les points d'une droite, les couples de réels $(x, y) \in \mathbb{R}^2$ pour représenter tous les points d'un plan et les triplets de réels $(x, y, z) \in \mathbb{R}^3$ pour représenter tous les points d'un espace.

De manière plus générale, étant donné un entier naturel non nul n, on appelle vecteur tout élément du produit cartésien $\mathbb{K}^n = \mathbb{K} \times \mathbb{K} \times \cdots \times \mathbb{K}$ (répété n fois). Ces vecteurs sont des listes ordonnées de n scalaires $x_1, x_2, \cdots, x_n$ et seront notés $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ ou plus simplement $x = (x_i)_{1 \leq i \leq n}$ et on dit que les x_i sont les composantes du vecteur x.

La représentation sous forme de vecteurs colonnes sera justifiée plus loin par l'utilisation du calcul matriciel.

On peut naturellement munir cet ensemble d'une opération interne d'addition notée $+$ et définie par :

$$\forall x = (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n,\ \forall y = (y_i)_{1 \leq i \leq n} \in \mathbb{K}^n,\ x + y = (x_i + y_i)_{1 \leq i \leq n}$$

On dit que cette opération est interne car elle associe à deux éléments x et y de $\mathbb{K}^n$ un élément $x + y$ de $\mathbb{K}^n$.

Des propriétés de l'addition des scalaires, on déduit facilement que cette opération d'addition vérifie les propriétés suivantes :

- (i) elle est commutative, ce qui signifie que pour tous vecteurs x et y, on a $x + y = y + x$;
- (ii) elle est associative, ce qui signifie que pour tous vecteurs x, y et z, on a $x + (y + z) = (x + y) + z$;
- (iii) le vecteur nul $0 = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ est un élément neutre pour cette addition, ce qui signifie que pour tout vecteur x on a $x + 0 = 0 + x = x$;

(iv) pour tout vecteur $x = (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n$, le vecteur $x' = (-x_i)_{1 \leq i \leq n}$ est tel que $x + x' = x' + x = 0$, on dit que x' est un opposé de x et on le note $-x$.

Tout cela se résume en disant que l'ensemble $\mathbb{K}^n$ muni de l'addition, que l'on note $(\mathbb{K}^n, +)$, est un groupe commutatif.

De même, on peut naturellement munir $\mathbb{K}^n$ d'une multiplication externe définie par :

$$\forall \lambda \in \mathbb{K},\ \forall x = (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n,\ \lambda \cdot y = (\lambda \cdot x_i)_{1 \leq i \leq n}$$

On dit que cette opération est externe car elle associe à un scalaire λ (en dehors de $\mathbb{K}^n$) et à un élément x de $\mathbb{K}^n$ un élément $\lambda \cdot x$ de $\mathbb{K}^n$.

On écrira plus simplement λx pour $\lambda \cdot x$.

Là encore des propriétés de l'addition et de la multiplication des scalaires, on déduit que cette opération externe vérifie les propriétés suivantes :

(v) pour tout scalaire λ et tous vecteurs x et y, on a $\lambda (x + y) = \lambda x + \lambda y$;

(vi) pour tous scalaires λ, μ et tout vecteur x, on a $(\lambda + \mu) x = \lambda x + \mu x$;

(vii) pour tous scalaires λ, μ et tout vecteur x, on a $\lambda (\mu x) = (\lambda \mu) x$;

$(viii)$ pour tout vecteur x, on a $1 \cdot x = x$.

Tout cela se résume en disant que l'ensemble $\mathbb{K}^n$ muni de cette addition interne et de cette multiplication externe, que l'on note $(\mathbb{K}^n, +, \cdot)$, est un $\mathbb{K}$-espace vectoriel, ou simplement un espace vectoriel, le corps $\mathbb{K}$ étant sous-entendu.

8.2 Définition d'un espace vectoriel réel ou complexe

De manière plus générale, on donne la définition suivante.

Définition 8.1 *On appelle $\mathbb{K}$-espace vectoriel tout ensemble non vide E muni d'une addition interne $(x, y) \in E \times E \mapsto x + y \in E$ et d'une multiplication externe $(\lambda, x) \in \mathbb{K} \times E \mapsto \lambda x \in E$ vérifiant les propriétés (i) à $(viii)$ précédentes.*

Pour simplifier, on dira espace vectoriel pour $\mathbb{K}$-espace vectoriel. Quand cela sera nécessaire, on précisera espace vectoriel réel (i. e. pour $\mathbb{K} = \mathbb{R}$) ou espace vectoriel complexe (i. e. pour $\mathbb{K} = \mathbb{C}$).

Les éléments d'un espace vectoriel sont appelés vecteurs.

Dans un espace vectoriel l'élément neutre pour l'addition est noté 0 et on dit que c'est le vecteur nul et le symétrique d'un vecteur x est noté $-x$ et on dit que c'est l'opposé de x.

On vérifie facilement que le neutre 0 est unique, c'est-à-dire que c'est l'unique élément e de E tel que $x + e = e + x = x$ pour tout $x \in E$ (on a $e = e + 0$ puisque 0 est neutre et $e + 0 = 0$ puisque e est neutre, donc $e = 0$), que pour tout $x \in E$ l'opposé $-x$ est unique (si x' est un autre opposé, de $x + x' = 0$, on déduit que $(-x) + (x + x') = x' = (-x) + 0 = -x$) et que tout élément de E est simplifiable, c'est-à-dire que pour tous x, y, z dans E l'égalité $x + y = x + z$ équivaut à $y = z$ (il suffit d'ajouter $-x$ aux deux membres de cette égalité).

Pour x, y dans E, la somme $x + (-y)$ est simplement notée $x - y$.

Exemple 8.1 *L'ensemble $\mathbb{C}$ des nombres complexes est un espace vectoriel réel et aussi un espace vectoriel complexe.*

Exemple 8.2 *On rappelle qu'une suite réelle est une application u définie sur $\mathbb{N}$ et à valeurs réelles. L'ensemble de toutes les suites réelles est un espace vectoriel réel.*

Exemple 8.3 *Plus généralement, étant donné une partie I non vide de $\mathbb{R}$, l'ensemble E de toutes les applications de I dans $\mathbb{R}$ (resp. dans $\mathbb{C}$) est un espace vectoriel réel (resp. complexe).*

Exemple 8.4 *L'ensemble noté $\mathbb{R}[x]$ des fonctions polynomiales réelles (on dira plus simplement polynômes réel), c'est-à-dire l'ensemble des fonctions P définies par $P(x) = \sum_{k=0}^{n} a_k x^k$ pour tout réel x, où les coefficients a_k sont réels, est un espace vectoriel réel. Même chose sur $\mathbb{C}$.*

Exemple 8.5 *L'ensemble $\mathcal{F}$ des polynômes trigonométriques, c'est-à-dire l'ensemble des fonctions P définies par $P(x) = \sum_{k=0}^{n} (a_k \cos(kx) + b_k \sin(kx))$ pour tout réel x, où les coefficients a_k et b_k sont réels, est un espace vectoriel réel.*

Exercice 8.1 *Montrer que dans un espace vectoriel E, l'égalité $\lambda x = 0$ où λ est un scalaire et x un vecteur est équivalente à $\lambda = 0$ ou $x = 0$.*

Solution 8.1 *Pour tout vecteur x, on a :*

$$0 \cdot x + x = 0 \cdot x + 1 \cdot x = (0+1)\,x = 1 \cdot x = x$$

et simplifiant par x (ce qui revient à ajouter $-x$ aux deux membres de cette égalité), on aboutit à $0 \cdot x = 0$.
De même, pour tout scalaire λ, on a :

$$\lambda \cdot 0 = \lambda \cdot (0+0) = \lambda \cdot 0 + \lambda \cdot 0$$

et simplifiant par $\lambda \cdot 0$, on aboutit à $\lambda \cdot 0 = 0$.
Supposons que $\lambda x = 0$. Si $\lambda = 0$ c'est terminé, sinon λ est inversible dans $\mathbb{K}$ et :

$$x = 1 \cdot x = \left(\frac{1}{\lambda}\lambda\right) x = \frac{1}{\lambda}(\lambda x) = \frac{1}{\lambda} 0 = 0.$$

Exercice 8.2 *Montrer que dans un espace vectoriel E, on a $(-1)\,x = -x$ pour tout vecteur x.*

Solution 8.2 *Pour tout vecteur x, on a :*

$$x + (-1)\,x = 1 \cdot x + (-1)\,x = (1-1)\,x = 0 \cdot x = 0$$

et $(-1)\,x = -x$ puisque l'opposé de x est unique.

Définition 8.2 *Soient E un espace vectoriel, n un entier naturel non nul et $x, y, x_1, \cdots, x_n$ des éléments de E.*
On dit que y est colinéaire à x s'il existe un scalaire λ tel que $y = \lambda x$.
Plus généralement, on dit que y est combinaison linéaire de $x_1, \cdots, x_n$ s'il existe des scalaires $\lambda_1, \cdots, \lambda_n$ tels que $y = \sum_{k=1}^{n} \lambda_k x_k$.

On peut remarquer que, par définition, un espace vectoriel est stable par combinaison linéaire, c'est-à-dire que si $x_1, \cdots, x_n$ sont dans E et $\lambda_1, \cdots, \lambda_n$ dans $\mathbb{K}$, alors la combinaison linéaire $\sum_{k=1}^{n} \lambda_k x_k$ est encore dans E.

En s'inspirant de la construction l'espace vectoriel $\mathbb{K}^n$ comme produit cartésien de p exemplaires de l'espace vectoriel $\mathbb{K}$, on vérifie facilement que le produit cartésien $F = F_1 \times \cdots \times F_p$ de p espaces vectoriels $F_1, \cdots, F_p$ est naturellement muni d'une structure d'espace vectoriel avec les lois définies par :

$$\begin{cases} x + y = (x_1 + y_1, \cdots, x_p + y_p) \\ \lambda x = (\lambda x_1, \cdots, \lambda x_p) \end{cases}$$

où $x = (x_1, \cdots, x_p)$, $y = (y_1, \cdots, y_p)$ sont deux éléments de F et λ un scalaire.

8.3 Sous-espaces vectoriels

Définition 8.3 *Soit E un espace vectoriel. On dit qu'une partie F de E est un sous-espace vectoriel de E si :*

1. *le vecteur 0 est dans F ;*
2. *pour tous vecteurs x, y dans F et tout scalaire λ, les vecteurs $x + y$ et λx sont dans F.*

L'appellation sous-espace vectoriel est justifiée par le résultat suivant.

Théorème 8.1 *Tout sous-espace vectoriel d'un espace vectoriel est un espace vectoriel.*

Démonstration. Soit F un sous-espace vectoriel d'un espace vectoriel E.
Comme F contient 0, il est non vide.
Le deuxième point de la définition nous dit que l'addition des vecteurs et la multiplication d'un vecteur par un scalaire restreintes à F y définissent bien respectivement une opération interne et externe.
L'addition des vecteurs qui est commutative sur E l'est en particulier sur F.
L'élément neutre 0 pour l'addition est bien dans F.
Tout vecteur $x \in F$ admet un opposé $-x \in E$ et en écrivant que $-x = (-1)\, x$, on voit que $-x$ est bien dans F.
Les propriétés (v) à $(viii)$ vérifiées dans E le sont en particulier dans F. ■
De manière équivalente, on peut dire qu'une partie F d'un espace vectoriel E est un sous-espace vectoriel si, et seulement si, F est non vide et pour tous vecteurs x, y dans F, tous scalaires λ, μ, le vecteur $\lambda x + \mu y$ est dans F.

Exercice 8.3 *Justifier l'affirmation précédente.*

Solution 8.3 *Laissée au lecteur.*

De manière plus générale, un sous-espace vectoriel d'un espace vectoriel est une partie non vide stable par combinaison linéaire.

