Rudiments pratiques d'algèbre linéaire

IUP GSI 1ère année et DEUG STPI 2ème année

François Dumas

Table des matières

Introduction

Ce document n'est pas un cours. C'est un outil de travail personnel destiné à aider les étudiantes et étudiants suivant l'enseignement d'algèbre linéaire commun au DEUG STPI deuxième année et à la première année d'IUP GSI.

Le cœur du programme de cet enseignement est la réduction des matrices carrées et ses applications, en particulier à la résolution des systèmes différentiels linéaires à coefficients constants. Aborder ces questions nécessite d'avoir assimilé un minimum de connaissances sur les notions de base de l'algèbre linéaire : espaces vectoriels de dimension finie, bases, applications linéaires, rang, déterminant.

Or les étudiantes et étudiants admis à ce cours en début d'année viennent d'horizons fort différents, ont reçu des formations diverses en mathématiques en général et en algèbre linéaire en particulier, et présentent donc des niveaux très hétérogènes au départ dans cette discipline.

Il est alors utile de consacrer les premières semaines du semestre à une “remise à niveau”, qui est pour certain(e)s une révision, pour d'autres un premier apprentissage, pour tous l'occasion de fixer dans un cadre bien défini et homogène les connaissances préalables nécessaires.

Le temps imparti à cette phase préliminaire étant bien bref au vu de l'étendue du champ à couvrir, des notes écrites ont pu sembler utiles. D'où l'existence de ce polycopié, et ses principales caractéristiques.

(i) Ce n'est pas un cours. Il ne contient pas tout ce qui est relatif aux sujets abordés. Il ne contient pas les démonstrations des résultats (sauf dans de très rares cas où connaître le principe de preuve est indispensable ou particulièrement éclairant). Il se limite aux concepts les plus fondamentaux, aux outils absolument nécessaires, et ne contient rien qui ne soit utile pour la suite. Il faut le voir comme un *vademecum*, un bagage minimal pour aborder les points du programme.

(ii) C'est un outil de travail personnel, d'où de nombreux commentaires et explications parallèles, pour aider celle ou celui qui découvre seul(e) ces nouvelles notions à en comprendre mieux les motivations et les principes. D'où la présence aussi d'exemples très simples et d'exercices d'application immédiate pour en vérifier la bonne assimilation.

(iii) C'est un document adapté à un type d'enseignement, de public et de volume horaire bien spécifiques, d'où certains partis pris dans la présentation et le choix de privilégier certains points de vue, l'insistance sur les savoir-faire et les applications.

Enfin bien sûr, ce n'est pas un document sous forme définitive. Il est susceptible d'évoluer avec le public auquel il est destiné et suivant les réactions qu'il recueille. Il contient certainement des coquilles, des erreurs, des points à améliorer ou modifier, et je serais heureux qu'on me les signale.

Je tiens déjà à remercier les étudiantes et étudiants de l'an passé pour leurs observations, Cyrille Ospel pour ses remarques et suggestions issues de son expérience lors des séances de TD qu'il a assurées ces dernières années, et François Martin qui a accepté d'en relire avec soin et pertinence la nouvelle version.

François Dumas (juillet 2003)

Chapitre 1

Espaces vectoriels $\mathbb{R}^n$ et $\mathbb{C}^n$

Pour une question de commodité, dans toute la suite, on désignera par la lettre $\mathbb{K}$ soit le corps $\mathbb{R}$ des nombres réels, soit le corps $\mathbb{C}$ des nombres complexes.

1.1 Structure d'espace vectoriel sur $\mathbb{K}^n$

1.1.1 Ensemble $\mathbb{K}^n$

Pour tout nombre entier $n \geq 1$, on note $\mathbb{K}^n$ l'ensemble des n-uplets de nombres de $\mathbb{K}$. Un tel n-uplet se note avec des parenthèses :

$$u = (x_1, x_2, \ldots, x_n) \;\text{ avec } x_i \in \mathbb{K} \text{ pour tout } 1 \leq i \leq n.$$

• Lorsque $n = 2$, on dit que $u = (x_1, x_2)$ est un couple ; on préfère alors souvent ne pas utiliser d'indice et noter $u = (x, y)$.
• Lorsque $n = 3$, on dit que $u = (x_1, x_2, x_3)$ est un triplet, et l'on peut noter aussi $u = (x, y, z)$.
• Remarquons que, pour $n = 1$, on a $\mathbb{K}^1 = \mathbb{K}$, et l'on supprime les parenthèses pour écrire $u = x_1$.
• Insistons bien sur le fait que la notion de n-uplet fait intervenir l'ordre dans lequel sont pris les nombres x_i, c'est-à-dire que :

$$[(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_n) \text{ dans } \mathbb{K}^n] \Leftrightarrow [x_i = y_i \text{ dans } \mathbb{K} \text{ pour tout } 1 \leq i \leq n].$$

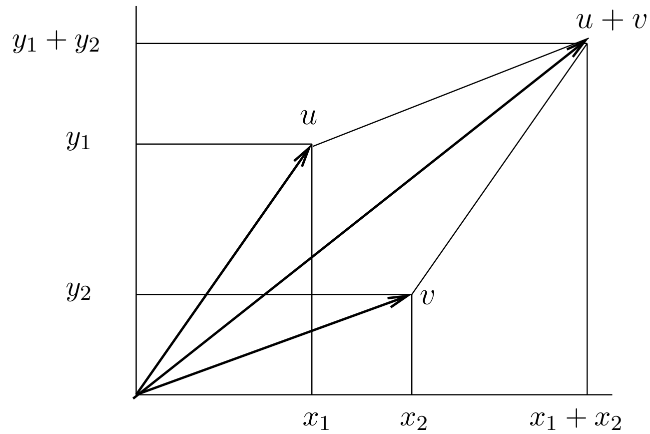
Par exemple $(x_1, x_2) \neq (x_2, x_1)$ dans $\mathbb{K}^2$ dès lors que $x_1 \neq x_2$ dans $\mathbb{K}$.
• Dans toute la suite, les éléments de $\mathbb{K}^n$, c'est-à-dire les n-uplets de nombres de $\mathbb{K}$, seront appelés des vecteurs. Cette terminologie (qui trouvera sa justification algébrique dans la suite) est bien sûr conforme à l'intuition que chacun a déjà en physique ou en géométrie, puisqu'on a l'habitude de repérer un vecteur du plan par un couple de nombres réels (cas $n = 2$) et un vecteur de l'espace par un triplet de nombres réels (cas $n = 3$).

1.1.2 Addition dans $\mathbb{K}^n$

a) Définition.
Fixons $n \geq 1$. On a une notion naturelle d'*addition* dans $\mathbb{K}^n$ définie à partir de l'addition des nombres dans $\mathbb{K}$, en additionnant terme à terme, c'est-à-dire :

$$\begin{aligned} &\text{si } u = (x_1, x_2, \ldots, x_n) \text{ et } v = (y_1, y_2, \ldots, y_n), \\ &\text{alors } \; u + v = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n). \end{aligned}$$

• Par exemple dans $\mathbb{R}^3$: $(1, -2, 0) + (-1, 4, 3) = (0, 2, 3)$.
• Remarquons que, pour $\mathbb{K} = \mathbb{R}$, cette addition correspond, via le passage aux coordonnées par rapport à une base du plan ($n = 2$) ou de l'espace ($n = 3$), à l'addition "géométrique" des vecteurs par la règle du parallélogramme :



b) Propriétés.
L'addition dans $\mathbb{K}^n$ hérite directement de certaines propriétés connues de l'addition des nombres dans $\mathbb{K}$.

• Par exemple, la somme de deux vecteurs quelconques vérifie

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (y_1, y_2, \dots, y_n) + (x_1, x_2, \dots, x_n)$$

puisque l'on a dans $\mathbb{K}$ l'égalité $x_i + y_i = y_i + x_i$ pour chaque $1 \leq i \leq n$. On traduit cette propriété en disant que l'addition dans $\mathbb{K}^n$ est *commutative.*

• De même elle est *associative*, [ce qui signifie que : $(u + v) + w = u + (v + w)$ pour tous u, v, w dans $\mathbb{K}^n$], simplement parce que l'addition dans $\mathbb{K}$ est elle-même associative.

• Le n-uplet $(0, 0, \dots, 0)$ est appelé le *vecteur nul* de $\mathbb{K}^n$. Il joue le rôle particulier d'élément *neutre* pour l'addition dans $\mathbb{K}^n$, ce qui signifie que l'on a pour tout vecteur :

$$(x_1, x_2, \dots, x_n) + (0, 0, \dots, 0) = (0, 0, \dots, 0) + (x_1, x_2, \dots, x_n) = (x_1, x_2, \dots, x_n).$$

Dans la suite, on désignera souvent par la lettre E l'ensemble $\mathbb{K}^n$, et on notera alors $0_E = (0, 0, \dots, 0)$.

• Enfin pour tout vecteur $u = (x_1, x_2, \dots, x_n)$, on peut considérer le vecteur *opposé* $-u$, défini par $-u = (-x_1, -x_2, \dots, -x_n)$, qui vérifie que $u + (-u) = (-u) + u = (0, 0, \dots, 0)$.

c) En résumé.
Notons pour simplifier $E = \mathbb{K}^n$. L'addition dans E vérifie les quatre propriétés suivantes :

(1) elle est associative : $\forall\, u, v, w \in E,\ u + (v + w) = (u + v) + w$;

(2) il existe $0_E \in E$, appelé le *vecteur nul* de E, tel que : $\forall\, u \in E,\ u + 0_E = 0_E + u = u$;

(3) pour tout $u \in E$, il existe un unique vecteur $-u \in E$, appelé l'*opposé* de u dans E, vérifiant $u + (-u) = (-u) + u = 0_E$;

(4) elle est commutative : $\forall\, u, v \in E,\ u + v = v + u$.

1.1.3 Produit externe d'un vecteur de $\mathbb{K}^n$ par un nombre de $\mathbb{K}$

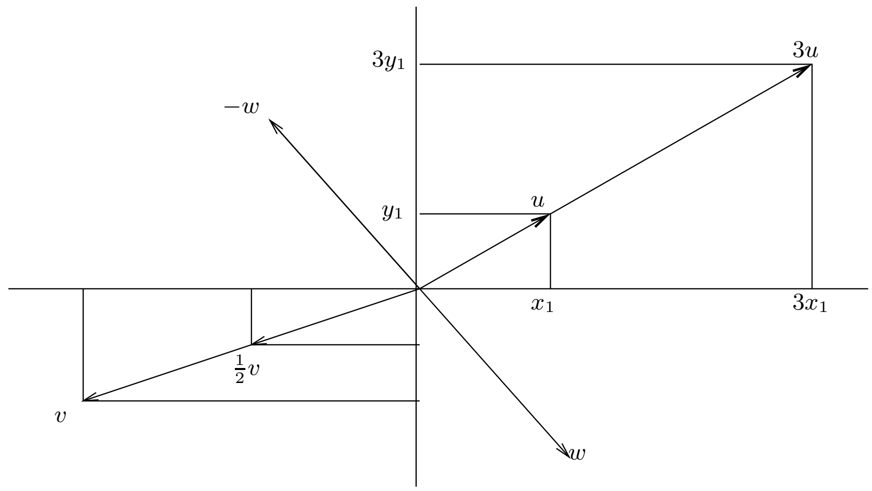
a) Définition.
On peut à partir de la multiplication des nombres dans $\mathbb{K}$ définir une autre opération sur $\mathbb{K}^n$, le produit d'un vecteur de $\mathbb{K}^n$ par un nombre de $\mathbb{K}$. Il s'agit donc d'un *produit externe*, sur $\mathbb{K}^n$ qui est défini de la façon suivante :

$$\text{si } u = (x_1, x_2, \dots, x_n) \in \mathbb{K}^n \text{ et } \alpha \in \mathbb{K}, \text{ on pose } \alpha.u = (\alpha x_1, \alpha x_2, \dots, \alpha x_n) \in \mathbb{K}^n.$$

Dans cette écriture, chaque αx_i est le produit (au sens de la multiplication ordinaire dans $\mathbb{K}$) des deux nombres α et x_i.

• Par exemple dans $\mathbb{R}^3$: $2.(-1, 4, 3) = (-2, 8, 6)$ et $-\frac{1}{3}(6, 0, -5) = (-2, 0, \frac{5}{3})$.

• Remarquons que, pour $\mathbb{K} = \mathbb{R}$, ce produit externe correspond, via le passage aux coordonnées par rapport à une base du plan ($n = 2$) ou de l'espace ($n = 3$), à l'opération "géométrique" usuelle du produit d'un vecteur u par un réel α (même direction que u, même sens ou sens opposé suivant le signe de α, norme multipliée par $|\alpha|$).



b) Propriétés.
Elles se déduisent immédiatement de celles de la multiplication (interne) dans $\mathbb{K}$, par exemple :
• si α et β sont deux nombres de $\mathbb{K}$ et $u = (x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$, on a :
$(\alpha + \beta).u = \big((\alpha + \beta)x_1, (\alpha + \beta)x_2, \ldots, (\alpha + \beta)x_n\big) = \big(\alpha x_1 + \beta x_1, \alpha x_2 + \beta x_2, \ldots, \alpha x_n + \beta x_n\big)$
$= \alpha.(x_1, x_2, \ldots, x_n) + \beta.(x_1, x_2, \ldots, x_n) = \alpha.u + \beta.u,$
$\alpha.(\beta.u) = \alpha.(\beta x_1, \beta x_2, \ldots, \beta x_n) = \big(\alpha(\beta x_1), \alpha(\beta x_2), \ldots, \alpha(\beta x_n)\big)$
$= \big(\alpha\beta x_1, \alpha\beta x_2, \ldots, \alpha\beta x_n\big) = (\alpha\beta).u.$
• D'autres propriétés s'obtiennent de la même façon, entre autres :
$1.u = u$ pour tout $u \in \mathbb{K}^n$,
$\alpha.(u + v) = \alpha.u + \alpha.v$ pour tout $\alpha \in \mathbb{K}$ et tous $u, v \in \mathbb{K}^n$.

c) En résumé.
Notons pour simplifier $E = \mathbb{K}^n$. Le produit externe d'un vecteur de E par un nombre de $\mathbb{K}$ vérifie les quatre propriétés suivantes :

(5) $\forall\ u \in E,\ \forall\ \alpha, \beta \in \mathbb{K},\ (\alpha + \beta).u = \alpha.u + \beta.u\,;$

(6) $\forall\ u, v \in E,\ \forall\ \alpha \in \mathbb{K},\ \alpha.(u + v) = \alpha.u + \alpha.v\,;$

(7) $\forall\ u \in E,\ \forall\ \alpha, \beta \in \mathbb{K},\ \alpha.(\beta.u) = (\alpha\beta).u\,;$

(8) $\forall\ u \in E,\ 1.u = u.$

1.1.4 Calcul dans l'espace vectoriel $\mathbb{K}^n$

a) Définition.
On résume les propriétés (1) à (4) de l'addition et les propriétés (5) à (8) du produit externe en disant que l'ensemble $E = \mathbb{K}^n$ muni de ces deux opérations est un *espace vectoriel* sur $\mathbb{K}$.
Dans cette terminologie algébrique, les éléments de E sont appelés (comme on l'a déja dit) des *vecteurs*. Les nombres de $\mathbb{K}$ sont appelés des *scalaires*.

b) D'autres règles de calcul dans l'espace vectoriel $\mathbb{K}^n$.
Il résulte immédiatement des propriétés (1) à (8) définissant la structure d'espace vectoriel de $E = \mathbb{K}^n$ que l'on a aussi dans E les règles de calcul ci-dessous :

$\forall\, u \in E,\ \forall\, \alpha \in \mathbb{K},\ (-\alpha).u = \alpha.(-u) = -\alpha.u\,;$

en particulier : $\forall\, u \in E,\ (-1).u = -u\,;$

$\forall\, u, v \in E,\ \forall\, \alpha, \beta \in \mathbb{K},\ \alpha.(u - v) = \alpha.u - \alpha.v\,;$

$\forall\, u, v \in E,\ \forall\, \alpha, \beta \in \mathbb{K},\ (\alpha - \beta).u = \alpha.u - \beta.u\,;$

ainsi que la propriété suivante, qui jouera un rôle important dans certains des raisonnements à venir :

$(\alpha.u = 0_E)$ si et seulement si $(\alpha = 0$ dans $\mathbb{K}$ ou $u = 0_E$ dans $E)$.

c) Combinaisons linéaires de deux vecteurs de $\mathbb{K}^n$.

• Les calculs que l'on aura à faire dans $\mathbb{K}^n$ combinent naturellement les deux opérations + et . de $\mathbb{K}^n$. Si u et v sont deux vecteurs de $\mathbb{K}^n$, on appelle combinaison linéaire de u et v tout vecteur de la forme $\alpha.u + \beta.v$, où α et β sont des scalaires.

• Par exemple pour $n = 3$, à partir des vecteurs $u = (1, -2, 3)$ et $v = (0, 5, -1)$, on peut fabriquer la combinaison linéaire $w = 2.u + 3.v = (2, -4, 6) + (0, 15, -3) = (2, 11, 3)$.

• Pour calculer plus commodément, on dispose parfois les vecteurs en colonnes. Par exemple pour $n = 4$:

$$\begin{pmatrix} 2 \\ -5 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ -4 \\ -2 \\ 1 \end{pmatrix}, \qquad 3.\begin{pmatrix} 1 \\ 0 \\ -2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ -6 \\ 9 \end{pmatrix}, \qquad -\begin{pmatrix} 1 \\ 8 \\ -1 \\ 0 \end{pmatrix} + \tfrac{1}{2}.\begin{pmatrix} 4 \\ 0 \\ -2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ -8 \\ 0 \\ \frac{3}{2} \end{pmatrix}.$$

1.1.5 Commentaire : notion générale d'espace vectoriel

• D'une façon générale, on appelle espace vectoriel sur $\mathbb{K}$, ou encore $\mathbb{K}$-espace vectoriel, tout ensemble non-vide E muni d'une addition et d'un produit externe satisfaisant les conditions (1) à (8). Il existe d'innombrables exemples de situations faisant intervenir la notion d'espace vectoriel (espaces de fonctions ou de suites en analyse, espaces de polynômes en algèbre, espaces des vecteurs géométrique du plan ou de l'espace en mécanique ou en physique en général,...).

• Les espaces $\mathbb{K}^n$ ne sont donc qu'un exemple d'espaces vectoriels. Mais, plus qu'un exemple, ils constituent un prototype au sens que, pour toute une classe très vaste d'espaces vectoriels (ceux que l'on appelle "de dimension finie") on peut toujours se ramener (en un sens précis que l'on verra plus loin) au cas de $\mathbb{K}^n$.

• En clair, les espaces vectoriels $\mathbb{K}^n$ sont les objets fondamentaux de tout notre travail à venir. Toutes les notions, définitions et théorèmes que l'on va formuler dans ce cours le seront pour les espaces $\mathbb{K}^n$, mais on pourra garder à l'esprit que les mêmes résultats sont toujours valables dans le contexte plus abstrait des espaces vectoriels de dimension finie, (et même pour certains dans le contexte des espaces vectoriels quelconques).

• On utilisera dans la suite l'abréviation $\mathbb{K}$-e.v. pour $\mathbb{K}$-espace vectoriel.

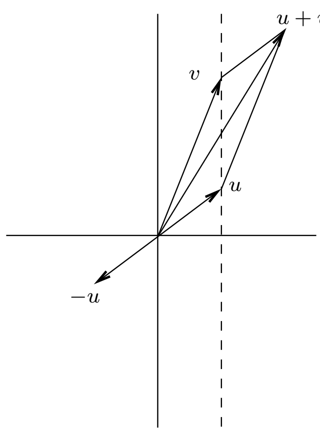
1.2 Sous-espaces vectoriels de $\mathbb{K}^n$

1.2.1 Exemples préliminaires

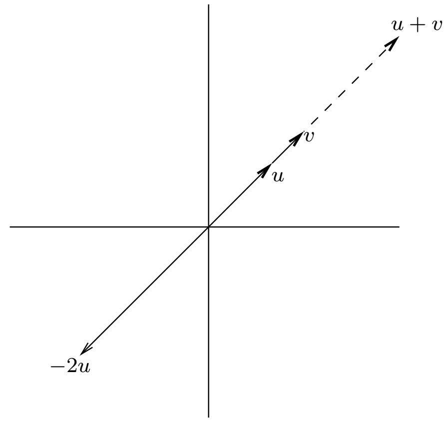
Plaçons-nous dans le $\mathbb{R}$-e.v. $E = \mathbb{R}^2$.

Considérons d'abord le sous-ensemble (ou la partie) F de E constituée des vecteurs (x, y) tels que $x = y$. Si $u = (x, x)$ et $v = (x', x')$ sont deux vecteurs quelconque de F, leur somme $u + v$ reste dans F, puisque $u + v = (x + x', x + x')$. On traduit cette propriété en disant que F est stable pour

l'addition de E. De même, quels que soient un vecteur $u = (x, x)$ de F et un scalaire $\alpha \in \mathbb{R}$, le vecteur $\alpha.u = (\alpha x, \alpha x)$ reste dans F. On traduit cette propriété en disant que F est stable aussi pour le produit externe.
Considérons maintenant le sous-ensemble G de E constituée des vecteurs (x, y) tels que $x = 1$. Si $u = (1, y)$ et $v = (1, y')$ sont deux vecteurs quelconques de G, leur somme $u + v$ est le vecteur $u + v = (2, y + y')$, qui n'est plus un vecteur de G. Donc G n'est pas stable pour l'addition de E. De même, le produit externe d'un vecteur $u = (1, y)$ par un scalaire $\alpha \in \mathbb{R}$ est $\alpha.u = (\alpha, \alpha y)$ qui n'est plus un élément de G dès lors que le réel α est distinct de 1. Donc G n'est pas non plus stable pour le produit externe.



L'ensemble des vecteurs (x, y) de $\mathbb{R}^2$ tels que $x = 1$ n'est stable ni par addition, ni par le produit externe



L'ensemble des vecteurs (x, y) de $\mathbb{R}^2$ tels que $x = y$ est stable par addition, et par le produit externe

A noter qu'un sous-ensemble de E peut être stable pour l'addition mais pas pour le produit externe ; c'est la cas par exemple du sous-ensemble des vecteurs (x, y) de E tels que $x \geq 0$.
L'objet de cette section est d'étudier, dans la cadre général où $E = \mathbb{K}^n$, les sous-ensembles de E qui respectent la structure de $\mathbb{K}$-e.v. de E, c'est-à-dire qui sont stables à la fois pour l'addition et le produit externe de E. On les appelle des sous-espaces vectoriels de E.

Dans toute la suite de cette section, on fixe $n \geq 1$ et on désigne par E le $\mathbb{K}$-e.v. $\mathbb{K}^n$.

1.2.2 Notion de sous-espace vectoriel

a) Définitions.
Soit F un sous-ensemble (une partie) non-vide de E. On dit que

- F est *stable pour l'addition* $+$ lorsque : $\forall\, u, v \in F,\ u + v \in F$.
- F est *stable pour le produit externe* . lorsque : $\forall\, u \in F,\ \forall\, \alpha \in \mathbb{K},\ \alpha.u \in F$.
- F est *stable par combinaison linéaire* lorsque F est stable à la fois pour $+$ et .

ce qui équivaut à : $\forall\, u, v \in F,\ \forall\, \alpha, \beta \in \mathbb{K},\ \alpha.u + \beta.v \in F$,

ou encore : $\forall\, u_1, u_2, \ldots, u_p \in F,\ \forall\, \alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbb{K},\ \alpha_1.u_1 + \alpha_2.u_2 + \cdots + \alpha_p.u_p \in F$.

b) Définition.
On appelle *sous-espace vectoriel* de E toute partie non-vide de E qui est stable à la fois pour l'addition et pour le produit externe, c'est-à-dire qui est stable par combinaison linéaire.

On utilisera dans la suite l'abréviation ss-e.v. pour sous-espace vectoriel.

b) Remarques.

• Soit F un ss-e.v. de E. Puisque F est stable pour l'addition de E, la restriction de cette addition à F définit une addition dans F. De même pour le produit externe. Les propriétés (1), (4), (5), (6), (7), (8) étant vraies pour des vecteurs quelconques de E, elles restent a fortiori vraies pour des vecteurs quelconques de F. Le vecteur nul 0_E appartient forcément à F puisque $0_E = 0.u \in F$ pour tout $u \in F$. De même, pour tout $u \in F$, son opposé $-u$ dans E appartient forcément à F puisqu'il vaut $-u = (-1).u$. En résumé, les restrictions à F de l'addition et du produit externe de E vérifient les 8 propriétés qui font de F un $\mathbb{K}$-e.v.

Ainsi, un ss-e.v. de E n'est autre qu'un sous-ensemble de E qui est lui-même un $\mathbb{K}$-e.v. pour les restrictions à F de l'addition et du produit externe de E.

• Il est clair que $\{0_E\}$ et E lui-même sont toujours des ss-e.v. de E.

• Comment s'y prendre pour vérifier qu'un sous-ensemble F de E est un ss-e.v. :

– d'abord, s'assurer que F est non-vide, et le plus simple pour cela est de vérifier que le vecteur nul 0_E appartient bien à F (a contrario, dès lors qu'un sous-ensemble ne contient pas le vecteur nul, on est sûr que ce n'est pas un ss-e.v. !)

– puis montrer que : $(\forall\ u, v \in F,\ u + v \in F)$ et $(\forall\ u \in F,\ \forall\ \alpha \in \mathbb{K},\ \alpha.u \in F)$; si les calculs sont simples, on peut vérifier en une seule fois que : $(\forall\ u, v \in F,\ \forall\ \alpha, \beta \in \mathbb{K}, \alpha.u + \beta.v \in F)$.

c) Exemples.

Dans $E = \mathbb{R}^3$, les parties $P = \{(x, 0, z) ; x \in \mathbb{R}, z \in \mathbb{R}\}$ et $D = \{(x, 0, x) ; x \in \mathbb{R}\}$ sont des ss-e.v. de E. De plus, D est un ss-e.v. de P.
En revanche, $A = \{(x, 1, x) ; x \in \mathbb{R}\}$ n'est pas un ss-e.v. car la somme de deux vecteurs $(x, 1, x)$ et $(x', 1, x')$ est $(x + x', 2, x + x')$ qui n'appartient pas à A. [On pourrait aussi remarquer que $0_E = (0, 0, 0) \notin A$, ce qui suffit à prouver que A n'est pas un ss-e.v. de E].

1.2.3 Intersection de sous-espaces vectoriels

a) Proposition. *Si F et G sont deux ss-e.v. de E, alors $F \cap G$ est un ss-e.v. de E.*

Preuve : c'est un simple exercice. A faire ! □

b) Exemple. Dans $E = \mathbb{R}^3$, $P = \{(x, 0, z) ; x \in \mathbb{R}, z \in \mathbb{R}\}$, $Q = \{(x, y, 0) ; x \in \mathbb{R}, y \in \mathbb{R}\}$ ont pour intersection $P \cap Q = \{(x, 0, 0) ; x \in \mathbb{R}\}$.

c) Remarques.

• Rappelons que l'intersection $F \cap G$ est l'ensemble des vecteurs de E qui appartiennent à la fois à F et à G.

