ALGÈBRE LINÉAIRE 2

Table des matières

Introduction 2
0. Algèbre linéaire dans K^n 5
1. Espaces vectoriels sur un corps 19
2. Applications linéaires 23
3. Sommes directes 31
4. Systèmes générateurs et libres 35
5. Bases 41
6. Dimension 45
7. Rang 49
8. Matrices d'applications linéaires 53
9. Dualité 58
10. Déterminants 64
11. Diagonalisation 74
12. Produit scalaire 78
13. Quelques exercices 84

Date: Version du 1er décembre 2013.

Introduction

La notion de *ligne droite* est la base de l'algèbre linéaire. On dit plus simplement *droite* et on précise parfois *droite affine*. En fait, sur une droite affine, on peut marquer des points mais aucun ne joue un rôle privilégié *a priori*. Quand on fixe une origine O, on parle de *droite vectorielle* et on écrit parfois $\overrightarrow{D}$. On peut alors considérer le *vecteur* $\overrightarrow{OP}$ joignant O à P. Et réciproquement, à tout vecteur u de $\overrightarrow{D}$ correspond un unique point P, extrémité du vecteur u d'origine O. On écrit parfois $P = O + u$.

[DESSIN]

Si on fixe un vecteur de base e, tout vecteur u de $\overrightarrow{D}$ est un multiple de e, disons qu'on a $u = x \times e$ avec x un réel. On dit que x est la *composante* du vecteur u dans la *base* e sur la droite $\overrightarrow{D}$. Si $P = O + u$, on dit alors que x est la *coordonnée* de P dans le *repère* (O, e).

[DESSIN]

On s'intéresse à certaines *transformations (dites affines) de la droite* comme les *translations* et les *homothéties*. Une telle transformation est dite *vectorielle* ou *linéaire* si elle fixe l'origine. Une translation correspond à une application de la forme $x \mapsto x + \alpha$. Elle n'est pas linéaire si $\alpha \neq 0$. Une homothétie vectorielle correspond à une application $x \mapsto \lambda x$ avec $\lambda \neq 0$.

[DESSIN]

On peut étudier les droites d'un même *plan* (affine) . On parle de *plan vectoriel* si on fixe une *origine* O. Les droites vectorielles sont celles qui passent par O. Par un point, il passe une unique droite parallèle à une autre. Il suit que toute droite affine D s'obtient par translation à partir d'une unique droite vectorielle $\overrightarrow{D}$ passant par O, appelée *direction* de D.

[DESSIN]

On voit assez rapidement qu'il faut fixer deux vecteurs de base e_1 et e_2 pour pouvoir atteindre n'importe quel point du plan à partir de O en ajoutant des multiples de ces vecteurs. Tout vecteur u du plan s'écrira donc de manière unique $x_1e_1 + x_2e_2$ et (x_1, x_2) seront les *composantes* de u dans la *base* (e_1, e_2). Comme ci-dessus, on dira aussi que (x_1, x_2) sont les *coordonnées* du point P, extrémité de u dans le *repère* (O, e_1, e_2).

[DESSIN]

Si on fixe un repère, une droite D du plan est caractérisée par son (une) *équation*

$$ax + by = \alpha$$

avec $a \neq 0$ ou $b \neq 0$. En d'autre termes, D correspond à l'ensemble des solutions (x, y) de cette équation. Et sa direction $\overrightarrow{D}$ est donnée par l'équation *homogène* $ax + by = 0$.

[DESSIN]

Intersecter deux droites revient à résoudre un *système*

$$(S) : \left\{ \begin{array}{l} ax + by = \alpha \\ cx + dy = \beta \end{array} \right. .$$

On trouve soit un point, ou rien, ou une droite. L'intersection des directions s'obtient en résolvant le système homogène

$$(\overrightarrow{S}) : \left\{ \begin{array}{l} ax + by = 0 \\ cx + dy = 0 \end{array} \right. .$$

correspondant.

[DESSIN]

On peut s'intéresser aux *transformations (dites affines) du plan* comme les translations, symétries, homothéties, rotations... Quitte à faire une translation, une telle transformation préserve l'origine et est donnée par une *transformation linéaire*

$$\varphi : (x, y) \mapsto (ax + by, cx + dy).$$

[DESSIN]

Le *noyau* d'une transformation linéaire est formé de tous les points qui sont envoyés sur l'origine. Cet ensemble correspond au *noyau*

$$\ker \varphi := \{(x, y) \in \mathbb{R}^2, \varphi(x, y) = (0, 0)\}$$

de φ. On retrouve l'ensemble des solutions du système homogène $(\overrightarrow{S})$.

On peut aussi regarder une *projection* du plan sur une droite Δ. Si on demande toujours que l'origine soit préservée, celle-ci correspond à une *forme linéaire*

$$f : (x, y) \to ax + by.$$

[DESSIN]

La droite $\overrightarrow{D}$ d'équation $ax + by = 0$ est alors le « noyau » de la projection.

On remarque aussi que l'application linéaire φ ci-dessus est donnée par un couple de formes linéaires (f, g). Et que les solutions de $(\overrightarrow{S})$ correspondent à l'intersection des droites données par les noyaux des formes linéaires.

On identifie ainsi implicitement la droite avec $\mathbb{R}$, le plan avec $\mathbb{R}^2$ et on fera de même avec les espaces de dimension supérieure. Les *sous-espaces (affines)* correspondent à des ensembles de solutions de systèmes linéaires. Les transformations géométriques correspondent à des applications linéaires, à translation près. Les problèmes géométriques deviennent des questions algébriques.

Mais on peut aller encore plus loin en utilisant la notation matricielle. On définit un vecteur « colonne » ou un vecteur « ligne » comme un tableau

$$\left[\begin{array}{c} a \\ b \end{array} \right] \quad \text{ou} \quad \left[\begin{array}{cc} a & b \end{array} \right].$$

On dit aussi matrice 2×1 et matrice 1×2. De même, une matrice 2×2 est un tableau
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$
On identifiera une matrice 1×1 avec le réel a correspondant mais on pourrait aussi écrire $[a]$. On peut ajouter des matrices de même type en additionnant terme à terme, par exemple
$$A := \begin{bmatrix} a & b \\ c & d \end{bmatrix}, B := \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} \quad \text{et alors} \quad A + B = \begin{bmatrix} a+a' & b+b' \\ c+c' & d+d' \end{bmatrix}.$$
On peut aussi multiplier une matrice par une constante, par exemple
$$A := \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{et alors} \quad \lambda A = \begin{bmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{bmatrix}.$$
Mais toute la magie tient dans le fait qu'on peut aussi multiplier un vecteur ligne par un vecteur colonne
$$\begin{bmatrix} \alpha & \beta \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \alpha a + \beta b$$
pour obtenir un scalaire. On peut alors étendre cette méthode pour multiplier une matrice 2×2 à droite par une ligne
$$\begin{bmatrix} \alpha & \beta \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \alpha a + \beta c & \alpha b + \beta d \end{bmatrix}$$
ou à gauche par une colonne
$$\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} \alpha c + \beta d \\ \gamma c + \delta d \end{bmatrix}$$
et on peut aussi multiplier les matrices 2×2 entre elles...

Avec ces conventions, on voit ainsi que le système (S) ci-dessus peut se réécrire
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$
ou encore
$$AX = B \quad \text{avec} \quad A := \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad X := \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{et} \quad B := \begin{bmatrix} \alpha \\ \beta \end{bmatrix}.$$
On a ainsi ramené formellement un problème en deux variables en un problème en une variable ! A est *matrice de l'application linéaire* φ ci-dessus. Étudier A revient à étudier φ.

0. Algèbre linéaire dans K^n

Nous allons formaliser un peu tout ce qui précède en oubliant l'origine géométrique des objets. On va aussi travailler en dimension quelconque. On va même remplacer $\mathbb{R}$ par corps quelconque K qui pourra aussi bien être celui des complexes $\mathbb{C}$, des rationnels $\mathbb{Q}$, le corps à deux éléments $\mathbf{F}_2$ ou encore le corps des fractions rationnelles $\mathbb{R}(X)$. Un élément de K est un *scalaire*.

Il faut faire attention que des notions comme $a \geq b$ ou $a/2$ n'ont pas toujours de sens dans ce corps K. Par exemple, on ne peut pas dire si i est plus grand que 1 dans $\mathbb{C}$ et on ne peut pas diviser par 2 dans $\mathbf{F}_2$. On ne peut pas non plus extraire de racine en général...

Définition 0.1. *Si $n \in \mathbb{N}$, alors K^n désigne l'ensemble des n-uples $(x_1, \ldots, x_n)$ d'éléments de K. Un tel n-uple s'appelle aussi un* vecteur *de K^n et on dit que x_i est sa i-ème* composante *(mais on peut aussi voir $(x_1, \ldots, x_n)$ comme un* point *de K^n auquel cas on parle de* coordonnées*).*

Dans K^n, on dispose d'une addition *(terme à terme)*

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

qui fournit une somme *et d'une* multiplication externe *(terme à terme)*

$$\lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n)$$

qui fournit un produit.

Par exemple, K^2 peut s'identifier au plan comme on l'a vu dans l'introduction, K^1 n'est autre que K lui même que l'on identifie à la droite et K^0 est le singleton $\{0\}$ (par convention) qu'on peut voir comme un point.

On notera parfois tout simplement $0 = (0, \ldots, 0)$ si on pense au vecteur (on met parfois une flèche) ou $O := (0, \ldots, 0)$ si on le voit comme un point.

En général, on écrira

$$e_i := (0, \ldots, 0, 1, 0, \ldots, 0)$$

avec le 1 à la i-ème place. On dit parfois que $(e_1, \ldots, e_n)$ est la *base canonique de K^n* (tout ce qui suit est relatif à cette base). Remarquons en particulier que si $u := (x_1, \ldots, x_n)$, on a

$$x_1 e_1 + \cdots + x_n e_n = (x_1, 0, \ldots, 0) + \ldots + (0, \ldots, 0, x_n) = (x_1, \ldots, x_n) = u.$$

Définition 0.2. *Si $m, n \in \mathbb{N}$, une* matrice $n \times m$ *est un tableau à n lignes et m colonnes*

$$A := [a_{ij}] := \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}.$$

On dit vecteur ligne *si $n = 1$ et* vecteur colonne *si $m = 1$. La* matrice *de*

$$u := (x_1, \ldots, x_n) \in K^n$$

est le vecteur colonne

$$[u] := \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Enfin, on dit matrice carrée *si $m = n$.*

Pour $m = 0$ ou $n = 0$, on considère par convention qu'il y a une seule matrice, la matrice nulle. Pour $m = 2$ et/ou $n = 2$, on retrouve les matrices à 2 lignes et/ou 2 colonnes introduites plus haut. Et pour $m = n = 1$, on peut identifier matrices et scalaires (et on le fera).

Si $m, n \in \mathbb{N}$, on note $M_{n\times m}(K)$ l'ensemble des matrices $n \times m$. On écrira simplement $M_n(K)$ si $m = n$.

Définition 0.3. *La* somme *de $A := [a_{ij}]$ et $B := [b_{ij}]$ est $A + B := [a_{ij} + b_{ij}]$ et le* produit *de λ par A est $\lambda A := [\lambda a_{ij}]$.*

Ici encore, ce n'est qu'une généralisation de ce qu'on a vu dans la section précédente. D'autre part, la proposition suivante montre que le passage d'un n-uple au vecteur colonne correspondant est compatible avec nos opérations.

Proposition 0.4. *Si $u, v \in K^n$, on a $[u + v] = [u] + [v]$. Et si $\lambda \in K$, on a $[\lambda u] = \lambda[u]$.*

Démonstration. Si $u := (x_1, \ldots, x_n), v := (y_1, \ldots, y_n) \in K^n$, on a par définition $u + v = (x_1 + y_1, \ldots, x_n + y_n)$ et donc

$$[u + v] := \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}.$$

D'autre part, on a

$$[u] = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad [v] = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad \text{et donc} \quad [u] + [v] = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}.$$

De même, si $u := (x_1, \ldots, x_n) \in K^n$ et $\lambda \in K$, on $\lambda u = (\lambda x_1, \ldots, \lambda x_n)$ et donc

$$[\lambda u] := \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{bmatrix} = \lambda[u].$$

□

La définition suivante peut sembler artificielle mais elle est explicitée par la proposition qui suit.

Définition 0.5. *Une* forme linéaire *sur K^n est une application $f : K^n \to K$ qui satisfait :*

i) *Si $u, v \in K^n$, on a $f(u + v) = f(u) + f(v)$.*
ii) *Si $u \in K^n$ et $\lambda \in K$, alors $f(\lambda u) = \lambda f(u)$*

On note parfois $\check{K}^n$ l'ensemble de toutes les formes linéaires sur K^n.

Proposition 0.6. *Une application $f : K^n \to K$ est une forme linéaire si et seulement si*
$$f(x_1, \ldots, x_n) = a_1x_1 + \cdots + a_nx_n$$
avec $a_1, \ldots, a_n \in K$. De plus, $a_1, \ldots, a_n$ sont alors uniques.

Démonstration. On se donne tout d'abord
$$f : (x_1, \ldots, x_n) \mapsto a_1x_1 + \cdots + a_nx_n.$$
Si $u := (\alpha_1, \ldots, \alpha_n)$ et $v := (\beta_1, \ldots, \beta_n)$, on a
$$u + v := (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n)$$
et donc
$$\begin{aligned} f(u+v) &= a_1(\alpha_1 + \beta_1) + \cdots + a_n(\alpha_n + \beta_n) \\ &= a_1\alpha_1 + \cdots + a_n\alpha_n + a_1\beta_1 + \cdots + a_n\beta_n = f(u) + f(v). \end{aligned}$$
Et de même, si $u := (\alpha_1, \ldots, \alpha_n)$ et $\lambda \in K$, on a
$$\begin{aligned} f(\lambda u) = f(\lambda\alpha_1, \ldots, \lambda\alpha_n) &= a_1\lambda\alpha_1 + \cdots + a_n\lambda\alpha_n \\ &= \lambda(a_1\alpha_1 + \cdots + a_n\alpha_n) = \lambda f(u). \end{aligned}$$
Réciproquement, on se donne une application f satisfaisant ces conditions et on pose pour tout $i = 1, \ldots, n$, $a_i := f(e_i)$. Rappelons que si $u := (x_1, \ldots, x_n)$, on peut aussi écrire $u = x_1e_1 + \cdots + x_ne_n$ et on voit alors que
$$\begin{aligned} f(u) &= f(x_1e_1 + \cdots + x_ne_n) = f(x_1e_1) + \cdots + f(x_ne_n) \\ &= x_1f(e_1) + \cdots + x_nf(e_n) = a_1x_1 + \cdots + a_nx_n. \end{aligned}$$
□

Définition 0.7. *La* matrice *de f est le vecteur ligne*
$$[f] := \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}.$$

Par exemple, les projections
$$p_i : u := (x_1, \ldots, x_n) \mapsto x_1$$
sont des formes linéaires. La matrice de p_i est le vecteur ligne $\begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix}$ avec 1 à la i-ème place.

Rappelons que si $f, g : I \to K$ sont deux applications d'un ensemble quelconque vers K, leur *somme* est l'application
$$f + g \quad : \quad \begin{aligned} I &\longrightarrow K \\ a &\longmapsto f(a) + g(a) \end{aligned}$$
et le *produit* de f par $\lambda \in K$ est l'application
$$\lambda f \quad : \quad \begin{aligned} I &\longrightarrow K \\ a &\longmapsto \lambda f(a) \end{aligned}$$
En d'autres termes, on a par définition
$$(f + g)(a) = f(a) + g(a) \quad \text{et} \quad (\lambda f)(a) = \lambda f(a).$$

On appliquera cela au cas $I = K^n$. En particulier, on voit que si f est la forme linéaire

$$u := (x_1, \ldots, x_n) \mapsto f(u) := a_1x_1 + \cdots + a_nx_n,$$

on a

$$(a_1p_1 + \cdots a_np_n)(u) = a_1p_1(u) + \cdots a_np_n(u) = a_1x_1 + \cdots a_nx_n = f(u).$$

Ceci étant vrai pour tout $u \in K^n$, on a donc $f := a_1p_1 + \cdots + a_np_n$. On dit que $(p_1, \ldots, p_n)$ est la *base canonique* de $\check{K}^n$.

On montre maintenant que les opérations sur les formes linéaires sont compatibles avec les opérations analogues sur les vecteurs lignes.

Proposition 0.8. *Si f et g sont des formes linéaires sur K^n, alors $f+g$ aussi et on a $[f+g] = [f]+[g]$. De même, si $\lambda \in K$, alors λf est aussi une forme linéaire sur K^n et $[\lambda f] = \lambda[f]$.*

Démonstration. On se donne donc deux formes linéaires

$$f : (x_1, \ldots, x_n) \mapsto a_1x_1 + \cdots + a_nx_n \quad \text{et} \quad g : (x_1, \ldots, x_n) \mapsto b_1x_1 + \cdots + b_nx_n$$

et on a donc, par définition,

$$(f+g)(x_1, \ldots, x_n) = f(x_1, \ldots, x_n) + g(x_1, \ldots, x_n)$$

$$= a_1x_1 + \cdots + a_nx_n + b_1x_1 + \cdots + b_nx_n = (a_1+b_1)x_1 + \cdots + (a_n+b_n)x_n$$

si bien que $f+g$ est linéaire et

$$[f+g] = \begin{bmatrix} a_1+b_1 & \cdots & a_n+b_n \end{bmatrix} = [f]+[g].$$

De même, si

$$f : (x_1, \ldots, x_n) \mapsto a_1x_1 + \cdots + a_nx_n$$

est une forme linéaire et $\lambda \in K$, on a

$$(\lambda f)(x_1, \ldots, x_n) = \lambda f(x_1, \ldots, x_n)$$

$$= \lambda(a_1x_1 + \cdots + a_nx_n) = \lambda a_1x_1 + \cdots + \lambda a_nx_n$$

si bien que λf est linéaire et

$$[\lambda f] = \begin{bmatrix} \lambda a_1 & \cdots & \lambda a_n \end{bmatrix} = \lambda[f].$$

□

On peut multiplier les matrices entre elles pourvu qu'elles aient les bonnes dimensions.

Définition 0.9. *Soient $A := [a_{ij}]$ une matrice $n \times m$ et $B := [b_{ij}]$ une matrice $m \times l$. Leur produit AB est la matrice $n \times l$ avec $AB = [c_{ij}]$ où*

$$c_{ij} := \sum_{k=1}^{m} a_{ik}b_{kj} := a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj}.$$

On a donc le diagramme suivant

$$\begin{array}{cc}
\text{(CARRÉ)} & \left.\overbrace{\begin{bmatrix} b_{11} & \cdots & b_{1l} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{ml} \end{bmatrix}}^{l}\right\} m \\
n\left\{\underbrace{\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}}_{m}\right. & \left.\underbrace{\begin{bmatrix} c_{11} & \cdots & c_{1l} \\ \vdots & & \vdots \\ c_{n1} & \cdots & c_{nl} \end{bmatrix}}_{l}\right\} n
\end{array}.$$

ou chaque entrée de la matrice produit s'obtient en multipliant la ligne correspondante par la colonne correspondante.

On peut aussi définir la multiplication en deux étapes, tout d'abord un vecteur ligne par un vecteur colonne :

$$\begin{bmatrix} a_1 & \cdots & a_m \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = a_1 b_1 + \cdots + a_m b_m,$$

puis en écrivant la première matrice en superposant ses lignes et la seconde en juxtaposant ses colonnes :

$$\begin{bmatrix} L_1 \\ \vdots \\ L_n \end{bmatrix} \begin{bmatrix} C_1 & \cdots & C_l \end{bmatrix} = [L_i C_j].$$

Aussi, on peut remarquer que l'on a toujours

$$\begin{bmatrix} a_1 & \cdots & a_m \end{bmatrix} \begin{bmatrix} L_1 \\ \vdots \\ L_m \end{bmatrix} = a_1 L_1 + \cdots + a_m L_m$$

et

$$\begin{bmatrix} C_1 & \cdots & C_m \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = b_1 C_1 + \cdots + b_m C_m.$$

On définit aussi les *puissances* d'une matrice carrée en posant

$$A^r = \underbrace{A \times \cdots \times A}, \quad r \text{ fois}.$$

La multiplication d'un vecteur ligne par un vecteur colonne correspond à l'action de la forme linéaire sur le n-uple comme on le voit ci-dessous.

Proposition 0.10. *Si f est une forme linéaire sur K^n et $u \in K^n$, on a $[f][u] = f(u)$.*

Démonstration. On se donne donc une forme linéaire f de matrice $[f] = [a_1 \cdots a_n]$ sur K^n et le vecteur $u = (x_1, \ldots, x_n) \in K^n$. Par définition, on a donc $f(u) = a_1x_1 + \cdots + a_nx_n$ et bien sûr aussi

$$[f][u] = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = a_1x_1 + \cdots + a_nx_n.$$

□

Définition 0.11. *Si f une forme linéaire sur K^n et $\alpha \in K$, on dit que*

$$(E) \quad f(u) = \alpha$$

est une équation linéaire. Le noyau *de f est*

$$\ker f := \{u \in K^n, \quad f(u) = 0\}.$$

On dit que

$$H := \{u \in K^n, \quad f(u) = \alpha\}$$

*est l'*ensemble des solutions *de (E) et que (E) est une* équation *de H. Si f est non nulle, on dit que H est un* hyperplan affine. *Lorsque $\alpha = 0$, on dit que $H = \ker f$ est un* hyperplan vectoriel.

Bien sûr, dans le cas $n = 2$, un hyperplan correspond à une droite du plan. Pour $n = 1$, c'est un point de la droite. Dans le cas $n = 3$, on trouve un plan dans l'espace, donné par son équation $ax + by + cz = \alpha$. En général, si f est donnée par

$$(x_1, \ldots, x_n) \mapsto a_1x_1 + \cdots + a_nx_n,$$

alors H est *l'ensemble des solutions* de l'équation

$$a_1x_1 + \cdots + a_nx_n = \alpha.$$

Définition 0.12. *Si $f_1, \ldots, f_k$ sont des formes linéaires sur K^n et $\alpha_1, \ldots, \alpha_k \in K$, on dit que*

$$(S) : \left\{ \begin{array}{ccc} f_1(u) & = & \alpha_1 \\ \vdots & & \vdots \\ f_k(u) & = & \alpha_k \end{array} \right.$$

est un système linéaire. *Il est dit* homogène *si $\alpha_1 = \cdots = \alpha_m = 0$. Une solution de (S) est un vecteur u qui satisfait toutes les équations. Deux systèmes linéaires sont* équivalents *s'ils ont même ensemble de solutions.*

Définition 0.13. *Une partie non vide E de K^n est un* sous-espace vectoriel *si :*

i) $E \neq \emptyset$

ii) *si $u, v \in E$, alors $u + v \in E$*

iii) *si $u \in E$ et $\lambda \in K$, alors $\lambda u \in E$.*

Proposition 0.14. *Une partie E de K^n est un sous-espace vectoriel si et seulement si c'est l'ensemble des solutions d'un système linéaire homogène.*

En d'autres termes, un sous-espace vectoriel est une intersection d'hyperplans vectoriels.

Par convention, K^n est un sous-espace vectoriel de lui même en tant qu'intersection vide. De même, $\{O\}$ est l'intersection de tous les hyperplans vectoriels (il suffit de prendre les hyperplans des coordonnées). Et bien sûr, tout hyperplan vectoriel est un sous-espace vectoriel.

Démonstration. Si E est l'intersection d'hyperplans vectoriels définis par des formes linéaires non nulles $f_1, \ldots, f_k$, on a $(0, \ldots, 0) \in E$ et donc $E \neq \emptyset$. De plus, si $u, v \in E$, on a alors pour tout $i = 1, \ldots, k$,

$$f_i(u + v) = f_i(u) + f_i(v) = 0 + 0 = 0.$$

et donc $u + v \in E$. De même, si $u \in E$ et $\lambda \in K$, on a pour tout $i = 1, \ldots, k$,

$$f_i(\lambda u) = \lambda f_i(u) = \lambda 0 = 0.$$

et donc $\lambda u \in E$.

Pour la réciproque, on renvoie sur 9.1. □

On dira aussi qu'une partie non vide E de K^n est un *sous-espace affine* si c'est l'ensemble des solutions d'un système linéaire (pas nécessairement homogène). En d'autres termes, un espace affine est une intersection non vide d'hyperplans affines. En particulier, tout hyperplan affine est un sous-espace affine. En général, un espace affine s'obtient en *translatant* un espace vectoriel.

On généralise maintenant la notion de somme et de produit par un scalaire aux applications d'un ensemble quelconque I vers K^n.

Définition 0.15. *Si $f, g : I \to K^n$ sont deux applications de I vers K^n, leur* somme *est l'application*

$$f + g \quad : \quad \begin{aligned} I &\longrightarrow K^n \\ a &\longmapsto f(a) + g(a). \end{aligned}$$

De même, le produit *de f par $\lambda \in K$ est l'application*

$$\lambda f \quad : \quad \begin{aligned} I &\longrightarrow K^n \\ a &\longmapsto \lambda f(a). \end{aligned}$$

Définition 0.16. *Si $f : I \to K^n$ est une application, ses* composantes *sont les applications $f_i := p_i \circ f : I \to K$.*

On voit donc que $f(a) = (f_1(a), \ldots, f_n(a))$ lorsque $a \in I$.

Lemme 0.17. *Si $f, g : I \to K^n$ sont deux applications de I vers K^n, les composantes de $f + g$ sont les sommes $f_i + g_i$ des composantes. Et si $\lambda \in K$, les composantes de λf sont les produits λf_i des composantes par λ.*

Démonstration. Facile □

Définition 0.18. *Une application $f : K^m \to K^n$ est* linéaire *si :*

i) *Si $u, v \in K^m$, on a $f(u + v) = f(u) + f(v)$.*

ii) *Si $u \in K^m$ et $\lambda \in K$, alors $f(\lambda u) = \lambda f(u)$*

Proposition 0.19. *Une application $f : K^m \to K^n$ est linéaire si et seulement si ses composantes sont des formes linéaires.*

Démonstration. On note comme d'habitude, pour tout $i = 1, \dots, n$, f_i la i-ème composante de f. Pour $u, v \in K^m$, on a

$$f(u+v) = (f_1(u+v), \dots, f_n(u+v))$$

et

$$\begin{aligned} f(u) + f(v) &= (f_1(u), \dots, f_n(u)) + (f_1(v), \dots, f_n(v)) \\ &= (f_1(u) + f_1(v), \dots, f_n(u) + f_n(v)). \end{aligned}$$

De même, si $u \in K^m$ et $\lambda \in K$, on a

$$f(\lambda u) = (f_1(\lambda u), \dots, f_n(\lambda u))$$

et

$$\lambda f(u) = \lambda(f_1(u), \dots, f_n(u)) = (\lambda f_1(u), \dots, \lambda f_n(u)).$$

On voit donc que f satisfait les conditions ci-dessus si et seulement si toutes ses composantes les satisfont aussi. $\square$

Dans le cas ou $n = 1$, une application linéaire n'est rien d'autre qu'une forme linéaire sur K^m. Pour $m = n = 2$, on retrouve les transformations de plan étudiées dans l'introduction.

Définition 0.20. *Soit $f : K^m \to K^n$ une application linéaire. Si pour tout $i = 1, \dots, n$, on a*

$$[f_i] =: \begin{bmatrix} a_{i1} & \cdots & a_{im} \end{bmatrix},$$

alors la matrice *de f est*

$$A := [f] := \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \in M_{n \times m}(K).$$

On peut remarquer que la matrice de f s'obtient en superposant les vecteurs lignes des f_i comme ceci :

$$[f] = \begin{bmatrix} [f_1] \\ \vdots \\ [f_n] \end{bmatrix}.$$

Si on note, comme d'habitude, $e_j := (0, \dots, 0, 1, 0, \dots, 0)$ avec le 1 à la j-ème place, on a

$$f(e_j) = (a_{1j}, \dots, a_{nj})$$

et on peut donc aussi écrire f en alignant les vecteurs colonnes des $f(e_j)$:

$$[f] = \begin{bmatrix} [f(e_1)] & \cdots & [f(e_m)] \end{bmatrix}.$$

Remarquons enfin que la matrice de l'identité est la *matrice unité*

$$I_n := \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}.$$

Définition 0.21. *Le* noyau *de* $f : K^m \to K^n$ *est*

$$\ker f = \{u \in K^m, f(u) = 0\}.$$

Attention, ici on a $0 = (0, \ldots, 0)$ et on voit donc que $f(u) = 0$ si et seulement si, pour tout $i = 1, \ldots, n$, on a $f_i(u) = 0$. En d'autres termes, le *noyau* de f est l'intersection des noyaux de ses composantes f_i : on a

$$\ker f = \bigcap_{i=1}^{n} \ker f_i.$$

Remarquons aussi qu'un sous-espace vectoriel E de K^n n'est autre que le noyau d'une application linéaire.

La compatibilité vue plus haut des opérations sur les matrices avec les formes linéaires s'étend en fait au cas des applications linéaires.

