

# 4

# Compléments d'algèbre linéaire

Hermann Günther Grassmann (1809 – 1877)

Nous continuons cette année avec des compléments en algèbre linéaire. Le but de ce chapitre est de comprendre les interactions entre sommes directes et endomorphismes, et les conséquences sur leurs écritures matricielles. Ce sera le fondement des techniques de réduction que nous verrons plus tard dans l'année. Nous y ajoutons l'interpolation de Lagrange, qui donne une seconde base naturelle de $\mathbb{K}_n[X]$ et qui reviendra dans l'étude des polynômes annulateurs.

**Programme officiel**

**Algèbre linéaire**

A – Compléments sur les espaces vectoriels, les endomorphismes et les matrices

a) Produit d'espaces vectoriels, somme de sous-espaces vectoriels

b) Matrices par blocs et sous-espaces stables

c) Trace

e) Interpolation de Lagrange



# 4 Compléments d'algèbre linéaire

## I. Produits et sommes d'espaces vectoriels

**Définition 4.1 (Produit d'espaces vectoriels)**

Soient $F_1, \dots, F_n$ des $\mathbb{K}$-espaces vectoriels. Le *produit* $F_1 \times \cdots \times F_n$ est l'ensemble

$$F_1 \times \cdots \times F_n = \{(x_1, \dots, x_n) \mid \forall i \in \{1, \dots, n\},\ x_i \in F_i\},$$

muni des lois composante par composante :

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n), \qquad \lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n).$$

C'est un $\mathbb{K}$-espace vectoriel.

**Proposition 4.2 (Dimension d'un produit)**

Si $F_1, \dots, F_n$ sont de dimension finie, alors $F_1 \times \cdots \times F_n$ est de dimension finie et

$$\dim(F_1 \times \cdots \times F_n) = \dim F_1 + \cdots + \dim F_n.$$

**Démonstration**

Si $\mathcal{B}_i = (e_1^{(i)}, \dots, e_{d_i}^{(i)})$ est une base de $F_i$, alors la famille des $(0, \dots, 0, e_k^{(i)}, 0, \dots, 0)$ pour $1 \leq i \leq n$ et $1 \leq k \leq d_i$ forme une base de $F_1 \times \cdots \times F_n$. Sa cardinal est $d_1 + \cdots + d_n$. □

**Exemple 4.3 (L'espace $\mathbb{K}^n$)**

En prenant $F_1 = \cdots = F_n = \mathbb{K}$, la définition redonne exactement

$$\underbrace{\mathbb{K} \times \cdots \times \mathbb{K}}_{n \text{ facteurs}} = \mathbb{K}^n.$$

La dimension d'un produit d'espaces de dimension finie est la somme de leurs dimensions (proposition précédente) ; comme $\dim \mathbb{K} = 1$, on obtient $\dim \mathbb{K}^n = n$. Les facteurs n'ont aucune raison d'être égaux : $\mathbb{R}^2 \times \mathbb{R}^3$ est un espace vectoriel de dimension 5, dont les éléments sont les couples $((x_1, x_2), (y_1, y_2, y_3))$. En concaténant les coordonnées, on l'identifie naturellement à $\mathbb{R}^5$.

**Exemple 4.4 (Une matrice est la donnée de ses colonnes)**

Soient $n, p \in \mathbb{N}^*$. À une matrice $M \in \mathcal{M}_{n,p}(\mathbb{K})$ on associe le $p$-uplet $(C_1, \dots, C_p)$ de ses colonnes, qui sont des éléments de $\mathcal{M}_{n,1}(\mathbb{K})$, identifié à $\mathbb{K}^n$. L'application

$$\Phi \colon \mathcal{M}_{n,p}(\mathbb{K}) \to (\mathbb{K}^n)^p, \qquad M \mapsto (C_1, \dots, C_p)$$

est linéaire, car l'addition et la multiplication par un scalaire se font coefficient par coefficient, donc colonne par colonne. Elle est bijective, car une matrice est

entièrement déterminée par ses colonnes et tout $p$-uplet de colonnes définit une matrice. C'est donc un isomorphisme, et on retrouve

$$\dim \mathcal{M}_{n,p}(\mathbb{K}) = p \times \dim(\mathbb{K}^n) = np.$$

**Exemple 4.5 (Solutions d'une équation différentielle)**

Soit $\omega > 0$ et soit $S$ l'ensemble des solutions sur $\mathbb{R}$ de l'équation différentielle $y'' + \omega^2 y = 0$ ; c'est un $\mathbb{R}$-espace vectoriel. On connaît la forme de ses éléments : ce sont exactement les fonctions $t \mapsto \lambda\cos(\omega t) + \mu\sin(\omega t)$, où $\lambda$ et $\mu$ décrivent $\mathbb{R}$. Autrement dit, l'application

$$\Psi : \mathbb{R}\times\mathbb{R} \to S, \qquad (\lambda,\mu) \mapsto (t \mapsto \lambda\cos(\omega t) + \mu\sin(\omega t))$$

est linéaire et surjective. Elle est aussi injective : si $\Psi(\lambda,\mu)$ est la fonction nulle, l'évaluation en 0 donne $\lambda = 0$, puis l'évaluation de la dérivée en 0 donne $\mu\omega = 0$, donc $\mu = 0$ puisque $\omega \neq 0$. C'est donc un isomorphisme, et

$$\dim S = \dim(\mathbb{R}\times\mathbb{R}) = 2.$$

Le produit traduit ici que se donner une solution revient à se donner deux nombres indépendants, les coefficients des deux solutions de référence $t \mapsto \cos(\omega t)$ et $t \mapsto \sin(\omega t)$.

**Définition 4.6 (Somme de sous-espaces vectoriels)**

Soient $F_1, \dots, F_n$ des sous-espaces vectoriels d'un $\mathbb{K}$-espace vectoriel $E$. La *somme* $F_1 + \cdots + F_n$ est le sous-espace vectoriel

$$F_1 + \cdots + F_n = \{x_1 + \cdots + x_n \mid (x_1,\dots,x_n) \in F_1 \times \cdots \times F_n\}.$$

C'est le plus petit sous-espace vectoriel de $E$ contenant tous les $F_i$.

**Exemple 4.7 (Deux plans de $\mathbb{R}^3$)**

Dans $E = \mathbb{R}^3$ rapporté à sa base canonique $(e_1, e_2, e_3)$, posons $F_1 = \mathrm{Vect}(e_1, e_2)$ et $F_2 = \mathrm{Vect}(e_2, e_3)$.
Tout vecteur $x = (x_1, x_2, x_3)$ s'écrit $x = (x_1, x_2, 0) + (0, 0, x_3)$, avec $(x_1, x_2, 0) \in F_1$ et $(0, 0, x_3) \in F_2$. Donc $F_1 + F_2 = \mathbb{R}^3$.
Cette écriture n'est pas la seule : on a aussi $x = (x_1, 0, 0) + (0, x_2, x_3)$. Le vecteur $e_2$ appartient en effet à la fois à $F_1$ et à $F_2$, et il se décompose au moins de deux façons : $e_2 = e_2 + 0 = 0 + e_2$. On notera par ailleurs que $\dim F_1 + \dim F_2 = 4$, alors que $\dim(F_1 + F_2) = 3$.

**Exemple 4.8 (Un Vect est une somme de droites)**

Soient $v_1, \dots, v_p$ des vecteurs de $E$ et, pour chaque $i$, $D_i = \mathrm{Vect}(v_i)$ la droite qu'il engendre. Par définition de la somme :

$$D_1 + \cdots + D_p = \left\{\lambda_1 v_1 + \cdots + \lambda_p v_p \mid (\lambda_1,\dots,\lambda_p) \in \mathbb{K}^p\right\} = \mathrm{Vect}(v_1,\dots,v_p).$$

Une somme de droites n'est donc rien d'autre qu'un espace engendré, et l'on retrouve

au passage que $\mathrm{Vect}(v_1, \dots, v_p)$ est le plus petit sous-espace vectoriel de $E$ contenant $v_1, \dots, v_p$. Par exemple, dans $\mathbb{K}[X]$ :

$$\mathbb{K}_n[X] = \mathrm{Vect}(1) + \mathrm{Vect}(X) + \cdots + \mathrm{Vect}(X^n).$$

**Exemple 4.9 (Matrices triangulaires)**

Soit $n \geq 2$. Notons $\mathcal{T}^+$ (resp. $\mathcal{T}^-$) l'ensemble des matrices triangulaires supérieures (resp. inférieures) de $\mathcal{M}_n(\mathbb{K})$, et $\mathcal{D}$ celui des matrices diagonales. Ce sont des sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{K})$.
Toute matrice $M = (m_{i,j})$ est somme de sa partie strictement supérieure, de sa diagonale et de sa partie strictement inférieure ; les deux premières donnent un élément de $\mathcal{T}^+$ et la dernière un élément de $\mathcal{T}^-$. Donc

$$\mathcal{T}^+ + \mathcal{T}^- = \mathcal{M}_n(\mathbb{K}).$$

Là encore la décomposition n'est pas unique, puisqu'on peut placer la diagonale de $M$ dans l'un ou l'autre terme ; d'ailleurs $\mathcal{T}^+ \cap \mathcal{T}^- = \mathcal{D} \neq \{0\}$.
En revanche, si l'on note $\mathcal{N}^+$ et $\mathcal{N}^-$ les ensembles de matrices triangulaires supérieures et inférieures *à diagonale nulle*, on a encore

$$\mathcal{N}^+ + \mathcal{D} + \mathcal{N}^- = \mathcal{M}_n(\mathbb{K}),$$

mais cette fois la décomposition est unique : le terme dans $\mathcal{N}^+$ est imposé par les coefficients $m_{i,j}$ pour $i < j$, celui dans $\mathcal{D}$ par les $m_{i,i}$ et celui dans $\mathcal{N}^-$ par les $m_{i,j}$ pour $i > j$.

