ALGÈBRE LINÉAIRE (programme de 2ème année PSI)

I. Produit et somme d'espaces vectoriels

1. Produit cartésien.

Si E et F sont deux $\mathbb{K}$-espaces vectoriels, l'ensemble $E \times F$ des couples (x, y) avec $x \in E$ et $y \in F$ est muni d'une structure de $\mathbb{K}$-espace vectoriel en posant

$$\forall (x, y) \in E \times F \quad \forall (x', y') \in E \times F \qquad (x, y) + (x', y') = (x + x', y + y') \qquad \text{(addition interne)}$$

$$\forall (x, y) \in E \times F \quad \forall \lambda \in \mathbb{K} \qquad \lambda\,(x, y) = (\lambda x, \lambda y) \qquad \text{(multiplication par un scalaire)}\ .$$

Les vérifications des axiomes de la structure d'espace vectoriel sont laissées à un très improbable lecteur (ce n'est pas très intéressant à écrire!).

De façon plus générale, si $E_1, \cdots, E_n$ sont des $\mathbb{K}$-espaces vectoriels, l'ensemble $E_1 \times \cdots \times E_n$, aussi noté $\prod_{k=1}^{n} E_k$, constitué des n-uplets $(x_1, \cdots, x_n)$ avec $x_k \in E_k$ pour tout $k \in [\![1, n]\!]$, est muni d'une structure de $\mathbb{K}$-espace vectoriel en posant

$$(x_1, \cdots, x_n) + (y_1, \cdots, y_n) = (x_1 + y_1, \cdots, x_n + y_n) \qquad \text{(addition interne)}$$

$$\lambda\,(x_1, \cdots, x_n) = (\lambda x_1, \cdots, \lambda x_n) \qquad \text{(multiplication par un scalaire)}\ .$$

Les quantificateurs ont été omis pour alléger un peu l'écriture.

<u>Proposition.</u> Si les espaces vectoriels E_k, $1 \leq k \leq n$, sont de dimension finie, alors leur produit cartésien est aussi de dimension finie, et $\boxed{\dim\Big(\prod_{k=1}^{n} E_k\Big) = \sum_{k=1}^{n} \dim(E_k)}$.

<u>Preuve succincte.</u> Montrons-le simplement dans le cas du produit cartésien de deux espaces vectoriels E et F, de dimensions n et p respectivement. Soit $\mathcal{B} = (e_1, \cdots, e_n)$ une base de E, soit $\mathcal{C} = (f_1, \cdots, f_p)$ une base de F. Un élément de $E \times F$ est un couple (x, y) avec $x = \sum_{i=1}^{n} x_i e_i \in E$ et $y = \sum_{j=1}^{p} y_j f_j \in F$. On peut le décomposer en

$$\textbf{(*)} \qquad (x, y) = x_1(e_1, 0_F) + \cdots + x_n(e_n, 0_F) + y_1(0_E, f_1) + \cdots + y_p(0_E, f_p)\ ,$$

ce qui prouve que la famille $\mathcal{F} = \big((e_1, 0_F), \cdots, (e_n, 0_F), (0_E, f_1), \cdots, (0_E, f_p)\big)$ est génératrice de l'espace vectoriel $E \times F$. Elle est d'autre part libre puisque, si la combinaison linéaire qui figure au second membre de la relation **(*)** *est nulle, alors $x = \sum_{i=1}^{n} x_i e_i = 0_E$ et $y = \sum_{j=1}^{p} y_j f_j = 0_F$, ce qui entraîne la nullité de tous les scalaires en utilisant la liberté des familles $\mathcal{B}$ et $\mathcal{C}$. L'espace vectoriel $E \times F$, qui admet donc une base constituée de $n + p$ éléments, est de dimension $n + p$, ce qu'il fallait prouver dans le cas de deux espaces. Le cas général est plus lourd à écrire, mais l'idée reste la même.*

<u>Exemple.</u> Soient F et G deux sous-espaces vectoriels d'un même espace vectoriel E. Soit l'application

$$\varphi : \begin{cases} F \times G & \to \quad E \\ (y, z) & \mapsto \ y + z \end{cases}.$$

Cette application est clairement linéaire (il suffit de l'écrire!), on a $\mathrm{Im}(\varphi) = F + G$ (c'est la définition de la somme de deux s.e.v.), et $\mathrm{Ker}(\varphi) = \big\{(x, -x)\ ;\ x \in F \cap G\big\}$ est isomorphe à $F \cap G$. Si E est de dimension finie, le théorème du rang appliqué à φ donne alors $\dim(F \times G) = \mathrm{rg}(\varphi) + \dim\big(\mathrm{Ker}(\varphi)\big)$, puis la **formule de Grassmann**

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

ainsi que le fait que F et G sont en somme directe ($F \cap G = \{0_E\}$) si et seulement si φ est injective, i.e. si et seulement si $\dim(F+G) = \dim(F) + \dim(G)$.

2. <u>Somme d'un nombre fini de sous-espaces vectoriels.</u>

a. <u>Définition.</u>

<u>Définition.</u> Soient $E_1, \cdots, E_m$ des sous-espaces vectoriels d'un espace vectoriel E. On appelle **somme** des E_i, $1 \leq i \leq m$, l'ensemble S des vecteurs x de E qu'il est possible d'écrire comme somme d'un vecteur x_1 de E_1, ..., d'un vecteur x_m de E_m:

$$S = \left\{ x \in E \mid \exists (x_1, \cdots, x_m) \in E_1 \times \cdots \times E_m \quad x = x_1 + \cdots + x_m \right\} .$$

<u>Proposition.</u> Cet ensemble S est un sous-espace vectoriel de E contenant chacun des sous-espaces E_i.
On le note $S = E_1 + \cdots + E_m$ ou encore $S = \sum_{i=1}^{m} E_i$.

<u>Preuve.</u> L'ensemble S est non vide puisqu'il contient le vecteur nul que l'on peut écrire $0_E = 0_E + \cdots + 0_E$, et il est stable par combinaisons linéaires puisque, avec des notations que le lecteur comprendra:

$$\lambda(x_1 + \cdots + x_m) + (y_1 + \cdots + y_m) = (\lambda x_1 + y_1) + \cdots + (\lambda x_m + y_m) .$$

Si $x_1 \in E_1$, l'écriture $x_1 = x_1 + 0_E + \cdots + 0_E$ montre que $x_1 \in S$, d'où l'inclusion $E_1 \subset S$, et on procède de même pour les autres.

<u>Remarque.</u> On peut préciser que ce sous-espace somme S est "le plus petit" sous-espace vectoriel de E (au sens de l'inclusion) contenant chacun des E_i. Il faut comprendre par là que, si V est un sous-espace vectoriel de E contenant chacun des E_i, alors V contient S. On peut dire aussi que S est "le plus petit" sous-espace vectoriel de E, au sens de l'inclusion, contenant la réunion $\bigcup_{i=1}^{n} E_i$, cette réunion n'étant pas elle-même, en général, un sous-espace vectoriel de E.

b. <u>Notion de somme directe.</u>

<u>Définition.</u> On dit que les sous-espaces $E_1, \cdots, E_m$ de E **sont en somme directe** lorsque tout vecteur x de la somme $S = E_1 + \cdots + E_m$ admet **une unique** décomposition en $x = x_1 + \cdots + x_m$ avec $x_1 \in E_1, \cdots, x_m \in E_m$.

