Algèbre linéaire, réduction

Olivier Fouquet

Un ingénieur vient écouter son meilleur ami, un mathématicien, donner une conférence sur les propriétés géométriques des espaces de dimension 23. A plusieurs reprises, l'orateur justifie l'une des ses assertions en mentionnant que “on voit bien que...” ou alors “qu'il est visuellement évident que...” A la fin de l'exposé, l'ingénieur demande à son ami comment il réussit à visualiser un espace de dimension 23. “C'est facile” lui répond ce dernier “je visualise en dimension n et ensuite je prends $n = 23$”.

Table des matières

1 Bases de l'algèbre linéaire

1.1 Espaces vectoriels

1.1.1 Motivation et premiers exemples

L'algèbre linéaire est l'étude abstraite des espaces vectoriels et des applications linéaires entre espaces vectoriels. A ce titre, il s'agit d'une généralisation et d'une formalisation de certains aspects de la géométrie classique. De manière informelle, un espace vectoriel est un ensemble non-vide d'éléments, que l'on appelle donc des vecteurs, qui vérifie les propriétés suivantes.

1. On peut additionner deux vecteurs de E. Plus précisément, si u et v sont deux vecteurs de l'espace vectoriel E, alors $u + v$ est un vecteur de l'espace vectoriel E.

2. On peut multiplier un vecteur de E par un scalaire. Plus précisément, si u est un vecteur de E et λ est un élément d'un corps K, alors λu est un vecteur de E.

Bien entendu, beaucoup de termes restent imprécis dans la définition ci-dessus, à commencer par *scalaire*. Afin de fixer les idées, nous encourageons le lecteur à supposer dorénavant que la lettre K désigne l'ensemble $\mathbb{Q}$ des nombres rationnels, l'ensemble $\mathbb{R}$ des nombres réels ou l'ensemble $\mathbb{C}$ des nombres complexes. Les lecteurs d'humeur audacieuse pourront explorer le texte dans le cas plus général où K est un corps quelconque, c'est-à-dire un anneau commutatif non nul dont tous les éléments non-nuls admettent un inverse multiplicatifs (ce qui est bien le cas de $\mathbb{Q}, \mathbb{R}$ et $\mathbb{C}$). Avant de donner une définition précise, explorons quelques exemples d'espaces vectoriels.

Exemples d'espaces vectoriels

1. L'ensemble $\mathbb{Q}$ est un $\mathbb{Q}$-espace vectoriel. En effet, l'addition de deux nombres rationnels u et v est bien un nombre rationnel $w = u + v$ et la multiplication d'un nombre rationnel u par un nombre rationnel λ (que l'on voit ici comme un scalaire) est bien un nombre rationnel $v = \lambda u$. De même, $\mathbb{R}$ est un $\mathbb{R}$-espace vectoriel et $\mathbb{C}$ est un $\mathbb{C}$-espace vectoriel. En revanche, $\mathbb{Q}$ n'est pas un $\mathbb{R}$-espace vectoriel car si $u \in \mathbb{Q}$ est un vecteur non-nul, alors $\sqrt{2}u$ n'est pas un nombre rationnel bien que $\sqrt{2} \in \mathbb{R}$ soit un scalaire. De même $\mathbb{R}$ n'est pas un $\mathbb{C}$-espace vectoriel car si $u \in \mathbb{R}$ est un vecteur non-nul, alors $i \cdot u$ n'est pas un nombre réel bien que $i \in \mathbb{C}$ soit un scalaire (ici i désigne une solution de l'équation $X^2 + 1 = 0$). Notons enfin que $\mathbb{R}$ est un $\mathbb{Q}$-espace vectoriel et que $\mathbb{C}$ est un $\mathbb{Q}$-espace vectoriel et un $\mathbb{R}$-espace vectoriel.

2. Le plan réel
$$\mathbb{R}^2 = \{(x, y) | x \in \mathbb{R},\ y \in \mathbb{R}\}$$
est un $\mathbb{R}$-espace vectoriel. Un vecteur de $\mathbb{R}^2$ est un vecteur au sens géométrique usuel du terme. En particulier, un vecteur $u \in \mathbb{R}^2$ a deux coordonnées et

peut s'écrire $u = \begin{pmatrix} a \\ b \end{pmatrix}$. Notons que le plan $\mathbb{R}^2$ est aussi un $\mathbb{Q}$-espace vectoriel, mais un $\mathbb{Q}$-espace vectoriel extrêmement compliqué, et que c'est un $\mathbb{C}$-espace vectoriel à condition de poser $(\alpha + \beta i) \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \alpha a - \beta b \\ \alpha b + \beta a \end{pmatrix}$. Le plan $\mathbb{R}^2$ muni de sa structure de $\mathbb{R}$-espace vectoriel est un bon modèle de la géométrie plane usuelle.

3. L'espace tridimensionnel réel
$$\mathbb{R}^3 = \{(x, y, z) | x \in \mathbb{R},\ y \in \mathbb{R},\ z \in \mathbb{R}\}$$
est un $\mathbb{R}$-espace vectoriel. Un vecteur de $\mathbb{R}^2$ est un vecteur de l'espace au sens géométrique usuel du terme. En particulier, un vecteur $u \in \mathbb{R}^3$ a trois coordonnées et peut s'écrire
$$u = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$
Tout comme $\mathbb{R}^2$, l'espace $\mathbb{R}^3$ est aussi un $\mathbb{Q}$-espace vectoriel, mais ce n'est pas un $\mathbb{C}$-espace vectoriel. L'espace $\mathbb{R}^3$ muni de sa structure de $\mathbb{R}$-espace vectoriel est un bon modèle de la géométrie usuelle de notre espace physique.

4. L'ensemble
$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 | x + y + z = 0 \right\} \subset \mathbb{R}^3$$
est un $\mathbb{R}$-espace vectoriel.

5. L'ensemble
$$\mathbb{R}^4 = \{(x_1, x_2, x_3, x_4) | \forall\ 1 \le i \le 4,\ x_i \in \mathbb{R}\}$$
est un $\mathbb{R}$-espace vectoriel. Contrairement à $\mathbb{R}$, $\mathbb{R}^2$ et $\mathbb{R}^3$, il n'admet pas d'interprétation géométrique évidente.

6. Plus généralement, si n est un entier strictement positif, l'ensemble
$$\mathbb{R}^n = \{(x_1, \cdots, x_n) | \forall\ 1 \le i \le n,\ x_i \in \mathbb{R}\}$$
est un $\mathbb{R}$-espace vectoriel.

7. Plus généralement encore, si K est un corps (par exemple $K = \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$) et si n est un entier strictement positif, l'ensemble
$$K^n = \{(x_1, x_2, \ldots, x_n) | \forall\ 1 \le i \le n,\ x_i \in K\}$$
est un K-espace vectoriel.

8. L'ensemble
$$\mathbb{R}[X] = \left\{ \sum_{n=0}^{\infty} a_n X^n | \forall n, a_n \in \mathbb{R},\ a_n = 0 \text{ pour presque tout } n \right\}$$
des polynômes à coefficients dans $\mathbb{R}$ est un $\mathbb{R}$-espace vectoriel.

9. Si n est un entier positif, l'ensemble

$$\mathbb{R}_n[X] = \left\{ \sum_{m=0}^{\infty} a_m X^m | \forall m, a_m \in \mathbb{R},\ \forall\ m > n,\ a_m = 0 \right\} \subset \mathbb{R}[X]$$

des polynômes de degré au plus n est un $\mathbb{R}$-espace vectoriel.

10. Plus généralement, si K est un corps, l'ensemble

$$K[X] = \left\{ \sum_{n=0}^{\infty} a_n X^n | a_n \in K,\ a_n = 0 \text{ pour presque tout } n \right\}$$

des polynômes à coefficients dans K est un K-espace vectoriel. L'ensemble

$$K_d[X] = \left\{ \sum_{n=0}^{\infty} a_n X^n | a_n \in K,\ a_n = 0, \forall\ n > d \right\}$$

des polynômes à coefficients dans K de degré au plus d est un K-espace vectoriel.

11. Si K est un corps (par exemple $K = \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$), l'ensemble

$$K[X] = \left\{ \sum_{n=0}^{\infty} a_n X^n | a_n \in K \right\}$$

des séries entières à coefficients dans K est un K-espace vectoriel.

12. L'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ est un $\mathbb{R}$-espace vectoriel.

13. Plus généralement, si U est un ensemble et n est un entier strictement positif, l'ensemble des fonctions de U vers $\mathbb{R}^n$ est un $\mathbb{R}$-espace vectoriel.

14. Plus généralement encore, si U est un ensemble, n est un entier strictement positif et K est un corps, l'ensemble des fonctions de U vers K^n est un K-espace vectoriel.

15. L'ensemble des fonctions dérivables de $\mathbb{R}$ dans $\mathbb{R}$ est un $\mathbb{R}$-espace vectoriel. Plus généralement, si n est un entier ou ∞ et U un intervalle de $\mathbb{R}$, l'ensemble des fonctions de U dans $\mathbb{R}$ de classe C^n est un $\mathbb{R}$-espace vectoriel.

16. L'ensemble des fonctions f de $\mathbb{R}$ dans $\mathbb{R}$ satisfaisant l'équation différentielle

$$f' + \cos(x) f = 0$$

est un $\mathbb{R}$-espace vectoriel.

17. Plus généralement, si les g_n sont des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, l'ensemble des fonctions f de $\mathbb{R}$ dans $\mathbb{R}$ satisfaisant l'équation différentielle

$$\sum_{n=0}^{d} g_n(x) \frac{d^n f}{dx^n} = 0$$

est un $\mathbb{R}$-espace vectoriel.

18. L'ensemble
$$M_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} | (a,b,c,d) \in \mathbb{R}^4 \right\}$$
des matrices de taille 2×2 à coefficients réels est un $\mathbb{R}$-espace vectoriel.

19. Plus généralement, si K est un corps et n est un entier strictement positif, l'ensemble
$$M_n(K) = \left\{ \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & a_{i,j} & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix} | \forall\, 1 \leq i,j \leq n, a_{i,j} \in K \right\}$$
des matrices de taille $n \times n$ à coefficients dans K est un K-espace vectoriel.

20. Plus généralement, si K est un corps et n, m sont des entiers strictement positifs, l'ensemble
$$M_{n,m}(K) = \left\{ \begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & a_{i,j} & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix} | \forall\, 1 \leq i \leq n, 1 \leq j \leq m, a_{i,j} \in K \right\}$$
des matrices de taille $n \times m$ à coefficients dans K est un K-espace vectoriel.

21. Plus généralement, si E est un K-espace vectoriel, l'ensemble $\mathrm{End}(E)$ des applications linéaires de E dans E est un K-espace vectoriel.

22. Plus généralement encore, si E et F sont des K-espaces vectoriels, l'ensemble $\mathrm{Hom}(E,F)$ des applications linéaires de E à valeurs dans F est un K-espace vectoriel.

23. Soit K un corps (par exemple $K = \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$). L'ensemble $\{0\}$ est un K-espace vectoriel que l'on appelle l'espace vectoriel nul.

Exemple d'ensemble qui ne sont pas des espaces vectoriels

1. L'ensemble vide $\varnothing$ n'est pas un espace vectoriel. En effet, un espace vide contient toujours au moins un élément, à savoir 0_E.

2. L'ensemble
$$C = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 | x^2 + y^2 = 1 \right\} \subset \mathbb{R}^2$$
n'est pas un espace vectoriel. En effet, $u = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $v = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ appartiennent à C mais $u + v = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ n'appartient pas à C.

3. L'ensemble
$$X = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 | x \geq 0, \; y \geq 0, z \geq 0 \right\}$$
n'est pas un $\mathbb{R}$-espace vectoriel. En effet, si $u \in X$ est un vecteur non-nul, alors $-u \notin X$ donc X n'est pas stable par multiplication par le scalaire -1.

4. L'ensemble
$$P = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{Q}^3 | x + y + z = 0 \right\} \subset \mathbb{R}^3$$
n'est pas un $\mathbb{R}$-espace vectoriel. En effet, si $u \in P$ est un vecteur non-nul, alors $\sqrt{2}u \notin P$ donc P n'est pas stable par multiplication par le scalaire $\sqrt{2}$.

5. L'ensemble des polynômes ne s'annulant pas en zéro. En effet, cet ensemble n'est pas stable par addition.

6. L'ensemble
$$D = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 | x = y \right\} \cup \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 | x = -y \right\} \subset \mathbb{R}^2$$
n'est pas un $\mathbb{R}$-espace vectoriel. En effet, $u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $v = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ appartiennent à D mais $u + v = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ n'appartient pas à D.

1.1.2 Définitions formelles

Soit K un corps.

Espaces vectoriels

Définition 1.1. *Un ensemble E est un K-espace vectoriel si et seulement s'il vérifie les propriétés suivantes.*

1. *Il existe une loi de composition interne*
$$\begin{aligned} + : \quad E \times E &\longrightarrow E \\ (u, v) &\longmapsto u + v \end{aligned}$$
faisant de E un groupe commutatif (ceci signifie que $+$ est associative, commutative, admet un élément neutre 0_E et que tout élément $u \in E$ admet un inverse $v = -u \in E$ tel que $u + v = 0_E$).

2. *Il existe une opération de multiplication par un scalaire*
$$\begin{aligned} \cdot : \quad K \times E &\longrightarrow E \\ (\lambda, u) &\longmapsto \lambda \cdot u \end{aligned}$$

distributive sur + (ceci signifie que $\lambda \cdot (u+v) = \lambda \cdot u + \lambda \cdot v$ et $(\lambda+\mu) \cdot u = \lambda \cdot u + \mu \cdot u$ pour tout $(\lambda, \mu, u, v,) \in K^2 \times E^2$), mixte associative (ceci signifie que $(\lambda\mu) \cdot u = \lambda \cdot (\mu \cdot u)$ pour tout $(\lambda, \mu, u) \in K^2 \times E$) et telle que $1 \cdot u = u$ pour tout $u \in E$.

Il résulte de la définition que $0 \cdot u = 0_E$ et que $\lambda \cdot 0_E = 0_E$ pour tout $(\lambda, u) \in K \times E$. En effet, $0 \cdot u = (0+0) \cdot u = 0 \cdot u + 0 \cdot u$ donc $0 \cdot u = 0_E$ en ajoutant $-0 \cdot u$ des deux côtés de l'équation et de même $\lambda \cdot 0_E = \lambda \cdot (0_E + 0_E) = \lambda \cdot 0_E + \lambda \cdot 0_E$ donc $\lambda \cdot 0_E = 0_E$. On en déduit que $(-1) \cdot u = -u$ pour tout $u \in E$ et donc que l'on peut confondre sans trop de risque d'erreur 0 et 0_E. Pour cette raison, on note à partir de maintenant λu pour $\lambda \cdot u$ et 0 pour 0_E.

Combinaisons linéaires Soit I un ensemble fini (par exemple $I = \{1, \cdots, n\}$), $(u_i)_{i \in I} \in E^I$ une famille de vecteurs d'un K-espace vectoriel E et $(\lambda_i)_{i \in I} \in K^I$ une famille de scalaires. Il résulte immédiatement des deux propriétés fondamentales de + et $\cdot$ que

$$v = \sum_{i \in I} \lambda_i u_i$$

est un vecteur de E. En effet, $\lambda_i u_i$ appartient à E pour tout $i \in I$ puis une récurrence immédiate montre que c'est également le cas de leur somme. Plus généralement, soit I un ensemble infini, $(u_i)_{i \in I} \in E^I$ une famille de vecteurs d'un K-espace vectoriel E et $(\lambda_i)_{i \in I} \in K^I$ une famille de scalaires. L'élément

$$v = \sum_{i \in I_0} \lambda_i u_i$$

est un vecteur de E pour tout sous-ensemble $I_0 \subset I$ de cardinal fini. On appelle une telle somme une K-combinaison linéaire de vecteurs de E indexée par I. Ce n'est pas une exagération de dire que l'algèbre linéaire est l'étude des combinaisons linéaires.

On prêtera attention au fait qu'une combinaison linéaire est toujours indexée par un ensemble fini même si l'ensemble I peut être infini.

Sous-espaces vectoriels Un sous-espace vectoriel E de F est un sous-ensemble d'un K-espace vectoriel F qui est aussi un espace vectoriel (pour les mêmes lois + et $\cdot$). Pour qu'un ensemble E soit un K-espace vectoriel, il suffit qu'il soit un sous-ensemble d'un K-espace vectoriel F et qu'il satisfasse de plus aux trois propriétés suivantes.

1. E est non-vide.
2. $\forall (u, v) \in E^2,\ (u+v) \in E$.
3. $\forall (\lambda, u) \in K \times E,\ \lambda u \in E$.

En conséquence, une intersection arbitraire de sous-espaces vectoriels d'un espace vectoriel F est un sous-espace vectoriel de F.

Espace vectoriel engendré Soit U un sous-ensemble d'un espace vectoriel E. On appelle espace vectoriel engendré par U et on désigne par $\operatorname{Vect} U$ l'intersection des sous-espaces vectoriels de E contenant U. Il résulte de cette définition et du paragraphe précédent que $\operatorname{Vect} U$ est un sous-espace vectoriel de E (et donc en particulier qu'il est non-vide même si U est vide). Si $(u_i)_{i\in I}$ est une famille de vecteurs de E indexée par un ensemble I, on note $\operatorname{Vect}(u_i)_{i\in I}$ pour $\operatorname{Vect}\{u_i|i \in I\}$. Si $U = \{u_1, \cdots, u_n\}$ est un ensemble fini de vecteurs de E, on note $\operatorname{Vect}(u_1, \cdots, u_n)$ pour $\operatorname{Vect} U$. L'ensemble des K-combinaisons linéaires de vecteurs de U forme un K-sous-espace vectoriel contenant U et contenu dans $\operatorname{Vect} U$. Par définition de $\operatorname{Vect} U$, $\operatorname{Vect} U$ est donc l'ensemble des K-combinaisons linéaires de vecteurs de U.

Familles génératrices, familles libres, dimension Une famille $(u_i)_{i\in I}$ de vecteurs d'un K-espace vectoriel E est dite génératrice si $\operatorname{Vect}(u_i)_{i\in I}$ est égal à E. Un K-espace vectoriel est dit de dimension finie s'il admet une famille génératrice indexée par un ensemble I de cardinal fini. Une famille générative est dite minimale si elle n'admet pas de sous-famille strict génératrice : pour tout $J \subsetneq I$, la famille $(u_i)_{i\in J}$ n'est pas génératrice. Une famille $(u_i)_{i\in I}$ de vecteurs d'un K-espace vectoriel E est dite libre si la seule K-combinaison linéaire des $(u_i)_{i\in I}$ égale à 0 est la combinaison dont tous les scalaires sont nuls. Une famille libre est dite maximale si elle ne peut être complétée en une famille libre : pour tout $J \supsetneq I$, $(u_i)_{i\in J}$ n'est pas libre. Une famille qui est à la fois libre et génératrice s'appelle une base de E (cette définition ne requiert pas que E soit de dimension finie mais en pratique, nous considérerons presque exclusivement des bases d'espaces vectoriels de dimension finie). Notons qu'une famille libre maximale est génératrice, donc une base, et qu'une famille génératrice minimale est libre, donc une base.

On prendra garde au fait qu'un ensemble E qui est un K-espace vectoriel et un L-espace vectoriel pour deux corps K et L distincts peut très bien être de dimension finie en tant que K-espace vectoriel mais ne pas être de dimension finie en tant que L-espace vectoriel. De même, une famille peut être génératrice du L-espace vectoriel E mais ne pas être génératrice du K-espace vectoriel E ou bien être libre dans le K-espace vectoriel E mais ne pas l'être dans le L-espace vectoriel E.

Propriété fondamentale de la dimension Dans un K-espace vectoriel de dimension finie, toutes les familles génératrices minimales, toutes les familles libres maximales et toutes les bases ont même cardinal. On appelle ce cardinal commun la dimension de l'espace vectoriel E sur K (comme noté plus haut, la dimension de E dépend de manière cruciale de K). Toute famille libre peut-être complétée en une base et on peut extraire de toute famille génératrice une base. Il en résulte que si $E \subset F$ sont des espaces vectoriels et si F est de dimension finie, alors E est de dimension finie, $\dim E \leq \dim F$ et $E = F$ si et seulement si $\dim E = \dim F$.

Somme directe Soit E, F deux sous-espaces vectoriels d'un espace vectoriel G. On désigne par $E+F$ l'espace vectoriel engendré par l'ensemble $E\cup F$. La notation $E+F$ est justifiée par le fait que $E + F$ est aussi l'espace vectoriel $\{u + v|u \in E,\ v \in F\}$.

$$E + F = \operatorname{Vect}(E \cup F) = \{u + v|u \in E,\ v \in F\}$$

Tout élément w de $E+F$ peut donc s'écrire $w=u+v$ avec $u \in E$ et $v \in F$. Lorsque pour tout $w \in E+F$, cette écriture est de plus unique, on dit que E et F sont en somme directe et on note $E \oplus F$. Les trois propriétés suivantes sont équivalentes.

1. $\forall\ w \in E+F$, l'écriture $w=u+v$ est unique.
2. L'écriture de 0 sous la forme $0=u+v$ est unique. Autrement dit, si $0=u+v$ avec $u \in E$ et $v \in F$, alors $u=v=0$.
3. $E \cap F=\{0\}$.

En pratique, c'est le plus souvent la dernière propriétés que l'on vérifie. Si E et F sont en somme directe et si de plus $E+F=G$, on dit que E et F sont supplémentaires et que F est un supplémentaire de E dans G (et réciproquement).

Si G est de dimension finie et si E et F sont supplémentaires, alors $\dim E+\dim F=\dim G$. Tout sous-espace vectoriel $E \subset G$ admet un supplémentaire. Un sous-espace vectoriel E admet en général plusieurs supplémentaires distincts. Par exemple, si u n'est pas de la forme $\begin{pmatrix}\alpha\\0\end{pmatrix}$ avec $\alpha \in \mathbb{R}$, alors $\mathrm{Vect}(u)$ est un supplémentaire de $\mathrm{Vect}\begin{pmatrix}1\\0\end{pmatrix}$ dans $\mathbb{R}^2$. Il faut donc toujours éviter de parler *du* supplémentaire de E sans plus de précisions.

La notion de supplémentaire n'a rien à voir avec la notion de complémentaire : pour commencer, le complémentaire d'un sous-espace vectoriel n'est jamais un sous-espace vectoriel car il ne contient pas 0 et de plus, un sous-ensemble admet un unique complémentaire alors qu'un sous-espace vectoriel admet en général plusieurs supplémentaires distincts.

Les notions de somme et de somme directe se généralisent à un nombre fini de sous-espaces vectoriels.

$$\sum_{i=1}^{n} E_i=\left\{\sum_{i=1}^{n} u_i | u_i \in E_i\right\}=\mathrm{Vect}\left(\bigcup_{i=1}^{n} E_i\right)$$

On dit que les E_i sont en somme directe et on note

$$E_1 \oplus \cdots \oplus E_n=\bigoplus_{i=1}^{n} E_i$$

si pour tout $u \in \sum_{i\in I} u_i$, l'écriture de u comme somme d'éléments des E_i est unique. On prendra garde qu'il ne suffit pas que l'intersection $E_i \cap E_j$ des espaces vectoriels deux à deux soit réduite à 0 pour que des espaces vectoriels soient en somme directe. La condition correcte est la suivante.

$$\forall j,\ E_j \cap \sum_{i\neq j} E_i=\{0\}$$

On peut formuler cette condition en disant que la seule combinaison linéaire nulle

$$\sum_{i=1}^{n} \lambda_i u_i,\ u_i \in E_i$$

est la combinaison linaire dont tous les λ_i sont nuls.

1.1.3 Exemples

Exemples d'espaces vectoriels de dimension finie et infinie L'espace vectoriel $\{0\}$ est de dimension 0. En effet, il est engendré par la famille vide.

Le $\mathbb{Q}$-espace vectoriel $\mathbb{Q}$, le $\mathbb{R}$-espace vectoriel $\mathbb{R}$, le $\mathbb{C}$-espace vectoriel $\mathbb{C}$, plus généralement le K-espace vectoriel K, plus généralement le K-espace vectoriel K^n pour $n \geq 1$ sont de dimension finie. Le K-espace vectoriel K^n est de dimension n et une de ses bases est la famille $(e_i)_{1\leq i\leq n}$ telle que toutes les coordonnées de e_j sont nulles sauf la j-ième qui est égale à 1.

Le $\mathbb{Q}$-espace vectoriel $\mathbb{R}$, le $\mathbb{Q}$-espace vectoriel $\mathbb{C}$, plus généralement les $\mathbb{Q}$-espaces vectoriels $\mathbb{R}^n$ et $\mathbb{C}^n$ pour $n \geq 1$ ne sont pas des espaces vectoriels de dimension finie.

Le $\mathbb{R}$-espace vectoriel $\mathbb{C}$ est de dimension finie égale à 2 et une de ses bases et la famille $(1, i)$. Réciproquement, le $\mathbb{C}$-espace vectoriel $\mathbb{R}^2$ est de dimension finie égale à 1 et une de ses bases est la famille (1). Plus généralement, tout $\mathbb{C}$-espace vectoriel E de dimension finie n et dont une base est $(b_j)_{j\in J}$ est un $\mathbb{R}$-espace vectoriel de dimension finie $2n$ et dont une base est $(b_j, ib_j)_{j\in J}$ (ici i désigne une solution de $X^2 + 1 = 0$ et non pas un indice).