**Remarque (Produit et somme)**

Les mots *produit* et *somme* désignent ici deux constructions de natures différentes. Le produit $F_1 \times \cdots \times F_n$ se fabrique à partir d'espaces vectoriels quelconques, sans qu'aucun lien ne les unisse : c'est un espace *nouveau*, dont les éléments sont des $n$-uplets, et dans lequel les composantes restent séparées. La somme $F_1 + \cdots + F_n$, elle, n'a de sens que si les $F_i$ sont des sous-espaces d'un même espace $E$ : c'est un *sous-espace de* $E$, dont les éléments sont des vecteurs de $E$, et où les composantes ont été additionnées, donc perdues de vue.
Le vocabulaire décrit la construction et non la dimension, ce qui peut surprendre : le produit est le produit cartésien des ensembles sous-jacents, mais sa dimension est la *somme* des dimensions ; la somme est formée des sommes de vecteurs, mais sa dimension est en général *plus petite* que la somme des dimensions, comme le montre l'exemple des deux plans de $\mathbb{R}^3$.
Les deux notions sont néanmoins reliées. Si les $F_i$ sont des sous-espaces d'un même $E$, l'application

$$\varphi : F_1 \times \cdots \times F_n \to E, \qquad (x_1, \dots, x_n) \mapsto x_1 + \cdots + x_n$$

est linéaire, et son image est exactement $F_1 + \cdots + F_n$ : la somme est l'image du produit par $\varphi$. Reste à savoir si $\varphi$ perd de l'information, c'est-à-dire si l'on peut remonter d'un vecteur de la somme aux composantes dont il provient. C'est l'objet de la définition suivante.

**Définition 4.10 (Somme directe)**

On dit que la somme $F_1 + \cdots + F_n$ est *directe*, et on note $F_1 \oplus \cdots \oplus F_n$ ou $\bigoplus_{i=1}^{n} F_i$, lorsque tout élément de la somme admet une *unique* décomposition :

$$\forall x \in F_1 + \cdots + F_n, \quad \exists!(x_1, \ldots, x_n) \in F_1 \times \cdots \times F_n, \quad x = x_1 + \cdots + x_n.$$

De manière équivalente, la somme est directe si et seulement si l'application

$$\varphi : F_1 \times \cdots \times F_n \longrightarrow F_1 + \cdots + F_n, \quad (x_1, \ldots, x_n) \mapsto x_1 + \cdots + x_n$$

est injective.

**Proposition 4.11 (Caractérisation d'une somme directe)**

Soient $F_1, \ldots, F_n$ des sous-espaces vectoriels de $E$. La somme $F_1 + \cdots + F_n$ est directe si et seulement si la seule décomposition du vecteur nul est la décomposition nulle, c'est-à-dire :

$$\forall (x_1, \ldots, x_n) \in F_1 \times \cdots \times F_n, \quad (x_1 + \cdots + x_n = 0_E) \implies (x_1 = \cdots = x_n = 0_E).$$

**Démonstration**

Supposons la somme directe. Soit $(x_1, \ldots, x_n)$ dans $F_1 \times \cdots \times F_n$ tel que $x_1 + \cdots + x_n = 0_E$. Le vecteur nul appartient à la somme et admet déjà la décomposition $0_E = 0_E + \cdots + 0_E$ ; par unicité, $(x_1, \ldots, x_n) = (0_E, \ldots, 0_E)$.
Réciproquement, supposons la condition vérifiée. Soit $x$ dans $F_1 + \cdots + F_n$. L'existence d'une décomposition est acquise par définition de la somme ; montrons son unicité. Soient donc $(x_1, \ldots, x_n)$ et $(y_1, \ldots, y_n)$ deux éléments de $F_1 \times \cdots \times F_n$ tels que

$$x = x_1 + \cdots + x_n = y_1 + \cdots + y_n.$$

En soustrayant :

$$(x_1 - y_1) + \cdots + (x_n - y_n) = 0_E,$$

et $x_i - y_i \in F_i$ pour tout $i$, puisque $F_i$ est un sous-espace vectoriel. L'hypothèse donne alors $x_i - y_i = 0_E$, c'est-à-dire $x_i = y_i$, pour tout $i$.
Notons que la condition de l'énoncé s'écrit exactement $\operatorname{Ker} \varphi = \{0\}$, où $\varphi$ est l'application introduite dans la définition : on retrouve que la somme est directe si et seulement si $\varphi$ est injective. □

**Remarque (Le cas de deux sous-espaces)**

Pour $n = 2$, on retrouve le critère vu en première année : la somme $F_1 + F_2$ est directe si et seulement si $F_1 \cap F_2 = \{0_E\}$.
En effet, supposons la somme directe et soit $x \in F_1 \cap F_2$. Alors $x + (-x) = 0_E$, avec $x \in F_1$ et $-x \in F_2$ : la proposition précédente donne $x = 0_E$. Réciproquement, supposons $F_1 \cap F_2 = \{0_E\}$ et soient $x_1 \in F_1$ et $x_2 \in F_2$ tels que $x_1 + x_2 = 0_E$. Alors $x_1 = -x_2$ appartient à la fois à $F_1$ et à $F_2$, donc $x_1 = 0_E$, puis $x_2 = 0_E$.
Ce critère ne se généralise pas : pour $n \geq 3$, il ne suffit pas que les $F_i$ soient deux à deux d'intersection nulle. Dans $E = \mathbb{R}^2$ rapporté à sa base canonique, prenons

$$D_1 = \operatorname{Vect}(e_1), \qquad D_2 = \operatorname{Vect}(e_2), \qquad D_3 = \operatorname{Vect}(e_1 + e_2).$$

Ces trois droites sont deux à deux distinctes, donc deux à deux d'intersection nulle.

Pourtant

$$e_1 + e_2 + (-(e_1 + e_2)) = 0_E$$

est une décomposition non nulle du vecteur nul : la somme n'est pas directe.

**Exemple 4.12 (Trois droites de $\mathbb{R}^3$)**

Dans $E = \mathbb{R}^3$, posons $u_1 = (1, 1, 0)$, $u_2 = (0, 1, 1)$, $u_3 = (1, 0, 1)$ et $D_i = \text{Vect}(u_i)$ pour $i \in \{1, 2, 3\}$.
Un élément de $D_i$ s'écrit $\lambda_i u_i$ avec $\lambda_i \in \mathbb{R}$. Supposons $\lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 u_3 = 0$, c'est-à-dire, en coordonnées :

$$(\lambda_1 + \lambda_3,\ \lambda_1 + \lambda_2,\ \lambda_2 + \lambda_3) = (0, 0, 0).$$

Les deux premières équations donnent $\lambda_2 = \lambda_3 = -\lambda_1$, et la troisième devient alors $-2\lambda_1 = 0$. Donc $\lambda_1 = 0$, puis $\lambda_2 = \lambda_3 = 0$, et les trois vecteurs $\lambda_i u_i$ sont nuls : la somme $D_1 + D_2 + D_3$ est directe.
Le calcul montre au passage que la famille $(u_1, u_2, u_3)$ est libre ; comptant trois vecteurs dans un espace de dimension 3, c'est une base, et

$$\mathbb{R}^3 = D_1 \oplus D_2 \oplus D_3.$$

Les trois droites sont deux à deux d'intersection nulle, mais ce n'est pas ce qui permet de conclure : c'est bien le calcul ci-dessus.

**Exemple 4.13 (Un plan et une droite de $\mathbb{R}^3$, de deux façons)**

Dans $E = \mathbb{R}^3$, notons $s(x, y, z) = x + y + z$ : c'est une forme linéaire non nulle. Posons

$$F = \text{Ker}\, s = \left\{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\right\} \qquad \text{et} \qquad D = \text{Vect}(w),\ \text{où } w = (1, 1, 1).$$

Ainsi $F$ est un plan, $D$ est une droite, et $s(w) = 3$.
*Par l'intersection.* Soit $v \in F \cap D$. Comme $v \in D$, il existe $t \in \mathbb{R}$ tel que $v = tw$. Comme $v \in F$, on a $0 = s(v) = t\, s(w) = 3t$, donc $t = 0$ et $v = 0$. Ainsi $F \cap D = \{0\}$ : la somme de deux sous-espaces d'intersection nulle est directe (remarque précédente).
*Par l'unicité de l'écriture.* Soient $v \in F$ et $t \in \mathbb{R}$ tels que $v + tw = 0$. En appliquant $s$ :

$$0 = s(v) + t\, s(w) = 0 + 3t,$$

donc $t = 0$, puis $v = 0$. La seule décomposition du vecteur nul est la décomposition nulle, donc la somme est directe (caractérisation d'une somme directe, proposition précédente).
Dans les deux cas, la somme est directe. Elle vaut d'ailleurs $\mathbb{R}^3$ : pour $v \in \mathbb{R}^3$, en posant $t = \frac{s(v)}{3}$, on a $s(v - tw) = s(v) - 3t = 0$, donc $v - tw \in F$ et $v = (v - tw) + tw$. Finalement $\mathbb{R}^3 = F \oplus D$.

