Sommaire

1 Espaces vectoriels

Sommaire

"A force de forger, on devient forgeron.
A force de faire des mathématiques,
on devient mathématicien.
A force de travailler, on réussit."

.

VOUS AVEZ DÉJÀ rencontré les vecteurs en physique et en géométrie élémentaire. L'ensemble des vecteurs du plan est muni de deux opérations naturelles : l'addition des vecteurs, et la multiplication d'un vecteur par un nombre réel. L'espace physique est lui aussi équipé de cette structure. Bien d'autres ensembles sont munis de deux opérations possédant des propriétés analogues à celles-ci : des ensembles de fonctions, de polynômes, de suites, etc. C'est cette structure commune, appelée structure d'espace vectoriel, que nous allons étudier dans ce chapitre. Compte tenu de l'origine concrète et géométrique que nous avons indiquée, il est conseillé de s'appuyer sur cette analogie pour nourrir son intuition dans un contexte plus général.

Dans ce chapitre, $\mathbb{K}$ désigne un corps commutatif, qui sera en pratique $\mathbb{R}$ ou $\mathbb{C}$.

1.1 Structure d'espace vectoriel

Avant de donner la définition théorique d'un espace vectoriel E sur le corps $\mathbb{K}$, nous allons traiter d'abord les deux exemples concrets du plan $\mathbb{R}^2$ et de l'espace physique $\mathbb{R}^3$ pour motiver l'étude générale. Ces deux exemples vont nous servir comme modèles pour définir la notion d'espace vectoriel dans un cadre abstrait.

Exemple 1.1.1. Considérons le plan $\mathbb{R}^2$ constitué des éléments de la forme $v = (x, y)$ avec x et y dans $\mathbb{R}$. On munit $\mathbb{R}^2$ de l'addition habituelle

$$\boxed{(x, y) + (x', y') = (x + x', y + y')}$$

pour tous (x, y) et (x', y') dans $\mathbb{R}^2$, et de la multiplication par les réels, appelée aussi *multiplication externe* ou encore *loi externe*, donnée par

$$\boxed{\alpha \cdot (x, y) = (\alpha x, \alpha y)}$$

pour tout $\alpha \in \mathbb{R}$ et $(x, y) \in \mathbb{R}^2$.

Alors l'ensemble $\mathbb{R}^2$, muni de ces deux opérations, satisfait les propriétés suivantes :

1. $v_1 + v_2 = (x, y) + (x', y') = (x + x', y + y') = (x' + x, y' + y) = v_2 + v_1$.
L'addition est donc commutative dans $\mathbb{R}^2$.

2. Soient $v_1 = (x, y), v_2 = (x', y'), v_3 = (x", y")$ dans $\mathbb{R}^2$.
Montrons que $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$.

$$\begin{aligned}
(v_1 + v_2) + v_3 &= [(x, y) + (x', y')] + (x", y") \\
&= (x + x', y + y') + (x", y") && \text{par définition de la loi } + \text{ dans } \mathbb{R}^2 \\
&= \big((x + x') + x", (y + y') + y"\big) && \text{par définition de la loi } + \text{ dans } \mathbb{R}^2 \\
&= \big(x + (x' + x"), y + (y' + y")\big) && \text{par l'associativité de l'addition dans } \mathbb{R} \\
&= (x, y) + (x' + x", y' + y") && \text{par définition de la loi } + \text{ dans } \mathbb{R}^2 \\
&= (x, y) + [(x', y') + (x", y")] && \text{par définition de la loi } + \text{ dans } \mathbb{R}^2 \\
&= v_1 + (v_2 + v_3).
\end{aligned}$$

Donc l'addition est associative dans $\mathbb{R}^2$.

3. Posons $0_{\mathbb{R}^2} = (0, 0)$. Alors $0_{\mathbb{R}^2} + v = (0, 0) + (x, y) = (0 + x, 0 + y) = (x, y) = v$ pour tout $v = (x, y)$ dans $\mathbb{R}^2$.
Donc $0_{\mathbb{R}^2}$ est l'élément neutre de l'addition dans $\mathbb{R}^2$.

4. Pout élément $v = (x, y) \in \mathbb{R}^2$, il existe $w = (a, b) \in \mathbb{R}^2$ tel que $v + w = 0_{\mathbb{R}^2}$.
En effet, $v + w = 0_{\mathbb{R}^2} \Leftrightarrow (x + a, y + b) \Leftrightarrow x + a = 0$ et $y + b = 0$
$\Leftrightarrow a = -x$ et $b = -y \Leftrightarrow w = (-x, -y)$.

Ainsi, tout élément $v = (x, y) \in \mathbb{R}^2$ est symétrisable et son élément symétrique est $w = (-x, -y)$, qu'on note $-v$.

Selon votre cours d'Algèbre du premier semestre SMIA1, les quatre premières propriétés précédentes nous disent simplement que $(\mathbb{R}^2, +)$ est un **groupe abélien**.

D'autre part, la multiplication externe satisfait les assertions suivantes :

5. Soient $\alpha \in \mathbb{R}$ et $v_1 = (x, y), v_2 = (x', y') \in \mathbb{R}^2$.
Vérifions que $\alpha \cdot (v_1 + v_2) = \alpha \cdot v_1 + \alpha \cdot v_2$.

$$\begin{aligned}
\alpha \cdot (v_1 + v_2) &= \alpha \cdot (x + x', y + y') \\
&= \big(\alpha(x + x'),\ \alpha(y + y')\big) && \text{par définition de la multiplication externe} \\
&= \big(\alpha x + \alpha x',\ \alpha y + \alpha y')\big) \\
&= \big(\alpha x, \alpha y\big) + \big(\alpha x', \alpha y'\big) && \text{par définition de la loi } + \text{ dans } \mathbb{R}^2 \\
&= \alpha \cdot v_1 + \alpha \cdot v_2. && \text{par définition de la multiplication externe}
\end{aligned}$$

6. Pour tous $\alpha, \beta \in \mathbb{R}$ et $v = (x, y) \in \mathbb{R}^2$, on a $(\alpha + \beta).v = \alpha.v + \beta.v$. En effet,

$$\begin{aligned}
(\alpha + \beta) \cdot v &= (\alpha + \beta) \cdot (x, y) \\
&= \big((\alpha + \beta)x,\ (\alpha + \beta)y\big) && \text{par définition de la multiplication externe} \\
&= \big(\alpha x + \beta x,\ \alpha y + \beta y\big) \\
&= \big(\alpha x, \alpha y\big) + \big(\beta x, \beta y\big) && \text{par définition de la loi } + \text{ dans } \mathbb{R}^2 \\
&= \alpha \cdot (x, y) + \beta \cdot (x, y) && \text{par définition de la multiplication externe} \\
&= \alpha \cdot v + \beta \cdot v.
\end{aligned}$$

7. Pour tous $\alpha, \beta \in \mathbb{R}$ et $v = (x, y) \in \mathbb{R}^2,\ \alpha \cdot \big(\beta \cdot v)\big) = (\alpha\beta) \cdot v$, car

$$\begin{aligned}
\alpha \cdot (\beta \cdot v)) &= \alpha \cdot (\beta x, \beta y) && \text{par définition de la multiplication externe} \\
&= \big(\alpha(\beta x),\ \alpha(\beta y)\big) && \text{par définition de la multiplication externe} \\
&= \big((\alpha\beta)x,\ (\alpha\beta)y\big) && \text{par associativité de la multiplication dans } \mathbb{R} \\
&= (\alpha\beta) \cdot (x, y) && \text{par définition de la multiplication externe} \\
&= (\alpha\beta) \cdot v.
\end{aligned}$$

8. Finalement, $1 \cdot v = v$ puisque $1 \cdot v = 1 \cdot (x, y) = (1 \times x, 1 \times y) = (x, y) = v$.

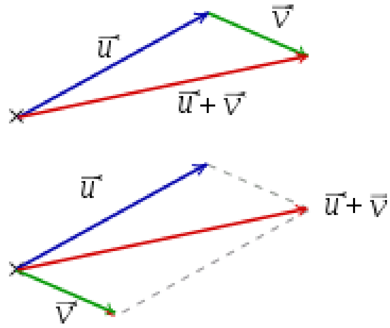

FIGURE 1.1: Somme de deux vecteurs $\overrightarrow{u}$ et $\overrightarrow{v}$

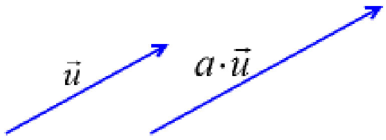

FIGURE 1.2: Multiplication d'un vecteur $\overrightarrow{u}$ par un scalaire a

Exemple 1.1.2. De même, l'ensemble $\mathbb{R}^3$ des éléments de la forme $v = (x, y, z)$ avec $x, y, z \in \mathbb{R}$ peut être équipé des deux opérations suivantes

$$\boxed{(x, y, z) + (x', y', z') = (x + x', y + y', z + z')}$$

$$\boxed{\alpha \cdot (x, y, z) = (\alpha x, \alpha y, \alpha z)}$$

pour tous (x, y, z) et (x', y', z') dans $\mathbb{R}^3$ et pour tout $\alpha \in \mathbb{R}$.
On vérifie de la même manière que les huit propriétés précédentes sont également valables dans $\mathbb{R}^3$. Ici, l'élément neutre du groupe abélien $(\mathbb{R}^3, +)$ est $0_{\mathbb{R}^3} = (0, 0, 0)$, tandis que l'élément symétrique de $v = (x, y, z)$ est donné par $-v = (-x, -y, -z)$.

Le plan $\mathbb{R}^2$ et l'espace $\mathbb{R}^3$ jouissent des mêmes propriétés structurelles par rapport aux opérations d'addition et de multiplication scalaire. Ces propriétés s'étendent à d'autres ensembles, en particulier des ensembles de polynômes, de fonctions ou de suites numériques. Ainsi, plutôt que d'étudier séparément et à plusieurs reprises les propriétés communes de ces differents ensembles, il est plus commode d'introduire un unique concept regroupant ces propriétés. C'est dans cette optique qu'est introduite la notion d'espace vectoriel.

Soit E un ensemble quelconque non vide. Rappelons qu'une *loi interne* sur E est la donnée d'une application de $E \times E$ dans E. Traditionnellement, l'image (ou le composé) d'un couple $(u, v) \in E \times E$ peut se noter $u * v$, $u \triangle v$, $u \top v, \ldots$
Mais dans le cadre des espaces vectoriels, comme pour les groupes abéliens, on

adopte la notation additive $u + v$. Le fait que $+$ soit une loi interne veut dire tout simplement :

$$\boxed{\forall u, v \in E \qquad u + v \in E}$$

Définition 1.1.3. *Soient $\mathbb{K}$ un corps commutatif ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) et E un ensemble. On appelle **loi externe sur E à scalaires dans** $\mathbb{K}$ une application de $\mathbb{K} \times E$ dans E. On note $\alpha \cdot v$ ou simplement αv l'image du couple $(\alpha, v) \in \mathbb{K} \times E$.*
Cela signifie que

$$\boxed{\forall \alpha \in \mathbb{K} \quad \forall v \in E \quad \alpha \cdot v \in E}$$

*Une loi externe s'appelle aussi **produit externe** ou encore **multiplication par un scalaire**.*

Exemple 1.1.4.

1. Considérons le corps $K = \mathbb{R}$. Pour $E = \mathbb{R}^2$ (respectivement $E = \mathbb{R}^3$), nous avons déjà considéré la loi externe définie par $\alpha \cdot (x, y) = (\alpha x, \alpha y)$ (respectivement $\alpha \cdot (x, y, z) = (\alpha x, \alpha y, \alpha z)$).

2. On peut définir une autre loi externe sur $E = \mathbb{R}^2$ à scalaires dans $\mathbb{K} = \mathbb{R}$ en posant $\alpha \cdot (x, y) = (\alpha x, 0)$. Mais on verra que ceci ne conviendra pas pour la structure d'espace vectoriel.

3. Si l'on prend $K = \mathbb{C}$ et $E = \mathbb{R}^3$, l'opération $\alpha \cdot (x, y, z) = (\alpha x, \alpha y, \alpha z)$ pour tout $\alpha \in \mathbb{C}$ et $(x, y, z) \in \mathbb{R}^3$ ne définit pas une loi externe, parce que $(\alpha x, \alpha y, \alpha z)$ n'appartient pas nécessairement à $\mathbb{R}^3$.

Définition 1.1.5. *Soit E un ensemble non vide muni d'une loi interne $(v, v') \in E \times E \longrightarrow v + v' \in E$, et muni d'une loi externe : $(\alpha, v) \in \mathbb{K} \times E \longrightarrow \alpha \cdot v \in E$. On dit que $(E, +, .)$ est un **espace vectoriel sur** $\mathbb{K}$ si :*

1. $(E, +)$ est un groupe abélien ;

2. La loi externe vérifie les axiomes suivants :

   (a) $\alpha \cdot (v_1 + v_2) = \alpha \cdot v_1 + \alpha \cdot v_2$

   (b) $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$

   (c) $\alpha \cdot (\beta \cdot v) = (\alpha\beta) \cdot v$

   (d) $1 \cdot v = v$

   pour tous $v, v_1, v_2 \in E$ et pour tous $\alpha, \beta \in \mathbb{K}$.

Vocabulaire. On dit aussi que E est un $\mathbb{K}$-**espace vectoriel**. Lorsqu'il n'y a aucune ambiguité sur le corps $\mathbb{K}$, on utilise simplement l'expression d'**espace vectoriel**.
Par commodité, les éléments d'un espace vectoriel E se notent $x, y, z, u, v, w, \ldots$ et s'appellent **vecteurs**, tandis que les éléments du corps $\mathbb{K}$ sont représentés par des lettres grecques $\alpha, \beta, \gamma, \lambda, \ldots$ et s'appellent **scalaires**.
L'élément neutre du groupe abélien $(E, +)$ sera noté 0_E, et l'élément symétrique d'un élément $v \in E$ s'écrira $-v$.

Des erreurs fréquentes à ne pas commettre : il ne faut pas confondre le scalaire $0 \in \mathbb{K}$ avec le vecteur nul $0_E \in E$. Aussi, ne pas confondre l'addition $+$ du corps $\mathbb{K}$ ($= \mathbb{R}$ ou $\mathbb{C}$) avec l'addition $+$ de l'espace vectoriel E. Ce sont deux lois internes différentes qu'on désigne de la même manière pour simplifier les notations.

Remarque 1.1.6. Soient v_1 et v_2 sont des vecteurs dans un espace vectoriel E. Nous utilisons la notation $v_1 - v_2$ au lieu de $v_1 + (-v_2)$, où $-v_2$ est l'élément symétrique de v_2 dans le groupe abélien $(E, +)$.

Il est inutile de s'inquiéter de la quantité de propriétés à vérifier dans la définition d'espace vectoriel. Dans tous les exemples que lon rencontrera, les opérations sont parfaitement naturelles et leurs propriétés évidentes, comme c'était le cas dans les deux exemples 1.1.1 et 1.1.2. On ne vérifie d'ailleurs presque jamais les huit propriétés de la définition (quatre axiomes pour le groupe abélien $(E, +)$ et quatre autres axiomes pour la loi externe). On reviendra à cette remarque quand nous aborderons la notion de sous-espace vectoriel qui simplifiera la tâche et nous permettra d'éviter la vérification de longs calculs élémentaires et inutiles.

La structure d'espace vectoriel réside donc dans le fait qu'on peut additionner et multiplier par un scalaire les vecteurs. Avec un peu d'effort d'abstraction, la notion d'espace vectoriel prendra son sens petit à petit, lorsque vous aurez traité de nombreux exemples. Suivant les espaces dans lesquels on se place, ces vecteurs peuvent être des nombres, des fonctions, des suites, etc.

Examinons maintenant des exemples classiques d'espaces vectoriels avec lesquels nous allons travailler tout au long de ce cours.

Exemple 1.1.7. Les exemples 1.1.1 et 1.1.2 nous disent simplement que les ensembles $\mathbb{R}^2$ et $\mathbb{R}^3$, munis des opérations naturelles précitées, sont des espaces vectoriels sur le corps $\mathbb{K} = \mathbb{R}$.

Exemple 1.1.8. Soient $n \geq 1$ un entier et $\mathbb{K}^n$ l'ensemble des n-uplets $(x_1, \ldots, x_n)$ où les $x_i \in \mathbb{K}$. Etant donnés deux éléments $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$

dans $\mathbb{K}^n$ et α dans $\mathbb{K}$, nous pouvons définir l'addition et la multiplication externe en posant

$$\boxed{(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)}$$

$$\boxed{\alpha(x_1, \ldots, x_n) = (\alpha x_1, \ldots, \alpha x_n)}$$

On montre alors aisément que $(\mathbb{K}^n, +, .)$ est un espace vectoriel sur $\mathbb{K}$. Les vérifications faciles se font exactement commme pour $\mathbb{R}^2$ et $\mathbb{R}^3$ dans les exemples 1.1.1 et 1.1.2. L'élément neutre pour l'addition est le vecteur nul $0_{\mathbb{K}^n} = (0, \ldots, 0) \in \mathbb{K}^n$. L'élément symétrique du vecteur $v = (x_1, \ldots, x_n)$ est le vecteur $-v = (-x_1, \ldots, -x_n)$. En particulier, pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, alors :

- $\mathbb{R}^n$ **est un espace vectoriel sur** $\mathbb{R}$,
- $\mathbb{C}^n$ **est un espace vectoriel sur** $\mathbb{C}$.

En particulier, pour $\mathbb{K} = \mathbb{R}$ et $n = 2$ ou 3, on retrouve les deux $\mathbb{R}$-espaces $\mathbb{R}^2$ et $\mathbb{R}^3$ traités dans les exemples 1.1.1 et 1.1.2.

Exemple 1.1.9. En prenant $n = 1$ dans l'exemple précédent, nous voyons que le corps $\mathbb{K}$ est lui même un espace vectoriel sur $\mathbb{K}$. Ici, la loi interne et la loi externe sont respectivement l'addition et la multiplication définies sur $\mathbb{K}$. Ceci nous fournit l'exemple le plus simple d'espace vectoriel sur $\mathbb{K}$. On ne peut pas, dans ce cas, faire la distinction entre les vecteurs et les scalaires. En particulier, $\mathbb{R}$ est un espace vectoriel sur $\mathbb{R}$, et $\mathbb{C}$ est un espace vectoriel sur $\mathbb{C}$.

Exemple 1.1.10. L'espace vectoriel des polynômes.
Soit $\mathbb{K}[X]$ l'ensemble des polynômes en l'indéterminée X à cœfficients dans $\mathbb{K}$. On le munit de la loi interne

$$(P, Q) \in \mathbb{K}[X] \times \mathbb{K}[X] \longrightarrow P + Q \in \mathbb{K}[X]$$

qui est l'addition des polynômes, et de la multiplication externe

$$(\alpha, P) \in \mathbb{K} \times \mathbb{K}[X] \longrightarrow \alpha P \in \mathbb{K}[X]$$

qui est la multiplication d'un polynôme par un élément de $\mathbb{K}$.

Alors $\mathbb{K}[X]$ est un espace vectoriel sur $\mathbb{K}$. En effet, on sait déjà (cours d'Algèbre SMIA1) que $(\mathbb{K}[X], +)$ est un groupe abélien. Nous laissons aux étudiants le soin de vérifier les quatre axiomes de la loi externe. Notons que les vecteurs ici sont juste des polynômes et ne peuvent donc pas être représentés géométriquement comme les vecteurs du plan $\mathbb{R}^2$ ou de l'espace $\mathbb{R}^3$.

L'Analyse nous fournit des exemples importants d'espaces vectoriels, à travers les fonctions et les suites. Ces espaces vectoriels s'appellent souvent *espaces fonctionnels.* En voici quelques exemples :

Exemple 1.1.11. L'espace vectoriel des applications.
Soit X un ensemble non vide et notons $\mathcal{F}(X, \mathbb{K})$ l'ensemble des applications de X vers $\mathbb{K}$. Cet ensemble s'écrit aussi $\mathbb{K}^X$. Equippons $\mathcal{F}(X, \mathbb{K})$ de l'addition et de la multiplication par un scalaire de la manière suivante :

- Pour toutes applications $f : X \longrightarrow \mathbb{K}$ et $g : X \longrightarrow \mathbb{K}$, la nouvelle application $f + g : X \longrightarrow \mathbb{K}$ est définie par

$$\boxed{\forall x \in X \quad (f+g)(x) = f(x) + g(x)}$$

Pour toute application $f : X \longrightarrow \mathbb{K}$ et pour tout scalaire $\alpha \in \mathbb{K}$, la nouvelle application $\alpha \cdot f : X \longrightarrow \mathbb{K}$ est donnée par

$$\boxed{\forall x \in X \quad (\alpha \cdot f)(x) = \alpha \times f(x)}$$

Alors l'ensemble $\mathcal{F}(X, \mathbb{K})$, muni de ces deux opérations, est un espace vectoriel sur $\mathbb{K}$. En effet, rappelons d'abord le fait important suivant qui sera utilisée à maintes reprises : pour deux applications $f : X \longrightarrow \mathbb{K}$ et $g : X \longrightarrow \mathbb{K}$, on a

$$\boxed{f = g \Leftrightarrow f(x) = g(x) \quad \forall x \in X}$$

1. Soient $f, g \in \mathcal{F}(X, \mathbb{K})$; alors $f + g = g + f$, car pour tout $x \in X$, $(f + g)(x) = f(x) + g(x)$. Or, l'addition dans $\mathbb{K}$ est commutative, donc $f(x) + g(x) = g(x) + f(x) = (g + f)(x)$. D'où $f + g = g + f$.

2. De même, l'associativité de l'addition dans $\mathcal{F}(X, \mathbb{K})$ suit immédiatement l'associativité de l'addition dans $\mathbb{K}$.

3. Définissons l'*application nulle* $\theta : X \longrightarrow \mathbb{K}$ par $\theta(x) = 0 \quad \forall x \in \mathbb{K}$. Alors pour tout $f \in \{(\mathbb{R}, \mathbb{R})$, nous avons $f + \theta = f$ puisque $(f + \theta)(x) = f(x) + \theta(x) = f(x) + 0 = f(x) \quad \forall x \in X$. Ainsi, l'application nulle θ joue le rôle de l'élément neutre $0_{\mathcal{F}(X,\mathbb{K})}$.

4. Soit $f \in \mathcal{F}(X, \mathbb{K})$ et cherchons $g \in \mathcal{F}(X, \mathbb{K})$ tel que $f + g = \theta$. Ceci équivaut à $f(x) + g(x) = 0 \quad \forall x \in X$. Donc l'élément symétrique g de f est l'application $g : X \longrightarrow \mathbb{K}$ définie par $g(x) = -f(x) \quad \forall x \in X$.

Par conséquent $(\mathcal{F}(X, \mathbb{K}), +)$ est un groupe abélien.

5. $\alpha(f+g) = \alpha f + \alpha g \quad \forall \alpha \in \mathbb{K},\ \forall f, g \in \mathcal{F}(X, \mathbb{K})$, car pour tout $x \in X$, on a $\big[\alpha(f+g)\big](x) = \alpha\big[(f+g)(x)\big] = \alpha\big[f(x)+g(x)\big]$. Or la multiplication est distributive par rapport à l'addition dans le corps $\mathbb{K}$. Donc $\alpha\big[f(x)+g(x)\big] = \alpha f(x) + \alpha g(x) = (\alpha f)(x) + (\alpha g)(x) = (\alpha f + \alpha g)(x)$. D'où $\big[\alpha(f+g)\big](x) = (\alpha f + \alpha g)(x) \quad \forall x \in X$, et par suite, $\alpha(f+g) = \alpha f + \alpha g$.

Les axiomes

6. $(\alpha+\beta)f = \alpha f + \beta f$ et

7. $\alpha(\beta f) = (\alpha\beta)f$

se démontrent de façon smilaire.

8. Enfin, $1 \cdot f = f \quad \forall f \in \mathcal{F}(X, \mathbb{K})$ parce que $(1 \cdot f)(x) = 1 \times f(x) = f(x)$ pour tout $x \in X$.

Observons que les vecteurs de $\mathcal{F}(X, \mathbb{K})$ sont les applications de X vers $\mathbb{K}$ et les scalaires sont les éléments du corps K.

Attention, il ne faut pas confondre une application $f : X \longrightarrow \mathbb{K}$, qui est un élément de $\mathcal{F}(X, \mathbb{K})$, avec sa valeur $f(x) \in \mathbb{K}$ en un élément x de X, qui est un scalaire.

Cas particulier : en prenant $X = \mathbb{K} = \mathbb{R}$, nous obtenons l'espace vectoriel $\mathcal{F}(\mathbb{R}, \mathbb{R}) = \mathbb{R}^{\mathbb{R}}$ des applications (ou fonctions) de $\mathbb{R}$ vers $\mathbb{R}$.

Exemple 1.1.12. L'espace vectoriel des suites.

Soit $\mathcal{U}$ l'ensemble des suites $(u_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathbb{K}$. Lorsque $\mathbb{K} = \mathbb{R}$ (respectivement $\mathbb{K} = \mathbb{C}$), il s'agit des suites réelles (respectivement complexes). On munit $\mathcal{U}$ des lois interne et externe suivantes

$$\boxed{(u_n)_{n\in\mathbb{N}} + (v_n)_{n\in\mathbb{N}} = (u_n + v_n)_{n\in\mathbb{N}} \qquad \alpha(u_n)_{n\in\mathbb{N}} = (\alpha u_n)_{n\in\mathbb{N}}}$$

pour toute suite $u = (u_n)_{n\in\mathbb{N}} \in \mathcal{U}$ et pour tout scalaire $\alpha \in \mathbb{K}$.

Alors l'ensemble $\mathcal{U}$, muni de ces deux opérations, est un espace vectoriel sur $\mathbb{K}$. Les vérifications sont encore une fois élémentaires, mais on peut le déduire immédiatement sans faire aucun calcul. En effet, remarquons qu'une suite $u = (u_n)_{n\in\mathbb{N}}$ dans $\mathcal{U}$ peut être regardée comme une application : $\mathbb{N} \longrightarrow \mathbb{K},\ n \mapsto u_n$. Par conséquent, $\mathcal{U} = \mathcal{F}(\mathbb{N}, \mathbb{K})$ est l'ensemble des applications de $\mathbb{N}$ dans $\mathbb{K}$. En outre, les lois interne et externe de $\mathcal{F}(\mathbb{N}, \mathbb{K})$ définies dans l'exemple 1.1.11 induisent celles de $\mathcal{U}$. Par conséquent, $\mathcal{U}$ est bien un espace vectoriel sur $\mathbb{K}$.

Exemple 1.1.13. L'espace vectoriel produit.

Soient $n \geq 1$ un entier et $E_1, E_2, \ldots, E_n$ des espaces vectoriels sur le même corps $\mathbb{K}$.

Nous allons généraliser l'exemple 1.1.8 en considérant cette fois-ci le produit catésien $E_1 \times E_2 \times \cdots \times E_n$, noté aussi $\prod_{i=1}^{n} E_i$, qui est l'ensemble des n-uplets $(u_1, \ldots, u_n)$ où u_i est un vecteur de E_i pour tout $i = 1, \ldots, n$. Nous munissons cet ensemble des lois interne et externe suivantes

$$\boxed{(u_1, \ldots, u_n) + (v_1, \ldots, v_n) = (u_1 + v_1, \ldots, u_n + v_n)}$$

$$\boxed{\alpha(u_1, \ldots, u_n) = (\alpha u_1, \ldots, \alpha u_n)}$$

pour tous $(u_1, \ldots, u_n)$ et $(v_1, \ldots, v_n)$ dans $\prod_{i=1}^{n} E_i$ et α dans $\mathbb{K}$.
Il est alors très facile d'établir que $\prod_{i=1}^{n} E_i$ est un espace vectoriel sur $\mathbb{K}$, appelé l'*espace vectoriel produit* des espaces vectoriels $E_1, \ldots, E_n$.
Le vecteur nul de $\prod_{i=1}^{n} E_i$ est $(0_{E_1}, \ldots, 0_{E_n})$, où 0_{E_i} est le vecteur nul de l'espace vectoriel E_i pour chaque i. L'élément symétrique du vecteur $u = (u_1, \ldots, u_n)$ est le vecteur $-u = (-u_1, \ldots, -u_n)$, où $-u_i$ désigne l'élément symétrique du vecteur u_i dans l'espace vectoriel E_i.
Cas particulier : lorsque les espaces $E_1, \ldots, E_n$ sont égaux, c'est-à-dire $E_1 = \cdots = E_n = E$, on obtient l'espace vectoriel produit $E^n = \underbrace{E \times E \times \cdots \times E}_{n \text{ fois}}$.

Il est important de souligner que la notion d'espace vectoriel dépend du corps $\mathbb{K}$. Pour s'en convaincre, considérons l'exemple suivant :

Exemple 1.1.14. On a déjà vu que $\mathbb{C}$ est un $\mathbb{C}$-espace vectoriel. Nous allons voir que $\mathbb{C}$ est aussi un $\mathbb{R}$-espace vectoriel. En effet, gardons l'addition usuelle ordinaire de $\mathbb{C}$ et considérons la loi externe de $\mathbb{C}$ sur le corps $\mathbb{R}$ donnée par : $(\lambda, x) \mapsto \lambda x$ pour tout vecteur $x \in \mathbb{C}$ et pour tout scalaire $\lambda \in \mathbb{R}$. Alors on vérifie aisément que cette loi externe satisfait les quatre axiomes demandés dans la définition d'espace vectoriel. Comme $(\mathbb{C}, +)$ est déjà un groupe abélien, c'est un espace vectoriel sur le corps $\mathbb{R}$.

Soit E un espace vectoriel sur $\mathbb{K}$. Comme conséquences de l'associativité et de la commutativité de l'addition des vecteurs, nous voyons que si $v_1, \ldots, v_n$ sont dans E, alors la somme $v_1 + \cdots + v_n$ est un vecteur bien défini de E, qui peut s'écrire aussi $\sum\limits_{i=1}^{n} v_i$. De plus, on peut vérifier aisément par récurrence :

1. $\alpha(v_1 + \ldots + v_n) = \alpha v_1 + \ldots + \alpha v_n$.
2. $(\alpha_1 + \ldots + \alpha_n)v = \alpha_1 v + \cdots \alpha_n v$.

Proposition 1.1.15. (Règles de calculs dans un espace vectoriel).
Soit E un espace vectoriel sur $\mathbb{K}$. On a :

1. $\forall x \in E \quad 0 \cdot x = 0_E$.

2. $\forall \alpha \in K \quad \alpha \cdot 0_E = 0_E$.

3. *Si $\alpha \cdot x = 0_E$, avec $\alpha \in \mathbb{K}$ et $x \in E$, alors $\alpha = 0$ ou $x = 0_E$.*

4. $\forall \alpha \in \mathbb{K} \quad \forall x \in E \quad -(\alpha \cdot x) = (-\alpha) \cdot x = \alpha \cdot (-x)$.

Démonstration. La preuve est aisée. On n'utilisera que les axiomes d'espace vectoriel vus dans la définition 1.1.3. On s'attachera à bien comprendre le sens des opérations et notations employées.

1. Soit $x \in E$. Montrons que $0 \cdot x = 0_E$. On a

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

c'est-à-dire

$$0 \cdot x = 0 \cdot x + 0 \cdot x.$$

Puisque le vecteur $0 \cdot x$ a pour élément symétrique $-(0 \cdot x)$ dans le groupe abélien $(E, +)$, il suffit alors d'additionner à gauche et à droite dans cette égalité le vecteur $-(0 \cdot x)$ pour simplifier cette égalité. On obtient alors (à l'aide de l'associativité et la commutativité de la loi $+$)

$$\underbrace{-(0 \cdot x) + 0 \cdot x}_{=0_E} = \underbrace{-(0 \cdot x) + 0 \cdot x}_{=0_E} + 0 \cdot x, \text{ d'où } 0_E = 0 \cdot x.$$

2. Soit $\alpha \in K$. Montrons que $\alpha \cdot 0_E = 0_E$. On a

$$\alpha \cdot 0_E = \alpha(0_E + 0_E) = \alpha \cdot 0_E + \alpha \cdot 0_E,$$

c'est-à-dire

$$\alpha \cdot 0_E = \alpha \cdot 0_E + \alpha \cdot 0_E.$$

En ajoutant $-(\alpha \cdot 0_E)$ à gauche et à droite dans cette dernière égalité, on obtient $0_E = \alpha \cdot 0_E$.

3. Soient $\alpha \in \mathbb{K}$ et $x \in E$ tels que $\alpha \cdot x = 0_E$. Montrons que $\alpha = 0$ ou $x = 0_E$. Si $\alpha = 0$, il n'y a rien à montrer. Si $\alpha \neq 0$, nous allons obtenir nécessairement $x = 0_E$. En effet, comme α est un scalaire non nul dans $\mathbb{K}$, on peut considérer son inverse α^{-1} dans le corps $\mathbb{K}$. Multiplions selon la loi externe l'égalié $\alpha \cdot x = 0_E$ à gauche par

le scalaire α^{-1} :
$\alpha^{-1} \cdot (\alpha \cdot x) = \alpha^{-1} \cdot 0_E$. Or, d'après la seconde propriété, $\alpha^{-1} \cdot 0_E = 0_E$. D'autre part,

$$\alpha^{-1} \cdot (\alpha \cdot x) = (\alpha^{-1}\alpha) \cdot x = 1 \cdot x = x.$$

Par conséquent, $x = 0_E$.

4. Soient $\alpha \in \mathbb{K}$ et $x \in E$. Commençons par montrer l'égalité $-(\alpha \cdot x) = (-\alpha) \cdot x$. D'après le première propriété, on a

$$\alpha \cdot x + (-\alpha) \cdot x = (\alpha + (-\alpha)) \cdot x = 0 \cdot x = 0_E,$$

c'est-à-dire $\alpha \cdot x + (-\alpha) \cdot x = 0_E$. Ceci signifie que le vecteur $(-\alpha) \cdot x$ est l'élément symétrique du vecteur $\alpha \cdot x$, c'est-à-dire, $(-\alpha) \cdot x = -(\alpha \cdot x)$.
Montrons maintenant l'égalité $\alpha \cdot (-x) = -(\alpha \cdot x)$. D'après la seconde propriété, on a

$$\alpha \cdot x + \alpha \cdot (-x) = \alpha \cdot (x + (-x)) = \alpha \cdot 0_E = 0_E,$$

c'est-à-dire

$$\alpha \cdot x + \alpha \cdot (-x) = 0_E.$$

On en déduit de la même manière que $\alpha \cdot (-x) = -(\alpha \cdot x)$. ■

Remarque 1.1.16. En prenant $\alpha = 1$ dans la propriété 4, on obtient l'égalité

$$(-1) \cdot x = -x \text{ pour tout vecteur } x \in E.$$

Mais on peut le vérifier directement de façon analogue en écrivant

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

et donc $-x = (-1) \cdot x$.

Remarque 1.1.17. Le fait que $\mathbb{K}$ soit un corps est ici utilisé pour être assuré de l'inversibilité pour la multiplication de tout scalaire non nul $\alpha \in \mathbb{K}$.

1.2 Sous-espaces vectoriels

Comme toujours, lorsqu'on définit une *structure* (groupe, anneau, corps, etc.), on introduit la notion de *sous-structure*. En suivant cet esprit, il est naturel d'introduire la notion de *sous-espace vectoriel*.

Définition 1.2.1. *Soit E un espace vectoriel sur $\mathbb{K}$ et F une partie de E. On dit que F est un* ***sous-espace vectoriel de*** *E si F est non vide et si F est lui-même un espace vectoriel sur $\mathbb{K}$ pour les opérations d'addition et de multiplication externe induites par E.*

La proposition suivante montre qu'il n'est pas nécessaire de vérifier tous les axiomes d'espace vectoriel pour montrer qu'une partie F de E est un sous-espace vectoriel de E. La stabilité par rapport à ces deux opérations d'addition et de multiplication externe suffit.

Proposition 1.2.2. *Soit F une partie d'un espace vectoriel E. Alors F est un sous-espace vectoriel de E si et seulement si :*

1. $0_E \in F$.

2. $\forall x \in F \quad \forall y \in F \quad x + y \in F$.

3. $\forall x \in F \quad \forall \alpha \in \mathbb{K} \quad \alpha \cdot x \in F$.

DÉMONSTRATION.

- Si F est un sous-espace vectoriel de E, c'est un espace vectoriel sur $\mathbb{K}$ pour l'addition et la multiplication externe héritées de E. En particulier, il doit être évidemment stable pour ces deux opérations. De plus, comme $(F, +)$ est un groupe, c'est un sous-groupe du groupe abélen $(E, +)$, et donc contient l'élément neutre 0_E de E. D'où F satisfait nécessairement les trois conditions 1, 2 et 3.
- Réciproquement, supposons que ces trois conditions sont satisfaites dans F, et montrons que $(F, +, \cdot)$ est un espace vectoriel sur $\mathbb{K}$:
  - D'abord, en prenant un vecteur arbitraire $w \in F$ et le scalaire $\alpha = 0$, on obtient par la condition 3 : $\alpha \cdot w = 0 \cdot w = 0_E \in F$.
  - Maintenant, la condition 3 implique que $\forall x \in F,\ -x = (-1) \cdot x \in F$. Donc F est un sous-groupe du groupe abélien $(E, +)$, et par suite, $(F, +)$ est lui-même un groupe abélien.
  - Finalement, les quatre axiomes (a), (b), (c) et (d) de la loi externe dans la définition 1.1.5 sont automatiquement valables dans F, puisqu'elles sont déjà vraies dans E.

    Par conséquent, F est bien un sous-espace vectoriel de E. ■

Plutôt que de séparer la stabilité des loi interne et externe dans la proposition ci-dessus, on utilise souvent la caractérisation suivante qui combine les deux conditions.

Proposition 1.2.3. *Soit F une partie d'un espace vectoriel E. Alors F est un sous-espace vectoriel de E si et seulement si :*

1. $0_E \in F$.

2. $\forall x, y \in F \quad \forall \alpha, \beta \in \mathbb{K} \quad \alpha \cdot x + \beta \cdot y \in F$.

DÉMONSTRATION.

- Si F est un sous-espace vectoriel de E, alors la proposition 1.2.2 nous dit que $\forall x, y \in F \quad \forall \alpha, \beta \in \mathbb{K} \quad \alpha \cdot x \in \mathbb{K}$ et $\beta \cdot y \in F$, donc la somme $\alpha \cdot x + \beta \cdot y \in F$.
- Réciproquement, si F satisfait les deux conditions 1 et 2, alors :
  - En prenant $\alpha = \beta = 1$ dans 2, nous obtenons $x + y \in F$ pour tous x, y dans F.
  - Soit $\alpha \in \mathbb{K}$ et $x \in F$. En choisissant $\beta = 0$ dans 2, il vient $\alpha \cdot x + 0 \cdot y \in F$, d'où $\alpha \cdot x \in F$.

    D'après la proposition précédente, F est un sous-espace vectoriel de E. ■

Remarque 1.2.4. On vérifie facilement les points suivants :

1. Les deux parties $\{0_E\}$ et E constituent deux sous-espaces vectoriels de E, appelés *sous-espaces triviaux* de E. Les autres sous-espaces de E sont dits *non triviaux.*

2. Si F est un sous-espace de E, alors le complémentaire C_E^F de F dans E n'est jamais un sous-espace de E. Il suffit de remarquer que $0_E \notin C_E^F$.

3. On peut étendre la seconde condition de la proposition 1.2.3 à n'importe quel nombre de vecteurs. Plus précisément, à l'aide d'un raisonnement par récurrence simple, on peut établir aisément que si F est un sous-espace vectoriel de E, alors

$$\boxed{\forall n \geq 1 \quad \forall x_1, \ldots, x_n \in F \quad \forall \alpha_1, \ldots, \alpha_n \in \mathbb{K} \quad \alpha_1 \cdot x_1 + \cdots + \alpha_n \cdot x_n \in F}$$

On se souviendra que, pour montrer qu'un ensemble muni d'une loi interne et d'une loi externe est un espace vectoriel, il est souvent commode de prouver que c'est un sous-espace vectoriel d'un espace vectoriel plus large qui est bien connu. C'est un moyen rapide qui évite des calculs lourds et inutiles.

Certains des espaces vectoriels donnés en exemples à la section précédente 1.1 sont peu utilisés car trop généraux. Par exemple, l'espace $\mathbb{K}[X]$ des polynômes de degré quelconque ou encore l'espace des suites à valeurs dans $\mathbb{C}$ sont des espaces trop vastes. Dans la pratique, on est amené à effectuer les calculs en ne manipulant que certains vecteurs de l'espace vectoriel considéré (bien que les calculs soient autorisés pour tous les éléments de l'espace), ce qui permet de travailler avec des sous-espaces vectoriels plus intéressants.

Exemple 1.2.5. Lorsqu'on décide de travailler seulement avec les polynômes de degré inférieur ou égal à un entier donné n, on considère alors le sous-ensemble $\mathbb{K}_n[X]$ de $\mathbb{K}[X]$ défini par $\mathbb{K}_n[X] = \{P \in \mathbb{K}[X] \ / \ \deg(P) \leq n\}$. Cet ensemble est formé par les polynômes P de la forme $P = a_0 + a_1X + \cdots + a_nX^n$ avec $a_0, a_1, \ldots, a_n \in \mathbb{K}$. Cet ensemble contient évidemment le polynôme nul $0_{\mathbb{K}[X]}$ (on rappelle la convention $\deg(0_{\mathbb{K}[X]}) = -\infty$). On constate que si $P = a_0 + a_1X + \cdots + a_nX^n$ et $Q = b_0 + b_1X + \cdots + b_nX^n$ sont deux polynômes dans $\mathbb{K}_n[X]$ et $\alpha \in \mathbb{K}$ alors $P + Q = (a_0 + b_0) + (a_1 + b_1)X + \cdots + (a_n + b_n)X^n$ et $\alpha P = (\alpha a_0) + (\alpha a_1)X + \cdots + (\alpha a_n)X^n$ sont dans $\mathbb{K}_n[X]$. Ainsi, $\mathbb{K}_n[X]$ est un sous-espace vectoriel de l'espace vectoriel $\mathbb{K}[X]$.

Exemple 1.2.6. Soit W le sous-ensemble de $\mathbb{R}^3$ défini par

$$W = \{(x_1, 2x_1 + x_2, x_1 - x_2) \ / \ x_1, x_2 \in \mathbb{R}\}.$$

Nous affirmons que W est un sous-espace vectoriel de $\mathbb{R}^3$. En effet, puisque

$$\begin{aligned}
\alpha(x_1, 2x_1 + x_2, x_1 - x_2) &= (\alpha x_1, \alpha(2x_1 + x_2), \alpha(x_1 - x_2)) \\
&= (\alpha x_1, 2(\alpha x_1) + \alpha x_2, \alpha x_1 - \alpha x_2) \\
&= (z_1, 2z_1 + z_2, z_1 - z_2) \qquad \text{avec } z_1 = \alpha x_1,\ z_2 = \alpha x_2,
\end{aligned}$$

alors W est stable pour la multiplication externe. Pour montrer que W est stable pour l'addition vectorielle, soient $u = (x_1, 2x_1+x_2, x_1-x_2)$ et $v = (y_1, 2y_1+y_2, y_1-y_2)$ des vecteurs de W. Alors

$$\begin{aligned}
u + v &= (x_1 + y_1, 2(x_1 + y_1) + (x_2 + y_2), (x_1 + y_1) - (x_2 + y_2)) \\
&= (t_1, 2t_1 + t_2, t_1 - t_2) \qquad \text{avec } t_1 = x_1 + y_1,\ t_2 = x_2 + y_2.
\end{aligned}$$

Exemple 1.2.7. Nous avons vu dans l'exemple 1.1.14 que, d'une part, $E = \mathbb{C}$ est un espace vectoriel sur le corps $\mathbb{C}$, et d'autre part, $E = \mathbb{C}$ est un espace vectoriel sur le corps $\mathbb{R}$. Ces deux structures d'espaces vectoriels sont différentes. Notons $i\,\mathbb{R} = \{ia \ / \ a \in \mathbb{R}\}$ l'ensemble des nombres complexes imaginaires purs. On a :

- $\mathbb{R}$ est un sous-espace vectoriel du $\mathbb{R}$-espace vectoriel $\mathbb{C}$, car :
  $\forall x, y \in \mathbb{R},\ \forall \alpha, \beta \in \mathbb{R},\ \alpha x + \beta y \in \mathbb{R}$.

- $i\mathbb{R}$ est aussi un sous-espace vectoriel du $\mathbb{R}$-espace vectoriel $\mathbb{C}$, puisque :
  $\forall x = ia \in \mathbb{R}$, $\forall y = ib \in i\,\mathbb{R}$ avec $a, b \in \mathbb{R}$ et $\forall \alpha, \beta \in \mathbb{R}$, $\alpha x + \beta y = i(\alpha a + \beta b) \in i\,\mathbb{R}$.
- Par contre, $\mathbb{R}$ n'est pas un sous-espace vectoriel du $\mathbb{C}$-espace vectoriel $\mathbb{C}$, puisque le produit αx d'un réel $x \in \mathbb{R}$ par un scalaire $\alpha \in \mathbb{C}$ n'est pas toujours dans $\mathbb{R}$.
- De même, $i\mathbb{R}$ n'est pas un sous-espace vectoriel du $\mathbb{C}$-espace vectoriel $\mathbb{C}$, parce que $\alpha = 1 + i \in \mathbb{C}$ et $x = i \in i\mathbb{R}$, mais $\alpha x = (1 + i)i = -1 + i \notin i\mathbb{R}$.

Etudions maintenant quelques exemples analytiques de sous-espaces.

Exemple 1.2.8. Soit $\mathcal{U}$ le $\mathbb{C}$-espace vectoriel des suites complexes $(u_n)_{n\in\mathbb{N}^*}$ que nous avons considéré dans l'exemple 1.1.12. Considérons le sous-ensemble F de $\mathcal{U}$ des suites $(u_n)_{n\in\mathbb{N}^*}$ satisfaisant la relation de récurrence

$$u_{n+2} = u_{n+1} + u_n \qquad \forall n \in \mathbb{N}^* \tag{1.1}$$

Alors F est un sous-espace vectoriel de $\mathcal{U}$. En effet, la suite nulle vérifie trivialement la relation (1.1). Soient (u_n) et (v_n) deux éléments de F et $\alpha, \beta \in \mathbb{K}$. Alors la suite $(w_n)_n = \alpha(u_n)_n + \beta(v_n)_n = (\alpha u_n + \beta v_n)_n$ appartient aussi à F puisque

$$\begin{aligned} w_{n+2} &= \alpha u_{n+2} + \beta v_{n+2} \\ &= \alpha(u_{n+1} + u_n) + \beta(v_{n+1} + v_n) \\ &= (\alpha u_{n+1} + \beta v_{n+1}) + (\alpha u_n + \beta v_n) \\ &= w_{n+1} + w_n. \end{aligned}$$

Observons, en particulier, que ce sous-espace F contient la célèbre *suite de Fibonacci* $(F_n)_{n\in\mathbb{N}^*}$ définie par $F_{n+2} = F_{n+1} + F_n$ avec $F_1 = F_2 = 1$.

Exemple 1.2.9. Soit $a < b$ deux réels et notons $\mathcal{C}([a, b], \mathbb{R})$ l'ensemble des fonctions $f : [a, b] \longrightarrow \mathbb{R}$ continues sur $[a, b]$. Alors $\mathcal{C}([a, b], \mathbb{R})$ est un espace vectoriel sur $\mathbb{R}$. En effet, si l'on prend $X = [a, b]$ et $\mathbb{K} = \mathbb{R}$ dans l'exemple 1.1.11, on peut considérer le $\mathbb{R}$-espace vectoriel $\mathcal{F}([a, b], \mathbb{R})$ des fonctions $f : [a, b] \longrightarrow \mathbb{R}$. Puisque $\mathcal{C}([a, b], \mathbb{R})$ est inclus dans $\mathcal{F}([a, b], \mathbb{R})$, il suffit de prouver que $\mathcal{C}([a, b], \mathbb{R})$ est un sous-espace vectoriel de $\mathcal{F}([a, b], \mathbb{R})$. Pour cela, il est bien connu en Analyse que si f et g sont des fonctions continues sur $[a, b]$ et $\alpha \in \mathbb{R}$, il en est de même pour les fonctions $f + g$ et αf.

Exemple 1.2.10. L'ensemble F des fonctions réélles de classe C^2 de $\mathbb{R}$ dans $\mathbb{R}$ vérifiant l'équation différentielle $2f''(x) - 3f'(x) + f(x) = 0$ est un sous-espace vectoriel de l'espace vectoriel $\mathcal{F}(\mathbb{R},\mathbb{R})$. En effet, si f et g sont deux éléments de F et $\alpha,\beta\in\mathbb{R}$, alors la fonction $\alpha f+\beta g$ est également de classe C^2. De plus,

$$\begin{aligned}2(\alpha f+\beta g)''(x)-3(f+g)'(x)+(f+g)(x)\\ &=2(\alpha f''(x)+\beta g''(x))-3(f'(x)+g'(x))+(f(x)+g(x))\\ &=\alpha(2f''(x)-3f'(x)+f(x))+\beta(2g''(x)-3g'(x)+g(x))\\ &=0,\end{aligned}$$

ce qui montre bien que $\alpha f+\beta g\in F$.

Exemple 1.2.11. Soit $E=\mathbb{R}$ et $F=]0,1[$. Il n'est pas pénible de vérifier que F n'est pas un sous-espace de E. Par exemple, les éléments $x=1/2$ et $y=3/4$ appartiennent à F, mais leur somme $x+y=1/2+3/4=5/4\notin F$.

On peut faire des opérations sur les sous-espaces vectoriels d'un espace vectoriel E pour en fabriquer d'autres. Voici une première méthode pour le faire :

Théorème 1.2.12. *L'intersection d'une famille quelconque de sous-espaces vectoriels de E est un sous-espace vectoriel de E.*

Démonstration. Soit $(F_i)_{i\in I}$ une famille de sous-espaces vectoriels de E et $F=\bigcap_{i\in I} F_i=\{x\in E\ /\ \forall i\in I: x\in F_i\}$ leur intersection. Comme $0_E\in F_i$ pour tout $i\in I$, alors $0_E\in F$. Soient $x,y\in F$ et $\alpha,\beta\in\mathbb{K}$. Pour tout $i\in I$, nous avons $x,y\in F_i$ et donc $\alpha x+\beta y\in F_i$ car F_i est un sous-espace vectoriel de E. Il en résulte que $\alpha x+\beta y\in F$, ce qui achève la preuve. ■

Remarque 1.2.13. Le résultat précédent est valable pour une famille finie ou infinie de sous-espaces vectoriels $(F_i)_{\in I}$. On travaillera souvent avec des intersections finies $F_1\cap F_2\cdots\cap F_n$, et notamment l'intersection $F_1\cap F_2$ de deux sous-espaces vectoriels. Remarquons que deux sous-espaces vectoriels F_1 et F_2 d'un même espace vectoriel E ne sont jamais disjoints, puisque $0_E\in F_1\cap F_2$.

Exemple 1.2.14. Soient $F=\{(0,y,z)\ /\ y,z\in\mathbb{R}\}$ et $G=\{(x,y,0)\ /\ x,y\in\mathbb{R}\}$. On vérifie aisément que F et G sont deux sous-espaces vectoriels de $\mathbb{R}^3$. Pour déterminer le sous-espace vectoriel $F\cap G$, prenons $u=(x,y,z)\in F\cap G$. Comme $u\in F$, il vient $x=0$ et puisque $u\in G$, alors $z=0$, d'où $u=(0,y,0)$. Ainsi

$$F\cap G=\{(0,y,0)\ /y\in\mathbb{R}\}.$$

Exemple 1.2.15. Soient F et G les sous-ensembles du $\mathbb{R}$-espace vectoriel $\mathcal{F}(\mathbb{R},\mathbb{R})$ constitués des fonctions respectivement paires et impaires :

$$F=\{f\in\mathcal{F}(\mathbb{R},\mathbb{R})\ /\ f(-x)=f(x)\ \forall x\in\mathbb{R}\},$$

$$G = \{f \in \mathcal{F}(\mathbb{R}, \mathbb{R}) \; / \; f(-x) = -f(x) \quad \forall x \in \mathbb{R}\}.$$

On vérifie facilement que F et G sont des sous-espaces vectoriels de $\mathcal{F}(\mathbb{R}, \mathbb{R})$. Soit $f \in F \cap G$. D'une part, $f \in F$ implque $f(-x) = f(x)$ pour tout $x \in \mathbb{R}$. D'autre part, $f \in G$ entraîne $f(-x) = -f(x)$ pour tout $x \in \mathbb{R}$. On en déduit que pour tout $x \in \mathbb{R}, f(x) = -f(x)$, d'où $f(x) = 0$ et par suite f est l'application nulle 0. On a donc $F \cap G = \{0\}$.

Exemple 1.2.16. Pour tout entier $n \geq 1$, on note $C^n(\mathbb{R}, \mathbb{R})$ l'ensemble des fonctions n fois dérivables de $\mathbb{R}$ dans $\mathbb{R}$. Il est facile de voir que $C^n(\mathbb{R}, \mathbb{R})$ est un sous-espace vectoriel de l'espace vectoriel $\mathcal{F}(\mathbb{R}, \mathbb{R})$. D'après la proposition précédente, l'intersection $\bigcap_{i \in I} C^n(\mathbb{R}, \mathbb{R})$ de tous ces sous-espaces $C^n(\mathbb{R})$ est un sous-espace vectoriel de $\mathcal{F}(\mathbb{R}, \mathbb{R})$. En outre, une fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$ appartient à $\bigcap_{i \in I} C^n(\mathbb{R}, \mathbb{R})$ si et seulement si f est n fois dérivable sur $\mathbb{R}$ pour tout $n \geq 1$. Par conséquent, cette intersection est précisément le sous-espace $C^\infty(\mathbb{R}, \mathbb{R})$ des fonctions infiniment dérivables de $\mathbb{R}$ dans $\mathbb{R}$.

L'exemple qui suit montre que, contrairement à l'intersection, l'union d'une famille de sous-espaces vectoriels de E n'est pas en général un sous-espace vectoriel de E. En fait, ceci provient tout simplement du fait que l'union d'une famille de sous-groupes d'un groupe G n'est pas en général un sous-groupe de G. Remarquer néanmoins que la loi externe ne pose pas de problème : il est facile de vérifier que l'union $F = \bigcup_{i \in I} F_i = \{x \in E \; / \; \exists i \in I : x \in F_i\}$ de sous-espaces vecoriels F_i est stable pour la loi externe.

Exemple 1.2.17. Considérons les deux sous-espaces vectoriels F et G d' l'espace vectoriel $\mathbb{R}^2$ définis par :

$$F = \{(x, 0) \; / \; x \in \mathbb{R}\}, \qquad G = \{(0, y) \; / \; y \in \mathbb{R}\}.$$

L'ensemble $F \cup G$ n'est pas un sous-espace vectoriel de $\mathbb{R}^2$. En effet, puisque $(1, 0) \in F$ et $(0, 1) \in G$, alors $(1, 0) \in F \cup G$ et $(0, 1) \in F \cup G$. Cependant, la somme $(1, 0) + (0, 1) = (1, 1) \notin F \cup G$ car $(1, 1) \notin F$ et $(1, 1) \notin G$.

Nous introduirons plus loin une notion de *somme* de sous-espaces vectoriels qui remplacera le concept de la réunion de sous-espaces vectoriels.

1.3 Parties génératrices, parties libres, bases

Dans cette section, nous allons voir comment une petite partie de vecteurs peut totalement déterminer la structure entière d'un espace vectoriel, et nous allons considérer les conséquences de cette observation.

1.3.1 Parties génératrices

Nous commençons maintenant par intoduire une méthode très importante de construction de sous-espaces vectoriels d'un espace vectoriel donné E. Pour cela, nous avons besoin d'abord de la notion suivante :

Définition 1.3.1. *Soient E un $\mathbb{K}$-espace vectoriel et $v_1, \ldots, v_n$ des vecteurs de E. On appelle **combinaison linéaire** de $v_1, \ldots, v_n$ tout vecteur x de E de la forme $x = \alpha_1 v_1 + \cdots + \alpha_n v_n$ avec $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$.*

Avant de continuer notre démarche, regardons quelques exemples spécifiques de combinaisons linéaires.

Exemple 1.3.2. **1.** $0_E = 0 \cdot v_1 + \cdots + 0 \cdot v_n$ est une combinaison linéaire particulière de $v_1, \ldots, v_n$.

2. Dans $\mathbb{R}^2$, tout vecteur $v = (x, y)$ est combinaison linéaire de $e_1 = (1, 0)$ et $e_2 = (0, 1)$, puisque $(x, y) = (x, 0 + (0, y) = x(1, 0) + y(0, 1) = xe_1 + ye_2$.

3. Plus généralement, dans l'espace vectoriel $\mathbb{K}^n$ sur $\mathbb{K}$, tout vecteur $x = (x_1, \ldots, x_n)$ peut se décomposer sous forme :

$$\begin{aligned} x &= (x_1, 0, \ldots, 0) + (0, x_2, 0, \ldots, 0) + \cdots + (0, \ldots, 0, x_n) \\ &= x_1(1, 0, \ldots, 0) + x_2(0, 1, 0, \ldots, 0) + \cdots + x_n(0, \ldots, 0, 1). \end{aligned}$$

Donc x est combinaison linéaire des vecteurs :

$$e_1 = (1, 0, \ldots, 0), \quad e_2 = (0, 1, 0, \ldots, 0), \quad \ldots e_n = (0, \ldots, 0, 1).$$

4. Dans le $\mathbb{R}$-espace vectoriel $\mathbb{C}$, tout complexe $z = a + ib$ est combinaison linéaire de 1 et i.

5. Tout polynôme $P = a_0 + a_1 X + \cdots + a_n X^n$ dans $\mathbb{K}_n[X]$ est combinaison linéaire des monômes $1, X, X^2, \ldots, X^n$.

6. Dans l'espace vectoriel $\mathcal{F}(\mathbb{R}, \mathbb{R})$ décrit dans l'exemple 1.1.11, la fonction $f : x \mapsto \cos(2x)$ est combinaison linéaire des fonctions $g : x \mapsto 1$ et $h : x \mapsto \cos^2(x)$, puisque $f = 2h - g$.

7. Dans $\mathbb{R}^3$, le vecteur $v = (3, -5, 0)$ est combinaison linéaire des deux vecteurs $v_1 = (1, -1, 2), v_2 = (0, 1, 3)$ parce que $v = 3v_1 - 2x_2$.

8. Dans $\mathbb{R}^3$, le vecteur $w = (1, 1, 1)$ n'est pas combinaison linéaire des vecteurs $u_1 = (1, -1, 2)$ et $u_2 = (0, 1, 3)$. En effet, si c'était le cas, il existerait deux scalaires α et β dans $\mathbb{R}$ tels que $w = \alpha u_1 + \beta u_2$. On obtient alors le système

$$\begin{cases} \alpha = 1 & (1) \\ -\alpha + \beta = 1 & (2) \\ 2\alpha + 3\beta = 1 & (3) \end{cases}$$

Or, les équations (1) et (2) donnent $\alpha = 1$ et $\beta = 2$. En remplaçant ces valeurs dans (3), on trouve $8 = 1$, ce qui est absurde. Le système n'a donc aucune solution (α, β).

Remarque 1.3.3. Soit F un sous-espace vectoriel d'un espace vectoriel E. On a vu dans la remarque 1.2.4 que si $v_1, \ldots, v_n \in F$, alors toutes les combinaisons linéaires $\alpha_1 \cdot v_1 + \cdots + \alpha_n \cdot v_n$ de $v_1, \ldots, v_n$ sont aussi dans F. En d'autres termes, F est *stable par combinaisons linéaires.*

Nous allons caractériser une famille de vecteurs d'un espace vectoriel dont les combinaisons linéaires permettent de construire tous les autres vecteurs de l'espace.

