Algèbre linéaire - Résumé

1. Définition

Définition 1. *On appelle* $\mathbb{K}$-espace vectoriel *un triplet* $(E, +, \cdot)$ *où* E *est un ensemble non vide, dont les éléments sont appelés* vecteurs, $+$ *une loi de composition interne, i.e. une application de* $E \times E$ *dans* E *qui, à un couple de vecteurs* (x, y)*, associe leur somme* $x + y$ *et* $\cdot$*, une loi externe, ici une application de* $\mathbb{K} \times E$ *dans* E *qui, à un couple* (λ, x)*, associe le vecteur* $\lambda \cdot x$*. L'ensemble* $\mathbb{K}$ *est muni lui aussi de deux lois de composition internes commutatives (l'addition et la multiplication) et forme ce que l'on appelle un* corps *(tout élément non nul y est inversible) qui, dans le cadre du programme, sera toujours* $\mathbb{R}$ *ou* $\mathbb{C}$*. Les éléments de* $\mathbb{K}$ *sont appelés* scalaires*. On suppose par ailleurs que les deux opérations ont les propriétés suivantes :*

(i) $\forall (x, y, z) \in E^3 \colon x + (y + z) = (x + y) + z$;

(ii) $\exists\, 0_E \in E, \forall x \in E \colon x + 0_E = x$;

(iii) $\forall (x, y) \in E^2 \colon x + y = y + x$;

(iv) $\forall x \in E, \exists y \in E \colon x + y = 0_E$;

(v) $\forall (\lambda, x, y) \in \mathbb{K} \times E^2 \colon \lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$;

(vi) $\forall (\lambda, \mu, x) \in \mathbb{K}^2 \times E \colon (\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$;

(vii) $\forall (\lambda, \mu, x) \in \mathbb{K}^2 \times E \colon \lambda \cdot (\mu \cdot x) = \lambda\mu \cdot x$;

(viii) $\forall x \in E \colon 1 \cdot x = x$.

Cette définition est de type axiomatique. Elle pourrait être (un peu) différente, mais elle est conçue pour que toutes les propriétés calculatoires des espaces vectoriels, notamment celles qui sont « évidentes », puissent se déduire des propriétés qu'elle impose. Donnons un exemple.

Proposition 1. *Soit* E *un* $\mathbb{K}$*-espace vectoriel et* $(\lambda, x) \in \mathbb{K} \times E$*. Alors,*

$$\lambda \cdot x = 0_E \iff \lambda = 0_{\mathbb{K}} \text{ ou } x = 0_E.$$

Preuve. $\boxed{\Leftarrow}$ Soit $x \in E$. D'après la propriété *(v)*, on a $\lambda \cdot 0_E = \lambda \cdot (0_E + 0_E) = \lambda \cdot 0_E + \lambda \cdot 0_E$, d'où $\lambda \cdot 0_E = 0_E$. De même, la propriété *(vi)* permet d'écrire que, si $x \in E$, alors $0_{\mathbb{K}} \cdot x = (0_{\mathbb{K}} + 0_{\mathbb{K}}) \cdot x = 0_{\mathbb{K}} \cdot x + 0_{\mathbb{K}} \cdot x$, d'où $0_{\mathbb{K}} \cdot x = 0_E$. Si vous êtes convaincu, c'est que vous n'avez pas remarqué que, dans les deux raisonnements, on a utilisé une simplification du type $x + z = x \Rightarrow z = 0_E$ et que nous n'avons pas démontré que c'était licite. De fait, par *(iv)*, soit $y \in E$ tel que $x + y = 0_E$. Alors, en ajoutant y à gauche et à droite de $x + z = x$, le membre de gauche devient $(x+z)+y = (z+x)+y = z+(x+y) = z+0_E = z$ (dans l'ordre, par *(iii)*, *(i)* et *(ii)*) et celui de droite devient 0_E, d'où le résultat.
$\boxed{\Rightarrow}$ Si $\lambda \cdot x = 0_E$ et $\lambda \neq 0_{\mathbb{K}}$, alors, en multipliant par λ^{-1}, il vient $\lambda^{-1} \cdot (\lambda \cdot x) = (\lambda^{-1}\lambda) \cdot x = 1 \cdot x = x$ par *(vii)* et *(viii)* pour le membre de gauche et $\lambda^{-1} \cdot 0_E = 0_E$ d'après $\Leftarrow$ pour le membre de droite. □

Exemple 1. $\mathbb{K}^n$, $\mathbb{K}_n[X]$, $\mathbb{K}[X]$, $\mathscr{M}_{p,q}(\mathbb{K})$, $\mathcal{F}(A, F)$ (ensemble des fonctions définies sur une partie A non vide d'un espace vectoriel normé et à valeurs dans un autre e.v.n. F), $\mathcal{C}^p(I, \mathbb{R})$, $\big\{(u_n)_n \in \mathbb{C}^n \,;\, \sum u_n \text{ ACV}\big\}$, $\mathcal{L}^1(I)$, sont des exemples courants d'espaces vectoriels.

2. Familles de vecteurs

Définition 2. *Soient* E *un e.v. et* $\mathcal{F} \subset E$ *une famille de vecteurs de* E*.*

(i) Une combinaison linéaire *de vecteurs de* E *est une somme* $\displaystyle\sum_{x \in \mathcal{F}} \lambda_x x$*, où* $\mathcal{F}$ *est une partie finie de* E *et* $(\lambda_x)_{x \in \mathcal{F}} \in \mathbb{K}^{\mathcal{F}}$*. Si* $\mathcal{F}$ *est infinie, on peut toujours écrire* $\displaystyle\sum_{x \in \mathcal{F}} \lambda_x x$*, où il est cette fois entendu que* $(\lambda_x)_{x \in \mathcal{F}} \in \mathbb{K}^{(\mathcal{F})}$*, c'est-à-dire que* $\#\{x \in \mathcal{F} \,;\, \lambda_x \neq 0_{\mathbb{K}}\}$ *est fini.*

(ii) Que $\mathcal{F}$ *soit fini ou non, on note* $\mathrm{Vect}(\mathcal{F})$ *l'ensemble des combinaisons linéaires de vecteurs de* $\mathcal{F}$*.*

(iii) $\mathcal{F} \subset E$ *est* génératrice *si tout vecteur de* E *s'écrit comme combinaison linéaire de vecteurs de* $\mathcal{F}$*.*

(iv) E *est de dimension finie s'il admet une famille génératrice finie.*

(v) $\mathcal{F} \subset E$ *est libre si, et seulement si (on donne les deux écritures)*

$$\forall (\lambda_x)_{x \in \mathcal{F}} \in \mathbb{K}^{(\mathcal{F})} \colon \sum_{x \in \mathcal{F}} \lambda_x x = 0_E \Longrightarrow \big[\forall x \in \mathcal{F} \colon \lambda_x = 0_{\mathbb{K}}\big];$$

$$\forall n \in \mathbb{N}^*, \forall (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{K}^n \colon \sum_{i=1}^{n} \lambda_i x_i = 0_E \Longrightarrow \big[\forall i \in [\![1, n]\!] \colon \lambda_i = 0_{\mathbb{K}}\big].$$

(vi) $\mathcal{F} \subset E$ *est* liée *si elle n'est pas libre.*

(vii) Une base de E *est une famille de* E *libre et génératrice.*

L'écriture abstraite $\sum_{x\in\mathcal{F}} \lambda_x x$ se justifie par le fait qu'à l'exception des bases en dimension finie (à cause des matrices d'applications linéaires), l'ordre n'a pas d'importance : une somme de vecteurs ne dépend pas de l'ordre dans lequel l'on procède à la somme, une famille de vecteurs est libre ou génératrice indépendamment de son indiciation, comme une somme de s.e.v. est indépendante de l'ordre dans lequel on somme les sous-espaces. On peut aussi bien sûr utiliser une écriture indiciée : on choisit une famille finie $(x_i)_{1\leqslant i\leqslant n} \in \mathcal{F}^n$ et l'on exprime une combinaison linéaire sous la forme $x = \sum_{i=1}^{n} \lambda_i x_i$ avec $(\lambda_i)_{1\leqslant i\leqslant n} \in \mathbb{K}^n$.

Notons que, par convention, on peut aussi définir des combinaisons linéaires de 0 vecteur par $\sum_{x\in\varnothing} x = 0_E$. Retenir qu'une combinaison linéaire est toujours finie (ne **jamais** écrire de séries sans structure d'e.v.n.) et que, pour montrer qu'une famille infinie est libre, il suffit d'en considérer une sous-famille finie arbitraire.

Une sous-famille d'une famille libre est libre. Une sur-famille d'une famille liée (resp. génératrice) est liée (resp. génératrice). Un singleton est libre si, et seulement si, le seul vecteur qui le compose est non nul. Une famille de deux vecteurs est liée si, et seulement si, ces deux vecteurs sont colinéaires. À partir de trois vecteurs, on ne peut en général déterminer sans calcul si la famille est libre ou non. Une famille est liée si, et seulement si, elle contient un vecteur s'exprimant comme combinaison linéaire des autres (en particulier, une famille contenant le vecteur nul est liée), mais ce n'est pas forcément le cas de tous les vecteurs qu'elle contient. Ainsi, si (u, v) est libre, $(u, 2u, v)$ est liée sans que v s'exprime comme combinaison linéaire de u et de $2u$. Si $\mathcal{L}$ est une famille libre de E et si $x \in E$, $\mathcal{L} \cup \{x\}$ est libre si, et seulement si, $x \notin \operatorname{Vect}(\mathcal{L})$. La théorie de la dimension repose sur le lemme suivant :

Lemme 1. *Si $\mathcal{G} \subset E$ est de cardinal n, et si $\mathcal{F} \subset \operatorname{Vect}(\mathcal{G})$ est telle que $\#\mathcal{F} \geqslant n+1$, alors $\mathcal{F}$ est liée.*

Démonstration. On procède par récurrence sur n. La propriété se montre facilement pour $n = 1$. Admettons qu'elle soit vraie à l'ordre $n-1$ et considérons $\mathcal{G} = \{e_1, e_2, \dots, e_n\}$ et $(x_1, x_2, \dots, x_{n+1}) \in \operatorname{Vect}(\mathcal{G})^{n+1}$ (comme toute sur-famille d'une famille liée est liée, il suffit de considérer une famille de $n+1$ vecteurs). On pose $\mathcal{G}' = \{e_1, e_2, \dots, e_{n-1}\}$.