**Exemple 4.14 (Une décomposition de $\mathcal{M}_n(\mathbb{K})$)**

Soit $n \geq 2$. Pour $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{K})$, notons $s(M) = \sum_{i=1}^{n} m_{i,i}$ la somme des coefficients diagonaux de $M$. C'est une forme linéaire sur $\mathcal{M}_n(\mathbb{K})$, et $s(M^\top) = s(M)$

puisque la transposition ne change pas la diagonale. Posons

$$\mathcal{A} = \left\{M \mid M^\top = -M\right\}, \qquad \mathcal{H} = \mathrm{Vect}(I_n), \qquad \mathcal{S}_0 = \left\{M \mid M^\top = M \text{ et } s(M) = 0\right\}.$$

Ce sont trois sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{K})$. Montrons que

$$\mathcal{M}_n(\mathbb{K}) = \mathcal{A} \oplus \mathcal{H} \oplus \mathcal{S}_0.$$

*La somme est directe.* Soient $A \in \mathcal{A}$, $\lambda \in \mathbb{K}$ et $S \in \mathcal{S}_0$ tels que $A + \lambda I_n + S = 0$. En transposant, et comme $A^\top = -A, I_n^\top = I_n$ et $S^\top = S$ :

$$-A + \lambda I_n + S = 0.$$

La différence des deux relations donne $2A = 0$, donc $A = 0$ ; leur somme donne $\lambda I_n + S = 0$. En appliquant $s$ à cette dernière égalité, et comme $s(I_n) = n$ et $s(S) = 0$, il vient $n\lambda = 0$, donc $\lambda = 0$ et enfin $S = 0$. La seule décomposition de la matrice nulle est la décomposition nulle, donc la somme est directe (caractérisation d'une somme directe).
*La somme vaut* $\mathcal{M}_n(\mathbb{K})$. Soit $M \in \mathcal{M}_n(\mathbb{K})$. Posons

$$A = \frac{M - M^\top}{2}, \qquad \lambda = \frac{s(M)}{n}, \qquad S = \frac{M + M^\top}{2} - \lambda I_n.$$

On vérifie que $A^\top = -A$, donc $A \in \mathcal{A}$, et que $S^\top = S$. De plus $s\left(\frac{M+M^\top}{2}\right) = s(M)$, donc $s(S) = s(M) - \lambda n = 0$ et $S \in \mathcal{S}_0$. Enfin $A + \lambda I_n + S = M$.
La forme linéaire $s$ est la *trace*, que nous retrouverons plus loin dans ce chapitre.

**Proposition 4.15 (Dimension d'une somme)**

Si $F_1, \ldots, F_n$ sont des sous-espaces vectoriels de dimension finie, alors

$$\dim(F_1 + \cdots + F_n) \leq \dim F_1 + \cdots + \dim F_n,$$

avec égalité si et seulement si la somme est directe.

**Démonstration**

L'application $\varphi$ ci-dessus est surjective par définition, donc par le théorème du rang :

$$\dim(F_1 + \cdots + F_n) = \dim(F_1 \times \cdots \times F_n) - \dim(\mathrm{Ker}\,\varphi) = \sum_{i=1}^{n} \dim F_i - \dim(\mathrm{Ker}\,\varphi).$$

L'inégalité en découle, et l'égalité a lieu si et seulement si $\mathrm{Ker}\,\varphi = \{0\}$, c'est-à-dire si et seulement si $\varphi$ est injective, autrement dit si la somme est directe. □

**Exemple 4.16 (Un cas d'égalité)**

Reprenons la décomposition $\mathcal{M}_n(\mathbb{K}) = \mathcal{A} \oplus \mathcal{H} \oplus \mathcal{S}_0$ obtenue plus haut. La somme étant directe, sa dimension est la somme des dimensions des sous-espaces (proposition précédente) :

$$n^2 = \dim \mathcal{M}_n(\mathbb{K}) = \dim \mathcal{A} + \dim \mathcal{H} + \dim \mathcal{S}_0.$$

Or $\dim \mathcal{A} = \frac{n(n-1)}{2}$ et $\dim \mathcal{H} = 1$. On en déduit

$$\dim \mathcal{S}_0 = n^2 - \frac{n(n-1)}{2} - 1 = \frac{n(n+1)}{2} - 1.$$

La proposition sert donc ici à *calculer* une dimension inconnue à partir de celles que l'on connaît.

**Exemple 4.17 (Un cas d'inégalité stricte)**

Reprenons les deux plans $F_1 = \mathrm{Vect}(e_1, e_2)$ et $F_2 = \mathrm{Vect}(e_2, e_3)$ de $\mathbb{R}^3$. On a vu que $F_1 + F_2 = \mathbb{R}^3$, donc

$$\dim(F_1 + F_2) = 3 < 4 = \dim F_1 + \dim F_2.$$

L'inégalité est stricte, ce qui confirme que la somme n'est pas directe : on avait effectivement exhibé deux décompositions du même vecteur.
De même, si $\mathcal{T}^+$ et $\mathcal{T}^-$ désignent les matrices triangulaires supérieures et inférieures de $\mathcal{M}_n(\mathbb{K})$, on a $\mathcal{T}^+ + \mathcal{T}^- = \mathcal{M}_n(\mathbb{K})$, et

$$\dim(\mathcal{T}^+ + \mathcal{T}^-) = n^2 < n(n+1) = \frac{n(n+1)}{2} + \frac{n(n+1)}{2} = \dim \mathcal{T}^+ + \dim \mathcal{T}^-.$$

L'écart vaut $n$, qui est exactement la dimension de $\mathcal{T}^+ \cap \mathcal{T}^-$, l'espace des matrices diagonales : c'est ce que prédit la formule de Grassmann, valable pour deux sous-espaces.

**Proposition 4.18 (Concaténation de familles libres)**

Soient $F_1, \dots, F_n$ des sous-espaces vectoriels de $E$. La somme $F_1 + \cdots + F_n$ est directe si et seulement si, pour toutes familles libres $\mathcal{L}_1$ de $F_1$, …, $\mathcal{L}_n$ de $F_n$, la famille concaténée $(\mathcal{L}_1, \dots, \mathcal{L}_n)$ est libre.

**Démonstration**

Supposons la somme directe. Soient $\mathcal{L}_1, \dots, \mathcal{L}_n$ des familles libres de $F_1, \dots, F_n$, et notons $\mathcal{L}_i = (e_{i,1}, \dots, e_{i,p_i})$. Considérons des scalaires $\lambda_{i,j}$ tels que

$$\sum_{i=1}^{n} \sum_{j=1}^{p_i} \lambda_{i,j} e_{i,j} = 0_E,$$

et posons $x_i = \sum_{j=1}^{p_i} \lambda_{i,j} e_{i,j}$, qui appartient à $F_i$. On a $x_1 + \cdots + x_n = 0_E$, donc $x_i = 0_E$ pour tout $i$ par la caractérisation d'une somme directe. Comme $\mathcal{L}_i$ est libre, les $\lambda_{i,j}$ sont tous nuls : la famille concaténée est libre.
Réciproquement, supposons la condition vérifiée. Soit $(x_1, \dots, x_n) \in F_1 \times \cdots \times F_n$ tel que $x_1 + \cdots + x_n = 0_E$. Pour chaque $i$, posons $\mathcal{L}_i = (x_i)$ si $x_i \neq 0_E$, et $\mathcal{L}_i$ la famille vide sinon : c'est une famille libre de $F_i$. La famille concaténée est donc libre. Or la somme de ses vecteurs est $x_1 + \cdots + x_n = 0_E$ : c'est une combinaison linéaire nulle dont tous les coefficients valent 1. La famille concaténée est donc vide, c'est-à-dire que $x_i = 0_E$ pour tout $i$, et la somme est directe. □

**Remarque**

Supposons $E = F_1 \oplus \cdots \oplus F_r$, et donnons-nous des bases $\mathcal{B}_1, \ldots, \mathcal{B}_r$ de $F_1, \ldots, F_r$. La famille concaténée $(\mathcal{B}_1, \ldots, \mathcal{B}_r)$ est libre d'après la proposition précédente. Elle est aussi génératrice : tout $x \in E$ s'écrit $x = x_1 + \cdots + x_r$ avec $x_i \in F_i$, et chaque $x_i$ est combinaison linéaire des vecteurs de $\mathcal{B}_i$. On obtient ainsi une base de $E$ en concaténant des bases de chacun des $F_i$.

**Définition 4.19 (Base adaptée à une décomposition)**

Soit $E = F_1 \oplus \cdots \oplus F_r$ une décomposition en somme directe. Une *base de $E$ adaptée à cette décomposition* est une base obtenue par concaténation de bases $\mathcal{B}_1, \ldots, \mathcal{B}_r$ des sous-espaces $F_1, \ldots, F_r$ :

$$\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_r).$$

**Exemple 4.20 (Une base adaptée à une décomposition de $\mathcal{M}_2(\mathbb{K})$)**

Reprenons $\mathcal{M}_2(\mathbb{K}) = \mathcal{A} \oplus \mathcal{H} \oplus \mathcal{S}_0$ dans le cas $n = 2$, et posons

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad K = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad L = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

La famille $(J)$ est une base de $\mathcal{A}$ et $(I_2)$ une base de $\mathcal{H}$. Enfin une matrice symétrique s'écrit $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$, et la condition $a + c = 0$ la met sous la forme $aK + bL$ : la famille $(K, L)$ est une base de $\mathcal{S}_0$.
Par concaténation, $\mathcal{B} = (J, I_2, K, L)$ est une base de $\mathcal{M}_2(\mathbb{K})$ adaptée à la décomposition. Elle compte bien 4 vecteurs, et l'on retrouve $\dim \mathcal{S}_0 = 2 = \frac{2\times 3}{2} - 1$.

**Exemple 4.21 (Des bases non adaptées)**

La base canonique $(E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2})$ de $\mathcal{M}_2(\mathbb{K})$ n'est pas adaptée à la décomposition précédente : aucune de ces quatre matrices n'appartient à $\mathcal{A}$, à $\mathcal{H}$ ou à $\mathcal{S}_0$.
Un échec peut être plus discret. La famille $(J, I_2, K, L + I_2)$ est encore une base de $\mathcal{M}_2(\mathbb{K})$ : on n'a modifié le dernier vecteur qu'en lui ajoutant un autre vecteur de la famille. Mais elle n'est pas adaptée, car $L + I_2$ n'est pas dans $\mathcal{S}_0$ : la somme de ses coefficients diagonaux vaut 2. Les deux derniers vecteurs ne forment donc pas une base de $\mathcal{S}_0$.