Définition 1.3.4. *Soient E un espace vectoriel sur $\mathbb{K}$ et $v_1, \ldots, v_n$ des vecteurs de E. On dit que les vecteurs $v_1, \ldots, v_n$* ***engendrent*** *E si tout élément de E est combinaison linéaire de $v_1, \ldots, v_n$, ce qui peut s'écrire :*

$$\forall x \in E \quad \exists \alpha_1, \ldots, \alpha_n \in \mathbb{K} : \quad x = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

Vocabulaire. *On dit aussi que $(v_1, \ldots, v_n)$ est une* ***famille génératrice de*** *E, ou que $(v_1, \ldots, v_n)$ est un* ***système de générateurs de*** *E. On dit également que $\{v_1, \ldots, v_n\}$ est une* ***partie génératrice de*** *E et que l'espace vectoriel E est* ***engendré par*** *$\{v_1, \ldots, v_n\}$.*

D'après l'exemple précédent, on peut déduire :

Exemple 1.3.5. **1.** Les vecteurs $e_1 = (1, 0, \ldots, 0),\ e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)$ engendrent l'espace $\mathbb{K}^n$.

2. $(1, X, X^2, \ldots, X^n)$ est une famille génératrice de $\mathbb{K}_n[X]$.

3. $(1, i)$ est une famille génératrice du $\mathbb{R}$-espace vectoriel $\mathbb{C}$.

4. Les vecteurs $v_1 = (1, -1, 2), v_2 = (0, 1, 3)$ n'engendrent pas $\mathbb{R}^3$, puisqu'il existe au moins un vecteur, à savoir $(1, 1, 1)$, qui n'est pas combinaison linéaire de v_1, v_2 (exemple 1.3.2 (8)).

Proposition 1.3.6. *Soit $\mathcal{F} = (v_1, \ldots, v_p)$ une famlle de vecteurs de E et $\mathcal{F}' = (v_1, \ldots, v_p, v_{p+1}, \ldots, v_n)$ une sur-famille de $\mathcal{F}$. Si $\mathcal{F}$ engendre E, alors $\mathcal{F}'$ engendre E.*

Autrement dit : toute sur-famille d'une famille génératrice de E est encore génératrice de E.

DÉMONSTRATION. Soit $x \in E$. Puisque $\mathcal{F}$ engendre E, il existe $\alpha_1, \ldots, \alpha_p$ dans $\mathbb{K}$ tels que $x = \alpha_1 v_1 + \cdots + \alpha_p v_p$. En notant $\alpha_{p+1} = \cdots = \alpha_n = 0$, on a alors $x = \alpha_1 v_1 + \cdots + \alpha_p v_p + \alpha_{p+1} v_{p+1} + \cdots + \alpha_n v_n$. Ceci montre que $\mathcal{F}'$ engendre aussi E. ∎

Remarque 1.3.7. La contraposée de cette proposition exprime que si $\mathcal{F}'$ n'engendre pas E, alors aucune sous-famille $\mathcal{F}$ de $\mathcal{F}'$ n'engendre E.

Lorsqu'une famille de vecteurs $\mathcal{F} = (v_1, \ldots, v_p)$ n'engendre pas l'espace vectoriel E tout entier, est-ce qu'elle peut engendrer un certain sous-espace vectoriel F de E ? Si un tel sous-espace F existe, tout élément de F sera combinaison linéaire de $v_1, \ldots, v_p$. Cette question nous mène à l'étude de l'ensemble des vecteurs qui peuvent être représentés comme combinaisons linéaires d'une famille donnée de vecteurs $v_1, \ldots, v_p$. On a le résultat suivant.

Théorème 1.3.8. *Soient $\mathcal{F} = \{v_1, \ldots, v_n\}$ une partie finie d'un espace vectoriel E. On note $\mathsf{Vect}(\mathcal{F})$, ou encore $\mathsf{Vect}(v_1, \cdots, v_n)$, l'ensemble de toutes les combinaisons linéaires des vecteurs $v_1, \ldots, v_n$:*

$$\boxed{\mathsf{Vect}(v_1, \cdots, v_n) = \{\alpha_1 v_1 + \cdots + \alpha_n v_n \ / \ \alpha_1, \ldots, \alpha_n \in \mathbb{K}\}}$$

On a les assertions suivantes :

1. $\mathsf{Vect}(\mathcal{F})$ *est un sous-espace vectoriel de E.*
2. $\mathsf{Vect}(\mathcal{F})$ *contient $\mathcal{F}$.*
3. *Si G est un sous-espace vectoriel de E contenant $\mathcal{F}$, alors G contient* $\mathsf{Vect}(\mathcal{F})$.

DÉMONSTRATION.

1. Le vecteur nul 0_E de E appartient à $\mathsf{Vect}(\mathcal{F})$ puisque $0_E = 0 \cdot v_1 + \ldots + 0 \cdot v_n$. Soient x et y deux vecteurs de $\mathsf{Vect}(\mathcal{F})$. Il existe des scalaires $\alpha_1, \ldots, \alpha_n$ et

$\beta_1, \ldots, \beta_n$ dans $\mathbb{K}$ tels que $x = \alpha_1 v_1 + \cdots + \alpha_p v_n$ et $y = \beta_1 v_1 + \cdots + \beta_p v_n$. Alors $x + y = (\alpha_1 + \beta_1)v_1 + \cdots + (\alpha_n + \beta_n)v_n \in \mathsf{Vect}(\mathcal{F})$.
De plus, pour tout $\lambda \in \mathbb{K}$, $\lambda x = (\lambda\alpha_1)v_1 + \ldots + (\lambda\alpha_n)v_n \in \mathsf{Vect}(\mathcal{F})$. D'où $\mathsf{Vect}(\mathcal{F})$ est un sous-espace vectoriel de E.

2. Il est clair que

$$\begin{array}{rcl} v_1 & = & 1 \cdot v_1 + 0 \cdot v_2 + \cdots + 0 \cdot v_n, \\ v_2 & = & 0 \cdot v_1 + 1 \cdot v_2 + 0 \cdot v_3 + \cdots + 0 \cdot v_n, \\ \vdots & & \quad\quad \vdots \\ v_n & = & 0 \cdot v_1 + 0 \cdot v_2 + \cdots + 0 \cdot v_{n-1} + 1 \cdot v_n. \end{array}$$

C'est-à-dire, $v_1, \ldots, v_n$ sont des combinaisons linéaires particulières de $v_1, \ldots, v_n$. D'où $v_1, \ldots, v_n$ appartiennent tous à $\mathsf{Vect}(v_1, \cdots, v_n)$.

3. Soit G un sous-espace vectoriel de E contenant $\mathcal{F} = \{v_1, \ldots, v_n\}$. Tout vecteur $x \in \mathsf{Vect}(v_1, \cdots, v_n)$ est de la forme $x = \alpha_1 v_1 + \cdots + \alpha_p v_n$. Comme $v_1 \in G, \ldots, v_n \in G$ et G est un sous-espace vectoriel de E, il vient $x \in G$. Nous avons ainsi prouvé que $\mathsf{Vect}(v_1, \cdots, v_n) \subseteq G$. ■

Remarque 1.3.9. - Nous pouvons formuler d'une autre manière le théorème précédent en disant que $\mathsf{Vect}(\mathcal{F})$ est le plus petit (pour l'inclusion) sous-espace vectoriel de E contenant $\mathcal{F}$.
- Si $\mathcal{F} = (v)$ contient un seul vecteur v, alors $\mathsf{Vect}(v) = \{\alpha v \ / \ \alpha \in \mathbb{K}\}$, qui se note aussi .
- En particulier, pour $x = 0_E$, nous avons $\mathsf{Vect}(0_E) = \{0_E\}$.
- La famille $\mathcal{F}$ peut être vide, et dans ce cas on écrit par convention $\mathsf{Vect}(\emptyset) = \{0_E\}$.

Définition 1.3.10. $\mathsf{Vect}(\mathcal{F})$ *s'appelle le* ***sous-espace vectoriel de*** E ***engendré par*** *la famille* $\mathcal{F}$. *On dit que les vecteurs* $v_1, \ldots, v_n$ ***engendrent*** *le sous-espace* $\mathsf{Vect}(\mathcal{F})$.
Le vocabulaire de la définition 1.3.4 est encore valable ici en remplaçant l'espace E par le sous-espace $\mathsf{Vect}(\mathcal{F})$. C'est-à-dire, on peut dire aussi que $(v_1, \ldots, v_n)$ *est une* ***famille génératrice de*** $\mathsf{Vect}(\mathcal{F})$*, etc.*

Remarque 1.3.11. - Soit $v_1, \ldots, v_n$ des vecteurs d'un espace vectoriel E et soit F un sous-espace vectoriel de E. Alors :

$$\boxed{v_1, \ldots, v_n \text{ engendrent } F \Leftrightarrow F = \mathsf{Vect}(v_1, \cdots, v_n)}$$

C'est en particulier vrai pour $F = E$ tout entier.

- Si les vecteurs $v_1, \ldots, v_n$ engendrent F, alors $v_1, \ldots, v_n$ doivent appartenir à F. L'exemple suivant illustre ce point crucial.

Exemple 1.3.12. Prenons le sous-espace $W = \{(x_1, 2x_1 + x_2, x_1 - x_2) \ / \ x_1, x_2 \in \mathbb{R}\}$ de $\mathbb{R}^3$ traité dans l'exemple 1.2.6. Cherchons une partie génératrice de W.

• Tout vecteur $x = (x_1, 2x_1 + x_2, x_1 - x_2) \in W$ s'écrit :

$$\begin{aligned} x &= (x_1, 2x_1, x_1) + (0, x_2, -x_2) \\ &= x_1(1, 2, 1) + x_2(0, 1, -1) \end{aligned}$$

qui est donc combinaison linéaire de $v_1 = (1, 2, 1)$ et $v_2 = (0, 1, -1)$. Remarquons que v_1 et v_2 appartiennent bien à W, puisque $v_1 = (z_1, 2z_1 + z_2, z_1 - z_2)$ avec $z_1 = 1, z_2 = 0$, tandis que $v_2 = (t_1, 2t_1 + t_2, t_1 - t_2)$ avec $t_1 = 0, t_2 = 1$. D'où $W = \mathsf{Vect}(v_1, v_2)$, c'est-à-dire, (v_1, v_2) est une famille génératrice de W.

• Décomposons maintenant le vecteur $x = (x_1, 2x_1 + x_2, x_1 - x_2) \in W$ d'une autre façon qui ne sera pas convenable :

$$\begin{aligned} x &= (x_1, 2x_1, 0) + (0, 0, x_1) + (0, x_2, 0) + (0, 0, -x_2) \\ &= x_1(1, 2, 0) + x_1(0, 0, 1) + x_2(0, 1, 0) + x_2(0, 0, -1). \end{aligned}$$

Ceci nous dit que x est aussi combinaison linéaire des vecteurs $u_1 = (1, 2, 0)$, $u_2 = (0, 0, 1)$, $u_3 = (0, 1, 0)$, $u_4 = (0, 0, -1)$. Cependant, les vecteurs u_1, u_2, u_3, u_4 n'appartiennent pas à W. En effet, si par exemple $u_1 \in W$, il existera $x_1, x_2 \in \mathbb{R}$ tels que $u_1 = (x_1, 2x_1 + x_2, x_1 - x_2)$, ce qui donne le système

$$\begin{cases} x_1 = 1 \\ 2x_1 + x_2 = 2 \\ x_1 - x_2 = 0 \end{cases}$$

qui est impossible (la seconde équation implique $x_2 = 0$ et la troisième équation donne $x_2 = 1$). Par conséquent, $W \neq \mathsf{Vect}(u_1, u_2, u_3, u_4)$. En fait, on a seulement l'inclusion stricte $W \subset \mathsf{Vect}(u_1, u_2, u_3, u_4)$, puisque tout vecteur de W est combinaison linéaire de u_1, u_2, u_3, u_4.

Proposition 1.3.13. *Pour tout sous-espace vectoriel G de E et pour toute famille $\mathcal{F} = (v_1, \ldots, v_n)$ de vecteurs de E, on a :*

$$\boxed{\mathsf{Vect}(\mathcal{F}) \subseteq G \Leftrightarrow \mathcal{F} \subseteq G}$$

DÉMONSTRATION. Si $\mathsf{Vect}(\mathcal{F}) \subseteq G$, alors puisque $\mathcal{F} \subseteq \mathsf{Vect}(\mathcal{F})$, il vient $\mathcal{F} \subseteq G$. La réciproque est exactement l'assertion 3 du théorème 1.3.8. ■

L'intérêt des familles génératrices réside dans le fait que, pour comparer des sous-espaces vectoriels d'un espace vectoriel E, il suffit de travailler avec les familles génératrices. En suivant cette approche, on peut énoncer le résultat ci-dessous.

Proposition 1.3.14. *Soient $\mathcal{F}$ et $\mathcal{F}'$ deux familles de vecteurs de E. On a :*

1. $\mathcal{F} \subseteq \mathcal{F}' \Rightarrow \mathsf{Vect}(\mathcal{F}) \subseteq \mathsf{Vect}(\mathcal{F}')$.

2. $\mathsf{Vect}(\mathcal{F}) \subseteq \mathsf{Vect}(\mathcal{F}') \Leftrightarrow$ *tout vecteur de $\mathcal{F}$ est combinaison linéaire de $\mathcal{F}'$.*

DÉMONSTRATION.

1. Soit le sous-espace vectoriel $G = \mathsf{Vect}(\mathcal{F}')$; on a donc $\mathcal{F}' \subseteq G$. Si $\mathcal{F} \subseteq \mathcal{F}'$, alors $\mathcal{F} \subseteq G$. Donc par la remarque 1.3.11, on tire que $\mathsf{Vect}(\mathcal{F}) \subseteq G$.

[On peut le montrer d'une autre manière en utilisant le raisonnement direct suivant : Si $\mathcal{F} \subseteq \mathcal{F}'$, on peut écrire $\mathcal{F} = (v_1, \ldots, v_p)$ et $\mathcal{F}' = (v_1, \ldots, v_p, v_{p+1}, \ldots, v_n)$. Tout vecteur $x \in \mathsf{Vect}(\mathcal{F})$ peut s'écrire sous la forme $x = \alpha_1 \cdot v_1 + \ldots + \alpha_p \cdot v_p$, et donc $x = \alpha_1 \cdot v_1 + \ldots + \alpha_p \cdot v_p + 0 \cdot v_{p+1} + \cdots + 0 \cdot v_n \in \mathsf{Vect}(\mathcal{F}')$. D'où $\mathsf{Vect}(\mathcal{F}) \subseteq \mathsf{Vect}(\mathcal{F}')$]

2.
- Supposons que $\mathsf{Vect}(\mathcal{F}) \subseteq \mathsf{Vect}(\mathcal{F}')$. Alors tout vecteur $x \in \mathsf{Vect}(\mathcal{F})$ appartient à $\mathsf{Vect}(\mathcal{F}')$. En particulier, puisque $\mathcal{F} \subseteq \mathsf{Vect}(\mathcal{F})$, alors tout vecteur $x \in \mathcal{F}$ appartient à $\mathsf{Vect}(\mathcal{F}')$, et c'est donc une combinaison linéaire de $\mathcal{F}'$.
- Réciproquement, supposons que tout vecteur de $\mathcal{F}$ est combinaison linéaire de $\mathcal{F}'$. Alors tout vecteur de $\mathcal{F}$ appartient à $\mathsf{Vect}(\mathcal{F}')$, c'est-à-dire, $\mathcal{F} \subseteq G = \mathsf{Vect}(\mathcal{F}')$. En appliquant la remarque 1.3.11, on obtient $\mathsf{Vect}(\mathcal{F}) \subseteq G = \mathsf{Vect}(\mathcal{F}')$.

  [Voici une autre preuve de cette réciproque : supposons que tout vecteur de $\mathcal{F}$ est combinaison linéaire de $\mathcal{F}'$. On peut donc écrire :

$$\begin{array}{rcl} v_1 & = & \alpha_{11}u_1 + \alpha_{12}u_2 + \cdots + \alpha_{1n}u_n \\ v_2 & = & \alpha_{21}u_1 + \alpha_{22}u_2 + \cdots + \alpha_{2n}u_n \\ \vdots & & \vdots \\ v_p & = & \alpha_{p1}u_1 + \alpha_{p2}u_2 + \cdots + \alpha_{pn}u_n \end{array}$$

Soit maintenant x un vecteur quelconque de $\mathsf{Vect}(\mathcal{F})$. Il existe des scalaires $\lambda_1, \dots, \lambda_p \in \mathbb{K}$ tels que $x = \lambda_1 v_1 + \cdots + \lambda_p v_p$. En remplaçant $v_1, \dots, v_p$ par leurs combinaisons linéaires de $u_1, \dots, u_n$, on voit facilement que x est combinaison linéaire de $u_1, \dots, u_n$:

$$x = \Big(\lambda_1\alpha_{11}+\lambda_2\alpha_{21}+\cdots+\lambda_p\alpha_{p1}\Big)u_1+\cdots+\Big(\lambda_1\alpha_{1n}+\lambda_2\alpha_{2n}+\cdots+\lambda_p\alpha_{pn}\Big)u_n.$$

Donc $x \in \mathsf{Vect}(\mathcal{F}')$. Ce raisonnement signifie qu'*une combinaison linéaire de combinaisons linéaires de $u_1, \dots, u_n$ est encore une combinaison linéaire de $u_1, \dots, u_n$.*] ■

Comme conséquence immédiate, on déduit :

Corollaire 1.3.15. *Soient $\mathcal{F}$ et $\mathcal{F}'$ deux familles de vecteurs de E. On a :*
$\mathsf{Vect}(\mathcal{F}) = \mathsf{Vect}(\mathcal{F}') \Leftrightarrow$ *tout vecteur de $\mathcal{F}$ est combinaison linéaire de $\mathcal{F}'$, et tout vecteur de $\mathcal{F}'$ est combinaison linéaire de $\mathcal{F}$.*

Exemple 1.3.16. Considérons dans $\mathbb{R}^3$ le vecteurs :

$$u = (1,0,1),\ v = (2,3,-1),\ x = (1,3,-2),\ y = (3,3,0),\ z = (0,-3,3).$$

Les deux familles $\{u, v\}$ et $\{x, y, z\}$ engendrent le même sous-espace vectoriel de $\mathbb{R}^3$. En effet,

- On remarque que $x = v - u,\ y = u + v,\ z = -v + 2u$, ce qui prouve que $\mathsf{Vect}(x, y, z) \subseteq \mathsf{Vect}(u, v)$.
- Réciproquement, puisque $u = \frac{1}{2}(y - x)$ et $v = 2x + z$, alors $\mathsf{Vect}(u, v) \subseteq \mathsf{Vect}(x, y, z)$.

En tenant compte des résultats précédents, on obtient en particulier la version suivante qui sera souvent utilisée dans la pratique :

Corollaire 1.3.17. *Soient E un espace vectoriel sur K et $v_1, \dots, v_p, v_{p+1}$ des vecteurs de E. Si v_{p+1} est combinaison linéaire de $v_1, \dots, v_p$, alors*

$$\mathsf{Vect}(v_1, \dots, v_p, v_{p+1}) = \mathsf{Vect}(v_1, \dots, v_p).$$

Autrement dit, le sous-espace vectoriel engendré par une famille de vecteurs est inchangé lorsqu'on augmente cette famille d'une combinaison linéaire de ses vecteurs.

Démonstration. Posons $\mathcal{F} = \{v_1, \dots, v_p\}$ et $\mathcal{F}' = \{v_1, \dots, v_p, v_{p+1}\}$. Puisque $\mathcal{F} \subseteq \mathcal{F}'$, la proposition 1.3.14 (1) implique $\mathsf{Vect}(\mathcal{F}) \subseteq \mathsf{Vect}(\mathcal{F}')$. Maintenant, comme tout vecteur de $\mathcal{F}'$ est combinaison linéaire de $\mathcal{F}$, on déduit de la proposition 1.3.14 (2) que $\mathsf{Vect}(\mathcal{F}') \subseteq \mathsf{Vect}(\mathcal{F})$. D'où le résultat. ■

Ce corollaire permet de réduire une famille génératrice $\mathcal{F}$ d'un sous-espace vectoriel F de E s'il y a "trop" de vecteurs dans $\mathcal{F}$. Plus précisément, si l'un des vecteurs de $\mathcal{F}$ est combinaison linéaire des autres vecteurs de $\mathcal{F}$, on peut l'éliminer de cette famille. Nous reviendrons plus loin à cette observation lorsque nous introduirons le *rang* d'une famille de vecteurs.

Remarque 1.3.18. Soit $\mathcal{F} = \{v_1, \cdots, v_n\}$ une partie génératrice d'un espace vectoriel E. Si $u \in E$, il s'écrit sous la forme d'une combinaison linéaire $u = \alpha_1 v_1 + \cdots + \alpha_n v_n$ de $v_1, \cdots, v_n$. Il n'est pas sûr que les scalaires $\alpha_1, \ldots, \alpha_n$ soient uniques.
Par exemple, $\{1, X, X^2, X^3, X^2 - X\}$ est une partie génératrice de $\mathbb{K}_n[X]$, car tout polynôme $P \in \mathbb{K}_n[X]$ s'écrit

$$P(X) = a_0 \cdot 1 + a_1 \cdot X + a_2 \cdot X^2 + a_3 \cdot X^3 + 0 \cdot (X^2 - X).$$

Le polynôme X^2 s'écrit de deux manières différentes comme combinaison linéaire de $\mathcal{F}$:

$$X^2 = 0\cdot 1 + 0\cdot X + 1\cdot X^2 + 0\cdot X^3 + 0\cdot(X^2 - X) = 0\cdot 1 + 1\cdot X + 0\cdot X^2 + 0\cdot X^3 + 1\cdot(X^2 - X).$$

Soit donc $\mathcal{F} = \{v_1, \cdots, v_n\}$ une partie génératrice d'un espace vectoriel E. Si l'on veut que tout vecteur $u \in E$ se décompose de manière unique comme combinaison linéaire $u = \alpha_1 v_1 + \cdots + \alpha_n v_n$ de $\mathcal{F}$, il faut que ceci soit en particulier vrai pour la décomposition du vecteur 0_E sur $v_1, \ldots, v_n$: $0_E = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Mais évidemment $0_E = 0 \cdot v_1 + \cdots + 0 \cdot v_n$. L'unicité de la décomposition de 0_E imposera donc nécessairement $\alpha_1 = \cdots = \alpha_n = 0$. Ceci nous mène de façon naturelle au concept de *partie libre*, qui est l'objet de la section suivante.

1.3.2 Parties libres

Soient E un espace vectoriel sur $\mathbb{K}$ et $\mathcal{F} = \{v_1, \ldots, v_p\}$ une partie finie de vecteurs de E. Le vecteur 0_E appartient évidemment au sous-espace vectoriel $\mathsf{Vect}(\mathcal{F})$ de E, puisque 0_E s'écrit comme combinaison linéaire triviale $0_E = 0 \cdot v_1 + \cdots + 0 \cdot v_n$. En connexion avec la remarque précédente, une question naturelle se pose : est-ce que le vecteur 0_E peut se décomposer d'une autre façon comme combinaison linéaire de $v_1, \ldots, v_n$? C'est-à-dire, est-ce qu'on peut trouver d'autres scalaires $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tels que $0_E = \alpha_1 v_1 + \cdots + \alpha_n v_n$? Ceci nous conduit à la définition suivante :

Définition 1.3.19. *Soient $v_1, \dots, v_p$ des vecteurs de E.*

1. *On dit que les vecteurs $v_1, \dots, v_n$ sont* ***linéairement indépendants*** *si la seule combinaison linéaire du vecteur 0_E sur $v_1, \dots, v_n$ est la combinaison linéaire triviale $0_E = 0 \cdot v_1 + \cdots + 0 \cdot v_n$ obtenue par les scalaires nuls. C'est-à-dire,*

$$\boxed{\forall \alpha_1, \dots, \alpha_n \in \mathbb{K} : \qquad \alpha_1 v_1 + \cdots + \alpha_n v_n = 0_E \implies \alpha_1 = \cdots = \alpha_n = 0}$$

   On dit aussi que $(v_1, \dots, v_n)$ est une ***famille libre*** *ou que $\{v_1, \dots, v_n\}$ est une* ***partie libre****.*

2. *Dans le cas contraire, on dit que les vecteurs $v_1, \dots, v_n$ sont* ***linéairement dépendants****. C'est-à-dire, s'il existe des scalaires <u>non tous nuls</u> $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ tels que $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_E$. C'est-à-dire,*

$$\boxed{\exists (\alpha_1, \dots, \alpha_n) \neq (0, \dots, 0) \text{ tel que } : \alpha_1 v_1 + \cdots + \alpha_n v_n = 0_E}$$

   On dit aussi que $(v_1, \dots, v_n)$ est une ***famille liée****, ou que $\{v_1, \dots, v_n\}$ est une* ***partie liée****.*
   Une telle égalité $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_E$ avec $(\alpha_1, \dots, \alpha_n) \neq (0, \dots, 0)$ est dite ***relation de liaison entre les vecteurs*** *$v_1, \dots, v_n$.*

Remarque 1.3.20. **1.** Dans la définition précédente, l'ordre des éléments $v_1, \dots, v_n$ n'a pas d'importance. Par exemple, si la famille $(v_1, v_2, \dots, v_n)$ est libre (respectivement liée), alors il en est de même pour la famille $(v_2, v_1, \dots, v_n)$

2. Dire que les scalaires $\alpha_1, \dots, \alpha_n$ sont <u>non tous nuls</u> signifie $(\alpha_1, \dots, \alpha_n) \neq (0, \dots, 0)$, c'est-à-dire, il existe au moins un scalaire α_i tel que $\alpha_i \neq 0$. Mais attention, ceci ne veut pas dire que les scalaires $\alpha_1, \dots, \alpha_n$ sont tous non nuls.

3. Ainsi, $(v_1, \dots, v_n\}$ est liée s'il existe des scalaires non tous nuls $\alpha_1, \dots, \alpha_n$ tels que $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_E$. Elle est libre lorsqu'il n'existe pas de scalaires non tous nuls $\alpha_1, \dots, \alpha_n$ tels que $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_E$, c'est-à-dire, si la seule façon d'obtenir le vecteur nul 0_E comme combinaison linéaire de ces vecteurs $v_1, \dots, v_n$ est de considérer la combinaison linéaire dont tous les scalaires sont nuls.

L'expérience montre que beaucoup d'étudiants trouvent des difficultés à comprendre les notions de partie libre et partie liée. Entraînez-vous à écrire correctement la définition afin de l'appliquer sans erreurs dans les exemples et les exercices.

Exemple 1.3.21. Dans le $\mathbb{C}$-espace vectoriel $E = \mathbb{C}^3$, les vecteurs $v_1 = (i, 0, 1), v_2 = (1, i, 0), v_3 = (0, 0, i)$ sont-ils linéairement indépendants ? Pour répondre à cette question, considérons trois scalaires $\alpha_1, \alpha_2, \alpha_3$ dans $\mathbb{C}$ tels que $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0_{\mathbb{C}^3}$. Alors on obtient le système linéaire suivant :

$$\begin{cases} \alpha_1 i + \alpha_2 = 0 & (1) \\ \alpha_2 i = 0 & (2) \\ \alpha_1 + \alpha_3 i = 0 & (3) \end{cases}$$

Il est facile de voir que ce système entraîne $\alpha_1 = \alpha_2 = \alpha_3 = 0$, c'est-à-dire que le système admet une seule solution $(\alpha_1, \alpha_2, \alpha_3) = (0, 0, 0)$. Par conséquent, les vecteurs v_1, v_2, v_3 sont linéairement indépendants.

Exemple 1.3.22. Etudions l'indépendance linéaire des vecteurs $x = (1, 3, -2),\ y = (3, 3, 0),\ z = (0, -3, 3)$ dans $\mathbb{R}^3$. Pour tous scalaires α, β, γ dans $\mathbb{R}$, l'équation $\alpha x + \beta y + \gamma z = 0_{\mathbb{R}^3}$ est équivalente au système :

$$\begin{cases} \alpha + 3\beta = 0 & (1) \\ 3\alpha + 3\beta - 3\gamma = 0 & (2) \\ -2\alpha + 3\gamma = 0 & (3) \end{cases}$$

En résolvant ce système, l'équation (1) donne $\alpha = -3\beta$, donc l'équation (3) implique $\gamma = \frac{2}{3}\alpha = \frac{2}{3}(-3\beta) = -2\beta$. Reportons ces résultats dans (2) : $3\alpha + 3\beta - 3\gamma = 3(-3\beta) + 3\beta - 3(-2\beta) = 0$. Il en résulte qu'il y a une infinité de solutions $(\alpha, \beta, \gamma) \in \mathbb{R}^3$: on peut donner à β une valeur arbitraire, et les scalaires α et γ sont donnés en fonction de β par les formules $\alpha = -3\beta$ et $\gamma = -2\beta$. Les scalaires α, β, γ ne sont donc pas nécessairement nuls. Il suffit de prendre par exemple $\beta = 1$, ce qui donne alors $\alpha = -3$ et $\gamma = -2$, pour trouver une combinaison linéaire nulle non triviale de x, y, z : $-3x + 1y - 2z = 0_{\mathbb{R}^3}$. Par conséquent, la partie $\{v_1, v_2, v_3\}$ est liée. L'égalité $-3x + 1y - 2z = 0_{\mathbb{R}^3}$ est une relation de liaison entre les vecteurs x, y et z. Observons que l'ensemble $\mathcal{S}$ des solutions du système précédent est donné par : $\mathcal{S} = \{(\alpha, \beta, \gamma) \in \mathbb{R}^3 \ / \ \alpha = -3\beta,\ \gamma = -2\beta\} = \{(-3\beta, \beta, -2\gamma) \ / \ \beta \in \mathbb{R}\}$. Il est possible de choisir une autre valeur de β, par exemple $\beta = 2$, ce qui implique $\alpha = -6$ et $\gamma = -4$. Dans ce cas, on obtient la relation de liaison $-6x + 2y - 4z = 0_{\mathbb{R}^3}$, qui est équivalente à la première relation de liaison $-3x + 1y - 2z = 0_{\mathbb{R}^3}$ (il suffit de multiplier cette dernière par 2). Notons que la solution triviale $(\alpha, \beta, \gamma) = (0, 0, 0)$ appartient évidemment à $\mathcal{S}$ (prendre $\beta = 0$) puisqu'on a toujours $0x + 0y + 0z = 0_{\mathbb{R}^3}$.

Exemple 1.3.23. Considérons les vecteurs $u = (1, 2, 1),\ v = (-1, 3, 4),\ w = (-4, 7, 11)$ du $\mathbb{R}$ espace vectoriel $\mathbb{R}^3$. La partie $\{u, v, w\}$ est liée, puisque $(-1)u + 3v + (-1)w = 0_{\mathbb{R}^3}$.

Remarque 1.3.24. Soit $\mathcal{F} = (v_1, \ldots, v_n)$ une famille de vecteurs de E.

1. Si la famille $\mathcal{F}$ contient le vecteur 0_E, alors elle est liée. Pour le voir, supposons par exemple que $v_1 = 0_E$. Alors $1 \cdot v_1 + 0 \cdot v_2 + \cdots + 0 \cdot v_n = 0_E$ avec les scalaires non tous nuls $\alpha_1 = 1, \alpha_2 = \ldots = \alpha_n = 0$.
2. Il en découle qu'une famille libre ne doit pas contenir le vecteur 0_E.
3. Un vecteur v est libre si et seulement si $v \neq 0_E$.
   En effet, si v est libre, alors $v \neq 0_E$ d'après ce qui précède. Réciproquement, si $v \neq 0_E$, alors pour tout $\alpha \in \mathbb{K}$, l'égalité $\alpha v = 0_E$ implique $\alpha = 0$ selon la proposition 1.1.15 (3).
4. Par contraposée, un vecteur $v \in E$ est lié si et seulement si $v = 0_E$.
5. La famille $\mathcal{F}$ peut contenir des vecteurs qui se répètent. Dans ce cas, $\mathcal{F}$ est liée. En effet, si par exemple $v_1 = v_2$, alors il suffit de poser $\alpha_1 = 1, \alpha_2 = -1$ et $\alpha_3 = \cdots = \alpha_n = 0$ pour avoir la relation $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0_E$ avec $(\alpha_1, \alpha_2, \ldots, \alpha_n) \neq (0, 0, \ldots, 0)$.
6. D'après ce qui précède, aucun vecteur ne peut se répèter dans une famille libre.

Exemple 1.3.25. **1.** Les vecteurs $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0), \ldots,$ $e_n = (0, \ldots, 0, 1)$ sont linéairement indépendants dans le $\mathbb{K}$-espace vectoriel $\mathbb{K}^n$.

En effet, soient $\alpha_1, \ldots, \alpha_n$ dans $\mathbb{K}$, et supposons $\alpha_1 e_1 + \cdots + \alpha_n e_n = 0_{\mathbb{K}^n}$. Puisque

$$\begin{aligned} \alpha_1 e_1 + \cdots + \alpha_n e_n &= (\alpha_1, 0, \ldots, 0) + (0, \alpha_2, 0, \ldots, 0) + \cdots + (0, \ldots, 0, \alpha_n) \\ &= (\alpha_1, \alpha_2, \ldots, \alpha_n), \end{aligned}$$

il vient $(\alpha_1, \alpha_2, \ldots, \alpha_n) = (0, 0, \ldots, 0)$, d'où $\alpha_1 = \cdots = \alpha_n = 0$.

2. Pour tout entier $n \in \mathbb{N}$, la famille $(1, X, X^2, \ldots, X^n)$ est libre dans $\mathbb{K}[X]$, car si $a_0, \ldots, a_n$ sont des scalaires dans $\mathbb{K}$ vérifiant $a_0 \cdot 1 + a_1 \cdot X + \cdots + a_n \cdot X^n = 0_{\mathbb{K}_n[X]}$, c'est-à-dire, $a_0 + a_1 X + \cdots + a_n X^n = 0_{\mathbb{K}_n[X]}$, alors $a_0 = a_1 = \cdots = a_n = 0$ (par définition du polynôme nul $0_{\mathbb{K}_n[X]}$).
3. Les vecteurs 1 et i sont linéairement indépendants dans le $\mathbb{R}$-espace vectoriel $\mathbb{C}$ parce que toute relation $a \cdot 1 + b \cdot i = 0$ avec $a, b \in \mathbb{R}$ entraîne $a = b = 0$ (les parties réelle et imaginaire d'un nombre complexe nul sont nuls).
4. Par contre, 1 et i ne sont pas linéairement indépendants dans le $\mathbb{C}$-espace vectoriel $\mathbb{C}$. Pour le voir, il suffit d'observer que $1 \cdot 1 + i \cdot i = 0$.

L'énoncé qui suit est l'analogue de la proposition 1.3.6 pour les familles libres :

Proposition 1.3.26. *Soit* $\mathcal{F} = (v_1, \ldots, v_p)$ *une famille de vecteurs de* E *et* $\mathcal{F}' = (v_1, \ldots, v_p, v_{p+1}, \ldots, v_n)$ *une sur-famille de* $\mathcal{F}$.

1. *Si* $\mathcal{F}'$ *est libre, alors* $\mathcal{F}$ *libre.*
   Autrement dit, toute sous-famille d'une famille libre est une famille libre.
2. *Si* $\mathcal{F}$ *est liée, alors* $\mathcal{F}'$ *liée.*
   C'est-à-dire, toute sur-famille d'une famille liée est une famille liée.

Démonstration. Supposons que $\mathcal{F}$ soit libre et montrons que F' est aussi libre. Soient $\alpha_1, \ldots, \alpha_p$ dans $\mathbb{K}$ tels que $\alpha_1 v_1 + \ldots + \alpha_p v_p = 0_E$. Alors $\alpha_1 v_1 + \ldots + \alpha_p v_p + 0 v_p + \ldots + 0 v_n = 0_E$. Puisque $(v_1, \ldots, v_p)$ est libre, tous les scalaires qui interviennent dans cette combinaison linéaire nulle sont nuls. D'où $\alpha_1 = \cdots = \alpha_n = 0$.
La seconde propriété s'obtient immédiatement dde la première par contraposition. ■

La définition d'une famille liée prend plus de sens si nous considérons la caractérisation suivante.

Proposition 1.3.27. *Soient* $v_1, \ldots, v_p$ *des vecteurs d'un espace vectoriel* E*. La famille* $(v_1, \ldots, v_p)$ *est liée si et seulement si l'un des vecteurs* v_i *est combinaison linéaire des autres.*
Dans ce cas, $\mathsf{Vect}(v_1, \ldots, v_p) = \mathsf{Vect}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_p)$.

Démonstration. • Si la famille est liée, il existe des scalaires non tous nuls $\alpha_1, \ldots, \alpha_p$ dans $\mathbb{K}$ tels que $\alpha_1 v_1 + \cdots + \alpha_p v_p = 0_E$. Il existe au moins un scalaire $\alpha_i \neq 0$. Il convient alors d'isoler le vecteur v_i comme suit :
$\alpha_i v_i = -\sum\limits_{j \neq i} \alpha_j v_j$, ce qui donne $v_i = \sum\limits_{j \neq i} \left(\dfrac{-\alpha_j}{\alpha_i} \right) v_j = \sum\limits_{j \neq i} \beta_j v_j$ avec $\beta_j = \dfrac{-\alpha_j}{\alpha_i}$ pour tout $j \neq i$. Donc v_i est combinaison linéaire des autres vecteurs $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$.
• Réciproquement, supposons que l'un des vecteurs v_i soit combinaison linéaire des autres. Alors il existe des scalaires $\lambda_1, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots, \lambda_p$ dans $\mathbb{K}$ tels que $v_i = \sum\limits_{j \neq i} \lambda_j v_j$, donc $v_i - \sum\limits_{j \neq i} \lambda_j v_j = 0_E$, ce qui prouve que $(v_1, \ldots, v_p)$ est liée, car le scalaire 1 affecté à v_i dans cette combinaison linéaire nulle est non nul.
• Dans ces conditions, le corollaire 1.3.17 implique que

$$\mathsf{Vect}(v_1, \ldots, v_p) = \mathsf{Vect}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_p).$$

■

Remarque 1.3.28. **1.** Avec les conditions de la proposition précédente, nous avons $\mathsf{Vect}(v_1, \dots, v_p) = \mathsf{Vect}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_p)$. Ceci permet donc de réduire la famille génératrice $v_1, \dots, v_p$ du sous-espace vectoriel $F = \mathsf{Vect}(v_1, \dots, v_p)$ en supprimant le vecteur v_i. Pour continuer à réduire cette famille génératrice, on pourra répéter ce processus en éliminant un vecteur de cette famille chaque fois qu'il est combinaison linéaire des autres.

2. Par contraposition de la proposition ci-dessus :
Une famille $(v_1, \dots, v_p)$ est libre si et seulement si aucun vecteur v_i de cette famille n'est combinaison linéaire des autres.

3. Si la famille $\mathcal{F}$ contient deux vecteurs v_i et v_j qui sont proportionnels (colinéaires) : $v_i = \lambda v_j$ avec $\lambda \in \mathbb{K}$, alors $\mathcal{F}$ est liée. En effet, il suffit d'observer que $v_i = 0v_1 + \cdots + \lambda v_j + \cdots + 0v_p$ est combinaison linéaire des autres.

4. Une famille libre ne doit donc jamais avoir deux vecteurs colinéaires.

5. Une famille (u, v) à deux vecteurs est liée si et seulement si u et v sont colinéaires. En effet, si u et v sont colinéaires, alors (u, v) sera évidemment liée. Réciproquement, si (u, v) est liée, la proposition 1.3.27 exprime que l'un des vecteurs de cette famille, par exemple u, est combinaison linéaire de l'autre, c'est-à-dire de v, c'est-à-dire, $u = \lambda v$ avec $\lambda \in \mathbb{K}$.

Exemple 1.3.29. **1.** Dans $\mathbb{R}^4$, les trois vecteurs $v_1 = (1, 1, 0, -1)$, $v_2 = (1, 2, 1, 0)$ et $v_3 = (3, 5, 2, -1)$ forment une partie liée, puisque l'un des vecteurs, à savoir le vecteur v_3, s'exprime comme combinaison linéaire des deux autres : $v_3 = v_1 + 2v_2$. Notons qu'on a aussi $v_1 = -2v_2 + v_3$, c'est-à- dire, v_1 est aussi combinaison linéaire des autres. Même chose pour $v_2 = \left(\frac{-1}{2}\right) v_1 + \frac{1}{2} v_3$.

2. Dans le $\mathbb{C}$-espace vectoriel $E = \mathbb{C}^4$, les vecteurs $v_1 = (1, 0, 1, 1), v_2 = (0, 2, 2i, 6)$ et $v_3 = (1, i, 0, 1+3i)$ sont linéairement dépendants puisque $v_1 + \frac{i}{2} v_2 - v_3 = 0_{\mathbb{C}^4}$.

3. La famille $(1, X, 1 - X^2, X^2, \dots, X^n)$ est liée dans $\mathbb{K}[X]$.

Remarque 1.3.30. Attention ! Si une famille $\mathcal{F}$ est liée, on ne peut choisir "au hazard" un vecteur de cette famille qui soit combinaison linéaire des autres. Par exemple, soient v_1 et v_2 deux vecteurs linéairement indépendants. Alors la famille $(v_1, -v_1, v_2)$ est liée, car v_1 est proportionnel à $-v_1$. Mais v_2 n'est pas combinaison linéaire de v_1 et $-v_1$, car si c'était le cas, il existerait γ et λ dans $\mathbb{K}$ tels que $v_2 = \gamma v_1 + \lambda(-v_1) = (\gamma - \lambda) v_1$, et donc (v_1, v_2) serait liée, absurde.

Lorsqu'on travaille dans des espaces vectoriels de fonctions, on est parfois amené à étudier l'indépendance linéaire d'une famille de fonctions. Il faut alors penser à utiliser des techniques d'analyse (passage aux limites, dérivation, etc.) si les fonctions considérées le permettent.

Exemple 1.3.31. Dans l'espace vectoriel $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$, on considère les fonctions f_n définies par $f_n(x) = e^{nx}$ $(n \in \mathbb{N}^*)$. Montrons par récurrence que pour tout $n \geq 1$, la partie $\{f_1, \dots, f_n\}$ est libre.

- La propriété est vraie pour $n = 1$, car la fonction f_1 est différente de la fonction nulle θ.
- Supposons la propriété vraie pour un entier n et montrons-la pour $n+1$. Soient $\alpha_1, \dots, \alpha_n, \alpha_{n+1}$ dans $\mathbb{R}$ tels que

$$\alpha_1 f_1 + \alpha_2 f_2 + \cdots + \alpha_n f_n + \alpha_{n+1} f_{n+1} = \theta. \tag{1.2}$$

Alors

$$\forall x \in \mathbb{R} \qquad \alpha_1 e^x + \alpha_2 e^{2x} + \cdots + \alpha_n e^{nx} + \alpha_{n+1} e^{(n+1)x} = 0. \tag{1.3}$$

Dérivons cette équation :

$$\forall x \in \mathbb{R} \qquad \alpha_1 e^x + 2\alpha_2 e^{2x} + \cdots + n\alpha_n e^{nx} + (n+1)\alpha_{n+1} e^{(n+1)x} = 0. \tag{1.4}$$

D'autre part, la multiplication de l'équation (1.3) par $n+1$ fournit

$$\forall x \in \mathbb{R} \qquad (n+1)\alpha_1 e^x + (n+1)\alpha_2 e^{2x} + \cdots + (n+1)\alpha_n e^{nx} + (n+1)\alpha_{n+1} e^{(n+1)x} = 0. \tag{1.5}$$

En effectuant la soustraction des deux dernières équations (1.4) et (1.5), on arrive à supprimer le dernier terme :

$$\forall x \in \mathbb{R} \qquad n\alpha_1 e^x + (n-1)\alpha_2 e^{2x} + \cdots + \alpha_n e^{nx} = 0. \tag{1.6}$$

Donc

$$\beta_1 f_1 + \beta_2 f_2 + \cdots + \beta_n f_n = \theta,$$

avec $\beta_1 = n\alpha_1,\ \beta_2 = (n-1)\alpha_2, \dots, \beta_n = \alpha_n$. Or, l'hypothèse de récurrence nous dit que $(f_1, \dots, f_n)$ est libre, donc $\beta_1 = \beta_2 = \cdots = \beta_n = 0$, ce qui entraîne $\alpha_1 = \cdots = \alpha_n = 0$. Finalement, nous remplaçons ces valeurs $\alpha_1 = \cdots = \alpha_n = 0$ dans (1.2) pour obtenir $\alpha_{n+1} f_{n+1} = \theta$, d'où $\alpha_{n+1} = 0$. ■

Nous achevons cette section par le résultat technique suivant qui sera employé à maintes reprises dans la théorie de la dimension pour démontrer qu'une famille de vecteurs est libre.

Théorème 1.3.32. *Soit $(v_1, \ldots, v_p)$ une famille <u>libre</u> de vecteurs d'un espace vectoriel E. Pour tout vecteur $v \in E$, on a :*

$$v \notin \mathsf{Vect}(v_1, \ldots, v_p) \Longrightarrow (v_1, \ldots, v_p, v) \text{ est libre.}$$

Démonstration. Supposons que $v \notin \mathsf{Vect}(v_1, \ldots, v_p)$ et montrons que $(v_1, \ldots, v_p, v)$ est libre. Soient $\alpha_1, \ldots, \alpha_p, \alpha_{p+1}$ dans $\mathbb{K}$ tels que

$$\alpha_1 v_1 + \cdots + \alpha_p v_p + \alpha v = 0_E. \tag{1.7}$$

On veut prouver que $\alpha_1 = \ldots = \alpha_p = \alpha = 0$. Commençons par montrer que $\alpha = 0$. Si $\alpha \neq 0$, alors

$$v = \left(\frac{-1}{\alpha}\right)(\alpha_1 v_1 + \cdots + \alpha_p v_p).$$

Donc $v \in \mathsf{Vect}(v_1, \ldots, v_p)$, ce qui est contradictoire avec l'hypothèse. Par conséquent, on a nécessairement $\alpha = 0$, et donc l'égalité (1.7) devient $\alpha_1 v_1 + \cdots + \alpha_p v_p = 0_E$. Maintenant, comme $(v_1, \ldots, v_p)$ est libre, alors $\alpha_1 = \cdots = \alpha_p = 0$. ■

En passant à la contraposée, et en appliquant le corollaire 1.3.17, le théorème précédent peut être formulé comme suit :

Corollaire 1.3.33. *Soit $(v_1, \ldots, v_p)$ une famille <u>libre</u> de vecteurs d'un espace vectoriel E. Pour tout vecteur $v \in E$, on a :*

$$(v_1, \ldots, v_p, v) \text{ est liée} \Longrightarrow v \in \mathsf{Vect}(v_1, \ldots, v_p).$$

Et dans ces conditions, $\mathsf{Vect}(v_1, \ldots, v_p, v) = \mathsf{Vect}(v_1, \ldots, v_p)$.

1.3.3 Bases

Si l'on se souvient de la géométrie usuelle du plan ou de l'espace, telle qu'elle a été étudiée dans le secondaire, il est clair que l'un des outils essentiels était la décomposition des "vecteurs" sur une "base". C'est cette idée qui va être développée dans le cas général.

Définition 1.3.34. *Soit E un espace vectoriel sur un corps K et $(a_1, \ldots, a_n)$ une partie de E. On dit que $(a_1, \ldots, a_n)$ est une **base** de E si $(a_1, \ldots, a_n)$ est une partie à la fois libre et génératrice de E.*

L'avantage des bases s'explique par le résultat fondamental ci-dessous :

Théorème 1.3.35. *$B = (a_1, \dots, a_n)$ est une base de E si et seulement si tout vecteur x de E s'écrit de manière unique comme combinaison linéaire de $a_1, \dots, a_n$:*

$$\boxed{\forall x \in E \ \ \exists! \ (\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n : x = \alpha_1 a_1 + \cdots + \alpha_n a_n}$$

Démonstration.

- Si B est une base de E, elle engendre E et donc tout vecteur de E est combinaison linéaire de B. Si un vecteur x de E admet deux écritures
$$x = \sum_{i=1}^{n} \alpha_i a_i = \sum_{i=1}^{n} \beta_i a_i,$$
avec $\alpha_i, \beta_i \in \mathbb{K}$, on obtient la relation
$$\sum_{i=1}^{n} (\alpha_i - \beta_i) a_i = 0_E.$$
Et comme B est libre, alors pour tout $i = 1, \dots, n : \alpha_i - \beta_i = 0$, ce qui donne $\alpha_i = \beta_i$. La décomposition de x est donc unique.
- Réciproquement, supposons que tout vecteur x de E s'écrit de manière unique comme combinaison linéaire des vecteurs de B. Alors en particulier B engendre E. Pour montrer qu'elle est libre, soit $\alpha_1 a_1 + \cdots + \alpha_n a_n = 0_E$ une combinaison linéaire nulle de $a_1, \dots, a_n$, avec les scalaires $\alpha_i \in \mathbb{K}$. Or, on a aussi la combinaison linéaire triviale $0a_1 + \cdots + 0a_n = 0_E$. Par conséquent, l'unicité de la décomposition du vecteur 0_E comme combinaison linéaire de $a_1, \dots, a_n$ assure que $\alpha_1 = \ldots = \alpha_n = 0$. Ainsi, B est une base de E. ■

Remarque 1.3.36. Dans la remarque 1.3.18, nous avons souligné que si $(a_1, \dots, a_n)$ est une famille génératrice de E, alors la décomposition d'un vecteur $x \in E$ comme combinaison linéaire de $(a_1, \dots, a_n)$ n'est pas nécessairement unique. Mais si, de plus, la famille $(a_1, \dots, a_n)$ est libre, le théorème précédent affirme l'unicité de cette décomposition de x.

Ceci motive la définition légitime suivante :

Définition 1.3.37. *Avec les notations du théorème 1.3.35, les scalaires $\alpha_1, \dots, \alpha_n$ s'appellent les **coordonnées** (ou **composantes**) du vecteur x dans la base B.*

L'ordre des éléments $a_1, \ldots, a_n$ d'une base $B = (a_1, \ldots, a_n)$ est très important. Il est clair que si l'on change l'ordre des vecteurs $a_1, a_2, \ldots, a_n$, la nouvelle famille, par exemple $B' = (a_2, a_1, \ldots, a_n)$ reste encore une base de E. Mais B' est différente de B et les coordonnées d'un vecteur x de E dans cette nouvelle base changent : si $x = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_n a_n$, alors $\alpha_1, \alpha_2, \ldots, \alpha_n$ sont les coordonnées de x dans la base B, tandis que $\alpha_2, \alpha_1, \ldots, \alpha_n$ deviennent les coordonnées de x dans la base B'. L'importance de l'ordre des vecteurs d'une base sera visible lorsqu'on étudiera au chapitre 3 la notion de matrice associée à une application linéaire dans des bases données.

Exemple 1.3.38. **1.** Nous avons déjà vu que les vecteurs $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)$ sont linéairement indépendants et forment une partie génératrice du $\mathbb{K}$-espace vectoriel $\mathbb{K}^n$. Donc $B = \{e_1, \ldots, e_n\}$ est une base de $\mathbb{K}^n$, appelée **base canonique** de $\mathbb{K}^n$.

De plus, tout vecteur $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ s'écrit $x = x_1 e_1 + \ldots + x_n e_n$. Par conséquent, les scalaires $x_1, \ldots, x_n$ sont les coordonnées du vecteur $x = (x_1, \ldots, x_n)$ de $\mathbb{K}^n$ par rapport à la base canonique B.

2. La famille $B = (1, X, X^2, \ldots, X^n)$ est une base de $\mathbb{K}_n[X]$, appelée aussi **base canonique** de $\mathbb{K}_n[X]$. Tout polynôme P de $\mathbb{K}_n[X]$ s'exprime de façon unique sous forme $P = a_0 1 + a_1 X + a_2 X^2 + \ldots + a_n X^n$. Par conséquent, les cœfficients $a_0, a_1, \ldots, a_n$ sont les coordonnées de P dans cette B.

3. Soit $a \in \mathbb{K}$ fixé. On rappelle que tout polynôme P de $\mathbb{K}_n[X]$ s'exprime aussi sous la forme

$$P = P(a) + \frac{P'(a)}{1!}(X-a) + \frac{P''(a)}{2!}(X-a)^2 + \ldots + \frac{P^{(n)}(a)}{n!}(X-a)^n$$

où, pour tout entier k compris entre 0 et n, $P^{(k)}(a)$ désigne la dérivée d'ordre k au point a de $P(X)$. C'est la formule de Taylor pour les polynômes que vous avez apprise en premier semestre SMIA1.

Ainsi, pour tout $a \in \mathbb{K}$, la famille $B' = (1, X-a, (X-a)^2, \ldots, (X-a)^n)$ est une famille génératrice de $\mathbb{K}_n[X]$.

Montrons qu'elle est libre. Soient $\alpha_0, \alpha_1, \ldots, \alpha_n$ dans $\mathbb{K}$ tels que

$$\alpha_0 + \alpha_1(X-a) + \alpha_2(X-a)^2 + \cdots + \alpha_n(X-a)^n = 0.$$

En prenant $X = a$, on obtient $\alpha_0 = 0$, ce qui donne

$$\alpha_1(X-a) + \alpha_2(X-a)^2 + \cdots + \alpha_n(X-a)^n = 0,$$

ou encore

$$(X-a)\left[\alpha_1 + \alpha_2(X-a) + \cdots + \alpha_n(X-a)^{n-1}\right] = 0.$$

Puisque l'anneau $\mathbb{K}[X]$ des polynômes est intègre et le polynôme $X-a$ est différent du polynôme nul $0_{\mathbb{K}[X]}$, il vient

$$\alpha_1 + \alpha_2(X-a) + \cdots + \alpha_n(X-a)^{n-1} = 0.$$

En remplaçant X par a, on trouve $\alpha_1 = 0$. En itérant ce procédé, on montre que $\alpha_0 = \alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$.

Par conséquent, $B' = (1, X-a, (X-a)^2, \ldots, (X-a)^n)$ est une autre base de $\mathbb{K}_n[X]$. Les coordonnées d'un polynôme $P \in \mathbb{K}[X]$ par rapport à cette base B' sont donc

$$P(a), \frac{P'(a)}{1!}, \frac{P''(a)}{2!}, \ldots, \frac{P^{(n)}(a)}{n!}.$$

4. Les vecteurs 1 et i constituent une base $B = (1, i)$ du $\mathbb{R}$-espace vectoriel $\mathbb{C}$. Les coordonnées d'un complexe $z = a + ib$ dans cette base sont ses parties réelle et imaginaire a et b.

5. Soient l'espace vectoriel $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$ et n un entier fixé. Considérons l'ensemble F des fonctions $f : \mathbb{R} \longrightarrow \mathbb{R}$ de la forme $f(x) = \alpha_1 e^x + \cdots + \alpha_n e^{nx}$ où $\alpha_1, \ldots, \alpha_n$ parcourent $\mathbb{R}$. Il est facile de voir que $F = \mathsf{Vect}(f_1, \ldots, f_n)$ où $f_1, \ldots, f_n$ sont les fonctions définies par $f_k(x) = e^{kx}$ $(1 \leq k \leq n)$. Donc $(f_1, \ldots, f_n)$ engendre F. Puisque $(f_1, \ldots, f_n)$ est libre (exemple 1.3.31), alors c'est une base de F.

6. Les familles $(v_1 = (1, -1, 2), v_2 = (0, 1, 3))$ et $(x = (1, 3, -2),\ y = (3, 3, 0),\ z = (0, -3, 3))$ ne sont pas des bases de $\mathbb{R}^3$. En effet, la première famille n'engendre pas $\mathbb{R}^3$ (exemple 1.3.5 (4)), tandis que la seconde est liée (exemple 1.3.22).

7. Reprenons l'exemple 1.2.6 concernant le sous-espace vectoriel W de $\mathbb{R}^3$ donné par $W = \{(x_1, 2x_1 + x_2, x_1 - x_2) \ / \ x_1, x_2 \in \mathbb{R}\}$. Cherchons une base de W. Pour ce faire, on commence d'abord par trouver une famille génératrice de W. Justement, on a déjà vu dans l'exemple 1.3.12 que W est engendré par la famille $(v_1 = (1, 2, 1), v_2 = (0, 1, -1))$. Il reste à prouver que cette famille est libre. Si $\alpha v_1 + \beta v_2 = 0_{\mathbb{R}^3}$, alors $(\alpha, 2\alpha + \beta, \alpha - \beta) = (0, 0, 0)$, ce qui entraîne aisément $\alpha = \beta = 0$. D'où $B = (v_1, v_2)$ est une base de W.

Les exemples précédents montrent qu'un espace vectoriel E peut avoir plusieurs bases différentes. Donc il est naturel de se demander s'il y a en général une base "préférée" à choisir parmi les autres bases de E. Pour l'espace vectoriel $\mathbb{K}^n$, il est souvent préférable de travailler avec la *base canonique* $(e_1, \ldots, e_n)$. De même, pour l'espace $\mathbb{K}_n[X]$, on choisit d'habitude la *base canonique* $(1, X, \ldots, X^n)$. Ces bases canoniques simplifient les calculs dans la plupart des situations. Cependant, il y a des occasions où d'autres bases serviront mieux certains propos spécifiques. C'est par exemple le cas pour la base $(1, X-a, (X-a)^2, \ldots, (X-a)^n)$ de $\mathbb{K}_n[X]$.

1.4 Espaces vectoriels de dimensions finies

Nous avons vu qu'un espace vectoriel peut avoir plusieurs bases différentes. Notre objectif dans cette section est de prouver l'existence des bases et de collecter certaines propriétés concernant le nombre de vecteurs d'une base. Pour cela, nous limitons notre étude au cas des espaces vectoriels *de dimensions finies.*

1.4.1 Généralités

Définition 1.4.1. *On dit qu'un espace vectoriel E sur $\mathbb{K}$ est* ***de dimension finie*** *s'il possède une famille génératrice finie. Dans le cas contraire, on dit que E est* ***de dimension infinie****.*

Voici quelques exemples classiques d'espaces vectoriels de dimensions finies :

Exemple 1.4.2. 1. Le $\mathbb{R}$-espace vectoriel $\mathbb{R}$ est de dimension finie : $\mathbb{R} = \mathsf{Vect}(1)$. En effet, ici $E = \mathbb{K} = \mathbb{R}$, et tout vecteur $x \in \mathbb{R}$ s'écrit trivialement $x = x.1$ comme produit externe du vecteur 1 par le scalaire x.

2. De même, le $\mathbb{C}$-espace vectoriel $\mathbb{C}$ est de dimension finie : $\mathbb{C} = \mathsf{Vect}(1)$.
3. D'une façon générale, le $\mathbb{K}$-espace vectoriel $\mathbb{K}$ est de dimension finie : $\mathbb{K} = \mathsf{Vect}(1)$. Mais on peut prendre n'importe quel élément $a \neq 0$ comme générateur, car pour tout $x \in \mathbb{K}$, on a $x = \lambda a$ avec $\lambda = \dfrac{x}{a} \in \mathbb{K}$. D'où $\mathbb{K} = \mathsf{Vect}(a)$.
4. Pour tout entier $n \geq 1$, l'espace vectoriel $\mathbb{K}^n = \mathsf{Vect}(\mathsf{e_1}, \dots, \mathsf{e_n})$ est de dimension finie. En particulier $\mathbb{R}^n$ est un espace vectoriel de dimension finie sur $\mathbb{R}$ et $\mathbb{C}^n$ est un espace vectoriel de dimension finie sur $\mathbb{C}$.
5. Le $\mathbb{R}$-espace vectoriel $\mathbb{C}$ est de dimension finie : $\mathbb{C} = \mathsf{Vect}(1, \mathsf{i})$. En effet, ici $E = \mathbb{C}$, $\mathbb{K} = \mathbb{R}$ et tout vecteur $x \in \mathbb{C}$ s'écrit $x = a.1 + b.i$ où $a, b \in \mathbb{R}$ sont respectivement la partie réelle et la partie imaginaire de z.
6. $\mathbb{K}_n[X]$ est de dimension finie, puisqu'il est engendré par la famille finie $(1, X, X^2, \dots, X^n)$.
7. Soit E un $\mathbb{K}$-espace vectoriel quelconque (non nécessairement de dimension finie) et soit $(v_1, \dots, v_p)$ une famille finie de vecteurs de E. Alors, par construction, le sous-espace $\mathsf{Vect}(\mathsf{v_1}, \dots, \mathsf{v_p})$ est toujours de dimension finie.