Si $(x_1, x_2, \dots, x_{n+1}) \in \operatorname{Vect}\left(\mathcal{G}'\right)^{n+1}$, on peut appliquer l'hypothèse de récurrence, qui montre que $(x_1, x_2, \dots, x_{n+1})$ est liée. Sinon, on peut écrire

$$\forall i \in [\![1, n+1]\!],\ \exists \big(\lambda_j^{(i)}\big)_{1\leqslant j\leqslant n} \in \mathbb{K}^n : x_i = \sum_{j=1}^{n} \lambda_j^{(i)} e_j,$$

où l'un au moins des scalaires $\lambda_n^{(i)}$ est non nul. Sans perte de généralité, supposons que $\lambda_n^{(n+1)} \neq 0_{\mathbb{K}}$. Alors,

$$\forall i \in [\![1, n]\!] : x_i' = x_i - \frac{\lambda_n^{(i)}}{\lambda_n^{(n+1)}} x_{n+1} = \sum_{j=1}^{n-1} \left(\lambda_j^{(i)} - \frac{\lambda_n^{(i)} \lambda_j^{(n+1)}}{\lambda_n^{(n+1)}} \right) e_j + \in \mathcal{G}'.$$

On peut appliquer l'hypothèse de récurrence à la famille $(x_i')_{1\leqslant i\leqslant n}$, qui est donc liée, et il existe ainsi, pour $i \in [\![1, n]\!]$, des réels μ_i non tous nuls tels que

$$0_E = \sum_{i=1}^{n} \mu_i x_i' = \sum_{i=1}^{n} \mu_i x_i - \left(\frac{1}{\lambda_n^{(n+1)}} \sum_{i=1}^{n} \mu_i \lambda_n^{(i)} \right) x_{n+1},$$

donc la famille $(x_1, x_2, \dots, x_{n+1})$ est liée, ce qui clôt la récurrence.

Corrélativement, deux bases d'un espace vectoriel de dimension finie ont même cardinal, ce qui permet de définir la dimension d'un e.v. de dimension finie comme le cardinal de n'importe laquelle de ses bases. Si $\dim E = n$, toute famille d'au moins $n+1$ vecteurs est liée.

Théorème 1. Théorème de la base incomplète (énoncé général).
Soient E un $\mathbb{K}$-e.v. de dimension finie, $\mathcal{L}$ une famille libre de vecteurs de E et $\mathcal{G}$ une famille génératrice de vecteurs de E. Alors, il existe une base $\mathcal{B}$ de E telle que $\mathcal{L} \subset \mathcal{B} \subset \mathcal{L} \cup \mathcal{G}$. Autrement dit, on peut compléter la famille libre $\mathcal{L}$ en une base $\mathcal{B}$ en piochant dans la famille génératrice $\mathcal{G}$.

La démonstration consiste à ajouter des vecteurs de $\mathcal{G} \setminus \operatorname{Vect}(\mathcal{L})$ à $\mathcal{L}$ tant que la famille ainsi construite n'est pas génératrice de E, le processus s'arrêtant en un temps fini en vertu du lemme 1.

Le cas $\mathcal{L} = \varnothing$ signifie que l'on peut extraire une base de toute famille génératrice ; le cas $\mathcal{G} = E$ assure que toute famille libre se complète en une base et, si $\mathcal{L} = \varnothing$ et $\mathcal{G} = E$, on obtient que tout espace de dimension finie admet une base.

Certains espaces ont une base particulière, dite *base canonique*. C'est le cas de $\mathbb{K}^n$, de $\mathscr{M}_{p,n}(\mathbb{K})$, de $\mathbb{K}[X]$, de $\mathbb{K}_n[X]$ et de $\mathbb{K}^{(\mathcal{F})}$. Cette propriété ne s'étend pas à leurs s.e.v ; un plan de $\mathbb{R}^3$ ne possède pas, en général, de base canonique.

3. Sous-espaces

3.1. Comment montrer qu'un s.e.v. est un s.e.v.

Théorème et définition 1. *Un sous-espace (vectoriel) F d'un espace vectoriel E est une partie de $(E,+,\cdot)$ sur laquelle les deux lois induisent par restriction une structure d'espace vectoriel. Si $F\subset E$, F est un s.e.v. de E si, et seulement si :*

(i) F n'est pas vide ;

(ii) F est stable par combinaison linéaire.

En pratique, un e.v. contenant toujours le vecteur nul, on vérifie le plus souvent, pour montrer que $F\neq\varnothing$, que $0_E\in F$, ce qui est en général trivial. Il ne faut pas oublier de le mentionner car, comme le montre l'exemple suivant, cette histoire de non-vacuité peut êtredéterminante.

Exemple 2. Soient $A\in\mathscr{M}_{n,p}(\mathbb{K})$ et $B\in\mathbb{K}^n$, considéré comme un vecteur colonne. Soient les ensembles $S=\{X\in\mathbb{K}^p\,;\ AX=B\}$ et $F=\{X-Y\,;\ (X,Y)\in S^2\}$. Alors, il est immédiat que F est stable par combinaisons linéaires, mais F est vide si S l'est, *i.e.* si les équations du système $AX=B$ sont incompatibles, ce qui peut arriver. Notons que si $S\neq\varnothing$, alors $F=\{X\in\mathbb{K}^p\,;\ AX=0_{\mathbb{K}^n}\}=\operatorname{Ker}A$.

Pour montrer la stabilité par combinaison linéaire, on peut montrer d'un coup que, pour tout $(\lambda,\mu)\in\mathbb{K}^2$ et pour tout couple $(x,y)\in F^2$, alors $\lambda x+\mu y\in F$ (on peut même se contenter du cas $\mu=1$), ou séparer la stabilité par addition et par multiplication par un scalaire. Des solutions alternatives pertinentes sont d'exprimer F comme un espace vectoriel engendré par une certaine famille de vecteurs ou d'écrire F comme le noyau ou l'image d'une application linéaire connue.

Définition 3. *Si $\mathcal{F}\subset E$, le sous-espace engendré par $\mathcal{F}$ est le plus petit s.e.v. de E contenant $\mathcal{F}$. Il est constitué des combinaisons linéaires des éléments de $\mathcal{F}$. On le note* $\operatorname{Vect}(\mathcal{F})$.

Exemple 3. — $\mathcal{T}_n^{++}(\mathbb{K})$, espace des matrices triangulaires supérieures à diagonale nulle à coefficients dans $\mathbb{K}$, est un espace vectoriel car $\mathcal{T}_n^{++}(\mathbb{K})=\operatorname{Vect}(E_{i,j}\,;\ 1\leqslant i<j\leqslant n)$.
— $\mathcal{S}_n(\mathbb{K})$ est un s.e.v. de $\mathscr{M}_n(\mathbb{K})$ car $\mathcal{S}_n(\mathbb{K})=\operatorname{Ker}\big(T-\mathrm{id}_{\mathscr{M}_n(\mathbb{K})}\big)=E_1(T)$, espace propre associé à la valeur propre 1 de l'application de transposition T, qui est linéaire.
— $\left\{f\in\mathcal{C}([0,1],\mathbb{R})\,;\ \displaystyle\int_0^1 f=0_{\mathbb{R}}\right\}$ est un s.e.v. de $\mathcal{C}([0,1],\mathbb{R})$ car c'est le noyau de l'intégrale sur $[0,1]$.

Un sous-espace F de E est dit *non trivial* s'il est distinct des deux sous-espaces dits *triviaux*, à savoir $\{0_E\}$ et E. En dimension finie, cela se traduit par $0<\dim F<\dim E$ car, si F et G sont des s.e.v. de E et si $F\subset G$, alors $F=G\iff\dim F=\dim G$.

Exercice. Si F et G sont des s.e.v. de E, $F\cup G$ est un s.e.v. de E si, et seulement si, $F\subset G$ ou $G\subset F$.

3.2. Produit d'espaces vectoriels.

Théorème et définition 2. *Si E_1, $E_2,\ldots$, E_p sont des e.v. sur un même corps $\mathbb{K}$, on munit l'ensemble $E=E_1\times E_2\times\cdots\times E_p$ d'une structure de $\mathbb{K}$-e.v. en posant :*

(i) $\forall x=(x_1,x_2,\ldots,x_p)\in E$, $\forall y=(y_1,y_2,\ldots,y_p)\in E\colon x+y=(x_1+y_1,x_2+y_2,\ldots,x_p+y_p)$;

(ii) $\forall\lambda\in\mathbb{K}$, $\forall x=(x_1,x_2,\ldots,x_p)\in E\colon \lambda x=(\lambda x_1,\lambda x_2,\ldots,\lambda x_p)$.