Exercice 8.4 *Justifier l'affirmation précédente.*

Solution 8.4 *Laissée au lecteur.*

Exemple 8.6 *Si E un espace vectoriel, alors $\{0\}$ et E sont des sous-espaces vectoriels de E.*

Exemple 8.7 $\mathbb{R}$ *et l'ensemble des imaginaires purs sont des sous-espaces vectoriels réels de* $\mathbb{C}$.

Exercice 8.5 *Montrer que l'intersection de deux sous-espaces vectoriels d'un espace vectoriel E est un sous-espace vectoriel. Qu'en est-il de la réunion ?*

Solution 8.5 *Soient F, G deux sous-espaces vectoriels de E. L'intersection $H = F \cap G$ contient 0 puisqu'ils sont dans F et G et pour tous x, y dans H, λ, μ dans $\mathbb{K}$, le vecteur $\lambda x + \mu y$ est dans F et G, donc dans H. En définitive, H est un sous-espace vectoriel de E.*
En général la réunion de deux sous-espaces vectoriels de E n'est pas un sous-espace vectoriel. Par exemple dans $\mathbb{R}^2$, les ensembles F et G définis par $F = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \mid x \in \mathbb{R} \right\}$ (l'axe des x) et $G = \left\{ \begin{pmatrix} 0 \\ y \end{pmatrix} \mid x \in \mathbb{R} \right\}$ (l'axe des y) sont des sous-espaces vectoriels, mais pas $F \cup G$ puisque $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ sont dans cette réunion, mais pas leur somme $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Exercice 8.6 *Soient F, G deux sous-espaces vectoriels de E. Montrer que $F \cup G$ est un sous-espace vectoriel de E si et seulement si $F \subset G$ ou $G \subset F$.*

Solution 8.6 *Si $F \subset G$ [resp. $G \subset F$], on a alors $F \cup G = G$ [resp. $F \cup G = F$] et c'est un sous-espace vectoriel de E.*
Réciproquement supposons que $F \cup G$ soit un sous-espace vectoriel de E. Si $F \not\subseteq G$ et $G \not\subseteq F$, il existe alors $x \in F \setminus G$ et $y \in G \setminus F$, donc comme x et y sont dans $F \cup G$, il en est de même de $x + y$, mais $x + y \in F$ entraîne $y = (x + y) - x \in F$, ce qui n'est pas et $x + y \in G$ entraîne $x = (x + y) - y \in G$, ce qui n'est pas, il y a donc une impossibilité et on a nécessairement $F \subset G$ ou $G \subset F$.

Exercice 8.7 *On désigne par $\mathcal{F}$ l'espace vectoriel des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ et par $\mathcal{P}$ [resp. $\mathcal{I}$] le sous-ensemble de $\mathcal{F}$ formés de toutes les fonctions paires [resp impaires] de $\mathbb{R}$ dans $\mathbb{R}$.*

1. *Montrer que $\mathcal{P}$ et $\mathcal{I}$ sont des sous-espaces vectoriels de $\mathcal{F}$.*
2. *Calculer $\mathcal{P} \cap \mathcal{I}$.*
3. *Montrer que toute fonction $f \in \mathcal{F}$, s'écrit de manière unique comme somme d'une fonction paire et d'une fonction impaire. Ce résultat se traduit en disant que $\mathcal{F}$ est somme directe des sous-espaces $\mathcal{P}$ et $\mathcal{I}$ et on note $\mathcal{F} = \mathcal{P} \oplus \mathcal{I}$.*

Solution 8.7

1. *La fonction nulle est à la fois paire et impaire donc dans $\mathcal{P}$ et dans $\mathcal{I}$. Si f, g sont deux fonctions paires [resp impaires], il en est alors de même de $f + g$ et de λf pour tout réel λ. Les ensembles $\mathcal{P}$ et $\mathcal{I}$ sont donc bien des sous-espaces vectoriels de $\mathcal{F}$.*
2. *Dire qu'une fonction f est dans $\mathcal{P} \cap \mathcal{I}$ signifie qu'elle est à la fois paire et impaire et donc que pour tout réel x, on a :*
$$f(x) = f(-x) = -f(x)$$
ce qui revient à dire que $f(x) = 0$. On a donc $\mathcal{P} \cap \mathcal{I} = \{0\}$.
3. *Pour toute fonction $f \in \mathcal{F}$, la fonction g [resp. h] définie sur $\mathbb{R}$ par :*
$$g(x) = \frac{f(x) + f(-x)}{2} \text{ [resp. } h(x) = \frac{f(x) - f(-x)}{2}\text{]}$$
est paire [resp. impaire] et $f = g + h$. Si (g', h') est un autre couple dans $\mathcal{P} \times \mathcal{I}$ tel que $f = g' + h'$, la fonction $g - g' = h' - h$ est dans $\mathcal{P} \cap \mathcal{I}$ donc nulle et $g = g'$, $h = h'$. Une telle écriture est donc unique.
Par exemple si f est la fonction exp, *les fonctions g et h sont les fonctions hyperboliques* ch *et* sh *définies par* $\operatorname{ch}(x) = \dfrac{e^x + e^{-x}}{2}$ *et* $\operatorname{sh}(x) = \dfrac{e^x - e^{-x}}{2}$.

Définition 8.4 *Dans l'espace $\mathbb{K}^n$, où n est un entier naturel non nul, on appelle droite vectorielle tout sous-ensemble de la forme :*
$$D = \{\lambda a \mid \lambda \in \mathbb{K}\}$$
où a est un vecteur non nul donné.

On notera $D = \mathbb{K}a$ une telle droite.
On a donc, en notant $a = (a_k)_{1 \leq k \leq n}$, pour tout $x = (x_k)_{1 \leq k \leq n} \in \mathbb{K}^n$:
$$(x \in D) \Leftrightarrow (\exists \lambda \in \mathbb{K} \mid \forall k \in \{1, 2, \cdots, n\}, \ x_k = \lambda a_k)$$

Une telle représentation est appelée représentation paramétrique de la droite D.

Cette définition correspond bien à la notion de droite vectorielle du plan $\mathbb{R}^2$ ou de l'espace $\mathbb{R}^3$ étudiée en Lycée.

On vérifie facilement qu'une droite de $\mathbb{K}^n$ est un sous-espace vectoriel.

Pour $n = 1$, $D = \mathbb{K}$ est la seule droite vectorielle puisque, pour tout $a \neq 0$, tout scalaire x peut s'écrire $x = \frac{x}{a} a = \lambda a$, donc $\mathbb{K} \subset \mathbb{K}a \subset \mathbb{K}$ et $\mathbb{K} = \mathbb{K}a$.

Définition 8.5 *Dans l'espace $\mathbb{K}^n$, où $n \geq 2$, on appelle plan vectoriel tout sous-ensemble de la forme :*

$$P = \left\{ \lambda a + \mu b \mid (\lambda, \mu) \in \mathbb{K}^2 \right\}$$

où a et b sont deux vecteurs non colinéaires donnés.

On notera $P = \mathbb{K}a \oplus \mathbb{K}b$ un tel plan.

On a donc, en notant $a = (a_k)_{1 \leq k \leq n}$ et $b = (b_k)_{1 \leq k \leq n}$, pour tout $x = (x_k)_{1 \leq k \leq n} \in \mathbb{K}^n$:

$$(x \in P) \Leftrightarrow (\exists (\lambda, \mu) \in \mathbb{K} \mid \forall k \in \{1, 2, \cdots, n\}, \ x_k = \lambda a_k + \mu b_k)$$

Une telle représentation est appelée représentation paramétrique du plan P.

Là encore cette définition correspond bien à la notion de plan vectoriel de l'espace $\mathbb{R}^3$ étudiée en Lycée.

On vérifie facilement qu'un plan de $\mathbb{K}^n$ est un sous-espace vectoriel.

Une partie finie d'un espace vectoriel E distincte de $\{0\}$ n'est pas un sous-espace vectoriel, mais à partir d'une partie finie de vecteurs de E, on peut engendrer un sous-espace vectoriel en s'inspirant des définitions de droites et plans.

Théorème 8.2 *Soient E un espace vectoriel et $x_1, \cdots, x_n$ des éléments de E. L'ensemble F de toutes les combinaisons linéaires de $x_1, \cdots, x_n$ est un sous-espace vectoriel de E.*

Démonstration. L'ensemble F contient $0 = \sum_{k=1}^{n} 0 \cdot x_k$ et pour $x = \sum_{k=1}^{n} \lambda_k x_k$, $y = \sum_{k=1}^{n} \mu_k x_k$ dans F et λ, μ dans $\mathbb{K}$, on a :

$$\lambda x + \mu y = \sum_{k=1}^{n} (\lambda \lambda_k + \mu \mu_k) x_k \in F$$

donc F est bien un sous-espace vectoriel de E. ■

Définition 8.6 *Avec les notations du théorème précédent, on dit que F est le sous-espace vectoriel de E engendré par $x_1, \cdots, x_n$ et on le note $F = \langle x_1, \cdots, x_n \rangle$, ou $F = \mathrm{Vect}\{x_1, \cdots, x_n\}$ ou encore $F = \sum_{k=1}^{n} \mathbb{K} x_k$.*

Remarque 8.1 *Si tous les x_k sont nuls, alors $F = \{0\}$.*

Exemple 8.8 *Dans l'espace $\mathbb{K}[x]$ des fonctions polynomiales, pour tout entier naturel non nul n, le sous-espace vectoriel engendré par $1, x, \cdots, x^n$ est formé de l'ensemble des polynômes de degré au plus égal à n, on le note $\mathbb{K}_n[x]$ ou $\mathbb{K}[x]_n$ (le cas $n = 0$ correspond aux polynômes constants).*

De manière un peu plus générale, on peut définir le sous-espace vectoriel d'un espace vectoriel E engendré par une famille X non vide de E (non nécessairement finie) comme l'ensemble $F = \mathrm{Vect}\,(X)$ (ou $F = \langle X \rangle$) de toutes les combinaisons linéaires d'éléments de X. Un vecteur x de E est donc dans $\mathrm{Vect}\,(X)$ si, et seulement si, il existe un entier $p \geq 1$, des vecteurs $x_1, \cdots, x_p$ dans X et des scalaires $\lambda_1, \cdots, \lambda_p$ tels que $x = \sum\limits_{k=1}^{p} \lambda_k x_k$.

Le théorème qui suit nous donne deux définitions équivalentes de $\mathrm{Vect}\,(X)$.

Théorème 8.3 *Si X est une partie non vide d'un espace vectoriel E,* $\mathrm{Vect}\,(X)$ *est l'intersection de tous les sous-espaces vectoriels de E qui contiennent X. C'est aussi le plus petit (pour l'ordre défini par l'inclusion) sous-espace vectoriel de E qui contient X, c'est-à-dire que* $\mathrm{Vect}\,(X)$ *contient X et est contenu dans tout sous-espace vectoriel de E qui contient X.*

On peut aussi définir des sous-espaces vectoriels de $\mathbb{K}^n$ en utilisant des équations linéaires (on verra plus loin, avec la notion de base, que cela est encore possible pour n'importe quel espace vectoriel).

Théorème 8.4 *Étant donnés un entier naturel non nul n et n scalaires non tous nuls $a_1, \cdots, a_n$, l'ensemble :*

$$F = \left\{ x = (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n \mid \sum_{k=1}^{n} a_k x_k = 0 \right\}$$

est un sous-espace vectoriel de $\mathbb{K}^n$.

Démonstration. Il suffit de vérifier. ■

Remarque 8.2 *En fait si tous les a_k sont nuls, F est encore défini et c'est $\mathbb{K}^n$ tout entier.*

Pour $n = 1$, on a $a_1 \neq 0$ et cet espace F est réduit à $\{0\}$.