Il ne ne faudrait pas croire que tous les espaces vectoriels sont de dimensions finies. L'exemple suivant illustre cette situation.

Exemple 1.4.3. L'espace vectoriel $\mathbb{K}[X]$ des polynômes est de dimension infinie sur $\mathbb{K}$. En effet, supposons par l'absurde qu'il est engendré par un nombre fini de polynômes $P_1, \ldots, P_r$. Soit n le plus haut degré des polynômes $P_1, \ldots, P_r$, et écrivons le polynôme X^{n+1} comme combinaison linéaire de $P_1, \ldots, P_r$:

$$X^{n+1} = \alpha_1 P_1 + \cdots + \alpha_r P_r. \tag{1.8}$$

En comparant les degrés dans cette égalité polynômiale (1.8), on aboutit à une contradiction : $\deg(X^{n+1}) = n+1$ et $\deg(\alpha_1 P_1 + \cdots + \alpha_r P_r) \leq n$.

On verra plus loin que l'espace vectoriel $\mathcal{F}(\mathbb{R}, \mathbb{R})$ est également de dimension infinie. En fait, il existe une théorie des espaces vectoriels de dimension infinie, mais cela dépasse le cadre de ce cours.

1.4.2 Existence des bases dans les espaces vectoriels de dimensions finies

Nous avons déjà donné des exemples de bases dans certains espaces vectoriels. Mais jusqu'à maintenant, nous n'avons pas encore prouvé l'existence d'une base dans un espace vectoriel quelconque. C'est le but du théorème suivant :

Théorème 1.4.4. *Soit E un espace vectoriel de dimension finie sur $\mathbb{K}$. Alors toute partie génératrice de E contient une base de E :*
Si $(a_1, \ldots, a_n)$ est une partie génératrice de E, on peut en extraire une base $(a_{j_1}, \ldots, a_{j_p})$ de E.

Démonstration. Si $E = \{0_E\}$, on dit par convention que $\varnothing$ est une base $\{0_E\}$. Prenons donc $E \neq \{0_E\}$ et raisonnons par récurrence sur n. Pour $n = 1$, $E = \mathsf{Vect}(a_1)$ avec $a_1 \neq 0_E$ et donc a_1 est libre. Ainsi le vecteur a_1 est une base de E. Supposons que l'énoncé du théorème est vrai pour tout espace vectoriel engendré par n vecteurs. Soit E un espace vectoriel engendré par $n+1$ vecteurs $a_1, \ldots, a_n, a_{n+1}$. Distinguons les deux cas suivants :

- Si le système $(a_1, \ldots, a_n, a_{n+1})$ est libre, c'est une base de E.
- S'il est lié, la proposition 1.3.27 entraîne l'existence d'un vecteur a_i de cette famille qui soit combinaison linéaire des autres et tel que

$$\mathsf{Vect}(a_1, \ldots, a_i, \ldots, a_{n+1}) = \mathsf{Vect}(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_{n+1}).$$

  Il s'ensuit que $E = \mathsf{Vect}(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_{n+1})$, c'est-à-dire que E est engendré par n vecteurs $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_{n+1}$. L'hypothèse de récurrence nous dit alors qu'il existe une base $(b_1, \ldots, b_p)$ de E extraite de la famille $(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_{n+1})$. ■

Nous pouvons maintenant formuler une conséquence immédiate du théorème 1.4.4.

Corollaire 1.4.5. *Tout espace vectoriel E de dimension finie sur un corps $\mathbb{K}$ admet au moins une base.*

Comme nous avons vu dans certains exemples, un espace vectoriel peut avoir plusieurs bases différentes. Une question naturelle s'impose, à savoir, quelles propriétés communes peuvent exister entre différentes bases de E ? Nous allons répondre à cette question en prouvant que toutes les bases d'un espace vectoriel de dimension finie possèdent le même nombre d'éléments. Pour cela, nous avons besoin de montrer un résultat préliminaire important qui va préparer le terrain. Sa démonstration technique demande un peu de concentration et de patience !

Lemme 1.4.6 (de Steinitz)**.** *Soit E un espace vectoriel de dimension quelconque sur $\mathbb{K}$ et $n \geq 1$ un entier. Alors $n+1$ combinaisons linéaires de n vecteurs de E sont toujours linéairement dépendants.*
Plus précisément, si les vecteurs $b_1, \ldots, b_{n+1}$ sont combinaisons linéaires des vecteurs $a_1, \ldots, a_n$, alors $b_1, \ldots, b_{n+1}$ sont linéairement dépendants.

Démonstration. La démonstration est basée sur un procédé d'élimination et s'effectue par récurrence sur l'entier n.

1. Pour $n = 1$, on a deux vecteurs b_1 et b_2 qui sont combinaisons linéaires du vecteur a_1. Il existe donc des scalaires α et β tels que $b_1 = \alpha a_1$ et $b_2 = \beta a_1$. Si $\alpha = 0$ ou $\beta = 0$, alors $b_1 = 0_E$ ou $b_2 = 0_E$, et donc (b_1, b_2) est liée. Si $\alpha \neq 0$ et $\beta \neq 0$, on a la relation de liaison non triviale $\beta b_1 - \alpha b_2 = 0_E$ qui montre que (b_1, b_2) est liée.
2. Supposons la propriété satisfaite pour $n-1$, c'est-à-dire, n combinaisons linéaires de $n-1$ vecteurs sont toujours linéairement dépendants, et démontrons-la pour $n+1$. Soit donc $n+1$ vecteurs $b_1, \ldots, b_n, b_{n+1}$ de E qui sont combinaisons linéaires de n vecteurs $a_1, \ldots, a_n$. On peut alors écrire pour tout $i \in \{1, \ldots, n+1\}$,
$$b_i = c_i + \alpha_i a_n \text{ avec } \alpha_i \in \mathbb{K} \text{ et } c_i \in \mathsf{Vect}(a_1, \ldots, a_{n-1}).$$
A ce niveau, deux cas se présentent :
   - Si tous les scalaires α_i son nuls $(1 \leq i \leq n+1)$, alors chaque $b_i \in \mathsf{Vect}(a_1, \ldots, a_{n-1})$. En particulier, les n vecteurs $b_1, \ldots, b_n$ sont combinaisons linéaires de $a_1, \ldots, a_{n-1}$. Nous appliquons alors l'hypothèse de récurrence pour déduire que les vecteurs $b_1, \ldots, b_n$ sont linéairement dépendants, et il en sera donc de même de la sur-famille $b_1, \ldots, b_n, b_{n+1}$.

- Sinon, au moins un des scalaires α_i est non nul. Quitte à renuméroter si nécessaire les vecteurs, on peut supposer (sans perdre la généralité) que $\alpha_1 \neq 0$ par exemple. Considérons maintenant les vecteurs $v_2, \ldots, v_{n+1}$ définis par

$$v_2 = \alpha_1 b_2 - \alpha_2 b_1, \ \ldots, v_{n+1} = \alpha_1 b_{n+1} - \alpha_{n+1} b_1,$$

c'est-à-dire,

$$v_i = \alpha_1 b_i - \alpha_i b_1 \ \ (i = 2, \ldots, n+1).$$

Il n'est pas difficile de remarquer que ces vecteurs $v_2, \ldots, v_{n+1}$ n'ont pas de composantes sur le vecteur a_n. En effet, pour tout $i = 2, \ldots, n+1$,

$$v_i = \alpha_1 b_i - \alpha_i b_1 = \alpha_1 \Big(c_i + \alpha_i a_n \Big) - \alpha_i \Big(c_1 + \alpha_1 a_n \Big) = \alpha_1 c_i - \alpha_i c_1.$$

Donc $v_i \in \mathsf{Vect}(a_1, \ldots, a_{n-1})$, car c_i et c_1 sont dans $\mathsf{Vect}(a_1, \ldots, a_{n-1})$.
Par conséquent, nous sommes en présence de n vecteurs $v_2, \ldots, v_{n+1}$ qui sont combinaisons linéaires de $n-1$ vecteurs $a_1, \ldots, a_{n-1}$. Par hypothèse de récurrence, ces vecteurs $v_2, \ldots, v_{n+1}$ sont linéairement dépendants. On en déduit l'existence de scalaires <u>non tous nuls</u> $\beta_2, \ldots, \beta_{n+1}$ tels que

$$\beta_2 v_2 + \cdots + \beta_{n+1} v_{n+1} = 0_E,$$

ce qui revient à dire que

$$\beta_2(\alpha_1 b_2 - \alpha_2 b_1) + \cdots + \beta_{n+1}(\alpha_1 b_{n+1} - \alpha_{n+1} b_1) = 0_E.$$

En regroupant les cœfficients en b_1, on obtient :

$$-\left(\beta_2 \alpha_2 + \cdots + \beta_{n+1} \alpha_{n+1}\right) b_1 + (\beta_2 \alpha_1) b_2 + \cdots + (\beta_{n+1} \alpha_1) b_{n+1} = 0_E.$$

Comme $\alpha_1 \neq 0$ et $(\beta_2, \ldots, \beta_{n+1}) \neq (0, \ldots, 0)$, alors $(\beta_2 \alpha_1, \ldots, \beta_{n+1} \alpha_1) \neq (0, \ldots, 0)$. Finalement, la famille $(b_1, b_2, \ldots, b_{n+1})$ est liée, ce qui achève la preuve. ■

Le lemme précédent est le point de départ de plusieurs résultats que nous démontrerons dans la suite. Commençons par le corollaire immédiat :

Corollaire 1.4.7. *Soit E un espace vectoriel engendré par une famille $(a_1, \ldots, a_n)$ à n éléments. Toute famille $(b_1, \ldots, b_p)$ d'éléments de E de cardinal $p > n$ est liée.*

Démonstration. Soit $(b_1, \ldots, b_n, b_{n+1}, \ldots, b_p)$ une famille de vecteurs de cardinal $p > n$. Selon le lemme de Steinitz, la famille $(b_1, \ldots, b_n, b_{n+1})$ est liée, car

$b_1, \ldots, b_n, b_{n+1}$ sont combinaisons linéaires de $a_1, \ldots, a_n$. On applique alors la proposition 1.3.26 (2) pour déduire que la sur-famille $(b_1, \ldots, b_n, b_{n+1}, \ldots, b_p)$ est liée. ■

Par contraposition, nous obtenons :

Corollaire 1.4.8. *Soit E un espace vectoriel de dimension finie sur K. Si $\mathcal{L} = (a_1, \ldots, a_p)$ est une famille libre de E et $\mathcal{G} = (b_1, \ldots, b_n)$ est une famille génératrice de E, alors $p \leq n$.*

DÉMONSTRATION. Par l'absurde, si $p > n$, alors le corollaire qui précède implique que la famille $\mathcal{L} = (a_1, \ldots, a_p)$ ne sera pas libre. ■

A présent, nous sommes en position de prouver le théorème principal suivant :

Corollaire 1.4.9. *Soit E un espace vectoriel de dimension finie sur K, et B_1, B_2 deux bases quelconques de E. Alors $\mathrm{card}(B_1) = \mathrm{card}(B_2)$.*

DÉMONSTRATION. Puisque B_1 est libre et B_2 engendre E, le corollaire 1.4.8 affirme que $\mathsf{card}(B_1) \leq \mathsf{card}(B_2)$. En inversant les rôles de B_1 et B_2, on obtient aussi $\mathsf{card}(B_2) \leq \mathsf{card}(B_1)$. La démonstration est terminée. ■

Le corollaire 1.4.9 justifie bien la définition suivante.

Définition 1.4.10. *Soit E un espace vectoriel de dimension finie sur $\mathbb{K}$. On appelle **dimension** de E sur $\mathbb{K}$, et l'on note $\dim_{\mathbb{K}} E$, le cardinal de n'importe quelle base de E.*
Lorsqu'il n'y a aucune confusion sur le corps $\mathbb{K}$, on peut noter simplement $\dim(E)$ la dimension de E sur $\mathbb{K}$.

Exemple 1.4.11. D'après l'exemple 1.4.2, on a immédiatement :

1. $\dim_{\mathbb{K}} \mathbb{K}^n = n$.
2. En particulier, $\dim_{\mathbb{K}}(\mathbb{K}) = 1$ et donc $\dim_{\mathbb{R}} \mathbb{R} = 1$, $\dim_{\mathbb{C}} \mathbb{C} = 1$.
3. Mais attention, $\dim_{\mathbb{R}} \mathbb{C} = 2$. Ceci montre la dimension d'un espace vectoriel E dépend du corps des scalaires $\mathbb{K}$.
4. $\dim_{\mathbb{K}} \mathbb{K}_n[X] = n + 1$.
5. Retournons à l'exemple 1.2.6 à propos du sous-espace W de $\mathbb{R}^3$ donné par
$$W = \{(x_1, 2x_1 + x_2, x_1 - x_2) \ / \ x_1, x_2 \in \mathbb{R}\}.$$

D'après l'exemple 1.3.38(7), la famille $B = (v_1 = (1,2,1), v_2 = (0,1,-1))$ est une base de W. D'où W est de dimension 2. Remarquons que 2 est le nombre de paramètres x_1, x_2 qui définissent le sous-espace W. Ces deux paramètres x_1 et x_2 varient de façons arbitraires dans $\mathbb{R}$ et n'ont aucune relation entre eux. Pour bien fixer les idées, prenons un autre exemple avec des paramètres dépendant les unes des autres :

6. Soit $F = \{(x,y,z) \in \mathbb{R}^3 \ / \ x + y - 3z = 0\}$. On vérifie aisément que F est un sous-espace vectoriel de $\mathbb{R}^3$ dont nous allons chercher la dimension. Tout vecteur $v = (x,y,z)$ de F vérifie $x + y - 3z = 0$. Ici, les trois paramètres x, y, z ne sont pas tous quelconques, car on peut écrire par exemple x en fonction de y, z : $x = -y + 3z$. On va donc éliminer un paramètre, par exemple x, et laisser seulement deux paramètres, par exemple y et z. Ainsi, $v = (-y + 3z, y, z)$ où y, z sont deux paramètres arbitraires. Montrons alors que F est de dimension 2. On sépare les paramètres y et z dans le vecteur v :

$$\begin{aligned} v &= (-y + 3z, y, z) \\ &= (-y, y, 0) + (3z, 0, z) \\ &= y(-1,1,0) + z(3,0,1), \end{aligned}$$

qui est donc combinaison linéaire des vecteurs $a_1 = (-1,1,0), a_2 = (3,0,1)$. On vérifie facilement que a_1 et a_2 appartiennt à F, car chacun de ces deux vecteurs satisfait l'équation $x + y - 3z = 0$. On montre ensuite que (a_1, a_2) est libre. D'où $B = (a_1, a_2)$ est une base de $\mathbb{R}^2$, et par suite, $\dim(F) = 2$.

Remarque 1.4.12. **1.** Lorsque l'espace vectoriel E est de dimension infinie, on écrit par convention $\dim_{\mathbb{K}} E = \infty$.

2. Le singleton $E = \{0_E\}$ ne possède pas de base, car le vecteur 0_E est lié. Par abus d'écriture, on prend $\dim\{0_E\} = 0$.

3. En adoptant le vocabulaire de la géométrie :

- Un espace vectoriel de dimension 1 est de la forme $D = \mathsf{Vect}(u)$ avec $u \neq 0_E$. On l'appelle **droite vectorielle** engendrée par u. On dit aussi que c'est la *droite vectorielle dirigée par* le vecteur u. Observons que cette droite D est dirigée aussi par n'importe quel vecteur non nul $v \in \mathsf{Vect}(u)$. En effet, v est de la forme $v = \lambda u$ avec $\lambda \in \mathbb{K}^*$, et donc $\mathsf{Vect}(v) = \mathsf{Vect}(u) = D$.
- Un espace vectoriel de dimension 2 s'écrit $P = \mathsf{Vect}(u, v)$ avec (u, v) libre. Il s'appelle **plan vectoriel** engendré par u et v.

- Dans un espace vectoriel E de dimension n, tout sous-espace vectoriel de dimension $n-1$ est appelé **hyperplan** de E. Par exemple, les hyperplans de $\mathbb{R}^3$ sont les plans vectoriels alors que les hyperplans de $\mathbb{R}^2$ correspondent aux droites vectorielles.

Le corollaire 1.4.8 permet aussi de déduire :

Corollaire 1.4.13. *Soit E un espace vectoriel de dimension finie n. Pour toute partie libre $\mathcal{L}$ de E et pour toute partie génératrice $\mathcal{G}$ de E, on a*

$$\operatorname{card}(\mathcal{L}) \leq n \leq \operatorname{card}(\mathcal{G})$$

Autrement dit :

1. *Toute famille libre de E a au plus n éléments,*
2. *Toute famille génératrice de E a au moins n éléments.*

DÉMONSTRATION. Soit B une base de E. Puisque $\dim(E) = n$, alors B contient exactement n vecteurs. D'abord on applique le corollaire 1.4.8 à la partie libre $\mathcal{L}$ et la partie génératrice B pour trouver $\mathsf{card}(\mathcal{L}) \leq n$. La seconde inégalité $n \leq \operatorname{card}(\mathcal{G})$ s'obtient de même en considérant la partie libre B et la partie génératrice $\mathcal{G}$. ∎

Le corollaire qui précède peut s'exprimer de façon équivalente :

Corollaire 1.4.14. *Soit E un espace vectoriel de dimension finie n.*

1. *Toute famille qui a plus de n éléments de E est liée,*
2. *Toute famille qui a moins de n éléments de E n'engendre pas E.*

Exemple 1.4.15. On a vu que, dans l'espace vectoriel $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$, pour tout $n \geq 1$, la famille de fonctions $\{f_1, \ldots, f_n\}$, avec $f_k(x) = e^{kx}$, est libre. On en déduit que l'espace vectoriel $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$ est de dimension infinie. En effet, s'il était de dimension finie n, alors toute famille libre dans E doit être de cardinal $\leq n$. Mais $(f_1, \ldots, f_n, f_{n+1})$ est libre ! Contradiction.

Voici un autre moyen pour fabriquer des bases d'un espace vectoriel :

Théorème 1.4.16 (Théorème de la base incomplète)**.** *Soit E un espace vectoriel de dimension finie n. Si $\mathcal{L} = (a_1, \ldots, a_p)$ est une famille libre de E avec $p < n$, alors on peut la compléter en une base $B = (a_1, \ldots, a_p, a_{p+1}, \ldots, a_n)$ de E.*

Le théorème de la base incomplète est très important en Algèbre Linéaire ; il est utilisé dans de nombreuses situations

Démonstration. Comme $p < n$, alors la famille $\mathcal{L}$ n'engendre pas E, d'après le corollaire 1.4.14. Il existe alors un vecteur de E, que nous notons a_{p+1}, qui n'est pas combinaison linéaire de $a_1, \ldots, a_p$. Selon le théorème 1.3.32, la famille $(a_1, \ldots, a_p, a_{p+1})$ sera libre. Si cette nouvelle famlle $(a_1, \ldots, a_p, a_{p+1})$ engendre E, alors c'est une base de E et $p+1 = n$. Si ce n'est pas le cas, il existe un vecteur a_{p+2} qui n'est pas combinaison linéaire de $a_1, \ldots, a_p, a_{p+1}$. Donc $(a_1, \ldots, a_p, a_{p+1}, a_{p+2})$ sera libre, en vertu du théorème 1.3.32. On répète ce processus autant de fois que possible jusqu'à obtenir finalement une famille $B = (a_1, \ldots, a_p, a_{p+1}, a_{p+2}, \ldots, a_n)$ qui est libre. Cette dernière famille est nécessairement génératrice de E, car dans le cas contraire, il existera un vecteur $v_{n+1} \in E$ qui n'est pas combinaison linéaire de B, et donc $(a_1, \ldots, a_n, a_{n+1})$ sera libre d'après le théorème 1.3.32 ; et ceci contredit le corollaire 1.4.13, puisque le cardinal d'une famille libre ne peut dépasser la dimension. Ainsi, B est une base de E. ■

Soit E un espace vectoriel de dimension n. D'après le corollaire 1.4.14, une partie libre de cardinal $< n$ ne peut pas être génératrice de E. De même, une partie génératrice de cardinal $> n$ ne peut être libre. Que peut-on dire maintenant d'une partie libre ou génératrice de E ayant exactement n éléments ? la réponse est donnée par le résultat suivant qui est très souvent utilisé pour simplifier les calculs lorsqu'on désire prouver qu'une famille de vecteurs est une base d'un espace vectoriel de dimension finie dont on connait déjà la dimension :

Corollaire 1.4.17. *Soit E un espace vectoriel de dimension n, et $(a_1, \ldots, a_n)$ une famille à n vecteurs de E. Alors :*

$$(a_1, \ldots, a_n) \text{ est une base de } E \iff (a_1, \ldots, a_n) \text{ est libre } \iff (a_1, \ldots, a_n) \text{ engendre } E.$$

Démonstration.

- Supposons que la famille $\mathcal{F} = (v_1, \ldots, v_n)$ est libre. Pour montrer qu'elle engendre E, prenons un vecteur quelconque v de E. D'après le corollaire 1.4.14, la famille $(v_1, \ldots, v_n, v)$ est forcément liée, car elle est de cardinal $n+1 > n$. D'après le corollaire 1.3.33, $v \in \mathsf{Vect}(v_1, \ldots, v_n)$. Ainsi, v est combinaison linéaire de $v_1, \ldots, v_n$. D'où $(v_1, \ldots, v_n)$ est une base de E.
- Si la famille $\mathcal{F} = (v_1, \ldots, v_n)$ est génératrice de E, le théorème 1.4.4 affirme que cette famille contient une base B de E. Puisque $\dim(E) = n$, cette base B admet obligatoirement n vecteurs. Par conséquent, cette base B ne peut être que la famille $\mathcal{F}$ elle-même. ■

Il est naturel de penser que tout sous-espace vectoriel d'un espace vectoriel de dimension finie est lui-même de dimension finie :

Théorème 1.4.18. *Soient E un espace vectoriel de dimension finie n et F un sous-espace vectoriel de E. Alors :*

1. *F est lui-même de dimension finie et*

$$\dim_{\mathbb{K}}(F) \leq \dim_{\mathbb{K}}(E).$$

2. *En particulier, si $\dim_{\mathbb{K}}(F) = \dim_{\mathbb{K}}(E)$, alors $F = E$.*

Démonstration.

1. Toute famille libre d'éléments de F est aussi une famille libre dans E, elle a donc moins de n éléments d'après le corollaire 1.4.13. Considérons une famille libre $\mathcal{L} = (v_1, \ldots, v_p)$ d'éléments de F, de cardinal maximal p. On a donc $p \leq n$ et toute famille libre de F est de cardinal $\leq p$. Soit v un vecteur quelconque de F. La famille $(v_1, \ldots, v_p, v)$ ne peut pas être libre, car son cardinal $p + 1 > p$. Elle est donc liée, donc le corollaire 1.3.33 entraîne que v est combinaison linéaire de $(v_1, \ldots, v_p)$. La famille $B_F = (v_1, \ldots, v_p)$ est donc génératrice de F : c'est une base de F.
2. Si $\dim(F) = \dim(E)$, alors la base $B_F = (v_1, \ldots, v_p)$ de F est de cardinal $p = n$. En particulier, B_F est une famille libre dans E de cardinal $n = \dim(E)$. D'après le corollaire 1.4.17, B_F est une base de E. Il en découle que tout vecteur $u \in E$ est combinaison linéaire de B_F, et donc $u \in F$. D'où, $E = F$. ■

La seconde assertion de ce théorème est souvent utilisée pour démontrer que deux $\mathbb{K}$-espaces vectoriels E et F de dimensions finies sont égaux, à condition bien sûr que F soit inclus dans E.

Exemple 1.4.19. Soit F le sous-ensemble de $\mathbb{R}^4$ défini par

$$F = \{(x_1, x_2, x_1 + x_2, 2x_1 + x_2) \ / \ x_1 \in \mathbb{R}, x_2 \in \mathbb{R}\}.$$

On montre facilement que F est un sous-espace vectoriel de $\mathbb{R}^4$. La représentation de F est dite paramétrique, car chaque vecteur $x = (x_1, x_2, x_1 + x_2, 2x_1 + x_2) \in F$ est déterminé à l'aide des paramètres x_1, x_2. En vertu du théorème 1.4.18, $\dim(F) \leq \dim(\mathbb{R}^4) = 4$.

- Pour trouver la dimension de F, on doit connaître une base de F. Pour ce faire, cherchons une famille génératrice de F.

Soit $x = (x_1, 3x_2, x_1 + x_2, 2x_1 + x_2) \in F$. On a :

$$\begin{aligned} x &= (x_1, 0, x_1, 2x_1) + (0, 3x_2, x_2, x_2) \\ &= x_1(1,0,1,2) + x_2(0,3,1,1) \end{aligned}$$

qui est donc combinaison linéaire de $a_1 = (1,0,1,2)$ et $a_2 = (0,3,1,1)$. Remarquons que a_1 et a_2 appartiennent bien à F, puisque $a_1 = (y_1, 3y_2, y_1 + y_2, 2y_1 + y_2)$ avec $y_1 = 1, y_2 = 0$, et $a_2 = (z_1, 3z_2, z_1 + z_2, 2z_1 + z_2)$ avec $z_1 = 0, t_2 = 1$. D'où $F = \mathsf{Vect}(a_1, a_2)$, c'est-à-dire, (a_1, a_2) est une famille génératrice de F.

Montrons maintenant que cette famille (a_1, a_2) est libre. Soient α_1 et α_2 dans $\mathbb{R}$ tels que $\alpha_1 a_1 + \alpha_2 a_2 = 0_{\mathbb{R}^4}$. Alors les calculs simples nous donnent $(\alpha_1, 3\alpha_2, \alpha_1 + \alpha_2, 2\alpha_1 + \alpha_2) = (0,0,0,0)$, d'où le système linéaire

$$\begin{cases} \alpha_1 = 0 \\ 3\alpha_2 = 0 \\ \alpha_1 + \alpha_2 = 2 \\ 2\alpha_1 + \alpha_2 = 0 \end{cases}$$

On obtient donc nécessairement $\alpha_1 = \alpha_2 = 0$. Notons qu'on pouvait remarquer simplement que a_1 et a_2 ne sont pas proportionnels, pour dire qu'ils sont linéairement indépendants. Par conséquent (a_1, a_2) est une base de F, ce qui entraîne $\dim(F) = 2$. Le sous-espace vectoriel F est donc un plan vectoriel de $\mathbb{R}^4$.

– La famille (a_1, a_2) est libre dans $\mathbb{R}^4$, mais ce n'est pas une base de $\mathbb{R}^4$. Complétons (a_1, a_2) en une base de $\mathbb{R}^4$, en nous inspirant du théorème de la base incomplète. Comme $\dim(\mathbb{R}^4) = 4$, il manque deux vecteurs dans (a_1, a_2). Mais comment compléter concrètement une famille libre en une base de E ? Pour satisfaire la curiosité des étudiants, nous affirmons qu'il y a une infinité de possibilités pour choisir deux vecteurs convenables a_3 et a_4 de $\mathbb{R}^4$. Prenons par exemple $a_3 = (0,0,1,0)$ et $a_4 = (0,0,0,1)$ issus de la base canonique de $\mathbb{R}^4$, et regardons si (a_1, a_2, a_3, a_4) est libre. Soient $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ dans $\mathbb{R}$ tels que $\alpha_1 a_1 + \alpha_2 a_2 + \alpha_3 a_3 + \alpha_4 a_4 = 0_E$. Alors le système

$$\begin{cases} \alpha_1 = 0 \\ 3\alpha_2 = 0 \\ \alpha_1 + \alpha_2 + \alpha_3 = 2 \\ 2\alpha_1 + \alpha_2 + \alpha_4 = 0 \end{cases}$$

donne aisément $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$. Par conséquent, (a_1, a_2, a_3, a_4) est libre ; et comme son cardinal 4 est égal à la dimension de l'espace $\mathbb{R}^4$, alors le corollaire 1.4.17 assure que (a_1, a_2, a_3, a_4) est une base de $\mathbb{R}^4$.

Nous achevons cette sous-section par l'étude de la dimension de l'espace vectoriel produit.

Exemple 1.4.20. Soient E_1 et E_2 deux espaces vectoriels de dimensions finies sur le même corps $\mathbb{K}$. Montrons que l'espace vectoriel produit $E_1 \times E_2$ est de dimension finie et que

$$\boxed{\dim(E_1 \times E_2) = \dim(E_1) + \dim(E_2)}$$

Pour cela, soient $B_1 = (u_1, \ldots, u_n)$ une base de E_1 et et $B_2 = (v_1, \ldots, v_p)$ une base de E_2. Construisons une base de $E_1 \times E_2$.

- On commence toujours par chercher une partie génératrice de $E_1 \times E_2$: Soit $x = (x_1, x_2)$ un vecteur quelconque de $E_1 \times E_2$. Comme B_1 engendre E_1 et B_2 engendre E_2, il existe des scalaires $\alpha_1, \ldots, \alpha_n$ et $\beta_1, \ldots, \beta_p$ dans $\mathbb{K}$ tels que $x_1 = \alpha_1 u_1 + \cdots + \alpha_n u_n$ et $x_2 = \beta_1 v_1 + \cdots + \beta_p v_p$. Il en résulte que :

$$\begin{aligned} x &= (\alpha_1 u_1 + \cdots + \alpha_n u_n,\ \beta_1 v_1 + \cdots + \beta_p v_p) \\ &= (\alpha_1 u_1 + \cdots + \alpha_n u_n, 0_{E_2}) + (0_{E_1}, \beta_1 v_1 + \cdots + \beta_p v_p) \\ &= (\alpha_1 u_1, 0_{E_2}) + \cdots + (\alpha_n u_n, 0_{E_2}) + (0_{E_1}, \beta_1 v_1) + \cdots + (0_{E_1}, \beta_p v_p) \\ &= \alpha_1 (u_1, 0_{E_2}) + \cdots + \alpha_n (u_n, 0_{E_2}) + \beta_1 (0_{E_1}, v_1) + \cdots + \beta_p (0_{E_1}, v_p). \end{aligned}$$

  Donc $B = \{(u_1, 0_{E_2}), \ldots, (u_n, 0_{E_2})\} \bigcup \{(0_{E_1}, v_1), \ldots, (0_{E_1}, v_p)\}$ est une partie génératrice de $E_1 \times E_2$.

- Montrons que cette partie B est libre. Soient $\alpha_1, \ldots, \alpha_n$ et $\beta_1, \ldots, \beta_p$ dans $\mathbb{K}$ tels que $\alpha_1(u_1, 0_{E_2}) + \cdots + \alpha_n(u_n, 0_{E_2}) + \beta_1(0_{E_1}, u_1) + \cdots + \beta_p(0_{E_1}, v_p) = 0_E = (0_{E_1}, 0_{E_2})$. En remontant dans les calculs précédents, il vient

$$(\alpha_1 u_1 + \cdots + \alpha_n u_n,\ \beta_1 v_1 + \cdots + \beta_n v_p) = (0_{E_1}, 0_{E_2}),$$

  d'où

$$\alpha_1 u_1 + \cdots + \alpha_n u_n = 0_{E_1} \quad \text{et} \quad \beta_1 v_1 + \cdots + \beta_p v_p = 0_{E_2}.$$

  Puisque B_1 est libre dans E_1 et B_2 est libre dans E_2, nous déduisons

$$\alpha_1 = \cdots = \alpha_n = 0 \quad \text{et} \quad \beta_1 = \cdots = \beta_p = 0.$$

Il en découle que B est une base de $E_1 \times E_2$, d'où $\dim(E_1 \times E_2) = \mathsf{card}(B) = n + p = \dim(E_1) + \dim(E_2)$.

On peut généraliser par récurrence la formule précédente à plusieurs $\mathbb{K}$-espaces vectoriels $E_1, E_2, \ldots, E_k$:

$$\boxed{\dim(E_1 \times E_2 \times \cdots \times E_k) = \dim(E_1) + \dim(E_2) + \cdots + \dim(E_k)}$$

Il suffit d'employer la formule $E_1 \times \cdots \times E_{k-1} \times E_k = (E_1 \times \cdots \times E_{k-1}) \times E_k$.

Exemple 1.4.21. Rappelons que $\dim_{\mathbb{K}}(\mathbb{K}) = 1$, et donc les seuls sous-espaces vectoriels de $\mathbb{K}$ sont les sous-espaces vectoriels triviaux $\{0_{\mathbb{K}}\}$ et $\mathbb{K}$. La formule précédente de la dimension d'un espace vectoriel produit permet de retrouver que $\dim_{\mathbb{K}}(\mathbb{K}^n) = n$. En effet,

$$\dim_{\mathbb{K}}(\mathbb{K}^n) = \dim_{\mathbb{K}}\Big(\underbrace{\mathbb{K} \times \cdots \times \mathbb{K}}_{n\,\texttt{fois}}\Big) = \dim_{\mathbb{K}}(\mathbb{K}) + \cdots + \dim_{\mathbb{K}}(\mathbb{K}) = 1 + \cdots + 1 = n.$$

1.4.3 Rang d'une famille finie de vecteurs

Définition 1.4.22. *Soient E un espace vectoriel sur K et $\mathcal{F}$ une famille finie de vecteurs de E. On appelle **rang** de $\mathcal{F}$, et l'on note $\mathsf{rg}(\mathcal{F})$, la dimension du sous-espace vectoriel $\mathsf{Vect}(\mathcal{F})$ de E engendré par $\mathcal{F}$.*
En notant $\mathcal{F} = (v_1, \ldots, v_p)$, on a :

$$\boxed{\mathsf{rg}(v_1, \ldots, v_p) = \dim\left(\mathsf{Vect}(v_1, \ldots, v_p)\right)}$$

Il est clair que l'ordre des vecteurs $v_1, \ldots, v_p$ dans la famille $(v_1, \ldots, v_p)$ n'a pas d'importance dans le calcul du rang de cette famille.

Remarque 1.4.23. Soient $\mathcal{F}$ et $\mathcal{F}'$ deux familles finies de vecteurs de E.
Si $\mathcal{F} \subseteq \mathcal{F}'$, alors $\mathsf{rg}(\mathcal{F}) \leq \mathsf{rg}(\mathcal{F}')$.
En effet, selon la proposition 1.3.14, l'inclusion $\mathcal{F} \subseteq \mathcal{F}'$ des familles implique l'inclusion $\mathsf{Vect}(\mathcal{F}) \subseteq \mathsf{Vect}(\mathcal{F}')$ des sous-espaces. Donc on peut dire que $\mathsf{Vect}(\mathcal{F})$ est un sous-espace vectoriel de l'espace vectoriel $\mathsf{Vect}(\mathcal{F}')$. Finalement, en appliquant le théorème 1.4.18, il vient $\dim(\mathsf{Vect}(\mathcal{F})) \leq \dim(\mathsf{Vect}(\mathcal{F}'))$. D'où le résultat.

Avant d'énoncer quelques propriétés du rang, traitons un exemple simple pour comprendre comment calculer le rang d'une famille.

Exemple 1.4.24. Dans l'espace vectoriel $E = \mathbb{R}^4$, déterminons le rang de la famille $\mathcal{F} = (v_1, v_2, v_3)$ où $v_1 = (1, 1, 0, -1)$, $v_2 = (1, 2, 1, 0)$ et $v_3 = (3, 5, 2, -1)$.

On a $\mathsf{rg}(v_1, v_2, v_3) = \dim(\mathsf{Vect}(v_1, v_2, v_3))$. Cherchons donc une base du sous-espace vecoriel $F = \mathsf{Vect}(v_1, v_2, v_3)$. Il est évident que la famille (v_1, v_2, v_3) engendre F ; mais est-elle libre ? Un calcul facile montre qu'elle est liée, puisque $v_3 = v_1 + 2v_2$. Donc (v_1, v_2, v_3) n'est pas une base de F. Mais le théorème 1.4.4 affirme que (v_1, v_2, v_3) contient une base de F. Plus précisément, puisque v_3 est combinaison linéaire de v_1 et v_2, le corollaire 1.3.17 nous permet d'éliminer le vecteur v_3 :

$$\mathsf{Vect}(v_1, v_2, v_3) = \mathsf{Vect}(v_1, v_2).$$

Ainsi, $\mathsf{rg}(v_1, v_2, v_3) = \dim(\mathsf{Vect}(v_1, v_2, v_3)) = \dim(\mathsf{Vect}(v_1, v_2)) = \mathsf{rg}(v_1, v_2)$. La sous-famille (v_1, v_2) engendre donc F. Or, il est facile de vérifier que (v_1, v_2) est libre, parce que les deux vecteurs v_1 et v_2 ne sont pas proportionnels. Par conséquent, (v_1, v_2) est une base de F. D'où $\mathsf{rg}(v_1, v_2, v_3) = \dim(F) = 2$.

Proposition 1.4.25. *Soient E un espace vectoriel sur $\mathbb{K}$ et $\mathcal{F} = (v_1, \ldots, v_p)$ une famille de p vecteurs de E. Alors :*

1. $\mathsf{rg}(v_1, \ldots, v_p) \leq p$. *C'est-à-dire, le rang d'une famille de p vecteurs est au plus égal à p.*
2. *Si E est de dimension finie, on a* $\mathsf{rg}(v_1, \ldots, v_p) \leq \dim(E)$.

DÉMONSTRATION. Notons $F = \mathsf{Vect}(\mathcal{F})$, et donc $\mathsf{rg}(\mathcal{F}) = \dim(F)$.

1. D'une part, $\mathcal{F} = (v_1, \ldots, v_p)$ est une famille génératrice du sous-espace vectoriel F ayant p éléments. D'autre part, le corollaire 1.4.13 atteste que le cardinal p de la famille génératrice $\mathcal{F}$ de l'espace vectoriel F est $\geq \dim(F)$. Il s'ensuit que $\mathsf{rg}(\mathcal{F}) \leq p$.
2. F étant un sous-espace vectoriel de E, le théorème 1.4.18 exige nécessairement $\dim(F) \leq \dim(E)$. Ceci entraîne $\mathsf{rg}(\mathcal{F}) \leq \dim(E)$. ■

La proposition précédente peut s'écrire :

$$\boxed{\mathsf{rg}(\mathcal{F}) \leq \min\Big(\mathsf{card}(\mathcal{F}), \dim(E)\Big)}$$

Proposition 1.4.26. *Soient E un espace vectoriel sur $\mathbb{K}$ et $\mathcal{F} = (v_1, \ldots, v_p)$ une famille de p vecteurs de E. Alors :*

1. $\mathsf{rg}(v_1, \ldots, v_p) = p \iff (v_1, \ldots, v_p)$ *est libre.*

2. $\mathsf{rg}(v_1, \ldots, v_p)$ *est égal au cardinal maximum d'une sous-famille libre de $\mathcal{F}$.*

Démonstration.

1. • Si $\mathcal{F} = (v_1, \ldots, v_p)$ est libre, comme elle engendre déjà le sous-espace $\mathsf{Vect}(\mathcal{F})$, c'est une base de $\mathsf{Vect}(\mathcal{F})$. D'où

$$\mathsf{rg}(\mathcal{F}) = \dim(\mathsf{Vect}(\mathcal{F})) = \mathsf{card}(\mathcal{F}) = p.$$

   • Réciproquement, si $\mathsf{rg}(\mathcal{F}) = p$, alors $\dim(\mathsf{Vect}(\mathcal{F})) = p$. Donc $\mathcal{F}$ est une famille génératrice de $\mathsf{Vect}(\mathcal{F})$ dont le cardinal p est égal à la dimension de $\mathsf{Vect}(\mathcal{F})$. Le corollaire 1.4.17. affirme alors que $\mathcal{F}$ est une base de $\mathsf{Vect}(\mathcal{F})$, et donc $\mathcal{F}$ est libre.

2. Puisque $\mathcal{F}$ est une famille génératrice du sous-espace $\mathsf{Vect}(\mathcal{F})$, le théorème 1.4.4 nous dit que $\mathcal{F}$ contient une base $B = (v_{i_1}, \ldots, v_{i_r})$ de $\mathsf{Vect}(\mathcal{F})$. Donc B est une sous-famille de $\mathcal{F}$ qui est une base de $\mathsf{Vect}(\mathcal{F})$. Il en résulte que $\mathsf{card}(B) = \dim(\mathsf{Vect}(\mathcal{F})) = \mathsf{rg}(\mathcal{F})$.
   Soit maintenant L une sous-famille libre quelconque de $\mathcal{F}$. Nous allons voir que $\mathsf{card}(L) \leq \mathsf{card}(B)$. En effet, par le corollaire 1.4.13, $\mathsf{card}(L) \leq \dim(\mathsf{Vect}(\mathcal{F}))$. D'où $\mathsf{card}(L) \leq \mathsf{rg}(\mathcal{F}) = \mathsf{card}(B)$. Par conséquent, parmi les sous-familles libres de $\mathcal{F}$, la sous-famille libre B admet le plus grand cardinal. ■

Remarque 1.4.27. Le rang de $(v_1, \ldots, v_p)$ est donc le nombre maximum d'éléments d'une famille libre extraite de $(v_1, \ldots, v_p)$.
Par conséquent, $\mathsf{rg}(v_1, \ldots, v_p) = r$ si et seulement si il existe une famille libre de r vecteurs extraite de $(v_1, \ldots, v_p)$, et si toute famille extraite de $(v_1, \ldots, v_p)$ de cardinal $> r$ est liée.
Enfin, si $\mathsf{rg}(v_1, \ldots, v_p) = r$, alors toute famille libre de r éléments extraite de $(v_1, \ldots, v_p)$ est une base du sous-espace vectoriel engendré par les vecteurs $v_1, \ldots, v_p$.

Exemple 1.4.28. Retournons à l'exemple ci-dessus dans lequel la famille $\mathcal{F} = (v_1, v_2, v_3)$ était liée. Donc n'importe quelle famille libre extraite de $\mathcal{F}$ est nécessairement de cardinal < 3. Comme (v_1, v_2) est une famille libre extraite de $\mathcal{F}$ qui est de cardinal 2, la proposition précédente nous confirme que $\mathsf{rg}(v_1, v_2, v_3) = 2$.
Dans cet exemple, on peut prendre aussi (v_2, v_3) (ou encore (v_1, v_3)) qui est aussi une famille libre extraite de $\mathcal{F}$ de cardinal maximum 2.

Exemple 1.4.29. Dans l'espace vectoriel $E = \mathbb{R}^3$, considérons la famille $\mathcal{F} = (v_1, v_2, v_3, v_4)$ avec $v_1 = (1, -1, 1)$, $v_2 = (-1, 1, -1)$, $v_3 = (0, 1, 1)$, $v_4 = (1, 0, 2)$. D'après la proposition 1.4.24,

$$\mathsf{rg}(\mathcal{F}) \leq \min\Big(\mathsf{card}(\mathcal{F}), \dim(\mathbb{R}^4)\Big) = \min(4, 3) = 3.$$

Le rang de $\mathcal{F}$ ne peut donc pas être égal à 4. Ceci peut être expliqué d'une autre manière sans faire aucun calcul : la famille $\mathcal{F}$ est liée car son cardinal 4 est supérieur à la dimension 3 de l'espace vectoriel $\mathbb{R}^3$. D'autre part, on remarque que $v_2 = -v_1$ et $v_4 = v_1 + v_3$. Donc $\mathsf{Vect}(v_1, v_2, v_3, v_4) = \mathsf{Vect}(v_1, v_3)$. Comme (v_1, v_3) est libre, c'est une famille libre extraite de $\mathcal{F}$ de cardinal maximum. D'où $\mathsf{rg}(v_1, v_2, v_3, v_4) = 2$. Observons qu'on ne peut pas extraire de $\mathcal{F}$ une famille libre de cardinal 3.

Nous retiendrons donc que la détermination du rang d'une famille finie $\mathcal{F}$ revient à extraire de $\mathcal{F}$ une sous-famille libre qui contient le maximum de vecteurs. Du point de vue théorique, cette méthode est intéressante ; mais pour le côté pratique, elle n'est pas très efficace, surtout quand le cardinal p de la famille est assez grand. Nous verrons au chapitre 3 un algorithme efficace, appelé *méthode de Gauss*, permettant de calculer plus facilement le rang d'une famille finie de vecteurs. C'est ce moyen systématique qui sera le plus souvent utilisé dans la pratique.

1.5 Somme de sous-espaces vectoriels

Nous avons déjà observé dans l'exemple 1.2.17 que la réunion de sous-espaces vectoriels d'un espace vectoriel E n'est pas en général un sous-espace vectoriel de E. En fait, pour contourner ce problème, la réunion devra être remplacée par le concept suivant :

Définition 1.5.1. *Soient $F_1, \cdots, F_p$ des sous-espaces vectoriels d'un espace vectoriel E. On appelle **somme** des sous-espaces vectoriels $F_1, \ldots, F_p$, et l'on note, $F_1 + \cdots + F_p$, ou encore $\sum\limits_{i=1}^{n} F_i$, l'ensemble de tous les vecteurs de E de la forme $x_1 + \cdots + x_p$, où $x_i \in F_i$ pour tout i.*

En d'autres termes,

$$\boxed{F_1 + \cdots + F_p = \{x_1 + \cdots + x_p \ / \ x_1 \in F_1, \ldots, x_p \in F_p\}}$$

Donc un vecteur x de E appartient à $F_1+\cdots+F_p$ si et seulement si il existe des vecteurs $x_1 \in F_1,\ldots,x_p \in F_p$ tels que $x = x_1+\cdots+x_p$.

En particulier, la somme F_1+F_2 de deux sous-espaces vectoriels de E est définie par $F_1+F_2 = \{x_1+x_2 \ / \ x_1 \in F_1, x_2 \in F_2\}$.

Par analogie avec le théorème 1.3.8, on va montrer que l'ensemble $F_1+\cdots+F_p$ est le plus petit sous-espace vectoriel de E contenant $F_1,\ldots,F_p$:

Proposition 1.5.2. *Soient $F_1,\ldots,F_p$ des sous-espaces vectoriels de E. Alors :*

1. $F_1+\cdots+F_p$ est un sous-espace vectoriel de E.

2. $F_1+\cdots+F_p$ contient $F_1,\ldots,F_p$.

3. Si G est un sous-espace vectoriel de E contenant $F_1,\ldots,F_p$, alors G contient $F_1+\cdots+F_p$.

DÉMONSTRATION.

1.
- D'abord $0_E = 0_E+\cdots+0_E \in F_1+\cdots+F_p$ car $0_E \in F_i$ pour tout i.
- Soient $x, y \in F_1+\cdots+F_p$ et $\alpha, \beta \in K$. Il existe des vecteurs $x_1 \in F_1,\cdots,x_p \in F_p$ et $y_1 \in F_1,\cdots,y_p \in F_p$ tels que $x = x_1+\cdots+x_p$ et $y = y_1+\cdots+y_p$. D'après les règles de calculs dans l'espace vectoriel E, $\alpha x+\beta y = (\alpha x_1+\beta y_1)+\cdots+(\alpha x_p+\beta y_p)$. Or, chaque F_i étant un sous-espace vectoriel de E, donc $\alpha x_i+\beta y_i \in F_i$. D'où $\alpha x+\beta y \in F_1+\cdots+F_p$, et par suite, $F_1+\cdots+F_p$ est bien un sous-espace vectoriel de E.

2. Pour tout $i = 1,\ldots,p$ et pour tout $x_i \in F_i$, on peut écrire $x_i = 0_E+\cdots+x_i+\cdots+0_E \in F_1+\cdots+F_p$, car $0_i \in F_1,\cdots,x_i \in F_i,\ldots,0_E \in F_p$. Donc $F_i \subseteq F_1+\cdots+F_p$. D'où $F_1+\cdots+F_p$ contient chacun des sous-espaces $F_1,\cdots,F_p$.

3. Soit G un sous-espace vectoriel de E contenant $F_1,\ldots,F_p$. Pour prouver que G contient $F_1+\cdots+F_p$, prenons un vecteur quelconque $x \in F_1+\cdots+F_p$. Il s'écrit sous forme $x = x_1+\cdots+x_p$ avec $x_1 \in F_1,\ldots,x_p \in F_p$. Puisque $F_1 \subseteq G,\ldots F_p \subseteq G$, alors $x_1 \in G,\ldots,x_p \in G$. Mais comme G est un sous-espace vectoriel de E, alors $x_1+\cdots+x_p \in G$, d'où $x \in G$. ■

La proposition précédente exprime que $F_1+\cdots+F_p$ est le plus petit sous-espace vectoriel de E contenant $F_1 \cup\cdots\cup F_p$.

Remarque 1.5.3. La commutativité et l'associativité de l'addition des vecteurs dans l'espace vecoriel E entraîne immédiatement la commutativité et l'associativité de la somme des sous-espaces vectoriels de E :

$$F_1 + F_2 = F_2 + F_1, \qquad (F_1 + F_2) + F_3 = F_1 + (F_2 + F_3).$$

Il est évident aussi que $F + \{0_E\} = F$ pour tout sous-espace F de E.
Mais attention, nous avons $F + F = F$. En effet, on a déjà $F \subseteq F + F$ car tout vecteur $x \in F$ s'écrit $x = x + 0_E \in F + F$. Réciprquement, si $x \in F + F$, il existe $x_1, x_2 \in F$ tels que $x = x_1 + x_2$. Comme F est un sous-espace de E, alors $x \in F$, d'où $F + F \subseteq F$.
Notons aussi que, par exemple, $F_1 + F_2 \subseteq F_1 + F_2 + F_3$, car tout vecteur $x \in F_1 + F_2$ s'écrit $x = x_1 + x_2$ avec $x_1 \in F_1$ et $x_2 \in F_2$, et donc $x = x_1 + x_2 + 0_E$ avec $0_E \in F_3$. Plus généralement, si $J \subseteq I = \{1, \ldots, p\}$, alors $\sum\limits_{i \in J} F_i \subseteq \sum\limits_{i \in I} F_i$.

Proposition 1.5.4. *Soit F_1 et F_2 deux sous-espaces de E. Si $A_1 = \{a_1, \ldots, a_n\}$ engendre F_1 et $A_2 = \{b_1, \ldots, b_m\}$ engendre F_2, alors $A_1 \cup A_2 = \{a_1, \ldots, a_n, b_1, \ldots, b_m\}$ engendre $F_1 + F_2$.*

DÉMONSTRATION. Si $x \in F_1 + F_2$, il existe $x_1, x_2 \in F$ tels que $x = x_1 + x_2$. Or, x_1 se décompose comme combinaison linéaire de éléments de A_1 : $x_1 = \alpha_1 a_1 + \cdots + \alpha_n a_n$, avec $\alpha_1, \ldots, \alpha_n \in K$. De même, $x_2 = \beta_1 b_1 + \cdots + \beta_m b_m$, avec $\beta_1, \ldots, \beta_m \in K$. D'où $x = \alpha_1 a_1 + \cdots + \alpha_n a_n + \beta_1 b_1 + \cdots + \beta_m b_m$ est combinaison linéaire des vecteurs de $A_1 \cup A_2$. ■

Remarque 1.5.5. Cette propriété s'étend facilement à un nombre quelconque p de sous-espaces vectoriels $F_1, \ldots, F_p$ de E :
Si $F_i = \mathsf{Vect}(A_i)$ pour tout $i = 1, \ldots, p$, alors $F_1 + \cdots + F_p = \mathsf{Vect}(A_1 \cup \cdots \cup A_p)$, ou de manière équivalente,

$$\boxed{\mathsf{Vect}(A_1 \cup \cdots \cup A_p) = \mathsf{Vect}(A_1) + \cdots + \mathsf{Vect}(A_p)}$$

Remarque 1.5.6. Soient $F_1, \ldots, F_p$ des sous-espaces vectoriels de E. Parfois, il se peut que le sous-espace $F_1 + \cdots + F_p$ soit égal à E tout entier. Cette situation s'écrit :

$$E = F_1 + \cdots + F_p \iff E \subseteq F_1 + \cdots + F_p$$

$$\iff \forall x \in E \quad \exists\, (x_1, \ldots, x_p) \in F_1 \times \cdots \times F_p : x = x_1 + \cdots + x_p.$$

En effet, on a toujours $F_1 + \cdots + F_p \subseteq E$ car $F_1 + \cdots + F_p$ est un sous-espace de E. Donc l'égalité $E = F_1 + \cdots + F_p$ est équivalente à l'inclusion $E \subseteq F_1 + \cdots + F_p$, c'est-à-dire, tout élément x de E peut s'écrire sous la forme $x_1 + \cdots + x_p$ avec $(x_1, \ldots, x_p) \in F_1 \times \cdots \times F_p$.

Nous voulons à présent généraliser pour les sous-espaces vectoriels la notion de vecteurs linéairement indépendants. Soient $F_1, \ldots, F_p$ des sous-espaces d'un espace vectoriel E. Nous savons que tout vecteur $x \in F_1 + \cdots + F_p$ peut s'écrire sous la forme $x = x_1 + \cdots + x_p$, avec $x_1 \in F_1, \ldots, x_p \in F_p$. Cette représentation de x n'est pas unique en général. Pour le voir, prenons l'exemple suivant avec $p = 2$ pour simplifier :

Exemple 1.5.7. Soient $E = \mathbb{R}^3$, $F_1 = \{(a, a, b) \ / \ a, b \in \mathbb{R}\}$, $F_2 = \{(c, d, d) \ / \ c, d \in \mathbb{R}\}$. Une vérification de routine montre que F_1 et F_2 sont des sous-espaces vectoriels de E. De plus, le vecteur $x = (1, 2, 1)$ appartient à $F_1 + F_2$ car $(1, 2, 1) = (1, 1, 0) + (0, 1, 1)$ avec $(1, 1, 0) \in F_1$ et $(0, 1, 1) \in F_2$. Mais ce même vecteur x s'écrit aussi d'une autre manière : $x = (0, 0, -1) + (1, 2, 2)$ avec $(0, 0, -1) \in F_1$ et $(1, 2, 2) \in F_2$.

Pour remédirer à cet inconvénient, nous devons introduire la notion de *somme directe* qui s'impose naturellement. Soulignons d'abord que le vecteur 0_E appartient toujours au sous-espace $F_1 + \cdots + F_p$ puisque $0_E = 0_E + \cdots + 0_E$. Peut-on avoir une autre décomposition $0_E = x_1 + \cdots + x_p$ avec $x_1 \in F_1, \ldots, x_p \in F_p$?

Définition 1.5.8. *Soient $F_1, \ldots, F_p$ des sous-espaces vectoriels de E. On dit que $F_1 + \cdots + F_p$ est une **somme directe** si la seule écriture du vecteur 0_E dans $F_1 + \cdots + F_p$ est l'écriture triviale $0_E = 0_E + \cdots + 0_E$. C'est-à-dire :*

$$\boxed{\forall x_1 \in F_1, \ldots, \forall x_p \in F_p : \quad x_1 + \cdots + x_p = 0_E \Longrightarrow x_1 = \cdots = x_p = 0_E}$$

Dans ce cas, la somme $F_1 + \cdots + F_p$ est notée $F_1 \oplus \cdots \oplus F_p$.

Exemple 1.5.9. Dans $E = \mathbb{R}^3$, considérons les sous-espaces vectoriels

$$F_1 = \{(\alpha, 0, 0) \ / \ \alpha \in \mathbb{R}\}, \ F_2 = \{(\beta, \beta, 0) \ / \ \beta \in \mathbb{R}\}, \ F_3 = \{(\gamma, \gamma, \gamma) \ / \ \gamma \in \mathbb{R}\}.$$

Montrons que la somme $F_1 + F_2 + F_3$ est directe. Pour ce faire, prenons $x_1 = (\alpha, 0, 0) \in F_1$, $x_2 = (\beta, \beta, 0) \in F_2$ et $x_3 = (\gamma, \gamma, \gamma) \in F_3$ tels que $x_1 + x_2 + x_3 = 0_{\mathbb{R}^3}$. Alors on obtient le système linéaire

$$\begin{cases} \alpha + \beta + \gamma = 0 \\ \qquad \beta + \gamma = 0 \\ \qquad\qquad \gamma = 0 \end{cases}$$

qui donne aisément $\alpha = \beta = \gamma = 0$. D'où $x_1 = x_2 = x_3 = (0, 0, 0) = 0_{\mathbb{R}^3}$.

La proposition ci-dessous caractérise les sommes directes et répond à l'exigence de l'unicité discutée précédemment :

Proposition 1.5.10. *Soient $F_1, \ldots, F_p$ des sous-espaces vectoriels de E. Alors la somme $F_1 + \cdots + F_p$ est directe si et seulement si tout vecteur $x \in F_1 + \cdots + F_p$ s'écrit de manière unique sous la forme $x = x_1 + \cdots + x_p$, avec $x_1 \in F_1, \ldots, x_p \in F_p$.*

Démonstration.

- Supposons que la somme $F_1 + \cdots + F_p$ est directe et soit $x \in F_1 + \cdots + F_p$. Si $x = x_1 + \cdots + x_p = y_1 + \cdots + y_p$, avec $x_1, y_1 \in F_1, \ldots, x_p, y_p \in F_p$, alors $(x_1 - y_1) + \cdots + (x_p - y_p) = 0_E$. Puisque $x_1 - y_1 \in F_1, \ldots, x_p - y_p \in F_p$, alors l'hypothèse implique $x_1 - y_1 = \cdots = x_p - y_p = 0_E$. D'où $x_1 = y_1, \ldots, x_p = y_p$.
- Réciproquement, supposons que tout vecteur $x \in F_1 + \cdots + F_p$ s'écrit de manière unique sous la forme $x = x_1 + \cdots + x_p$, avec $x_1 \in F_1, \ldots, x_p \in F_p$. Alors, en particulier pour le vecteur $x = 0_E$, il s'écrit dans $F_1 + \cdots + F_p$ sous la forme unique $0_E = 0_E + \cdots + 0_E$. Par conséquent, la somme $F_1 + \cdots + F_p$ est directe.

■

Ainsi, à titre d'exemple, cette proposition nous dit que pour les deux sous-espaces vectoriels F_1 et F_2 de l'exemple 1.5.7, la somme $F_1 + F_2$ n'est pas directe.

Dans le cas de deux sous-espaces F_1 et F_2, on a la caractérisation suivante de la somme directe $F_1 \oplus F_2$ à l'aide de l'intersection $F_1 \cap F_2$:

Proposition 1.5.11. *Soient F_1 et F_2 deux sous-espaces vectoriels de E.*
La somme $F_1 + F_2$ est directe $\iff$ $F_1 \cap F_2 = \{0_E\}$.

Démonstration.

- Supposons que la somme $F_1 + F_2$ est directe et montrons que $F_1 \cap F_2 = \{0_E\}$. Soit $x \in F_1 \cap F_2$. On a $x + (-x) = 0_E$ avec $x \in F_1$ et $-x \in F_2$. Or, la seule écriture de 0_E dans la somme $F_1 \oplus F_2$ est $0_E = 0_E + 0_E$. On en déduit que $x = 0_E$ (et $-x = 0_E$). D'où $F_1 \cap F_2 = \{0_E\}$.
- Supposons $F_1 \cap F_2 = \{0_E\}$ et montrons que la somme $F_1 + F_2$ est directe. Soit $x_1 \in F_1$ et $x_2 \in F_2$ tels que $x_1 + x_2 = 0_E$. Alors $x_1 = -x_2 \in F_1 \cap F_2$, ce qui donne $x_1 = x_2 = 0_E$.