Le K-espace vectoriel $K_n[X]$ des polynômes à coefficients dans K de degré au plus n est un K-espace vectoriel de dimension finie $n+1$ (attention au décalage) dont une base est $(1, X, X^2, \cdots, X^n)$. Le K-espace vectoriel $K[X]$ des polynômes à coefficients dans K n'est pas de dimension finie (en effet, il contient des sous-espaces vectoriels de dimension n pour tout $n \in \mathbb{N}$, par exemple les $K_n[X]$).

Le $\mathbb{R}$-espace vectoriel $C^\infty(I, \mathbb{R})$ des fonctions C^∞ d'un intervalle ouvert non-vide $I \subset \mathbb{R}$ vers $\mathbb{R}$ est de dimension infinie. Il en est donc de même pour le $\mathbb{R}$-espace vectoriel $C^n(I, \mathbb{R})$ des fonctions C^n pour tout $n \in \mathbb{N}$ (car une fonction C^∞ est C^n pour tout $n \in \mathbb{N}$). Une famille libre infinie de $C^\infty(I, \mathbb{R})$ est donnée par la famille des $(e^{nx})_{n\in\mathbb{N}}$ et plus généralement par la famille des $(e^{\alpha_i x})_{i\in I}$ pour tout I qui n'est pas de cardinal fini et toute famille $(\alpha_i)_{i\in I}$ de réels tous distincts.

L'ensemble

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 | x + y + z = 0 \right\} \subset \mathbb{R}^3$$

est un $\mathbb{R}$-espace vectoriel de dimension finie égale à 2 dont une base est $\left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right)$.

L'ensemble

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 | x + y + z = 0,\ 3x + 2y + z = 0 \right\} \subset \mathbb{R}^3$$

est un $\mathbb{R}$-espace vectoriel de dimension finie égale à 1 dont une base est $\left(\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right)$.

L'ensemble

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 | x+y+z=0,\ 3x+2y+z=0, x+y=0 \right\} \subset \mathbb{R}^3$$

est un $\mathbb{R}$-espace vectoriel de dimension finie égale à 0. L'ensemble

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 | x+y+z=0,\ 3x+2y+z=0, 2x+y=0 \right\} \subset \mathbb{R}^3$$

est un $\mathbb{R}$-espace vectoriel de dimension finie égale à 1 dont une base est $\left(\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right)$.

Pour $n \geq 1$, l'ensemble $M_n(K)$ des matrices de taille $n \times n$ à coefficients dans K est un K-espace vectoriel de dimension finie égale à n^2 dont une base est la famille $(e_{ij})_{1 \leq i,j \leq n}$ formée des matrices e_{ij} dont toutes les coordonnées sont nulles sauf celle à la ligne i et à la colonne j qui est égale à 1. Plus généralement, pour n, m deux entiers strictement positifs, l'ensemble $M_{n,m}(K)$ des matrices de taille $n \times m$ à coefficients dans K est un K-espace vectoriel de dimension finie égale à nm dont une base est la famille $(e_{ij})_{1 \leq i \leq n, 1 \leq j \leq m}$ formée des matrices e_{ij} dont toutes les coordonnées sont nulles sauf celle à la ligne i et à la colonne j qui est égale à 1. Plus généralement, l'ensemble $\mathrm{Hom}(E, F)$ des applications linéaires d'un K-espace vectoriel E vers un K-espace vectoriel F est un espace vectoriel de dimension finie si et seulement si E et F sont de dimension finie. Dans ce cas, $\dim \mathrm{Hom}(E, F) = \dim E \times \dim F$.

1.2 Applications linéaires

1.2.1 Définition et premières propriétés

Définition Soit K un corps et E, F deux K-espaces vectoriels. Une application linéaire

$$f : E \longrightarrow F$$

est une application de E vers F vérifiant les propriétés suivantes.

1. Pour tout $(u, v) \in E^2$, $f(u + v) = f(u) + f(v)$.
2. Pour tout $(\lambda, u) \in K \times E$, $f(\lambda u) = \lambda f(u)$.

Il résulte de ces deux propriétés que $f(0_E) = 0_F$ et que si

$$v = \sum_{i \in I} \lambda_i u_i$$

est une combinaison linéaire des vecteurs u_i, alors

$$f(v) = \sum_{i \in I} \lambda_i f(u_i)$$

et $f(v)$ est donc une combinaison linéaire des $f(u_i)$ indexée par les mêmes scalaires λ_i.

De même qu'un espace vectoriel est de manière informelle un ensemble dans lequel on peut faire des combinaisons linéaires, une application linéaire est de manière informelle une application qui préserve les combinaisons linéaires. Il résulte de la compatibilité de f avec les combinaisons linéaires que f est uniquement déterminée par la donnée de l'image par f d'une famille génératrice de E et donc en particulier d'une base de E.

L'ensemble des applications linéaires de E vers F est noté $\mathrm{Hom}(E, F)$ et l'ensemble des applications linéaires de E vers lui-même est noté $\mathrm{End}(E)$. Soit E, F et G des K-espaces vectoriels. La composée d'une application linéaire de E vers F et d'une application linéaire de F vers G est une application linéaire de E vers G. En particulier, la composition $\circ$ est une loi de composition interne

$$\begin{aligned} \circ : \quad \mathrm{End}(E) \times \mathrm{End}(E) &\longrightarrow \mathrm{End}(E) \\ (f, g) &\longmapsto f \circ g \end{aligned}$$

de $\mathrm{End}(E)$. Cette loi de composition interne est associative et admet un élément neutre, à savoir l'identité Id, mais n'est pas commutative sauf si E est de dimension finie égale à 1. Pour $n \in \mathbb{N}$, on note f^n la composée de f avec elle-même n fois (ceci implique en particulier que $f^0 = \mathrm{Id}$ pour tout $f \in \mathrm{End}(E)$).

Noyau, image, rang L'ensemble $\{u \in E | f(u) = 0\}$ des vecteurs de E d'image égal à 0 est un sous-espace vectoriel de E que l'on appelle le noyau de f et que l'on note $\ker f$. L'ensemble $\{v | \exists u \in E, f(u) = v\}$ des vecteurs de F qui sont image d'un vecteur de E est un sous-espace vectoriel de F que l'on appelle l'image de f et que l'on note $\mathrm{im}\, f$. Lorsque la dimension de $\mathrm{im}\, f$ est finie, on l'appelle le rang de f. Lorsque E est de dimension finie, $\ker f$ et $\mathrm{im}\, f$ sont de dimension finie et $\dim \ker f + \dim \mathrm{im}\, f = \dim E$.

Isomorphisme Pour vérifier qu'une application linéaire est injective, il suffit de montrer que son noyau est réduit à 0. Une application linéaire bijective s'appelle un isomorphisme d'espaces vectoriels. Deux K-espaces vectoriels E et F sont dits isomorphes, ce que l'on note $E \simeq F$, s'il existe un isomorphisme d'espaces vectoriels entre eux. La propriété d'être isomorphe est une relation d'équivalence sur l'ensemble des espaces vectoriels. Deux espaces vectoriels de dimension finie sont isomorphes si et seulement s'ils ont la même dimension. Le quotient de l'ensemble des espaces vectoriels de dimension finie par la relation d'équivalence *être isomorphe* est donc l'ensemble $\mathbb{N}$.

Soit E, F deux K-espaces vectoriels de dimension finite et soit $f \in \mathrm{Hom}(E, F)$. Il résulte du théorème de rang que deux quelconques des trois propriétés suivantes impliquent les autres.

1. L'application f est injective.

2. L'application f est surjective.

3. L'application f est un isomorphisme d'espaces vectoriels.

4. La dimension de E est égale à la dimension de F.

L'ensemble des applications de $\mathrm{End}(E)$ qui sont inversibles (et qui sont donc des isomorphismes d'espaces vectoriels de E vers lui-même) s'appelle le groupe linéaire général de E et est noté $\mathrm{GL}(E)$. Il est muni d'une structure de groupe non-commutatif sauf si E est de dimension finie égale à 1. On note $\mathrm{GL}_n(K)$ le groupe linéaire général de K^n. Notons que $\mathrm{GL}_1(K)$ n'est rien d'autre que le groupe $K^\times$ des éléments non-nuls de K

1.2.2 Matrices

Matrice d'une application linéaire Soit E, F et G des K-espaces vectoriels de dimension finie n, m et p respectivement. Soit $f \in \mathrm{Hom}(E, F)$ et $g \in \mathrm{Hom}(F, G)$ deux applications linéaires. Soit $(a_i)_{1\leq i\leq n}$, $(b_i)_{1\leq i\leq m}$ et $(c_i)_{1\leq i\leq p}$ des bases de E, F et G respectivement. La matrice $M(f) \in M_{m,n}(K)$ de f relativement aux bases (a_i) et (b_i) est la matrice $(m(f)_{ij})_{1\leq i\leq m, 1\leq j\leq n}$ telle que

$$f(a_j) = \sum_{i=1}^{m} m(f)_{ij} b_i$$

pour tout $1 \leq j \leq n$. De même, la matrice $M(g) \in M_{p,m}(K)$ de g relativement aux bases (b_i) et (c_i) est la matrice $(m(g)_{ij})_{1\leq i\leq p, 1\leq j\leq m}$ telle que

$$g(b_j) = \sum_{i=1}^{p} m(g)_{ij} b_i$$

pour tout $1 \leq j \leq m$. La matrice $M(g \circ f) \in M_{p,n}(K)$ de $g \circ f \in \mathrm{Hom}(E, G)$ vérifie alors

$$m(g \circ f)_{ij} = \sum_{k=1}^{m} m(g)_{ik} m(f)_{kj}$$

pour tout $1 \leq i \leq p$ et tout $1 \leq j \leq n$.

Produit matriciel Il résulte du calcul précédent et des propriétés de la composition que l'ensemble $M_{n,m}(K) \times M_{m,p}(K)$ est muni d'une multiplication

$$\begin{array}{rccl} \cdot : & M_{n,m}(K) \times M_{m,p}(K) & \longrightarrow & M_{n,p}(K) \\ & (a_{ij})_{1\leq i\leq n, 1\leq j\leq m}(b_{ij})_{1\leq i\leq m, 1\leq j\leq p} & \longmapsto & \left(\displaystyle\sum_{k=1}^{m} a_{ik} b_{kj}\right)_{1\leq i\leq n, 1\leq j\leq p} \end{array}$$

et que $M_n(K)$ est muni d'une loi de composition interne

$$\begin{array}{rccl} \cdot : & M_n(K) \times M_n(K) & \longrightarrow & M_n(K) \\ & (a_{ij})(b_{ij}) & \longmapsto & \left(\displaystyle\sum_{k=1}^{n} a_{ik} b_{kj}\right) \end{array}$$

associative, admettant comme élément neutre la matrice $\mathrm{Id}_n = (\delta_{ij})$ et non-commutative sauf si $n = 1$.

Notons que l'application

$$\begin{aligned} \mathrm{ev}_1 : \quad \mathrm{Hom}(K, K^n) &\longrightarrow K^n \\ f &\longmapsto f(1) \end{aligned}$$

est injective et est donc un isomorphisme d'espaces vectoriels entre les espaces vectoriels $\mathrm{Hom}(K, K^n)$ et K^n car ils sont tous deux de dimension finie n. En utilisant cet isomorphism, on peut identifier les vecteurs de K^n aux vecteurs de $\mathrm{Hom}(K, K^n)$ et donc aux vecteurs de $M_{n,1}(K)$. On peut donc en particulier multiplier une matrice $M \in M_{n,m}(K)$ et un vecteur $u \in K^m \simeq \mathrm{Hom}(K, K^m) \simeq M_{m,1}(K)$ pour obtenir un vecteur $Mu \in K^n$.

Lorsque $E = F$ et $(a_i) = (b_i)$, on dit que $M(f)$ est la matrice de f relative à la base (a_i) pour dire que $M(f)$ est la matrice de f relative à la base (a_i) et à la base (a_i).

Matrice inversible Une matrice $M \in M_n(K)$ telle qu'il existe une matrice N vérifiant $MN = \mathrm{Id}_n$ s'appelle une matrice inversible. Une telle matrice est la matrice d'une application $f \in \mathrm{GL}_n(K)$ relative à un certain choix de base et réciproquement, la matrice d'une application de $\mathrm{GL}_n(K)$ est inversible. Il en résulte que N est unique et vérifie $NM = \mathrm{Id}_n$. L'ensemble des matrices inversibles de taille $n \times n$ à coefficients dans K est noté $\mathrm{GL}_n(K)$.

Changement de bases Soit $f \in \mathrm{End}(E)$ un endomorphisme d'un K-espace vectoriel de dimension finie n. Soit $M = (a_{ij})$ la matrice de f relative à une base $(b_i)_{1 \leq i \leq n}$. Soit $(c_i)_{1 \leq i \leq n}$ une autre base de E. Pour tout $1 \leq j \leq n$, on peut exprimer c_j comme combinaison linéaire des b_i.

$$c_j = \sum_{i=1}^{n} p_{ij} b_i$$

Ceci définit une matrice $P = (p_{ij}) \in M_n(K)$ qui est la matrice dans la base (b_i) de l'application linéaire p qui envoie b_i sur c_i pour tout i. L'application linéaire p est inversible d'inverse p^{-1} égale à l'application linéaire qui envoie c_i sur b_i pour tout i. De

$$\begin{aligned} p^{-1}(c_j) &= \sum_{i=1}^{n} p_{ij} p^{-1}(b_i) \\ &= \sum_{i=1}^{n} \sum_{k=1}^{n} p_{ij} \beta_{ki} b_k = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} \beta_{ki} p_{ij} \right) b_k \\ &= b_j, \end{aligned}$$

on déduit que la matrice (β_{ki}) de p^{-1} relative à la base (b_i) est la matrice P^{-1}. Donc

$$\begin{aligned} f(c_j) &= \sum_{i=1}^n p_{ij} f(b_i) = \sum_{i=1}^n p_{ij} \sum_{k=1}^n a_{ki} b_k \\ &= \sum_{i=1}^n p_{ij} \sum_{k=1}^n a_{ki} \sum_{s=1}^n \beta_{sk} c_s = \sum_{s=1}^n \left(\sum_{i=1}^n \sum_{k=1}^n \beta_{sk} a_{ki} p_{ij} \right) c_s. \end{aligned}$$

Le coefficient s, j de la matrice de f relative à la base (c_i) est donc $\sum_{i=1}^n \sum_{k=1}^n \beta_{sk} a_{ki} p_{ij}$, qui est aussi le coefficient s, j de la matrice $P^{-1}MP$.

En conclusion, si M est la matrice de f dans la base (b_i) et si P est la matrice de l'écriture des c_i dans la base (b_i), alors $P^{-1}MP$ est l'écriture de la matrice de f dans la base (c_i).

Notons que si $MP = PM$ ou si $P^{-1}M = MP^{-1}$, alors $P^{-1}MP = M$ et la matrice de f dans la base (b_i) est égale à la matrice de f dans la base (c_i). C'est la cas en particulier pour l'application nulle, dont la matrice est $(0)_{ij}$ dans toutes les bases ; de l'identité, dont la matrice est $(\delta_{ij})_{ij}$ dans toutes les bases et de $\lambda\,\mathrm{Id}$ dont la matrice est $(\lambda\delta_{ij})_{ij}$ dans toutes les bases.

1.3 Systèmes d'équations linéaires

1.3.1 Interprétation linéaire des systèmes d'équations

Reformulation théorique Supposons nous donnés n scalaires $(u_i)_{1\leq i\leq n} \in K^n$ et nm scalaires $(a_{i,j})_{1\leq i\leq n, 1\leq j\leq m} \in K^{nm}$. Une observations fondamentale de l'algèbre linéaire est que la résolution du système d'équations linéaires à n équations et m inconnues

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,m}x_m = u_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,m}x_m = u_2 \\ \cdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,m}x_m = u_n \end{cases} \tag{1.3.1}$$

est équivalente d'après la formule du produit matriciel à la résolution de l'équation

$$AX = U$$

où A est la matrice $(a_{ij}) \in M_{n,m}(K)$, X est le vecteur $(x_i) \in M_{m,1}(K)$ et U est le vecteur $(u_i) \in M_{m,1}(K)$. Après un choix de bases $(b_i)_{1\leq i\leq m}$ de K^m et $(c_i)_{1\leq i\leq n}$ de K^n, la résolution de l'équation $AX = U$ est elle-même équivalente à la caractérisation de $f^{-1}(\{u\})$ pour $u \in K^n$ un vecteur de matrice U dans la base (b_i) et $f : K^m \longrightarrow K^n$ une application linéaire de matrice A relativement aux bases (b_i) et (c_i).

Inversement, si $f \in \mathrm{Hom}(E, F)$ est une application linéaire entre K-espaces vectoriels de dimension finie, le problème de déterminer le noyau de f, donc de caractériser $f^{-1}(\{0\})$, ou plus généralement celui de caractériser $f^{-1}(\{v\})$ pour $v \in F$ ou encore celui de caractériser l'image de f, donc de caractériser l'ensemble de $v \in F$ tel que $f^{-1}(\{v\})$ soit non-vide, devient équivalent après un choix de bases de E et F à l'étude d'une équation matricielle de type $AX = U$ et donc à la résolution d'un système d'équations linéaires.

Lorsque u est le vecteur nul, Il résulte directement des définitions et du théorème du rang que le système (1.3.1) admet une unique solution si et seulement si f est injective et donc si et seulement si $m \leq n$ (autrement dit, il y a moins d'inconnues que d'équations) et f est de rang m (autrement dit, l'image d'une base de K^m par f est une base d'un sous-espace vectoriel de dimension m). On en déduit que l'équation $f(v) = u$ admet une unique solution si et seulement si f est injective et u appartient à l'image de f. En vertu du théorème de rang, f est injective si et seulement si elle est surjective lorsque $n = m$. Dans ce cas, la seconde condition pour que $f(v) = u$ admette une unique solution découle donc de la première. Si le système (1.3.4) a autant d'équations que d'inconnues, alors le fait qu'il admette une unique solution ou non ne dépend donc pas des u_i mais seulement des $a_{i,j}$, et plus précisément uniquement du fait que la matrice A est inversible ou non.

Résolutions théorique de systèmes d'équations linéaires Pour résoudre un système linéaire, la méthode pratique la plus efficace est celle dite du pivot de Gauss, qui consiste à utiliser un coefficient non-nul d'une ligne pour annuler les coefficients correspondant sur toutes les autres lignes. Formellement, ceci revient à appliquer successivement l'opération élémentaire suivante sur les lignes du système : remplacer la paire de lignes (L_i, L_j) par la paire de lignes $(L_i, L_j - \frac{a_{i,j}}{a_{i,i}} L_i)$. Cette opération transforme un système d'équations linéaires en un système équivalent (c'est-à-dire ayant le même ensemble de solutions) car si les (x_i) sont solutions des deux équations (L_i, L_j), alors ils sont solutions des deux équations $(L_i, L_j - \frac{a_{i,j}}{a_{i,i}} L_i)$ et réciproquement, s'ils sont solutions de la paire d'équation $(L_i, L_j - \frac{a_{i,j}}{a_{i,i}} L_i)$, ils sont solutions de la paire d'équation $(L_i, L_j - \frac{a_{i,j}}{a_{i,i}} L_i + \frac{a_{i,j}}{a_{i,i}} L_i) = (L_i, L_j)$.

Si l'on traduit le problème de la résolution de système d'équations linéaires en problème de la résolution d'équation matricielle $AX = U$, on peut donner la reformulation suivante de l'opération élémentaire et plus généralement de la résolution d'un système d'équations linéaires. Une manière de transformer l'équation matricielle

$$AX = U \tag{1.3.2}$$

en une équation équivalente est de multiplier à gauche les deux membres de l'équation par une matrice inversible commune P pour obtenir l'équation

$$PAX = PU. \tag{1.3.3}$$

Les équations (1.3.2) et (1.3.3) sont équivalentes car on passe de l'une à l'autre simplement en multipliant par P ou P^{-1} à gauche. La matrice P étant inversible, elle est carrée et puisqu'on peut la multiplier à gauche avec A et U, c'est une matrice de $M_m(K)$.

Vue sous cette angle, l'opération élémentaire de résolution d'un système d'équations linéaires se décompose en trois sous-opérations :

1. Échanger deux lignes L_i et L_j.

2. Multiplier la ligne L_i par un scalaire λ non-nul.

3. Remplacer la ligner L_j par la ligne $L_j - L_i$.

En termes matriciels, ces trois opérations s'interprètent de la manière suivante.

1. Multiplier à gauche par la matrice inversible $E_{ij} = (p_{kl})_{1\leq k,l\leq m}$ définie par

$$a_{kl} = \begin{cases} 1 \text{ si } k = l \text{ et } k \neq i, j \\ 1 \text{ si } k = i \text{ et } l = j \text{ ou } k = j \text{ et } l = i \\ 0 \text{ sinon.} \end{cases}$$

2. Multiplier à gauche par la matrice inversible $M_{i,\lambda} = (p_{kl})_{1\leq k,l\leq m}$ définie par

$$p_{kl} = \begin{cases} 1 \text{ si } k = l \text{ et } k \neq i \\ \lambda \text{ si } k = l \text{ et } k = i \\ 0 \text{ sinon.} \end{cases}$$

3. Multiplier à gauche par la matrice inversible $S_{ij} = (p_{kl})_{1\leq k,l\leq m}$ définie par

$$p_{kl} = \begin{cases} 1 \text{ si } k = l \\ 1 \text{ si } k = j \text{ et } l = i \\ 0 \text{ sinon.} \end{cases}$$

Si on applique successivement ces opérations, on obtient donc une équation matricielle équivalente

$$BX = V$$

où $B = (b_{ij})$ est sous-forme échelonnée réduite, ce qui signifie que tous ses coefficients sont nuls sauf éventuellement les b_{ii} pour $1 \leq i \leq d \leq \min(m, n)$.

1.3.2 Exemples de résolutions pratique de systèmes d'équations linéaires

Dans cette sous-section, on suppose que $K = \mathbb{R}$ (le choix de $K = \mathbb{Q}$ ou $\mathbb{C}$ ne changerait rien aux calculs suivants).

Exemple élémentaire Considérons le système suivant

$$\begin{cases} 4y + 2z = -2 \\ 5x + 3y + 2z = 0 \\ x + y + z = 3 \end{cases}$$

qui correspond à l'équation matricielle

$$\begin{pmatrix} 0 & 4 & 2 \\ 5 & 3 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 3 \end{pmatrix}$$

Multiplions à gauche par E_{13} puis par $M_{1,-1/5}S_{1,2}M_{1,-5}$ pour obtenir

$$\begin{pmatrix}1 & 1 & 1\\ 5 & 3 & 2\\ 0 & 4 & 2\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{pmatrix}3\\ 0\\ -2\end{pmatrix}, \quad \begin{pmatrix}1 & 1 & 1\\ 0 & -2 & -3\\ 0 & 4 & 2\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{pmatrix}3\\ -15\\ -2\end{pmatrix}.$$

Multiplions à gauche par $M_{2,-1/4}S_{2,3}M_{2,2}$ puis par $M_{3,-1/4}S_{2,1}M_{2,-1}$ pour obtenir

$$\begin{pmatrix}1 & 1 & 1\\ 0 & 1 & 3/2\\ 0 & 0 & -4\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{pmatrix}3\\ 15/2\\ -32\end{pmatrix}, \quad \begin{pmatrix}1 & 0 & -1/2\\ 0 & 1 & 3/2\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{pmatrix}-9/2\\ 15/2\\ 8\end{pmatrix}.$$

Finalement, multiplions à gauche par $M_{3,2}S_{3,1}M_{3,1/2}$ puis par $M_{3,-2/3}S_{3,2}M_{3,-3/2}$ pour obtenir

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 3/2\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{pmatrix}-1/2\\ 15/2\\ 8\end{pmatrix}, \quad \begin{pmatrix}1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{pmatrix}-1/2\\ -9/2\\ 8\end{pmatrix}$$

et finalement le fait que le système

$$\begin{cases}4y + 2z = -2\\ 5x + 3y + 2z = 0\\ x + y + z = 3\end{cases}$$

admet comme unique solution

$$\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{pmatrix}-1/2\\ -9/2\\ 8\end{pmatrix}.$$

Exemple avec un paramètre Considérons le système suivant

$$\begin{cases}x - y + z = 1\\ -x + y + z = 0\end{cases}$$

qui correspond à l'équation matricielle

$$\begin{pmatrix}1 & -1 & 1\\ -1 & 1 & 1\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{pmatrix}1\\ 0\end{pmatrix}.$$

Multiplions à gauche par $M_{3,1/2}S_{1,2}$ puis par $M_{2,-1}S_{2,1}M_{2,-1}$ pour obtenir

$$\begin{pmatrix}1 & -1 & 1\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{pmatrix}1\\ 1/2\end{pmatrix}, \quad \begin{pmatrix}1 & -1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{pmatrix}1/2\\ 1/2\end{pmatrix}$$

et le fait que l'ensemble des solutions réelles du système

$$\begin{cases}x - y + z = 1\\ -x + y + z = 0\end{cases}$$

est l'ensemble

$$\left\{\begin{pmatrix}1/2 + y\\ y\\ 1/2\end{pmatrix} |y \in \mathbb{R}\right\}$$

Exemple détermination du noyau et de l'image Soit $f \in \mathrm{End}(\mathbb{R}^2)$ l'application dont la matrice dans la base canonique de $\mathbb{R}^2$ est la matrice

$$M = \begin{pmatrix} 6 & 9 \\ 10 & 15 \end{pmatrix}.$$

Un vecteur $u \in \mathbb{R}^2$ est dans le noyau de f si et seulement si $MU = 0$ pour $U = \begin{pmatrix} x \\ y \end{pmatrix}$ le vecteur de $M_{2,1}(\mathbb{R})$ dont les coordonnées sont les coordonnées de u dans la base canonique de $\mathbb{R}^2$. En multipliant l'équation $MU = 0$ par $M_{1,-1/10}S_{1,2}M_{1,-5/3}$, on obtient l'équation matricielle

$$\begin{pmatrix} 1 & 3/2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0.$$

Les vecteurs du noyau de f sont donc les vecteurs vérifiant $x = -3y/2$ pour $y \in \mathbb{R}$. Cet espace vectoriel est de dimension 1. D'après le théorème du rang, l'image de l'application f est donc de dimension 1 et f n'est donc pas surjective. Un vecteur v de coordonnées (v_1, v_2) dans la base canonique de $\mathbb{R}^2$ est dans l'image de f si et seulement si l'équation matricielle $MU = V$ admet une solution pour $V = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$. En multipliant à nouveau par $M_{1,-1/10}S_{1,2}M_{1,-5/3}$, on obtient l'équation matricielle

$$\begin{pmatrix} 1 & 3/2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} v_1/6 \\ v_2 - 5v_1/3 \end{pmatrix}.$$

Cette équation matricielle admet des solutions si et seulement si $v_2 - 5v_1/3 = 0$ et donc l'image de f est le sous-espace vectoriel $\{(v_1, 5v_1/3) \in \mathbb{R}^2 | v_1 \in \mathbb{R}\}$ qui est bien de dimension 1.