Proposition 0.22. *Soit* $f : K^m \to K^n$ *une application linéaire.*

i) *Si* $u \in K^m$, *on a* $[f][u] = [f(u)]$.
ii) *Si* $g : K^m \to K^n$ *est une autre application linéaire,* $f + g$ *est linéaire et* $[f + g] = [f] + [g]$.
iii) *Si* $\lambda \in K$, *alors* λf *est linéaire et* $[\lambda f] = \lambda[f]$.
iv) *Si* $g : K^l \to K^m$ *est une autre application linéaire, alors* $f \circ g$ *est linéaire et* $[f \circ g] = [f][g]$.

Démonstration. Cela se fait encore en se ramenant aux composantes. Tout d'abord, on a si $u \in K^m$,

$$[f][u] = \begin{bmatrix} [f_1] \\ \vdots \\ [f_n] \end{bmatrix} [u] = \begin{bmatrix} f_1(u) \\ \vdots \\ f_n(u) \end{bmatrix} = [f(u)]$$

Ensuite, si $g : K^m \to K^n$ est une autre application linéaire, alors $f + g$ est linéaire car ses composantes $f_i + g_i$ le sont et on a

$$[f + g] = \begin{bmatrix} [f_1 + g_1] \\ \vdots \\ [f_n + g_n] \end{bmatrix} = \begin{bmatrix} [f_1] \\ \vdots \\ [f_n] \end{bmatrix} + \begin{bmatrix} [g_1] \\ \vdots \\ [g_n] \end{bmatrix} = [f] + [g]$$

De même, si $\lambda \in K$, alors λf est linéaire car ses composantes λf_i le sont et on a

$$[\lambda f] = \begin{bmatrix} [\lambda f_1] \\ \vdots \\ [\lambda f_n] \end{bmatrix} = \lambda \begin{bmatrix} [f_1] \\ \vdots \\ [f_n] \end{bmatrix} = \lambda[f]$$

On termine par la composition avec une application $g : K^l \to K^m$. On suppose tout d'abord que $n = 1$ si bien que f est une forme linéaire

$$(x_1, \ldots, x_m) \mapsto a_1 x_1 + \cdots + a_m x_m.$$

et on a donc pour tout $u \in K^l$,

$$(f \circ g)(u) = f(g(u)) = a_1 g_1(u) + \cdots + a_m g_m(u) = (a_1 g_1 + \cdots + a_m g_m)(u).$$

En d'autres termes, $f \circ g$ est une forme linéaire et $f \circ g = a_1 g_1 + \cdots + a_m g_m$. On a donc bien

$$[f \circ g] = [a_1 g_1 + \cdots + a_m g_m] = a_1[g_1] + \cdots + a_m[g_m]$$
$$= \begin{bmatrix} a_1 & \cdots & a_m \end{bmatrix} \begin{bmatrix} [g_1] \\ \vdots \\ [g_m] \end{bmatrix} = [f][g].$$

En général, on remarque que les composantes de $f \circ g$ sont

$$p_i \circ (f \circ g) = (p_i \circ f) \circ g = f_i \circ g$$

et il suit que $f \circ g$ sera linéaire et que

$$[f \circ g] = \begin{bmatrix} [f_1 \circ g] \\ \vdots \\ [f_n \circ g] \end{bmatrix} = \begin{bmatrix} [f_1][g] \\ \vdots \\ [f_n][g] \end{bmatrix} = \begin{bmatrix} [f_1] \\ \vdots \\ [f_n] \end{bmatrix} [g] = [f][g].$$

□

Corollaire 0.23. *On considère*

i) *le système*

$$(S) : \left\{ \begin{array}{ccc} a_{11}x_1 + \cdots + a_{1m}x_m & = & b_1 \\ \vdots & & \vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m & = & b_n \end{array} \right. ,$$

ii) *l'application linéaire*

$$\begin{array}{ccc} K^m & \xrightarrow{f} & K^n \\ (x_1, \ldots, x_m) & \longmapsto & (a_{11}x_1 + \cdots + a_{1m}x_m, \ldots, a_{n1}x_1 + \cdots + a_{nm}x_m) \end{array}$$

ainsi que les vecteurs $u := (x_1, \ldots, x_n)$ *et* $v := (b_1, \ldots, b_n)$.

iii) *les matrices*

$$A := \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}, \quad B := \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \quad \text{et} \quad X := \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}.$$

Alors, les assertions suivantes sont équivalentes

i) $(x_1, \ldots, x_n)$ *est solution du système* (S)

ii) $f(u) = v$.

iii) $AX = B$

Démonstration. Le deux premières conditions sont clairement équivalentes et il résulte de la proposition précédente qu'elles sont aussi équivalentes à la troisième. □

Définition 0.24. *Avec les notations du corollaire, on dit que la matrice* $n \times (m+1)$ *obtenue en* concaténant A *et* B, *c'est à dire*

$$A := \begin{bmatrix} a_{11} & \cdots & a_{1m} & b_1 \\ \vdots & & \vdots & \\ a_{n1} & \cdots & a_{nm} & b_n \end{bmatrix},$$

est la matrice du système (S).

Définition 0.25. *Une matrice $A \in M_n(K)$ est* inversible *s'il existe une matrice que l'on note $A^{-1} \in M_n(K)$ telle que $AA^{-1} = A^{-1}A = I$.*

Remarquons que A^{-1}, si elle existe, est unique (exercice !). Mais la matrice

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

par exemple, n'est pas inversible (exercice !), contrairement à

$$A := \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}$$

dont l'inverse est

$$A^{-1} = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix} \text{(exercice !)}.$$

Proposition 0.26. *Soit $f : K^n \to K^n$ une application linéaire. Alors, f est bijective si et seulement si $[f]$ est inversible et dans ce cas, f^{-1} est linéaire et $[f^{-1}] = [f]^{-1}$.*

Démonstration. Si f est bijective, il existe pour tout $j = 1, \ldots, n$ un unique vecteur $v_j := (b_{1j}, \ldots, b_{nj})$ tel que $f(v_j) = e_i$ et on peut donc considérer la matrice $[b_{ij}]$. Si g est l'application linéaire correspondante, on a $g(e_j) = v_j$ pour tout $j = 1, \ldots, n$. Si $u := (a_1, \ldots, a_n) \in K^n$, on sait que $u = a_1e_1 + \cdots + a_ne_n$ et il suit que

$$\begin{aligned} g(u) &= g(a_1e_1 + \cdots + a_ne_n) \\ &= a_1g(e_1) + \cdots + a_ng(e_n) = a_1v_1 + \cdots + a_nv_n \end{aligned}$$

si bien que

$$\begin{aligned} (f \circ g)(u) &= f(a_1v_1 + \cdots + a_nv_n) \\ &= a_1f(v_1) + \cdots + a_nf(v_n) = a_1e_1 + \cdots + a_ne_n = u. \end{aligned}$$

On voit donc que g n'est autre que l'application réciproque f^{-1} de f qui est donc bien linéaire, et comme $f^{-1} \circ f = f \circ f^{-1} = \mathrm{Id}_{K^n}$, on a $[f][f^{-1}] = [f^{-1}][f] = I_n$.

Réciproquement, si $[f]$ est inversible, on considère l'application linéaire g dont la matrice est $[f]^{-1}$. On a alors $[f \circ g] = [f][g] = [f][f]^{-1} = I_n$ et il suit que $f \circ g = \mathrm{Id}_{K^n}$ et de la même manière, on obtient que $g \circ f = \mathrm{Id}_{K^n}$. Il suite que g est la réciproque de f et donc que f est bijective. $\square$

Si A est une matrice inversible et $AX = B$, alors $X = A^{-1}B$. Donc pour résoudre un système carré d'ordre n,

$$(S) : \left\{ \begin{array}{ccc} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ \vdots & & \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n & = & b_n \end{array} \right. ,$$

et si on sait que sa matrice $A := [a_{ij}]$ est inversible et que $A^{-1} = [c_{ij}]$, on aura une unique solution donnée par

$$(S) : \left\{ \begin{array}{ccc} x_1 & = & c_{11}b_1 + \cdots + c_{1n}b_n \\ \vdots & & \vdots \\ x_n & = & c_{n1}b_1 + \cdots + c_{nn}b_n \end{array} \right. .$$

En pratique, ça ne nous avance pas à grand chose car on ne sait pas vraiment inverser des matrices sans résoudre de système *avant*.

Mais par exemple, pour résoudre le système

$$\left\{ \begin{array}{rcl} 5x+2y & = & 2 \\ 2x+y & = & 1 \end{array} \right. ,$$

qui peut s'écrire $AX = B$, il suffit de calculer

$$X = A^{-1}B = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

La dernière partie de cette section est consacrée à la *méthode du pivot de Gauss* pour résoudre les systèmes linéaires.

Définition 0.27. *Les* opérations élémentaires *sur les lignes d'une matrice sont les suivantes*

i) *Échanger deux lignes*

ii) *Multiplier une ligne par une constante non nulle*

iii) *Ajouter à une ligne un multiple d'une autre ligne*

Le but (ultime) de la méthode de Gauss est d'effectuer une suite d'opérations élémentaires sur les lignes d'une matrice afin de transformer celle ci en une matrice dite échelonnée :

$$\begin{bmatrix}
0 & \cdots & 0 & 1 & * & \cdots & * & 0 & & 0 & * & \cdots & * \\
\vdots & & \vdots & 0 & \cdots & \cdots & 0 & 1 & & \vdots & \vdots & & \vdots \\
\vdots & & \vdots & \vdots & & & \vdots & 0 & \ddots & 0 & \vdots & & \vdots \\
\vdots & & \vdots & \vdots & & & \vdots & \vdots & & 1 & * & \cdots & * \\
\vdots & & \vdots & \vdots & & & \vdots & \vdots & & 0 & \cdots & \cdots & 0 \\
\vdots & & \vdots & \vdots & & & \vdots & \vdots & & \vdots & & & \vdots \\
0 & \cdots & 0 & 0 & \cdots & \cdots & 0 & 0 & & 0 & \cdots & \cdots & 0
\end{bmatrix}$$

sur laquelle on peut « lire » les solutions. Par exemple, avec un système de trois équations à 3 inconnues, on essaiera de faire apparaître successivement

$$\begin{bmatrix}
1 & & 0 & \rightarrow & 0 & \alpha \\
\downarrow & & & & \uparrow & \\
0 & & 1 & & 0 & \beta \\
\downarrow & \nearrow & \downarrow & & \uparrow & \\
0 & & 0 & \leftarrow & 1 & \gamma
\end{bmatrix}.$$

auquel cas, le système correspondant aura l'unique solution $x = \alpha$, $y = \beta$ et $z = \gamma$.

On peut aussi utiliser la méthode du pivot de Gauss pour trouver l'inverse d'une matrice.

Définition 0.28. *Une matrice élémentaire est une matrice obtenue en effectuant une opération élémentaire sur les lignes d'une matrice unité I.*

Par exemple, les matrices 2×2 élémentaires sont les matrices de la forme

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} \lambda & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ \lambda & 1 \end{bmatrix}, \begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix}.$$

Proposition 0.29. *Effectuer une opération élémentaire sur les lignes d'une matrice revient à multiplier à gauche par la matrice élémentaire correspondante.*

Démonstration. C'est un peu laborieux mais allons y. En général, les lignes de la matrice associée à une application linéaire f sont les vecteurs lignes $[f_1], \ldots, [f_n]$ ou $f_1, \ldots, f_n$ désignent les composantes de f. En particulier, comme les composantes de Id_{K^n} sont les projections $p_1, \ldots, p_n$, on voit que les lignes de I_n sont les matrices $[p_1], \ldots, [p_n]$. Et comme on a, pour tout $i = 1, \ldots n$, $p_i \circ f = f_i$, on voit que $[p_i][f] = [f_i]$. En d'autres termes, la i-ème ligne d'une matrice s'obtient en multipliant à gauche par la i-ème ligne de la matrice unité.

On considère alors les différents types d'opération élémentaire. Si on échange $[p_i]$ et $[p_j]$, ça va bien échanger $[f_i]$ et $[f_j]$. Et si on multiplie $[p_i]$ par λ, ça aura clairement le même effet sur $[f_i]$. Enfin, si on ajoute $\lambda[p_j]$ à $[p_i]$ et qu'on multiplie ensuite par f, on obtient bien

$$\begin{aligned}([p_i] + \lambda[p_j])[f] = [p_i + \lambda p_j][f] = [(p_i + \lambda p_j) \circ f] = [p_i \circ f + \lambda p_j \circ f] \\ = [p_i \circ f] + \lambda[p_j \circ f] = [f_i] + \lambda[f_j].\end{aligned}$$

□

Par exemple, on a

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix},$$

$$\begin{bmatrix} \lambda & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \lambda a & \lambda b \\ c & d \end{bmatrix},$$

et

$$\begin{bmatrix} 1 & 0 \\ \lambda & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c + \lambda a & d \end{bmatrix}.$$

Proposition 0.30. *Effectuer une opération élémentaire sur les lignes (de la matrice) d'un système ne change pas l'ensemble des solutions du système.*

Démonstration. Il s'agit essentiellement de montrer que chacune de ces opérations est réversible. Si on a échangé deux lignes, il suffit de les échanger à nouveau. Si on a multiplié une ligne par une constante non nulle λ, il suffit de multiplier celle-si par $1/\lambda$ (d'où l'importance de l'hypothèse $\lambda \neq 0$). Enfin, si on a rajouté à la j-ème ligne le produit par λ de la i-ème, il suffit de retrancher ce même multiple, c'est à dire, ajouter le produit par $-\lambda$ de la i-ème ligne à la j-ème. □

Proposition 0.31. *Si on applique à I les mêmes opérations élémentaires, et dans le même ordre, que celles qui permettent de transformer A en I, on obtient A^{-1}.*

Bien sûr, cela ne peut fonctionner que si la matrice A est inversible.

Démonstration. On suppose donc que l'on obtient I en multipliant à gauche A successivement par des matrices élémentaires $E_1, \ldots, E_r$ et on note B la matrice obtenue en multipliant I successivement par les mêmes matrices dans le même ordre. En d'autres termes, on a

$$E_r \cdots E_2 E_1 A = I \quad \text{et} \quad B := E_r \cdots E_2 E_1 I$$

si bien que $B := E_r \cdots E_2 E_1$ et donc $BA = I$. On remarque que les matrices élémentaires sont inversibles (leur inverse étant la matrice correspondant à l'opération élémentaire inverse). Il suit que B est inversible et comme on a $BA = I$, A est aussi inversible d'inverse B. $\square$

En pratique, on applique la méthode du pivot de Gauss à la matrice $[A\ I]$ obtenue en concaténant A et I. Par exemple, pour inverser la matrice

$$A := \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix},$$

on retranche deux fois la première ligne à la seconde dans

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

pour obtenir

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

et on ajoute deux fois la seconde à la première avant de multiplier la ligne du bas par -1, ce qui donne

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

et on a donc

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

1. Espaces vectoriels sur un corps

On fixe un corps K (qu'on pourra prendre égal à $\mathbb{R}$ ou $\mathbb{C}$).

Définition 1.1. *Un* espace vectoriel *sur* K *est un ensemble* E *muni d'une* addition

$$\begin{aligned} E \times E &\longrightarrow E \\ (u,v) &\longmapsto u+v \end{aligned}$$

fournissant une somme *telle que*

i) *si* $u, v \in E$, *alors* $u + v = v + u$ *(commutativité)*
ii) *si* $u, v, w \in E$, *alors* $(u + v) + w = u + (v + w)$ *(associativité)*
iii) *il existe* $0 \in E$ *tel que pour tout* $u \in E$, $u + 0 = u$ *(zéro)*
iv) *si* $u \in E$, *il existe* $-u \in E$ *tel que* $u + (-u) = 0$ *(opposé)*

et d'une multiplication externe

$$\begin{aligned} K \times E &\longrightarrow E \\ (\lambda,u) &\longmapsto \lambda u \end{aligned}$$

fournissant un produit *telle que*

v) *si* $u \in E$, *alors* $1u = u$
vi) *si* $u \in E$ *et* $\lambda, \mu \in K$, *alors* $(\lambda + \mu)u = \lambda u + \mu u$
vii) *si* $u \in E$ *et* $\lambda, \mu \in K$, *alors* $\lambda(\mu u) = (\lambda\mu)u$
viii) *si* $u, v \in E$ *et* $\lambda \in K$, *alors* $\lambda(u + v) = \lambda u + \lambda v$.

Les éléments de E s'appelleront des *vecteurs*. On écrit parfois 0_E pour préciser qu'il s'agit du zéro de E (on verra plus tard que le zéro est unique).

Notations : si $u, v, w \in E$, on écrit $u + v + w := (u + v) + w$; si $u, v \in E$, on écrit $u - v := u + (-v)$ (on verra plus tard que l'opposé est unique) ; si $u \in E$ et $\lambda, \mu \in K$, on écrit $\lambda\mu u := (\lambda\mu)u$.

Nos premiers exemples viennent de la section précédente.

Tout d'abord, K^n (avec les opérations définies plus haut) est un espace vectoriel comme on le vérifie aisément. On a

$$0_{K^n} = (0, \dots, 0) \quad \text{et} \quad -(x_1, \dots, x_n) = (-x_1, \dots, -x_n).$$

Plus généralement, tout sous-espace vectoriel de K^n (défini par exemple comme l'ensemble des solutions d'un système linéaire homogène), muni des lois induites, est un espace vectoriel. Cela vaut en particulier pour les hyperplans et donc pour les droites du plan.

L'espace $\check{K}^n$ des formes linéaires sur K (définition ci-dessus) est aussi un espace vectoriel (pour les opérations définies ci-dessus). Plus généralement, l'ensemble $L(K^m, K^n)$ de toutes les applications linéaires de K^m vers K^n est un espace vectoriel : le zéro est l'application nulle $u \mapsto 0$ et l'opposée d'un application f est l'application $-f : u \mapsto -f(u)$; de même pour $M_{n\times m}(K)$ et donc en particulier, pour l'ensemble des vecteurs lignes ou des vecteurs colonnes de longueur donnée. Le zéro est la matrice dont toutes les entrées sont nulles et l'opposé d'une matrice est la matrice obtenue en prenant les opposés de toutes les entrées.

Si I est un ensemble quelconque, l'ensemble K^I des applications de I dans K (muni des opérations définies plus haut) est un espace vectoriel. Si I est un intervalle de $\mathbb{R}$, les fonctions continues à valeurs dans $\mathbb{R}$ définies sur I forment un espace vectoriel $\mathcal{C}(I)$ pour les lois induites. Il en va de même de l'ensemble $\mathcal{C}^1(I)$ des fonctions continûment différentiables lorsque I est ouvert. On peut aussi regarder les fonctions absolument intégrables, etc.

Les familles d'éléments de K indexées par un ensemble I forment aussi un espace vectoriel que l'on identifiera à K^I : il revient au même de se donner la famille $(x_i)_{i\in I}$ ou l'application

$$\begin{aligned} I &\longrightarrow K \\ i &\longmapsto x_i \end{aligned}$$

L'addition et la multiplication se font terme à terme. Comme cas particulier, on trouve l'espace vectoriel $K^{\mathbb{N}}$ des suites indexées par $\mathbb{N}$. On peut aussi regarder les suites convergentes dans $\mathbb{R}$, de même que les suites absolument sommables, etc. On peut aussi considérer les suites de Cauchy sur $\mathbb{Q}$ qui forment un espace vectoriel sur $\mathbb{Q}$.

Les polynômes sur K forment aussi un espace vectoriel $K[X]$.

On peut considérer $\mathbb{C}$ comme espace vectoriel sur lui même, sur $\mathbb{R}$ ou sur $\mathbb{Q}$.

Lemme 1.2. *Soit E un espace vectoriel sur K.*

i) *Il existe un unique élément $0 \in E$ satisfaisant la condition donnée.*
ii) *Étant donné un élément $u \in E$ il existe un unique élément $-u$ satisfaisant la condition donnée.*
iii) *Si $u, v \in E$, la condition $u + v = u$ implique $v = 0$.*

Démonstration. Notons dès à présent que la troisième assertion implique la première. Notons aussi que si $u, v \in E$, alors

$$v = v + 0 = v + (u + (-u)) = (v + u) + (-u) = (u + v) + (-u).$$

On montre maintenant la seconde assertion : si $u + v = 0$, alors

$$v = 0 + (-u) = (-u) + 0 = -u.$$

Et on montre aussi la troisième : si $u + v = u$, alors

$$v = u + (-u) = 0.$$

□

Proposition 1.3. *Soit E un espace vectoriel sur K.*

i) *Si $u \in E$ et $\lambda \in K$, alors*
$$\lambda u = 0 \Leftrightarrow (\lambda = 0 \text{ ou } u = 0),$$
ii) *si $u \in E$ et $\lambda \in K$, alors $\lambda(-u) = (-\lambda)u = -(\lambda u)$,*
iii) *si $u, v \in E$ et $\lambda \in K$, alors $\lambda(u - v) = \lambda u - \lambda v$,*
iv) *si $u \in E$ et $\lambda, \mu \in K$, alors $(\lambda - \mu)u = \lambda u - \mu u$,*
v) *si $u, v \in E$ et $\lambda \in K^\times$ sont tels que $\lambda u = \lambda v$, alors $u = v$.*

Démonstration. On a
$$0u = (0+0)u = 0u + 0u$$
et donc $0u = 0$. On montre de même que $\lambda 0 = 0$. Réciproquement, supposons que $\lambda u = 0$ avec $\lambda \neq 0$. Alors,
$$u = 1u = (\lambda^{-1}\lambda)u = \lambda^{-1}(\lambda u) = \lambda^{-1}0 = 0.$$
D'où la première assertion.

On a
$$\lambda u + \lambda(-u) = \lambda(u + (-u)) = \lambda 0 = 0$$
et donc $\lambda(-u) = -(\lambda u)$. De même, on a
$$\lambda u + (-\lambda)u = (\lambda + (-\lambda))u = 0u = 0$$
et donc $(-\lambda)u = -(\lambda u)$, ce qui nous donne la seconde assertion.

On poursuit. On a
$$\lambda(u - v) = \lambda(u + (-v)) = \lambda u + \lambda(-v) = \lambda u + (-\lambda v) = \lambda u - \lambda v$$
et de la même manière,
$$(\lambda - \mu)u = \lambda u + (-\mu)u = \lambda u + (-\mu u) = \lambda u - \mu u.$$

Enfin, si $\lambda u = \lambda v$, alors
$$\lambda(u - v) = \lambda u - \lambda v = 0$$
et si $\lambda \neq 0$, on a nécessairement $u - v = 0$ et alors $u = v$. □

On peut maintenant définir la notion de sous-espace vectoriel d'un espace vectoriel quelconque.

Définition 1.4. *Un* sous espace vectoriel *d'un espace vectoriel E est une partie F de E telle que*

i) $F \neq \emptyset$

ii) *si $u, v \in F$, alors $u + v \in F$*

iii) *si $\lambda \in K$ et $u \in F$, alors $\lambda u \in F$.*

Par exemple, tout espace vectoriel E est un sous-espace vectoriel de lui même et la partie réduite au vecteur nul $\{0\}$ est un sous-espace vectoriel de E.

On dispose bien sûr des sous-espaces vectoriels de K^n décrits dans la section précédente, et donc en particulier des droites vectorielles du plan et plus généralement des hyperplans.

Si on regarde nos exemples d'espaces vectoriels, on voit que si I est un intervalle de $\mathbb{R}$, alors $\mathcal{C}^1(I)$ est un sous-espace vectoriel de $\mathcal{C}(I)$ qui est lui même un sous-espace vectoriel de $\mathbb{R}^I$. Ou encore, on voit que les suites convergentes forment un sous-espace vectoriel de $\mathbb{R}^{\mathbf{N}}$. Aussi, les polynômes de degré au plus N forment un sous-espace vectoriel $K[X]_{\leq N}$ de $K[X]$.

Mais il faut être prudent car, dans $\mathbb{R}^2$, la droite d'équation $x + y = 1$ n'est pas un sous-espace vectoriel, le cercle $x^2 + y^2 = 1$ non plus et l'équation $xy = 0$ ne définit pas un sous-espace vectoriel, c'est l'union de deux droites. Attention cependant, les solutions de $x^2 + y^2 = 0$ forment bien un sous-espace vectoriel de $\mathbb{R}^2$: c'est tout simplement $\{(0,0)\}$. Mais on obtient pas un sous-espace vectoriel de $\mathbb{C}^2$ (vu comme espace vectoriel sur $\mathbb{C}$) car on a $(1,i)$ et $(i,1)$ dedans mais pas $(1+i,i+1)$).

Notons enfin que sur $\mathbf{F}_2$, la troisième condition est toujours satisfaite si les deux premières le sont.

Proposition 1.5. *Soit E un espace vectoriel sur K.*

i) *Si F est un sous-espace vectoriel de E, alors l'addition et la multiplication par les scalaires font de F un espace vectoriel sur K.*

ii) *Un sous ensemble F de E est un sous espace vectoriel si et seulement si*
   (a) $0 \in F$
   (b) *lorsque $u, v \in F$ et $\lambda, \mu \in K$, on a $\lambda u + \mu v \in F$.*

Démonstration. Pour la première assertion, toutes les propriété sont héritées de E à part les deux qui font intervenir la notion d'existence. Plus précisément, il suffit de montrer que le vecteur nul est dans F et que l'opposé d'un élément de F est toujours dans F. Or, si $u \in F$, on a bien $-u = (-1)u \in F$ et il suit, comme F est non vide, que l'on a aussi $0 = u + (-u) \in F$.

Montrons maintenant la seconde assertion. On vient de voir que si F est un sous-espace vectoriel de E, la condition a) est toujours satisfaite et on montre facilement que la condition b) l'est aussi : si $u, v \in F$ et $\lambda, \mu \in K$, on a $\lambda u \in F$ et $\mu v \in F$ si bien que $\lambda u + \mu v \in F$. Réciproquement, si ces conditions sont satisfaites, on voit que F est non vide et les deux autres propriétés s'obtiennent à partir de la condition b) en faisant tout d'abord $\lambda = \mu = 1$ pour la première puis $\mu = 0$ pour l'autre. $\square$

2. Applications linéaires

Nous pouvons maintenant considérer des applications linéaires entre différents espaces vectoriels (sur un même corps K) :

Définition 2.1. *Une application $f : E \to F$ entre deux espaces vectoriels sur K est* linéaire *si elle satisfait :*

i) *Si $u, v \in E$, on a $f(u+v) = f(u) + f(v)$.*

ii) *Si $u \in E$ et $\lambda \in K$, alors $f(\lambda u) = \lambda f(u)$*

Leur ensemble se note $\mathcal{L}(E, F)$.

Si $F = E$, on dit que f est un endomorphisme *de E (ou un* opérateur *sur E) et on note $L(E)$ leur ensemble. Si $F = K$, on dit que f est une* forme linéaire *sur E et on note $\check{E}$ leur ensemble, aussi appelé* dual *de E. Une application linéaire bijective est un* isomorphisme. *Un endomorphisme qui est aussi un isomorphisme est un* automorphisme. *Leur ensemble se note $GL(E)$.*

En général, on dispose toujours de l'identité

$$Id_E \quad : \quad \begin{array}{c} E \longrightarrow E \\ x \longmapsto x \end{array}$$

qui est un automorphisme et de l'application nulle

$$0_{E,F} \quad : \quad \begin{array}{c} E \longrightarrow F \\ x \longmapsto 0 \end{array}$$

qui est l'unique application linéaire constante.

On définit une *homothétie* comme une application de E dans lui même de la forme $u \longmapsto ku$ avec $k \neq 0, 1$. C'est un automorphisme. Aussi, si F est un sous-espace vectoriel de E, l'application d'inclusion $F \hookrightarrow E$ est linéaire.

Un endomorphisme f de E est une *symétrie* si $f \circ f = \mathrm{Id}$ et une *projection* si $f \circ f = f$. Dans le plan, on peut considérer par exemple la symétrie par rapport à l'origine ou a une droite vectorielle, ou la projection sur une droite vectorielle par rapport à une autre.

Nous avons vu que les applications linéaires $f : K^m \to K^n$ sont les applications de la forme

$$(x_1, \ldots, x_m) \mapsto (a_{11}x_1 + \cdots + a_{1m}x_m, \ldots, a_{n1}x_1 + \cdots + a_{nm}x_m).$$

En d'autre terme, on a une bijection

$$\begin{array}{c} L(K^m, K^n) \xrightarrow{\simeq} M_{n\times m}(K) \\ f \longmapsto [f] \end{array}$$

Bien sûr, cela s'applique en particulier aux formes linéaires.

Comme cas particulier, on peut regarder les rotations

$$(x, y) \mapsto (x \cos\theta - y \sin\theta, x \sin\theta + y \cos\theta).$$

de $\mathbb{R}^2$. Si on identifie comme d'habitude $\mathbb{R}^2$ et $\mathbb{C}$, une rotation n'est autre que l'homothétie $z \mapsto e^{i\theta}z$ de $\mathbb{C}$ vu comme espace vectoriel sur lui même.

Mais les applications de $\mathbb{R}^2$ dans lui même :

$$(x,y) \mapsto (x+1, y-1) \quad \text{et} \quad (x,y) \mapsto (x^2, y^2),$$

par exemple, ne sont pas linéaires.