Exemple 1.5.12. Dans $E = \mathbb{R}^3$, considérons les sous-espaces vectoriels $F_1 = \{(\alpha, 0, 0) \ / \ \alpha \in \mathbb{R}\}$, $F_2 = \{(\beta, \beta, 0) \ / \ \beta \in \mathbb{R}\}$. Pour montrer que la somme $F_1 + F_2$ est directe, il suffit de prouver que $F_1 \cap F_2 = \{0_E\}$. Soit donc un vecteur $x \in F_1 \cap F_2$. Il existe α et β dans $\mathbb{R}$ tels que $x = (\alpha, 0, 0) = (\beta, \beta, 0)$. Donc $\alpha = \beta$ et $0 = \beta$, d'où $x = (0, 0, 0) = 0_E$. Il s'ensuit que $F_1 \cap F_2 = \{0_E\}$.

Remarque 1.5.13. La proposition 1.5.11 ne se généralise pas pour $p \geq 3$. C'est-à-dire, si $F_1, \ldots, F_p$ sont des sous-espaces vectoriels de E avec $p \geq 3$, le fait que la somme $F_1 + \cdots + F_p$ soit directe n'est pas équivalente à la condition $F_1 \cap \ldots \cap F_p = \{0_E\}$.
En effet, prenons les trois sous-espaces vectoriels de $E = \mathbb{R}^3$ définis par

$$F_1 = \{(\alpha, 0, 0) \ / \ \alpha \in \mathbb{R}\}, \ F_2 = \{(\beta, \beta, 0) \ / \ \beta \in \mathbb{R}\}, \ F_3 = \{(2a + b, a, b) \ / \ a, b \in \mathbb{R}\}.$$

Il est clair que $F_1 \cap F_2 \cap F_3 = \{0_{\mathbb{R}^3}\}$, car $F_1 \cap F_2 \cap F_3 \subseteq F_1 \cap F_2$ et on a déjà vu dans l'exemple 1.5.12 que $F_1 \cap F_2 = \{0_E\}$.
Cependant, la somme $F_1 + F_2 + F_3$ n'est pas directe, puisque $(1, 0, 0) + (1, 1, 0) + (-2, -1, 0) = (0, 0, 0)$ avec $(1, 0, 0) \in F_1, (1, 1, 0) \in F_2, (-2, -1, 0) \in F_3$.

En fait, nous disposons pour $p \geq 3$ du critère suivant :

Théorème 1.5.14. *Soient $F_1, \ldots, F_p$ des sous-espaces vectoriels de E. Alors*

$$\textit{La somme } F_1 + \cdots + F_p \textit{ est directe} \iff \forall i \in \{2, \ldots, p\}, \ F_i \cap (F_1 + \cdots + F_{i-1}) = \{0_E\}.$$

Démonstration.

- Supposons que la somme $F_1 + \cdots + F_p$ est directe. Soit $i \in \{2, \ldots, p\}$ et $x \in F_i \cap (F_1 + \cdots + F_{i-1})$. Alors $x \in F_i$ et il existe $(x_1, \ldots, x_{i-1}) \in F_1 \times \cdots \times F_{i-1}$ tel que $x = x_1 + \cdots + x_{i-1}$. En notant $x_i = -x \in F_i, \ x_{i+1} = 0_E \in F_{i+1}, \ldots, x_p = 0_E \in F_p$, nous avons évidemment $x_1 + \cdots + x_p = 0_E$. Il découle de la somme directe $F_1 \oplus \cdots \oplus F_p$ que tous les vecteurs $x_1, \ldots, x_p$ valent 0_E. En particulier $x = -x_i = 0_E$. D'où $F_i \cap (F_1 + \cdots + F_{i-1}) = \{0_E\}$.
- Réciproquement, supposons que pour tout $i \in \{2, \ldots, p\}$, $F_i \cap (F_1 + \cdots + F_{i-1}) = \{0_E\}$. Montrons que la somme $F_1 + \cdots + F_p$ est directe. Soient $x_1 \in F_1, \ldots, x_p \in F_p$ tels que $x_1 + \cdots + x_p = 0_E$. On a $x_p \in F_p$ et $x_p = -x_1 - \cdots - x_{p-1} \in F_1 + \cdots + F_{p-1}$. Donc $x_p \in F_p \cap (F_1 + \cdots + F_{p-1})$, d'où $x_p = 0_E$.
  On a donc $x_1 + \cdots + x_{p-1} = 0_E$. En reprenant le même raisonnement pour $x_{p-1} = -x_1 - \cdots - x_{p-2} \in F_{p-1} \cap (F_1 + \cdots + F_{p-2})$, on obtient $x_{p-1} = 0_E$. En réitérant, on trouve finalement $x_p = x_{p-1} = \cdots = x_2 = x_1 = 0_E$. ■

Remarque 1.5.15. Soient $F_1, \dots, F_p$ des sous-espaces vectoriels de E et supposons que leur somme $F_1 + \cdots + F_p$ est directe. Alors pour tous $i \neq j$ dans $\{1, \dots, p\}$, on a $F_i \cap F_j = \{0_E\}$. En effet, si $i > j$, alors $F_j \subseteq F_1 + \cdots + F_j + \cdots + F_{i-1}$, et donc $F_i \cap F_j \subseteq F_i \cap (F_1 + \cdots + F_{i-1}) = \{0_E\}$. Même raisonnement si $i < j$ en échangeant les rôles de i et j.
On en déduit aussi que $F_1 \cap \ldots \cap F_p = \{0_E\}$. La réciproque est fausse, comme on a déjà expliqué plus haut.

Définition 1.5.16. *Soient $F_1, \dots, F_p$ des sous-espaces vectoriels de E. On dit que E **est la somme directe** de $F_1, \dots, F_p$, et l'on écrit $E = F_1 \oplus \cdots \oplus F_p$, si :*

- *$E = F_1 + \cdots + F_p$.*
- *La somme $F_1 + \cdots + F_p$ est directe.*

Exemple 1.5.17. Reprenons l'exemple 1.5.9 et montrons que l'espace vectoriel $E = \mathbb{R}^3$ est la somme directe de F_1, F_2, F_3. Nous savons déjà que la somme $F_1+F_2+F_3$ est directe. Il reste à prouver que $E = F_1 + F_2 + F_3$, qui se ramène à l'inclusion $E \subseteq F_1 + F_2 + F_3$. Soit donc un vecteur quelconque $x = (a, b, c) \in \mathbb{R}^3$ et cherchons s'il existe des vecteurs $x_1 = (\alpha, 0, 0) \in F_1,\ x_2 = (\beta, \beta, 0) \in F_2$ et $x_3 = (\gamma, \gamma, \gamma) \in F_3$ tels que $x = x_1 + x_2 + x_3$. Ceci équivaut au système linéaire

$$\begin{cases} \alpha + \beta + \gamma = a \\ \qquad \beta + \gamma = b \\ \qquad\qquad \gamma = c \end{cases}$$

à trois inconnues α, β, γ qu'il faut trouver en fonction des paramètres donnés a, b, c. On trouve facilement $\gamma = c,\ \beta = b - c,\ \alpha = a - b$. Donc $x_1 = (a - b, 0, 0), x_2 = (b - c, b - c, 0)$ et $x_3 = (c, c, c)$ existent. D'où $E = F_1 + F_2 + F_3$, et par suite, $E = F_1 \oplus F_2 \oplus F_3$.

Exemple 1.5.18. Soient E un espace vectoriel de dimension n sur $\mathbb{K}$ et $B = (a_1, \dots, a_n)$ une base de E. Considérons les droites vectorielles $F_i = \mathsf{Vect}(v_i)$ pour $i = 1, \dots, n$. Alors $E = F_1 \oplus \cdots \oplus F_n$.
En effet, puisque B engendre E, alors tout vecteur $x \in E$ peut s'écrire sous forme $x = \alpha_1 a_1 + \cdots + \alpha_n a_n$ avec $\alpha_1, \dots, \alpha_n \in \mathbb{K}$. Donc $x = x_1 + \cdots + x_n$ avec $x_i = \alpha_i a_i \in F_i$ pour tout $i = 1, \dots, n$. D'où $E \subseteq F_1 + \cdots + F_n$, et par suite, $E = F_1 + \cdots + F_n$.
D'autre part, soient $x_1 \in F_1, \dots, x_n \in F_n$ tels que $x_1 + \cdots + x_n = 0_E$. Chaque $x_i \in F_i$ est de la forme $x_i = \lambda_i a_i$. Donc $\lambda_1 a_1 + \cdots + \lambda_n a_n = 0_E$. Or, B est libre, d'où $\lambda_1 = \ldots = \lambda_n = 0$. Ainsi, $x_1 = \ldots = x_n = 0_E$, ce qui implique que la somme $F_1 + \cdots + F_n$ est directe. En conclusion, $E = F_1 \oplus \cdots \oplus F_n$.

Proposition 1.5.19. *Soient $F_1, \dots, F_p$ des sous-espaces vectoriels de E. On a :*
$E = F_1 \oplus \cdots \oplus F_p$ si et seulement si tout vecteur $x \in E$ s'écrit de manière unique sous la forme $x = x_1 + \cdots + x_p$ avec $x_1 \in F_1, \dots, x_p \in F_p$.

Démonstration.

- Si $E = F_1 \oplus \cdots \oplus F_p$, alors $E = F_1 + \cdots + F_p$ et la somme $F_1 + \cdots + F_p$ est directe. Il suffit alors d'appliquer la proposition 1.5.10 pour déduire que tout vecteur $x \in E$ s'écrit de manière unique sous la forme $x = x_1 + \cdots + x_p$ avec $x_1 \in F_1, \dots, x_p \in F_p$.
- Réciproquement, supposons que tout vecteur $x \in E$ s'écrit de manière unique sous la forme $x = x_1 + \cdots + x_p$ avec $x_1 \in F_1, \dots, x_p \in F_p$. Donc, en particulier, tout vecteur $x \in E$ appartient à la somme $F_1 + \cdots + F_p$, d'où $E = F_1 + \cdots + F_p$. Il découle de la proposition 1.5.10 que tout vecteur $x \in F_1 + \cdots + F_p$ s'écrit de manière unique sous la forme $x = x_1 + \cdots + x_p$ avec $x_1 \in F_1, \dots, x_p \in F_p$. Donc la somme $F_1 + \cdots + F_p$ est directe, d'où $E = F_1 \oplus \cdots \oplus F_p$. ■

Il y a une autre caractérisation importante des sommes directes lorsqu'on travaille à l'aide des bases :

Proposition 1.5.20. *Soient E un espace vectoriel de dimension finie et $F_1, \dots, F_p$ des sous-espaces vectoriels de E, de bases respectives $B_1, \dots, B_p$. On a :*
$E = F_1 \oplus \cdots \oplus F_p$ si et seulement si $B_1, \dots, B_p$ sont disjoints deux à deux et $B_1 \cup \ldots \cup B_p$ est une base de E.

Démonstration.

- Si $E = F_1 \oplus \cdots \oplus F_p$, alors la remarque 1.5.15 entraîne que $F_i \cap F_j = \{0_E\}$ pour tous indices $i \neq j$. Par conséquent $B_i \cap B_j = \emptyset$, car le vecteur nul 0_E ne peut appartenir à une base. Donc $B_1, \dots, B_p$ sont disjoints deux à deux.
  Montrons que la famille $B = B_1 \cup \ldots \cup B_p$ est une base de E. Soit $x \in E$. Puisque $E = F_1 \oplus \cdots \oplus F_p$, nous savons par la proposition 1.5.19 que x se décompose de façon unique sous forme $x = x_1 + \cdots + x_p$ avec $x_i \in F_i$ pour tout $i = 1, \dots, p$. De plus, comme B_i est une base de F_i, alors chaque x_i s'écrit de manière unique comme combinaison linéaire des éléments de B_i. Par conséquent, x peut s'écrire de manière unique comme combinaison linéaire des éléments de B. Ceci montre, via le théorème 1.3.35, que B est une base de E.

– Réciproquement, supposons que $B_1, \dots, B_p$ sont disjoints deux à deux et que $B = B_1 \cup \dots \cup B_p$ est une base de E. Montrons que $E = F_1 \oplus \cdots \oplus F_p$. Puisque $B = B_1 \cup \dots \cup B_p$ est une base de E, alors tout vecteur $x \in E$ se décompose d'une seule manière comme combinaison linéaire des éléments de B. Donc tout vecteur $x \in E$ se décompose d'une seule manière sous forme $x = x_1 + \cdots + x_p$ avec $x_i \in F_i$ pour tout indice i. Il résulte de la proposition 1.5.19 que $E = F_1 \oplus \cdots \oplus F_p$. ■

Corollaire 1.5.21. *Soient E un espace vectoriel de dimension finie et $F_1, \dots, F_p$ des sous-espaces vectoriels de E.*
Si $E = F_1 \oplus \cdots \oplus F_p$, alors $\dim(E) = \dim(F_1) + \cdots + \dim(F_p)$.

Démonstration. Soient $B_1, \dots, B_p$ des bases respectives de $F_1, \dots, F_p$. D'après la proposition précédente, $B_1 \cup \dots \cup B_p$ est une base de E. Donc $\dim(E) = \mathsf{card}(B_1 \cup \dots B_p)$. Mais $B_1, \dots, B_p$ sont disjointes deux à deux, d'où $\mathsf{card}(B_1 \cup \dots B_p) = \mathsf{card}(B_1) + \cdots + \mathsf{card}(B_p)$. Par suite, $\dim(E) = \dim(F_1) + \cdots + \dim(F_p)$. ■

Exemple 1.5.22. Dans l'exemple 1.5.17, on a montré que $E = F_1 \oplus F_2 \oplus F_3$. Nous allons le prouver d'une autre manière en nous servant de la proposition précédente. Cherchons donc des bases de F_1, F_2, F_3. Il est très facile de voir que

$$F_1 = \{(\alpha, 0, 0) \ / \ \alpha \in \mathbb{R}\} = \mathsf{Vect}(u) \text{ avec } u = (1, 0, 0),$$

$$F_2 = \{(\beta, \beta, 0) \ / \ \beta \in \mathbb{R}\} = \mathsf{Vect}(v), \text{ avec } v = (1, 1, 0),$$

$$F_3 = \{(\gamma, \gamma, \gamma) \ / \ \gamma \in \mathbb{R}\} = \mathsf{Vect}(w), \text{ avec } w = (1, 1, 1).$$

Comme les vecteurs u, v, w sont distincts, il suffit donc de montrer que $B = (u, v, w)$ est une base de $E = \mathbb{R}^3$. Pour cela, puisque $\mathsf{card}(B) = 3 = \dim(\mathbb{R}^3)$, on a besoin de vérifier seulement que B est libre. Soient donc $\alpha, \beta, \gamma \in \mathbb{R}$ tels que $\alpha u + \beta v + \gamma w = 0_{\mathbb{R}^3}$. En développant les calculs, on obtient le système linéaire

$$\begin{cases} \alpha + \beta + \gamma = 0 \\ \beta + \gamma = 0 \\ \gamma = 0 \end{cases}$$

qui est le même système obtenu dans l'exemple 1.5.9. Donc $\alpha = \beta = \gamma = 0$. Ainsi, B est une base de E, et par suite, on retrouve que $E = F_1 \oplus F_2 \oplus F_3$.

Le cas $p = 2$ nous conduit à une définition importante :

Définition 1.5.23. *Soient F_1 et F_2 deux sous-espaces vectoriels de E. On dit que F_1 et F_2 sont* ***supplémentaires*** *dans E si $E = F_1 \oplus F_2$. On dit aussi que F_2 est un supplémentaire de F_1 dans E (et que F_1 est un supplémentaire de F_2 dans E).*

En vertu des résultats précédents, on peut énoncer immédiatement :

Proposition 1.5.24. *Soient F_1 et F_2 deux sous-espaces vectoriels de E. Les assertions suivantes sont équivalentes :*

1. *F_1 et F_2 sont supplémentaires dans E.*

2. *$E = F_1 + F_2$ et $F_1 \cap F_2 = \{0_E\}$.*

3. *Tout vecteur $x \in E$ s'écrit de manière unique sous la forme $x = x_1 + x_2$ avec $x_1 \in F_1, x_2 \in F_2$.*

4. *Si B_1 est une base de F_1 et B_2 est une base de F_2, alors B_1 et B_2 sont disjoints et $B_1 \cup B_2$ est une base de E.*

Théorème 1.5.25. *Soit E un espace vectoriel de dimension finie n. Tout sous-espace vectoriel F de E admet au moins un supplémentaire.*

Démonstration. Soit $B_F = (a_1, \dots, a_k)$ une base de F. Cette famille B_F est libre dans E, donc son cardinal $k \leq n = \dim(E)$. On complète alors B_F en une base $B = (a_1, \dots, a_k, a_{k+1}, \dots, a_n)$ de E. Considérons le sous-espace $G = \mathsf{Vect}(a_{k+1}, \dots, a_n)$. Il est clair que la famille $B_G = (a_{k+1}, \dots, a_n)$ est une base de G, car elle est libre et génératrice de G. De plus, B_F et B_G sont disjoints et $B_F \cup B_G = B$ est une base de E. D'après la proposition 1.5.20, on conclut que $E = F_1 \oplus F_2$.

Corollaire 1.5.26. *Soient F_1 et F_2 deux sous-espaces vectoriels de E. Les assertions suivantes sont équivalentes :*

1. *F_1 et F_2 sont supplémentaires dans E.*

2. *$F_1 \cap F_2 = \{0_E\}$ et $\dim(E) = \dim(F_1) + \dim(F_2)$.*

3. *$F_1 + F_2 = E$ et $\dim(E) = \dim(F_1) + \dim(F_2)$.*

⚠ Pour $p = 2$, il ne faut pas confondre supplémentaire et complémentaire ! En effet, si F est un sous-espace vectoriel de E, alors le complémentaire de F dans E est l'ensemble $E \backslash F$ des éléments de E qui n'appartiennent pas à F. Et ce complémentaire n'est pas un sous-espace de E car il ne contient pas 0_E. D'autre part, un supplémentaire de F dans E n'est pas unique.

Terminons ce chapitre par un exemple consacré aux polynômes :

Exemple 1.5.27. Considérons l'espace vectoriel $\mathbb{K}[X]$ des polynômes, et $\mathbb{K}_n[X]$ son sous-espace vectoriel formé des polynômes de degré $\leq n$.

- Pour tout polynôme $P \in \mathbb{K}[X]$, on note $(P) = \{AP \;/\; A \in \mathbb{K}[X]\}$ l'ensemble des multiples de P. C'est un sous-espace vectoriel de $\mathbb{K}[X]$. En effet, puisque $AP + BP = (A + B)P \in (P)$ et $\lambda(AP) = (\lambda A)P \in (P)$ pour tous éléments AP et BP dans (P) et pour tout scalaire $\lambda \in \mathbb{K}$, alors (P) est stable pour l'addition et la multiplication externe.
- Supposons maintenant que P soit de degré $n + 1$. Alors les sous-espaces vectoriels (P) et $\mathbb{K}_n[X]$ sont supplémentaires dans $\mathbb{K}[X]$: $\mathbb{K}[X] = (P) \oplus \mathbb{K}_n[X]$. En effet,

  * Si $B \in (P) \cap \mathbb{K}_n[X]$, alors il existe $A \in \mathbb{K}[X]$ tel que $B = AP$ et $\deg(B) \leq n$. Si $A \neq 0$, alors $\deg(B) = \deg(A) + \deg(P) \geq n + 1$, ce qui contredit l'hypothèse $\deg(P) = n + 1$. Par conséquent, $A = 0$, d'où $B = 0$, et par suite $(P) \cap \mathbb{K}_n[X] = \{0\}$.
  * Montrons maintenant que $\mathbb{K}[X] = (P) + \mathbb{K}_n[X]$. Soit $B \in \mathbb{K}[X]$ et effectuons la division euclidienne de B par P : il existe un couple (unique) $(A, R) \in \mathbb{K}[X] \times \mathbb{K}[X]$ tel que $B = AP + R$ avec $\deg(R) < n + 1$. Comme $AP \in (P)$ et $R \in \mathbb{K}_n[X]$, il en résulte que $\mathbb{K}[X] = (P) + \mathbb{K}_n[X]$. Ainsi, $\mathbb{K}[X] = (P) \oplus \mathbb{K}_n[X]$.

2 Applications linéaires

Sommaire

"Ce n'est pas parce que c'est difficile qu'on ne travaille pas, mais c'est parce qu'on ne travaille pas que ça devient difficile."

.

En premier semestre, vous avez déjà pris connaissance des notions d'homomorphismes de groupes et d'anneaux. Ce sont des applications qui respectent les structures de ces objets. Par analogie, il existe une notion similaire d'homomorphismes pour les espaces vectoriels, appelés souvent applications linéaires, et qui respectent la structure algébrique des espaces vectoriels, c'est à dire, non seulement la structure de groupe additif sous-jacent, mais aussi la loi externe. Vu que les espaces vectoriels ont pour motivation géométrique, il ne faudrait pas s'étonner de voir que plusieurs applications linéaires ont pour origine les transformations en géométrie (rotations, symétries, homothéties, etc.).

2.1 Généralités

Soient E et F deux espaces vectoriels sur le même corps $\mathbb{K}$. Chacun de ces deux espaces vectoriels est équipé d'une loi interne et d'une loi externe.

Définition 2.1.1. *Soient E et F deux espaces vectoriels sur le même corps $\mathbb{K}$ et $f : E \longrightarrow F$ une application. On dit que f est **linéaire** si pour tous vecteurs $u, v \in E$ et pour tout scalaire $\alpha \in \mathbb{K}$, on a :*

1. $f(u+v) = f(u) + f(v)$.

2. $f(\alpha \cdot u) = \alpha \cdot f(u)$.

<u>**Vocabulaire**</u> : Soit $f : E$ une application linéaire.

1. Si $E = F$, on dit que f est un **endomorphisme** de E.

2. Si $f : E \longrightarrow F$ est une application linéaire bijective, on dit que f est un **isomorphisme** de E dans F.

3. Un endomorphisme $f : E \longrightarrow E$ qui est bijectif est appelé un **automorphisme** de E. C'est aussi un isomorphisme de E dans E.

4. Si $F = \mathbb{K}$, on dit que $f : E \longrightarrow \mathbb{K}$ est une **forme linéaire** sur E.

Remarque 2.1.2.1. La nécessité que E et F soient des espaces vectoriels sur le même corps $\mathbb{K}$ apparaît clairement dans la définition d'une application linéaire $f : E \longrightarrow F$. Pour mettre en évidence le corps $\mathbb{K}$, on dit parfois que f est une application **$\mathbb{K}$-linéaire**.

2. Comme conséquence immédiate de la définition, nous voyons que $f(0_E) = 0_F$. En effet, en prenant $u \in E$, on a $f(0_E) = f(0_K \cdot u) = 0_K \cdot f(u) = 0_F$. Ceci peut s'expliquer aussi en se rappelant que l'image de l'élément neutre 0_E du groupe de départ $(E, +)$ par un homomorphisme de groupes $f : (E, +) \longrightarrow (F, +)$ est l'élément neutre 0_F du groupe d'arrivée $(F, +)$.

On peut utiliser la proposition suivante pour montrer qu'une application est linéaire :

Proposition 2.1.3. *Soit $f : E \longrightarrow F$ une application. Alors :*

f est linéaire $\iff$ $f(\alpha u+\beta v) = \alpha f(u)+\beta f(v)$ pour tous $u, v \in E$ et $\alpha, \beta \in K$.

Démonstration.

– Si f est linéaire, alors les deux conditions de la définition précédente impliquent $f(\alpha u + \beta v) = f(\alpha u) + f(\beta v) = \alpha f(u) + \beta f(v)$.

– Réciproquement, supposons que l'application f vérifie la relation $f(\alpha u+\beta v) = \alpha f(u)+\beta f(v)$ pour tous $u, v \in E$ et $\alpha, \beta \in K$, et montrons que f est linéaire. D'abord, en prenant $\alpha = \beta = 1$, on obtient $f(u + v) = f(u) + f(v)$. Ensuite, pour $\alpha \in K$ et $\beta = 0_K$, il vient $f(\alpha u) = f(\alpha u + 0_K v) = \alpha f(u) + 0_K f(v) = \alpha f(u)$. ■

Remarque 2.1.4. Il est également aisé de vérifier qu'une application $f : E \longrightarrow F$ est linéaire si et seulement si $f(\alpha u + v) = \alpha f(u) + f(v)$ pour tous $u, v \in E$ et $\alpha \in K$.

La proposition ci-dessous s'établit facilement par récurrence.

Proposition 2.1.5. Soit $f : E \longrightarrow E$ une application linéaire. Pour tout entier $n \geq 1$, pour tous vecteurs $v_1, \cdots, v_n$ dans E et pour tous scalaires $\alpha_1, \cdots, \alpha_n$ dans $\mathbb{K}$, on a

$$\boxed{f(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_1 f(v_1) + \cdots + \alpha_n f(v_n)}$$

Démonstration. La propriété est déjà vérifiée pour $n = 1$ et $n = 2$. Si elle est vraie pour n, alors, pour tous vecteurs $v_1, \cdots, v_n, v_{n+1}$ dans E et pour tous scalaires $\alpha_1, \cdots, \alpha_n, \alpha_{n+1}$ dans $\mathbb{K}$,

$$\begin{aligned}
f(\alpha_1 v_1 + \cdots + \alpha_n v_n + \alpha_{n+1} v_{n+1}) &= f\Big((\alpha_1 v_1 + \cdots + \alpha_n v_n) + \alpha_{n+1} v_{n+1}\Big) \\
&= f(\alpha_1 v_1 + \cdots + \alpha_n v_n) + \alpha_{n+1} f(v_{n+1}) \\
&= \Big(\alpha_1 f(v_1) + \cdots + \alpha_n f(v_n)\Big) + \alpha_{n+1} f(v_{n+1}) \\
&= \alpha_1 f(v_1) + \cdots + \alpha_n f(v_n) + \alpha_{n+1} f(v_{n+1})
\end{aligned}$$

Avant de continuer l'exploration des applications linéaires, il est nécessaire de faire une pause et de prendre connaissance de quelques exemples importants. Il nous semble en effet fondamental d'assimiler ces notions pour pouvoir bien suivre le reste du chapitre.

Exemple 2.1.6. Soit $a \in \mathbb{K}$. L'application $f : \mathbb{K} \longrightarrow \mathbb{K}, x \mapsto ax$ est une application linéaire car pour tous $x, x' \in \mathbb{K}$ et pour tous $\alpha, \beta \in \mathbb{K}$,

$$f(\alpha x + \beta x') = a(\alpha x + \beta x') = \alpha(ax) + \beta(ax') = \alpha f(x) + \beta f(x').$$

Exemple 2.1.7. Soit E un espace vectoriel sur K.

- Soit $\lambda \in K$ un scalaire fixé et $h_\lambda : E \longrightarrow E$ l'application définie par $h_\lambda(v) = \lambda v$ pour tout $v \in E$. Alors f est une application linéaire, appelée **homothétie de rapport** λ.
- En particulier, pour $\lambda = 0$, l'application nulle $\theta = h_0 : E \longrightarrow E,\ v \mapsto 0_E$ sont bien des applications linéaires.
- De même, pour $\lambda = 1$, on obtient l'application identité $\mathtt{id}_E = h_1 : E \longrightarrow E,\ v \mapsto v$, qui est aussi linéaire.

Exemple 2.1.8. Soit $f : E \longrightarrow F$ une application linéaire et S un sous-espace vectoriel de E. On rappelle que la restriction $f_{|S} : S \longrightarrow F$ de f sur S est définie par $f_{|S}(x) = f(x)$ pour tout vecteur $x \in S$. Il est alors clair que $f_{|S}$ est aussi une application linéaire.

Exemple 2.1.9. Considérons les espaces vectoriels $E = \mathbb{R}^3$ et $F = \mathbb{R}^2$ sur le corps $K = \mathbb{R}$. Soit l'application $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ définie par

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

Alors f est linéaire. En effet, pour tous vecteurs $u = (x, y, z)$ et $v = (x', y', z')$ dans $\mathbb{R}^3$ et pour tout scalaire α dans $\mathbb{R}$,

$$\begin{array}{rcl}
f(u+v) & = & f(x+x', y+y', z+z') \\
& = & \Big((x+x') + 2(y+y') + (z+z'), (y+y') - 3(z+z')\Big) \\
& = & \Big((x+2y+z) + (x'+2y'+z'), (y-3z) + (y'-3z')\Big) \\
& = & (x+2y+z, y-3z) + (x'+2y'+z', y'-3z') \\
& = & f(u) + f(v).
\end{array}$$

Et

$$\begin{array}{rcl}
f(\alpha u) & = & f(\alpha x, \alpha y, \alpha z) \\
& = & \Big(\alpha x + 2(\alpha y) + \alpha z, \alpha y - 3(\alpha z)\Big) \\
& = & \Big(\alpha(x+2y+z), \alpha(y-3z)\Big) \\
& = & \alpha(x+2y+z, y-3z) = \alpha f(u).
\end{array}$$

Exemple 2.1.10. L'application $\psi : \mathbb{R}^2 \longrightarrow \mathbb{C}$ définie par $\psi(x, y) = x + iy$ est une application linéaire du $\mathbb{R}$-espace vectoriel $\mathbb{R}^2$ dans le $\mathbb{R}$-espace vectoriel $\mathbb{C}$. En effet, pour tous vecteurs $u = (x, y)$ et $v = (x', y')$ dans $\mathbb{R}^2$,

$$\begin{array}{rcl}
\psi(u+v) & = & \psi(x+x', y+y') \\
& = & (x+x') + i(y+y') = (x+iy) + (x'+iy') \\
& = & \psi(x, y) + \psi(x', y') = \psi(u) + \psi(v).
\end{array}$$

D'autre part, pour tout vecteur $u = (x, y) \in \mathbb{R}^2$ et pour tout scalaire $\alpha \in \mathbb{R}$,

$$\begin{array}{rcl}
\psi(\alpha u) & = & \psi(\alpha x, \alpha y) \\
& = & \alpha x + i(\alpha y) = \alpha(x+iy) \\
& = & \alpha\psi(x, y) = \alpha\psi(u)).
\end{array}$$

On rappelle que $\mathbb{C}$, qui est un $\mathbb{C}$-espace vectoriel, possède aussi une structure de $\mathbb{R}$-espace vectoriel. Dans l'exemple ci-dessus, nous avons privilégié ici la structure de $\mathbb{R}$-espace vectoriel car, d'après la définition d'une application linéaire, les espaces vectoriels de départ et d'arrivée E et F doivent être définis sur le même corps $\mathbb{K}$.

L'analyse met à notre disposition plusieurs exemples concrets d'applications linéaires.

Exemple 2.1.11. Soit $\mathbb{K}[X]$ l'espace vectoriel sur $\mathbb{K}$ des polynômes.

1. **Dérivation**. L'application dérivée (ou de différentiation) $D : K[X] \longrightarrow K[X]$ définie par $D(P) = P' = \sum_{i=1}^{n} ia_iX^{i-1}$ pour tout polynôme $P = \sum_{i=0}^{n} a_iX^i$ est un endomorphisme de $\mathbb{K}[X]$. En effet, pour tous polynômes $P, Q \in \mathbb{K}[X]$ et pour tous scalaires $\alpha, \beta \in \mathbb{K}$,
$$(\alpha \cdot P + \beta \cdot Q)' = \alpha \cdot P' + \beta \cdot Q'.$$
2. **Intégration**. L'application d'intégration $I : K[X] \longrightarrow K[X]$ déterminée par
$$I(P) = \int_0^X P(t)dt = \sum_{i=0}^{n} \frac{a_i}{i+1}X^{i+1}$$
pour tout polynôme $P = \sum_{i=0}^{n} a_iX^i$ est aussi un endomorphisme de $\mathbb{K}[X]$.

Voici à présent quelques exemples d'applications non linéaires :

Exemple 2.1.12. Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ l'application définie par $f(x, y) = xy$. Alors f n'est pas linéaire, car si $u = (x, y) \in \mathbb{R}^2$ et $\alpha \in \mathbb{R}$:
$$f(\alpha u) = f(\alpha(x, y)) = f(\alpha x, \alpha y) = (\alpha x)(\alpha y) = \alpha^2(xy) = \alpha^2 f(x, y) = \alpha^2 f(u).$$
Ainsi, $f(\alpha u) \neq \alpha f(u)$ en général. Il suffit de choisir $\alpha \neq 1$ et un vecteur $u \in \mathbb{R}^2$ tel que $f(u) \neq 0_{\mathbb{R}^2}$, par exemple $u = (2, 1)$.

Exemple 2.1.13. Soit E un espace vectoriel et v_0 un vecteur de E non nul. On appelle **translation de vecteur** v_0 l'application $f_{v_0} : E \longrightarrow E$ définie par $f_{v_0}(v) = v + v_0 \quad \forall v \in E$. Il n'est pas difficile de se convaincre que cette application n'est pas linéaire. En effet, pour tous vecteurs $u, v \in E$, $f(u+v) = (u + v) + v_0$ et $f(u) + f(v) = (u + v_0) + (v + v_0) = (u + v) + 2v_0$. Ainsi, $f_{v_0}(u + v) \neq f_{v_0}(u) + f_{v_0}(v)$.

2.2 Image directe et image réciproque d'un sous-espace

Soient E et F deux ensembles non vides et $f : E \longrightarrow F$ une application quelconque. Rappelons du cours de SMIA1 que l'**image directe** $f(S)$ d'une partie $S \subseteq E$ est la partie de F définie par :

$$\boxed{f(S) = \{f(x) \ / \ x \in S\}}$$

Donc pour tout élément y de F, on a :

$$\boxed{y \in f(S) \iff \exists\, x \in S : y = f(x)}$$

D'autre part, si $T \subseteq F$ est une partie de F, l'**image réciproque** $f^{-1}(T)$ de T est donnée par :

$$\boxed{f^{-1}(T) = \{x \in E \ / \ f(x) \in T\}}$$

Ainsi, pour tout élément x de E, on a :

$$\boxed{x \in f^{-1}(T) \iff f(x) \in T}$$

En particulier, si $T = \{b\}$ contient un seul élément $b \in F$, alors

$$\boxed{f^{-1}(\{b\}) = \{x \in E \ / \ f(x) = b\}}$$

qui se note aussi $f^{-1}(b)$, et qui est l'ensemble de tous les antécédents de b.

⚠ Lorsqu'on écrit $f^{-1}(T)$, il ne faut pas penser que la notation f^{-1} signifie ici l'application réciproque de f. D'ailleurs, f n'est pas forcément bijective. Cependant, si f est maintenant bijective, alors on peut faire la confusion sans problème, puisque dans cas, l'élément b de F admet un seul antécédent x dans E.

Dans le contexte des espaces vectoriels, que peut-on dire de plus à propos des parties $f(S)$ et $f^{-1}(T)$ lorsque f est linéaire, S est un sous-espace de E et T est un sous-espace de F ? La réponse est fournie dans ce qui va suivre :

Théorème 2.2.1. *Soit $f : E \longrightarrow F$ une application linéaire.*

1. *Si S est un sous-espace vectoriel de E, son image directe $f(S)$ est un sous-espace vectoriel de F.*

2. *En particulier, $f(E)$ est un sous-espace vectoriel de F.*

DÉMONSTRATION.

1. – D'abord, puisque S est un sous-espace vectoriel de E, alors $0_E \in S$. Donc $f(0_E) \in f(S)$. Or, $f(0_E) = 0_F$, d'où $0_F \in f(S)$.

– Soient α, β dans $\mathbb{K}$ et y, y' dans $f(S)$. Il existe $x, x' \in S$ tels que $f(x) = y$ et $f(x') = y'$. Il vient $\alpha y + \beta y' = \alpha f(x) + \beta f(x') = f(\alpha x + \beta x') \in f(S)$ car $\alpha x + \beta x' \in S$.

2. Remarquer que $S = E$ est lui même un sous-espace vectoriel de E, et appliquer l'assertion 1. ■

Définition 2.2.2. *Le sous-espace vectoriel $f(E)$ de F s'appelle* ***image*** *de f et se note* Im(f) *:*

$$\boxed{\mathsf{Im}(f) = \{f(x) \ / \ x \in E\}}$$

En d'autres termes, pour tout $y \in F$:

$$\boxed{y \in \mathsf{Im}(f) \iff \textit{il existe } x \in E \textit{ tel que } y = f(x)}$$

Théorème 2.2.3. *Soit $f : E \to F$ une application linéaire. Alors :*

1. *Si T est un sous-espace vectoriel de F, son image réciproque $f^{-1}(T)$ est un sous-espace vectoriel de E.*

2. *En particulier, $f^{-1}(\{0_F\})$ est un sous-espace vectoriel de E.*

DÉMONSTRATION.

1. – Puisque $f(0_E) = 0_F \in T$, alors $0_E \in f^{-1}(T)$.

– Si $\alpha, \beta \in K$ et $x, x' \in f^{-1}(T)$, alors $f(\alpha x + \beta x') = \alpha f(x) + \beta f(x') \in T$ parce que $f(x)$ et $f(x')$ appartiennent au sous-espace T de F.

2. Appliquer la première assertion au sous-espace vectoriel $S = \{0_E\}$ de E. ■

Définition 2.2.4. *Le sous-espace vectoriel $f^{-1}(0_F)$ de F s'appelle* ***noyau*** *de f et s'écrit* $\ker(f)$ *:*

$$\boxed{\ker(f) = \{x \in E \ / \ f(x) = 0_F\}}$$

Donc pour tout $x \in E$,

$$\boxed{x \in \ker(f) \iff f(x) = 0_F}$$

Exemple 2.2.5. Soit λ dans K et h_λ l'homothétie de E de rapport λ définie par $h_\lambda(x) = \lambda x$. Le noyau $\ker(h_\lambda)$ est le sous-espace vectoriel de E formé par les vecteurs $x \in E$ vérifiant $\lambda x = 0_E$. Si $\lambda \neq 0$, alors nécessairement $x = 0_E$, donc $\ker(h_\lambda) = \{0_E\}$. Si $\lambda = 0$, alors h_0 est l'application nulle $x \mapsto 0_E$ qui a évidemment pour noyau $\ker(\theta) = \ker(h_0) = E$.
D'autre part, $\mathsf{Im}(h_\lambda)$ est le sous-espace vectoriel de E constitué des vecteurs de la forme $h_\lambda(x) = \lambda x$, où x parcourt E. Si $\lambda = 0$, alors $\mathsf{Im}(h_0) = \{0_E\}$. Maintenant, si $\lambda \neq 0$, alors $\mathsf{Im}(h_\lambda) = E$, car tout vecteur $y \in E$ peut s'écrire $y = \lambda\left(\frac{1}{\lambda}y\right) = h_\lambda(x)$ avec $x = \frac{1}{\lambda}y \in E$.

Le noyau $\ker(f)$ et l'image $\mathsf{Im}(f)$ sont des sous-espaces remarquables qui nous donnent certaines informations sur l'application linéaire f. En particulier, il est intéressant de savoir si une application linéaire f est surjective ou injective. Les résultats qui suivent caractérisent ces propriétés.

Proposition 2.2.6. *Soit $f : E \to F$ une application linéaire. Alors :*
(i) f est surjective $\iff$ $\mathsf{Im}(f) = F$.
(ii) f est injective $\iff$ $\ker(f) = \{0_E\}$.

Démonstration. L'assertion (i) résulte immédiatement de la définition d'une application surjective.
(ii) Nous savons que $\ker(f)$ est un sous-espace vectoriel de E, et donc il contient toujours le vecteur nul 0_E de E. Si f est injective, alors pour tout $x \in \ker(f)$, on a $f(x) = 0_F = f(0_E)$, d'où $x = 0_E$, et par suite $\ker(f) = \{0_E\}$.
Réciproquement, supposons que $\ker(f) = \{0_E\}$ et montrons que f est injective. Si $x, x' \in E$ vérifient $f(x) = f(x')$, alors $f(x - x') = f(x) - f(x') = 0_F$, ce qui donne $x - x' \in \ker(f)$. Or, par hypothèse, $\ker(f)$ ne contient que le vecteur nul 0_E de E. D'où $x - x' = 0_E$, ce qui donne $x = x'$. ∎

Exemple 2.2.7. Soit h_λ l'homothétie de E de rapport λ. D'après l'exemple précédent, h_0 n'est ni injective ni surjective, ce qui est normal puisque h_0 est l'application nulle θ. Si $\lambda \neq 0$, alors h_λ est à la fois injective et surjective, donc c'est un automorphisme de E.

Exemple 2.2.8. Soit $\psi : \mathbb{R}^2 \longrightarrow \mathbb{C}$ l'application linéaire définie par $\psi(x,y) = x+iy$ que nous avons considérée dans l'exemple 2.1.9. Alors ψ est injective.
En effet, pour tout vecteur $u=(x,y)\in\mathbb{R}^2$, on a
$\psi(u)=0_{\mathbb{R}^2} \iff \psi(x,y)=(0,0) \iff x+iy=0 \iff x=y=0 \iff u=(0,0)=0_{\mathbb{R}^2}$.
Donc $\ker(\psi)=\{0_{\mathbb{R}}\}$, ce qui prouve bien que ψ est injective.
Remarquer qu'il est aussi très facile de vérifier directement que ψ est injective en utilisant la définition :
$\psi(x,y)=\psi(x',y') \iff x+iy=x'+iy' \iff x=x'$ et $y=y' \iff (x,y)=(x',y')$.
L'application ψ est également surjective puisqu'à tout nombre complexe $z\in\mathbb{C}$ on peut associer le couple $(x,y)=(\mathtt{Re}(z),\mathtt{Im}(z))\in\mathbb{R}^2$ tel que $z=x+iy=\psi(x,y)$. L'application ψ est donc bijective. C'est un $\mathbb{R}$-isomorphisme de $\mathbb{R}^2$ dans $\mathbb{C}$.

2.3 Détermination d'une application linéaire dans une base

Nous avons déjà vu qu'une base d'un espace vectoriel E determine la structure complète de E. En suivant ce contexte, nous allons voir que les images des vecteurs d'une base par une application linéaire determine entièrement l'application linéaire en question.
Soit $f : E \longrightarrow F$ une application linéaire et soit $B=(a_1,\dots,a_n)$ une base de E. Si nous connaissons seulement les images $f(a_1),\dots,f(a_n)$ des éléments $a_1,\dots,a_n$ de la base B, alors nous pourrons calculer l'image d'un vecteur quelconque $v\in E$. En effet, il suffit de décomposer v comme combinaison linéaire $v=\alpha_1a_1+\cdots+\alpha_na_n$ des éléments de B. Alors on aura $f(v)=f(\alpha_1a_1+\cdots+\alpha_na_n)=\alpha_1f(a_1)+\cdots\alpha_nf(a_n)$. En d'autres termes, l'application linéaire f est parfaitement déterminée par la connaissance des images $f(a_1),\dots,f(a_n)$ des éléments d'une base B donnée de E.
Cette observation met à notre disposition un outil important pour fabriquer des applications linéaires :

Théorème 2.3.1. *Soient E et F deux espaces vectoriels sur le corps K et $B=(a_1,\dots,a_n)$ une base de E. Alors pour toute famille $(b_1,\dots,b_n)$ de n vecteurs de F, il existe une unique application linéaire $f : E\longrightarrow F$ telle que $f(a_i)=b_i$ pour tout $i\in\{1,\dots,n\}$. Plus précisément, f est donnée par $f(v)=\alpha_1b_1+\cdots\alpha_nb_n$ pour tout vecteur $v=\alpha_1a_1+\cdots+\alpha_na_n\in E$.*

Démonstration. *Unicité de f* : Supposons qu'il existe une application linéaire $f : E \longrightarrow F$ qui envoie chaque a_i vers b_i, et cherchons à trouver l'expression de $f(v)$ pour tout vecteur $v \in E$. Puisque B est une base de E, le vecteur v se décompose de manière unique comme combinaison linéaire $v = \sum_{i=1}^{n} \alpha_i a_i$ des éléments de B. Alors on aura

$$f(v) = f\left(\sum_{i=1}^{n} \alpha_i a_i\right) = \sum_{i=1}^{n} \alpha_i f(a_i) = \sum_{i=1}^{n} \alpha_i b_i.$$

Nous avons donc trouvé comment est définie l'application f cherchée.
Existence de f : Il reste à montrer que l'application f ainsi définie est linéaire. Elle est d'abord bien définie grâce à l'unicité de la décomposition de v. De plus, il est clair que $f(a_i) = b_i$ pour tout $i = 1, \ldots, n$. Montrons qu'elle est linéaire. Soient $x = x_1 a_1 + \cdots + x_n a_n$ et $x' = x'_1 a_1 + \cdots + x'_n a_n$ deux vecteurs de E et α, β deux scalaires de $\mathbb{K}$. On a alors, par définition de f, $f(x) = x_1 b_1 + \cdots + x_n b_n$ et $f(x') = x'_1 b_1 + \cdots + x'_n b_n$. D'autre part, $\alpha x + \beta x' = (\alpha x_1 + \beta x'_1) a_1 + \cdots + (\alpha x_n + \beta x'_n) a_n$, d'où $f(\alpha x + \beta x') = (\alpha x_1 + \beta x'_1) b_1 + \cdots + (\alpha x_n + \beta x'_n) b_n = \alpha(x_1 b_1 + \cdots + x_n b_n) + \beta(x'_1 b_1 + \cdots + x'_n b_n) = \alpha f(x) + \beta f(x')$. ■

Le résultat précédent explique la différence entre une application quelconque d'un ensemble vers un ensemble et une application linéaire. Pour définir une application linéaire, il suffit de la définir sur une base l'espace de départ. Intuitivement, cela veut dire qu'étant donnés deux espaces vectoriels E et F sur K, et une base $B = (a_1, \ldots, a_n)$ de E, on peut construire des applications linéaires de E vers F arbitraires. Il suffit de choisir des vecteurs $b_1, \ldots, b_n$ dans F qui seront les images respectives de $a_1, \ldots, b_n$. Ensuite, on calcule l'image $f(v)$ d'un vecteur quelconque $v \in E$ en le décomposant dans la base B. Les considérations précédentes impliquent immédiatement la remarque qui suit :
Une conséquence du théorème précédent est qu'une application linéaire est entièrement déterminée par les images des vecteurs d'une base. Pour s'en convaincre, énonçons l'observation suivante.

Lemme 2.3.2. *Soient $f : E \longrightarrow F$ et $g : E \longrightarrow F$ deux applications linéaires et $B = (a_1, \ldots, a_n)$ une base de E. Alors*

$$f = g \iff f(a_i) = g(a_i) \ \ \forall i \in \{1, \ldots, n\}.$$

Ceci exprime que, pour montrer que deux applications linéaires sont égales, il n'est pas nécessaire de le prouver pour tout $x \in E$, mais seulement pour les vecteurs $a_1, \dots, a_n$ d'une base de E.

En particulier, $f = 0 \iff f(a_i) = 0_F$ pour tout $i = 1, \dots, n$.

Dans le théorème 2.3.1, les vecteurs $b_1, \dots, b_n$ ne forment pas nécessairement une base de F. En fait, l'espace vectoriel F n'est pas nécessairement de même dimension n que l'espace E. Ce genre de questions sera traité plus loin.

Exemple 2.3.3. Déterminer l'unique application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^2$ telle que

$$f(1,0,0) = (1,1),\ f(0,1,0) = (0,1),\ f(0,0,1) = (1,0).$$

Posons $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$ et $b_1 = (1,1)$, $b_2 = (0,1)$, $b_3 = (1,0)$. Comme $B = (e_1, e_2, e_3)$ est une base de $\mathbb{R}^3$, le théorème 2.3.1 assure l'existence d'une application linéaire unique $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ telle que $f(e_i) = b_i$ $(1 \leq i \leq 3)$. Soit $u = (x, y, z)$ un vecteur quelconque de $\mathbb{R}^3$. Pour calculer son image $f(u)$, il suffit de le décomposer dans la base B, soit $u = xe_1 + ye_2 + ze_3$. Donc $f(u) = xf(e_1) + yf(e_2) + zf(e_3) = xb_1 + yb_2 + zb_3 = x(1,1) + y(0,1) + z(1,0) = (x+z, x+y)$.

Exemple 2.3.4. Soit $n \in \mathbb{N}$. Montrer qu'il existe une unique application linéaire $f : \mathbb{K}_n[X] \longrightarrow \mathbb{K}_{n+1}[X]$ telle que

$$f(X^i) = (X-1)^{i+1} \text{ pour tout } i \in \{0, 1, \dots, n\}.$$

Ceci résulte simplement du fait que la famille $(X^i)_{0 \leq i \leq n}$ est une base de $\mathbb{K}_n[X]$. Pour tout polynôme $P = \sum_{i=0}^{n} a_i X^i$, $\quad f(P) = \sum_{i=0}^{n} a_i f(X^i) = \sum_{i=0}^{n} a_i (X-1)^{i+1}$.

La proposition suivante caractérise la forme générale d'une application linéaire entre deux espaces de dimensions finies.

Proposition 2.3.5. *Soient E et F deux espaces vectoriels sur le corps K, $B = (e_1, \dots, e_n)$ une base de E et $B' = (e'_1, \dots, e'_p)$ une base de F. Une application $f : E \longrightarrow F$ est linéaire si et seulement si, il existe $n \times p$ scalaires (a_{ij}) $(1 \leq i \leq n,\ 1 \leq j \leq p)$ tels que, pour tout vecteur $x = x_1e_1 + x_2e_2 + \cdots + x_ne_n$ de E, $f(x) = (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)\, e'_1 + (a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n)\, e'_2 + \cdots + (a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pn}x_n)\, e'_p$.*

Démonstration.

- Supposons f linéaire. Les vecteurs $f(e_1), \dots, f(e_n)$ étant dans F, ils s'écrivent chacun comme combinaison linéaire unique des vecteurs de B'. Donc pour chaque $i = 1, \dots, n$, il existe des scalaires $a_{1i}, a_{2i}, \dots, a_{pi}$ dans K tels que

$$f(e_i) = a_{1i}e'_1 + a_{2i}e'_2 + \cdots + a_{pi}e'_p.$$

Soit maintenant $x = x_1e_1 + x_2e_2 + \cdots + x_ne_n$ un vecteur arbitraire de E. La linéarité de f entraîne $f(x) = x_1f(e_1) + x_2f(e_2) + \cdots + x_nf(e_n)$

$$= x_1\left(a_{11}e'_1 + a_{21}e'_2 + \cdots + a_{p1}e'_p\right) + x_2\left(a_{12}e'_1 + a_{22}e'_2 + \cdots + a_{p2}e'_p\right) + \dots$$

$$+x_n\left(a_{1n}e'_1 + a_{2n}e'_2 + \cdots + a_{pn}e'_p\right)$$

En regroupant convenablement les cœfficents associés aux vecteurs $e'_1, \dots, e'_p$, on obtient
$f(x) = (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)\,e'_1 + (a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n)\,e'_2 + \cdots + (a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pn}x_n)\,e'_p$.

– Soit $f : E \longrightarrow F$ une application définie par l'expression donnée dans l'énoncé de la proposition. Montrons qu'elle est linéaire. Soient les vecteurs $x = x_1e_1 + \cdots + x_ne_n$ et $y = y_1e_1 + \cdots + y_ne_n$ de E et α, β des scalaires de $\mathbb{K}$. Alors

$$\alpha x + \beta y = (\alpha x_1 + \beta y_1)e_1 + \cdots + (\alpha x_n + \beta y_n)e_n.$$

Donc

$$f(\alpha x + \beta y) = \Big[a_{11}(\alpha x_1 + \beta y_1) + \cdots + a_{1n}(\alpha x_n + \beta y_n)\Big]e'_1$$
$$+ \cdots + \Big[a_{p1}(\alpha x_1 + \beta y_1) + \cdots + a_{pn}(\alpha x_n + \beta y_n)\Big]e'_p.$$

Et l'on vérifie bien que $f(\alpha x+\beta y) = \alpha f(x)+\beta f(y)$, ce qui complète la preuve.
■

Ainsi, si E et F sont des espaces vectoriels de dimensions finies, il devient très facile de savoir si une application $f : E \longrightarrow F$ est linéaire. Il suffit de regarder si elle a la forme précédente. A titres d'exemples, en jetant un coup d'œil sur les applications données dans les exemples 2.1.8 et 2.1.11, il est immédiat que la première application doit être linéaire, contrairement à la seconde.

Une autre conséquence immédiate et que toute application linéaire de $\mathbb{K}$ dans $\mathbb{K}$ est nécessairement de la forme $x \mapsto ax$, où a est un élément fixé de $\mathbb{K}$. En effet, en posant $a = f(1)$, il résulte que $f(x) = f(x \cdot 1) = x \cdot f(1) = xa$ pour tout $x \in \mathbb{K}$.
Nous reviendrons dans le chapitre des matrices à l'expression d'une application linéaire donnée dans la proposition 2.3.5.

Il est naturel de se demander si les images $b_1 = f(a_1), \ldots, b_n = f(a_n)$ des vecteurs $a_1, \ldots, a_n$ forment une base de F. Pour cela, commençons par le fait suivant :

Proposition 2.3.6. *Soit $f : E \longrightarrow F$ une application linéaire et $B = (a_1, \ldots, a_n)$ une base de E. Alors $(f(a_1), \ldots, f(a_n))$ est une partie génératrice de $\mathsf{Im}(f)$: $\mathsf{Im}(f) = Vect(f(a_1), \ldots, f(a_n))$.*

Démonstration. D'abord, il est clair que les vecteurs $f(a_1), \ldots, f(a_n)$ appartiennent à $\mathsf{Im}(f)$. Soit y un vecteur quelconque de $Im(f)$. Il existe $x \in E$ tel que $y = f(x)$. Décomposons $x = \alpha_1 a_1 + \cdots + \alpha_n a_n$ dans la base $(a_1, \ldots, a_n)$. Il vient $y = f(\alpha_1 a_1 + \cdots + \alpha_n a_n) = \alpha_1 f(a_1) + \cdots + \alpha_n f(a_n)$, d'où le résultat voulu. ■

Remarque 2.3.7. Avec un raisonnement analogue, on peut affirmer d'une façon générale que si $f : E \longrightarrow F$ est une application linéaire et $(v_1, \ldots, v_p)$ est une famille finie de vecteurs de E, alors :

$$\boxed{f\left(\mathsf{Vect}(v_1, \ldots, v_p)\right) = \mathsf{Vect}(f(v_1), \ldots, f(v_p))}$$

Ceci découle de l'égalité $f(\alpha_1 v_1 + \cdots + \alpha_p v_p) = \alpha_1 f v_1) + \cdots + \alpha_p f(v_p)$ pour tous scalaires $\alpha_1, \ldots, al_p$ dans $\mathbb{K}$.

Lemme 2.3.8. *Soit $f : E \longrightarrow F$ une application linéaire.*

1. *Si $(a_1, \ldots, a_n)$ est une partie liée de E, alors $(f(a_1), \ldots, f(a_n))$ est une partie liée de F.*
2. *Si $(f(a_1), \ldots, f(a_n))$ est une partie libre de F, alors $(a_1, \ldots, a_n)$ est une partie libre de E.*

Démonstration. (a) Par hypothèse, il existe des scalaires non tous nuls $\alpha_1, \ldots, \alpha_n \in K$ tels que $\alpha_1 a_1 + \cdots + \alpha_n a_n = 0_E$. Il en résulte que $\alpha_1 f(a_1) + \cdots + \alpha_n f(a_n) = f(\alpha_1 a_1 + \cdots + \alpha_n a_n) = f(0_E) = 0_F$. Donc $(f(a_1), \ldots, f(a_n))$ est une partie liée de F.
L'assertion (b) n'est autre que la contraposée de a). ■

Voici maintenant un critère pratique utilisant les bases pour savoir si une application linéaire est injective, surjective, bijective.

Théorème 2.3.9. *Soient $f : E \longrightarrow F$ une application linéaire et $B = (a_1, \ldots, a_n)$ une base de E. On note $f(B) = (f(a_1), \ldots, f(a_n))$.*

1. *f est injective* $\Longleftrightarrow$ *$f(B)$ est libre dans F.*
2. *f est surjective* $\Longleftrightarrow$ *$f(B)$ est génératrice de F.*
3. *f est bijective* $\Longleftrightarrow$ *$f(B)$ est une base de F.*

Démonstration.

1. $\Rightarrow$) Supposons que f est injective et soient $\lambda_1, \ldots, \lambda_n$ dans K tels que $\lambda_1 f(a_1) + \cdots + \lambda_n f(a_n) = 0_F$. Alors $f(\lambda_1 a_1 + \cdots + \lambda_n a_n) = 0_F$, ce qui signifie $\lambda_1 a_1 + \cdots + \lambda_n a_n \in \ker(f)$. Or, f est injective, donc $\ker(f) = \{0_E\}$. Par conséquent $\lambda_1 a_1 + \cdots + \lambda_n a_n = 0_E$. Mais comme $(a_1, \ldots, a_n)$ est libre dans E, on déduit que $\lambda_1 = \cdots = \lambda_n = 0$. Ainsi, $f(B)$ est libre dans F.
   $\Leftarrow$) Supposons que $f(B)$ est libre. Pour montrer que f est injective, il suffit de prouver que $Ker(f) = \{0_E\}$. Soit donc un vecteur $x \in \ker(f)$, c'est-à-dire, $x \in E$ et $f(x) = 0_F$. Il existe $\alpha_1, \ldots, \alpha_n \in K$ tels que $x = \alpha_1 a_1 + \cdots + \alpha_n a_n$, donc $0_F = f(x) = \alpha_1 f(a_1) + \cdots + \alpha_n f(a_n)$. Puisque $f(B)$ est libre dans F, on obtient $\alpha_1 = \cdots = \alpha_n = 0$. D'où $x = 0_E$, et par suite, $\ker(f) = \{0_E\}$.
2. On sait que $Im(f) = Vect(f(a_1), \ldots, f(a_n))$. Donc
   f est surjective $\Leftrightarrow$ $Im(f) = F \Leftrightarrow Vect(f(a_1), \ldots, f(a_n)) = F \Leftrightarrow f(B)$ engendre F.
3. Se déduit directement de a) et b). ■

La proposition qui suit montre que l'injectivité ou la surjectivité d'une application linéaire a une influence sur les dimensions des espaces vectoriels E et F en question.

Corollaire 2.3.10. *Soient E, F deux espaces vectoriels de dimensions finies sur K et $f : E \longrightarrow F$ une application linéaire. Alors*

1. *f est injective* $\Rightarrow \dim(E) \leq \dim(F)$.
2. *f est surjective* $\Rightarrow \dim(E) \geq \dim(F)$.
3. *f est bijective* $\Rightarrow \dim(E) = \dim(F)$.

Démonstration. Soit $B = (a_1, \ldots, a_n)$ une base de E et notons $f(B) = (f(a_1), \ldots, f(a_n))$.

* Si f est injective, alors $f(B)$ est libre, donc card(f(B))$\leq \dim(F)$. D'où $n = \dim(E) \leq \dim(F)$.

∗ Si f est surjective, alors $f(B)$ est une partie génératrice de F. Ceci entraîne $\mathsf{card}(f(B)) \geq \dim(F)$, d'où $\dim(E) \geq \dim(F)$.

∗ La troisième assertion découle immédiatement des deux premières. ■

Remarque 2.3.11. La contraposée du corollaire ci-dessus exprime que si E et F sont deux espaces vectoriels de dimensions finies sur K, et si $f : E \longrightarrow F$ est une application linéaire, alors :

1. $\dim(E) > \dim(F) \Rightarrow f$ n'est pas injective.
2. $\dim(E) < \dim(F) \Rightarrow f$ n'est pas surjective.
3. $\dim(E) \neq \dim(F) \Rightarrow f$ n'est pas bijective.

Dans le cas spécial où E et F ont même dimension sur K, on obtient la situation avantageuse suivante :

Corollaire 2.3.12. *Soient E F deux espaces vectoriels de dimensions finies sur K, avec $\dim(E) = \dim(F)$, et soit $f : E \longrightarrow F$ une application linéaire. Alors : f est injective $\iff$ f est surjective $\iff$ f est bijective.*

Démonstration. Soit $B = (a_1, \ldots, a_n)$ une base de E et notons $f(B) = (f(a_1), \ldots, f(a_n))$.