1.3.3 Un problème mystérieux

Considérons le système d'équations linéaires à deux équations et deux inconnues suivant.

$$\begin{cases} ax + by = u \\ cx + dy = v \end{cases} \tag{1.3.4}$$

On sait bien comment résoudre un tel système. Néanmoins, procéder à une résolution générale du système sans hypothèse ancillaire sur les coefficients amène à faire une série de remarques intrigantes qui nous amèneront vers les parties les plus complexes, les plus ardues mais aussi les plus profondes de ce cours.

Une solution du système (1.3.4) est une solution du système

$$\begin{cases} ax + by = u \\ (ad - bc)y = av - uc \end{cases} \tag{1.3.5}$$

qui est aussi une solution du système

$$\begin{cases} (ad - bc)x = ud - bv \\ cx + dy = v \end{cases} \tag{1.3.6}$$

et réciproquement une solution communes des systèmes (1.3.5) et (1.3.6) est une solution du système (1.3.4). Déterminer l'ensemble des solutions du système (1.3.4) est donc équivalent à déterminer l'ensemble des solutions communes aux systèmes (1.3.5) et (1.3.6). Il se présente alors la disjonction de cas suivante.

1. Soit $ad - bc = 0$. Il se présente alors la disjonction de cas suivante.
   (a) Soit l'un des termes $av - uc$ et $ud - bv$ est non-nul. Le système (1.3.4) n'admet alors aucune solution.
   (b) Soit les deux termes $av - uc$ et $ud - bv$ sont nuls. Il se présente alors la disjonction de cas suivante.
      i. Soit tous les coefficients a, b, c, d sont nuls. Le système (1.3.4) admet alors l'ensemble K^2 ou l'ensemble $\varnothing$ selon que u et v sont nuls ou non.
      ii. Soit l'un des coefficients a, b, c, d est non-nul et l'on peut sans perte de généralité supposer que c'est a. Le système (1.3.4) admet alors l'ensemble
      $$\left\{ \left(\frac{u - by}{a}, y \right) | y \in K \right\}$$
      comme ensemble de solutions.
2. Soit $ad - bc \neq 0$. Le système (1.3.4) admet alors l'unique solution
   $$\left(\frac{ud - bv}{ad - bc}, \frac{av - uc}{ad - bc} \right)$$
   comme ensemble de solutions.

Il y plusieurs observations à faire sur cette résolution. Tout d'abord, elle est beaucoup plus difficile à mener en toute généralité qu'on pourrait le croire : non seulement le nombre de disjonction de cas à traiter est considérable mais de plus il n'est pas si aisé de maintenir la symétrie du problème, c'est-à-dire de ne pas supposer sans raison que l'un des coefficients joue un rôle particulier (par exemple en choisissant un pivot). On tremble à l'idée de généraliser cette résolution à un système de trois équations à trois inconnues, sans parler du cas général de n équations à n inconnues.

Lorsque $K = \mathbb{R}$, les disjonctions de cas intervenant et les différents ensembles de solutions auquel elles mènent admettent une interprétation géométrique. En effet, si l'un au moins des coefficients (α, β) est non-nul, l'équation $\alpha x + \beta y = \gamma$ est l'équation d'une droite du plan $\mathbb{R}^2$. Les disjonctions de cas ci-dessus ont donc l'interprétation géométrique suivante : ou bien les deux équations du système (1.3.4) sont effectivement des équations de droites, auquel cas le système admet une unique solution, une infinité de solution ou aucune solution selon que ces deux droites sont non-parallèles, parallèles et confondues ou parallèles et non confondues, ou bien l'une au moins des deux équations est dégénérée et dans ce cas on est ramené à étudier le cas d'une unique droite, ou de deux équations dégénérées. Bien que cette interprétation soit commode pour se souvenir des différentes possibilités et pour avoir une intuition de la dimension de l'espace vectoriel des solutions, à nouveau, on tremble à l'idée de la généraliser dans des espaces à 3, 4 ou n dimensions, surtout si l'on a pleine conscience qu'une telle généralisation impliquera de savoir répondre par exemple à la question des intersections possibles entre un espace tridimensionnel et un plan dans un espace à 4 dimensions. Notons aussi que la solution algébrique était valable sur tout corps K. Déjà en dimension 2 mais pour $K = \mathbb{C}$, on est amené à tenter d'interpréter géométriquement des espaces loin d'être triviaux.

Avant que l'algèbre linéaire ne vienne à notre rescousse, observons un dernier fait mystérieux. Si l'on considère la fonction

$$\begin{array}{rccl} \psi_2 : & K^4 & \longrightarrow & K \\ & (\alpha, \beta, \gamma, \delta) & \longmapsto & \alpha\delta - \beta\gamma \end{array}$$

alors non seulement le système (1.3.4) admet une solution unique si $\psi_2(a, b, c, d) \neq 0$ mais cette solution est donnée par

$$\left(\frac{\psi_2(u, b, v, d)}{\psi_2(a, b, c, d)}, \frac{\psi_2(a, u, c, v)}{\psi_2(a, b, c, d)} \right).$$

Le rêve serait de définir une telle fonction ψ_n pour les systèmes de n équations à n inconnues : il suffirait ensuite d'appliquer la fonction ψ_n pour savoir si un système admet une solution unique ou non et pour déterminer sa solution unique le cas échéant. Comme souvent en mathématiques, ce rêve est appelé à devenir réalité mais comme trop souvent dans la vie, les rêves réalisés perdent parfois le parfum d'utopie qui faisaient leur charme. Mais n'allons pas trop vite.

2 Réduction des endomorphismes

2.1 Sous-espaces stables, sous-espaces propres

2.1.1 Définition

Soit E un K-espace vectoriel et $f \in \mathrm{End}(E)$ un endomorphisme de K. Un sous-espace vectoriel $F \subset E$ de E est stable par f si $f(u)$ appartient à F pour tout $u \in F$, donc si f restreint à F est un endomorphisme de F. Un endomorphisme f admet toujours au moins deux sous-espaces vectoriels stables : son noyau $\ker f$ et son image $\mathrm{im}\, f$.

Soit $\lambda \in K$. Un sous-espace vectoriel $F \subset E$ est un sous-espace propre pour le scalaire λ si pour tout $u \in F_\lambda$, $f(u) = \lambda u$. Un sous-espace vectoriel $F \subset E$ est un sous-espace propre s'il existe $\lambda \in K$ tel que F soit un sous-espace propre pour le scalaire λ. Un sous-espace propre est nécessairement un sous-espace stable et réciproquement, un sous-espace stable de dimension au plus 1 est nécessairement un sous-espace propre. Le noyau de f est un sous-espace propre pour le scalaire 0. Si G_λ et F_λ sont deux sous-espace propres pour le même scalaire λ, alors $G_\lambda + F_\lambda$ est un sous-espace propre pour le scalaire λ. Pour tout $\lambda \in K$, il existe donc un sous-espace propre maximal pour le scalaire λ, à savoir l'espace vectoriel engendré par l'union des tous les sous-espaces propres pour le scalaire λ.

Un vecteur $u \in E$ est un vecteur propre de f pour la valeur propre λ si $u \neq 0$ et si $f(u) = \lambda u$. Un vecteur $u \in E$ est un vecteur propre de f s'il existe $\lambda \in K$ tel que u soit un vecteur propre de f pour la valeur propre λ. Un scalaire $\lambda \in K$ est une valeur propre de f s'il existe un vecteur propre de f pour la valeur propre λ. De manière équivalente, un scalaire λ est une valeur propre de f si le sous-espace propre maximal de f pour λ est de dimension strictement positive. De même, un vecteur $U \in M_{1,n}(K)$ est un vecteur propre de $M \in M_n(K)$ si M est la matrice d'une application linéaire f relative à la base canonique de K^n et que u est un vecteur propre de f et $\lambda \in K$ est une valeur propre de M s'il existe un vecteur U tel que U soit un vecteur propre de M pour la valeur propre λ. C'est aussi par définition $\ker(f - \lambda \,\mathrm{Id})$. Lorsque l'on fait référence à l'espace propre de f pour λ, on fait référence à son espace propre maximal $\ker(f - \lambda \,\mathrm{Id})$.

Si $E \subset F$ est un sous-espace stable pour f et g, alors c'est aussi un sous-espace stable pour $h \in \mathrm{Vect}(f, g) \subset \mathrm{End}(F)$. De même, si $E \subset F$ est un sous-espace propre pour le scalaire λ pour f et g, alors c'est aussi un sous-espace propre pour $h \in \mathrm{Vect}(f, g) \subset \mathrm{End}(F)$ (mais en général pas pour le même scalaire).

Notons la légère différence qui existe entre sous-espace propre pour le scalaire λ et sous-espace propre pour la valeur propre λ : dans le second cas, le sous-espace doit être de dimension strictement positive alors que dans le premier, il peut être réduit à l'espace vectoriel nul. L'intérêt d'introduire cette distinction est de se dispenser dans la suite de la vérification que les sous-espaces considérés sont bien non-nuls.

2.1.2 Familles de vecteurs propres

Proposition 2.1. *Soit $(E_{\lambda_s})_{s\in S}$ une famille de sous-espaces vectoriels propres pour les valeurs propres $(\lambda_s)_{s\in S}$. Supposons les valeurs propres $(\lambda_s)_{s\in S}$ distinctes deux à deux ; c'est-à-dire que $\lambda_s \neq \lambda_t$ si $s \neq t$. Les sous-espaces E_{λ_s} sont alors en somme directe.*

De manière équivalent et explicite, soit $(u_i)_{i\in I}$ une famille de vecteurs de E propres pour $f \in \mathrm{End}(E)$ pour les valeurs propres $(\lambda_i)_{i\in I}$ respectivement. On suppose que I est l'union disjointe d'ensembles $(J_s)_{s\in S}$ tels que $\lambda_i \neq \lambda_j$ si $i \in J_s$ et $j \notin J_s$. S'il existe une combinaison linéaire nulle

$$\sum_{i\in I} \alpha_i u_i = 0$$

dont tous les coefficients ne sont pas nuls, alors il existe $s \in S$ et une combinaison linéaire nulle

$$\sum_{i\in J_s} \beta_i u_i = 0 \tag{2.1.1}$$

dont tous les coefficients ne sont pas nuls. En particulier, si tous les scalaires λ_i sont distincts, la famille $(u_i)_{i\in I}$ est libre.

Démonstration. Si la deuxième propriété est vraie, toute famille $(u_s)_{s\in S}$ avec $u_s \in E_{\lambda_s}$ est une famille libre. Si

$$\sum_{s\in S} \alpha_s u_s$$

est une combinaison linéaire nulle des $u_s \in E_{\lambda_s}$, alors tous les α_s sont nuls et donc les E_{λ_s} sont en somme directe. Il suffit donc de montrer la deuxième propriété.

Supposons qu'il existe une combinaison linéaire des (u_i) qui soit nulle mais dont tous les coefficients ne sont pas nuls. Il existe alors une combinaison linéaire nulle des u_i avec I de cardinal minimal

$$\sum_{i\in I}\alpha_i u_i = \sum_{s\in S}\sum_{i\in J_s}\alpha_i u_i = 0 \tag{2.1.2}$$

et telle que les α_i soient tous non-nuls. Soit $j \in I$ un indice et s_j l'indice du sous-ensemble J_s tel que $j \in J_{s_j}$. En appliquant f, on obtient

$$\sum_{i\in I}\alpha_i \lambda_i u_i = \sum_{s\in S}\sum_{i\in J_s}\alpha_i \lambda_s u_i = 0. \tag{2.1.3}$$

En soustrayant l'équation (2.1.3) de l'équation (2.1.2) multiplié par λ_j, on obtient

$$\sum_{s\in S}\sum_{i\in J_s}\alpha_i(\lambda_j - \lambda_s)u_i = 0. \tag{2.1.4}$$

Par définition, $\lambda_s = \lambda_j$ lorsque s appartient à J_{s_j} donc la combinaison linéaire nulle (2.1.4) des u_i a tous ses coefficients indexés par les éléments de J_{s_j} nuls. Par définition de la combinaison linéaire indexée par les α_i, celle-ci est indexée par un ensemble de cardinal minimal, donc tous les $\alpha_i(\lambda_j - \lambda_s)$ sont nuls pour $s \neq s_j$. Comme $\lambda_j \neq \lambda_s$ lorsque $s \notin J_{s_j}$, $\alpha_i = 0$ pour tout $i \notin J_{s_j}$. Comme les α_i sont tous non-nuls, $J_s = J$ et donc il existe une combinaison linéaire nulle des $(u_i)_{i\in J_s}$ dont tous les coefficients ne sont pas nuls.

Supposons maintenant que les λ_i soient tous distincts et soit

$$\sum_{i\in I}\alpha_i u_i = 0$$

une combinaison linéaire nulle des u_i. Les J_s étant de cardinal 1, si $\alpha_i \neq 0$, alors il existe d'après la première partie de la preuve un scalaire nul β_i tel que $\beta_i u_i = 0$. Ceci contredit le fait que u_i soit un vecteur propre, donc un vecteur non-nul. La combinaison linéaire indexée par les α_i est donc nulle, et la famille $(u_i)_{i\in I}$ est donc libre. □

Corollaire 2.2. *Une application linéaire $f \in \mathrm{End}(E)$ d'un espace vectoriel E de dimension finie d admet au plus d valeurs propres distinctes.*

Démonstration. En effet, une famille de vecteurs propres pour des valeurs propres distinctes est une famille libre d'après la proposition précédente et est donc de cardinal inférieur à la dimension de E. □

2.1.3 Lien avec les matrices triangulaires supérieures

Soit $f \in \mathrm{End}(E)$ un endomorphisme d'un K-espace vectoriel de dimension finie n. On dit que la matrice $M = (a_{ij})_{1\leq i,j\leq n}$ de f dans une base $B = (b_1, \cdots, b_n)$ est triangulaire supérieure si $a_{ij} = 0$ lorsque $i > j$ (cela signifie visuellement que tous les coefficients de M en dessous de la diagonale sont nuls). D'un point de vue algébrique, cela signifie que $f(b_i)$ appartient à $\mathrm{Vect}(b_j)_{1\leq j\leq i}$ pour tout $1 \leq i \leq n$. En particulier, le sous-espace $\mathrm{Vect}(b_j)_{1\leq j\leq i}$ est stable par f pour tout $1 \leq i \leq n$.

Proposition 2.3. *Si la matrice $M = (a_{ij})_{1\leq i,j\leq n}$ d'un endomorphisme f est une matrice triangulaire supérieure, alors les valeurs propres de f sont les scalaires a_{ii}.*

Démonstration. Soit λ une valeur propre de f. Alors le noyau de $f - \lambda\,\mathrm{Id}$ contient un vecteur v non-nul dans son noyau. Posons $g = f - \lambda\,\mathrm{Id}$. La matrice $N = (b_{ij})_{1\leq i,j\leq n}$ de g dans la base B est également triangulaire supérieure et il suffit de montrer qu'il existe un i tel que $b_{ii} = 0$. Le vecteur v s'écrit

$$v = \sum_{i=1}^{n} \alpha_i b_i.$$

Soit k le plus grand entier tel que $\alpha_k \neq 0$. De

$$f(v) = \sum_{i=1}^{n} \alpha_i f(b_i) = 0$$

on déduit que

$$f(b_k) = -\frac{1}{\alpha_k} \sum_{i=1}^{k-1} \alpha_{k-1} f(b_i).$$

D'après l'hypothèse que la matrice N est triangulaire supérieure, pour $1 \leq i \leq k-1$, chaque $f(b_i)$ appartient à $\mathrm{Vect}(b_1, \cdots, b_{k-1})$. Donc $f(b_k)$ appartient à $\mathrm{Vect}(b_1, \cdots, b_{k-1})$ et le coefficient b_{kk} est nul.

Réciproquement, fixons un i et posons $\lambda = a_{ii}$. Il suffit de montrer que $g = f - \lambda\,\mathrm{Id}$ n'est pas injective. Soit $N = (b_{ij})_{1\leq i,j\leq n}$ la matrice de g dans la base B. Alors $b_{ii} = 0$. Le sous-espace $F = \mathrm{Vect}(b_1 n \cdots, b_i)$ est stable par f donc stable par g et il suffit donc de montrer que g n'est pas injective après restriction à F. Les vecteurs $g(b_k)$ pour $k < i$ sont dans le sous-espace $\mathrm{Vect}(b_1, \cdots, b_k) \subsetneq F$ d'après l'hypothèse que la matrice de g est triangulaire supérieure. Le vecteur $g(b_i)$ est également dans ce sous-espace car $b_{ii} = 0$. Donc l'image de F par g est strictement incluse dans F donc g restreinte à F n'est pas injective. □

2.2 Polynômes d'endomorphismes

2.2.1 Révision sur les polynômes

Soit K un corps. L'ensemble des polynômes à coefficients dans K est l'ensemble

$$K[X] = \left\{ \sum_{n=0}^{\infty} a_n X^n | a_n \in K,\ \text{Tous les } a_n \text{ sont nuls sauf éventuellement un nombre fini} \right\}$$

Le polynôme dont tous les coefficients sont nuls s'appelle le polynôme nul et il est noté 0. Plus généralement, on identifie le polynôme dont tous les coefficients sont nuls sauf le coefficient a_0 à l'élément a_0 du corps K.

L'ensemble des polynômes est muni d'une loi de composition interne

$$\begin{aligned} + : \quad & K[X] \times K[X] && \longrightarrow K[X] \\ & \left(\sum_{n=0}^{\infty} a_n X^n, \sum_{n=0}^{\infty} b_n X^n\right) && \longmapsto \sum_{n=0}^{\infty} (a_n + b_n) X^n \end{aligned}$$

qui est associative, commutative et qui admet 0 comme élément neutre. Tout polynôme admet un (unique) inverse pour l'addition, ce qui fait donc de $K[X]$ un groupe commutatif. L'ensemble des polynômes est muni d'une loi de composition interne

$$\begin{aligned} \cdot : \quad & K[X] \times K[X] && \longrightarrow K[X] \\ & \left(\sum_{n=0}^{\infty} a_n X^n, \sum_{n=0}^{\infty} b_n X^n\right) && \longmapsto \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k b_{n-k}\right) X^n \end{aligned}$$

qui est associative, commutative, distributive sur l'addition et qui admet le polynôme 1 comme élément unité. La donnée de $K[X]$ muni des lois $+$ et $\cdot$ fait de l'ensemble des polynômes à coefficients dans K un K-espace vectoriel qui est également un anneau commutatif.

Sur $K[X]$ est défini une application

$$\begin{aligned} \deg : \quad & K[X] && \longrightarrow \{-\infty\} \cup \mathbb{N} \\ & \sum_{n=0}^{\infty} a_n X^n && \longmapsto \begin{cases} -\infty \text{ si } a_n = 0 \text{ pour tout } n, \\ \min\{n \in \mathbb{N} | a_n \neq 0\} \text{ sinon.} \end{cases} \end{aligned}$$

Si $(P, Q) \in K[X]^2$ sont deux polynômes, $\deg(PQ) = \deg(P) + \deg(Q)$ et $\deg(P + Q) \leq \max\{\deg P, \deg Q\}$. Les polynômes de degré 0 sont les éléments non-nuls de K et sont exactement les éléments inversibles de $K[X]$ pour la loi $\cdot$.

Proposition 2.4. *Soit $P \in \mathbb{R}[X]$ un polynôme non-nul à coefficients réels et de degré $d \geq 0$. Il existe des réels $(\lambda_i)_{i \in I_1}$ et des couples de réels $(\alpha_i, \beta_i)_{i \in I_2}$ vérifiant $\alpha_i^2 - 4\beta_i < 0$ indexés par des ensembles de cardinaux finis (éventuellement nuls) I_1 et I_2 respectivement ainsi qu'un réel α tels que P s'écrive*

$$P = \alpha \prod_{i \in I_1} (X - \lambda_i) \prod_{i \in I_2} (X^2 + \alpha_i X + \beta_i). \tag{2.2.1}$$

De plus, cette écriture est unique à l'ordre des λ_i et des (α_i, β_i) près et $d = |I_1| + 2|I_2|$.

Soit $P \in \mathbb{C}[X]$ un polynôme non-nul à coefficients complexes de degré d. Il existe d complexes $(\lambda_i)_{1 \leq i \leq d}$ et un complexe α tels que P s'écrive

$$P = \alpha \prod_{i=1}^{d} (X - \lambda_i). \tag{2.2.2}$$

De plus, cette écriture est unique à l'ordre des $(\lambda_i)_{1 \leq i \leq d}$ près.

Cette proposition n'est vrai que pour les corps $\mathbb{R}$ et $\mathbb{C}$. Par exemple, le polynôme $X^3 - 2 \in \mathbb{Q}[X]$ n'admet ni factorisation de la forme (2.2.2), ni factorisation de la forme (2.2.1).

2.2.2 Polynômes d'endomorphismes

Soit f un endomorphisme de E un K-espace vectoriel. Alors pour tout $n \in \mathbb{N}$ et tout $\lambda \in K$, f^n est un endomorphisme de E. Pour tout polynôme

$$P = \sum_{n=0}^{\infty} a_n X^n \in K[X],$$

l'application

$$P(f) = \sum_{n=0}^{\infty} a_n f^n$$

est un endomorphisme de E (ici, on rappelle que P étant un polynôme, il existe un entier N tel que $a_n = 0$ pour tout $n > N$). On appelle un tel endomorphisme un polynôme d'endomorphismes (en f).

Un polynôme d'endomorphismes en f n'est rien d'autre qu'une combinaison linéaires de la familles de vecteurs $(f^n)_{n\in\mathbb{N}}$, qui est une famille de vecteurs du K-espace vectoriel $\mathrm{End}(E)$, donc l'ensemble des polynômes d'endomorphismes en f est le sous-espace vectoriel $\mathrm{Vect}(f^n)_{n\in\mathbb{N}}$.

La composition $\circ$ est une loi de composition interne de $\mathrm{End}(E)$ associative, distributive sur l'addition et commutative lorsqu'elle est restreinte à deux éléments de la forme f^m pour $m \in \mathbb{N}$. Elle est donc commutative lorsqu'elle est restreinte à $\mathrm{Vect}(f^n)_{n\in\mathbb{N}}$. Sur ce sous-espace vectoriel, la composition se comporte donc comme la multiplication de $K[X]$. En particulier, toute relation d'égalité algébrique entre polynômes reste valable pour les polynômes d'endomorphismes en f à condition de remplacer le scalaire λ par l'endomorphisme $\lambda\,\mathrm{Id}$.

Un exemple important est donné par la factorisation de polynômes. Si le polynôme $P \in K[X]$ admet une factorisation

$$P = \sum_{n=0}^{\infty} a_n X^n = \prod_{i\in I} (X - \lambda_i),$$

ou plus généralement

$$P = \prod_{j\in J} P_j$$

où les P_j sont des polynômes de $K[X]$ (le cas précédent correspondant à la situation où tous les P_j sont unitaires de degré 1) alors le polynôme d'endomorphismes

$$P(f) = \sum_{n=0}^{\infty} a_n f^n$$

admet une factorisation semblable :

$$P(f) = \prod_{i\in I} (f - \lambda_i\,\mathrm{Id}), P(f) = \prod_{j\in J} P_j(f) \tag{2.2.3}$$

Proposition 2.5. *Soit E un $\mathbb{C}$-espace vectoriel de dimension finie, f un endomorphisme de E et $P \in \mathbb{C}[X]$ un polynôme non-nul de degré d. Il existe d complexes $(\lambda_i)_{1\leq i\leq d}$ et un scalaire α tels que $P(f)$ s'écrive*

$$P(f) = \alpha\prod_{i=1}^{d}(f - \lambda_i \operatorname{Id}). \tag{2.2.4}$$

Démonstration. Ceci résulte de la proposition 2.4 et de la factorisation (2.2.3). $\square$

Notons qu'un polynôme d'endomorphismes peut tout à fait être l'endomorphisme nul sans pour autant que tous ses coefficients soient nuls. Par exemple, le polynôme d'endomorphismes $f - \operatorname{Id}$ est nul si $f = \operatorname{Id}$ et $f^2 - \operatorname{Id}$ est nul si $f = \operatorname{Id}$, $f = -\operatorname{Id}$ ou si $f \in \operatorname{End}(\mathbb{R}^2)$ a pour matrice relativement à un choix de base la matrice $\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$. Il s'ensuit que contrairement au cas de la factorisation (2.2.2), la factorisation (2.2.4) n'est pas unique en générale.