Pour tout $i\in[\![1,p]\!]$, l'application $\theta_i\colon E_i\to E$ définie par $\theta_i(x)=\big(0_{E_1},\ldots,0_{E_{i-1}},x,0_{E_{i+1}},\ldots,0_{E_p}\big)$ est une application injective, dite injection canonique *de E_i dans E. On obtient une base de E par concaténation des images par les injections canoniques de bases des espaces E_i. Ainsi, si tous les E_i sont de dimension finie, alors, pour $n_i=\dim E_i$, on a $\dim E=n_1+n_2+\cdots+n_p$.*

L'ordre d'un produit n'est pas tout à fait indifférent : $E\times F$ n'est pas le même espace que $F\times E$, même s'ils sont isomorphes. La définition précédente est exprimée formellement. Voilà ce que cela donne pour $p=2$:

Si F et G sont des $\mathbb{K}$-e.v., et si $E = F \times G$, les éléments de E sont les couples (x, y) avec $x \in F$ et $y \in G$. On a $0_E = (0_F, 0_G)$, $\lambda(x_1, y_1) + \mu(x_2, y_2) = (\lambda x_1 + \mu x_2, \lambda y_1 + \mu y_2)$ et, si $(f_1, \dots, f_n)$ et $(g_1, \dots, g_m)$ sont des bases de F et de G respectivement, $\mathcal{B} = \big((f_1, 0_G), \dots, (f_n, 0_G), (0_F, g_1), \dots, (0_F, g_m)\big)$ forme une base de E. Attention : la famille $\{(f_i, g_j)\,;\ (i,j) \in [\![1,n]\!] \times [\![1,m]\!]\}$ n'est pas une base de E ! On a par exemple $(f_1, g_1) - (f_2, g_1) - (f_1, g_2) + (f_2, g_2) = 0_E$. La dimension du produit est la somme des dimensions et non leur produit. On pourra s'en souvenir en pensant à l'identité relativement évidente $\mathbb{K}^n \times \mathbb{K}^m \simeq \mathbb{K}^{n+m}$.

3.3. Somme de sous-espaces vectoriels.

Théorème et définition 3. *Soient E un e.v. et F_1, F_2,..., F_p des s.e.v. de E. Soit ϕ l'application définie sur $F_1 \times F_2 \times \cdots \times F_p$ l'application qui, à $(x_1, x_2, \dots, x_p)$, associe $x_1 + x_2 + \cdots + x_p$. Alors,*

(i) La somme $F_1 + F_2 + \cdots + F_p = \displaystyle\sum_{k=1}^{p} F_k$ *est, par définition,* $\mathrm{Im}(\phi) = \mathrm{Vect}\big(F_1 \cup F_2 \cup \cdots \cup F_p\big)$.

(ii) Si ϕ est injective, on dit que la somme est directe et on la note $F_1 \oplus F_2 \oplus \cdots \oplus F_p$ *ou* $\displaystyle\bigoplus_{k=1}^{p} F_k$.

(iii) Définition récursive de la somme directe : la somme est directe si, et seulement si,

$$F_1 + F_2 + \cdots + F_{p-1} = F_1 \oplus F_2 \oplus \cdots \oplus F_{p-1} \quad \text{et} \quad \big(F_1 + F_2 + \cdots + F_{p-1}\big) \cap F_p = \{0_E\}.$$

(iv) $\dim\big(F_1 + F_2 + \cdots + F_p\big) \leqslant \dim F_1 + \dim F_2 + \cdots + \dim F_p$ *avec égalité ssi la somme est directe.*

(v) Si la somme est directe, on en obtient une base (en dimension finie) par concaténation des bases des F_k. Une telle base est dite adaptée à la décomposition en somme directe.

(vi) Réciproquement, soient $\mathcal{L}$ une famille libre finie et une partition $\mathcal{L} = \mathcal{L}_1 \uplus \mathcal{L}_2 \uplus \cdots \uplus \mathcal{L}_p$. Alors, les espaces $F_k = \mathrm{Vect}(\mathcal{L}_k)$ *sont en somme directe.*

— Les s.e.v. $F_1, F_2, \dots, F_p$ sont en somme directe si, et seulement si, toute famille $(x_1, x_2, \dots, x_p)$ avec $x_i \in F_i \setminus \{0_E\}$ est libre.
— Les s.e.v. $F_1, F_2, \dots, F_p$ sont en somme directe si, et seulement si, pour tout $(x_1, x_2, \dots, x_p) \in F_1 \times F_2 \times \cdots F_p$, $x_1 + x_2 + \cdots + x_p = 0_E$ entraîne (et est alors équivalent à) $x_1 = x_2 = \cdots = x_p = 0_E$.
— Les s.e.v. $F_1, F_2, \dots, F_p$ sont en somme directe si, et seulement si, tout vecteur $z \in F_1 + F_2 + \cdots + F_p$ se décompose *de manière unique* sous la forme $z = x_1 + x_2 + \cdots + x_p$.
— Une somme de s.e.v., comme une somme de vecteurs, ne dépend pas de l'ordre dans lequel on considère la somme, le fait qu'elle soit directe ou non n'en dépend pas non plus. On pourra donc éventuellement utiliser des écritures abstraites comme $\displaystyle\sum_{F \in \mathcal{F}} F$ ou $\displaystyle\bigoplus_{F \in \mathcal{F}} F$, où $\mathcal{F}$ est une famille finie de s.e.v. de E.
— Toute somme extraite d'une somme directe est directe.
— Si, pour tout i, G_i est un s.e.v. de F_i et si les F_i sont en somme directe, alors les G_i le sont aussi.
— Si $\mathcal{F}$ et $\mathcal{G}$ sont des sous-ensembles de E, alors $\mathrm{Vect}(\mathcal{F} \cup \mathcal{G}) = \mathrm{Vect}(\mathcal{F}) + \mathrm{Vect}(\mathcal{G})$.
— Pour parler d'une somme de s.e.v., il faut évidemment que tous les s.e.v. soient des s.e.v. d'un *même* e.v. (alors que, pour le produit, la seule condition est que chaque e.v. du produit soit défini sur le même corps $\mathbb{K}$).
— Cas particulier de s.e.v. de dimension 1, appelés *droites vectorielles* : $\mathrm{Vect}(x_1, x_2, \dots, x_p) = \displaystyle\sum_{k=1}^{p} \mathrm{Vect}(x_k)$; la somme est directe si, et seulement si, la famille $(x_1, x_2, \dots, x_p)$ est libre.
— Cas particulier de deux sous-espaces de dimension finie supplémentaires : on dit que F et G, s.e.v. de E sont supplémentaires si $E = F \oplus G$. C'est le cas si deux des trois propositions suivantes sont vérifiée, la troisième l'étant alors automatiquement.

(i) $E = F + G$, *i.e.* tout vecteur de E s'écrit comme la somme d'un vecteur de F et d'un vecteur de G.

(ii) $F \cap G \subset \{0\}$, *i.e.* $F + G = F \oplus G$. De manière équivalente la décomposition $e = f + g$, si elle existe, est unique. Notons qu'en fait, $F \cap G = \{0\}$, l'inclusion réciproque étant toujours vérifiée car tout s.e.v. contient le vecteur nul.

(iii) $\dim E = \dim F + \dim G$.

Si F est un s.e.v. de E, il admet une infinité de supplémentaires, sauf si $F = E$ (unique supplémentaire $\{0_E\}$ ou si $F = \{0_E\}$ (unique supplémentaire E). Supposons $E = F \oplus G_1 = F \oplus G_2$. Si $G_1 \subset G_2$, alors $G_1 = G_2$; si G_1 est de dimension finie, alors G_2 également et $\dim G_1 = \dim G_2$. — La formule sur les dimensions est un cas particulier de la *formule de Graßmann* :

$$\dim(F+G) = \dim F + \dim G - \dim(F \cap G),$$

dont on notera l'analogie avec la formule de dénombrement $\#(F \cup G) = \#F + \#G - \#(F \cap G)$.
— La caractérisation de la supplémentation de deux s.e.v. en dimension finie s'étend à p sous-espaces et deux des trois propositions suivantes entraînent la troisième :

(i) $E = F_1 + F_2 + \cdots + F_p$ (ce qui est équivalent à la surjectivité de ϕ) ;

(ii) la somme $F_1 + F_2 + \cdots + F_p$ est directe (ce qui est équivalent à l'injectivité de ϕ) ;

(iii) $\dim E = \dim F_1 + \dim F_2 + \cdots + \dim F_p$ (ce qui est équivalent à $\dim \prod_{i=1}^{p} F_i = \dim E$).

— Extension de la définition récursive. Supposons que $E = \sum_{i=1}^{p} F_i$ et, pour $i \in [\![1,p]\!]$, $F_i = \sum_{j=1}^{m_i} G_j^{(i)}$, alors

$$E = \bigoplus_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant m_i}} G_j^{(i)} \iff \left[\forall i \in [\![1,p]\!] : F_i = \bigoplus_{j=1}^{m_i} G_j^{(i)} \quad \& \quad E = \bigoplus_{i=1}^{p} G_i \right].$$

— Une idée fausse : si une somme de p s.e.v. est directe, alors l'intersection des s.e.v. formant cette somme est réduite au vecteur nul, mais la réciproque est fausse pour $p \geqslant 3$. C'est la grande différence entre le cas d'une somme deux s.e.v. et le cas général : pour deux s.e.v., montrer que $F_1 + F_2 = F_1 \oplus F_2$ revient à montrer que $F_1 \cap F_2 = \{0_E\}$. Pour p sous-espaces, il faut montrer qu'une famille $(x_1, x_2, \ldots, x_p)$ de vecteurs non nuls avec $x_i \in F_i$ est libre.

3.4. **Hyperplans.** Il est utile de noter que si F est un s.e.v. de E et si $x \in E$, alors se présentent deux cas :
— si $x \in F$, alors $F + \mathrm{Vect}(x) = F$;
— si $x \notin F$, alors $F + \mathrm{Vect}(x) = F \oplus \mathrm{Vect}(x)$.

Théorème et définition 4. *Soient E un $\mathbb{K}$-e.v. et H un s.e.v. de E distinct de E. Alors, H est un* hyperplan *de E si l'une des conditions équivalentes suivantes est vérifiée :*

(i) il existe une droite vectorielle D telle que $E = H \oplus D$;

(ii) H est le noyau d'une forme linéaire non nulle $u \in \mathcal{L}(E, \mathbb{K})$. Cette forme linéaire est alors unique à un coefficient multiplicatif non nul près ;

(iii) $\dim H = \dim E - 1$ (en dimension finie).