1. Si f est injective, alors $f(B)$ est libre dans F. Mais $\dim(F) = n$ et $f(B)$ est de cardinal n, donc $f(B)$ est une base de F. D'où f est bijective.
2. Si f est surjective, $f(B)$ engendre F. Puisque $\dim(F) = n$ et $f(B)$ est de cardinal n, alors $f(B)$ est une base de F, et par suite, f est bijective. ■

Ce corollaire important est très souvent utilisé dans les exercices pour démontrer qu'une application linéaire est bijective, pourvu que la condition que $\dim_K(E) = \dim_K(F)$ soit vérifiée, ce qui est en particulier le cas pour un endomorphisme ($E = F$). Il permet de simplifier les calculs et vérifier par exemple seulement que $\ker(f) = \{0_E\}$ (injectivité de f).

Nous proposerons ultérieurement une seconde démonstration de ce dernier résultat concernant le cas où $\dim(E) = \dim(F)$ quand nous aborderons le théorème du rang.

2.4 Opérations sur les applications linéaires

2.4.1 L'espace vectoriel $\mathcal{L}_{\mathbb{K}}(E, F)$

Soient E et F deux espaces vectoriels sur le même corps $\mathbb{K}$. Notons $\mathcal{L}_{\mathbb{K}}(E, F)$ l'ensemble des applications linéaires de E dans F. Lorsque $E = F$, on notera $\mathcal{L}_{\mathbb{K}}(E, E) = \mathcal{L}_{\mathbb{K}}(E)$, qui est l'ensemble des endomorphismes de E. Pour $F = \mathbb{K}$, l'ensemble $\mathcal{L}(E, \mathbb{K})$ des formes linéaires sur E s'appelle l'**espace dual** de E ou simplement le dual de E et est noté E^*.

Nous allons définir dans $\mathcal{L}_{\mathbb{K}}(E, F)$ une opération d'addition et une multiplication externe sur le corps $\mathbb{K}$ qui font de $\mathcal{L}_{\mathbb{K}}(E, F)$ un espace vectoriel sur $\mathbb{K}$.

- **Somme de deux applications linéaires :** Soient f et g deux applications linéaires de E dans F. On définit l'application $f+g : E \to F$ par : $(f+g)(x) = f(x) + g(x)$ pour tout $x \in E$.
- **Produit d'une application linéaire par un scalaire :** On définit le produit externe $\lambda.f$ d'une application linéaire $f \in \mathcal{L}_{\mathbb{K}}(E, F)$ par un scalaire $\lambda \in \mathbb{K}$ en posant : $(\lambda \cdot f)(x) = \lambda \cdot f(x)$ pour tout $x \in E$.

Proposition 2.4.1. *Soient E et F deux espaces vectoriels sur $\mathbb{K}$.*
a) Pour toutes applications linéaires f et g de E dans F, et pour tout scalaire λ dans $\mathbb{K}$, les applications $f + g$ et λf sont linéaires.
b) Muni de ces opérations, l'ensemble $\mathcal{L}_{\mathbb{K}}(E, F)$ est un espace vectoriel sur $\mathbb{K}$.

Démonstration. a) Pour tous vecteurs $x, y \in E$ et pour tous scalaires $\alpha, be \in K$, $(f + g)(\alpha x + \beta y) = f(\alpha x + \beta y) + g(\alpha x + \beta y) = (\alpha f(x) + \beta f(y)) + (\alpha g(x) + \beta g(y)) = \alpha(f(x)+g(x))+\beta(f(y)+g(y)) = \alpha(f+g)(x)+\beta(f+g)(y)$.
b) Montrons que $\mathcal{L}_{\mathbb{K}}(E, F)$ est un sous-espace vectoriel de l'espace vectoriel $\mathcal{F}_{\mathbb{K}}(E, F)$ des applications de E dans F. D'abord, l'application nulle $\theta : E \mapsto F, x \mapsto 0_F$ est linéaire, donc $\theta \in \mathcal{L}_{\mathbb{K}}(E, F)$. L'assertion a) achève donc la preuve. ∎

Pour montrer que deux applications $f : E \longrightarrow F$ et $g : E \longrightarrow F$ sont égales, il faut prouver que $f(x) = g(x)$ $\forall x \in E$. Il s'agit là de la définition générale de l'égalité de deux applications d'un ensemble dans un autre.

Nous verrons plus loin (chapitre 3) que si les dimensions de E et F sont finies, la dimension de $\mathcal{L}_{\mathbb{K}}(E, F)$ est finie et est égale au produit $(\dim E)(\dim F)$.

En prenant le cas particulier $E = F$ dans la proposition 2.4.1, on obtient :

Corollaire 2.4.2. *Soit E un espace vectoriel sur $\mathbb{K}$. Alors l'ensemble $\mathcal{L}_{\mathbb{K}}(E)$ des endomorphismes de E est un espace vectoriel sur $\mathbb{K}$.*

2.4.2 Composition des applications linéaires

Soient $f : E \longrightarrow F$ et $g : F \longrightarrow G$ deux applications où l'ensemble d'arrivée F de f coïncide avec l'ensemble de départ de g. On rappelle que l'*application composée* $g \circ f : E \longrightarrow G$ est définie par $(g \circ f)(x) = g[f(x)]$ pour tout élément $x \in E$.

Proposition 2.4.3. *Soient E, F et G des espaces vectoriels sur le même corps $\mathbb{K}$. Si $f : E \longrightarrow F$ et $g : F \longrightarrow G$ sont des applications linéaires, alors $g \circ f : E \longrightarrow G$ est aussi une application linéaire.*

Démonstration. Soient $x, y \in E$ et $\alpha, \beta \in \mathbb{K}$. En utilisant la linéarité de f et g, on a

$$\begin{aligned}
(g \circ f)(\alpha x + \beta y) &= g\left(f(\alpha x + \beta y)\right) && \text{(par définition de la loi } \circ) \\
&= g\left(\alpha f(x) + \beta f(y)\right) && \text{(par linéarité de } f) \\
&= \alpha g(f(x)) + \beta g(f(y)) && \text{(par linéarité de } g) \\
&= \alpha (g \circ f)(x) + \beta (g \circ f)(y) && \text{(par définition de la loi } \circ) \blacksquare
\end{aligned}$$

Cette proposition se traduit en écrivant :

$$\boxed{\forall f \in \mathcal{L}(E, F),\ \forall g \in \mathcal{L}(F, G),\ g \circ f \in \mathcal{L}(E, G)}$$

En particulier, si $E = F = G$, alors :

$$\boxed{\forall f, g \in \mathcal{L}(E) : \quad g \circ f \in \mathcal{L}(E)}$$

Si $f \in \mathcal{L}_K(E)$, on pose $f^0 = id_E$ et par récurrence $f^{n+1} = f^n \circ f$ pour $n \in \mathbb{N}^*$, c'est-à-dire, $f^n = f \circ f \circ \cdots \circ f$ (n fois). Il est alors clair que $f^n \circ f^p = f^p \circ f^n = f^{n+p}$ pour tous entiers $n, p \in \mathbb{N}^*$.

Notons que la composition $\circ$ des applications étant toujours associative, elle le reste évidemment lorsqu'on travaille sur les applications linéaires.

Par contre, la loi $\circ$ n'est pas en général commutative : il existe des endomorphismes $f, g \in \mathcal{L}_{\mathbb{K}}(E)$ tels que $g \circ f \neq f \circ g$.

A présent, nous allons voir que la composition des applications linéaires est distributive à gauche et à droite par rapport à l'addition :

Proposition 2.4.4. *Soient E, F et G des espaces vectoriels sur $\mathbb{K}$. Pour toutes applications linéaires $f, f' \in \mathcal{L}(E, F)$ et $g, g' \in \mathcal{L}(F, G)$, et pour tous scalaires $\alpha, \beta \in \mathbb{K}$, on a :*
a) $(g + g') \circ f = g \circ f + g' \circ f$.
b) $g \circ (f + f') = g \circ f + g \circ f'$.

Démonstration. a) Soit x un vecteur quelconque de E. On a :
a) $[(g + g') \circ f](x) = (g + g')[f(x)] = g[f(x)] + g'[f(x)] = (g \circ f)(x) + (g' \circ f)(x) = [(g \circ f) + (g' \circ f)](x)$. On en déduit que les applications $(g + g') \circ f$ et $(g \circ f) + (g' \circ f)$ sont égales.
b) $[g \circ (f + f')](x) = g[(f + f')(x)] = g[f(x) + f'(x)] = g[f(x)] + g[f'(x)] = (g \circ f)(x) + (g \circ f')(x) = [(g \circ f) + (g \circ f')](x)$. D'où l'égalité désirée. ∎

Remarquons que dans la propriété a), la linéarité des applications f, f', g, g' n'a pas été utilisée, tandis que dans b), nous avons tenu compte uniquement que g est linéaire.

Proposition 2.4.5. *Soient E, F et G des espaces vectoriels sur $\mathbb{K}$. Pour toutes appications linéaires $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$ et pour tout scalaire $\alpha \in \mathbb{K}$, $(\alpha g) \circ f = g \circ (\alpha f) = \alpha(g \circ f)$.*

Démonstration. Les vérifications sont automatiques, puisque les expressions :
$[\alpha(g \circ f)](x) = \alpha[(g \circ f)(x)] = \alpha[g(f(x))] = \alpha(g \circ f)(x) = \alpha[g(f(x))]$,
$[(\alpha g) \circ f](x) = (\alpha g)(f(x)) = \alpha[g(f(x))]$,
$[g \circ (\alpha f)](x) = g[(\alpha f)(x)] = g[\alpha f(x)] = \alpha[g(f(x)]$
coïncident pour tout vecteur $x \in E$. ∎

Observons que pour montrer la formule $(\alpha g) \circ f = \alpha(gf)$, on n'a pas utilisé de linéarité. La formule $g \circ (\alpha f) = \alpha(g \circ f)$ a nécessité la linéarité de g mais pas de f.

2.4.3 Isomorphismes

Soit $f : E \longrightarrow F$ une application quelconque (non nécessairement linéaire). On rappelle que si f est bijective, il existe une application $f^{-1} : F \longrightarrow E$ définie par :

$$\boxed{f^{-1}(y) = x \iff f(x) = y \qquad \text{(pour tous } x \in E \text{ et } y \in F)}$$

De plus, f^{-1} est aussi bijective et vérifie :

$$\boxed{f \circ f^{-1} = \mathtt{id}_F, \quad f^{-1} \circ f = \mathtt{id}_E}$$

Cette bijection f^{-1} s'appelle la **bijection réciproque** de f.
Maintenant, que peut-on dire de plus de f^{-1} si f est linéaire ?

Proposition 2.4.6. *Si $f : E \longrightarrow F$ un isomorphisme, alors sa bijection réciproque $f^{-1} : F \to E$ est aussi un isomorphisme.*

Démonstration. On sait que f^{-1} est bijective. Montrons qu'elle est linéaire. Soient $y, y' \in F$ et $\alpha, \beta \in K$. Il existe un unique vecteur $x \in E$ tel que $f(x) = y$ et il existe un unique vecteur $x' \in E$ tel que $f(x') = y'$. On a donc $x = f^{-1}(y)$ et $x' = f^{-1}(y')$. Etablissons que $f^{-1}(\alpha y + \beta y') = \alpha f^{-1}(y) + \beta f^{-1}(y')$. Puisque f est linéaire, $f(\alpha x + \beta x') = \alpha f(x) + \beta f(x') = \alpha y + \beta y'$. Le vecteur $\alpha x + \beta x'$ de E est donc l'antécédent par la bijection f du vecteur $\alpha y + \beta y'$ de F. Ainsi, $f^{-1}(\alpha y + \beta y') = \alpha x + \beta x' = \alpha f^{-1}(y) + \beta f^{-1}(y')$, ce qui montre bien que f^{-1} est linéaire. ∎

Proposition 2.4.7. *Soient E, F, G des espaces vectoriels sur K.*
Si $f : E \longrightarrow F$ et $g : F \longrightarrow G$ sont des isomorphisme, alors $g \circ f : E \longrightarrow G$ est aussi un isomorphisme. En outre,

$$\boxed{(g \circ f)^{-1} = f^{-1} \circ g^{-1}}$$

Démonstration. D'abord, $g \circ f$ est linéaire car f et g sont linéaires. D'après le cours SMIA1 d'algèbre, nous savons que la composée de deux bijections est aussi une bijection; donc $g \circ f$ est bijective. Enfin,

$$(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ (f \circ f^{-1}) \circ g^{-1} = g \circ \mathtt{id}_F \circ g^{-1} = g \circ g^{-1} = \mathtt{id}_G.$$

Et on vérifie de même que

$$f^{-1} \circ g^{-1}) \circ (g \circ f) = \mathtt{id}_E.$$

D'où l'égalité $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. ■

Il ne faut pas s'étonner de la formule $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$, où l'ordre est inversé. La loi $\circ$ n'étant pas commutative, en général $(g \circ f)^{-1} \neq g^{-1} \circ f^{-1}$. On a $g^{-1} : G \longrightarrow F$ et $f^{-1} : F \longrightarrow E$, donc $f^{-1} \circ g^{-1} : G \longrightarrow E$ est bien définie. Par contre, $g^{-1} \circ f^{-1}$ n'est même pas définie si $E \neq G$. Donnons un exemple concret de la vie quotidienne pour se détendre. Avant de sortir le matin de chez soi pour aller à la faculté des sciences de Tétouan, il faut d'abord mettre les chaussettes puis ensuite les chaussures. Mais en rentrant le soir à la maison, content d'avoir compris les cours et TD de SMIA2, on inverse l'ordre des choses : on commence d'abord par enlever les chaussures avant d'enlever les chaussettes ! Donc la formule précédente est tout à fait naturelle.

Définition 2.4.8. *Soient E et F deux espaces vectoriels sur le même corps K. S'il existe un isomorphisme $f : E \longrightarrow F$, on dit que E est **isomorphe** à F, et l'on écrit $E \cong F$.*

Nous voyons immédiatement que pour tous espaces vectoriels E, F, G sur $\mathbb{K}$:

* $E \cong E$ (*réflexivité*),
* Si $E \cong F$, alors $F \cong E$ (*symétrie*),
* Si $E \cong F$ et $F \cong G$, alors $E \cong G$ (*transitivité*).

En effet, puisque l'application identité $\mathtt{id}_E : E \longrightarrow E$ est un isomorphisme, alors $E \cong E$. Maintenant, si $E \cong F$, il existe un isomorphisme $f : E \longrightarrow F$, donc $f^{-1} : F \longrightarrow E$ est un isomorphisme, d'où $F \cong E$. Finalement, la troisième propriété résulte de la proposition 2.4.6.
Les trois assertions ci-dessus expriment que la relation $E \cong F$ est une **relation d'équivalence**. Il est ainsi légitime de dire que E et F sont **isomorphes** lorsque $E \cong F$.

Etudions maintenant l'effet d'isomorphisme sur les dimensions. En général, un isomorphisme conserve plusieurs propriétés d'un espace vectoriel vers un autre. Qu'en est-il pour les dimensions ?

Proposition 2.4.9. *Soient E et F deux espaces vectoriels de dimensions finies sur $\mathbb{K}$. Alors E est isomorphe à F si et seulement si* $\dim_{\mathbb{K}}(E) = \dim_{\mathbb{K}}(F)$.

Démonstration.

* Si E est isomorphe à F, il existe un isomorphisme f de E sur F. D'après le corollaire 2.3.9, la bijectivité de f entraîne $\dim(F) = \dim(E)$.

* Réciproquement, supposons que $\dim(E) = \dim(F)$ que nous notons n et construisons un isomorphisme f de E dans F. Pour ce faire, soient $B = (a_1, \ldots, a_n)$ une base de E et $B' = (b_1, \ldots, b_n)$ une base de F. Selon le théorème 2.3.1, il existe une application linéaire (unique) $f : E \longrightarrow F$ telle que $f(a_i) = b_i \ \ \forall i \in \{1, \ldots, n\}$. On a donc $f(B) = B'$, c'est-à-dire, l'image par f d'une base de E est une base de F. Ainsi, f est bijective, d'après le théorème 2.3.8, et donc f est un isomorphisme. Ceci montre bien que E et F sont isomorphes. ∎

Remarque 2.4.10. Dans la preuve précédente :

1. L'isomorphisme f est défini tout simplement par :

$$\forall x = \alpha_1 a_1 + \cdots + \alpha_n a_n \in E, \ \ f(x) = \alpha_1 b_1 + \cdots + \alpha_n b_n.$$

L'isomorphisme réciproque $f^{-1} : F \longrightarrow E$ de f est déterminé par

$$f^{-1}(\alpha_1 b_1 + \cdots + \alpha_n b_n) = \alpha_1 a_1 + \cdots + \alpha_n a_n \quad \forall y = \alpha_1 b_1 + \cdots + \alpha_n b_n \in F.$$

2. Il est aussi très facile de montrer que f est bijective en utilisant le noyau et l'image de f. En effet, si $x = \alpha_1 a_1 + \cdots + \alpha_n a_n \in \ker(f)$, alors

$$0_F = f(\alpha_1 a_1 + \cdots + \alpha_n a_n = \alpha_1 b_1 + \cdots + \alpha_n b_n,$$

donc $\alpha_i = 0 \ \forall i$. D'où $x = 0_E$ et par suite $\ker(f) = \{0_E\}$, ce qui prouve l'injectivité de f. Quant à la surjectivité de f, elle provient du fait que $\mathsf{Im}(f) = \mathsf{Vect}(f(a_1), \ldots, f(a_n)) = \mathsf{Vect}(b_1, \ldots, b_n) = F$.
Vous pouvez également vérifier que $\forall y \in F$, $\exists! x \in E$ tel que $y = f(x)$.

i Il y a une ressemblance entre la proposition 2.4.9 et le résultat suivant bien connu dans la théorie des ensembles finis : Soient E et F deux ensembles finis. Alors il existe une bijection f entre E et F si et seulement si E et F ont le même cardinal.

Comme conséquence, nous avons le corollaire suivant.

Corollaire 2.4.11. *Tout espace vectoriel E de dimension finie n sur $\mathbb{K}$ est isomorphe à $\mathbb{K}^n$.*

Démonstration. Résulte de la proposition 2.4.8, car $\dim(E) = \dim(\mathbb{K}^n)$. ∎

Exemple 2.4.12. Soit $\psi : \mathbb{R}^2 \longrightarrow \mathbb{C}$ l'application linéaire définie par $\psi(x,y) = x+iy$ que nous avons traitée dans l'exemple 2.2.8. Comme ψ est un isomorphisme, alors les deux $\mathbb{R}$-espaces vectoriels $\mathbb{R}^2$ et $\mathbb{C}$ sont isomorphes. De plus, on a bien $\dim_{\mathbb{R}} \mathbb{R}^2 = \dim_{\mathbb{R}} \mathbb{C} = 2$.

2.4.4 Le groupe linéaire $GL_{\mathbb{K}}(E)$

E étant un espace vectoriel, désignons par $GL_{\mathbb{K}}(E)$ l'ensemble des automorphismes de E, c'est-à-dire l'ensemble des applications linéaires bijectives de E sur E. On sait que la loi $\circ$ est associative. De plus, il est immédiat que :

$*$ l'application identique $\mathtt{id}_E$ est un automorphisme de E ; c'est l'élément neutre de $GL_{\mathbb{K}}(E)$.

$*$ l'application composée $g \circ f$ de deux automorphismes f et g de E est un automorphisme de E. En d'autres termes :

$$\forall f, g \in GL_{\mathbb{K}}(E), \quad g \circ f \in GL_{\mathbb{K}}(E).$$

$*$ l'application réciproque f^{-1} d'un automorphisme f de E est un automorphisme de E. C'est-à-dire,

$$\forall f \in GL_{\mathbb{K}}(E), \quad f^{-1} \in GL_{\mathbb{K}}(E).$$

Ces arguments permettent d'énoncer :

Proposition 2.4.13. *L'ensemble $GL_{\mathbb{K}}(E)$ des automorphismes de E est un groupe pour la composition des applications, appelé **groupe linéaire** de E.*

Remarque 2.4.14. Rappelons que pour tout λ dans $\mathbb{K}$, l'homothétie de rapport λ est l'endomorphisme $h_\lambda : E \longrightarrow E, x \mapsto \lambda x$. En particulier, l'homothétie h_1 de rapport 1 est l'endomorphisme identité id_E. Soient h_{λ_1} et h_{λ_2} deux homothéties de rapports respectifs λ_1 et λ_2. Alors leur composée $h_{\lambda_1} \circ h_{\lambda_2}$ est l'homothétie $h_{\lambda_1\lambda_2}$ de rapport $\lambda_1\lambda_2$, car pour tout $x \in E$:

$$(h_{\lambda_1} \circ h_{\lambda_2})(x) = h_{\lambda_1}(h_{\lambda_2}(x)) = h_{\lambda_1}(\lambda_2 x) = \lambda_1(\lambda_2 x) = (\lambda_1\lambda_2)x = h_{\lambda_1\lambda_2}(x).$$

D'autre part, toute homothétie h_λ de rapport $\lambda \neq 0$ est bijective et son application réciproque h_λ^{-1} est l'homothétie $h_{\frac{1}{\lambda}}$, puisque $h_\lambda \circ h_{\frac{1}{\lambda}} = h_{\lambda\frac{1}{\lambda}} = h_1 = \mathtt{id}_E$.
Il résulte de ces propriétés que l'ensemble $\mathcal{H}$ des homothéties de E de rapports λ non nuls est un sous-groupe du groupe linéaire $GL_{\mathbb{K}}(E)$ de E.
Remarquons que $h_{\lambda_1} \circ h_{\lambda_2} = h_{\lambda_2} \circ h_{\lambda_1}$, et donc le groupe $\mathcal{H}$ est commutatif, malgré le fait que le groupe $GL_{\mathbb{K}}(E)$ n'est pas commutatif.

2.4.5 L'algèbre $\mathcal{L}_{\mathbb{K}}(E)$

Le comportement sympathique de la composition des applications linéaires vis-à-vis de l'addition vu dans la section 2.4.2 nous conduit au résultat suivant lorsqu'on travaille avec les endomorphismes ($E = F = G$):

Proposition 2.4.15. $(\mathcal{L}_{\mathbb{K}}(E), +, \circ)$ *est un anneau unitaire.*

Démonstration. $(\mathcal{L}_K(E), +)$ est déjà un groupe abélien, car $(\mathcal{L}_K(E), +, \cdot)$ est un espace vectoriel sur $\mathbb{K}$. De plus, nous avons vu que la loi $\circ$ est associative, distributive à droite et à gauche par rapport à l'addition. D'où $(\mathcal{L}_{\mathbb{K}}(E), +, \circ)$ est un anneau. C'est même un anneau unitaire, puisqu'il possède un élément neutre $\mathtt{id}_E$ pour la loi $\circ$. ∎

Notons que l'anneau $\mathcal{L}_K(E)$ n'est pas en général commutatif. Nous reviendrons sur ce point dans un prochain chapitre quand nous étudierons l'anneau des matrices carrées.

Remarque 2.4.16. Soit $(A, +, \times)$ un anneau commutatif unitaire quelconque d'élément unité 1_A. Puisque $(A, +)$ est un groupe, tout élément $x \in A$ est symétrisable (pour l'addition) : $\exists x' \in A$ tel que : $x + x' = 0_E$. Par contre, les éléments de A ne sont pas nécessairement tous inversibles (pour la multiplication) (il suffit de penser à l'anneau $(\mathbb{Z}, +, \times)$, à titre d'exemple, où 1 et -1 sont les seuls éléments inversibles). Néanmoins, nous rappelons que l'ensemble $\mathcal{U}(A)$ des éléments inversibles de A forment un groupe multiplicatif $(\mathcal{U}(A), \times)$. En revenant à notre anneau $(\mathcal{L}_{\mathbb{K}}(E), +, \circ)$ associé à un espace vectoriel E sur $\mathbb{K}$, il est important de souligner que le groupe des éléments inversibles de $(\mathcal{L}_{\mathbb{K}}(E), +, \circ)$ est précisément le groupe linéaire $GL_{\mathbb{K}}(E)$.

Nous venons de voir que $\mathcal{L}_{\mathbb{K}}(E)$ est muni de deux lois internes $+$ et $\circ$ et d'une loi externe telles que

* $(\mathcal{L}_{\mathbb{K}}(E), +, .)$ est un espace vectoriel sur $\mathbb{K}$,
* $(\mathcal{L}_{\mathbb{K}}(E), +, \circ)$ est un anneau,
* $\alpha(g \circ f) = (\alpha g) \circ f = g \circ (\alpha f)$ pour tous $f, g \in \mathcal{L}_{\mathbb{K}}(E)$ et pour tout $\alpha \in \mathbb{K}$.

Pour ces trois raisons, on dit que $(\mathcal{L}_{\mathbb{K}}(E), +, \cdot, \circ)$ est une **algèbre** sur $\mathbb{K}$, conformément à la définition plus générale suivante :

Définition 2.4.17. *Une structure d'algèbre sur $\mathbb{K}$ sur un ensemble A est la donnée de deux lois de composition internes, notées habituellement $+$ et $\times$, et d'une loi externe sur $\mathbb{K}$ telle que :*

1. $(A, +, .)$ *est un espace vectoriel sur* $\mathbb{K}$.

2. $(A, +, \times)$ *est un anneau.*

3. $\alpha(ab) = (\alpha a) \times b = a \times (\alpha b)$ *pour tous* $a, b \in A$ *et* $\alpha \in \mathbb{K}$.

Remarque 2.4.18. Soit $(A, +, \cdot, \times)$ une algèbre sur $\mathbb{K}$. Si l'anneau $(A, +, \times)$ est unitaire, c'est-à-dire, si la loi interne $\times$ possède un élément neutre (noté 1_A), on dit que $(A, +, \cdot, \times)$ une algèbre *unitaire*. De même, si la loi $\times$ est commutative, l'algèbre est dite *commutative*.

Exemple 2.4.19. L'exemple le plus simple d'algèbre sur $\mathbb{K}$ est l'algèbre $(\mathbb{K}, +, \times, .)$. Cette algèbre est évidemment commutative et unitaire.

Exemple 2.4.20. Nous venons juste de voir que si E est un espace vectoriel sur $\mathbb{K}$, alors $(\mathcal{L}_{\mathbb{K}}(E), +, \cdot, \circ)$ est une algèbre unitaire non commutative.

Exemple 2.4.21. Un autre exemple très important d'algèbre est donné par $(\mathbb{K}[X], +, \times, .)$, l'algèbre des polynômes à une indéterminée sur le corps $\mathbb{K}$. En effet, il est aisé de voir que les trois conditions de la définition d'algèbre sont vérifiées. Cette algèbre est commutative unitaire.

Exemple 2.4.22. L'ensemble $(\mathcal{F}(\mathbb{R}, \mathbb{R}), +, \cdot, \times)$ des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ est une algèbre commutative unitaire.

Nous verrons au chapitres suivant un autre exemple fondamental d'algèbre formé par les matrices carrées.
D'autre part, à l'instar des structures de groupes, anneaux, corps et espaces vectoriels, les notions de sous-algèbre et d'homomorphisme d'algèbres se définissent de manière analogues et naturelles et jouent des rôles importants :

Définition 2.4.23. *Soient $(A, +, \times, .)$ et $(B, +, \times, .)$ deux algèbres sur $\mathbb{K}$ et $f : A \longrightarrow B$ une application. On dit que f est un* ***homomorphisme d'algèbres*** *si f est à la fois un homomorphisme d'anneaux et une application linéaire, c'est-à-dire :*

$$f(a+b) = f(a) + f(b), \; f(a \times b) = f(a) \times f(b), \; f(\alpha a) = \alpha f(a)$$

pour tous $a, b \in A$ et $\alpha \in \mathbb{K}$.

Ainsi, un homomorphisme d'algèbres conserve à la fois les structures d'espaces vectoriels et les structures d'anneaux. On pourra donc utiliser toutes les propriétés connues des homomorphismes d'anneaux et des applications linéaires. On définit de même un **isomorphisme d'algèbres** comme étant un homomorphisme d'algèbres qui est bijectif.

Exemple 2.4.24. L'application $f : \mathbb{K}[X] \longrightarrow \mathbb{K}$ définie par $f(P) = P(0)$ est un homomorphisme d'algèbres de $(\mathbb{K}[X], +, \times, .)$ dans $(\mathbb{K}, +, \times, .)$. En effet, $f(P+Q) = (P+Q)(0) = P(0)+Q(0) = f(P)+f(Q)$, $f(P\times Q) = (P\times Q)(0) = P(0) \times Q(0) = f(P) \times f(Q)$ et $f(\alpha P) = (\alpha P)(0) = \alpha P(0) = \alpha f(P)$.

Exemple 2.4.25. Considérons $\mathbb{C}$ comme espace vectoriel sur le corps $\mathbb{K} = \mathbb{R}$. Alors l'application de conjugaison $\mathbb{C} \longrightarrow \mathbb{C}$, $z \mapsto \overline{z}$ est un homomorphisme d'algèbres. Il suffit de remarquer que $\overline{z + z'} = \overline{z}+\overline{z'}$, $\overline{z \times z'} = \overline{z}\times\overline{z'}$ et $\overline{\alpha z} = \alpha\overline{z}$ pour tous nombres complexes $z, z' \in \mathbb{C}$ et pour tout réel $\alpha \in \mathbb{R}$.
En fait, c'est même un isomorphisme d'algèbres.

Définition 2.4.26. *Soit $(A, +, \times, .)$ une algèbre sur $\mathbb{K}$ et S une partie non vide de A. On dit que S est une **sous-algèbre** de A si S est elle-même une algèbre sur $\mathbb{K}$ pour les lois induites par celles de A.*

Par analogie avec les sous-groupes, sous-anneaux et sous-espaces vectoriels, pour montrer qu'une partie non vide S est une sous-algèbre d'une algèbre $(A, +, \times, .)$, on ne vérifie pas toutes les conditions de la définition précédente, mais seulement quelques-unes, car la plupart d'entre elles sont déjà satisfaites dans A. Plus précisément, S doit être simultanément un sous-espace vectoriel de l'espace vectoriel $(A, +, .)$ et un sous-anneau de l'anneau $(A, +, \times)$. Ceci se traduit par :

Proposition 2.4.27. *Soit $(A, +, \times, .)$ une algèbre sur K et S une partie de A. Alors S est une* sous-algèbre *de A si et seulement si les conditions suivantes sont satisfaites :*

1. $0_A \in S$.

2. *pour tous* $a, b \in S$, $a + b \in S$.

3. *pour tous* $a \in S$ *et* $\alpha \in K$, $\alpha a \in S$.

4. *pour tous* $a, b \in S$, $a \times b \in S$.

Démonstration. Les trois premières conditions traduisent le fait que S est un sous-espace vectoriel de $(A, +, .)$. Il en découle que $(S, +, .)$ est aussi un espace vectoriel sur K, et en particulier, S est un sous-groupe du groupe $(A, +)$. La quatrième condition montre que S est stable pour la multiplication, et donc aussi un sous-anneau de l'anneau $(A, +, \times)$. ■

Exemple 2.4.28. $\mathbb{R}$ est une sous-algèbre de l'algèbre $(\mathbb{C}, +, \times, .)$, où $\mathbb{C}$ est considéré comme espace vectoriel sur $\mathbb{R}$.

Exemple 2.4.29. $\mathbb{K}$ est une sous-algèbre de l'algèbre $(\mathbb{K}[X], +, \times, .)$.

Exemple 2.4.30. Par contre, $\mathbb{K}_n[X]$ n'est pas une sous-algèbre de $\mathbb{K}[X]$. C'est un sous-espace de $\mathbb{K}[X]$ mais ce n'est pas un sous-anneau de $\mathbb{K}[X]$. C'est la condition 4 de la proposition 2.4.26 qui n'est pas valable.

Exemple 2.4.31. L'ensemble $\mathcal{C}(\mathbb{R}, \mathbb{R})$ des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$ est une sous-algèbre de l'algèbre $(\mathcal{F}(\mathbb{R}, \mathbb{R})$ des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. En effet, on sait que c'est un sous-espace vectoriel de l'espace vectoriel $\mathcal{F}(\mathbb{R}, \mathbb{R})$. En outre, il est stable pour la multiplication puisque le produit de deux fonctions continues est continue.

2.5 Théorème du noyau

Le théorème suivant joue un rôle central en algèbre linéaire. Il est souvent intitulé **théorème du noyau** ou **théorème du rang**.

Théorème 2.5.1. *Soit $f : E \to F$ une application linéaire avec E de dimension finie. Alors* $\ker(f)$ *et* $\mathsf{Im}(f)$ *sont de dimensions finies, et l'on a :*

$$\boxed{\dim E = \dim \ker(f) + \dim \mathsf{Im}(f)}$$

Démonstration. D'abord, puisque $\ker(f)$ est un sous-espace vectoriel de E, alors $\ker(f)$ est aussi de dimension finie. Soit $B_1 = (e_1, \ldots, e_p)$ une base de $\ker(f)$. D'après le théorème de la base incomplète, on peut compléter B_1 pour former une base $B = (e_1, \ldots, e_p, a_1, \ldots, a_q)$ de E. Montrons que la famille $B_2 = (f(a_1), \ldots, f(a_q))$ est une base de $\mathsf{Im}(f)$, ce qui entraînera que sa dimension est finie :

* Soient $\alpha_1, \ldots, \alpha_q \in \mathbb{K}$ tels que $\alpha_1 f(a_1) + \cdots + \alpha_q f(a_q) = 0_F$. Alors

$$f(\alpha_1 a_1 + \cdots + \alpha_q a_q) = 0_F$$

et donc
$$\alpha_1 a_1 + \cdots + \alpha_q a_q \in \ker(f).$$
Puisque $B_1 = (e_1, \ldots, e_p)$ est une base de $\ker(f)$, il existe des scalaires $\beta_1, \ldots, \beta_p \in K$ tels que
$$\alpha_1 a_1 + \cdots + \alpha_q a_q = \beta_1 e_1 + \cdots + \beta_p e_p,$$
soit
$$\beta_1 e_1 + \cdots + \beta_p e_p - \alpha_1 a_1 - \cdots - \alpha_q a_q = 0_F.$$
Comme B est libre, tous les scalaires $\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q$ sont nuls. Ainsi, B_2 est libre.

* Montrons maintenant que B_2 engendre $\mathsf{Im}(f)$. Pour tout $y \in \mathsf{Im}(f)$, il existe $x \in E$ tel que $y = f(x)$. Or, B est une base de E, donc on peut écrire le vecteur x comme combinaison linéaire
$$x = \beta_1 e_1 + \cdots + \beta_p e_p + \alpha_1 a_1 + \cdots + \alpha_q a_q.$$
Il vient
$$y = f(x) = \beta_1 f(e_1) + \cdots + \beta_p f(e_p) + \alpha_1 f(a_1) + \cdots + \alpha_q f(a_q).$$
Mais les vecteurs $e_1, \ldots, e_p$ appartiennent à $\ker(f)$, donc
$$f(e_1) = \cdots = f(e_p) = 0_F.$$
Par conséquent, $y = \beta_1 f(e_1) + \cdots + \beta_p f(e_p)$ est combinaison linéaire des éléments de B_2. D'où B_2 engendre $\mathsf{Im}(f)$, et par suite, c'est une base de $\mathsf{Im}(f)$.

Finalement, pour achever la preuve, il suffit d'écrire
$$\dim E = \mathsf{card}(B) = p + q = \mathsf{card}(B_1) + \mathsf{card}(B_2) = \dim\ker(f) + \dim \mathsf{Im}(f).$$
■

Remarque 2.5.2. Bien que $\dim E = \dim\ker(f) + \dim \mathtt{Im}(f)$, en général les sous-espaces $\ker(f)$ et $\mathtt{Im}(f)$ ne sont pas supplémentaires dans E.
En effet, d'abord $\mathtt{Im}(f)$ est un sous-espace de F, tandis que $\ker(f)$ est un sous-espace vectoriel de E.
Puis, même si $E = F$, les sous-espaces $\ker(f)$ et $\mathtt{Im}(f)$ peuvent ne pas être supplémentaires dans E, comme le montre l'exemple suivant dans lequel le théorème du noyau sera vérifié.

Exemple 2.5.3. Soit f l'application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^2$ définie par $f(x,y) = (x+y, -x-y)$. Pour tout vecteur $u = (x,y) \in \mathbb{R}^2$:
$u \in \ker(f) \iff x+y = 0$ et $2x+2y = 0 \iff x+y=0 \iff y=-x \iff u = (x,-x) \iff u \in \mathsf{Vect}(1,-1)$.
Ceci nous dit que $\ker(f) = \mathsf{Vect}\{(1,-1)\}$ est la droite vectorielle de $\mathbb{R}^2$ engendrée par le vecteur $(1,-1)$.
D'autre part, $\mathsf{Im}(f) = \{f(x,y) \ / \ (x,y) \in \mathbb{R}^2\} = \{(x+y,-x-y) \ / \ (x,y) \in \mathbb{R}^2\} = \{(z,-z) \ / \ z \in \mathbb{R}\} = \mathsf{Vect}\{(1,-1)\}$.
Par conséquent, les sous-espaces $\ker(f)$ et $\mathtt{Im}(f)$ sont égaux, et ne peuvent donc pas être supplémentaires.
On vérifie bien le théorème du noyau : $2 = 1+1$.

Définition 2.5.4. *Soient E et F deux espaces vectoriels de dimensions finies sur $\mathbb{K}$ et $f \in \mathcal{L}_{\mathbb{K}}(E,F)$. L'entier naturel* $\dim \mathsf{Im}(f)$ *s'appelle **rang** de f et se note* $\mathsf{rg}(f)$ *:*

$$\boxed{\mathsf{rg}(f) = \dim \mathsf{Im}(f) = \dim(E) - \dim(\ker(f))}$$

La formule précédente s'écrit donc

$$\boxed{\dim E = \dim \ker(f) + \mathsf{rg}(f)}$$

Voici quelques propriétés élémentaires du rang d'une application linéaire :

Proposition 2.5.5. *Soit $f : E \longrightarrow F$ une application linéaire. On a*

1. $\mathsf{rg}(f) \leq \dim(E)$ *et* $\mathsf{rg}(f) \leq \dim(F)$.
2. *Pour toute base $B = (a_1, \ldots, a_n)$ de E,* $\mathsf{rg}(f) = \mathsf{rg}(f(a_1), \ldots, f(a_n))$.

Démonstration.

1. Puisque $\mathsf{Im}(f)$ est un sous-espace de F, il est immédiat que $\mathsf{rg}(f) = \dim \mathsf{Im}(f)) \leq \dim(F)$.
   D'autre part, $\mathsf{rg}(f) = \dim(E) - \dim \ker(f) \leq \dim(E)$.
2. Si $B = (a_1, \ldots, a_n)$ est une base de E, alors $\mathsf{Im}(f) = \mathsf{Vect}(f(a_1), \ldots, f(a_n))$.
   D'où $\mathsf{rg}(f) = \dim(\mathsf{Vect}(f(a_1), \ldots, f(a_n))) = \mathsf{rg}((f(a_1), \ldots, f(a_n))$. ∎

A titre d'exemple, le rang de l'application linéaire de l'exemple précédent est 1.

Une autre façon de prouver l'inégalité $\mathsf{rg}(f) \leq \dim(E)$ consiste à utiliser la propriété 2 : $\mathsf{rg}(f) = \mathsf{rg}(f(a_1), \dots, f(a_n))$, où $(a_1, \dots, a_n)$ est une base de E. Comme $(f(a_1), \dots, f(a_n))$ est une famille de n vecteurs de F, alors $\mathsf{rg}(f(a_1), \dots, f(a_n)) \leq n$, puisque le rang d'une famille de vecteurs est toujours inférieur ou égal à son cardinal (voir proposition 1.4.24).

Remarque 2.5.6. Soit $f \in \mathcal{L}(E, F)$. Alors : $\mathsf{rg}(f) = 0 \iff f = 0$.
En effet, si $f = 0$, alors $\mathsf{Im}(f) = \{0_F\}$, d'où $\mathsf{rg}(f) = \dim(\{0_F\}) = 0$.
Réciproquement, si $\mathsf{rg}(f) = 0$, alors $\dim \mathsf{Im}(f) = 0$, ce qui implique $\mathsf{Im}(f) = \{0_F\}$. Donc pour tout $x \in E, f(x) = 0_F$, c'est-à-dire $f = 0$.

Nous nous intéressons maintenant à la question suivante : quand-est ce que les inégalités $\mathsf{rg}(f) \leq \dim(E)$ et $\mathsf{rg}(f) \leq \dim(F)$ deviennent des égalités ?
En connexion avec cette question, nous allons voir que la connaissance du rang d'une application linéaire f permet de savoir si f est injective, surjective, bijective.

Proposition 2.5.7. *Soient E et F deux espaces vectoriels de dimensions finies sur K, et $f \in \mathcal{L}(E, F)$. On a :*

(a) *f est injective* $\iff \mathsf{rg}(f) = \dim(E)$.

(b) *f est surjective* $\iff \mathsf{rg}(f) = \dim(F)$.

(c) *f est bijective* $\iff \mathsf{rg}(f) = \dim(E) = \dim(F)$

Démonstration. En utilisant le théorème du rang :

(a) f est injective $\iff \ker(f) = \{0_E\} \iff \dim(\ker(f)) = 0 \iff \mathsf{rg}(f) = \dim(E) - 0 = \dim(E)$.

(b) f est surjective $\iff \mathsf{Im}(f) = F$. Or, $\mathsf{Im}(f)$ est un sous-espace vectoriel de F, donc $\mathsf{Im}(f) = F \iff \dim \mathsf{Im}(f) = \dim(F) \iff \mathsf{rg}(f) = \dim(F)$.

(c) est une conséquence immédiate des deux premières propriétés. ■

Remarque 2.5.8. On peut démontrer la proposition précédente d'une manière différente :
f est injective $\iff (f(a_1), \dots, f(a_n))$ est libre dans $F \iff \mathsf{rg}(f(a_1), \dots, f(a_n)) = n \iff \mathsf{rg}(f) = n = \dim(E)$.

f est surjective $\iff$ $(f(a_1), \ldots, f(a_n))$ engendre F $\iff$ $\mathsf{rg}((f(a_1), \ldots, f(a_n)) = \dim(F)$.

La proposition précédente permet de caractériser les automorphismes :

Corollaire 2.5.9. *Soit E un espace vectoriel de dimension n sur $\mathbb{K}$, et $f \in \mathcal{L}(E)$ un endomorphisme de E. On a :*

$$\boxed{f \text{ est bijective} \iff \mathsf{rg}(f) = n}$$

Par définition, une application linéaire $f : E \longrightarrow F$ est bijective si elle est à la fois injective et surjective. Cependant, nous avons déjà vu dans le corollaire 2.3.11 que si E et F sont de dimensions finies avec $\dim(E) = \dim(F)$, chacune de ces deux conditions implique l'autre. Dans un but pédagogique, voici une autre preuve de ce résultat, comme conséquence de la proposition précédente :

Corollaire 2.5.10. *Soient E et F deux espaces vectoriels de dimensions finies sur K, avec $\dim(E) = \dim(F)$, et soit $f \in \mathcal{L}_K(E, F)$. Alors :*
f est injective $\Leftrightarrow$ f est surjective $\Leftrightarrow$ f est bijective.

Démonstration. f est injective $\iff$ $\mathsf{rg}(f) = \dim(E)$ $\iff$ $\mathsf{rg}(f) = \dim(F)$ $\iff$ f est surjective. ■

Le corollaire précédent est en particulier valable pour un endomorphisme $f \in \mathcal{L}_K(E)$ de E lorsque E est de dimension finie. Dans la pratique, c'est plutôt cette version qui est souvent utilisée. Au risque de nous faire répéter, énonçons clairement ce résultat pour insister sur son extrême importance :

Corollaire 2.5.11. *Soient E un espace vectoriel de dimension finie sur $\mathbb{K}$ et $f \in \mathcal{L}_{\mathbb{K}}(E)$ un endomorphisme de E. Alors :*
f est injective $\iff$ f est surjective $\iff$ f est bijective.

Pour terminer cette section, appliquons le théorème du noyau aux formes linéaires :

Remarque 2.5.12. Soit E un espace vectoriel de dimension n sur $\mathbb{K}$, et considérons une forme linéaire non nulle $f : E \longrightarrow \mathbb{K}$. Puisque $\mathsf{Im}(f)$ est un sous-espace vectoriel de $\mathbb{K}$, $\dim_K \mathsf{Im}(f) \leq \dim_{\mathbb{K}}(\mathbb{K}) = 1$. Or, $f \neq 0$, alors $\mathsf{Im}(f) \neq \{0_E\}$, d'où $\dim \mathsf{Im}(f) = 1$ nécessairement. On en déduit que

$\dim(\ker(f)) = n-1$. Rappelons qu'un sous-espace vectoriel de E de dimension $n-1$ est appelé un *hyperplan* de E. Donc le noyau d'une forme linéaire non nulle est toujours un hyperplan.

Exemple 2.5.13. Soient $a, b \in \mathbb{R}$ et considérons l'application $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x,y) = ax + by$ pour tout vecteur $(x,y) \in \mathbb{R}^2$. On vérifie facilement que c'est une forme linéaire sur $\mathbb{R}^2$. Elle est non nulle si $(a,b) \neq (0,0)$. Dans ce cas, son noyau, qui est l'ensemble $\ker(f) = \{(x,y) \in \mathbb{R}^2 \ / \ ax + by = 0\}$, est un hyperplan de $\mathbb{R}^2$. C'est donc un espace vectoriel de dimension 1.

2.6 Projecteurs et symétries

Comme leurs noms indiquent, les concepts de projections et de symétries proviennent de la géométrie. Commençons par introduire la notion suivante :

Définition 2.6.1. *Si $E = F_1 \oplus F_2$ est une décomposition d'un espace vectoriel E en deux sous-espaces supplémentaires F_1 et F_2, tout vecteur $x \in E$ possède une unique décomposition sous la forme $x = x_1 + x_2$, avec $x_1 \in F_1$ et $x_2 \in F_2$. On appelle* ***projection sur** F **parallèlement à** G *l'application $p_1 : E \longrightarrow E$, $x = x_1 + x_2 \mapsto p_1(x) = x_1$. On dit que le vecteur x_1 est la* ***projection de** x **sur** F_1 **parallèlement à** F_2.*
De même, on définit de manière symétrique la ***projection** p_2 **sur** F_2 **parallèlement à** F_1 *: $p_2 : E \longrightarrow E$, $x = x_1 + x_2 \mapsto p_2(x) = x_2$.*

Proposition 2.6.2. *Soient E un espace vectoriel sur K, et F_1, F_2 deux sous-espaces vectoriels de E tels que $E = F_1 \oplus F_2$. La projection p_1 sur F_1 parallèlement à F_2 est un endomorphisme de E vérifiant :*

$$\boxed{p_1 \circ p_1 = p_1, \qquad \ker(p_1) = F_2, \qquad Im(p_1) = F_1.}$$

Démonstration. • Soient $x, y \in E$ et $\alpha, \beta \in K$. Il existe $(x_1, x_2) \in F_1 \times F_2$ et $(y_1, y_2) \in F_1 \times F_2$ tels que $x = x_1 + x_2$ et $y = y_1 + y_2$. Alors $\alpha x + \beta y = \alpha(x_1 + x_2) + \beta(y_1 + y_2) = (\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2)$ avec $\alpha x_1 + \beta y_1 \in F_1$ et $\alpha x_2 + \beta y_2 \in F_2$. Donc $p_1(\alpha x + \beta y) = \alpha x_1 + \beta y_1 = \alpha p_1(x) + \beta p_1(y)$.
• Pour tout vecteur $x = x_1 + x_2 \in E$, on a $p_1 \circ p_1)(x) = p_1(p_1(x)) = p_1(x_1)$. Or, l'écriture de x_1 dans la somme directe $F_1 \oplus F_2 = E$ est évidemment $x_1 = x_1 + 0_E$, car $x_1 \in F_1$ et $0_E \in F_2$. D'où $p_1(x_1) = x_1 = p_1(x)$, et par suite, $p_1 \circ p_1 = p_1$.

• Pour tout $x = x_1 + x_2 \in E$, on a $p_1(x) = 0_E \Leftrightarrow x_1 = 0_E \Leftrightarrow x = x_2 \Rightarrow x \in F_2$, donc $\ker(p_1) \subseteq F_2$. D'autre part, tout vecteur $x \in F_2$ se décompose sous forme $x = 0_E + x$ avec $0_E \in F_1$ et $x \in F_2$, c'est-à-dire, $x = x_1 + x_2$ avec $x_1 = 0_E$ et $x_2 = x$. Donc $p_1(x) = x_2 = 0_E$, d'où $F_2 \subseteq \ker(p_1)$.
Ainsi, $\ker(p_1) = F_2$.
Montrons que $Im(p_1) = F_1$. D'abord, pour tout $x \in E$, $p_1(x) = x_1 \in F_1$, d'où $Im(p_1) \subseteq F_1$.
D'autre part, pour tout x de F_1, on a $x = x + 0_E$ avec $x \in F_1$ et $0_E \in G$, d'où $x = x_1 + x_2$ avec $x_1 = x$ et $x_2 = 0_E$. Ceci implique $p_1(x) = x$, et donc $x \in Im(p_1)$.
Ainsi, $Im(p_1) = F_1$. ■

Remarque 2.6.3. Avec les notations ci-dessus, on a

$$\mathsf{Im}(p_1) = \{x \in E \ / \ p_1(x) = x\}$$

En effet, si $x \in Im(p_1)$, il existe $y \in E$ tel que $x = p_1(y)$. Donc $p_1(x) = p_1(p_1(y)) = (p_1 \circ p_1)(y) = p_1(y) = x$, d'où $Im(p_1) \subseteq \{x \in E \ / \ p_1(x) = x\}$.
Si maintenant $x \in E$ avec $p_1(x) = x$, alors il est clair que $x \in Im(p_1)$, d'où $\{x \in E \ / \ p_1(x) = x\} \subseteq Im(p_1)$.

Remarque 2.6.4. 1) Avec les notations ci-dessus, on a évidemment la somme directe $E = \mathsf{Im}(p_1) \oplus Ker(p_1)$.
2) En échangeant les rôles des sous-espaces F_1 et F_2, il est clair que la projection sur F_2 parallèlement à F_1 vérifie aussi $p_2 \circ p_2 = p_2$, $\ker(p_2) = F_2$, $Im(p_2) = F_2 \oplus F_1$.

Proposition 2.6.5. *Soit $E = F_1 \oplus F_2$, p_1 (resp. p_2) la projection sur F_1 (resp. sur F_2) parallèlement à F_2 (resp. à F_1). On a*

$$id_E = p_1 + p_2, \quad p_1 \circ p_2 = p_2 \circ p_1 = 0.$$

DÉMONSTRATION. 1) Pour tout $x = x_1 + x_2 \in E$, $(p_1 + p_2)(x) = p_1(x) + p_2(x) = x_1 + x_2 = x$, donc $id_E = p_1 + p_2$.
2) $(p_1 \circ p_2)(x) = p_1(p_2(x)) = p_1(x_2) = 0_E$ car $p_2(x) \in F_2 = \ker(p_1)$, d'où $p_1 \circ p_2 = 0$. De façon analogue on vérifie que $p_2 \circ p_1 = 0$. ■

Nous venons de prouver que si $E = F_1 \oplus F_2$, alors la projection p_1 sur F_1 parallèlement à F_2 vérifie $p_1 \circ p_1 = p_1$. Nous étudions maintenant la réciproque de cette propriété.

Définition 2.6.6. *Soit E un espace vectoriel sur k et $f \in \mathcal{L}_K(E)$ un endomorphisme de E. On dit que f est un* ***projecteur*** *si $p \circ p = p$.*

Théorème 2.6.7. *Soit p un projecteur de E. On a :*

(a) $E = Im(p) \oplus \ker(p)$.

(b) p est la projection sur $Im(p)$ parallèlement à $\ker(p)$.

Démonstration. (a) • Soit $x \in \ker(p) \cap Im(p)$. Alors $p(x) = 0_E$, et il existe $y \in E$ tel que $x = p(y)$. Donc $0_E = p(x) = p(p(y)) = (p \circ p)(y) = p(y) = x$. Ceci montre que $\ker(p) \cap Im(p) = \{0_E\}$.
• On a déjà l'inclusion triviale $Im(p) + \ker(p) \subseteq E$. Montrons l'inclusion réciproque $E \subseteq Im(p) + \ker(p)$. Pour ce faire, pour tout x dans E, écrivons $x = p(x) + (x - p(x))$. Il est clair que $p(x) \in Im(p)$. De plus, $x - p(x) \in \ker(p)$, car $p(x - p(x)) = p(x) - (p \circ p)(x) = 0_E$. D'où $E = Im(p) + \ker(p)$.
(b) Puisque, pour tout x de E, $x = p(x) + (x - p(x))$ avec $p(x) \in Im(p)$ et $x - p(x) \in \ker(p)$, on déduit que p est la projection sur $Im(p)$ parallèlement à $\ker(p)$. ■

Proposition 2.6.8. *Soit p un projecteur de E, et $q = id_E - p$. On a :*

(a) q est un projecteur de E

(b) $p \circ q = q \circ p = 0$.

(c) q est la projection de E sur $\ker(p)$ *parallèlement à $Im(p)$.*

(d) $\ker(q) = Im(p)$, $Im(q) = \ker(p)$.

Démonstration. (a) L'application q est clairement un endomorphisme de E. Elle vérifié $q \circ q = (id_E - p) \circ (id_E - p) = id_E \circ id_E - id_E \circ p - p \circ id_E + p \circ p = id_E - 2p + p = id_E - p = q$, donc q est aussi un projecteur de E.
(b) $p \circ q = p \circ (id_E - p) = p \circ id_E - p \circ p = p - p = 0$.
(c) Soit x un vecteur quelconque de E ; on a vu que sa décomposition selon la somme directe $E = Im(p) \oplus \ker(p)$ est $x = p(x) + (x - p(x))$ avec $p(x) \in Im(p)$ et $x - p(x) \in \ker(p)$. Il en découle que q est la projection de E sur $\ker(p)$ parallèlement à $Im(p)$. ■

Si E est de dimension finie et p est un projecteur de E. Alors la somme directe $E = Im(p) \oplus \ker(p)$ nous permet de vérifier automatiquement le théorème du rang $\dim E = \dim Im(p) + \dim \ker(p)$.

Une remarque s'impose à ce niveau : En général, un endomorphisme $f \in \mathcal{L}_K(E)$ vérifiant $E = Im(f) \oplus \ker(f)$ n'est pas nécessairement un projecteur.

Définition 2.6.9. *Soit $E = F_1 \oplus F_2$ une décomposition d'un espace vectoriel E en deux sous-espaces supplémentaires F_1 et F_2. On appelle* ***symétrie par rapport à*** *F_1* ***parallèlement à*** *F_2 l'application $s_1 : E \longrightarrow E,\ x = x_1 + x_2 \mapsto s_1(x) = x_1 - x_2$. On dit que le vecteur x_1 est le symétrique de x par rapport à F_1 parallèlement à F_2.*
De même, on définit de manière symétrique la symétrie s_2 par rapport à F_2 parallèlement à F_1 : $s_2 : E \longrightarrow E,\ x = x_1 + x_2 \mapsto s_2(x) = x_2 - x_1$.

Remarque 2.6.10. Il est clair que $s_1 = p_1 - p_2$, car $s_1(x) = x_1 - x_2 = p_1(x) - p_2(x_2) = (p_1 - p_2)(x)$ pour tout vecteur $x = x_1 + x_2 \in E$.
De plus, puisque $\mathtt{id}_E = p_1 + p_2$, alors on obtient

$$p_1 = \frac{1}{2}(s_1 + \mathtt{id}_E), \quad s_1 = 2p_1 - \mathtt{id}_E$$

Proposition 2.6.11. *Soit $E = F_1 \oplus F_2$. La symétrie s_1 par rapport à F_1 parallèlement à F_2 est un automorphisme de E vérifiant :*

$$\boxed{s_1 \circ s_1 = id_E}$$

Démonstration.

- Si $s_1(x) = 0_E$, alors $x_1 - x_2 = 0_E$ et donc $x_1 = x_2 \in F_1 \cap F_2$. D'où $x_1 = x_2 = 0_E$, et par suite, $x = 0_E$. Ainsi, $\ker(s_1) = \{0_E\}$, donc s_1 est injective. D'autre part, il est clair de voir que $\mathsf{Im}(f) = E$, d'où s_1 est surjective.
- $(s_1 \circ s_1)(x) = s_1(x_1 - x_2) = s_1(x_1 + (-x_2)) = x_1 - (-x_2) = x_1 + x_2 = x$ pour tout $x = x_1 + x_2 \in E$. Il en résulte que $s_1 \circ s_1 = \mathtt{id}_E$.

On aurait pu montré la surjectivité de s_1 à l'aide de la formule $s_1 \circ s_1 = id_E$. En effet, ceci veut dire que s_1 est égale à son propre inverse s_1^{-1}.

Et voilà, nous arrivons à la fin de ce présent chapitre. Maintenant vous avez bien assimilé les bases pour pouvoir manipuler les espaces vectoriels et les applications linéaires. Préparez-vous maintenant au prochain chapitre sur les matrices. Vous êtes prêts !

3 Matrices

Sommaire

"A ne rien faire, on apprend bien peu. Le prix du succès est beaucoup moins élevé que celui de l'échec."

.

En mathématiques, les matrices servent à interpréter en termes calculatoires les résultats théoriques de l'algèbre linéaire que nous avons étudiées dans les espaces vectoriels et les applications linéaires. Toutes les disciplines étudiant des phénomènes linéaires utilisent les matrices. En particulier, elles sont présentes en Analyse, Géométrie, Physique, etc.

3.1 L'espace vectoriel $\mathcal{M}_{n,p}(\mathbb{K})$ des matrices de type (n, p)

Soit $\mathbb{K}$ un corps commutatif qui se sera en pratique $\mathbb{R}$ ou $\mathbb{C}$. Soient deux entiers naturels non nuls n et p.

Définition 3.1.1. *On appelle* `matrice de type` (n, p) *à cœfficients dans* $\mathbb{K}$ *un tableau rectangulaire à* n *lignes et* p *colonnes de la forme*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix}$$

où les $a_{ij} \in \mathbb{K}$ *s'appellent les* ***cœfficients*** *de la matrice* A.
Le symbole $\mathcal{M}_{n,p}(\mathbb{K})$ *désigne l'ensemble des matrices de type* (n, p) *à cœfficients dans* $\mathbb{K}$.

Chaque cœfficient a_{ij} est situé à l'intersection de la ligne numéro i avec la colonne numéro j. Puisqu'il y a au total $n \times p$ cœfficients a_{ij} dans la matrice A, on dit aussi que A est une matrice *de taille* $n \times p$. La matrice A pourra être écrite de façon abrégée :

$$A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$$

ou simplement

$$A = (a_{ij})$$

s'il n'y a pas de confusion sur sa taille.
Ainsi, pour une matrice $A = (a_{ij})$, le premier indice i correspond au numéro de la ligne, le second indice j est le numéro de la colonne.

Définition 3.1.2. *Deux matrices $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ et $B = (b_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ de même type (n, p) sont dites* ***égales*** *si $a_{ij} = b_{ij}$ quels que soient les indices $i \in \{1, \ldots, n\}$ et $j \in \{1, \ldots, p\}$.*

Exemple 3.1.3. Considérons les cas particuliers suivants :

(a) $n = 3$, $p = 2$: Une matrice A de type $(3, 2)$ à cœfficients dans $\mathbb{K}$ est de la forme

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

(b) $p = 1$: Une matrice de type $(n, 1)$ s'écrit

$$A = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}$$

On l'appelle **matrice colonne** (ou **uni-colonne**).

(c) $n = 1$. Une matrice de type $(1, p)$ a la forme $A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1p} \end{pmatrix}$ et s'appelle **matrice ligne** (ou **uni-ligne**).

(d) Une matrice A de type $(1, 1)$ contient un seul cœficient a_{11}. On peut donc considérer la matrice A comme un scalaire de $\mathbb{K}$. Ceci permet d'identifier l'ensemble $\mathcal{M}_{1,1}(\mathbb{K})$ avec $\mathbb{K}$.