**Remarque (Construction inverse : partitionner une base)**

La définition précédente part d'une décomposition pour fabriquer une base. On peut faire le chemin inverse.
Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de $E$, que l'on découpe en $r$ paquets consécutifs $\mathcal{B}_1, \ldots, \mathcal{B}_r$, et posons $F_i = \mathrm{Vect}(\mathcal{B}_i)$. Alors

$$E = F_1 \oplus \cdots \oplus F_r,$$

et $\mathcal{B}$ est adaptée à cette décomposition.
En effet, tout $x \in E$ s'écrit comme combinaison linéaire des vecteurs de $\mathcal{B}$ ; en regroupant les termes paquet par paquet, on obtient $x = x_1 + \cdots + x_r$ avec $x_i \in F_i$, donc $E = F_1 + \cdots + F_r$. Et si $x_1 + \cdots + x_r = 0_E$ avec $x_i \in F_i$, en écrivant chaque $x_i$ dans

$\mathcal{B}_i$ on obtient une combinaison linéaire nulle des vecteurs de $\mathcal{B}$ : tous les coefficients sont nuls puisque $\mathcal{B}$ est libre, donc chaque $x_i$ est nul. La somme est bien directe.

Par exemple, dans $E = \mathbb{K}_3[X]$ muni de la base $(1, X, X^2, X^3)$, prenons les deux paquets $(1, X^2)$ et $(X, X^3)$. En posant

$$P = \mathrm{Vect}(1, X^2) \qquad \text{et} \qquad I = \mathrm{Vect}(X, X^3),$$

on obtient $\mathbb{K}_3[X] = P \oplus I$ : tout polynôme de degré au plus 3 s'écrit de manière unique comme somme d'un polynôme pair et d'un polynôme impair. La base $(1, X^2, X, X^3)$ est adaptée à cette décomposition.

**Définition 4.22 (Base adaptée à un sous-espace)**

Soit $F$ un sous-espace vectoriel de $E$. Une *base de $E$ adaptée à $F$* est une base de $E$ adaptée à une décomposition $E = F \oplus G$, où $G$ est un supplémentaire de $F$ dans $E$.

**Remarque**

Une base $(e_1, \ldots, e_n)$ de $E$ est adaptée à $F$ si et seulement si $(e_1, \ldots, e_p)$ est une base de $F$, avec $p = \dim F$. Le sens direct découle de la définition. Réciproquement, si c'est le cas, la remarque précédente appliquée aux paquets $(e_1, \ldots, e_p)$ et $(e_{p+1}, \ldots, e_n)$ donne $E = F \oplus \mathrm{Vect}(e_{p+1}, \ldots, e_n)$, et la base est adaptée à cette décomposition. Une telle base existe toujours : il suffit de compléter une base de $F$ en une base de $E$ par le théorème de la base incomplète.

**Exemple 4.23 (Une base adaptée à un plan de $\mathbb{R}^3$)**

Reprenons le plan $F = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$ et la droite $D = \mathrm{Vect}(w)$, avec $w = (1, 1, 1)$, pour lesquels $\mathbb{R}^3 = F \oplus D$. Les vecteurs $f_1 = (1, -1, 0)$ et $f_2 = (0, 1, -1)$ appartiennent à $F$ et ne sont pas colinéaires ; comme $\dim F = 2$, la famille $(f_1, f_2)$ est une base de $F$. La famille $(f_1, f_2, w)$ est donc une base de $\mathbb{R}^3$ adaptée à $F$.
La famille $(f_1, f_2, e_1)$, où $e_1 = (1, 0, 0)$, est aussi une base de $\mathbb{R}^3$ adaptée à $F$ : comme $e_1 \notin F$, elle est libre, et elle est adaptée à la décomposition $\mathbb{R}^3 = F \oplus \mathrm{Vect}(e_1)$. En revanche, la base canonique $(e_1, e_2, e_3)$ n'est pas adaptée à $F$, puisque $e_1 \notin F$.



# 4 Compléments d'algèbre linéaire

## II. Sous-espaces stables

**Définition 4.24 (Sous-espace stable)**

Soit $u \in \mathcal{L}(E)$ et $F$ un sous-espace vectoriel de $E$. On dit que $F$ est *stable par* $u$ si $u(F) \subset F$, c'est-à-dire :

$$\forall x \in F, \quad u(x) \in F.$$

**Remarque**

Les sous-espaces $\{0_E\}$, $E$, $\operatorname{Ker} u$ et $\operatorname{Im} u$ sont toujours stables par $u$. En effet, $u(0_E) = 0_E \in \{0_E\}$ ; $u(E) \subset E$ par définition ; si $x \in \operatorname{Ker} u$ alors $u(x) = 0_E \in \operatorname{Ker} u$ ; si $y \in \operatorname{Im} u$ alors $y = u(x)$ pour un certain $x \in E$ et $u(y) = u(u(x)) \in \operatorname{Im} u$.

**Exemple 4.25 (Un sous-espace stable)**

Soit $E = \mathbb{K}_3[X]$ et $u : P \mapsto P'$ la dérivation, qui est bien un endomorphisme de $E$ puisque $\deg P' \leq \deg P$. Posons $F = \mathbb{K}_1[X]$.
Soit $P \in F$, que l'on écrit $P = a + bX$. Alors $u(P) = b$, qui est encore un polynôme de degré au plus 1 : $u(P) \in F$. Donc $F$ est stable par $u$.
Plus généralement, $\mathbb{K}_p[X]$ est stable par $u$ pour tout $p \leq 3$, car dériver fait baisser le degré.

**Exemple 4.26 (Un sous-espace non stable)**

Soit $E = \mathbb{R}^2$ et $r$ la rotation d'angle $\frac{\pi}{2}$, donnée par $r(x, y) = (-y, x)$. La droite $D = \operatorname{Vect}(e_1)$ n'est pas stable par $r$ :

$$r(e_1) = e_2 \notin D.$$

Il se trouve qu'*aucune* droite de $\mathbb{R}^2$ n'est stable par $r$. En effet, supposons $\operatorname{Vect}(v)$ stable, avec $v \neq 0$. Alors $r(v) \in \operatorname{Vect}(v)$, donc il existe $\lambda \in \mathbb{R}$ tel que $r(v) = \lambda v$. En appliquant $r$ à nouveau, $r(r(v)) = \lambda^2 v$. Or $r \circ r = -\mathrm{Id}$, puisque $r(r(x, y)) = r(-y, x) = (-x, -y)$. Il vient $\lambda^2 v = -v$, et comme $v \neq 0$, $\lambda^2 = -1$ : impossible dans $\mathbb{R}$.

**Proposition 4.27 (Stabilité par un endomorphisme qui commute)**

Soient $u, v \in \mathcal{L}(E)$ tels que $u \circ v = v \circ u$. Alors $\operatorname{Ker} u$ et $\operatorname{Im} u$ sont stables par $v$.

**Démonstration**

Soit $x \in \operatorname{Ker} u$. Alors

$$u\left(v(x)\right) = (u \circ v)(x) = (v \circ u)(x) = v\left(u(x)\right) = v(0_E) = 0_E,$$

donc $v(x) \in \operatorname{Ker} u$ : le noyau de $u$ est stable par $v$.

Soit $y \in \operatorname{Im} u$, et soit $x \in E$ tel que $y = u(x)$. Alors

$$v(y) = v\left(u(x)\right) = u\left(v(x)\right) \in \operatorname{Im} u,$$

donc $\operatorname{Im} u$ est stable par $v$. □

**Remarque**

Le cas $v = u$ redonne la remarque précédente, puisque $u$ commute avec lui-même.

**Exemple 4.28 (Dérivation et translation)**

Sur $E = \mathbb{K}_3[X]$, considérons

$$u : P \mapsto P' \qquad \text{et} \qquad v : P(X) \mapsto P(X+1).$$

Ces deux endomorphismes commutent : pour tout $P$,

$$(u \circ v)(P) = \frac{\mathrm{d}}{\mathrm{d}X}\big(P(X+1)\big) = P'(X+1) = (v \circ u)(P).$$

Or $\operatorname{Ker} u = \operatorname{Vect}(1)$, car les polynômes de dérivée nulle sont les constantes, et $\operatorname{Im} u = \mathbb{K}_2[X]$. Comme $u$ et $v$ commutent, le noyau et l'image de $u$ sont stables par $v$ (proposition précédente) : $\operatorname{Vect}(1)$ et $\mathbb{K}_2[X]$ sont stables par $v$, ce que l'on vérifie directement : $v$ laisse fixe toute constante, et ne change pas le degré d'un polynôme.

**Définition 4.29 (Endomorphisme induit)**

Si $F$ est stable par $u$, alors l'application $u|_F : F \to F$ définie par $u|_F(x) = u(x)$ pour tout $x \in F$ est un endomorphisme de $F$, appelé *endomorphisme induit par $u$ sur $F$*.

**Exemple 4.30 (L'endomorphisme induit par la dérivation)**

Reprenons $E = \mathbb{K}_3[X]$, $u : P \mapsto P'$ et $F = \mathbb{K}_1[X]$, stable par $u$. L'endomorphisme induit est

$$u|_F : \mathbb{K}_1[X] \to \mathbb{K}_1[X], \qquad a + bX \mapsto b.$$

C'est un endomorphisme de $F$, et non de $E$ : c'est un objet différent de $u$. Sa matrice dans la base $(1, X)$ de $F$ est

$$\operatorname{Mat}_{(1,X)}(u|_F) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

On peut lire sur $u|_F$ des propriétés qui lui sont propres : il n'est pas injectif, son noyau étant $\operatorname{Vect}(1)$, et son image est également $\operatorname{Vect}(1)$.
L'hypothèse de stabilité est indispensable pour former $u|_F$. Si l'on prend $G = \operatorname{Vect}(X^2)$, la restriction de $u$ à $G$ envoie $X^2$ sur $2X$, qui n'est pas dans $G$ : on obtient une application de $G$ dans $E$, mais pas un endomorphisme de $G$.