Pour $n = 2$, on a $(a_1, a_2) \neq (0, 0)$ et supposant par exemple que $a_2 \neq 0$, l'équation $a_1 x_1 + a_2 x_2 = 0$ équivaut à $x_2 = -\dfrac{a_1}{a_2} x_1$, ce qui signifie que F est l'ensemble des vecteurs de la forme :

$$x = x_1 \begin{pmatrix} 1 \\ -\frac{a_1}{a_2} \end{pmatrix}$$

où x_1 décrit $\mathbb{K}$, ce qui équivaut encore à dire que F est l'ensemble des vecteurs de la forme :

$$x = \frac{x_1}{a_2} \begin{pmatrix} a_2 \\ -a_1 \end{pmatrix} = \lambda \begin{pmatrix} a_2 \\ -a_1 \end{pmatrix}$$

où λ décrit $\mathbb{K}$. En définitive F est la droite engendrée par $\begin{pmatrix} a_2 \\ -a_1 \end{pmatrix}$.

Pour $n = 3$, on a $(a_1, a_2, a_3) \neq (0, 0, 0)$ et supposant par exemple que $a_3 \neq 0$, l'équation $a_1 x_1 + a_2 x_2 + a_3 x_3 = 0$ équivaut à $x_2 = -\dfrac{a_1}{a_3} x_1 - \dfrac{a_2}{a_3} x_2$, ce qui signifie que F est l'ensemble des vecteurs de la forme :

$$x = x_1 \begin{pmatrix} 1 \\ 0 \\ -\frac{a_1}{a_3} \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ -\frac{a_2}{a_3} \end{pmatrix}$$

où (x_1, x_2) décrit $\mathbb{K}^2$, ce qui équivaut encore à dire que F est l'ensemble des vecteurs de la forme :

$$x = \frac{x_1}{a_3}\begin{pmatrix} a_3 \\ 0 \\ -a_1 \end{pmatrix} + \frac{x_2}{a_3}\begin{pmatrix} 0 \\ a_3 \\ -a_2 \end{pmatrix} = \lambda \begin{pmatrix} a_3 \\ 0 \\ -a_1 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ a_3 \\ -a_3 \end{pmatrix}$$

où (λ, μ) décrit $\mathbb{K}^2$, les vecteurs $\begin{pmatrix} a_3 \\ 0 \\ -a_1 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ a_3 \\ -a_3 \end{pmatrix}$ étant non colinéaires puisque $a_3 \neq 0$.

En définitive F est le plan engendré par $\begin{pmatrix} a_3 \\ 0 \\ -a_1 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ a_3 \\ -a_3 \end{pmatrix}$.

De manière générale, on donne la définition suivante.

Définition 8.7 *Étant donné un entier naturel non nul n, on appelle hyperplan de $\mathbb{K}^n$ tout sous-espace vectoriel de la forme :*

$$H = \left\{ x = (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n \mid \sum_{k=1}^{n} a_k x_k = 0 \right\}$$

où les scalaires $a_1, \cdots, a_n$ ne sont pas tous nuls.

On dit aussi que H est l'hyperplan d'équation $\sum\limits_{k=1}^{n} a_k x_k = 0$.

Dans l'espace $\mathbb{R}^3$ l'intersection de deux plans vectoriels distincts est une droite.

Plus généralement, on a le résultat suivant.

Théorème 8.5 *Étant donnés deux entiers naturels non nuls n et p, des scalaires $a_{1,1}, \cdots, a_{1,n}$, $\cdots$, $a_{p,1}, \cdots, a_{p,n}$, l'ensemble :*

$$F = \left\{ x = (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n \mid \sum_{k=1}^{n} a_{i,k} x_k = 0 \ \ (1 \leq i \leq p) \right\}$$

est un sous-espace vectoriel de $\mathbb{K}^n$.

On dit que F est le sous-espace vectoriels de $\mathbb{K}^n$ d'équations linéaires :

$$\left\{ \begin{array}{c} \sum\limits_{k=1}^{n} a_{1,k} x_k = 0 \\ \vdots \\ \sum\limits_{k=1}^{n} a_{p,k} x_k = 0 \end{array} \right.$$

8.4 Applications linéaires

Pour ce paragraphe, on désigne par E et F deux espaces vectoriels. On note 0 le vecteur nul de E et aussi celui de F. En toute rigueur il faudrait noter 0_E et 0_F ces vecteurs nuls, mais en fonction du contexte on sait en général de quel vecteur nul il s'agit.

Définition 8.8 *On dit qu'une application u de E dans F est linéaire (ou que c'est un morphisme d'espaces vectoriels) si pour tous vecteurs x, y de E et tout scalaire λ, on a :*

$$\begin{cases} u(x+y) = u(x) + u(y) \\ u(\lambda x) = \lambda u(x) \end{cases}$$

Remarque 8.3 *Une application linéaire u de E dans F transforme 0_E en 0_F et l'opposé de x dans E en l'opposé de $u(x)$ dans F. En effet, on a :*

$$u(0) = u(0+0) = u(0) + u(0)$$

donc $u(0) = 0$ et :

$$0 = u(0) = u(x + (-x)) = u(x) + u(-x)$$

et donc $u(-x) = -u(x)$.

Exemple 8.9 *Pour tout scalaire λ, l'application $u : x \mapsto \lambda x$ est linéaire de E dans E. On dit que c'est l'homothétie de rapport λ. Pour $\lambda = 1$, cette application est l'application identité et on la note Id_E ou Id.*

Exemple 8.10 *Pour tout entier j compris entre 1 et n, l'application u définie sur $\mathbb{K}^n$ par $u(x) = x_j$, si $x = (x_i)_{1 \leq i \leq n}$ est linéaire de E dans $\mathbb{K}$. On dit que c'est la j-ième projection canonique.*

Exemple 8.11 *Une translation de vecteur non nul $x \mapsto x + a$, définie de E dans E, n'est pas linéaire.*

Exemple 8.12 *Étant donné un intervalle réel I, la dérivation $f \mapsto f'$ est linéaire de l'espace vectoriel des fonctions dérivables de I dans $\mathbb{R}$ dans l'espace vectoriel des fonctions de I dans $\mathbb{R}$.*

On notera $\mathcal{L}(E, F)$ (ou plus précisément $\mathcal{L}_{\mathbb{K}}(E, F)$) l'ensemble de toutes les applications linéaires de E dans F.

Si u et v sont deux applications linéaires de E dans F, $u + v$ est l'application définie sur E par :

$$\forall x \in E,\ (u+v)(x) = u(x) + v(x)$$

et pour tout scalaire λ, λu est l'application définie sur E par :

$$\forall x \in E,\ (\lambda u)(x) = \lambda u(x).$$

Il est facile de vérifier que $u + v$ et λu sont aussi des application linéaires de E dans F. On a donc ainsi défini une addition interne sur $\mathcal{L}(E, F)$ et une multiplication externe. Le résultat qui suit se démontre alors facilement.

Théorème 8.6 *L'ensemble $\mathcal{L}(E, F)$ de toutes les applications linéaires de E dans F muni de ces deux opérations est un espace vectoriel.*

Dans le cas où $F = E$, on note plus simplement $\mathcal{L}(E)$ pour $\mathcal{L}(E, F)$.

Les éléments de $\mathcal{L}(E)$ sont aussi appelés endomorphismes de E.

La composition des applications permet aussi de construire des applications linéaires à partir d'applications linéaires données.

Théorème 8.7 *Soient E, F, G des espaces vectoriels, u une application linéaire de E dans F et v une application linéaire de F dans G. La composée $v \circ u$ est alors une application linéaire de E dans G.*

Démonstration. Il suffit de vérifier. ■

Théorème 8.8 *Si u est une application linéaire de E dans F, on a alors pour tout entier naturel non nul n, tous vecteurs $x_1, \cdots, x_n$ de E et tous scalaires $\lambda_1, \cdots, \lambda_n$:*

$$u\left(\sum_{k=1}^{n} \lambda_k x_k\right) = \sum_{k=1}^{n} \lambda_k u(x_k).$$

Démonstration. On procède par récurrence pour $n \geq 1$. Pour $n = 1$, on a bien $u(\lambda x) = \lambda u(x)$ pour tout scalaire λ et tout vecteur x par définition d'une application linéaire.
Pour $n = 2$, toujours par définition d'une application linéaire, on a pour tous vecteurs x, y et tous scalaires λ, μ :

$$u(\lambda x + \mu y) = u(\lambda x) + u(\mu y) = \lambda u(x) + \mu u(y).$$

Supposant le résultat acquis au rang $n \geq 2$, on se donne $n+1$ vecteurs $x_1, \cdots, x_n, x_{n+1}$ et $n+1$ scalaires $\lambda_1, \cdots, \lambda_n, \lambda_{n+1}$ et on a :

$$\begin{aligned} u\left(\sum_{k=1}^{n+1} \lambda_k x_k\right) &= u\left(\sum_{k=1}^{n} \lambda_k x_k\right) + u(\lambda_{n+1} x_{n+1}) \\ &= \sum_{k=1}^{n} \lambda_k u(x_k) + \lambda_{n+1} u(x_{n+1}) \\ &= \sum_{k=1}^{n+1} \lambda_k u(x_k) \end{aligned}$$

■

Définition 8.9 *Soit u une application linéaire de E dans F.*
Le noyau de u est l'ensemble :

$$\ker(u) = \{x \in E \mid u(x) = 0\}$$

et l'image de u l'ensemble :

$$\operatorname{Im}(u) = \{u(x) \mid x \in E\}.$$

Théorème 8.9 *Le noyau d'une application linéaire u de E dans F est un sous-espace vectoriel de E et son image un sous-espace vectoriel de F.*

Démonstration. On a vu que $\ker(u)$ contient 0 et pour x, y dans $\ker(u)$, λ, μ dans $\mathbb{K}$, on a :

$$u(\lambda x + \mu y) = \lambda u(x) + \mu u(y) = 0$$

ce qui signifie que $\lambda x + \mu y \in \ker(u)$. Donc $\ker(u)$ est bien un sous-espace vectoriel de E.

De manière analogue, en utilisant la linéarité de u, on montre que $\operatorname{Im}(u)$ est un sous-espace vectoriel de F. ■

Théorème 8.10 *Soit u une application linéaire de E dans F.*

1. *L'application u est injective si, et seulement si,* $\ker(u)$ *est réduit à* $\{0\}$.
2. *L'application u est surjective si, et seulement si,* $\mathrm{Im}(u) = F$.

Démonstration.

1. Supposons u injective. Pour tout $x \in \ker(u)$, on a $u(x) = u(0)$ et nécessairement $x = 0$ puisque u est injective. Donc $\ker(u) = \{0\}$.
   Réciproquement, supposons que $\ker(u) = \{0\}$. Si x, y dans E sont tels que $u(x) = u(y)$, on a alors $u(x-y) = u(x) - u(y) = 0$, c'est-à-dire que $x - y \in \ker(u)$ et donc $x = y$.
2. Ce résultat est en fait valable pour toute application de E dans F (la linéarité de u n'intervient pas ici).

■

Les applications linéaires de E dans $\mathbb{K}$ ont un statut particulier.

Définition 8.10 *On appelle forme linéaire sur E toute application linéaire de E dans $\mathbb{K}$.*

Exemple 8.13 *Étant donné un segment $I = [a, b]$ non réduit à un point, l'application $f \mapsto \int_a^b f(x)\,dx$ est une forme linéaire sur l'espace vectoriel des fonctions continues de I dans $\mathbb{R}$.*

Exercice 8.8 *Montrer qu'une forme linéaire φ sur E non identiquement nulle est surjective.*

Solution 8.8 *Dire que $\varphi \neq 0$ signifie qu'il existe un vecteur $x_0 \in E$ tel que $\lambda = \varphi(x_0) \neq 0$. Pour tout scalaire y, on peut alors écrire :*

$$y = \frac{y}{\lambda}\lambda = \frac{y}{\lambda}\varphi(x_0) = \varphi\left(\frac{y}{\lambda}x_0\right)$$

soit $y = \varphi(x)$ avec $x = \frac{y}{\lambda}x_0 \in E$, ce qui signifie que φ est surjective.