<u>Caractérisation.</u>Les sous-espaces $E_1, \cdots, E_m$ de E sont en somme directe si et seulement si

$$\forall (x_1, \cdots, x_m) \in E_1 \times \cdots \times E_m \qquad x_1 + \cdots + x_m = 0_E \Longrightarrow x_1 = \cdots = x_m = 0_E$$

(la seule façon de décomposer le vecteur nul en une composante x_1 appartenant à E_1, ..., une composante x_m appartenant à E_m, est de prendre toutes les composantes nulles).

<u>Preuve.</u> Si les E_i sont en somme directe, il suffit d'appliquer la définition avec $x = 0_E$ pour obtenir cette propriété caractéristique. Réciproquement, supposons vérifiée cette propriété caractéristique, soit $x \in S = E_1 + \cdots + E_m$, supposons qu'il admette deux décompositions $x = x_1 + \cdots + x_m$ et $x = y_1 + \cdots + y_m$. En soustrayant les deux relations, on a $0_E = (x_1 - y_1) + \cdots + (x_m + y_m)$. L'unicité supposée de la décomposition du vecteur nul montre alors que $x_i - y_i = 0_E$ pour tout i, donc $x_i = y_i$ pour tout i. La somme est donc bien directe.

Remarque. On pourra noter une certaine ressemblance avec la notion de famille libre de vecteurs.

Lorsque la somme $S = \sum_{i=1}^{m} E_i$ est directe, on la note

$$S = E_1 \oplus \cdots \oplus E_m \quad \text{ou encore} \quad S = \bigoplus_{i=1}^{m} E_i \ .$$

ATTENTION! Dans le cas de **deux** sous-espaces F et G de E, ils sont en somme directe si et seulement si $F \cap G = \{0_E\}$. Il ne faut pas chercher à généraliser cela. Ainsi, dès que $m \geq 3$, la condition que les intersections deux à deux des sous-espaces E_i $(1 \leq i \leq m)$ soient réduites à $\{0_E\}$ $(E_i \cap E_j = \{0_E\}$ si $i \neq j)$ ne suffit pas à garantir que la somme est directe! Pour le voir, il suffit de considérer dans $E = \mathbb{R}^2$ interprété comme "le plan" trois droites vectorielles distinctes, par exemple l'axe des abscisses (noté F), l'axe des ordonnées (noté G), la première bissectrice (notée H). Pour formaliser un peu plus, notons (e_1, e_2) la base canonique de $\mathbb{R}^2$, alors $F = \text{Vect}(e_1)$, $G = \text{Vect}(e_2)$, $H = \text{Vect}(e_1 + e_2)$. On a clairement $F \cap G = G \cap H = H \cap F = \{0_E\}$, mais F, G, H ne sont pas en somme directe puisque le vecteur $e_1 + e_2$ peut se décomposer de différentes façons suivant la somme $F + G + H$, par exemple sous la forme $e_1 + e_2 + 0_E$ ou encore sous la forme $0_E + 0_E + (e_1 + e_2)$.

Remarque. Si les sous-espaces E_i, $1 \leq i \leq m$ sont en somme directe, alors toute sous-famille est encore en somme directe.

Preuve. Il suffit de montrer que le caractère direct de la somme persiste si l'on ôte le dernier sous-espace. Montrons donc que $E_1, \cdots, E_{m-1}$ sont en somme directe: si le vecteur nul se décompose en $0_E = x_1 + \cdots + x_{m-1}$ avec $x_i \in E_i$, $1 \leq i \leq m-1$, alors on peut aussi écrire $0_E = x_1 + \cdots + x_{m-1} + 0_E$ avec $0_E \in E_m$. Comme on a supposé $E_1, \cdots, E_m$ en somme directe, on déduit bien que toutes les composantes x_i sont nulles.

Remarque:"associativité de la somme directe". Si les m sous-espaces E_i, $1 \leq i \leq m$, sont en somme directe, et si les deux sous-espaces $S = \bigoplus_{i=1}^{m} E_i$ et E_{m+1} sont eux aussi en somme directe, alors les $m+1$ sous-espaces $E_1, \cdots, E_m, E_{m+1}$ sont en somme directe, propriété que l'on pourra résumer en

$$\left(\bigoplus_{i=1}^{m} E_i \right) \oplus E_{m+1} = \bigoplus_{i=1}^{m+1} E_i \ .$$

Preuve. Supposons $0_E = x_1 + \cdots + x_m + x_{m+1}$ avec $x_i \in E_i$ $(1 \leq i \leq m+1)$. Cette égalité s'écrit aussi $0_E = (x_1 + \cdots + x_m) + x_{m+1}$ avec $x_1 + \cdots + x_m \in S$ et $x_{m+1} \in E_{m+1}$. Ces deux derniers sous-espaces étant en somme directe, on a donc $x_{m+1} = 0_E$ et $x_1 + \cdots + x_m = 0_E$. Enfin, les m sous-espaces E_i, $1 \leq i \leq m$, étant eux aussi en somme directe, la dernière égalité montre que $x_i = 0_E$ pour tout $i \in [\![1, m]\!]$.

Exercice I.2.1. Montrer que m sous-espaces $E_1, \cdots, E_m$ de E sont en somme directe si et seulement si

$$\forall i \in [\![2, m]\!] \qquad E_i \cap \left(\sum_{j=1}^{i-1} E_j \right) = \{0_E\} \ . \qquad \square$$

Exercice I.2.2. Montrer que m sous-espaces $E_1, \cdots, E_m$ de E sont en somme directe si et seulement si

$$\forall i \in [\![1,m]\!] \qquad E_i \ \cap \ \Big(\sum_{j\neq i} E_j\Big) = \{0_E\} \ . \qquad \square$$

Exercice I.2.3. Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel E, on introduit les sous-espaces vectoriels (certains peuvent être réduits à $\{0_E\}$ mais cela n'a pas d'importance):

$$F = \mathrm{Ker}(u - \mathrm{id}_E)\ , \quad G = \mathrm{Ker}(u - 2\ \mathrm{id}_E)\ , \quad H = \mathrm{Ker}(u - 3\ \mathrm{id}_E)\ .$$

Montrer que F, G, H sont en somme directe. $\square$

c. Décomposition d'un espace en somme directe (HP).

Dans ce paragraphe, nous étudions le cas d'un espace vectoriel E qui se décompose en somme directe d'un nombre fini de sous-espaces E_i, $1 \leq i \leq m$, on a donc $E = E_1 \oplus \cdots \oplus E_m$.