2.3 Réduction des endomorphismes

2.3.1 Motivation et exemples

Motivation La vie est belle dans les espaces vectoriels de dimension 1 : tous les vecteurs s'écrivent λu pour un certain u et donc pour connaître entièrement une application linéaire f, il suffit de connaître $f(u)$. Le but de la réduction des endomorphismes est de se ramener, autant que faire se peut, à cette situation. En principe, l'idée est simple. Soit E un K-espace vectoriel de dimension finie n et soit $f \in \operatorname{End}(E)$ un endomorphisme de E. L'espace E admet une décomposition

$$E = E_1 \oplus \cdots \oplus E_n = \bigoplus_{i=1}^{n} E_i$$

en somme directe de sous-espaces vectoriels de dimension 1. Pour chacun de ces sous-espaces, f est potentiellement très simple et pour étudier f, il suffirait donc de l'étudier séparément pour chaque E_i puis de rassembler nos connaissances. Le problème est qu'il n'y a aucune raison que les E_i soient des espaces stables par f et donc f peut très bien permuter les E_i.

L'idéal serait donc qu'il existe une décomposition de E

$$E = \bigoplus_{\lambda} E_\lambda$$

en somme directe de sous-espaces stables par f. La matrice de f relativement à une base de E construite en choisissant des bases des E_λ serait alors une matrice par blocs. L'utopie serait carrément qu'il existe une telle décomposition en sous-espaces stables de dimension 1. Ainsi, l'étude de f se ramènerait à l'étude de f restreinte à E_λ, qui serait triviale car un sous-espace stable de dimension 1 est un sous-espace propre. La matrice de f relativement à une base de E construite en choisissant des bases de E_λ serait alors une matrice par blocs de taille 1, donc une matrice dont tous les coefficients seraient nuls sauf éventuellement les coefficients diagonaux. On imagine l'intérêt de travailler avec de telles matrices pour ce qui est de la résolution des systèmes linéaires ou pour le calcul matriciel plus généralement.

L'utopie n'est pas de ce monde, mais une partie substantielle du programme esquissé ci-dessus peut être réalisée.

Exemples Dans ce paragraphe et sauf mention contraire, K désigne un corps égal à $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$, E est un K-espace vectoriel de dimension finie n strictement positive et f est un endomorphisme de E.

1. L'espace E admet une décomposition en sous-espaces stables par f de dimension 1 lorsque f est l'application nulle, lorsque f est l'identité et plus généralement lorsque $f = \lambda\,\mathrm{Id}$ pour $\lambda \in K$. En fait, toute décomposition de E en somme directe de sous-espaces vectoriels de dimension 1 convient. La matrice d'un tel f relativement au choix d'une base quelconque est donc toujours diagonale (avec respectivement des 0, des 1 et des λ sur la diagonale).

2. Soit f dont la matrice dans la base canonique de $\mathbb{R}^2$ est $\begin{pmatrix} 3 & 1 \\ 0 & -5 \end{pmatrix}$. Dans la base, $((1,0),(1,-8))$ la matrice de f est $\begin{pmatrix} 3 & 0 \\ 0 & -5 \end{pmatrix}$ donc
$$\mathbb{R}^2 = \mathrm{Vect}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \oplus \mathrm{Vect}\begin{pmatrix} 1 \\ -8 \end{pmatrix}$$
est une décomposition de $\mathbb{R}^2$ en sous-espaces propres de f de dimension 1. Il est aisé de vérifier qu'à l'ordre près, cette décomposition est la seule décomposition en sous-espaces propres de f de dimension 1.

3. Soit f l'application linéaire dont la matrice dans la base canonique de K^2 est $\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$. Le vecteur $u = (1,1)$ est dans le noyau de f donc $\dim\ker f \geq 0$. Comme f n'est pas l'application nulle, $\dim\mathrm{im}\, f$ est strictement positif. D'après le théorème de rang, $\dim\mathrm{im}\, f + \dim\ker f = 2$ donc $\dim\ker f = 1$ et $\ker f = \mathrm{Vect}\, u$. Supposons que $\mathbb{R}^2$ s'écrive $E_1 \oplus E_2$ où E_1 et E_2 sont des espaces vectoriels non-nuls stables par f. Ils sont alors tous deux de dimension 1 donc des espaces propres pour f pour des valeurs propres λ_1 et λ_2. Si $\lambda_1\lambda_2 \neq 0$, alors le rang de f est égal à 2, ce qui est une contradiction. Donc l'une des valeurs propres, disons λ_1, est nulle et E_1 est donc égal à $\ker f$. Si $u = (x,y)$ est un vecteur non-nul de E_2, alors u est un vecteur propre donc il existe un scalaire λ tel que $f(u) = \lambda$ et donc le système
$$\begin{cases} x - y = \lambda x \\ x - y = \lambda y \end{cases} \quad \text{ou encore} \quad \begin{cases} (1-\lambda)x - y = 0 \\ x + (-1-\lambda)y = 0 \end{cases}$$
admet une solution non-nulle. Ce système admet une solution non-nulle si et seulement si $(1-\lambda)(-1-\lambda)+1 = \lambda^2$ est nul, donc si et seulement si $\lambda = 0$. Mais

l'espace propre correspondant à la valeur propre 0 est E_1 et nous obtenons donc une contradiction. Il n'est pas possible de décomposer K^2 en deux sous-espaces stricts stables sous l'action de f.

4. Soit f l'application linéaire dont la matrice dans la base canonique de $\mathbb{R}^2$ est $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Supposons que f admette un vecteur propre $u = (x, y)$ pour la valeur propre $\lambda \in \mathbb{R}$. Alors le système

$$\begin{cases} -y = \lambda x \\ x = \lambda y \end{cases} \quad \text{ou encore} \quad \begin{cases} -\lambda x - y = 0 \\ x - \lambda y = 0 \end{cases}$$

admet une solution non-nulle, et donc $\lambda^2 + 1 = 0$. C'est absurde. Donc f n'admet aucune valeur propre et il n'est pas possible de décomposer $\mathbb{R}^2$ en deux sous-espaces stricts stables par f.

5. Soit f l'application linéaire dont la matrice dans la base canonique de $\mathbb{C}^2$ est $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. L'application f admet un vecteur propre $u = (x, y)$ pour la valeur propre $\lambda \in \mathbb{C}$ si et seulement si le système

$$\begin{cases} -y = \lambda x \\ x = \lambda y \end{cases} \quad \text{ou encore} \quad \begin{cases} -\lambda x - y = 0 \\ x - \lambda y = 0 \end{cases}$$

admet une solution non-nulle, donc si et seulement si $\lambda^2 + 1 = 0$ et donc si et seulement si $\lambda = \pm i$. Donc $\mathbb{C}^2$ admet deux sous-espaces stables par f : le sous-espace propre E_i correspondant à la valeur propre i et E_{-i} correspondant à la valeur propre $-i$. Ces deux sous-espaces étant en somme directe, $\mathbb{C}^2 = E_i \oplus E_{-i}$ et l'on peut décomposer $\mathbb{C}^2$ en deux sous-espaces stricts stables par f.

Notons que les deux derniers exemples montrent que l'existence d'une décomposition de K^n en sous-espaces stables par f peut dépendre de K mais que l'exemple 3 montre que pour certaines applications linéaires, une telle décomposition est impossible quelque soit le corps considéré.

2.3.2 Coeur et nilespace

L'observation fondamentale qui inspire cette sous-section est le fait que si $f \in \mathrm{End}(E)$ alors il existe non seulement deux sous-espaces stables par f (cela nous le savons déjà : le noyau et l'image de f vérifient cette propriété) mais il existe même deux sous-espaces stables par f dont le somme directe est E. Autrement dit, il existe un sous-espace vectoriel stable admettant un supplémentaire stable.

Définition 2.6. *Soit E un K-espace vectoriel de dimensions finie et $f \in \mathrm{End}(E)$ une application linéaire. Le nilespace $N(f)$ de f est l'ensemble*

$$\{u \in E | \exists n \in \mathbb{N},\ f^n(u) = 0\}.$$

Le coeur $C(f)$ est l'ensemble

$$\{u \in E | \forall n \in \mathbb{N},\ \exists v_n,\ f^n(v_n) = u\}.$$

Si $f \in \mathrm{End}(E)$, $n \in \mathbb{N}$ et $u \in E$ vérifient que $f^n(u) = 0$, alors $f^{n+1}(u) = 0$ et plus généralement $f^{n+m}(m) = 0$ pour $m \in \mathbb{N}$. Donc $\ker f^n \subset \ker f^{n+m}$ pour tout $(n, m) \in \mathbb{N}^2$ et le nilespace de f est l'union des sous-espaces vectoriels $(\ker f^n)_{n \in \mathbb{N}}$. De même, si $f^n(v) = u$ alors $f^m(f^{n-m}(v)) = u$ pour tout $m \leq n$ donc $\mathrm{im}\, f^n \subset \mathrm{im}\, f^{n-m}$ et le coeur de f est l'intersection des sous-espaces vectoriels $(\mathrm{im}\, f^n)_{n \in \mathbb{N}}$.

Lemme 2.7. *Le coeur et le nilespace de f sont des sous-espaces vectoriels stables par f.*

Démonstration. Le coeur de f est un sous-espace vectoriel car c'est une intersection de sous-espaces vectoriels. Soit $(u, v, \lambda) \in N(f)$. Alors il existe $(n, m) \in \mathbb{N}^2$ tel que $f^n(u) = f^m(v) = 0$. Donc $f^{n+m}(u + v)$ et $f^n(\lambda u)$ sont nuls. Donc $u + v$ et λu sont dans $N(f)$, qui est donc bien un sous-espace vectoriel. Montrons tout d'abord que $N(f)$ est un sous-espace vectoriel stable par f. Soit $u \in N(f)$. Il existe alors $n > 0$ tel que $f^n(u) = 0$. Alors $f^{n-1}(f(u)) = 0$ donc $f(u)$ appartient à $N(f)$. Donc $N(f)$ est stable par f. Montrons que $C(f)$ est un sous-espace vectoriel stable par f. Soit $u \in C(f)$ et $n \in \mathbb{N}$. Il existe $u_n \in E$ tel que $f^n(u_n) = u$. Alors $f^n(f(u_n)) = f(u)$ donc il existe $v = f(u_n) \in E$ tel que $f^n(v) = f(u)$. Donc $C(f)$ est stable par f. □

Remarquons que si f est surjectif, alors $\mathrm{im}(f) = \mathrm{im}(f^0)$ et le coeur de f est égal à E. De même, si f est injectif, alors $\ker(f) = \ker(f^0)$ et le nilespace de f est nul. Ce résultat se généralise dans le lemme suivant.

Lemme 2.8. *Soit $n \in \mathbb{N}$ un entier tel que* $\ker(f^n) = \ker(f^{n+1})$. *Alors $N(f) =$* $\ker(f^n)$*. De même, si $n \in \mathbb{N}$ est un entier tel que* $\mathrm{im}(f^n) = \mathrm{im}(f^{n+1})$*, alors $C(f) =$* $\mathrm{im}(f^n)$*.*

Démonstration. Soit N un entier tel que $\ker f^N = \ker f^{N+1}$. Montrons par récurrence que $\ker f^{N+m} = \ker f^N$ pour tout $m \in \mathbb{N}$. C'est vrai par définition si $m = 0$. Supposons que $\ker f^{N+m} = \ker f^N$ et soit $v \in \ker f^{N+m+1}$. Alors $f(v)$ appartient à $\ker f^{N+m}$ donc $f(v)$ appartient à $\ker f^N$ donc v appartient à $\ker f^{N+1}$ donc v appartient à $\ker f^N$. Donc $\ker f^{N+m} = \ker f^N$ pour tout $m \in \mathbb{N}$ par récurrence. En appliquant ce résultat au plus petit n vérifiant la propriété, on obtient l'assertion sur $N(f)$.

Soit maintenant N un entier tel que $\mathrm{im}\, f^N = \mathrm{im}\, f^{N+1}$. Montrons par récurrence que $\mathrm{im}\, f^{N+m} = \mathrm{im}\, f^N$ pour tout $m \in \mathbb{N}$. C'est vrai par définition si $m = 0$. Supposons maintenant $m \leq 0$ et que $\mathrm{im}\, f^{N+m} = \mathrm{im}\, f^N$ et soit $u \in \mathrm{im}\, f^{N+m}$. Il existe $v \in E$ tel que $f^{N+m}(v) = u$. Soit $w = f^N(v)$. Alors w est aussi dans l'image de f^{N+1} par hypothèse sur N donc il existe $z \in E$ tel que $f^{N+1}(z) = w$. Donc $f^{N+m+1}(z) = f^m(w) = f^{N+m}(v) = u$. Donc u appartient à $\mathrm{im}\, f^{N+m+1}$. Donc $\mathrm{im}\, f^{N+m} = \mathrm{im}\, f^N$ pour tout $m \in \mathbb{N}$ par récurrence. En appliquant ce résultat au plus petit n vérifiant la propriété, on obtient l'assertion sur $C(f)$. □

Ce lemme implique que l'union croissante des noyaux des f^n est strictement croissante ou constante à partir d'un certain rang et que l'intersection décroissante des images des f^n est strictement décroissante ou constante à partir d'un certain rang. En général, les deux comportements sont possibles mais lorsque l'espace E est de dimension finie, alors il ne peut exister de suite infinie d'inclusion strictement décroissante de sous-espaces vectoriels et c'est donc la seconde alternative qui prévaut.

Lemme 2.9. *Soit E un K-espace vectoriel de dimensions finie d et $f \in \mathrm{End}(E)$ une application linéaire. Il existe un entier $N \leq d$ tel que* $\ker f^N = N(f)$ *et* $\mathrm{im}\, f^N = C(f)$.

Démonstration. D'après le lemme 2.8, la suite des sous-espaces vectoriels

$$\ker \mathrm{Id} \subset \ker f \subset \cdots \subset \ker f^n \subset \cdots \subset E$$

est strictement croissante ou stationnaire à partir d'un certain rang. La suite d'entiers

$$0 \leq \dim \ker f \leq \cdots \leq \dim \ker f^n \leq \cdots \leq n$$

est bornée donc c'est la deuxième alternative qui prévaut et il existe $N \in \mathbb{N}$ tel que $N(f) = \ker f^N$. Pour tout $n \in \mathbb{N}$, $\dim \mathrm{im}\, f^n = d - \dim \ker f^n$ donc la dimension de $\mathrm{im}\, f^n$ est également stationnaire à partir de N. Le sous-espace vectoriel $\mathrm{im}\, f^N$ est alors égal à $C(f)$. $\square$

Proposition 2.10. *Soit E un K-espace vectoriel de dimensions finie et $f \in \mathrm{End}(E)$ une application linéaire. Le nilespace $N(f)$ et le coeur $C(f)$ de f sont des sous-espaces vectoriels stables par f supplémentaires.*

Démonstration. Nous savons déjà que $N(f)$ et $C(f)$ sont des sous-espaces vectoriels stables par le lemme 2.7. Pour montrer que $E = C(f) \oplus N(f)$, il suffit de montrer que

$$\dim C(f) + \dim N(f) = \dim E \tag{2.3.1}$$

et que $C(f) \cap N(f) = \{0\}$. L'égalité (2.3.1) vient de ce qu'il existe $N \in \mathbb{N}$ tel que $C(f) = \mathrm{im}\, f^N$, $N(f) = \ker f^N$ d'après le lemme 2.9 et du théorème de rang. Soit donc $u \in C(f)$ non-nul. Soit $n \geq N$. Il existe $v \in E$ tel que $f^n(v) = u \neq 0$. Donc v n'appartient pas à $\ker f^N = N(f)$. Donc v n'apparient pas à $N(f)$ donc $f^m(v) \neq 0$ pour tout $m \in \mathbb{N}$ et de même $f^m(u) \neq 0$ pour tout $m \in \mathbb{N}$. Donc u n'appartient pas à $N(f)$. $\square$

Remarquons que $N(f)$ et $C(f)$ n'ont aucune raison d'être en somme directe lorsque E est de dimension infinie. Considérons par exemple la dérivation D comme application linéaire sur l'espace vectoriel $\mathbb{C}[X]$ des polynômes à coefficients complexes. Pour tout polynôme P, il existe $n \in \mathbb{N}$ tel que $D^n(P) = 0$ (on peut par exemple prendre $n = \deg P + 1$) donc le nilespace de la dérivation est $\mathbb{C}[X]$ tout entier. Pour tout polynôme P et tout $n \in \mathbb{N}$, il existe un polynôme Q_n tel que $D^n(Q_n) = P$ (à savoir le polynôme obtenu en prenant n fois la primitive de P) donc le coeur de la dérivation est également $\mathbb{C}[X]$ tout entier.

2.3.3 Espaces propres généralisés

Définition 2.11. *Soit E un K-espace vectoriel de dimension finie, $f \in \mathrm{End}(E)$ un endomorphisme et $\lambda \in K$ un scalaire. L'espace propre généralisé E_λ de f pour le scalaire λ est le nilespace de l'endomorphisme $f - \lambda\,\mathrm{Id}$.*

Un espace propre généralisé est donc en particulier un sous-espace vectoriel stable par f admettant un supplémentaire stable par f. Si n est un entier strictement positif, l'application linéaire $(f - \lambda \operatorname{Id})^n$ est injective si et seulement si l'application linéaire $f - \lambda \operatorname{Id}$ est injective. Un espace propre généralisé pour le scalaire λ est donc non-nul si et seulement si λ est une valeur propre de f.

Lemme 2.12. *Soit (λ, μ) deux valeurs propres distinctes de $f \in \operatorname{End}(E)$. Alors le nilespace $N(f - \lambda \operatorname{Id})$ de $f - \lambda \operatorname{Id}$ est inclus dans $C(f - \mu \operatorname{Id})$. En particulier, des sous-espaces propres généralisés pour des valeurs propres distinctes sont en somme directe.*

Démonstration. Le sous-espace $N(f - \lambda \operatorname{Id})$ est stable par f et stable par Id donc est stable par $f - \mu \operatorname{Id}$. Montrons que l'application $f - \mu \operatorname{Id}$ restreinte à $N(f - \lambda \operatorname{Id})$ est injective (donc inversible). Soit donc g l'application linéaire $f - \mu \operatorname{Id}$ restreinte à $N(f - \lambda \operatorname{Id})$ et soit u dans le noyau de g. Le vecteur u est dans $N(f - \lambda \operatorname{Id})$ donc il existe un entier $n \in \mathbb{N}$ tel que $(f - \lambda \operatorname{Id})^n(u) = 0$ et il est dans $\ker g$ donc $f(u) = \mu u$. Le polynôme d'endomorphismes de l'espace vectoriel $\operatorname{Vect}(u)$ égal à f est donc égal au polynôme d'endomorphismes $\mu \operatorname{Id}$. Donc

$$(f - \lambda \operatorname{Id})^n(u) = (\mu \operatorname{Id} - \lambda \operatorname{Id})^n(u) = (\mu - \lambda)^n u = 0.$$

Le scalaire $\mu - \lambda$ est non-nul donc u est nul. Donc $f - \mu \operatorname{Id}$ est injective après restriction à $N(f - \lambda \operatorname{Id})$. Donc $f - \mu \operatorname{Id}$ est un isomorphisme de $N(f - \lambda \operatorname{Id})$ donc $N(f - \lambda \operatorname{Id}) \subset C(f - \mu \operatorname{Id})$.

Soit maintenant $(E_{\lambda_i})_{i \in I}$ une famille d'espaces propres généralisés pour des valeurs propres distinctes et soit

$$\sum_{i \in J} \alpha_i u_i = 0$$

une combinaison linéaire nulle de vecteurs $u_i \in E_{\lambda_i}$ indexée par $J \subset I$. Soit $j \in J$ et $i \neq j$. Le vecteur u_i appartient au nilespace de $f - \lambda_i \operatorname{Id}$ donc au coeur de $f - \lambda_j \operatorname{Id}$. Donc

$$\alpha_j u_j + \sum_{i \neq j} \alpha_i u_i = 0$$

est une écriture de 0 comme somme d'un élément de $N(f - \lambda_j \operatorname{Id})$ et d'un élément de $C(f - \lambda_j \operatorname{Id})$. Ces deux espaces étant en somme directe, $\alpha_j u_j$. La combinaison linéaire est donc nulle et les (E_{λ_i}) sont en somme directe. $\square$

Il s'en suit qu'un endomorphisme d'un espace de dimension d a au plus d espaces propres généralisés non nuls.

Proposition 2.13. *Soit E un K-espace vectoriel de dimension finie et $f \in \operatorname{End}(E)$ un endomorphisme. Supposons que f admette une valeur propre $\lambda \in K$ et soit E_λ l'espace propre généralisé de f pour λ. Il existe alors une décomposition*

$$E = E_\lambda \oplus F$$

en sous-espaces vectoriels stables par f.

Sous l'hypothèse que f admet une valeur propre, on a donc réussi à réduire l'étude de f à l'étude de f restreint au sous-espace propre généralisé pour λ et au sous-espace vectoriel F.

Démonstration. Le scalaire λ est une valeur propre de f donc $\ker(f - \lambda \operatorname{Id}) \neq \{0\}$ donc $E_\lambda \neq \{0\}$. On peut alors prendre pour F le coeur de $f - \lambda \operatorname{Id}$. □

Théorème 2.1. *Soit E un $\mathbb{C}$-espace vectoriel de dimension finie et $f \in \operatorname{End}(E)$ un endomorphisme. Il existe alors une décomposition*

$$E = \bigoplus_{\lambda} E_\lambda$$

en sous-espaces propres généralisés de f, donc en particulier en sous-espaces stables par f.

Si E est un $\mathbb{C}$-espace vectoriel, on a donc réduit l'étude de f à l'étude de f restreint à ses sous-espaces propres généralisés.

Démonstration. Démontrons l'assertion par récurrence sur la dimension de E. L'assertion est vraie si $\dim E = 0$ ou 1. Supposons maintenant qu'elle soit vraie pour tous les $\mathbb{C}$-espaces vectoriels de dimension finie strictement positive au plus $n \in \mathbb{N}$ et considérons un espace vectoriel E de dimension $n + 1$.

L'espace vectoriel $\operatorname{End}(E)$ est alors de dimension $(n+1)^2$ donc la famille $(f^s)_{s \in \mathbb{N}}$ de vecteurs de $\operatorname{End}(E)$ est liée. Soit donc

$$\sum_{s \in S} \alpha_s f^s = 0 \tag{2.3.2}$$

une combinaison linéaire nulle des $(f^s)_{s \in \mathbb{N}}$ dont tous les coefficients ne sont pas nuls et soit

$$P = \sum_{s \in S} \alpha_s X^s$$

le polynôme correspondant. Ce polynôme est un polynôme non-nul à coefficient de $\mathbb{C}[X]$ donc il admet une factorisation

$$P = \alpha \prod_{s \in S} (X - \lambda_s).$$

avec S non-vide. Le polynôme d'endomorphismes $P(f)$, qui est nul d'après (2.3.2), admet donc d'après la proposition 2.5 une factorisation similaire

$$P(f) = \alpha \prod_{s \in S} (f - \lambda_s \operatorname{Id}) = 0.$$

L'endomorphisme $P(f)$ est nul et n'est donc pas injectif. Une composition d'applications injectives étant injective, l'un au moins des endomorphismes $f - \lambda_s \operatorname{Id}$, disons

$f - \lambda \operatorname{Id}$, n'est pas injectif. L'application f admet donc un vecteur propre pour la valeur propre λ et l'espace propre généralisé E_λ est donc non-nul. D'après la proposition 2.13, l'espace E admet donc une décomposition

$$E = E_\lambda \oplus F$$

avec F stable par f et de dimension strictement inférieure à celle de E. D'après l'hypothèse de récurrence, F admet une décomposition

$$F = \sum_\mu E_\mu$$

en somme directe de sous-espaces propres généralisés. Donc E admet une décomposition

$$E = E_\lambda \oplus \bigoplus_\mu E_\mu$$

en sous-espaces propres généralisés. Par récurrence, tout espace vectoriel complexe de décomposition finie admet donc une décomposition en sous-espaces propres généralisés. □

Soit $(b_i)_{1 \leq i \leq n}$ une base de E formée à partir de bases des sous-espaces propres généralisés E_λ. La matrice de f relativement à $(b_i)_{1 \leq i \leq n}$ est alors diagonale par blocs.

2.3.4 Polynôme caractéristique

Définition 2.14. *La multiplicité géométrique d'une valeur propre d'un endomorphisme f d'un K-espace vectoriel de dimension finie est l'entier* $\dim N(f - \lambda \operatorname{Id})$.

Il résulte du théorème 2.1 que la somme des multiplicités géométriques des valeurs propres d'un endomorphisme d'un $\mathbb{C}$-espace vectoriel de dimension finie d est égale à d.

Définition 2.15. *Soit E un $\mathbb{C}$-espace vectoriel de dimension finie n strictement positive et $f \in \operatorname{End}(E)$ un endomorphisme de E. Le spectre* $\operatorname{Spec}(f)$ *de f est l'ensemble des valeurs propres de f comptées avec leur multiplicité géométrique.*

Le cardinal de $\operatorname{Spec}(f)$ est donc n. L'endomorphisme f est bijectif si et seulement si $0 \notin \operatorname{Spec}(f)$.