Exercice 8.9 *On se donne un intervalle réel I non réduit à un point et on désigne par E l'ensemble de toutes les fonctions dérivables de I dans $\mathbb{R}$ et par F l'ensemble de toutes les fonctions de I dans $\mathbb{R}$.*

1. *Montrer que E est un espace vectoriel.*
2. *Déterminer le noyau de l'application linéaire $u : f \mapsto f'$ où f' est la fonction dérivée de f.*

Solution 8.9 *Laissée au lecteur.*

Exercice 8.10 *Soit u une application linéaire de E dans E (i. e. un endomorphisme de E). Montrer que :*

$$\mathrm{Im}(u) \subset \ker(u) \Leftrightarrow u \circ u = 0.$$

Solution 8.10 *Si* $\mathrm{Im}(u) \subset \ker(u)$, *on a alors $u(x) \in \ker(u)$ pour tout $x \in E$ et $u(u(x)) = 0$, ce qui signifie que $u \circ u = 0$.*
Réciproquement si $u \circ u = 0$, pour tout $y = u(x) \in \mathrm{Im}(u)$, on a $u(y) = u(u(x)) = u \circ u(x) = 0$, ce qui signifie que $y \in \ker(u)$ et $\mathrm{Im}(u) \subset \ker(u)$.

Exercice 8.11 *On appelle projecteur de E tout endomorphisme p de E tel que $p \circ p = p$.*

1. *Montrer que* $\mathrm{Im}\,(p)$ *est l'ensemble des vecteurs invariants de* p.
2. *Montrer que* $\mathrm{Im}\,(p) \cap \ker(p) = \{0\}$.
3. *Montrer que tout vecteur* $x \in E$ *s'écrit de manière unique comme somme d'un vecteur de* $\ker(p)$ *et d'un vecteur de* $\mathrm{Im}\,(p)$ *(on dit que* E *est somme directe de* $\ker(p)$ *et* $\mathrm{Im}\,(p)$, *ce qui se note* $E = \ker(p) \oplus \mathrm{Im}\,(p)$*).*
4. *Montrer que si* p *est un projecteur, il en est alors de même de* $q = Id_E - p$ *et on a* $\ker(q) = \mathrm{Im}\,(p)$ *et* $\mathrm{Im}\,(q) = \ker(p)$.

Solution 8.11 *Laissée au lecteur.*

On rappelle que si u est une bijection de E sur F, elle admet alors une application réciproque notée u^{-1} et définie par :

$$\left(y \in F \text{ et } x = u^{-1}(y)\right) \Leftrightarrow (x \in E \text{ et } y = u(x)).$$

Cette application u^{-1} est aussi l'unique application de F dans E telle que $u \circ u^{-1} = Id_F$ et $u^{-1} \circ u = Id_E$.

Dans le cas où u est linéaire, il en est de même de u^{-1}. En effet si y, y' sont deux éléments de F, ils s'écrivent de manière unique $y = u(x)$, $y' = u(x')$ et on a :

$$\begin{aligned} u^{-1}(y + y') &= u^{-1}(u(x) + u(x')) \\ &= u^{-1}(u(x + x')) = x + x' = u^{-1}(y) + u^{-1}(y') \end{aligned}$$

et pour tout scalaire λ :

$$u^{-1}(\lambda y) = u^{-1}(\lambda u(x)) = u^{-1}(u(\lambda x)) = \lambda x = \lambda u^{-1}(y).$$

Définition 8.11 *On appelle isomorphisme de* E *sur* F *toute application linéaire bijective de* E *sur* F.

Dans le cas où $E = F$, un isomorphisme de E sur E est appelé automorphisme de E.

On note $GL(E)$ l'ensemble de tous les automorphismes de E et on dit que $GL(E)$ est le groupe linéaire de E (l'appellation groupe sera justifiée plus loin).

Exercice 8.12 *L'ensemble* $GL(E)$ *est-il un espace vectoriel ?*

Solution 8.12 *Non, sauf dans le cas où* $E = \{0\}$.

8.5 La base canonique de $\mathbb{K}^n$ et expression matricielle des applications linéaires de $\mathbb{K}^n$ dans $\mathbb{K}^m$

Tout vecteur $x = (x_i)_{1 \le i \le n}$ de $\mathbb{K}^n$ s'écrit $\sum_{k=1}^{n} x_k e_k$, où on a noté, pour tout entier k compris entre 1 et n, e_k le vecteur dont toutes les composantes sont nulles sauf la k-ième qui vaut 1. L'espace vectoriel $\mathbb{K}^n$ est donc engendré par les vecteurs $e_1, e_2, \cdots, e_n$. On dit que le système $(e_1, e_2, \cdots, e_n)$, que l'on note plus simplement $(e_k)_{1 \le k \le n}$, est un système générateur de $\mathbb{K}^n$. De plus par définition du produit cartésien $\mathbb{K}^n$ une telle écriture est unique, ce qui se traduit en disant que le système $(e_k)_{1 \le k \le n}$ est un système libre.

On dit que $(e_k)_{1\leq k\leq n}$ est la base canonique de $\mathbb{K}^n$. Nous définirons plus loin la notion de base d'un espace vectoriel.

Si m est un autre entier naturel non nul, on note $(f_k)_{1\leq k\leq m}$ la base canonique de $\mathbb{K}^m$.

Étant donnée une application linéaire u de E dans F, on a pour tout vecteur $x = \sum\limits_{j=1}^{n} x_j e_j$ de $\mathbb{K}^n$:

$$u(x) = u\left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{j=1}^{n} x_j u(e_j)$$

et chacun des vecteurs $u(e_j)$, pour j compris entre 1 et n, étant dans $\mathbb{K}^m$, il s'écrit :

$$u(e_j) = \sum_{i=1}^{m} a_{ij} f_i$$

On a donc en définitive :

$$u(x) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} f_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij} x_j\right) f_i$$

ce qui signifie que les composantes du vecteurs $u(x)$ sont données par :

$$y_i = \sum_{j=1}^{n} a_{ij} x_j \ \ (1 \leq i \leq m) \tag{8.1}$$

Par exemple pour $n = m = 2$, on a :

$$\begin{cases} u(e_1) = a_{11} f_1 + a_{21} f_2 \\ u(e_2) = a_{12} f_1 + a_{22} f_2 \end{cases}$$

et :

$$u(x) = y_1 f_1 + y_2 f_2$$

avec :

$$\begin{cases} y_1 = a_{11} x_1 + a_{12} x_2 \\ y_2 = a_{21} x_1 + a_{22} x_2 \end{cases} \tag{8.2}$$

L'application linéaire u est donc uniquement déterminée par les 4 scalaires a_{11}, a_{12}, a_{21} et a_{22}. On stocke ces scalaires dans un tableau à 2 lignes et 2 colonnes :

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

où la première colonne $\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ est le vecteur $u(e_1)$ et la deuxième colonne $\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$ le vecteur $u(e_2)$. Un tel tableau est appelé matrice à 2 lignes et 2 colonnes ou plus simplement matrice 2×2.

On traduit les deux égalités de (8.2) en utilisant le produit matriciel :

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

De manière générale, une application linéaire u de $\mathbb{K}^n$ dans $\mathbb{K}^m$ est donc uniquement déterminée par la matrice A à m lignes et n colonnes :

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

où la colonne numéro j, pour j compris entre 1 et n, est le vecteur :

$$u(e_j) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

Les égalités (8.1) se traduisent alors par le produit matriciel :

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$$

qui s'effectue comme suit :

$$y_i = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \sum_{j=1}^{n} a_{ij} x_j$$

pour tout i compris entre 1 et m.

Tout cela est compacté en :

$$y = u(x) \Leftrightarrow y = Ax$$

et on dit que A est la matrice de u dans les bases canoniques de $\mathbb{K}^n$ et $\mathbb{K}^m$. Si $n = m$, on dit simplement que A est la matrice de $u \in \mathcal{L}(\mathbb{K}^n)$ dans la base canonique de $\mathbb{K}^n$.

Exercice 8.13 *Soit $u \in \mathcal{L}(\mathbb{K}^3)$ de matrice $A = \begin{pmatrix} 4 & 2 & -4 \\ -6 & -4 & 6 \\ -1 & -1 & 1 \end{pmatrix}$ dans la base canonique (e_1, e_2, e_3). Déterminer le noyau u.*

Solution 8.13 *L'image du vecteur $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ est le vecteur :*

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 4 & 2 & -4 \\ -6 & -4 & 6 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4x_1 + 2x_2 - 4x_3 \\ -6x_1 - 4x_2 + 6x_3 \\ -x_1 - x_2 + x_3 \end{pmatrix}$$

Dire que $x \in \mathbb{K}^3$ est dans le noyau de u équivaut à dire que ses composantes sont solutions du système linéaire de 3 équations à 3 inconnues :

$$\left\{\begin{array}{l} 4x_1 + 2x_2 - 4x_3 = 0 \\ -6x_1 - 4x_2 + 6x_3 = 0 \\ -x_1 - x_2 + x_3 = 0 \end{array}\right.$$

L'équation (3) donne $x_3 = x_1 + x_2$ qui reporté dans la première donne $x_2 = 0$ et $x_1 = x_3$. Réciproquement tout vecteur vérifiant ces conditions est solution du système linéaire. Le noyau de u est donc :

$$\ker(u) = \left\{ \begin{pmatrix} x_1 \\ 0 \\ x_1 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = x_1 (e_1 + e_3) \mid x \in \mathbb{K} \right\}$$

c'est donc la droite vectorielle engendré par le vecteur $e_1 + e_3$.

8.6 Matrices réelles ou complexes

On note $\mathcal{M}_{m,n}(\mathbb{K})$ l'ensemble de toutes les matrices à m lignes et n colonnes et à coefficients dans $\mathbb{K}$.

Une matrice $\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$ dans $\mathcal{M}_{m,n}(\mathbb{K})$ sera notée $A = ((a_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$, le premier indice i étant le numéro de ligne et le deuxième j, le numéro de colonne.

Pour $n = m$, on note $\mathcal{M}_n(\mathbb{K})$ l'ensemble $\mathcal{M}_{n,n}(\mathbb{K})$ et dit que c'est l'ensemble des matrices carrées d'ordre n à coefficients dans $\mathbb{K}$.

8.6.1 Opérations sur les matrices

Théorème 8.11 *Si u et v sont deux applications linéaires de $\mathbb{K}^n$ dans $\mathbb{K}^m$ ayant pour matrices respectives $A = ((a_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ et $B = ((b_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ dans les bases canoniques, alors l'application linéaire $u + v$ a pour matrice dans ces bases canoniques, la matrice $((a_{i,j} + b_{ij}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$. Et pour tout scalaire λ, la matrice de λu dans les bases canoniques est la matrice $((\lambda a_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$.*

Démonstration. Résulte de :

$$\begin{aligned} (u+v)(e_j) &= u(e_j) + v(e_j) \\ &= \sum_{i=1}^{m} a_{ij} f_i + \sum_{i=1}^{m} b_{ij} f_i \\ &= \sum_{i=1}^{m} (a_{ij} + b_{ij}) f_i \ \ (1 \leq j \leq n) \end{aligned}$$

et de :

$$(\lambda u)(e_j) = \lambda u(e_j) = \lambda \sum_{i=1}^{m} a_{ij} f_i = \sum_{i=1}^{m} \lambda a_{ij} f_i \ \ (1 \leq j \leq n)$$

■

On définit donc naturellement la somme de deux matrices $n \times m$ et le produit d'une telle matrice par un scalaire par :

$$A + B = ((a_{i,j} + b_{ij}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \text{ et } \lambda A = ((\lambda a_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$$

en utilisant les notations précédentes.