• De la proposition a) ci-dessus, on déduit immédiatement que l'intersection d'un nombre fini quelconque de ss-e.v. de E est un ss-e.v.

• En revanche, on peut aisément constater que la réunion $F \cup G$, (qui est l'ensemble des vecteurs de E qui appartiennent à F ou à G), n'est en général pas un ss-e.v. de E.
Par exemple, dans l'exemple b) ci-dessus prenons $u = (1, 0, 1)$ et $v = (1, 1, 0)$; on a $u \in P$ donc $u \in P \cup Q$, $v \in Q$ donc $v \in P \cup Q$, et pourtant $u + v = (2, 1, 1)$ n'est ni dans P ni dans Q, d'où $u + v \notin P \cup Q$; ce qui prouve que $P \cup Q$ n'est pas stable pour l'addition.
La bonne notion pour remplacer la réunion dans le cadre des sous-espaces vectoriels est celle de somme, introduite au paragraphe suivant.

1.2.4 Somme de sous-espaces vectoriels

a) Notation. Si F et G sont deux ss-e.v. de E, on note $F+G$ le sous-ensemble de E formé des vecteurs qui se décomposent en la somme d'un vecteur de F et d'un vecteur de G.

Donc : $F+G=\{u\in E\,;\,\exists\ v\in F,\ \exists\ w\in G,\ u=v+w\}=\{v+w\,;\,v\in F,w\in G\}$.

En d'autres termes : $\forall\ u\in E,\ (u\in F+G)\Leftrightarrow(\exists\ v\in F,\ \exists\ w\in G,\ u=v+w)$.

b) Proposition et définition. *Soient F et G deux ss-e.v. de E ; alors :*

(i) *$F+G$ est un ss-e.v. de E, appelé le sous-espace somme de F et de G,*
(ii) *F et G sont des ss-e.v. de $F+G$.*

Preuve : (i) est un simple exercice. Pour (ii), on remarque d'abord que tout $v\in F$ s'écrit $v=v+0_E$ avec $0_E\in G$, ce qui prouve que $v\in F+G$, et permet de conclure que $F\subset F+G$. Le reste de la preuve est laissé en exercice. A faire ! □

c) Remarque.
Il est clair que $F+G$ contient $F\cup G$. On peut montrer plus précisément que $F+G$ est le ss-e.v. engendré par $F\cup G$, c'est-à-dire que tout ss-e.v. contenant le sous-ensemble $F\cup G$ contient le ss-e.v. $F+G$. Ainsi, on pallie le fait que $F\cup G$ n'est en général pas un ss-e.v. en considérant l'ensemble en général plus grand $F+G$, qui lui est un ss-e.v. et qui est le plus petit ss-e.v. qui contient $F\cup G$ (au sens précisé ci-dessus).

1.2.5 Somme directe et sous-espaces vectoriels supplémentaires

a) Définition et notation. Soient F et G deux ss-e.v. de E.

On dit qu'ils sont *supplémentaires* dans E lorsque tout vecteur de E se décompose *de façon unique* en la somme d'un vecteur de F et d'un vecteur de G.
On dit aussi dans ce cas que E est *somme directe* de F et G et on note : $E=F\oplus G$.

Ainsi : $(E=F\oplus G)\Leftrightarrow(\forall\ u\in E,\ \exists!\ v\in F,\ \exists!\ w\in G,\ u=v+w)$.

La proposition suivante donne une autre caractérisation d'une somme directe.

b) Proposition. *Soient F et G deux ss-e.v. de E ; alors :*

$$(E=F\oplus G)\Leftrightarrow(E=F+G\ \text{ et }\ F\cap G=\{0_E\}).$$

Preuve : supposons $E=F+G$ et $F\cap G=\{0_E\}$. Soit $u\in E$ quelconque.
Parce que $E=F+G$, il existe $v\in F$ et $w\in G$ tels que $u=v+w$. Pour montrer que cette décomposition est unique, supposons que l'on ait aussi $u=v'+w'$ avec $v'\in F,w'\in G$. Alors $v+w=v'+w'$ implique $v-v'=w'-w$.
Mais $v-v'\in F$ (différence de deux vecteurs du ss-e.v. F) et de même $w'-w\in G$.
Ainsi $v-v'=w'-w\in F\cap G$, qui est réduit à $\{0_E\}$.
Donc $v-v'=w'-w=0_E$, c'est-à-dire $v=v'$ et $w=w'$.
On a montré : $(\forall\ u\in E,\ \exists!\ v\in F,\ \exists!\ w\in G,\ u=v+w)$. On conclut : $E=F\oplus G$.
La réciproque est à peu près évidente, et laissée en exercice au lecteur. □

b) Exemple. Dans $E=\mathbb{R}^3$, $P=\{(x,0,z)\,;\,x\in\mathbb{R},z\in\mathbb{R}\}$ et $S=\{(0,y,0)\,;\,y\in\mathbb{R}\}$ sont des ss-e.v. supplémentaires. On note : $\mathbb{R}^3=P\oplus S$.

1.3 Engendrement et familles génératrices

1.3.1 Exemples préliminaires

• Supposons ici que $E = \mathbb{R}^3$. Considérons les deux vecteurs $u = (0, 1, 2)$ et $v = (-1, 3, 1)$. On peut à partir de ces deux vecteurs en fabriquer "beaucoup" d'autres en utilisant l'addition et le produit externe, c'est-à-dire en prenant toutes les combinaisons linéaires $\alpha.u + \beta.v$ où α et β décrivent $\mathbb{R}$.

• Peut-on obtenir ainsi tous les vecteurs de E ? La réponse est non. En effet, le vecteur $w = (1, 1, 1)$ par exemple ne peut jamais s'écrire $\alpha.u + \beta.v$. Il suffit pour s'en convaincre de remarquer que l'on aurait sinon $(1, 1, 1) = (-\beta, \alpha+3\beta, 2\alpha+\beta) = (1, 1, 1)$, d'où $\beta = -1$, $\alpha-3 = 1$ et $2\alpha-1 = 1$, qui n'a pas de solution dans $\mathbb{R}$. On traduit cette propriété en disant que la famille de vecteurs $\mathcal{X} = (u, v)$ n'engendre pas E (ou n'est pas une famille génératrice de E).

• Que peut-on dire alors de l'ensemble des vecteurs de E qui sont de la forme $\alpha.u + \beta.v$? Cet ensemble F est égal à $\{(-\beta, \alpha + 3\beta, 2\alpha + \beta) ; \alpha, \beta \in \mathbb{R}\}$, et il est facile de montrer que c'est un ss-e.v. de E. On traduit cette propriété en disant que la famille de vecteurs $\mathcal{X} = (u, v)$ engendre F (ou est une famille génératrice de E).

Dans toute la suite de cette section, on fixe $n \geq 1$ et on désigne par E le $\mathbb{K}$-e.v. $\mathbb{K}^n$.

1.3.2 Combinaisons linéaires et sous-espace vectoriel engendré

Fixons un entier $p \geq 1$. Soit $\mathcal{X} = (u_1, u_2, \dots, u_p)$ une famille finie de p vecteurs de E.

a) Définition. On appelle *combinaison linéaire* des vecteurs de $\mathcal{X}$ tout vecteur de E de la forme :

$$\alpha_1.u_1 + \alpha_2.u_2 + \cdots + \alpha_p.u_p, \text{ avec } \alpha_1, \alpha_2, \dots, \alpha_p \text{ des scalaires de } \mathbb{K}.$$

Les α_i sont appelés les coefficients de la combinaison linéaire.

On utilisera dans la suite l'abréviation c.l. pour combinaison linéaire.

On utilisera aussi la notation globale habituelle : $\alpha_1.u_1 + \alpha_2.u_2 + \cdots + \alpha_p.u_p = \sum_{i=1}^{p} \alpha_i.u_i$.

b) Notation. On note $\operatorname{Vect}\mathcal{X}$ l'ensemble des combinaisons linéaires des vecteurs de $\mathcal{X}$. Donc :

$$\forall\, v \in E,\ (v \in \operatorname{Vect}\mathcal{X}) \Leftrightarrow (\exists\ \alpha_1, \alpha_2, \dots, \alpha_p \in \mathbb{K},\ v = \alpha_1.u_1 + \alpha_2.u_2 + \cdots + \alpha_p.u_p)$$

c) Proposition et définition. *Pour toute famille finie $\mathcal{X}$ de vecteurs de E,* $\operatorname{Vect}\mathcal{X}$ *est un ss-e.v. de E, appelé le ss-e.v. engendré par $\mathcal{X}$.*

Preuve : c'est un simple exercice. A faire ! □

d) Proposition.
Pour toute famille finie $\mathcal{X}$ de vecteurs de E, le ss-e.v. $\operatorname{Vect}\mathcal{X}$ *contient $\mathcal{X}$, et c'est le plus petit ss-e.v. de E qui contient $\mathcal{X}$, [au sens que tout ss-e.v. de E contenant $\mathcal{X}$ contient* $\operatorname{Vect}\mathcal{X}$*].*

Preuve : soit u_j l'un des vecteurs de $\mathcal{X}$. On écrit u_j comme la combinaison linéaire
$u_j = 0.u_1 + \cdots + 0.u_{j-1} + 1.u_j + 0.u_{j+1} + \cdots + 0.u_p$. Donc $u_j \in \operatorname{Vect}\mathcal{X}$. Ce qui prouve $\mathcal{X} \subset \operatorname{Vect}\mathcal{X}$. De plus, soit F un ss-e.v. de E tel que $\mathcal{X} \subset F$. Parce que F est stable par c.l., toute c.l. de vecteurs de F appartient encore à F. En particulier, comme $\mathcal{X} \subset F$, toute c.l. de vecteurs de $\mathcal{X}$ appartient à F. Donc $\operatorname{Vect}\mathcal{X} \subset F$. □

1.3.3 Famille génératrice

Soit F un ss-e.v. de E. Intuitivement, et comme on l'a vu à l'exemple préliminaire ci-dessus, dire qu'une famille finie $\mathcal{X}$ de p vecteurs engendre F (ou est génératrice de F) signifie que l'on peut fabriquer tous les vecteurs de F simplement par c.l. à partir des p vecteurs de $\mathcal{X}$. Précisons :

a) Définition.
Soit F un ss-e.v. de E. Soit $\mathcal{X} = (u_1, u_2, \ldots, u_p)$ une famille finie de p vecteurs de E. On dit que $\mathcal{X}$ *engendre* F, ou que $\mathcal{X}$ est une *famille génératrice* de F, lorsque le ss-e.v. de E engendré par $\mathcal{X}$ est égal à F.

Ceci équivaut donc à : $\operatorname{Vect} \mathcal{X} = F$,
ou encore à : tout vecteur de F est c.l. des vecteurs de $\mathcal{X}$,
ou encore à : $\forall\, v \in F,\ \exists\, \alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbb{K},\ v = \alpha_1.u_1 + \alpha_2.u_2 + \cdots + \alpha_p.u_p.$

b) Remarques.
• Par définition même, pour toute famille finie $\mathcal{X}$ de vecteurs de E, la famille $\mathcal{X}$ est génératrice du ss-e.v. $F = \operatorname{Vect} \mathcal{X}$.

• Il est clair que, si $\mathcal{X} = (u_1, u_2, \ldots, u_p)$ est une famille génératrice de F telle que le vecteur u_p soit c.l. des autres vecteurs $u_1, u_2, \ldots, u_{p-1}$, alors la famille $\mathcal{X}' = (u_1, u_2, \ldots, u_{p-1})$ est encore génératrice de F.

• La définition a) ci-dessus s'applique en particulier au cas où le ss-e.v. F de E que l'on considère est E lui-même. Dans ce cas, il est courant (bien qu'imprudent) d'utiliser l'expression "famille génératrice" toute seule (sans préciser ensuite de quel ss-e.v. on parle). Ainsi, dire que $\mathcal{X} = (u_1, u_2, \ldots, u_p)$ est une famille génératrice (sous-entendu de E) signifie que tout vecteur de E est c.l. des vecteurs de $\mathcal{X}$.

1.4 Dépendance linéaire, familles liées, familles libres

1.4.1 Exemples préliminaires

• Supposons ici que $E = \mathbb{R}^3$. Considerons les vecteurs $u = (0, 1, 2)$, $v = (-1, 3, 1)$ et $s = (-2, 7, 4)$. Il est clair que $u + 2.v + (-1).s = (0, 0, 0)$. On a donc une combinaison linéaire de ces 3 vecteurs qui est nulle, sans que les 3 coefficients (à savoir $1, 2, -1$) soient simultanément nuls. On traduit cette propriété en disant que la famille $\mathcal{X} = (u, v, s)$ est liée. Cela permet d'exprimer l'un des vecteurs comme une combinaison linéaire des deux autres, par exemple $s = u + 2.v$. Géométriquement, cela signifie que le vecteur s est dans le plan déterminé par les vecteurs u et v.

• Prenons maintenant $u = (0, 1, 2)$, $v = (-1, 3, 1)$ et $t = (0, 1, 0)$. Supposons qu'une combinaison linéaire $\alpha.u + \beta.v + \gamma.t$ soit nulle. On aurait donc : $(-\beta, \alpha + 3\beta + \gamma, 2\alpha + \beta) = (0, 0, 0)$. On en tire d'abord $\beta = 0$, puis $\gamma = 0$ et $\alpha = 0$. La seule combinaison linéaire de ces 3 vecteurs qui soit nulle est celle dont les coefficients valent tous les 3 zéro. On traduit cette propriété en disant que la famille $\mathcal{X}' = (u, v, t)$ est libre. Géométriquement, cela signifie que les vecteurs u, v, t ne sont pas situés dans un même plan de l'espace.

Dans toute la suite de cette section, on fixe $n \geq 1$ et on désigne par E le $\mathbb{K}$-e.v. $\mathbb{K}^n$.

1.4.2 Famille liée, famille libre

Soit $\mathcal{X} = (u_1, u_2, \ldots, u_p)$ une famille finie de vecteurs de E.
La question est de savoir s'il peut exister une c.l. de ces vecteurs qui soit nulle.

Bien sûr, si l'on prend tous les α_i égaux à 0, on a trivialement $0.u_1 + 0.u_2 + \cdots + 0.u_p = 0_E$. On ne retient donc pour la suite que les cas non-triviaux, c'est-à-dire où les α_i sont *non tous nuls.*
Rappelons que dire que les scalaires $\alpha_1, \alpha_2, \ldots, \alpha_p$ sont *non tous nuls* signifie que *l'un au moins* des α_i est non-nul (ce qui n'exclut pas que certains des α_i puissent être nuls ; mais pas tous...). Ceci équivaut aussi à dire que $(\alpha_1, \alpha_2, \ldots, \alpha_p) \neq (0, 0, \ldots, 0)$ dans $\mathbb{K}^p$.

a) Définition.
On dit que la famille $\mathcal{X} = (u_1, u_2, \ldots, u_p)$ est *liée* lorsqu'il existe une c.l. des vecteurs de $\mathcal{X}$, à coefficients non tous nuls, qui est nulle dans E.
Ceci équivaut donc à :

$$\exists\ \alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbb{K},\ \ \alpha_1.u_1 + \alpha_2.u_2 + \cdots + \alpha_p.u_p = 0_E \text{ et } (\alpha_1, \alpha_2, \ldots, \alpha_p) \neq (0, 0, \ldots, 0).$$

Une telle relation s'appelle une *relation de dépendance linéaire* entre les vecteurs de $\mathcal{X}$.

b) Définition.
On dit que la famille $\mathcal{X} = (u_1, u_2, \ldots, u_p)$ est *libre* lorsqu'elle n'est pas liée.
Ceci signifie donc qu'il n'existe pas de relation de dépendance linéaire à coefficients non tous nuls entre les vecteurs de $\mathcal{X}$,
c'est-à-dire que la seule c.l. nulle des vecteurs de $\mathcal{X}$ est celle où tous les coefficients sont nuls, ce qui se traduit encore par :

$$\forall\ \alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbb{K}, \big(\ \alpha_1.u_1 + \alpha_2.u_2 + \cdots + \alpha_p.u_p = 0_E \implies \alpha_1 = \alpha_2 = \cdots = \alpha_p = 0\ \big).$$

On dit aussi dans ce cas que les vecteurs de $\mathcal{X}$ sont *linéairement indépendants.* On regroupe dans la proposition suivante quelques propriétés utiles des familles libres ou liées.

c) Proposition. *Les familles finies de vecteurs de E vérifient les propriétés suivantes :*

(i) *Toute sous-famille d'une famille libre est libre.*

(ii) *Toute famille qui admet une sous-famille liée est liée.*

(iii) *Toute famille dont l'un des vecteurs est 0_E est liée.*

(iv) *Une famille formée d'un seul vecteur est liée si et seulement si ce vecteur est nul (donc est libre si et seulement si ce vecteur est non-nul).*

(v) *Une famille est liée si et seulement s'il existe un de ses vecteurs qui est combinaison linéaire des autres.*

(vi) *Une famille de deux vecteurs est liée si et seulement si l'un des deux est produit de l'autre par un scalaire.*

Preuve : résulte directement des définitions ; cf. ouvrage de référence. □

d) Remarques.
• Précisons ce que l'on entend par sous-famille dans la proposition ci-dessus : si $\mathcal{X} = (u_1, u_2, \ldots, u_p)$ est une famille finie de vecteurs de E, on appelle sous-famille de $\mathcal{X}$ toute famille de q vecteurs (avec $q \leq p$) choisis parmi les vecteurs de $\mathcal{X}$.

• Il est clair par ailleurs que toute famille qui admet une sous-famille génératrice d'un ss-e.v. F de E est elle-même génératrice de F.

1.5 Bases et dimension

Dans toute cette section, on fixe $n \geq 1$ et on désigne par E le $\mathbb{K}$-e.v. $\mathbb{K}^n$.

1.5.1 Bases et dimension de $\mathbb{K}^n$

a) <u>Définition.</u> On dit qu'une famille $\mathcal{X} = (u_1, u_2, \ldots, u_p)$ est une *base* de E lorsqu'elle est à la fois libre et génératrice de E.

b) <u>Théorème.</u> *Soit $\mathcal{X} = (u_1, u_2, \ldots, u_p)$ une famille finie de vecteurs de E. Les deux assertions suivantes sont équivalentes :*

(i) *$\mathcal{X}$ est une base de E ;*

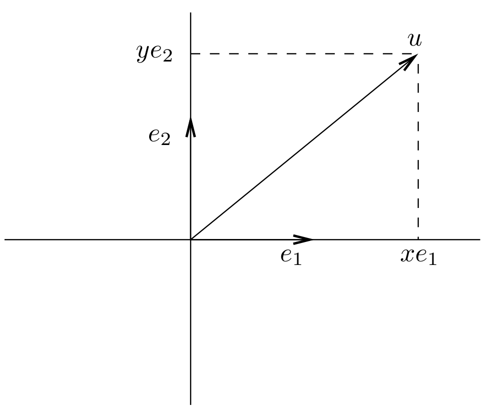
(ii) *Pour tout vecteur $v \in E$, il existe des scalaires $\alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbb{K}$* UNIQUES*, appelés les composantes de v dans la base $\mathcal{X}$, tels que $v = \alpha_1.u_1 + \alpha_2.u_2 + \cdots + \alpha_p.u_p$.*

Preuve : supposons que l'on a (i). Soit $v \in E$. Parce que $\mathcal{X}$ est génératrice, il existe $\alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbb{K}$ tels que $v = \sum_{i=1}^{p} \alpha_i.u_i$. Le problème est l'unicité. Supposons donc que l'on a aussi $v = \sum_{i=1}^{p} \beta_i.u_i$. Par différence membre à membre, on déduit : $\sum_{i=1}^{p}(\alpha_i - \beta_i).u_i = 0_E$. Comme $\mathcal{X}$ est libre, cela implique que, pour tout $1 \leq i \leq p$, on a $\alpha_i - \beta_i = 0$, et donc $\alpha_i = \beta_i$. D'où l'unicité voulue, ce qui prouve (ii).

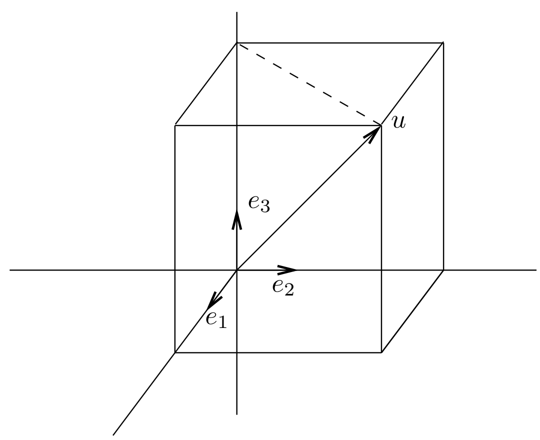
Réciproquement, supposons (ii). Il est clair que cela implique que $\mathcal{X}$ est génératrice. Pour montrer que $\mathcal{X}$ est libre, donnons-nous des scalaires $\alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbb{K}$ tels que $\sum_{i=1}^{p} \alpha_i.u_i = 0_E$. On peut donc écrire : $\sum_{i=1}^{p} \alpha_i.u_i = \sum_{i=1}^{p} 0.u_i$, et alors l'unicité supposée dans (i) implique que $\alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$. Ceci prouve que $\mathcal{X}$ est libre, et donc finalement $\mathcal{X}$ est une base de E, ce qui achève la preuve. $\square$

c) <u>Exemple fondamental : base canonique.</u>

Dans $\mathbb{R}^2$, la base canonique est (e_1, e_2) avec $e_1 = (1, 0)$ et $e_2 = (0, 1)$.
Dans $\mathbb{R}^3$, la base canonique est (e_1, e_2, e_3) avec $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ et $e_3 = (0, 0, 1)$.
Ce sont les bases que l'on utilise classiquement en géométrie ou en physique pour repérer un vecteur par ses composantes :



Tout vecteur u de $\mathbb{R}^2$ s'écrit de façon unique $u = xe_1 + ye_2$



Tout vecteur u de $\mathbb{R}^3$ s'écrit de façon unique $u = xe_1 + ye_2 + ze_3$

D'une façon générale, pour $n \geq 1$ quelconque, considérons dans $E = \mathbb{K}^n$ la famille $\mathcal{X} = (e_1, e_2, \ldots, e_n)$ où :

$$e_1 = (1,0,0,0,\ldots,0), \quad e_2 = (0,1,0,0,,\ldots,0), \quad e_3 = (0,0,1,\ldots,0),$$
$$\ldots,$$
$$e_{n-1} = (0,0,\ldots,0,1,0), \quad e_n = (0,0,\ldots,0,1).$$

Il est facile de vérifier que $\mathcal{X}$ ainsi définie est une base de $\mathbb{K}^n$ (le faire en exercice).

Cette base est appelée la *base canonique* de $\mathbb{K}^n$.

Elle a par définition la propriété caractéristique suivante :
pour tout vecteur $v = (x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$, les composantes de v dans la base canonique sont précisément $(x_1, x_2, \ldots, x_n)$.

d) Théorème et définition.

Toutes les bases du $\mathbb{K}$-e.v. $E = \mathbb{K}^n$ sont formées d'exactement n vecteurs.

On dit que le $\mathbb{K}$-e.v. E est de dimension n, et l'on note $\dim E = n$.

Preuve : admis ; cf. ouvrage de référence. □

e) Remarque.

On synthétise dans le théorème suivant diverses propriétés des bases d'un $\mathbb{K}$-e.v. de dimension finie.

Concernant les deux premières assertions, on retiendra qu'une base est une *famille libre maximale*, ou encore une *famille génératrice minimale*.

Les assertions (3) et (4) concernent quant à elles les familles dont le nombre de vecteurs est exactement égal à la dimension du $\mathbb{K}$-e.v. dans lequel on travaille. Elles constituent un critère d'usage très fréquent et précieux dans la pratique, qu'il est impératif de bien connaître !

f) Théorème. *Soit $\mathcal{X} = (u_1, u_2, \ldots, u_p)$ une famille finie de p vecteurs de $E = \mathbb{K}^n$. On a :*

(1) *Si $\mathcal{X}$ est libre, alors $p \leq n$; (et donc : si $p > n$, alors $\mathcal{X}$ est liée).*

(2) *Si $\mathcal{X}$ est génératrice, alors $p \geq n$.*

(3) *Si $p = n$ et si $\mathcal{X}$ est libre, alors $\mathcal{X}$ est une base de E.*

(4) *Si $p = n$ et si $\mathcal{X}$ est génératrice, alors $\mathcal{X}$ est une base de E.*

Preuve : résulte directement des définitions ; cf. ouvrage de référence. □

1.5.2 Bases et dimension des sous-espaces vectoriels

a) Définition. Soit F un ss-e.v. de $E = \mathbb{K}^n$. Une base de F est une famille finie $\mathcal{X}$ de p vecteurs de F qui est à la fois :
– génératrice de F (c'est-à-dire que tout vecteur de F est c.l. des vecteurs de $\mathcal{X}$).
– libre (et donc $p \leq n$ d'après le point (1) du théorème f) ci-dessus).

b) Remarque. En particulier, puisque $\mathcal{X}$ est toujours une famille génératrice de $\operatorname{Vect} \mathcal{X}$, on retiendra : *toute famille libre $\mathcal{X}$ de vecteurs de E est une base du ss-e.v.* $\operatorname{Vect} \mathcal{X}$ *engendré par $\mathcal{X}$.*

c) Théorème et définition. *Soit F un ss-e.v. de $E = \mathbb{K}^n$. On a alors les propriétés suivantes :*

(1) *Toutes les bases de F sont formées du même nombre de vecteurs, que l'on appelle la dimension de F, et que l'on note* $\dim F$.

(2) *on a toujours :* $0 \leq \dim F \leq \dim E = n$.

(3) $\dim F = n$ *si et seulement si* $F = E$.

(4) $\dim F = 0$ *si et seulement si* $F = \{0_E\}$.

Preuve : admis ; cf. ouvrage de référence. □

d) Remarques et terminologie.

• Attention : un ss-e.v. F de E, (et en particulier E lui-même), admet une infinité de bases, mais toutes sont formées du même nombre p de vecteurs égal à la dimension de F (et en particulier $p = n$ si $F = E$). On verra plus loin comment on passe de l'une à l'autre.

• Un ss-e.v. de dimension 1 s'appelle une *droite vectorielle*. Un ss-e.v. de dimension 2 s'appelle un *plan vectoriel*.

• Le point (4) du théorème c) ci-dessus concerne le cas dégénéré où $F = \{0_E\}$. On ne peut pas trouver de famille libre contenant des vecteurs de F puisque 0_E est le seul vecteur de F et que toute famille contenant 0_E est liée [cf. 1.4.2.c.(iii)]. C'est donc par pure convention, et pour assurer la cohérence de certaines formules sur les dimensions que l'on verra plus loin, que l'on pose $\dim\{0_E\} = 0$.

e) Une méthode pratique.

Soit $\mathcal{X} = (u_1, u_2, \dots, u_p)$ une famille finie de p vecteurs non tous nuls de E.
Notons $F = \operatorname{Vect} \mathcal{X}$ le ss-e.v. de E engendré par $\mathcal{X}$.
En d'autres termes, $\mathcal{X}$ est une famille génératrice de F.
D'après le point (2) de 1.5.1.f, on a $0 < \dim F \leq p$.