Définition 2.16. *Soit E un $\mathbb{C}$-espace vectoriel de dimension finie n strictement positive et $f \in \operatorname{End}(E)$ un endomorphisme de E. Le polynôme caractéristique $\mu_f \in \mathbb{C}_n[X]$ de f est le polynôme*

$$\mu_f = \prod_{\lambda \in \operatorname{Spec}(f)} (X - \lambda)$$

Par construction, μ_f est un polynôme unitaire de degré n à coefficients dans $\mathbb{C}$. D'après la remarque suivant la définition 2.15, f est un isomorphisme si et seulement le terme constant de μ_f est non-nul ou encore si et seulement si $\mu_f(0) \neq 0$.

Théorème 2.2 (Théorème de Cayley-Hamilton). *Soit f un endomorphisme d'un $\mathbb{C}$-espace vectoriel E de dimension finie et $\mu_f \in \mathbb{C}[X]$ son polynôme caractéristique. Le polynôme d'endomorphismes $\mu_f(f)$ est alors le polynôme nul.*

Démonstration. D'après le théorème 2.1, E admet une décomposition

$$E = \bigoplus_{\lambda} E_\lambda$$

en sous-espaces propres généralisés. Les sous-espaces E_λ sont stables par f, donc par l'endomorphisme $\mu_f(f)$. Pour montrer que $\mu_f(f)$ est l'endomorphisme nul, il suffit donc de montrer que $\mu_f(f)$ est l'endomorphisme nul après restriction à chaque E_λ. Soit donc λ une valeur propre de f de multiplicité géométrique m et $u \in E_\lambda$. Par définition de la multiplicité géométrique, $(f - \lambda \operatorname{Id})^m(u) = 0$ donc

$$\left(\prod_{\lambda' \neq \lambda} (f - \lambda \operatorname{Id})\right) \circ (f - \lambda \operatorname{Id})^m(u) = 0.$$

Donc $\mu_f(f)(u) = 0$ et $\mu_f(f)$ est bien l'endomorphisme nul. □

2.3.5 Réduction des endomorphismes diagonalisables

Soit E un $\mathbb{C}$-espace vectoriel de dimension finie et $f \in \operatorname{End}(E)$. D'après le théorème 2.1, E se décompose en somme de sous-espaces propres généralisés pour f.

Définition 2.17. *Un endomorphisme f d'un K-espace vectoriel E de dimension finie est diagonalisable si E admet une décomposition en sous-espaces propres*

$$E = \bigoplus_{\lambda} \ker(f - \lambda \operatorname{Id}).$$

Relativement à une base formée de bases des $\ker(f - \lambda \operatorname{Id})$, la matrice de f est diagonale avec les $\lambda \in \operatorname{Spec}(f)$ sur la diagonale. Ceci explique le terme diagonalisable.

Proposition 2.18. *Soit f un endomorphisme d'un $\mathbb{C}$-espace vectoriel E de dimension finie. Soit $P \in \mathbb{C}[X]$ tel que $P(f)$ soit l'endomorphisme nul. Alors le spectre de f est inclus dans l'ensemble des racines de P. L'endomorphisme f est diagonalisable si et seulement s'il existe un polynôme P de la forme*

$$P = \prod_{i=1}^{d} (X - \lambda_i)$$

avec $\lambda_i \neq \lambda_j$ si $i \neq j$ tel que $P(f)$ soit l'endomorphisme nul.

Démonstration. Soit $\lambda \in \operatorname{Spec} f$ et u un vecteur propre de f pour la valeur propre λ. Alors

$$P(f)(u) = \left(\sum_{i=0}^{d} \alpha_i \lambda^i\right) u = 0$$

et $u \neq 0$ donc λ est une racine de P. Supposons que f soit diagonale. Alors E s'écrit comme la somme directe des $\ker(f - \lambda \operatorname{Id})$ donc

$$\prod_{\lambda \in \operatorname{Spec} f} (X - \lambda)$$

est l'endomorphisme nul de E. Supposons réciproquement qu'il existe un polynôme P dont toutes les racines sont distinctes tel que $P(f)$ soit l'endomorphisme nul. D'après le théorème 2.1, E s'écrit comme la somme directe des espaces propres généralisés de f. Soit E_λ un des espaces propres généralisés de f. Si $\mu \neq \lambda$ est une racine de P, alors $f - \mu \operatorname{Id}$ est un endomorphisme injectif de E_λ donc

$$\prod_{\mu \neq \lambda} (f - \mu \operatorname{Id})$$

est un endomorphisme injectif de E_λ. Or $P(f)$ est l'endomorphisme nul de E_λ donc $f - \lambda \operatorname{Id}$ est l'endomorphisme nul de E_λ donc $E_\lambda = \ker(f - \lambda \operatorname{Id})$ et f est diagonalisable. □

Corollaire 2.19. *Supposons que f admette $\dim E$ valeurs propres distinctes. Alors f est diagonalisable.*

Démonstration. En effet, le polynôme caractéristique μ_f de f est alors un polynôme à racines simples tel que $\mu_f(f) = 0$. On applique la proposition 2.18. □

2.3.6 Réduction des endomorphismes nilpotents

Définition 2.20. *Un endomorphisme f d'un K-espace vectoriel est nilpotent s'il existe $n \in \mathbb{N}$ tel que $f^n = 0$.*

De manière équivalente, un endomorphisme de E est nilpotent si son nilespace est l'espace E tout entier. Soit E un $\mathbb{C}$-espace vectoriel de dimension finie n et soit $f \in \operatorname{End}(E)$. Supposons f nilpotent et soit u un vecteur propre de f pour la valeur propre λ (un tel vecteur existe d'après le théorème 2.1). Alors $f^n(u) = \lambda^n u$ donc $\lambda = 0$. Toutes les valeurs propres d'un endomorphisme nilpotent sont donc nulles. Réciproquement, si E est de dimension d et si $\operatorname{Spec}(f) = \{0^{(d)}\}$, alors E est le nilespace de f donc f est nilpotent.

On appelle ordre de nilpotence de f le plus petit entier $n \in \mathbb{N}$ tel que $f^n = 0$ et ordre de nilpotence de $u \in E$ le plus petit entier m tel que $f^m(u) = 0$. Il résulte immédiatement de ces définitions que l'ordre de nilpotence de u est plus petit que l'ordre de nilpotence de f et qu'il existe un vecteur $u_{\max}$ dont l'ordre de nilpotence est égal à l'ordre de nilpotence de f lorsque E est de dimension finie. L'ensemble des éléments d'ordre de nilpotence au plus d est un sous-espace vectoriel de f. Enfin, nous savons que l'ordre de nilpotence de f est plus petit que la dimension de E lorsque celle-ci est finie.

Lemme 2.21. *Des vecteurs non-nuls d'ordre de nilpotence distincts forment une famille libre. En particulier, si u est un vecteur d'ordre de nilpotence $d > 0$, alors la famille $(f^i(u))_{0 \leq i \leq d-1}$ est libre.*

Démonstration. Soit $(u_i)_{i\in I}$ des vecteurs non nuls d'ordre de nilpotence d_i avec $d_i \neq d_j$ si $i \neq j$. Supposons qu'il existe une combinaison linéaire nulle des u_i dont tous les coefficients ne sont pas nuls. Il en existe alors une

$$\sum_{i=1}^{n} \alpha_i u_i = 0$$

de cardinal minimal. Soit d_j le maximum des d_i. Alors

$$f^{d_j-1}\left(\sum_{i=1}^{n} \alpha_i u_i\right) = \sum_{i=1}^{n} \alpha_i f^{d_j-1}(u_i) = \alpha_j f^{d_j-1}(u_j) = 0$$

et ${}^{d_j-1}(u_j) \neq 0$ donc $\alpha_j = 0$. Contradiction.

La deuxième assertion découle de la première et de l'observation que $f(u)$ est d'ordre de nilpotence d si et seulement si u est d'ordre de nilpotence $d+1$. □

Proposition 2.22. *Soit f un endomorphisme nilpotent de E un K-espace vectoriel de dimension finie n. Il existe une base $(v_1, \cdots, v_n)$ telle que $f(v_i)$ appartienne à* $\mathrm{Vect}(v_1, \cdots, v_{i-1})$ *pour tout* $1 \leq i \leq n$.

Démonstration. On définit le procédé suivant. L'étape zéro est un choix d'une base du noyau de f. Si l'étape $s \geq 0$ est terminée et si la famille construite ne contient que des éléments d'ordre de nilpotence s, on la complète si possible par une famille libre maximale d'éléments d'ordre de nilpotence $s+1$. D'après le lemme 2.21, la famille obtenue est libre et engendre le sous-espace vectoriel des éléments d'ordre de nilpotence au plus $d+1$. Le procédé termine donc avec une famille libre maximale, donc une base de E, vérifiant la propriété voulue. □

Dans la base de la proposition précédente, la matrice de f est triangulaire supérieure avec des 0 sur la diagonale. On pourra remarquer que la démonstration montre en fait une propriété plus forte que celle de l'énoncé. Ceci inspire la proposition suivante, nettement plus forte.

Proposition 2.23. *Soit f un endomorphisme nilpotent de E un K-espace vectoriel de dimension finie. Il existe des vecteurs $(v_1, \cdots, v_s)$ d'ordres de nilpotence respectifs $(d_1, \cdots, d_s)$ tels que*

$$E = \bigoplus_{i=1}^{s} \mathrm{Vect}(f^n(v_i))_{n\in\mathbb{N}}$$

ou, de manière équivalente, tels que

$$(v_1, \cdots, f^{d_1-1}(v_1), v_2, \cdots, f^{d_2-1}(v_2), \cdots, v_s, \cdots, f^{d_s-1}(v_s))$$

soit une base de E et tels que $(f^{d_i-1}(v_i))_{1\leq i\leq s}$ soit une base de $\ker f$.

Démonstration. Soit E est de dimension 1. L'assertion est alors vraie. Soit E est de dimension strictement supérieur à 1 et on peut supposer par récurrence que l'assertion est vraie pour les endomorphismes nilpotents des espaces de dimensions strictement inférieure. Le noyau de f est non-nul donc $\mathrm{im}\, f$ est de dimension r strictement inférieur à E et est stable par f. D'après l'hypothèse de récurrence, il existe donc une base de $\operatorname{im} f$ de la forme $(v_1, \cdots, f^{d_1-1}(v_1), v_2, \cdots, f^{d_2-1}(v_2), \cdots, v_t, \cdots, f^{d_t-1}(v_t))$ avec $(f^{d_i-1}(v_i))_{1\leq i\leq t}$ une based de $\ker f \cap \operatorname{im} f$. Choisissons des $(u_i)_{1\leq i\leq t}$ tels que $f(u_i) = v_i$ pour tout $1 \leq i \leq t$ (ce qui est possible car les v_i appartiennent à $\operatorname{im} f$). Pour tout $1 \leq i \leq d$, l'ordre de nilpotence de u_i est alors l'ordre de nilpotence de v_i plus 1 donc est égal à $d_i + 1$. Soit $(w_i)_{1\leq i\leq k}$ une base d'un supplémentaire de $\operatorname{im} f \cap \ker f$ dans $\ker f$.

Rebaptisons la famille des

$$(u_1, \cdots, f^{d_1}(u_1), \cdots, u_t, \cdots, f^{d_t}(u_t), w_1, \cdots, w_k)$$

en

$$(z_1, \cdots, f^{\delta_1-1}(z_1), \cdots, z_\ell, \cdots, f^{\delta_\ell-1}(z_\ell))_{1\leq\ell\leq m}.$$

Ici, $\delta_j = d_j$ si z_j est l'un des u_i et δ_j si z_j est un des w_i et $m = \dim \operatorname{im} f + k = r + k$. Montrons que

$$Z = (z_1, \cdots, f^{\delta_1-1}(z_1), \cdots, z_\ell, \cdots, f^{\delta_\ell-1}(z_\ell))_{1\leq\ell\leq m}$$

forme une base de E vérifiant la propriété exigée. Par construction, les éléments de la base d'ordre de nilpotence 1 forment une base de $\ker f \cap \operatorname{im} f$ s'ils sont de la forme $f^{d_j-1}(v_j)$ et une base d'un supplémentaire de $\ker f \cap \operatorname{im} f$ dans $\ker f$ s'ils sont de la forme w_j. Ensemble, ils forment donc bien une base de $\ker f$. Il suffit donc de montrer que la famille Z est libre et de cardinal égal à la dimension de E.

Soit

$$\sum_{i=1}^{m}\sum_{j=0}^{\delta_j-1}\alpha_{i,j} f^j(z_i) = 0 \tag{2.3.3}$$

une combinaison linéaire nulle des vecteurs de Z. En appliquant f on obtient

$$0 = \sum_{i=1}^{m}\sum_{j=0}^{\delta_j-1}\alpha_{i,j} f^{j+1}(z_i) = \sum_{i=1}^{r}\sum_{j=0}^{d_j-1}\alpha_{i,j} f^j(v_i)$$

car les w_i sont dans le noyau de f. D'après l'hypothèse de récurrence, la famille des $f^j(v_i)$ est une base donc la combinaison linéaire

$$\sum_{i=1}^{r}\sum_{j=0}^{d_j-1}\alpha_{i,j} f^j(v_i)$$

a tous ses coefficients nuls. Les seuls $\alpha_{i,j}$ éventuellement non-nuls sont donc ceux pour $r < i \leq m$, auquel cas $\delta_j = 1$. La combinaison linéaire nulle (2.3.3) devient donc

$$\sum_{i=1}^{k}\alpha_i w_i = 0.$$

Mais les w_i forment une famille libre, donc la combinaison linéaire (2.3.3) a tous ses coefficients nuls. Reste à démontrer que Z est de cardinal $\dim E$. Par construction, il y $\dim \operatorname{im} f$ vecteurs de la forme $f^i(v_j)$ dans Z à quoi il faut ajouter les $\dim \ker f - \dim(\ker f \cap \dim \operatorname{im} f)$ vecteurs w_i et les vecteurs u_i. Or, les vecteurs $f^{d_i}(u_i)$ forment une base de $\operatorname{im} f \cap \ker f$. Le cardinal de Z est donc $\dim \operatorname{im} f + \dim \ker f$, donc $\dim E$ par le théorème du rang. $\square$

On appelle base adaptée à f une base comme dans la proposition 2.23. Les sous-espaces engendrés par $(f^i(u_j))$ et par (w_k) sont stables par f (par construction pour les premiers, parce que les w_k sont dans le noyau de f pour le second). La matrice de f dans une base adaptée est donc diagonale par blocs. Certains des blocs sont nuls, les autres peuvent être réarrangés de telle sorte que la matrice de f ait tous ses coefficients a_{ij} nuls sauf ceux avec $j = i + 1$ qui sont égaux à 1.

2.3.7 Réduction des endomorphismes de $\mathbb{C}^2$

Soit E un $\mathbb{C}$-espace vectoriel de dimension 2 et $f \in \mathrm{End}(E)$. L'espace E admet une décomposition en sous-espaces propres généralisés

$$E = \bigoplus_{\lambda} E_\lambda.$$

Il n'y a donc que deux possibilités : ou bien f admet deux valeurs propres distinctes, auquel cas f est diagonalisable d'après la sous-section 2.3.5 et il existe donc une base dans laquelle sa matrice est la matrice $\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$, ou bien la multiplicité géométrique de l'unique valeur propre λ de f est égale à 2, auquel cas E est le noyau de $(f - \lambda\,\mathrm{Id})^2$ et l'application $g = f - \lambda\,\mathrm{Id}$ est donc nilpotente. Dans le second cas, il existe d'après la proposition 2.23 une base adaptée $(g(v), v)$ de E relativement à laquelle la matrice de f est $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. La matrice de f dans cette base est donc $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.

2.3.8 Théorème de réduction de Jordan

En rassemblant les résultats des sous-sections précédentes, on arrive au théorème suivant.

Théorème 2.3. *[Théorème de réduction de Jordan] Soit E un $\mathbb{C}$-espace vectoriel de dimension finie strictement positive et $f \in \mathrm{End}(E)$ un endomorphisme de E de spectre S. Il existe une décomposition de E*

$$E = \bigoplus_{\lambda \in S} E_\lambda$$

en sous-espaces stables généralisés et des choix de bases de chacun des E_λ tels que la matrice de f relativement à ce choix de base soit diagonale par blocs de taille $m \times m$ de la forme

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda & 1 & & \cdots & 0 \\ 0 & 0 & \lambda & & \ddots & \cdots & 0 \\ \vdots & \vdots & & & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & & \lambda & 1 \\ 0 & 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

c'est-à-dire tels que

$$a_{ij} = \begin{cases} \lambda \ si \ i = j \\ 1 \ si \ j = i+1 \\ 0 \ sinon. \end{cases}$$

Démonstration. D'après le théorème 2.1, il suffit de montrer qu'il existe une base de E_λ telle que la matrice de f restreinte à E_λ soit diagonale par blocs de la forme

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda & 1 & & \cdots & 0 \\ 0 & 0 & \lambda & \ddots & \cdots & 0 \\ \vdots & \vdots & & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & & \lambda & 1 \\ 0 & 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}.$$

L'endomorphisme $f - \lambda \,\mathrm{Id}$ est un endomorphisme nilpotent de E_λ donc admet une base adaptée d'après la proposition 2.23 telle que sa matrice dans cette base soit diagonale par blocs de la forme

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & & \cdots & 0 \\ 0 & 0 & 0 & \ddots & \cdots & 0 \\ \vdots & \vdots & & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & & 0 & 1 \\ 0 & 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}.$$

Dans cette base, la matrice de $\lambda \,\mathrm{Id}$ est la matrice $(\lambda \delta_{ij})$ car cette matrice est la matrice de $\lambda \,\mathrm{Id}$ dans toutes les bases. La matrice de f est donc la matrice

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda & 1 & & \cdots & 0 \\ 0 & 0 & \lambda & \ddots & \cdots & 0 \\ \vdots & \vdots & & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & & \lambda & 1 \\ 0 & 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

comme voulu. □

En utilisant la proposition 2.22 plutôt que la proposition 2.23, on obtient le théorème suivant qui est un peu moins précis mais souvent suffisant pour les applications.

Théorème 2.4. *Soit E un $\mathbb{C}$-espace vectoriel de dimension finie strictement positive et $f \in \mathrm{End}(E)$ un endomorphisme de E de spectre S. Il existe une décomposition de E*

$$E = \bigoplus_{\lambda \in S} E_\lambda$$

en sous-espaces stables généralisés et des choix de bases de chacun des E_λ tels que la matrice de f relativement à ce choix de base soit diagonale par blocs de taille $m \times m$

de la forme

$$\begin{pmatrix} \lambda & * & * & \cdots & \cdots & * \\ 0 & \lambda & * & & \cdots & * \\ 0 & 0 & \lambda & \ddots & \cdots & * \\ \vdots & \vdots & & \ddots & \ddots & * \\ 0 & 0 & \cdots & & \lambda & * \\ 0 & 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

c'est-à-dire tels que

$$a_{ij} = \begin{cases} \lambda \text{ si } i = j \\ 0 \text{ si } i > j. \end{cases}$$

Démonstration. Même démonstration que pour le théorème 2.3 mais en invoquant la proposition 2.22 plutôt que la proposition 2.23. □

2.3.9 Trace et déterminant d'un endomorphisme

Définition 2.24. *Soit E un $\mathbb{C}$-espace vectoriel de dimension finie n strictement positif et $f \in \mathrm{End}(E)$ un endomorphisme de spectre S. La trace* $\mathrm{tr}(f)$ *de f est la somme des valeurs propres de f comptées avec multiplicité et le déterminant* $\det(f)$ *de f est le produit des valeurs propres de f comptées avec multiplicité.*

Si

$$\mu_f = \sum_{s=0}^{n} a_s X^s = \prod_{\lambda \in S} (X - \lambda)$$

est le polynôme caractéristique μ_f, la trace de f est égale à $-a_{n-1}$ et le déterminant de f est $(-1)^n a_0$.

Lemme 2.25. *Un endomorphisme f est bijectif si et seulement si* $\det(f) \neq 0$*. Si* $\det(f) \neq 0$*, l'inverse de f est un polynôme d'endomorphismes en f.*

Démonstration. Un endomorphisme est bijectif si et seulement si 0 n'appartient pas à son spectre. D'après le théorème de Cayley-Hamilton, le polynôme d'endomorphismes

$$\mu_f(f) = \sum_{s=0}^{n} a_s f^s$$

est nul. Lorsque $\det(f) \neq 0$, il en résulte que

$$\mathrm{Id} = -\frac{1}{a_0} f \left(\sum_{s=0}^{n-1} a_{s+1} f^s \right)$$

et donc que l'inverse de f est un polynôme d'endomorphismes. □

Définition 2.26. *Soit $M = (a_{ij}) \in M_n(K)$ une matrice. La trace* $\mathrm{tr}(M)$ *de M est la somme de ses coefficients diagonaux.*

$$\mathrm{tr}(M) = \sum_{i=1}^{n} a_{ii}$$

Proposition 2.27. *Soit $A = (a_{ij}) \in M_n(K)$ et $B = (b_{ij}) \in M_n(K)$ deux matrices. Alors*
$$\mathrm{tr}(AB) = \mathrm{tr}(BA).$$

Démonstration. Soit $AB = (c_{ij})$ et $BA = (d_{ij})$. Alors
$$\begin{aligned}\mathrm{tr}(AB) &= \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}b_{ki} \\ &= \sum_{k=1}^{n}\sum_{i=1}^{n} b_{ki}a_{ik} = \sum_{k=1} d_{kk} = \mathrm{tr}(BA).\end{aligned}$$
□

Corollaire 2.28. *Soit f un endomorphisme d'un K-espace vectoriel. Le scalaire* $\mathrm{tr}(f)$ *défini comme étant la trace de la matrice de f relativement à un choix quelconque de base est bien défini et est égal à la trace de f lorsque $K = \mathbb{C}$.*

Démonstration. Soit en effet $M(f)$ et $N(f)$ les matrices de f relativement à deux choix de bases. Alors $N(f) = P^{-1}M(f)P$ pour P une matrice de passage. Donc $\mathrm{tr}(N(f)) = \mathrm{tr}(P^{-1}M(f)P) = \mathrm{tr}(M(f)P^{-1}P) = \mathrm{tr}(M(f))$. Si $K = \mathbb{C}$, il existe une base dans laquelle les termes diagonaux de la matrice de f sont les valeurs propres de f comptées avec multiplicité. La dernière assertion est donc vraie. □

Ce corollaire démontre en particulier que si f est un endomorphisme d'un $\mathbb{C}$-espace vectoriel dont la matrice dans une certaine base est à coefficients rationnels ou réels, alors il en est de même de sa trace. Ceci n'était pas évident en utilisant la définition de la trace comme somme des valeurs propres de f et n'est pas évident non plus pour le déterminant de f.

Proposition 2.29. *Soit f un endomorphisme de $\mathbb{C}^2$ dont la matrice dans une base est $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Alors la trace de f est $a+d$ et le déterminant de f est $ad-bc$. Autrement dit, le polynôme caractéristique de f est le polynôme*
$$\mu_f = X^2 - (a+d)X + (ad-bc).$$

Démonstration. L'endomorphisme f admet λ comme valeur propre si et seulement s'il existe un vecteur $u = \begin{pmatrix} x \\ y \end{pmatrix}$ non-nul propre pour la valeur propre λ donc si et seulement si le système
$$\begin{cases} ax + by = \lambda x \\ cx + dy = \lambda y \end{cases} \Leftrightarrow \begin{cases} (a-\lambda)x + by = 0 \\ cx + (d-\lambda)y = 0 \end{cases} \tag{2.3.4}$$
admet une solution non-nulle donc si et seulement si
$$(a-\lambda)(d-\lambda) - bc = \lambda^2 - (a+d)\lambda + (ad-bc) = 0.$$
Les valeurs propres de f, qui sont par définition les racines du polynôme unitaire μ_f, sont donc les racines du polynôme unitaire $\lambda^2 - (a+d)\lambda + (ad-bc)$. Ces deux polynômes unitaires sont donc égaux. Toutes les assertions de la proposition en résultent. □

2.3.10 Réduction des endomorphismes rationnels ou réels

Dans cette sous-section, on suppose que $K = \mathbb{Q}$ ou que $K = \mathbb{R}$ et que f est un endomorphisme d'un K-espace vectoriel E de dimension finie strictement positive. Nous avons vu que dans ce cas, il est tout à fait possible que f n'admette pas de vecteur propre et donc qu'il ne puisse être réduit comme dans les sous-sections précédentes. Néanmoins, f s'étend par linéarité dans ce cas à un endomorphisme du $\mathbb{C}$-espace vectoriel F de même dimension que E obtenu en considérant les combinaisons linéaires complexes d'une base de E. En tant qu'endomorphisme de F, f admet une réduction de Jordan. Dans la démonstration de cette réduction, nous avons utilisé deux ingrédients fondamentaux :

1. Le fait que F se décompose en sous-espaces propres généralisés.
2. Le fait qu'un endomorphisme nilpotent admet une base adaptée.

Le second de ces deux faits est général et demeure donc vraie pour E. Le premier en revanche utilise de manière cruciale le fait qu'un polynôme à coefficients complexes admet une racine complexe, mais aucune autre propriété du corps $\mathbb{C}$. Il résulte de cette discussion que f admet une décomposition de Jordan sur E à la condition suffisante (et bien sûr nécessaire) que toutes ses valeurs propres en tant qu'endomorphisme de F soit dans K.