Si I est un ensemble et $x \in I$, on peut regarder l'évaluation en x,

$$\begin{array}{rcl} K^I & \longrightarrow & K \\ \varphi & \longmapsto & \varphi(x) \end{array}.$$

C'est une forme linéaire. Par restriction, c'est aussi le cas des applications d'évaluation $\mathcal{C}(I) \to \mathbb{R}$ ou $\mathcal{C}^1(I) \to \mathbb{R}$ si I est un intervalle (ouvert) de $\mathbb{R}$.

Alternativement, si on voit une application sur I comme une famille indexée par I, cela revient à envoyer une famille sur sa i-ème composante. Par exemple, on peut considérer la forme linéaire sur $K^{\mathbf{N}}$ qui envoie la suite $(x_n)_{n\in\mathbb{N}}$ sur x_0, ou sur un autre terme... Plus subtilement, on peut considérer la forme linéaire qui envoie une suite convergente sur sa limite. Ou même l'application qui envoie une suite de Cauchy dans $\mathbb{Q}$ sur sa limite dans $\mathbb{R}$, application qui est donc $\mathbb{Q}$-linéaire.

Deux exemples fondamentaux viennent de l'analyse, il s'agit de la dérivation

$$\begin{array}{rcl} d : \mathcal{C}^1(I) & \longrightarrow & \mathcal{C}(I) \\ f & \longmapsto & f' \end{array}$$

qui est une application linéaire et de l'intégration

$$\begin{array}{rcl} \mathcal{C}([a,b]) & \longrightarrow & \mathbb{R} \\ f & \longmapsto & \int_a^b f(x)dx \end{array}$$

qui est une forme linéaire. Si on fixe $x_0 \in I$, on peut aussi considérer la forme linéaire

$$\begin{array}{rcl} d : \mathcal{C}^1(I) & \longrightarrow & \mathbb{R} \\ f & \longmapsto & f'(x_0) \end{array}$$

ou l'application linéaire

$$\begin{array}{rcl} \mathcal{C}(I) & \longrightarrow & \mathcal{C}^1(I) \\ f & \longmapsto & (x \to \int_{x_0}^x f(t)dt) \end{array}.$$

Pour conclure, remarquons que si on voit $\mathbb{C}$ comme un espace vectoriel sur $\mathbb{R}$, alors la conjugaison complexe

$$\begin{array}{rcl} \mathbb{C} & \xrightarrow{\sigma} & \mathbb{C} \\ z & \longmapsto & \overline{z} \end{array}$$

est un automorphisme (c'est une symétrie). Mais cette application n'est plus linéaire si on voit $\mathbb{C}$ comme espace vectoriel sur lui même. Enfin, pour qu'une application linéaire entre espaces vectoriels sur $\mathbf{F}_2$ soit linéaire, il suffit qu'elle satisfasse la première condition.

Proposition 2.2. *Soit $f : E \to F$ une application linéaire. On a alors $f(0) = 0$ et, si $u, v \in E$, $f(u - v) = f(u) - f(v)$.*

Démonstration. On a
$$f(0) + f(0) = f(0 + 0) = f(0)$$
et donc $f(0) = 0$.

De même, si $u, v \in E$, on a
$$f(u - v) + f(v) = f(u - v + v) = f(u)$$
et donc $f(u - v) = f(u) - f(v)$. □

Proposition 2.3. *Une application $f : E \to F$ entre deux espaces vectoriels sur K est linéaire si et seulement si pour tous $u, v \in E, \lambda, \mu \in K$, on a*
$$f(\lambda u + \mu v) = \lambda f(u) + \mu f(v).$$

Démonstration. Supposons que $f : E \to F$ est linéaire et soient $u, v \in E, \lambda, \mu \in K$. On a alors
$$f(\lambda u + \mu v) = f(\lambda u) + f(\mu v) = \lambda f(u) + \mu f(v).$$
Inversement, si cette condition est satisfaite, on voit que f est linéaire en prenant d'une part $\lambda = \mu = 1$, et d'autre part $\mu = 0$. □

Proposition 2.4.

i) *Si $f : E \to F$ et $g : F \to G$ sont deux applications linéaires, alors gof est aussi linéaire.*

ii) *Si $f : E \to F$ est un isomorphisme entre deux espaces vectoriels, alors f^{-1} est aussi un isomorphisme.*

Démonstration. La première assertion se vérifie aisément. On se donne $u, v \in E, \lambda, \mu \in K$, et on vérifie que
$$\begin{aligned}(g \circ f)(\lambda u + \mu v) &= g(f(\lambda u + \mu v)) = g(\lambda f(u) + \mu f(v)) \\ &= \lambda g(f(u)) + \mu g(f(v))) = \lambda (g \circ f)(u) + \mu (g \circ f)(v).\end{aligned}$$
La seconde n'est pas beaucoup plus difficile. On sait déjà que f^{-1} est bijective (par définition). On se donne $u, v \in F, \lambda, \mu \in K$, et on veut montrer que
$$f^{-1}(\lambda u + \mu v) = \lambda f^{-1}(u) + \mu f^{-1}(v).$$
Comme f est bijective, il suffit de montrer que
$$\lambda u + \mu v = f(\lambda f^{-1}(u) + \mu f^{-1}(v))$$
et un rapide calcul du second membre nous donne bien
$$f(\lambda f^{-1}(u) + \mu f^{-1}(v)) = \lambda f(f^{-1}(u)) + \mu f(f^{-1}(v)) = \lambda u + \mu v.$$
□

Proposition 2.5. *Soit $f : E \to F$ une application linéaire. Alors,*

i) *Si G est un sous-espace vectoriel de E, alors*
$$f(G) := \{v \in F, \exists u \in G, f(u) = v\}$$
est un sous-espace vectoriel de F.

ii) *Si H est un sous-espace vectoriel de F, alors*
$$f^{-1}(H) := \{u \in E, f(u) \in H\}$$
est un sous-espace vectoriel de E.

Démonstration. On se donne tout d'abord un sous-espace G de E ainsi que $v, v' \in f(G)$ et $\lambda, \lambda' \in K$. Par définition, on peut trouver $u, u' \in G$ tels que $f(u) = v$ et $f(u') = v'$. Comme G est un sous-espace vectoriel de E, on a donc $\lambda u + \lambda' u' \in G$ et il suit que

$$\lambda v + \lambda' v' = \lambda f(u) + \lambda' f(u') = f(\lambda u + \lambda' u') \in f(G).$$

On remarque aussi que $0 = f(0) \in f(G)$ car $0 \in G$.

Seconde assertion maintenant : on suppose que H est un sous-espace vectoriel de F et on se donne $u, u' \in f^{-1}(H)$ ainsi que $\lambda, \lambda' \in K$. On a donc $f(u), f(u') \in H$ et il suit que

$$f(\lambda u + \lambda' u') = \lambda f(u) + \lambda' f(u') \in H$$

si bien que $\lambda u + \lambda' u' \in f^{-1}(H)$. Là encore, comme $f(0) = 0 \in H$, on a bien $0 \in f^{-1}(H)$. □

Définition 2.6. *Si $f : E \to F$ est une application linéaire, son* image *est*

$$\operatorname{Im} f := f(E) = \{v \in F, \exists u \in E, f(u) = v\}$$

et son noyau *est*

$$\ker f := f^{-1}(\{0\}) := \{u \in E, f(u) = 0\}.$$

Par exemple, le noyau de l'identité (ou même d'une homothétie) est réduit à $\{0\}$ et son image est l'espace tout entier. Par contre, c'est le noyau de l'application nulle $0_{E,F} : E \to F$ qui est E tout entier et son image qui est réduite à $\{0\}$.

Le noyau de l'application linéaire

$$(x_1, \ldots, x_m) \mapsto (a_{11}x_1 + \cdots + a_{1m}x_m, \ldots, a_{n1}x_1 + \cdots + a_{nm}x_m).$$

de K^m vers K^n est l'ensemble des solutions du système homogène

$$(S) : \left\{ \begin{array}{ccc} a_{11}x_1 + \cdots + a_{1m}x_m & = & 0 \\ \vdots & & \vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m & = & 0 \end{array} \right. .$$

Le noyau de la dérivation $d : \mathcal{C}^1(I) \to \mathcal{C}(I)$ est l'ensemble des applications constantes et son image est $\mathcal{C}(I)$ tout entier. Enfin, le noyau de l'application

$$\begin{array}{ccc} \mathcal{C}(I) & \longrightarrow & \mathcal{C}^1(I) \\ f & \longmapsto & (x \mapsto \int_{x_0}^{x} f(t)dt) \end{array} .$$

est nul et son image est formé des applications continûment dérivables qui s'annulent en x_0. Ce dernier espace n'est d'ailleurs rien d'autre que le noyau de l'évaluation en x_0.

Proposition 2.7. *Soit $f : E \to F$ une application linéaire. Alors,*

i) $\ker f$ *est un sous-espace vectoriel de E et* $\operatorname{im} f$ *est un sous-espace vectoriel de F.*

ii) *f est injective si et seulement si* $\ker f = 0$ *et f est surjective si et seulement si* $\operatorname{Im} f = F$.

Démonstration. La première assertion est un cas particulier de la proposition précédente. Par définition, f est surjective si et seulement si $\mathrm{Im}(f) = F$. Si $u, v \in E$ satisfont, $f(u) = f(v)$, alors

$$f(u - v) = f(u) - f(v) = 0$$

et il suit que $u - v \in \ker f$. On en déduit que si $\ker f = 0$, alors $u - v = 0$ et donc $u = v$. Cela montre que f est injective. Réciproquement, si $u \in \ker f$, alors $f(u) = 0 = f(0)$ et donc, si f est injective, on a bien $u = 0$. □

Définition 2.8. *Le noyau d'une forme linéaire non nulle est un* hyperplan.

Nous aurons besoin plus tard du résultat suivant :

Lemme 2.9. *Soient $f : E \to F$ et $g : F \to G$ deux applications linéaires. On a $\ker(f) \subset \ker(g \circ f)$ avec égalité si g est injective. On a $\mathrm{Im}(g \circ f) \subset \mathrm{Im}(g)$ avec égalité si f est surjective.*

Démonstration. Tout d'abord, si $u \in \ker f$, on a

$$(g \circ f)(u) = g(f(u)) = g(0) = 0$$

et donc $u \in \ker(g \circ f)$. Réciproquement, si $u \in \ker(g \circ f)$, on a $0 = (g \circ f)(u) = g(f(u))$ et donc $f(u) \in \ker g$. On voit donc que si g est injective, $f(u) = 0$ et donc $u \in \ker f$.

Montrons maintenant l'autre assertion. Si $w \in \mathrm{im}(g \circ f)$, alors il existe $u \in E$ tel que $(g \circ f)(u) = w$ et on a donc $g(f(u)) = w$ si bien que $w \in \mathrm{im}(g)$. Réciproquement, si $w \in \mathrm{im}(g)$ on peut trouver $v \in F$ tel que $g(v) = w$. Si f est surjective, on peut aussi trouver $u \in E$ tel que $f(u) = v$. On a donc $(g \circ f)(u) = g(f(u)) = g(v) = w$ et $w \in \mathrm{im}(g \circ f)$. □

Proposition 2.10. *Soit E un espace vectoriel sur K.*

i) *Si I est un ensemble, l'ensemble E^I des applications de I dans E est un espace vectoriel si on le munit de*

$$(f + g)(x) = f(x) + g(x) \text{ et } (\lambda f)(x) = \lambda f(x).$$

ii) *Si F est un autre espace vectoriel, l'ensemble $\mathcal{L}(E, F)$ est un sous-espace vectoriel de F^E.*

Démonstration. On vérifie aisément la première assertion. En effet, toutes les propriétés se vérifient terme à terme. Par exemple, le zéro est l'application nulle.

Nous allons être plus sérieux pour la seconde assertion. Tout d'abord, $\mathcal{L}(E, F)$ est non vide car l'application nulle est linéaire.

Il faut montrer ensuite que si f et g sont linéaires, alors $f + g$ aussi. On se donne donc $u, v \in E$ et $\lambda, \mu \in K$ et on calcule

$$(f + g)(\lambda u + \mu v) = f(\lambda u + \mu v) + g(\lambda u + \mu v)$$

$$= \lambda f(u) + \mu f(v) + \lambda g(u) + \mu g(v) = \lambda f(u) + \lambda g(u) + \mu f(v) + \mu g(v)$$

$$= \lambda(f(u) + g(u)) + \mu(f(v) + g(v)) = \lambda(f + g)(u)) + \mu(f + g)(v)).$$

Vérifions finalement que si f est linéaire, alors pour tout $\lambda \in K$, λf est linéaire. Il vaut mieux procéder en deux étapes pour éviter les confusions. On doit montrer que l'on a toujours

$$(\lambda f)(u+v) = (\lambda f)(u) + (\lambda f)(v)$$

ainsi que

$$(\lambda f)(\mu u) = \mu(\lambda f)(u).$$

Or par définition, on a

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

Et de même pour l'autre égalité :

$$(\lambda f)(\mu u) = \lambda f(\mu u) = \lambda\mu f(u) = \mu\lambda f(u) = \mu(\lambda f)(u).$$

□

Proposition 2.11. *Soit E un espace vectoriel sur K.*

i) *L'espace $L(E)$ est une algèbre pour l'addition, la multiplication externe et la composition : cela signifie que c'est un espace vectoriel et que, de plus, on a*
   (a) *Si $f, g \in L(E)$, alors $g \circ f \in L(E)$*
   (b) *Si $f, g, h \in L(E)$, alors $h \circ (g \circ f) = (h \circ g) \circ f$*
   (c) *Si $f \in L(E)$, alors $f \circ \mathrm{Id}_E = \mathrm{Id}_E \circ f = f$*
   (d) *Si $f, g, h \in L(E)$, alors $(f+g)\circ h = f\circ h + g\circ h$ et $f\circ(g+h) = f\circ g + f\circ h$*
   (e) *Si $f, g \in L(E)$ et $\lambda \in K$, alors $\lambda(f \circ g) = (\lambda f)\circ g = f \circ (\lambda g)$.*

ii) *L'ensemble GL(E) est un groupe pour la composition : cela signifie que*
   (a) *Si $f, g \in GL(E)$, alors $g \circ f \in GL(E)$*
   (b) *Si $f, g, h \in GL(E)$, alors $h \circ (g \circ f) = (h \circ g) \circ f$.*
   (c) *Si $f \in GL(E)$, alors $f \circ \mathrm{Id}_E = \mathrm{Id}_E \circ f = f$*
   (d) *Si $f \in GL(E)$, alors $f^{-1} \in GL(E)$.*

Démonstration. On s'attaque à la première assertion. On sait déjà que $L(E)$ est un espace vectoriel. On va passer les autres propriétés dans l'ordre. On a déjà vu la première qui dit que la composée de deux applications linéaires est linéaire. La seconde et la troisième propriété sont déjà vraies pour des applications quelconques. Les deux dernières nécessitent une vérification. On se donne $u \in E$ et on vérifie que

$$[(f+g)\circ h](u) = (f+g)(h(u)) = f(h(u)) + g(h(u))$$
$$= (f\circ h)(u) + (g \circ h)(u) = [f\circ h + g\circ h](u),$$

puis que

$$[f\circ(g+h)](u) = f((g+h)(u)) = f(g(u)+h(u)) = f(g(u)) + f(h(u))$$
$$= (f\circ g)(u) + (f\circ h)(u) = [f\circ g + f\circ h](u),$$

et enfin que

$$[\lambda(f\circ g)](u) = \lambda(f\circ g)(u) = \lambda f(g(u))$$
$$= (\lambda f)(g(u)) = [(\lambda f)\circ g](u)$$

d'une part mais aussi

$$[\lambda(f \circ g)](u) = f(\lambda g(u)) = f((\lambda g)(u)) = [f \circ (\lambda g)](u).$$

La seconde assertion n'apporte rien de nouveau puisque l'on sait que la composée de deux bijections est une bijection et que l'on a déjà vu que la réciproque d'une application linéaire bijective est aussi linéaire (et bijective).

□

Corollaire 2.12. i) *Si $n, m \in \mathbb{N}$, alors $M_{n\times m}(K)$ est un espace vectoriel sur K pour l'addition et la multiplication externe.*

ii) *Si $n \in \mathbb{N}$, alors $M_n(K)$ est une algèbre pour l'addition, la multiplication externe et la multiplication interne. Cela signifie que c'est un espace vectoriel et que, de plus, on a*
   (a) *Si $A, B, C \in M_n(K)$, alors $A(BC) = (AB)C$*
   (b) *Si $A \in M_n(K)$, alors $AI = IA = A$*
   (c) *Si $A, B, C \in M_n(K)$, alors $(A + B)C = AC + BC$ et $A(B + C) = AB + AC$*
   (d) *Si $A, B \in M_n(K)$ et $\lambda \in K$, alors $\lambda(AB) = (\lambda A)B = A(\lambda B)$.*

iii) *L'ensemble $GL_n(K)$ des matrices carrées inversibles d'ordre n est un groupe pour la multiplication. Cela signifie que*
   (a) *Si $A, B \in GL_n(K)$, alors $AB \in GL_n(K)$*
   (b) *Si $A, B, C \in GL_n(K)$, alors $A(BC) = (AB)C$*
   (c) *Si $A \in GL_n(K)$, alors $AI = IA = A$*
   (d) *Si $A \in GL_n(K)$, alors $AA^{-1} = I$ et $A^{-1}A = I$.*

Démonstration. Tout cela s'obtient par transport de structure via les bijections

$$L(K^m, K^n) \simeq M_{n\times m}(K), \quad L(K^n) \simeq M_n(K) \quad \text{et} \quad GL(K^n) \simeq GL_n(K)$$

qui sont compatibles avec l'addition, la multiplication externe et la multiplication interne (appelée composition pour les applications). □

Il suit que l'on peut calculer formellement dans $M_n(K)$ mais il faut se méfier car l'identité remarquable

$$(A + B)^2 = A^2 + 2AB + B^2$$

n'est pas toujours valide (contre-exemple ?). Mais par contre, on a toujours

$$(I + A)^k = \sum_{i=0}^{N} \binom{k}{i} A^i$$

si bien que pour calculer

$$\begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^7,$$

on écrit notre matrice sous la forme $I + A$ avec

$$A := \begin{bmatrix} 0 & 0 & 5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

et comme $A^2 = 0$, toutes ses puissances seront nulles, et on en déduit que notre matrice vaut

$$I + 7A := \begin{bmatrix} 1 & 0 & 35 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

3. Sommes directes

Proposition 3.1. *Si E_1 et E_2 sont deux espaces vectoriels, les lois sur $E_1 \times E_2$ définies par*

$$(u_1, u_2) + (v_1, v_2) := (u_1 + v_1, u_2 + v_2)$$

et

$$\lambda(u_1, u_2) := (\lambda u_1, \lambda u_2)$$

font de $E_1 \times E_2$ un espace vectoriel et les projections

$$\begin{aligned} E_1 \times E_2 &\longrightarrow E_2 \\ (u_1, u_2) &\longmapsto u_1 \end{aligned}$$

et

$$\begin{aligned} E_1 \times E_2 &\longrightarrow E_2 \\ (u_1, u_2) &\longmapsto u_2 \end{aligned}$$

sont des applications linéaires.

Démonstration. Toutes les propriétés d'un espace vectoriel se vérifient composante par composante et on voit en particulier que le zéro de $E \times F$ est $(0, 0)$ et que l'opposé d'un vecteur (u_1, u_2) sera le vecteur $(-u_1, -u_2)$.

De même, on voit facilement que les projections sont des applications linéaires. Pour la forme, on peut traiter le cas de la première que l'on appelera p_1. On se donne donc $(u_1, u_2), (v_1, v_2) \in E \times F$ ainsi que $\lambda, \mu \in K$. On calcule d'une part

$$\lambda p_1(u_1, u_2) + \mu p_1(v_1, v_2) = \lambda u_1 + \mu v_1$$

et d'autre part

$$p_1(\lambda(u_1, u_2) + \mu(v_1, v_2)) = p_1(\lambda u_1 + \mu v_1, \lambda u_2 + \mu v_2) = \lambda u_1 + \mu v_1.$$

□

Définition 3.2. *On dit alors que l'espace vectoriel $E_1 \times E_2$ est le produit des espaces vectoriels E_1 et E_2.*

On définit plus généralement le produit

$$E_1 \times E_2 \times \cdots \times E_n$$

d'un nombre quelconque d'espaces vectoriels (même infini d'ailleurs). Comme cas particulier, on retrouve

$$K^n := \underbrace{K \times \cdots \times K}_{n}.$$

Proposition 3.3. *Soient E_1 et E_2 deux sous-espaces vectoriels d'un espace vectoriel E. Alors,*

i) *L'ensemble*

$$E_1 + E_2 = \{u_1 + u_2, \quad u_1 \in E_1 \text{ et } u_2 \in E_2\}$$

est le plus petit sous-espace vectoriel de E contenant E_1 et E_2.

ii) *L'ensemble $E_1 \cap E_2$ est le plus grand sous-espace vectoriel de E contenu dans E_1 et E_2.*

iii) *L'application*

$$\begin{tikzcd}[row sep=tiny] \Phi : & E_1 \times E_2 \arrow[r] & E \\ & (u_1, u_2) \arrow[r, mapsto] & u_1 + u_2 \end{tikzcd}$$

est linéaire et son image est $E_1 + E_2$.

iv) *L'application*

$$\begin{tikzcd}[row sep=tiny] E_1 \cap E_2 \arrow[r] & E_1 \times E_2 \\ u \arrow[r, mapsto] & (u, -u) \end{tikzcd}$$

induit un isomorphisme entre $E_1 \cap E_2$ *et* $\ker \Phi$.

Démonstration. On vérifie d'abord la troisième assertion. La linéarité de l'application Φ se vérifie aisément : si $(u_1, u_2), (v_1, v_2) \in E_1 \times E_2$, et $\lambda, \mu \in K$, on a

$$\Phi(\lambda(u_1, u_2) + \mu(v_1, v_2)) = \Phi(\lambda u_1 + \mu v_1, \lambda u_2 + \mu v_2) = \lambda u_1 + \mu v_1 + \lambda u_2 + \mu v_2.$$

Et d'autre part,

$$\lambda\Phi(u_1, u_2) + \mu\Phi(v_1, v_2) = \lambda(u_1 + u_2) + \mu(v_1 + v_2) = \lambda u_1 + \lambda u_2 + \mu v_1 + \mu v_2.$$

Et l'image de Φ est par définition $E_1 + E_2$.

En particulier, on voit que que $E_1 + E_2$ est un sous-espace vectoriel de E. De plus, il contient bien E_1 et E_2. D'autre part, tout sous-espace vectoriel F contenant E_1 et E_2 contient aussi $E_1 + E_2$: si $u_1 \in E_1$ et $u_2 \in E_2$, alors u_1 et u_2 sont tous les deux dans F et comme F est un sous-espace vectoriel, on a $u_1 + u_2 \in F$. La première assertion est ainsi démontrée.

De même, pour la seconde assertion, on vérifie que $E_1 \cap E_2$ est un sous-espace vectoriel de E : Bien sûr, $0 \in E_1 \cap E_2$ et si $u, v \in E_1 \cap E_2$ et $\lambda, \mu \in K$, alors, $\lambda u + \mu v \in E_1$ car E_1 est un sous-espace vectoriel de E et, de même, $\lambda u + \mu v \in E_2$ car E_2 si bien que $\lambda u + \mu v \in E_1 \cap E_2$. Et cette intersection est, par définition, la plus grande partie de E contenue dans E_1 et E_2.

Finalement, l'application

$$\begin{tikzcd}[row sep=tiny] E_1 \cap E_2 \arrow[r, "\iota"] & E_1 \times E_2 \\ u \arrow[r, mapsto] & (u, -u) \end{tikzcd}$$

est linéaire : Si $u, v \in E_1 \cap E_2$ et $\lambda, \mu \in K$, alors, $\iota(\lambda u + \mu v) = (\lambda u + \mu v, -\lambda u - \mu v) = \lambda(u, -u) + \mu(v, -v)$. Elle est injective car son noyau est nul : si $(u, -u) = (0, 0)$, alors, en particulier, $u = 0$. De plus, son image est formée des couples (u, v) avec $u \in E_1, v \in E_2$ et $u + v = 0$, c'est à dire $\ker \Phi$. □

Par exemple, on peut considérer le plan E_1 donné par $x_1 = x_4 = 0$ et le plan E_2 donné par $x_1 + x_4 = x_2 + x_3 = 0$ dans K^4. Leur intersection est la droite donnée par $x_1 = x_2 + x_3 = x_4 = 0$ dirigée par le vecteur $(0, 1, -1, 0)$ et leur somme est l'hyperplan d'équation $x_1 + x_4 = 0$.

En général, on peut encore, on peut considérer des sommes

$$E_1 + E_2 + \cdots + E_n$$

ou des intersections

$$E_1 \cap E_2 \cap \cdots \cap E_n$$

de plus de deux sous-espaces (et même d'une infinité !). Nous n'en aurons probablement pas besoin.

Proposition 3.4. *Soient E_1 et E_2 deux sous-espace vectoriels d'un espace vectoriel E. Alors, les conditions suivantes sont équivalentes :*

i) *Tout élément de E s'écrit de manière unique sous la forme $u+v$ avec $u \in E_1$ et $v \in E_2$.*

ii) *On a* $\begin{cases} E = E_1 + E_2 \\ E_1 \cap E_2 = 0 \end{cases}$

iii) *L'application*

$$\Phi : \quad \begin{array}{rcl} E_1 \times E_2 & \longrightarrow & E \\ (u, v) & \longmapsto & u + v \end{array}$$

est un isomorphisme.

Démonstration. On fait une démonstration circulaire.

Il est clair que la première condition implique que

$$E = E_1 + E_2.$$

De plus, sous cette condition, si $u \in E_1 \cap E_2$, on a

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

et l'unicité implique que $u = 0$.

La seconde condition implique la troisième grâce à la proposition précédente. En effet, on sait que Φ est linéaire, que son noyau est isomorphe à $E_1 \cap E_2$ et donc réduit à 0, et que son image est $E_1 + E_2$, c'est à dire E. L'application est donc linéaire, injective et surjective, c'est à dire un isomorphisme.

Enfin, si la troisième condition est satisfaite, alors Φ est bijective et la première condition en découle immédiatement. □

Définition 3.5. *On dit alors que E_1 et E_2 sont* supplémentaires *ou que E est* somme directe *de E_1 et E_2 et on écrit $E = E_1 \oplus E_2$.*

Par exemple, si D et Δ sont deux droites (vectorielles) *distinctes* de K^2, on a toujours $K^2 = D \oplus \Delta$. De même, si H est un plan (vectoriel) de K^3 et D une droite (vectorielle) *non contenue dans H*, alors $K^3 = H \oplus D$. Enfin, si $\mathcal{F}$ est l'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, $\mathcal{P}$ le sous-espace des fonctions paires et $\mathcal{I}$ celui des fonctions impaires, on a $\mathcal{F} = \mathcal{P} \oplus \mathcal{I}$. On peut aussi considérer les matrices symétriques et antisymétriques.

Proposition 3.6. *Soient E_1 et E_2 deux sous-espaces vectoriels supplémentaires d'un espace vectoriel E. Alors, les* projections

$$E \simeq E_1 \times E_2 \to E_1 \quad \text{et} \quad E \simeq E_1 \times E_2 \to E_2$$

sont des applications linéaires surjectives de noyaux respectifs E_2 et E_1.

Démonstration. Les applications sont bien linéaires comme composées d'applications linéaires.

De plus, on peut décrire la première projection p_1 comme suit : si $u \in E$, on peut écrire de manière unique $u = u_1 + u_2$ avec $u_1 \in E_1$ et $u_2 \in E_2$ et alors, $p_1(u) = u_1$.

En particulier, on voit que si $p_1(u) = 0$, alors $u_1 = 0$ et donc $u = u_2 \in E_2$. Réciproquement, si $u \in E_2$, on écrit $u = 0 + u$ avec $0 \in E_1$ et $u \in E_2$ et on a donc bien $p_1(u) = 0$ si bien que $u \in \ker p_1$. On montre aussi que p_1 est surjective car $p_1(u) = u$ si $u \in E_1$.

Les assertions relatives à la seconde projection se montrent de la même manière. $\square$

Dans l'example ci-dessus $K^2 = D \oplus \Delta$, les projections sont respectivement les projections sur une des droites parallèlement à l'autre. Et on a une description analogue pour l'exemple $K^3 = H \oplus D$. Enfin, pour $\mathcal{F} = \mathcal{P} \oplus \mathcal{I}$, les projections sont les applications $\mathcal{F} \to \mathcal{P}$ et $\mathcal{F} \to \mathcal{I}$ qui associent à une fonction ses parties paire et impaire. De même pour les matrices symétriques et antisymétriques.

4. Systèmes générateurs et libres

Définition 4.1. *Soit E un espace vectoriel sur K. Si $u_1, \dots, u_n \in E$ et $\lambda_1, \dots, \lambda_n \in K$, on dit que*

$$u = \lambda_1 u_1 + \cdots + \lambda_n u_n$$

est une combinaison linéaire *de $u_1, \dots, u_n$ et que*

$$\lambda_1, \dots, \lambda_n$$

sont les coefficients. *On dit que la combinaison linéaire est* triviale *si tous les coefficients sont nuls.*

Par convention, la combinaison linéaire vide vaut 0. Une combinaison linéaire d'un vecteur isolé est tout simplement un multiple de ce vecteur. Enfin, si $u, v \in E$, une combinaison linéaire de u et v est un vecteur de la forme $\lambda u + \mu v$ avec $\lambda, \mu \in K$.