**Proposition 4.31 (Matrice d'un endomorphisme avec un sous-espace stable)**

Soit $u \in \mathcal{L}(E)$ et $E = F \oplus G$ avec $F$ stable par $u$. Si $\mathcal{B} = (\mathcal{B}_F, \mathcal{B}_G)$ est une base adaptée

à cette décomposition, alors

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} A & C \\ 0 & D \end{pmatrix},$$

où $A = \mathrm{Mat}_{\mathcal{B}_F}(u|_F) \in \mathcal{M}_p(\mathbb{K})$, $D \in \mathcal{M}_q(\mathbb{K})$ et $C \in \mathcal{M}_{p,q}(\mathbb{K})$, avec $p = \dim F$ et $q = \dim G$.

**Démonstration**

Soit $\mathcal{B}_F = (e_1, \dots, e_p)$ et $\mathcal{B}_G = (e_{p+1}, \dots, e_n)$. Pour $1 \leq j \leq p$, on a $e_j \in F$ et $F$ est stable par $u$, donc $u(e_j) \in F$. Les coordonnées de $u(e_j)$ dans la base $\mathcal{B}$ ont donc des composantes nulles sur les vecteurs $e_{p+1}, \dots, e_n$ : le bloc en bas à gauche est nul. Le bloc $A$ est la matrice de $u|_F$ dans la base $\mathcal{B}_F$. □

**Remarque**

La base $\mathcal{B}$ de la proposition est une base de $E$ adaptée à $F$. Réciproquement, toute base de $E$ adaptée à $F$ est adaptée à une décomposition $E = F \oplus G$ : la matrice de $u$ dans une telle base a donc la forme précédente.

**Remarque**

L'écriture $\begin{pmatrix} A & C \\ 0 & D \end{pmatrix}$ est une écriture *par blocs* : on découpe la matrice de taille $n = p+q$ selon ses $p$ premières lignes et ses $q$ dernières, et selon ses $p$ premières colonnes et ses $q$ dernières. On obtient quatre sous-matrices rectangulaires, dont chacune est notée par une seule lettre ; ici, 0 désigne la matrice nulle de $\mathcal{M}_{q,p}(\mathbb{K})$. C'est une manière pratique de présenter une matrice.

**Exemple 4.32 (Une matrice triangulaire par blocs)**

Reprenons $E = \mathbb{K}_3[X]$ et la dérivation $u : P \mapsto P'$. Le sous-espace $F = \mathbb{K}_1[X]$ est stable par $u$, et

$$G = \mathrm{Vect}(X^2, X^3)$$

en est un supplémentaire, puisque $(1, X)$ et $(X^2, X^3)$ partitionnent la base $(1, X, X^2, X^3)$ de $E$. Cette base est donc adaptée à la décomposition $E = F \oplus G$. Or

$$u(1) = 0, \qquad u(X) = 1, \qquad u(X^2) = 2X, \qquad u(X^3) = 3X^2,$$

d'où, en écrivant ces images en colonnes :

$$\mathrm{Mat}_{\mathcal{B}}(u) = \left(\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ \hline 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Le bloc en bas à gauche est bien nul, et le bloc en haut à gauche est la matrice de $u|_F$ dans la base $(1, X)$, calculée plus haut.
Le bloc en haut à droite, lui, n'est pas nul : c'est que $G$ n'est pas stable par $u$, puisque $u(X^2) = 2X \notin G$.

**Théorème 4.33 (Matrice diagonale par blocs)**

Soit $u \in \mathcal{L}(E)$ et $E = F_1 \oplus \cdots \oplus F_r$ avec tous les $F_i$ stables par $u$. Dans une base $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_r)$ adaptée à cette décomposition, la matrice de $u$ est diagonale par blocs :

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{pmatrix}, \quad \text{où } A_i = \mathrm{Mat}_{\mathcal{B}_i}(u|_{F_i}).$$

**Démonstration**

Pour $j$ correspondant à un vecteur de $\mathcal{B}_k$, on a $u(e_j) \in F_k$ car $F_k$ est stable. Or la base $\mathcal{B}$ est adaptée à la décomposition, donc les coordonnées de $u(e_j)$ sont nulles en dehors du bloc correspondant à $F_k$. On obtient ainsi une matrice diagonale par blocs. □

**Exemple 4.34 (Une matrice diagonale par blocs)**

Reprenons $E = \mathbb{K}_3[X]$, mais cette fois avec la dérivée seconde $v : P \mapsto P''$, et la décomposition en polynômes pairs et impairs

$$E = \mathcal{P} \oplus \mathcal{I}, \qquad \mathcal{P} = \mathrm{Vect}(1, X^2), \qquad \mathcal{I} = \mathrm{Vect}(X, X^3),$$

obtenue en partitionnant la base $(1, X, X^2, X^3)$.
Dériver deux fois fait baisser le degré de deux : la parité est conservée. Les deux sous-espaces sont donc stables par $v$, ce que confirme le calcul :

$$v(1) = 0, \qquad v(X^2) = 2, \qquad v(X) = 0, \qquad v(X^3) = 6X.$$

Dans la base adaptée $\mathcal{B} = (1, X^2, X, X^3)$ :

$$\mathrm{Mat}_{\mathcal{B}}(v) = \left(\begin{array}{cc|cc} 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{array}\right),$$

diagonale par blocs, les deux blocs étant les matrices de $v|_{\mathcal{P}}$ dans $(1, X^2)$ et de $v|_{\mathcal{I}}$ dans $(X, X^3)$.
La dérivation $u : P \mapsto P'$, elle, échange les deux sous-espaces au lieu de les stabiliser : $u(X^2) = 2X \in \mathcal{I}$. Sa matrice dans la même base n'est pas diagonale par blocs.



# 4 Compléments d'algèbre linéaire

## III. Matrices par blocs

On considère dans cette section des matrices à coefficients dans $\mathbb{K}$, découpées en blocs compatibles.

**Proposition 4.35 (Opérations par blocs)**

Soient $A$ et $B$ deux matrices découpées en blocs de tailles compatibles.

1. **Addition :** $(A + B)$ se calcule bloc par bloc.
2. **Multiplication :** si $A = (A_{ik})$ et $B = (B_{kj})$, alors $AB = (C_{ij})$ avec $C_{ij} = \sum_k A_{ik}B_{kj}$, la formule usuelle appliquée aux blocs.
3. **Transposition :** si $A = (A_{ij})$, alors $A^\top = (A_{ji}^\top)$.

**Exemple 4.36 (Addition par blocs)**

Découpons deux matrices de $\mathcal{M}_3(\mathbb{K})$ de la même manière, en blocs de tailles 2 et 1 :

$$A = \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 3 & 4 & 1 \\ \hline 0 & 1 & 2 \end{array}\right), \qquad B = \left(\begin{array}{cc|c} 0 & 1 & 1 \\ 1 & 0 & 0 \\ \hline 2 & 2 & 3 \end{array}\right).$$

Chaque bloc de $A + B$ s'obtient en ajoutant les blocs correspondants :

$$A + B = \left(\begin{array}{cc|c} 1 & 3 & 1 \\ 4 & 4 & 1 \\ \hline 2 & 3 & 5 \end{array}\right).$$

Il est essentiel que les *mêmes* découpages soient employés pour $A$ et pour $B$ : sans cela, les blocs à additionner n'auraient pas les mêmes tailles.
On en déduit par exemple que, pour un découpage fixé, l'ensemble des matrices diagonales par blocs est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ : ajouter deux telles matrices, ou en multiplier une par un scalaire, laisse nuls les blocs hors de la diagonale.

**Exemple 4.37 (Multiplication par blocs)**

Le seul accord exigé porte sur les indices sommés : le découpage des *colonnes* de $A$ doit être celui des *lignes* de $B$. Les découpages des lignes de $A$ et des colonnes de $B$ sont libres, et se retrouvent tels quels dans $AB$.

Prenons $A \in \mathcal{M}_{3,4}(\mathbb{K})$, lignes découpées en $1+2$ et colonnes en $1+3$, et $B \in \mathcal{M}_{4,3}(\mathbb{K})$, lignes découpées en $1+3$ et colonnes en $2+1$ :

$$A = \left(\begin{array}{c|ccc} 2 & 1 & 0 & 1 \\ \hline 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 1 \end{array}\right), \qquad B = \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 0 \\ \hline 2 & 0 & 1 \\ 1 & 1 & 0 \end{array}\right).$$

Le découpage $1+3$ est commun aux colonnes de $A$ et aux lignes de $B$ : les produits $A_{ik}B_{kj}$ ont bien un sens. En revanche $A$ est découpée en $2 \times 2$ blocs de tailles $1 \times 1$, $1 \times 3$, $2 \times 1$ et $2 \times 3$, et $B$ en blocs de tailles $1 \times 2$, $1 \times 1$, $3 \times 2$ et $3 \times 1$ : aucune symétrie là-dedans.
Le produit $AB$ est de taille $3 \times 3$, avec les lignes découpées en $1+2$ (comme $A$) et les colonnes en $2+1$ (comme $B$) :

$$AB = \left(\begin{array}{cc|c} 3 & 2 & 2 \\ 5 & 0 & 3 \\ \hline 3 & 2 & 1 \end{array}\right).$$

Vérifions par exemple le bloc en bas à gauche, de taille $2 \times 2$ :

$$A_{21}B_{11} + A_{22}B_{21} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 2 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 4 & 0 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 3 & 2 \end{pmatrix},$$

ce qui est bien le bloc annoncé. Attention à l'ordre des facteurs : les blocs sont des matrices, qui ne commutent pas, et qui n'ont même pas ici des tailles permettant de les multiplier dans l'autre sens.
Un cas particulier utile : si $M$ et $N$ sont deux matrices carrées triangulaires supérieures par blocs, découpées de la même façon,

$$M = \begin{pmatrix} A & C \\ 0 & D \end{pmatrix}, \quad N = \begin{pmatrix} A' & C' \\ 0 & D' \end{pmatrix}, \qquad \text{alors} \qquad MN = \begin{pmatrix} AA' & AC' + CD' \\ 0 & DD' \end{pmatrix}.$$

Le produit reste triangulaire par blocs, et les blocs diagonaux se multiplient entre eux.