On vérifie alors facilement le résultat suivant.

Théorème 8.12 *L'ensemble $\mathcal{M}_{m,n}(\mathbb{K})$ des matrices à m lignes et n colonnes est un espace vectoriel.*

On note 0 la matrice nulle, c'est-à-dire l'élément de $\mathcal{M}_{m,n}(\mathbb{K})$ dont toutes les composantes sont nulles et pour toute matrice $A = ((a_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$, on note $-A = ((-a_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ l'opposé de A.

En explicitant la matrice d'une composée de deux applications linéaires on définira le produit de deux matrices.

On se donne donc une application linéaire v de $\mathbb{K}^n$ dans $\mathbb{K}^m$ et une application linéaire u de $\mathbb{K}^m$ dans $\mathbb{K}^r$ de matrices respectives $B = ((b_{ij}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ dans $\mathcal{M}_{m,n}(\mathbb{K})$ et $A = ((a_{i,j}))_{\substack{1 \leq i \leq r \\ 1 \leq j \leq m}}$ dans $\mathcal{M}_{r,m}(\mathbb{K})$.

On note toujours $(e_k)_{1 \leq k \leq n}$ la base canonique de $\mathbb{K}^n$, $(f_k)_{1 \leq k \leq m}$ celle de $\mathbb{K}^m$ et $(g_k)_{1 \leq k \leq r}$ est celle de $\mathbb{K}^r$.

La matrice $C = ((c_{i,j}))_{\substack{1 \leq i \leq r \\ 1 \leq j \leq n}} \in \mathcal{M}_{r,n}(\mathbb{K})$ de $u \circ v \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^r)$ est obtenu en calculant les composantes des vecteurs $u \circ v(e_j)$, pour j compris entre 1 et n, dans la base $(g_k)_{1 \leq k \leq r}$.

On a :

$$u \circ v(e_j) = u(v(e_j)) = u\left(\sum_{k=1}^{m} b_{kj} f_k\right) = \sum_{k=1}^{m} b_{kj} u(f_k)$$

avec, pour k compris entre 1 et m :

$$u(f_k) = \sum_{i=1}^{r} a_{ik} g_i$$

ce qui donne :

$$u \circ v(e_j) = \sum_{k=1}^{m} b_{kj} \left(\sum_{i=1}^{r} a_{ik} g_i\right) = \sum_{i=1}^{r} \left(\sum_{k=1}^{m} a_{ik} b_{kj}\right) g_i$$

et signifie que les coefficients de la matrice C sont donnés par :

$$c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj} \ \ (1 \leq i \leq r, \ 1 \leq j \leq n).$$

Au vu de ce résultat, on donne la définition suivante.

Définition 8.12 *Étant données une matrice $A = ((a_{i,j}))_{\substack{1 \leq i \leq r \\ 1 \leq j \leq m}} \in \mathcal{M}_{r,m}(\mathbb{K})$ et une matrice $B = ((b_{ij}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathcal{M}_{m,n}(\mathbb{K})$, le produit AB de A par B est la matrice $C = ((c_{i,j}))_{\substack{1 \leq i \leq r \\ 1 \leq j \leq n}}$ de $\mathcal{M}_{r,n}(\mathbb{K})$ définie par :*

$$c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj} \ \ (1 \leq i \leq r, \ 1 \leq j \leq n).$$

Et nous venons de montrer le résultat suivant.

Théorème 8.13 *Si v est une application linéaire de $\mathbb{K}^n$ dans $\mathbb{K}^m$ de matrice $B = ((b_{ij}))_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$ dans $\mathcal{M}_{m,n}(\mathbb{K})$ et u une application linéaire de $\mathbb{K}^m$ dans $\mathbb{K}^r$ de matrice $A = ((a_{i,j}))_{\substack{1\leq i\leq r \\ 1\leq j\leq m}}$ dans $\mathcal{M}_{r,m}(\mathbb{K})$ dans les bases canoniques, alors la matrice dans les bases canoniques de l'application linéaire $u \circ v$ de $\mathbb{K}^n$ dans $\mathbb{K}^r$ est la matrice produit $C = AB$.*

Il est important de remarquer que l'on ne peut définir le produit AB que si le nombre de colonnes de la matrice A est égal au nombre de lignes de la matrice B, ce produit est donc défini de $\mathcal{M}_{r,m}(\mathbb{K}) \times \mathcal{M}_{m,n}(\mathbb{K})$ dans $\mathcal{M}_{r,n}(\mathbb{K})$.

Exercice 8.14 *Soient $A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 2 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & -1 \\ 0 & 2 \\ 2 & 3 \end{pmatrix}$. Calculer AB et BA.*

Solution 8.14 *On a :$AB = \begin{pmatrix} 7 & 12 \\ 1 & 8 \end{pmatrix}$ et $BA = \begin{pmatrix} 2 & 0 & 2 \\ -2 & 4 & 2 \\ -1 & 10 & 9 \end{pmatrix}$.*

Exercice 8.15 *Pour tout réel θ, on désigne par M_θ la matrice réelle :*

$$M_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

Montrer que pour tous réels θ et θ', on a $M_\theta M_{\theta'} = M_{\theta'} M_\theta = M_{\theta+\theta'}$.

Solution 8.15 *Laissée au lecteur.*

Exercice 8.16 *On se place dans l'espace $\mathcal{M}_2(\mathbb{K})$ des matrices carrées d'ordre 2 où on note $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ la matrice identité d'ordre 2.*

1. *Donner des exemples de matrices A et B telles que $AB = 0$ et $BA \neq 0$.*
2. *Montrer que si $AB = I_2$, alors $BA = I_2$.*

Solution 8.16 *Laissée au lecteur.*

Exercice 8.17 *Déterminer dans la base canonique (e_1, e_2) de $\mathbb{R}^2$ la matrice de l'endomorphisme u (s'il existe) tel que $u(e_1) = ae_1 - e_2$ et $u \circ u = u$ où a est un réel donné.*

Solution 8.17 *Laissée au lecteur.*

Exercice 8.18 *Donner une condition nécessaire et suffisante portant sur les réels a et b pour que l'endomorphisme u de $\mathbb{R}^2$ de matrice $A = \begin{pmatrix} a+1 & a \\ b & b+1 \end{pmatrix}$ soit un automorphisme.*

Solution 8.18 *Laissée au lecteur.*

Exercice 8.19 *Déterminer, par leur matrice dans la base canonique, tous les endomorphismes non nuls de $\mathbb{K}^2$ tels que $\mathrm{Im}(u) \subset \ker(u)$. Si u est un tel endomorphisme donner la matrice de $v = Id_E + u$ et montrer que c'est un automorphisme de E.*

Solution 8.19 *Laissée au lecteur.*

L'opération de multiplication des matrices est une opération interne sur l'espace $\mathcal{M}_n(\mathbb{K})$ des matrices carrées d'ordre n vérifiant les propriétés suivante :

– elle est associative, c'est-à-dire que pour toutes matrices A, B, C dans $\mathcal{M}_n(\mathbb{K})$, on a $A(BC) = (AB)C$;
– elle est distributive par rapport à l'addition, c'est-à-dire que pour toutes matrices A, B, C dans $\mathcal{M}_n(\mathbb{K})$, on a $A(B+C) = AB + AC$;
– la matrice

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

est l'élément neutre pour ce produit, c'est-à-dire que $A \cdot I_n = I_n \cdot A = A$ pour toute matrice A dans $\mathcal{M}_n(\mathbb{K})$.

Ces propriétés ajoutées à celle de l'addition des matrices se traduisent en disant que $(\mathcal{M}_n(\mathbb{K}), +, .)$ est un anneau unitaire.

L'associativité du produit matriciel dans $\mathcal{M}_n(\mathbb{K})$ permet de définir les puissances successives d'une matrice A par la relation de récurrence :

$$\begin{cases} A^0 = I_n \\ \forall p \in \mathbb{N},\ A^{p+1} = A^p A = A A^p. \end{cases}$$

Exercice 8.20 *Calculer A^n pour tout entier naturel n, où* $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

Solution 8.20 *On a $A^0 = I_3$, $A^1 = A$ et :*

$$A^2 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix},\ A^3 = \begin{pmatrix} 1 & 3 & 6 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}.$$

En supposant que, pour $n \geq 1$, on a $A^n = \begin{pmatrix} 1 & n & a_n \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}$, *où a_n est un entier à déterminer, on a :*

$$A^{n+1} = \begin{pmatrix} 1 & n & a_n \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & n+1 & n + a_n + 1 \\ 0 & 1 & n+1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & n & a_{n+1} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}$$

avec :

$$a_{n+1} = a_n + n + 1.$$

La suite $(a_n)_{n\geq 1}$ est donc définie par la relation de récurrence :

$$\begin{cases} a_1 = 1 \\ \forall n \geq 1,\ a_{n+1} = a_n + n + 1 \end{cases}$$

ce qui donne :

$$a_n = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

(vérification par récurrence sur $n \geq 1$). On a donc, pour tout $n \geq 0$:

$$A^n = \begin{pmatrix} 1 & n & \frac{n(n+1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}$$

Remarque 8.4 *Le produit des matrices dans $\mathcal{M}_n(\mathbb{K})$ n'est pas commutatif. Par exemple pour $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$, on a :*

$$AB = \begin{pmatrix} 5 & 11 \\ 11 & 25 \end{pmatrix} \neq BA = \begin{pmatrix} 10 & 14 \\ 14 & 20 \end{pmatrix}.$$

On dira que deux matrices A et B dans $\mathcal{M}_n(\mathbb{K})$ commutent si $AB = BA$.

8.6.2 Matrices inversibles

Définition 8.13 *On dit qu'une matrice A dans $\mathcal{M}_n(\mathbb{K})$ est inversible s'il existe une matrice A' dans $\mathcal{M}_n(\mathbb{K})$ telle que $AA' = A'A = I_n$. On dit alors que A' est un inverse de A.*

Si une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est inversible son inverse est alors unique. En effet si A'' est un autre inverse de A, on a :

$$A'' = A''I_n = A''(AA') = (A''A)A' = I_nA' = A'.$$

On note A^{-1} l'inverse de A quand il existe.

On peut aussi remarquer qu'une matrice inversible A n'est jamais nulle puisque $AA^{-1} = I_n \neq 0$.

Exercice 8.21 *Montrer que si la matrice $A \in \mathcal{M}_n(\mathbb{K})$ a une colonne [resp. une ligne] nulle, alors elle n'est pas inversible.*

Solution 8.21 *En notant $C_1, \cdots, C_n$ [resp. $L_1, \cdots, L_n$] les colonne [resp. lignes] de A, on a pour toute matrice $A' \in \mathcal{M}_n(\mathbb{K})$:*

$$A'A = (A'C_1, \cdots, A'C_n)$$

et pour $C_j = 0$, la colonne j de $A'A$ est nulle, ce qui interdit l'égalité $A'A = I_n$.
Pour ce qui est des lignes, on écrit que :

$$AA' = \begin{pmatrix} L_1A' \\ \vdots \\ L_nA' \end{pmatrix}.$$

Exercice 8.22 *Montrer que pour tout réel θ, la matrice $M_\theta = \begin{pmatrix} \cos(\theta) & -\cos(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ est inversible d'inverse $M_{-\theta}$.*

Solution 8.22 *Laissée au lecteur.*

Théorème 8.14 *Si $A \in \mathcal{M}_n(\mathbb{K})$ est inversible, alors A^{-1} est aussi inversible et $(A^{-1})^{-1} = A$. Le produit de deux matrices inversibles A et B est inversible et $(AB)^{-1} = B^{-1}A^{-1}$.*

Démonstration. Le premier point résulte de la définition et le deuxième de :

$$(AB)\left(B^{-1}A^{-1}\right) = A\left(BB^{-1}\right)A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$$

et :

$$\left(B^{-1}A^{-1}\right)(AB) = B^{-1}\left(A^{-1}A\right)B = B^{-1}I_nB = B^{-1}B = I_n.$$

■

Par récurrence, on déduit que le produit de p matrices inversibles $A_1, \cdots, A_p$ est inversible avec $(A_1 \cdots A_p)^{-1} = A_p^{-1} \cdots A_1^{-1}$.