(e) $n = p$. Une matrice A de type (n, n) s'écrit

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

Dans ce cas, on dit que A est une **matrice carrée d'ordre** n. L'ensemble $\mathcal{M}_{n,n}(\mathbb{K})$ de ces matrices carrées d'ordre n est désigné simplement par $\mathcal{M}_n(\mathbb{K})$. Dans ce cas, les coefficients a_{ii} avec $i = 1, \ldots, n$ sont dits **cœfficients diagonaux**, et la ligne oblique contenant les cœfficients diagonaux $a_{11}, \ldots, a_{nn}$ est appelée la **diagonale** de la matrice A.

(f) La matrice carrée H_n définie par le cœfficient général $a_{ij} = 1/(i+j-1)$ s'appelle la *matrice de Hilbert*, en hommage au mathématicien allemand David Hilbert qui l'étudia en 1894. Par exemple, pour $n = 4$:

$$H_4 = \begin{pmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1/3 & 1/4 & 1/5 \\ 1/3 & 1/4 & 1/5 & 1/6 \\ 1/4 & 1/5 & 1/6 & 1/7 \end{pmatrix}$$

Remarque 3.1.4. Soit une matrice $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$. Le cœfficient général de A est a_{ij}, il se trouve en position (i,j). La matrice A est donc une suite finie de np termes a_{ij} $(i = 1, \ldots, n,\ \ j = 1, \ldots, p)$. Rappelons que pour les suites numériques $u = (u_n)_{n\in\mathbb{N}}$, l'entier n est un *indice muet*, c'est-à-dire que sa notation n'a aucune importance : la suite $u = (u_n)_{n\in\mathbb{N}}$ peut s'écrire aussi $u = (u_m)_{m\in\mathbb{N}} = (u_i)_{i\in\mathbb{N}}$. De même pour une matrice $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$, les indices i et j sont muets : il est possible de changer les notations des indices $i \in \{1, \ldots, n\}$ et $j \in \{1, \ldots, p\}$ pour écrire par exemple $A = (a_{kl})$. L'essentiel est que le premier indice k soit compris entre 1 et n et que le second indice l doive varier entre 1 et p. Nous aurons besoin de cette observation quand nous aborderons la multiplication des matrices.

Remarque 3.1.5. Pour effectuer certaines opérations, il peut être utile de travailler sur le système des lignes ou des colonnes d'une matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$. On pourra alors l'écrire sous une des formes suivantes :

$* \ A = \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix}$ où $L_i = \begin{pmatrix} a_{i1} & a_{i2} & \ldots & a_{ip} \end{pmatrix}$ est la ligne i de la matrice A.

$* \ A = \begin{pmatrix} C_1 & C_2 & \ldots & C_p \end{pmatrix}$ où $C_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$ est la colonne j de A.

Nous allons maintenant considérer la question d'arithmétique des matrices. Plus précisément, il est raisonnable de munir l'ensemble $\mathcal{M}_{n,p}(K)$ de deux opérations naturelles, une *addition* et une *multiplication externe*, de la façon suivante : pour ajouter deux matrices A et B de même type, il suffit d'ajouter les cœfficients correspondants ; et pour multiplier une matrice A

par un scalaire λ, on multiplie tous les cœfficients de A par λ. En d'autres termes :

$$\begin{pmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1p} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{np} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & \dots & a_{1p}+b_{1p} \\ \vdots & & \vdots \\ a_{n1}+b_{n1} & \dots & a_{np}+b_{np} \end{pmatrix},$$

$$\lambda \cdot \begin{pmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \dots & \lambda a_{1p} \\ \vdots & & \vdots \\ \lambda a_{n1} & \dots & \lambda a_{np} \end{pmatrix}$$

pour toutes matrices $A = (a_{ij})$ et $B = (b_{ij})$ dans $\mathcal{M}_{n,p}(\mathbb{K})$ et pour tout scalaire λ dans $\mathbb{K}$. Ces opérations s'écrivent en notation simplifiée :

$$\boxed{(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij}), \qquad \lambda \cdot (a_{ij}) = (\lambda a_{ij})}$$

Autrement dit, si $A = (a_{ij})$ et $B = (b_{ij})$ sont dans $\mathcal{M}_{n,p}(\mathbb{K})$, alors $A + B = (c_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$ avec $c_{ij} = a_{ij} + b_{ij}$. De même, si $\lambda \in \mathbb{K}$, alors $\lambda \cdot A = (d_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$ avec $d_{ij} = \lambda a_{ij}$.
Les opérations d'addition et de multiplication des matrices satisfont les lois usuelles d'arithmétique qu'on peut regrouper dans la proposition qui suit.

Proposition 3.1.6. *L'ensemble $\mathcal{M}_{n,p}(\mathbb{K})$ est un espace vectoriel sur $\mathbb{K}$.*

Démonstration. La vérification de cette proposition est très facile :

∗ L'élément neutre de l'addition est la matrice dont tous les cœfficients a_{ij} sont nuls. Elle s'appelle **matrice nulle de type** (n, p) et se note $\mathbf{O}_{n,p}$ ou $\mathbf{O}$.

∗ L'élément symétrique de la matrice $A = (a_{ij})$ est la matrice $-A = (-a_{ij})$.

∗ La commutativité de l'addition des matrices $A + B = B + A$ s'obtient immédiatement de la commutativité de l'addition des scalaires dans $\mathbb{K}$:

$$(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij}) = (b_{ij} + a_{ij}) = (b_{ij}) + (a_{ij})$$

∗ De façon analogue, l'associativité des matrices $(A+B)+C = A+(B+C)$ provient de l'associativité de l'addition des scalaires dans $\mathbb{K}$.

Ainsi, $(\mathcal{M}_{n,p}(\mathbb{K}), +)$ est un groupe abélien.
Les quatre axiomes de la loi externe s'établissent aussi de manière aisée :

∗ $\alpha \cdot (A + B) = \alpha + \alpha \cdot B$;

∗ $(\alpha + \beta) = \alpha + \beta \cdot B$;

∗ $\alpha \cdot (\beta \cdot A) = (\alpha\beta) \cdot A$;

$*$ $1 \cdot A = A$.

pour toutes matrices $A = (a_{ij})$ et $B = (b_{ij})$ dans $\mathcal{M}_{n,p}(\mathbb{K})$ et pour tous scalaires α et β dans $\mathbb{K}$.
Par exemple, soient $A = (a_{ij})_{i,j} \in \mathcal{M}_{n,p}(\mathbb{K})$, $B = (b_{ij})_{i,j} \in \mathcal{M}_{n,p}(\mathbb{K})$ et $\alpha \in \mathbb{K}$ et vérifions la première assertion :

$$\begin{aligned}
\alpha(A+B) &= \alpha\left((a_{ij}) + (b_{ij})\right) \\
&= \alpha(a_{ij} + b_{ij}) \\
&= (\alpha(a_{ij} + b_{ij})) \\
&= (\alpha a_{ij} + \alpha b_{ij}) \\
&= (\alpha a_{ij}) + (\alpha b_{ij}) \\
&= \alpha\,(a_{ij}) + \alpha\,(b_{ij}) \\
&= \alpha A + \lambda B
\end{aligned}$$

Tous les calculs se font donc facilement de cette manière en se ramenant aux calculs sur le corps $\mathbb{K}$ via les cœfficients des matrices. ■

Après avoir prouvé que $\mathcal{M}_{n,p}(\mathbb{K})$ est un espace vectoriel sur $\mathbb{K}$, il est temps d'étudier sa dimension. Puisqu'une matrice $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$ possède np cœfficients a_{ij} indépendants les uns des autres, il est naturel de deviner que sa dimension serait np. Vérifions notre intuition dans le cas particulier $n = 2$, $p = 3$ pour mieux comprendre la situation. Cherchons donc une base de $\mathcal{M}_{2,3}(\mathbb{K})$:
Soit donc $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$ une matrice de type $(2,3)$. Décomposons A sous forme :

$$\begin{aligned}
A \quad = \quad & a_{11}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + a_{12}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + a_{13}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \\
+ \quad & a_{21}\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + a_{22}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} + a_{23}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}$$

Par conséquent, les six matrices

$$E_{11} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ E_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ E_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

$$E_{21} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},\ E_{22} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},\ E_{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

engendrent l'espace vectoriel $\mathcal{M}_{2,3}(\mathbb{K})$.
Toute matrice $A = (a_{ij})$ s'écrit donc comme combinaison linéaire

$$A = a_{11}E_{11} + a_{12}E_{12} + a_{13}E_{13} + a_{21}E_{21} + a_{22}E_{22} + a_{23}E_{23}.$$

De plus, ces six matrices $E_{11}, E_{12}, E_{13}, E_{21}, E_{22}, E_{23}$ sont linéairement indépendantes. En effet, prenons une combinaison linéaire nulle

$$a_{11}E_{11} + a_{12}E_{12} + a_{13}E_{13} + a_{21}E_{21} + a_{22}E_{22} + a_{23}E_{23} = \mathbf{O}$$

avec les a_{ij} dans $\mathbb{K}$. Alors les calculs précédents nous donnent en remontant :

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

et par suite, tous les scalaires a_{ij} sont nuls.
Ainsi, les six matrices $E_{11}, E_{12}, E_{13}, E_{21}, E_{22}, E_{23}$ forment une base de $\mathcal{M}_{2,3}(\mathbb{K})$, appelée la **base canonique** de $\mathcal{M}_{2,3}(\mathbb{K})$. Il en résulte que $\mathcal{M}_{2,3}(\mathbb{K})$ est de dimension $6 = 2 \times 3$ sur $\mathbb{K}$.
La matrice E_{ij} est celle dont tous les coefficients sont nuls sauf celui en position (i, j), qui vaut 1.

Nous pouvons étendre cette situation au cas général :

Théorème 3.1.7. *L'espace vectoriel $\mathcal{M}_{n,p}(\mathbb{K})$ est de dimension np sur $\mathbb{K}$. Une base de $\mathcal{M}_{n,p}(\mathbb{K})$ est formée par les matrices E_{ij} avec $i = 1, \dots, n$ et $j = 1, \dots, p$ telle que E_{ij} est la matrice dont tous les cœfficients sont nuls sauf celui de la ligne i et la colonne j, qui vaut 1. Cette base $(E_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ s'appelle la **base canonique** de $\mathcal{M}_{n,p}(\mathbb{K})$.*

Démonstration. Il suffit d'adapter la preuve correspondante à $n = 2, p = 3$. Considérons donc la famille $(E_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ des matrices $E_{ij} \in \mathcal{M}_{n,p}(\mathbb{K})$ telle que E_{ij} est la matrice dont tous les cœfficients sont nuls sauf celui en position (i, j), qui vaut 1. Pour toute matrice $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$, on décompose A sous la forme :

$$A = (a_{ij}) = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{ij}E_{ij} = \sum_{\substack{1\leq i\leq n\\ 1\leq j\leq p}} a_{ij}E_{ij}. \tag{3.1}$$

Ainsi, les np matrices $(E_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ constituent une famille génératrice de l'espace vectoriel $\mathcal{M}_{n,p}(\mathbb{K})$.
En outre, cette famille est libre. En effet, si $(a_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ sont np scalaires dans $\mathbb{K}$ vérifiant $\sum\limits_{\substack{1\leq i\leq n\\ 1\leq j\leq p}} a_{ij}E_{ij} = \mathbf{O}$, alors l'égalité (3.1) montre que la matrice (a_{ij}) est nulle, et donc tous les scalaires a_{ij} sont nuls.
Il s'ensuit que la famille $(E_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ est une base de $\mathcal{M}_{n,p}(\mathbb{K})$. Puisque cette famille contient exactement np matrices E_{ij}, alors la dimension de $\mathcal{M}_{n,p}(\mathbb{K})$ vaut np. ∎

Remarque 3.1.8. La matrice élémentaire $E_{ij} \in \mathcal{M}_{n,p}(\mathbb{K})$ peut s'écrire : $E_{ij} = (a_{kl})$ avec

$$a_{kl} = \begin{cases} 1 & \text{si } (k,l) = (i,j) \\ 0 & \text{sinon} \end{cases}$$

3.2 Matrice d'une application linéaire

Dans cette section nous allons voir qu'à toute application linéaire, on peut associer une matrice.
Soient E et F deux espaces vectoriels finies sur le même corps $\mathbb{K}$ et $f : E \longrightarrow F$ une application linéaire. En vue de l'étude des matrices, il est préférable de noter p la dimension de E et n celle de F. Soient donc $\mathcal{B} = (e_1, \ldots, e_p)$ une base de E et $\mathcal{B}' = (e'_1, \ldots, e'_n)$ une base de F.
Nous avons déjà vu au chapitre précédent que l'application linéaire f de E dans F est parfaitement déterminée par son action sur la base B, c'est-à-dire par la donnée des p vecteurs $f(e_1), \ldots, f(e_p)$, donc par la donnée des coordonnées de ces p vecteurs dans la base B' de F.
Posons donc

$$\left\{\begin{array}{lcl} f(e_1) & = & a_{11}e'_1 + a_{21}e'_2 + \cdots + a_{n1}e'_n = \sum_{i=1}^{n} a_{i1}e'_i \\ f(e_2) & = & a_{12}e'_1 + a_{22}e'_2 + \cdots + a_{n2}e'_n = \sum_{i=1}^{n} a_{i2}e'_i \\ \vdots & \vdots & \qquad\qquad\qquad \vdots \\ f(e_p) & = & a_{1p}e'_1 + a_{2p}e'_2 + \cdots + a_{np}e'_n = \sum_{i=1}^{n} a_{ip}e'_i \end{array}\right. \tag{3.2}$$

avec les a_{ij} dans $\mathbb{K}$. C'est-à-dire,

$$f(e_j) = a_{1j}e'_1 + a_{2j}e'_2 + \cdots + a_{nj}e'_n = \sum_{i=1}^{n} a_{ij}e'_i \qquad \text{pour tout } j = 1, \ldots, p.$$

L'application linéaire f est donc complètement déterminée si on connaît les np constantes a_{ij} $(1 \leq i \leq n,\ 1 \leq j \leq p)$.
Dans la pratique, ces np coordonnées a_{ij} seront présentées sous la forme d'une matrice à n lignes et à p colonnes. La première colonne étant formée des coordonnées $\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}$ du vecteur $f(e_1)$ dans la base B' que nous écrivons verticalement, la seconde colonne contient les coordonnées $\begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}$ du vecteur $f(e_2)$, et ainsi de suite jusqu'à la $p^{\text{ème}}$ colonne constituée par les

coordonnées $\begin{pmatrix} a_{1p} \\ a_{2p} \\ \vdots \\ a_{np} \end{pmatrix}$ du vecteur $f(e_p)$. On forme ainsi la matrice

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix}$$

qui se note $M(f, \mathcal{B}, \mathcal{B}')$ et qui contient les informations à propos de l'application linéaire f. Cette discussion nous conduit à la définition suivante :

Définition 3.2.1. *Avec les hypothèses ci-dessus, on appelle* ***matrice de l'application linéaire*** $f : E \longrightarrow F$ ***dans les bases*** $B = (e_1, \dots, e_p)$ *et* $B' = (e'_1, \dots, e'_n)$, *la matrice* $M(f, \mathcal{B}, \mathcal{B}')$ *de type* (n, p) *dont la* $j^{\text{ème}}$ *colonne est constituée des coordonnées du vecteur* $f(e_j)$ *dans la base* B' *:*

$$M(f, \mathcal{B}, \mathcal{B}') = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix}$$

Exemple 3.2.2. Soit $f : \mathbb{R}^2 \to \mathbb{R}^3$ l'application définie par

$$f(x, y) = (2x + 5y, -4x + 3y, -x + y).$$

Il est aisé de voir que f est linéaire. Soient $\mathcal{B} = (e_1, e_2)$ et $\mathcal{B}' = (e'_1, e'_2, e'_3)$ les bases canoniques respectives des espaces vectoriels $\mathbb{R}^2$ et $\mathbb{R}^3$. On a

$$f(e_1) = f(1, 0) = (2, -4, -1), \quad f(e_2) = f(0, 1) = (5, 3, 1).$$

Les coordonnées de $f(e_1)$ et $f(e_2)$ vont s'écrire verticalement pour former la matrice $M(f, \mathcal{B}, \mathcal{B}')$ de f relativement aux bases B et B' :

$$M(f, \mathcal{B}, \mathcal{B}') = \begin{pmatrix} 2 & 5 \\ -4 & 3 \\ -1 & 1 \end{pmatrix}$$

qui est de type $(3, 2)$.

Exemple 3.2.3. Considérons l'application linéaire dérivée $D : \mathbb{K}_3[X] \longrightarrow \mathbb{K}_2[X]$ donnée par $D(P) = P'$. Rapportons les espaces vectoriels $\mathbb{K}_3[X]$ et $\mathbb{K}_2[X]$ à leurs bases canoniques respectives $\mathcal{B} = (1, X, X^2, X^3)$ et $\mathcal{B}' = (1, X, X^2)$. On obtient facilement

$$M(D, \mathcal{B}, \mathcal{B}') = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

Attention ! Dans la matrice $M(f, \mathcal{B}, \mathcal{B}')$, la base $\mathcal{B}$ est la base de l'espace de départ E, et $\mathcal{B}'$ est la base de l'espace d'arrivée F. De plus, cette matrice $M(f, \mathcal{B}, \mathcal{B}')$ se remplit et se lit par colonnes. Parfois, en corrigeant les copies des épreuves, les enseignants sont surpris de voir cette matrice $M(f, \mathcal{B}, \mathcal{B}')$ remplie horizontalement, ce qui est faux !

Remarque 3.2.4. Soit $f : E \longrightarrow F$ une application linéaire et supposons connue sa matrice $M(f, \mathcal{B}, \mathcal{B}') = (a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le p}}$ dans les bases $\mathcal{B} = (e_1, \ldots, e_p)$ et $\mathcal{B}' = (e'_1, \ldots, e'_n)$ respectives de E et F. En lisant verticalement les colonnes, on tire

$$f(e_j) = \sum_{i=1}^{n} a_{ij} e_i \quad \text{pour tout } j = 1, \ldots, p. \tag{3.3}$$

Pour calculer maintenant l'image $f(x)$ d'un vecteur quelconque $x \in E$, il suffit de décomposer x dans la base $\mathcal{B}$: $x = x_1 e_1 + x_2 e_2 + \cdots + x_p e_p$, puis d'effectuer un calcul linéaire :

$$\begin{aligned}
f(x) &= x_1 f(e_1) + x_2 f(e_2) + \cdots + x_p f(e_p) \\
&= x_1 \Big(a_{11} e'_1 + a_{21} e'_2 + \cdots + a_{n1} e'_n \Big) + x_2 \Big(a_{12} e'_1 + a_{22} e'_2 + \cdots + a_{n2} e'_n \Big) \\
&\quad + \cdots + x_p \Big(a_{1p} e'_1 + a_{2p} e'_2 + \cdots + a_{np} e'_n \Big) \\
&= \Big(x_1 a_{11} + x_2 a_{12} + \cdots + x_p a_{1p} \Big) e'_1 + \Big(x_1 a_{21} + x_2 a_{22} + \cdots + x_p a_{2p} \Big) e'_2 \\
&\quad + \cdots + \Big(x_1 a_{n1} + x_2 a_{n2} + \cdots + x_p a_{np} \Big) e'_n
\end{aligned}$$

Définition 3.2.5. *Lorsque f est un endomorphisme d'un espace vectoriel E de dimension n, on travaille en général avec la même base $\mathcal{B}$ au départ et à l'arrivée (bien que ce n'est pas toujours le cas). On note alors*

$$M(f, \mathcal{B}) = M(f, \mathcal{B}, \mathcal{B}) \in \mathcal{M}_n(\mathbb{K})$$

*et on parle dans ce cas de la **matrice de l'endomorphisme f dans la base** $\mathcal{B}$, qui est une matrice carrée d'ordre n.*

Exemple 3.2.6. Soit E un espace vectoriel de dimension 3 et $f \in \mathcal{L}(E)$ l'endomorphisme de E dont la matrice par rapport à la base $\mathcal{B} = (e_1, e_2, e_3)$ est

$$M(f, \mathcal{B}) = \begin{pmatrix} 1 & -1 & 4 \\ -2 & 0 & 3 \\ 2 & 1 & 0 \end{pmatrix}$$

Soit $x = x_1 e_1 + x_2 e_2 + x_3 e_3$ un vecteur quelconque de E et calculons son image $f(x)$:

$$\begin{aligned}
f(x) &= x_1 f(e_1) + x_2 f(e_2) + x_3 f(e_3) \\
&= x_1(e_1 - 2e_2 + 2e_3) + x_2(-e_1 + e_3) + x_3(4e_1 + 3e_2) \\
&= (x_1 - x_2 + 4x_3) e_1 + (-2x_1 + 3x_3) e_2 + (2x_1 + x_2) e_3
\end{aligned}$$

Ainsi, par exemple, $f(5e_1 + e_2 - 2e_3) = -4e_1 - 16e_2 + 11e_3$.

Exemple 3.2.7. Soit E un espace vectoriel de dimension n muni d'une base $\mathcal{B} = (e_1, \ldots, e_n)$.

(a) Soit $\lambda \in \mathbb{K}$. Rappelons que l'homothétie de rapport λ est l'endomorphisme $h_\lambda : E \longrightarrow E$ donné par $h_\lambda(x) = \lambda x$ pour tout $x \in E$. Puisque $h_\lambda(e_i) = \lambda e_i$ pour tout $i = 1, \ldots, n$, alors la matrice de f dans la base $\mathcal{B}$ est très facile à écrire :

$$M(h_\lambda, \mathcal{B}) = \begin{pmatrix} \lambda & 0 & \ldots & 0 \\ 0 & \lambda & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda \end{pmatrix}$$

Une telle matrice est dite **matrice scalaire**.

(b) En particulier, pour $\lambda = 1$, il s'agit de l'endomorphisme identité $h_1 = \mathtt{id}_E$ qui a pour matrice

$$M(\mathtt{id}_E, \mathcal{B}) = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix},$$

notée I_n et appelée **matrice unité d'ordre** n.
On remarque que toute matrice scalaire est de la forme λI_n avec $\lambda \in \mathbb{K}$.

(c) Notons que si nous travaillons avec deux bases différentes $\mathcal{B}$ et $\mathcal{B}'$ de E, alors $M(\mathtt{id}_E, \mathcal{B}, \mathcal{B}') \neq I_n$.
En effet, prenons $E = \mathbb{R}^2$ et soient $\mathcal{B} = ((2,1), (-1,0))$ et $\mathcal{B}' = (e_1, e_2)$ la base canonique de $\mathbb{R}^2$. On a

$$\mathtt{id}_E(2,1) = (2,1) = 2e_1 + e_2 \quad \text{et} \quad \mathtt{id}_E(-1,0) = (-1,0) = e_1 + 0e_2.$$

D'où

$$M(\mathtt{id}_E, \mathcal{B}, \mathcal{B}') = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} \neq I_2.$$

Il est important d'insister sur le fait que la matrice qui représente une application linéaire dépend des bases choisies ! Si nous changeons les bases des espaces considérées, nous obtenons des matrices différentes. Il est aussi important de souligner que la matrice d'une application linéaire dépend aussi de l'ordre dans lequel les éléments de la base sont écrits. Plus tard, nous nous intéresserons aux relations entre les matrices d'une application linéaire obtenues à partir de différents choix de bases.

Nous avons déjà considéré la somme $f + g$ et le produit externe αf d'applications linéaires $f, g : E \longrightarrow F$. Une question naturelle est de comparer les matrices de $f + g$ et αf avec celles de f et g :

Proposition 3.2.8. *Soient E et F deux espaces vectoriels de dimensions finies sur $\mathbb{K}$, de bases respectives $\mathcal{B}$ et $\mathcal{B}'$. Pour toutes applications linéaires $f, g : E \longrightarrow F$ et pour tout scalaire $\alpha \in \mathbb{K}$, on a :*

$$\boxed{M(f+g,\mathcal{B},\mathcal{B}') = M(f,\mathcal{B},\mathcal{B}') + M(g,\mathcal{B},\mathcal{B}') \qquad M(\alpha f,\mathcal{B},\mathcal{B}') = \alpha M(f,\mathcal{B},\mathcal{B}')}$$

Démonstration. Posons $\mathcal{B} = (e_1, \dots, e_p)$, $\mathcal{B}' = (e'_1, \dots, e'_n)$ et notons

$$M(f,\mathcal{B},\mathcal{B}') = (a_{ij}) \qquad \text{et} \qquad M(g,\mathcal{B},\mathcal{B}') = (b_{ij}).$$

D'après la formule (3.3), on a

$$f(e_j) = \sum_{i=1}^{n} a_{ij} e_i \qquad \text{et} \qquad g(e_j) = \sum_{i=1}^{n} b_{ij} e_i \quad \text{pour tout } j = 1, \dots, p.$$

Par conséquent,

$$(f+g)(e_j) = f(e_j) + g(e_j) = \sum_{i=1}^{n} a_{ij} e_i + \sum_{i=1}^{n} b_{ij} e_i = \sum_{i=1}^{n} (a_{ij} + b_{ij}) e_i.$$

Ceci donne

$$M(f+g,\mathcal{B},\mathcal{B}') = (a_{ij} + b_{ij}) = (a_{ij}) + (b_{ij}) = M(f,\mathcal{B},\mathcal{B}') + M(g,\mathcal{B},\mathcal{B}').$$

D'autre part,

$$(\alpha f)(e_j) = \alpha f(e_j) = \alpha \left(\sum_{i=1}^{n} a_{ij} e_i \right) = \sum_{i=1}^{n} (\alpha a_{ij}) e_i.$$

D'où

$$M(\alpha f,\mathcal{B},\mathcal{B}') = (\alpha a_{ij}) = \alpha (a_{ij}) = \alpha M(f,\mathcal{B},\mathcal{B}'). \blacksquare$$

Comme conséquence, nous pouvons énoncer :

Théorème 3.2.9. *Soient E et F deux espaces vectoriels de dimensions finies sur $\mathbb{K}$ et de bases respectives $\mathcal{B}$ et $\mathcal{B}'$. On note $\dim E = p$ et $\dim F = n$. L'application*

$$\begin{array}{rccl} \varphi : & \mathcal{L}_{\mathbb{K}}(E,F) & \longrightarrow & \mathcal{M}_{n,p}(\mathbb{K}) \\ & f & \mapsto & M(f,\mathcal{B},\mathcal{B}') \end{array}$$

est un isomorphisme d'espaces vectoriels.

Démonstration. Nous devons prouver les points suivants :

* L'application $\varphi : \mathcal{L}_{\mathbb{K}}(E,F) \longrightarrow \mathcal{M}_{n,p}(\mathbb{K})$ est un homomorphisme d'espaces vectoriels, c'est-à-dire, une application linéaire. En effet, $\forall f, g \in \mathcal{L}_{\mathbb{K}}(E,F)$ et $\forall \alpha \in \mathbb{K}$, la proposition précédente entraîne

$$\varphi(f+g) = M(f+g,\mathcal{B},\mathcal{B}') = M(f,\mathcal{B},\mathcal{B}') + M(g,\mathcal{B},\mathcal{B}') = \varphi(f) + \varphi(g),$$

$$\varphi(\alpha f) = M(\alpha f,\mathcal{B},\mathcal{B}') = \alpha M(f,\mathcal{B},\mathcal{B}') = \alpha \varphi(f).$$

* L'application φ est injective. Pour ce faire, soient f et g dans $\mathcal{L}_{\mathbb{K}}(E,F)$ telles que $\varphi(f) = \varphi(g)$. Alors $M(f,B,B') = M(g,B,B')$. En posant $M(f,B,B') = (a_{ij})$ et $M(g,B,B') = (b_{ij})$, on obtient $a_{ij} = b_{ij}$ pour tous $i = 1,\dots,n$ et $j = 1,\dots,p$. D'autre part, selon la relation (3.3), on a $f(e_j) = \sum\limits_{i=1}^{n} a_{ij}e_i$ et $g(e_j) = \sum\limits_{i=1}^{n} b_{ij}e_i$. On en déduit que $f(e_j) = g(e_j)$ pour tout $j = 1,\dots,p$. D'après le lemme 2.3.2, il vient $f = g$.

* L'application φ est surjective. En effet, soit $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix}$ une matrice quelconque dans $\mathcal{M}_{n,p}(\mathbb{K})$ et cherchons une application linéaire $f \in \mathcal{L}_{\mathbb{K}}(E,F)$ vérifiant $\varphi(f) = A$, c'est-à-dire, $M(f,\mathcal{B},\mathcal{B}') = A$. Il suffit de choisir f de sorte que les coordonnées de chaque vecteur $f(e_j)$ dans la base B' soient situées sur la $j^{\text{ème}}$ colonne de A. En d'autres termes, soit $f : E \longrightarrow F$ l'unique application linéaire telle que

$$f(e_j) = \sum_{i=1}^{n} a_{ij}e_i \qquad \text{pour tout } j = 1,\dots,p.$$

Une telle application linéaire f existe grâce au théorème 2.3.1. Il est clair que cette application linéaire f vérifie $M(f,\mathcal{B},\mathcal{B}') = (a_{ij}) = A$.
Ainsi, φ est bijective, et c'est donc un isomorphisme d'espaces vectoriels.
■

Remarque 3.2.10. Dans la démonstration précédente, après avoir prouvé que l'application φ est linéaire, on aurait pu déduire son injectivité en montrant que $\ker(\varphi) = \{0\}$. Pour vérifier ceci, soit $f \in \mathcal{L}_{\mathbb{K}}(E,F)$ telle que $\varphi(f) = 0$. Alors $M(f,\mathcal{B},\mathcal{B}') = 0$. En posant $M(f,\mathcal{B},\mathcal{B}') = (a_{ij})$, il en découle que tous les a_{ij} sont nuls. Or, par la formule (3.3), $f(e_j) = \sum\limits_{i=1}^{n} a_{ij}e_i$, d'où $f(e_j) = 0_F$ pour tout $j = 1,\dots,p$. Par conséquent, $f = 0$. D'où $\ker(\varphi) = \{0\}$, et par suite, φ est injective.

Remarque 3.2.11. · La bijectivité de φ dans le théorème précédent nous dit qu'étant données des bases fixées B et B' de E et F respectivement, toute matrice $A \in \mathcal{M}_{np}(\mathbb{K})$ peut être considérée comme la matrice $M(f,\mathcal{B},\mathcal{B}')$ d'une application linéaire unique $f : E \longrightarrow F$:

$$\boxed{\forall A \in \mathcal{M}_{n,p}(\mathbb{K}) \quad \exists! f \in \mathcal{L}_{\mathbb{K}}(E,F) : A = M(f,\mathcal{B},\mathcal{B}')}$$

· Souvent, il y a un avantage de travailler avec les espaces $E = \mathbb{K}^p$ et

$F = \mathbb{K}^n$ munis de leurs bases canoniques respectives $\mathcal{B}_p$ et $\mathcal{B}_n$:

$$\boxed{\forall A \in \mathcal{M}_{n,p}(\mathbb{K}) \quad \exists! f \in \mathcal{L}_{\mathbb{K}}(\mathbb{K}^p, \mathbb{K}^n) : A = M(f, \mathcal{B}_p, \mathcal{B}_n)}$$

Dans le cas particulier des endomorphismes ($E = F$), le théorème (3.2.9) devient :

Corollaire 3.2.12. *Soit E un espace vectoriel de dimension finie n sur $\mathbb{K}$ et $\mathcal{B}$ une base de E. L'application*

$$\begin{array}{rccc} \varphi : & \mathcal{L}_{\mathbb{K}}(E) & \longrightarrow & \mathcal{M}_n(\mathbb{K}) \\ & f & \mapsto & M(f, \mathcal{B}) \end{array}$$

est un isomorphisme d'espaces vectoriels.

Remarque 3.2.13. · Ce corollaire implique que si E est un espace vectoriel de dimension n muni d'une base $\mathcal{B}$, alors :

$$\boxed{\forall A \in \mathcal{M}_n(\mathbb{K}) \quad \exists! f \in \mathcal{L}_{\mathbb{K}}(E) : A = M(f, \mathcal{B})}$$

· En particulier, pour $E = \mathbb{K}^n$ muni de sa base canonique $\mathcal{B}_n$:

$$\boxed{\forall A \in \mathcal{M}_n(\mathbb{K}) \quad \exists! f \in \mathcal{L}_{\mathbb{K}}(\mathbb{K}^n) : A = M(f, \mathcal{B}_n)}$$

Revenons à présent au théorème (3.2.9) pour déduire la dimension de l'espace $\mathcal{L}_{\mathbb{K}}(E, F)$ à l'aide de la dimension de $\mathcal{M}_{n,p}(\mathbb{K})$:

Corollaire 3.2.14. *Soient E et F deux espaces vectoriels de dimensions finies sur $\mathbb{K}$. Alors l'espace vectoriel $\mathcal{L}_{\mathbb{K}}(E, F)$ est de dimension finie, et l'on a*

$$\boxed{\dim \mathcal{L}_{\mathbb{K}}(E, F) = (\dim E)(\dim F)}$$

DÉMONSTRATION. Soient $p = \dim E$ et $n = \dim F$. Il suffit d'appliquer la proposition 2.4.9 aux espaces vectoriels isomorphes $\mathcal{L}_{\mathbb{K}}(E, F) \cong \mathcal{M}_{n,p}(\mathbb{K})$, sachant déjà que la dimension de $\mathcal{M}_{n,p}(\mathbb{K})$ est np. ■

Remarque 3.2.15. Lorsque $E = F$, on obtient $\dim \mathcal{L}_{\mathbb{K}}(E) = (\dim E)^2$.

Dans le corollaire précédent, nous avons déterminé la dimension de $\mathcal{L}_{\mathbb{K}}(E, F)$ grâce à celle de $\mathcal{M}_{n,p}(\mathbb{K})$. Autrement dit, les propriétés des matrices peuvent fournir immédiatement celles des applications linéaires.

Réciproquement, nous verrons dans la section suivante qu'à partir des propriétés connues sur les opérations algébriques pour les applications linéaires (voir chapitre 2), on déduit des propriétés sur les opérations algébriques pour les matrices.

Remarque 3.2.16. Les espaces vectoriels $\mathcal{L}_{\mathbb{K}}(E,F)$ et $\mathcal{M}_{n,p}(\mathbb{K})$ sont isomorphes ; mais cet isomorphisme $\varphi : \mathcal{L}_{\mathbb{K}}(E,F) \longrightarrow \mathcal{M}_{n,p}(\mathbb{K})$ dépend des bases $\mathcal{B}$ et $\mathcal{B}'$ choisies dans E et F. En réalité, on devrait le noter

$$\varphi_{\mathcal{B},\mathcal{B}'} : \mathcal{L}_{\mathbb{K}}(E,F) \longrightarrow \mathcal{M}_{n,p}(\mathbb{K})$$

3.3 Multiplication des matrices

3.3.1 Produit de deux matrices

Nous avons déjà appris comment additionner les matrices de même taille et multiplier une matrice par un scalaire. Nous pouvons définir maintenant le produit de deux matrices, à condition que leurs tailles soient compatibles. En fait, la question de multiplication de deux matrices est un peu plus compliquée. La définition suivante du produit de deux matrices sera justifiée plus loin quand on abordera le calcul de la matrice de la composée de deux application linéaires.

Définition 3.3.1. *Soient $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{jk}) \in \mathcal{M}_{p,q}(\mathbb{K})$. On appelle produit AB la matrice (c_{ik}) dans $\mathcal{M}_{n,q}(\mathbb{K})$ dont les cœfficients sont donnés par :*

$$\boxed{c_{ik} = \sum_{j=1}^{p} a_{ij}\ b_{jk} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{ip}b_{pk}}$$

pour tous $i = 1,\ldots,n$ et $k = 1,\ldots,q$.

Remarque 3.3.2. **(a)** Le produit AB de deux matrices est défini seulement quand le nombre de colonnes de la première matrice A est égal au nombre de lignes de la seconde matrice B. Dans ce cas, la matrice AB a le même nombre de lignes que A et le même nombre de colonnes que B :

$$\boxed{\texttt{type}\ (n,\mathbf{p}) \times \texttt{type}\ (\mathbf{p},q) = \texttt{type}\ (n,q)}$$

(b) En particulier, Si $A \in \mathcal{M}_{1,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,1}(\mathbb{K})$, alors $AB \in \mathcal{M}_{1,1}(\mathbb{K})$, c'est-à-dire, AB est un scalaire c_{11} égal à :

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \end{pmatrix} \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{p1} \end{pmatrix} = a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1p}b_{p1}.$$

(c) Soient $A = \begin{pmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix} \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = \begin{pmatrix} b_{11} & \dots & b_{1q} \\ \vdots & & \vdots \\ b_{p1} & \dots & b_{pq} \end{pmatrix} \in \mathcal{M}_{p,q}(\mathbb{K})$.

Pour une vision plus claire du produit $AB = (c_{ik}) \in \mathcal{M}_{n,q}(\mathbb{K})$, remarquons d'abord que la ligne i de A et la colonne k de B sont respectivement

$$L_i = \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{ip} \end{pmatrix} \in \mathcal{M}_{1,p}(\mathbb{K}) \quad \text{et} \quad C_k = \begin{pmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{pk} \end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{K})$$

La multiplication L_iC_k de ces deux matrices est possible :

$$L_iC_k = \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{ip} \end{pmatrix} \begin{pmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{pk} \end{pmatrix} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{ip}b_{pk} = c_{ik}.$$

Nous pouvons alors interpréter la règle de multiplication des matrices en disant que **le cœfficient c_{ik} du produit AB qui se trouve à l'intersection de la ligne i et de la colonne k est égal au produit matriciel de la ligne i de A par la colonne k de B.**
Le produit ainsi défini est parfois appelé "*produit ligne-colonne*".
En notant $L_1, \dots, L_n$ les lignes de A et $C_1, \dots, C_q$ les colonnes de B, on peut donc écrire $AB = (c_{ik})$ avec $c_{ik} = L_iC_k$. D'où

$$AB = \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix} \begin{pmatrix} C_1 & C_2 & \dots & C_q \end{pmatrix} = \begin{pmatrix} L_1C_1 & L_1C_2 & \dots & L_1C_q \\ L_2C_1 & L_2C_2 & \dots & L_2C_q \\ \vdots & \vdots & & \vdots \\ L_nC_1 & L_nC_2 & \dots & L_nC_q \end{pmatrix}$$

Exemple 3.3.3. Considérons les matrices

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{pmatrix}$$

à cœfficients dans $\mathbb{K}$. Puisque A est de type $(3, \mathbf{4})$ et B est de type $(\mathbf{4}, 2)$, alors le produit AB existe et c'est une matrice de type $(3, 2)$. Calculons le produit

$$AB = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{pmatrix}.$$

La formule $c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + a_{i3}b_{3k} + a_{i4}b_{4k}$ $(1 \leq i \leq 3,\ 1 \leq k \leq 2)$ donne

$$AB = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} + a_{14}b_{41} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} + a_{14}b_{42} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} + a_{24}b_{41} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} + a_{24}b_{42} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} + a_{34}b_{41} & a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} + a_{34}b_{42} \end{pmatrix}$$

Exemple 3.3.4. En appliquant la règle pratique de multiplication des matrices, on obtient :

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

Remarque 3.3.5. Lorsque deux matrices A et B sont carrées d'ordre n, on peut toujours effectuer leur produit AB qui est aussi une matrice carrée d'ordre n. Ceci montre que l'ensemble $\mathcal{M}_n(\mathbb{K})$ est stable pour la multiplication des matrices. C'est pour cette raison nous réserverons plus loin une section sur les matrices carrées pour en dire davantage.
Notons que si $n \neq p$, on ne pourra jamais multiplier deux matrices A et B dans $\mathcal{M}_{n,p}(\mathbb{K})$, et donc il n'est pas question de parler de stabilité de la multiplication dans l'ensemble $\mathcal{M}_{n,p}(\mathbb{K})$.
Observons également que si A et B sont des matrices, il est possible que AB puisse être définie et BA n'existe pas. Pour illustrer, il suffit de revenir à l'exemple 3.3.4.
Et même si AB et BA existent, elles n'ont pas nécessairement la même taille. Prendre par exemple $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,n}(\mathbb{K})$. Dans ce cas, $AB \in \mathcal{M}_n(\mathbb{K})$ et $BA \in \mathcal{M}_p(\mathbb{K})$.

3.3.2 Matrice d'un composé d'applications linéaires

La définition apparemment étrange de la multiplication des matrices n'était pas choisie au hasard, mais plutôt pour assurer certaines conditions algébriques importantes. En particulier, elle respecte la connexion naturelle entre les matrices et les applications linéaires.
Plus précisément, on se donne trois espaces vectoriels E, F, G de dimensions finies sur $\mathbb{K}$, avec $\dim E = q$, $\dim F = p$, $\dim G = n$. Fixons une base $\mathcal{B} = (e_1, \ldots, e_q)$ de E, une base $\mathcal{B}' = (e'_1, \ldots, e'_p)$ de F et une base $\mathcal{B}'' = (e''_1, \ldots, e''_n)$ de G.

Soient f une application linéaire de E dans F, et $g : F \longrightarrow G$ une application linéaire de F dans G :

$$E \xrightarrow{f} F \xrightarrow{g} G$$

On a ainsi une application linéaire $g \circ f : E \longrightarrow G$. Considérons les deux matrices $M(f, \mathcal{B}, \mathcal{B}') \in \mathcal{M}_{p,q}(\mathbb{K})$ et $M(g, \mathcal{B}', \mathcal{B}'') \in \mathcal{M}_{n,p}(\mathbb{K})$. On se propose de chercher la matrice $M(g \circ f, \mathcal{B}, \mathcal{B}'') \in \mathcal{M}_{n,q}(\mathbb{K})$ de l'application linéaire $g \circ f : E \longrightarrow G$ relativement aux bases $\mathcal{B}$ et $\mathcal{B}''$:

Théorème 3.3.6. *Avec les notations précédentes, on a*

$$\boxed{M(g \circ f, \mathcal{B}, \mathcal{B}'') = M(g, \mathcal{B}', \mathcal{B}'') \times M(f, \mathcal{B}, \mathcal{B}')}$$

Ce théorème affirme que, dans des bases convenables, la matrice de la composée de deux applications linéaires est le produit des matrices de ces applications linéaires

DÉMONSTRATION. Posons $A = M(g, \mathcal{B}', \mathcal{B}'') = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$, $B = M(f, \mathcal{B}, \mathcal{B}') = (b_{jk}) \in \mathcal{M}_{p,q}(\mathbb{K})$ et $C = M(g \circ f, \mathcal{B}, \mathcal{B}'') \in \mathcal{M}_{n,q}(\mathbb{K})$. Il est clair que le produit AB existe et $AB \in \mathcal{M}_{n,q}(\mathbb{K})$. Montrons que $AB = C$. Soit $k \in \{1, \dots, q\}$. Par définition de B, on a

$$f(e_k) = \sum_{j=1}^{p} b_{jk}\ e'_j.$$

D'où

$$(g \circ f)(e_k) = g(f(e_k)) = g\left(\sum_{j=1}^{p} b_{jk}\ e'_j\right) = \sum_{j=1}^{p} b_{jk}\ g(e'_j).$$

Or, par définition de A, pour tout $j \in \{1, \dots, p\}$, on a

$$g(e'_j) = \sum_{i=1}^{n} a_{ij}\ e''_i.$$

Donc

$$\begin{aligned}(g \circ f)(e_k) &= \sum_{j=1}^{p} b_{jk}\ g(e'_j) = \sum_{j=1}^{p} b_{jk}\left(\sum_{i=1}^{n} a_{ij}\ e''_i\right) \\ &= \sum_{j=1}^{p}\sum_{i=1}^{n} b_{jk}\ a_{ij}\ e''i = \sum_{i=1}^{n}\left(\sum_{j=1}^{p} a_{ij}\ b_{jk}\right) e''_i\end{aligned}$$

C'est-à-dire

$$(g \circ f)(e_k) = \sum_{i=1}^{n} c_{ik}\ e''_i$$

avec

$$c_{ik} = \sum_{j=1}^{p} a_{ij}\ b_{jk}. \tag{3.4}$$

Il en découle que $C = M(g \circ f, \mathcal{B}, \mathcal{B}'') = (c_{ik})$.
D'autre part, la relation (3.4) nous dit que $C = AB$. D'où le résultat désiré. ■

Exemple 3.3.7. Pour les étudiants qui trouvent certaines difficultés à manipuler les sommations et les indices, je leur propose de reprendre la démonstration précédente dans le cas particulier $\dim E = q = 3$, $\dim F = p = 2$, $\dim G = n = 4$. Cela permettra de comprendre mieux les calculs ci-dessus. Dans ce cas, $\mathcal{B} = (e_1, e_2, e_3)$, $\mathcal{B}' = (e'_1, e'_2)$, $\mathcal{B}'' = (e''_1, e''_2, e''_3, e''_4)$,

$$A = M(g, \mathcal{B}', \mathcal{B}'') = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{pmatrix}, \ B = M(f, \mathcal{B}, \mathcal{B}') = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}$$

Calculer le produit AB. Ensuite, déterminer la matrice $C = M(g \circ f, \mathcal{B}, \mathcal{B}'')$ en calculant d'abord les vecteurs $(g \circ f)(e_1), (g \circ f)(e_2), (g \circ f)(e_3)$ en fonction de $e''_1, e''_2, e''_3, e''_4$. Comparer cette matrice avec AB.
Allez-y, secouez vous et terminer l'exemple.

Lorsqu'on travaille avec les endomorphismes ($E = F$), on obtient le cas spécial :

Corollaire 3.3.8. *Soit E un espace vectoriel de dimension finie sur $\mathbb{K}$ et B une base de E. Pour tous endomorphismes f et g de E, on a*

$$\boxed{M(g \circ f, \mathcal{B}) = M(g, \mathcal{B}) \times M(f, \mathcal{B})}$$

3.3.3 Propriétés algébriques du produit matriciel

Le théorème 3.3.6 va nous permettre de montrer que les propriétés du produit matriciel sont celles de la composition des applications linéaires : associativité, distributivité à droite et à gauche par rapport à l'addition, non commutativité, etc. Les résultats de la section 2.4.2 (Composition des applications linéaires) du chapitre précédent ont tous une traduction matricielle :

Proposition 3.3.9. *Pour toutes matrices A, B, C et pour tout scalaire $\lambda \in \mathbb{K}$, on a les égalités*

(a) $A(B + C) = AB + AC$

(b) $(B + C)A = BA + CA$

(c) $(AB)C = A(BC)$

(d) $\lambda(AB) = (\lambda A)B = A(\lambda B)$

lorsque les produits des matrices considérées sont définis.

Démonstration. Pour démontrer cette proposition, nous avons deux possibilités. Soit nous appliquons la formule du produit matriciel, soit nous utilisons le calcul linéaire en nous ramenant aux applications linéaires. La deuxième possibilité est plus simple et plus constructive. Mais nous allons utiliser les deux méthodes pour que les choses soient complètes :

(a) Calcul linéaire :
Pour montrer l'égalité matricielle $A(B + C) = AB + AC$, on va se ramener à l'égalité $h \circ (f + g) = h \circ f + h \circ g$ des applications linéaires associées. Mais tout d'abord, préparons les choses convenablement :
Soient $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B, C \in \mathcal{M}_{p,q}(\mathbb{K})$. On a $AB, AC \in \mathcal{M}_{n,q}(\mathbb{K})$, donc $AB + AC \in \mathcal{M}_{n,q}(\mathbb{K})$. D'autre part, comme $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B + C \in \mathcal{M}_{p,q}(\mathbb{K})$, alors $A(B + C) \in \mathcal{M}_{n,p}(\mathbb{K})$. Les deux matrices $A(B+C)$ et $AB+AC$ ont donc la même taille. Montrons qu'elles sont égales.
Considérons les trois espaces vectoriels $\mathbb{K}^n$, $\mathbb{K}^p$, $\mathbb{K}^q$ munis de leurs bases canoniques respectives $\mathcal{B}_n$, $\mathcal{B}_p, \mathcal{B}_q$.
Puisque $A \in \mathcal{M}_{n,p}(\mathbb{K})$, la remarque 3.2.11 assure l'existence d'une application linéaire unique $h : \mathbb{K}^p \longrightarrow \mathbb{K}^n$ telle que $A = M(h, \mathcal{B}_p, \mathcal{B}_n)$.
De même, comme $B, C \in \mathcal{M}_{p,q}(\mathbb{K})$, il existe des applications linéaires uniques $f, g : \mathbb{K}^q \longrightarrow \mathbb{K}^p$ telles que $B = M(f, \mathcal{B}_q, \mathcal{B}_p)$ et $C = M(g, \mathcal{B}_q, \mathcal{B}_p)$.
Après avoir fait ces présentations, les calculs vont être très simples :

$$\begin{aligned}
A \times (B + C) &= M(h, \mathcal{B}_p, \mathcal{B}_n) \times \Big(M(f, \mathcal{B}_q, \mathcal{B}_p) + M(g, \mathcal{B}_q, \mathcal{B}_p)\Big) \\
&= M(h, \mathcal{B}_p, \mathcal{B}_n) \times M(f + g, \mathcal{B}_q, \mathcal{B}_p) \\
&= M(h \circ (f + g), \mathcal{B}_q, \mathcal{B}_n) \\
&= M(h \circ f + h \circ g, \mathcal{B}_q, \mathcal{B}_n) \\
&= M(h \circ f, \mathcal{B}_q, \mathcal{B}_n) + M(h \circ g, \mathcal{B}_q, \mathcal{B}_n) \\
&= M(h, \mathcal{B}_p, \mathcal{B}_n) \times M(f, \mathcal{B}_q, \mathcal{B}_p) + M(h, \mathcal{B}_p, \mathcal{B}_n) \times M(g, \mathcal{B}_q, \mathcal{B}_p) \\
&= A \times B + A \times C
\end{aligned}$$

(b) Calcul matriciel :
Pour que les produits intervenant dans l'égalité $(B+C)A = BA+CA$ existent, il faut maintenant prendre $B, C \in \mathcal{M}_{n,p}(\mathbb{K})$ et $A \in \mathcal{M}_{p,q}(\mathbb{K})$.
Posons

$$B = (b_{ij})_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq p}} \quad C = (c_{ij})_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq p}} \quad A = (a_{jk})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}}$$

On a

$$B+C=(t_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq p}}\qquad\text{avec}\quad t_{ij}=b_{ij}+c_{ij}.$$

Donc

$$(B+C)A=(d_{ik})_{\substack{1\leq i\leq n\\1\leq k\leq q}}\qquad\text{avec}\quad d_{ik}=\sum_{j=1}^{p}t_{ij}\,a_{jk}=\sum_{j=1}^{p}b_{ij}\,a_{jk}+\sum_{j=1}^{p}c_{ij}\,a_{jk}\tag{3.5}$$

D'autre part, on a

$$BA=(s_{ik})_{\substack{1\leq i\leq n\\1\leq k\leq q}}\qquad\text{avec}\quad s_{ik}=\sum_{j=1}^{p}b_{ij}\,a_{jk},$$

$$CA=(u_{ik})_{\substack{1\leq i\leq n\\1\leq k\leq q}}\qquad\text{avec}\quad u_{ik}=\sum_{j=1}^{p}c_{ij}\,a_{jk}.$$

Donc

$$BA+CA=(v_{ik})_{\substack{1\leq i\leq n\\1\leq k\leq q}}\text{ avec }v_{ik}=s_{ik}+u_{ik}=\sum_{j=1}^{p}b_{ij}\,a_{jk}+\sum_{j=1}^{p}c_{ij}\,a_{jk}\tag{3.6}$$

En comparant (3.5) et (3.6), on conclut que $(B+C)A=BA+CA$.

(c) Considérons trois matrices $A\in\mathcal{M}_{n,p}(\mathbb{K}),\ B\in\mathcal{M}_{p,q}(\mathbb{K}),\ C\in\mathcal{M}_{q,r}(\mathbb{K})$. Remarquons que les matrices $(AB)C$ et $A(BC)$ sont toutes les deux de type (n,r). Pour prouver l'associativité $(AB)C=A(BC)$, nous allons utiliser les deux méthodes :

- Calcul matriciel : Posons

$$A=(a_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq p}},\quad B=(b_{jk})_{\substack{1\leq j\leq p\\1\leq k\leq q}},\quad C=(c_{kl})_{\substack{1\leq k\leq q\\1\leq l\leq r}}$$

On a

$$AB=(x_{ik})_{\substack{1\leq i\leq n\\1\leq k\leq q}}\qquad\text{avec}\quad x_{ik}=\sum_{j=1}^{p}a_{ij}\,b_{jk}$$

Donc

$$(AB)C=(y_{il})_{\substack{1\leq i\leq n\\1\leq l\leq r}}\quad\text{avec}\quad y_{il}=\sum_{k=1}^{q}x_{ik}\,c_{kl}=\sum_{k=1}^{q}\left(\sum_{j=1}^{p}a_{ij}\,b_{jk}\right)c_{kl}\tag{3.7}$$

De manière similaire,

$$BC=(z_{jl})_{\substack{1\leq j\leq p\\1\leq l\leq r}}\qquad\text{avec}\quad z_{jl}=\sum_{k=1}^{q}b_{jk}\,c_{kl}$$

Donc

$$A(BC)=(w_{il})_{\substack{1\leq i\leq n\\1\leq l\leq r}}\quad\text{avec}\quad w_{il}=\sum_{j=1}^{p}a_{ij}\,z_{jl}=\sum_{j=1}^{p}a_{ij}\left(\sum_{k=1}^{q}b_{jk}\,c_{kl}\right)\tag{3.8}$$

D'après (3.7) et (3.8), on a pour tous $i = 1, \dots, n$ et $l = 1, \dots, r$:

$$y_{il} = \sum_{k=1}^{q} \sum_{j=1}^{p} a_{ij}\, b_{jk}\, c_{kl} \qquad \text{et} \qquad w_{il} = \sum_{j=1}^{p} \sum_{k=1}^{q} a_{ij}\, b_{jk}\, c_{kl}$$

Mais ces deux doubles sommations sont égales, d'où $y_{il} = w_{il}$ pour tous $i = 1, \dots, n$ et $l = 1, \dots, r$. En conclusion, $(AB)C = A(BC)$.

- Calcul linéaire :

  Considérons les quatre espaces vectoriels $\mathbb{K}^n$, $\mathbb{K}^p$, $\mathbb{K}^q$, $\mathbb{K}^r$ munis de leurs bases canoniques respectives $\mathcal{B}_n$, $\mathcal{B}_p, \mathcal{B}_q, \mathcal{B}_r$.
  Puisque $A \in \mathcal{M}_{n,p}(\mathbb{K})$, $B \in \mathcal{M}_{p,q}(\mathbb{K})$, $C \in \mathcal{M}_{q,r}(\mathbb{K})$, il existe des applications linéaires uniques

$$\mathbb{K}^r \xrightarrow{h} \mathbb{K}^q \xrightarrow{g} \mathbb{K}^p \xrightarrow{f} \mathbb{K}^n$$

  telles que

$$A = M(f, \mathcal{B}_p, \mathcal{B}_n), \quad B = M(g, \mathcal{B}_q, \mathcal{B}_p), \quad C = M(h, \mathcal{B}_r, \mathcal{B}_q).$$

  On en déduit :

$$\begin{aligned}
(A \times B) \times C &= \Big(M(f, \mathcal{B}_p, \mathcal{B}_n) \times M(g, \mathcal{B}_q, \mathcal{B}_p)\Big) \times M(h, \mathcal{B}_r, \mathcal{B}_q) \\
&= M(f \circ g, \mathcal{B}_q, \mathcal{B}_n) \times M(h, \mathcal{B}_r, \mathcal{B}_q) \\
&= M((f \circ g) \circ h, \mathcal{B}_r, \mathcal{B}_n) \\
&= M(f \circ (g \circ h), \mathcal{B}_r, \mathcal{B}_n) \\
&= M(f, \mathcal{B}_p, \mathcal{B}_n) \times M(g \circ h, \mathcal{B}_r, \mathcal{B}_p) \\
&= M(f, \mathcal{B}_p, \mathcal{B}_n) \times \Big(M(g, \mathcal{B}_q, \mathcal{B}_p) \times M(h, \mathcal{B}_r, \mathcal{B}_q)\Big) \\
&= A \times (B \times C)
\end{aligned}$$

  Notons que l'outil-clef dans cette méthode linéaire est l'associativité des applications : $(f \circ g) \circ h = f \circ (g \circ h)$, qui est vraie non seulement pour les applications linéaires, mais pour les applications quelconques. ■

(d) Nous laissons aux étudiants le soin de vérifier la formule $\lambda(AB) = (\lambda A)B = A(\lambda B)$, soit par un calcul matriciel simple, soit par un calcul linéaire utilisant la formule $\lambda(g \circ f) = (\lambda g) \circ f = g \circ (\lambda f)$ vue au chapitre 2. ■

Attention ! Le produit de deux matrices n'est pas commutatif :
En général, $AB \neq BA$.
Si $A = 0$ ou $B = 0$ alors $AB = 0$. Mais la réciproque n'est pas vraie :
Si $AB = 0$, on ne peut pas conclure en général que $A = 0$ ou $B = 0$.
De même, si $AB = AC$, il n'est pas vrai en général que $B = C$.

La multiplication des matrices obéit aux règles habituelles de l'arithmétique. Cependant, elle n'est pas commutative, car, si l'on revient aux applications linéaires, on sait que $g \circ f \neq f \circ g$ en général. L'exemple suivant précise ce point :

Exemple 3.3.10. Soient les deux matrices $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ et $B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$.

$AB = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$ et $BA = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}$

Donc $AB \neq BA$.

Exemple 3.3.11. Le produit AB de deux matrices non nulles A et B peut être égal à la matrice nulle. Autrement dit, il y a des *diviseurs de zéro* pour la multiplication des matrices :

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} -16 & 2 \\ 8 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Si $AB = AC$ et $A \neq 0$, on n'a pas en général $B = C$. En effet, $AB = AC$ entraîne $A(B - C) = 0$. Or, d'après l'exemple précédent, ceci n'implique pas nécessairement $B - C = 0$. Pour s'en convaincre, voici un exemple précis :

Exemple 3.3.12. Considérons les trois matrices

$$A = \begin{pmatrix} -3 & 2 \\ -6 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} -1 & 2 \\ 3 & -2 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 4 \\ 6 & 1 \end{pmatrix}$$

Nous vous laissons le soin de vérifier que

$$AB = AC = \begin{pmatrix} 9 & -10 \\ 18 & -20 \end{pmatrix}$$

Proposition 3.3.13. *La matrice unité I_n d'ordre n vérifie les conditions suivantes :*

- $I_n \times A = A \qquad \forall A \in \mathcal{M}_{n,p}(\mathbb{K})$.
- $B \times I_n = B \qquad \forall B \in \mathcal{M}_{p,n}(\mathbb{K})$.