Cela signifie concrètement que tout vecteur x de E **admet un unique décomposition** en $x = x_1 + \cdots + x_m$ avec $x_i \in E_i$ pour tout $i \in [\![1,m]\!]$. Les vecteurs x_i sont les **composantes** de x suivant cette décomposition -à ne pas confondre avec le terme de **coordonnées** (dans une base), qui sont des scalaires-. Pour tout $i \in [\![1,m]\!]$, on peut donc considérer l'application p_i qui, à tout vecteur x de E, associe sa composante x_i dans cette décomposition. Cette application est linéaire puisque, si x et y se décomposent respectivement en $x = x_1+\cdots+x_m$ et $y = y_1+\cdots+y_m$, alors $\lambda x+y$ se décompose en $\lambda x+y = (\lambda x_1+y_1)+\cdots+(\lambda x_m+y_m)$. Si on considère p_i comme allant de E vers E, c'est alors un endomorphisme de E, i.e. $p_i \in \mathcal{L}(E)$.

Propriétés. L'endomorphisme p_i est le projecteur sur E_i parallèlement à $\bigoplus_{j\neq i} E_j$. On a les relations

$$(\mathbf{1}): \qquad \forall (i,j) \in [\![1,m]\!]^2 \qquad i \neq j \Longrightarrow p_i \circ p_j = 0\ .$$

$$(\mathbf{2}): \qquad \sum_{i=1}^{m} p_i = \mathrm{id}_E\ .$$

Preuve. *Pour abréger un peu, nous noterons $E'_i = \bigoplus_{j\neq i} E_j$. Par associativité de la somme directe, on a bien $E = E_i \oplus E'_i$ et, si la décomposition d'un vecteur x de E suivant $E = \bigoplus_{j=1}^{m} E_j$ est $x = \sum_{j=1}^{m} x_j$, alors pour tout indice i fixé entre 1 et m, l'écriture $x = x_i + \Big(\sum_{j\neq i} x_j\Big)$ est la décomposition de x suivant $E = E_i \oplus E'_i$, donc $x_i = p_i(x)$ est bien le projeté de x sur E_i parallèlement à E'_i, ce qui confirme la nature géométrique de p_i. La relation* **(2)** *est seulement la traduction de l'identité $x = \sum_{i=1}^{m} x_i = \sum_{i=1}^{m} p_i(x)$ pour tout x. Enfin, le fait que p_i soit le projecteur sur E_i parallèlement à E'_i entraîne que $\mathrm{Im}(p_i) = E_i$ et $\mathrm{Ker}(p_i) = E'_i = \bigoplus_{k\neq i} E_k$. Si i et j sont deux indices distincts, on a alors $E_j \subset E'_i$, soit $\mathrm{Im}(p_j) \subset \mathrm{Ker}(p_i)$, ce qui traduit l'identité* **(1)**: $p_i \circ p_j = 0$.

Les projecteurs p_i, $1 \leq i \leq m$, sont appelés **projecteurs associés à la décomposition** de E en somme directe des E_i, $1 \leq i \leq m$.

<u>**Exercice I.2.4.**</u> Soit E un $\mathbb{K}$-espace vectoriel, soient $p_1, \cdots, p_m$ des endomorphismes de E vérifiant les relations

$$p_i \circ p_j = 0 \quad \text{si} \quad i \neq j \qquad \text{et} \qquad \sum_{i=1}^{m} p_i = \mathrm{id}_E \; .$$

Montrer que $E = \bigoplus_{i=1}^{m} \operatorname{Im} p_i$. Interpréter géométriquement les p_i. $\square$

Voici maintenant un théorème qui dit qu'une application linéaire de E vers un espace vectoriel F est entièrement déterminée par ses restrictions aux m sous-espaces E_i de la décomposition. Plus précisément,

<u>**Théorème.**</u> **Soient $E_1, \cdots, E_m$ des sous-espaces vectoriels de E tels que $E = \bigoplus_{i=1}^{m} E_i$, soit pour tout $i \in [\![1, m]\!]$ une application linéaire u_i de E_i vers un espace vectoriel F. Alors il existe une unique application linéaire u de E vers F telle que $u|_{E_i} = u_i$ pour tout i.**

<u>*Preuve.*</u> *Procédons par analyse-synthèse.*
Analyse: si u existe, alors pour tout x de E, on peut écrire $x = \sum_{i=1}^{m} p_i(x)$, où les p_i sont les projecteurs associés à la décomposition $E = \bigoplus_{i=1}^{m} E_i$. Par linéarité de u, on a nécessairement

$$u(x) = u\Big(\sum_{i=1}^{m} p_i(x)\Big) = \sum_{i=1}^{m} u\big(p_i(x)\big) = \sum_{i=1}^{m} u_i\big(p_i(x)\big)$$

puisque $p_i(x) \in E_i$ et que u doit coïncider avec u_i sur chaque sous-espace E_i. Ceci détermine u et prouve son unicité "sous réserve d'existence".

Synthèse: Soit l'application $u : E \to F$ définie par $\forall x \in E \quad u(x) = \sum_{i=1}^{m} u_i\big(p_i(x)\big)$.
De la linéarité des u_i et des p_i, on déduit la linéarité de u. Enfin, soit $j \in [\![1, m]\!]$, si $x \in E_j$, on a $x = p_j(x)$ puis $u(x) = \sum_{i=1}^{m} u_i\big(p_i \circ p_j(x)\big) = u_j\big(p_j(x)\big) = u_j(x)$, en utilisant le fait que $p_i \circ p_j = 0$ si $i \neq j$ et $p_j \circ p_j = p_j$ (projecteur). Donc u coïncide avec u_j sur chaque sous-espace E_j, autrement dit ce choix de u "convient", on a donc prouvé l'existence.

d. <u>Cas de la dimension finie.</u>

<u>**Théorème.**</u> **Soient $E_1, \cdots, E_m$ des sous-espaces vectoriels de dimension finie d'un espace vectoriel E. On a alors**

$$\dim\Big(\sum_{i=1}^{m} E_i\Big) \leq \sum_{i=1}^{m} \dim(E_i)$$

avec égalité si et seulement si la somme est directe.

Preuve. Considérons l'application $\varphi : \begin{cases} E_1 \times \cdots \times E_m & \to E \\ (x_1, \cdots, x_m) & \mapsto x_1 + \cdots + x_m \end{cases}$.