En fait, il existe alors une sous-famille de $\mathcal{X}$ (non-unique) qui est une base de F, et il est utile dans la pratique de connaître le raisonnement suivant, qui permet d'en exhiber une explicitement.

– Si dès le départ $\mathcal{X}$ est libre, c'est fini : $\mathcal{X}$ est une base de F [cf. 1.5.2.a] et $\dim F = p$.
– Si $\mathcal{X}$ est liée, alors l'un des vecteurs de $\mathcal{X}$ est c.l. des autres [cf. 1.4.2.c.(5)]. En notant $\mathcal{X}'$ la famille obtenue en enlevant ce vecteur à $\mathcal{X}$, on a alors $\operatorname{Vect} \mathcal{X} = \operatorname{Vect} \mathcal{X}'$, [cf. 1.3.3.b], de sorte que $\mathcal{X}'$ reste une famille génératrice de F (avec strictement moins de vecteurs que $\mathcal{X}$).
  – Si $\mathcal{X}'$ est libre, c'est fini : $\mathcal{X}'$ est une base de F.
  – Si $\mathcal{X}'$ est liée, alors l'un des vecteurs de $\mathcal{X}'$ est c.l. des autres. De même que ci-dessus, la famille $\mathcal{X}''$ obtenue en enlevant ce vecteur à $\mathcal{X}'$ reste une famille génératrice de F.

On réitère ainsi le raisonnement avec des familles successives (extraites chacune de la précédente), qui restent génératrices de F, dont le nombre de vecteurs diminue strictement, jusqu'à obtenir une première sous-famille libre, qui est alors une base de F.

1.5.3 Bases et somme directe

a) Remarque préliminaire.

Si F et G sont deux ss-e.v. de E, il est clair par définition même que l'on obtient une famille génératrice du sous-espace somme $F + G$ en réunissant une famille génératrice de F et une famille génératrice de G. La proposition suivante précise et complète cette observation.

b) Proposition. *Soient F et G deux ss-e.v. de E. Notons $p = \dim F$ et $q = \dim G$.*
Les assertions suivantes sont équivalentes.

(i) $E = F \oplus G$.

(ii) *Pour toute base $\mathcal{X}$ de F et toute base $\mathcal{Y}$ de G, la famille formée en adjoignant les p vecteurs de $\mathcal{X}$ et les q vecteurs de $\mathcal{Y}$ est une base de E.*

(iii) *Il existe une base $\mathcal{X}$ de F et une base $\mathcal{Y}$ de G telles que la famille formée en adjoignant les p vecteurs de $\mathcal{X}$ et les q vecteurs de $\mathcal{Y}$ soit une base de E.*

Preuve : cf. ouvrage de référence. □

c) Corollaire. *Si F et G sont deux ss-e.v. de E tels que $E = F \oplus G$, alors :* $\dim E = \dim F + \dim G$.

Preuve : résulte directement de la proposition précédente. □

Ce résultat est en fait un cas particulier de l'énoncé plus général suivant :

d) Proposition. *Pour tous ss-e.v. F et G de E, on a la relation :*

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

Preuve : cf. ouvrage de référence. □

e) Un complément important.

• La notion de somme directe que l'on a étudiée pour deux ss-e.v. (cf. 1.2.4 et 1.5.3) se généralise à un nombre fini quelconque de ss-e.v. Des ss-e.v. $F_1, F_2, \dots, F_m$ de E sont dits supplémentaires, et l'on note alors $E = F_1 \oplus F_2 \oplus \cdots \oplus F_m$, lorsque :

$$\forall v \in E,\ \exists!\ u_1 \in F_1,\ \exists!\ u_2 \in F_2, \dots,\ \exists!\ u_m \in F_m,\ v = u_1 + u_2 + \cdots + u_m.$$

• On peut alors montrer (comme pour la proposition 1.5.3.b) que :
$E = F_1 \oplus F_2 \oplus \cdots \oplus F_m$ si et seulement si une base de E s'obtient en adjoignant entre elles une base de chacun des ss-e.v. F_i.

• En revanche le critère 1.2.5.b n'est valable (du moins sous cette forme) que pour 2 ss-e.v.

1.5.4 Théorème dit de la base incomplète

a) Remarque préliminaire.
Considérons dans $E = \mathbb{K}^n$ une famille libre $\mathcal{X}$, formée de p vecteurs.
Parce que $\mathcal{X}$ est libre, on sait d'après le point (1) de 1.5.1.f que l'on ne peut pas avoir $p > n$.
Si $p = n$, on sait d'après le point (3) de 1.5.1.f que $\mathcal{X}$ est une base.
Le cas restant $p < n$ fait l'objet du théorème suivant :

b) Théorème.
Considérons dans $E = \mathbb{K}^n$ une famille $\mathcal{X} = (u_1, u_2, \dots, u_p)$ de p vecteurs de E avec $p < n$.
Si $\mathcal{X}$ est libre, alors il existe dans E des vecteurs $v_{p+1}, v_{p+2}, \dots, v_n$ tels que la famille $(u_1, \dots, u_p, v_{p+1}, \dots, v_n)$ soit une base de E.

Preuve : cf. ouvrage de référence. □

c) Commentaire.
Ainsi, toute famille libre est une "base en puissance" au sens que, si elle est "trop petite" pour être une base, on peut toujours lui ajouter des vecteurs pour construire une base.
Attention : cette façon de compléter n'est pas unique.
Attention aussi : cet énoncé n'a de sens que parce que l'on part d'une famille $\mathcal{X}$ qui est libre (il est trivialement faux sinon d'après le point (ii) de 1.4.2.c).

d) Corollaire.
Tout ss-e.v. F de E admet (au moins) un supplémentaire dans E.

Preuve : résulte immédiatement du théorème b) ci-dessus et de la proposition 1.5.3.b. □

Là encore, il convient d'insister sur le fait qu'il n'y a pas du tout unicité. C'est une erreur souvent rencontrée que de parler "du" supplémentaire d'un ss-e.v. dans E alors qu'il faut parler "d'un" supplémentaire.

1.6 Exercices sur le chapitre 1

Exercice 1. Lesquels des trois sous-ensembles suivants du $\mathbb{R}$-e.v. $\mathbb{R}^3$ sont-ils des ss-e.v. ? (Justifier les réponses par une démonstration.)

$$F = \{(x, y, z) \in \mathbb{R}^3\,;\, 2x + 3y - z = 0\},$$
$$F' = \{(x, y, z) \in \mathbb{R}^3\,;\, 2x + 3y - z = 1\},$$
$$F'' = \{(x, y, z) \in \mathbb{R}^3\,;\, 2x^2 + 3y^2 - z^2 = 0\}.$$

Exercice 2. Montrer que, pour chacun des systèmes d'équations homogènes suivants, l'ensemble des solutions est un ss-e.v. de $\mathbb{R}^3$ dont on déterminera une base :

$$(S_1)\quad \left\{\begin{array}{rcrcrcl} x & + & 3y & - & z & = & 0 \\ 2x & + & 5y & + & z & = & 0 \\ & & 2y & - & z & = & 0 \end{array}\right. , \qquad (S_2)\quad \left\{\begin{array}{rcrcrcl} x & - & y & + & 2z & = & 0 \\ 2x & - & 3y & + & z & = & 0 \\ -x & & & - & 5z & = & 0 \end{array}\right. ,$$

$$(S_3)\quad \left\{\begin{array}{rcrcrcl} x & + & y & - & z & = & 0 \\ -x & - & y & + & z & = & 0 \\ 2x & + & 2y & - & 2z & = & 0 \end{array}\right. .$$

Exercice 3. Dans le $\mathbb{R}$-e.v. $\mathbb{R}^3$, on considère les sous-ensembles :

$$F = \{(x, y, z) \in \mathbb{R}^3\,;\, x - 2y + z = 0\},$$
$$G = \{(x, y, z) \in \mathbb{R}^3\,;\, x + y + 2z = 0\},$$
$$H = \{(x, y, z) \in \mathbb{R}^3\,;\, \exists\, \lambda \in \mathbb{R},\ (x, y, z) = \lambda.(1, 1, 2)\}.$$

a) Vérifier que F, G, H sont des ss-e.v. de $\mathbb{R}^3$.

b) Déterminer $F + G$ et $F \cap G$.

c) Montrer que $\mathbb{R}^3 = F \oplus H = G \oplus H$. (Donc F et G sont deux supplémentaires de H.)

Exercice 4. Dans le $\mathbb{R}$-e.v. $\mathbb{R}^3$, on considère les sous-ensembles :

$$P_1 = \{(x, 0, z)\,;\, x, z \in \mathbb{R}\}, \quad \text{et} \quad P_2 = \{(x, y, x)\,;\, x, y \in \mathbb{R}\},$$

ainsi que l'ensemble D des solutions du système homogène : $(S)\quad \left\{\begin{array}{rcrcrcl} x & + & y & - & z & = & 0 \\ 2x & & & - & 2z & = & 0 \\ -x & + & 3y & + & z & = & 0 \end{array}\right. .$

Montrer que P_1 et P_2 sont des ss-e.v. de $\mathbb{R}^3$ et comparer $P_1 \cap P_2$ avec D.

Exercice 5. Dans le $\mathbb{R}$-e.v. $\mathbb{R}^4$, soit F l'ensemble des solutions de l'équation $2x - y + z - 2t = 0$. On définit les vecteurs $u = (1, 1, 1, 1)$, $v = (0, 1, 1, 0)$, $w = (1, 0, -2, 0)$ et $s = (1, 0, 0, 1)$.

a) Montrer que la famille de vecteurs $\mathcal{C} = (u, v, w, s)$ est une famille génératrice de F.

b) Est-ce une famille libre ?

c) En déduire une base $\mathcal{B}$ de F.

d) Compléter $\mathcal{B}$ en une base de $\mathbb{R}^4$ et en déduire un ss-e.v. G supplémentaire de F dans $\mathbb{R}^4$.

Exercice 6.

a) Soient dans le $\mathbb{R}$-e.v. $\mathbb{R}^4$ les vecteurs $u = (1, 5, -2, 0)$ et $v = (0, -2, \frac{3}{2}, \frac{7}{2})$. Montrer que $w = (1, 8, -\frac{17}{4}, -\frac{21}{4})$ appartient à $\mathrm{Vect}(u, v)$.

b) Soient dans le $\mathbb{R}$-e.v. $\mathbb{R}^3$ les vecteurs $a = (2, 3, -1)$, $b = (1, -1, -2)$, $c = (3, 7, 0)$, $d = (5, 0, -7)$ et $e = (0, 0, 1)$. Montrer que $\mathrm{Vect}(a, b) = \mathrm{Vect}(c, d)$. Déterminer une base de $\mathrm{Vect}(a, b) \cap \mathrm{Vect}(e, b)$.

Exercice 7. Dans le $\mathbb{R}$-e.v. $\mathbb{R}^4$, on considère les vecteurs

$$u_1 = \begin{pmatrix}1\\0\\2\\1\end{pmatrix}, \quad u_2 = \begin{pmatrix}0\\-2\\3\\1\end{pmatrix}, \quad u_3 = \begin{pmatrix}0\\0\\1\\2\end{pmatrix}, \quad u_4 = \begin{pmatrix}1\\-2\\4\\0\end{pmatrix}, \quad u_5 = \begin{pmatrix}0\\0\\0\\0\end{pmatrix}, \quad u_6 = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}.$$

Pour chacune des familles de vecteurs suivantes, indiquer si elle est libre ou liée en précisant avec soin l'argument le plus rapide pour conclure : $\mathcal{A} = (u_1, u_2, u_3, u_4, u_6)$, $\mathcal{B} = (u_1, u_2, u_5)$, $\mathcal{C} = (u_1, u_2, u_3, u_4)$, $\mathcal{D} = (u_1, u_2, u_3, u_6)$, $\mathcal{E} = (u_1, u_2, u_3)$, $\mathcal{F} = (u_1)$.

Exercice 8. Dans le $\mathbb{R}$-e.v. $\mathbb{R}^4$, on considère les vecteurs

$$u_1 = \begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \; u_2 = \begin{pmatrix}1\\0\\-1\\1\end{pmatrix}, \; u_3 = \begin{pmatrix}1\\0\\0\\1\end{pmatrix}, \; u_4 = \begin{pmatrix}0\\2\\1\\-2\end{pmatrix}, \; u_5 = \begin{pmatrix}-2\\1\\0\\1\end{pmatrix}, \; u_6 = \begin{pmatrix}-2\\3\\1\\-1\end{pmatrix}, \; u_7 = \begin{pmatrix}-4\\4\\1\\0\end{pmatrix}.$$

On note $F = \mathrm{Vect}(u_1, u_2, u_3, u_4)$ et $G = \mathrm{Vect}(u_5, u_6, u_7)$.

a) Déterminer une base de F puis une base de G.

b) Peut-on avoir $\mathbb{R}^4 = F \oplus G$?

c) Déterminer $F + G$ et en déduire $\dim(F \cap G)$.

d) Montrer que (u_4, u_5, u_6) est liée ; donner une base de $F \cap G$.

Exercice 9. Dans le $\mathbb{R}$-e.v. $\mathbb{R}^4$, on considère les vecteurs

$$v_1 = \begin{pmatrix}1\\2\\0\\1\end{pmatrix}, \; v_2 = \begin{pmatrix}1\\0\\2\\1\end{pmatrix}, \; v_3 = \begin{pmatrix}2\\0\\4\\2\end{pmatrix}, \; w_1 = \begin{pmatrix}1\\2\\1\\0\end{pmatrix}, \; w_2 = \begin{pmatrix}-1\\1\\1\\1\end{pmatrix}, \; w_3 = \begin{pmatrix}2\\-1\\0\\1\end{pmatrix}, \; w_4 = \begin{pmatrix}2\\2\\2\\2\end{pmatrix}.$$

a) Montrer que les familles (v_1, v_2), (w_1, w_2, w_3) et (v_1, v_2, w_1, w_2) sont libres.

b) Déterminer une base de $F = \mathrm{Vect}(v_1, v_2, v_3)$ et une base d'un supplémentaire dans $\mathbb{R}^4$ de F.

c) Déterminer une base de $G = \mathrm{Vect}(w_1, w_2, w_3, w_4)$.

d) Déterminer le sous-espace somme $F + G$.

e) Montrer que $v_1 + v_2 \in F \cap G$, et donner une base de $F \cap G$.

Chapitre 2

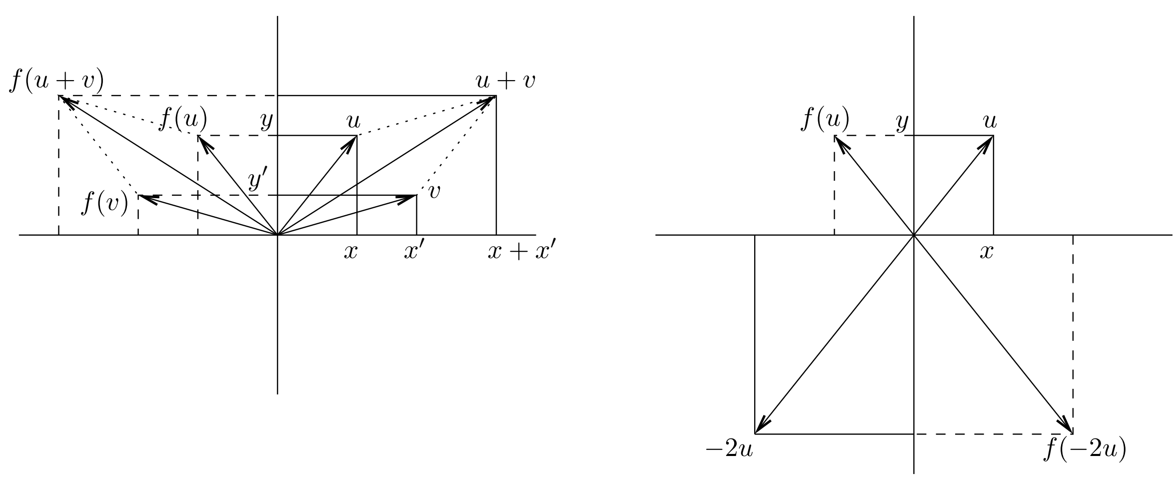
Applications linéaires

2.1 Endomorphismes de $\mathbb{K}^n$

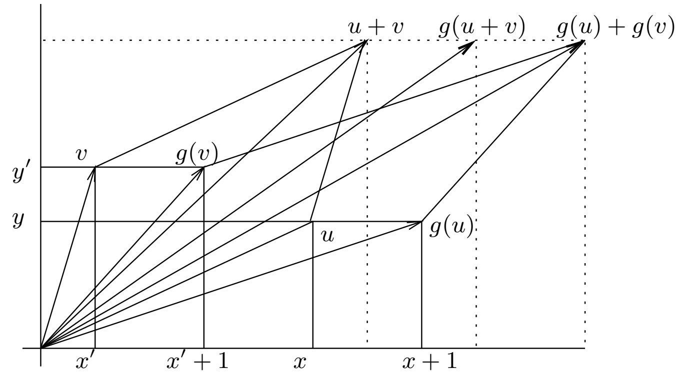
2.1.1 Exemples préliminaires

Plaçons-nous dans le $\mathbb{R}$-e.v. $E = \mathbb{R}^2$.

Considérons d'abord l'application f de E dans E consistant à envoyer un vecteur quelconque $u = (x, y)$ sur le vecteur $f(u) = (-x, y)$. Quels que soient $u = (x, y)$ et $v = (x', y')$, on a $f(u) = (-x, y)$ et $f(v) = (-x', y')$, de sorte $f(u)+f(v) = (-x-x', y+y')$. Comme par ailleurs $u+v = (x+x', y+y')$, on a donc : $f(u + v) = f(u) + f(v)$. En d'autres termes, l'image par f de la somme de deux vecteurs est égal à la somme des images par f de ces deux vecteurs. De la même façon, on a pour tout vecteur $u = (x, y)$ et tout scalaire $\alpha \in \mathbb{R}$, les égalités : $f(\alpha.u) = f(\alpha x, \alpha y) = (-\alpha x, \alpha y) = \alpha(-x, y) = \alpha.f(u)$. On traduit ces deux propriétés en disant que f est une application linéaire.



Considérons maintenant l'application g de E dans E consistant à envoyer un vecteur quelconque $u = (x, y)$ sur le vecteur $g(u) = (x + 1, y)$. Quels que soient $u = (x, y)$ et $v = (x', y')$, on a $g(u) = (x + 1, y)$ et $g(v) = (x' + 1, y')$, de sorte $g(u) + g(v) = (x + x' + 2, y + y')$, alors que $g(u + v) = g(x + x', y + y') = (x + x' + 1, y + y')$. Cette application g n'est donc pas linéaire.



L'objet de cette section est d'étudier, dans la cadre général où $E = \mathbb{K}^n$, les applications de E dans E qui respectent la structure de $\mathbb{K}$-e.v. de E, c'est-à-dire qui sont linéaires. On les appelle des endomorphismes de E.

Dans toute la suite de cette section, on fixe $n \geq 1$ et on désigne par E le $\mathbb{K}$-e.v. $\mathbb{K}^n$.

2.1.2 Notion d'endomorphisme

a) Définition.
Soit f une application $E \to E$. On dit que f est *linéaire*, ou que f est un *endomorphisme* du $\mathbb{K}$-e.v. E, lorsqu'elle vérifie les deux conditions :

$$\forall\, u, v \in E,\ f(u+v) = f(u) + f(v) \qquad \text{et} \qquad \forall\, u \in E,\ \forall\, \alpha \in \mathbb{K},\ f(\alpha.u) = \alpha.f(u).$$

b) Remarques.
• Les deux conditions de la définition peuvent être regroupées en une seule :

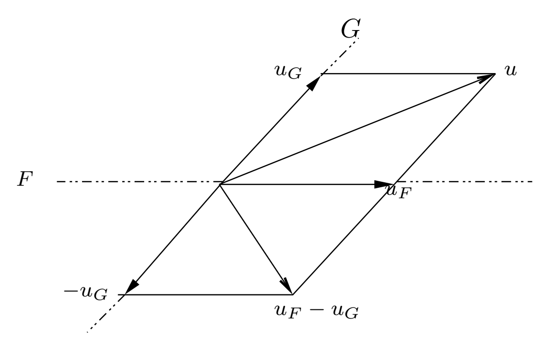
$$\big(f : E \to E \ \text{ est linéaire }\big) \ \Leftrightarrow\ \big(\forall\, u, v \in E,\ \forall\, \alpha, \beta \in \mathbb{K},\ f(\alpha.u + \beta.v) = \alpha.f(u) + \beta.f(v)\big),$$

qui s'étend à des combinaisons linéaires d'un nombre fini quelconque de vecteurs de E.

• Il est clair qu'une application linéaire $f : E \to E$ vérifie toujours $f(0_E) = 0_E$.

c) Notations. On note $\operatorname{End} E$ ou $\mathcal{L}(E)$ l'ensemble des endomorphismes de E.

d) Exemples.
• Soient F et G deux ss-e.v. de E tels que $E = F \oplus G$. On sait qu'alors tout $u \in E$ se décompose de façon unique en $u = u_F + u_G$ avec $u_F \in F$ et $u_G \in G$. L'application $E \to E$ définie par $u \mapsto u_F$ est linéaire, et est appelée la *projection* vectorielle sur F parallèlement à G. L'application $E \to E$ définie par $u \mapsto u_F - u_G$ est linéaire, et est appelée la *symétrie* vectorielle par rapport à F parallèlement à G.



• Pour tout $\lambda \in \mathbb{K}$ fixé, l'application $E \to E$ définie par $u \mapsto \lambda.u$ est un endomorphisme de E appelé l'*homothétie* vectorielle de rapport λ.

• En particulier, l'application qui à tout $u \in E$ associe le vecteur nul 0_E est un endomorphisme de E, appelé endomorphisme nul, que l'on notera O.

• En particulier aussi, l'application qui à tout $u \in E$ associe le vecteur u lui-même est un endomorphisme de E, appelé l'identité (ou application identique) de E, noté id_E.

• Soient a, b, c, d des scalaires quelconques. L'application $(x, y) \mapsto (ax + by, cx + dy)$ de $\mathbb{K}^2$ dans $\mathbb{K}^2$ est un endomorphisme.

De même, l'application $(x, y, z) \mapsto (ax + a'y + a''z, bx + b'y + b''z, cx + c'y + c''z)$ de $\mathbb{K}^3$ dans $\mathbb{K}^3$ est un endomorphisme, quel que soient les scalaires $a, a', a'', b, b', b'', c, c', c''$.

Cet exemple se généralise bien sûr à tout $\mathbb{K}^n$, et on verra plus loin que l'on peut toujours se ramener via les composantes par rapport à une base à ce type de transformations.

e) Une observation pratique importante.
Un endomorphisme f de E est entièrement déterminé par les images par f des vecteurs d'une base quelconque de E.

En effet, soit $\mathcal{B} = (u_1, u_2, \ldots, u_n)$ une base de E. Si v est un vecteur quelconque de E décomposé en $v = \sum_{i=1}^n \alpha_i u_i$ dans la base $\mathcal{B}$, avec $\alpha_1, \ldots, \alpha_n$ dans $\mathbb{K}$, son image $f(v)$ est égale à $f(v) = \sum_{i=1}^n \alpha_i f(u_i)$. Il suffit donc de connaître les vecteurs $f(u_1), \ldots, f(u_n)$ pour savoir calculer l'image par f de tout vecteur de E.

En particulier, si g est un endomorphisme de E qui coïncide avec f sur les vecteurs de la base $\mathcal{B}$ [ie. $f(u_i) = g(u_i)$ pour tout $1 \leq i \leq n$], alors $f = g$ [ie. $f(v) = g(v)$ pour tout $v \in E$].

2.1.3 Injectivité, noyau

a) Définition.
Soit f un endomorphisme de E. On appelle *noyau* de f le sous-ensemble de E formé des vecteurs $u \in E$ tels que $f(u) = 0_E$. On le note $\operatorname{Ker} f$. On a donc :

$$\operatorname{Ker} f = \{u \in E\,;\, f(u) = 0_E\}.$$

b) Proposition. *Pour tout endomorphisme f de E,* $\operatorname{Ker} f$ *est un ss-e.v. de E.*

Preuve : c'est un simple exercice. A faire ! □

c) Proposition. *Un endomorphisme f de E est injectif si et seulement si* $\operatorname{Ker} f = \{0_E\}$.

Preuve : Rappelons d'abord que, par définition :

$$f \text{ est injectif} \quad \Leftrightarrow \quad \forall\, u, v \in E,\ [\ f(u) = f(v) \Rightarrow u = v\ ].$$

Donnons maintenant une preuve du théorème, intéressante car elle est typique des arguments de linéarité : pour montrer une propriété concernant tous les vecteurs de E (ici l'injectivité), il suffit de la montrer pour 0_E.
Supposons que f est injective. Soit alors $u \in \operatorname{Ker} f$. Donc $f(u) = 0_E$. Mais (cf. 2.1.2.d) on a aussi $f(0_E) = 0_E$. Donc $f(u) = f(0_E)$, et l'injectivité implique $u = 0_E$. Ceci prouve bien que $\operatorname{Ker} f = \{0_E\}$.
Réciproquement, supposons que $\operatorname{Ker} f = \{0_E\}$. Donnons-nous $u, v \in E$ tels que $f(u) = f(v)$. Alors, $0_E = f(u) - f(v) = f(u - v)$ par linéarité. Donc $u - v \in \operatorname{Ker} f$, c'est-à-dire $u - v = 0_E$, et donc $u = v$. Ceci prouve l'injectivité de f. □

d) Dans la pratique.
Pour déterminer le noyau d'un endomorphisme $f : E \to E$ explicitement défini par des formules donnant les composantes de $f(u)$ en fonction de celles de u (relativement à une base de E), il suffit de résoudre les équations traduisant l'égalité $f(u) = 0_E$. On obtient ainsi un système d'équations du ss-e.v. $\operatorname{Ker} f$, ce qui permet ensuite d'en trouver une base (voir plus loin 2.3.3).

2.1.4 Surjectivité, image, rang

a) Définition.
Soit f un endomorphisme de E. On appelle *image* de f le sous-ensemble de E formé des vecteurs $u' \in E$ qui sont l'image par f d'un vecteur (au moins) de E. On le note $\operatorname{Im} f$. On a donc :

$$\operatorname{Im} f = \{u' \in E\,;\, \exists\, u \in E,\ u' = f(u)\} = \{f(u)\,;\, u \in E\}.$$

b) Proposition. *Pour tout endomorphisme f de E, $\operatorname{Im} f$ est un ss-e.v. de E.*

Preuve : c'est un simple exercice. A faire ! □

c) Proposition. *Un endomorphisme f de E est surjectif si et seulement si $\operatorname{Im} f = E$.*

Preuve : il suffit de rappeler que, par définition :

$$f \text{ est surjectif} \quad \Leftrightarrow \quad \forall\, u' \in E,\ \exists\, u \in E,\ u' = f(u).$$

La preuve de la proposition est alors immédiate. □

d) Une observation pratique importante.
Soit f un endomorphisme de E. Alors les images par f des vecteurs d'une base quelconque de E forment une famille génératrice de $\operatorname{Im} f$.