Exercice 8.23 *Soit $P \in \mathcal{M}_n(\mathbb{K})$ inversible. Montrer que $A \in \mathcal{M}_n(\mathbb{K})$ est inversible si, et seulement si, AP [resp. PA] est inversible.*

Solution 8.23 *Le théorème précédent nous dit que la condition est nécessaire.*
Réciproquement si AP [resp. PA] est inversible, alors $A = (AP)P^{-1}$ [resp. $A = P^{-1}(PA)$] est inversible.

Théorème 8.15 *Un endomorphisme u de $\mathbb{K}^n$ est bijectif si, et seulement si, sa matrice A dans la base canonique est inversible et dans ce cas A^{-1} est la matrice de u^{-1} dans la base canonique.*

Démonstration. Supposons u bijectif et notons A' la matrice de u^{-1} dans la base canonique de $\mathbb{K}^n$. De $u \circ u^{-1} = u^{-1} \circ u = Id$, on déduit que $AA' = A'A = I_n$, ce qui signifie que A est inversible d'inverse A'.

Réciproquement supposons A inversible et désignons par u' l'endomorphisme de $\mathbb{K}^n$ de matrice A^{-1} dans la base canonique. La matrice de $u \circ u'$ [resp. de $u' \circ u$] est $AA^{-1} = I_n$ [resp. $A^{-1}A = I_n$], donc $u \circ u' = I_d$ [resp. de $u' \circ u = I_d$] et u est surjective [resp. injective]. L'endomorphisme u est donc bijectif. ■

Exercice 8.24 *On désigne par $(e_i)_{1 \leq i \leq n}$ la base canonique de $\mathbb{K}^n$ avec $n \geq 2$ et pour tout entier j compris entre 1 et n, par M_i la matrice :*

$$M_j = (0, \cdots, 0, e_1, 0, \cdots, 0)$$

la colonne e_1 étant en position j.
Montrer que pour tout scalaire λ et tout entier j compris entre 2 et n, la matrice :

$$P_j(\lambda) = I_n + \lambda M_j$$

est inversible et déterminer son inverse.

Solution 8.24 *Soit, pour j et λ fixés, u l'endomorphisme de $\mathbb{K}^n$ canoniquement associé à $P_j(\lambda)$. Il est défini par :*

$$u(e_k) = \begin{cases} \lambda e_1 + e_j \text{ si } k = j \\ e_k \text{ si } k \neq j \end{cases}$$

Cet endomorphisme est inversible d'inverse u' défini par :

$$u'(e_k) = \begin{cases} -\lambda e_1 + e_j \text{ si } k = j \\ e_k \text{ si } k \neq j \end{cases}$$

donc $P_j(\lambda)$ est inversible d'inverse $P_j(-\lambda)$.

Les matrices $P_j(\lambda)$ sont des matrices de transvection (paragraphe 10.1). On peut vérifier que la multiplication à gauche par une matrice de transvection $P_j(\lambda)$ a pour effet de remplacer la ligne L_1 de A par $L_1 + \lambda L_j$, les autres lignes étant inchangées et la multiplication à droite par une matrice de transvection $P_j(\lambda)$ a pour effet de remplacer la colonne C_j par $C_j + \lambda C_1$, les autres colonnes étant inchangées (théorème 10.1).

Théorème 8.16 *Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est inversible si, et seulement si, l'unique solution du système linéaire $Ax = 0$ est $x = 0$.*

Démonstration. Si A est inversible et x est solution de $Ax = 0$, on a alors $A^{-1}Ax = x = 0$, donc 0 est l'unique solution de $Ax = 0$.

Pour la réciproque, on procède par récurrence sur $n \geq 1$.

On désigne par $(e_i)_{1\leq i\leq n}$ la base canonique de $\mathbb{K}^n$

Pour $n = 1$, $A = (a)$ est inversible si, et seulement si, $a \neq 0$ et le résultat est évident.

Supposons le résultat acquis pour $n - 1 \geq 1$ et soit $A \in \mathcal{M}_n(\mathbb{K})$ telle que $x = 0$ soit l'unique solution du système linéaire $Ax = 0$.

La première colonne de A n'est pas nulle puisque c'est $C_1 = Ae_1$ avec $e_1 \neq 0$, il existe donc un indice j tel que $a_{j1} \neq 0$. Montrer que A est inversible équivaut à montrer que $\dfrac{1}{a_{j1}}A$ est inversible, ce qui nous ramène à $a_{j1} = 1$. Si $a_{11} = 0$, alors $j \geq 2$ et il est équivalent de montrer que la matrice $P_j(1)A$ (notations de l'exercice 8.24) est inversible, ce qui nous ramène à $a_{11} = 1$ ($P_j(1)A$ se déduit de A en ajoutant la ligne j à la ligne 1). Il est encore équivalent de montrer que $AP_2(-a_{12})$ est inversible, ce qui ramène à $a_{12} = 0$ et multipliant à droite par $P_j(-a_{1j})$ pour $2 \leq j \leq n$, on se ramène à $a_{1j} = 0$. En résumé, il suffit de considérer le cas où :

$$A = \begin{pmatrix} 1 & 0 \\ c & B \end{pmatrix}$$

avec $0 \in \mathcal{M}_{1,n-1}(\mathbb{K})$, $c \in \mathcal{M}_{n-1,1}(\mathbb{K})$ et $B \in \mathcal{M}_{n-1}(\mathbb{K})$. Si $x' \in \mathbb{K}^{n-1} \setminus \{0\}$ est solution de $Bx' = 0$, alors $x = \begin{pmatrix} 0 \\ x' \end{pmatrix} \in \mathbb{K}^n \setminus \{0\}$ est solution de $Ax = 0$, ce qui contredit l'hypothèse de départ. Donc $x' = 0$ est l'unique solution de $Bx' = 0$ et B' est inversible. En posant $A' = \begin{pmatrix} 1 & 0 \\ d & B^{-1} \end{pmatrix}$ où $d \in \mathcal{M}_{n-1,1}(\mathbb{K})$ est à préciser, on a :

$$AA' = \begin{pmatrix} 1 & 0 \\ c & B \end{pmatrix}\begin{pmatrix} 1 & 0 \\ d & B^{-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ c + Bd & I_{n-1} \end{pmatrix}$$

et :

$$A'A = \begin{pmatrix} 1 & 0 \\ d & B^{-1} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ c & B \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ d + B^{-1}c & I_{n-1} \end{pmatrix}$$

soit $AA' = A'A = I_n$ en prenant $d = -B^{-1}c$. La matrice A est donc inversible. ■

Corollaire 8.1 *Un endomorphisme u de $\mathbb{K}^n$ est bijectif si, et seulement si, son noyau est réduit à $\{0\}$.*

Démonstration. Si A est la matrice canoniquement associée à u, on a $u(x) = Ax$ et $\ker(u) = \{0\}$ équivaut à dire que 0 est l'unique solution de $Ax = 0$, ce qui équivaut à A inversible encore équivalent à dire que u est un isomorphisme. ■

Corollaire 8.2 *Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est inversible d'inverse $A' \in \mathcal{M}_n(\mathbb{K})$ si, et seulement si, $A'A = I_n$ [resp. $AA' = I_n$].*

Démonstration. Supposons que $A'A = I_n$. Si x est solution de $Ax = 0$, on a alors $x = I_n x = A'(Ax) = A'0 = 0$ et A est inversible avec $A^{-1} = (A'A)A^{-1} = A'$.

Si $AA' = I_n$, la matrice A' est alors inversible d'inverse A, donc $A = (A')^{-1}$ est inversible d'inverse A'. ■

Exercice 8.25 *Montrer que si $A \in \mathcal{M}_n(\mathbb{K})$ a une colonne [resp. une ligne] nulle, alors elle n'est pas inversible.*

Solution 8.25 *Si la colonne j [resp. la ligne i] de A est nulle, alors pour toute matrice $A' \in \mathcal{M}_n(\mathbb{K})$, la matrice $A'A$ [resp. AA'] a sa colonne j [resp. sa ligne i] nulle. En conséquence il ne peut exister de matrice $A' \in \mathcal{M}_n(\mathbb{K})$ telle que $A'A = I_n$ [resp. $AA' = I_n$], ce qui signifie que A n'est pas inversible.*

Montrer qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est inversible et calculer son inverse revient à montrer que pour tout vecteur $y \in \mathbb{K}^n$ le système linéaire de n équation à n inconnues $Ax = y$ a une unique solution et à exprimer cette solution x en fonction de y. Nous verrons plus loin comment l'algorithme de Gauss nous permet d'effectuer une telle résolution.

Exercice 8.26 *Montrer que la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ est inversible et déterminer son inverse.*

Solution 8.26 *Il s'agit de résoudre, pour $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ donné dans $\mathbb{K}^2$, le système :*

$$\left\{ \begin{array}{l} x_1 + 2x_2 = y_1 \\ 3x_1 + 4x_2 = y_2 \end{array} \right.$$

Multipliant la première équation par 3 et retranchant la deuxième équation au résultat obtenu, on a $2x_2 = 3y_1 - y_2$, soit $x_2 = \frac{3}{2}y_1 - \frac{1}{2}y_2$ qui reporté dans la première équation donne $x_1 = -y_1 + y_2$. On a donc :

$$\left\{ \begin{array}{l} x_1 = -2y_1 + y_2 \\ x_2 = \frac{3}{2}y_1 - \frac{1}{2}y_2 \end{array} \right.$$

ce qui signifie que A est inversible et que :

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

De manière plus générale, on a le résultat suivant.

Théorème 8.17 *Une matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{K})$ est inversible si, et seulement si, $ad - bc \neq 0$ et dans ce cas son inverse est donné par :*

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Démonstration. Pour tous scalaires a, b, c, d, on a :

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = (ad - bc) I_2$$

Si $ad - bc \neq 0$, cela s'écrit $AA' = I_2$ avec $A' = \dfrac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, ce qui signifie que A est inversible d'inverse A'. Réciproquement si A est inversible, on a :

$$\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = A^{-1}\left(A\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\right) = (ad - bc)\, A^{-1}$$

donc $ad - bc \neq 0$ (puisque $A \neq 0$) et $A^{-1} = \dfrac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. ■

Exercice 8.27 *Soit* $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. *Pour tout entier* $n \geq 1$, *on désigne par* A^n *la matrice* $A \cdot A \cdot \cdots \cdot A$, *ce produit ayant* n *termes. On note* $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ *et par convention* $A^0 = I_2$.

1. *Calculer* A^2 *et* A^3.
2. *Montrer que pour tout entier* $n \geq 1$, *la matrice* A^n *est de la forme :*
$$A^n = \begin{pmatrix} a_n + (-1)^n & a_n \\ a_n & a_n + (-1)^n \end{pmatrix}$$
où a_n *est un entier.*
3. *Calculer* $A^2 - 2A - 3I_2$.
4. *Montrer que* A *est inversible et calculer* A^{-1}.
5. *Montrer que pour tout entier* $n \geq 2$, *il existe un polynôme* Q_n *et deux entiers* α_n *et* β_n *tels que :*
$$X^n = Q_n(X)\left(X^2 - 2X - 3\right) + \alpha_n X + \beta_n. \tag{8.3}$$
6. *En évaluant* (8.3) *en* -1 *et* 3 *déterminer* α_n *et* β_n.
7. *En déduire* A^n *pour tout* $n \geq 2$.