Démonstration. Pour toutes matrices $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{ij}) \in \mathcal{M}_{p,n}(\mathbb{K})$, les calculs directs montrent clairement que :

$$I_n A = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} = A,$$

$$BI_n = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{p1} & b_{p2} & \dots & b_{pn} \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{p1} & b_{p2} & \dots & b_{pn} \end{pmatrix} = B$$

On peut le voir d'une autre manière en se ramenant aux applications linéaires. Pour cela, soient les espaces vectoriels $\mathbb{K}^n$ et $\mathbb{K}^p$ rapportés à leurs bases canoniques respectives $\mathcal{B}_n$ et $\mathcal{B}_p$. On sait que $I_n = M(\mathtt{id}_{\mathbb{K}^n}, \mathcal{B}_n) = M(\mathtt{id}_{\mathbb{K}^n}, \mathcal{B}_n, \mathcal{B}_n)$. Comme $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,n}(\mathbb{K})$, il existe des applications linéaires uniques $f : \mathbb{K}^p \longrightarrow \mathbb{K}^n$ et $g : \mathbb{K}^n \longrightarrow \mathbb{K}^p$ telles que $A = M(f, \mathcal{B}_p, \mathcal{B}_n)$ et $B = M(g, \mathcal{B}_n, \mathcal{B}_p)$. Il en découle :

$$I_n \times A = M(\mathtt{id}_{\mathbb{K}^n}, \mathcal{B}_n, \mathcal{B}_n) \times M(f, \mathcal{B}_p, \mathcal{B}_n) = M(\mathtt{id}_{\mathbb{K}^n} \circ f, \mathcal{B}_p, \mathcal{B}_n) = M(f, \mathcal{B}_p, \mathcal{B}_n) = A.$$

$$B \times I_n = M(g, \mathcal{B}_n, \mathcal{B}_p) \times M(\mathtt{id}_{\mathbb{K}^n}, \mathcal{B}_n, \mathcal{B}_n) = M(g \circ \mathtt{id}_{\mathbb{K}^n}, \mathcal{B}_n, \mathcal{B}_p) = M(g, \mathcal{B}_n, \mathcal{B}_p) = B.$$

■

Exemple 3.3.14. Considérons la matrice $A = \begin{pmatrix} 10 & 0 \\ -3 & 8 \\ -1 & 11 \\ 7 & -4 \end{pmatrix}$. Alors

$$I_4 A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 10 & 0 \\ -3 & 8 \\ -1 & 11 \\ 7 & -4 \end{pmatrix} = \begin{pmatrix} 10 & 0 \\ -3 & 8 \\ -1 & 11 \\ 7 & -4 \end{pmatrix},$$

$$AI_2 = \begin{pmatrix} 10 & 0 \\ -3 & 8 \\ -1 & 11 \\ 7 & -4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 10 & 0 \\ -3 & 8 \\ -1 & 11 \\ 7 & -4 \end{pmatrix}$$

3.3.4 Ecriture matricielle d'une application linéaire

Soient E et F deux espaces vectoriels de dimensions respectives p et n sur $\mathbb{K}$, $\mathcal{B} = (e_1, \dots, e_p)$ une base de E et $\mathcal{B}' = (e'_1, \dots, e'_n)$ une base de F. Considérons une application linéaire $f : E \longrightarrow F$ et $A = M(f, \mathcal{B}, \mathcal{B}') = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$ sa matrice dans les bases $\mathcal{B}$ et $\mathcal{B}'$. On rappelle que chaque colonne numéro j de la matrice A contient les coordonnées du vecteur $f(e_j)$ dans la base $\mathcal{B}'$:

$$f(e_j) = a_{1j} e'_1 + a_{2j} e'_2 + \dots + a_{nj} e'_n = \sum_{i=1}^{n} a_{ij} e'_i \qquad \text{pour tout } j = 1, \dots, p.$$

De plus, selon la remarque (3.2.4), pour tout vecteur $x = x_1e_1 + x_2e_2 + \cdots + x_pe_p \in E$, son image $y = f(x) \in F$ est donnée par :

$$\begin{aligned} f(x) &= x_1 f(e_1) + x_2 f(e_2) + \cdots + x_p f(e_p) \\ &= x_1\Big(a_{11}e'_1 + a_{21}e'_2 + \cdots + a_{n1}e'_n\Big) + x_2\Big(a_{12}e'_1 + a_{22}e'_2 + \cdots + a_{n2}e'_n\Big) \\ &\quad + \cdots + x_p\Big(a_{1p}e'_1 + a_{2p}e'_2 + \cdots + a_{np}e'_n\Big) \\ &= \Big(x_1a_{11} + x_2a_{12} + \cdots + x_pa_{1p}\Big)e'_1 + \Big(x_1a_{21} + x_2a_{22} + \cdots + x_pa_{2p}\Big)e'_2 \\ &\quad + \cdots + \Big(x_1a_{n1} + x_2a_{n2} + \cdots + x_pa_{np}\Big)e'_n \end{aligned}$$

Donc $y = f(x) = y_1e'_1 + y_2e'_2 + \cdots + y_ne'_n$ avec

$$\left\{\begin{array}{lll} y_1 & = & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p \\ y_2 & = & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p \\ \vdots & \vdots & \qquad\qquad \vdots \\ y_n & = & a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{np}x_p \end{array}\right. \tag{3.9}$$

c'est-à-dire :

$$y_i = \sum_{j=1}^{p} a_{ij}x_j \qquad i = 1, \ldots, n.$$

Ecrivons les coordonnées de $x \in E$ dans la base $\mathcal{B}$ sous la forme d'une matrice uni-colonne à p lignes :

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{K}).$$

La matrice X s'appelle la **matrice-colonne des composantes de** x **dans la base** $\mathcal{B}$ de E.
De même, représentons les coordonnées de $y = f(x) \in F$ dans la base $\mathcal{B}'$ sous la forme d'une matrice uni-colonne à n lignes :

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{K}),$$

qui est la matrice-colonne des comoposantes de y dans la base $\mathcal{B}'$ de F.
Les relations (3.9) s'écrivent :

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1p} \\ a_{21} & a_{22} & \ldots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$$

D'où le résultat suivant :

Proposition 3.3.15. *Avec les hypothèses et notations précédentes, l'égalité $y = f(x)$ se traduit matriciellement par :*

$$\boxed{Y = AX}$$

Il s'agit ici du produit matriciel de la matrice A de type (n, p) par la matrice colonne X de type $(p, 1)$. Le résultat est la matrice Y de type $(n, 1)$.

Exemple 3.3.16. Soient E et F deux espaces vectoriels de dimensions respectives $p = 2$ et $n = 3$ sur $\mathbb{K}$, $\mathcal{B} = (e_1, e_2)$ une base de E et $\mathcal{B}' = (e'_1, e'_2, e'_3)$ une base de F. Soit l'application linéaire $f : E \longrightarrow F$ telle que sa matrice dans les bases $\mathcal{B}$ et $\mathcal{B}'$ soit

$$A = M(f, \mathcal{B}, \mathcal{B}') = \begin{pmatrix} -3 & 1 \\ 0 & 5 \\ -1 & 2 \end{pmatrix}$$

Soit $x = x_1e_1 + x_2e_2$ un vecteur quelconque de E et $y = f(x) = y_1e'_1 + y_2e'_2 + y_3e'_3$ son image dans F. La formule matricielle $Y = AX$ s'écrit

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -3 & 1 \\ 0 & 5 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -3x_1 + x_2 \\ 5x_2 \\ -x_1 + 2x_2 \end{pmatrix}$$

. On en déduit que

$$f(x_1e_1 + x_2e_2) = (-3x_1 + x_2)e'_1 + (5x_2)e'_2 + (-x_1 + 2x_2)e'_3$$

On peut trouver le noyau et l'image de f :

$$x = x_1e_1 + x_2e_2 \in \ker(f) \iff f(x) = 0_F \iff \left\{ \begin{array}{rcl} -3x_1 + x_2 & = & 0 \\ 5x_2 & = & 0 \\ -x_1 + 2x_2 & = & 0 \end{array} \right.$$

$\iff x_1 = x_2 = 0 \iff x = 0_E$.

Donc $\ker(f) = \{0_E\}$ et par suite f est injective. D'autre part, $\mathsf{Im}(f) = \mathsf{Vect}\{f(e_1), f(e_2)\}$. Or les vecteurs $f(e_1) = (3, 0, -1)$ et $f(e_2) = (1, 5, 2)$ ne sont pas colinéaires, donc ils sont linéairement indépendants, et donc ils forment une base de $\mathsf{Im}(f)$. Par conséquent, $\mathsf{Im}(f) \neq F$, d'où f n'est pas surjective. Ceci est logique puisque $\dim E \neq \dim F$ entraîne que f ne peut être un isomorphisme. Néanmoins, si nous remplaçons l'espace d'arrivée F par le plan vectoriel $\mathsf{Im}(f)$, alors l'application linéaire $f : E \longrightarrow \mathsf{Im}(f)$ devient maintenant un isomorphisme, car elle transforme la base (e_1, e_2) de E en une base $(f(e_1), f(e_2))$ de $\mathsf{Im}(f)$. Ainsi, $E \cong \mathsf{Im}(f)$.

Remarque 3.3.17. Soient les espaces vectoriels $\mathbb{K}^p$ et $\mathbb{K}^n$ équipés de leurs bases canoniques respectives $\mathcal{B}_p = (e_1, \dots, e_p)$ et $\mathcal{B}_n = (e'_1, \dots, e'_n)$. Etant donné une matrice $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$, considérons l'application

$f : \mathbb{K}^p \longrightarrow \mathbb{K}^n$ qui à tout vecteur $x = \sum_{i=1}^{p} x_i e_i$ de $\mathbb{K}^p$ associe le vecteur $y = f(x) = \sum_{i=1}^{n} y_i e'_i$ de $\mathbb{K}^n$ tel que $Y = AX$, c'est-à-dire :

$$y_i = \sum_{j=1}^{p} a_{ij} x_j \qquad i = 1, \ldots, n.$$

Alors f est une application linéaire et A est la matrice de f relativement aux bases $\mathcal{B}_p$ et $\mathcal{B}_n$: $A = M(f, \mathcal{B}_p, \mathcal{B}_n)$. On dit que f est l'application linéaire *canoniquement associée* à la matrice A. On identifie l'espace $\mathbb{K}^p$ à $\mathcal{M}_{p,1}(\mathbb{K})$ et l'espace $\mathbb{K}^n$ à $\mathcal{M}_{n,1}(\mathbb{K})$, ce qui revient à écrire les vecteurs de ces espaces sous forme de vecteurs colonnes. Par conséquent, on peut écrire $f : \mathbb{K}^p \longrightarrow \mathbb{K}^n$, $X \mapsto AX$. Lorsqu'il n'y aura pas d'ambiguïté, il nous arrivera de "confondre" la matrice A et l'application linéaire f canoniquement associée. On parlera alors du noyau de A et l'image de A :

$$\ker(A) = \{X \in \mathbb{K}^p \ / \ AX = 0\}, \qquad \mathsf{Im}(A) = \{AX \ / \ X \in \mathbb{K}^p\}.$$

On peut dire aussi que $\mathsf{Im}(A)$ est le sous-espace vectoriel de $\mathbb{K}^n$ engendré par les colonnes de la matrice A.

Remarque 3.3.18. Si $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$ est une matrice qui vérifie la condition : $AX = 0 \quad \forall X \in \mathcal{M}_{n,1}(\mathbb{K})$, alors nécessairement $A = 0$.
En effet, cela signifie que l'application linéaire canoniquement associée $f : \mathbb{K}^p \longrightarrow \mathbb{K}^n$ vérifie $f(x) = 0_{\mathbb{K}_n} \quad \forall x \in \mathbb{K}^p$. Donc $f = 0$, d'où $A = 0$.

On peut retrouver ceci matriciellement en prenant $X = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$, où 1 se trouve sur la ligne i de X et tous les autres cœfficients nuls :

$$AX = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1p} \\ a_{21} & a_{22} & \ldots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{np} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix}$$

Par conséquent, $AX = 0$ implique que la colonne i de A est nulle. En faisant varier $i \in \{1, \ldots, p\}$, on obtient que toutes les colonnes de A sont nulles, d'où la matrice A est nulle.
On en déduit que si $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$ sont deux matrices vérifiant $AX = BX \quad \forall X \in \mathcal{M}_{n,1}(\mathbb{K})$, alors $A = B$.

3.4 L'algèbre $\mathcal{M}_n(\mathbb{K})$ des matrices carrées d'ordre n

Nous savons que l'ensemble $\mathcal{M}_n(\mathbb{K})$ des matrices carrées d'ordre n à cœfficients dans $\mathbb{K}$ a déjà une structure d'espace vectoriel sur $\mathbb{K}$. De plus, les règles de calculs pour l'addition et la multiplication que nous avons démontrées dans la sous-section 3.3.3 sont partout définies dans $\mathcal{M}_n(\mathbb{K})$. Ainsi, la multiplication est une loi de composition interne dans $\mathcal{M}_n(\mathbb{K})$. Elle est associative, distributive à droite et à gauche par rapport à l'addition. Elle possède un élément neutre I_n, la matrice unité d'ordre n (canoniquement associée à l'endomorphisme identité $\mathtt{id}_{\mathbb{K}^n}$ de $\mathbb{K}^n$) :

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

Il en découle que $(\mathcal{M}_n(\mathbb{K}), +, \times)$ est un anneau unitaire non commutatif. De plus, la multiplication vérifie

$$\forall \lambda \in \mathbb{K} \quad \forall A, B \in \mathcal{M}_n(\mathbb{K}) \qquad \lambda(AB) = (\lambda A)B = A(\lambda B).$$

Ces arguments justifient le théorème suivant :

Théorème 3.4.1. $(\mathcal{M}_n(\mathbb{K}), +, \cdot, \times)$ *est une algèbre unitaire non commutative*

D'autre part, soit E un espace vectoriel de dimension finie n sur $\mathbb{K}$ et de base $\mathcal{B}$. D'après le corollaire (3.2.12), l'application $\varphi : (\mathcal{L}_{\mathbb{K}}(E), +, \cdot) \longrightarrow (\mathcal{M}_n(\mathbb{K}), +, \cdot)$ définie par $\varphi(f) = M(f, \mathcal{B})$ est un isomorphisme d'espaces vectoriels. En outre, le corollaire 3.3.8 nous dit que

$$\varphi(g \circ f) = M(g \circ f, \mathcal{B}) = M(g, \mathcal{B}) \times M(f, \mathcal{B}) = \varphi(g) \times \varphi(f).$$

Donc $\varphi : (\mathcal{L}_{\mathbb{K}}(E), +, \circ) \longrightarrow (\mathcal{M}_n(\mathbb{K}), +, \times)$ est aussi un homomorphisme d'anneaux. En combinant ces considérations, on peut énoncer :

Théorème 3.4.2. *Soit E un espace vectoriel de dimension finie n sur $\mathbb{K}$ et de base $\mathcal{B}$. L'application*

$$\begin{array}{rccc} \varphi : & (\mathcal{L}_{\mathbb{K}}(E), +, \cdot, \circ) & \longrightarrow & (\mathcal{M}_n(\mathbb{K}), +, \cdot, \times) \\ & f & \mapsto & M(f, \mathcal{B}) \end{array}$$

est un isomorphisme d'algèbres.
Cet isomorphisme φ dépend évidemment de la base $\mathcal{B}$ choisie.

Ce théorème est en particulier vrai pour l'espace vectoriel $E = \mathbb{K}^n$ muni de sa base canonique $\mathcal{B}_n$. Donc l'application $\varphi : (\mathcal{L}_{\mathbb{K}}(\mathbb{K}^n), +, \cdot, \circ) \longrightarrow$

$(\mathcal{M}_n(\mathbb{K}), +, \cdot, \times)$ donnée par $\varphi(f) = M(f, \mathcal{B}_n)$ est un isomorphisme d'algèbres. Les propriétés des algèbres $(\mathcal{L}_{\mathbb{K}}(\mathbb{K}^n), +, \cdot, \circ)$ et $(\mathcal{M}_n(\mathbb{K}), +, \cdot, \times)$ sont donc identiques. En particulier, dès que $n \geq 2$, ces algèbres ne sont ni commutatives, ni intègres. Ces algèbres sont de dimension n^2. Il est clair que dans le cas trivial $n = 1$, $\mathcal{M}_1(\mathbb{K})$ s'identifie à $\mathbb{K}$, qui est donc une algèbre commutative intègre.

Remarque 3.4.3. • Si A est une matrice carrée d'ordre n, on peut considérer ses puissances A^i pour $i \geq 0$. Elles sont définies par $A^0 = I_n$ et $A^i = \underbrace{AA \times A}_{i fois}$. Il est clair que $A^iA^j = A^{i+j} = A^jA^i$ pour tous entiers $i, j \geq 0$. Donc, malgré le fait que la multiplication des matrices n'est pas commutative, les puissances d'une même matrice A commutent entre elles.
On a aussi la formule $(A^i)^j = A^{ij}$ pour tous $i, j \geq 0$.
• Si A et B sont deux matrices carrées d'ordre n, on n'a pas en général la formule $(AB)^i = A^iB^i$ pour $i \geq 1$. Ceci provient du fait que $(AB)^i = \underbrace{(AB)(AB)\ldots(AB)}_{i fois}$.
Par contre, si $AB = BA$, ce qui est possible pour certaines matrices, alors en permutant de proche en proche A et B dans cette expression, on trouve finalement $(AB)^i = (\underbrace{AA\ldots A)}_{i fois})(\underbrace{BB\ldots B)}_{i fois}) = A^iB^i$.

• De même, $(A + B)^2 = (A + B)(A + B) = A(A + B) + B(A + B) = A^2 + AB + BA + B^2 \neq A^2 + 2AB$ en général, et l'on a égalité si $AB = BA$.

La *formule du binôme de Newton* pour les matrices se démontre exactement comme dans un anneau unitaire quelconque lorsque les deux éléments commutent :

Proposition 3.4.4 (Formule du binôme de Newton)**.** *Soient A et B deux matrices dans $\mathcal{M}_n(\mathbb{K})$ avec $AB = BA$. Alors pour tout entier $p \in \mathbb{N}^*$,*

$$\boxed{(A + B)^p = \sum_{k=0}^{p} \mathsf{C}_p^k \, A^k \, B^{p-k}}$$

où $\mathsf{C}_p^k = \dfrac{p!}{k!\,(p-k)!}$ *sont les cœfficients binômiaux.*

Démonstration. Par récurrence sur $p \geq 1$:

· Pour $p = 1$, nous avons

$$\sum_{k=0}^{1} \mathsf{C}_1^k A^k B^{1-k} = \mathsf{C}_1^0 A^0 B^1 + \mathsf{C}_1^1 A^1 B^0 = B + A = A + B.$$

La formule est vraie pour $p = 1$.

- Supposons que la formule soit vraie pour un entier $p \geq 1$ et montrons-la pour $p + 1$. On a

$$(A+B)^{p+1} = (A+B)(A+B)^p = (A+B)\sum_{k=0}^{p} \mathsf{C}_p^k A^k B^{p-k}.$$

En utilisant la distributivité du produit par rapport à l'addition et le fait que $AB = BA$, nous obtenons

$$(A+B)^{p+1} = \sum_{k=0}^{p} \mathsf{C}_p^k A^{k+1} B^{p-k} + \sum_{k=0}^{p} \mathsf{C}_p^k A^k B^{p+1-k}.$$

En posant $l = k + 1$ dans la première somme, on a $k = l - 1$, d'où

$$(A+B)^{p+1} = \sum_{l=1}^{p+1} \mathsf{C}_p^{l-1} A^l B^{p+1-l} + \sum_{k=0}^{p} \mathsf{C}_p^k A^k B^{p+1-k}.$$

L'indice de sommation étant muet, nous pouvons regrouper les deux sommes (attention aux termes extrêmes) :

$$(A+B)^{p+1} = \mathsf{C}_p^p A^{p+1} + \sum_{k=1}^{p} \left(\mathsf{C}_p^{k-1} + \mathsf{C}_p^k\right) A^k B^{p+1-k} + \mathsf{C}_p^0 B^{p+1}.$$

Or, $\mathsf{C}_p^p = 1$ et $\mathsf{C}_p^0 = 1$, ce qui donne

$$(A+B)^{p+1} = A^{p+1} + \sum_{k=1}^{p} \left(\mathsf{C}_p^{k-1} + \mathsf{C}_p^k\right) A^k B^{p+1-k} + B^{p+1}.$$

Nous utilisons alors la **formule du triangle de Pascal** $\boxed{\mathsf{C}_p^{k-1} + \mathsf{C}_p^k = \mathsf{C}_{p+1}^k}$ pour obtenir :

$$(A+B)^{p+1} = A^{p+1} + \sum_{k=1}^{p} \mathsf{C}_{p+1}^k A^k B^{p+1-k} + B^{p+1}.$$

Finalement, en faisant entrer les deux termes isolés dans la somme, nous déduisons :

$$(A+B)^{p+1} = \sum_{k=0}^{p+1} \mathsf{C}_{p+1}^k A^k B^{p+1-k}.$$

■

⚠ Il faut noter qu'il est obligatoire que les matrices A et B commutent pour pouvoir utiliser la formule du binôme de Newton. En particulier, ceci est valable si l'une des deux matrices est la matrice I_n.

Traitons l'exemple qui suit comme application de la formule du binôme de Newton :

Exemple 3.4.5. On considère la matrice $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. On va calculer A^p pour tout entier $p \geq 2$.

Remarquons tout d'abord que $A = N + I_3$ avec $N = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

Calculons ensuite les puissances successives de la matrice N.

On a $N^0 = I_3$, $N^1 = N$, $N^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ et $N^3 = O$.

Par suite, $N^k = O$ pour tout entier $k \geq 3$.

Puisque $I_3 N = N I_3$, les matrices N et I_3 commutent et nous pourrons donc utiliser la formule du binôme de Newton pour calculer $A^p = (N + I_3)^p$.

On a donc, compte tenu des résultats obtenus sur les puissances successives de la matrice N :

$$A^p = (N + I_3)^p = \sum_{k=0}^{p} \mathsf{C}_p^k \; N^k \; I_3^{p-k} = \mathsf{C}_p^0 \; N^0 + \mathsf{C}_p^1 \; N + \mathsf{C}_p^2 \; N^2 \; \forall p \geq 2.$$

De plus, comme $\mathsf{C}_p^0 = 1,\ \mathsf{C}_p^1 = p,\ \mathsf{C}_p^2 = \frac{p(p-1)}{2}$, on en déduit que :

$$A^p = I_3 + pN + \frac{p(p-1)}{2} N^2,$$

soit

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

D'où

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

Si nous voulons définir les puissances A^i avec des exposants i négatifs, il faut introduire d'abord la notion de matrice *inversible*. Pour cela, puisque $(\mathcal{M}_n(\mathbb{K}), +, \times)$ est un anneau unitaire, d'élément neutre I_n pour la multiplication, il est intéressant d'étudier les éléments inversibles de cet anneau.

Définition 3.4.6. *On dira qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est* ***inversible*** *si elle est inversible comme élément de l'anneau $(\mathcal{M}_n(\mathbb{K}), +, \times)$, c'est-à-dire, s'il existe une matrice $B \in \mathcal{M}_n(\mathbb{K})$ telle que $AB = BA = I_n$.*
Dans ce cas, la matrice B est unique et appelée l' ***inverse*** *de A, et notée A^{-1}.*

Une matrice inversible doit être nécessairement une matrice carrée.

Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée. Il est intéressant de savoir si elle est inversible, et dans ce cas, nous serons contents de calculer son inverse A^{-1}. Mais comment le faire ? Nous allons voir qu'il y a plusieurs méthodes. Mais pour l'instant, nous pouvons seulement appliquer la définition de l'inversiblité :

Exemple 3.4.7. Pour simplifier, prenons des matrices d'ordre 2 :

- Soit $A = \begin{pmatrix} 1 & 2 \\ 4 & 8 \end{pmatrix}$ et cherchons s'il existe une matrice $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ telle que $AB = I_2$, c'est-à-dire,
$$\begin{pmatrix} 1 & 2 \\ 4 & 8 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
Ceci donne deux systèmes linéaires
$$\left\{ \begin{array}{rcl} a + 2c & = & 1 \\ 4a + 8c & = & 0 \end{array} \right. \qquad \left\{ \begin{array}{rcl} b + 2d & = & 0 \\ 4b + 8d & = & 1 \end{array} \right.$$
Il est clair que chacun des ces deux systèmes est impossible. Donc une telle matrice B n'existe pas. D'où A n'est pas inversible.

- Considérons maintenant la matrice $A' = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ et cherchons s'il existe une matrice $B' = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ telle que $A'B' = I_2$. Les deux systèmes s'écrivent
$$\left\{ \begin{array}{rcl} a - c & = & 1 \\ a + c & = & 0 \end{array} \right. \qquad \left\{ \begin{array}{rcl} b - d & = & 0 \\ b + d & = & 1 \end{array} \right.$$
La résolution de ces deux systèmes fournit $B' = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$. D'où A' est inversible et $(A')^{-1} = B'$.

D'après l'exemple précédent, l'inversiblité d'une matrice d'ordre 2 nécessite la résolution de deux systèmes linéaires à deux inconnues. En général, pour une matrice A carrée d'ordre n, il faut résoudre n systèmes linéaires à n inconnues, et donc c'est trop long ! Par conséquent, cette méthode est à éviter. On donnera d'autres méthodes plus intéressantes.

Soit E un espace vectoriel de dimension n et de base $\mathcal{B}$. En raison de l'isomorphisme φ précédemment établi entre les anneaux $(\mathcal{L}_{\mathbb{K}}(E), +, \circ)$ et $(\mathcal{M}_n(\mathbb{K}), +, \times)$, il est naturel que les éléments inversibles de l'anneau $(\mathcal{M}_n(\mathbb{K}), +, \times)$, c'est-à-dire les matrices inversibles d'ordre n, correspondent aux éléments inversibles de l'anneau $(\mathcal{L}_{\mathbb{K}}(E), +, \circ)$, c'est-à-dire aux automorphismes de E :

Proposition 3.4.8. *Soient E un espace vectoriel de dimension n, $\mathcal{B}$ une base de E, $f \in \mathcal{L}(E)$ un endomorphisme de E et $A = M(f, \mathcal{B})$ la*

matrice de f dans la base $\mathcal{B}$. On a :
La matrice A est inversible si et seulement si l'endomorphisme f est un automorphisme de E.
Dans ce cas, $A^{-1} = M(f^{-1}, \mathcal{B})$.

DÉMONSTRATION. Montrons les deux implications séparément :

$\Longleftarrow$) Si f est un automorphisme de E, on sait que f^{-1} est un automorphisme de E qui vérifie $f \circ f^{-1} = f^{-1} \circ f = \mathtt{id}_E$. Donc

$$M(f \circ f^{-1}, \mathcal{B}) = M(f^{-1} \circ f, \mathcal{B}) = M(\mathtt{id}_E, \mathcal{B}),$$

c'est-à-dire,

$$M(f, \mathcal{B}) \times M(f^{-1}, \mathcal{B}) = M(f^{-1}, \mathcal{B}) \times M(f, \mathcal{B}) = I_n.$$

Par conséquent, en considérant la matrice $B = M(f^{-1}, \mathcal{B}) \in \mathcal{M}_n(\mathbb{K})$, on obtient $AB = BA = I_n$, d'où A est inversible, d'inverse $A^{-1} = B = M(f^{-1}, \mathcal{B})$.

$\Longrightarrow$) Si A est inversible, il existe $B \in \mathcal{M}_n(\mathbb{K})$ telle que $AB = BA = I_n$.
Soit $g \in \mathcal{L}(E)$ l'unique endomorphisme de E associé à la matrice B dans la base $\mathcal{B}$: $M(g, \mathcal{B}) = B$. Alors :

$$AB = M(f, \mathcal{B}) \times M(g, \mathcal{B}) = M(f \circ g, \mathcal{B}),$$

$$BA = M(g, \mathcal{B}) \times M(f, \mathcal{B}) = M(g \circ f, \mathcal{B}).$$

Donc les égalités $AB = BA = I_n$ entraînent

$$M(f \circ g, \mathcal{B}) = M(g \circ f, \mathcal{B}) = M(\mathtt{id}_E, \mathcal{B}).$$

D'après la bijection $\varphi : \mathcal{L}_{\mathbb{K}}(E) \longrightarrow \mathcal{M}_n(\mathbb{K})$, il vient $f \circ g = g \circ f = \mathtt{id}_E$. D'où f est un automorphisme de E, d'inverse $f^{-1} = g$. ■

Remarque 3.4.9. L'ingrédient-clef de la démonstration ci-dessus est l'isomorphisme d'anneaux $\varphi : (\mathcal{L}_{\mathbb{K}}(E), +, \circ) \longrightarrow (\mathcal{M}_n(\mathbb{K}), +, \times)$. En réalité, la proposition précédente est un cas particulier d'un résultat plus général sur les anneaux, que vous pouvez démontrer aisément :
Soient $(\mathcal{A}, +, \times)$ et $(\mathcal{A}', +, \times)$ deux anneaux unitaires et supposons qu'il existe un isomorphisme d'anneaux $T : (\mathcal{A}, +, \times) \longrightarrow (\mathcal{A}', +, \times)$. Soit $b = f(a) \in \mathcal{A}'$ avec $a \in \mathcal{A}$. Alors : b est inversible dans l'anneau $(\mathcal{A}', +, \times)$ si et seulement si a est inversible dans l'anneau $(\mathcal{A}, +, \times)$. Dans ce cas, l'inverse de b dans $\mathcal{A}'$ est l'image par T de l'inverse de a dans $\mathcal{A}$, c'est-à-dire, $b^{-1} = T(a^{-1})$.
Donc l'isomorphisme T envoie les éléments inversibles de l'anneau $(\mathcal{A}, +, \times)$

vers les éléments inversibles de l'anneau $(\mathcal{A}', +, \times)$.
On remarque de plus, que si $\mathcal{U}$ (respectivement $\mathcal{U}'$) désigne le groupe multiplicatif des éléments inversibles de $\mathcal{A}$ (respectivement de $\mathcal{A}'$), alors l'application $a \in \mathcal{U} \longrightarrow T(a) \in \mathcal{U}'$ est un isomorphisme du groupe $(\mathcal{U}, \times)$ vers le groupe $(\mathcal{U}', \times)$.

Remarque 3.4.10. D'après la proposition 3.4.8, une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est inversible si et seulement si l'endomorphisme f de $E = \mathbb{K}^n$ canoniquement associé à A est bijective. Mais selon le corollaire 2.5.11 du chapitre 2, il suffit que f soit injective, ou encore que $\ker(f) = \{0_E\}$. Par conséquent, A est inversible si et seulement si :

$$\forall X \in \mathcal{M}_{n,1}(\mathbb{K}),\ AX = 0 \Longrightarrow X = 0$$

Nous verrons plus loin d'autres caractérisations de l'inversibilité d'une matrice carrée faisant intervenir le rang.

La proposition 3.4.8 nous fournit une méthode pour savoir si une matrice est inversible, et aussi pour calculer son inverse dans le cas échéant :

Exemple 3.4.11. Retrouvons les résultats de l'exemple 3.4.7 à la lumière de la proposition 3.4.8 :
Soit l'espace vectoriel $E = \mathbb{R}^2$ muni de sa base canonique $\mathcal{B} = (e_1, e_2)$.

On sait qu'il existe un endomorphisme unique f de E tel que $M(f, \mathcal{B}) = A = \begin{pmatrix} 1 & 2 \\ 4 & 8 \end{pmatrix}$. Donc

$$f(e_1) = e_1 + 4e_2, \quad f(e_2) = 2e_1 + 8e_2.$$

Comme $f(e_2) = 2f(e_1)$, alors f transforme la base $\mathcal{B} = (e_1, e_2)$ de E en une famille liée $(f(e_1), f(e_2))$. D'où f n'est pas un automorphisme, et par suite, la matrice A n'est pas inversible.

Pour la matrice $A' = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$, il existe un endomorphisme unique g de E tel que $M(f, \mathcal{B}) = A'$. On a ainsi

$$g(e_1) = e_1 + e_2, \quad g(e_2) = -e_1 + e_2.$$

Puisque les vecteurs $g(e_1)$ et $g(e_2)$ ne sont pas colinéaires, ils sont linéairement indépendants, et forment donc une base $\mathcal{B}'$ de E. Il en découle que g est un automorphisme de E, ce qui entraîne que la matrice A' est inversible.
Pour calculer son inverse $(A')^{-1}$, on applique la formule $(A')^{-1} =$

$M(g^{-1}, \mathcal{B})$. Ceci revient à exprimer $g^{-1}(e_1)$ et $g^{-1}(e_2)$ en fonction de e_1, e_2. Pour cela, écrivons

$$\begin{cases} g(e_1) &= e_1 + e_2 \\ g(e_2) &= -e_1 + e_2 \end{cases}$$

La somme de ces deux équations donne $g(e_1) + g(e_2) = 2e_2$. Donc $e_2 = \frac{1}{2}(g(e_1) + g(e_2)) = g(\frac{1}{2}e_1 + \frac{1}{2}e_2)$. D'où

$$g^{-1}(e_2) = \frac{1}{2}e_1 + \frac{1}{2}e_2.$$

La première équation entraîne $e_1 = g(e_1) - e_2 = g(e_1) - g(\frac{1}{2}e_1 + \frac{1}{2}e_2) = g(e_1 - \frac{1}{2}e_1 - \frac{1}{2}e_2) = g(\frac{1}{2}e_1 - \frac{1}{2}e_2)$. D'où

$$g^{-1}(e_1) = \frac{1}{2}e_1 - \frac{1}{2}e_2.$$

Finalement, en écrivant verticalement les coordonnées de $g^{-1}(e_1)$ et $g^{-1}(e_2)$ dans la base $\mathcal{B}$, on retrouve

$$(A')^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

.

En revenant à notre isomorphisme d'anneaux $\varphi : (\mathcal{L}_{\mathbb{K}}(E), +, \circ) \longrightarrow (\mathcal{M}_n(\mathbb{K}), +, \times)$, où E un espace vectoriel de dimension n et $\mathcal{B}$ est une base de E, on peut énoncer :

Proposition 3.4.12. **(a)** *L'ensemble $GL_n(\mathbb{K})$ des matrices inversibles de $\mathcal{M}_n(\mathbb{K})$ est un groupe multiplicatif qu'on appelle* ***groupe linéaire d'ordre*** n ***du corps*** $\mathbb{K}$. .

(b) *Soient E un espace vectoriel de dimension n, $\mathcal{B}$ une base de E et $GL_{\mathbb{K}}(E)$ le groupe linéaire de E. Alors l'application $f \in GL_{\mathbb{K}}(E) \mapsto M(f, \mathcal{B}) \in GL_n(\mathbb{K})$ est un isomorphisme du groupe $(GL_{\mathbb{K}}(E), \circ)$ vers le groupe $(GL_n(\mathbb{K}), \times)$.*

Démonstration. Elle se déduit immédiatement de la dernière observation de la remarque précédente. Mais nous pouvons facilement fournir une démonstration directe :

(a) $(GL_n(\mathbb{K}), \times)$ est un groupe car :

(b) $I_n \in GL_n(\mathbb{K})$ puisque I_n est inversible ($I_n I_n = I_n$).

(c) Pour toutes matrices A et B dans $GL_n(\mathbb{K})$,

$$(AB)(B^{-1}A^{-1}) = I_n, \qquad (B^{-1}A^{-1})(AB) = I_n,$$

donc AB est inversible, d'inverse $(AB)^{-1} = B^{-1}A^{-1}$. D'où $AB \in GL_n(\mathbb{K})$.

(d) Pour toute matrice A dans $GL_n(\mathbb{K})$, $A^{-1}A = AA^{-1} = I_n$, donc $A^{-1} \in GL_n(\mathbb{K})$.

(e) La multiplication des matrices est *a fortiori* associative dans le sous-ensemble $GL_n(\mathbb{K})$ de $\mathcal{M}_n(\mathbb{K})$.

(f) L'application $f \mapsto M(f, \mathcal{B})$ est un homomorphisme du groupe $(GL_{\mathbb{K}}(E), \circ)$ vers le groupe $(GL_n(\mathbb{K}), \times)$, puisque $M(g \circ f, \mathcal{B}) = M(g, \mathcal{B}) \times M(f, \mathcal{B})$. De plus, elle est bijective d'après la proposition (3.4.8). Les groupes linéaires $(GL_{\mathbb{K}}(E), \circ)$ et $(GL_n(\mathbb{K}), \times)$ sont donc isomorphes. ∎

Remarque 3.4.13. **(a)** Le groupe linéaire $GL_n(\mathbb{K})$ est non commutatif dès que $n \geq 2$.

(b) Si A et B sont deux matrices inversibles d'ordre n, alors AB l'est aussi. De plus, $(AB)^{-1} = B^{-1}A^{-1}$.
Ce résultat se généralise à un nombre quelconque de matrices inversibles $A_1, \ldots, A_k$ d'ordre n : $(A_1 \ldots A_k)^{-1} = A_k^{-1} \ldots A_1^{-1}$. C'est une propriété des éléments inversibles de tout anneau unitaire.

(c) Si A est inversible, on définit ses puissances avec des exposants négatifs en posant pour tout $i > 0$:

$$A^{-i} = \left(A^{-1}\right)^i = \underbrace{A^{-1}A^{-1} \times A^{-1}}_{i fois}$$

Il n'est pas difficile d'observer que pour tout $i = 0, 1, 2, \ldots$, la matrice A^i est aussi inversible et

$$\left(A^i\right)^{-1} = A^{-i} = \left(A^{-1}\right)^i \tag{3.10}$$

En effet, en utilisant de proche en proche l'égalité $AA^{-1} = I_n$, on obtient

$$A^i \left(A^{-1}\right)^i = \left(\underbrace{AA \times A}_{i fois}\right)\left(\underbrace{A^{-1}A^{-1} \times A^{-1}}_{i fois}\right) = I_n$$

On peut également établir 3.10 à l'aide d'une simple récurrence sur l'entier $i \geq 0$.

Exemple 3.4.14. Soit $A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$. De l'exemple 3.4.7, on sait que A est inversible et $A^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$.

Donc

$$A^{-2} = \left(A^{-1}\right)^2 = \frac{1}{4}\begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix},$$

$$A^{-3} = \left(A^{-1}\right)^3 = \left(A^{-1}\right)^2 A^{-1} = \frac{1}{8}\begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

$$= \frac{1}{8}\begin{pmatrix} -2 & 2 \\ -2 & -2 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix}$$

En général, dans un anneau unitaire non commutatif $(\mathcal{A}, +, \times)$, il n'est pas toujours vrai que $ab = 1_{\mathcal{A}}$ implique $ba = 1_{\mathcal{A}}$. Pour montrer qu'un élément $a \in \mathcal{A}$ est inversible, il faut qu'il existe $b \in \mathcal{A}$ vérifiant les deux conditions. Cependant, nous allons voir que l'une de ces deux conditions suffit si nous nous plaçons dans l'anneau $\mathcal{M}_n(\mathbb{K})$ des matrices carrées d'ordre n, ou dans l'anneau $(\mathcal{L}_{\mathbb{K}}(E), +, \times)$ des endomorphismes d'un espace vectoriel E de dimension finie :

Théorème 3.4.15. *Soit E un espace vectoriel de dimension finie n sur $\mathbb{K}$ et $f, g \in \mathcal{L}(E)$ deux endomorphismes de E.*
Si $g \circ f = \mathtt{id}_E$ alors f et g sont bijectives, $f \circ g = \mathtt{id}_E$ et $f^{-1} = g$.

Démonstration.

Pour montrer que f est bijective, commençons par établir qu'elle est injective. Pour tous x et x' dans E, si $f(x) = f(x')$, alors $g(f(x)) = g(f(x'))$, c'est-à-dire $(g \circ f)(x) = (g \circ f)(x')$, d'où $x = x'$. Ainsi, f est injective.
D'après le théorème 2.5.11 du chapitre 2, l'endomorphisme injectif f en dimension finie est nécessairement bijectif ; c'est donc un automorphisme de E.

Pour montrer que g est bijective, prouvons d'abord qu'elle est surjective. Pour tout $y \in E$, on a

$$y = \mathtt{id}_E(y) = (g \circ f)(y) = g(f(y)) = g(x) \quad \text{avec } x = f(y) \in E.$$

D'où g est surjective. En appliquant le même théorème 2.5.11 du chapitre 2, on déduit que g est bijective.

Montrons maintenant que $f \circ g = \mathtt{id}_E$. Considérons l'automorphisme réciproque f^{-1} de f. En composant à droite par f^{-1} les deux membres de l'égalité $g \circ f = \mathtt{id}_E$, on obtient $(g \circ f) \circ f^{-1} = \mathtt{id}_E \circ f^{-1}$, donc $g \circ (f \circ f^{-1}) = f^{-1}$, ce qui donne $g = f^{-1}$.
Or, on sait que $f \circ f^{-1} = \mathtt{id}_E$, d'où $f \circ g = \mathtt{id}_E$. ■

Remarque 3.4.16. Le théorème précédent n'est plus valable si l'espace vectoriel E n'est pas de dimension finie. En effet, soit, par exemple, $E = \mathbb{R}[X]$ et f, g les endomorphismes de E définies par

$$f(P) = P' = a_1 + 2a_2X + \cdots + na_nX^{n-1},$$

$$g(P) = \int_0^X P(t)dt = a_0X + \frac{a_1}{2}X^2 + \cdots + \frac{a_n}{n+1}X^{n+1}$$

pour tout polynôme $P = a_0 + a_1X + \cdots + a_nX^n \in \mathbb{R}[X]$ (voir exemple 2.1.11, chapitre 2).
On vérifie immédiatement que $f \circ g = \mathtt{id}_E$. Cependant, $g \circ f \neq \mathtt{id}_E$, puisque $(g \circ f)(1) = g(f(1)) = g(0) = 0 \neq 1$.
Donc g est inversible à gauche dans l'anneau $\mathcal{L}(E)$ mais pas à droite, et f est inversible à droite, mais pas à gauche.
Cela vient du fait que f est surjective mais pas pas injective ($\ker(f) = \mathbb{R}$ est l'ensemble des polynômes constants) tandis que g est injective mais pas surjective (1 n'a pas d'antécédent par g).

On peut énoncer les conséquences immédiates du théorème précédent :

Corollaire 3.4.17. *Soit E un espace vectoriel de dimension finie n sur $\mathbb{K}$ et $f \in \mathcal{L}(E)$ un endomorphisme de E. Les conditions suivantes sont équivalentes :*

(a) *f est bijectif.*

(b) *f est inversible à gauche pour la loi $\circ$:* $\exists g \in \mathcal{L}(E)$ tel que $g \circ f = \mathtt{id}_E$.

(c) *f est inversible à droite pour la loi $\circ$:* $\exists h \in \mathcal{L}(E)$ tel que $f \circ h = \mathtt{id}_E$.

Dans ce cas, $g = h = f^{-1}$.

Démonstration. Il est clair que $1 \Longrightarrow 2$ et $1 \Longrightarrow 3$, car si f est bijective, sa bijection réciproque f^{-1} est un endomorphisme de E qui vérifie $f^{-1} \circ f = f \circ f^{-1} = \mathtt{id}_E$.
Les implications $2 \Longrightarrow 1$ et $3 \Longrightarrow 1$ sont conséquences du lemme précédent. ∎

Ce corollaire admet la version matricielle suivante :

Corollaire 3.4.18. *Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée d'ordre n. Les conditions suivantes sont équivalentes :*

(a) *A est inversible.*

(b) *A est inversible à gauche :* $\exists\ B \in \mathcal{M}_n(\mathbb{K})$ tel que $BA = I_n$.

(c) *A est inversible à droite :* $\exists\ C \in \mathcal{M}_n(\mathbb{K})$ tel que $AC = I_n$.

Dans ce cas, $B = C = A^{-1}$.

Démonstration. Les implications $1 \Longrightarrow 2$ et $1 \Longrightarrow 3$ sont évidentes.
Montrons maintenant :

$2 \Longrightarrow 1)$ Supposons qu'il existe $B \in \mathcal{M}_n(\mathbb{K})$ tel que $BA = I_n$.
Soit l'espace vectoriel $E = \mathbb{K}^n$ équipé de la base canonique $\mathcal{B}$. Soient f et g les endomorphismes de E associés respectivement aux matrices A et B dans la base $\mathcal{B}$, c'est-à-dire : $A = M(f, \mathcal{B})$ et $B = M(g, \mathcal{B})$. On a

$$M(\mathtt{id}_E, \mathcal{B}) = I_n = BA = M(g, \mathcal{B}) \times M(f, \mathcal{B}) = M(g \circ f, \mathcal{B}),$$

d'où $\mathtt{id}_E = g \circ f$. D'après le théorème 3.4.15, on a aussi $\mathtt{id}_E = f \circ g$. On en déduit :

$$I_n = M(\mathtt{id}_E, \mathcal{B}) = M(f \circ g, \mathcal{B}) = M(f, \mathcal{B}) \times M(g, \mathcal{B}) = AB.$$

Il s'ensuit que $BA = AB = I_n$, et par suite A est inversible et $B = A^{-1}$.

Pour l'implication $3 \Longrightarrow 1)$, on applique un raisonnement similaire :
Supposons qu'il existe $C \in \mathcal{M}_n(\mathbb{K})$ tel que $AC = I_n$. Soit $h \in \mathcal{L}(E)$ l'endomorphisme de E tel que $C = M(h, \mathcal{B})$. On a

$$M(\mathtt{id}_E, \mathcal{B}) = I_n = AC = M(f, \mathcal{B}) \times M(h, \mathcal{B}) = M(f \circ h, \mathcal{B}),$$

d'où $\mathtt{id}_E = f \circ h$. D'après le théorème 3.4.15, on a aussi $\mathtt{id}_E = h \circ f$. On en déduit :

$$I_n = M(\mathtt{id}_E, \mathcal{B}) = M(h \circ f, \mathcal{B}) = M(h, \mathcal{B}) \times M(f, \mathcal{B}) = CA.$$

Il résulte que $AC = CA = I_n$, d'où A est inversible et $C = A^{-1}$. ■

$\boxed{\textbf{Calcul pratique de } A^{-1}}$:
Une autre méthode pour calculer l'inverse A^{-1} d'une matrice inversible $A \in \mathcal{M}_n(\mathbb{K})$ consiste à utiliser l'égalité $Y = AX$ établie dans la proposition 3.3.15, où $X, Y \in \mathcal{M}_{n,1}(\mathbb{K})$. En effet, si f est l'endomorphisme de $E = \mathbb{K}^n$ canoniquement associé à la matrice A, alors A est inversible si et seulement si f est bijectif si et seulement si $\forall Y \in \mathcal{M}_{n,1}(\mathbb{K}),\ \exists\ X \in \mathcal{M}_{n,1}(\mathbb{K}) : Y = AX$. Dans ce cas, on résout un système linéaire pour exprimer X en fonction de Y, car si A est inversible, alors : $Y = AX \iff X = A^{-1}Y$.
Voyons ceci sur l'exemple suivant :

Exemple 3.4.19. Montrons que la matrice $A = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 0 & 3 \\ -2 & 3 & 0 \end{pmatrix}$ est inversible et calculons A^{-1}.

En notant $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$, on a :

$$Y = AX \iff \left\{ \begin{array}{lcll} x_1 + x_2 + 2x_3 & = & y_1 & (1) \\ 3x_1 \qquad + 3x_3 & = & y_2 & (2) \\ -2x_1 + 3x_2 & = & y_3 & (3) \end{array} \right.$$

Pour que A soit inversible, il faut que ce système admet toujours une et une seule solution $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{K})$ quelque soit $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{K})$. Pour résoudre ce système, les équations (3) et (2) donnent respectivement :

$$x_2 = \frac{1}{3}(y_3 + 2x_1), \quad x_3 = \frac{1}{3}(y_2 - 3x_1). \tag{3.11}$$

En reportant ces valeurs dans (1), on obtient :

$$x_1 + \frac{1}{3}(y_3 + 2x_1) + \frac{2}{3}(y_2 - 3x_1) = y_1,$$

d'où

$$\boxed{x_1 = -3y_1 + 2y_2 + y_3}$$

Par conséquent, (3.11) implique :

$$\boxed{x_2 = -2y_1 + \frac{4}{3}y_2 + y_3, \quad x_3 = 3y_1 - \frac{5}{3}y_2 - y_3}$$

Il en découle que A est inversible. De plus, ces résultats s'écrivent matriciellement :

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -3 & 2 & 1 \\ -2 & \frac{4}{3} & 1 \\ 3 & -\frac{5}{3} & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

D'où $A^{-1} = \begin{pmatrix} -3 & 2 & 1 \\ -2 & \frac{4}{3} & 1 \\ 3 & -\frac{5}{3} & -1 \end{pmatrix}$.

Attention ! L'erreur est humaine et tout le monde peut se tromper dans les calculs. Pour être sûr des calculs, vérifier toujours l'égalité $AA^{-1} = I_n$. Ici, on vérifie bien que

$$\begin{pmatrix} 1 & 1 & 2 \\ 3 & 0 & 3 \\ -2 & 3 & 0 \end{pmatrix} \begin{pmatrix} -3 & 2 & 1 \\ -2 & \frac{4}{3} & 1 \\ 3 & -\frac{5}{3} & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Exemple 3.4.20. Pour étudier l'inversiblité de la matrice $A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$, on résout le système $\left\{ \begin{array}{rcl} x_2 + x_3 + x_4 & = & y_1 \\ x_1 \quad + x_3 + x_4 & = & y_2 \\ x_1 + x_2 \quad + x_4 & = & y_3 \\ x_1 + x_2 + x_3 \quad & = & y_4 \end{array} \right.$

Après calculs, on trouve :

$$\left\{ \begin{array}{rcl} x_1 & = & \frac{1}{3}(-2y_1 + y_2 + y_3 + y_4) \\ x_2 & = & \frac{1}{3}(y_1 - 2y_2 + y_3 + y_4) \\ x_3 & = & \frac{1}{3}(y_1 + y_2 - 2y_3 + y_4) \\ x_4 & = & \frac{1}{3}(y_1 + y_2 + y_3 - 2y_4) \end{array} \right.$$

D'où

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

Calculer l'inverse d'une matrice est une tâche ardue à la main dès lors qu'on aborde les matrices 3×3, et la difficulté croît avec la taille

Le calcul de l'inverse d'une matrice inversible est l'un des problèmes importants de l'algèbre linéaire. Les méthodes classiques se ramènent en général à résoudre un ensemble de systèmes linéaires. On verra au dernier chapitre une méthode pratique pour résoudre les systèmes linéaires, appelée *méthode de Gauss*.
Précisons que pour des matrices de grande taille, il y a des logiciels de calcul d'inverse des matrices inversibles. A l'heure actuelle où les ordinateurs sont de plus en plus puissants, il est bénéfique d'utiliser des programmes informatiques qui calculent l'inverse d'une matrice : MAPLE, MATHEMATICA, MATLAB, etc. Néanmoins, la taille de la matrice est limitée par la taille de la mémoire de l'ordinateur !
Les performances d'un algorithme se mesurent par sa vitesse de calcul.

Il existe des matrices particulières qui sont très faciles à manipuler (produit, inverse, etc.). C'est le cas par exemple des matrices dites diagonales :

Définition 3.4.21. *On appelle* ***matrice diagonale*** *une matrice carrée $D = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ dont les cœfficients non situés sur la diagonale sont nuls : $a_{ij} \neq 0$ pour $i \neq j$.*

Elle s'écrit donc $D = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$.

En posant $\lambda_i = a_{ii}$ pour $1 \leq i \leq n$, on note $D = \mathsf{diag}(\lambda_1, \dots, \lambda_n)$. En particulier, la matrice unité d'ordre n est diagonale : $I_n = \mathsf{diag}(1, \dots, 1)$.

Remarquons que le produit de deux matrices diagonales est encore une matrice diagonale :

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}\begin{pmatrix} \gamma_1 & 0 & \dots & 0 \\ 0 & \gamma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \gamma_n \end{pmatrix} = \begin{pmatrix} \lambda_1\gamma_1 & 0 & \dots & 0 \\ 0 & \lambda_2\gamma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n\gamma_n \end{pmatrix},$$

c'est-à-dire,

$$\boxed{\mathsf{diag}(\lambda_1,\ldots,\lambda_n)\times\mathsf{diag}(\gamma_1,\ldots,\gamma_n)=\mathsf{diag}(\lambda_1\gamma_1,\ldots,\lambda_n\gamma_n)}$$

On en déduit que

$$\boxed{\Big(\mathsf{diag}(\lambda_1,\ldots,\lambda_n)\Big)^k=\mathsf{diag}(\lambda_1^k,\ldots,\lambda_n^k)\qquad\forall\, k\geq 1}$$

Notons aussi que si $\lambda_1\neq 0,\ldots,\lambda_n\neq 0$, alors

$$\mathsf{diag}(\lambda_1,\ldots,\lambda_n)\times\mathsf{diag}\left(\lambda_1^{-1},\ldots,\lambda_n^{-1}\right)=\mathsf{diag}(1,\ldots,1)=I_n,$$

donc la matrice $\mathsf{diag}(\lambda_1,\ldots,\lambda_n)$ est inversible et son inverse est

$$\boxed{\Big(\mathsf{diag}(\lambda_1,\ldots,\lambda_n)\Big)^{-1}=\mathsf{diag}\left(\lambda_1^{-1},\ldots,\lambda_n^{-1}\right)}$$

3.5 Changement de bases

Pour travailler dans un espace vectoriel E de dimension finie sur $\mathbb{K}$, on utilise souvent une base et les coordonnées des vecteurs dans cette base. Mais puisqu'il y a plusieurs bases de E, il y a des moments où l'on souhaite changer de base pour faciliter certains calculs. Dans ce cas, les coordonnées d'un vecteur vont changer d'une base à une autre. De même, puisque la matrice d'une application linéaire dépend des bases choisies, cette matrice change également quand on change les bases.
Dans cette section, nous étudions ce phénomène de changement de bases et ses effets sur les coordonnées d'un vecteur et sur la matrice d'une application linéaire.

3.5.1 Matrice de passage

Soit E un espace vectoriel de dimension finie n sur $\mathbb{K}$ et soient $\mathcal{B}=(e_1,\ldots,e_n)$ et $\mathcal{B}'=(e'_1,\ldots,e'_n)$ deux bases de E. Puisque $\mathcal{B}$ est une base de E, chaque vecteur e'_j se décompose de façon unique comme combinaison linéaire des vecteurs $e_1,\ldots,e_n$ de $\mathcal{B}$:

$$\left\{\begin{array}{lcl} e'_1 &=& a_{11}\ e_1+a_{21}e_2+\cdots+a_{n1}e_n=\sum\limits_{i=1}^n a_{i1}e_i\\ e'_2 &=& a_{12}e_1+a_{22}e_2+\cdots+a_{n2}e_n=\sum\limits_{i=1}^n a_{i2}e_i\\ \vdots & \vdots & \qquad\qquad\vdots\\ e'_n &=& a_{1n}e_1+a_{2n}e_2+\cdots+a_{nn}e_n=\sum\limits_{i=1}^n a_{in}e_i \end{array}\right. \tag{3.12}$$

avec les a_{ij} dans $\mathbb{K}$ pour $i, j \in \{1, \dots, n\}$. C'est-à-dire,

$$e'_j = a_{1j}e_1 + a_{2j}e_2 + \cdots + a_{nj}e_n = \sum_{i=1}^{n} a_{ij}e_i \qquad \text{pour tout } j = 1, \dots, n.$$

Les n^2 cœfficients a_{ij} vont former une matrice qui jouera un rôle central dans le procédé de changement de bases :

Définition 3.5.1. *On appelle **matrice de passage** (ou **matrice de changement de base**) de la base $\mathcal{B}$ à la base $\mathcal{B}'$ la matrice P dont les colonnes sont formées par les coordonnées des vecteurs $e'_1, \dots, e'_n$ dans la base $\mathcal{B}$:*

$$P = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

*Dans ce cas, on dit que $\mathcal{B}$ est l'**ancienne base** de E et que $\mathcal{B}'$ est la **nouvelle base**.*

Cette matrice de passage P **se remplit par colonnes**, chaque colonne j représente le vecteur e'_j de la nouvelle base $\mathcal{B}'$ exprimé par ses coordonnées dans l'ancienne base $\mathcal{B}$.

Exemple 3.5.2. Soit E un espace vectoriel de dimension 3 et $\mathcal{B} = (e_1, e_2, e_3)$ une base de E. Considérons les trois vecteurs e'_1, e'_2, e'_3 de E donnés par

$$e'_1 = e_1 - e_3, \quad e'_2 = e_2 + e_3, \quad e'_3 = e_1 + e_3.$$

On vérifie par des calculs simples que la famille $\mathcal{B}' = (e'_1, e'_2, e'_3)$ est libre, et donc c'est une base de E. La matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$ est

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

Remarque 3.5.3. Soit l'endomorphisme identité $\mathtt{id}_E : E \longrightarrow E, x \mapsto x$. On rappelle que $M(\mathtt{id}_E, \mathcal{B}) = M(\mathtt{id}_E, \mathcal{B}, \mathcal{B}) = I_n$ et $M(\mathtt{id}_E, \mathcal{B}') = M(\mathtt{id}_E, \mathcal{B}', \mathcal{B}') = I_n$.
Munissons maintenant l'espace de départ E de la base $\mathcal{B}'$ et l'espace d'arrivée E de la base $\mathcal{B}$, que nous schématisons par $\mathtt{id}_E : (E, \mathcal{B}') \longrightarrow (E, \mathcal{B})$. Alors la matrice $M(\mathtt{id}_E, \mathcal{B}', \mathcal{B})$ a pour colonnes les coordonnées des vecteurs $\mathtt{id}_E(e'_1) = e'_1, \dots, \mathtt{id}_E(e'_n) = e'_n$ dans la base $\mathcal{B}$. Donc cette matrice $M(\mathtt{id}_E, \mathcal{B}', \mathcal{B})$ est identique à P :

$$\boxed{P = M(\mathtt{id}_E, \mathcal{B}', \mathcal{B})}$$

Mais attention à l'ordre des bases !

Proposition 3.5.4. *La matrice de passage P de la base $\mathcal{B}$ à la base $\mathcal{B}'$ est inversible, et son inverse P^{-1} est égale à la matrice de passage de la base $\mathcal{B}'$ à la base $\mathcal{B}$.*

Démonstration. D'après le théorème 3.3.6, on a

$$M(\mathtt{id}_E, \mathcal{B}', \mathcal{B}) \times M(\mathtt{id}_E, \mathcal{B}, \mathcal{B}') = M(\mathtt{id}_E \circ \mathtt{id}_E, \mathcal{B}, \mathcal{B}) = M(\mathtt{id}_E, \mathcal{B}) = I_n$$

On peut vérifier aussi que

$$M(\mathtt{id}_E, \mathcal{B}, \mathcal{B}') \times M(\mathtt{id}_E, \mathcal{B}', \mathcal{B}) = M(\mathtt{id}_E \circ \mathtt{id}_E, \mathcal{B}', \mathcal{B}') = M(\mathtt{id}_E, \mathcal{B}') = I_n,$$

mais ceci n'était pas nécessaire en vertu du corollaire 3.4.18. Donc la matrice $P = M(\mathtt{id}_E, \mathcal{B}', \mathcal{B})$ est inversible et son inverse est $P^{-1} = M(\mathtt{id}_E, \mathcal{B}, \mathcal{B}')$. Or, la remarque ci-dessus dit que $M(\mathtt{id}_E, \mathcal{B}, \mathcal{B}')$ est la matrice de passage de la base $\mathcal{B}'$ à la base $\mathcal{B}$. D'où le résultat désiré. ■

Pour ne pas se tromper entre les matrices P et P^{-1}, on insiste que la matrice de passage P contient en colonnes les coordonnées des vecteurs de la nouvelle base $\mathcal{B}'$ exprimées dans l'ancienne base $\mathcal{B}$. La matrice P^{-1} admet comme colonnes les coordonnées des vecteurs de l'ancienne base $\mathcal{B}$ exprimées dans la nouvelle base $\mathcal{B}'$.