En effet, soit $\mathcal{B} = (u_1, u_2, \ldots, u_n)$ une base de E. On peut donc considérer la famille $\mathcal{C} = (f(u_1), f(u_2), \ldots, f(u_n))$ de vecteurs de E. Soit $v' \in \operatorname{Im} f$ quelconque. Il existe donc au moins un vecteur $v \in E$ tel que $v' = f(v)$. Ce vecteur v se décompose dans la base $\mathcal{B}$: il existe certains $\alpha_1, \ldots, \alpha_n$ uniques dans $\mathbb{K}$ tels que $v = \sum_{i=1}^n \alpha_i u_i$. Donc, par linéarité, $v' = f(v) = f(\sum_{i=1}^n \alpha_i u_i) = \sum_{i=1}^n \alpha_i f(u_i)$ apparaît comme une c.l. des vecteurs de $\mathcal{C}$.

Ainsi, il est facile dans la pratique de connaître une famille génératrice de $\operatorname{Im} f$. On peut alors en déduire une base de $\operatorname{Im} f$ (par exemple avec la méthode exposée en 1.5.2.e), et éventuellement des équations de $\operatorname{Im} f$ (voir plus loin en 2.3.3).

e) Définition. Soit f un endomorphisme de E. On appelle *rang* de f, noté $\operatorname{rg} f$, la dimension du ss-e.v. $\operatorname{Im} f$ de E. On a donc :

$$0 \leq \operatorname{rg} f = \dim \operatorname{Im} f \leq \dim E = n$$

Le théorème suivant (très important !) porte le nom de "formule du rang".

f) Théorème (important). *Pour tout endomorphisme f de E, on a :*

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

Preuve. cf. ouvrage de référence. □

2.1.5 Bijectivité, action sur les bases

a) Rappels.
• Par définition, une application $f : E \to E$ est dite *bijective* de E sur E lorsqu'elle est à la fois injective et surjective, ce qui équivaut à : $\forall\, u' \in E,\ \exists!\, u \in E,\ u' = f(u)$.

• Si f est bijective de E sur E, il existe une bijection de E sur E, appelée la *bijection réciproque* de f, notée f^{-1}, qui vérifie $f \circ f^{-1} = \mathrm{id}_E$ et $f^{-1} \circ f = \mathrm{id}_E$.
Elle est définie par :

pour tout $u' \in E$, le vecteur $u = f^{-1}(u')$ est l'unique vecteur de E tel que $f(u) = u'$.

b) Proposition.
Soit f un endomorphisme de E. Si f est bijective de E sur E, alors sa bijection réciproque f^{-1} est elle-même linéaire, donc un endomorphisme de E.

Preuve : cf. ouvrage de référence. □

c) Définition, notation.

• Un endomorphisme de E qui est bijectif de E sur E s'appelle un *automorphisme* de E. D'après le b) ci-dessus, sa bijection réciproque f^{-1} est alors aussi un automorphime de E.

• On note $GL(E)$ l'ensemble des automorphismes de E. On l'appelle le *groupe linéaire* de E, (terminologie justifiée au paragraphe suivant).

• On verra dans les chapitres ultérieurs plusieurs critères pratiques (rang, déterminant,...) pour reconnaître en dimension finie si un endomorphisme donné est ou non un automorphisme. Une première caractérisation concerne l'action sur les bases de E, et fait l'objet du paragraphe suivant.

d) Proposition.

Soit $\mathcal{B} = (u_1, u_2, \dots, u_n)$ une base de E.
Soit $\mathcal{C} = (u'_1, u'_2, \dots, u'_n)$ une famille finie quelconque de n vecteurs de E.
Alors il existe un unique endomorphisme f de E tel que :

$$f(u_1) = u'_1,\ f(u_2) = u'_2,\ \dots,\ f(u_n) = u'_n.$$

En outre : – *(f injective de E dans E) $\Leftrightarrow$ ($\mathcal{C}$ est une famille libre dans E) ;*
– *(f surjective de E sur E) $\Leftrightarrow$ ($\mathcal{C}$ est une famille génératrice de E) ;*
– *(f bijective de E sur E) $\Leftrightarrow$ ($\mathcal{C}$ est une base dans E).*

Preuve : repose sur les calculs des 2.1.2.e et 2.1.4.d précédents ; cf. ouvrage de référence. □

e) Corollaire.

Soit f en endomorphisme de E. Les assertions suivantes sont équivalentes :

(i) *f est un automorphisme, [ou encore f est bijective, ou encore $f \in GL(E)$] ;*

(ii) *pour toute base $\mathcal{B}$ de E, les images par f des vecteurs de $\mathcal{B}$ forment une base de E.*

(iii) *il existe une base $\mathcal{B}$ de E telle que les images par f des vecteurs de $\mathcal{B}$ forment une base de E.*

Preuve : conséquence immédiate du d) précédent. □

N.B. On retient cet énoncé (très utile dans la pratique) sous la forme : *un endomorphisme est un automorphisme si et seulement s'il transforme une base en une base.*

Le théorème ci-dessous, très utile dans la pratique, met en évidence qu'alors qu'a priori une application $E \to E$ est bijective si et seulement si elle est à la fois injective et surjective, dans le cas d'un endomorphisme, une seule des deux propriétés suffit, l'autre étant alors automatiquement vérifiée.

f) Théorème (très utile dans la pratique). Soit f un endomorphisme de $E = \mathbb{K}^n$. On a :

$$f \text{ injective} \ \Leftrightarrow\ f \text{ surjective} \ \Leftrightarrow\ f \text{ bijective} ;$$

Preuve : donnons deux méthodes de démonstration.

Première méthode. Supposons f injective. D'après 2.1.3.c, on a alors $\operatorname{Ker} f = \{0_E\}$. Donc d'après 2.1.4.f, on a $\dim \operatorname{Im} f = n$. Ceci implique d'après 1.5.2.c.(3) que $E = \operatorname{Im} f$. On conclut avec 2.1.4.c que f est surjective. La réciproque s'obtient avec les mêmes arguments.

Seconde méthode. Supposons f injective. Prenons $\mathcal{B} = (u_1, u_2, \dots, u_n)$ une base de E et notons $\mathcal{C} = (u'_1, u'_2, \dots, u'_n)$ la famille de n vecteurs de E définie par $f(u_1) = u'_1,\ f(u_2) = u'_2,\ \dots,\ f(u_n) = u'_n$. Parce que f est supposée injective, on sait d'après la proposition d) ci-dessus que $\mathcal{C}$ est libre. D'après 1.5.1.f.(3), comme $\mathcal{C}$ est formée de n vecteurs, il en résulte que $\mathcal{C}$ est aussi génératrice de E, d'où f surjective en réappliquant la proposition d). La réciproque s'obtient avec les mêmes arguments. □

2.1.6 Algèbre des endomorphismes, groupe linéaire

a) Somme et produit externe.

• Si f et g sont deux endomorphismes de E, on définit l'application somme $f+g : E \to E$ par :

$$\forall\ u \in E,\ (f+g)(u) = f(u) + g(u).$$

Il est facile de vérifier que $f+g$ est encore un endomorphisme de E.

• Si f est un endomorphisme de E, et $\alpha \in \mathbb{K}$ un scalaire, on définit l'application $\alpha.f : E \to E$ par :

$$\forall\ u \in E,\ (\alpha.f)(u) = \alpha.f(u).$$

Il est facile de vérifier que $\alpha.f$ est encore un endomorphisme de E.

• On montre facilement (laissé en exercice) que :

pour ces lois + et . l'ensemble $\operatorname{End} E$ *des applications linéaires* $E \to E$ *est un* $\mathbb{K}$*-e.v.*

b) Composition.

• Rappelons que, pour deux applications $f : E \to E$ et $g : E \to E$, on définit la composée $g \circ f : E \to E$ par : $\forall\ u \in E, (g \circ f)(u) = g(f(u))$.

• On montre alors facilement (laissé en exercice) la proposition suivante :

Si f et g sont deux endomorphismes de E, alors $g \circ f$ est encore un endomorphisme de E.

c) Algèbre des endomorphismes.

D'après le a) ci-dessus, on sait déjà que $\operatorname{End} E$ est un $\mathbb{K}$-e.v.

De plus, d'après le b), on a : $\forall\ f, g \in \operatorname{End} E,\ \ g \circ f \in \operatorname{End} E$.

Cela signifie que la composition $\circ$ définit une loi interne (une opération, un produit) dans $\operatorname{End} E$.

Elle est bien sûr associative : $\forall\ f, g, h \in \operatorname{End} E,\ \ f \circ (g \circ h) = (f \circ g) \circ h = f \circ g \circ h$.

L'application identique id_E est neutre pour cette loi : $\forall\ f \in \operatorname{End} E,\ \ f \circ \mathrm{id}_E = \mathrm{id}_E \circ f = f$.

C'est un simple exercice de montrer que l'on a aussi :

$\forall\ f, g, h \in \operatorname{End} E,\ \ f \circ (g+h) = f \circ g + f \circ h$ et $(f+g) \circ h = f \circ h + g \circ h$,

$\forall\ f, g \in \operatorname{End} E,\ \forall\ \alpha \in \mathbb{K},\ \ \alpha.(f \circ g) = (\alpha.f) \circ g = f \circ (\alpha.g)$.

On résume la conjonction de toutes ces propriétés en disant que :

$\operatorname{End} E$ *muni des trois lois +, . et $\circ$ est une* $\mathbb{K}$*-algèbre.*

Attention, la loi $\circ$ n'est en général pas commutative dans $\operatorname{End} E$: on peut avoir $f \circ g \neq g \circ f$.

d) Groupe linéaire.

Restreignons-nous ici aux endormophismes de E qui sont bijectifs, c'est-à-dire qui sont des automorphismes de E. Leur ensemble $GL(E)$ défini en 2.1.5.c vérifie :

(1) la loi $\circ$ est interne dans $GL(E)$: $\forall\ f, g \in GL(E),\ \ f \circ g \in GL(E)$;

(2) la loi $\circ$ est associative dans $GL(E)$: $\forall\ f, g, h \in GL(E),\ \ f \circ (g \circ h) = (f \circ g) \circ h$;

(3) l'application id_E est neutre pour $\circ$ dans $GL(E)$: $\forall\ f \in GL(E),\ \ f \circ \mathrm{id}_E = \mathrm{id}_E \circ f = f$;

(4) pour tout $f \in GL(E)$, il existe $f^{-1} \in GL(E)$ tel que $f \circ f^{-1} = f^{-1} \circ f = \mathrm{id}_E$.

En effet : Le point (1) résulte du b) ci-dessus et du fait général que la composée de deux bijections est une bijection. Les points (2) et (3) sont des propriétés générales de la loi $\circ$. Le point (4) n'est autre que la proposition 2.1.5.b. □

On résume la conjonction de ces quatre propriétés en disant que :

l'ensemble $GL(E)$ des automorphismes de E est un groupe pour la loi $\circ$
(que l'on l'appelle le groupe linéaire de E).

Attention ! En général, le groupe $GL(E)$ n'est pas commutatif : on peut avoir $f \circ g \neq g \circ f$.

2.2 Applications linéaires de $\mathbb{K}^n$ dans $\mathbb{K}^p$

2.2.1 Notion d'application linéaire

a) Définition, notation, remarques.

• Soient $E = \mathbb{K}^n$ et $E' = \mathbb{K}^p$, avec n, p deux entiers naturels. Soit f une application $E \to E'$. On dit que f est *linéaire* (ou que f est un *morphisme de* $\mathbb{K}$*-e.v.*), lorsqu'elle vérifie les deux conditions :

$$\forall\ u, v \in E,\ f(u+v) = f(u) + f(v) \qquad \text{et} \qquad \forall\ u \in E,\ \forall\ \alpha \in \mathbb{K},\ f(\alpha.u) = \alpha.f(u).$$

• On note $\mathcal{L}(E, E')$ l'ensemble des applications linéaires de E dans E'.

• Les endomorphismes de E étudiés en détail à la section précédente ne sont donc qu'un cas particulier (celui où $E = E'$) de telles applications. La plupart des remarques que l'on a faites pour les endomorphismes s'étendent immédiatement à ce contexte un peu plus général. Entre autres :

- les deux conditions de la définition peuvent être regroupées en une seule :
$$\forall\ u, v \in E,\ \forall\ \alpha, \beta \in \mathbb{K},\ f(\alpha.u + \beta.v) = \alpha.f(u) + \beta.f(v),$$
qui s'étend à des combinaisons linéaires d'un nombre fini quelconque de vecteurs de E.

- une application linéaire $f : E \to E'$ vérifie toujours $f(0_E) = 0_{E'}$.

- une application linéaire $f : E \to E'$ est entièrement déterminée par les images par f des vecteurs d'une base quelconque de E ; en particulier, si $\mathcal{B} = (u_1, u_2, \ldots, u_n)$ est une base de E, il suffit de connaître les composantes des vecteurs $f(u_1), f(u_2), \ldots, f(u_n)$ dans une base $\mathcal{C} = (v_1, v_2, \ldots, v_p)$ de E' pour savoir calculer l'image par f de n'importe quel vecteur de E.

• On trouvera dans les exercices de ce chapitre des exemples concrets d'applications linéaires.

b) Somme et produit externe d'application linéaires.
Soient $E = \mathbb{K}^n$ et $E' = \mathbb{K}^p$.

• Si f et g sont deux applications linéaires $E \to E'$, l'application somme $f + g : E \to E'$ définie par : $(f+g)(u) = f(u) + g(u)$ pour tout $u \in E$ est encore linéaire.

• Si f est une application linéaire $E \to E'$, et si $\alpha \in \mathbb{K}$ est un scalaire, l'application $\alpha.f : E \to E'$ définie par : $(\alpha.f)(u) = \alpha.f(u)$ pour tout $u \in E$ est encore linéaire.

• On montre facilement (laissé en exercice) la proposition suivante :

pour ces lois + et . l'ensemble $\mathcal{L}(E, E')$ des applications linéaires $E \to E'$ est un $\mathbb{K}$-e.v.

c) Composition d'applications linéaires.
Soient $E = \mathbb{K}^n$, $E' = \mathbb{K}^p$, $E'' = \mathbb{K}^q$.

• Rappelons que, pour deux applications $f : E \to E'$ et $g : E' \to E''$, on définit la composée $g \circ f : E \to E''$ par : $\forall\ u \in E, (g \circ f)(u) = g(f(u))$.

• On montre alors facilement (laissé en exercice) la proposition suivante :

Si $f : E \to E'$ et $g : E' \to E''$ sont linéaires, alors $g \circ f : E \to E''$ est linéaire.

d) Remarque importante.
Les trois lois (somme, produit externe par un scalaire, produit de composition) définies ci-dessus pour les applications linéaires correspondent exactement aux lois que l'on définira pour les matrices dans les chapitres suivants, (via la correspondance entre applications linéaires et matrices en dimension finie, que l'on précisera alors). Le cas des endomorphismes (que l'on a détaillé en 2.1.6.c) correspondra au cas particulier des matrices carrées.

2.2.2 Image directe ou réciproque d'un sous-espace vectoriel

a) Définition et proposition.
Soient $E = \mathbb{K}^n$, $E' = \mathbb{K}^p$, et f une application linéaire $f : E \to E'$.
(i) Pour tout ss-e.v. F' de E', on appelle *image réciproque* de F' par f l'ensemble :

$$f^{-1}(F') = \{u \in E\,;\, f(u) \in F'\}.$$

C'est un ss-e.v. de E.
(ii) Pour tout ss-e.v. F de E, on appelle *image directe* de F par f l'ensemble :

$$f(F) = \{u' \in E'\,;\, \exists\, u \in F,\ u' = f(u)\} = \{f(u)\,;\, u \in F\}.$$

C'est un ss-e.v. de E'.

Preuve : évidente, en utilisant simplement la définition d'un ss-e.v. et la linéarité de f. □

b) Cas particulier du noyau.
Soient $E = \mathbb{K}^n$, $E' = \mathbb{K}^p$, et f une application linéaire $f : E \to E'$. Si l'on choisit pour ss-e.v. F' de E' le sous-espace nul $\{0_{E'}\}$, on obtient pour ss-e.v. $f^{-1}(F')$ le *noyau* de f :

$$\operatorname{Ker} f = \{u \in E\,;\, f(u) = 0_{E'}\}.$$

Exactement comme en 2.1.3.c pour le cas particulier des endomorphismes, on peut montrer que :

$$f \text{ est injective de } E \text{ dans } E' \qquad \Leftrightarrow \qquad \operatorname{Ker} f = \{0_E\}.$$

c) Cas particulier de l'image.
Soient $E = \mathbb{K}^n$, $E' = \mathbb{K}^p$, et f une application linéaire $f : E \to E'$. Si l'on choisit pour ss-e.v. F de E le sous-espace E lui-même, on obtient pour ss-e.v. $f(F)$ l'*image* de f :

$$\operatorname{Im} f = \{u' \in E'\,;\, \exists\, u \in E,\ u' = f(u)\} = \{f(u)\,;\, u \in E\}.$$

Exactement comme en 2.1.4.d, on peut encore montrer que : *les images par f des vecteurs d'une base quelconque de E forment une famille génératrice du ss-e.v.* $\operatorname{Im} f$ *de* E'.
Et comme en 2.1.4.c pour le cas particulier des endomorphismes, on a aussi :

$$f \text{ est surjective de } E \text{ dans } E' \qquad \Leftrightarrow \qquad \operatorname{Im} f = E'.$$

d) Formule du rang.
Les résultats présentés pour les endomorphisme aux paragraphes 2.1.4.e et 2.1.4.f s'étendent de la façon suivante aux applications linéaires $f : E \to E'$.

Soient $E = \mathbb{K}^n$, $E' = \mathbb{K}^p$, et f une application linéaire $f : E \to E'$.
• *Définition.* On appelle *rang* de f, noté $\operatorname{rg} f$, la dimension du ss-e.v. $\operatorname{Im} f$ de E'. On a donc :

$$0 \leq \operatorname{rg} f = \dim \operatorname{Im} f \leq \dim E' = p.$$

• *Théorème.* On a : $\quad \dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim E = n.$

Preuve. cf. ouvrage de référence. □

• *Remarque.* Attention, le théorème vu précédemment en 2.1.5.f n'est pas vrai en général, mais seulement lorsque $n = p$.

e) Isomorphismes.
Soient $E = \mathbb{K}^n$, $E' = \mathbb{K}^p$, et f une application linéaire $f : E \to E'$.
• On dit que f est un isomorphisme de E sur E' lorsque f est bijective de E sur E'.

• Si f est un isomorphisme de E sur E', alors $n = p$.
  En effet : la bijectivité de f implique que $\operatorname{Ker} f = \{0_E\}$ et $\operatorname{Im} f = E'$, d'où $\dim \operatorname{Ker} f = 0$ et $\dim \operatorname{Im} f = p$, et la formule du rang ci-dessus devient $0 + p = n$.

On est donc dans ce cas ramené à un automorphisme de $\mathbb{K}^n$, au sens de 2.1.5.c.

2.3 Formes linéaires et équations des sous-espaces vectoriels

2.3.1 Formes linéaires et hyperplans

a) Définitions.
Soit $n \geq 1$ et $E = \mathbb{K}^n$. Comme on sait que $\mathbb{K}$ est lui-même un $\mathbb{K}$-e.v. (de dimension 1, cf. 1.5.1.d), on peut considérer les applications linéaires de E dans $\mathbb{K}$.

• On appelle *forme linéaire sur E* toute application linéaire de E dans $\mathbb{K}$.

Une forme linéaire sur E est donc une application $f : E \to \mathbb{K}$ qui, à tout vecteur $u \in E$ associe un scalaire $f(u)$ de telle sorte que : $\forall\ u, v \in E, \forall\ \alpha, \beta \in \mathbb{K},\ f(\alpha.u + \beta.v) = \alpha.f(u) + \beta.f(v)$.

• En particulier, on appelle *forme linéaire nulle* l'application $f : E \to \mathbb{K}$ définie par $f(u) = 0$ pour tout $u \in E$.

• On note E^* l'ensemble des formes linéaires sur E. Par définition, E^* n'est autre que $\mathcal{L}(E, \mathbb{K})$ avec la notation introduite en 2.2.1.a. C'est donc un $\mathbb{K}$-e.v. (cf. 2.2.1.b) ; on l'appelle l'espace dual de E.

La proposition suivante donne une description explicite des formes linéaires sur E.

b) Proposition.
Soit $E = \mathbb{K}^n$. Pour toute application $f : E \to \mathbb{K}$, les deux conditions suivantes sont équivalentes :

(i) *f est une forme linéaire sur E,*

(ii) *il existe dans $\mathbb{K}$ des scalaires uniques $a_1, a_2, \ldots, a_n$ tels que, pour tout $u = (x_1, x_2, \ldots, x_n) \in E$, on ait : $f(u) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$.*

En outre, f est la forme linéaire nulle si et seulement si $a_1 = a_2 = \ldots = a_n = 0$.

Preuve : notons $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ la base canonique de E.
Pour tout $u = x_1.e_1 + x_2.e_2 + \cdots + x_n.e_n$, on a $f(u) = x_1f(e_1) + x_2f(e_2) + \cdots + x_nf(e_n)$ pour toute $f \in E^*$. L'équivalence voulue s'en déduit avec $a_1 = f(e_1), a_2 = f(e_2), \ldots, a_n = f(e_n)$. □

c) Remarque.
Soient $E = \mathbb{K}^n$, et $f \in E^*$ une forme linéaire sur E.
D'après 2.2.2.d, on a $\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim E = n$.
Mais ici $\operatorname{Im} f$ est un ss-e.v. de $\mathbb{K}$, et le $\mathbb{K}$-e.v. $\mathbb{K}$ est de dimension 1. Donc d'après 1.5.2.c.(2), deux cas seulement sont possibles :
• ou bien, $\dim \operatorname{Im} f = 1$, donc $\operatorname{Im} f = \mathbb{K}$, c'est-à-dire f est surjective. Dans ce cas, $\operatorname{Ker} f$ est de dimension $n - 1$.
• ou bien, $\dim \operatorname{Im} f = 0$, donc $\operatorname{Im} f = \{0\}$, c'est-à-dire f est la forme linéaire nulle. Dans ce cas, $\operatorname{Ker} f = E$ (de dimension n).

d) Exemples.
Soit (a, b) fixé $\neq (0, 0)$ dans $\mathbb{R}^2$. L'ensemble des couples (x, y) du plan $\mathbb{R}^2$ vérifiant $ax + by = 0$ est le noyau de la forme linéaire non-nulle $f : (x, y) \mapsto ax + by$, donc est un ss-e.v. de dimension 1. On retrouve la notion classique d'équation d'une droite vectorielle dans le plan $\mathbb{R}^2$.

De même, l'ensemble des triplets (x, y, z) de l'espace $\mathbb{R}^3$ vérifiant $ax + by + cz = 0$, avec (a, b, c) fixé $\neq (0, 0, 0)$, est un plan vectoriel dans l'espace $\mathbb{R}^3$.

e) Définition.
Dans le $\mathbb{K}$-e.v. $E = \mathbb{K}^n$, on appelle *hyperplan vectoriel* de E tout ss-e.v. qui est de dimension $n - 1$.
Donc, dans l'espace $E = \mathbb{R}^3$, les hyperplans vectoriels de E sont les plans vectoriels ;
Dans le plan $E = \mathbb{R}^2$, les hyperplans vectoriels de E sont les droites vectorielles.

f) Théorème.
Pour tout ss-e.v. H de $E = \mathbb{K}^n$, les deux conditions suivantes sont équivalentes :
(i) *H est un hyperplan vectoriel de E,*
(ii) *il existe une forme linéaire non-nulle $f \in E^*$ (non unique) telle que $H = \operatorname{Ker} f$.*

Preuve : le fait que (ii) $\Rightarrow$ (i) a été montré plus haut au point c). La réciproque repose sur l'utilisation du théorème de la base incomplète (cf. ouvrage de référence pour plus de détails). □

2.3.2 Equations linéaires homogènes

a) Définitions. On appelle *équation linéaire homogène* à n inconnues et à coefficients dans $\mathbb{K}$ toute relation de la forme :

$$(*) \qquad a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0, \quad \text{avec } a_1, a_2, \ldots, a_n \text{ des scalaires fixés dans } \mathbb{K}.$$

Tout vecteur $(x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$ tel que l'égalité $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$ soit vraie dans $\mathbb{K}$ s'appelle une *solution* de $(*)$ dans $\mathbb{K}$.

L'ensemble des solutions de $(*)$ est toujours un ss-e.v. de $\mathbb{K}^n$, comme le montre la proposition suivante.

b) Proposition.
L'ensemble des solutions de l'équation $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$

- *est égal à $\mathbb{K}^n$ tout entier dans le cas trivial où tous les coefficients a_i sont nuls* [ie. $(a_1, a_2, \ldots, a_n) = (0, 0, \ldots, 0)$] ;
- *est un hyperplan vectoriel de $\mathbb{K}^n$ si les coefficients a_i sont non tous nuls* [ie. $(a_1, a_2, \ldots, a_n) \neq (0, 0, \ldots, 0)$, ie. *l'un au moins des a_i est non-nul*].

Preuve : D'après 2.3.1.b, l'ensemble de solutions cherché est le noyau de la forme linéaire f sur $\mathbb{K}^n$ définie par $f(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$. D'où le résultat d'après 2.3.1.c. □

c) Définition. Deux équations $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$ et $a'_1x_1 + a'_2x_2 + \cdots + a'_nx_n = 0$ sont dites *équivalentes* lorsqu'elles ont le même ss-e.v. de solutions.

d) Proposition.
Deux équations linéaires homogènes $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$ et $a'_1x_1 + a'_2x_2 + \cdots + a'_nx_n = 0$ telles que $(a_1, a_2, \ldots, a_n) \neq (0, 0, \ldots, 0) \neq (a'_1, \ldots, a'_n)$ sont équivalentes si et seulement si :

$$\exists\, \lambda \in \mathbb{K},\ \lambda \neq 0,\ \text{tel que } a'_1 = \lambda a_1,\ a'_2 = \lambda a_2,\ \ldots,\ a'_n = \lambda a_n.$$

Preuve : un sens est clair ; pour l'autre, cf. ouvrage de référence. □

e) Corollaire de synthèse. (Ce qu'il faut retenir) *Soit $E = \mathbb{K}^n$.*

(1) *Pour tout hyperplan vectoriel H de E, il existe des scalaires $a_1, a_2, \ldots, a_n$ non tous nuls tels que H soit exactement l'ensemble des solutions de l'équation $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$. On dit que cette équation est une équation de H (dans la base canonique de E).*

(2) *Réciproquement, quels que soient $a_1, a_2, \ldots, a_n$ des scalaires non tous nuls, le sous-ensemble H de E formé des solutions de l'équation $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$ est un hyperplan vectoriel de E.*

(3) *Les équations $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$ et $a'_1x_1 + a'_2x_2 + \cdots + a'_nx_n = 0$, avec $(a_1, a_2, \ldots, a_n) \neq (0, 0, \ldots, 0) \neq (a'_1, \ldots, a'_n)$, sont deux équations d'un même hyperplan vectoriel de E (dans la base canonique de E) si et seulement si*

$$\exists\, \lambda \in \mathbb{K}, \lambda \neq 0,\ \text{tel que } a'_1 = \lambda a_1,\ a'_2 = \lambda a_2,\ \ldots,\ a'_n = \lambda a_n.$$

Preuve : résulte de la combinaison des résultats précédents. □

f) Exemples. Reprenons à la lumière de ces énoncés les exemples 2.3.1.d.