**Exemple 4.38 (Transposition par blocs)**

Rien n'oblige les blocs à être carrés, ni les découpages des lignes et des colonnes à coïncider. Prenons $M \in \mathcal{M}_{3,4}(\mathbb{K})$, dont on découpe les lignes en $1+2$ et les colonnes en $1+3$ :

$$M = \left(\begin{array}{c|ccc} 1 & 2 & 3 & 4 \\ \hline 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{array}\right) = \begin{pmatrix} A & C \\ B & D \end{pmatrix},$$

où $A$ est de taille $1 \times 1$, $C$ de taille $1 \times 3$, $B$ de taille $2 \times 1$ et $D$ de taille $2 \times 3$. La transposée est de taille $4 \times 3$, avec les lignes découpées en $1+3$ et les colonnes en $1+2$ :

$$M^\top = \left(\begin{array}{c|cc} 1 & 5 & 9 \\ \hline 2 & 6 & 10 \\ 3 & 7 & 11 \\ 4 & 8 & 12 \end{array}\right) = \begin{pmatrix} A^\top & B^\top \\ C^\top & D^\top \end{pmatrix}.$$

Deux choses se produisent en même temps : chaque bloc est transposé, et les blocs hors de la diagonale échangent leurs positions. C'est indispensable pour que les tailles soient cohérentes : $C$ est de taille $1 \times 3$, donc $C^\top$ est de taille $3 \times 1$ et ne peut se placer qu'en bas à gauche.
En particulier, pour une matrice carrée, la transposée d'une matrice triangulaire supérieure par blocs est triangulaire inférieure par blocs, et la transposée d'une matrice diagonale par blocs est diagonale par blocs.

**Proposition 4.39 (Déterminant par blocs)**

Soit $M$ une matrice carrée triangulaire par blocs :

$$M = \begin{pmatrix} A_1 & * & \cdots & * \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & A_r \end{pmatrix},$$

où les $A_i$ sont des matrices carrées. Alors

$$\det M = \det A_1 \cdot \det A_2 \cdots \det A_r.$$

**Remarque (Le cas triangulaire inférieur)**

Le même résultat vaut pour une matrice triangulaire *inférieure* par blocs :

$$M = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ * & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ * & \cdots & * & A_r \end{pmatrix} \qquad \Longrightarrow \qquad \det M = \det A_1 \cdot \det A_2 \cdots \det A_r.$$

En effet, $M^\top$ est alors triangulaire *supérieure* par blocs, de blocs diagonaux $A_1^\top, \ldots, A_r^\top$, comme on l'a vu plus haut. La proposition s'y applique, et comme une matrice et sa transposée ont même déterminant :

$$\det M = \det M^\top = \det A_1^\top \cdots \det A_r^\top = \det A_1 \cdots \det A_r.$$

Une matrice diagonale par blocs relève évidemment des deux cas.

**Exemple 4.40 (Un déterminant par blocs)**

Considérons

$$M = \left(\begin{array}{cc|cc} 1 & 2 & 9 & 0 \\ 3 & 4 & 1 & 2 \\ \hline 0 & 0 & 5 & 6 \\ 0 & 0 & 7 & 8 \end{array}\right).$$

Elle est triangulaire supérieure par blocs, avec $A_1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ et $A_2 = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$. Donc

$$\det M = \det A_1 \cdot \det A_2 = (4-6)(40-42) = (-2) \times (-2) = 4.$$

En particulier $M$ est inversible, sans qu'il ait fallu développer un déterminant $4 \times 4$. Le bloc en haut à droite n'intervient pas dans le calcul.
Cette formule se combine avec la forme de la matrice d'un endomorphisme dans une base adaptée à un sous-espace stable (voir la proposition page 12) : si $E = F \oplus G$ avec $F$ stable par $u$, la matrice de $u$ dans une base adaptée est $\begin{pmatrix} A & C \\ 0 & D \end{pmatrix}$ avec $A = \mathrm{Mat}_{\mathcal{B}_F}(u|_F)$, donc

$$\det u = \det(u|_F) \cdot \det D.$$

**Remarque**

L'hypothèse « triangulaire par blocs » ne peut pas être supprimée. Pour une matrice découpée en quatre blocs quelconques, il n'existe pas de formule analogue à celle du déterminant $2 \times 2$ : en général,

$$\det \begin{pmatrix} A & C \\ B & D \end{pmatrix} \neq \det A \cdot \det D - \det C \cdot \det B.$$

Par exemple, avec des blocs $2 \times 2$, prenons $A = D = 0$ et $C = B = I_2$ :

$$M = \left(\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right).$$

Le membre de droite vaut $0 \times 0 - 1 \times 1 = -1$. Or $M$ est la matrice de la permutation qui échange les vecteurs $e_1$ et $e_3$ d'une part, $e_2$ et $e_4$ d'autre part : c'est un produit de deux transpositions, donc $\det M = 1$.



# 4

# Compléments d'algèbre linéaire

## IV. Trace

**Définition 4.41 (Trace d'une matrice)**

Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$. La *trace* de $A$ est

$$\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

**Proposition 4.42 (Propriétés de la trace)**

Soient $A, B \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$.

1. **Linéarité :** $\operatorname{tr}(\lambda A + B) = \lambda \operatorname{tr}(A) + \operatorname{tr}(B)$.
2. $\operatorname{tr}(A^\top) = \operatorname{tr}(A)$.
3. Si $A$ est diagonale par blocs $A = \operatorname{diag}(A_1, \dots, A_r)$, alors $\operatorname{tr}(A) = \operatorname{tr}(A_1) + \cdots + \operatorname{tr}(A_r)$.

**Démonstration**

Notons $A = (a_{ij})$ et $B = (b_{ij})$.
*Linéarité.* Le coefficient d'indice $(i, i)$ de $\lambda A + B$ vaut $\lambda a_{ii} + b_{ii}$. En sommant sur $i$ :

$$\operatorname{tr}(\lambda A + B) = \sum_{i=1}^{n} (\lambda a_{ii} + b_{ii}) = \lambda \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = \lambda \operatorname{tr}(A) + \operatorname{tr}(B).$$

*Transposition.* Le coefficient d'indice $(i, j)$ de $A^\top$ est $a_{ji}$. Sur la diagonale, $i = j$ : la transposition ne change aucun coefficient diagonal, donc $\operatorname{tr}(A^\top) = \operatorname{tr}(A)$.
*Matrices diagonales par blocs.* Les coefficients diagonaux de $A$ sont exactement, dans cet ordre, ceux de $A_1$, puis ceux de $A_2$, etc. : les blocs situés hors de la diagonale n'apportent aucun coefficient diagonal. En découpant la somme selon les blocs, $\operatorname{tr}(A) = \operatorname{tr}(A_1) + \cdots + \operatorname{tr}(A_r)$. □

**Exemple 4.43 (Calculs de trace)**

Soit

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \qquad \text{de sorte que} \qquad \operatorname{tr}(A) = 1 + 5 + 9 = 15.$$

Seuls les coefficients diagonaux comptent : les six autres n'interviennent pas, et $A^\top$ a la même trace.
La linéarité permet de conclure sans écrire la matrice. Par exemple, comme $\operatorname{tr}(I_3) = 3$ :

$$\operatorname{tr}(A - 5I_3) = \operatorname{tr}(A) - 5\operatorname{tr}(I_3) = 15 - 15 = 0.$$

Retrancher $\frac{\operatorname{tr}(A)}{3} I_3$ ramène donc la trace à zéro.

Enfin, pour une matrice diagonale par blocs, il suffit d'ajouter les traces des blocs :

$$N = \left(\begin{array}{cc|cc} 1 & 2 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ \hline 0 & 0 & 5 & 6 \\ 0 & 0 & 7 & 8 \end{array}\right), \qquad \operatorname{tr}(N) = (1+4) + (5+8) = 18.$$

**Remarque (La trace est une forme linéaire)**

La première propriété dit exactement que

$$\operatorname{tr} : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$$

est une application linéaire à valeurs dans $\mathbb{K}$, c'est-à-dire une *forme linéaire* sur $\mathcal{M}_n(\mathbb{K})$. Elle n'est pas nulle, puisque $\operatorname{tr}(I_n) = n \neq 0$.
Son noyau

$$\mathcal{H}_0 = \{M \in \mathcal{M}_n(\mathbb{K}) \mid \operatorname{tr}(M) = 0\}$$

est donc un hyperplan de $\mathcal{M}_n(\mathbb{K})$ : c'est un sous-espace vectoriel de dimension $n^2 - 1$. Le calcul de l'exemple précédent se lit alors comme une décomposition : pour $M \in \mathcal{M}_n(\mathbb{K})$,

$$M = \frac{\operatorname{tr}(M)}{n} I_n + \left(M - \frac{\operatorname{tr}(M)}{n} I_n\right),$$

où le second terme est de trace nulle. La somme est directe, car une matrice $\lambda I_n$ de trace nulle vérifie $n\lambda = 0$, donc $\lambda = 0$. Ainsi

$$\mathcal{M}_n(\mathbb{K}) = \operatorname{Vect}(I_n) \oplus \mathcal{H}_0,$$

et l'on retrouve $\dim \mathcal{H}_0 = n^2 - 1$. C'est la forme linéaire notée $s$ dans la première section qui apparaissait déjà là : le sous-espace $\mathcal{S}_0$ y était l'intersection de $\mathcal{H}_0$ avec l'espace des matrices symétriques.