Si H est un hyperplan de E, alors, pour tout $x \in E \setminus H$, $E = H \oplus \mathrm{Vect}(x)$ et, corrélativement, tout supplémentaire de H est une droite vectorielle. Corrélativement, si H' est un s.e.v. de E tel que $H \subset H'$, alors $H' = H$ ou $H' = E$.

Si E est de dimension finie (seul cas au programme), les propriétés ci-dessus sont évidentes.

Proposition 2. *Soit $(e_1, e_2, \ldots, e_n)$ une base de E. Si $(a_i)_{1 \leqslant i \leqslant n} \in \mathbb{K}^n \setminus \{0_{\mathbb{K}^n}\}$, l'ensemble des vecteurs $x = \sum_{i=1}^{n} x_i e_i$ vérifiant l'équation $a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0$ décrit un hyperplan de E. Réciproquement, tout hyperplan de E admet une équation de ce type, unique à une constante multiplicative non nulle près.*

Les équations d'hyperplans sont uniques à un coefficient multiplicatif près dans une base donnée. Elles dépendent évidemment de la base. De fait, les x_i de l'équation sont relatifs au fait que, si $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ est une base de E, chaque vecteur $x \in E$ peut s'écrire de manière unique sous la forme $x = x_1 e_1 + \cdots + x_n e_n$. En dimension 2 et 3, on retrouve les équations des droites (vectorielles) du plan euclidien et des plans (vectoriels) de l'espace euclidien.

Un exemple classique d'hyperplan de $\mathscr{M}_n(\mathbb{K})$ est l'ensemble des matrices de trace nulle. C'est le noyau de la forme linéaire trace et il est d'équation $a_{1,1} + a_{2,2} + \cdots + a_{n,n} = 0$ où la matrice générique A est repérée par ses coefficients $a_{i,j}$ dans la base canonique de $\mathscr{M}_n(\mathbb{K})$.

En général, les équations du type $a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = a_{n+1}$ décrivent les hyperplans affines (penser au droites du plan et aux plans de l'espace euclidiens ; la notion générale d'espace affine est hors programme). Là aussi, l'équation est unique à un facteur multiplicatif non nul près.

4. Calcul matriciel

4.1. **Opérations.**

Définition 4. *Soient n et p deux entiers strictement positifs. On appelle matrice à n lignes et p colonnes (ou matrice de taille (n,p) un tableau de scalaires $A=(a_{i,j})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}}$. Les opérations terme à terme sur les scalaires en font un espace vectoriel, noté $\mathscr{M}_{n,p}(\mathbb{K})$, de base dite canonique $(E_{i,j})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}}$ lue par ordre lexicographique des indices (i.e. $E_{1,1}$, $E_{1,2}$, …, $E_{1,p}$, $E_{2,1}$, …, $E_{n,p}$), où $E_{i,j}$ est la matrice dont le coefficient situé à la i-ème ligne et la j-ème colonne vaut 1, les autres valant 0. On note $L_i(A)$ la i-ème ligne de A et $C_j(A)$ sa j-ème colonne. On note aussi parfois $a_{i,j}=A|_{i,j}$. Une matrice de taille $(1,1)$ est librement identifiée à son unique coefficient. On note $\mathscr{M}_n(\mathbb{K})=\mathscr{M}_{n,n}(\mathbb{K})$.*

— Le produit de matrices obéit à des règles simples faciles à mémoriser en songeant au jeu des dominos. En particulier, $E_{i,j}E_{k,\ell}=\delta_{j,k}E_{i,\ell}$, où $\delta_{j,k}$ est le *symbole de Kronecker*, valant 1 si $j=k$ et 0 sinon.

— Si $A\in\mathscr{M}_{n,p}(\mathbb{K})$ et $X\in\mathscr{M}_{p,1}(\mathbb{K})$ (vecteur colonne de $\mathbb{K}^p$), $AX=\displaystyle\sum_{j=1}^{p}x_jC_j(A)$ ou, ce qui est équivalent, $AX|_i=L_i(A)X$.

— Plus généralement, le produit définit une application de $\mathscr{M}_{n,p}(\mathbb{K})\times\mathscr{M}_{p,q}(\mathbb{K})$ dans $\mathscr{M}_{n,q}(\mathbb{K})$. Les manières équivalentes de décrire ce produit, toutes utiles, sont :

$$(AB)|_{i,j}=\sum_{k=1}^{p}a_{i,k}b_{k,j}=L_i(A)C_j(B),\quad L_i(AB)=L_i(A)B,\quad C_j(AB)=AC_j(B).$$

— Cette « règle des dominos » permet d'écrire assez facilement des produits quelconques. Ainsi, pour une suite finie de matrices $A_k=\left(a_{i,j}^{(k)}\right)_{1\leqslant i,j\leqslant n}\in\mathscr{M}_n(\mathbb{K})$, si $B=A_1\times A_2\times\cdots\times A_p$, alors

$$B|_{i,j}=\sum_{(k_1,k_2,\ldots,k_{p-1})\in[\![1,n]\!]^{p-1}}a_{i,k_1}^{(1)}a_{k_1,k_2}^{(2)}\cdots a_{k_{p-2},k_{p-1}}^{(p-1)}a_{k_{p-1},j}^{(p)}.$$

— La formule du binôme de Newton s'applique à la somme de deux matrices **si elles commutent**.
— Si A est inversible et si $AB=AC$, alors $B=C$.

Théorème et définition 5. *Si $A\in\mathscr{M}_{n,p}(\mathbb{K})$, on appelle* rang de A *et on note* $\mathrm{rg}(A)$ *la dimension de l'espace vectoriel engendré par ses colonnes. C'est aussi celle de l'espace vectoriel engendré par ses lignes.*

Ainsi, $\mathrm{rg}(A)\leqslant\min(n,p)$. De plus, $\mathrm{rg}(A)=\mathrm{rg}(A^{\mathsf{T}})$ et, par le théorème du rang, $\mathrm{rg}(AB)\leqslant\min(\mathrm{rg}\,A,\mathrm{rg}\,B)$.

Théorème et définition 6. *Une matrice $A\in\mathscr{M}_n(\mathbb{K})$ est inversible si l'une des conditions équivalentes suivantes est vérifiée :*

(i) il existe $B\in\mathscr{M}_n(\mathbb{K})$ telle que $AB=I_n$ (inversibilité à droite) ;
(ii) il existe $C\in\mathscr{M}_n(\mathbb{K})$ telle que $CA=I_n$ (inversibilité à gauche) ;
(iii) il existe $D\in\mathscr{M}_n(\mathbb{K})$ telle que $AD=DA=I_n$ (inversibilité tout court) ;
(iv) $\mathrm{rg}(A)=n$ *;*
(v) $\det A\neq 0$.

Si A est inversible, les différents inverses sont uniques et égaux entre eux. On note $A^{-1}=B=C=D$. La notion d'inversibilité existe aussi en dimension infinie (pour les endomorphismes, par exemple). Il n'y a alors plus équivalence entre inversibilité, inversibilité à droite et inversibilité à gauche.

Si $(A,B)\in\mathscr{M}_n(\mathbb{K})^2$, alors AB est inversible si, et seulement si, A et B le sont, et l'on a $(AB)^{-1}=B^{-1}A^{-1}$. L'ensemble des matrices $M\in\mathscr{M}_n(\mathbb{K})$ inversibles est noté $\mathrm{GL}_n(\mathbb{K})$ et forme un groupe pour la multiplication, de neutre I_n. Seules les matrices carrées sont susceptibles d'être inversibles. Les matrices rectangulaires peuvent avoir un inverse à droite (si $p>n$) ou à gauche (si $p<n$), mais pas d'inverse. Il peut être pratique de connaître l'inverse d'une matrice de taille 2 :

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

On divise par le déterminant $ad-bc$, nécessairement non nul, on intervertit les éléments diagonaux et l'on change le signe des éléments antidiagonaux. Il existe une formule générale pour les matrices de taille n utilisant la *comatrice*, mais elle est hors programme en PSI (les ex-MPSI peuvent l'utiliser).

4.2. **Déterminant et trace.** Le déterminant et la trace dont définis sur des matrices carrées. La trace est la somme des éléments diagonaux d'une matrice. C'est une application linéaire et elle vérifie $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, d'où l'on déduit facilement que $\operatorname{tr}(P^{-1}AP) = \operatorname{tr}(A)$. Attention : que $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ n'entraîne pas une commutativité générale : ainsi, $\operatorname{tr}(ABC) \neq \operatorname{tr}(ACB)$ en général.

Théorème et définition 7. ***1.*** *(PCSI) Il existe une unique application* $\det : \mathscr{M}_n(\mathbb{K}) \to \mathbb{K}$ *ayant les propriétés suivantes :*

(i) l'application det *est linéaire par rapport à chaque colonne de* M *;*

(ii) l'application det *est* antisymétrique *par rapport aux colonnes de* M*, ce qui signifie que la permutation de deux colonnes multiplie le déterminant par* -1 *; elle est de manière équivalente* alternée*, ce qui signifie que si deux colonnes sont égales, le déterminant est nul ;*

(iii) $\det I_n = 1$.

2. *(MPSI) Si* $A = (a_{i,j})_{1\leqslant i,j\leqslant n}$*, on a* $\det A = \displaystyle\sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$.