• Dans $E = \mathbb{R}^2$, les équations d'une droite vectorielle sont de la forme $ax + by = 0$ avec $(a, b) \neq (0, 0)$. De plus $ax + by = 0$ et $a'x + b'y = 0$ sont équations de la même droite si et seulement si $(a', b') = \lambda.(a, b)$ pour un certain $\lambda \in \mathbb{K}$ non-nul.

• Dans $E = \mathbb{R}^3$, les équations d'un plan vectoriel sont de la forme $ax + by + cz = 0$ avec $(a, b, c) \neq (0, 0, 0)$. De plus $ax+by+cz = 0$ et $a'x+b'y+c'z = 0$ sont équations du même plan si et seulement si $(a', b', c') = \lambda.(a, b, c)$ pour un certain $\lambda \in \mathbb{K}$ non-nul.

• Toujours dans $E = \mathbb{R}^3$ muni de la base canonique considérons :
– un plan vectoriel H d'équation $ax + by + cz = 0$ avec $(a, b, c) \neq (0, 0, 0)$,
– et un plan vectoriel H' d'équation $a'x + b'y + c'z = 0$ avec $(a', b', c') \neq (0, 0, 0)$.

L'ensemble F des vecteurs de E dont les composantes sont solutions de l'une et de l'autre de ces deux équations est donc le ss-e.v. $F = H \cap H'$.

D'une part $\dim F \leq \dim H = \dim H' = 2$.

D'autre part on ne peut pas avoir $\dim F = 0$, car sinon 1.5.3.d donnerait $\dim(H+H') = 2+2-0 = 4$, ce qui est impossible dans le $\mathbb{K}$-e.v. E de dimension 3.

Deux cas seulement sont donc possibles :

– ou bien $\dim F = 2$, c'est-à-dire $F = H = H'$,
ce qui équivaut (cf. ci-dessus) à l'existence d'un $\lambda \in \mathbb{K}$ non-nul tel que $(a', b', c') = \lambda.(a, b, c)$,

– ou bien $\dim F = 1$, lorsqu'il n'existe pas de tel λ.

Ainsi, en dimension 3, l'ensemble des solutions communes de deux équations linéaires homogènes “non multiples l'une de l'autre” est une droite vectorielle.

Cet exemple est un cas particulier de la notion de système introduit ci-dessous.

2.3.3 Systèmes homogènes d'équations linéaires

a) Définitions.
• On appelle *système linéaire homogène de p équations à n inconnues, à coefficients dans* $\mathbb{K}$, la conjonction de p équations de la forme $(*)$ au sens de 2.3.3.a. On note un tel système :

$$(S_0) \left\{ \begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + & a_{1,n}x_n & = & 0 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + & a_{2,n}x_n & = & 0 \\ \cdots & & & & & & & & \\ a_{p,1}x_1 & + & a_{p,2}x_2 & + & \cdots & + & a_{p,n}x_n & = & 0 \end{array} \right. ,$$

où les $a_{i,j}$ sont de scalaires fixés $(1 \leq i \leq p,\ 1 \leq j \leq n)$.

• Une *solution* de (S_0) dans $\mathbb{K}$ est un vecteur de $\mathbb{K}^n$ qui est solution de chacune des p équations qui composent le système.

Il est clair que l'ensemble des solutions d'un tel système linéaire homogène est un ss-e.v. de $\mathbb{K}^n$, puisque c'est l'intersection des ss-e.v. de solutions (hyperplans ou $\mathbb{K}^n$ tout entier) de chacune des p équations (cf. 2.3.2.b et 1.2.3.a).

En particulier, l'ensemble des solutions d'un système linéaire homogène n'est jamais vide, puisque le vecteur nul $(0, 0, \ldots, 0)$ est toujours solution.

• Deux tels systèmes linéaires homogènes sont *équivalents* lorsqu'ils ont même ss-e.v. de solutions.

b) Equations linéaires d'un sous-espace vectoriel.
On peut montrer réciproquement que tout ss-e.v. F de $E = \mathbb{K}^n$ est (de façon non unique) l'intersection d'un nombre fini p d'hyperplans vectoriels (on peut même choisir $p = n - \dim F$; cf. ouvrage de référence). Il existe donc un système (S_0) de p équations linéaires homogènes à n inconnues dans $\mathbb{K}$, tel que F soit exactement le ss-e.v. formé des vecteurs de E qui sont solutions de (S_0).

c) Remarque.
On étudiera plus loin dans le cours certaines méthodes générales pour résoudre de tels systèmes (et même des systèmes non homogènes, c'est-à-dire avec second membre non-nul). Pour l'instant, rappelons simplement la proposition suivante, (que l'on retrouvera ultérieurement dans un contexte plus général), et qui est l'un des arguments les plus précieux dans la pratique.

d) Proposition. (Méthode pratique de résolution)
Soit (S_0) un système linéaire homogène.

(1) *Si l'on permute entre elles deux équations de (S_0), on obtient un système équivalent.*
(2) *Si l'on multiplie par un scalaire non-nul une des équations de (S_0), on obtient un système équivalent.*
(3) *Si l'on ajoute à l'une des équations de (S_0) une combinaison linéaire des autres, on obtient un système équivalent.*

Preuve : (1) et (2) sont clairs ; le (3) est un peu plus délicat : cf. ouvrage de référence, chapitre sur la méthode de Gauss. $\square$

e) Exemple.
Soient α et β deux paramètres réels. On considère dans $\mathbb{R}$ le système linéaire homogène :

$$(S_0)\ \left\{\begin{array}{ccccccc} x & + & \beta y & + & \alpha z & = & 0 \qquad (1)\\ x & + & y & + & \alpha z & = & 0 \qquad (2)\\ x & + & y & + & z & = & 0 \qquad (3)\end{array}\right.$$

D'après la proposition d) ci-dessus, en remplaçant (1) par (1) – (2) et (2) par (2) – (3), on obtient le système équivalent :

$$(S'_0)\ \left\{\begin{array}{ccc} (\beta-1)y & = & 0 \qquad (1')\\ (\alpha-1)z & = & 0 \qquad (2')\\ x + y + z & = & 0 \qquad (3)\end{array}\right.$$

On raisonne donc en distinguant les trois cas suivants :
1er cas : $\alpha \neq 1$ et $\beta \neq 1$
$(1')$ a pour seule solution $y = 0$ et $(2')$ a pour seule solution $z = 0$. Alors (3) devient $x = 0$. L'ensemble des solutions de (S_0) est donc le sous-espace nul $\{(0, 0, 0)\}$.
2ème cas : $\alpha = 1$ et $\beta \neq 1$
$(1')$ a pour seule solution $y = 0$ et tout $z \in \mathbb{R}$ est solution de $(2')$. Alors (3) devient $x + z = 0$. L'ensemble des solutions de (S_0) est le sous-espace $F = \{(-z, 0, z)\,;\ z \in \mathbb{R}\}$; c'est donc la droite vectorielle dont un vecteur de base est $u = (-1, 0, 1)$.
3ème cas : $\alpha \neq 1$ et $\beta = 1$
$(2')$ a pour seule solution $z = 0$ et tout $y \in \mathbb{R}$ est solution de $(1')$. Alors (3) devient $x + y = 0$. L'ensemble des solutions de (S_0) est le sous-espace $F' = \{(-y, y, 0); y \in \mathbb{R}\}$; c'est donc la droite vectorielle dont un vecteur de base est $u' = (-1, 1, 0)$.
4ème cas : $\alpha = 1$ et $\beta = 1$
Tout $y \in \mathbb{R}$ est solution de $(1')$ et tout $z \in \mathbb{R}$ est solution de $(2')$. Le système équivaut à la seule équation $x + y + z = 0$. L'ensemble des solutions de (S_0) est le sous-espace $H = \{(-y - z, y, z)\,;\ y, z \in \mathbb{R}\}$; c'est donc le plan vectoriel dont une base est (u, u').

2.4 Exercices sur le chapitre 2

Exercice 1.
Soit $f : \mathbb{R}^3 \to \mathbb{R}^3$ définie par $f(x,y,z) = (3x+3z,\ -2x+y-4z,\ x+3y-5z)$.

a) Montrer que f est un endomorphisme de $\mathbb{R}^3$.

b) Calculer les images par f des vecteurs e_1, e_2, e_3 de la base canonique de $\mathbb{R}^3$.

c) Calculer l'image par f du vecteur $u = (-1,2,3)$.

d) Déterminer $\operatorname{Ker} f$ et $\operatorname{Im} f$; que peut-on en déduire pour f ?

Exercice 2.
Soit $f : \mathbb{R}^3 \to \mathbb{R}^3$ définie par $f(x,y,z) = (2x-y+3z,\ x+y-2z,\ 4x+y-z)$.

a) Montrer que f est un endomorphisme de $\mathbb{R}^3$.

b) Déterminer $\operatorname{Ker} f$ et $\operatorname{Im} f$; que peut-on en déduire pour f ?

Exercice 3.
Pour chacune des applications linéaires suivantes, indiquer si elle est injective, surjective, bijective.

a) $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(x,y) = (2x+y,\ x-3y)$.

b) $f : \mathbb{R}^2 \to \mathbb{R}^3$ définie par $f(x,y) = (2y,\ x+y,\ -x+y)$.

c) $f : \mathbb{R}^3 \to \mathbb{R}^3$ définie par $f(x,y,z) = (y+z,\ x+y-z,\ -x+y)$.

Exercice 4.
Soient $\mathcal{B} = (e_1,e_2,e_3,e_4)$ et $\mathcal{C} = (u_1,u_2,u_3,u_4,u_5,u_6)$ les bases canoniques de $\mathbb{R}^4$ et $\mathbb{R}^6$ respectivement. Soit f l'application linéaire de $\mathbb{R}^4$ dans $\mathbb{R}^6$ définie par :

$f(e_1) = -u_1-u_2-u_3,\ f(e_2) = u_1-u_4-u_5,\ f(e_3) = u_2+u_4-u_6,\ f(e_4) = u_3+u_5+u_6.$

a) Montrer que $(u_1,u_2,u_4,f(e_1),f(e_2),f(e_3))$ est une base de $\mathbb{R}^6$.

b) Même question pour $(u_1,u_3,u_4,f(e_1),f(e_2),f(e_3))$.

c) Déterminer une base de $\operatorname{Im} f$; donner deux supplémentaires de $\operatorname{Im} f$ dans $\mathbb{R}^6$.

d) Déterminer une base de $\operatorname{Ker} f$; donner un supplémentaire de $\operatorname{Ker} f$ dans $\mathbb{R}^4$.

Exercice 5.
Soient E un $\mathbb{R}$-e.v. de dimension 2 et $\mathcal{B} = (u_1,u_2)$ une base de E. Soit f l'endomorphisme de E définie par $f(u_1) = 3.u_1 - u_2$ et $f(u_2) = u_1 + u_2$.

a) Pour tout vecteur $u \in E$ de composantes (x,y) dans la base $\mathcal{B}$, exprimer en fonction de x,y les composantes (x',y') du vecteur $f(u)$ dans la base $\mathcal{B}$.

b) Déterminer $\operatorname{Ker} f$: que peut-on en déduire ?

c) Montrer que $v_1 = (2,1)$ et $v_2 = (-1,1)$ forment une base $\mathcal{B}'$.

d) Pour tout vecteur $u \in E$ de composantes (a,b) dans la base $\mathcal{B}'$, exprimer en fonction de a,b les composantes (a',b') du vecteur $f(u)$ dans la base $\mathcal{B}'$.

Exercice 6.
Pour chacune des applications linéaires f suivantes, déterminer $\operatorname{Ker} f$ et $\operatorname{Im} f$, et indiquer si f est injective, surjective, bijective.

a) $f : \mathbb{R}^3 \to \mathbb{R}^3$ définie par $f(x,y,z) = (2x-y-z,\ x-2y+z,\ x+y-2z)$.

b) $f : \mathbb{R}^3 \to \mathbb{R}^3$ définie par $f(x,y,z) = (-x+y+z,\ x-y+z,\ x+y-z)$.

c) $f : \mathbb{R}^4 \to \mathbb{R}^3$ définie par $f(x,y,z,t) = (x-y+z+t,\ x+2z-t,\ x+y+3z-3t)$.

Exercice 7.

a) Déterminer l'unique forme linéaire $f : \mathbb{R}^2 \to \mathbb{R}$ vérifiant $f(1,1) = 3$ et $f(0,1) = -2$.

b) Déterminer toutes les formes linéaires $f : \mathbb{R}^3 \to \mathbb{R}$ vérifiant $f(1,1,0) = f(0,0,1) = 0$ et $f(1,0,1) \neq 0$.

c) Déterminer l'unique forme linéaire $f : \mathbb{R}^3 \to \mathbb{R}$ vérifiant $f(1,1,1) = -2$, $f(1,-1,1) = 2$ et $f(0, \frac{1}{2}, 1) = 1$. Donner une base du noyau de f et une base d'un supplémentaire de ce noyau.

Exercice 8.

Dans le $\mathbb{R}$-e.v. $\mathbb{R}^4$, on considère $H = \{(x,y,z,t) \in \mathbb{R}^4 \,;\, x + y + z + t = 0\}$.

a) Montrer que H est un hyperplan de $\mathbb{R}^4$ et déterminer toutes les formes linéaires sur $\mathbb{R}^4$ dont H est le noyau. Donner une base de H.

b) Montrer que la droite vectorielle D engendrée par $u = (1,1,1,1)$ est un supplémentaire de H dans $\mathbb{R}^4$.

Chapitre 3

Matrices

3.1 Calcul matriciel

3.1.1 Notion de matrice

a) Définition.
Soient n et p deux entiers ≥ 1.
Une *matrice* à p *lignes* et n *colonnes* à coefficients dans $\mathbb{K}$ est une application :

$$A : \{1, 2, \ldots, p\} \times \{1, 2, \ldots, n\} \to \mathbb{K}$$

qui, à tout couple (i, j), associe un nombre $a_{i,j}$ de $\mathbb{K}$.
On la note sous la forme d'un tableau à p lignes et n colonnes :

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \ldots & a_{1,n-1} & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \ldots & a_{2,n-1} & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \ldots & a_{3,n-1} & a_{3,n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{p,1} & a_{p,2} & a_{p,3} & \ldots & a_{p,n-1} & a_{p,n} \end{pmatrix},$$

ou encore de façon générique :

$$A = (a_{i,j})_{1 \leq i \leq p,\, 1 \leq j \leq n}.$$

Le premier indice est l'indice de ligne, le second l'indice de colonne.
Les termes $a_{i,i}$ sont appelés les termes diagonaux de A.

b) Notation et terminologie.
On note $\mathcal{M}_{p,n}(\mathbb{K})$ l'ensemble des matrices à p lignes et n colonnes à coefficients dans $\mathbb{K}$.
Pour $p = 1$, une matrice de $\mathcal{M}_{1,n}(\mathbb{K})$ s'appelle une *matrice ligne* (c'est un n-uplet).
Pour $n = 1$, une matrice de $\mathcal{M}_{p,1}(\mathbb{K})$ s'appelle une *matrice colonne* (c'est un p-uplet).
Pour $p = n$, une matrice de $\mathcal{M}_{n,n}(\mathbb{K})$ s'appelle une *matrice carrée d'ordre* n.

c) Exemples.

$$\begin{pmatrix} i & -1 \\ 0 & i \end{pmatrix} \in \mathcal{M}_{2,2}(\mathbb{C}), \qquad \begin{pmatrix} 2 & \sqrt{5} & -7 \\ 1 & -2 & -2 \\ 3 & 0 & 9 \\ 0 & 1 & \frac{1}{4} \end{pmatrix} \in \mathcal{M}_{4,3}(\mathbb{R}), \qquad \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} \in \mathcal{M}_{3,3}(\mathbb{K}),$$

$$\begin{pmatrix} 5 \\ 1 \\ 0 \end{pmatrix} \in \mathcal{M}_{3,1}(\mathbb{R}), \qquad \left(\sqrt{3}, -1, 0\right) \in \mathcal{M}_{1,3}(\mathbb{R}).$$

3.1.2 Addition et produit externe

a) Addition des matrices. On peut additionner terme à terme (ie. coefficient par coefficient) deux matrices ayant le même nombre p de lignes et le même nombre n de colonnes ; on obtient une matrice somme à p lignes et n colonnes :

$$\text{si } A = (a_{i,j})_{1\le i\le p,\,1\le j\le n} \text{ et si } B = (b_{i,j})_{1\le i\le p,\,1\le j\le n},$$
$$\text{alors } A + B = (a_{i,j} + b_{i,j})_{1\le i\le p,\,1\le j\le n}.$$

b) Produit externe d'une matrice par un scalaire. On peut multiplier une matrice A par un scalaire quelconque α en multipliant par α chaque coefficient de A (au sens du produit dans $\mathbb{K}$) ; on obtient une matrice du même nombre de lignes et de colonnes que A :

$$\text{si } A = (a_{i,j})_{1\le i\le p,\,1\le j\le n} \text{ et si } \alpha \in \mathbb{K},$$
$$\text{alors } \alpha.A = (\alpha a_{i,j})_{1\le i\le p,\,1\le j\le n}.$$

c) Matrice nulle. On appelle *matrice nulle* dans $\mathcal{M}_{p,n}(\mathbb{K})$ la matrice $O_{p,n}$ de $\mathcal{M}_{p,n}(\mathbb{K})$ dont tous les coefficients sont nuls. C'est bien sûr l'élément neutre de l'addition :

$$A + O_{p,n} = O_{p,n} + A = A \text{ pour toute } A \in \mathcal{M}_{p,n}(\mathbb{K}).$$

d) Opposée d'une matrice. Si $A = (a_{i,j})_{1\le i\le p,\,1\le j\le n}$ est une matrice à p lignes et n colonnes, on appelle *matrice opposée* de A, notée $-A$, la matrice à p lignes et n colonnes dont le coefficient sur la i-ième ligne et la j-ième colonne est l'opposé de $a_{i,j}$. En particulier :

$$-A = (-1).A \quad \text{et} \quad A + (-A) = (-A) + A = O_{p,n}.$$

On notera $A - B = A + (-B)$ pour deux matrices quelconques dans $\mathcal{M}_{p,n}(\mathbb{K})$.

e) Exemples.

$$3.\begin{pmatrix}5 & 1 & 2\\ 8 & 9 & 3\\ 1 & -6 & 1\end{pmatrix} - \begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix} = \begin{pmatrix}15 & 3 & 6\\ 24 & 27 & 9\\ 3 & -18 & 3\end{pmatrix} - \begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix} = \begin{pmatrix}14 & 3 & 6\\ 24 & 27 & 8\\ 3 & -19 & 3\end{pmatrix},$$

$$\begin{pmatrix}1\\3\\5\end{pmatrix} + 2.\begin{pmatrix}1\\0\\1\end{pmatrix} = \begin{pmatrix}1\\3\\5\end{pmatrix} + \begin{pmatrix}2\\0\\2\end{pmatrix} = \begin{pmatrix}3\\3\\7\end{pmatrix}$$

f) Règles de calcul. Il est facile de vérifier que l'addition et le produit externe vérifient les propriétés notées (1) à (8) en 1.1.2.c et en 1.1.3.c, c'est-à-dire que l'ensemble $E = \mathcal{M}_{p,n}(\mathbb{K})$ est un espace vectoriel sur $\mathbb{K}$, au sens défini en 1.5. On a en particulier les règles de calcul suivantes, pour toutes matrices A, B, C dans $\mathcal{M}_{p,n}(\mathbb{K})$ et tous scalaires $\alpha, \beta \in \mathbb{K}$:

$$\begin{array}{lll} A + B = B + A & A + (B + C) = (A + B) + C & A - B = -(B - A)\\ (\alpha + \beta).A = \alpha.A + \beta.A & \alpha.(A + B) = \alpha.A + \alpha.B & \alpha.(\beta.A) = (\alpha\beta).A\\ 1.A = A & (-\alpha).A = \alpha.(-A) = -\alpha.A & \alpha.(A - B) = \alpha.A - \alpha.B \end{array}.$$

3.1.3 Produit matriciel

La définition du produit de deux matrices peut sembler a priori un peu artificielle, mais son sens s'éclaircira plus loin en relation avec le produit de composition des applications linéaires.

a) Définition.
Remarquons avant tout que le produit d'une matrice A par une matrice B n'est défini que lorsque le nombre de colonnes de A est égal au nombre de lignes de B.

Soient donc : $A = (a_{i,j})_{1\le i\le p,\,1\le j\le n} \in \mathcal{M}_{p,n}(\mathbb{K})$ et $B = (b_{k,l})_{1\le k\le n,\,1\le l\le q} \in \mathcal{M}_{n,q}(\mathbb{K})$;

le produit $C = A \times B$, noté aussi $C = AB$, est la matrice $(c_{i,l})_{1\le i\le p,\,1\le l\le q} \in \mathcal{M}_{p,q}(\mathbb{K})$ dont le terme général $c_{i,l}$ est défini par :

$$c_{i,l} = a_{i,1}b_{1,l} + a_{i,2}b_{2,l} + \cdots + a_{i,n-1}b_{n-1,l} + a_{i,n}b_{n,l} = \sum_{k=1}^{n} a_{i,k}b_{k,l}.$$

Ainsi, pratiquement, le coefficient de AB sur la i-ième ligne et la l-ième colonne s'obtient en faisant la somme des n produits terme à terme de la i-ième ligne de A par la l-ième colonne de B.

$$\begin{pmatrix} * & * & \dots & * \\ * & * & \dots & * \\ \boxed{a_{i,1}} & \boxed{a_{i,2}} & \dots & \boxed{a_{i,n}} \\ * & * & \dots & * \\ * & * & \dots & * \end{pmatrix} \times \begin{pmatrix} * & * & \boxed{b_{1,l}} & * & * \\ * & * & \boxed{b_{2,l}} & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ * & * & \boxed{b_{n,l}} & * & * \end{pmatrix} = \begin{pmatrix} * & * & \vdots & * & * \\ * & * & \vdots & * & * \\ \dots & \dots & \boxed{c_{i,l}} & \dots & \dots \\ * & * & \vdots & * & * \\ * & * & \vdots & * & * \end{pmatrix}.$$

En résumé :

$$\underbrace{(a_{i,j})_{1\le i\le p,\,1\le j\le n}}_{\in\mathcal{M}_{p,n}(\mathbb{K})} \times \underbrace{(b_{k,l})_{1\le k\le n,\,1\le l\le q}}_{\in\mathcal{M}_{n,q}(\mathbb{K})} = \underbrace{\Big(\sum_{k=1}^{n} a_{i,k}b_{k,l}\Big)_{1\le i\le p,\,1\le l\le q}}_{\in\mathcal{M}_{p,q}(\mathbb{K})}.$$

b) <u>Exemples.</u>

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} \alpha & \beta & \gamma \\ \alpha' & \beta' & \gamma' \\ \alpha'' & \beta'' & \gamma'' \end{pmatrix} = \begin{pmatrix} \alpha + 2\alpha' + 3\alpha'' & \beta + 2\beta' + 3\beta'' & \gamma + 2\gamma' + 3\gamma'' \\ 4\alpha + 5\alpha' + 6\alpha'' & 4\beta + 5\beta' + 6\beta'' & 4\gamma + 5\gamma' + 6\gamma'' \end{pmatrix};$$

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} a+g & b+h & c+i \\ d & e & f \\ g & h & i \end{pmatrix}; \qquad \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b & a+c \\ d & e & d+f \\ g & h & g+i \end{pmatrix};$$

$$\begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} \begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} a+4b \\ c+4d \\ e+4f \end{pmatrix}; \qquad \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} = \begin{pmatrix} a+2c+3e & b+2d+3f \end{pmatrix}.$$

c) <u>Remarques.</u>

• Le produit AB peut exister sans que le produit BA soit défini ; et même si les deux existent, ils ne sont en général pas égaux (cf. les exemples 2 et 3 précédents).

• Le produit de matrices, *lorsqu'il est défini*, vérifie les propriétés d'associativité $[A(BC) = (AB)C]$ et de distributivité sur l'addition $[AB + AC = A(B + C)$ et $BA + CA = (B + C)A]$.

3.1.4 Algèbre des matrices carrées

a) <u>Notations.</u> Soit n un entier ≥ 1 fixé.
On note $\mathcal{M}_n(\mathbb{K})$ l'ensemble des matrices carrées d'ordre n à coefficients dans $\mathbb{K}$.
D'après la définition 3.1.3.a, pour toutes matrices carrées A et B d'ordre n, les produits AB et BA sont tous les deux définis ; donc le produit matriciel définit un produit interne (une multiplication) dans $\mathcal{M}_n(\mathbb{K})$.

b) <u>Matrice identité.</u>
On note I_n et on appelle matrice identité d'ordre n la matrice dont les coefficients sont égaux à 1 sur la diagonale et à 0 partout ailleurs. Donc :

$$I_n = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} = (\delta_{i,j})_{1\leq i,j\leq n}$$

avec la notation usuelle : $\delta_{i,j} = 1$ si $i = j$, et $\delta_{i,j} = 0$ si $i \neq j$.

On déduit immédiatement de 3.1.3.a que I_n est neutre pour le produit matriciel dans $\mathcal{M}_n(\mathbb{K})$:

$$\forall\ A \in \mathcal{M}_n(\mathbb{K}),\ A \times I_n = I_n \times A = A.$$

c) Matrice nulle.

On note simplement O_n la matrice nulle dans $\mathcal{M}_n(\mathbb{K})$; rappelons (cf. 3.1.2.d) que tous ses coefficients sont nuls, et qu'elle est donc élément neutre pour l'addition dans $\mathcal{M}_n(\mathbb{K})$:

$$\forall\ A \in \mathcal{M}_n(\mathbb{K}),\ A + O_n = O_n + A = A.$$

On déduit immédiatement de 3.1.3.a que le produit de n'importe quelle matrice $A \in \mathcal{M}_n(\mathbb{K})$ par la matrice nulle O_n est égale à la matrice nulle :

$$\forall\ A \in \mathcal{M}_n(\mathbb{K}),\ A \times O_n = O_n \times A = O_n.$$

Mais attention ! Le produit de deux matrices dans $\mathcal{M}_n(\mathbb{K})$ peut être nul sans qu'aucune des deux ne soit nulle. Par exemple :

$$\begin{pmatrix} 2 & 0 \\ 3 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{et pourtant } \begin{pmatrix} 2 & 0 \\ 3 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},\ \begin{pmatrix} 0 & 0 \\ 4 & 1 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

On traduit cette propriété en disant que $\mathcal{M}_n(\mathbb{K})$ n'est *pas intègre* pour le produit matriciel.

d) Structure de $\mathbb{K}$-algèbre.

On sait déjà que $\mathcal{M}_n(\mathbb{K})$ est un $\mathbb{K}$-e.v. (pour l'addition et le produit externe par un scalaire, cf. 3.1.2.f) ; de plus, on a vu que le produit matriciel (interne) dans $\mathcal{M}_n(\mathbb{K})$ vérifie les propriétés suivantes :

– $\forall\ A, B, C \in \mathcal{M}_n(\mathbb{K}),\ A(BC) = (AB)C$ (associativité),

– $\forall\ A \in \mathcal{M}_n(\mathbb{K}),\ A \times I_n = I_n \times A = A$ (élément neutre),

– $\forall\ A, B, C \in \mathcal{M}_n(\mathbb{K}),\ A(B + C) = AB + AC$ et $(A + B)C = AC + BC$,

– $\forall\ A, B \in \mathcal{M}_n(\mathbb{K}),\ \forall\ \alpha \in \mathbb{K},\ (\alpha.A)B = A(\alpha.B) = \alpha.(AB)$.