Cette application est clairement linéaire, et $\mathrm{Im}(\varphi) = \sum_{i=1}^{m} E_i$ *par définition de la somme des sous-espaces* E_i*. Le théorème du rang appliqué à* φ *donne déjà l'inégalité voulue puisque*

$$\dim\Big(\prod_{i=1}^{m} E_i\Big) = \sum_{i=1}^{m} \dim(E_i) = \mathrm{rg}(\varphi) + \dim\big(\mathrm{Ker}(\varphi)\big) \geq \mathrm{rg}(\varphi) = \dim\Big(\sum_{i=1}^{m} E_i\Big).$$

De plus, il y a égalité si et seulement si $\mathrm{Ker}(\varphi) = \{(0_E, \cdots, 0_E)\}$*, c'est-à-dire si et seulement si la seule façon de décomposer le vecteur nul est de prendre toutes les composantes nulles, ce qui est la caractérisation des sommes directes.*

Si, maintenant, E est un espace vectoriel de dimension finie n se décomposant en une somme directe de m sous-espaces: $E = \bigoplus_{i=1}^{m} E_i$, on introduit la notion de **base** de E **adaptée à cette décomposition**. On appelle ainsi toute base de E obtenue par concaténation (juxtaposition) d'une base de E_1, une base de E_2, ..., une base de E_m. Pour détailler un peu plus (mais cela oblige à alourdir les notations), notons d_1, $\cdots$, d_m les dimensions respectives des sous-espaces E_1, $\cdots$, E_m. On a alors $n = \sum_{i=1}^{m} d_i$ puisque E est la somme **directe** des E_i. Pour tout $i \in [\![1, m]\!]$, soit $\mathcal{B}_i = (e_1^{(i)}, \cdots, e_{d_i}^{(i)})$ une base de E_i. La famille de vecteurs $\mathcal{B}$ obtenue par concaténation des bases $\mathcal{B}_i$, $1 \leq i \leq m$ est alors une base de E adaptée à la décomposition $E = \bigoplus_{i=1}^{m} E_i$. On a donc $\mathcal{B} = (e_1^{(1)}, \cdots, e_{d_1}^{(1)}, e_1^{(2)}, \cdots, e_{d_2}^{(2)}, \cdots\cdots, e_1^{(m)}, \cdots, e_{d_m}^{(m)})$.
Le lecteur vérifiera facilement que $\mathcal{B}$ *est effectivement une base de* E*.*
Inversement, si E est de dimension finie n, par partition d'une base $\mathcal{B} = (e_1, \cdots, e_n)$ de E, on obtient une décomposition de E en somme directe. Précisons: si on se donne m entiers naturels non nuls d_1, $\cdots$, d_m dont la somme vaut n, notons E_1 le sous-espace de E engendré par les d_1 premiers vecteurs de la base $\mathcal{B}$, puis E_2 le sous-espace engendré par les d_2 suivants, ..., enfin E_m le sous-espace engendré par les d_m derniers, on a alors $E = \bigoplus_{i=1}^{m} E_i$.

Preuve. Bôf, j'ai pas envie!

Enfin, si E est un espace vectoriel de dimension n, et F un sous-espace vectoriel de E de dimension p (avec $1 \leq p \leq n-1$), on appelle **base** de E **adaptée à** F toute base $\mathcal{B} = (e_1, \cdots, e_n)$ de E dont les p premiers vecteurs appartiennent à F. Ces p premiers vecteurs constituent alors une base $\mathcal{C} = (e_1, \cdots, e_p)$ de F: *la famille est libre puisqu'elle est extraite d'une base de* E*, et vu son cardinal c'est bien une base de* F*.*

Remarque. Avec les notations ci-dessus, une base de E adaptée à F existe toujours: en effet, le sous-espace F admet une base $(e_1, \cdots, e_p)$, qui est alors une famille libre dans E, que l'on peut compléter (théorème de la base incomplète) en une base $(e_1, \cdots, e_p, e_{p+1}, \cdots, e_n)$ de E, qui est alors adaptée à F.

II. Matrices et endomorphismes.

1. Matrices par blocs.

Il arrivera qu'une matrice soit définie **par blocs**, il serait assez fastidieux d'examiner tous les cas de figure possibles, nous nous limiterons donc pour l'essentiel aux définitions en quatre blocs: $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ avec $A \in \mathcal{M}_{p,r}(\mathbb{K})$, $B \in \mathcal{M}_{p,s}(\mathbb{K})$,$C \in \mathcal{M}_{q,r}(\mathbb{K})$ et $D \in \mathcal{M}_{q,s}(\mathbb{K})$ où p, q, r, s sont quatre entiers naturels non nuls. Ainsi, $M \in \mathcal{M}_{p+q,r+s}(\mathbb{K})$. Une telle écriture est compatible avec les opérations de la structure d'espace vectoriel dans $\mathcal{M}_{p+q,r+s}(\mathbb{K})$, i.e. si une autre matrice M' se décompose en quatre blocs avec les mêmes formats que ci-dessus: $M' = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix}$, et si α est un scalaire, on a

$$\alpha M + M' = \begin{pmatrix} \alpha A + A' & \alpha B + B' \\ \alpha C + C' & \alpha D + D' \end{pmatrix} .$$

Notons aussi, pour la transposition, que $M^\top = \begin{pmatrix} A^\top & C^\top \\ B^\top & D^\top \end{pmatrix} \in \mathcal{M}_{r+s,p+q}(\mathbb{K})$.

L'opération la plus intéressante est le produit, mais il ne peut s'effectuer que s'il y a une certaine compatibilité entre les formats.

Proposition. Soient A, B, C, D, E, F, G, H huit matrices de formats (k, n), (k, p), (l, n), (l, p), (n, q), (n, r), (p, q) et (p, r) respectivement, où k, l, n, p, q, r sont des entiers naturels non nuls. On a alors

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} E & F \\ G & H \end{pmatrix} = \begin{pmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{pmatrix} .$$

Ces matrices se multiplient donc "par blocs" de la même façon que si les blocs étaient des coefficients scalaires.

Ne me demandez surtout pas une preuve de cette formule, c'est horrible à écrire (il faut faire des décalages d'indices dans tous les sens) et cela n'a vraiment aucun intérêt!

Exercice II.1.1. Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée, soit $M = \begin{pmatrix} A & A \\ 0_n & A \end{pmatrix} \in \mathcal{M}_{2n}(\mathbb{K})$. Calculer M^k pour tout entier naturel k. $\square$

On s'intéressera tout particulièrement aux matrices **diagonales par blocs**, c'est-à-dire de la forme $M = \begin{pmatrix} A & 0_{n,p} \\ 0_{p,n} & D \end{pmatrix}$ avec $A \in \mathcal{M}_n(\mathbb{K})$ et $D \in \mathcal{M}_p(\mathbb{K})$, ou plus généralement de la forme

$$M = \mathrm{diag}(D_1, \cdots, D_k) = \begin{pmatrix} D_1 & & (0) \\ & \ddots & \\ (0) & & D_k \end{pmatrix} \in \mathcal{M}_{n_1+\cdots+n_k}(\mathbb{K})$$

avec $D_1 \in \mathcal{M}_{n_1}(\mathbb{K})$, ..., $D_k \in \mathcal{M}_{n_k}(\mathbb{K})$.