**Théorème 4.44**

Pour toutes matrices $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,n}(\mathbb{K})$,

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

**Démonstration**

On calcule directement :

$$\operatorname{tr}(AB) = \sum_{i=1}^{n} (AB)_{ii} = \sum_{i=1}^{n} \sum_{k=1}^{p} a_{ik} b_{ki} = \sum_{k=1}^{p} \sum_{i=1}^{n} b_{ki} a_{ik} = \sum_{k=1}^{p} (BA)_{kk} = \operatorname{tr}(BA).$$

L'échange des sommes est licite car ce sont des sommes finies. □

**Corollaire 4.45 (Invariance par similitude)**

Deux matrices semblables ont même trace. En effet, si $B = P^{-1}AP$ alors

$$\operatorname{tr}(B) = \operatorname{tr}(P^{-1}(AP)) = \operatorname{tr}((AP)P^{-1}) = \operatorname{tr}(A).$$

**Exemple 4.46 (La trace d'un projecteur est son rang)**

Soit $E$ un $\mathbb{K}$-espace vectoriel de dimension finie $n$ et $p \in \mathcal{L}(E)$ un *projecteur*, c'est-à-dire un endomorphisme vérifiant $p \circ p = p$. Posons $r = \operatorname{rg} p = \dim(\operatorname{Im} p)$.
*Une décomposition adaptée à $p$.* Les sous-espaces $\operatorname{Im} p$ et $\operatorname{Ker} p$ sont stables par $p$. Mieux, on les connaît complètement : si $y \in \operatorname{Im} p$, il existe $x$ tel que $y = p(x)$, donc

$$p(y) = p\left(p(x)\right) = p(x) = y,$$

et l'endomorphisme induit $p|_{\operatorname{Im} p}$ est l'identité de $\operatorname{Im} p$ ; quant à $p|_{\operatorname{Ker} p}$, c'est l'endomorphisme nul. Par ailleurs $E = \operatorname{Im} p \oplus \operatorname{Ker} p$ : tout $x \in E$ s'écrit

$$x = p(x) + \left(x - p(x)\right),$$

avec $p(x) \in \operatorname{Im} p$ et $p\left(x - p(x)\right) = p(x) - p(x) = 0$, et si $y \in \operatorname{Im} p \cap \operatorname{Ker} p$ alors $y = p(y) = 0$ d'après ce qui précède.
*La matrice dans une base adaptée.* Prenons une base $\mathcal{B}$ adaptée à cette décomposition. Les deux sous-espaces étant stables, la matrice de $p$ est diagonale par blocs, les blocs étant ceux des endomorphismes induits :

$$\operatorname{Mat}_{\mathcal{B}}(p) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

où le bloc nul est de taille $(n-r) \times (n-r)$. Notons $J_r$ cette matrice ; sa trace vaut $r$.
*Conclusion.* Soit $\mathcal{B}'$ une base *quelconque* de $E$. Les matrices $\operatorname{Mat}_{\mathcal{B}'}(p)$ et $J_r$ représentent le même endomorphisme dans deux bases : elles sont semblables, donc de même trace. Ainsi

$$\operatorname{tr}\left(\operatorname{Mat}_{\mathcal{B}'}(p)\right) = r = \operatorname{rg} p \qquad \text{pour toute base } \mathcal{B}'.$$

Autrement dit, le rang d'un projecteur se lit sur la trace de n'importe laquelle de ses matrices, sans avoir à chercher une base adaptée.
Par exemple, l'endomorphisme de $\mathbb{R}^2$ de matrice $A = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ dans la base canonique vérifie $A^2 = A$ : c'est un projecteur, et $\operatorname{tr}(A) = 1$. Son rang vaut donc 1, ce que confirme la lecture de $A$, dont les deux colonnes sont égales.

**Définition 4.47 (Trace d'un endomorphisme)**

Soit $u \in \mathcal{L}(E)$ avec $E$ de dimension finie. La *trace de $u$* est définie par $\operatorname{tr}(u) = \operatorname{tr}(\operatorname{Mat}_{\mathcal{B}}(u))$, où $\mathcal{B}$ est une base quelconque de $E$. Cette définition ne dépend pas du choix de $\mathcal{B}$ grâce à l'invariance par similitude.

**Proposition 4.48 (Propriétés de la trace d'un endomorphisme)**

Soient $u, v \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$.

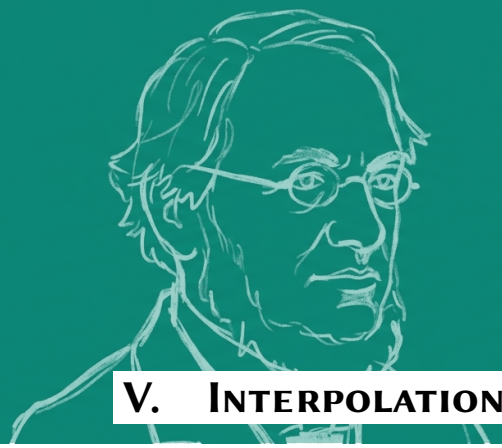
1. **Linéarité :** $\operatorname{tr}(\lambda u + v) = \lambda \operatorname{tr}(u) + \operatorname{tr}(v)$.
2. $\operatorname{tr}(u \circ v) = \operatorname{tr}(v \circ u)$.

**Démonstration**

La linéarité découle de celle de la trace matricielle et de la linéarité de l'application $u \mapsto \mathrm{Mat}_{\mathcal{B}}(u)$. Pour la seconde propriété, on fixe une base $\mathcal{B}$ et on écrit :

$$\mathrm{tr}(u \circ v) = \mathrm{tr}(\mathrm{Mat}_{\mathcal{B}}(u) \cdot \mathrm{Mat}_{\mathcal{B}}(v)) = \mathrm{tr}(\mathrm{Mat}_{\mathcal{B}}(v) \cdot \mathrm{Mat}_{\mathcal{B}}(u)) = \mathrm{tr}(v \circ u).$$

□



# 4

# Compléments d'algèbre linéaire

## V. Interpolation de Lagrange

Dans toute cette section, on fixe $n+1$ scalaires *distincts* $a_0, a_1, \dots, a_n \in \mathbb{K}$.

**Théorème 4.49 (Interpolation de Lagrange)**

Soient $a_0, \dots, a_n \in \mathbb{K}$ distincts et $b_0, \dots, b_n \in \mathbb{K}$. Il existe un unique polynôme $P \in \mathbb{K}_n[X]$ (de degré au plus $n$) tel que

$$\forall i \in \{0, \dots, n\}, \quad P(a_i) = b_i.$$

**Démonstration**

Considérons l'application d'évaluation :

$$\Phi : \mathbb{K}_n[X] \to \mathbb{K}^{n+1}, \quad P \mapsto (P(a_0), P(a_1), \dots, P(a_n)).$$

C'est une application linéaire entre deux espaces de même dimension $n+1$. Il suffit de montrer que $\Phi$ est injective. Soit $P \in \operatorname{Ker}\Phi$ : alors $P$ s'annule en les $n+1$ points distincts $a_0, \dots, a_n$ et $\deg P \leq n$. Un polynôme de degré au plus $n$ ayant $n+1$ racines est le polynôme nul. Donc $\operatorname{Ker}\Phi = \{0\}$ et $\Phi$ est un isomorphisme. □

**Exemple 4.50 (Une interpolation à la main)**

Cherchons le polynôme $P$ de degré au plus 2 tel que

$$P(-1) = 3, \qquad P(0) = 1, \qquad P(1) = 1.$$

Le théorème d'interpolation de Lagrange en garantit l'existence et l'unicité, les trois points $-1$, 0 et 1 étant distincts. Écrivons $P = \gamma + \beta X + \alpha X^2$ ; les trois conditions donnent le système

$$\begin{cases} \gamma - \beta + \alpha = 3, \\ \gamma = 1, \\ \gamma + \beta + \alpha = 1. \end{cases}$$

La deuxième équation donne $\gamma = 1$. En soustrayant les deux autres, on obtient $2\beta = -2$, donc $\beta = -1$, puis $\alpha = 1$. Ainsi

$$P = X^2 - X + 1,$$

ce que l'on vérifie : $P(-1) = 3$, $P(0) = 1$ et $P(1) = 1$.
La contrainte sur le degré est essentielle pour l'unicité. Sans elle, il y a une infinité de solutions : pour tout $\lambda \in \mathbb{K}$, le polynôme $P + \lambda(X+1)X(X-1)$ prend encore les valeurs voulues, puisque le second terme s'annule en $-1$, 0 et 1.

**Définition 4.51 (Polynômes de Lagrange)**

Les *polynômes d'interpolation de Lagrange* associés aux points $a_0, \dots, a_n$ sont les polynômes

$$L_i = \prod_{\substack{j=0 \\ j \neq i}}^{n} \frac{X - a_j}{a_i - a_j}, \qquad i \in \{0, \dots, n\}.$$

On vérifie immédiatement que $L_i(a_j) = \delta_{ij}$ (symbole de Kronecker).