Comme exemple, on voit que le vecteur $(2, 3, 2) \in K^3$ est combinaison linéaire de $(1, 0, 1)$ et $(0, 1, 0)$. Contrairement au vecteur $(1, 2, 3)$ par exemple.

Les définitions de sous-espaces vectoriels et d'applications linéaires se généralisent à des combinaisons linéaires générales.

Proposition 4.2.

i) *Une partie F d'un espace vectoriel E est un sous-espace vectoriel si et seulement si elle est stable par combinaison linéaire : Pour tout $u_1, \dots, u_n \in F$ et tout $\lambda_1, \dots, \lambda_n \in K$, on a*

$$\lambda_1 u_1 + \cdots + \lambda_n u_n \in F.$$

ii) *Une application $\varphi : E \to F$ est linéaire si et seulement si elle préserve les combinaisons linéaires : Pout tout $u_1, \dots, u_n \in E$ et tout $\lambda_1, \dots, \lambda_n \in K$, on a*

$$f(\lambda_1 u_1 + \cdots + \lambda_n u_n) = \lambda_1 f(u_1) + \cdots + \lambda_n f(u_n).$$

Démonstration. Les conditions sont clairement suffisantes car il suffit de considérer le cas $n = 2$ (et aussi le cas $n = 0$ pour s'assurer que $0 \in F$ dans la première assertion).

Et on démontre aisément par récurrence sur n que celles-ci sont nécessaires. On sait que $0 \in F$ dans le premier cas et que $f(0) = 0$ dans le second. On suppose la condition satisfaite à l'ordre n. Dans le premier cas, on aura bien

$$\lambda_1 u_1 + \cdots + \lambda_{n+1} u_{n+1} = (\lambda_1 u_1 + \cdots + \lambda_n u_n) + \lambda_{n+1} u_{n+1} \in F.$$

Et dans le second cas

$$f(\lambda_1 u_1 + \cdots + \lambda_{n+1} u_{n+1}) = f((\lambda_1 u_1 + \cdots + \lambda_n u_n) + \lambda_{n+1} u_{n+1})$$

$$= f(\lambda_1 u_1 + \cdots + \lambda_n u_n) + \lambda_{n+1} f(u_{n+1}) = \lambda_1 f(u_1) + \cdots + \lambda_n f(u_n) + \lambda_{n+1} f(u_{n+1}).$$

□

Notation : Soient $\mathcal{S}$ est une partie d'un espace vectoriel E, on note $\mathrm{Vect}(\mathcal{S})$ l'ensemble de toutes les combinaisons linéaires d'éléments de E.

Proposition 4.3.

i) *Si $\mathcal{S}$ et $\mathcal{T}$ deux parties d'un espace vectoriel E. Si $\mathcal{S} \subset \mathcal{T}$, alors* $\mathrm{Vect}(\mathcal{S}) \subset \mathrm{Vect}(\mathcal{T})$.

ii) *Soit $f : E \to F$ une application linéaire et $\mathcal{S}$ une partie de E, alors $f(\mathrm{Vect}(\mathcal{S})) = \mathrm{Vect}(f(\mathcal{S}))$.*

Démonstration. Exercice □

Proposition 4.4. *Soit E un espace vectoriel.*

i) *Toute intersection de sous-espaces vectoriels de E est un sous-espace vectoriel de E.*

ii) *Si $\mathcal{S}$ est une partie de E, il existe un plus petit sous-espace vectoriel de E contenant $\mathcal{S}$ (c'est l'intersection des sous-espace vectoriels contenant $\mathcal{S}$) : c'est $\mathrm{Vect}(\mathcal{S})$.*

Démonstration. La première assertion se vérifie aisément : si $(E_i)_{i\in I}$ est une famille de sous-espaces vectoriels de E, on a bien sûr, $0 \in \cap_{i\in I}E_i$ et si $u, v \in \cap_{i\in I}E_i$ et $\lambda, \mu \in K$, on a pour tout $i \in I$, $\lambda u + \mu v \in E_i$ si bien que $\lambda u + \mu v \in \cap_{i\in I}E_i$.

L'existence et la première caractérisation d'un plus petit sous-espace vectoriel contenant une partie $\mathcal{S}$ donnée en résultent formellement : en effet, il suffit de prendre l'intersection de tous les sous-espaces vectoriels qui contiennent $\mathcal{S}$.

Il reste à montrer la dernière assertion. On montre tout d'abord que l'ensemble F des combinaisons linéaires d'éléments de $\mathcal{S}$ est un sous-espace vectoriel : c'est assez pénible à écrire mais il est clair que la somme de deux combinaisons linéaires est encore une combinaison linéaire et qu'il en va de même pour le produit par un scalaire.

Enfin, le sous-espace vectoriel F contient $\mathcal{S}$ et il résulte de la proposition précédente que c'est le plus petit. □

Définition 4.5. *On dit alors que F est le* sous espace engendré *par $\mathcal{S}$ ou que $\mathcal{S}$ est un* système générateur *de F. Si $F = E$, on dit tout simplement que $\mathcal{G}$ est un système générateur.*

On voit donc qu'une partie $\mathcal{G}$ d'un espace vectoriel E est un système générateur si et seulement si $E = \mathrm{Vect}(\mathcal{G})$.

Dans K, toute partie contenant au moins un vecteur non nul est un système générateur. En d'autres termes, seuls $\emptyset$ et $\{0\}$ ne sont pas générateurs.

Dans K^2, un système générateur contient au moins deux éléments non nuls. Mais ce n'est pas suffisant, par exemple $\{(2, -4), (-3, 6)\}$ n'est pas générateur dans $\mathbb{R}^2$. Par contre, $\{(1, 0), (0, 1)\}$ est bien un système générateur de $\mathbb{R}^2$ et plus généralement, la base canonique de K^n est un système générateur.

Les projections sur les axes forment un système générateur de $\check{K}^n$.

Un système générateur $\mathcal{G}$ de $K[X]$ est toujours une partie infinie. Il suffit qu'il y ait au moins un polynôme de chaque degré dans $\mathcal{G}$ pour que le système soit générateur.

Pour finir, notons que dans l'espace vectoriel $\{0\}$, tout système est générateur (il n'y en a que deux : $\emptyset$ et $\{0\}$) et que tout espace vectoriel E possède au moins un système générateur : E lui même.

Proposition 4.6. *Soit $\mathcal{G}$ un système générateur d'un espace vectoriel E et $\mathcal{S}$ une partie de E contenant $\mathcal{G}$. Alors, $\mathcal{S}$ est aussi un système générateur de E.*

Démonstration. On a $E = \text{Vect}(\mathcal{G}) \subset \text{Vect}(\mathcal{S}) \subset E$ et donc $\text{Vect}(\mathcal{S}) \subset E$. $\square$

Proposition 4.7. *Soit $\mathcal{G}$ un système générateur d'un espace vectoriel E et $f : E \to F$ linéaire. Alors $f(\mathcal{G})$ est un système générateur de* $\text{Im} f$. *En particulier, f est surjective si et seulement si $f(\mathcal{G})$ est un système générateur de F.*

Démonstration. On a $\text{Im} f = f(E) = f(\text{Vect}(\mathcal{G})) = \text{Vect}(f(\mathcal{G}))$. $\square$

Définition 4.8. *Soit F un sous-espace vectoriel d'un espace vectoriel E. Une* paramétrisation linéaire *de F est une application linéaire $\phi : K^n \to E$ telle que* $F := \text{im}\phi$.

Par exemple, si H est l'hyperplan d'équation $x+y+z=0$ dans $\mathbb{R}^3$, on peut prendre la paramétrisation

$$\phi : \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^3 \\ (\lambda, \mu) & \longmapsto & (\lambda + \mu, -\lambda, -\mu). \end{array}$$

Notons qu'on peut définir plus généralement, une paramétrisation pas nécessairement linéaire d'une partie F d'un ensemble E comme une application de K^n vers E dont l'image est F : par exemple

$$\begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R}^2, \\ t & \longmapsto & (\cos t, \sin t) \end{array}$$

est une paramétrisation du cercle S^1.

Proposition 4.9. *Si E est un espace vectoriel sur K et $u_1, \dots, u_n \in E$, alors l'application*

$$\phi : \begin{array}{ccc} K^n & \longrightarrow & E \\ (\lambda_1, \dots, \lambda_n) & \longmapsto & \lambda_1 u_1 + \cdots + \lambda_n u_n \end{array}$$

est une paramétrisation (linéaire) du sous-espace F engendré par $u_1, \dots, u_n$.

Démonstration. Par définition, l'image de ϕ est exactement le sous-espace engendré par $u_1, \dots, u_n$. Il faut juste s'assurer que ϕ est bien linéaire. En d'autres termes, il faut vérifier que si on se donne $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n, \lambda, \mu \in K$, on a bien

$$\lambda(\lambda_1 u_1 + \cdots + \lambda_n u_n) + \mu(\mu_1 u_1 + \cdots \mu_n u_n)$$

$$= (\lambda\lambda_1 + \mu\mu_1)u_1 + \cdots + (\lambda\lambda_n + \mu\mu_n)u_n.$$

$\square$

Définition 4.10. *Soit E un espace vectoriel sur K. On dit que $u_1, \dots, u_n \in E$ sont* linéairement dépendants *si il existe une combinaison linéaire non triviale qui s'annule :*

$$\lambda_1 u_1 + \cdots + \lambda_n u_n = 0.$$

(c'est une relation de dépendance linéaire*). Sinon, on dit que $u_1, \dots, u_n$ sont* linéairement indépendants.

Par exemple, dans K, deux vecteurs quelconques sont toujours dépendants. En général, un vecteur tout seul est indépendant si et seulement il est non nul. De toutes manières, si l'un des vecteurs $u_1, \ldots, u_n$ est nul, ils sont obligatoirement dépendants.

Dans K^2, deux vecteurs u et v sont dépendants si et seulement si l'un est multiple de l'autre, par exemple $(2, -4)$ et $(-3, 6)$ sont dépendants dans $\mathbb{R}^2$. Par contre, $(1, 0)$ et $(0, 1)$ sont linéairement indépendants et plus généralement, les vecteurs de la base canonique de K^n sont linéairement indépendants.

Notons aussi que les projections sur les axes sont des formes linéaires linéairement indépendantes. Enfin, des polynômes non nuls de degrés distincts sont linéairement indépendants.

Proposition 4.11. *Des vecteurs $u_1, \ldots, u_n$ d'un espace vectoriel E sont linéairement dépendants si et seulement si l'un d'entre eux est combinaison linéaire des autres.*

Démonstration. En effet, s'ils sont linéairement dépendants, on peut écrire

$$\lambda_1 u_1 + \cdots + \lambda_n u_n = 0$$

avec au moins un des coefficients non nuls. Quitte à les réordonner, on peut supposer que $\lambda_n \neq 0$ et on aura alors

$$u_n = -\frac{\lambda_1}{\lambda_n} u_1 - \cdots - \frac{\lambda_{n-1}}{\lambda_n} u_{n-1} = 0$$

Réciproquement, si on suppose que l'un des vecteurs est combinaison linéaire des autres, quitte à les renommer, on peut supposer que c'est u_n et que l'on peut donc écrire $u_n = \lambda_1 u_1 + \cdots \lambda_{n-1} u_{n-1}$. Il suit que

$$\lambda_1 u_1 + \cdots + \lambda_{n-1} u_n + (-1) u_n = 0$$

avec, bien sûr, $-1 \neq 0$.

□

On a besoin de généraliser cette notion à des familles éventuellement infinies.

Définition 4.12. *Une famille $(u_i)_{i \in I}$ d'éléments d'un espace vectoriel E est un* système lié *si il existe $i_1, \ldots, i_n$ distincts dans I tels que les vecteurs $u_{i_1}, \ldots, u_{i_n}$ soient linéairement dépendants. Sinon, on dit que la famille $(u_i)_{i \in I}$ est* libre.

Par exemple, la famille $(1, X, X^2, X^3, \ldots)$ de $K[X]$ est un système libre : en effet, supposons qu'une combinaison linéaire $\lambda_1 X^{i_1} + \cdots \lambda_n X^{i_n}$ avec des exposants distincts s'annule. Et soit i_k le plus grand exposant. Alors, le polynôme $P := \lambda_1 X^{i_1} + \cdots + \lambda_n X^{i_n}$ est le polynôme nul mais son degré est i_k. Contradiction.

En fait, il suffit qu'il y ait au plus un polynôme de chaque degré dans $\mathcal{L}$ pour que ce soit un système libre dans $K[X]$.

De même, les fonctions $x \mapsto x^\alpha$ pour $\alpha \in \mathbb{R}$ forment une famille libre de $\mathcal{C}(\mathbb{R}_{>0})$. On le montre ainsi : supposons que la fonction $\lambda_1 x^{\alpha_1} + \cdots \lambda_n x^{\alpha_n}$ soit identiquement nulle. On dérive et on multiplie par x pour obtenir que $\lambda_1 \alpha_1 x^{\alpha_1} + \cdots \lambda_n \alpha_n x^{\alpha_n}$ est aussi identiquement nulle. On multiplie la première par α_n et on retranche la seconde pour trouver que $\lambda_1(\alpha_n - \alpha_1)x^{\alpha_1} + \cdots + \lambda_{n-1}(\alpha_n - \alpha_{n-1})x^{\alpha_{n-1}}$ est identiquement nulle. On peut ainsi procéder par récurrence sur n.

La plupart des autres exemples importants résultent en fait de la prochaine proposition.

Avant de poursuivre, je souhaite cependant faire la remarque suivante : un système générateur est un ensemble mais un système libre (ou lié) est une famille. Dans un ensemble, les éléments ne sont pas ordonnés mais on ne peut pas les répéter. Dans une famille, c'est le contraire. Étant donné une famille $(u_i)_{i\in I}$, on peut considérer son support $\{u_i, i \in I\}$ qui est un ensemble. De même, tout ensemble S définit une famille $(u)_{u\in S}$ indexée par lui même. Par exemple, le support de la famille (infinie) $(u_i = (-1)^i)_{i\in\mathbb{N}}$ est l'ensemble

$$S = \{-1, 1\} = \{x \in \mathbb{R}, x^2 = 1\}$$

auquel on associe la famille (finie) $(u_1 := 1, u_{-1} := -1)$ indexée par l'ensemble $\{-1, 1\}$. En pratique, on mélangera allègrement les deux notions mais il faudra faire attention...

Proposition 4.13. *Des vecteurs $u_1, \ldots, u_n$ d'un espace vectoriel E sont dépendants (resp. indépendants) si et seulement si le système $(u_1, \ldots, u_n)$ est lié (resp. libre).*

Démonstration. En effet, dire que le système est lié signifie qu'on peut trouver une combinaison linéaire non triviale d'une *partie* des vecteurs qui s'annule. Il suffit de compléter avec des coefficients nuls pour obtenir une combinaison linéaire non triviale de *tous* les vecteurs qui s'annule. □

Proposition 4.14. *Soit $\mathcal{L}$ un système libre d'un espace vectoriel E et $\mathcal{S}$ une partie de E contenue dans $\mathcal{L}$. Alors, $\mathcal{S}$ est aussi un système libre de E.*

Démonstration. Clair : toute combinaison linéaire d'éléments de $\mathcal{S}$ est en particulier combinaison linéaire d'éléments de $\mathcal{L}$. □

Proposition 4.15. *Soit $\mathcal{L}$ un système libre d'un espace vectoriel E et $f : E \to F$ une application linéaire et injective. Alors, $f(\mathcal{L})$ est un système libre de F.*

Démonstration. On écrit $\mathcal{L} =: (u_i)_{i\in I}$ et on remarque que si $f(\mathcal{L})$ est lié, on peut trouver une combinaison linéaire non-triviale nulle

$$\lambda_{i_1} f(u_{i_1}) + \cdots + \lambda_{i_n} f(u_{i_n}) = 0$$

Comme f est linéaire, on a

$$f(\lambda_{i_1} u_{i_1} + \cdots + \lambda_{i_n} u_{i_n}) = 0.$$

Et comme f est injective, on a nécessairement

$$\lambda_{i_1} u_{i_1} + \cdots + \lambda_{i_n} u_{i_n} = 0.$$

□

Remarquons que si $\mathcal{S}$ est un système lié de E et $f : E \to F$ une application linéaire, alors $f(\mathcal{S})$ est aussi lié.

Proposition 4.16. *Des vecteurs $u_1, \ldots, u_n$ d'un espace vectoriel E sont linéairement indépendants si et seulement si l'application*

$$\begin{aligned} K^n &\xrightarrow{\phi} E \\ (\lambda_1, \ldots, \lambda_n) &\longmapsto \lambda_1 u_1 + \cdots + \lambda_n u_n \end{aligned}$$

est injective.

Démonstration. Dire que ϕ est injective signifie que son noyau est nul, c'est à dire que lorsque

$$\lambda_1 u_1 + \cdots + \lambda_n u_n = 0,$$

on a nécessairement

$$\lambda_1 = \cdots = \lambda_n = 0.$$

□

5. Bases

Définition 5.1. *Une* base *d'un espace vectoriel est un système libre et générateur.*

Il faut être prudent car, comme on l'a déjà fait remarquer, un système libre est une famille et un système générateur est un ensemble. En général, on voit plutôt une base comme une famille libre dont le support est un ensemble générateur.

La seule base de l'espace nul est l'ensemble vide. Tout vecteur non nul de K est une base de K. Les vecteurs $(1,0)$ et $(0,1)$, forment une base de $\mathbb{R}^2$ mais il en va de même de $(2,3)$ et $(3,2)$.

En général, la base canonique $(e_1,\ldots,e_n)$ est une base de K^n (on rappelle que e_i est le vecteur dont toutes les composantes sont nulles sauf la i-ème qui vaut 1).

Les projections sur les axes forment une base de $\check{K}^n$. Le système $(1,X,X^2,X^3,\ldots)$ est une base de $K[X]$. On peut aussi considérer la base $(e_{i,j})$ de $M_{n\times m}(K)$ ou $e_{i,j}$ est la matrice qui vaut zéro partout sauf à l'intersection de la i-ème ligne et de la j-ème colonne ou elle vaut 1. Dans ces trois cas, on parle aussi de *base canonique*.

Remarquons qu'une base $\mathcal{B}$ de $K[X]$ est toujours infinie. D'autre part, il suffit qu'il y ait exactement un polynôme de chaque degré dans $\mathcal{B}$ pour que ce soit une base.

Proposition 5.2. *Le système $\mathcal{S}$ est une base de E si et seulement si tout élément de E s'écrit de manière unique comme combinaison linéaire d'éléments de $\mathcal{S}$.*

Démonstration. Si $\mathcal{S}$ est une base, c'est un système générateur et donc tout élément de E s'écrit comme combinaison linéaire d'éléments de $\mathcal{S}$. De plus, si on peut écrire

$$u = \lambda_1 u_1 + \cdots + \lambda_n u_n$$

et

$$u = \mu_1 u_1 + \cdots + \mu_n u_n$$

avec $u_1,\ldots,u_n \in \mathcal{S}$ distincts, alors

$$(\lambda_1 - \mu_1)u_1 + \cdots + (\lambda_n - \mu_n)u_n = 0.$$

Le fait que $\mathcal{S}$ est libre entraîne que

$$\lambda_1 - \mu_1 = \cdots = \lambda_n - \mu_n = 0$$

et on a donc

$$\lambda_1 = \mu_1, \ldots, \lambda_n = \mu_n.$$

Réciproquement, notre condition implique trivialement que $\mathcal{S}$ est générateur et il faut s'assurer qu'il est libre. Or, si $u_1,\ldots,u_n \in \mathcal{S}$ sont tels que

$$\lambda_1 u_1 + \cdots + \lambda_n u_n = 0$$

alors, comme on peut aussi écrire

$$0u_1 + \cdots + 0u_n = 0,$$

on a nécessairement

$$\lambda_1 = \cdots = \lambda_n = 0.$$

$\square$

Définition 5.3. *Si* $(u_1, \ldots, u_n)$ *est une base de* E *et*

$$u = \lambda_1 u_1 + \cdots + \lambda_n u_n,$$

on dit que $\lambda_1, \ldots, \lambda_n$ *sont les* composantes *de* u.

Par exemple, les composantes du vecteur $(1, 4)$ dans la base canonique $((1, 0), (0, 1))$ de $\mathbb{R}^2$ sont 1 et 4. Mais dans la base $((2, 3), (3, 2))$, ce sont 2 et -1. En général, les composantes du vecteur $(x_1, \ldots, x_n) \in K^n$ dans la base canonique sont $x_1, \ldots, x_n$. Et les composantes de la forme linéaire

$$(x_1, \ldots, x_n) \mapsto a_1 x_1 + \cdots a_n x_n$$

dans la base canonique de $\check{K}^n$ sont $a_1, \ldots, a_n$. Attention à l'ordre : les composantes du polynôme $a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0$ dans la base canonique $(1, X, X^2, \ldots)$ sont $a_0, a_1, a_2, \ldots$.

Proposition 5.4. *Soit* $f : E \to F$ *une application linéaire et* $\mathcal{B}$ *une base de* E. *Alors,* f *est un isomorphisme (resp. surjective, resp. injective) si et seulement si* $f(\mathcal{B})$ *est une base (resp. un système générateur, resp. un système libre) de* F.

Démonstration. C'est une conséquence des résultats analogues sur les systèmes générateurs et libres à part le fait que si $f(\mathcal{B})$ est libre, alors f est injective. Mais ce n'est pas difficile : en effet, supposons que $u \in E$ satisfasse $f(u) = 0$ et écrivons $u = \lambda_1 u_1 + \cdots + \lambda_n u_n$ comme combinaison linéaire d'éléments (distincts) de $\mathcal{B}$. On aura alors

$$\lambda_1 f(u_1) + \cdots + \lambda_n f(u_n) = f(\lambda_1 u_1 + \cdots + \lambda_n u_n) = f(u) = 0$$

et comme $f(u_1), \ldots, f(u_n)$ sont linéairement indépendants, on aura bien $\lambda_1 = \cdots = \lambda_n = 0$ et donc $u = 0$. $\square$

Il est important dans cette dernière démonstration de regarder l'image de la famille $\mathcal{B}$ et non pas l'image de l'ensemble. Par exemple, avec l'application

$$\begin{array}{ccc} \mathbb{R}^2 & \xrightarrow{f} & \mathbb{R}^2 \\ (x, y) & \longmapsto & (x + y, 0) \end{array},$$

l'image de la base canonique est réduite à $\{e_1\}$ qui est libre mais f n'est pas injective !

Proposition 5.5. *Les vecteurs* $u_1, \ldots, u_n$ *forment une base de* E *si et seulement si l'application*

$$\begin{array}{ccc} K^n & \xrightarrow{\Phi} & E \\ (\lambda_1, \ldots, \lambda_n) & \longmapsto & \lambda_1 u_1 + \cdots + \lambda_n u_n \end{array}$$

est un isomorphisme.

Démonstration. Résulte des résultats analogues sur les systèmes générateurs et libres. $\square$

Proposition 5.6. *Soient* $(u_1, \ldots, u_n)$ *une base de* E *et* $v_1, \ldots, v_n \in F$. *Alors, il existe une unique application linéaire* $f : E \to F$ *telle que*

$$f(u_1) = v_1, \ldots, f(u_n) = v_n.$$

De plus, $v_1, \ldots, v_n$ sont linéairement indépendants (resp. générateurs, resp. forment une base de F) si et seulement si f est injective, (resp. surjective, resp. bijective).

Démonstration. On se donne donc une base $(u_1, \ldots, u_n)$ de E et
$$v_1, \ldots, v_n \in F.$$
Soit $f : E \to F$ une application linéaire telle que
$$f(u_1) = v_1, \ldots, f(u_n) = v_n.$$
Tout $u \in E$ s'écrit de manière unique
$$u =: \lambda_1 u_1 + \cdots + \lambda_n u_n$$
et on a donc
$$f(u) = \lambda_1 v_1 + \cdots + \lambda_n v_n.$$
D'où l'unicité. Réciproquement, on peut toujours définir f comme ceci et vérifier qu'elle est bien linéaire : En effet, si on a aussi
$$u' =: \lambda'_1 u_1 + \cdots + \lambda'_n u_n$$
et $\mu, \mu' \in K$, on doit vérifier que
$$f(\mu u + \mu' u') = \mu f(u) + \mu' f(u').$$
Or, on a
$$\mu u + \mu' u' =: (\mu\lambda_1 + \mu'\lambda'_1)u_1 + \cdots + (\mu\lambda_n + \mu'\lambda'_n)u_n$$
et donc
$$f(\mu u + \mu' u') =: (\mu\lambda_1 + \mu'\lambda'_1)v_1 + \cdots + (\mu\lambda_n + \mu'\lambda'_n)v_n.$$
Et c'est bien la même chose que
$$\mu f(u) + \mu' f(u') = \mu(\lambda_1 v_1 + \cdots + \lambda_n v_n) + \mu'(\lambda'_1 v_1 + \cdots + \lambda'_n v_n).$$
Les autres assertions sont simplement un rappel.

□

Théorème 5.7. *(Théorème de la base incomplète) Si $\mathcal{L}$ est un système libre contenu dans un système générateur $\mathcal{G}$, il existe une base $\mathcal{B}$ de E contenue dans $\mathcal{G}$ et contenant $\mathcal{L}$.*

Démonstration. On considère l'ensemble des systèmes libres contenus dans $\mathcal{G}$ qui contiennent $\mathcal{L}$. Et on lui applique le *lemme de Zorn* : comme, trivialement, toute union croissante de systèmes libre est libre, il existe un système libre maximal $\mathcal{B}$ contenu dans $\mathcal{G}$ et contenant $\mathcal{L}$. Montrons que $\mathcal{B}$ est une base. Comme $\mathcal{G}$ est générateur, il suffit de montrer que tout $u \in \mathcal{G}$ est combinaison linéaire d'éléments de $\mathcal{B}$. La maximalité de $\mathcal{B}$ implique que, soit $u \in \mathcal{B}$ auquel cas on a gagné, soit $\mathcal{B} \cup \{u\}$ est lié. Dans ce dernier cas, on peut trouver une combinaison linéaire non triviale nulle
$$\lambda u + \lambda_1 u_1 + \cdots + \lambda_n u_n = 0$$
avec $u_1, \ldots, u_n \in \mathcal{B}$ distincts. Comme $\mathcal{B}$ est libre, on ne peut pas avoir $\lambda = 0$. On voit donc que
$$u = -\frac{\lambda_1}{\lambda} u_1 - \cdots - \frac{\lambda_n}{\lambda} u_n.$$

□

Corollaire 5.8. i) *Tout espace vectoriel possède une base.*

ii) *Tout système générateur contient une base.*

iii) *Tout système libre est contenu dans une base.*

Démonstration. C'est immédiat car $\emptyset$ est libre et E est générateur. $\square$

Considérons par exemple les vecteurs $u := (1, -1, 0)$ et $v := (1, 0, -1)$ de $\mathbb{R}^3$. Ils sont indépendants et on peut donc considérer le système libre $\mathcal{L} := (u, v)$. D'autre part, on sait que (e_1, e_2, e_3) est une base de $\mathbb{R}^3$ et est donc générateur. Il suit que le système $\mathcal{G} := (u, v, e_1, e_2, e_3)$ est générateur et on a $\mathcal{L} \subset \mathcal{G}$. Le théorème nous dit donc qu'il existe une base $\mathcal{B}$ de $\mathbb{R}^3$ avec $\mathcal{L} \subset \mathcal{B} \subset \mathcal{G}$. Ça fait donc un nombre fini de possibilités que l'on peut tester. En fait, on peut prendre $\mathcal{B} := (u, v, e_1)$.