On rencontrera aussi des matrices **triangulaires par blocs**, c'est-à-dire de la forme $M = \begin{pmatrix} A & B \\ 0_{p,n} & D \end{pmatrix}$, ou bien $M = \begin{pmatrix} A & 0_{n,p} \\ C & D \end{pmatrix}$ avec $A \in \mathcal{M}_n(\mathbb{K})$, $B \in \mathcal{M}_{n,p}(\mathbb{K})$, $C \in \mathcal{M}_{p,n}(\mathbb{K})$, $D \in \mathcal{M}_p(\mathbb{K})$. Dans ces configurations, les blocs diagonaux seront toujours des matrices carrées.

Proposition. **Si** $M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$ **ou** $M = \begin{pmatrix} A & 0 \\ C & D \end{pmatrix}$ **est triangulaire par blocs, les blocs diagonaux** A **et** D **étant des matrices carrées, alors** $\det(M) = \det(A) \cdot \det(D)$**.**

cf. poly de cours sur les déterminants.

2. Sous-espaces stables.

La définition est banale: si u est un endomorphisme d'un espace vectoriel E, et si F est un sous-espace de E, on dit que F est **stable** par u si on a $u(F) \subset F$, i.e. tout vecteur de F a son image dans F.

Si le sous-espace F de E est stable par l'endomorphisme u, on peut définir une application $u_F : \begin{cases} F \to F \\ x \mapsto u_F(x) = u(x) \end{cases}$, qui est clairement linéaire et qui est donc un endomorphisme de F, on l'appelle **endomorphisme de** F **induit par** u.

Remarque. Cette notion est un peu différente de la **restriction** $u|_F$ de u à F qui est, elle, une application linéaire allant de F vers E (et que l'on peut définir sans que F soit stable par u). On a donc

$$u_F \in \mathcal{L}(F) \quad \text{alors que} \quad u|_F \in \mathcal{L}(F, E)\ .$$

Proposition. **Soient** f **et** g **deux endomorphismes d'un espace vectoriel** E **qui commutent (**$f \circ g = g \circ f$**). Alors le noyau et l'image de** f **sont stables par** g**.**

Preuve. • *Montrons que* $\mathrm{Ker}(f)$ *est stable par* g*: soit* $x \in \mathrm{Ker}(f)$*, il faut montrer que* $g(x) \in \mathrm{Ker}(f)$*. Or,* $f\big(g(x)\big) = g\big(f(x)\big) = g(0_E) = 0_E$*, ce qui conclut.*

• *Montrons que* $\mathrm{Im}(f)$ *est stable par* g*: soit* $y \in \mathrm{Im}(f)$*, il faut montrer que* $g(y) \in \mathrm{Im}(f)$*. Or,* y *admet un antécédent* x *par* f*, i.e.* $y = f(x)$*, donc* $g(y) = g\big(f(x)\big) = f\big(g(x)\big) \in \mathrm{Im}(f)$*, ce qu'il fallait prouver.*

Exercice II.2.1. Soit $p \in \mathcal{L}(E)$ un projecteur, soit $g \in \mathcal{L}(E)$. Montrer que g commute avec p si et seulement si les sous-espaces $\mathrm{Ker}(p)$ et $\mathrm{Im}(p)$ sont stables par g. $\square$

Faisons maintenant le lien entre les sous-espaces stables et les matrices par blocs ayant des formes particulières.

• Tout d'abord, si $\dim(E) = n$, si F est un sous-espace vectoriel de E de dimension p, si $\mathcal{B} = (e_1, \cdots, e_p, e_{p+1}, \cdots, e_n)$ est une base de E adaptée à F, enfin si u est un endomorphisme de E, alors le s.e.v. F est stable par u si et seulement si la matrice de u relativement à $\mathcal{B}$ est **triangulaire supérieure par blocs** avec les formats suivants:

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} A & B \\ 0_{n-p,p} & D \end{pmatrix}\ ,\ A \in \mathcal{M}_p(\mathbb{K})\ ,\ B \in \mathcal{M}_{p,n-p}(\mathbb{K})\ ,\ D \in \mathcal{M}_{n-p}(\mathbb{K})\ .$$

En effet, la stabilité de F *par* u *signifie que les images par* u *des vecteurs* $e_1, \cdots, e_p$ *qui engendrent* F *appartiennent encore à* F*, c'est-à-dire ont leurs* $n-p$ *dernières coordonnées nulles dans la base* $\mathcal{B}$*. Et cette condition suffit par linéarité de* u*: si* $u(e_1), \cdots, u(e_p)$ *appartiennent à* F*, alors* F *est stable par* u*.*

Dans le contexte ci-dessus, le bloc A en haut à gauche représente alors l'endomorphisme induit u_F dans la base $\mathcal{C} = (e_1, \cdots, e_p)$ de F, i.e. $A = \mathrm{Mat}_{\mathcal{C}}(u_F)$.

• Si un espace vectoriel E de dimension finie n se décompose en une somme directe $E = \bigoplus_{i=1}^{m} E_i$, avec $\dim(E_i) = d_i$ pour tout $i \in [\![1, m]\!]$, si $\mathcal{B}$ est une base de E adaptée à cette décomposition, enfin si u est un endomorphisme de E, alors u laisse stable chacun des E_i, $1 \leq i \leq m$, si et seulement si la matrice de u relativement à $\mathcal{B}$ est **diagonale par blocs** de la forme $\mathrm{Mat}_{\mathcal{B}}(u) = \mathrm{diag}(D_1, \cdots, D_m)$, où $D_i \in \mathcal{M}_{d_i}(\mathbb{K})$ pour tout i. Dans ce cas, la base $\mathcal{B}$ est obtenue par concaténation de $\mathcal{B}_1, \cdots, \mathcal{B}_m$, où chaque $\mathcal{B}_i$ est une base de E_i, et le i-ème bloc diagonal D_i représente l'endomorphisme de E_i induit par u dans la base $\mathcal{B}_i$, $1 \leq i \leq m$.

Exercice II.2.2. Avec les notations du paragraphe ci-dessus, montrer que l'ensemble des endomorphismes de E laissant stable chacun des E_i, $1 \leq i \leq m$, est un sous-espace vectoriel de $\mathcal{L}(E)$. Quelle est sa dimension ? $\square$

Exercice II.2.3. Soit E un espace vectoriel de dimension n, soit $\mathcal{B} = (e_1, \cdots, e_n)$ une base de E, soit $u \in \mathcal{L}(E)$, soit $M = \mathrm{Mat}_{\mathcal{B}}(u)$ la matrice de u relativement à la base $\mathcal{B}$. Montrer que la matrice M est triangulaire supérieure si et seulement si, pour tout $k \in [\![1, n]\!]$, le sous-espace $F_k = \mathrm{Vect}(e_1, \cdots, e_k)$ est stable par u. Énoncer une condition semblable pour que M soit triangulaire inférieure. $\square$

Exercice II.2.4. Soit E un espace vectoriel de dimension n, soit $p \in \mathcal{L}(E)$ un projecteur de rang r. On définit le **commutant** de p, noté $\mathcal{C}(p)$, comme étant l'ensemble des endomorphismes de E qui commutent avec p:

$$\mathcal{C}(p) = \left\{ f \in \mathcal{L}(E) \mid f \circ p = p \circ f \right\} .$$

Montrer que $\mathcal{C}(p)$ est un sous-espace vectoriel de $\mathcal{L}(E)$. Quelle est sa dimension ? *On utilisera l'exercice II.2.1.* $\square$

3. Trace.

Définition. Si $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ est une matrice carrée, on définit sa **trace** comme étant la somme de ses coefficients diagonaux:

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

Voici quelques propriétés élémentaires.