La définition vue en PCSI n'est pas importante. Il importe en revanche de bien connaître les propriétés ci-dessous.
— La n-linéarité entraîne que $\det(\lambda A) = \lambda^n \det A$.
— Le déterminant est multiplicatif : $\det(AB) = (\det A)(\det B)$.
— Le déterminant caractérise l'inversibilité : $\det A \neq 0 \iff A \in \mathrm{GL}_n(\mathbb{K})$.
— $\det(A^{\mathsf{T}}) = \det A$.
— Développement par rapport à une ligne ou à une colonne. Soit $A_{i,j}$ la matrice extraite de $A \in \mathscr{M}_n(\mathbb{K})$ obtenue en en supprimant la i-ème ligne et la j-ème colonne. Alors,

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det A_{i,j} = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det A_{i,j}.$$

— Le déterminant est un invariant de similitude, ce qui permet de définir le déterminant d'un endomorphisme.
— Le déterminant d'une matrice triangulaire est le produit de ses éléments diagonaux.
— Déterminant de Vandermonde :

$$\begin{vmatrix} 1 & 1 & \cdots & \cdots & 1 \\ x_1 & x_2 & \cdots & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & \cdots & x_n^2 \\ \vdots & \vdots & & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{1\leqslant i<j\leqslant n} (x_j - x_i).$$

Application 1. La famille des suites $\big((z^n)_{n\in\mathbb{N}}\big)_{z\in\mathbb{C}}$ est libre.

Soient $P = \displaystyle\prod_{k=1}^{p} (X - \theta_k) = \sum_{i=0}^{p} a_i X^i$ un polynôme scindé à racines simples et $\mathcal{U}$ l'ensemble des suites vérifiant la relation de récurrence $u_{n+p} = -\displaystyle\sum_{i=0}^{p-1} a_i u_{n+i}$. Alors, $\mathcal{U}$ est un espace vectoriel de dimension p (ce qui reste vrai sans l'hypothèse de simplicité des racines de P), dont les suites $\big((\theta_k^n)_{n\in\mathbb{N}}\big)_{1\leqslant k\leqslant p}$ forment une base.

4.3. **Opérations élémentaires.** Elles sont de trois types :
— Les *transvections* sont des opérations du type $L_i \longleftarrow L_i + \lambda L_j$ ou $C_i \longleftarrow C_i + \lambda C_j$ avec $i \neq j$. Elles conservent le déterminant. La première revient à multiplier à gauche par $I_n + \lambda E_{i,j}$, la deuxième à multiplier à droite par $I_n + \lambda E_{j,i}$.
— Les *affinités*, ou *dilatations* sont des opérations du type $L_i \longleftarrow \lambda L_i$ ou $C_i \longleftarrow \lambda C_i$ avec $\lambda \neq 0$. Elles multiplient le déterminant par λ et reviennent à multiplier la matrice par $I_n + (\lambda - 1)E_{i,i}$. On peut réaliser plusieurs affinités en même temps en multipliant par une matrice diagonale. Ainsi, pour $D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, on a $C_j(AD) = \lambda_j C_j(A)$ et $L_i(DA) = \lambda_i L_i(A)$ pour $(i,j) \in [\![1,n]\!]$.
— Les *transpositions* sont des opérations du type $L_i \longleftrightarrow L_j$ ou $C_i \longleftrightarrow C_j$. Elles changent le signe du déterminant et reviennent à multiplier la matrice par $I_n + E_{i,j} + E_{j,i} - E_{i,i} - E_{j,j}$.

Les multiplications ci-dessus partent du calcul de base $(AE_{i,j})|_{k,\ell} = \delta_{j,\ell}a_{k,i}$, soit $C_j(AE_{i,j}) = C_i(A)$, les autres colonnes étant nulles. De même, $(E_{i,j}A)_{|k,\ell} = \delta_{i,k}a_{j,\ell}$, soit $L_i(E_{i,j}A) = L_j(A)$, les autres lignes étant nulle.
Toutes les opérations élémentaires conservent le rang. Les opérations élémentaires **sur les lignes** conservent en plus le noyau et l'image. Elles sont associées aux matrices creuses inversibles particulières décrites ci-dessus et reviennent à multiplier à gauche par ces matrices pour les opérations sur les lignes et à droite pour les opérations sur les colonnes. Les techniques de pivot et de mise sous forme échelonnée reposent sur des opérations élémentaires. Celles-ci fournissent également un algorithme efficace de calcul de l'inverse et du déterminant.
On enchaîne souvent plusieurs opérations élémentaires. En particulier, la produit à gauche par une matrice diagonale multiplie chaque ligne par le coefficient diagonal correspondant (idem sur les lignes en multipliant à droite).

4.4. **Matrices par blocs.** Les matrices par blocs sont définies dans la propriété ci-dessous, avec leur produit.

Proposition 3. ***1.*** *Sous réserve de compatibilité dimensionnelle, les règles de calculs de la somme et du produit sont les mêmes sur les blocs que sur les coefficients. Plus précisément,*

$$
\begin{array}{cc}
 & \begin{array}{cccc} \overset{p_1}{\longleftrightarrow} & \overset{p_2}{\longleftrightarrow} & \dots & \overset{p_s}{\longleftrightarrow} \end{array} \\
\begin{array}{c} n_1 \updownarrow \\ n_2 \updownarrow \\ \vdots \\ \vdots \\ n_r \updownarrow \end{array} &
\left(\begin{array}{c|c|c|c}
A_{1,1} & A_{1,2} & \dots & A_{1,s} \\ \hline
A_{2,1} & A_{2,2} & \dots & A_{2,s} \\ \hline
\vdots & \vdots & \ddots & \vdots \\ \hline
\vdots & \vdots & \ddots & \vdots \\ \hline
A_{r,1} & A_{r,2} & \dots & A_{r,s}
\end{array}\right)
\end{array}
\begin{array}{cc}
\begin{array}{ccccc} \overset{q_1}{\longleftrightarrow} & \overset{q_2}{\longleftrightarrow} & \dots & \dots & \overset{q_t}{\longleftrightarrow} \end{array} & \\
\left(\begin{array}{c|c|c|c|c}
B_{1,1} & B_{1,2} & \dots & \dots & B_{1,t} \\ \hline
B_{2,1} & B_{2,2} & \dots & \dots & B_{2,t} \\ \hline
\vdots & \vdots & \ddots & \ddots & \vdots \\ \hline
B_{s,1} & B_{s,2} & \dots & \dots & B_{s,t}
\end{array}\right) &
\begin{array}{c} \updownarrow p_1 \\ \updownarrow p_2 \\ \vdots \\ \updownarrow p_s \end{array}
\end{array}
=
\begin{array}{cc}
\begin{array}{ccccc} \overset{q_1}{\longleftrightarrow} & \overset{q_2}{\longleftrightarrow} & \dots & \dots & \overset{q_t}{\longleftrightarrow} \end{array} & \\
\left(\begin{array}{c|c|c|c|c}
C_{1,1} & C_{1,2} & \dots & \dots & C_{1,t} \\ \hline
C_{2,1} & C_{2,2} & \dots & \dots & C_{2,t} \\ \hline
\vdots & \vdots & \ddots & \ddots & \vdots \\ \hline
\vdots & \vdots & \ddots & \ddots & \vdots \\ \hline
C_{s,1} & C_{s,2} & \dots & \dots & C_{s,t}
\end{array}\right) &
\begin{array}{c} \updownarrow n_1 \\ \updownarrow n_2 \\ \vdots \\ \vdots \\ \updownarrow n_r \end{array}
\end{array}
$$

avec $A_{i,k} \in \mathscr{M}_{n_i,p_k}(\mathbb{K})$*,* $B_{k,j} \in \mathscr{M}_{p_k,q_j}(\mathbb{K})$ *et* $C_{i,j} = \displaystyle\sum_{k=1}^{s} A_{i,k}B_{k,j} \in \mathscr{M}_{n_i,q_j}(\mathbb{K})$*.*

2. *Si A est triangulaire par blocs, c'est-à-dire si $(n_1, n_2, \dots, n_r) = (p_1, p_2, \dots, p_s)$, et si $A_{i,j} = 0_{\mathscr{M}_{n_i,n_j}(\mathbb{K})}$ pour tout $i < j$ (triangulaire inférieure par blocs) ou pour tout $i > j$ (triangulaire supérieure par blocs),* $\det(A) = \displaystyle\prod_{k=1}^{r} \det(A_{k,k})$ *(noter que les blocs diagonaux sont alors bien carrés).*

En pratique, les matrices par blocs considérées (prenons le cas de A), vérifient $r = s$ et $(n_1, n_2, \dots, n_r) = (p_1, p_2, \dots p_r)$ et les blocs diagonaux sont ainsi carrés (mais les autres ne le sont pas nécessairement). En particulier, les matrices définies par blocs sont utiles pour exprimer le stabilité de certains sous-espaces et pour les problématiques de réduction spécifiques plus compliquées que la diagonalisation (réduction des isométries ou des endomorphismes antisymétriques, par exemple). On peut réaliser des opérations élémentaires par blocs, qui rassemblent une série d'opérations élémentaires usuelles.

Exemple 4. Soit $A \in \mathscr{M}_n(\mathbb{K})$.

$$\left(\begin{array}{c|c} 0 & A \\ \hline A & 0 \end{array}\right) \overset{}{\mathcal{L}_1 \leftrightarrow \mathcal{L}_2} = (-1)^n \left(\begin{array}{c|c} A & 0 \\ \hline 0 & A \end{array}\right) = (-1)^n (\det A)^2.$$

La transposition par blocs $\mathcal{L}_1 \leftrightarrow \mathcal{L}_2$ est équivalente aux n transpositions $L_i \leftrightarrow L_{i+n}$, d'où la contribution de $(-1)^n$.

Multiplier par une matrice diagonale par blocs (carrés !) opère sur les blocs de la matrice multipliée comme la multiplication par une matrice diagonale opère sur les coefficients.

5. Applications linéaires

5.1. **Définition.**

Définition 5. *Soient E et F deux e.v. sur le même corps $\mathbb{K}$. Une application linéaire de E dans F est une application $f \colon E \to F$ telle que,*

$$\forall (x, y) \in E^2,\ \forall (\lambda, \mu) \in \mathbb{K}^2 \colon\ f(\lambda x + \mu y) = \lambda f(x) + \mu f(y).$$

On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F, c'est un espace vectoriel. En dimension finie, $\dim \mathcal{L}(E, F) = \dim E \times \dim F$.