Proposition 5.9. *Tout sous-espace vectoriel de E possède un supplémentaire dans E.*

Démonstration. On se donne donc un sous-espace vectoriel F de E, on choisit une base $\mathcal{C}$ de F et on la complète en une base $\mathcal{B}$ de E. On note $\mathcal{D} := \mathcal{B} \setminus \mathcal{C}$ et G le sous-espace engendré par $\mathcal{D}$. Il faut montrer que $E = F \oplus G$. Or tout $u \in E$ s'écrit de manière unique comme combinaison linéaire d'éléments de $\mathcal{B}$, c'est à dire comme somme d'une combinaison linéaire d'éléments de $\mathcal{C}$ et d'une combinaison linéaire d'éléments de $\mathcal{D}$. Donc tout élément de E s'écrit bien de manière unique comme somme d'un élément de F et d'un élément de G. $\square$

6. Dimension

Lemme 6.1 (Lemme de Steinitz)**.** *Soit $n \geq 1$. Soit $(e_1, \cdots, e_n)$ une famille de n éléments de E. Soit $(u_1, \cdots, u_{n+1})$une famille d'éléments de* $\mathrm{Vect}(e_1, \cdots, e_n)$*. Alors la famille $(u_1, \cdots, u_{n+1})$ est liée.*

Démonstration. La preuve se fait par récurrence sur n. Pour $n = 1$, la vérification est facile. Soit $n \geq 1$ supposons que la propriété est vraie au rang n. Soit alors une famille $(u_j)_{j=1,\cdots,n+2}$ contenue dans $\mathsf{Vect}(e_1, \cdots, e_{n+1})$. Nous voulons montrer qu'elle est liée. On peut écrire :

$$\begin{aligned} u_1 &= a_{1,1}e_1 + \cdots + a_{1,n+1}e_{n+1}, \\ u_2 &= a_{2,1}e_1 + \cdots + a_{2,n+1}e_{n+1}, \\ &\cdots \\ u_{n+2} &= a_{n+2,1}e_1 + \cdots + a_{n+2,n+1}e_{n+1}. \end{aligned}$$

Deux cas se présentent. Si, pour tout j, on a : $a_{j,1} = 0$, on constate que la famille (u_j) est contenue dans l'espace engendré par $e_2, \cdots, e_{n+1}$ qui est de dimension n. On en déduit aisément par récurrence que (u_j) est liée. Sinon, il existe j_0 tel que $a_{j_0,1} \neq 0$ et on utilise l'algorithme du pivot de Gauss. Il est clair que la famille $\left(u_j - \frac{a_{j,1}}{a_{j_0,1}} u_{j_0}\right)_{j=1,\cdots,n+2; j \neq j_0}$ appartient à l'espace engendré par $e_2, \cdots, e_{n+1}$. Par récurrence, on en déduit que cette famille est liée et la conclusion s'ensuit aisément. □

On en déduit le lemme suivant.

Lemme 6.2. *Soit $(u_1, \ldots, u_n)$ un système générateur d'un espace vectoriel E et $(v_1, \ldots, v_n)$ un système libre de E. Alors $(v_1, \ldots, v_n)$ est une base.*

Théorème 6.3. *Deux bases d'un même espace vectoriel ont même nombre d'éléments (fini ou infini).*

Démonstration. C'est une conséquence du lemme précédent ! □

En fait, dans un espace vectoriel engendré par une famille finie d'éléments, on peut montrer qu'il y a toujours une base sans utiliser le lemme de Zorn.

Proposition 6.4. *Soit E un espace vectoriel engendré par une famille finie. Alors E possède une base.*

Définition 6.5. *Le nombre d'éléments n d'une base d'un espace vectoriel E est la* dimension *de E. On écrit $\dim E = n$. Un espace de dimension 1 (resp. 2) est* une droite *(resp.* un plan*).*

L'espace nul $\{0\}$ est de dimension nulle car sa base est l'ensemble vide $\emptyset$ qui n'a aucun élément. Plus généralement, K^n est de dimension n car sa base canonique possède n éléments et il en va de même de $\check{K}^n$. Et $K[X]$ est de dimension infinie alors que $\dim K[X]_{\leq n} = n + 1$. Enfin, $\mathcal{C}(I)$ est de dimension infinie (en général).

Un hyperplan de K^2 n'est autre qu'une droite, en effet, c'est un ensemble de la forme

$$\{(x, y) \in K^2, \quad ax + by = 0\},$$

avec a, b non tous les deux nuls, dont une base est $(b, -a)$. Et réciproquement, la droite engendrée par un vecteur non nul (a, b) est l'hyperplan d'équation $bx - ay = 0$.

Proposition 6.6. *Si E est un espace vectoriel de dimension n, alors toute base a n éléments, tout système générateur a au moins n éléments et tout système libre a au plus n éléments.*

Les vecteurs

$$(3, 5, 7, 11), (5, 7, 11, 13), (7, 11, 13, 17), (11, 13, 17, 19), (13, 17, 19, 23)$$

de $\mathbb{R}^4$ sont ils linéairement indépendants ? Non car, dans un espace de dimension 4, 5 vecteurs sont toujours dépendants ! Inutile de calculer à moins d'avoir besoin d'une relation de dépendance linéaire.

Proposition 6.7. *Soit E un espace vectoriel de dimension finie n et $\mathcal{S}$ une partie à n éléments de E. Alors $\mathcal{S}$ est libre si et seulement si $\mathcal{S}$ est une base si et seulement si $\mathcal{S}$ est générateur.*

Démonstration. Si $\mathcal{S}$ est libre, elle est contenue dans une base (qui contient n éléments !). De même, si $\mathcal{S}$ est générateur, il contient une base à n éléments : il est donc égal à cette base. $\square$

En pratique, pour montrer que 3 vecteurs de $\mathbb{R}^3$ forment une base, il suffit de montrer qu'ils sont linéairement indépendants *ou* générateurs !

Proposition 6.8. *Si F est un sous-espace vectoriel d'un espace vectoriel E, alors*

$$\dim F \leq \dim E.$$

Si de plus, on a $\dim F = \dim E < \infty$, *alors $E = F$.*

Démonstration. On complète une base de F qui est un système libre de E en une base de E. Bien sûr, si E et F ont même dimension finie, ils vont avoir même base et donc être engendrés par les mêmes éléments. Ils seront donc égaux. $\square$

Par exemple, on voit que les seuls sous-espaces vectoriels de K sont $\{0\}$ et K car $\dim K = 1$. De même, outre $\{0\}$ et K^2, les seuls sous-espaces vectoriels de K^2 sont les droites vectorielles. Où encore, outre $\{0\}$ et K^3, les seuls sous-espaces vectoriels de K^3 sont les droites et les plans vectoriels.

Montrons par exemple que l'hyperplan H d'équation $x + y + z = 0$ dans $\mathbb{R}^3$ a pour base $u := (1, -1, 0)$ et $v := (1, 0, -1)$. Tout d'abord, u et v ne sont pas multiples l'un de l'autre. Il sont donc indépendants. Donc l'espace F engendré par u et v est de dimension 2. L'hyperplan H ne contient pas $(1, 0, 0)$ et donc $H \neq \mathbb{R}^3$. Donc, H est de dimension strictement inférieure à 3. D'autre part, comme u et v sont dans H, on a $F \subset H$. On a donc nécessairement $F = H$. Sinon, F aurait une dimension strictement inférieure à H, c'est à dire au plus 1.

Cet argument s'étend pour montrer qu'un hyperplan de $\mathbb{R}^3$ n'est autre qu'un plan.

Proposition 6.9. *Soit $f : E \to F$ une application linéaire. Alors, si f est injective (resp. surjective, resp. bijective), on a*

$$\dim E \leq \dim F$$

$$\text{(\textit{resp.} } \dim F \leq \dim E,$$

$$\textit{resp. } \dim E = \dim F).$$

Démonstration. Si f est injective, l'image d'une base de E est un système libre de F qui a même nombre d'éléments et qui est contenu dans une base de F. Et si f est surjective, l'image d'une base de E est un système générateur de F qui contient une base de F. □

L'application linéaire

$$\begin{array}{rcl} \mathbb{R}^4 & \longrightarrow & \mathbb{R}^3 \\ (x,y,z,t) & \longmapsto & (3x+5y, 7y+11z, 13z+17t) \end{array}$$

est elle injective ? Non, parce que $4 > 3$! Et l'application linéaire

$$\begin{array}{rcl} \mathbb{R}^4 & \longrightarrow & \mathbb{R}^5 \\ (x,y,z,t) & \longmapsto & (3x+5y, 7y+11z, 13z+17t, 19t+23x, 27x+31y) \end{array}.$$

Est elle surjective ? Non, parce que $4 < 5$!

Proposition 6.10. *Deux espaces vectoriels de même dimension finie sont isomorphes. En particulier, tout espace vectoriel sur K de dimension finie n est isomorphe à K^n.*

Démonstration. En effet, si on se donne une base de chaque, il existe une unique application linéaire qui envoie la première sur la seconde. Et c'est un isomorphisme.

□

Définition 6.11. *Des vecteurs sont* colinéaires *s'ils appartiennent à une même droite et* coplanaires *s'ils appartiennent à un même plan.*

Par exemple, les vecteurs $(2, -4)$ et $(-3, 6)$ sont colinéaires.

Proposition 6.12. *Deux vecteurs sont linéairement dépendants si et seulement si ils sont colinéaires. De même, trois vecteurs sont linéairement dépendants si et seulement si ils sont coplanaires.*

Démonstration. Plus généralement, n vecteurs sont dépendants si et seulement si l'espace engendré est de dimension strictement inférieure à n et cette dernière condition est équivalente au fait d'appartenir tous à un sous-espace de dimension strictement inférieure à n. On applique ça à $n = 2$ et $n = 3$. □

Mais attention, trois vecteurs non colinéaires deux à deux peuvent cependant être coplanaires. Prendre par exemple $(1, 0, 0), (0, 1, 0), (1, 1, 0)$ dans $\mathbb{R}^3$.

Proposition 6.13. *Un sous-espace vectoriel H d'un espace vectoriel E est un hyperplan si et seulement si il existe une droite D de E telle que*

$$E = H \oplus D.$$

Et ceci est alors vrai pour toute droite D non contenue dans H.

Démonstration. Si H est un hyperplan, on peut écrire $H = \ker f$ avec $f : E \to K$ linéaire non nulle. Il existe donc $u \in E$ tel que $f(u) \neq 0$ et on notera D la droite dirigée par u. Notons que, réciproquement, si D est une droite non contenue dans H et dirigée par un vecteur u, alors $f(u) \neq 0$.

On a bien sûr $H \cap D = \{0\}$ parce-que c'est un sous-espace vectoriel de D (qui a dimension 1) distinct de D (car u n'est pas dedans). Et on peut facilement voir que tout $v \in E$ s'écrit comme somme d'un élément de H et d'un élément de D :

$$v = (v - \frac{f(v)}{f(u)}u) + \frac{f(v)}{f(u)}u.$$

Réciproquement, si $E = H \oplus D$, alors H est le noyau de l'application composée de la projection sur D et d'un isomorphisme quelconque entre D et K. C'est donc bien un hyperplan. $\square$

7. Rang

Proposition 7.1 (Dimension d'une somme directe). *Soit E un espace vectoriel et E_1 et E_2 deux sous-espaces vectoriels tels que $E = E_1 \oplus E_2$. Alors, on a :*

$$\dim E = \dim E_1 + \dim E_2.$$

Démonstration. Soient $\mathcal{B}_1$ et $\mathcal{B}_2$ des bases de E_1 et E_2 respectivement. Il est aisé de voir que $\mathcal{B}_1 \cup \mathcal{B}_2$ est une base de E. □

Nous en déduisons facilement que :

Proposition 7.2. *Soient E_1 et E_2 deux espaces vectoriels. Alors, on a :*

$$\dim(E_1 \times E_2) = \dim E_1 + \dim E_2.$$

Cela permet de redémontrer par récurrence sur n que $\dim K^n = n$.

Théorème 7.3 (Théorème du rang). *Si $f : E \to F$ est une application linéaire, alors*

$$\dim E = \dim \ker f + \dim \operatorname{Im} f.$$

Démonstration. Notons $E_1 = \ker f$ et considérons un supplémentaire E_2 de E_1 dans E :

$$E = E_1 \oplus E_2.$$

On constate que la restriction de f à E_2 et à valeurs dans $\operatorname{Im} f$. Cette restriction est un isomorphisme de sorte que $\dim E_2 = \dim \operatorname{Im} f$. La conclusion s'ensuit aisément. □

La définition ci-dessous justifie la terminologie « Théorème du rang » :

Définition 7.4.

i) *Le* rang *d'une partie $\mathcal{S}$ d'un espace vectoriel E est la dimension de l'espace engendré par ce système. On la note* $\mathrm{rg}(\mathcal{S})$.

ii) *Le* rang *d'une application linéaire f est la dimension de* $\mathrm{Im} f$. *On le note* $\mathrm{rg}(f)$.

iii) *Si A est la matrice d'une application linéaire $f : K^m \to K^n$ (dans la base canonique), le* rang *de A est le rang de f. On le note* $\mathrm{rg}(A)$.

iv) *Le* rang *d'un système linéaire*

$$(S) : \left\{ \begin{array}{rcl} a_{11}x_1 + \cdots + a_{1m}x_m & = & b_1 \\ \vdots & & \vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m & = & b_n \end{array} \right. ,$$

est le rang de la matrice associée. On le note $\mathrm{rg}(S)$.

Par exemple, le rang de $\mathcal{S} := \{(1, -1, 0), (0, 1, -1), (-1, 0, 1)\} \subset \mathbb{R}^3$ est 2 car l'espace engendré est le plan d'équation $x + y + z = 0$.

Les deux premières définitions ne sont pas indépendantes : si $\mathcal{B}$ est une base de E, alors le rang de f est égal au rang de $f(\mathcal{B})$ puisque ce dernier est générateur de

$\mathrm{Im}(f)$. Par exemple, l'application

$$\begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ (x,y,z) & \longmapsto & (x-z, y-x, z-y) \end{array}$$

est de rang 2 car l'image de la base canonique de $\mathbb{R}^3$ est justement l'ensemble $\mathcal{S}$ ci-dessus.

On voit ainsi que le rang d'une matrice est le rang du système formé par ses vecteurs colonnes. Dans notre exemple, on a

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

Enfin, dans le cas d'un système linéaire homogène de rang r, le théorème du rang nous dit que l'espace des solutions est de dimension $n - r$.

Par exemple,

$$(S) : \left\{ \begin{array}{rcl} x - z & = & 0 \\ y - x & = & 0 \\ z - x & = & 0 \end{array} \right. ,$$

est de rang 2 et l'ensemble des solutions est donc une droite $(3 - 2 = 1)$. C'est bien sûr la droite de direction $(1, 1, 1)$.

Proposition 7.5. *Soient $f : E \to F$ et $g : F \to G$ deux applications linéaires. Si f est surjective, on a $\mathrm{rg}(g \circ f) = \mathrm{rg}(g)$ et si g est injective, on a $\mathrm{rg}(g \circ f) = \mathrm{rg}(f)$.*

Démonstration. On sait que si f est surjective, on a $\mathrm{Im}(g\circ f) = \mathrm{Im}(g)$ et la première assertion en découle. De même, si g est injective, on utilise le théorème du rang et le fait que $\ker(g \circ f) = \ker(f)$. □

On utilisera ce résultat essentiellement pour des applications linéaires bijectives.

Corollaire 7.6. *Soit $A \in M_{n\times m}(K)$. Si $P \in GL_n(K)$, alors $\mathrm{rg}(PA) = \mathrm{rg}(A)$. De même, si $Q \in GL_m(K)$, alors $\mathrm{rg}(AQ) = \mathrm{rg}(A)$.*

Démonstration. On écrit A comme la matrice d'une application f et P comme la matrice d'une application φ (dans la base canonique). Comme P est inversible, on sait que φ est bijective et donc en particulier injective. Il suit que $\mathrm{rg}(\varphi \circ f) = \mathrm{rg} f$ et comme la matrice de la composée $\varphi \circ f$ n'est autre que le produit PA, on a bien $\mathrm{rg}(PA) = \mathrm{rg}(A)$. Et de même pour l'autre égalité. □

Corollaire 7.7. *Les opérations élémentaires sur les lignes ou sur les colonnes d'une matrice ne changent pas son rang.*

Démonstration. En effet, on sait qu'une faire une opération élémentaire revient à multiplier par une matrice élémentaire, qui est inversible. □

En fait, le rang de la matrice est le nombre de lignes non nulles après avoir appliqué la méthode du pivot de Gauss.

Proposition 7.8 (Relation de Grassmann). *Si E_1 et E_2 sont deux sous-espaces vectoriels de E, alors*

$$\dim E_1 + \dim E_2 = \dim(E_1 + E_2) + \dim(E_1 \cap E_2).$$

Démonstration. On sait que l'application

$$\begin{array}{rrcl} \Phi : & E_1 \times E_2 & \longrightarrow & E \\ & (u_1, u_2) & \longmapsto & u_1 + u_2 \end{array}$$

est linéaire, que son image est $E_1 + E_2$ et que son noyau est isomorphe à $E_1 \cap E_2$. On a donc, grâce au théorème du rang,

$$\dim(E_1 \times E_2) = \dim(E_1 + E_2) + \dim E_1 \cap E_2.$$

□

On peut en déduire par exemple que deux plans de $\mathbb{R}^3$ ont une droite en commun : en effet, si H et H' désignent ces plans, on a

$$\dim(H \cap H') = 2 + 2 - \dim(H + H') \geq 4 - 3 = 1$$

Corollaire 7.9. *Soient E_1 et E_2 deux sous-espaces vectoriels d'un espace vectoriel E de dimension finie. Alors*

$$E = E_1 \oplus E_2 \Leftrightarrow \begin{cases} \dim E = \dim E_1 + \dim E_2 \\ E_1 \cap E_2 = \{0\}. \end{cases}$$

Démonstration. On connaît déjà ce résultat avec la première condition remplacée par $E = E_1 + E_2$. Il suffit donc de montrer que si $E_1 \cap E_2 = \{0\}$, on a

$$E = E_1 + E_2 \Leftrightarrow \dim E = \dim E_1 + \dim E_2.$$

Grâce à la relation de Grassmann, on a $\dim E_1 + \dim E_2 = \dim(E_1 + E_2)$. Et comme $E_1 + E_2$ est un sous-espace vectoriel de E qui est de dimension finie, on a bien

$$E = E_1 + E_2 \Leftrightarrow \dim E = \dim E_1 + \dim E_2.$$

□

Par exemple, on voit que le supplémentaire d'une droite dans un plan est une autre droite. Et dans l'espace, c'est un plan ne contenant pas cette droite.

Proposition 7.10. *Un sous-espace vectoriel H d'un espace vectoriel E de dimension finie n est un hyperplan si et seulement si* $\dim H = n - 1$.

Démonstration. On sait que H est un hyperplan si et seulement si il existe une droite D avec $E = H \oplus D$. En général, on sait qu'il existe toujours un supplémentaire D pour H et on a alors $\dim D + \dim H = \dim E$. Comme $\dim E = n$, on voit que $\dim D = 1$ si et seulement si $\dim H = n - 1$. □

Une autre application du théorème du rang est la suivante.

Proposition 7.11. *Soit E un espace vectoriel de dimension finie n et f un endomorphisme de E. Alors f est injective si et seulement si f est surjective si et seulement si f est bijective si et seulement si* $\mathrm{rg}(f) = n$.

Démonstration. Par définition, f est surjective si et seulement si $\mathrm{Im}(f) = E$ et comme $\dim E = n$ est finie, cela signifie que $\mathrm{rg}(f) = \dim \mathrm{Im}(f) = \dim E = n$. Maintenant, grâce au théorème du rang, on a $\mathrm{rg}(f) = n - \dim \ker f$ et on voit que $\mathrm{rg}(f) = n$ si et seulement si $\dim \ker f = 0$, c'est à dire $\ker f = \{0\}$, ce qui signifie bien que f est injective.

□

Corollaire 7.12.
i) *Une matrice carrée A d'ordre n est inversible si et seulement si* $\mathrm{rg}(A) = n$.
ii) *Si A et B sont deux matrices carrées d'ordre n telles que $AB = I$, alors A et B sont inversibles et inverses l'une de l'autre.*

Démonstration. Pour la première assertion, on écrit A comme matrice d'une application f dans la base canonique et on traduit le résultat précédent.

Pour la seconde assertion, on écrit A et B comme matrices d'applications f et g respectivement si bien que AB est la matrice de $f \circ g$. Notre hypothèse nous dit que $f \circ g = \mathrm{Id}_E$ et implique en particulier que f est surjective. Celle-ci est donc bijective si bien que A est inversible. On en déduit que

$$B = A^{-1}AB = A^{-1}I = A^{-1}$$

comme annoncé. Il suit formellement que B aussi est inversible et que $B^{-1} = A$. □

8. Matrices d'applications linéaires

Dans cette section, tous les espaces vectoriels sont supposés de dimension finie. On commence par les vecteurs colonnes.

Définition 8.1. *Soit E un espace vectoriel muni d'une base*

$$\mathcal{B} := (u_1, \ldots, u_n).$$

Si $u \in E$ a pour coordonnées $(a_1, \ldots, a_n)$ dans cette base, le vecteur colonne *associé à u est*

$$[u]_{\mathcal{B}} := \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

Par exemple, le vecteur colonne associé à $(X-3)^2$ dans la base canonique de $K[X]_{\leq 2}$ est

$$\begin{bmatrix} 9 \\ -6 \\ 1 \end{bmatrix}.$$

Et le vecteur colonne associé à $(2, 0)$ dans la base $\{(1, 1), (1, -1)\}$ de $\mathbb{R}^2$ est

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

On rappelle qu'il revient au même de se donner une base $\mathcal{B}$ d'un espace vectoriel E ou une paramétrisation (bijective) $\Phi : K^n \to E$: si $(e_1, \ldots, e_n)$ désigne la base canonique de K^n et $\mathcal{B} := (u_1, \ldots, u_n)$, on a tout simplement $\Phi(e_i) = u_i$.

Lemme 8.2. *Si $\Phi : K^n \to E$ est la paramétrisation d'un espace vectoriel E correspondant à une base $\mathcal{B}$ de E, alors pour tout $u \in E$, on a $[u]_{\mathcal{B}} = [\Phi^{-1}(u)]$.*

Démonstration. Par définition, si $(e_1, \ldots, e_n)$ désigne la base canonique de K^n et $\mathcal{B} := (u_1, \ldots, u_n)$ alors, $\Phi(e_i) = u_i$ pour tout i. Donc, si $u := a_1u_1 + \cdots + a_nu_n$, on a bien $\Phi^{-1}(u) = a_1e_1 + \cdots a_ne_n$. □

On passe maintenant aux applications linéaires.

Définition 8.3. *Soient E et F deux espaces vectoriels munis de bases*

$$\mathcal{B} := (u_1, \ldots, u_m)$$

et $\mathcal{C}$ respectivement. Alors, la matrice d'une application linéaire *$f : E \to F$ est*

$$[f]_{\mathcal{B}}^{\mathcal{C}} := \big[[f(u_1)]_{\mathcal{C}} \cdots [f(u_m)]_{\mathcal{C}}\big].$$

Lorsque $E = F$ et $\mathcal{B} = \mathcal{C}$, on écrit tout simplement $[f]_{\mathcal{B}}$.

Par exemple, dans la base canonique, la matrice de l'application linéaire

$$\begin{aligned} \mathbb{R}[X]_{\leq 2} &\longrightarrow R[X]_{\leq 2} \\ P &\longmapsto P' \end{aligned}$$

est

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

Le lemme suivant permet de ramener toutes les questions sur les matrices d'applications linéaires à des espaces de la forme K^n.

Lemme 8.4. *Soit $f : E \to F$ une application linéaire, $\Phi : K^m \to E$ la paramétrisation associée à une base $\mathcal{B}$ de E et $\Psi : K^n \to F$ la paramétrisation associée à une base $\mathcal{C}$ de F. On a alors*

$$[f]_{\mathcal{B}}^{\mathcal{C}} = [\Psi^{-1} \circ f \circ \Phi].$$

Démonstration. Si $(e_1, \ldots, e_m)$ désigne la base canonique de K^m et $\mathcal{B} := (u_1, \ldots, u_m)$, il s'agit de montrer que pour tout i, on a

$$[(\Psi^{-1} \circ f \circ \Phi)(e_i)] = [f(u_i)]_{\mathcal{C}},$$

ce qui résulte immédiatement du lemme précédent. □

Proposition 8.5. *Soient E et F deux espaces vectoriels de dimension m et n, respectivement, munis de bases $\mathcal{B}$ et $\mathcal{C}$, respectivement. Alors, l'application*

$$\begin{aligned} L(E,F) &\longrightarrow M_{n\times m}(K) \\ f &\longmapsto [f]_{\mathcal{B}}^{\mathcal{C}} \end{aligned}$$

est un isomorphisme d'espaces vectoriels.

Démonstration. Si $\Phi : K^m \to E$ est la paramétrisation associée à $\mathcal{B}$ et $\Psi : K^n \to F$ la paramétrisation associée à $\mathcal{C}$, notre application se décompose en deux isomorphismes

$$\begin{array}{ccccc} L(E,F) & \xrightarrow{\simeq} & L(K^m, K^n) & \xrightarrow{\simeq} & M_{n\times m}(K). \\ f & \longmapsto & \Psi^{-1} \circ f \circ \Phi & & \\ & & g & \longmapsto & [g] \end{array}$$

En fait, on sait déjà que la seconde application est un isomorphisme. Comme la première est clairement bijective, d'inverse $g \mapsto \Psi \circ g \circ \Phi^{-1}$, il faut juste s'assurer qu'elle est linéaire, c'est à dire l'identité

$$\Psi^{-1} \circ (\lambda f + \mu g) \circ \Phi = \lambda(\Psi^{-1} \circ f \circ \Phi) + \mu(\Psi^{-1} \circ g \circ \Phi).$$

□

Corollaire 8.6. *Si E et F sont deux espaces vectoriels, on a*

$$\dim L(E,F) = \dim E \times \dim F.$$

Démonstration. □

En particulier, $\dim \check{E} = \dim E$ et $\dim L(E) = (\dim E)^2$.

Proposition 8.7. *Si $\mathcal{B}$ et $\mathcal{C}$ sont des bases d'espaces vectoriels E et F respectivement, si $u \in E$ et si $f : E \to F$ est linéaire, on a*

$$[f(u)]_{\mathcal{C}} = [f]_{\mathcal{B}}^{\mathcal{C}}[u]_{\mathcal{B}}.$$

Si $\mathcal{B}$, $\mathcal{C}$ et $\mathcal{D}$ sont des bases de E, F et G respectivement, et si $f : E \to F$ et $g : F \to G$ sont linéaires, on a

$$[g \circ f]_{\mathcal{B}}^{\mathcal{D}} = [g]_{\mathcal{C}}^{\mathcal{D}}[f]_{\mathcal{B}}^{\mathcal{C}}.$$

Démonstration. À nouveau, on se ramène au cas de la base canonique en utilisant des paramétrisations. Avec nos notations habituelles, la première assertion se ramènera donc à vérifier que

$$[\Psi^{-1}(f(u))] = [\Psi^{-1} \circ f \circ \Phi][\Phi^{-1}(u)],$$

et comme nous connaissons déjà le résultat pour la base canonique, on est réduit à

$$\Psi^{-1}(f(u)) = (\Psi^{-1} \circ f \circ \Phi)(\Phi^{-1}(u)).$$

La seconde assertion se démontre de la même manière en faisant intervenir trois paramétrisations. □

Corollaire 8.8. *Soit E un espace vectoriel de dimension n muni d'une base $\mathcal{B}$. Alors,*

i) *L'application*

$$\begin{aligned} L(E) &\longrightarrow M_n(K) \\ f &\longmapsto [f]_{\mathcal{B}} \end{aligned}$$

est un isomorphisme d'algèbres : C'est un isomorphisme d'espaces vectoriels et on a en plus, pour $f, g \in L(E)$,

$$[g \circ f]_{\mathcal{B}} = [g]_{\mathcal{B}}[f]_{\mathcal{B}}.$$

ii) *L'application*

$$\begin{aligned} GL(E) &\longrightarrow GL_n(K) \\ f &\longmapsto [f]_{\mathcal{B}} \end{aligned}$$

est un isomorphisme de groupes : C'est une bijection et on a en plus, pour $f, g \in L(E)$,

$$[g \circ f]_{\mathcal{B}} = [g]_{\mathcal{B}}[f]_{\mathcal{B}}.$$

Démonstration. Cela résulte des résultats analogues pour les bases canoniques en passant par les paramétrisations. □

Définition 8.9. *Soit E un espace vectoriel muni de deux bases $\mathcal{B}$ et $\mathcal{B}'$. La* matrice de passage *de $\mathcal{B}$ à $\mathcal{B}'$ est $[\mathrm{Id}_E]_{\mathcal{B}'}^{\mathcal{B}}$.*

Attention à l'ordre : on munit l'espace de *départ* de la « nouvelle » base $\mathcal{B}'$ et l'espace d'*arrivée* de l' « ancienne » base $\mathcal{B}$. En pratique, on écrit les vecteurs de la nouvelle base dans l'ancienne.

Pour donner un exemple, il faut deux bases du même espace vectoriel. Par exemple, on peut prendre pour $\mathcal{B}$ la base canonique de $\mathbb{R}^2$ et $\mathcal{B}' := \{(1,1),(1,-1)\}$. La matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est alors

$$P := \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Mais la matrice de passage dans l'autre sens, de $\mathcal{B}'$ à $\mathcal{B}$ est

$$P^{-1} := \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

Ceci est général et on a le résultat suivant :

Proposition 8.10. i) *Une matrice de passage est inversible. Plus précisément, l'inverse de la matrice de passage d'une base $\mathcal{B}$ à une base $\mathcal{B}'$ est la matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$.*

ii) *Soit E un espace vectoriel muni de deux bases $\mathcal{B}$ et $\mathcal{B}'$ et F un espace vectoriel muni de deux bases $\mathcal{C}$ et $\mathcal{C}'$. On note P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ et Q la matrice de passage de $\mathcal{C}$ à $\mathcal{C}'$. Soit $f : E \to F$ une application linéaire, A sa matrice dans les bases $\mathcal{B}$ et $\mathcal{C}$ et A' sa matrice dans les bases $\mathcal{B}'$ et $\mathcal{C}'$. On a alors,*

$$A' = Q^{-1}AP.$$

Démonstration. La première assertion est immédiate (vérifier). Pour la seconde, on remarque que

$$f := \mathrm{Id}_F \circ f \circ \mathrm{Id}_E$$

et donc que

$$[f]_{\mathcal{B}'}^{\mathcal{C}'} := [\mathrm{Id}_F]_{\mathcal{C}}^{\mathcal{C}'}[F]_{\mathcal{B}}^{\mathcal{C}}[\mathrm{Id}_E]_{\mathcal{B}'}^{\mathcal{B}}.$$

□

Définition 8.11. *Si A et A' sont deux matrices $n \times m$ telles qu'il existe des matrices inversibles P et Q avec $A' = Q^{-1}AP$, on dit que A et A' sont* équivalentes.