**Proposition 4.52 (Base de Lagrange)**

La famille $(L_0, L_1, \dots, L_n)$ est une base de $\mathbb{K}_n[X]$, appelée *base de Lagrange* associée aux points $a_0, \dots, a_n$. La décomposition d'un polynôme $P \in \mathbb{K}_n[X]$ dans cette base est :

$$P = \sum_{i=0}^{n} P(a_i)\, L_i.$$

**Démonstration**

La famille $(L_0, \dots, L_n)$ contient $n+1$ polynômes de $\mathbb{K}_n[X]$, espace de dimension $n+1$. Il suffit de montrer qu'elle est libre. Si $\sum_{i=0}^{n} \lambda_i L_i = 0$, alors en évaluant en $a_k$ on obtient $\lambda_k = 0$ pour tout $k$. La formule de décomposition se vérifie en constatant que $P$ et $\sum P(a_i)L_i$ sont deux polynômes de degré au plus $n$ qui coïncident aux $n+1$ points $a_0, \dots, a_n$. □

**Exemple 4.53 (Décomposition dans la base de Lagrange)**

Reprenons les points $a_0 = -1$, $a_1 = 0$ et $a_2 = 1$. Les polynômes de Lagrange associés sont

$$L_0 = \frac{(X-0)(X-1)}{(-1-0)(-1-1)} = \frac{X^2 - X}{2},$$
$$L_1 = \frac{(X+1)(X-1)}{(0+1)(0-1)} = 1 - X^2,$$
$$L_2 = \frac{(X+1)(X-0)}{(1+1)(1-0)} = \frac{X^2 + X}{2}.$$

On vérifie bien que chacun vaut 1 en son point et 0 aux deux autres.
Décomposons $P = X^2 - X + 1$. Ses valeurs aux trois points sont 3, 1 et 1, calculées plus haut. Comme $P \in \mathbb{K}_2[X]$, ses coordonnées dans la base de Lagrange sont ses valeurs aux points $a_i$, c'est-à-dire $P = \sum_{i=0}^{2} P(a_i)L_i$ (proposition précédente), d'où directement

$$P = 3L_0 + L_1 + L_2,$$

sans aucun système à résoudre. Vérifions :

$$3 \cdot \frac{X^2 - X}{2} + (1 - X^2) + \frac{X^2 + X}{2} = \frac{4X^2 - 2X}{2} + 1 - X^2 = 2X^2 - X + 1 - X^2 = X^2 - X + 1.$$

Toute la différence avec l'exemple précédent est là : dans la base canonique, trouver les coordonnées demande de résoudre un système ; dans la base de Lagrange, les coordonnées *sont* les valeurs de $P$ aux points $a_i$.

**Remarque (La somme des polynômes de Lagrange)**

En appliquant la formule de décomposition au polynôme constant $P = 1$, qui est bien de degré au plus $n$ et vaut 1 en chaque $a_i$, on obtient

$$\sum_{i=0}^{n} L_i = 1.$$

Sur l'exemple précédent :

$$\frac{X^2 - X}{2} + (1 - X^2) + \frac{X^2 + X}{2} = X^2 + 1 - X^2 = 1.$$

Le même procédé appliqué à $P = X$ donne $\sum_{i=0}^{n} a_i L_i = X$, et ainsi de suite pour chaque $X^k$ avec $k \leq n$.

**Remarque (Le rôle de l'hypothèse « $a_i$ distincts »)**

Tout ce qui précède s'effondre si deux des points coïncident.
D'abord, les polynômes $L_i$ ne sont même pas définis : le dénominateur $\prod_{j \neq i}(a_i - a_j)$ s'annule.
Ensuite, le problème d'interpolation lui-même perd ses deux propriétés. Supposons par exemple $a_0 = a_1$. Si $b_0 \neq b_1$, aucun polynôme ne convient, puisqu'il faudrait que $P$ prenne deux valeurs distinctes au même point : l'existence tombe. Et si $b_0 = b_1$, la condition d'indice 1 répète celle d'indice 0 : il ne reste que $n$ conditions pour $n + 1$ coefficients, et l'unicité tombe.
Concrètement, avec $a_0 = a_1 = 0$, $a_2 = 1$ et $b_0 = b_1 = b_2 = 1$, tous les polynômes

$$1 + \lambda(X^2 - X), \qquad \lambda \in \mathbb{K},$$

sont de degré au plus 2 et valent 1 en 0 et en 1.
C'est bien ce que dit la démonstration du théorème : l'application d'évaluation $\Phi$ n'est injective que parce qu'un polynôme non nul de degré au plus $n$ ne peut pas avoir $n + 1$ racines *distinctes*.

**Définition 4.54 (Matrice de Vandermonde)**

La *matrice de Vandermonde* associée aux scalaires $a_1, \ldots, a_n$ est

$$V(a_1, \ldots, a_n) = \begin{pmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{pmatrix}.$$

C'est la matrice de l'application d'évaluation $(P(a_1), \ldots, P(a_n))$ dans les bases canoniques.

**Théorème 4.55 (Déterminant de Vandermonde)**

$$\det V(a_1, \ldots, a_n) = \prod_{1 \leq i < j \leq n} (a_j - a_i).$$

En particulier, la matrice est inversible si et seulement si les $a_i$ sont deux à deux distincts.

**Démonstration**

On procède par récurrence sur $n$.
**Initialisation :** pour $n = 1$, $V(a_1) = (1)$ et $\det V = 1$, le produit vide valant 1.
**Hérédité :** supposons la formule vraie pour $n-1$ scalaires. On considère le déterminant de $V(a_1, \dots, a_n)$ comme un polynôme en $a_n$, en développant par rapport à la dernière ligne. On obtient un polynôme $Q(X)$ de degré au plus $n-1$ en $X = a_n$. En remplaçant $a_n$ par $a_i$ pour $i < n$, deux lignes de la matrice deviennent identiques, donc $Q(a_i) = 0$ pour $i = 1, \dots, n-1$. Ainsi $Q$ possède $n-1$ racines et est de degré au plus $n-1$, d'où :

$$Q(X) = c \prod_{i=1}^{n-1} (X - a_i).$$

Le coefficient dominant $c$ est le cofacteur de $a_n^{n-1}$ dans la dernière ligne, à savoir $\det V(a_1, \dots, a_{n-1})$. Par hypothèse de récurrence :

$$c = \prod_{1 \le i < j \le n-1} (a_j - a_i).$$

Donc

$$\det V(a_1, \dots, a_n) = Q(a_n) = \prod_{1 \le i < j \le n-1} (a_j - a_i) \cdot \prod_{i=1}^{n-1} (a_n - a_i) = \prod_{1 \le i < j \le n} (a_j - a_i).$$

□

**Remarque (Vandermonde et l'interpolation)**

La matrice de Vandermonde est celle de l'application d'évaluation $\Phi$ dans les bases canoniques. Le théorème précédent fournit donc une seconde démonstration du théorème d'interpolation : chercher $P = c_0 + c_1 X + \cdots + c_n X^n$ tel que $P(a_i) = b_i$, c'est résoudre le système linéaire

$$V(a_0, \dots, a_n) \begin{pmatrix} c_0 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} b_0 \\ \vdots \\ b_n \end{pmatrix},$$

qui admet une unique solution si et seulement si $\det V \neq 0$, c'est-à-dire si et seulement si les $a_i$ sont deux à deux distincts.
L'hypothèse de la section trouve ainsi sa traduction matricielle : si $a_i = a_j$ avec $i \neq j$, les lignes $i$ et $j$ de $V$ sont identiques, le déterminant est nul, et le système n'a pas de solution ou en a une infinité selon le second membre — exactement ce que l'on avait observé.
Sur l'exemple des points $-1$, $0$ et $1$ :

$$V(-1, 0, 1) = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \qquad \det V = (0+1)(1+1)(1-0) = 2 \neq 0,$$

et le système d'inconnues $(c_0, c_1, c_2)$ et de second membre $(3, 1, 1)$ est celui que nous avions résolu à la main, de solution $(c_0, c_1, c_2) = (1, -1, 1)$.

**Exemple 4.56 (Liberté de la famille des exponentielles)**

Soient $a_1, \dots, a_n$ des réels deux à deux distincts. Montrons que la famille des fonctions

$$f_i : t \mapsto \mathrm{e}^{a_i t}, \qquad i \in \{1, \dots, n\},$$

est libre dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$.
Soient $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tels que $\lambda_1 f_1 + \cdots + \lambda_n f_n$ soit la fonction nulle, c'est-à-dire

$$\forall t \in \mathbb{R}, \quad \sum_{i=1}^{n} \lambda_i \mathrm{e}^{a_i t} = 0.$$

Posons $x_i = \mathrm{e}^{a_i}$, de sorte que $\mathrm{e}^{a_i t} = x_i^t$. En évaluant la relation précédente en $t = 0, 1, \dots, n-1$, on obtient

$$\forall j \in \{0, \dots, n-1\}, \quad \sum_{i=1}^{n} \lambda_i \, x_i^j = 0.$$

C'est un système linéaire de $n$ équations aux $n$ inconnues $\lambda_1, \dots, \lambda_n$, dont la matrice a pour coefficient $x_i^j$ en ligne $j$ et colonne $i$ : c'est la transposée de $V(x_1, \dots, x_n)$.
Or les $x_i$ sont deux à deux distincts, puisque exp est injective sur $\mathbb{R}$ et que les $a_i$ le sont. Par la formule du déterminant de Vandermonde (théorème précédent) :

$$\det V^\top(x_1, \dots, x_n) = \det V(x_1, \dots, x_n) = \prod_{1 \le i < k \le n} (x_k - x_i) \neq 0.$$

Le système est de Cramer et homogène : sa seule solution est $\lambda_1 = \cdots = \lambda_n = 0$. La famille est donc libre.