On résume toutes ces propriétés du produit et le fait que $\mathcal{M}_n(\mathbb{K})$ est un $\mathbb{K}$-e.v. (voir 3.1.2.f) en disant que : l'ensemble des matrices carrées $\mathcal{M}_n(\mathbb{K})$ est une $\mathbb{K}$-algèbre.

On verra plus loin en 3.2.3.c que cette structure d'algèbre est liée à celle de l'algèbre $\mathcal{L}(E)$ des endomorphismes de $E = \mathbb{K}^n$, vue en 2.1.6.

e) Mise en garde.

Il résulte de d) ci-dessus que plusieurs des règles de calcul sur les matrices carrées sont formellement analogues à celles dont on a l'habitude pour le calcul algébrique dans $\mathbb{R}$ ou $\mathbb{C}$ (factorisation, développement, simplification,...) mais on prendra garde que le produit des matrices présente deux particularités (qui dans la pratique changent beaucoup de choses dans les raisonnements...) :

• le produit dans $\mathcal{M}_n(\mathbb{K})$ n'est pas commutatif : on peut avoir $AB \neq BA$; [cf. 3.1.3.c].

• $\mathcal{M}_n(\mathbb{K})$ n'est pas intègre pour le produit matriciel : on peut avoir $AB = O_n$ avec $A \neq O_n$ et $B \neq O_n$; [cf. 3.1.4.c].

3.1.5 Matrices carrées inversibles, groupe linéaire

a) Définition.
Une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ est dite *inversible* lorsqu'il existe dans $\mathcal{M}_n(\mathbb{K})$ une matrice, notée A^{-1}, vérifiant : $AA^{-1} = A^{-1}A = I_n$.
On peut montrer qu'un tel inverse, lorsqu'il existe, est nécessairement unique ; on dira donc que A^{-1} est *la* matrice inverse de A.
On peut aussi montrer qu'il suffit que $AA^{-1} = I_n$ ou que $A^{-1}A = I_n$ pour conclure que A est inversible d'inverse A^{-1}.

b) Exemples.
• La matrice $A = \begin{pmatrix} 2 & 1 & 0 \\ 3 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ est inversible, d'inverse $A^{-1} = \begin{pmatrix} 3 & -2 & 1 \\ -5 & 4 & -2 \\ 1 & -1 & 1 \end{pmatrix}$. Il suffit de calculer le produit de ces deux matrices et de vérifier qu'il vaut I_3.

• En revanche, la matrice $B = \begin{pmatrix} 3 & 6 \\ 2 & 4 \end{pmatrix}$ n'est pas inversible dans $\mathcal{M}_2(\mathbb{K})$. En effet, sinon il existerait une matrice $\begin{pmatrix} x & y \\ z & t \end{pmatrix}$ telle que $\begin{pmatrix} 3 & 6 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. En identifiant la première colonne dans chaque membre, on aurait $3x + 6z = 1$ et $2x + 4z = 0$, ce qui est incompatible.

• On verra plus tard des méthodes concrètes d'une part pour déterminer si une matrice carrée donnée est ou non inversible, d'autre part pour calculer son inverse lorsqu'il existe.

c) Groupe linéaire.
Le sous-ensemble de $\mathcal{M}_n(\mathbb{K})$ formé des matrices inversibles est noté $GL(n, \mathbb{K})$ ou $GL_n(\mathbb{K})$.
On a alors la proposition suivante :

(i) $I_n \in GL(n, \mathbb{K})$ et $I_n^{-1} = I_n$,

(ii) $\forall\, A, B \in GL(n, \mathbb{K}),\ AB \in GL(n, \mathbb{K})$ et $(AB)^{-1} = B^{-1}A^{-1}$, (← *attention à l'ordre*)

(iii) $\forall\, A \in GL(n, \mathbb{K}),\ A^{-1} \in GL(n, \mathbb{K})$ et $(A^{-1})^{-1} = A$.

Preuve : les points (i) et (iii) sont clairs. Pour le (ii), calculons : $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$; ceci prouve que $(AB)^{-1} = B^{-1}A^{-1}$. □

Les 3 assertions ci-dessus et l'associativité du produit matriciel montrent que $GL(n, \mathbb{K})$ est ce qu'on l'on appelle un groupe, dit *groupe linéaire des matrices carrées inversibles d'ordre* n.

Il est bien sûr intimement lié au groupe linéaire $GL(E)$ des automorphismes du $\mathbb{K}$-e.v. $E = \mathbb{K}^n$ défini en 2.1.6.d, ceci via la correspondance entre matrices carrées et endomorphismes de E déjà évoquée en 3.1.4.d, comme on le verra en détail en 3.2.3.c.

3.1.6 Matrices équivalentes, matrices carrées semblables

a) Définition.
Deux matrices A et B de $\mathcal{M}_{p,n}(\mathbb{K})$ sont dites *équivalentes* lorsqu'il existe deux matrices carrées inversibles $P \in GL(p, \mathbb{K})$ et $Q \in GL(n, \mathbb{K})$ vérifiant $A = PBQ$.

b) Définition.
Deux matrices carrées A et B de $\mathcal{M}_n(\mathbb{K})$ sont dites *semblables* lorsqu'il existe une matrice carrée inversible $P \in GL(n, \mathbb{K})$ vérifiant $A = PBP^{-1}$.

On notera $A \sim B$ pour signifier que A est semblable à B.

c) Proposition.

La relation "être semblable" est une relation d'équivalence dans $\mathcal{M}_n(\mathbb{K})$, ce qui signifie, que :

– $\forall\ A \in \mathcal{M}_n(\mathbb{K}),\ A \sim A$ [réflexivité]

– $\forall\ A, B \in \mathcal{M}_n(\mathbb{K}),\ \big[A \sim B \Rightarrow B \sim A\big]$ [symétrie]

– $\forall\ A, B, C \in \mathcal{M}_n(\mathbb{K}),\ \big[(A \sim B)$ et $(B \sim C)\big] \Rightarrow (A \sim C)$ [transitivité]

Preuve : c'est un simple exercice reposant sur la proposition 3.1.5.c ; à faire ! □

d) Remarques.

• La relation définie en a) est aussi une relation d'équivalence. Le sens profond de ces deux relations sera mis en évidence dans les chapitres suivants.

• L'observation suivante, bien qu'élémentaire, aura plus loin d'importantes applications :

$$\forall\ A, B \in \mathcal{M}_n(\mathbb{K}),\ \forall\ P \in GL(n, \mathbb{K}),\quad \big[(A = PBP^{-1}) \Rightarrow (\forall\ m \in \mathbb{N},\ A^m = PB^mP^{-1}\big].$$

En effet : $A^m = (PBP^{-1})^m = (PBP^{-1})(PBP^{-1})(PBP^{-1}) \dots (PBP^{-1})$
$= PB(P^{-1}P)B(P^{-1}P)B(P^{-1} \dots P)BP^{-1} = PBI_nBI_nB \dots BP^{-1} = PB^mP^{-1}$.

3.1.7 Transposition

a) Définition.

Pour toute matrice $A \in \mathcal{M}_{p,n}(\mathbb{K})$, on appelle *transposée* de A la matrice ${}^tA \in \mathcal{M}_{n,p}(\mathbb{K})$ obtenue en échangeant les lignes et les colonnes de A. En d'autres termes :

si $A = (a_{i,j})_{1 \le i \le p,\, 1 \le j \le n} \in \mathcal{M}_{p,n}(\mathbb{K})$, alors ${}^tA = (a'_{i,j})_{1 \le i \le n,\, 1 \le j \le p} \in \mathcal{M}_{p,n}(\mathbb{K})$
est définie par $a'_{i,j} = a_{j,i}$ pour tous $1 \le i \le n$, $1 \le j \le p$.

b) Exemples. ${}^t\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$; ${}^t(\alpha\ \beta\ \gamma\ \delta) = \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix}$; ${}^t\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{pmatrix}$.

c) Proposition. *On a les règles de calcul suivantes sur la transposition :*

• $\forall\ A \in \mathcal{M}_{p,n}(\mathbb{K}),\ {}^t({}^tA) = A$;

• $\forall\ A, B \in \mathcal{M}_{p,n}(\mathbb{K}),\ \forall\ \alpha \in \mathbb{K},\ {}^t(A + B) = {}^tA + {}^tB$ et ${}^t(\alpha.A) = \alpha.{}^tA$;

• $\forall\ A \in \mathcal{M}_{p,n}(\mathbb{K}),\ \forall\ B \in \mathcal{M}_{n,q}(\mathbb{K}),\ {}^t(AB) = {}^tB{}^tA$ (*attention à l'ordre*) ;

• $\forall\ A \in GL(n, \mathbb{K}),\ {}^tA \in GL(n, \mathbb{K})$ et $({}^tA)^{-1} = {}^t(A^{-1})$.

Preuve : c'est un exercice technique ; à faire ! □

3.1.8 Quelques types particuliers de matrices carrées

Soit $A = (a_{i,j})_{1 \le i,j \le n} \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée quelconque.

a) A est dite *diagonale* lorsque tous les termes non diagonaux sont nuls (ie. $a_{i,j} = 0$ pour tous $1 \le i \ne j \le n$).

Il est clair que, si A et B sont deux matrices diagonales, le produit AB est diagonal, et égal au produit BA.

$$AB = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & a_n \end{pmatrix} \begin{pmatrix} b_1 & 0 & \dots & 0 \\ 0 & b_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & b_n \end{pmatrix} = \begin{pmatrix} a_1b_1 & 0 & \dots & 0 \\ 0 & a_2b_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & a_nb_n \end{pmatrix} = BA.$$

b) A est dite *scalaire* lorsque elle est diagonale et que de plus tous les termes diagonaux sont égaux, (ie. A est de la forme $\alpha.I_n$). Elle commute alors avec toute $B \in \mathcal{M}_n(\mathbb{K})$.

c) A est dite *triangulaire supérieure* lorsque tous les termes en-dessous de la diagonale sont nuls (c'est-à-dire $a_{i,j} = 0$ pour tous $1 \leq j < i \leq n$). Par exemple, $A = \begin{pmatrix} 3 & 4 & -7 & 9 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & -1 \end{pmatrix}$.

Le produit de deux matrices triangulaires supérieures est encore triangulaire supérieure.

d) A est dite *symétrique* lorsque ${}^tA = A$; A est dite *antisymétrique* lorsque ${}^tA = -A$.

Par exemple : $\begin{pmatrix} 3 & 4 & -7 & 9 \\ 4 & 1 & 2 & 0 \\ -7 & 2 & 5 & 8 \\ 9 & 0 & 8 & -1 \end{pmatrix}$ est symétrique, et $\begin{pmatrix} 0 & 4 & -7 & 9 \\ -4 & 0 & 2 & -5 \\ 7 & -2 & 0 & 8 \\ -9 & 5 & -8 & 0 \end{pmatrix}$ est antisymétrique.

3.1.9 Trace d'une matrice carré

a) Définition. La trace d'une matrice carrée A, notée $\operatorname{tr} A$, est la somme de ses coefficients diagonaux. Donc :

$$\text{si } A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{K}), \quad \operatorname{tr} A = a_{1,1} + a_{2,2} + \cdots + a_{n,n} = \sum_{i=1}^{n} a_{i,i} \in \mathbb{K}.$$

b) Proposition. *On a les propriétés suivantes de la trace :*

(i) $\forall\, A \in \mathcal{M}_n(\mathbb{K}),\ \operatorname{tr}({}^tA) = \operatorname{tr} A$;

(ii) $\forall\, A, B \in \mathcal{M}_n(\mathbb{K}),\ \forall\, \alpha \in \mathbb{K},\ \operatorname{tr}(A + B) = \operatorname{tr} A + \operatorname{tr} B$ et $\operatorname{tr}(\alpha.A) = \alpha \operatorname{tr} A$.

(iii) $\forall\, A, B \in \mathcal{M}_n(\mathbb{K}),\ \operatorname{tr}(AB) = \operatorname{tr}(BA)$.

(iv) $\forall\, A \in \mathcal{M}_n(\mathbb{K}),\ \forall\, P \in GL(n, \mathbb{K}),\ \operatorname{tr}(PAP^{-1}) = \operatorname{tr}(A)$.

Preuve : les points (i) et (ii) sont évidents. Le point (iii) ne repose que la formule définissant le produit matriciel et la définition de la trace : écrire le calcul explicite en exercice. Le point (iv) découle alors de (iii) ; il suffit d'écrire que $\operatorname{tr}(PAP^{-1}) = \operatorname{tr}[(PA)P^{-1}] = \operatorname{tr}[P^{-1}(PA)] = \operatorname{tr}[P^{-1}PA] = \operatorname{tr} A$. □

c) Remarque. Le point (iv) signifie que deux matrices carrées semblables ont toujours la même trace. Ainsi la trace est le premier "invariant" que l'on rencontre (on en verra d'autres : le rang, le déterminant, le polynôme caractéristique..), qui caractérise non pas une matrice carrée mais la classe de toutes les matrices carrées qui lui sont semblables. (Il en résulte, cf. section 3.2 suivante, que c'est une notion qui est en fait attachée à un endomorphisme). La trace joue en particulier un rôle important dans l'étude des valeurs propres, comme on le verra plus loin dans le cours.

d) Remarque. Attention de ne pas faire dire aux points (iii) et (iv) plus que ce qu'ils disent : on n'a pas en général $\operatorname{tr}(AB) = \operatorname{tr}(A) \operatorname{tr}(B)$.

3.2 Matrices et applications linéaires

3.2.1 Matrice d'une application linéaire par rapport à des bases

a) Exemple préliminaire.
Soit $\mathcal{B} = (e_1, e_2)$ la base canonique de $\mathbb{K}^2$, avec donc $e_1 = (1, 0)$ et $e_2 = (0, 1)$, et soit $\mathcal{C} = (v_1, v_2, v_3)$ la base canonique de $\mathbb{K}^3$, avec $v_1 = (1, 0, 0)$, $v_2 = (0, 1, 0)$ et $v_3 = (0, 0, 1)$.

Soit f une application linéaire $\mathbb{K}^2 \to \mathbb{K}^3$. D'après 2.2.1.a, f est entièrement déterminée par les images de e_1 et e_2. Choisissons par exemple $f(e_1) = (2,1,0) = 2.v_1 + v_2$ et $f(e_2) = (-1,2,3) = -v_1 + 2.v_2 + 3.v_3$. Pour tout vecteur $(x,y) \in \mathbb{K}^2$, on peut alors calculer $f(x,y) = f(x.e_1 + y.e_2) = x.f(e_1) + y.f(e_2) = x.(2,1,0) + y.(-1,2,3) = (2x-y, x+2y, 3y)$.

On peut donc écrire matriciellement : $$\begin{pmatrix} 2x-y \\ x+2y \\ 3y \end{pmatrix} = \underbrace{\begin{pmatrix} 2 & -1 \\ 1 & 2 \\ 0 & 3 \end{pmatrix}}_{\text{soit } A} \begin{pmatrix} x \\ y \end{pmatrix}.$$

La matrice A s'appelle la matrice de f par rapport aux bases $\mathcal{B}$ et $\mathcal{C}$. Elle définit complètement l'application f puisqu'elle permet de calculer l'image par f de tout vecteur de $\mathbb{K}^2$. Les deux colonnes de A sont formées des composantes dans la base $\mathcal{C}$ des vecteurs $f(e_1)$ et $f(e_2)$.

b) Définition.
Soient $\mathcal{B} = (e_1, \dots, e_n)$ une base de $E = \mathbb{K}^n$ et $\mathcal{C} = (v_1, \dots, v_p)$ une base de $E' = \mathbb{K}^p$.
Soit $f \in \mathcal{L}(E, E')$ une application linéaire de E dans E'.
On appelle matrice de f par rapport aux bases $\mathcal{B}$ et $\mathcal{C}$, notée $M_{\mathcal{BC}}(f)$, la matrice à p lignes et n colonnes dont les colonnes sont formées des composantes des vecteurs $f(e_1), \dots, f(e_n)$ décomposés dans la base $\mathcal{C}$.

$$M_{\mathcal{BC}}(f) = \begin{array}{c} \begin{pmatrix} * & * & \dots & * \\ * & * & \dots & * \\ \vdots & \vdots & & \vdots \\ * & * & \dots & * \end{pmatrix} \\ \begin{array}{cccc} \uparrow & \uparrow & & \uparrow \\ f(e_1) & f(e_2) & \dots & f(e_n) \end{array} \end{array} \begin{array}{l} \leftarrow v_1 \\ \leftarrow v_2 \\ \quad\vdots \\ \leftarrow v_p \\ \\ \\ \end{array}$$

Dans le cas particulier où $E = E'$ et $\mathcal{B} = \mathcal{C}$, on note simplement $M_{\mathcal{B}}(f)$ pour $M_{\mathcal{BB}}(f)$.

c) Proposition.
On reprend les données et notations précédentes ; on note $A = M_{\mathcal{BC}}(f)$. Soit u un vecteur quelconque de E. Si l'on note X la matrice colonne des n composantes de u dans la base $\mathcal{B}$, alors la matrice colonne Y des p composantes de son image $f(u)$ dans la base $\mathcal{C}$ est donnée par le produit matriciel :
$$Y = AX.$$

Preuve : il suffit de rédiger dans le cadre général les calculs détaillés sur l'exemple a) ci-dessus ; cf. ouvrage de référence. □

d) Exemple. Soient $E = \mathbb{K}^4$ et $E' = \mathbb{K}^3$, $\mathcal{B} = (e_1, e_2, e_3, e_4)$ une base de E et $\mathcal{C} = (v_1, v_2, v_3)$ une base de E'. Soit $f \in \mathcal{L}(E, E')$ définie par :
$$f(e_1) = v_1 + v_2 + v_3,\ f(e_2) = 2.v_1 - v_2,\ \ f(e_3) = 4.v_1 + v_3,\ \ f(e_4) = -v_1 + v_2 - 2.v_3.$$
Donc :
$$A = M_{\mathcal{BC}}(f) = \begin{pmatrix} 1 & 2 & 4 & -1 \\ 1 & -1 & 0 & 1 \\ 1 & 0 & 1 & -2 \end{pmatrix}.$$

De plus, si $u \in E$ a pour composantes (x,y,z,t) dans la base $\mathcal{B}$, les composantes (x',y',z') de $f(u)$ dans la base $\mathcal{C}$ sont :
$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 2 & 4 & -1 \\ 1 & -1 & 0 & 1 \\ 1 & 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \quad \text{donc} \quad \begin{cases} x' = x + 2y + 4z - t \\ y' = x - y + t \\ z' = x + z - 2t \end{cases}$$

e) Exemple. Soit f l'endomorphisme de $\mathbb{K}^3$ qui, à tout vecteur $u = (x,y,z)$ associe $f(u) = (x',y',z')$ défini par :
$$x' = x - y + 2z,\ \ y' = -x + y + z,\ \ z' = 3x + y - z.$$

Alors, en notant $\mathcal{B} = (e_1, e_2, e_3)$ la base canonique de $\mathbb{K}^3$, on a :

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & 1 \\ 3 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \text{donc} \quad M_{\mathcal{B}}(f) = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & 1 \\ 3 & 1 & -1 \end{pmatrix},$$

et donc : $\quad f(e_1) = e_1 - e_2 + 3e_3, \;\; f(e_2) = -e_1 + e_2 + e_3, \;\; f(e_3) = 2e_1 + e_2 - e_3.$

3.2.2 Propriétés de la correspondance entre matrices et applications linéaires

a) Remarque fondamentale.

(i) Pour une application linéaire donnée $f \in \mathcal{L}(E, E')$, sa matrice $A = M_{\mathcal{BC}}(f)$ par rapport à des $\mathcal{B}$ de E et $\mathcal{C}$ de E' contient toutes les informations définissant f :

1. d'une part A donne explicitement les images par f des vecteurs de $\mathcal{B}$, (ce qui d'après 2.2.1.a suffit à déterminer f) ;
2. d'autre part A permet d'après la proposition 3.2.1.c de calculer directement l'image par f de tout vecteur u de E (en donnant les composantes de $f(u)$ dans la base $\mathcal{C}$ en fonction des composantes de u dans la base $\mathcal{B}$).

(ii) Réciproquement, si $A \in \mathcal{M}_{p,n}(\mathbb{K})$ est une matrice donnée à p lignes et n colonnes, elle peut toujours être considérée comme la matrice d'une certaine application linéaire de $\mathbb{K}^n$ dans $\mathbb{K}^p$ par rapport à des bases données.

(iii) En particulier, pour toute matrice $A \in \mathcal{M}_{p,n}(\mathbb{K})$ il existe une unique application linéaire f de $\mathbb{K}^n$ dans $\mathbb{K}^p$ telle que $A = M_{\mathcal{BC}}(f)$ pour la base canonique $\mathcal{B}$ de $\mathbb{K}^n$ et la base canonique $\mathcal{C}$ de $\mathbb{K}^p$. Cette application f est celle qui, à tout vecteur $u = (x_1, \ldots, x_n) \in \mathbb{K}^n$, associe le vecteur $f(u) = (y_1, \ldots, y_p) \in \mathbb{K}^p$ défini par :

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{p,1} & a_{p,2} & \ldots & a_{p,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n \\ \vdots \\ a_{p,1}x_1 + a_{p,2}x_2 + \cdots + a_{p,n}x_n \end{pmatrix}.$$

b) Proposition.

Soient $\mathcal{B} = (e_1, \ldots, e_n)$ une base de $E = \mathbb{K}^n$ et $\mathcal{C} = (v_1, \ldots, v_p)$ une base de $E' = \mathbb{K}^p$. Quelles que soient f et g des applications linéaires de E dans E', et pour tout scalaire $\alpha \in \mathbb{K}$, on a dans $\mathcal{M}_{p,n}(\mathbb{K})$:

$$M_{\mathcal{BC}}(f + g) = M_{\mathcal{BC}}(f) + M_{\mathcal{BC}}(g) \quad \text{et} \quad M_{\mathcal{BC}}(\alpha.f) = \alpha.M_{\mathcal{BC}}(f).$$

Preuve : découle de la définition des lois + et . dans $\mathcal{M}_{p,n}(\mathbb{K})$ vue en 3.1.2, de la définition des lois + et . dans $\mathcal{L}(E, E')$ vue en 2.2.1.b, et de la définition 3.2.1.b qui permet de passer de l'une à l'autre. Pour plus de détails, cf. ouvrage de référence. $\square$

c) Théorème.

Soient $E = \mathbb{K}^n$ et $\mathcal{B}$ une base de E. Soient $E' = \mathbb{K}^p$ et $\mathcal{C}$ une base de E'. Soient $E'' = \mathbb{K}^q$ et $\mathcal{D}$ une base de E''. Quelles que soient f une application linéaire de E dans E', et g une application linéaire de E' dans E'', on a dans $\mathcal{M}_{q,n}(\mathbb{K})$:

$$\underbrace{M_{\mathcal{BD}}(g \circ f)}_{\in \mathcal{M}_{q,n}(\mathbb{K})} = \underbrace{M_{\mathcal{CD}}(g)}_{\in \mathcal{M}_{q,p}(\mathbb{K})} \times \underbrace{M_{\mathcal{BC}}(f)}_{\in \mathcal{M}_{p,n}(\mathbb{K})}.$$

Preuve : découle de la définition du produit matriciel vue en 3.1.3.a, de la définition de la loi $\circ$ pour les applications linéaires vue en 2.2.1.c, et de la définition 3.2.1.b qui permet de passer de l'une à l'autre. Pour plus de détails, cf. ouvrage de référence. $\square$

3.2.3 Cas particulier des matrices carrées et des endomorphismes

a) Proposition.

Soient $E = \mathbb{K}^n$ et $\mathcal{B}$ une base de E. On a :

(1) $\forall\ f, g \in \mathcal{L}(E),\ \forall\ \alpha \in \mathbb{K},\ M_{\mathcal{B}}(f+g) = M_{\mathcal{B}}(f) + M_{\mathcal{B}}(g)$ *et* $M_{\mathcal{B}}(\alpha.f) = \alpha.M_{\mathcal{B}}(f)$;

(2) $\mathrm{Mat}_{\mathcal{B}}(O) = O_n$;

(3) $\forall\ f, g \in \mathcal{L}(E),\ M_{\mathcal{B}}(g \circ f) = M_{\mathcal{B}}(g) \times M_{\mathcal{B}}(f)$;

(4) $\mathrm{Mat}_{\mathcal{B}}(\mathrm{id}_E) = I_n$.

Preuve : (1) (2) (3) sont des cas particuliers de 3.2.2.b et 3.2.2.c pour $n = p = q$ et $\mathcal{B} = \mathcal{C} = \mathcal{D}$. Le (4) est clair par définition même de $\mathrm{Mat}_{\mathcal{B}}(f)$ pour $f = \mathrm{id}_E$. □

b) Théorème.

Soit $E = \mathbb{K}^n$. Pour tout endomorphisme f de E, les conditions suivantes sont équivalentes :

(i) $f \in GL(E)$; (ie. f est bijective)

(ii) *il existe une base $\mathcal{B}$ de E telle que* $M_{\mathcal{B}}(f) \in GL(n, \mathbb{K})$; (ie. $M_{\mathcal{B}}(f)$ est inversible)

(iii) *pour toutes bases $\mathcal{B}$ et $\mathcal{C}$ de E, on a* $M_{\mathcal{BC}}(f) \in GL(n, \mathbb{K})$.

En outre, dans ce cas, on a dans $GL(n, \mathbb{K})$ les égalités :

$$\big(M_{\mathcal{BC}}(f)\big)^{-1} = M_{\mathcal{CB}}(f^{-1}), \quad \text{et en particulier :} \quad \big(M_{\mathcal{B}}(f)\big)^{-1} = M_{\mathcal{B}}(f^{-1}).$$

Preuve : il est clair que (iii) implique (ii). Les deux autres implications découlent directement du théorème 3.2.2.c avec $n = p = q$ et $\mathcal{B} = \mathcal{C} = \mathcal{D}$, et du point (4) de la proposition précédente. □

c) Interprétation.

Soient $E = \mathbb{K}^n$ et $\mathcal{B}$ une base *fixée* de E.

Considérons l'application Φ de $\mathcal{L}(E)$ dans $\mathcal{M}_n(\mathbb{K})$ qui, à tout endomorphisme f associe sa matrice $M_{\mathcal{B}}(f)$ par rapport à $\mathcal{B}$.

1. On sait que, pour l'addition et le produit externe, $\mathcal{L}(E)$ et $\mathcal{M}_n(\mathbb{K})$ sont des $\mathbb{K}$-e.v. comme on l'a vu en 2.1.6.a et en 3.1.2.f. On traduit alors le point (1) de la proposition 3.2.3.a en disant que l'application Φ est linéaire.