Une application linéaire est également appelée homomorphisme d'espace vectoriel*. Un* endomorphisme *est une application linéaire de E dans E, un* isomorphisme *est une application linéaire bijective et un* automorphisme *est à la fois un endomorphisme et un isomorphisme.*

En particulier, $f(0_E) = f(0_{\mathbb{K}}.0_E) = 0_{\mathbb{K}} f(0_E) = 0_F$. Cela essentiellement pour rappeler que $f(0_E) = 0_F$ est une **propriété** des applications linéaires et non une partie de leur définition. La considération du vecteur nul est utile pour montrer que l'on est en présence d'un s.e.v. ; elle ne sert à rien (sinon à perdre des points...) pour montrer qu'une application est linéaire.

Deux sous-espaces sont naturellement associés à une application **linéaire** $f \in \mathcal{L}(E, F)$:
— son noyau, $\operatorname{Ker} f = \{x \in E \ ;\ f(x) = 0_F\}$, s.e.v. de E, qui est réduit à $\{0_E\}$ si, et seulement si, f est injective (c'est une proposition) ;
— son image, $\operatorname{Im}(f) = \{f(x) \ ;\ x \in E\}$, s.e.v. de F, qui est égal à F si, et seulement si, f est surjective (c'est une définition).

— Une application $f \in \mathcal{L}(E, F)$ est entièrement déterminée par l'image d'une base de E ; plus précisément, si $(e_1, e_2, \ldots, e_n)$ est une base de E et si $(u_1, u_2, \ldots, u_n)$ est une famille de vecteurs de F, il existe une unique application linéaire $f \colon E \to F$ telle que, pour tout $i \in [\![1, n]\!]$, $f(e_i) = u_i$; f est injective ssi $\big(f(e_1), f(e_2), \ldots, f(e_n)\big)$ est libre et surjective ssi $\big(f(e_1), f(e_2), \ldots, f(e_n)\big)$ est génératrice (cela vaut alors pour l'image de toute base de E).
— Plus généralement, si $E = \bigoplus_{k=1}^{p} F_k$, l'application $f \longmapsto \big(f_{|F_1}, f_{|F_2}, \ldots, f_{|F_p}\big)$ induit un isomorphisme de $\mathcal{L}(E, G)$ sur $\prod_{k=1}^{p} \mathcal{L}(F_k, G)$, ce qui revient à dire que f est entièrement déterminée par les $f_{|F_k}$; f est injective si, et seulement si, les s.e.v. $\operatorname{Im}(f_{|F_1})$ sont en somme directe et si les $f_{|F_k}$ sont toutes injectives ; f est surjective si, et seulement si, $f(F_1) + f(F_2) + \cdots + f(F_p) = E$.

Théorème 2. Théorème du rang. *Soit* $f \in \mathcal{L}(E, F)$.
1. *Si* $E = \operatorname{Ker}(f) \oplus G$, *alors* $f_{|G}$ *réalise un isomorphisme de* G *sur* $\operatorname{Im}(f)$.
2. *E est de dimension finie si, et seulement si,* $\operatorname{Ker} f$ *et* $\operatorname{Im} f$ *le sont et l'on a alors* $\dim E = \dim \operatorname{Ker} f + \dim \operatorname{Im} f$.

En particulier, une application linéaire contracte les dimensions : la dimension de l'image est au plus égale à celle de l'espace de départ et il y a égalité si, et seulement si, l'application est injective.
En général, $\operatorname{Ker} f$ et $\operatorname{Im} f$ ne sont pas supplémentaires.

Si $f \in \mathcal{L}(E, F)$ avec $\dim E = \dim F < \infty$ (en particulier, si f est un endomorphisme d'un e.v. de dimension finie), alors

$$f \text{ est injective} \iff f \text{ est surjective} \iff f \text{ est bijective.}$$

5.2. **Lien entre matrices et applications linéaires.**

Définition 6. ***1.*** *Si* $f \in \mathcal{L}(E, F)$, *avec* E *et* F *de dimensions finies respectives* p *et* n, *si* $\mathcal{B}_E$ *est une base de* E *et* $\mathcal{B}_F$ *une base de* F, *la matrice* $A = \operatorname{mat}_{\mathcal{B}_F, \mathcal{B}_E}(f)$ *est la matrice de* $\mathscr{M}_{n,p}(\mathbb{K})$ *dont les colonnes sont les coordonnées dans* $\mathcal{B}_F$ *de l'image par* f *des vecteurs de* $\mathcal{B}_E$, *pris dans l'ordre.*
2. *Si* $E = F$, *on choisit en général une seule base et l'on note* $\operatorname{mat}_{\mathcal{B}}(f)$ *la matrice dont la colonne d'indice* j *est le vecteur des coordonnées de l'image par* f *du* $j^{\text{ème}}$ *vecteur de* $\mathcal{B}$.
3. *À une matrice* $A \in \mathscr{M}_{n,p}(\mathbb{K})$ *est associée* canoniquement *une application linéaire de* $\mathbb{K}^p$ *dans* $\mathbb{K}^n$, *celle ayant pour matrice* A *dans les bases canoniques de* $\mathbb{K}^p$ *et* $\mathbb{K}^n$ *; elle est donné par* $\phi_A \colon X \longmapsto AX$.

Il n'y a pas de notation universelle pour la matrice d'une application linéaire dans deux bases. On a pris ici le parti d'écrire $\operatorname{mat}_{\mathcal{B}_F, \mathcal{B}_E}(f) \in \mathscr{M}_{n,p}(\mathbb{K})$, avec la base de l'espace d'arrivée avant celle de l'espace de départ. Cette écriture est compatible avec la taille de la matrice (qui, elle, est universelle), n étant la dimension de F et p celle de E. Cette convention respecte aussi l'écriture par domino des changements de base (*cf.* formule (1) *infra*).
Si X est le vecteur colonne des coordonnées de $E \ni x$ dans la base $\mathcal{B}_E$, $A = \operatorname{mat}_{\mathcal{B}_E, \mathcal{B}_F}(f)$ et Y le vecteur colonne des coordonnées de $E \ni y$ dans la base $\mathcal{B}_F$, alors $Y = AX = \phi_A(X) \iff y = f(x)$.

Exemple 5. L'application $\psi \colon \mathbb{R}_2[X] \to \mathbb{R}_2[X]$ définie par $\psi(P) = (X-1)P' + 2P(X+1)$ est linéaire et vérifie $\psi(1) = 2$, $\psi(X) = 3X + 1$ et $\psi(X^2) = 4X^2 + 2X + 2$. On calcule (par exemple),

$$\psi(X^2 - 2X + 3) = (X-1)(2X-2) + 2[(X+1)^2 - 2(X+1) + 3] = 4X^2 - 4X + 6 \qquad \& \qquad \begin{pmatrix} 2 & 1 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ -4 \\ 4 \end{pmatrix}.$$

Définition 7. *Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases d'un même espace E, la matrice de changement de base $P_{\mathcal{B}\to\mathcal{B}'}$ est la matrice dont les colonnes expriment, dans l'ordre, les coordonnées des vecteurs de $\mathcal{B}'$ (la nouvelle base) dans $\mathcal{B}$ (l'ancienne).*

— Si $\mathcal{B}$, $\mathcal{B}'$ et $\mathcal{B}''$ sont trois bases de E, $P_{\mathcal{B}\to\mathcal{B}''} = P_{\mathcal{B}\to\mathcal{B}'} \times P_{\mathcal{B}'\to\mathcal{B}''}$.
— Si X et X' sont les vecteurs colonnes des coordonnées de $x \in E$ dans $\mathcal{B}$ et $\mathcal{B}'$ respectivement, alors $X = P_{\mathcal{B}\to\mathcal{B}'}X'$.
— Pour les matrices d'applications linéaires $f \in \mathcal{L}(E,F)$, on a (noter l'écriture en domino)

(1) $$\mathrm{mat}_{\mathcal{B}'_F,\mathcal{B}'_E}(f) = P^{-1}_{\mathcal{B}_F\to\mathcal{B}'_F}\,\mathrm{mat}_{\mathcal{B}_F,\mathcal{B}_E}(f)P_{\mathcal{B}_E\to\mathcal{B}'_E} = P_{\mathcal{B}'_F\to\mathcal{B}_F}\,\mathrm{mat}_{\mathcal{B}_F,\mathcal{B}_E}(f)P_{\mathcal{B}_E\to\mathcal{B}'_E}.$$

On peut retrouver où placer quelle matrice de passage en s'arrangeant pour que le produit soit valide ; la preuve s'obtient en utilisant les formules de changement de base pour les vecteurs.
— En particulier, pour un endomorphisme,

$$\mathrm{mat}_{\mathcal{B}'}(f) = P^{-1}_{\mathcal{B}\to\mathcal{B}'}\,\mathrm{mat}_{\mathcal{B}}(f)P_{\mathcal{B}\to\mathcal{B}'} = P_{\mathcal{B}'\to\mathcal{B}}\,\mathrm{mat}_{\mathcal{B}}(f)P_{\mathcal{B}\to\mathcal{B}'}.$$

— Soient $f \in \mathcal{L}(E,F)$ et $g \in \mathcal{L}(F,G)$; soient $\mathcal{B}_E$, $\mathcal{B}_F$ et $\mathcal{B}_G$ des bases de E, F et G respectivement. Alors,

$$\mathrm{mat}_{\mathcal{B}_G,\mathcal{B}_E}(g\circ f) = \mathrm{mat}_{\mathcal{B}_G,\mathcal{B}_F}(g)\,\mathrm{mat}_{\mathcal{B}_F,\mathcal{B}_E}(f).$$

5.3. **Relation de similitude.**

Définition 8. *Deux matrices A et B de $\mathscr{M}_n(\mathbb{K})$ sont semblables s'il existe une matrice inversible P telle que $B = P^{-1}AP$. On appelle* invariant de similitude *une propriété ou une quantité matricielle stable par similitude.*

La relation de similitude est une relation d'équivalence. Elle prend tout son sens en considérant les endomorphismes : deux matrices sont semblables si, et seulement si, elles représentent le même endomorphisme. Autrement dit, A et B sont semblables s'il existe $\mathcal{B}$ et $\mathcal{B}'$ deux bases de $\mathbb{K}^n$ et $f \in \mathcal{L}(\mathbb{K}^n)$ tels que $A = \mathrm{mat}_{\mathcal{B}}(f)$ et $B = \mathrm{mat}_{\mathcal{B}'}(f)$.