Solution 8.27

1. *On a :*
$$A^2 = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix},\ A^3 = \begin{pmatrix} 13 & 14 \\ 14 & 13 \end{pmatrix}$$
2. *C'est vrai pour* $n = 1$ *avec* $a_1 = 2$ *et en supposant le résultat acquis pour* n, *on a :*
$$\begin{aligned} A^{n+1} &= \begin{pmatrix} a_n + (-1)^n & a_n \\ a_n & a_n + (-1)^n \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 3a_n + (-1)^n & 3a_n + 2(-1)^n \\ 3a_n + 2(-1)^n & 3a_n + (-1)^n \end{pmatrix} \\ &= \begin{pmatrix} 3a_n + 2(-1)^n + (-1)^{n+1} & 3a_n + 2(-1)^n \\ 3a_n + 2(-1)^n & 3a_n + 2(-1)^n + (-1)^{n+1} \end{pmatrix} \end{aligned}$$
puisque :
$$2(-1)^n + (-1)^{n+1} = (-1)^n(2 - 1) = (-1)^n$$
3. *On a* $A^2 - 2A - 3I_2 = 0$.
4. *On a :*
$$A^{-1} = \frac{1}{3}\begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix}$$
par calcul direct ou avec la question précédente :
$$A^{-1} = \frac{1}{3}(A - 2I_2).$$

5. *Pour* $n = 2$, *on a :*
$$X^2 = \left(X^2 - 2X - 3\right) + 2X + 3$$
donc $Q_2 = 1$ *et* $(\alpha_2, \beta_2) = (2, 3)$. *Supposant le résultat acquis pour* $n \geq 2$, *on a :*
$$\begin{aligned} X^{n+1} &= XQ_n(X)\left(X^2 - 2X - 3\right) + \alpha_n X^2 + \beta_n X \\ &= XQ_n(X)\left(X^2 - 2X - 3\right) + \alpha_n\left(\left(X^2 - 2X - 3\right) + 2X + 3\right) + \beta_n X \\ &= (\alpha_n + XQ_n)(X)\left(X^2 - 2X - 3\right) + (2\alpha_n + \beta_n)X + 3\alpha_n \end{aligned}$$
soit le résultat au rang $n + 1$.

6. -1 *et* 3 *sont les racines de* $X^2 - 2X - 3$, *donc :*
$$\begin{cases} -\alpha_n + \beta_n = (-1)^n \\ 3\alpha_n + \beta_n = 3^n \end{cases}$$
et résolvant le système :
$$\begin{cases} \alpha_n = \dfrac{3^n - (-1)^n}{4} \\ \beta_n = \dfrac{3^n + 3(-1)^n}{4} \end{cases}$$

7. *On a :*
$$\begin{aligned} A^n &= \alpha_n A + \beta_n I_2 = \frac{1}{4}\left(\left(3^n - (-1)^n\right)A + \left(3^n + 3(-1)^n\right)I_2\right) \\ &= \frac{1}{4}\begin{pmatrix} 3^n - (-1)^n + 3^n + 3(-1)^n & 2\left(3^n - (-1)^n\right) \\ 2\left(3^n - (-1)^n\right) & 3^n - (-1)^n + 3^n + 3(-1)^n \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 3^n + (-1)^n & 3^n - (-1)^n \\ 3^n - (-1)^n & 3^n + (-1)^n \end{pmatrix} = \begin{pmatrix} a_n + (-1)^n & a_n \\ a_n & a_n + (-1)^n \end{pmatrix} \end{aligned}$$
avec :
$$a_n = \frac{3^n - (-1)^n}{2} \text{ et } a_n + (-1)^n = \frac{3^n + (-1)^n}{2}.$$

8.6.3 Déterminant d'une matrice d'ordre 2

Définition 8.14 *Le déterminant d'une matrice* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{K})$ *est le scalaire :*
$$\det(A) = ad - bc.$$

Ce déterminant est aussi noté :
$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

Le théorème 8.17 nous dit qu'une matrice $A \in \mathcal{M}_2(\mathbb{K})$ est inversible si, et seulement si, son déterminant est non nul. Nous généraliserons plus loin cette définition du déterminant.

Avec le théorème qui suit on résume les propriétés fondamentales du déterminant des matrices d'ordre 2.

Théorème 8.18 *On désigne par* A, B *des matrices d'ordre* 2.

1. $\det(I_2) = 1$.

2. *Pour tout scalaire* λ, *on a*, $\det(\lambda A) = \lambda^2 \det(A)$.
3. $\det(AB) = \det(A)\det(B)$.
4. *Si* A *est inversible, alors* $\det(A^{-1}) = \dfrac{1}{\det(A)}$.
5. *Si l'une des lignes [resp. des colonnes] de* A *est nulle, alors* $\det(A) = 0$.
6. *Si* A' *est la matrice déduite de* A *en permutant les deux lignes [resp. les deux colonnes], alors* $\det(A') = -\det(A)$.

Démonstration. On note $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ et $B = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$.

1. Il suffit de vérifier.
2. On a $\lambda A = \begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix}$ et :
$$\det(\lambda A) = \lambda^2 ad - bc = \lambda^2 \det(A).$$
3. On a :
$$AB = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}$$
et :
$$\begin{aligned}\det(AB) &= (aa' + bc')(cb' + dd') - (ab' + bd')(ca' + dc') \\ &= aa'cb' + aa'dd' + bc'cb' + bc'dd' - ab'ca' - ab'dc' - bd'ca' - bd'dc' \\ &= aa'dd' + bc'cb' - ab'dc' - bd'ca' \\ &= ad(a'd' - b'c') - bc(a'd' - b'c') \\ &= (ad - bc)(a'd' - b'c') = \det(A)\det(B).\end{aligned}$$
4. Dans le cas où A est inversible, on $AA^{-1} = I_2$ et :
$$\det(A)\det\left(A^{-1}\right) = \det\left(AA^{-1}\right) = \det(I_2) = 1,$$
ce qui donne $\det(A^{-1}) = \dfrac{1}{\det(A)}$.
5. Résulte de la définition.
6. On a $A' = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$ [resp. $A' = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$] et :
$$\det(A') = cb - ad = -\det(A).$$

■

8.6.4 Transposée d'une matrice

Définition 8.15 *Si* $A = ((a_{ij}))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$ *est une matrice à* n *lignes et* m *colonnes, la transposée de* A *est la matrice à* m *lignes et* n *colonnes* ${}^tA = \left(\left(a'_{ij}\right)\right)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ *où* $a'_{ij} = a_{ji}$.

La transposée d'un vecteur colonne $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ est le vecteur ligne ${}^tX = (x_1, x_2, \cdots, x_n)$.

En représentant $A = ((a_{ij}))_{\substack{1\leq i\leq n \\ 1\leq j\leq m}}$ sous forme de lignes $A = \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix}$ [resp. de colonnes $M = (C_1, C_2, \cdots, C_m)$] où :

$$L_i = (a_{i1}, a_{i2}, \cdots, a_{im}) \text{ [resp. } C_j = \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{n,j} \end{pmatrix}]$$

est la ligne numéro i [resp. la colonne numéro j] de M, on a :

$${}^tA = \left({}^tL_1, {}^tL_2, \cdots, {}^tL_n\right) \text{ [resp. } {}^tA = \begin{pmatrix} {}^tC_1 \\ {}^tC_2 \\ \vdots \\ {}^tC_m \end{pmatrix}]$$

Exemple 8.14 *La transposée d'une matrice carrée triangulaire supérieure [resp. inférieure] est triangulaire inférieure [resp. supérieure].*

Définition 8.16 *On dit qu'une matrice carrée $A = ((a_{ij}))_{1\leq i,j\leq n}$ est symétrique si elle est égale à sa transposée, ce qui revient à dire que $a_{ij} = a_{ji}$ pour tous i, j compris entre 1 et n.*

Définition 8.17 *On dit qu'une matrice carrée $A = ((a_{ij}))_{1\leq i,j\leq n}$ est anti-symétrique si ${}^tA = -A$, ce qui revient à dire que $a_{ij} = -a_{ji}$ pour tous i, j compris entre 1 et n.*

Remarque 8.5 *Une matrice anti-symétrique a tous ses termes diagonaux nulles. En effet, pour tout i compris entre 1 et n, on $a_{ii} = -a_{ii}$ et en conséquence, $a_{ii} = 0$.*

Théorème 8.19 *Pour toutes matrices $A = ((a_{i,j}))_{\substack{1\leq i\leq n \\ 1\leq j\leq m}}$ et $B = ((b_{ij}))_{\substack{1\leq i\leq n \\ 1\leq j\leq m}}$ dans $\mathcal{M}_{n,m}(\mathbb{K})$ et tout scalaire λ, on a :*

$${}^t\left({}^tA\right) = A,\ {}^t(A+B) = {}^tA + {}^tB,\ {}^t(\lambda A) = \lambda\, {}^tA$$

Pour toutes matrices $A = ((a_{i,j}))_{\substack{1\leq i\leq p \\ 1\leq j\leq n}}$ dans $\mathcal{M}_{p,n}(\mathbb{K})$ et $B = ((b_{ij}))_{\substack{1\leq i\leq n \\ 1\leq j\leq m}}$ dans $\mathcal{M}_{n,m}(\mathbb{K})$:

$${}^t(AB) = {}^tB\, {}^tA$$

Si $A \in \mathcal{M}_n(\mathbb{K})$ est inversible, alors ${}^tA$ est aussi inversible et :

$$\left({}^tA\right)^{-1} = {}^t\left(A^{-1}\right)$$

Démonstration. On a ${}^tA = \left(\left(a'_{ij} \right) \right)_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$ avec $a'_{ij} = a_{ji}$ et ${}^t\left({}^tA\right) = {}^tA' = \left(\left(a''_{ij} \right) \right)_{\substack{1\leq i\leq n \\ 1\leq j\leq m}}$ avec $a''_{ij} = a'_{ji} = a_{ij}$. Donc ${}^t\left({}^tA\right) = A$.

Les résultats sur les combinaisons linéaires de matrices sont évidents.

Les coefficients de $C = AB$ sont les $c_{ij} = \sum\limits_{k=1}^{n} a_{ik}b_{kj}$ et ceux de ${}^tC$ les :

$$c'_{ij} = c_{ji} = \sum_{k=1}^{n} a_{jk}b_{ki} = \sum_{k=1}^{n} b'_{ik}a'_{kj}$$

et on reconnaît là les coefficients de ${}^tB\ {}^tA$.

Si $A \in \mathcal{M}_n(\mathbb{K})$ est inversible, on a :

$$I_n = {}^tI_n = {}^t\left(AA^{-1}\right) = {}^t\left(A^{-1}\right)\ {}^tA$$

ce qui signifie que ${}^tA$ est inversible avec $\left({}^tA\right)^{-1} = {}^t\left(A^{-1}\right)$. ■

Cette notion de matrice transposée nous sera utile lors de l'étude des formes bilinéaires et quadratiques.

8.6.5 Trace d'une matrice carrée

Définition 8.18 *La trace d'une matrice carrée* $A = ((a_{i,j}))_{1\leq i,j\leq n} \in \mathcal{M}_n(\mathbb{K})$ *est le scalaire :*

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii}$$

(somme des termes diagonaux).

Exemple 8.15 *La trace de la matrice identité* I_n *est* $\mathrm{Tr}(I_n) = n$.