3 Déterminants

A la fin de la section précédente, nous avons vu que la trace d'un endomorphisme d'un espace vectoriel complexe admettant une matrice à coefficients réels était réelle et avons soulevé une question similaire pour le déterminant. Nous avons aussi vu que le déterminant d'un endomorphisme de $\mathbb{C}^2$ de matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est $ad - bc$, qui est exactement la quantité qui apparaît dans la résolution des systèmes d'équations linéaires. L'objectif de cette section est de généraliser ces propriétés à la dimension n.

3.1 Formes linéaires, multilinéaires

3.1.1 Formes linéaires

Définition 3.1. *Soit E un K-espace vectoriel. Une forme linéaire de E est un élément de* $\mathrm{Hom}(E, K)$*, c'est-à-dire une application linéaire à coefficients dans K.*

La projection selon la première coordonnées dans la base canonique de K^n et plus généralement la projection selon une coordonnée quelconque sont des formes linéaires de K^n. L'application qui a un polynôme associe son n-ième coefficient est une forme linéaire de $K[X]$. L'évaluation $x \mapsto f(x)$ en $x \in U$ fixé est une forme linéaire de l'espace vectoriel $C^n(U, \mathbb{R})$ (pour U un ouvert de $\mathbb{R}$).

Proposition 3.2. *L'ensemble E^* des formes linéaires de E est un espace vectoriel. Supposons E de dimension finie égale à n et soit $(b_i)_{1\leq i\leq n}$ une base de E. Il existe une unique famille $(b_i^*)_{1\leq i\leq n}$ de formes linéaires vérifiant*

$$\forall (i,j)\in\{1,\cdots,n\}^2,\ b_i^*(b_j)=\delta_{ij}=\begin{cases}1 \text{ si } i=j\\ 0 \text{ sinon.}\end{cases}$$

La famille $(b_i^*)_{1\leq i\leq n}$ est une base de E^*. En particulier, l'espace vectoriel E^* est de dimension n. L'application

$$\begin{aligned}\mathrm{ev}:\ E &\longrightarrow E^{**}\\ u &\longmapsto (f\mapsto f(u))\end{aligned}$$

est un isomorphisme d'espaces vectoriels entre E et E^{**}.

Démonstration. L'ensemble E^* est $\mathrm{Hom}(E,K)$ donc est un espace vectoriel et il est de dimension n lorsque $\dim E=n$. Soit $i\in\{1,\cdots,n\}$. La forme linéaire b_i^* est définie en posant

$$b_i^*\left(\sum_{j=1}^n \alpha_j b_j\right)=\alpha_j.$$

Le fait que (b_i) soit une base implique alors que b_i^* est bien défini (car un vecteur u admet une unique décomposition comme combinaison linéaire des b_i) et la définition d'une combinaison linéaire montre alors que b_i^* est une application linéaire. Par construction, b_i^* vérifie bien $b_i^*(b_j)=\delta_{ij}$ pour tout $j\in\{1,\cdots,n\}$. Pour montrer que les b_i^* forment une base, il suffit de montrer qu'il forme une famille libre. Soit donc

$$f=\sum_{i=1}^n \alpha_i b_i^*=0$$

une combinaison linéaire nulle des b_i^*. Alors $f(e_j)=\alpha_j=0$ donc tous les α_i sont nuls.

Considérons enfin l'application

$$\begin{aligned}\mathrm{ev}:\ E &\longrightarrow E^{**}\\ u &\longmapsto (f\mapsto f(u))\end{aligned}$$

de E dans E^{**}. Soit $(u,v,\lambda)\in E^2\times K$ et $f\in E^*$. Alors $\mathrm{ev}(u+v)(f)=f(u+v)=f(u)+f(v)=(\mathrm{ev}(u)+\mathrm{ev}(v))(f)$ donc $\mathrm{ev}(u+v)=\mathrm{ev}(u)+\mathrm{ev}(v)$. De même $\mathrm{ev}(\lambda u)(f)=f(\lambda u)=\lambda f(u)=\lambda\,\mathrm{ev}(u)(f)$ donc $\mathrm{ev}(\lambda u)=\lambda\,\mathrm{ev}(u)$. Donc ev est une application linéaire. Soit $u\in E$ non nul. Alors u appartient à une base donc il existe une forme linéaire u^* vérifiant $u^*(u)=1$ donc $\mathrm{ev}(u)(u^*)\neq 0$ donc $\mathrm{ev}(u)\neq 0$. Par contraposition, ev est donc injective. Les dimensions de E et E^{**} étant égale, ev est un isomorphisme d'espaces vectoriels. □

L'espace vectoriel E s'appelle l'espace vectoriel dual de E et la base (b_i^*) s'appelle la base duale de (b_i). Notons que la construction de b_i^* dépend du choix de tous les b_j, si bien qu'à strictement parler on devrait plutôt écrire $b_i^{*,(b_j)}$. Le noyau d'une forme linéaire d'un espace vectoriel de dimension finie n est un espace vectoriel de dimension $n-1$ d'après le théorème du rang et s'appelle un hyperplan de E (par analogie avec les plans dans $\mathbb{R}^3$).

3.1.2 Formes bilinéaires

Définition 3.3. *Soit E un K-espace vectoriel. Une forme bilinéaire de E est une application*

$$(\cdot|\cdot): \ E \times E \longrightarrow K$$

qui est linéaire par rapport aux deux variables ; c'est-à-dire que pour tout $(u, v) \in E^2$, les applications $(u|\cdot)$ et $(\cdot|v)$ sont des formes linéaires.

Une forme bilinéaire est symétrique lorsque $(u|v) = (v|u)$ pour tout $(u, v) \in K^2$. Elle est antisymétrique lorsque $(u|v) = -(v|u)$ pour tout $(u, v) \in K^2$. Elle est définie lorsque $(u|u) \neq 0$ pour tout $u \in K$.

3.1.3 Formes multilinéaires

Définition 3.4. *Soit E un K-espace vectoriel. Une forme n-linéaire de E est une application*

$$\psi(\cdot, \cdots, \cdot): \ E^n \longrightarrow K$$

qui est linéaire par rapport aux n variables ; c'est-à-dire que pour tout j et pour tout $(u_i)_{1 \leq i \leq n-1} \in E^{n-1}$, l'application

$$\begin{aligned} \psi_j : \quad E \ &\longrightarrow K \\ u_j \ &\longmapsto \psi(u_1, \cdots, u_j, \cdots, u_{n-1}) \end{aligned}$$

est une forme linéaire. Une forme multilinéaire est une forme n-linéaire pour un certain n.

Une forme n-linéaire ψ est symétrique lorsque pour tout $(i, j) \in \{1, \cdots, n\}^2$ et toute famille $(u_s)_{1 \leq s \leq n}$ de E,

$$\psi(u_1, \cdots, u_i, \cdots, u_j, \cdots, u_n) = \psi(u_1, \cdots, u_j, \cdots, u_i, \cdots, u_n).$$

Une forme n-linéaire ψ est antisymétrique lorsque pour tout $(i, j) \in \{1, \cdots, n\}^2$ et toute famille $(u_s)_{1 \leq s \leq n}$ de E,

$$\psi(u_1, \cdots, u_i, \cdots, u_j, \cdots, u_n) = -\psi(u_1, \cdots, u_j, \cdots, u_i, \cdots, u_n). \tag{3.1.1}$$

Une forme n-linéaire est alternée lorsque $\psi(u_1, \cdots, u_n) = 0$ dès lors que $u_i = u_j$ pour $i \neq j$.

Lemme 3.5. *Une forme n-linéaire alternée est antisymétrique. Si $K = \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$, une forme multilinéaire antisymétrique est alternée.*

Démonstration. Soit ψ une forme multilinéaire. Supposons ψ alternée et fixons une paire d'indices (i, j). La forme bilinéaire ϕ obtenue en fixant toutes les vecteurs sauf ceux d'indices i ou j est alternée, et il suffit de montrer qu'elle est antisymétrique. Or

$$0 = \phi(u + v, u + v) = \phi(u, v) + \phi(v, u)$$

par bilinéarité de ϕ. Supposons réciproquement ψ antisymétrique. L'équation (3.1.1) appliquée à $\psi(u_1, \cdots, u_i, \cdots, u_i, \cdots, u_n)$ implique

$$\psi(u_1, \cdots, u_i, \cdots, u_i, \cdots, u_n) = -\psi(u_1, \cdots, u_i, \cdots, u_i, \cdots, u_n)$$

et donc $2\psi(u_1, \cdots, u_i, \cdots, u_i, \cdots, u_n) = 0$. Si $K = \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$, ceci entraîne

$$\psi(u_1, \cdots, u_i, \cdots, u_i, \cdots, u_n) = 0.$$

□

Proposition 3.6. *Soit $K = \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$. Soit E un K-espace vectoriel de dimension finie n strictement positive. Soit ψ une forme n-linéaire alternée. Soit $(u_i)_{1 \leq i \leq n}$ une famille de vecteurs de E. Alors $\psi((u_i)) \neq 0$ seulement si (u_i) est une base.*

Les formes n-linéaires alternées non-nulles (si elles existent) ont donc la faculté de détecter les bases.

Démonstration. Il suffit de montrer que $\psi((u_i)) \neq 0$ seulement si (u_i) est libre. Montrons la contraposée et supposons que la famille (u_i) est liée. Il existe alors une combinaison linéaire nulle

$$\sum_{i=1}^{n} \alpha_i u_i = 0$$

des u_i dont tous les coefficients ne sont pas nuls et dont on peut donc supposer sans perte de généralité que les coefficients α_j est égal à 1. Donc

$$\psi((u_i)) = \psi\left(u_1, \cdots, -\sum_{i \neq j} \alpha_i u_u, \cdots, u_n\right) = -\sum_{i \neq j} \alpha_i \psi(u_1, \cdots, u_i, \cdots, u_n)$$

Pour tout $i \neq j$, deux des vecteurs de $(u_1, \cdots, u_i, \cdots, u_n)$ sont égaux car u_i est en position j. Donc $\psi(u_1, \cdots, u_i, \cdots, u_n) = 0$ et il en est donc de même pour $\psi((u_i))$. □

Cette proposition admet une réciproque qui est néanmoins nettement plus difficile à démontrer. Commençons par une définition.

Définition 3.7. *Soit $\sigma \in \mathfrak{S}_n$ une permutation, c'est-à-dire une bijection de $\{1, \cdots, n\}$ vers lui-même. Soit $n(\sigma)$ le nombre d'échange nécessaires pour transformer le n-uplet $(\sigma(1), \sigma(2), \cdots, \sigma(n))$ en $(1, 2, \cdots, n)$ si l'on procède de la manière suivante : si les j premiers éléments sont dans l'ordre voulu, on échange si nécessaire $j+1$ et $\sigma(j+1)$. Soit $\varepsilon(\sigma)$ l'entier $(-1)^{n(\sigma)}$.*

Par exemple, si σ est la permutation telle que $(\sigma(1), \sigma(2), \sigma(3))$ soit égal à $(2, 3, 1)$, alors le procédé de la définition 3.7 transforme $(2, 3, 1)$ en $(1, 3, 2)$ puis en $(1, 2, 3)$ donc nécessite deux échanges. Donc $\varepsilon(\sigma) = (-1)^2 = 1$. Si maintenant σ est la permutation $(1, 3, 5, 2, 4)$, le procédé de la définition 3.7 transforme $(1, 3, 5, 2, 4)$ en $(1, 2, 5, 3, 4)$ puis en $(1, 2, 3, 5, 4)$ puis en $(1, 2, 3, 4, 5)$ donc nécessite 3 échanges. Donc $\varepsilon(\sigma) = (-1)^3 = -1$. On peut démontrer que $\varepsilon(\sigma)$ ne dépend en fait pas de la façon de réarranger le n-uplet $(\sigma(1), \cdots, \sigma(n))$ mais nous n'aurons pas besoin de ce fait.

Proposition 3.8. *Soit $K = \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$. Soit E un K-espace vectoriel de dimension finie n strictement positif. Soit ψ une forme n-linéaire alternée. Si ψ est non-nulle, alors $\psi((b_i)) \neq 0$ si et seulement si (b_i) est une base. Plus précisément, si (b_i) est une base, l'application $\psi \mapsto \psi((b_i))$ est une injection de l'ensemble des formes n-linéaires alternées vers K.*

Démonstration. Supposons ψ non-nulle. Il existe alors une famille (c_i) de vecteurs telle que $\psi((c_i)) \neq 0$ et cette famille est nécessairement une base. Soit (b_i) une autre famille de vecteurs. Nous savons déjà que $\psi((b_i)) \neq 0$ seulement si (b_i) est une base. Réciproquement, supposons que (b_i) soit une base. Pour tout j, c_j peut alors s'écrire

$$c_j = \sum_{i=1}^{n} \gamma_{ij} b_i$$

dans la base des (b_i). Donc

$$\psi((c_i)) = \psi\left(\sum_{i=1}^{n} \gamma_{i1} b_i, \cdots, \sum_{i=1}^{n} \gamma_{ij} b_j, \cdots, \sum_{i=1}^{n} \gamma_{in} b_i\right).$$

Développons cette expression en utilisant la linéarité par rapport à chaque variable et en indexant les indices eux-mêmes.

$$\psi((c_i)) = \psi\left(\sum_{i_1=1}^{n} \gamma_{i_1 1} b_{i_1}, \cdots, \sum_{i_j=1}^{n} \gamma_{i_j j} b_{i_j}, \cdots, \sum_{{}_n i=1}^{n} \gamma_{i_n n} b_{i_n}\right)$$

Donc

$$\psi((c_i)) = \sum_{i_1=1}^{n} \gamma_{i_1 1} \psi\left(b_{i_1}, \sum_{i_2=1}^{n} \gamma_{i_2 2} b_{i_2}, \cdots, \sum_{i_j=1}^{n} \gamma_{i_j j} b_{i_j}, \cdots \sum_{i_n=1}^{n} \gamma_{i_n n} b_i\right)$$

et plus généralement

$$\psi((c_i)) = \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} \left(\prod_{j=1}^{n} \gamma_{i_j j} \psi((b_{i_j}))\right). \tag{3.1.2}$$

La forme n-linéaire ψ étant alternée, $\psi((b_{i_j})) = 0$ sauf si tous les i_j sont distincts, ou de manière équivalente sauf si $j \mapsto i_j$ est une bijection σ de l'ensemble $\{1, \cdots, n\}$ vers lui-même. Les termes non-nuls de la somme du membre de droite de (3.1.2) sont donc les termes de la forme

$$\prod_{j=1}^{n} \gamma_{\sigma(j)j} \psi((b_{\sigma(j)j}))$$

où γ est une bijection de $\{1, \cdots, n\}^2$. Notons $\mathfrak{S}_n$ l'ensemble de ces bijections. Alors

$$\psi((c_i)) = \sum_{\sigma \in \mathfrak{S}_n} \prod_{j=1}^{n} \gamma_{\sigma(j)j} \psi((b_{\sigma(j)})).$$

Maintenant, on peut réordonner les $b_{\sigma(j)}$ en permutant, si nécessaire, $b_{\sigma(1)}$ avec b_1, puis $b_{\sigma(2)}$ avec b_2 et ainsi de suite. Comme dans la définition 3.7, notons $\varepsilon(\sigma) \in \{-1, 1\}$ le signe tel que

$$\psi((b_{\sigma(j)})) = \varepsilon(\sigma)\psi((b_j))$$

par le procédé de cette définition. On obtient finalement

$$\psi(c_1, c_2, \cdots, c_n) = \left(\sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\prod_{j=1}^{n} \gamma_{\sigma(j)j}\right) \psi(b_1, b_2, \cdots, b_n).$$

Donc $\psi(b_1, \cdots, b_n) \neq 0$.

Pour montrer, la dernière assertion, il suffit de remarquer que si $\psi((b_i)) = \phi((b_i))$ alors

$$\psi((c_i)) = 0 = \phi((c_i))$$

pour toute famille liée (c_i) tandis que

$$\psi((c_i)) = \left(\sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\prod_{j=1}^{n} \gamma_{\sigma(j)j}\right) \psi(b_1, b_2, \cdots, b_n) = \left(\sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\prod_{j=1}^{n} \gamma_{\sigma(j)j}\right) \phi(b_1, b_2, \cdots, b_n)$$

et

$$\phi((c_i)) = \left(\sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\prod_{j=1}^{n} \gamma_{\sigma(j)j}\right) \phi(b_1, b_2, \cdots, b_n)$$

pour toute famille libre (c_i). □

3.2 Déterminants

Nous sommes en position de définir le déterminant d'une matrice carrée et d'un endomorphisme.

3.2.1 Déterminant d'une matrice

Soit n un entier strictement positif, K le corps $\mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$ et soit $M = (a_{ij}) \in\in M_n(K)$.

Définition 3.9. *La déterminant de M est le scalaire*

$$\det(M) = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\prod_{j=1}^{n} a_{\sigma(j)j}.$$

Le déterminant d'une famille (b_i) de vecteurs de K^n est le déterminant de la matrice dont les vecteurs colonnes sont les b_i.

Exemples :

1. Soit $M = a \in M_1(K) \simeq K$. Alors $\det(M) = a$.

2. Soit $M = (m_{ij}) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(K)$. Alors

$$\det(M) = m_{\sigma(1)1}m_{\sigma(2)2} - m_{\tau(1)1}m_{\tau(2)2}$$

pour σ égal à l'identité et τ l'unique autre bijection de $\{1, 2\}$ vers lui-même, à savoir $\tau(1) = 2$ et $\tau(2) = 1$. Donc

$$\det(M) = ad - bc.$$

3. Soit $M = (\delta_{ij})$ la matrice de l'identité. La seule bijection σ de $\{1, \cdots, n\}$ vers lui-même telle que

$$\prod_{j=1}^{n} \delta_{\sigma(j)j} \neq 0$$

est l'identité et dans ce cas, ce produit vaut 1 donc $\det(M) = 1$. Plus généralement, le déterminant de $M = (\lambda\delta_{ij})$ est égal à λ^n. Plus généralement encore, si M est une matrice diagonale avec des λ_i sur la diagonale, alors

$$\det(M) = \prod_{i=1}^{n} \lambda_i.$$

4. Plus généralement encore, la seule permutation qui vérifie $\sigma(i) \leq i$ pour tout $i \in \{1, \cdots, n\}^2$ ou $\sigma(i) \geq i$ pour tout $i \in \{1, \cdots, n\}^2$ est l'identité donc

$$\det(M) = \prod_{i=1}^{n} \lambda_i$$

si M est triangulaire supérieure ou triangulaire inférieure avec des λ_i sur la diagonale.

3.2.2 Propriétés fondamentales du déterminant

Définition 3.10. *Soit n un entier strictement supérieur à 1. Soit $A = (a_{ij}) \in M_n(K)$. Pour $(k, \ell) \in \{1, \cdots, n\}^n$, la matrice $A_{k\ell}$ est la matrice $(a_{ij})_{i \neq k, j \neq \ell} \in M_{n-1}(K)$.*

Théorème 3.1. *Soit n un entier strictement positif. L'application ϕ_n qui à $A \in M_n(K)$ associe a si $n = 1$ et $A = (a)$ et qui associe à A*

$$\phi_n(A) = \sum_{i=1}^{n} (-1)^{i+1} a_{1i} \phi_{n-1}(A_{1i})$$

si $n > 1$ et $A = (a_{ij})$ est une forme multilinéaire alternée égale à 1 en la matrice de l'identité. De plus, $\phi_n = \det$ pour tout $n > 0$.

Démonstration. Nous démontrons toutes les assertions par récurrence. Elles sont toutes vraies si $n = 1$. Supposons maintenant que n soit strictement supérieur à 1 et que toutes les assertions au sujet de ϕ_{n-1} soient vraies pour toutes les matrices $M \in M_{n-1}(K)$. Supposons toutes les colonnes fixées sauf la colonne j qui est égale au vecteur u. Alors

$$\phi_n(u_1, \cdots, \lambda u_j, \cdots, u_n) = \phi(\lambda_j A) = \sum_{i=1}^{n} (-1)^{i+1} \lambda_{ij} a_{1i} \phi_{n-1}(\lambda_j A_{1i}) = \lambda \phi_n(A)$$

où $\lambda_j A$ est la matrice égale à A sauf en la colonne j où elle est égale à λu_j et où $\lambda_{ij} = 1$ sauf en $i = j$ ou $\lambda_{ij} = \lambda$. De même, soit v un vecteur.

$$\phi_n(u_1, \cdots, u + v, \cdots, u_n) = \sum_{i=1}^{n} (-1)^{i+1} (a+b)_{1i} \phi_{n-1}((A+B)_{1i}) = \phi_n(A) + \phi_n(B)$$

où B est la matrice égale à A sauf en la colonne j où elle est égale à v. L'application ϕ_n est donc multilinéaire. Enfin, si deux colonnes (disons k et ℓ) sont égales, alors les colonnes k et ℓ sont encore égales dans A_{ij} pour tout ij donc $\phi_{n-1}(A_{ij}) = 0$. Il en est donc de même pour $\phi_n(A)$ et ϕ_n est alternée. Enfin

$$\phi_n(\mathrm{Id}_n) = \phi_{n-1}(\mathrm{Id}_{n-1}) + \sum_{i=2}^{n} (-1)^{i+1} a_{1i} \phi_{n-1}(A_{1i})$$

et $a_{1i} = 0$ pour $i > 1$ donc $\phi_n(\mathrm{Id}_n) = \phi(\mathrm{Id}_{n-1}) = 1$.

D'après la proposition 3.8, il existe une unique forme n-linéaire alternée égale à 1 en la base canonique de K^n et cette forme vérifie

$$\phi_n(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}$$

si A est la matrice d'une famille (b_i) dans la base canonique. Donc $\phi_n = \det$. $\square$

Exemple : Le déterminant de la matrice

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

est égal à $aei - afh - bdi + bfg + cdh - ceg$.

Nous avons maintenant deux définitions potentielles du déterminant d'une matrice carrée. D'une part on peut utiliser la définition 3.9, ou bien on peut choisir un endomorphisme de $\mathbb{C}^n$ dont la matrice dans la base canonique de $\mathbb{C}^n$ est M et calculer $\det(f)$. Les calculs précédents montrent que ces deux définitions coincident lorsque M est une matrice triangulaire supérieure ou bien lorsque $n = 2$. Le théorème suivant assure qu'il en est toujours ainsi.

Théorème 3.2. *L'application*

$$\begin{aligned} \det : \quad M_n(K) &\longrightarrow K \\ (a_{ij}) &\longmapsto \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1} a_{\sigma(i)i} \end{aligned}$$

vue comme application de $(K^n)^n$ *est l'unique forme* n*-linéaire alternée égale à 1 en la matrice de l'identité. Cette application vérifie* $\det(AB) = \det(A)\det(B)$ *pour tout* $(A, B) \in M_n(K)$ *et* $\det(A^{-1}) = (\det(A))^{-1}$ *pour tout* $A \in \mathrm{GL}_n(K)$*. Soit* f *un endomorphisme de* $\mathbb{C}^n$ *et* $M(f)$ *sa matrice relativement à un choix quelconque de base. Alors* $\det(f)$ *est égal à* $\det(M(f))$*.*

Démonstration. D'après le théorème 3.1, l'application det est une forme n-linéaire alternée valant 1 en la base canonique de K^n. D'après la proposition 3.8, c'est donc l'unique telle forme.

Soit $(A, B) \in M_n(K)^2$. Soit A n'est pas inversible et alors AB ne l'est pas non plus et donc $\det(AB)$ et $\det(A)\det(B)$ sont tous les deux nuls. Soit A est inversible. L'application

$$\begin{aligned} \det(A-)/\det(A) : \quad M_n(K) &\longrightarrow K \\ B &\longmapsto \det(AB)/\det(A) \end{aligned}$$

est alors une forme n-linéaire alternée non-nulle de valeur 1 en l'identité I_n. D'après la proposition 3.8, les formes linéaires alternées $\det(A-)/\det(A)$ et det sont donc égales.

Soit $A \in \mathrm{GL}_n(K)$. Alors $\det(AA^{-1}) = 1$ et $\det(AA^{-1}) = \det(A)\det(A^{-1})$ donc $\det(A^{-1}) = (\det(A))^{-1}$. En particulier, $\det(M) = \det(P^{-1}MP)$ pour toute matrice inversible P et donc le déterminant le scalaire $\det(M(f))$ est bien défini pour tout $f \in \mathrm{End}(K^n)$. Soit $M(f)$ la matrice relativement à la base telle que $M(f)$ soit triangulaire supérieure de diagonale égale au spectre de f. Le le déterminant de $M(f)$ est alors égal au déterminant de f. □

Ce théorème met en particulier sur le plan le déterminant et la trace d'un endomorphisme.