2. On sait aussi que le $\mathbb{K}$-e.v. $\mathcal{L}(E)$ avec de plus la loi $\circ$ est une $\mathbb{K}$-algèbre (cf. 2.1.6.c), et que le $\mathbb{K}$-e.v. $\mathcal{M}_n(\mathbb{K})$ avec de plus le produit matriciel est aussi une $\mathbb{K}$-algèbre (cf. 3.1.4.d). On traduit alors la proposition 3.2.3.a en disant que l'application Φ est un morphisme d'algèbres.

3. Comme il résulte clairement des remarques 3.2.2.a que Φ est bijective, on dit que Φ est un isomorphisme d'algèbres ; ceci se traduit concrètement par le fait que toute notion ou propriété relative aux endomorphismes de $\mathbb{K}^n$ a sa traduction directe en terme de matrices dans $\mathcal{M}_n(\mathbb{K})$, et réciproquement.

4. En particulier, le théorème 3.2.3.b s'interprète en disant que la restriction de Φ aux automorphismes de E est un isomorphisme de groupes du groupe linéaire $GL(E)$ des automorphismes de E sur le groupe linéaire $GL(n, \mathbb{K})$ des matrices inversibles (d'où le même vocabulaire).

Observons enfin que la correspondance Φ ci-dessus est définie en travaillant avec une base fixée. La question de savoir quelles matrices sont associées à un endomorphisme suivant le choix de la base avec laquelle on travaille est de toute autre nature, directement liée à la notion de matrices semblables définie en 3.1.6.b, et sera traitée à la prochaine section 3.3.

3.2.4 Application au calcul de l'inverse d'une matrice carrée inversible

a) Principe de la méthode.
Soit $A = (a_{i,j})_{1\leq i,j\leq n}$ une matrice carrée dans $\mathcal{M}_n(\mathbb{K})$.
D'après 3.2.2.a.(iii), il existe un endomorphisme unique $f \in \mathcal{L}(E)$ tel que $A = M_{\mathcal{B}}(f)$ pour $E = \mathbb{K}^n$ et $\mathcal{B}$ la base canonique de E. Cet endomorphisme f est celui qui, à tout vecteur $u = (x_1, \ldots, x_n) \in \mathbb{K}^n$, associe le vecteur $f(u) = (y_1, \ldots, y_n) \in \mathbb{K}^n$ défini par :

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \ldots & a_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n \\ \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n \end{pmatrix}.$$

Supposons que l'on parvienne à montrer qu'il existe des scalaires $(b_{i,j})_{1\leq i,j\leq n}$ tels que le système

$$\begin{cases} y_1 = a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n \\ y_2 = a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n \\ \cdots \\ y_n = a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n \end{cases}$$

soit équivalent à

$$\begin{cases} x_1 = b_{1,1}y_1 + b_{1,2}y_2 + \cdots + b_{1,n}y_n \\ x_2 = b_{2,1}y_1 + b_{2,2}y_2 + \cdots + b_{2,n}y_n \\ \cdots \\ x_n = b_{n,1}y_1 + b_{n,2}y_2 + \cdots + b_{n,n}y_n \end{cases}.$$

On pose alors $B = (b_{i,j})_{1\leq i,j\leq n} \in \mathcal{M}_n(\mathbb{K})$,
on appelle g l'endomorphisme de $\mathbb{K}^n$ tel que $B = M_{\mathcal{B}}(g)$,
de sorte que : $\quad \forall\, X, Y \in \mathcal{M}_{n,1}(\mathbb{K}),\ [Y = AX \Leftrightarrow X = BY]$,
ou encore : $\quad \forall\, u, v \in \mathbb{K}^n,\ [v = f(u) \Leftrightarrow u = g(v)]$.
On conclut que $f \in GL(\mathbb{K}^n)$ et $g = f^{-1}$,
et donc, d'après 3.2.3.b : $\quad A \in GL(n, \mathbb{K})$ et $B = A^{-1}$.

b) Exemple.
Soit $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$. L'endomorphisme f de $\mathbb{R}^3$ dont A est la matrice par rapport à la base canonique est celui qui, à un triplet quelconque $u = (x_1, x_2, x_3)$, associe $f(u) = (y_1, y_2, y_3)$ défini par $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. On a :

$$\begin{cases} y_1 = x_1 + 2x_2 \\ y_2 = x_2 + x_3 \\ y_3 = -x_1 + x_3 \end{cases} \Leftrightarrow \begin{cases} y_1 = x_1 + 2x_2 \\ y_2 = x_2 + x_3 \\ y_3 + y_1 = 2x_2 + x_3 \end{cases} \Leftrightarrow \begin{cases} y_1 = x_1 + 2x_2 \\ y_2 = x_2 + x_3 \\ y_3 + y_1 - 2y_2 = -x_3 \end{cases}$$

qui est encore équivalent, en remontant à partir de la dernière équation à :

$$\begin{cases} x_3 = -y_1 + 2y_2 - y_3 \\ x_2 = y_2 - x_3 \\ x_1 = y_1 - 2x_2 \end{cases}, \quad \text{c'est-à-dire :} \begin{cases} x_3 = -y_1 + 2y_2 - y_3 \\ x_2 = y_1 - y_2 + y_3 \\ x_1 = -y_1 + 2y_2 - 2y_3 \end{cases},$$

que l'on écrit :

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 & 2 & -2 \\ 1 & -1 & 1 \\ -1 & 2 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

On conclut que $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}$ est inversible et que $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & 2 & -2 \\ 1 & -1 & 1 \\ -1 & 2 & -1 \end{pmatrix}$.

c) Remarques.
Dans la pratique, il est prudent (bien que non nécessaire sur le plan logique) de vérifier le calcul en s'assurant que le produit de la matrice A de départ par la matrice A^{-1} que l'on vient de trouver est bien égal à la matrice identité.

On verra plus loin dans le cours d'une part une méthode permettant de tester si une matrice carrée est ou non inversible (par le déterminant), d'autre part une méthode beaucoup plus conceptuelle (utilisant aussi les déterminants) donnant une formule générale pour l'inverse (lorsqu'il existe). Cependant la méthode détaillée au a) du calcul de l'inverse d'une matrice carrée inversible par inversion formelle d'un système d'équations linéaires, (qui consiste finalement à montrer qu'une matrice est inversible en trouvant son inverse), reste dans bien des cas concrets la plus rapide et efficace dans la pratique.

3.3 Changements de base

3.3.1 Matrice d'une famille finie de vecteurs dans une base

a) Définition.
Soient $E = \mathbb{K}^n$ et $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ une base de E. Soit $\mathcal{C} = (v_1, v_2, \ldots, v_p)$ une famille finie de p vecteurs de E. On appelle matrice de la famille $\mathcal{C}$ dans la base $\mathcal{B}$ la matrice à n lignes et p colonnes dont les colonnes sont constituées des composantes des vecteurs de $\mathcal{C}$ décomposés dans la base $\mathcal{B}$. C'est donc la matrice :

$$A = \begin{array}{cc} \begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,p} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \ldots & a_{n,p} \end{pmatrix} & \begin{matrix} \leftarrow e_1 \\ \leftarrow e_2 \\ \vdots \\ \leftarrow e_n \end{matrix} \\ \begin{matrix} \uparrow & \quad \uparrow & & \quad \uparrow \\ v_1 & \quad v_2 & \ldots & \quad v_p \end{matrix} & \end{array}$$

où l'on a noté $v_i = a_{1,i}.e_1 + a_{2,i}.e_2 + \cdots + a_{n,i}.e_n$ pour tout $1 \leq i \leq p$.

b) Réciproquement.
Toute matrice $A = (a_{i,j})_{1 \leq i \leq n\,,\,1 \leq j \leq p}$ donnée dans $\mathcal{M}_{n,p}(\mathbb{K})$ peut toujours être considérée comme la matrice dans la base canonique $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ de $\mathbb{K}^n$ de la famille $\mathcal{C} = (v_1, v_2, \ldots, v_p)$ de vecteurs de $\mathbb{K}^n$ définis par les *colonnes* de A :

$$v_1 = (a_{1,1}, a_{2,1}, \ldots, a_{n,1}), \;\; v_2 = (a_{1,2}, a_{2,2}, \ldots, a_{n,2}), \;\; \ldots, \;\; v_p = (a_{1,p}, a_{2,p}, \ldots, a_{n,p}).$$

3.3.2 Matrice de passage

On va appliquer la définition précédente avec $p = n$ pour donner un critère permettant de reconnaître si une famille de n vecteurs dans $\mathbb{K}^n$ est ou non une base.

a) Proposition et définition.
Soit $E = \mathbb{K}^n$. Soit $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ une base de E. Soit $\mathcal{B}' = (e'_1, e'_2, \ldots, e'_n)$ une famille de n vecteurs de E. Soit P la matrice de la famille $\mathcal{B}'$ dans la base $\mathcal{B}$, au sens de la définition ci-dessus. Alors :

$\mathcal{B}'$ est une base de E si et seulement si P est inversible.

Dans ce cas, P est appelée la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$.

On a donc :

$$P = \text{Pass}_{(\mathcal{B}\to\mathcal{B}')} = \begin{array}{cc} \begin{pmatrix} * & * & \dots & * \\ * & * & \dots & * \\ \vdots & \vdots & & \vdots \\ * & * & \dots & * \end{pmatrix} & \begin{array}{l} \leftarrow e_1 \\ \leftarrow e_2 \\ \quad\vdots \\ \leftarrow e_n \end{array} \\ \begin{array}{cccc} \uparrow & \uparrow & & \uparrow \\ e'_1 & e'_2 & \dots & e'_n \end{array} & \end{array}$$

Preuve : d'après la définition 3.2.1.b, P n'est autre que la matrice par rapport à la base $\mathcal{B}$ de l'endomorphisme f de E défini par $f(e_i) = e'_i$ pour tout $1 \leq i \leq n$. Dès lors, P inversible équivaut à f bijectif d'après 3.2.3.b, et donc à $\mathcal{B}'$ base de E d'après 2.1.5.e. □

b) Remarques.

On peut remarquer que, d'après la définition 3.2.1.b, on a : $\text{Pass}_{(\mathcal{B}\to\mathcal{B}')} = M_{\mathcal{B}'\mathcal{B}}(\text{id}_E)$. Il est clair aussi que, si P est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, alors P^{-1} est la matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$.

3.3.3 Première formule de changement de bases

Elle a pour objet d'exprimer les composantes d'un vecteur dans une base en fonction des composantes du même vecteur dans une autre base à l'aide de la matrice de passage d'une base à l'autre. Plus précisément :

Proposition.

Soit $E = \mathbb{K}^n$. Soient $\mathcal{B} = (e_1, e_2, \dots, e_n)$ et $\mathcal{B}' = (e'_1, e'_2, \dots, e'_n)$ deux bases de E. Soit $P \in GL(n, \mathbb{K})$ la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Si, pour un vecteur u quelconque de E, on note X la matrice colonne des composantes de u dans la base $\mathcal{B}$, et X' la matrice colonne des composantes de u dans la base $\mathcal{B}'$, alors on a :

$$X = PX'.$$

Preuve : la formule $X = PX'$ ne fait que regrouper sous une forme “compacte” un calcul essentiellement élémentaire. Afin d'en faciliter la lecture, on va le détailler dans le cas particulier $n = 3$, et on laisse en exercice (le faire !) la rédaction de ces mêmes calculs pour n quelconque.

Avec les données de l'énoncé, notons : $P = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}$, $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $X' = \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}$.

$$\begin{aligned} u &= x'_1.e'_1 + x'_2.e'_2 + x'_3.e'_3 \quad \text{par définition de } X' \\ &= x'_1.(a_{1,1}.e_1 + a_{2,1}.e_2 + a_{3,1}.e_3) + x'_2(a_{1,2}.e_1 + a_{2,2}.e_2 + a_{3,2}.e_3) \\ &\qquad + x'_3(a_{1,3}.e_1 + a_{2,3}.e_2 + a_{3,3}.e_3) \quad \text{par définition de } P \\ &= (x'_1 a_{1,1} + x'_2 a_{1,2} + x'_3 a_{1,3}).e_1 + (x'_1 a_{2,1} + x'_2 a_{2,2} + x'_3 a_{2,3}).e_2 + (x'_1 a_{3,1} + x'_2 a_{3,2} + x'_3 a_{3,3}).e_3 \\ &\qquad \text{d'après les règles de calculs dans } \mathbb{K}^n. \end{aligned}$$

Comme par ailleurs $u = x_1.e_1 + x_2.e_2 + x_3.e_3$ par définition de X, on en déduit par unicité des composantes de u dans la base $\mathcal{B}$ que :

$$\begin{cases} x_1 = x'_1 a_{1,1} + x'_2 a_{1,2} + x'_3 a_{1,3} \\ x_2 = x'_1 a_{2,1} + x'_2 a_{2,2} + x'_3 a_{2,3}, \\ x_3 = x'_1 a_{3,1} + x'_2 a_{3,2} + x'_3 a_{3,3} \end{cases} \quad \text{donc} \quad \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{X} = \underbrace{\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}}_{P} \underbrace{\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}}_{X'}. \qquad \square$$

3.3.4 Seconde formule de changement de bases

Elle a pour objet d'exprimer la matrice d'une application linéaire par rapport à des bases données en fonction de la matrice de la même application linéaire par rapport à deux nouvelles bases, et des deux matrices de passage correspondantes. Plus précisément :

a) Proposition.
Soient $E = \mathbb{K}^n$ et $E' = \mathbb{K}^p$.
Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E et $P \in GL(n, \mathbb{K})$ la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$.
Soient $\mathcal{C}$ et $\mathcal{C}'$ deux bases de E' et $Q \in GL(p, \mathbb{K})$ la matrice de passage de $\mathcal{C}$ à $\mathcal{C}'$.
Soit $f \in \mathcal{L}(E, E')$ une application linéaire de E dans E'.
Alors on a dans $\mathcal{M}_{p,n}(\mathbb{K})$ l'égalité :

$$M_{\mathcal{B}'\mathcal{C}'}(f) = Q^{-1}\, M_{\mathcal{B}\mathcal{C}}(f)\, P,$$

de sorte que les deux matrices $M_{\mathcal{B}'\mathcal{C}'}(f)$ et $M_{\mathcal{B}\mathcal{C}}(f)$ sont équivalentes (au sens de 3.1.6.a).

Preuve : Notons $A = M_{\mathcal{B}\mathcal{C}}(f)$ et $A' = M_{\mathcal{B}'\mathcal{C}'}(f)$. Pour $u \in E$ quelconque, posons :
X la matrice colonne des n composantes de u dans la base $\mathcal{B}$,
X' la matrice colonne des n composantes de u dans la base $\mathcal{B}'$,
Y la matrice colonne des p composantes de $f(u)$ dans la base $\mathcal{C}$,
Y' la matrice colonne des p composantes de $f(u)$ dans la base $\mathcal{C}'$.
Comme $P = \mathrm{Pass}_{(\mathcal{B}\to\mathcal{B}')}$ et $Q = \mathrm{Pass}_{(\mathcal{C}\to\mathcal{C}')}$, on a d'après 3.3.3 : $X = PX'$ et $Y = QY'$.
De plus, d'après 3.2.1.c, on a : $Y = M_{\mathcal{B}\mathcal{C}}(f)X = AX$ et $Y' = M_{\mathcal{B}'\mathcal{C}'}(f)X' = A'X'$.
On déduit de ces quatre égalités que : $A'X' = Y' = Q^{-1}Y = Q^{-1}AX = Q^{-1}APX'$.
L'égalité $A'X' = (Q^{-1}AP)X'$ obtenue étant établie pour tout $u \in E$, et donc pour toute matrice colonne X', on conclut que $A' = Q^{-1}AP$. □

b) Exemple.
Soit $\mathcal{B} = (e_1, e_2, e_3)$ la base canonique de $\mathbb{R}^3$. Soit $\mathcal{C} = (v_1, v_2)$ la base canonique de $\mathbb{R}^2$. Soit $f \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^2)$ définie par sa matrice $A = M_{\mathcal{B}\mathcal{C}}(f) = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 1 & -3 \end{pmatrix}$. Posons :

$$\begin{cases} e'_1 &= e_1 + 2.e_2 \\ e'_2 &= e_1 + e_2 + e_3 \\ e'_3 &= 2.e_2 + e_3 \end{cases} \text{dans } \mathbb{R}^3, \quad \text{et} \quad \begin{cases} v'_1 &= 2.v_1 + v_2 \\ v'_2 &= 3.v_1 + 2.v_2 \end{cases} \text{dans } \mathbb{R}^2.$$

On vérifie aisément (laissé en exercice) que $\mathcal{B}' = (e'_1, e'_2, e'_3)$ est une base de $\mathbb{R}^3$ et que $\mathcal{C}' = (v'_1, v'_2)$ est une base de $\mathbb{R}^2$. Par définition de la notion de matrice de passage, on a :

$$P = \mathrm{Pass}_{(\mathcal{B}\to\mathcal{B}')} = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{et} \quad Q = \mathrm{Pass}_{(\mathcal{C}\to\mathcal{C}')} = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}.$$

On calcule : $Q^{-1} = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}$, puis : $A' = M_{\mathcal{B}'\mathcal{C}'}(f) = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 2 & 1 & -3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} =$
$\begin{pmatrix} -4 & -5 & 9 \\ 3 & 3 & -6 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -14 & 0 & -1 \\ 9 & 0 & 0 \end{pmatrix}$.
Donc f, qui était définie par : $f(e_1) = v_1 + 2.v_2$, $f(e_2) = -v_1 + v_2$ et $f(e_3) = -3.v_2$,
est aussi définie par : $f(e'_1) = -14.v'_1 + 9.v'_2$, $f(e'_2) = 0_{\mathbb{R}^2}$ et $f(e'_3) = -v'_1$.

c) Corollaire. (Cas particulier des endomorphismes)
Soit $E = \mathbb{K}^n$. Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E et $P \in GL(n, \mathbb{K})$ la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Soit $f \in \mathcal{L}(E)$ un endomorphisme de E. Alors on a dans $\mathcal{M}_n(\mathbb{K})$ les égalités :

$$M_{\mathcal{B}'}(f) = P^{-1}\, M_{\mathcal{B}}(f)\, P, \qquad M_{\mathcal{B}}(f) = P\, M_{\mathcal{B}'}(f)\, P^{-1},$$

de sorte que les deux matrices $M_{\mathcal{B}'}(f)$ et $M_{\mathcal{B}}(f)$ sont semblables (au sens de 3.1.6.b).

Preuve : On applique a) au cas particulier où $E' = E$, $\mathcal{C} = \mathcal{B}$, $\mathcal{C}' = \mathcal{B}'$, $Q = P$. □

Ce corollaire est particulièrement utile dans la pratique, et à bien connaître. On en verra dans les chapitres suivants de nombreuses applications.

3.4 Exercices sur le chapitre 3

Exercice 1.

Soit f l'endomorphisme de $\mathbb{R}^3$ dont la matrice par rapport à la base canonique est $A = \begin{pmatrix} 1 & 1 & -1 \\ \frac{1}{2} & -3 & -\frac{1}{2} \\ -1 & -1 & 1 \end{pmatrix}$.

Déterminer son image et son noyau, et montrer qu'ils sont supplémentaires.

Exercice 2.
Soit f l'endomorphisme de $\mathbb{R}^2$ défini par $f(e_1) = 3.e_1 - e_2$ et $f(e_2) = e_1 + e_2$, où $\mathcal{B} = (e_1, e_2)$ est la base canonique de $\mathbb{R}^2$. Donner la matrice A de f par rapport à $\mathcal{B}$. Montrer que f est un automorphisme de $\mathbb{R}^2$. Déterminer f^{-1}.

Exercice 3.
Donner la matrice par rapport à la base canonique de l'endomorphisme f de $\mathbb{R}^3$ défini par :

$$f(x, y, z) = (y + 3z,\ x + y - z,\ -x + 2y).$$

L'endomorphisme f est-il un automorphisme ? Si oui, expliciter l'automorphisme f^{-1}.

Exercice 4.
Soient $\mathcal{B} = (e_1, e_2, e_3)$ la base canonique de $\mathbb{R}^3$ et $\mathcal{B}' = (u_1, u_2, u_3)$ où $u_1 = (0, 1, -1)$, $u_2 = (1, 0, -1)$ et $u_3 = (1, 1, 1)$.
a) Montrer que $\mathcal{B}'$ est une base de $\mathbb{R}^3$. Calculer l'inverse de la matrice P de passage de $\mathcal{B}$ à $\mathcal{B}'$.
b) Soit f l'endomorphisme de $\mathbb{R}^3$ dont la matrice par rapport à $\mathcal{B}'$ est $A' = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.
Calculer la matrice A de f par rapport à $\mathcal{B}$.
c) Calculer $(A')^n$ pour tout entier n, et en déduire A^n.

Exercice 5.
Soit f l'endomorphisme de $\mathbb{R}^3$ qui, à tout vecteur $u = (x, y, z)$, associe $f(u) = (x', y', z')$ défini par :

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

a) Ecrire la matrice A de f par rapport à la base canonique.
b) Déterminer $\operatorname{Ker} f$ et $\operatorname{Im} f$.
c) Montrer que les vecteurs $u_1 = (1, 0, 1)$, $u_2 = (1, 1, 0)$ et $u_3 = (0, 1, 1)$ forment une base $\mathcal{B}'$ de $\mathbb{R}^3$.
d) Calculer la matrice de f par rapport à $\mathcal{B}'$.

Exercice 6.
Déterminer, en discutant suivant les valeurs du paramètre réel α, le noyau et l'image de l'endomorphisme f_α de $\mathbb{R}^3$ dont la matrice par rapport à la base canonique est $A_\alpha = \begin{pmatrix} 1 & 1 & \alpha \\ 1 & \alpha & 1 \\ \alpha & 1 & 1 \end{pmatrix}$. Pour quelles valeurs de α est-ce un automorphisme ?

Exercice 7.
Montrer que $A = \begin{pmatrix} 0 & 1 & -1 \\ -3 & 4 & -3 \\ -1 & 1 & 0 \end{pmatrix}$ est inversible et calculer A^{-1}. Calculer par ailleurs $A^2 - 3A + 2I_3$ et retrouver le résultat.

Exercice 8.
On note E le $\mathbb{R}$-e.v. $\mathbb{R}^3$ et $\mathcal{B} = (e_1, e_2, e_3)$ sa base canonique. Soit f l'endomorphisme de E défini par :

$$f(e_1) = e_1 + 2e_3, \quad f(e_2) = 4e_1 + 3e_2 + 9e_3, \quad f(e_3) = -2e_1 - e_2 - 4e_3.$$

a) Ecrire la matrice A de f par rapport à $\mathcal{B}$. Calculer A^2 et A^3. En déduire que A est inversible et calculer A^{-1}.

b) Montrer que $D = \{u \in E\,;\, f(u) = u\}$ est une droite vectorielle de E. On notera v_1 un vecteur non-nul quelconque de D.

c) Montrer que $P = \{(x, y, z) \in E\,;\, x - 2y = 0\}$ est un plan vectoriel. En donner une base (v_2, v_3).

d) Montrer que $E = D \oplus P$ et donner la matrice A' de f par rapport à la base $\mathcal{B}' = (v_1, v_2, v_3)$.

e) On pose $u_1 = e_1 + e_2 + 2e_3$, $u_2 = f(u_1)$ et $u_3 = f(u_2)$. Montrer que $\mathcal{B}'' = (u_1, u_2, u_3)$ est une base de E et donner la matrice A'' de f par rapport à $\mathcal{B}''$.

Exercice 9.
Soient $E = \mathbb{R}^3$ et $E' = \mathbb{R}^4$. Soient $\mathcal{B} = (e_1, e_2, e_3)$ une base de E et $\mathcal{C} = (u_1, u_2, u_3, u_4)$ une base de E'. On définit les familles $\mathcal{B}' = (e'_1, e'_2, e'_3)$ de vecteurs de E et $\mathcal{C}' = (u'_1, u'_2, u'_3, u'_4)$ de vecteurs de E' par :

$$\begin{cases} e'_1 = e_1 + e_3 \\ e'_2 = e_1 + 2e_2 + e_3\,, \\ e'_3 = -e_1 + e_2 + e_3 \end{cases} \qquad \begin{cases} u'_1 = u_1 + u_2 + u_3 + u_4 \\ u'_2 = u_1 - u_2 + u_3 + u_4 \\ u'_3 = u_1 + u_2 + u_3 - u_4 \\ u'_4 = -u_1 + u_2 + u_3 - u_4 \end{cases}.$$

a) Montrer que $\mathcal{B}'$ est une base de E et $\mathcal{C}'$ est une base de E'. Expliciter la matrice de passage P de $\mathcal{B}$ à $\mathcal{B}'$ et la matrice de passage Q de $\mathcal{C}$ à $\mathcal{C}'$.

b) Soit f l'application linéaire de E dans E' dont la matrice par rapport aux bases $\mathcal{B}$ et $\mathcal{C}$ est :

$$A = \mathrm{Mat}_{\mathcal{B}\mathcal{C}}(f) = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}.$$

Calculer la matrice $A' = \mathrm{Mat}_{\mathcal{B}'\mathcal{C}'}(f)$ de f par rapport aux bases $\mathcal{B}'$ et $\mathcal{C}'$.

Exercice 10.
Soit f l'endomorphisme de $\mathbb{C}^3$ dont la matrice par rapport à la base canonique $\mathcal{B} = (e_1, e_2, e_3)$ est

$$A = \mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} -2i & -2i & 1 \\ 2(1+i) & 1+2i & -1+i \\ 1+2i & 1+i & -1 \end{pmatrix}.$$

a) En posant $u_1 = e_1 - e_2$, $u_2 = -ie_2 + e_3$ et $u_3 = e_1 - e_2 + ie_3$, montrer que la famille $\mathcal{B}' = (u_1, u_2, u_3)$ est une base de $\mathbb{C}^3$; déterminer la matrice de passage P de $\mathcal{B}$ à $\mathcal{B}'$ et son inverse P^{-1}.

b) Calculer la matrice $A' = \mathrm{Mat}_{\mathcal{B}'}(f)$ de f par rapport à la base $\mathcal{B}'$.

c) En déduire que f est un automorphisme de $\mathbb{C}^3$, et déduire du calcul de $(A')^{-1}$ le calcul de A^{-1}.

Chapitre 4

Rang

4.1 Divers points de vue sur la notion de rang

4.1.1 Rang d'une famille finie de vecteurs

a) Définition.
Soit $\mathcal{C} = (u_1, u_2, \ldots, u_n)$ une famille finie de n vecteurs dans un $\mathbb{K}$-e.v. E. On appelle *rang de la famille* $\mathcal{C}$ la dimension du ss-e.v. $\operatorname{Vect} \mathcal{C}$ engendré par $\mathcal{C}$ (cf. 1.3.2).
On note $\operatorname{rg} \mathcal{C}$ le rang de $\mathcal{C}$.

b) Remarque. Avec les notations précédentes, on a : $0 \leq \operatorname{rg} \mathcal{C} \leq n$.
En effet, si l'on note $F = \operatorname{Vect} \mathcal{C}$ et $m = \operatorname{rg} \mathcal{C} = \dim F$, toute base de F est formée de m vecteurs et, comme $\mathcal{C}$ est une famille génératrice de F, on a $m \leq n$ d'après 1.5.1.f.(2).