Théorème 8.20 *L'application trace est linéaire de* $\mathcal{M}_n(\mathbb{K})$ *dans* $\mathbb{K}$ *(on dit que c'est une forme linéaire) et pour toutes matrices* A, B *dans* $\mathcal{M}_n(\mathbb{K})$, *on a* $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

Exercice 8.28 *Montrer qu'une matrice et sa transposée ont même trace.*

Exercice 8.29 *Calculer* $\mathrm{Tr}\left({}^tAA\right)$ *pour* $A = ((a_{i,j}))_{1\leq i,j\leq n} \in \mathcal{M}_n(\mathbb{K})$.

8.7 Systèmes d'équations linéaires

On se donne une matrice $A = ((a_{ij}))_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$ dans $\mathcal{M}_{m,n}(\mathbb{K})$, un vecteur $b = (b_i)_{1\leq i\leq m}$ dans $\mathbb{K}^m$ et on s'intéresse au système linéaire $Ax = b$ d'inconnue $x = (x_i)_{1\leq i\leq n}$ dans $\mathbb{K}^n$. Un tel système s'écrit :

$$\left\{ \begin{array}{c} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{array} \right.$$

Pour $b = 0$, un tel système a au moins $x = 0$ comme solution. Le système $Ax = 0$ est le système homogène associé au système $Ax = b$.

En notant, pour j compris entre 1 et n, $C_j = (a_{ij})_{1\leq i\leq m}$ la colonne numéro j de la matrice A, résoudre le système $Ax = b$ revient à trouver tous les scalaires $x_1, \cdots, x_n$ tels que :

$$x_1C_1 + x_2C_2 + \cdots + x_nC_n = b$$

Dans le cas où le nombre d'inconnues n est strictement supérieur au nombre d'équations m, le système homogène $Ax = 0$ a une infinité de solutions.

Théorème 8.21 *Pour toute matrice* $A = ((a_{ij}))_{\substack{1\leq i\leq m\\1\leq j\leq n}} \in \mathcal{M}_{m,n}(\mathbb{K})$ *avec* $n > m$, *le système homogène* $Ax = 0$ *a une infinité de solutions.*

Démonstration. On sait déjà que $x = 0$ est solution.

On désigne par $B = \begin{pmatrix} A \\ 0_{n-m,n} \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$ la matrice carrée d'ordre n ayant pour m premières lignes celles de A, les suivantes étant nulles. Pour toute matrice $B' \in \mathcal{M}_n(\mathbb{K})$ la matrice BB' a sa dernière ligne nulle et ne peut donc égaler I_n, ce qui signifie que la matrice B n'est pas inversible. Il existe donc $x \in \mathbb{K}^n \setminus \{0\}$ tel que $Bx = \begin{pmatrix} Ax \\ 0_{n-m,n} \end{pmatrix} = 0$ (théorème 8.16). Le vecteur x est donc solution non nulle de $Ax = 0$ et la droite dirigée par x nous donne une infinité de solutions de ce système linéaire. ■

Exercice 8.30 *Résoudre le système linéaire :*

$$\begin{cases} 2x + y - z = 0 \ (1) \\ x + y + z = 0 \ (2) \end{cases}$$

Solution 8.28 *On élimine l'inconnue* z *en additionnant les deux équations, ce qui donne* $3x + 2y = 0$, *soit* $y = -\frac{3}{2}x$ *qui reporté dans* (1) *nous donne* $z = \frac{1}{2}x$. *Une solution de ce système est donc de la forme* $X = \frac{x}{2}u$ *où* u *est le vecteur* $u = \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}$ *et* x *un scalaire. Réciproquement l'égalité* $Au = 0$ *nous dit que tout vecteur* X *colinéaire à* u *est solution du système.*
En définitive l'ensemble des solutions de ce système est la droite $D = \mathbb{K}u$ *dirigée par* u.

Théorème 8.22 *Soit* A *dans* $\mathcal{M}_n(\mathbb{K})$. *Le système* $Ax = b$ *a une unique solution dans* $\mathbb{K}^n$ *pour tout vecteur* b, *si et seulement si, la matrice* A *est inversible.*

Démonstration. Supposons A inversible. Le vecteur $A^{-1}b$ est solution de $Ax = b$ et si x est solution de ce système, on a alors $A^{-1}(Ax) = A^{-1}b$, soit $x = A^{-1}b$. Notre système a bien une unique solution.

Réciproquement, supposons que, pour tout $b \in \mathbb{K}^n$ le système $Ax = b$ a unique solution. En désignant par $(e_j)_{1\leq j\leq n}$ la base canonique de $\mathbb{K}^n$ et, pour tout j compris entre 1 et n, par C_j la solution de $Ax = e_j$, la matrice $A' = (C_1, \cdots, C_n)$ est telle que :

$$AA' = (AC_1, \cdots, AC_n) = (e_1, \cdots, e_n) = I_n$$

ce qui signifie que A est inversible d'inverse A'. ■

La méthode des pivots de Gauss peut être utilisée pour résoudre un tel système. Nous décrivons dans un premier temps cette méthode sur un exemple avec l'exercice qui suit.

Exercice 8.31 *Résoudre le système linéaire :*

$$\begin{cases} x + y + z = 3 \ (1) \\ 2x + y + z = 4 \ (2) \\ x - y + z = 1 \ (3) \end{cases}$$

Solution 8.29 *La première étape consiste à éliminer x dans les équations (2) et (3). Pour ce faire on remplace l'équation (2) par $(2)-2(1)$ et l'équation (3) par $(3)-(1)$, ce qui donne le système :*

$$\begin{cases} x+y+z=3 \ (1) \\ -y-z=-2 \ (2) \\ -2y=-2 \ (3) \end{cases}$$

La deuxième étape consiste à éliminer y dans l'équation (3) en remplaçant cette équation par $(3)-2(2)$, ce qui donne :

$$\begin{cases} x+y+z=3 \ (1) \\ -y-z=-2 \ (2) \\ 2z=2 \ (3) \end{cases}$$

Le système obtenue est alors un système triangulaire et il se résout en remontant les équations, ce qui donne :

$$\begin{cases} z=1 \\ y=2-z=1 \\ x=3-y-z=1 \end{cases}$$

8.8 Sommes et sommes directes de sous-espaces vectoriels

On se donne pour ce paragraphe un espace vectoriel E.

Définition 8.19 *Soient F et G deux sous-espaces vectoriels de E. On dit que E est somme des espaces F et G si l'application :*

$$\begin{array}{cccc} \varphi & F\times G & \rightarrow & E \\ & (x,y) & \mapsto & x+y \end{array}$$

est surjective et on note alors $E=F+G$.
Si cette application est bijective, on dit que E est somme directe des espaces F et G et on note $E=F\oplus G$.

Dire que $E=F+G$ signifie donc que l'on peut écrire tout vecteur $x\in E$ sous la forme $x=y+z$, où $y\in F$ et $z\in G$ et dire que $E=F\oplus G$ signifie donc que l'on peut écrire de manière unique tout vecteur $x\in E$ sous la forme $x=y+z$, où $y\in F$ et $z\in G$.

Théorème 8.23 *Soient F et G deux sous-espaces vectoriels de E. On a $E=F\oplus G$ si, et seulement si, $E=F+G$ et $F\cap G=\{0\}$.*

Démonstration. Supposons que $E=F\oplus G$, on a alors $E=F+G$ et tout $x\in F\cap G$ s'écrit $x=x+0=0+x$ avec $(x,0)$ et $(0,x)$ dans $F\times G$, ce qui impose $x=0$ puisque φ est bijective.

Réciproquement supposons que $E=F+G$ et $F\cap G=\{0\}$. Si $x\in E$ s'écrit $x=y+z=y'+z'$, avec y,y' dans F et z,z' dans G, on a alors $y-y'=z'-z\in F\cap G$, donc $y-y'=z-z'=0$ et $(y,z)=(y',z')$. La somme est donc directe. ■

Définition 8.20 *On dit que deux sous-espaces vectoriels F et G de E sont supplémentaires, si $E=F\oplus G$. On dit aussi que F est un supplémentaire de G ou que G est un supplémentaire de F dans E.*

Remarque 8.6 *E est l'unique supplémentaire de $\{0\}$, mais un sous-espace vectoriel F de E distinct de $\{0\}$ et de E admet une infinité de supplémentaires. Il suffit de considérer deux droites de $\mathbb{R}^2$ dirigées par deux vecteurs non colinéaires pour s'en convaincre.*

On peut définir la somme ou la somme directe de p sous-espaces de E comme suit.

Définition 8.21 *Soient $p \geq 2$ un entier et $F_1, \cdots, F_p$ des sous-espaces vectoriels de E. On dit que E est somme des espaces $F_1, \cdots, F_p$ si l'application :*

$$\begin{array}{cccc} \varphi & F_1 \times \cdots \times F_p & \to & E \\ & (x_1, \cdots, x_p) & \mapsto & x_1 + \cdots + x_p \end{array}$$

est surjective et on note alors $E = F_1 + \cdots + F_p$ ou de manière plus compacte $E = \sum_{k=1}^{p} F_k$. Si cette application est bijective, on dit que E est somme directe des espaces $F_1, \cdots, F_p$ et on note $E = F_1 \oplus \cdots \oplus F_p$ ou $E = \bigoplus_{k=1}^{p} F_k$.

En fait la somme de deux sous-espaces vectoriels F et G peut se définir par :

$$F + G = \{y + z \mid y \in F \text{ et } z \in G\}.$$

Il est facile de vérifier que $F + G$ est un sous-espace vectoriel de E. Ce sous-espace n'est en général pas égal à E.

On peut aussi définir $F + G$ comme le sous-espace vectoriel de E engendré par la réunion $F \cup G$ (qui en général n'est pas un espace vectoriel).

Théorème 8.24 *Si F, G sont deux sous-espaces vectoriels de E, alors la somme $F + G$ est le sous-espace vectoriel de E engendré par $F \cup G$.*

Démonstration. Dire que $x \in \operatorname{Vect}(F \cup G)$ équivaut à dire qu'il s'écrit $x = \sum_{k=1}^{p} \lambda_k x_k$ où les x_k sont des éléments de $F \cup G$ et les λ_k des scalaires. En séparant les x_k qui sont dans F deux ceux qui sont dans G, cette somme peut s'écrire $x = y + z$ avec $y \in F$ et $z \in G$, ce qui signifie que $x \in F + G$.

Réciproquement $x \in F + G$ s'écrit $x = y + z$ avec $(y, z) \in F \times G$ et il est bien dans $\operatorname{Vect}(F \cup G)$. ■

De manière un peu plus générale, on a le résultat suivant.

Théorème 8.25 *Si $F_1, \cdots, F_p$ sont des sous-espaces vectoriels de E, alors la somme $\sum_{k=1}^{p} F_k$ est le sous-espace vectoriel de E engendré par $\bigcup_{k=1}^{p} F_k$.*

Exercice 8.32 *Montrer que si φ est une forme linéaire non nulle sur E, il existe alors un vecteur non nul a dans E tel que :*

$$E = \ker(\varphi) \oplus \mathbb{K}a.$$

Solution 8.30 *La forme linéaire φ étant non nulle, on peut trouver un vecteur a dans E tel que $\varphi(a) \neq 0$. Ce vecteur a est nécessairement non nul. Pour tout vecteur x dans E, le vecteur $h = x - \frac{\varphi(x)}{\varphi(a)} a$ est dans le noyau de φ et en écrivant que $x = h + \frac{\varphi(x)}{\varphi(a)} a$ on déduit que $E = \ker(\varphi) + \mathbb{K}a$. Si x est dans $\ker(\varphi) \cap \mathbb{K}a$ on a alors $x = \lambda a$ et $\lambda \varphi(a) = \varphi(x) = 0$ avec $\varphi(a) \neq 0$ ce qui entraîne $\lambda = 0$ et $x = 0$. On a donc $\ker(\varphi) \cap \mathbb{K}a = \{0\}$ et $E = \ker(\varphi) \oplus \mathbb{K}a$.*