Le rang, la trace, le déterminant, l'inversibilité sont des invariants de similitude. On en verra d'autres dans le chapitre sur la réduction.

5.4. **Sous-espaces stables.**

Définition 9. *Soit $f \in \mathcal{L}(E)$. On dit qu'un s.e.v. F de E est* stable *par f si $f(F) \subset F$, i.e. si $\forall x \in F : f(x) \in F$.*

Proposition 4. *1. Une somme de s.e.v. stables par un endomorphisme f est un s.e.v. stable par f.*
2. Une intersection de s.e.v. stables par f est un s.e.v. stable par f.
3. Si $f \in \mathcal{L}(E)$, alors $\mathrm{Ker}(f)$ et $\mathrm{Im}(f)$ sont stables par f.
4. Soient deux endomorphismes f et g de E tels que $f\circ g = g\circ f$. Alors, $\mathrm{Ker}(f)$ et $\mathrm{Im}(f)$ sont stables par g.

La proposition 4 a l'air anecdotique, mais elle est au cœur d'un thème classique d'exercices : la *coréduction*.

Définition 10. *Soient $f \in \mathcal{L}(E)$ et F un sous-espace vectoriel de E. On définit par* restriction *l'application linéaire $f_{|F} \in \mathcal{L}(F,E)$ par $\forall u \in F : f_{|F}(u) = f(u)$. Si F est stable par f, alors $\mathrm{Im}\left(f_{|F}\right) \subset F$ et l'on peut définir l'endomorphisme de F induit par f sur F, que l'on notera $f_{\|F}$. On parle d'*endomorphisme induit.

L'endomorphisme induit par f sur F n'est pas exactement sa restriction car l'espace d'arrivée change. Si $E = F\oplus G$ et si $\mathcal{B}$ est une base adaptée à cette décompostion, concaténée de $\mathcal{B}_F$ et de $\mathcal{B}_G$, on peut décomposer par blocs la matrice $\mathrm{mat}_{\mathcal{B}}(f) = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)$. Alors, F est stable par f si, et seulement si, $C = 0_{\mathscr{M}_{\dim G,\dim F}(\mathbb{K})}$. On a alors $\mathrm{mat}_{\mathcal{B}_F}(f_{\|F}) = A$ et $\mathrm{mat}_{\mathcal{B},\mathcal{B}_F}(f_{|F}) = \begin{pmatrix} A \\ 0 \end{pmatrix}$. *Attention :* la notion d'endomorphisme induit par f sur F n'a de sens que si F est stable par f. La notation $f_{\|F}$ n'est pas universelle et doit être redéfinie en cas d'utilisation. Sans hypothèse de stabilité, si $p = p_{F//G}$ est le projecteur sur F parallèlement à G, alors $A = \mathrm{mat}_{\mathcal{B}_F}\left((p\circ f)_{\|F}\right)$ et $B = \mathrm{mat}_{\mathcal{B}_F,\mathcal{B}_G}\left((\mathrm{id}_E - p)\circ f\right)$.

Plus généralement, on a vu que si $E = F_1 \oplus F_2 \oplus \cdots \oplus F_p$, alors $f \in \mathcal{L}(E)$ est entièrement déterminée par ses restrictions $f_i = f_{|F_i}$ (si $\dim F_i = 1$ pour tout i, on retrouve le fait qu'une application linéaire est entièrement déterminée par l'image d'une base). Dans une base adaptée à cette décomposition, $\mathrm{mat}(f)$ est

— diagonale par blocs si, et seulement si, tous les F_i sont stables par f ;
— triangulaire supérieure par blocs si, et seulement si, $F_1 \oplus F_2 \oplus \cdots \oplus F_k$ est stable par f pour tout $k \in [\![1, p-1]\!]$;
— triangulaire inférieure par blocs si, et seulement si, $F_k \oplus F_{k+1} \oplus \cdots \oplus F_p$ est stable par f pour tout $k \in [\![2, p]\!]$.

6. Polynômes

6.1. **Espaces de polynômes (rappels).** On rappelle que, pour $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, on note $\mathbb{K}[X]$ le $\mathbb{K}$-espace vectoriel des polynômes à coefficients dans $\mathbb{K}$ et, pour $n \in \mathbb{N}$, $\mathbb{K}_n[X]$ celui des polynômes de degré au plus n. Alors, $\mathbb{K}[X]$ est de dimension $\aleph_0$ de base canonique $(X^k)_{k \geqslant 0}$ et $\mathbb{K}_n[X]$ est de dimension $n+1$ de base canonique $(X^k)_{0 \leqslant k \leqslant n}$.

Si $\mathcal{P} \subset \mathbb{K}[X]$, la famille $\mathcal{P}$ est dite *échelonnée en degré* si la restriction de l'application deg à $\mathcal{P}$ est injective (tous les polynômes de $\mathcal{P}$ sont de degrés distincts). Une famille échelonnée en degré est libre. Si $\mathcal{P}$ est échelonnée en degré et si $\{\deg(P), P \in \mathcal{P}\} = \mathbb{N}$, alors $\mathcal{P}$ est une base de $\mathbb{K}[X]$. Si $\{\deg(P), P \in \mathcal{P}\} = [\![0, n]\!]$, $\mathcal{P}$ est une base de $\mathbb{K}_n[X]$.

Proposition 5 (Formule de Taylor)**.** *Pour tout $P \in \mathbb{K}[X]$ et tout $a \in \mathbb{K}$, on a $P = \displaystyle\sum_{k=0}^{\deg(P)} \frac{P^{(k)}(a)}{k!}(X-a)^k$.*

On dit que a est *racine d'ordre m de P* si $(X-a)^m || P$, c'est-à-dire si P se factorise par $(X-a)^m$ mais pas par $(X-a)^{m+1}$. De manière équivalent, $P(a) = P'(a) = \cdots = P^{(m-1)}(a)$ et $P^{(m)}(a) \neq 0$.

Théorème 3 (Division euclidienne)**.** *Si $A \in \mathbb{K}[X]$ est un polynôme de degré $d \geqslant 1$, pour tout $P \in \mathbb{K}[X]$, il existe un unique couple $(Q, R) \in \mathbb{K}[X] \times \mathbb{K}_{d-1}[X]$ tel que $P = AQ + R$.*

Théorème 4 (Théorème de factorisation)**.** ***1.*** *Si $P \in \mathbb{C}[X]$ est non nul, il existe une factorisation unique (à l'ordre près) $P = c \displaystyle\prod_{k=1}^{r} (X - a_i)^{n_i}$ avec $c \in \mathbb{C}^*$, $(a_1, a_2, \ldots, a_r) \in \mathbb{C}^m$ et $(n_1, n_2, \ldots, n_r) \in \mathbb{N}^{*m}$. Autrement dit, tout polynôme complexe non constant est* scindé *(*théorème de d'Alembert-Gauß *ou* théorème fondamental de l'algèbre*).*
2. *Si $P \in \mathbb{R}[X]$ est non nul, il existe une unique factorisation $P = c \displaystyle\prod_{k=1}^{r} (X - a_i)^{n_i} \times \prod_{j=1}^{s} (X^2 + p_j X + q_j)^{m_j}$ avec $c \in \mathbb{R}^*$, $(a_1, a_2, \ldots, a_r) \in \mathbb{R}^m$ et $(n_1, n_2, \ldots, n_r) \in \mathbb{N}^{*m}$, $(p_j, q_j) \in \mathbb{R}^2$ pour tout $j \in [\![1, s]\!]$ avec $p_j^2 - 4q_j < 0$ et $(m_1, m_2, \ldots, m_s) \in \mathbb{N}^{*s}$.*

6.2. **Interpolation de Lagrange.**

Théorème et définition 8. ***1.*** *Soit $(x_0, x_1, \ldots, x_n) \in \mathbb{K}^{n+1}$ une famille de scalaires <u>deux à deux distincts</u>. Alors, l'application $\varphi \colon P \longmapsto \big(P(x_0), P(x_1), \ldots, P(x_n)\big)$ réalise un isomorphisme de $\mathbb{K}_n[X]$ sur $\mathbb{K}^{n+1}$ et la matrice de φ dans les bases canoniques de $\mathbb{K}^{n+1}$ et de $\mathbb{K}_n[X]$ est la matrice de Vandermonde $(x_i^j)_{0 \leqslant i,j \leqslant n}$.*
2. *L'image réciproque de la base canonique de $\mathbb{K}^{n+1}$ par φ forme* les polynômes d'interpolation de Lagrange. *Explicitement, pour $i \in [\![0, n]\!]$, on a $L_i = \displaystyle\prod_{\substack{0 \leqslant j \leqslant n \\ j \neq i}} \frac{X - x_j}{(x_i - x_j)}$. Alors, pour $(a_0, a_1, \ldots, a_n) \in \mathbb{K}^{n+1}$, l'unique polynôme P de $\mathbb{K}_n[X]$ tel que $P(x_i) = a_i$ pour tout $i \in [\![0, n]\!]$ est $P = \displaystyle\sum_{i=0}^{n} a_i L_i$. En particulier, pour $a_i = 1$ pour tout i, $\displaystyle\sum_{i=0}^{n} L_i = 1$.*

— On sait donc donner une expression explicite d'un polynôme dans différentes bases : $P = \displaystyle\sum_{k=0}^{\deg(P)} \frac{P^{(k)}(a)}{k!}(X-a)^k$ (Taylor) et $P = \displaystyle\sum_{k=0}^{n} P(x_i) L_i$ (Lagrange).
— Notons $\mathcal{L} = (L_0, L_1, \ldots, L_n)$. C'est une base de $\mathbb{K}_n[X]$ comme image réciproque d'une base par un isomorphisme. Les formules de changement de base donnent

$$I_n = \mathrm{mat}_{\mathrm{can}, \mathcal{L}}(\varphi) = P_{\mathrm{can} \to \mathrm{can}} \times \mathrm{mat}_{\mathrm{can}, \mathrm{can}}(\varphi) \times P_{\mathrm{can} \to \mathcal{L}} = (x_i^j)_{0 \leqslant i,j \leqslant n} \times L,$$

où $C_j(L)$ est le vecteur des coordonnées de L_j dans la base canonique des polynômes. Cela donne une expression explicite de l'inverse d'une matrice de Vandermonde.