Si A et A' sont deux matrices carrées d'ordre n telles qu'il existe une matrice inversible P avec $A' = P^{-1}AP$, on dit qu'elles sont semblables.

On peut montrer que deux matrices sont équivalentes si et seulement si elles ont même rang (dans un sens c'est clair, non). Par contre le fait d'être semblable est une propriété très forte.

Regardons par exemple l'application linéaire

$$\begin{array}{ccc} \mathbb{R}^2 & \xrightarrow{f} & \mathbb{R}^2 \\ (x,y) & \longmapsto & (x+2y, 2x+y) \end{array}$$

Sa matrice dans la base canonique est

$$A := \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}.$$

Il suit que sa matrice dans la base $\mathcal{B}' := \{(1,1),(1,-1)\}$ est

$$A' = P^{-1}AP = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}.$$

C'est pratique pour calculer les puissances de A :

$$A^n = PA'^n P^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3^n & 0 \\ 0 & (-1)^n \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{3^n+(-1)^n}{2} & \frac{3^n-(-1)^n}{2} \\ \frac{3^n-(-1)^n}{2} & \frac{3^n+(-1)^n}{2} \end{bmatrix}.$$

Application : la double suite récurrente

$$\begin{cases} u_{n+1} & = & u_n + 2v_n \\ v_{n+1} & = & 2u_n + v_n \end{cases}$$

avec $u_0 = 3$ et $v_0 = 2$? On écrit

$$\begin{bmatrix} u_{n+1} \\ v_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} u_n \\ v_n \end{bmatrix}$$

ou encore

$$X_{n+1} = AX_n \quad \text{avec} \quad X_n := \begin{bmatrix} u_n \\ v_n \end{bmatrix} \quad \text{et} \quad X_0 := \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

On a donc

$$X_n := A^n X_0 = \begin{bmatrix} \frac{3^n+(-1)^n}{2} & \frac{3^n-(-1)^n}{2} \\ \frac{3^n-(-1)^n}{2} & \frac{3^n+(-1)^n}{2} \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{5.3^n+(-1)^n}{2} \\ \frac{5.3^n-(-1)^n}{2} \end{bmatrix}$$

si bien que

$$u_n = \frac{5.3^n + (-1)^n}{2} \quad \text{et} \quad v_n = \frac{5.3^n - (-1)^n}{2}.$$

9. Dualité

Proposition 9.1. *Soit F un sous-espace vectoriel de dimension m d'un espace vectoriel E de dimension finie n et $r := n-m$. Alors, il existe $f_1, \dots, f_r \in \check{E}$ linéairement indépendantes telles que*

$$F = \{u \in E, \quad f_1(u) = \cdots = f_r(u) = 0\}$$

Démonstration. Soit G un supplémentaire de F dans E et $p : E \to G$ la projection. Soit $\phi : K^r \simeq G$ une paramétrisation bijective (donné par une base de G). On regarde la composée $f := \phi^{-1} \circ p$ qui est donnée par r formes linéaires $f_1, \dots, f_r$. Comme F est le noyau de f, on a bien l'égalité annoncée. Le fait que $f_1, \dots, f_r$ sont linéairement indépendantes résulte du fait que q est surjectif grace au corollaire 9.13 que nous démontrerons plus bas. □

En particulier, on voit qu'un sous-espace vectoriel de K^n est toujours l'ensemble des solutions d'un système homogène.

Proposition 9.2. *Si $\mathcal{B} := (u_1, \dots, u_n)$ est une base d'un espace vectoriel E, alors*

$$\check{\mathcal{B}} := (\check{u}_1, \dots, \check{u}_n),$$

où $\check{u}_i$ est l'unique forme linéaire telle que

$$\check{u}_i(u_j) = \delta_{ij} = \begin{cases} 1 & \text{si} \quad i = j \\ 0 & \text{sinon} \end{cases},$$

est une base de $\check{E}$. De plus, on a toujours

$$f = f(u_1)\check{u}_1 + \cdots + f(u_n)\check{u}_n$$

et

$$u = \check{u}_1(u)u_1 + \cdots + \check{u}_n(u)u_n.$$

Démonstration. On montre d'abord que si $f \in \check{E}$, alors

$$f = f(u_1)\check{u}_1 + \cdots + f(u_n)\check{u}_n.$$

Pour cela, il suffit d'appliquer chaque membre à u_i pour $i = 1, \dots, n$. On en déduit que $(\check{u}_1, \dots, \check{u}_n)$ est générateur et donc une base car $\dim \check{E} = \dim E$. Pour démontrer la seconde formule, on écrit

$$u = \lambda_1 u_1 + \cdots \lambda_n u_n$$

et on calcule $\check{u}_i(u) = \lambda_i$. □

On voit donc que les composantes de u sont les $\check{u}_i(u)$.

Définition 9.3. *Avec les notations de la proposition, on dit que $\mathcal{B}^*$ est la* base duale *de $\mathcal{B}$.*

Par exemple, la base duale de la base canonique $(e_1, \dots, e_n)$ de K^n est la base canonique $(p_1, \dots, p_n)$ de $\check{K}^n$ (formée des projections sur les axes). Aussi, la base duale de la base canonique de $\mathbb{R}[X]_{\leq n}$ est formée des formes $P \mapsto \frac{P^{(i)}(0)}{i!}$.

Proposition 9.4. *Soit E un espace vectoriel sur K. Si $u \in E$, l'application*

$$\begin{aligned} \check{E} &\xrightarrow{\varphi_u} K \\ f &\longmapsto f(u) \end{aligned}$$

est une forme linéaire sur $\check{E}$. L'application induite

$$\begin{aligned} E &\longrightarrow \check{\check{E}} \\ u &\longmapsto \varphi_u \end{aligned}$$

est linéaire. Si de plus, $\dim E < \infty$, *c'est un isomorphisme.*

Démonstration. Les deux première assertions se vérifient aisément : la première résulte de l'identité

$$(\lambda f + \mu g)(u) = \lambda f(u) + \mu g(u)$$

et la seconde de l'identité

$$f(\lambda u + \mu v) = \lambda f(u) + \mu f(v).$$

Si de plus, $\dim E < \infty$, on peut choisir une base $(u_1, \ldots, u_n)$ de E. Comme nos espaces ont même dimension finie, il suffit pour conclure de montrer que notre application est injective. Si u est dans le noyau, alors pour tout $i = 1, \ldots n$, on a $\check{u}_i(u) = \varphi_u(\check{u}_i) = 0$. Il suit que

$$u = \check{u}_1(u)u_1 + \cdots + \check{u}_n(u)u_n = 0.$$

□

Définition 9.5. *Si $f : E \to F$ est une application linéaire, l'*application duale *est*

$$\begin{aligned} \check{f} : \check{F} &\longrightarrow \check{E} \\ g &\longrightarrow g \circ f. \end{aligned}$$

Proposition 9.6. *Si $f : E \to F$ est une application linéaire alors $\check{f}$ aussi.*

Démonstration. On a toujours

$$\check{f}(\lambda g + \mu h) = (\lambda g + \mu h) \circ f = \lambda g \circ f + \mu h \circ f = \lambda \check{f}(g) + \mu \check{f}(h)$$

si bien que $\check{f}$ est bien linéaire. □

Définition 9.7. *La* transposée *de la matrice*

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}$$

est la matrice

$${}^tA := \begin{bmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{1m} & \cdots & a_{nm} \end{bmatrix}.$$

Par exemple, la transposée de

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

est la matrice

$${}^tA := \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ 2 & 1 \end{bmatrix}.$$

Proposition 9.8. *Si E et F sont de dimension finie, avec bases $\mathcal{B}$ et $\mathcal{C}$ respectivement, et si $A := [f]_{\mathcal{B}}^{\mathcal{C}}$, alors $[\check{f}]_{\check{\mathcal{C}}}^{\check{\mathcal{B}}} = {}^tA$.*

Démonstration. Si on écrit $\mathcal{B} = (u_1, \ldots, u_n)$ et $\mathcal{C} = (v_1, \ldots, v_m)$, on a

$$\check{f}(\check{v}_i)(u_j) = (\check{v}_i \circ f)(u_j) = \check{v}_i(f(u_j))$$

et l'assertion en résulte : la j-ème composante de $\check{f}(\check{v}_i)$ est identique à la i-ème composante de $f(u_j)$. $\square$

Proposition 9.9. *Si E et F sont deux espaces vectoriels, l'application*

$$\begin{aligned} L(E,F) &\longrightarrow L(\check{F},\check{E}) \\ f &\longmapsto \check{f} \end{aligned}$$

est linéaire. C'est un isomorphisme si E et F sont de dimension finie.

Démonstration. On vérifie facilement que l'on a toujours

$$(\lambda_1\check{f}_1 + \lambda_2\check{f}_2) = (\lambda_1 f_1 + \lambda_2 f_2)\check{}.$$

Pour la seconde assertion, on peut fixer des bases et on voit immédiatement que si ${}^tA = 0$, alors $A = 0$. Il suit que le noyau de l'application est nul, donc que celle ci est injective et elle est donc bijective car les deux espaces ont même dimension. $\square$

Proposition 9.10. *Si $f : E \to F$ est une application linéaire, alors* $\mathrm{rg}(\check{f}) = \mathrm{rg}(f)$.

Démonstration. On note r le rang de f et on se donne une base $u_{r+1}, \ldots u_n$ de $\ker f$ que l'on complète en une base $u_1, \ldots, u_n$ de E. Il suit du théorème du rang que les vecteurs $v_1 := f(u_1), \ldots, v_r := f(u_r)$ sont linéairement indépendants et on complète en une base $v_1, \ldots, v_m$ de F. Nous savons que $\mathrm{Im}(\check{f})$ est engendré par les formes linéaires $\check{f}(\check{v}_j)$ pour tout $j = 1, \ldots m$. Nous allons les calculer.

Pour tout $i = 1, \ldots, n$, on a $\check{f}(\check{v}_j)(u_i) = \check{v}_j(f(u_i))$ qui est nul si $i > r$ car alors $f(u_i) = 0$ et égal à $\check{v}_j(v_i)$ si $i \leq r$, c'est à dire 0 si $i \neq j$ et 1 sinon. Pour résumer, on a $\check{f}(\check{v}_j) = \check{u}_j$ si $j \leq r$ et 0 sinon. On voit donc que $\mathrm{Im}(\check{f})$ est engendré par $\check{u}_1, \ldots, \check{u}_r$ et a donc même dimension r que $\mathrm{Im}(f)$. $\square$

Corollaire 9.11. *Si $A \in M_{n\times m}(K)$, on a* $\mathrm{Rg}({}^tA) = \mathrm{Rg}(A)$.

En particulier, on voit que le rang d'une matrice est égal au rang du système des ses vecteurs lignes.

Corollaire 9.12. *Soit $f : E \to F$ une application linéaire entre espaces vectoriels de dimension finie. Alors, f est injective (resp. surjective, resp. bijective) si et seulement si $\check{f}$ est surjective (resp. injective, resp. bijective).*

Démonstration. En effet, on sait que f est injective si et seulement si $\dim E = \mathrm{Rg}(f)$, ce qui est équivalent à $\dim \check{E} = \mathrm{Rg}(\check{f})$ qui veut dire que $\check{f}$ est surjective car $\check{E}$ est l'espace d'arrivée de $\check{f}$. On démontre de même la seconde assertion et la troisième en découle. $\square$

Nous obtenons maintenant le résultat dont nous avions besoin pour conclure la démonstration de la proposition 9.1.

Corollaire 9.13. *Soient E un espace vectoriel et $f_1, \ldots, f_r \in \check{E}$. Alors, $f_1, \ldots, f_r$ sont linéairement indépendants si et seulement si l'application*

$$\begin{array}{rcl} E & \xrightarrow{f} & K^r \\ u & \longmapsto & (f_1(u), \ldots f_r(u)) \end{array}$$

est surjective. Ils sont générateurs si et seulement si l'application est injective. Enfin, c'est une base de $\check{E}$ si et seulement si l'application est bijective.

Démonstration. Par définition, si p_i désigne la projection sur le i-ème facteur, alors $\check{f}(p_i) = f_i$. Comme les projections forment une base de $\check{K}^r$, on voit que $f_1, \ldots, f_r$ sont linéairement indépendants (resp. générateurs, resp. une base) si et seulement si $\check{f}$ est injective (resp. surjective, resp. bijective) et on vient de voir que c'est équivalent à f surjective (resp. injective, resp. bijective). $\square$

Proposition 9.14. *Soit E un espace vectoriel, $\mathcal{S}$ un ensemble de formes linéaires et*

$$F := \{u \in E, \quad \forall f \in \mathcal{S}, f(u) = 0\}.$$

Alors, $\dim F + \mathrm{Rg}(\mathcal{S}) = \dim E$.

Démonstration. On remarque tout d'abord que cet énoncé ne dépend que du sous-espace $\mathrm{Vect}(\mathcal{S}) \subset \check{E}$. On peut supposer qu'il est de dimension finie car sinon, E sera de dimension infinie. On peut donc supposer que $\mathcal{S} = (f_1, \ldots, f_r)$ avec $f_1, \ldots, f_r$ linéairement indépendantes. Et on applique le théorème du rang à l'application $f : E \to K^r$ correspondante, qui est surjective par le corollaire 9.13. $\square$

Proposition 9.15. i) *Si $f : E \to F$ et $g : F \to G$ sont deux applications linéaires, alors, $(g \circ f)\check{} = \check{f} \circ \check{g}$.*

ii) *Si $A \in M_{n\times m}(K)$ et $B \in M_{m\times l}(K)$ alors ${}^t(AB) = {}^tB{}^tA$.*

Démonstration. La seconde assertion résulte immédiatement de la première et celle ci se vérifie aisément : si $h \in \check{G}$, on a

$$(g \circ f)\check{}\,(h) = h \circ g \circ f$$

et

$$(\check{f} \circ \check{g})(h) = \check{f}(\check{g}(h)) = \check{g}(h) \circ f = h \circ g \circ f.$$

$\square$

On finit avec une définition.

Définition 9.16. *Une matrice $A \in M_n(K)$ est* symétrique *(resp.* antisymétrique*) si elle satisfait ${}^tA = A$ (resp. ${}^tA = -A$).*

On peut montrer que les matrices symétriques (resp. antisymétriques) forment un sous-espace vectoriel $SM_n(K)$ (resp. $AM_n(K)$) de $M_n(K)$ et que

$$M_n(\mathbb{R}) = SM_n(\mathbb{R}) \oplus AM_n(\mathbb{R}).$$

La décomposition est donnée par

$$A = \frac{A + {}^tA}{2} + \frac{A - {}^tA}{2}.$$

Application : chercher les carrés magiques

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & j \end{bmatrix}$$

c'est à dire, tels que la sommes des termes des 3 lignes, des 3 colonnes et des 2 diagonales soient toutes égales à une même quantité m. On doit donc résoudre un système de 8 équations à 9 inconnues et un paramètre. On remarque d'abord que si on a 2 carrés magiques de sommes m et m', la différence est un carré magique de somme $m - m'$. Il suffit donc d'ajouter un carré magique particulier de somme m comme

$$A = \begin{bmatrix} \frac{m}{3} & \frac{m}{3} & \frac{m}{3} \\ \frac{m}{3} & \frac{m}{3} & \frac{m}{3} \\ \frac{m}{3} & \frac{m}{3} & \frac{m}{3} \end{bmatrix} = \frac{m}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

à un carré magique général de somme 0 pour les trouver tous. On est donc ramené au cas $m = 0$. On remarque ensuite que le transposé d'un carré magique est un carré magique. Il suffit donc de chercher les carrés magiques symétriques, les carrés magiques antisymétriques et de les ajouter les uns aux autres. Un carré magique symétrique est de la forme

$$A = \begin{bmatrix} a & b & c \\ b & e & f \\ c & f & j \end{bmatrix}.$$

Ça fait un système homogène de 5 équations à 6 inconnues. Il faut le résoudre. Si on ajoute toutes les entrées de la matrice, on trouve que $a + 2b + 2c + e + j + 2f = 0$, et puisque $a + e + j = 0$, on en déduit que $b + c + f = 0$. Comme $a + b + c = 0$, on trouve $f = a$ et on peut réécrire la matrice

$$A = \begin{bmatrix} a & b & c \\ b & e & a \\ c & a & j \end{bmatrix}.$$

Sur chaque ligne, la somme des termes doit être la même si bien que $c = e$ et $b = j$. On a donc

$$A = \begin{bmatrix} a & j & e \\ j & e & a \\ e & a & j \end{bmatrix}.$$

avec les conditions que les diagonales sont nulles.

On a donc $e = 0$ et $j = -a$ si bien que

$$A = \begin{bmatrix} a & -a & 0 \\ -a & 0 & a \\ 0 & a & -a \end{bmatrix} = a \begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}.$$

On a ainsi tous les carrés magiques symétriques de somme nulle. C'est plus facile pour les carrés antisymétriques puisqu'on aura

$$A = \begin{bmatrix} 0 & b & c \\ -b & 0 & f \\ -c & -f & 0 \end{bmatrix}.$$

avec les conditions $f = b = -c$ si bien que

$$A = \begin{bmatrix} 0 & b & -b \\ -b & 0 & b \\ b & -b & 0 \end{bmatrix} = b \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}.$$

On voit donc que les carrés magiques de somme m forment un espace *affine* de dimension 2 : tout carré magique s'écrit de manière unique sous a forme

$$A = a \begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} + \frac{m}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} a + \frac{m}{3} & -a + b + \frac{m}{3} & -b + \frac{m}{3} \\ -a - b + \frac{m}{3} & \frac{m}{3} & a + b + \frac{m}{3} \\ b + \frac{m}{3} & a - b + \frac{m}{3} & -a + \frac{m}{3} \end{bmatrix}.$$

Par exemple, on prend $m = 12$ et $a = 1$ et $b = 2$ et on trouve

$$A = \begin{bmatrix} 5 & 5 & 2 \\ 1 & 4 & 7 \\ 6 & 3 & 3 \end{bmatrix}.$$

Bien sûr, si on veut des solutions en entiers naturels, il faut que $m = 3n$ et que $|a| + |b| \leq n$.

10. Déterminants

On rappelle que le *groupe symétrique* $\mathfrak{S}_n$ est l'ensemble de toutes les *permutations* σ de l'ensemble $\{1,\ldots,n\}$ (bijections de l'ensemble sur lui même). On note une permutation sous forme d'une matrice à deux lignes

$$\begin{bmatrix} 1 & \ldots & i & \cdots & n \\ \sigma(1) & \ldots & \sigma(i) & \cdots & \sigma(n) \end{bmatrix}$$

Une permutation qui échange simplement deux entiers distincts i est j est une *transposition* et on la note $(i \quad j)$. Plus généralement, une *permutation circulaire* (ou *cycle*) de longueur k envoie i_1 sur i_2 puis i_2 sur i_3 et ainsi de suite jusqu'à i_k qui est envoyé sur i_1. On la note $(i_1 \quad \cdots \quad i_k)$. Toute permutation peut s'obtenir en composant des transpositions. Toute permutation est composée d'un certains nombre de cycles disjoints.

La *signature* $\epsilon(\sigma)$ d'une transposition $(i \quad j)$ est -1 par définition. Plus généralement, la signature d'un cycle de longueur k est $(-1)^{k-1}$. En fait, on définit la signature d'une permutation quelconque σ comme $(-1)^i$ ou i est le nombre d'*inversions* de la permutation (chaque fois que $\sigma(j) < \sigma(i)$ alors que $i < j$). En fait ϵ est l'unique application de $\mathfrak{S}_n$ sur $\{\pm 1\}$ qui satisfait $\epsilon(\sigma\tau) = \epsilon(\sigma)\epsilon(\tau)$ (pour $n > 1$).

Nous aurons aussi besoin de considérer le *groupe alterné* $\mathfrak{A}_n$ de toutes les permutations *paires*, c'est à dire, dont la signature est 1. Remarquons que toutes les permutations impaires, dont la signature est -1, s'obtiennent en multipliant une transposition fixée τ par une permutation paire. On résume ça avec la formule $S_n = \mathfrak{A}_n \coprod \tau\mathfrak{A}_n$.

Dans tout ce chapitre, K sera égal à $\mathbb{R}$ ou $\mathbb{C}$ et E un K-ev de dimension n.

10.1. **Formes p-linéaires et alternées sur E.**

Définition 10.1 (Formes p-linéaires sur E). *Soient p et n deux entiers tels que : $1 \leqslant p \leqslant n$. On dit que*

$$\phi : \left\{ \begin{array}{ccc} E \times \cdots \times E & \to & K \\ (u_1, \cdots, u_p) & \mapsto & \phi(u_1, \cdots, u_p) \end{array} \right.$$

est une forme p-linéaire sur E lorsque pour tout $(u_1, \cdots, u_p) \in E^p$ et pour tout $i \in \{1, \cdots, p\}$, l'application partielle :

$$\phi_i : \left\{ \begin{array}{ccc} E & \to & K \\ x & \mapsto & \phi(u_1, \cdots, \underset{i}{x}, \cdots, u_p) \end{array} \right.$$

est une forme linéaire.

Il ne faut pas confondre forme bilinéaire et forme linéaire.

En effet, soit φ une forme bilinéaire sur E et soit $u_1, u_2, v_1, v_2 \in E$.

Si φ est bilinéaire, alors

$\varphi(u_1 + u_2, v_1 + v_2) = \varphi(u_1, v_1 + v_2) + \varphi(u_2, v_1 + v_2) = \varphi(u_1, v_1) + \varphi(u_1, v_2) + \varphi(u_2, v_1) + \varphi(u_2, v_2)$

Si φ est linéaire, alors $\varphi(u_1 + u_2, v_1 + v_2) = \varphi(u_1, v_1) + \varphi(u_2, v_2)$.

Donnons quelques exemples.

Exemples
Prenons $n = p = 2$, $E = K^2$ et en notant $u = (u_1, u_2)$ et $v = (v_1, v_2)$:

- $\phi_1(u, v) = u_1 v_2 - u_2 v_1$
- $\phi_2(u, v) = u_1 v_1 + u_2 v_2$
- $\phi_3(u, v) = \phi_1(u, v) + \phi_2(u, v)$

Proposition 10.2. *L'ensemble des formes p-linéaires sur E est un K-espace vectoriel.*

Définition 10.3 (Formes p-linéaires alternées)**.** *On dit qu'une forme p-linéaire ϕ est alternée lorsque : $\phi(u_1, \cdots, u_p) = 0$ dès qu'il existe $i_1 \neq i_2$ tels que $u_{i_1} = u_{i_2}$.*
On note $\mathcal{A}_p(E)$ l'ensemble des formes p-linéaires et alternées.

Proposition 10.4. *$\mathcal{A}_p(E)$ est un K-espace vectoriel.*

Exemples
Reprendre les exemples précédents et reconnaître les formes alternées.

Proposition 10.5. *Soit ϕ une forme p-linéaire sur E.*
- *ϕ est alternée si et seulement si, pour tout $(u_1, \cdots, u_p) \in E^p$ et pour tous $i, j \in \{1, \cdots, p\}$ tels que $i \neq j$, on a*

$$\phi(u_1, \cdots, \underset{i}{u_i}, \cdots, \underset{j}{u_j}, \cdots u_p) = -\phi(u_1, \cdots, \underset{i}{u_j}, \cdots, \underset{j}{u_i}, \cdots u_p).$$

- $\phi \in \mathcal{A}_p(E) \Longleftrightarrow \forall (u_1, \cdots, u_p) \in E^p \quad \forall \sigma \in \mathfrak{S}_p \quad \phi(u_{\sigma(1)}, \cdots, u_{\sigma(p)}) = \varepsilon(\sigma)\phi(u_1, \cdots, u_p).$

Proposition 10.6 (Familles liées et formes alternées)**.** *Soit $\phi \in \mathcal{A}_p(E)$ et $(u_1, \cdots u_p)$ une famille liée. Alors $\phi(u_1, \cdots, u_p) = 0$.*

En particulier, une famille qui contient 0 vérifie cela.

10.2. **Formes n-linéaires et alternées sur E.** À partir de maintenant, nous ne nous intéressons plus qu'à $\mathcal{A}_n(E)$.

10.2.1. *Développement suivant une base.* Notons $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ une base de E et considérons $\phi \in \mathcal{A}_n(E)$. On cherche à "calculer" $\phi(u_1, \cdots, u_n)$ à l'aide de $\mathcal{B}$. Écrivons pour chaque j :

$$u_j = \sum_{i=1}^{n} u_{i,j} e_i.$$

On écrit donc :

$$\phi(u_1, \cdots, u_n) = \phi\left(\sum_{i_1=1}^{n} u_{i_1,1} e_{i_1}, \sum_{i_2=1}^{n} u_{i_2,2} e_{i_2}, \cdots, \sum_{i_n=1}^{n} u_{i_n,n} e_{i_n}\right).$$

En utilisant la n-linéarité de ϕ, on a :

$$\phi(u_1, \cdots, u_n) = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_n=1}^{n} u_{i_1,1} u_{i_2,2} \cdots u_{i_n,n} \phi(e_{i_1}, e_{i_2}, \cdots, e_{i_n}).$$

Comme ϕ est alternée, cette somme ne comporte en fait que les termes pour lesquels les $(i_j)_{1\leqslant j\leqslant n}$ sont deux à deux distincts. Pour un tel $(i_j)_{1\leqslant j\leqslant n}$, il existe une unique permutation $\sigma \in \mathfrak{S}_n$ définie par $\sigma(j) = i_j$. Ainsi, on peut écrire :

$$\phi(u_1, \cdots, u_n) = \sum_{\sigma\in\mathfrak{S}_n} u_{\sigma(1),1}u_{\sigma(2),2}\cdots u_{\sigma(n),n}\phi(e_{\sigma(1)}, e_{\sigma(2)}, \cdots, e_{\sigma(n)}).$$

Or, ϕ est alternée, donc :

$$\forall\sigma \in \mathfrak{S}_n \quad \phi(e_{\sigma(1)}, e_{\sigma(2)}, \cdots, e_{\sigma(n)}) = \varepsilon(\sigma)\phi(e_1, e_2, \cdots, e_n).$$

Proposition 10.7. *Soit $\mathcal{B} = (e_i)_{1\leqslant i\leqslant n}$ une base de E et $\phi \in \mathcal{A}_n(E)$. Alors, pour tout $(u_1, \cdots u_n) \in E^n$ et en posant :*

$$u_j = \sum_{i=1}^{n} u_{i,j}e_i,$$

on a la formule suivante :

$$\phi(u_1, \cdots, u_n) = \left(\sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)u_{\sigma(1),1}u_{\sigma(2),2}\cdots u_{\sigma(n),n}\right)\phi(e_1, e_2, \cdots, e_n).$$

10.2.2. *Introduction du déterminant.* La dernière proposition mène à la définition suivante :

Définition 10.8 (Déterminant)**.** *Soit $\mathcal{B} = (e_i)_{1\leqslant i\leqslant n}$ une base de E. Pour tout $(u_1, \cdots u_n) \in E^n$ et en posant :*

$$u_j = \sum_{i=1}^{n} u_{i,j}e_i,$$

on définit :

$$\det_{\mathcal{B}}(u_1, \cdots, u_n) = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)u_{\sigma(1),1}u_{\sigma(2),2}\cdots u_{\sigma(n),n}.$$

Proposition 10.9. *Avec les notations précédentes, on a : $\det_{\mathcal{B}} \in \mathcal{A}_n(E)$. De plus, cette forme est non nulle et $\mathcal{A}_n(E)$ est une droite vectorielle. On l'appelle* ***déterminant dans la base $\mathcal{B}$****.*

Démonstration. Il y a plusieurs choses à vérifier.
Montrons déjà que $\det_{\mathcal{B}}$ est une **forme n-linéaire**.
Fixons $(u_2, \cdots u_n) \in E^{n-1}$ et montrons par exemple que l'application $\phi_1 : E \to K$ définie par :

$$u_1 \mapsto \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)u_{\sigma(1),1}u_{\sigma(2),2}\cdots u_{\sigma(n),n}$$

est une forme linéaire. Considérons donc $\phi_1(u_1 + \lambda u_1')$ avec u_1, u_1' dans E et λ dans K. Il vient :

$$\phi_1(u_1 + \lambda u_1') = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)(u_{\sigma(1),1} + \lambda u'_{\sigma(1),1})u_{\sigma(2),2}\cdots u_{\sigma(n),n}.$$

Il ne reste alors plus qu'à développer !
Montrons que $\det_{\mathcal{B}}$ est **alternée**.
Fixons $\tau \in \mathfrak{S}_n$; on a :

$$\det_{\mathcal{B}}(u_{\tau(1)}, \cdots, u_{\tau(n)}) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) u_{\tau\sigma(1),1} u_{\tau\sigma(2),2} \cdots u_{\tau\sigma(n),n}$$
$$= \sum_{\gamma \in \mathfrak{S}_n} \varepsilon(\tau^{-1}\gamma) u_{\gamma(1),1} u_{\gamma(2),2} \cdots u_{\gamma(n),n} = \varepsilon(\tau) \det_{\mathcal{B}}(u_1, \cdots, u_n).$$

Vérifions enfin que cette forme est non triviale.
On a :

$$\det_{\mathcal{B}}(e_1, \cdots, e_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) e_{\sigma(1),1} e_{\sigma(2),2} \cdots e_{\sigma(n),n}.$$

Le seul terme non nul de la somme est obtenu pour $\sigma = \mathrm{Id}$ et vaut 1 ! On a donc montré que :

$$\det_{\mathcal{B}}(e_1, \cdots, e_n) = 1.$$

Le fait que $\mathcal{A}_n(E)$ est une droite vectorielle suit alors de la propostion 10.7. $\square$

Proposition 10.10. *Soit $\mathcal{B}$ une base de E et $\phi \in \mathcal{A}_n(E)$. Alors, on dispose de l'identité :*

$$\phi = \phi(\mathcal{B}) \det_{\mathcal{B}}.$$

10.3. **Propriétés du déterminant, applications.**

10.3.1. *Propriétés élémentaires du déterminant.* Soit $\mathcal{B}$ une base de E. Soient $(u_1, \cdots u_n) \in E^n$, $v \in E$ et $\lambda \in K$. Alors, nous avons montré :

i) $\det_{\mathcal{B}} \mathcal{B} = 1$,

ii) $\forall i \in E_n, \quad \det_{\mathcal{B}}(u_1, \cdots, u_i + \lambda v, \cdots, u_n) = \det_{\mathcal{B}}(u_1, \cdots, u_i, \cdots, u_n) + \lambda \det_{\mathcal{B}}(u_1, \cdots, v, \cdots, u_n)$,

iii) Pour tout $\tau \in \mathfrak{S}_n$, $\det_{\mathcal{B}}(u_{\tau(1)}, \cdots, u_{\tau(n)}) = \varepsilon(\tau) \det_{\mathcal{B}}(u_1, \cdots, u_n)$.