Linéarité. L'application $\mathrm{tr} : \begin{cases} \mathcal{M}_n(\mathbb{K}) & \to & \mathbb{K} \\ A & \mapsto & \mathrm{tr}(A) \end{cases}$ **est une forme linéaire sur** $\mathcal{M}_n(\mathbb{K})$.

C'est immédiat. On a donc $\mathrm{tr}(\lambda A + B) = \lambda\,\mathrm{tr}(A) + \mathrm{tr}(B)$, pour $\lambda \in \mathbb{K}$ et A, $B \in \mathcal{M}_n(\mathbb{K})$.

Proposition. Une matrice et sa transposée ont la même trace: $\mathrm{tr}(A^\top) = \mathrm{tr}(A)$.

C'est immédiat.

Proposition. Si $A \in \mathcal{M}_n(\mathbb{K})$ **et** $B \in \mathcal{M}_n(\mathbb{K})$**, alors** $\mathrm{tr}(AB) = \mathrm{tr}(BA)$. **Et, plus généralement, si** $A \in \mathcal{M}_{p,q}(\mathbb{K})$ **et** $B \in \mathcal{M}_{q,p}(\mathbb{K})$**, alors** $\mathrm{tr}(AB) = \mathrm{tr}(BA)$**.**

Preuve. Posons $C = AB$ et $D = BA$ et notons avec des lettres minuscules les coefficients de chaque matrice notée en majuscules. Avec $A \in \mathcal{M}_{p,q}(\mathbb{K})$ et $B \in \mathcal{M}_{q,p}(\mathbb{K})$, les matrices $C = AB$ et $D = BA$ sont carrées d'ordres p et q respectivement. On a alors

$$
\begin{aligned}
\operatorname{tr}(AB) &= \sum_{i=1}^{p} c_{i,i} = \sum_{i=1}^{p}\Big(\sum_{j=1}^{q} a_{i,j}b_{j,i}\Big) = \sum_{i=1}^{p}\Big(\sum_{j=1}^{q} b_{j,i}a_{i,j}\Big) \\
&= \sum_{j=1}^{q}\Big(\sum_{i=1}^{p} b_{j,i}a_{i,j}\Big) = \sum_{j=1}^{q} d_{j,j} = \operatorname{tr}(BA)\ .
\end{aligned}
$$

Cette relation permet, lorsque les formats des matrices sont compatibles, de faire des permutations circulaires pour exprimer la trace d'un produit de plusieurs matrices. Par exemple si A, B, C sont carrées de même format, on a

$$\operatorname{tr}(ABC) = \operatorname{tr}\big(A(BC)\big) = \operatorname{tr}\big((BC)A\big) = \operatorname{tr}(BCA)$$

et aussi $\operatorname{tr}(ABC) = \operatorname{tr}(CAB)$ de façon similaire, mais **attention!**, on a en général $\operatorname{tr}(ACB) \neq \operatorname{tr}(ABC)$.

En voici une conséquence fondamentale.

Invariance par similitude. Deux matrices semblables ont la même trace.

Preuve. Si $A \in \mathcal{M}_n(\mathbb{K})$ et $B \in \mathcal{M}_n(\mathbb{K})$ sont semblables, alors il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $B = P^{-1}AP$, on en déduit que

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

Cette dernière propriété permet de définir la trace d'un endomorphisme en dimension finie:

Définition. Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel E de dimension finie n, soit $\mathcal{B}$ une base de E. La trace de la matrice représentant u dans la base $\mathcal{B}$ ne dépend pas du choix de la base $\mathcal{B}$, on l'appelle alors **trace de l'endomorphisme** u, on la note $\operatorname{tr}(u)$.

Preuve. En effet, si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de E, les matrices $M = \mathrm{Mat}_{\mathcal{B}}(u)$ et $M' = \mathrm{Mat}_{\mathcal{B}'}(u)$ sont semblables, donc elles ont la même trace. Cela a donc un sens de poser $\operatorname{tr}(u) = \operatorname{tr}\big(\mathrm{Mat}_{\mathcal{B}}(u)\big)$, où $\mathcal{B}$ est une quelconque base de E.

Les propriétés de la trace d'un endomorphisme se déduisent de celles vues pour la trace d'une matrice.

Linéarité. Si E est un espace vectoriel de dimension finie, l'application $\operatorname{tr} : \begin{cases} \mathcal{L}(E) \to \mathbb{K} \\ u \mapsto \operatorname{tr}(u) \end{cases}$ est une forme linéaire sur $\mathcal{L}(E)$. On a donc $\operatorname{tr}(\lambda u + v) = \lambda \operatorname{tr}(u) + \operatorname{tr}(v)$.

Proposition. Si u et v sont deux endomorphismes d'un espace vectoriel E de dimension finie, alors $\operatorname{tr}(u \circ v) = \operatorname{tr}(v \circ u)$.

Exemple. Trace d'un projecteur. Soit p un projecteur dans un espace vectoriel E de dimension finie n, soit r son rang. On sait que $E = \mathrm{Im}(p) \oplus \mathrm{Ker}(p)$. Si $\mathcal{B} = (e_1, \cdots, e_r, e_{r+1}, \cdots, e_n)$ est une base de E adaptée à cette décomposition, les vecteurs de $\mathrm{Im}(p)$ étant invariants par p, on a $\mathrm{Mat}_{\mathcal{B}}(p) = J_r = \begin{pmatrix} I_r & 0_{r,n-r} \\ 0_{n-r,r} & 0_{n-r} \end{pmatrix}$, donc $\operatorname{tr}(p) = r = \mathrm{rg}(p)$.

La trace d'un projecteur (en dimension finie) est égale à son rang.

Exercice II.3.1. Soit E un espace vectoriel de dimension n, soit $\mathcal{B} = (e_1, \cdots, e_n)$ une base de E, soit $\mathcal{B}^* = (e_1^*, \cdots, e_n^*)$ sa base duale (e_i^* est la i-ème forme linéaire coordonnée relativement à la base $\mathcal{B}$). Montrer que, pour tout endomorphisme u de E, on a

$$\mathrm{tr}(u) = \sum_{i=1}^{n} e_i^*\big(u(e_i)\big) . \qquad \square$$

Exercice II.3.2. Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice, soit l'application $\tau_A : \begin{cases} \mathcal{M}_n(\mathbb{K}) & \to & \mathbb{K} \\ M & \mapsto & \mathrm{tr}(AM) \end{cases}$.