3.2.3 Développement en ligne et en colonne

Définition 3.11. *La transposée* ${}^tA$ *d'une matrice* $A = (a_{ij}) \in M_{n,m}(K)$ *est la matrice* $({}^ta_{ij}) = (a_{ji}) \in M_{m,n}(K)$*.*

Lemme 3.12. *Soit* $(A, B) \in M_n(K)^2$*. Alors* ${}^t(AB) = {}^tB\,{}^tA$*. En particulier, si* $A \in M_n(K)$ *est inversible alors* ${}^tA$ *est inversible et* $({}^tA)^{-1} = {}^t(A^{-1})$*.*

Démonstration. Soit $A = (a_{ij})$, $B = (b_{ij})$ et $AB = (c_{ij})$.

$$(c_{ji}) = \sum_{k=1}^{n} a_{jk}b_{ki} = \sum_{k=1}^{n} b_{ki}a_{jk} = \sum_{k=1}^{n} {}^tb_{ik}\ {}^ta_{kj} = (b_{ji})(a_{ji})$$

Si A est inversible, alors $AA^{-1} = \mathrm{Id}_n$ donc ${}^t(A^{-1})\,{}^tA = {}^t\mathrm{Id}_n = \mathrm{Id}_n$ donc ${}^tA$ est inversible et $({}^tA)^{-1} = {}^t(A^{-1})$. □

Proposition 3.13. *Soit* $A \in M_n(\mathbb{C})$*. Alors* $\det(A) = \det({}^tA)$*.*

Démonstration. Soit f l'endomorphisme représenté par A dans la base canonique de $\mathbb{C}^n$. Il existe une base de $\mathbb{C}^n$ dans laquelle la matrice $M = P^{-1}AP$ de f est triangulaire supérieure. Si l'on nomme λ_i les éléments de la diagonale, $\det(A)$ est égal au produit des λ_i. La matrice $N = {}^tM$ est alors triangulaire inférieure avec les λ_i sur la diagonale et donc $\det(N) = \det(M)$. Par ailleurs, $N = {}^tP^{-1}(P^{-1}AP) = {}^tP\,{}^tA\,{}^tP^{-1}$ donc $\det(N) = \det({}^tA)$. □

Corollaire 3.14. *Pour toute ligne* i

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) \tag{3.2.1}$$

et pour toute colonne j

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}). \tag{3.2.2}$$

On appelle l'équation (3.2.1) le développement de $\det(A)$ relativement à la ligne i et l'équation (3.2.2) le développement de $\det(A)$ relativement à la colonne j

Démonstration. Le développement en colonne de $\det(A)$ relativement à la colonne j est le développement en ligne de $\det({}^tA) = \det(A)$ relativement à la ligne j. Il suffit donc de démontrer la première assertion. Elle est vraie pour $i = 1$ d'après le théorème 3.1. Supposons qu'elle soit vraie pour $i - 1 \geq 1$ et soit B la matrice de la famille de vecteurs égale à la famille des vecteurs de colonne de ${}^tA$ mais avec b_{i-1} et b_i permutés. Alors $\det(A) = \det({}^tA) = -\det(B) = -\det({}^tB)$. D'après l'hypothèse de récurrence, on peut calculer $\det(B)$ par un développement selon la ligne $i-1$ donc

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

Donc

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

□

3.3 Lien avec l'inversion et les systèmes linéaires

Définition 3.15. *Soit* $A = (a_{ij}) \in M_n(K)$ *une matrice. La co-matrice* A^* *de* A *est la matrice* a^* *définie par*

$$a^*_{ij} = (-1)^{i+j} \det(A_{ij}).$$

Le corollaire 3.14 a la conséquence suivante.

Proposition 3.16. *Soit $A \in M_n(K)$ une matrice. Alors A est inversible si et seulement si $\det(A) \neq 0$. Dans ce cas, son inverse est la matrice*

$$\frac{1}{\det(A)}\,{}^tA^* = (b_{ij})$$

définie par

$$b_{ij} = (-1)^{i+j}\frac{\det(A_{ji})}{\det(A)}$$

Démonstration. En effet, le terme (i,j) de $A\,{}^tA^*$ est donné par

$$\sum_{k=1}^{n}(-1)^{j+k}a_{ik}\det(A_{jk})$$

donc est d'après le corollaire 3.14 égal au déterminant de la matrice dont toutes les lignes sont égales à celles de A sauf la ligne j qui est remplacé par la ligne i de A (développé selon la ligne i). Lorsque $i \neq j$, cette matrice a deux lignes identiques et son déterminant est donc nul. Lorsque $i = j$, cette matrice est égale à A et son déterminant est donc égal à $\det(A)$. On a donc bien $b_{ij} = \delta_{ij}$ et $A^{-1} = {}^tA^*/\det(A)$. □

Exemple : Soit M la matrice

$$\begin{pmatrix} 1 & 2 & 3 \\ 5 & 0 & -1 \\ 2 & 2 & 2 \end{pmatrix}.$$

Alors $\det(M) = 8$ donc M est inversible et son inverse est égal à

$$M^{-1} = \frac{1}{8}\begin{pmatrix} 2 & 2 & -2 \\ -12 & -4 & 16 \\ 10 & 2 & -10 \end{pmatrix}.$$

Soit N la matrice

$$\begin{pmatrix} 1 & -2 & -2 & 0 \\ -1 & 2 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

Alors $\det(N) = 2$ donc N est inversible et son inverse est égal à

$$N^{-1} = \frac{1}{2}\begin{pmatrix} 0 & 2 & 4 & 0 \\ 0 & 2 & 2 & 0 \\ -1 & -1 & 0 & 0 \\ 0 & -2 & -4 & 2 \end{pmatrix}.$$

Corollaire 3.17. *Le système d'équations linéaires à n équations et n inconnues*

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = u_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = u_2 \\ \cdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n = u_n \end{cases} \tag{3.3.1}$$

admet une unique solution si et seulement si $\det(a_{ij}) \neq 0$*. Lorsque* $\det(a_{ij}) \neq 0$*, l'unique solution* (x_i) *de* (3.3.1) *est*

$$x_i = \frac{1}{\det(A)} \sum_{k=1}^{n} (-1)^{i+k} \det(A_{ki}) u_k.$$

Donc $x_i = \det(b_{k\ell})$ *pour*

$$b_{k\ell} = \begin{cases} a_{k\ell} \text{ si } \ell \neq i \\ u_k \text{ si } \ell = i \end{cases}.$$

Démonstration. Le système (3.3.1) admet une unique solution si et seulement si l'équation matricielle

$$AX = U$$

admet une unique solution donc si et seulement si $A = (a_{ij})$ est une matrice inversible donc si et seulement si $\det(a_{ij}) \neq 0$. Lorsque A est inversible, l'unique solution de $AX = U$ est $X = A^{-1}U$ donc $X = (x_i)$ avec

$$x_i = \frac{1}{\det(A)} \sum_{k=1}^{n} (-1)^{i+k} \det(A_{ki}) u_k$$

d'après la proposition 3.16. D'après le corollaire 3.14, x_i est donc égal au déterminant de la matrice dont toutes les colonnes sont égales à A sauf la colonne i qui est égale à U (développé selon la colonne i) divisé par le déterminant de A. □

Exemple : Soit (S) le système d'équations linéaires

$$\begin{cases} 3x_1 + 3x_2 + 3x_3 = u_1 \\ -2x_1 - x_2 - 2x_3 = u_2 \\ 3x_1 + 3x_2 + 2x_3 = u_3 \end{cases}$$

correspondant à l'équation $AX = U$. Alors $\det(a_{ij}) = -3$ donc (S) admet une unique solution donnée par

$$\begin{cases} x_1 = -\frac{1}{3}(4u_1 + 3u_2 - 3u_3) \\ x_2 = \frac{1}{3}(2u_1 + 3u_2) \\ x_3 = u_1 - u_3 \end{cases}.$$

Il est important de remarquer que le déterminant est un outil théorique intéressant pour résoudre les systèmes mais qu'en pratique il ne faut jamais l'utiliser. En effet, la méthode de résolution nécessite de calculer $n + 1$ déterminants pour résoudre un système de n équations, et ceci requiert toujours beaucoup plus de calculs que la méthode par réduction.

3.4 Lien avec la réduction

Définition 3.18. *Une fonction polynomiale P à valeurs dans K est une fonction de K vers K telle qu'il existe des scalaires $(a_s)_{0\leq s\leq n}$ tels que*

$$P(\lambda) = \sum_{s=0}^{n} a_s \lambda^s$$

pour tout $\lambda \in K$. Une fonction polynomiale matricielle M est une fonction de K vers $M_n(K)$ telle qu'il existe des fonctions polynomiales $(P_{ij})_{1\leq i,j\leq n}$ telles que

$$M(\lambda) = \left(P_{ij}(\lambda)\right)_{ij} \in M_n(K)$$

pour tout $\lambda \in K$.

A tout polynôme $P \in K[X]$ est associé la fonction polynomiale $\lambda \mapsto P(\lambda)$ que l'on note également P par abus de notation.

Proposition 3.19. *La fonction polynomiale $\lambda \mapsto \det(\lambda \operatorname{Id} - f)$ est égale au polynôme caractéristique μ_f de f évalué en λ. En particulier, si $K = \mathbb{C}$ alors*

$$\det(\lambda \operatorname{Id} - f) = \prod_{\mu \in \operatorname{Spec}(f)} (\lambda - \mu)$$

Démonstration. Développons le déterminant $\det(\lambda \operatorname{Id} - f)$ (ou bien en utilisant la définition, ou bien selon la première colonne). Il apparaît que le seul terme en λ^n du développement est de coefficient 1 (cela se voit par récurrence si le développement est selon la première colonne ou en remarquant que c'est le coefficient correspondant à la bijection identité de $\{1, \cdots, n\}$ si on utilise la définition). La fonction polynomiale $\det(\lambda \operatorname{Id} - f)$ est donc une fonction polynomiale unitaire de degré n nulle exactement lorsque λ est une racine de μ_f. Elle est donc égale à $\mu_f(\lambda)$. □

Corollaire 3.20. *Soit $f \in \operatorname{End}(E)$ un endomorphisme d'un K-espace vectoriel de dimension finie n dont la matrice M dans une base de E est à coefficients dans $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$. Alors le polynôme caractéristique de f est un polynôme à coefficients dans $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$.*

Démonstration. En effet, la fonction polynomiale matricielle $\lambda \mapsto \lambda \operatorname{Id} - M$ est alors à valeurs dans $M_n(A)$ pour A égal à $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$ et $\det(\lambda \operatorname{Id} - M)$ est alors un polynôme à coefficients dans A. □

Bien que cette proposition donne un outil théorique intéressant, il est important de garder à l'esprit qu'il ne faut en général pas tenter de déterminer les valeurs propres d'un endomorphisme de cette façon : non seulement cela demande beaucoup plus de calculs que de déterminer les espaces propres en résolvant les systèmes linéaires associés, mais en plus ou bien l'endomorphisme en question est un endomorphisme d'espaces vectoriels de dimension 1 ou 2 (auquel cas il existe une formule directe donnant le polynôme caractéristique) ou bien c'est un endomorphisme d'espaces vectoriels de dimension $n \geq 3$ et alors on ne connaît en général pas de méthode efficace pour trouver les racines de ce polynôme.

4 Espaces vectoriels euclidiens

4.1 Définitions et premières propriétés

4.1.1 Produit scalaire

Définition 4.1. *Soit E un $\mathbb{R}$-espace vectoriel. Une forme bilinéaire $(\cdot|\cdot)$ sur E est positive si $(u|u) \geq 0$ pour tout $u \in E$. Elle est définie si $(u|u) = 0$ seulement si $u = 0$.*

Un espace vectoriel euclidien E est un $\mathbb{R}$-espace vectoriel de dimension finie n muni d'une forme $(\cdot|\cdot)$ bilinéaire symétrique, définie et positive. La forme bilinéaire symétrique définie positive d'un espace vectoriel euclidien est appelé le produit scalaire euclidien de cet espace. Le produit scalaire d'un vecteur $u \in E$ avec lui-même est un réel positif dont la racine carrée est appelée la norme de ce vecteur et est notée $\|u\|$. Par définition, la norme d'un vecteur non-nul est un réel strictement positif.

Exemple : L'exemple fondamental d'espace euclidien est l'espace vectoriel $\mathbb{R}^n$ muni de la forme $(\cdot|\cdot)$ définie par

$$(u|v) = \sum_{i=1}^{n} u_i v_i$$

où

$$u = \sum_{i=1}^{n} u_i e_i, \ v = \sum_{i=1}^{n} v_i e_i$$

et $(e_i)_{1\leq i\leq n}$ est la base canonique de $\mathbb{R}^n$. Cet exemple est tellement important que nous le promouvons au rang de théorème de ce cours.

Théorème 4.1. *Soit $n \geq 1$. L'espace $\mathbb{R}^n$ muni de la forme*

$$\begin{aligned}(\cdot|\cdot) : \quad \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ ((u_i), (v_i)) &\longmapsto \sum_{i=1}^{n} u_i v_i\end{aligned}$$

est un espace vectoriel euclidien.

Démonstration. Il suffit de montrer que $(\cdot|\cdot)$ est une forme bilinéaire symétrique définie positive. Soit donc $(u, v, w, \lambda) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$. Alors

$$(u + v|w) = \sum_{i=1}^{n} (u_i + v_i) w_i = \sum_{i=1}^{n} u_i w_i + \sum_{i=1}^{n} v_i w_i = (u|w) + (v|w)$$

et

$$(\lambda u|v) = \sum_{i=1}^{n} \lambda u_i v_i = \lambda (u|v)$$

donc $(\cdot|\cdot)$ est une forme bilinéaire. De plus,

$$(u|v) = \sum_{i=1}^{n} u_i v_i = \sum_{i=1}^{n} v_i u_i = (v|u)$$

donc $(\cdot|\cdot)$ est symétrique. De

$$(u|u) = \sum_{i=1}^{n} u_i^2,$$

on déduit que $(u|u) \geq 0$ et que $(u|u) = 0$ seulement si tous les u_i sont nuls, et donc seulement si $u = 0$. $\square$

Le produit scalaire de $\mathbb{R}^n$ du théorème 4.1 est appelé le produit scalaire canonique de $\mathbb{R}^n$. Notons tout de même que le produit scalaire canonique de $\mathbb{R}^n$ n'est pas le seul produit scalaire défini sur $\mathbb{R}^n$ qui fasse de $\mathbb{R}^n$ un espace vectoriel euclidien. A titre d'exemple, on peut considérer

$$\begin{array}{rrl} <\cdot,\cdot>: & \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow \mathbb{R} \\ & ((u_i),(v_i)) & \longmapsto \displaystyle\sum_{i=1}^{n} 2^i u_i v_i \end{array}$$

et vérifier qu'il s'agit bien d'une forme bilinéaire symétrique définie positive. Un exemple essentiellement différent est la forme bilinéaire

$$(P,Q) \mapsto \int_0^1 P(x)Q(x)dx$$

sur l'espace vectoriel $\mathbb{R}_n[X]$.

La proposition suivante montre que le calcul du produit scalaire de deux vecteurs peut se faire uniquement à partir de leurs normes et de la norme de leur somme.

Proposition 4.2. *Soit E un espace euclidien et soit $(u,v) \in E^2$. Alors*

$$(u|v) = \frac{\|u+v\|^2 - \|u\|^2 - \|v\|^2}{2}.$$

Démonstration. Par bilinéarité de $(\cdot|\cdot)$, la norme de $u+v$ vérifie

$$(u+v|u+v) = (u|u) + 2(u|v) + (v|v) = \|u\|^2 + \|v\|^2 + 2(u|v)$$

et donc

$$(u|v) = \frac{\|u+v\|^2 - \|u\|^2 - \|v\|^2}{2}.$$

$\square$

4.1.2 L'inégalité de Cauchy-Schwarz

Théorème 4.2. *Soit E un espace euclidien. Pour tout $(u,v) \in E^2$,*

$$|(u|v)| \leq \|u\|\|v\| \tag{4.1.1}$$

et cette inégalité est une égalité si et seulement si (u,v) est une famille liée.

Démonstration. Soit $(u,v) \in E^2$. Supposons tout d'abord que l'un des vecteurs u ou v est nul. Alors les deux membres de l'inégalité (4.1.1) sont nuls donc l'inégalité est vraie, est en fait une inégalité et (u,v) est bien une famille liée. Maintenant, on suppose que l'un des deux vecteurs, disons v, est non-nul. Considérons la fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$ définie par

$$f(t) = (u+tv|u+tv).$$

Alors

$$f(t) = (u|u) + 2t(v|u) + t^2(v|v)$$

donc f est une fonction polynomiale du second degré exactement. La forme bilinéaire $(\cdot|\cdot)$ est définie positive donc $f(t) \geq 0$ pour tout $t \in \mathbb{R}$ et $f(t) = 0$ si et seulement si $u + tv = 0$. Donc $4(u|v)^2 - 4(u|u)(v|v) \leq 0$ et $4(u|v)^2 - 4(u|u)(v|v) = 0$ si et seulement s'il existe $t \in \mathbb{R}$ tel que $u + tv = 0$; c'est-à-dire si et seulement si (u,v) est une famille liée. En prenant la racine carrée, on obtient bien

$$|(u|v)| \leq \|u\|\|v\|$$

avec égalité si et seulement si (u,v) est une famille liée. □

Corollaire 4.3. *Soit $(u,v) \in E^2$. Alors*

$$\|u+v\| \leq \|u\| + \|v\|$$

avec égalité si et seulement si (u,v) est une famille liée.

Démonstration. Soit $(u,v) \in E^2$. Par définition de la norme, linéarité de $(\cdot|\cdot)$ et par l'inégalité (4.1.1) de Cauchy-Schwarz, on obtient

$$\begin{aligned}
\|u+v\|^2 &= (u+v|u+v) \\
&= (u|u) + (v|v) + 2(u|v) \\
&\leq \|u\|^2 + \|v\|^2 + 2\|u\|\|v\| \\
&= (\|u\| + \|v\|)^2
\end{aligned}$$

et on conclut en prenant les racines carrées des deux membres. Ces inégalités sont des égalités si et seulement si l'inégalité de Cauchy-Schwarz est une égalité donc si et seulement si (u,v) est une famille liée. □

4.1.3 Bases orthogonales, orthonormalisation

Définition 4.4. *Une famille $(b_i)_{1 \leq i \leq m}$ d'un espace vectoriel euclidien de produit scalaire $(\cdot|\cdot)$ est orthogonale si $(e_i|e_j) = 0$ si $i \neq j$. Elle est normale si $(e_i|e_i) = 1$ pour tout i. Elle est orthonormale si elle est orthogonale et normale, donc si $(e_i|e_j) = \delta_{ij}$ pour tout $1 \leq i,j \leq m$.*

La base canonique est orthonormale pour le produit scalaire canonique de $\mathbb{R}^n$. Soit $(b_i)_{1 \leq i \leq n}$ une base orthonormale de l'espace vectoriel euclidien E et soit

$$u = \sum_{i=1}^{n} \alpha_i b_i \in E.$$

Alors

$$\|u\|^2 = (u|u) = \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j(b_i|b_j) = \sum_{i=1}^{n}\alpha_i^2 \tag{4.1.2}$$

par bilinéarité de $(\cdot|\cdot)$ et orthonormalité de la base (b_i). Si $u \in E$ est un vecteur non-nul, alors $(u|u)$ est strictement positif et le vecteur

$$v = \frac{1}{\sqrt{(u|u)}}u$$

est donc bien défini. De

$$(v|v) = (\frac{1}{\sqrt{(u|u)}}u|\frac{1}{\sqrt{(u|u)}}u) = \frac{1}{(u|u)}(u|u) = 1,$$

et de $(\lambda v|\lambda v) = \lambda^2(v|v) = \lambda^2$, on déduit qu'il existe exactement deux vecteurs de $\mathrm{Vect}(u)$ de norme 1 : à savoir v et $-v$. Le processus qui associé à $u \in E$ non-nul le vecteur

$$v = \frac{1}{\|u\|}u$$

s'appelle la normalisation et l'on dit que v est le normalisé de u.

Lemme 4.5. *Soit $(b_1, \cdots, b_m)$ une famille orthonormale d'un espace euclidien E et soit*

$$u = \sum_{i=1}^{m}\alpha_i b_i \in \mathrm{Vect}(b_i).$$

Alors $\alpha_i = (u|b_i)$ pour tout i et

$$u = \sum_{i=1}^{m}(u|b_i)b_i.$$

En particulier, une famille orthogonale de vecteurs non nuls est libre.

Démonstration. Le vecteur u s'écrit de manière unique sous la forme

$$u = \sum_{i=1}\alpha_i b_i.$$

Par bilinéarité de $(\cdot|\cdot)$ et orthonormalité de la base (b_i),

$$(u|b_j) = \sum_{i=1}^{n}\alpha_i(b_i|b_j) = \alpha_j$$

pour tout j. La deuxième assertion découle de la première en normalisant les vecteurs de la famille orthogonale, ce qui est possible car ils sont non-nuls par hypothèse. □

Proposition 4.6. *Soit E un espace vectoriel euclidien et $(v_1, \cdots, v_m)$ une famille libre de E. Il existe alors une base orthonormale $(b_1, \cdots, b_m)$ de $\mathrm{Vect}(v_1, \cdots, v_m)$.*

Démonstration. S'il existe une base orthogonale comme dans la proposition, en normaliser les vecteurs produira une base orthonormale comme dans la proposition et la proposition sera donc démontrée. Il suffit donc de montrer qu'il existe une base orthogonale $(b_1, \cdots, b_m)$ de $\mathrm{Vect}(v_1, \cdots, v_m)$. Démontrons le par récurrence sur m. Soit $m = 1$ et la condition d'orthogonalité est vide, donc satisfaite. Supposons maintenant que l'espace vectoriel engendré par une familles libre de cardinal au plus $m \geq 1$ admette une base orthonormale et soit $(v_1, \cdots, v_{m+1})$ une famille libre. Il existe une base orthonormale (b_i) de $\mathrm{Vect}(v_1, \cdots, v_m)$ par l'hypothèse de récurrence donc il suffit de montrer qu'il existe un vecteur non-nul b_{m+1} orthogonal aux b_i tel que v_{m+1} appartienne à $\mathrm{Vect}(b_i)_{1 \leq i \leq m+1}$. Soit

$$b_{m+1} = v_{m+1} - \sum_{i=1}^{m} (v_{m+1}|b_i) b_i.$$

. Alors v_{m+1} appartient à $\mathrm{Vect}(b_1, \cdots, b_{m+1})$. La famille des v_i étant libre, v_{m+1} n'appartient pas à $\mathrm{Vect}(v_1, \cdots, v_m)$ donc à $\mathrm{Vect}(b_1, \cdots, b_m)$ donc b_{m+1} est non-nul. Soit $1 \leq j \leq m$ un entier.

$$(b_{m+1}|b_j) = (v_{m+1}|b_j) - \sum_{i=1}^{m} (v_{m+1}|b_i)(b_i|b_j) = (v_{m+1}|b_j) - (v_{m+1}|b_j) = 0$$

par bilinéarité de $(\cdot|\cdot)$ et orthonormalité de la famille $(b_i)_{1 \leq i \leq m}$. Donc b_{m+1} est orthogonal à tous les b_i. □

La famille orthonormale que l'on obtient à partir de la famille libre (v_i) par le procédé de la proposition précédente s'appelle la famille de Gram-Schmidt associée à (v_i).

4.1.4 Orthogonal, supplémentaire orthogonal

Définition 4.7. *Deux sous-espaces vectoriels F et G d'un espace vectoriel euclidien E sont orthogonaux si $(u|v) = 0$ pour tout $(u, v) \in F \times G$. Ils sont supplémentaires orthogonaux s'ils sont supplémentaires et orthogonaux. L'orthogonal $F^\perp$ d'un sous-espace vectoriel est*

$$\{v \in E | \forall\ u \in F,\ (u|v) = 0\}.$$

Tout sous-espace vectoriel $G \subset E$ orthogonal à F est par définition inclus dans $F^\perp$.

Proposition 4.8. *Deux espaces vectoriels orthogonaux sont en somme directe. Un sous-espace vectoriel F admet comme unique supplémentaire orthogonal $F^\perp$.*

Démonstration. Soit E un espace euclidien de dimension n et soit F, G deux sous-espaces. Supposons F et G orthogonaux. Soit $u \in F \cap G$. Alors $(u|u) = 0$ donc $u = 0$. Par définition, $F^\perp$ est l'intersection sur $u \in F$ des noyaux des formes linéaires $(u|\cdot)$.

$$F^\perp = \bigcap_{u \in F} \ker(u|\cdot)$$

Donc $F^\perp$ est un sous-espace vectoriel de E orthogonal à F par construction. D'après la première assertion, $F^\perp$ est donc en somme directe avec F et il suffit pour conclure de montrer que $F + F^\perp = E$. Soit $(b_i)_{1\leq i\leq m}$ une base de F et (b_i, c_i) une base de E obtenue en complétant la famille des (b_i). Le procédé d'orthnormalisation de la proposition 4.6 produit alors une base (u_i, v_i) orthonormale de E telle que la famille $(u_i)_{1\leq i\leq m}$ soit une base de F. Il s'ensuit que $u \in E$ s'écrit

$$u = \sum_{i=1}^{m} \alpha_i u_i + \sum_{i=m+1}^{n} \alpha_i v_i = v + w$$

avec $v \in F$ et $w \in F^\perp$ et donc que $F + F^\perp = E$. □

En particulier, l'orthogonal de E tout entier est $\{0\}$.

Corollaire 4.9. *Soit $F \subset E$ un sous-espace vectoriel d'un espace euclidien. Alors $(F^\perp)^\perp = F$.*

Démonstration. En effet, F et $(F^\perp)^\perp$ sont tous deux l'unique supplémentaire orthogonal de $F^\perp$. □

Soit $F \subset E$ un sous-espace vectoriel d'un espace euclidien. A la décomposition

$$E = F \oplus F^\perp$$

est associé un projecteur $p \in \mathrm{End}(E)$ d'image F et de noyau $F^\perp$. Si $u \in E$ s'écrit (nécessairement de manière unique) $u = v + w$ avec $v \in F$ et $w \in F^\perp$, alors $p(u)$ est égal à v. L'endomorphisme p vérifie donc $p^2 = p$ et est en fait l'unique endomorphisme de E vérifiant $p^2 = p$, de noyau $F^\perp$ et d'image F. On appelle cet endomorphisme le projecteur orthogonal de E sur F.

4.2 Endomorphisme adjoint

4.2.1 Définition

Soit E un espace vectoriel euclidien de produit scalaire $(\cdot|\cdot)$.