10.4. **Déterminant et liberté.**

La proposition suivante explique comment on passe du déterminant dans une base au déterminant dans une autre base.

Proposition 10.11. *Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E. On dispose de la formule suivante :*

$$\det_{\mathcal{B}'} = \det_{\mathcal{B}'}(\mathcal{B}) \det_{\mathcal{B}}.$$

Corollaire 10.12. *Soit $\mathcal{B}$ une base de E.*

Une famille $(u_1, \cdots, u_n)$ est une base de E si et seulement si $\det_{\mathcal{B}}(u_1, \cdots, u_n) \neq 0$.

10.4.1. *Déterminant d'un endomorphisme et d'une matrice carrée.*

Déterminant d'un endomorphisme. Soit $u \in \mathcal{L}(E)$. Nous allons donner un sens à $\det u$ indépendamment d'une quelconque base. Considérons l'application :

$$l_u : \left\{ \begin{array}{ccc} \mathcal{A}_n(E) & \to & \mathcal{A}_n(E) \\ \phi & \mapsto & \phi_u \end{array} \right. ,$$

où ϕ_u est la forme n-linéaire et alternée définie par :

$$\phi_u(v_1, \cdots, v_n) = \phi(u(v_1), \cdots, u(v_n)).$$

Lemme 10.13. *Avec les notations précédentes, l_u est une application linéaire.*

Proposition 10.14. *Pour tout $u \in \mathcal{L}(E)$, il existe un unique élément de K noté $\det u$ tel que pour tout $(v_1, \cdots, v_n) \in E^n$ et pour tout $\phi \in \mathcal{A}_n(E)$:*

$$\phi(u(v_1), \cdots, u(v_n)) = \det u \, \phi(v_1, \cdots, v_n).$$

Démonstration. Comme l_u est un endomorphisme de $\mathcal{A}_n(E)$ qui est de dimension 1, on en déduit que l_u est une homothétie. $\det u$ est alors ce rapport d'homothétie. □

Définition 10.15 (Déterminant d'un endomorphisme)**.** *Soit $u \in \mathcal{L}(E)$.*

*$\det u$ est appelé **déterminant de** u.*

Proposition 10.16. *Soit $u \in \mathcal{L}(E)$ et $\lambda \in K$. On dispose des propriétés suivantes :*

i) $\det \mathrm{Id}_E = 1$,

ii) $\det(\lambda u) = \lambda^n \det u$.

Proposition 10.17 (Multiplicativité du déterminant)**.** *Soient u et v dans $\mathcal{L}(E)$. Alors :*

$$\det u \circ v = \det u \det v.$$

Proposition 10.18 (Déterminant d'un endomorphisme et base)**.** *Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E et $u \in \mathcal{L}(E)$. Alors,*

$$\det_{\mathcal{B}} u(\mathcal{B}) = \det_{\mathcal{B}'} u(\mathcal{B}') = \det u.$$

Cette proposition signifie que pour calculer le déterminant d'un endomorphisme, il suffit de prendre n'importe quelle base et de calculer le déterminant des images par u des vecteurs de cette base dans cette base.

Corollaire 10.19. *Soit $u \in \mathcal{L}(E)$.*

Alors $u \in GL(E)$ si et seulement si $\det u \neq 0$, auquel cas $\det u^{-1} = \frac{1}{\det u}$.

Définition 10.20 (Groupe spécial linéaire)**.** *On pose*

$$SL(E) = \{u \in GL(E) : \det u = 1\}.$$

Proposition 10.21. *$(SL(E), \circ)$ est un sous-groupe de $(GL(E), \circ)$ et il est appelé **groupe spécial linéaire** de E.*

Déterminant d'une matrice carrée, méthodes de calcul.

Définitions et propriétés.

Définition 10.22 (Déterminant d'une matrice)**.** *Étant donnée une matrice* $U = (u_{i,j}) \in M_n(K)$ *et considérant la* ***base canonique*** $\mathcal{B} = (e_i)$ *de* $E = K^n$*, on introduit les vecteurs de* K^n *définis pour chaque* $1 \leqslant j \leqslant n$ *:*

$$C_j = \sum_{i=1}^{n} u_{i,j} e_i.$$

Les (C_j) *sont les "colonnes" de la matrice* U*. Le déterminant de la matrice* U*, noté* $\det U$*, est le déterminant de* $(C_1, \cdots C_n)$ *dans la base canonique de* K^n*.*

Proposition 10.23. *Soit* $\mathcal{B}$ *une base de* E*. Soit* $(x_1, \cdots x_n) \in E^n$*. Alors :*

$$\det \operatorname{Mat}_{\mathcal{B}}(x_1, \cdots, x_n) = \det_{\mathcal{B}}(x_1, \cdots, x_n).$$

Proposition 10.24. *Soit* $\mathcal{B}$ *une base de* E *et* $u \in \mathcal{L}(E)$*. On pose* $U = \operatorname{Mat}_{\mathcal{B}} u$*. Alors :*

$$\det U = \det u.$$

On en déduit :

Corollaire 10.25. *Soit* $U \in M_n(K)$ *et* $\lambda \in K$*. On dispose des propriétés suivantes :*

i) $\det I_n = 1,$

ii) $\det(\lambda U) = \lambda^n \det U.$

Corollaire 10.26. *Si* U *et* V *sont deux matrices de* $M_n(K)$*, alors*

$$\det(UV) = \det U \det V.$$

Corollaire 10.27. *Soit* $U \in M_n(K)$*.*

Alors $U \in GL(n, K)$ *si et seulement si* $\det U \neq 0$*, auquel cas* $\det U^{-1} = \frac{1}{\det U}$*.*

Cela permet de montrer que si $U \in M_n(K)$ et $P \in GL_n(K)$, alors

$$\det(PUP^{-1}) = \det U.$$

Comment en donner une autre preuve ?

Proposition 10.28. *Soit* $U = (u_{i,j}) \in M_n(K)$*. Considérons* ${}^tU$ *sa transposée, i.e la matrice de terme général* $v_{i,j} = u_{j,i}$*. Alors*

$$\det U = \det {}^tU.$$

Exercice 1. *Soit* $\mathcal{B}$ *une base de* E *et* $u \in \mathcal{L}(E)$*. On note* $M = \operatorname{Mat}_{\mathcal{B}} u$*. Notant* $\mathcal{B}^*$ *la base duale de* $\mathcal{B}$*, montrer que :*

$${}^tM = \operatorname{Mat}_{\mathcal{B}^*} {}^tu.$$

En déduire que $\det u = \det {}^tu$*.*

Définition 10.29. *On pose :*

$$SL_n(K) = \{M \in GL_n(K) : \det M = 1\}.$$

$SL_n(K)$ *est appelé* ***groupe spécial linéaire*** *d'ordre* n *sur* K*.*

Introduisons une notation commode. Le déterminant d'une matrice $U = (u_{i,j})$ sera noté :

$$\det U = \begin{vmatrix} u_{1,1} & u_{1,2} & \cdots & u_{1,n} \\ u_{2,1} & u_{2,2} & \cdots & u_{2,n} \\ \vdots & \cdots & \cdots & \vdots \\ u_{n,1} & u_{n,2} & \cdots & u_{n,n} \end{vmatrix}.$$

Calcul pratique : la méthode de Gauss.
Nous allons voir que l'algorithme de Gauss permet de se ramener à des calculs de déterminants plus simples.

Lemme 10.30. *Soit $U \in M_n(K) = (u_{i,j})$. On pose $u_{1,1} = a$ et on suppose que $u_{i,1} = 0$ pour tout $i \geqslant 2$. Alors,*

- *soit $a = 0$ et $\det U = 0$,*
- *soit $a \neq 0$ et $\det U = a \det V$ où $V \in M_{n-1}(K) = (v_{i,j})$ est définie par : $v_{i,j} = u_{i+1,j+1}$ pour tout $1 \leqslant i \leqslant n$ et $1 \leqslant j \leqslant n$.*

Démonstration. Si $a = 0$, il s'agit de se souvenir que le déterminant d'un système lié est nul, en particulier un système qui contient 0.
Supposons donc que a n'est pas nul. La linéarité par rapport au premier vecteur colonne donne :

$$\det U = a \begin{vmatrix} 1 & u_{1,2} & \cdots & u_{1,n} \\ 0 & u_{2,2} & \cdots & u_{2,n} \\ \vdots & \cdots & \cdots & \vdots \\ 0 & u_{n,2} & \cdots & u_{n,n} \end{vmatrix}$$

Nous voulons calculer le déterminant de la matrice :

$$W = \begin{bmatrix} 1 & u_{1,2} & \cdots & u_{1,n} \\ 0 & u_{2,2} & \cdots & u_{2,n} \\ \vdots & \cdots & \cdots & \vdots \\ 0 & u_{n,2} & \cdots & u_{n,n} \end{bmatrix}.$$

Nous revenons maintenant à la définition du déterminant :

$$\det W = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) w_{\sigma(1),1} w_{\sigma(2),2} \cdots w_{\sigma(n),n}.$$

Examinons chaque terme de la somme.
Si $\sigma(1) \neq 1$, alors $w_{\sigma(1),1} = 0$.
Sinon, $\sigma(1) = 1$ et $w_{1,1} = 1$.
On est donc amené à réécrire la somme sous la forme :

$$\det W = \sum_{\sigma \in \mathfrak{S}_n : \sigma(1) = 1} \varepsilon(\sigma) w_{\sigma(2),2} \cdots w_{\sigma(n),n}.$$

Notons

$$\mathrm{Fix}(1) = \{\sigma \in \mathfrak{S}_n : \sigma(1) = 1\}.$$

Il est clair que l'application de Fix(1) vers $\mathfrak{S}(\{2, \cdots, n\})$ définie par :

$$\sigma \mapsto \sigma_{|\{2,\cdots,n\}} = \tilde{\sigma}$$

est bijective. La somme peut donc être réécrite :

$$\det W = \sum_{\tilde{\sigma} \in \mathfrak{S}(\{2,\cdots,n\})} \varepsilon(\tilde{\sigma}) w_{\tilde{\sigma}(2),2} \cdots w_{\tilde{\sigma}(n),n},$$

car $\varepsilon(\sigma) = \varepsilon(\tilde{\sigma})$. Ainsi, on reconnaît $\det V$. □

Lemme 10.31. *Donnons nous une matrice $U = (u_{i,j})$ triangulaire supérieure, i.e telle que $u_{i,j} = 0$ pour $j > i$. Alors*

$$\det U = \prod_{i=1}^{n} u_{i,i}.$$

Exemples
Calculer les déterminants suivants par la méthode de Gauss :

$$\Delta_1 = \begin{vmatrix} 1 & 0 & -1 & 2 \\ 0 & 3 & 0 & 1 \\ -2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{vmatrix},$$

$$\Delta_2 = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ -2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 3 \end{vmatrix},$$

$$\Delta_3 = \begin{vmatrix} 3 & 0 & 1 & 3 \\ 1 & -1 & 0 & 1 \\ -2 & -1 & 0 & 0 \\ 1 & 1 & 2 & 1 \end{vmatrix}.$$

Qu'en conclut-on sur l'inversibilité des matrices correspondantes ?

Développement par rapport à une ligne et une colonne.
Ce qui suit va exploiter la propriété de n-linéarité du déterminant. Tout ce que nous allons dire sur les colonnes est valable aussi pour les lignes, c'est pourquoi, nous ne nous occupons que de développement par rapport aux colonnes.

Lemme 10.32 (Développement par rapport à la j-ème colonne)**.** *Soit $A = (a_{i,j}) \in M_n(K)$. On écrit $A = (C_1, \cdots, C_n)$ où les C_j sont les colonnes de A. On fixe j un indice de colonne. Alors*

$$\det A = \sum_{i=1}^{n} a_{i,j} \det(C_1, \cdots, \underset{j}{e_i}, \cdots, C_n).$$

*On note $A_{i,j} = \det(C_1, \cdots, \underset{j}{e_i}, \cdots, C_n)$; ce nombre est appelé **cofacteur** de $a_{i,j}$.*

Définition 10.33 (Mineurs)**.** *Soit $A = (a_{i,j}) \in M_n(K)$. On note toujours $A_{i,j}$ le cofacteur de $a_{i,j}$. On considère de plus la sous-matrice de A obtenue en supprimant la i-ème ligne et la j-ème colonne notée $M_{i,j} \in M_{n-1}(K)$. Son déterminant, noté $m_{i,j}$ est appelé **mineur** associé à $a_{i,j}$.*

Exemples
Pour une matrice d'ordre 3, un développement suivant la première colonne mène à

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - d \begin{vmatrix} b & c \\ h & i \end{vmatrix} + g \begin{vmatrix} b & c \\ e & f \end{vmatrix}.$$

Il faut toujours développer selon la colonne ou la ligne qui comporte le plus de zéros !

$$\begin{vmatrix} 2 & 4 & 5 \\ 1 & 7 & -3 \\ -1 & -2 & -6 \end{vmatrix} \overset{L_1 \leftarrow L_1+2L_3}{=} \begin{vmatrix} 0 & 0 & -7 \\ 1 & 7 & -3 \\ -1 & -2 & -6 \end{vmatrix} = -7 \begin{vmatrix} 1 & 7 \\ -1 & -2 \end{vmatrix} = -7(-2+7) = -35.$$

Lemme 10.34 (Relations entre mineurs et cofacteurs)**.** *Avec les notations précédentes, on a :*

$$A_{i,j} = (-1)^{i+j} m_{i,j}.$$

Exemples
$M = \begin{pmatrix} \underset{+}{a} & \underset{-}{b} \\ \underset{-}{c} & \underset{+}{d} \end{pmatrix}$, il vient $A_{11} = d$, $A_{12} = -c$, $A_{21} = -b$ et $A_{22} = a$.

Pour une matrice d'ordre 3, on a le schéma : $M = \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$

Définition 10.35 (Comatrice)**.** *Soit* $A = (a_{i,j}) \in M_n(K)$*. On appelle* ***comatrice*** *de* A *la matrice des cofacteurs* $(A_{i,j})$*. Elle est notée* $\mathrm{com}(A)$.

Proposition 10.36 (Formule de la comatrice)**.** *Soit* $A = (a_{i,j}) \in M_n(K)$*. Avec les notations introduites, on a :*

$$\tilde{A}A = A\tilde{A} = (\det A)\,\mathrm{I}_n,$$

o ‘u $\tilde{A} = {}^t\mathrm{com}(A)$.

Corollaire 10.37. *En particulier, quand* A *est inversible, son inverse est donnée par*

$$A^{-1} = \frac{1}{\det A}\,{}^t\mathrm{com}(A).$$

On s'intéresse enfin à la résolution de $AX = b$ avec A une matrice inversible.

Proposition 10.38 (Formules de Cramer)**.** *Soient* $A = (a_{i,j}) \in GL_n(K)$ *et* $b \in K^n$*. Alors, l'équation*

$$AX = b$$

a une unique solution $X = (X_1, \cdots, X_n)$ *qui est donnée par :*

$$X_i = \frac{1}{\det A} \det(C_1, \cdots, \underset{i}{b}, \cdots, C_n),$$

pour tout $1 \leqslant i \leqslant n$.

Ces formules sont utilisables en dimension 2, voire 3, mais deviennent largement fastidieuses en dimension supérieure ! La méthode la plus fiable et la plus rapide reste le pivot de Gauss !

Exemples

- Le système $\left\{\begin{array}{l} ax+by=e \\ cx+dy=f \end{array}\right.$ est de Cramer si et seulement si $\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$, auquel cas la solution est donnée par les formules
$$x = \frac{\begin{vmatrix} e & b \\ f & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} \text{ et } y = \frac{\begin{vmatrix} a & e \\ c & f \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}.$$
- Résolution du système $\left\{\begin{array}{l} x+y+z=1 \\ 2x+y-z=0 \\ x+3y=3 \end{array}\right.$.
Le déterminant du système vaut $\begin{vmatrix} 1 & 1 & 1 \\ 2 & 1 & -1 \\ 1 & 3 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 3 & 2 & 0 \\ 1 & 3 & 0 \end{vmatrix} = \begin{vmatrix} 3 & 2 \\ 1 & 3 \end{vmatrix} = 9-2=7$. On obtient $x = \dfrac{\begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 3 & 3 & 0 \end{vmatrix}}{7} = -\dfrac{3}{7}$, $y = \dfrac{\begin{vmatrix} 1 & 1 & 1 \\ 2 & 0 & -1 \\ 1 & 3 & 0 \end{vmatrix}}{7} = \dfrac{8}{7}$ et $z = \dfrac{\begin{vmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 3 & 3 \end{vmatrix}}{7} = \dfrac{2}{7}$.

11. Diagonalisation

On suppose que les espaces vectoriels sont de *dimension finie* dans ce qui suit.

Définition 11.1. *Une matrice* $A := [a_{ij}]$ *est* diagonale *si* $a_{ij} = 0$ *pour* $j \neq i$.

Définition 11.2. *Soit* f *un endomorphisme d'un espace vectoriel* E. *Si*

$$f(u) = \lambda u \quad \text{avec} \quad \lambda \in K \quad \text{et} \quad \in E \setminus \{0\}$$

on dit que λ *est une* valeur propre *pour* f *et que* u *est un* vecteur propre *pour* f.

Exemples :

i) L'application nulle n'a qu'une valeur propre : 0.
ii) L'identité n'a qu'une valeur propre : 1.
iii) Une projection a (au plus) deux valeurs propre : 0 et 1.
iv) Une symétrie a (au plus) deux valeur propres : 1 et -1.
v) L'homothétie de rapport λ a une seule valeur propre : λ.
vi) La rotation d'angle θ n'a pas de valeur propre réelle sauf si θ est l'angle plat (mais une valeur propre complexe $e^{i\theta}$).
vii) l'application linéaire

$$\begin{array}{ccc} \mathbb{R}^2 & \xrightarrow{f} & \mathbb{R}^2 \\ (x, y) & \longmapsto & (x + 2y, 2x + y) \end{array}$$

satisfait $f(1, 1) = (3, 3) = 3(1, 1)$ et $f(1, -1) = (-1, 1) = -(1, -1)$ si bien que 3 et -1 sont des valeurs propres.

Proposition 11.3. *Soit* f *un endomorphisme d'un espace vectoriel* E *et* $\mathcal{B}$ *une base de* E. *Alors,* $[f]_{\mathcal{B}}$ *est diagonale si et seulement si* $\mathcal{B}$ *est formée de vecteurs propres.*

Définition 11.4. *On dit alors que* f *est* diagonalisable.

Proposition 11.5. *Soit* f *un endomorphisme d'un espace vectoriel* E, $\lambda \in K$ *et*

$$E_\lambda := \ker(\lambda \mathrm{Id}_E - f).$$

Alors, $E_\lambda \neq \{0\}$ *si et seulement si* λ *est une valeur propre de* f *et alors, les vecteurs propres associés à* λ *sont les vecteurs non nuls de* E_λ.

Démonstration. En effet, dire que $u \in E_\lambda$ signifie que $(\lambda \mathrm{Id}_E - f)(u) = 0$ ou encore que $\lambda u = f(u)$. □

Définition 11.6. *Si* λ *est une valeur propre pour* f, *on dit que* E_λ *est un* sous-espace propre *pour* f.

Par exemple, si p est la projection sur F de direction G, le sous-espace propre associé à 0 est G et celui associé à 1 est F. Les sous-espaces propres de l'application $(x, y) \mapsto (x + 2y, 2x + y)$ sont les droites d'équations $x - y = 0$ et $x + y = 0$.

Proposition 11.7. *Soit* f *un endomorphisme d'un espace vectoriel* E. *Alors* f *est bijectif si et seulement si* 0 *n'est pas valeur propre pour* f. *Dans, ce cas, si* λ *est une valeur propre pour* f, *alors* λ^{-1} *est une valeur propre pour* f^{-1} *avec même sous-espace propre.*

Démonstration. Dire que 0 est valeur propre pour f signifie qu'il existe $u \neq 0$ tel que $f(u) = 0$ c'est à dire que $\ker f = 0$, ce qui signifie que f n'est pas bijective puisque $\dim E < \infty$.

Si f est bijective et u est un vecteur propre de f pour une valeur propre λ, on a $f(u) = \lambda u$ si bien que $u = f^{-1}(\lambda u)$ et comme f^{-1} est linéaire, $u = \lambda f^{-1}(u)$ et donc finalement $\lambda^{-1}u = f^{-1}(u)$. □

Proposition 11.8. *Soit f un endomorphisme d'un espace vectoriel de dimension finie E et λ une valeur propre de f. Si $k \in \mathbb{N}$, alors λ^k est valeur propre de f^k. Cela est toujours valide pour $k < 0$ si f est bijective.*

Démonstration. On sait déjà que $1 = \lambda^0$ est valeur propre pour $\mathrm{Id}_E = f^0$. On montre ensuite l'assertion par récurrence sur $k > 0$. On suppose donc que λ est valeur propre de f et que $f^k(u) = \lambda^k u$. On en déduit que

$$f^{k+1}(u) = f^k(f(u)) = f^k(\lambda u) = \lambda f^k(u) = \lambda\lambda^k u = \lambda^{k+1}u.$$

Enfin, si f est bijective, on sait que λ^{-1} est valeur propre pour f^{-1} avec même sous-espace propre que λ. D'autre part, on a $\lambda^{-k} = (\lambda^{-1})^k$ et $f^{-k} = (f^{-1})^k$.

□

C'est comme ça qu'on détermine les valeurs propres d'une symétrie ($s^2 = \mathrm{Id}$) ou d'une projection ($p^2 = p$) par exemple.

Définition 11.9. *Si f est un endomorphisme d'un espace vectoriel E, alors*

$$P := \det(X\mathrm{Id}_E - f) \in K[X]$$

est le polynôme caractéristique *de f.*

Pour l'endomorphisme $(x, y) \mapsto (x + 2y, 2x + y)$, par exemple, on trouve

$$\begin{vmatrix} X-1 & -2 \\ -2 & X-1 \end{vmatrix} = (X-1)^2 - 4 = (X-3)(X+1).$$

Le célèbre *théorème de Cayley-Hamilton* dit que l'on a toujours $P(f) = 0$. Dans notre exemple, on aura donc $f^2 - 2f + 3\mathrm{Id} = 0$. Nous n'aurons pas le temps de discuter ce théorème.

Proposition 11.10. *Soient f un endomorphisme d'un espace vectoriel E et P son polynôme caractéristique. Alors, $\lambda \in K$ est valeur propre de f si et seulement si λ est racine de P.*

Démonstration. En effet, dire que λ est racine de P signifie que $\det(\lambda\mathrm{Id}_E - f) = 0$, ce qui signifie que $\lambda\mathrm{Id}_E - f$ n'est pas bijective ou encore que son noyau E_λ n'est pas nul. □

Théorème 11.11. *Soit f un endomorphisme d'un espace vectoriel de dimension finie E, $\lambda_1, \ldots, \lambda_k$ ses valeurs propres et $E_1, \ldots, E_k$ les sous-espaces propres. Alors, $\dim E_1 + \cdots + \dim E_k \leq \dim E$ avec égalité si et seulement si f est diagonalisable. C'est le cas en particulier si $\dim E = n$ et qu'il y a n valeurs propres distinctes.*

Lemme 11.12. *Si $F_i := E_1 + \cdots + E_i$, alors $F_{i+1} = F_i \oplus E_{i+1}$. De plus, F_i possède une base formée de vecteurs propres.*

Démonstration. Pour la première assertion, il suffit bien sûr de montrer que $F_i \cap E_{i+1} = 0$ et on va en fait montrer que $F_j \cap E_{i+1} = 0$ pour $j \leq i$. On procède par récurrence sur $i + j$. Si $u \in E_{i+1}$, on a par définition $f(u) = \lambda_{i+1} u$ mais si on a aussi $u \in F_j$, on peut écrire $u = u_1 + \cdots + u_j$ avec $u_1 \in E_1, \ldots, u_j \in E_j$ et on a donc aussi

$$f(u) = \lambda_1 u_1 + \cdots + \lambda_j u_j.$$

On en déduit que

$$((\lambda_{i+1} - \lambda_1)u_1 + \cdots + (\lambda_{i+1} - \lambda_{j-1})u_{j-1}) + (\lambda_{i+1} - \lambda_j)u_j = 0.$$

Par récurrence, on voit que les deux termes sont nuls et en particulier que $(\lambda_{i+1} - \lambda_j)u_j = 0$, ce qui montre que $u_j = 0$ et donc que $u \in F_{j-1}$ et on a fini (par récurrence).

La seconde assertion se démontre grace à la première par récurrence sur i puisqu'une base de F_{i+1} est composé d'une base de F_i et d'une base de E_{i+1}. □

Démonstration. Il résulte du lemme que $\dim E_1 + \cdots + \dim E_k = \dim F_k \leq \dim E$ avec égalité si et seulement si $F_k = E$. Et dans ce dernier cas, E aura une base formée de vecteurs propres. Réciproquement, si $\mathcal{B}$ est une base de E formée de vecteurs propres, alors $\mathcal{B} \subset \cup E_i \subset F_k$ et c'est donc aussi une base de F_k qui est donc nécessairement égal à E.

Bien sûr, si on a n valeurs propres distinctes, alors $k = n$ et comme chaque $E_i \neq 0$, on a $\dim F_n \geq n$ et donc $E = F_n$. □

Remarquons que la dernière condition n'est pas nécessaire car l'identité à une seule valeur propre, quelle que soit la dimension de E mais est pourtant diagonalisable, et même diagonale ! Le phénomène analogue se produit plus généralement avec les projections ou les symétries par exemple.

Corollaire 11.13. *Si le polynôme caractéristique de f a toutes ses racines dans K et que celles-ci sont distinctes, alors f est diagonalisable.*

Démonstration. On a vu que ce sont justement les valeurs propres de f. □

Remarquons que la première condition est toujours satisfaite si $K = \mathbf{C}$.

Définition 11.14. *Soit $A \in M_n(K)$ et f est l'endomorphisme de K^n qui lui est associé. Alors, le* polynôme carcactéristique *de A, les* valeurs propres *de A, les* sous-espaces propres *de A ainsi que les* vecteurs propres *de A sont ceux de f. On dit que A est* diagonalisable *si f l'est.*

Le polynôme caractéristique de A est $P = \det(XI_n - A) \in K[X]$. Les valeurs propres de A sont les racines λ de P dans K. Les vecteurs propres u associés à λ correspondent aux vecteurs colonnes U tels que $AU = \lambda U$. La matrice A est diagonalisable si elle est semblable à une matrice diagonale ($A = PDP^{-1}$ avec $P \in GL_n(K)$ et D diagonale).