Exemple 3.5.5. Retournons à l'exemple 3.5.2. On déduit de la proposition précédente que la matrice $P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix}$ est inversible et que son inverse P^{-1} est la matrice de passage de la base $\mathcal{B}' = (e'_1, e'_2, e'_3)$ à la base $\mathcal{B} = (e_1, e_2, e_3)$. Si nous voulons donc calculer P^{-1} de cette manière, il suffit de trouver les coordonnées des vecteurs e_1, e_2, e_3 dans la base $\mathcal{B}'$. Pour cela, considérons le système

$$\begin{cases} e'_1 &= e_1 - e_3 & (1) \\ e'_2 &= e_2 + e_3 & (2) \\ e'_3 &= e_1 + e_3 & (3) \end{cases}$$

où les inconnues sont les vecteurs e_1, e_2, e_3 qu'il faut calculer en fonction des vecteurs e'_1, e'_2, e'_3. Pour résoudre ce système vectoriel, on peut se permettre d'additionner les équations et de multiplier une équation par un scalaire. Par exemple, $(1) + (3)$ donne $e'_1 + e'_3 = 2e_1$, d'où $e_1 = \frac{1}{2}e'_1 + \frac{1}{2}e'_3$. Ensuite, en faisant $(3) - (1)$, on obtient $e_3 = -\frac{1}{2}e'_1 + \frac{1}{2}e'_3$. Finalement, l'équation (2) implique $e_2 = e'_2 - e_3$, d'où $e_2 = \frac{1}{2}e'_1 + e'_2 - \frac{1}{2}e'_3$. Par conséquent,

$$P^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Remarque 3.5.6. Il est évident que la matrice de passage de la base $\mathcal{B}$ à la même base $\mathcal{B}$ est la matrice unité I_n.
D'autre part, si P est la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$ et Q est la matrice de passage de la base $\mathcal{B}'$ à la base $\mathcal{B}''$, alors la matrice de passage R de la base $\mathcal{B}$ à la base $\mathcal{B}''$ vaut PQ. En effet, $P = M(\mathtt{id}_E, \mathcal{B}', \mathcal{B})$, $Q = M(\mathtt{id}_E, \mathcal{B}'', \mathcal{B}')$, $R = M(\mathtt{id}_E, \mathcal{B}'', \mathcal{B})$, d'où

$$\begin{aligned} PQ = M(\mathtt{id}_E, \mathcal{B}', \mathcal{B}) \times M(\mathtt{id}_E, \mathcal{B}'', \mathcal{B}') &= M(\mathtt{id}_E \circ \mathtt{id}_E, \mathcal{B}'', \mathcal{B}) \\ &= M(\mathtt{id}_E, \mathcal{B}'', \mathcal{B}) = R. \end{aligned}$$

3.5.2 Changement de coordonnées d'un vecteur

On se pose ici le problème suivant : connaissant les coordonnées $x_1, \ldots, x_n$ d'un vecteur $x \in E$ dans la base $\mathcal{B} = (e_1, \ldots, e_n)$, comment trouver ses coordonnées $x'_1, \ldots, x'_n$ dans la base $\mathcal{B}' = (e'_1, \ldots, e'_n)$?
En utilisant les relations 3.12, on a

$$\begin{aligned} x &= x_1 e_1 + x_2 e_2 + \cdots + x_n e_n \\ &= x'_1 e'_1 + x'_2 e'_2 + \cdots + x'_n e'_n \\ &= x'_1(a_{11} e_1 + \cdots + a_{n1} e_n) + x'_2(a_{12} e_1 + \cdots + a_{n2} e_n) + \cdots + x'_n(a_{1n} e_1 + \cdots + a_{nn} e_n) \\ &= (a_{11} x'_1 + a_{12} x'_2 + \cdots + a_{1n} x'_n) e_1 + \cdots + (a_{n1} x'_1 + a_{n2} x'_2 + \cdots + a_{nn} x'_n) e_n \end{aligned}$$

On en déduit :

$$\left\{ \begin{array}{lll} x_1 & = & a_{11} x'_1 + a_{12} x'_2 + \cdots + a_{1n} x'_n = \sum_{j=1}^{n} a_{1j} x'_j \\ x_2 & = & a_{21} x'_1 + a_{22} x'_2 + \cdots + a_{2n} x'_n = \sum_{j=1}^{n} a_{2j} x'_j \\ \vdots & \vdots & \qquad\qquad \vdots \\ x_n & = & a_{n1} x'_1 + a_{n2} x'_2 + \cdots + a_{nn} x'_n = \sum_{j=1}^{n} a_{nj} x'_j \end{array} \right. \tag{3.13}$$

ou encore

$$x_i = \sum_{j=1}^{n} a_{ij} x'_j \qquad j = 1, \ldots, n$$

le système (3.13) s'écrit matriciellement

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}$$

Par conséquent, en posant

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{et} \qquad X' = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix},$$

on obtient $X = PX'$. On peut donc énoncer :

Proposition 3.5.7. *La matrice colonne X des anciennes coordonnées d'un vecteur $x \in E$ est égale au produit de la matrice de passage P par la matrice colonne X' des nouvelles coordonnées de x :*

$$\boxed{X = PX'}$$

Remarque 3.5.8. Nous pouvons montrer beaucoup plus rapidement la formule $X = PX'$ en utilisant l'égalité $P = M(\mathtt{id}_E, \mathcal{B}', \mathcal{B})$.
En effet, Rappelons la formule $Y = AX$ de la proposition 3.3.15 qui traduit matriciellement l'égalité $y = f(x)$, où f est une application linéaire. Or, ici l'application linéaire f est l'endomorphisme identité $\mathtt{id}_E : (E, \mathcal{B}') \longrightarrow (E, \mathcal{B})$, et donc la matrice A n'est autre que la matrice $M(\mathtt{id}_E, \mathcal{B}', \mathcal{B}) = P$. Par conséquent, l'égalité $Y = AX$ s'écrit ici $X = PX'$.

Exemple 3.5.9. En revenant à l'exemple 3.5.2, pour tout vecteur $x = x_1e_1 + x_2e_2 + x_3e_3 = x'_1e'_1 + x'_2e'_2 + x'_3e'_3 \in E$, la relation $X = PX'$ s'écrit

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}$$

D'où

$$\begin{cases} x_1 &= & x'_1 + x'_3 \\ x_2 &= & x'_2 \\ x_3 &= & -x'_1 + x'_2 + x'_3 \end{cases}$$

Remarque 3.5.10. Dans un changement de bases pour un vecteur, la formule $X = PX'$ exprime donc naturellement les *anciennes coordonnées* (coordonnées de x dans la base $\mathcal{B}$) en fonction des *nouvelles coordonnées* (coordonnées de x dans la base $\mathcal{B}'$). Si l'on désire exprimer les nouvelles coordonnées de x en fonction des anciennes coordonnées de x, il suffit de multiplier à gauche par P^{-1} l'égalité $X = PX'$, ce qui donne $P^{-1}X = P^{-1}(PX') = (P^{-1}P)X' = I_nX' = X'$. D'où la formule

$$\boxed{X' = P^{-1}X}$$

Notons que cette dernière formule nécessite le calcul de l'inverse P^{-1} de la matrice P.

Exemple 3.5.11. En reprenant les données de l'exemple précédent, puisqu'on a déjà calculé $P^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$, alors la relation $X' =$

$P^{-1}X$ donne

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

D'où

$$\begin{cases} x'_1 &= \frac{1}{2}(x_1 + x_2 - x'_3) \\ x'_2 &= x_2 \\ x'_3 &= \frac{1}{2}(x_1 - x_2 + x_3) \end{cases}$$

Rappelons que le calcul de l'inverse d'une matrice est équivalent à la résolution d'un système linéaire. Par conséquent, une autre manière de calculer l'inverse P^{-1} de P est de résoudre l'équation $X = PX'$ pour exprimer X en fonction de X'. C'est la même chose qu'on a fait dans la section précédente dans l'équation $Y = AX$ équivalente à $X = A^{-1}Y$. Expliquons ceci sur un exemple :

Exemple 3.5.12. Soit $P = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 0 \\ 0 & 1 & 2 \end{pmatrix}$. L'égalité $X = PX'$ se traduit par le système

$$\begin{cases} x_1 &= x'_1 + 2x'_2 + 3x'_3 \\ x_2 &= 2x'_1 + 3x'_2 \\ x_3 &= x'_2 + 2x'_3 \end{cases}$$

La résolution de ce système donne, après calculs nécessaires,

$$\begin{cases} x'_1 &= \frac{1}{4}(6x_1 - x_2 - 9x'_3) \\ x'_2 &= \frac{1}{4}(-4x_1 + 2x_2 + 6x'_3) \\ x'_3 &= \frac{1}{4}(2x_1 - x_2 - 1x_3) \end{cases}$$

Donc

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 6 & -1 & -9 \\ -4 & 2 & 6 \\ 2 & -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Il en résulte que $P^{-1} = \frac{1}{4} \begin{pmatrix} 6 & -1 & -9 \\ -4 & 2 & 6 \\ 2 & -1 & -1 \end{pmatrix}$

3.5.3 Changement de la matrice d'une application linéaire

Le résultat ci-dessous étudie l'action d'un changement de bases sur la matrice d'une application linéaire.

Proposition 3.5.13. *Soient E et F deux espaces vectoriels sur $\mathbb{K}$, $\dim(E) = p$, $\dim F = n$. Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E, et $\mathcal{C}, \mathcal{C}'$ deux bases de F. Soit $f : E \longrightarrow F$ une application linéaire, $A = M(f, \mathcal{B}, \mathcal{C})$, $A' =$*

$M(f, \mathcal{B}', \mathcal{C}')$. Soient $P \in Gl_p(\mathbb{K})$ la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ et $Q \in Gl_n(\mathbb{K})$ la matrice de passage de $\mathcal{C}$ à $\mathcal{C}'$. Alors :

$$\boxed{A' = Q^{-1}AP}$$

DÉMONSTRATION. Posons $\mathcal{B} = (e_1, \ldots, e_p)$, $\mathcal{B}' = (e'_1, \ldots, e'_p)$, $\mathcal{C} = (u_1, \ldots, u_n)$, $\mathcal{C}' = (u'_1, \ldots, u'_n)$. Soit $x = x_1e_1 + \ldots + x_pe_p = x'_1e'_1 + \cdots + x'_pe'_p \in E$. Posons $y = f(x) = y_1u_1 + \cdots + y_nu_n = y_1u_1 + \cdots + y_nu_n \in F$. Considérons les matrices colonnes

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}, X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_p \end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{K}), \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, Y' = \begin{pmatrix} y'_1 \\ \vdots \\ y'_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{K})$$

L'égalité $y = f(x)$ se traduit matriciellment par

$$Y = AX, \quad Y' = A'X'.$$

D'autre part, les formules de changement de bases dans E et F sont

$$X = PX', \quad Y = QY'.$$

Donc

$$QY' = Y = AX = A(PX') = APX',$$

ce qui implique

$$Y' = Q^{-1}APX'$$

Mais comme $Y' = A'X'$, il vient $A' = Q^{-1}AP$. ■

Remarque 3.5.14. Avec les notations précédentes, on a $P = M(\mathtt{id}_E, \mathcal{B}', \mathcal{B})$, $Q = M(\mathtt{id}_F, \mathcal{C}', \mathcal{C})$ et donc $Q^{-1} = M(\mathtt{id}_F, \mathcal{C}, \mathcal{C}')$.
Une autre démonstration plus courte du résultat ci-dessus consiste à considérer les applications linéaires

$$E \xrightarrow{id_E} E \xrightarrow{f} F \xrightarrow{id_F} F$$

L'égalité $f = \mathtt{id}_F \circ f \circ \mathtt{id}_E$ possède la traduction matricielle :

$$\begin{aligned} A' &= M(f, \mathcal{B}', \mathcal{C}') = M(\mathtt{id}_F \circ f \circ \mathtt{id}_E, \mathcal{B}', \mathcal{C}') \\ &= M(\mathtt{id}_F, \mathcal{C}, \mathcal{C}') \times M(f, \mathcal{B}, \mathcal{C}) \times M(\mathtt{id}_E, \mathcal{B}', \mathcal{B}) \\ &= Q^{-1} \times A \times P \end{aligned}$$

Exemple 3.5.15. Soit $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ l'application linéaire définie par $f(x_1, x_2, x_3) = (2x_1 + x_2 + 4x_3, x_2 - x_3)$. On munit les espaces vectoriels $E=\mathbb{R}^3$ et $F=\mathbb{R}^2$ de leurs bases canoniques $\mathcal{B}=(e_1,e_2,e_3)$ et $\mathcal{C}=(u_1,u_2)$. Alors la matrice de f dans ces bases est

$$A=M(f,\mathcal{B},\mathcal{C})=\begin{pmatrix}2&1&4\\0&1&-1\end{pmatrix}$$

Soient maintenant les nouvelles bases $\mathcal{B}'=(e'_1,e'_2,e'_3)$ de E et $\mathcal{C}'=(u'_1,u'_2)$ de F avec

$$e'_1=(1,0,-1),e'_2=(0,1,1),e'_3=(1,0,1),\qquad u'_1=(1,1),u'_2=(-1,1)$$

La matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$ de E est

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

D'autre part, la matrice de passage de la base $\mathcal{C}$ à la base $\mathcal{C}'$ de F est $Q=\begin{pmatrix}1&-1\\1&1\end{pmatrix}$. De plus, d'après l'exemple 3.4.7,

$$Q^{-1}=\begin{pmatrix}\frac{1}{2}&\frac{1}{2}\\-\frac{1}{2}&\frac{1}{2}\end{pmatrix}=\frac{1}{2}\begin{pmatrix}1&1\\-1&1\end{pmatrix}$$

Par conséquent, la matrice de f dans les bases $\mathcal{B}'$ et $\mathcal{C}'$ est

$$A'=M(f,\mathcal{B}',\mathcal{C}')=Q^{-1}AP=\frac{1}{2}\begin{pmatrix}1&1\\-1&1\end{pmatrix}\begin{pmatrix}2&1&4\\0&1&-1\end{pmatrix}\begin{pmatrix}1&0&1\\0&1&0\\-1&1&1\end{pmatrix}$$

Finalement, les calculs fournissent $A'=\frac{1}{2}\begin{pmatrix}-1&5&6\\3&-5&-7\end{pmatrix}$.

La proposition précédente nous amène à la définition naturelle suivante :

Définition 3.5.16. *Soient deux matrices $A,B\in\mathcal{M}_{n,p}(\mathbb{K})$. On dit que A est **équivalente** à B s'il existe deux matrices inversibles $P\in GL_p(\mathbb{K})$ et $Q\in GK_n(\mathbb{K})$ telles que $B=Q^{-1}AP$.*
D'après la proposition précédente, il est équivalent de dire que A et B représentent la même application linéaire dans des bases différentes.

Remarque 3.5.17. • En notant $Q'=Q^{-1}$, on peut dire que A est équivalente à B s'il existe deux matrices inversibles $P\in GL_p(\mathbb{K})$ et $Q'\in GL_n(\mathbb{K})$ telles que $B=Q'AP$.
• La relation (A *est équivalente à* B) est une relation d'équivalence dans $\mathcal{M}_{n,p}(\mathbb{K})$. En effet :
- Réflexivité : $\forall A\in\mathcal{M}_{n,p}(\mathbb{K}),\ A=I_nAI_p$. - *Symétrie* : S'il existe $P\in GL_p(\mathbb{K})$ et $Q\in GL_n(\mathbb{K})$ telles que $B=Q^{-1}AP$, alors $A=$

$QBP^{-1} = (Q')^{-1}BP'$ avec $Q' = Q^{-1} \in GL_n(\mathbb{K})$ et $P' = P^{-1} \in GL_p(\mathbb{K})$. - *Transitivité* : Si A est équivalente à B et B est équivalente à C, il existe $P_1, P_2 \in GL_p(\mathbb{K})$ et $Q_1, Q_2 \in GL_n(\mathbb{K})$ telles que $B = Q_1^{-1}AP_1$ et $C = Q_2^{-1}BP_2$. Alors

$$C = Q_2^{-1}(Q_1^{-1}AP_1)P_2 = (Q_2^{-1}Q_1^{-1})A(P_1P_2) = (Q_1Q_2)^{-1}A(P_1P_2) = Q_3^{-1}AP_3$$

avec $P_3 = P_1P_2 \in GL_p(\mathbb{K})$ $Q_3 = Q_1Q_2 \in GL_n(\mathbb{K})$. D'où A est équivalente à C.

Lorsque quon a compris ce qui se passe quand les deux espaces vectoriels E et F sont différents, cest facile de particulariser au cas $E = F$, $\mathcal{B} = \mathcal{C}, \mathcal{B}' = \mathcal{C}'$ et $f : E \longrightarrow E$. On déduit immédiatement :

Corollaire 3.5.18. *Soient E un espace vectoriel de dimension finie sur $\mathbb{K}$, $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E et P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Soit f un endomorphisme de E, $A = M(f, \mathcal{B})$, $A' = M(f, \mathcal{B}')$. Alors :*

$$\boxed{A' = P^{-1}AP}$$

Définition 3.5.19. *Soient $A, A' \in \mathcal{M}_n(\mathbb{K})$. On dit que A est semblable à A', et on note $A \sim A'$, s'il existe une matrice inversible $P \in Gl_n(\mathbb{K})$ telle que $A' = P^{-1}AP$.*
Il est équivalent de dire que A et A' représentent le même endomorphisme f dans des bases différentes.

Remarque 3.5.20. La relation $\sim$ est aussi une relation d'équivalence dans $\mathcal{M}_n(\mathbb{K})$. En effet,

Réflexivité : $\forall A \in \mathcal{M}_n(\mathbb{K}),\ A \sim A$ car $A = I_n^{-1}AI_n$.

Symétrie : Si $A \sim A'$, il existe $P \in Gl_n(\mathbb{K})$ telle que $A' = P^{-1}AP$, alors $A = PA'P^{-1} = Q^{-1}A'Q$ avec $Q = P^{-1} \in Gl_n(\mathbb{K})$. Donc $A' \sim A$.

Transitivité : Si $A \sim A'$ et $A' \sim A''$, il existe $P, Q \in Gl_n(\mathbb{K})$ telles que $A' = P^{-1}AP$ et $A'' = Q^{-1}A'Q$. Alors

$$A'' = Q^{-1}A'Q = Q^{-1}(P^{-1}AP)Q = Q^{-1}P^{-1}APQ = (PQ)^{-1}A(PQ),$$

c'est-à-dire, $A'' = R^{-1}AR$ avec $R = PQ \in Gl_n(\mathbb{K})$. D'où $A \sim A''$.

Remarque 3.5.21. Soulignons la formule importante suivante :

$$\boxed{(P^{-1}AP)^k = P^{-1}A^kP \qquad \forall\, k \geq 0}$$

puisque :

$$\begin{aligned}
\left(P^{-1}AP\right)^k &= (P^{-1}AP)(P^{-1}AP)\ldots(P^{-1}AP) \\
&= P^{-1}A\underbrace{(PP^{-1})}_{I_n}A\underbrace{(PP^{-1})}_{I_n}A\ldots A\underbrace{(PP^{-1})}_{I_n}AP \\
&= P^{-1}\underbrace{(AA\ldots A)}_{k fois}P \\
&= P^{-1}A^kP
\end{aligned}$$

Par conséquent,

$$\boxed{A' = P^{-1}AP \implies (A')^k = P^{-1}A^kP \quad \text{et} \quad A^k = P(A')^kP^{-1} \qquad \forall\, k \geq 0}$$

Le procédé de changement de bases est coûteux en calculs et pas toujours très efficace. Changer de bases a pour but de simplifier la matrice d'une application linéaire, et si on aboutit à une matrice diagonale cela vaut le coup. Dans beaucoup de situations, il est demandé de connaître les puissances A^k d'une matrice carrée A. En général, il est difficile de prévoir quelle sera la forme de A^k. Néanmoins, pour une matrice diagonale $D = \mathsf{diag}(\lambda_1, \ldots, \lambda_n)$, nous avons déjà vu que D^k se calcule très facilement : $D^k = \mathsf{diag}(\lambda_1^k, \ldots, \lambda_n^k)$. Par conséquent, si un changement de bases nous fournit une matrice diagonale $D = P^{-1}AP$, alors les puissances A^k peuvent être obtenues à l'aide des puissances D^k grâce à la formule $A^k = PD^kP^{-1}$. Traitons l'exemple suivant pour illustrer :

Exemple 3.5.22. Soit f l'endomorphisme de $E = \mathbb{R}^3$ représenté par la matrice $A = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}$ dans la base canonique $\mathcal{B} = (e_1, e_2, e_3)$. Considérons la nouvelle base $\mathcal{B}' = (e'_1, e'_2, e'_3)$ de E avec

$$e'_1 = (1,1,1),\ e'_2 = (1,0,-1),\ e'_3 = (0,1,-1).$$

Pour calculer la matrice $D = M(f, \mathcal{B}')$ de f dans la nouvelle base $\mathcal{B}'$, on peut employer la formule $D = P^{-1}AP$ qui nécessite le calcul de P^{-1}. Mais nous allons voir qu'il est plus rapide ici de calculer la matrice $D = M(f, \mathcal{B}')$ en utilisant directement la définition de cette matrice, c'est-à-dire, en calculant les coordonnées de $f(e'_1), f(e'_2), f(e'_3)$ dans la

base $\mathcal{B}' = (e'_1, e'_2, e'_3)$.
Pour cela, en appliquant la relation $Y = AX$, on obtient :

$$\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} = - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \text{donc } f(e'_1) = -e'_1.$$

$$\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \text{donc } f(e'_2) = 2e'_2.$$

$$\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix} = 2 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad \text{donc } f(e'_3) = 2e'_3.$$

Par conséquent, $M(f, \mathcal{B}') = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, qui est une matrice plus simple que la matrice A, puisqu'elle est diagonale. En particulier cette matrice $M(f, \mathcal{B}')$ est inversible, car ses éléments diagonaux $-1, 2, 2$ sont tous non nuls. Ceci implique que l'application linéaire f est bijective. Et Comme cette propriété est propre à l'application linéaire f, elle ne dépend pas de la base dans laquelle on calcule sa matrice. Et donc en particulier, cela signifie que la matrice A est inversible.
La matrice de passage P de la base $\mathcal{B}$ à la base $\mathcal{B}'$ est $P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix}$.
En notant $D = M(f, \mathcal{B}')$, on a $D = P^{-1}AP$ et $A = PDP^{-1}$. Si nous voulons calculer les puissances A^k de A, il suffit d'appliquer la formule $A^k = PD^kP^{-1}$, et ceci va maintenant nécessiter le calcul de la matrice P^{-1}. Sans faire les détails des calculs, on trouve $P^{-1} = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ -1 & 2 & -1 \end{pmatrix}$.
Puisque

$$D^k = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}^k = \begin{pmatrix} (-1)^k & 0 & 0 \\ 0 & 2^k & 0 \\ 0 & 0 & 2^k \end{pmatrix},$$

il vient

$$A^k = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} (-1)^k & 0 & 0 \\ 0 & 2^k & 0 \\ 0 & 0 & 2^k \end{pmatrix} \cdot \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ -1 & 2 & -1 \end{pmatrix}$$

$$= \frac{1}{3} \begin{pmatrix} (-1)^k + 2^{k+1} & (-1)^k - 2^k & (-1)^k - 2^k \\ (-1)^k - 2^k & (-1)^k + 2^{k+1} & (-1)^k - 2^k \\ (-1)^k - 2^{k+1} + 2^k & (-1)^k - 2^{k+1} + 2^k & (-1)^k + 2^k + 2^k \end{pmatrix}$$

$$= \frac{1}{3}\begin{pmatrix} (-1)^k + 2^{k+1} & (-1)^k - 2^k & (-1)^k - 2^k \\ (-1)^k - 2^k & (-1)^k + 2^{k+1} & (-1)^k - 2^k \\ (-1)^k - 2^k & (-1)^k - 2^k & (-1)^k + 2^{k+1} \end{pmatrix}.$$

Pour s'assurer que les calculs sont corrects, prendre par exemple $k = 0$, $k = 1$ et vérifier qu'on a bien $A^0 = I_3$, $A^1 = A$.

3.6 Transposée, rang et trace d'une matrice

3.6.1 Transposée d'une matrice

Définition 3.6.1. *Soit* $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix}$ *une matrice de type* (n, p). *On appelle* ***transposée*** *de* A, *la matrice notée* ${}^tA$ *de type* (p, n) *qui est définie par :*

$${}^tA = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} \in \mathcal{M}_{p,n}(\mathbb{K})$$

Donc si $A = (a_{ij})_{\substack{1\leq i\leq n \\ 1\leq j\leq p}}$, *alors*

$${}^tA = (b_{ij})_{\substack{1\leq i\leq p \\ 1\leq j\leq n}} \quad \text{avec } b_{ij} = a_{ij} \quad \forall i \in \{1, \dots, p\},\ \forall j \in \{1, \dots, n\}$$

Autrement dit, la matrice ${}^tA$ *s'obtient en échangeant les lignes et les colonnes de* A. *C'est-à-dire, les lignes de* A *deviennent les colonnes de* ${}^tA$ *(et les colonnes de* A *deviennent les lignes de* ${}^tA$*).*

Exemple 3.6.2. La transposée de la matrice $A = \begin{pmatrix} 3 & 2 & 0 \\ 4 & 7 & -2 \end{pmatrix}$ qui est de type $(2, 3)$ est la matrice ${}^tA = \begin{pmatrix} 3 & 4 \\ 2 & 7 \\ 0 & -2 \end{pmatrix}$ qui est de type $(3, 2)$.

Exemple 3.6.3. La transposée d'une matrice carrée d'ordre n est une matrice carrée d'ordre n.

En particulier, pour la matrice unité $I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$, il est facile de voir que

$$\boxed{{}^tI_n = I_n}$$

Exemple 3.6.4. la transposée d'une matrice-ligne est une matrice-colonne et réciproquement :

$$^t\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix}, \qquad {}^t\begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Remarque 3.6.5. Soit $A = (a_{ij})_{\substack{1\le i\le n\\1\le j\le p}}$ une matrice de type (n,p). Puisque les notations des indices i et j n'ont aucune importance (indices *muets*), on peut écrire sa transposée sous forme ${}^tA = (a'_{ji})_{\substack{1\le j\le p\\1\le i\le n}}$ avec $a'_{ji} = a_{ij}$ pour tous indices $j \in \{1,\dots,p\}$ et $i \in \{1,\dots,n\}$.

Voici quelques propriétés immédiates de la transposée :

Proposition 3.6.6. *On a :*

(a) $\forall A \in \mathcal{M}_{n,p}(\mathbb{K}),\ \ {}^t({}^tA) = A$.

(b) $\forall A, B \in \mathcal{M}_{n,p}(\mathbb{K}),\ \forall \alpha,\beta \in \mathbb{K}\quad {}^t(\alpha A + \beta B) = \alpha\,{}^tA + \beta\,{}^tB$.
Donc l'application $t : \mathcal{M}_{n,p}(\mathbb{K}) \longrightarrow \mathcal{M}_{p,n}(\mathbb{K})$ *qui à chaque matrice* A *associe sa transposée* ${}^tA$ *est linéaire.*

DÉMONSTRATION. :

(a) Immédiat, car si $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$, alors ${}^tA = (a'_{ij}) \in \mathcal{M}_{p,n}(\mathbb{K})$ avec $a'_{ij} = a_{ji}$. Donc ${}^t\left({}^tA\right) = (a''_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$ avec $a''_{ij} = a'_{ji} = a_{ij}$, ce qui donne ${}^t({}^tA) = A$.

(b) En notant $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$, $B = (b_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$, on a $\alpha A + \beta B = (c_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$ avec $c_{ij} = \alpha a_{ij} + \beta b_{ij}$. Donc

$${}^t(\alpha A+\beta B) = (c'_{ij}) \in \mathcal{M}_{p,n}(\mathbb{K}) \text{ avec } c'_{ij} = c_{ji} = \alpha a_{ji}+\beta b_{ji}. \quad (3.14)$$

Ecrivons maintenant ${}^tA = (a'_{ij}) \in \mathcal{M}_{p,n}(\mathbb{K})$ avec $a'_{ij} = a_{ji}$ et ${}^tB = (b'_{ij}) \in \mathcal{M}_{p,n}(\mathbb{K})$ avec $b'_{ij} = b_{ji}$. Donc

$$\alpha\,{}^tA+\beta\,{}^tB = (d_{ij}) \in \mathcal{M}_{p,n}(\mathbb{K}) \text{ avec } d_{ij} = \alpha a'_{ij}+\beta b'_{ij} = \alpha a_{ji}+\beta b_{ji}. \quad (3.15)$$

En comparant (3.14) et (3.15), on conclut au résultat désiré. ■

Proposition 3.6.7. *La transposée satisfait aussi les propriétés intéressantes suivantes :*

i. $\forall A \in \mathcal{M}_{n,p}(\mathbb{K})$, $\forall B \in \mathcal{M}_{p,r}(\mathbb{K})$:

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

ii. $\forall A \in GL_n(\mathbb{K})$, ${}^tA \in GL_n(\mathbb{K})$ *et l'on a :*

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

DÉMONSTRATION.

i. Notons $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$, $B = (b_{jk}) \in \mathcal{M}_{p,r}(\mathbb{K})$. On a ${}^tA = (a'_{ji}) \in \mathcal{M}_{p,n}(\mathbb{K})$ avec $a'_{ji} = a_{ij}$ et ${}^tB = (b'_{kj}) \in \mathcal{M}_{r,p}(\mathbb{K})$ avec $b'_{kj} = b_{jk}$. Le produit $({}^tB)({}^tA)$ existe et $({}^tB)({}^tA) = (c'_{ki}) \in \mathcal{M}_{r,n}(\mathbb{K})$ avec

$$c'_{ki} = \sum_{j=1}^{p} b'_{kj}a'_{ji} = \sum_{j=1}^{p} b_{jk}a_{ij} = \sum_{j=1}^{p} a_{ij}b_{jk}. \tag{3.16}$$

D'autre part, $AB = (d_{ik}) \in \mathcal{M}_{n,r}(\mathbb{K})$ avec $d_{ik} = \sum_{j=1}^{p} a_{ij}b_{jk}$. Donc ${}^t(AB) = (d'_{ki}) \in \mathcal{M}_{r,n}(\mathbb{K})$ avec

$$d'_{ki} = c_{ik} = \sum_{j=1}^{p} a_{ij}b_{jk}. \tag{3.17}$$

En comparant (3.16) et (3.17), on obtient $c'_{ki} = d'_{ki}$ pour tous $k \in \{1, \ldots, r\}$ et $i \in \{, \ldots, n\}$. D'où ${}^t(AB) = ({}^tB)({}^tA)$.

ii. Soit $A \in GL_n(\mathbb{K})$, c'est-à-dire, A est une matrice inversible d'ordre n. Puisque $AA^{-1} = I_n$, alors ${}^t(AA^{-1}) =^t I_n = I_n$. Mais ${}^t(AA^{-1}) =^t (A^{-1})({}^tA)$, d'où ${}^t(A^{-1})({}^tA) = I_n$. Par conséquent, la matrice ${}^tA$ est inversible et son inverse est $({}^tA)^{-1} = {}^t(A^{-1})$. ■

Exemple 3.6.8. Soit la matrice $A = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 0 & 3 \\ -2 & 3 & 0 \end{pmatrix}$. Nous avons déjà montré dans l'exemple 3.4.14 que A est inversible et que $A^{-1} = \begin{pmatrix} -3 & 2 & 1 \\ -2 & \frac{4}{3} & 1 \\ 3 & -\frac{5}{3} & -1 \end{pmatrix}$. D'après la proposition précédente, la transposée ${}^tA = \begin{pmatrix} 1 & 3 & -2 \\ 1 & 0 & 3 \\ 2 & 3 & 0 \end{pmatrix}$ est inversible et que son inverse est

$$({}^tA)^{-1} = {}^t(A^{-1}) = \begin{pmatrix} -3 & -2 & 3 \\ 2 & \frac{4}{3} & -\frac{5}{3} \\ 1 & 1 & -1 \end{pmatrix}$$

3.6.2 Rang d'une matrice

Nous avons déjà défini la notion de rang pour une famille de vecteurs et pour une application linéaire : le rang d'une famille de vecteurs $(v_1, \dots, v_k)$ est la dimension du sous espace vectoriel $\mathsf{Vect}(v_1, \dots, v_k)$ qu'elle engendre, et le rang d'une application linéaire f est la dimension de son image $\mathsf{Im}(f)$. Ces deux concepts sont liés : Soient E et F deux espaces vectoriels sur $\mathbb{K}$, et $f : E \longrightarrow F$ une application linéaire. Donc si $\mathcal{B} = (e_1, \dots, e_p)$ est une base de E, le rang de f est aussi le rang de la famille $(f(e_1), \dots, f(e_p))$ et ce, quelle que soit la base $\mathcal{B} = (e_1, \dots, e_p)$ de E. Ce rang ne dépend pas non plus de la base $\mathcal{B}' = (e'_1, \dots, e'_n)$ de F dans laquelle on écrit la famille à l'arrivée : c'est le rang des vecteurs colonnes de la matrice de f, quelles que soient les bases par rapport auxquelles on écrit cette matrice.
En particulier, si $E = \mathbb{K}^p$ et $F = \mathbb{K}^n$, alors $f(e_1), \dots, f(e_p)$ sont les vecteurs colonnes de la matrice de f relativement aux bases canoniques de $\mathbb{K}^p$ et $\mathbb{K}^n$. Nous sommes ainsi amenés à introduire la définition suivante.

Définition 3.6.9. *Soit $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$ une matrice de type (n, p). On appelle **rang** de A, et on note $\mathsf{rg}(A)$, le rang de ses vecteurs colonnes dans $\mathbb{K}^n$.*

Ainsi, en notant

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix}, \text{ on a } \mathsf{rg}(A) = \mathsf{rg}(C_1, \dots, C_p) \text{ où}$$

$$C_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \ C_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \dots, C_p = \begin{pmatrix} a_{1p} \\ a_{2p} \\ \vdots \\ a_{np} \end{pmatrix}$$

sont les colonnes de A.
Donc le rang de A est la dimension du sous-espace vectoriel de $\mathbb{K}^n$ engendré par $C_1, \dots, C_p$. Comme il y a p vecteurs, alors $\mathsf{rg}(C_1, \dots, C_p) \leq p$. Puisque ces vecteurs $C_1, \dots, C_p$ appartiennent à $\mathbb{K}^n$, on a aussi $\mathsf{rg}(C_1, \dots, C_p) \leq n$.
D'autre part, D'après le chapitre 1, on sait que le rang d'une famille finie $\mathcal{F}$ de vecteurs est égal au nombre maximal de vecteurs linéairement indépendants extraits de cette famille $\mathcal{F}$. Nous pouvons donc énoncer :

Proposition 3.6.10. *Pour toute matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$, on a :*

i. $\mathsf{rg}(A) \leq \min(p, n)$.

ii. *Le rang de la matrice A est le nombre maximal de vecteurs colonnes de A qui sont linéairement indépendants.*

Exemple 3.6.11. Soit la matrice $A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \end{pmatrix}$. On a $\mathsf{rg}(A) \leq 2$. On observe que les deux premiers vecteurs colonnes $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ne sont pas colinéaires, donc linéairement indépendants. Ainsi, le rang de A est égal à 2.

Exemple 3.6.12.

$$\mathsf{rg}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 1, \qquad \mathsf{rg}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 2, \qquad \mathsf{rg}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 3.$$

Définition 3.6.13. *Soient E un espace vectoriel de dimension n sur $\mathbb{K}$, $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E et $\mathcal{F} = (v_1, \ldots, v_p)$ une famille de p vecteurs de E. Pour chaque $j \in \{1, \ldots, p\}$, soient $(a_{1j}, a_{2j}, \ldots, a_{nj})$ les coordonnées de v_j dans la base $\mathcal{B}$:*

$$v_j = a_{1j}e_1 + a_{2j}e_2 + \cdots + a_{nj}e_n = \sum_{i=1}^{n} a_{ij}e_i.$$

La matrice $(a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1p} \\ a_{21} & a_{22} & \ldots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{np} \end{pmatrix}$ de $\mathcal{M}_{n,p}(\mathbb{K})$ s'appelle la ***matrice de la famille** $\mathcal{F} = (v_1, \ldots, v_p)$ **dans la base** $\mathcal{B}$, et se note $\mathsf{Mat}_{\mathcal{B}}(\mathcal{F})$ ou $\mathsf{Mat}_{\mathcal{B}}(v_1, \ldots, v_p)$. C'est la matrice dont les colonnes sont les coordonnées des vecteurs $v_1, \ldots, v_p$ dans la base $\mathcal{B}$.*

Exemple 3.6.14. Soit E un espace vectoriel de dimension 2 muni d'une base $\mathcal{B} = (e_1, e_2)$ et soient les vecteurs $v_1 = e_1 + 7e_2$, $v_2 = 4e_2$, $v_3 = -e_1 + e_2$, $v_4 = 5e_1 - 2e_2$. La matrice de la famille (v_1, v_2, v_3, v_4) dans la base $\mathcal{B}$ est $\mathsf{Mat}_{\mathcal{B}}(v_1, v_2, v_3, v_4) = \begin{pmatrix} 1 & 0 & -1 & 5 \\ 7 & 2 & 1 & -2 \end{pmatrix}$.

Proposition 3.6.15. *Soient E un espace vectoriel de dimension n sur $\mathbb{K}$, $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E et $\mathcal{F} = (v_1, \ldots, v_p)$ une famille de p vecteurs de E. En notant $A = \mathsf{Mat}_{\mathcal{B}}(v_1, \ldots, v_p)$, on a*

$$\mathsf{rg}(v_1, \ldots, v_p) = \mathsf{rg}(A)$$

C'est-à-dire, le rang d'une famille finie $\mathcal{F} = (v_1, \ldots, v_p)$ de vecteurs de E est égal au rang de la matrice de $\mathcal{F}$ dans n'importe quelle base $\mathcal{B}$ de E.

Démonstration. Soit $f : \mathbb{K}^n \longrightarrow E$ l'application définie par $f(x_1, \dots, x_n) = \sum_{i=1}^{n} x_i e_i$. Il est facile de voir que f est linéaire. De plus, $\ker(f) = \{0_{\mathbb{K}^n}\}$, car si $f(x_1, \dots, x_n) = 0_E$ alors $\sum_{i=1}^{n} x_i e_i = 0_E$, donc $x_1 = \cdots = x_n = 0$, d'où $(x_1, \dots, x_n) = (0, \dots, 0) = 0_{\mathbb{K}^n}$. Ainsi, f est injective, et comme $\dim \mathbb{K}^n = \dim E = n$, alors f est bijective (On peut le voir d'une autre manière en remarquant que f transforme la base canonique de $\mathbb{K}^n$ en la base $\mathcal{B} = (e_1, \dots, e_n)$ de E : $f(1, 0, \dots, 0) = e_1, \dots, f(0, \dots, 0, 1) = e_n$). Donc f est un isomorphisme de $\mathbb{K}^n$ dans E. En outre, f transforme les vecteurs colonnes $C_1, \dots, C_p$ de la matrice $A = \mathsf{Mat}_{\mathcal{B}}(v_1, \dots, v_p)$ (qui sont dans $\mathbb{K}^n$) en les vecteurs $v_1, \dots, v_p$ de E : $f(C_1) = v_{1,l}\, dots, f(C_p) = v_p$. Il en résulte que

$$f\left(\mathsf{Vect}(C_1, \dots, C_p)\right) = \mathsf{Vect}(f(C_1), \dots, f(C_p)) = \mathsf{Vect}(v_1, \dots, v_p).$$

Ainsi, les deux espaces vectoriels $\mathsf{Vect}(C_1, \dots, C_p)$ et $\mathsf{Vect}(v_1, \dots, v_p)$ sont isomorphes (prendre la restriction $f : \mathsf{Vect}(C_1, \dots, C_p) \longrightarrow \mathsf{Vect}(v_1, \dots, v_p)$ qui est à la fois injective et surjective). Il en découle que $\dim\left(\mathsf{Vect}(C_1, \dots, C_p)\right) = \dim\left(\mathsf{Vect}(v_1, \dots, v_p)\right)$. D'où $\mathsf{rg}(A) = \mathsf{rg}(v_1, \dots, v_p)$. ■

Exemple 3.6.16. Dans l'exemple précédent, pour calculer le rang de la famille de vecteurs (v_1, v_2, v_3, v_4) de E, il suffit de trouver le rang de la matrice $\begin{pmatrix} 1 & 0 & -1 & 5 \\ 7 & 2 & 1 & -2 \end{pmatrix}$, qui vaut 2.

La proposition suivante relie le rang d'une matrice avec le rang d'une application linéaire associée :

Proposition 3.6.17. *Soient E, F deux espaces vectoriels sur $\mathbb{K}$, $\mathcal{B}$ une base de E et $\mathcal{B}'$ une base de F. Soit $f : E \longrightarrow F$ une application linéaire et $M(f, \mathcal{B}, \mathcal{B}')$ sa matrice dans les bases $\mathcal{B}$ et $\mathcal{B}'$). On a :*

$$\boxed{\mathsf{rg}(f) = \mathsf{rg}\left(M(f, \mathcal{B}, \mathcal{B}')\right)}$$

Démonstration. Posons $\dim E = p$, $\dim F = n$, $\mathcal{B} = (e_1, \dots, e_p)$, $\mathcal{B}' = (e'_1, \dots, e'_n)$. Les colonnes de la matrice $M(f, \mathcal{B}, \mathcal{B}')$ sont les coordonnées de $f(e_1), \dots, f(e_p)$ dans la base $\mathcal{B}' = (e'_1, \dots, e'_n)$. Donc

$$M(f, \mathcal{B}, \mathcal{B}') = \mathsf{Mat}_{\mathcal{B}'}(f(e_1), \dots, f(e_p)),$$

ce qui donne

$$\mathsf{rg}\left(M(f, \mathcal{B}, \mathcal{B}')\right) = \mathsf{rg}\left(\mathsf{Mat}_{\mathcal{B}'}(f(e_1), \dots, f(e_p))\right).$$

Or, d'après la proposition 3.6.15,

$$\mathsf{rg}\left(\mathsf{Mat}_{\mathcal{B}'}(f(e_1),\ldots,f(e_p))\right) = \mathsf{rg}(f(e_1),\ldots,f(e_p)).$$

Il s'ensuit :

$$\begin{aligned}\mathsf{rg}\left(\mathsf{Mat}_{\mathcal{B}'}(f(e_1),\ldots,f(e_p))\right) &= \dim\left(\mathsf{Vect}(f(e_1),\ldots,f(e_p))\right)\\ &= \dim(\mathsf{Im}(f)) = \mathsf{rg}(f).\end{aligned}$$

D'où $\mathsf{rg}\left(M(f,\mathcal{B},\mathcal{B}')\right) = \mathsf{rg}(f)$. ■

Exemple 3.6.18. Retournons à l'exemple 3.5.15 de l'application linéaire $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ définie par $f(x_1,x_2,x_3) = (2x_1 + x_2 + 4x_3, x_2 - x_3)$. On a vu que la matrice de f par rapport aux bases canoniques $\mathcal{B}$ et $\mathcal{C}$ de $\mathbb{R}^3$ et $\mathbb{R}^2$ est $A = M(f,\mathcal{B},\mathcal{C}) = \begin{pmatrix} 2 & 1 & 4\\ 0 & 1 & -1\end{pmatrix}$.
On en déduit que $\mathsf{rg}(f) = \mathsf{rg}(A) = 2$, puisque les deux premiers vecteurs colonnes sont linéairement indépendants.
D'autre part, on a trouvé que la matrice de f dans les nouvelles bases $\mathcal{B}'$ et $\mathcal{C}'$ de $\mathbb{R}^3$ et $\mathbb{R}^2$ est $A' = M(f,\mathcal{B}',\mathcal{C}') = \frac{1}{2}\begin{pmatrix} -1 & 5 & 6\\ 3 & -5 & -7\end{pmatrix}$.
On vérifie bien qu'on a aussi $\mathsf{rg}(A') = 2$, puisque le rang de f est le rang de sa matrice dans n'importe quelles bases choisies.

Calculer le rang d'une application linéaire, ou d'une famille de vecteurs, revient à calculer le rang d'une matrice.

La connaissance du rang fournit un critère d'inversibilité pour une matrice carrée :

Proposition 3.6.19. *Pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$, on a :*

$$\boxed{A \text{ est inversible} \iff \mathsf{rg}(A) = n}$$

Démonstration. Soit f l'endomorphisme de $E = \mathbb{K}^n$ représenté par la matrice A dans la base canonique $\mathcal{B} = (e_1,\ldots,e_n)$ de $\mathbb{K}^n$. On a $\mathsf{rg}(f) = \mathsf{rg}(A)$. De plus, on rappelle que la matrice A est inversible si et seulement si l'application linéaire associée f est bijective. Or, selon le chapitre 2, f est bijective si et seulement si $\mathsf{Im}(f) = E$, c'est-à-dire si $\mathsf{rg}(f) = n$. D'où, A est inversible si et seulement si $\mathsf{rg}(A) = n$. ■

Le rang d'une matrice est celui des applications linéaires qu'elle représente, qui ne dépend pas des bases. Si deux matrices représentent la même application dans des bases différentes, elles auront nécessairement même rang. Rappelons que deux matrices sont équivalentes si elles représentent la même application linéaire dans deux bases différentes. Par conséquent, deux matrices équivalentes ont même rang. Nous allons démontrer la réciproque. Pour ce faire, nous prouvons d'abord le résultat important suivant :

Théorème 3.6.20. *Soient E, F deux espaces vectoriels sur $\mathbb{K}$, $\dim E = p$, $\dim F = n$, et soit $f : E \longrightarrow F$ une application linéaire. Alors il existe une base $\mathcal{B}$ de E et une base $\mathcal{C}$ de F telles que*

$$M(f,\mathcal{B},\mathcal{C}) = \left(\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right) = \left(\begin{array}{cccc|ccc} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ \hline 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{array}\right), \quad (3.18)$$

où r est le rang de f et I_r est la matrice unité d'ordre r.

Démonstration. Puisque $r = \mathsf{rg}(f) = \dim \mathsf{Im}(f)$, le théorème du rang donne $\dim\ker(f) = \dim E - \dim \mathsf{Im}(f) = p - r$. Considérons une base de $\ker(f)$, que nous notons $(e_{r+1}, \dots, e_p)$. Cette famille étant libre dans E, d'après le théorème de la base incomplète, il existe des vecteurs $e_1, \dots, e_r$ de E tels que $\mathcal{B} = (e_1, \dots, e_r, e_{r+1}, \dots, e_p)$ soit une base de E.
On a donc $\mathsf{Im}(f) = \mathsf{Vect}\{f(e_1), \dots, f(e_r), f(e_{r+1}), \dots, f(e_p)\}$. Mais comme
$f(e_{r+1}) = \dots = f(e_p) = 0_F$, alors $\mathsf{Im}(f) = \mathsf{Vect}\{f(e_1), \dots, f(e_r)\}$. Ceci montre que la famille $(f(e_1), \dots, f(e_r))$ engendre $\mathsf{Im}(f)$. Or, le cardinal de cette famille est $r = \dim \mathsf{Im}(f)$, donc $(f(e_1), \dots, f(e_r))$ est nécessairement une base de $\mathsf{Im}(f)$, d'où elle est libre dans F. (On peut prouver directement que $(f(e_1), \dots, f(e_r))$ est libre dans F : si $\alpha_1 f(e_1) + \dots + \alpha_r f(e_r) = 0_F$, alors $f(\alpha_1 e_1 + \dots + \alpha_r e_r) = 0_F$, donc $\alpha_1 e_1 + \dots + \alpha_r e_r \in \ker(f)$. Alors il existe $\alpha_{r+1}, \dots, \alpha_p$ dans $\mathbb{K}$ tels que $\alpha_1 e_1 + \dots + \alpha_r e_r = \alpha_{r+1} e_{r+1} + \dots + \alpha_p e_p$. D'où $\alpha_1 e_1 + \dots + \alpha_r e_r - \alpha_{r+1} e_{r+1} - \dots - \alpha_p e_p = 0_E$, et par suite, $\alpha_1 = \dots = \alpha_r = \alpha_{r+1} = \dots = \alpha_p = 0$).
Notons $e'_1 = f(e_1), \dots, e'_r = f(e_r)$. D'après le théorème de la base incomplète, on peut compléter la famille libre $(e'_1, \dots, e'_r)$ de F en une base $\mathcal{C} = (e'_1, \dots, e'_r, e'_{r+1}, \dots, e'_n)$ de F.
Puisque $f(e_1) = e'_1, \dots, f(e_r) = e'_r, f(e_{r+1}) = 0_F, \dots, f(e_p) = 0_F$, la matrice $M(f,\mathcal{B},\mathcal{C})$ de f dans les base $\mathcal{B}$ et $\mathcal{C}$ s'écrit sous la forme (3.18). ∎

Remarque 3.6.21. Dans le théorème précédent, puisque $\dim E = p$ et $\dim F = n$, alors la matrice $M(f, \mathcal{B}, \mathcal{C})$ est de type (n, p), c'est-à-dire, elle contient au total n lignes et p colonnes. Or, la matrice unité I_r possède r lignes et r colonnes avec $r \leq \min(p, n)$. Donc pour être plus précis dans la formule (3.18), il faut écrire :

$$M(f, \mathcal{B}, \mathcal{C}) = \left(\begin{array}{c|c} I_r & O_{r,p-r} \\ \hline O_{n-r,r} & O_{n-r,p-r} \end{array}\right),$$

où $O_{i,j}$ désigne la matrice nulle de type (i, j).
Autrement dit, on écrit d'abord la matrice I_r, puisque on complète cette matrice par $n - r$ lignes nulles et $p - r$ colonnes nulles.

Corollaire 3.6.22. *Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $r = \mathsf{rg}(A)$. Alors il existe deux matrices inversibles $P \in GL_p(\mathbb{K})$ et $Q \in GL_n(\mathbb{K})$ telles que*

$$Q^{-1}AP = \left(\begin{array}{c|c} I_r & O_{r,p-r} \\ \hline O_{n-r,r} & O_{n-r,p-r} \end{array}\right)$$

Autrement dit, toute matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ de rang r est équivalente à la matrice $\left(\begin{array}{c|c} I_r & O_{r,p-r} \\ \hline O_{n-r,r} & O_{n-r,p-r} \end{array}\right)$.

Démonstration. Soient les espaces vectoriels $E = \mathbb{K}^p$ et $F = \mathbb{K}^n$ équipés de leurs bases canoniques respectives $\mathcal{B}_p$ et $\mathcal{B}_n$. Puisque $A \in \mathcal{M}_{n,p}(\mathbb{K})$, il existe une application linéaire unique $f : E \longrightarrow F$ telle que $M(f, \mathcal{B}_p, \mathcal{B}_n) = A$.
D'après le théorème précédent, il existe une nouvelle base $\mathcal{B}$ de E et une nouvelle base $\mathcal{C}$ de F telles que

$$M(f, \mathcal{B}, \mathcal{C}) = \left(\begin{array}{c|c} I_r & O_{r,p-r} \\ \hline O_{n-r,r} & O_{n-r,p-r} \end{array}\right)$$

D'autre part, d'après la formule de changement de bases, il existe des matrices inversibles $P \in GL_p(\mathbb{K})$ et $Q \in GL_n(\mathbb{K})$ telles que

$$M(f, \mathcal{B}, \mathcal{C}) = Q^{-1}M(f, \mathcal{B}_p, \mathcal{B}_n)P = Q^{-1}AP.$$

D'où le résultat. ■

Corollaire 3.6.23. *Deux matrices de sont équivalentes si et seulement si elles ont même rang.*

Démonstration. Soient $A, B \in \mathcal{M}_{n,p}(\mathbb{K})$.

(c) Si A et B sont équivalentes, elles représentes la même application linéaire $f : \mathbb{K}^p \longrightarrow \mathbb{K}^n$ dans des bases différentes. Donc $\mathsf{rg}(A) = \mathsf{rg}(f)$ et $\mathsf{rg}(B) = \mathsf{rg}(f)$, ce qui donne $\mathsf{rg}(A) = \mathsf{rg}(B)$.

(d) Réciproquement, supposons que $\mathsf{rg}(A) = \mathsf{rg}(B)$, que nous notons r. Selon le corollaire 3.6.22, chacune des matrices A et B est équivalente à la matrice $\left(\begin{array}{c|c} I_r & O_{r,p-r} \\ \hline O_{n-r,r} & O_{n-r,p-r} \end{array}\right)$. D'où, par transitivité, A est équivalente à B. ∎

On peut maintenant étudier le rang de la transposée d'une matrice :

Corollaire 3.6.24. *Pour toute matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$, on a*

$$\boxed{\mathsf{rg}({}^tA) = \mathsf{rg}(A)}$$

Démonstration. Soit $r = \mathsf{rg}(A)$ et notons

$$J_{n,p,r} = \left(\begin{array}{c|c} I_r & O_{r,p-r} \\ \hline O_{n-r,r} & O_{n-r,p-r} \end{array}\right) \in \mathcal{M}_{n,p}(\mathbb{K}).$$

Il existe des matrices inversibles $P \in GL_p(\mathbb{K})$ et $Q \in GL_n(\mathbb{K})$ telles que $Q^{-1}AP = J_{n,p,r}$, donc $A = QJ_{n,p,r}P^{-1}$. En appliquant la transposée, on obtient

$${}^tA = {}^t\left(QJ_{n,p,r}P^{-1}\right) = {}^t\left(P^{-1}\right)\,{}^t(J_{n,p,r})\,{}^tQ.$$

Or, il est facile de constater que

$${}^tJ_{n,p,r} = {}^t\left(\begin{array}{c|c} I_r & O_{r,p-r} \\ \hline O_{n-r,r} & O_{n-r,p-r} \end{array}\right) = \left(\begin{array}{c|c} I_r & O_{r,n-r} \\ \hline O_{p-r,r} & O_{p-r,n-r} \end{array}\right) = J_{p,n,r} \in \mathcal{M}_{p,n}(\mathbb{K})$$

(car ${}^tI_r = I_r$). De plus, ${}^t\left(P^{-1}\right) = ({}^tP)^{-1}$. D'où

$${}^tA = ({}^tP)^{-1}J_{p,n,r}^t Q = (P')^{-1}J_{p,n,r}Q',$$

avec $P' = {}^tP \in GL_p(\mathbb{K})$ et $Q' = {}^tQ \in GL_n(\mathbb{K})$. Il en découle que ${}^tA$ est équivalente à la matrice $J_{p,n,r}$. Or, le rang de $J_{p,n,r}$ est r car les r premières colonnes de $J_{p,n,r}$ sont linéairement indépendantes. Le corollaire précédent affirme donc que $\mathsf{rg}({}^tA) = \mathsf{rg}(J_{p,n,r}) = r$. D'où $\mathsf{rg}({}^tA) = \mathsf{rg}(A)$. ∎

3.6.3 Trace d'une matrice et d'un endomorphisme

Définition 3.6.25. *Soit $A = (a_{ij})$ une matrice carrée d'ordre n. La **trace** de A, notée $\mathsf{tr}(A)$, est la somme des éléments diagonaux de A :*

$$\boxed{\mathsf{tr}(A) = \sum_{i=1}^{n} a_{ii}}$$

Exemple 3.6.26. La trace de la matrice $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 5 & 6 & 7 \end{pmatrix}$ est $\mathsf{tr}(A) = 1 + 4 + 7 = 12$.

Exemple 3.6.27. La trace de la matrice nulle est 0. La trace de la matrice unité I_n est n.

Observons que la trace d'une matrice $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ est égale à la trace de sa transposée puisque les éléments diagonaux $a_{11}, \dots, a_{nn}$ de A sont aussi les éléments diagonaux de la transposée.

Proposition 3.6.28. *L'application* $\mathsf{tr} : \mathcal{M}_n(\mathbb{K}) \longrightarrow \mathbb{K}$ *qui à chaque matrice A associe sa trace* $\mathsf{tr}(A)$ *est linéaire.*
C'est donc une forme linéaire de $\mathcal{M}_n(\mathbb{K})$.

Démonstration. Soient $A = (a_{ij})$ et $B = (b_{ij})$ dans $\mathcal{M}_n(\mathbb{K})$, et soient α, β dans $\mathbb{K}$. Alors $\alpha A + \beta B = (c_{ij})$ avec $c_{ij} = \alpha a_{ij} + \beta b_{ij}$. Donc

$$\mathsf{tr}(\alpha A+\beta B) = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n}(\alpha a_{ii}+\beta b_{ii}) = \alpha \sum_{i=1}^{n} a_{ii}+\beta \sum_{i=1}^{n} b_{ii} = \alpha\mathsf{tr}(A)+\beta\mathsf{tr}(B)$$

■

Nous savons qu'en général, $AB \neq BA$ pour A et B dans $\mathcal{M}_n(\mathbb{K})$. Néanmoins, nous avons :

Théorème 3.6.29. *Soient A et B deux matrices carrées d'ordre n. On a*

$$\boxed{\mathsf{tr}(AB) = \mathsf{tr}(BA)}$$

Démonstration. En posant $A = (a_{ij})$ et $B = (b_{jk})$, on a $AB = (c_{ik})$ avec $c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}$. Donc les éléments diagonaux de AB sont les cœfficients $c_{ii} = \sum_{j=1}^{n} a_{ij}b_{ji}$. D'où

$$\mathsf{tr}(AB) = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}b_{ji}.$$

D'autre part, $BA = (d_{ik})$ avec $d_{ik} = \sum_{j=1}^{n} b_{ij}a_{jk}$. Par conséquent, les éléments diagonaux de BA sont les cœfficients $d_{ii} = \sum_{j=1}^{n} b_{ij}a_{ji}$. D'où

$$\mathsf{tr}(BA) = \sum_{i=1}^{n} d_{ii} = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij}a_{ji}.$$

On peut échanger les rôles de i et j, car ces indices sont muets; ce qui donne $\mathsf{tr}(AB) = \mathsf{tr}(BA)$. ■

Corollaire 3.6.30. *Deux matrices semblables ont la même trace.*

Démonstration. Soient A et B deux matrices semblables dans $\mathcal{M}_n(\mathbb{K})$. Il existe une matrice inversible $P \in \mathcal{M}_n(\mathbb{K})$ telle que $B = P^{-1}AP$. Il en résulte que $\mathsf{tr}(B) = \mathsf{tr}(P^{-1}AP) = \mathsf{tr}\left(P^{-1}(AP)\right) = \mathsf{tr}\left((AP)P^{-1}\right) = \mathsf{tr}\left(APP^{-1}\right) = \mathsf{tr}(A)$. ■

Si deux matrices n'ont pas la même trace, alors elles ne peuvent pas être semblables.

Exemple 3.6.31. Soient les deux matrices $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 5 & 6 & 7 \end{pmatrix}$ et $A = \begin{pmatrix} 2 & 1 & 3 \\ 4 & 0 & 5 \\ 6 & 5 & 7 \end{pmatrix}$. Puisque $\mathsf{tr}(A) = 12$ et $\mathsf{tr}(B) = 9$, ces deux matrices ne sont pas semblables.

Remarque 3.6.32. Contrairement à ce qu'on pourrait penser, la réciproque du corollaire 3.6.30 n'est pas vraie. Plus explicitement, deux matrices ayant la même trace ne sont pas en général semblables.
Il suffit de considérer les matrices $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ et $A = \begin{pmatrix} 2 & 1 \\ 4 & 0 \end{pmatrix}$.
On a $\mathsf{tr}(I_2) = \mathsf{tr}(A) = 2$, mais I_2 et A ne sont pas semblables. En effet, si elles étaient semblables, il existerait une matrice inversible P d'ordre 2 telle que $A = P^{-1}I_2P$, donc $A = I_2$, ce qui est faux.

Soit E un espace vectoriel de dimension finie n sur $\mathbb{K}$ et $\mathcal{B}$, $\mathcal{B}'$ deux bases de E. Si f est un endomorphisme de E, les matrices $A = M(f, \mathcal{B})$ et $A' = M(f, \mathcal{B}')$ sont semblables : $A' = P^{-1}AP$, où P est la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$. Par conséquent, $\mathsf{tr}(A) = \mathsf{tr}(A')$. Cet argument justifie la définition suivante :

Définition 3.6.33. *Soit f un endomorphisme d'un espace vectoriel E de dimension finie. La* ***trace*** *de f est définie comme la trace de sa matrice $M(f, \mathcal{B})$ dans n'importe quelle base $\mathcal{B}$ de E.*

Un exercice classique. *Montrer qu'il n'existe pas de matrices A, B dans $\mathcal{M}_n(\mathbb{K})$ telles que $AB - BA = I_n$.*
Supposons par l'absurde l'existence de telles matrices A et B. En appliquant la trace, on obtient $\mathsf{tr}(AB - BA) = \mathsf{tr}(I_n) = n$.
Or, $\mathsf{tr}(AB - BA) = \mathsf{tr}(AB) - \mathsf{tr}(BA) = 0$, d'où $0 = n$, ce qui est impossible.