Proposition 4.10. *Soit E un espace vectoriel euclidien. L'application*

$$\begin{array}{rccl} (|\cdot): & E & \longrightarrow & E^* \\ & u & \longmapsto & (\cdot|u) \end{array}$$

est un isomorphisme d'espace vectoriels. En particulier, pour toute forme linéaire ϕ, il existe un unique vecteur $v_\phi \in E$ tel que $\phi = (\cdot|v_\phi)$.

Démonstration. Par bilinéarité de $(\cdot|\cdot)$, l'application $(|\cdot)$ est linéaire. Comme E et E^* ont même dimension, il suffit de démontrer que $(|\cdot)$ est une injection et donc que son noyau est nul. Mais $(\cdot|u)$ est la forme linéaire nulle seulement si $u = 0$ car $(\cdot|\cdot)$ est définie. La deuxième assertion est une reformulation de la première. □

Théorème 4.3. *Soit $f \in \mathrm{End}(E)$ un endomorphisme de E. Il existe une unique endomorphisme $f^* \in \mathrm{End}(E)$ tel que $(f(u)|v) = (u|f^*(v))$ pour tout $(u,v) \in E^2$.*

Démonstration. Soit $v \in E$. L'application $u \mapsto (f(u)|v)$ est une forme linéaire donc il existe d'après la proposition 4.10 un unique vecteur w tel que $(f(u)|v) = (u|w)$ pour tout $u \in E$. Posons

$$f^* : \begin{array}{rcl} E & \longrightarrow & E \\ v & \longmapsto & w \end{array}$$

Il suffit de démontrer que f^* est bien une application linéaire. Soit donc $(u, v_1, v_2) \in E^3$ et $\lambda \in \mathbb{R}$. Alors

$$(f(u)|v_1 + v_2) = (f(u)|v_1) + (f(u)|v_2) = (u|f^*(v_1)) + (u|f^*(v_2)) = (u|f^*(v_1) + f^*(v_2))$$

donc $f^*(v_1 + v_2) = f^*(v_1) + f^*(v_2)$. De même

$$(f(u), \lambda v_1) = \lambda(f(u), v_1) = \lambda(u, f^*(v_1)) = (u|\lambda f^*(v_1))$$

donc $f^*(\lambda v_1) = \lambda f^*(v_1)$. Donc f^* est un endomorphisme de E. □

Définition 4.11. *L'endomorphisme f^* est appelé l'adjoint de f. Un endomorphisme est dit auto-adjoint s'il est égal à son adjoint, c'est-à-dire si $f^* = f$. Un endomorphisme est dit orthogonal s'il est inversible d'inverse son adjoint, c'est-à-dire si $f^* = f^{-1}$.*

Exemples :

1. L'identité est un endomorphisme auto-adjoint et orthogonal.

2. Les endomorphismes de la forme $\lambda\,\mathrm{Id}$ sont auto-adjoints. Ils sont orthogonaux si et seulement si $\lambda \in \{\pm 1\}$.

3. Un endomorphisme dont la matrice dans la base canonique de $\mathbb{R}^n$ est diagonale est auto-adjoint pour le produit scalaire canonique. Un tel endomorphisme est orthogonal si et seulement si toutes ses valeurs propres sont égales à 1 ou -1.

4. Soit $F \subset E$ un sous-espace vectoriel d'un espace euclidien et soit p le projecteur orthogonal de E sur F. Soit $(u,v) \in E^2$ qui s'écrivent respectivement $u = p(u) + x$ et $v = p(v) + y$ avec $(x,y) \in (F^\perp)^2$. Alors $(p(u)|v) = (p(u)|p(v)) = (u|p(v))$ car $(x|p(v))$ et $(p(u)|y)$ sont nuls. L'endomorphisme p est donc auto-adjoint.

5. L'endomorphisme dont la matrice dans la base canonique de $\mathbb{R}^2$ est $\begin{pmatrix} 3 & 2 \\ 2 & -5 \end{pmatrix}$ est auto-adjoint pour le produit scalaire canonique (mais pas orthogonal).

6. L'endomorphisme dont la matrice dans la base canonique de $\mathbb{R}^2$ est $\begin{pmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{pmatrix}$ est orthogonal pour le produit scalaire canonique (mais pas auto-adjoint).

L'opération d'adjonction est une involution linéaire.

Proposition 4.12. *L'application*

$$\begin{aligned} {}^* : \quad \mathrm{End}(E) &\longrightarrow \mathrm{End}(E) \\ f &\longmapsto f^* \end{aligned}$$

est un isomorphisme de l'espace vectoriel $\mathrm{End}(E)$ *qui est son propre inverse. Si* $(f, g) \in \mathrm{End}(E)^2$*, alors* $(fg)^* = g^* f^*$*. En particulier, la composé de deux endomorphismes auto-adjoints qui commutent est un endomorphisme auto-adjoint. Si* f *est auto-adjoint, tout polynôme d'endomorphisme en* f *est auto-adjoint.*

Démonstration. Soit $(f, g, \lambda) \in \mathrm{End}(E)^2 \times \mathbb{R}$ et soit $(u, v) \in E^2$. Alors

$$\begin{aligned} ((f+g)(u)|v) &= (f(u) + g(u)|v) = (f(u)|v) + (g(u)|v) \\ &= (u|f^*(v)) + (u|g^*(v)) = (u|f^*(v) + g^*(v)) \end{aligned}$$

donc $(f+g)^* = f^* + g^*$. De même

$$(\lambda f(u)|v) = \lambda(f(u)|v) = \lambda(u|f^*(v)) = (u|\lambda f^*(v))$$

donc $(\lambda f)^* = \lambda f^*$. Donc * est une application linéaire de $\mathrm{End}(E)$ vers lui-même.

Pour montrer que * est un isomorphisme, il suffit donc de montrer que c'est une injection. Soit donc f dans le noyau de * et $u \in E$. Alors $(u|f^*(f(u))) = 0$ donc $(f(u)|f(u)) = 0$ donc $f(u) = 0$. Donc f est l'endomorphisme nul. Enfin, soit $(u, v) \in E^2$. Alors $(f^{**}(u)|v) = (u|f^*(v)) = (f(u)|v)$ donc $f^{**}(u) - f(u)$ est dans l'orthogonal de E, donc est nul. Donc $f^{**} = f$.

Soit maintenant $(f, g) \in \mathrm{End}(E)^2$ et $(u, v) \in E^2$. Alors $(f(g(u))|v) = (g(u)|f^*(v)) = (u|g^*(f^*(v)))$ donc $(fg)^* = g^* f^*$. Si f et g sont auto-adjoints et commutent, alors $(fg)^* = g^* f^* = gf = fg$ donc fg est auto-adjoint. Un polynôme d'endomorphismes en f est une combinaison linéaire de puissances de f donc est auto-adjoint si f l'est par l'assertion précédente et par linéarité de l'adjonction. □

Notons que f est orthogonal si et seulement si $(f(u)|f(v)) = (u|v)$ pour tout $(u, b) \in E^2$, donc si et seulement si f préserve le produit scalaire.

Proposition 4.13. *Un endomorphisme* f *est orthogonal si et seulement s'il préserve la norme ; c'est-à-dire si et seulement si* $\|f(u)\| = \|u\|$ *pour tout* $u \in E$*. En particulier,* f *est orthogonal si et seulement si l'image d'une base orthonormale par* f *est une base orthonormale.*

Démonstration. Soit $u \in E$ et $f \in \mathrm{End}(E)$. Si f est orthogonal, alors

$$\|f(u)\| = (f(u)|f(u)) = (u|u) = \|u\|$$

donc f préserve la norme. Réciproquement, supposons que f préserve la norme et soit $(u, v) \in E^2$. Alors

$$(u|v) = \frac{\|u+v\|^2 - \|u\|^2 - \|v\|^2}{2}$$

d'après la proposition 4.2 donc

$$(f(u)|f(v)) = \frac{\|f(u+v)\|^2 - \|f(u)\|^2 - \|f(v)\|^2}{2} = \frac{\|u+v\|^2 - \|u\|^2 - \|v\|^2}{2} = (u|v)$$

par linéarité et orthogonalité de f.

Si f est orthogonal et si (e_i) est une base orthonormale, alors $(f(e_i)|f(e_j)) = (e_i|e_j) = \delta_{ij}$ donc la famille $(f(e_i))$ est une base orthonormale. Réciproquement, si f envoie la base orthonormale (e_i) sur une base orthonormale $(f(e_i))$, considérons

$$u = \sum_{i=1}^{n} \alpha_i e_i \in E.$$

Alors $\|u\|^2 = \sum_{i=1}^{n} \alpha_i^2$ par le calcul de l'équation (4.1.2). Par ailleurs

$$f(u) = \sum_{i=1}^{n} \alpha_i f(e_i)$$

par linéarité de f et $(f(e_i))$ est une base orthonormale donc $\|f(u)\|^2 = \sum_{i=1}^{n} \alpha_i^2$ à nouveau par l'équation (4.1.2). Donc $\|f(u)\| = \|u\|$ et f est donc orthogonal d'après la première partie de la preuve. □

4.2.2 Matrice adjointe, matrice orthogonale

Proposition 4.14. *La matrice de l'endomorphisme adjoint f^* d'un endomorphisme f d'un espace euclidien E relativement à une base orthonormale est la transposée de la matrice de l'endomorphisme f relativement à cette base.*

Démonstration. Soit $(b_i)_{1 \leq i \leq n}$ une base orthonormale de l'espace euclidien E, soit $f \in \mathrm{End}(E)$ un endomorphisme de E et soit $M = (a_{ij})$ la matrice de f relativement à la base (b_i). Notons $M^* = (a_{ij}^*)$ la matrice de l'adjoint f^* de f relativement à la base (b_i). Par définition, on a alors

$$f(b_j) = \sum_{i=1}^{n} a_{ij} b_i$$

et donc $a_{ij} = (f(b_j)|b_i)$. De même

$$f^*(b_j) = \sum_{i=1}^{n} a_{ij}^* b_i$$

et $a_{ij}^* = (f^*(b_j)|b_i)$. Par définition de l'endomorphisme adjoint de f, on a également $(f(b_j)|b_i) = (b_j|f^*(b_i)) = (f^*(b_i)|b_j)$ donc $a_{ij} = a_{ji}^*$ et $M^* = {}^tM$. nous avons montré la proposition suivante. □

Corollaire 4.15. *Un endomorphisme f de matrice $M = (a_{ij})_{1 \leq i,j \leq n}$ dans une base orthonormale est auto-adjoint si et seulement si M est égale à sa transposée, ou encore si et seulement si $a_{ij} = a_{ji}$ pour tout $1 \leq i,j \leq n$. Il est orthogonal si et seulement s'il est inversible et ${}^tM = M^{-1}$ ou encore si ${}^tMM = \mathrm{Id}$.*

4.3 Réduction des endomorphismes auto-adjoints

4.3.1 Valeur propre

Soit E un espace euclidien et f un endomorphisme de E. L'espace vectoriel E étant un $\mathbb{R}$-espace vectoriel, la réduction de Jordan ne s'applique pas à f et nous ignorons donc si f admet une valeur propre (et de fait, il peut très bien ne pas en admettre). La proposition suivante assure que les endomorphismes auto-adjoints admettent une valeur propre.

Proposition 4.16. *Soit $f \in \operatorname{End}(E)$ un endomorphisme auto-adjoint d'un espace euclidien E. Alors f admet un vecteur propre v pour la valeur propre $\lambda \in \mathbb{R}$.*

Démonstration. Nous savons (par exemple par le théorème de Cayley-Hamilton) qu'il existe un polynôme unitaire $P \in \mathbb{R}_n[X]$ qui annule f. D'après la proposition 2.4, P admet une factorisation

$$P = \prod_{i=1}^{m}(X - \lambda_i)\prod_{i=1}^{p}(X^2 + \alpha_i X + \beta_i)$$

avec $\alpha_i^2 - 4\beta_i < 0$ pour tout i. Donc l'endomorphisme

$$P(f) = \prod_{i=1}^{m}(f - \lambda_i \operatorname{Id})\prod_{i=1}^{p}(f^2 + \alpha_i f + \beta_i \operatorname{Id})$$

est l'endomorphisme nul de E, donc n'est en particulier pas injectif. Nous voulons montrer que l'un des endomorphismes $f - \lambda_i \operatorname{Id}$ n'est pas injectif. Il suffit donc de démontrer que les endomorphismes $f^2 + \alpha_i f + \beta_i \operatorname{Id}$ sont injectifs pour tout $1 \leq i \leq p$. Soit donc $u \in E$ un vecteur non-nul. Alors

$$\begin{aligned}((f^2 + \alpha_i f + \beta_i \operatorname{Id})(u)|u) &= (f^2(u)|u) + \alpha_i(f(u)|u) + \beta_i(u|u)\\ &= \|f(u)\|^2 + \alpha_i(f(u)|u) + \beta_i\|u\|^2\end{aligned}$$

D'après l'inégalité (4.1.1) de Cauchy-Schwarz

$$-(f(u)|u) \leq \|f(u)\|\|u\|$$

donc

$$|\alpha_i|(f(u)|u) \geq -|\alpha_i|\|f(u)\|\|u\|$$

donc

$$\begin{aligned}((f^2 + \alpha_i f + \beta_i \operatorname{Id})(u)|u) &\geq \|f(u)\|^2 - |\alpha_i|\|f(u)\|\|u\| + \beta_i\|u\|^2\\ &= \left(\|f(u)\| - \frac{|\alpha_i|}{2}\|u\|\right)^2 + \left(\beta_i - \frac{\alpha_i^2}{4}\right)\|u\|^2\end{aligned}$$

Donc $((f^2 + \alpha_i f + \beta_i \operatorname{Id})(u)|u)$ est la somme d'un carré et du produit du réel strictement positif $\|u\|^2$ et du réel strictement positif $\beta_i - \frac{\alpha_i^2}{4}$. Donc $((f^2 + \alpha_i f + \beta_i \operatorname{Id})(u)|u)$ est strictement positif et $(f^2 + \alpha_i f + \beta_i \operatorname{Id})(u)$ est donc non-nul. $\square$

Proposition 4.17. *Soit u et v deux vecteurs propres d'un endomorphisme auto-adjoint f pour des valeurs propres distinctes. Alors $(u|v) = 0$. Le noyau et l'image de f sont orthogonaux*

Démonstration. En effet $(f(u)|v) = \lambda(u|v) = (u|f(v)) = \mu(u|v)$ donc $(\lambda - \mu)(u|v) = 0$ donc $(u|v) = 0$. Soit maintenant u dans le noyau et v dans l'image de f. Il existe alors $w \in E$ tel que $v = f(w)$. Donc $(u|v) = (u|f(w)) = (f(u)|w) = (0|w) = 0$. $\square$

4.3.2 Réduction des endomorphismes auto-adjoints

Nous donnons deux démonstrations légèrement différentes du théorème fondamental suivant.

Théorème 4.4. *Soit E un espace euclidien et f un endomorphisme auto-adjoint de E. Alors il existe une base orthonormale de E formée de vecteurs propres de f. De manière équivalente, f est diagonalisable dans une base orthonormale.*

Démonstration. Par le théorème de réduction de Jordan, un endomorphisme est diagonalisable si et seulement ses espaces propres sont égaux à ses espaces propres généralisés. Soit E_λ un espace propre généralisé de f. Alors E_λ est un espace vectoriel euclidien et $f - \lambda \operatorname{Id}$ est un endomorphisme auto-adjoint nilpotent de $f - \lambda \operatorname{Id}$. Il suffit donc de démontrer qu'un endomorphisme auto-adjoint nilpotent d'un espace vectoriel euclidien est l'endomorphisme nul. Soit donc f un endomorphisme auto-adjoint nilpotent d'un espace vectoriel euclidien F et soit $u \in F$. Il existe alors un (unique) $n \in \mathbb{N}$ tel que $f^n(u) = 0$ et $f^{n-1}(u) \neq 0$. Supposons $n > 0$. Le vecteur $v = f^{n-1}(u)$ est alors à la fois dans le noyau de f et dans l'image de f. D'après la proposition 4.17, v vérifie donc $(v|v) = 0$ donc est nul, ce qui contredit la définition de n. Donc $n = 0$ donc $f(u) = 0$. Donc f est l'endomorphisme nul.

Donc E admet une décomposition $E = \bigoplus_\lambda \ker(f - \lambda \operatorname{Id})$. Chacun des espaces $\ker(f - \lambda \operatorname{Id})$ admet une base orthonormale d'après la proposition 4.6. L'union de ces bases est une base de E orthonormale de vecteurs propres pour f. $\square$

Démonstration. Montrons que E admet une base orthonormale de vecteurs propres pour f par récurrence sur la dimension de E. Soit E est de dimension 1 et n'importe quelle base convient. Maintenant on suppose l'assertion vraie pour tous les endomorphismes auto-adjoints d'espaces euclidiens de dimension au plus n et on suppose que E est de dimension $n + 1$. D'après la proposition 4.16, f admet un vecteur propre v que l'on peut normaliser afin qu'il soit de norme 1. Soit $u \in \operatorname{Vect}(v)^\perp$. Alors $(f(u)|v) = (u|f(v)) = (u|\lambda v) = \lambda(u|v) = 0$ donc $\operatorname{Vect}(v)^\perp$ est stable par f. Donc E admet une décomposition en sous-espaces stables

$$E = \operatorname{Vect}(v)^\perp \oplus \operatorname{Vect}(v).$$

La restriction de f à $\operatorname{Vect}(v)^\perp$ est diagonalisable en base orthonormale par hypothèse de récurrence, donc f est diagonalisable et la base obtenue en ajoutant v à la base de $\operatorname{Vect}(v)^\perp$ est bien orthonormale. $\square$

5 Réduction de quelques endomorphismes importants

5.1 Projecteur

Dans cette sous-section K désigne $\mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$ et E désigne un K-espace vectoriel de dimension finie .

5.1.1 Définition et réduction

Définition 5.1. *Un endomorphisme p de E est un projecteur s'il vérifie $p^2 = p$.*

On dit également que p est idempotent, ou un idempotent.

Proposition 5.2. *Soit p un projecteur. Il existe deux sous-espaces vectoriels F et G tels que $E = F \oplus G$, F est l'image de p, G est le noyau de p et si $u \in E$ s'écrit $u = v + w$ avec $v \in F$ et $w \in G$ alors $p(u) = v$.*

Démonstration. Le polynôme $X^2 - X$ s'écrit $X^2 - X = X(X-1)$ dans $K[X]$ et annule p donc l'endomorphisme p est diagonalisable et son spectre est inclus dans $\{0, 1\}$ d'après la proposition 2.18. Posons $F = \ker(p - \mathrm{Id})$ et $G = \ker p$. Alors E s'écrit $E = F \oplus G$. L'endomorphisme p étant diagonalisable, G est également le nilespace de p et la décomposition $E = F \oplus G$ est aussi la décomposition $E = C(p) \oplus N(p)$ donc F est l'image de p (et même son coeur). La dernière assertion découle des précédentes en calculant $p(u)$. $\square$

On dit également que p est la projection sur F parallèlement à G.

5.1.2 Projection orthogonale

Lorsque E est un $\mathbb{R}$-espace vectoriel euclidien et lorsque p est de plus auto-adjoint, alors p est une projection orthogonale. Ceci est équivalent à $F = G^\perp$ et à $F \subset G^\perp$.

5.2 Symétrie

5.2.1 Définition et réduction

Définition 5.3. *Un endomorphisme s de E est une symétrie s'il vérifie $s^2 = \mathrm{Id}$.*

Proposition 5.4. *Soit s une symétrie. Il existe deux sous-espaces vectoriels F et G tels que $E = F \oplus G$, F est l'espace propre de s pour la valeur propre 1, G est l'espace propre pour la valeur propre -1. La projection sur F parallèlement à G est égale à $(\mathrm{Id} + s)/2$ et $s = 2p - \mathrm{Id}$.*

Démonstration. Le polynôme $X^2 - 1$ s'écrit $X^2 - 1 = (X-1)(X+1)$ dans $K[X]$ et annule s donc l'endomorphisme s est diagonalisable et son spectre est inclus dans $\{-1, 1\}$ d'après la proposition 2.18. Donc E s'écrit $E = E_1 \oplus E_{-1}$ avec $E_{\pm 1}$ l'espace propre de s pour la valeur propre ± 1. Soit p l'endomorphisme $(\mathrm{Id} + s)/2$. Alors $p^2 = p$ donc p est un projecteur, $s = 2p - \mathrm{Id}$, E_1 est l'image de p et E_{-1} est le noyau de p. □

On dit également que s est la symétrie par rapport à G parallèlement à F.

5.2.2 Symétrie orthogonale

Supposons de plus que E est un $\mathbb{R}$-espace vectoriel euclidien. Soit s une symétrie de E telle que $F = G^\perp$ (ou de manière équivalente telle que $F \subset G^\perp$). Alors s est auto-adjoint. Soit $(u, v) \in E^2$. Alors $(s(u)|s(v)) = (u|s^2(v)) = (u|v)$. Donc s est également orthogonal. Réciproquement, si s est un endomorphisme auto-adjoint et symétrique alors $s^* = s$ et $s^* = s^{-1}$ donc $s = s^{-1}$ donc $s^2 = \mathrm{Id}$ et s est une symétrie auto-adjointe. Les symétries orthogonales, c'est-à-dire les symétries par rapport à G parallèlement à $G^\perp$, sont donc les endomorphismes orthogonaux auto-adjoints.

5.3 Rotation

Dans cette sous-section, E est l'espace euclidien $\mathbb{R}^2$ muni du produit scalaire canonique.

5.3.1 Définition et réduction

Définition 5.5. *Un endomorphisme r de $\mathbb{R}^2$ est une rotation s'il est orthogonal et de déterminant 1.*

Proposition 5.6. *Soit r une rotation. Il existe $\theta \in [0, 2\pi]$ telle que la matrice de r dans dans la base canonique de $\mathbb{R}^2$ soit*

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Démonstration. Soit en effet $u = r(e_1) = \alpha e_1 + \beta e_2$. Alors $\|u\| = 1$ car r est orthogonal donc $\alpha^2 + \beta^2 = 1$ donc α appartient à $[-1, 1]$ donc il existe $\theta \in [0, 2\pi]$ tel que $\alpha = \cos\theta$. De $\beta^2 = 1 - \alpha^2$, on déduit que $\sin\theta = \beta$ ou $\sin\theta = -\beta$. En remplaçant si nécessaire θ par $-\theta$ (ce qui ne change pas $\cos\theta$), on peut assurer que $\sin\theta = \beta$. Si $r(e_2) = \gamma e_1 + \delta e_2$, alors $(\gamma, \delta) \in [-1, 1]^2$ vérifient

$$\begin{cases} \alpha\delta - \beta\gamma = 1 \\ \gamma^2 + \delta^2 = 1 \\ \alpha\gamma + \beta\delta = 0. \end{cases}$$

La seule solution de ces système est $(\gamma, \delta) = (-\sin\theta, \cos\theta)$. □

5.4 Endomorphisme d'ordre fini

Dans cette sous-section, K est égal à $\mathbb{C}$ et E est donc un $\mathbb{C}$-espace vectoriel de dimension finie.

5.4.1 Définition et réduction

Définition 5.7. *Un endomorphisme f de E est d'ordre fini s'il existe $n > 0$ tel que $f^n = \mathrm{Id}$.*

Proposition 5.8. *Soit f un endomorphisme f d'ordre fini. Alors f est diagonalisable et son spectre est inclus dans $\{\zeta \in \mathbb{C} | \zeta^n = 1\}$.*

Démonstration. Soit $\lambda \in \mathbb{C}$ une racine du polynôme $X^n - 1$. Alors $\lambda \neq 0$ donc $n\lambda^{n-1} \neq 0$ donc λ n'est pas une racine du polynôme nX^{n-1} donc λ n'est pas une racine double du polynôme $X^n - 1$. Le polynôme $X^n - 1$ est donc à racines simples et annule f. Donc l'endomorphisme f est diagonalisable et son spectre est inclus dans $\{\zeta \in \mathbb{C} | \zeta^n = 1\}$ d'après la proposition 2.18. $\square$

5.4.2 Permutation

Un endomorphisme f de E est un endomorphisme de permutation s'il existe une base $(e_i)_{1 \leq i \leq n}$ et une bijection σ de $\{1, \cdots, n\}$ dans lui-même telle que la matrice de f dans la base (e_i) soit la matrice $(\delta_{i\sigma(i)})_{1 \leq i \leq n}$; ou de manière équivalente si la matrice de f dans la base (e_i) a tous ses coefficients nuls sauf exactement un coefficient par ligne et par colonne, qui est égal à 1.

Proposition 5.9. *Un endomorphisme de permutation est diagonalisable.*

Démonstration. Soit f un endomorphisme de permutation. D'après la proposition 5.8, il suffit de montrer qu'il existe $n \in \mathbb{N}$ non-nul tel que $f^n = \mathrm{Id}$. Soit $1 \leq i \leq n$. L'ensemble $\{f^n(e_i) | n \in \mathbb{N}\}$ est inclus dans l'ensemble $\{e_j | 1 \leq j \leq n\}$ donc il existe deux entiers distincts a_i et b_i tels que $f^{a_i}(e_i) = f^{b_i}(e_i)$. Il existe donc un entier $n_i \in \mathbb{N}$ non-nul (par exemple $a_i - b_i$ ou son inverse) tel que $f^{n_i}(e_i) = e_i$. Soit n le produit sur i de tous les n_i. Alors $f^n(e_j) = e_j$ pour tout $1 \leq j \leq n$ donc $f^n = \mathrm{Id}$. $\square$