Proposition 11.15. *Soit A la matrice d'un endomorphisme f d'un espace vectoriel E dans une base $\mathcal{B}$. Alors, le polynôme caractéristique de A est identique au polynôme caractéristique de f, les valeurs propres de A sont identiques aux valeurs propres de F et enfin, A est diagonalisable si et seulement si f est diagonalisable.*

Démonstration. La première assertion résulte du fait que le déterminant d'un endomorphisme ne dépend pas du choix de la base. La seconde en découle. Enfin, dire que f est diagonalisable signifie qu'il existe une base dans laquelle sa matrice est diagonale, c'est à dire, une matrice de passage P telle que $D := P^{-1}AP$ soit diagonale. $\square$

Proposition 11.16. *Si P est le polynôme caractéristique de $A \in M_n(K)$, alors*

$$P = X^n - tX^{n-1} + \cdots + (-1)^n\delta$$

ou $\delta = \det A$ et $t = \mathrm{tr}A$ est la somme des éléments diagonaux.

Démonstration. Clairement, on a $P(0) = \det(-A) = (-1)^n \det A$. Pour la trace, c'est un peu plus compliqué et il faut regarder ce qui apparaît dans le coefficient de X^{n-1} quand on calcule le déterminant. $\square$

Notons en particulier que si toutes les valeurs propres de A sont distinctes et dans K, alors $\det A$ est égal au produit de ces valeurs propres et que $\mathrm{tr}A$ est égal à la somme. C'est vrai plus généralement si on les prend avec leur multiplicité et dans un corps suffisamment grand.

Quelques exemples pour clore ce chapitre : on considère les matrices suivantes de $M_2(\mathbb{R})$:

$$A := \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad B := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad C := \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} et \quad \text{et} \quad D := \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

i) La matrice A a pour polynôme caractéristique $X^2 - 2X - 3$ et donc deux valeurs propres distinctes 3 et -1. Elle est diagonalisable.

ii) La matrice B a pour polynôme caractéristique $X^2 + 1$ et n'a donc aucune valeur propre. Elle n'est pas diagonalisable.

iii) La matrice C a pour polynôme caractéristique $X^2 - 2X + 1$ et a donc 1 pour unique valeur propre. Et le sous espace propre est l'axe des x ! Elle n'est donc pas diagonalisable.

iv) La matrice D a pour polynôme caractéristique $X^2 + 2X + 1$ et a donc -1 pour unique valeur propre. Mais cette fois ci, elles est diagonalisable (et même diagonale).

12. Produit scalaire

Définition 12.1. *La matrice d'une forme bilinéaire $\varphi : E \times E \to K$ dans une base $\mathcal{B} := (u_1, \ldots, u_n)$ de l'espace vectoriel E est*

$$[\varphi]_{\mathcal{B}} := [\varphi(u_i, u_j)].$$

Par exemple, l'application

$$\begin{aligned} \mathbb{R}^2 \times \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ ((a,b),(c,d)) &\longrightarrow ac + bd \end{aligned}$$

a pour matrice $I := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ dans la base canonique. Et le déterminant

$$\begin{aligned} \mathbb{R}^2 \times \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ ((a,b),(c,d)) &\longrightarrow ad - bc \end{aligned}$$

a pour matrice $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

Proposition 12.2. *Si φ est une forme bilinéaire sur E et $\mathcal{B}$ une base E, on a pour tout $u, v \in E$,*

$$\varphi(u,v) = {}^t[u]_{\mathcal{B}}[\varphi]_{\mathcal{B}}[v]_{\mathcal{B}}.$$

Démonstration. Pour alléger les notations, on omet le nom de la base dans les indices. On peut écrire

$$[u] = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} \quad \text{et} \quad [v] = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix}.$$

Par bilinéarité, il suit que dans la base $\mathcal{B} = (u_1, \ldots, u_n)$, on a

$$\varphi(u,v) = \varphi(\lambda_1 u_1 + \cdots + \lambda_n u_n, \mu_1 u_1 + \cdots + \mu_n u_n)$$

n'est autre que la somme de tous les $\lambda_i \mu_j \varphi(u_i, u_j)$ pour $i, j = 1, \ldots, n$. D'autre part, si on pose $b_{ij} := \varphi(u_i, u_j)$, on voit que ${}^t[u][\varphi]$ est le vecteur ligne dont la j-ème composante est $\lambda_1 b_{1j} + \cdots + \lambda_n b_{nj}$. Pour obtenir${}^t[u][\varphi]t[v]$, on multiplie par μ_j et on somme sur j. On trouve bien la somme de tous les $\lambda_i b_{ij} \mu_j$. □

Définition 12.3. *Soit E un espace vectoriel sur $\mathbb{R}$. Une forme bilinéaire*

$$(u,v) \longmapsto \langle u, v \rangle$$

sur E est symétrique *si pour tout $u, v \in E$, on a*

$$\langle v, u \rangle = \langle u, v \rangle.$$

Et bien sûr, linéaire par rapport à la seconde variable signifie que l'on a toujours

$$\langle u, \lambda_1 v_1 + \lambda_2 v_2 \rangle = \lambda_1 \langle u, v_1 \rangle + \lambda_2 \langle u, v_2 \rangle.$$

Par symétrie, elle sera alors automatiquement linéaire aussi par rapport à la première variable.

On voit que le premier exemple ci-dessus est symétrique mais pas le second. En fait, une application bilinéaire est symétrique si et seulement si sa matrice dans une base quelconque est symétrique.

Définition 12.4. *Une forme bilinéaire symétrique $\langle -,- \rangle$ sur un espace vectoriel E est* définie positive *ou encore, un* produit scalaire, *si on a la propriété*

$$\langle u,u \rangle > 0 \Leftrightarrow u \neq 0.$$

L'exemple ci-dessus est bien un produit scalaire car on a bien $a^2 + b^2 > 0$ si et seulement si $(a,b) \neq (0,0)$.

On admettra le théorème suivant qui est bien pratique :

Théorème 12.5. *Une forme bilinéaire symétrique est un produit scalaire si et seulement si les valeurs propres de sa matrice dans une base quelconque sont strictement positives.*

Démonstration. Admis. □

12.1. **Produit scalaire et inégalité de Cauchy-Schwarz.**

Définition 12.6. *Soit E un espace vectoriel sur $\mathbb{R}$. Une* norme *sur E est une application $\| - \| : E \to \mathbb{R}$ telle que*

i) *Si $u \in E$, alors $\|u\| > 0 \Leftrightarrow u \neq 0$*
ii) *Si $u, v \in E$ alors $\|u+v\| \leq \|u\| + \|v\|$*
iii) *si $\lambda \in \mathbb{R}$ et $u \in E$, alors $\|\lambda u\| = |\lambda| \|u\|$*

Sur $\mathbb{R}^2$, on a par exemple les normes $\|(x,y)\|_1 = |x| + |y|$ ou $\|(x,y)\|_2 = \sqrt{x^2+y^2}$ ou encore $\|(x,y)\|_\infty = \max(|x|,|y|)$ comme on peut aisément le vérifier.

Proposition 12.7. *Soit $\langle -,- \rangle$ un produit scalaire. Alors,*

i) *La fonction*

$$\| - \| : E \to \mathbb{R}, u \longmapsto \sqrt{\langle u,u \rangle}$$

est une norme sur E et on a toujours

$$\langle u,v \rangle = \frac{\|u+v\|^2 - \|u\|^2 - \|v\|^2}{2}$$

ii) *(Inégalité de Cauchy-Schwarz) On a toujours*

$$\langle u,v \rangle \leq \|u\| \|v\|.$$

Démonstration. On montre d'abord l'inégalité de Cauchy-Schwartz. On peut supposer $u \neq 0$. On regarde le polynôme

$$\|\lambda u + v\|^2 = \|u\|^2 |\lambda|^2 + 2\lambda \langle u,v \rangle + \|v\|^2.$$

Celui-ci est toujours positif (ou nul). Son discriminant est donc négatif (ou nul) et on a donc

$$\langle u,v \rangle^2 - \|u\|^2 \|v\|^2 \leq 0.$$

Il suit que

$$\langle u,v \rangle \leq \|u\| \|v\|.$$

On vérifie ensuite que l'on a bien une norme. La première et la dernière condition résultent directement des définitions. La seconde est conséquence de l'inégalité de Cauchy-Schwarz : on a d'une part

$$\|u+v\|^2 = \|u\|^2 + 2\langle u, v\rangle + \|v\|^2,$$

et d'autre part

$$(\|u\| + \|v\|)^2 = \|u\|^2 + 2\|u\|^2\|v\|^2 + \|v\|^2.$$

Enfin, pour la formule, il suffit de développer le membre de droite. □

L'inégalité de Cauchy-Schwarz est très importante car elle permet de définir l'*angle* entre deux vecteurs non nuls u et v comme l'unique $\theta \in [0, \pi]$ tel que

$$\cos\theta = \frac{\langle u, v\rangle}{\|u\|\|v\|}.$$

Définition 12.8. *Un* espace vectoriel euclidien *est un espace vectoriel de dimension finie sur* $\mathbb{R}$ *muni d'un produit scalaire* $\langle -, -\rangle$.

Sur $\mathbb{R}^n$, le produit scalaire standard est

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n)\rangle = x_1y_1 + \cdots + x_ny_n$$

et la norme associée est

$$\|(x_1, \ldots, x_n)\| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

12.2. **Orthogonalité.**

Définition 12.9. *Soit E un espace vectoriel euclidien. Deux vecteurs u et v sont* orthogonaux *si $\langle u, v\rangle = 0$ et on écrit alors $u \perp v$.*

Par exemple, dans $\mathbb{R}^2$ avec le produit scalaire habituel, on a donc $(1, 0)\perp(0, 1)$.

Définition 12.10. *Soit E un espace vectoriel euclidien.*

i) *Deux parties $\mathcal{S}$ et $\mathcal{T}$ de E sont* orthogonales *si pour tout $u \in \mathcal{S}$ et $v \in \mathcal{T}$, on a $u\perp v$. On écrira alors $\mathcal{S}\perp\mathcal{T}$*
ii) *La* partie orthogonale *à une partie $\mathcal{S}$ de E est*

$$\mathcal{S}^\perp := \{u \in E, \quad u\perp\mathcal{S}\}.$$

On voit donc que $\mathcal{S}\perp\mathcal{T}$ si et seulement si $\mathcal{T} \subset \mathcal{S}^\perp$.

Par exemple, dans $\mathbb{R}^2$, l'axe des x est l'orthogonal à l'axe des y.

Proposition 12.11. *Soit E un espace vectoriel euclidien. Alors,*

i) *Si $\mathcal{S}$ est une partie de E, on a $\mathcal{S} \subset \mathcal{S}^{\perp\perp}$.*
ii) *Si $\mathcal{S} \subset \mathcal{T} \subset E$, alors $\mathcal{T}^\perp \subset \mathcal{S}^\perp \subset E$.*

Démonstration. Si $u \in \mathcal{S}$ et $v \in \mathcal{S}^\perp$, alors $u\perp v$ et par symétrie $v\perp u$. Il suit que $u \in \mathcal{S}^{\perp\perp}$ et on a donc bien $\mathcal{S} \subset \mathcal{S}^{\perp\perp}$. Supposons maintenant que $\mathcal{S} \subset \mathcal{T}$ et donnons nous $u \in \mathcal{T}^\perp$ et $v \in S$. On a bien évidemment $v \in T$ et donc $u\perp v$, ce qui montre que $u \in \mathcal{S}^\perp$. □

Définition 12.12. *Un système $\mathcal{S}$ d'éléments non-nuls de E est* orthogonal *si pour tout $u \neq v$ dans $\mathcal{S}$, on a $u \perp v$. Il est* orthonormal *si de plus, pour tout $u \in \mathcal{S}$, on a $\|u\| = 1$.*

Par exemple, le système $((1,0),(0,1))$ est orthonormal dans $\mathbb{R}^2$ muni du produit scalaire standard contrairement à $((1,1),(1,-1))$ qui est cependant orthogonal.

Proposition 12.13 (Base duale et produit scalaire)**.** *Une base $\mathcal{B} := (u_1, \ldots, u_n)$ est orthonormale si et seulement si pour tout $i = 1, \ldots, n$, on a $\check{u}_i = \langle u_i, - \rangle$.*

Démonstration. Pour que deux applications linéaires coincident, il suffit qu'elles prennent les mêmes valeurs sur tous les vecteurs d'une base. Ici, cela se traduit par $\check{u}_i(u_j) = \langle u_i, u_j \rangle$ pour tout $i, j = 1, \ldots, n$, c'est à dire $\langle u_i, u_j \rangle = 0$ pour $i \neq j$ et $\langle u_i, u_i \rangle = 1$ pour tout i. Où encore, $u_i \perp u_j$ pour $i \neq j$ et $\|u_i\| = 1$ pour tout i. $\square$

Proposition 12.14 (Formes linéaires et produit scalaire)**.** *Soit E un espace vectoriel euclidien. Si $u \in E$, l'application*

$$\begin{aligned} E &\xrightarrow{\langle u, - \rangle} \mathbb{R} \\ v &\longmapsto \langle u, v \rangle \end{aligned}$$

est une forme linéaire. De plus, l'application

$$\begin{aligned} E &\longrightarrow \check{E} \\ u &\longmapsto \langle u, - \rangle \end{aligned}$$

est un isomorphisme d'espace vectoriel.

Démonstration. La première assertion exprime simplement la linéarité par rapport à la première variable. Et le fait que la seconde application est linéaire exprime la linéarité par rapport à l'autre variable. Enfin, comme les espaces ont même dimension, pour montrer qu'on a un isomorphisme, il suffit de vérifier que le noyau est trivial. Or si l'application $\langle u, - \rangle$ est nulle, on a en particulier $\|u\|^2 = \langle u, u \rangle = 0$ et donc $u = 0$. $\square$

Par exemple, dans $\mathbb{R}^n$ avec le produit scalaire standard, on voit que $\langle e_i, u \rangle = p_i(u)$ et donc que l'isomorphisme envoie la base canonique de $\mathbb{R}^n$ sur le base canonique de $\check{\mathbb{R}}^n$. En d'autres termes, elle transforme un vecteur colonne en vecteur ligne.

Proposition 12.15 (Procédé de Schmidt)**.** *Soit une base $(u_1, \ldots, u_m)$ de E et on pose*

$$v_1 := u_1, \quad v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\|v_1\|^2} v_1, \quad v_3 = u_3 - \frac{\langle u_3, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle u_3, v_2 \rangle}{\|v_2\|^2} v_2, \quad \ldots$$

et

$$w_1 := \frac{v_1}{\|v_1\|}, \quad w_2 := \frac{v_2}{\|v_2\|}, \quad \ldots.$$

La famille $(w_j)_{1 \leq j \leq n}$ est une base de E et, pour tout $k \in \{1, \cdots, n\}$, on a aussi $\mathrm{vect}(u_1, \cdots, u_k) = \mathrm{vect}(w_1, \cdots, w_k)$.

Par exemple, si on veut une base orthonormale du plan d'équation $x + y + z = 0$, on part de la base formée de $(1, -1, 0)$ et $(1, 0, -1)$ et on remplace le second vecteur par

$$(1,0,-1) - \frac{\langle (1,-1,0),(1,0,-1)\rangle}{\|(1,-1,0)\|^2}(1,-1,0) = (1,0,-1) - \frac{1}{2}(1,-1,0) = (\frac{1}{2},\frac{1}{2},-1)$$

pour obtenir une base orthogonale puis on normalise en divisant respectivement par

$$\|v_1\| = \sqrt{2} \quad \text{et} \quad \|v_1\| = \sqrt{\frac{3}{2}}$$

pour trouver

$$(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 0) \quad \text{et} \quad (\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{6}, -\frac{\sqrt{6}}{3}).$$

Proposition 12.16. *Soit $(u_1, \dots, u_n)$ une base orthogonale d'un espace affine euclidien. Alors,*

i) *Les coordonnées de $v \in E$ sont*

$$\frac{\langle v, u_1\rangle}{\|u_1\|^2}, \quad \dots, \quad \frac{\langle v, u_n\rangle}{\|u_n\|^2}.$$

ii) *Si*

$$v = \lambda_1 u_1 + \cdots + \lambda_n u_n$$

et

$$w = \mu_1 u_1 + \cdots + \mu_n u_n,$$

alors

$$\langle w, z\rangle = \lambda_1\mu_1\|u_1\|^2 + \cdots + \lambda_n\mu_n\|u_n\|^2.$$

Démonstration. Un rapide calcul nous donne la seconde formule et on en déduit la première, car alors

$$\langle v, u_i\rangle = \lambda_i\|u_i\|^2.$$

□

Proposition 12.17. *Soit E un espace vectoriel euclidien. Alors,*

i) *Si F est un sous-espace vectoriel de E, on a*

$$E = F \oplus F^\perp \text{ et } F^{\perp\perp} = F$$

et si $(u_1, \dots, u_m)$ une base orthonormale de F, alors, la projection orthogonale sur F est donnée par

$$u \mapsto \langle u, u_1\rangle u_1 + \cdots + \langle u, u_m\rangle u_m.$$

De plus, la distance entre x et F est $\|x_{F^\perp}\|$.

ii) *Si F et G sont deux sous-espace vectoriel de E, alors*

$$(F + G)^\perp = F^\perp \cap G^\perp$$

et

$$(F \cap G)^\perp = F^\perp + G^\perp.$$

Démonstration. Si F est un sous-espace vectoriel de E, il est clair que $F \cap F^{\perp} = 0$. De plus, comme on vient de le voir,

$$\dim E = \dim F^{\perp} + \dim F.$$

On a donc bien

$$E = F \oplus F^{\perp}.$$

On voit aussi que $\dim F^{\perp\perp} = \dim F$ et comme $F \subset F^{\perp\perp}$, on a nécessairement $F^{\perp\perp} = F$.

Montrons maintenant la seconde assertion. On a $F, G \subset F + G$ et donc $(F+G)^{\perp} \subset F^{\perp}, G^{\perp}$ par renversement d'inclusion. Si H est un sous-espace vectoriel de E tel que $H \subset F^{\perp}, G^{\perp}$, alors $F, G \subset H^{\perp}$ et donc $F + G \subset H^{\perp}$ si bien que $H \subset (F+G)^{\perp}$. $\square$

12.3. **Endomorphismes symétriques et diagonalisation.** Dans cette section, E est un espace euclidien.

Définition 12.18. *On dit que $f \in L(E)$ est un endomorphisme symétrique si et seulement si :*

$$\langle f(x), y\rangle = \langle x, f(y)\rangle, \quad \forall (x,y) \in E^2.$$

Proposition 12.19. *Soit $f \in L(E)$. f est symétrique si et seulement si, pour toute base orthonormée, la matrice de f est symétrique.*

Proposition 12.20. *Soit $f \in L(E)$ symétrique. Soit F un sous-espace vectoriel de E stable par f, c'est à dire tel que $f(F) \subset F$. Alors $f(F^{\perp}) \subset F^{\perp}$.*

Théorème 12.21. *Soit $f \in L(E)$ symétrique. Alors f est diagonalisable (sur $\mathbb{R}$) dans une base orthonormée.*

Démonstration. On remarque d'abord un fait général. Soit A la matrice de f dans une base orthonormée de E. A peut être considérée comme une matrice à coefficients dans $\mathbb{C}$. Elle admet donc une valeur propre $\lambda \in \mathbb{C}$ associée à un vecteur propre $X_0 \in \mathbb{C}^n : AX_0 = \lambda X_0$.

À présent, montrons que λ est réel. On observe que :

$${}^tYAX = {}^tXAY, \quad \forall X, Y \in \mathbb{C}^n.$$

On trouve alors ${}^t\overline{X_0}AX_0 = \lambda \|X_0\|^2$. Comme A est symétrique, le membre de droite est réel et λ est donc réel. Quitte alors à prendre la partie réelle de X_0 ou sa partie imaginaire, on trouve un vecteur propre réel. Il existe donc $x_0 \in E \setminus \{0\}$ et $\lambda \in \mathbb{R}$ tels que :

$$f(x_0) = \lambda x_0.$$

On pose $F = \mathrm{vect} x_0$. F est stable par f et il en est de même pour son orthogonal $F^{\perp}$. Ainsi, on peut considérer l'endomorphisme de $F^{\perp}$ induit par f et noté $f_{|F^{\perp}}$ et qui est encore symétrique. Le résultat s'ensuit par récurence.

$\square$

13. Quelques exercices

Exercice 2. *Soient F et G deux sous-espaces vectoriels d'un espace vectoriel E. Montrer que si $F \cup G$ est un sous-espace vectoriel de E, alors $F \subset G$ ou $G \subset F$.*

Exercice 3. *Soient F, G et H trois sous-espaces vectoriels d'un espace vectoriel E avec $F \subset G$. Montrer que*

$$F + (H \cap G) = (F + H) \cap G.$$

Exercice 4. *Soient F, G et H trois sous-espaces vectoriels d'un espace vectoriel E. A-t-on*

$$(F + H) \cap G = (F \cap G) + (H \cap G)?$$

Exercice 5. *Soit $f : E \to F$ une application linéaire. Montrer que l'application $G \to f(G)$ est une bijection de l'ensemble des sous-espaces vectoriels de E qui contiennent* $\ker f$ *sur l'ensemble des sous-espaces vectoriels de* $\mathrm{Im} f$.

Exercice 6. *On rappelle qu'un endomorphisme p de E est une* projection *si $p^2 := p \circ p = p$. Montrer que p est un projecteur si et seulement si $q := \mathrm{Id}_E - p$ en est un. Montrer qu'alors,* $\ker p = \mathrm{Im} q$, $\mathrm{Im} p = \ker q$ *et* $E = \ker p \oplus \mathrm{Im} p$.

Exercice 7. *Soient p et q deux projecteurs de E. Montrer que* $\mathrm{Im} p \subset \mathrm{Im} q$ *si et seulement si $q \circ p = p$.*

Exercice 8. *($1 + 1 \neq 0$) On rappelle qu'un endomorphisme s de E est une* symétrie *si* $s^2 = \mathrm{Id}_E$. *Montre que s est une symétrie si et seulement si $p = \frac{s+\mathrm{Id}_E}{2}$ est un projecteur. En déduire qu'alors*

$$E = \ker(s - Id_E) \oplus \ker(s + Id_E).$$

Exercice 9. *Montrer qu'une homothétie est une application linéaire distincte de l'identité qui laisse les droites invariantes.*

Exercice 10. *Montrer que les sous-espaces vectoriels de $\mathbb{R}^4$ engendrés par $u := (2, 3, -1, 0)$ et $v := (-3, 1, 0, 2)$, d'une part et $u' := (-5, 9, -2, 6)$ et $v' := (5, 2, -1, -2)$, d'autre part, sont identiques*

Exercice 11. *Soit E l'espace vectoriel des applications $[0, \pi] \to \mathbb{R}$. Montrer que*

$$\mathcal{S} := (1, \sin x, \sin 2x, \sin 3x, \ldots)$$

est un système libre de E.

Exercice 12. *Montrer que l'application*

$$M_n(\mathbb{R}) \to M_n(\mathbb{R}), A \mapsto {}^tA$$

est une symétrie vectorielle.

Exercice 13. *On rappelle qu'une matrice* symétrique *(resp.* antisymétrique*) est une matrice A telle que*

$${}^tA = A \text{ (resp. } {}^tA = -A).$$

Leur ensemble se note $SM_n(\mathbb{R})$ (resp. $AM_n(\mathbb{R})$). Montrer que $SM_n(\mathbb{R})$ et $AM_n(\mathbb{R})$ sont des sous-espaces vectoriels de $M_n(\mathbb{R})$ et que

$$M_n(\mathbb{R}) = SM_n(\mathbb{R}) \oplus AM_n(\mathbb{R}).$$

Exercice 14. *On note C (resp. C_k) l'ensemble des carrés magiques 3-3 (resp. de trace k), c'est à dire, les $A \in M_3(\mathbb{R})$ telles que les sommes des éléments des lignes, des colonnes et des diagonales soient toutes égales (resp. égales à k).*

i) *Montrer que C_k est non-vide et que si $M, N \in C_k$, alors $M - N \in C_0$.*

ii) *Montrer que*
$$C_0 = SC_0 \oplus AC_0$$
avec
$$SC_0 = C_0 \cap SM_3(\mathbb{R}) \text{ et } AC_0 = C_0 \cap AM_3(\mathbb{R}).$$

iii) *Déterminer une base de SC_0 puis une base de AC_0. En déduire une base de C_0 puis une base de C.*

iv) *Trouver un carré magique de trace 27 dont toutes les entrées sont distinctes.*

Exercice 15. *Soit E l'espace des polynômes de degré au plus 3 sur $\mathbb{R}$. Soient f_1, f_2, f_3 et f_4 les formes linéaires qui envoient P sur $P(0), P(1), P'(0)$ et $P'(1)$ respectivement. Montrer que c'est une base de $\check{E}$. Déterminer la base de E dont c'est la base duale.*

Exercice 16. *Soit $A \in M_n(\mathbb{R})$ telle que $A^2+A+I = 0$. Montrer que A est inversible.*

Exercice 17. *Soit $N \in M_n(\mathbb{R})$ telle que $N^r = 0$. Montrer que $A := I + N$ est inversible.*

Exercice 18. *Calculer A^n pour*

$$A := \begin{bmatrix} 1 & -1 \\ 1 & -3 \end{bmatrix}.$$

(On pourra chercher une relation linéaire entre les premières puissances de A, puis chercher le reste dans la division de X^n par le polynôme correspondant)

Exercice 19. *Même question avec*
$$A := \begin{bmatrix} -15 & 10 & 8 \\ -8 & 6 & 4 \\ -24 & 15 & 13 \end{bmatrix}.$$

Exercice 20. *Résoudre le système*

$$\begin{cases} u_{n+1} & = & u_n - v_n \\ v_{n+1} & = & u_n + 3v_n \end{cases}$$

avec conditions initiales $u_0 = 1$ et $v_0 = 2$.

Exercice 21. *Soit $A \in M_n(\mathbb{R})$ telle que $A^2 + A + I = 0$. Calculer $\det A$. Même question avec $A^2 - A + I = 0$.*

Exercice 22. *Montrer que si n est impair, il n'existe pas de $A \in M_n(\mathbb{R})$ telle que $A^2 + I = 0$. Même question avec $A^2 - \sqrt{2}A + I = 0$.*

Exercice 23. *Montrer que si $A \in M_n(\mathbb{C})$, alors $\det \bar{A} = \overline{\det A}$. En déduire que si $A, B \in M_n(\mathbb{R})$ satisfont $AB = BA$, alors $\det(A^2 + B^2) \geq 0$.*

Exercice 24. *Montrer que si n est impair et $A \in M_n(\mathbb{R})$ antisymétrique, alors* $\det A = 0$.

Exercice 25. *Calculer pour tout $n \geq 2$*

$$\Delta_n := \begin{vmatrix} 0 & \cdots & \cdots & 0 & a_n \\ x & \ddots & & \vdots & a_{n-1} \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & x & a_1 \end{vmatrix}$$

puis

$$\Gamma_n := \begin{vmatrix} x & 0 & \cdots & 0 & a_n \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & x & a_1 \\ a_n & \cdots & \cdots & a_1 & x \end{vmatrix}.$$

Exercice 26. *Calculer*

$$\Delta_n := \begin{vmatrix} 1+x^2 & -x & 0 & \cdots & 0 \\ -x & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -x \\ 0 & \cdots & 0 & -x & 1+x^2 \end{vmatrix}$$

Exercice 27. *Montrer que si $P \in \mathbb{R}[X]$ est tel que* $\deg P < n$, *et*

$$A = \begin{bmatrix} P(x) & \cdots & P(x+n) \\ \vdots & & \vdots \\ P(x+n) & \cdots & P(x+n+m) \end{bmatrix},$$

alors $\det A = 0$. *Calculer* $\det B$ *avec*

$$B = \begin{bmatrix} 0 & \cdots & n^2 \\ \vdots & & \vdots \\ n^2 & \cdots & (n+m)^2 \end{bmatrix}.$$

Exercice 28. *Soient $\lambda_1, \ldots, \lambda_n$ et $a \neq b$ fixés. Calculer*

$$\Delta := \begin{vmatrix} \lambda_1 & b & \cdots & b \\ a & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ a & \cdots & a & \lambda_n \end{vmatrix}.$$

On pourra montrer que

$$\Delta(X) := \begin{vmatrix} \lambda_1 + X & b + X & \cdots & b + X \\ a + X & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & b + X \\ a + X & \cdots & a + X & \lambda_n + X \end{vmatrix}$$

est un polynôme de degré au plus 1*, calculer* $\Delta(-a)$ *et* $\Delta(-b)$ *et en déduire* $\Delta = \Delta(0)$*.*

Exercice 29. *On considère, pour* $a \in \mathbb{R}$*, l'application linéaire*

$$f_a : \mathbb{R}[X]_{\leq 3} \to \mathbb{R}[X]_{\leq 3}, P \mapsto P(X + a)$$

ainsi que sa matrice M_a *dans la base canonique. Montrer que* M_a *est inversible et calculer* M_a^n *pour* $n \in \S$*. On pourra d'abord remarquer que pour* $a, b \in \mathbb{R}$*, on a* $f_{a+b} = f_a \circ f_b$*.*

Exercice 30. *Montrer que* f *est un projecteur si et seulement si* f *est diagonalisable et ses valeurs propres* $\in \{0, 1\}$*.*

Exercice 31. *Calculer le polynôme caractéristique de*

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

Cette matrice est-elle diagonalisable ?