Vérifier que τ_A est une forme linéaire sur $\mathcal{M}_n(\mathbb{K})$. Calculer $\tau_A(E_{i,j})$, où $E_{i,j}$ est une matrice élémentaire. En déduire que l'application $\mathcal{T} : \begin{cases} \mathcal{M}_n(\mathbb{K}) & \to & \big(\mathcal{M}_n(\mathbb{K})\big)^* \\ A & \mapsto & \tau_A \end{cases}$ est un isomorphisme, puis que toute forme linéaire sur $\mathcal{M}_n(\mathbb{K})$ est de la forme $M \mapsto \mathrm{tr}(AM)$, où $A \in \mathcal{M}_n(\mathbb{K})$ est une matrice fixée. $\square$

III. Polynômes d'endomorphismes et de matrices.

1. Définition. Propriétés opératoires.

Si u est un endomorphisme d'un $\mathbb{K}$-espace vectoriel E, et si $P = \sum_{k=0}^{d} a_k X^k \in \mathbb{K}[X]$ est un polynôme, on peut considérer l'endomorphisme v de E tel que $v = \sum_{k=0}^{d} a_k u^k$, obtenu par substitution de l'endomorphisme u à l'indéterminée X dans l'écriture formelle du polynôme P. *On rappelle que, par convention,* $u^0 = \mathrm{id}_E$. Cet endomorphisme v est noté $P(u)$. Ainsi,

$$P(u) = \sum_{k=0}^{d} a_k u^k = a_0 \,\mathrm{id}_E + a_1 u + \cdots + a_d u^d .$$

On dit souvent que $P(u)$ est un "polynôme de l'endomorphisme u" mais **attention** à ne pas se méprendre: $P(u)$ n'est pas un polynôme, c'est un endomorphisme de E!

Attention! Si x est un vecteur de E, on peut donc appliquer l'endomorphisme $P(u)$ à ce vecteur x, ce qui donne un vecteur que l'on notera $P(u)(x)$, et **surtout pas** $P\big(u(x)\big)$, ce dernier parenthésage n'ayant aucun sens! Dans le même ordre d'idées, la notation $u^2(x)$ est cohérente et correspond à $(u \circ u)(x) = u\big(u(x)\big)$, tandis que la notation $u(x)^2$ est aberrante!

De même, si $A \in \mathcal{M}_n(\mathbb{K})$ est une matrice carrée d'ordre n, et si $P = \sum_{k=0}^{d} a_k X^k \in \mathbb{K}[X]$ est un polynôme, on peut considérer la matrice $B \in \mathcal{M}_n(\mathbb{K})$ telle que $B = \sum_{k=0}^{d} a_k A^k$, obtenue par substitution de la matrice A à l'indéterminée X dans l'écriture formelle du polynôme P. *On rappelle que, par convention,* $A^0 = I_n$. Cette matrice B est notée $P(A)$. Ainsi,

$$P(A) = \sum_{k=0}^{d} a_k A^k = a_0 I_n + a_1 A + \cdots + a_d A^d .$$

On dit souvent que $P(A)$ est un "polynôme de la matrice A" mais **attention** à ne pas se méprendre: $P(A)$ n'est pas un polynôme, c'est une matrice carrée d'ordre n.

<u>Proposition.</u> Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel E. L'application

$$\Phi_u : \begin{cases} \mathbb{K}[X] & \to & \mathcal{L}(E) \\ P & \mapsto & P(u) \end{cases}$$

est linéaire. *La preuve est peu intéressante!*

On a donc, pour tous $P \in \mathbb{K}[X]$, $Q \in \mathbb{K}[X]$ et $\lambda \in \mathbb{K}$, $\Phi_u(\lambda P + Q) = \lambda\, \Phi_u(P) + \Phi_u(Q)$, soit

$$(\lambda P + Q)(u) = \lambda\, P(u) + Q(u)\ .$$

De même, avec des matrices carrées au lieu des endomorphismes,

<u>Proposition.</u> Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée d'ordre n. L'application

$$\Phi_A : \begin{cases} \mathbb{K}[X] & \to & \mathcal{M}_n(\mathbb{K}) \\ P & \mapsto & P(A) \end{cases}$$

est linéaire. *La preuve est peu intéressante!*

On a donc, pour tous $P \in \mathbb{K}[X]$, $Q \in \mathbb{K}[X]$ et $\lambda \in \mathbb{K}$, $\Phi_A(\lambda P + Q) = \lambda\, \Phi_A(P) + \Phi_A(Q)$, soit

$$(\lambda P + Q)(A) = \lambda\, P(A) + Q(A)\ .$$

On a aussi des propriétés opératoires concernant les produits de polynômes.

<u>Proposition.</u> Soit $u \in \mathcal{L}(E)$, où E est un $\mathbb{K}$-espace vectoriel, soient P et Q deux polynômes de $\mathbb{K}[X]$, on a alors l'égalité

$$(PQ)(u) = P(u) \circ Q(u)\ .$$

De même, si $A \in \mathcal{M}_n(\mathbb{K})$, on a $(PQ)(A) = P(A) \cdot Q(A)$.

<u>Preuve.</u> Bôf!

Comme $\mathbb{K}[X]$ est un anneau commutatif (i.e. la multiplication des polynômes est commutative), on déduit que

$$P(u) \circ Q(u) = (PQ)(u) = (QP)(u) = Q(u) \circ P(u)\ .$$

De même, avec $A \in \mathcal{M}_n(\mathbb{K})$, on a $P(A) \cdot Q(A) = (PQ)(A) = (QP)(A) = Q(A) \cdot P(A)$.

Ainsi, **deux polynômes d'un même endomorphisme u commutent**. De même, **deux polynômes d'une même matrice A commutent.**

<u>Remarque.</u> Si A et B sont deux matrices carrées semblables, alors pour tout polynôme $P \in \mathbb{K}[X]$, les matrices $P(A)$ et $P(B)$ sont elles aussi semblables, et toujours avec la même matrice de passage: *en effet, partons de la relation $B = S^{-1}AS$ avec $S \in \mathrm{GL}_n(\mathbb{K})$. Une récurrence immédiate donne $B^k = S^{-1}A^kS$ pour tout $k \in \mathbb{N}$, puis si $P = \sum_{k=0}^{d} a_kX^k$ est un polynôme,*

$$P(B) = \sum_{k=0}^{d} a_kB^k = \sum_{k=0}^{d} a_k\, S^{-1}A^kS = S^{-1}\left(\sum_{k=0}^{d} a_kA^k\right)S = S^{-1}\, P(A)\, S\ .$$

2. Polynômes annulateurs.

Définition. Soit $u \in \mathcal{L}(E)$, où E est un $\mathbb{K}$-espace vectoriel. On appelle **polynôme annulateur** de u tout polynôme $P \in \mathbb{K}[X]$ tel que $P(u)$ soit l'endomorphisme nul.

Définition. Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée d'ordre n. On appelle **polynôme annulateur** de A tout polynôme $P \in \mathbb{K}[X]$ tel que $P(A)$ soit la matrice nulle.