— Si $Q=\prod_{i=0}^{n}(X-x_i)$, la relation $\sum_{i=0}^{n}L_i=1$ donne $\frac{1}{Q}=\sum_{i=0}^{n}\frac{L_i}{Q}=\sum_{i=0}^{n}\frac{\alpha_i}{X-x_i}$ avec $\alpha_i\prod_{\substack{0\leqslant j\leqslant n\\ j\neq i}}(x_i-x_j)=1$, ce qui donne l'existence de la décomposition en éléments simples de Q, l'unicité étant assurée par la liberté de $(L_i)_{0\leqslant i\leqslant n}$.

— Toujours avec $Q=\prod_{i=0}^{n}(X-x_i)$, le reste de la division euclidienne de P par Q est le polynôme d'interpolation de Lagrange de P en les racines de Q.

6.3. **Polynômes d'endomorphismes et de matrices.**

Théorème et définition 9. ***1.*** *Pour E un $\mathbb{K}$-e.v., $f\in\mathcal{L}(E)$ et $m\in\mathbb{N}$, $f^m=\overbrace{f\circ\cdots\circ f}^{\#m}$ est, par définition, la composition répétée m fois de f. Par convention, $f^0=\mathrm{id}_E$. Cela permet de définit une application $\Phi_f\colon\mathbb{K}[X]\to\mathcal{L}(E)$ par $\Phi_f(P)=P(f)$. Alors,*
- *— $\forall(P,Q)\in\mathbb{K}[X]^2$, $\forall(\lambda,\mu)\in{}^2$, $\forall f\in\mathcal{L}(E)\colon(\lambda P+\mu Q)(f)=\lambda P(f)+\mu Q(f)$;*
- *— $\forall(P,Q)\in\mathbb{K}[X]^2$, $\forall f\in\mathcal{L}(E)\colon(PQ)(f)=P(f)\circ Q(f)=Q(f)\circ P(f)=(QP)(f)$.*

2. *On note $\mathbb{K}[f]=\left\{P(f)\,;\ P\in\mathbb{K}[X]\right\}$ l'image de Φ_f. C'est un s.e.v. de $\mathcal{L}(E)$ dont les éléments commutent, en particulier avec f.*
3. *Un polynôme annulateur de f est un polynôme P tel que $P(f)=0_{\mathcal{L}(E)}$. L'ensemble des polynômes annulateurs de f, noyau de Φ_f, est un s.e.v. de $\mathbb{K}[X]$.*

Si $f\in\mathcal{L}(E)$, f admet un polynôme annulateur non nul si, et seulement si, $\mathbb{K}[f]$ est de dimension finie. Si $\dim E=n$, la famille $(f^k;0\leqslant k\leqslant n^2)$ est une famille liée de $\mathcal{L}(E)$ (qui est de dimension n^2), donc tout endomorphisme $f\in\mathcal{L}(E)$ admet un polynôme annulateur non nul de degré au plus n^2 (le théorème de Cayley-Hamilton montre que l'on peut remplacer n^2 par n). En dimension infinie, il peut ne pas y avoir de polynôme annulateur non nul ; prendre par exemple l'opérateur de dérivation sur $\mathbb{K}[X]$.

Exemple 6. — Un projecteur est un endomorphisme dont X^2-X est polynôme annulateur.
— Une symétrie et un endomorphisme dont X^2-1 est polynôme annulateur.
— Un endomorphisme f est nilpotent s'il admet un polynôme annulateur de la forme X^m ; la plus petite valeur de m est alors appelée *indice de nilpotence* de f. Par exemple, sur $\mathbb{K}_n[X]$, X^{n+1} est polynôme annulateur de l'opérateur de dérivation.

Proposition 6. *Si f et g sont des endomorphismes du $\mathbb{K}$-e.v. E et P et Q des éléments de $\mathbb{K}[X]$:*
- *(i) si $P(f)=0$, alors $QP(f)=PQ(f)=0$;*
- *(ii) $\mathrm{Ker}(P(f))$ et $\mathrm{Im}(P(f))$ sont stables par f ;*
- *(iii) si $PQ(f)=0$ et si $Q(f)$ est inversible, alors $P(f)=0$;*
- *(iv) si F est un s.e.v. de E stable par f, alors F est stable par $P(f)$. Si $P(f)=0_{\mathcal{L}(E)}$, alors $P(f_{\|F})=0_{\mathcal{L}(F)}$;*
- *(v) si $f\circ g=g\circ f$ (en part. si $f=g$), alors $P(f)\circ Q(g)=Q(g)\circ P(f)$ et $\mathrm{Ker}(P(f))$ et $\mathrm{Im}(P(f))$ sont stables par $Q(g)$.*

Par définition, si $P=\sum_{k=0}^{d}a_kX^k$, alors $P(f)=\sum_{k=0}^{d}a_kf^k$. Un polynôme annulateur dont le coefficient constant est non nul montre l'inversibilité de l'endomorphisme et en donne l'inverse : on vérifie que si $P(f)=0_{\mathcal{L}(E)}$ avec $P(0)\neq 0_{\mathbb{K}}$, alors $f^{-1}=Q(f)$ où Q est donné par $Q=\frac{P(0)-P}{a_0X}$ (ce résultat est facile à retrouver et ne doit pas être appris par cœur).

Les définitions relatives aux endomorphismes et la proposition 6 s'étendent sans difficulté aux matrices carrées. Les polynômes s'appliquent aux matrices carrées avec la convention $A^0=I_n$ pour $A\in\mathscr{M}_n(\mathbb{K})$. On démontre facilement la proposition suivante :

Proposition 7. *Si A et B appartiennent à $\mathscr{M}_n(\mathbb{K})$ et si $P \in \mathbb{K}[X]$, alors,*

(i) si $\mathrm{mat}_{\mathcal{B}}(f) = A$, alors $\mathrm{mat}_{\mathcal{B}}(P(f)) = P(A)$;

(ii) $P(A^{\mathsf{T}}) = P(A)^{\mathsf{T}}$;

(iii) si $B = Q^{-1}AQ$, alors $P(B) = Q^{-1}P(A)Q$.

Si $A \in \mathscr{M}_n(\mathbb{K})$ admet un polynôme annulateur scindé à racines simples et $m \in \mathbb{N}$, la division euclidienne de X^m par P_0 s'écrit $X^m = Q_m P_0 + R_m$ avec $\deg R_m < \deg P_0$, R_m étant le polynôme d'interpolation de Lagrange de X^m en les racines de P_0. Alors, $A^m = R_m(A)$, ce qui permet donc de calculer explicitement A^m pour $m \in \mathbb{N}$.

7. Systèmes linéaires

Définition 11. *Un système linéaire est une équation de la forme $AX = B$, où $A \in \mathscr{M}_{n,p}(\mathbb{K})$ et $B \in \mathscr{M}_{n,1}(\mathbb{K}) \simeq \mathbb{K}^n$ sont donnés, l'inconnue étant $X \in \mathscr{M}_{p,1}(\mathbb{K}) \simeq \mathbb{K}^p$. Le* rang *du système est par définition* $\mathrm{rg}(A)$.

Autrement dit, en termes de $\phi_A \in \mathscr{L}(\mathbb{K}^p, \mathbb{K}^n)$ l'endomorphisme canoniquement associé à la matrice A, on recherche l'image réciproque $\phi_A^{-1}(\{B\})$. Si $B \notin \mathrm{Im}(\phi_A)$, alors le système n'a pas de solution. Dans le cas contraire, si X_0 est solution du système, l'ensemble des solutions est l'espace affine $X_0 + \mathrm{Ker}(\phi_A)$, qui est de dimension $p - \mathrm{rg}(A)$. On parle alors de *système compatible*, ce qui est équivalent à $B \in \mathrm{Vect}\big(C_j(A)\, 1 \leqslant j \leqslant p\big)$ ou encore au fait que la *matrice augmentée* obtenue à partir de A en lui adjoignant B comme colonne surnuméraire, a même rang que A. On dispose également du principe de superposition des solutions : si $AX_0 = B_0$ et $AX_1 = B_1$, alors $A(X_0 + X_1) = B_0 + B_1$. La résolution effective se fait par la méthode du *pivot de Gauß*.

Si $A \in \mathscr{M}_n(\mathbb{R})$, le système admet une unique solution, si, et seulement si, $A \in \mathbf{GL}_n(\mathbb{K})$, donnée par $X = A^{-1}B$. On parle alors de *système de Cramer*.

Proposition 8 (HP)**.** *Si $AX = B$ est un système de Cramer, son unique solution $X = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}^{\mathsf{T}}$ est donnée par $x_k = \dfrac{\det(A_k)}{\det(A)}$, où A_k est la matrice obtenue à partir de A en substituant B à $C_k(A)$.*