Exemples. Un endomorphisme $u \in \mathcal{L}(E)$ est un projecteur si et seulement s'il admet $X^2 - X$ comme polynôme annulateur. C'est une symétrie si et seulement s'il admet $X^2 - 1$ comme polynôme annulateur. L'endomorphisme u est nilpotent si et seulement s'il admet un polynôme annulateur de la forme X^k avec $k \in \mathbb{N}^*$.

Exercice III.2.1.(archi-méga classique). Soit E un espace vectoriel de dimension n, soit $u \in \mathcal{L}(E)$ un endomorphisme nilpotent. Montrer que le polynôme X^n annule u. $\square$

Proposition. Soit $u \in \mathcal{L}(E)$ où E est un $\mathbb{K}$-espace vectoriel. On note $\mathcal{I}_u$ l'ensemble des polynômes annulateurs de u, soit $\mathcal{I}_u = \{P \in \mathbb{K}[X] \mid P(u) = 0\}$. Alors $\mathcal{I}_u$ est un sous-espace vectoriel de $\mathbb{K}[X]$, et on a la propriété

$$\forall P \in \mathcal{I}_u \quad \forall Q \in \mathbb{K}[X] \qquad PQ \in \mathcal{I}_u \ ,$$

que l'on peut énoncer en disant: "tout multiple d'un polynôme annulateur est encore annulateur". On peut dire la même chose de l'ensemble $\mathcal{I}_A$ des polynômes annulateurs d'une matrice carrée A.

Preuve. En fait, $\mathcal{I}_u = \mathrm{Ker}(\Phi_u)$, où Φ_u est l'application linéaire introduite au paragraphe précédent, et le noyau d'une application linéaire est bien un s.e.v. de l'espace de départ. De même, $\mathcal{I}_A = \mathrm{Ker}(\Phi_A)$.

Enfin, si $P \in \mathcal{I}_u$ et si Q est un polynôme quelconque, alors

$$(PQ)(u) = P(u) \circ Q(u) = 0_{\mathcal{L}(E)} \circ Q(u) = 0_{\mathcal{L}(E)} \ ,$$

ce qui prouve bien que $PQ \in \mathcal{I}_u$.

Exercice III.2.2. Soit $A \in \mathcal{M}_n(\mathbb{K})$. Montrer que la famille $(I_n, A, A^2, \cdots, A^{n^2})$ est liée dans $\mathcal{M}_n(\mathbb{K})$. En déduire que A admet un polynôme annulateur non nul. $\square$

Exercice III.2.3. Soit la matrice $A = \begin{pmatrix} 2 & 4 & 0 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{pmatrix}$. Calculer A^n, puis $P(A)$ où $P \in \mathbb{R}[X]$ est un polynôme quelconque. *On exprimera les coefficients de la matrice $P(A)$ à l'aide des nombres $P(2)$, $P'(2)$ et $P''(2)$.* Préciser l'ensemble $\mathcal{I}_A$ des polynômes annulateurs de A. $\square$

Exercice III.2.4. Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel E, soit $d \in \mathbb{N}^*$. On suppose que u admet un polynôme annulateur de degré d, mais qu'il n'en admet pas de degré $d-1$. On note $\mathbb{K}[u]$ l'ensemble des endomorphismes de E qui sont des polynômes en u, i.e. $\mathbb{K}[u] = \{P(u)\ ;\ P \in \mathbb{K}[X]\}$. Montrer que $\mathbb{K}[u] = \mathrm{Vect}(\mathrm{id}_E, u, \cdots, u^{d-1})$ et que $\dim\big(\mathbb{K}[u]\big) = d$. $\square$

3. Applications des polynômes annulateurs.

• Soit $A \in \mathcal{M}_n(\mathbb{K})$. On suppose que A admet un polynôme annulateur P tel que $P(0) \neq 0$, i.e. le terme constant du polynôme P est non nul. Alors A est inversible, et on peut exprimer A^{-1} comme un polynôme de la matrice A.

En effet, posons $P = a_0 + a_1 X + \cdots + a_d X^d$ avec $a_0 \neq 0$ et $a_d \neq 0$. La relation $P(A) = 0$, à savoir $a_0 I_n + a_1 A + \cdots + a_d A^d = 0$, peut se mettre sous la forme

$$I_n = -\frac{1}{a_0}\left(a_1 I_n + a_2 A + \cdots + a_d A^{d-1}\right) A \;,$$

donc A est inversible et $A^{-1} = Q(A)$ avec $Q = -\dfrac{1}{a_0}\left(a_1 + a_2 X + \cdots + a_d X^{d-1}\right)$.

• Soit $A \in \mathcal{M}_n(\mathbb{K})$, soit $P \in \mathbb{K}[X]$ un polynôme annulateur de A de degré d. Si $F \in \mathbb{K}[X]$ est un polynôme quelconque, on peut poser la division euclidienne de F par P, cela s'écrit $F = PQ + R$ avec $\deg(R) < d$. En substituant la matrice A à l'indéterminée X dans cette identité polynomiale, on obtient la relation (entre matrices)

$$F(A) = P(A) \cdot Q(A) + R(A) = R(A) \quad \text{puisque} \quad P(A) = 0 \;.$$

Comme $R \in \mathbb{K}_{d-1}[X]$, on voit que toute matrice qui est un polynôme de la matrice A est en fait combinaison linéaire de I_n, A, A^2, $\cdots$, A^{d-1}. L'ensemble, noté $\mathbb{K}[A]$, des matrices "polynômes de la matrice A" est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$, par exemple parce que c'est l'image de l'application linéaire Φ_A introduite au paragraphe 1, et il est de dimension au plus d puisqu'il coïncide finalement avec $\mathrm{Vect}(I_n, A, A^2, \cdots, A^{d-1})$.

De plus, si le polynôme annulateur P admet d racines distinctes (i.e. s'il est scindé à racines simples), alors on peut (en résolvant un système linéaire, *cf.* exercice **III.3.1.** ci-dessous) expliciter pour tout $k \in \mathbb{N}$ le reste R_k de la division euclidienne du polynôme X^k par P, et donc expliciter en fonction de l'entier naturel k les puissances A^k de la matrice A, puisque $A^k = R_k(A)$.

Exercice III.3.1. Soit $A = \begin{pmatrix} 1 & 0 & -3 \\ 0 & -2 & 0 \\ -3 & 0 & 1 \end{pmatrix}$. Trouver un polynôme annulateur de A de degré deux, chercher ses racines. En déduire l'expression de A^k pour tout entier naturel k. Montrer que A est inversible et exprimer A^{-1} comme polynôme de la matrice A. $\square$

Exercice III.3.2. Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ vérifie la relation

$$(A - I_n)^2 (A - 2I_n) = 0 \;.$$

Pour tout $k \in \mathbb{N}$, exprimer A^k en fonction de l'entier k et des matrices I_n, A et A^2. $\square$