4.1.2 Rang d'une application linéaire

a) Définition.
Soient $E = \mathbb{K}^n$ et $E' = \mathbb{K}^p$, et $f \in \mathcal{L}(E, E')$ une application linéaire de E dans E'. On appelle *rang de l'application linéaire* f la dimension du ss-e.v. $\operatorname{Im} f$ de E' (cf. 2.2.2.d).
On note $\operatorname{rg} f$ le rang de f.

b) Remarque. Avec les notations précédentes, on a : $0 \leq \operatorname{rg} f \leq p$.
En effet, $\operatorname{Im} f$ est un ss-e.v. de E', d'où $\dim \operatorname{Im} f \leq \dim E'$ d'après 1.5.2.c.(2).

4.1.3 Rang d'un système d'équations linéaires homogènes

a) Définition.
Soit (S_0) un système linéaire homogène de p équations à n inconnues, à coefficients dans $\mathbb{K}$, au sens de 2.3.3.a. On appelle *rang du système* (S_0) l'entier naturel :

$$\operatorname{rg}(S_0) = n - (\text{la dimension du ss-e.v. des solutions de } (S_0) \text{ dans } \mathbb{K}^n).$$

b) Remarque. Il est clair que $0 \leq \operatorname{rg}(S_0) \leq n$.

4.1.4 Rang d'une matrice

Soit $A = \begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{p,1} & a_{p,2} & \ldots & a_{p,n} \end{pmatrix} \in \mathcal{M}_{p,n}(\mathbb{K})$ une matrice à p lignes et n colonnes.

a) Premier point de vue.

On associe à la matrice A la famille des n vecteurs $u_1, u_2, \ldots, u_n$ de $\mathbb{K}^p$ formée par les n colonnes de A. Donc A est la matrice de la famille $\mathcal{C}$ de vecteurs de $\mathbb{K}^p$ dans la base canonique au sens de 3.3.1. On appelle alors *rang de la matrice* A, noté $\operatorname{rg} A$, le rang de cette famille $\mathcal{C}$ au sens du 4.1.1 ci-dessus.
On a déjà remarqué en 4.1.1.b que : $\operatorname{rg} A = \operatorname{rg} \mathcal{C} \leq n$.
Mais comme ici $\operatorname{Vect} \mathcal{C}$ est un ss-e.v. de $\mathbb{K}^p$, on a aussi $\operatorname{rg} \mathcal{C} = \dim \operatorname{Vect} \mathcal{C} \leq \dim \mathbb{K}^p = p$.
En résumé : $\operatorname{rg} A \leq n$ et $\operatorname{rg} A \leq p$.

b) Deuxième point de vue.

D'après 3.2.2.a, il existe une unique application linéaire f de $\mathbb{K}^n$ dans $\mathbb{K}^p$ telle que A soit la matrice de f par rapport aux bases canoniques de $\mathbb{K}^n$ et $\mathbb{K}^p$. On appelle alors *rang de la matrice* A, noté $\operatorname{rg} A$, le rang de cette application linéaire f au sens du 4.1.2.a ci-dessus.
On a déjà remarqué en 4.1.2.b que : $\operatorname{rg} A = \operatorname{rg} f \leq p$.
Mais comme ici $\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim \mathbb{K}^n = n$ d'après 2.2.2.d, on a aussi $\operatorname{rg} f = \dim \operatorname{Im} f \leq \dim \mathbb{K}^n = n$.
En résumé : $\operatorname{rg} A \leq n$ et $\operatorname{rg} A \leq p$.

c) Troisième point de vue.

On peut associer à A le système linéaire homogène de p équations à n inconnues suivant :

$$(S_0) \begin{cases} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + & a_{1,n}x_n & = & 0 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + & a_{2,n}x_n & = & 0 \\ \cdots \\ a_{p,1}x_1 & + & a_{p,2}x_2 & + & \cdots & + & a_{p,n}x_n & = & 0 \end{cases} \quad \text{c'est-à-dire encore} \quad A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

On dit que A est la matrice du système (S_0). On appelle alors *rang de la matrice* A, noté $\operatorname{rg} A$, le rang de ce système (S_0) au sens du 4.1.3.a ci-dessus.
Il est clair que le $\mathbb{K}$-e.v. des solutions dans $\mathbb{K}^n$ de ce système (S_0) n'est autre que le noyau de l'application linéaire $f \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^p)$ introduite au b) précédent.

d) Quatrième point de vue.

On appelle *sous-matrice de* A, ou *matrice extraite de* A, toute matrice à p' lignes et n' colonnes (avec $p' \leq p$ et $n' \leq n$) obtenue en supprimant $p - p'$ lignes et $n - n'$ colonnes de A, en laissant à la même position les termes restants.
On appelle alors *rang de la matrice* A, noté $\operatorname{rg} A$, le maximum des ordres des matrices carrées extraites de A qui sont inversibles.
En d'autres termes, $\operatorname{rg} A = r$ signifie qu'il existe une matrice extraite de A qui appartient à $GL(r, \mathbb{K})$, et que toute matrice carrée d'ordre $m > r$ extraite de A est non-inversible.

e) Théorème de synthèse.

Pour toute matrice $A \in \mathcal{M}_{p,n}(\mathbb{K})$*, les quatre façons de définir le rang de* A *exposées ci-dessus coïncident, et définissent donc le même entier naturel* $\operatorname{rg} A$*, qui vérifie* $\operatorname{rg} A \leq p$ *et* $\operatorname{rg} A \leq n$.

Preuve : cf. ouvrage de référence. □

4.1.5 Un exemple numérique

Soit $A = \begin{pmatrix} 1 & 2 & 1 & 3 & 3 \\ 0 & -1 & -1 & -1 & -2 \\ 2 & 0 & -2 & 2 & -2 \\ 1 & 0 & 1 & 1 & 1 \end{pmatrix}$, matrice à 4 lignes et 5 colonnes à coefficients réels.

a) Premier point de vue.

On introduit la famille $\mathcal{C} = (u_1, u_2, u_3, u_4, u_5)$ formée des 5 vecteurs de $\mathbb{R}^4$ correspondant aux colonnes de A. Donc :

$u_1 = (1, 0, 2, 1), \quad u_2 = (2, -1, 0, 0), \quad u_3 = (1, -1, -2, 1), \quad u_4 = (3, -1, 2, 1), \quad u_5 = (3, -2, -2, 1).$

On a : $u_4 = u_1 + u_2$ et $u_5 = u_2 + u_3$, de sorte que $\operatorname{Vect} \mathcal{C} = \operatorname{Vect}(u_1, u_2, u_3)$. On montre ensuite (par un calcul facile laissé au lecteur...) que (u_1, u_2, u_3) est libre. Donc (u_1, u_2, u_3) est une base de $\operatorname{Vect} \mathcal{C}$. Ceci prouve que $\dim \operatorname{Vect} \mathcal{C} = 3$. On conclut que $\operatorname{rg} A = 3$.

b) Troisième point de vue.

On introduit le système homogène de 4 équations à 5 inconnues (notées ici x, y, z, t, s) dont A est la matrice :

$$(S_0) \left\{ \begin{array}{rrrrrl} x & +\ 2y & +\ z & +\ 3t & +\ 3s = 0 & (1) \\ & -\ y & -\ z & -\ t & -\ 2s = 0 & (2) \\ 2x & & -\ 2z & +\ 2t & -\ 2s = 0 & (3) \\ x & & +\ z & +\ t & +\ s = 0 & (4) \end{array} \right.$$

En utilisant 2.3.3.d.(3), on obtient un système équivalent en remplaçant (3) par $(3) - 2 \times (1)$ et en remplaçant (4) par $(4) - (1)$:

$$(S_0) \Leftrightarrow \left\{ \begin{array}{rrrrrl} x & +\ 2y & +\ z & +\ 3t & +\ 3s = 0 & (1) \\ & -\ y & -\ z & -\ t & -\ 2s = 0 & (2) \\ & -\ 4y & -\ 4z & -\ 4t & -\ 8s = 0 & (3') \\ & -\ 2y & & -\ 2t & -\ 2s = 0 & (4') \end{array} \right., \quad \text{puis} \quad (S_0) \Leftrightarrow \left\{ \begin{array}{rrrrr} x & +\ 2y & +\ z & +\ 3t & +\ 3s = 0 \\ & y & +\ z & +\ t & +\ 2s = 0 \\ & y & & +\ t & +\ s = 0 \end{array} \right.,$$

en remarquant que $(3')$ et (2) sont équivalentes, et en simplifiant (2) par -1 et $(4')$ par -2 [d'après 2.3.3.d.(2)]. On en tire :

$$(S_0) \Leftrightarrow \left\{ \begin{array}{l} y = -t - s \\ z = -t - 2s - y = -t - 2s - (-t - s) = -s \\ x = -3t - 3s - z - 2y = -3t - 3s - (-s) - 2(-t - s) = -t \end{array} \right.$$

L'ensemble des solutions de (S_0) est le ss-e.v. $F_0 = \{(-t, -t - s, -s, t, s)\,;\, t, s \in \mathbb{R}\}$ de $\mathbb{R}^5$, c'est-à-dire le plan vectoriel dont une base est (u, v) avec $u = (-1, -1, 0, 1, 0)$ et $v = (0, -1, -1, 0, 1)$. Le rang de (S_0) est égal à $5 - \dim F_0 = 5 - 2 = 3$. On conclut que $\operatorname{rg} A = 3$.

c) Deuxième point de vue.

Soit f l'application linéaire de $\mathbb{R}^5$ dans $\mathbb{R}^4$ telle que A soit la matrice de f par rapport aux bases canoniques. Pour tout vecteur $u = (x, y, z, t, s) \in \mathbb{R}^5$, on a clairement :

$$u \in \operatorname{Ker} f \quad \Leftrightarrow \quad f(u) = (0, 0, 0, 0) \quad \Leftrightarrow \quad A \begin{pmatrix} x \\ y \\ z \\ t \\ s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad u \text{ est solution de } (S_0).$$

Comme on l'a vu au b) ci-dessus, le ss-e.v. des solutions de (S_0) est de dimension 2. Donc $\dim \operatorname{Ker} f = 2$. Comme $\dim \operatorname{Im} f + \dim \operatorname{Ker} f = \dim \mathbb{R}^5 = 5$, on en déduit $\operatorname{rg} f = \dim \operatorname{Im} f = 5 - 2 = 3$. On conclut que $\operatorname{rg} A = 3$.

d) Quatrième point de vue.

Comme $\operatorname{rg} A \leq$ nombre de lignes et $\leq$ nombre de colonnes, on sait déjà ici que $\operatorname{rg} A \leq 4$.

A est de rang 4 si et seulement s'il existe une matrice carrée d'ordre 4 extraite de A et inversible. Or, il existe ici 5 matrices carrées d'ordre 4 extraites de A. (On les obtient en supprimant l'une des colonnes de A). Notons-les :

$$A_1 = \begin{pmatrix} 2 & 1 & 3 & 3 \\ -1 & -1 & -1 & -2 \\ 0 & -2 & 2 & -2 \\ 0 & 1 & 1 & 1 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1 & 1 & 3 & 3 \\ 0 & -1 & -1 & -2 \\ 2 & -2 & 2 & -2 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \qquad A_3 = \begin{pmatrix} 1 & 2 & 3 & 3 \\ 0 & -1 & -1 & -2 \\ 2 & 0 & 2 & -2 \\ 1 & 0 & 1 & 1 \end{pmatrix},$$

$$A_4 = \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & -1 & -1 & -2 \\ 2 & 0 & -2 & -2 \\ 1 & 0 & 1 & 1 \end{pmatrix}, \qquad A_5 = \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & -1 & -1 & -1 \\ 2 & 0 & -2 & 2 \\ 1 & 0 & 1 & 1 \end{pmatrix}.$$

Désignons comme en a) par u_1, u_2, u_3, u_4, u_5 les vecteurs colonnes de A dans $\mathbb{R}^4$.

La matrice A_1 est celle de la famille (u_2, u_3, u_4, u_5). Comme on a $u_5 = u_2 + u_3$, cette famille est liée, donc d'après 3.3.2 la matrice A_1 n'est pas inversible.

Par le même raisonnement, A_4 n'est pas inversible car $u_5 = u_2 + u_3$, A_5 et A_3 ne sont pas inversibles car $u_4 = u_1 + u_2$, et A_2 n'est pas inversible car $u_5 = u_3 + u_4 - u_1$.
On a ainsi vérifié que toutes les matrices carrées d'ordre 4 extraites de A sont non-inversibles, ce qui prouve que $\operatorname{rg} A < 4$.
On regarde ensuite les matrices carrées d'ordre 3 extraites de A. Si l'une au moins est inversible, cela suffit à prouver que $\operatorname{rg} A = 3$. Or, la matrice obtenue en supprimant dans A la 1ère ligne et les deux dernières colonnes est : $M = \begin{pmatrix} 0 & -1 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & 1 \end{pmatrix}$, et il est facile de vérifier (soit par la méthode vue en 3.2.4, soit en vérifiant que son déterminant est non-nul comme on le verra plus loin dans le cours) qu'elle est inversible.

Conclusion : on a montré que toutes les matrices carrées d'ordre 4 extraites de A sont non-inversibles, et qu'il existe une matrice carrée d'ordre 3 extraite de A qui est inversible ; ceci prouve que $\operatorname{rg} A = 3$.

4.1.6 Matrices carrées de rang maximal

On s'intéresse ici au cas particulier d'une matrice A carrée ($n = p$). On a donc toujours $\operatorname{rg} A \leq n$. Le théorème suivant exprime plusieurs caractérisations des situations où $\operatorname{rg} A$ prend la plus grande valeur possible, c'est-à-dire $\operatorname{rg} A = n$.

a) Théorème.

Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée. Soient :
$\mathcal{C}$ la famille de n vecteurs de $\mathbb{K}^n$ telle que A soit la matrice de $\mathcal{C}$ dans la base canonique (cf. 3.3.1),
f l'endomorphisme de $\mathbb{K}^n$ tel que A soit la matrice de f par rapport à la base canonique (cf. 3.2.1),
(S_0) le système linéaire homogène tel que A soit la matrice de (S_0) (cf. 4.1.4.c).
On sait que : $\operatorname{rg} A = \operatorname{rg} \mathcal{C} = \operatorname{rg} f = \operatorname{rg}(S_0) \leq n$. *Les assertions suivantes sont alors équivalentes :*

(i) $\operatorname{rg} A = \operatorname{rg} \mathcal{C} = \operatorname{rg} f = \operatorname{rg}(S_0) = n$;
(ii) *la famille $\mathcal{C}$ est une base de $\mathbb{K}^n$;*
(iii) *l'endomorphisme f est bijectif,*
c'est-à-dire f est un automorphisme de $\mathbb{K}^n$, ou encore $f \in GL(\mathbb{K}^n)$;
(iv) *le système (S_0) admet pour unique solution le vecteur nul de $\mathbb{K}^n$;*
(v) *la matrice A est inversible, c'est-à-dire $A \in GL(n, \mathbb{K})$.*

Preuve : cet énoncé ne fait que recapituler divers critères vus précédemment. En effet :
$\operatorname{rg} \mathcal{C} = n$ si et seulement si $\operatorname{Vect} \mathcal{C} = \mathbb{K}^n$, c'est-à-dire si et seulement si $\mathcal{C}$ est une famille génératrice de $\mathbb{K}^n$; mais cela équivaut à (ii) d'après 1.5.1.f.(4).
$\operatorname{rg} f = n$ si et seulement si $\operatorname{Im} f = \mathbb{K}^n$, c'est-à-dire si et seulement si f est surjective ; mais la surjectivité de f équivaut à (iii) d'après 2.1.5.f.
$\operatorname{rg}(S_0) = n$ équivaut à dire que le ss-e.v. de $\mathbb{K}^n$ formé des solutions de (S_0) est de dimension $n - n = 0$, et donc à (iv). [On pourrait aussi utiliser la dernière remarque de 4.1.4.c pour déduire que (iv) équivaut à $\operatorname{Ker} f = \{0_{\mathbb{K}^n}\}$, c'est-à-dire à f injective d'après 2.1.3.c, et donc à (iii) d'après 2.1.5.f].
Enfin, en utilisant le 4.1.4.d, $\operatorname{rg} A = n$ signifie simplement que A est elle-même inversible, c'est-à-dire (v). □

b) Remarque. A l'opposé, la situation où $\operatorname{rg} A$ est minimal, c'est-à-dire $\operatorname{rg} A = 0$, correspond au cas où A est la matrice nulle (cf. 3.1.2.d), c'est-à-dire où f est l'endomorphisme nul de $\mathbb{K}^n$ (cf. 2.1.2.d), ou encore où tous les vecteurs de la famille $\mathcal{C}$ sont égaux au vecteur nul de $\mathbb{K}^n$. Dans ce cas, $\operatorname{Vect} \mathcal{C} = \operatorname{Im} f = \{0_{\mathbb{K}^n}\}$ et le ss-e.v. des solutions du système (S_0), qui est égal à $\operatorname{Ker} f$, est $\mathbb{K}^n$ tout entier.

4.1.7 Deux propriétés du rang

a) Proposition.

Toute matrice est de même rang que sa transposée (définie en 3.1.7).

Preuve : cf. ouvrage de référence. □

b) Théorème.

Deux matrices de $\mathcal{M}_{p,n}(\mathbb{K})$ ont même rang si et seulement si elles sont équivalentes (au sens défini en 3.1.6).

Preuve : cf. ouvrage de référence. □

La méthode de calcul effectif du rang présentée à la section suivante est une conséquence pratique importante de ces deux propriétés.

4.2 Une méthode de calcul

4.2.1 Rang et transformations élémentaires

a) Théorème.

Soit $A \in \mathcal{M}_{p,n}(\mathbb{K})$; le rang de A est le même que celui de la matrice obtenue à partir de A en appliquant l'une des six transformations élémentaires suivantes :

1. *échanger deux colonnes ;*
2. *multiplier une colonne par un scalaire non-nul ;*
3. *ajouter à une colonne une combinaison linéaire des autres colonnes ;*
4. *échanger deux lignes ;*
5. *multiplier une ligne par un scalaire non-nul ;*
6. *ajouter à une ligne une combinaison linéaire des autres lignes.*

Preuve : les points 1, 2 et 3 sont clairs en interprétant A comme la matrice de la famille des n vecteurs colonnes dans $\mathbb{K}^p$ (cf. 4.1.4.a). Dès lors, 4, 5 et 6 s'en déduisent avec 4.1.7.a. On peut aussi interpréter A comme matrice d'un système linéaire (cf. 4.1.4.c) et rappeler la proposition 2.3.3.d.□

b) Méthode pratique.

Soit $A \in \mathcal{M}_{p,n}(\mathbb{K})$. En appliquant les transformations élémentaires détaillées ci-dessus, on se ramène à une matrice A' *échelonnée*, c'est-à-dire dont le nombre de zéros en début de chaque ligne augmente strictement de ligne en ligne. Cette matrice est de même rang que A d'après le théorème a) ci-dessus. Mais le rang de A' est facile à calculer, soit en utilisant le quatrième point de vue (cf. 4.1.4.d), soit en utilisant les déterminants comme on va le voir au chapitre suivant.

4.2.2 Un exemple numérique.

Soit $A = \begin{pmatrix} 1 & a+1 & 2a^2 & 2a^2+2 \\ -1 & a-1 & a-a^2 & 3a-a^2-2 \\ 0 & a & a^2 & a^2+a \\ 1 & a+1 & 2a & 2a+2 \end{pmatrix} \in \mathcal{M}_{4,4}(\mathbb{R})$, où $a \in \mathbb{R}$ fixé.

On garde L_1 et L_3, on remplace L_2 par $L_2 + L_1$, et L_4 par $L_4 - L_1$, on obtient :

$$A_1 = \begin{pmatrix} 1 & a+1 & 2a^2 & 2a^2+2 \\ 0 & 2a & a+a^2 & 3a+a^2 \\ 0 & a & a^2 & a^2+a \\ 0 & 0 & 2a-2a^2 & 2a-2a^2 \end{pmatrix}.$$

On garde L_1, L_2 et L_4, on remplace L_3 par $2L_3$ puis par $2L_3 - L_2$, on obtient :

$$A_2 = \begin{pmatrix} 1 & a+1 & 2a^2 & 2a^2+2 \\ 0 & 2a & a+a^2 & 3a+a^2 \\ 0 & 0 & a^2-a & a^2-a \\ 0 & 0 & 2(a-a^2) & 2(a-a^2) \end{pmatrix}.$$

On garde L_1, L_2, L_3, on remplace L_4 par L_4+2L_3, on obtient : $A_3 = \begin{pmatrix} 1 & a+1 & 2a^2 & 2a^2+2 \\ 0 & 2a & a+a^2 & 3a+a^2 \\ 0 & 0 & a^2-a & a^2-a \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

Finalement, $\operatorname{rg} A = \operatorname{rg} A_3 = \begin{cases} 1 & \text{si } a = 0 \\ 2 & \text{si } a = 1 \\ 3 & \text{si } a \notin \{0,1\} \end{cases}$.

4.3 Exercices sur le chapitre 4

Exercice 1.
On considère dans le $\mathbb{R}$-e.v. $\mathbb{R}^5$ les six vecteurs :

$$u_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ -1 \\ 0 \end{pmatrix} \quad u_2 = \begin{pmatrix} 3 \\ 4 \\ 6 \\ 0 \\ 1 \end{pmatrix} \quad u_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \\ 4 \end{pmatrix} \quad u_4 = \begin{pmatrix} 3 \\ 3 \\ 3 \\ -1 \\ -3 \end{pmatrix} \quad u_5 = \begin{pmatrix} -4 \\ -3 \\ -2 \\ 3 \\ 7 \end{pmatrix} \quad u_6 = \begin{pmatrix} 4 \\ 5 \\ 8 \\ -1 \\ 1 \end{pmatrix}.$$

Calculer le rang de la famille $\mathcal{C}$ formée de ces six vecteurs, et donner une base du ss-e.v. de $\mathbb{R}^5$ qu'ils engendrent.

Exercice 2.
Calculer le rang du système linéaire homogène (à coefficients et inconnues complexes) :

$$\begin{cases} x + y + z + t + s = 0 \\ x + iy - z - it + s = 0 \\ x - y + z - t + s = 0 \\ x - iy - z + it + s = 0 \end{cases}$$

Exercice 3.
Calculer le rang de l'application linéaire de $\mathbb{R}^4$ dans $\mathbb{R}^5$ définie par :

$$f(x,y,z,t) = (x-y,\ y-z,\ z-t,\ t,\ 7x-3y-4z+5t).$$

Exercice 4.
Calculer le rang de chacune des deux matrices :

$$A = \begin{pmatrix} 1 & 3 & -2 & -1 \\ 2 & 5 & -2 & 1 \\ 1 & 1 & 6 & 13 \\ -2 & -6 & 8 & 10 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 3 & 2 & 12 & 1 \\ 6 & 5 & 34 & 3 \\ -12 & -3 & -14 & -3 \\ 9 & 4 & 20 & 2 \\ 3 & 8 & 44 & 0 \end{pmatrix}.$$

Exercice 5.
Les matrices carrées suivantes sont-elles de rang maximal ? Si oui, calculer leur inverse :

$$A = \begin{pmatrix} 3 & 2 \\ 8 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 1 \\ 1 & 3 & 3 & 2 \\ 1 & 4 & 3 & 3 \end{pmatrix}.$$

Exercice 6.
Calculer, en discutant suivant les valeurs du paramètre réel a, le rang de la famille (u_1, u_2, u_3, u_4) de vecteurs de $\mathbb{R}^4$ définie par :

$u_1 = (1, -1, 0, 1), \qquad u_2 = (a+1,\ a-1,\ a,\ a+1), \qquad u_3 = (2a^2,\ a-a^2,\ a^2,\ 2a),$
$u_4 = (2a^2+2,\ 3a-a^2-2,\ a^2+a,\ 2a+2).$

Exercice 7.
Les réels a et b étant deux paramètres fixés, on considère l'endomorphisme f de $\mathbb{R}^4$ qui, à tout vecteur (x, y, z, t), associe $f(x, y, z, t) = (x', y', z', t')$ défini par :

$$\begin{cases} x' = x + 2y - z - at \\ y' = x + 3y + (a-3)z - (a-b)t \\ z' = 3x + 6y + (a-b-3)z + (b-3a-2)t \\ t' = -x - y + (b-1)z + (2a-b+2)t \end{cases}$$

a) Montrer que f est un automorphisme si et seulement si $a \neq b$.
b) Déterminer $\operatorname{Ker} f$ et $\operatorname{Im} f$ dans le cas où $a = b \neq 2$.
c) Déterminer $\operatorname{Ker} f$ et $\operatorname{Im} f$ dans le cas où $a = b = 2$.
d) En déduire le rang de f suivant les valeurs de a et b.

Exercice 8.
Les réels a et b étant deux paramètres fixés, on considère l'endomorphisme f de $\mathbb{R}^4$ dont la matrice par rapport à la base canonique est :

$$A = \begin{pmatrix} 1 & 1 & -1 & 3 \\ 3 & 2 & -1 & 6 \\ a & -1 & b & a-b-1 \\ -a+b-2 & -a-3 & 2a-b+7 & -3a+b-10 \end{pmatrix}$$

a) Déterminer, en discutant suivant les valeurs de a et b, le rang de f.
b) Déterminer $\operatorname{Ker} f$ et $\operatorname{Im} f$:
lorsque $a = b = 1$;
lorsque $a = 0$ et $b = 2$;
lorsque $a = -1$ et $b = 1$.

Chapitre 5

Déterminant

5.1 Cas particulier des matrices 2×2

a) Dfinition.
Pour toute matrice carrée $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{K})$, on appelle *déterminant* de A, noté $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$, ou encore $\det A$, le scalaire défini par :

$$\det A = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \in \mathbb{K}.$$

b) Quelques règles de calcul.
On peut alors faire les observations suivantes :

• Si on échange les deux lignes ou les deux colonnes, on multiplie le déterminant par -1 :

$$\begin{vmatrix} c & d \\ a & b \end{vmatrix} = cb - da = -(ad - bc) = -\begin{vmatrix} a & b \\ c & d \end{vmatrix} \quad \text{et} \quad \begin{vmatrix} b & a \\ d & c \end{vmatrix} = bc - ad = -(ad - bc) = -\begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

• Le déterminant est linéaire par rapport à chaque ligne ou chaque colonne :

$$\begin{vmatrix} \lambda a+\mu a' & b \\ \lambda c+\mu c' & d \end{vmatrix} = (\lambda a + \mu a')d - b(\lambda c + \mu c') = \lambda(ad - bc) + \mu(a'd - bc') = \lambda \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \mu \begin{vmatrix} a' & b \\ c' & d \end{vmatrix},$$

de même : $\begin{vmatrix} a & \lambda b+\mu b' \\ c & \lambda d+\mu d' \end{vmatrix} = \lambda \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \mu \begin{vmatrix} a & b' \\ c & d' \end{vmatrix}$,

et encore : $\begin{vmatrix} \lambda a+\mu a' & \lambda b+\mu b' \\ c & d \end{vmatrix} = \lambda \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \mu \begin{vmatrix} a' & b' \\ c & d \end{vmatrix}$, et $\begin{vmatrix} a & b \\ \lambda c+\mu c' & \lambda d+\mu d' \end{vmatrix} = \lambda \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \mu \begin{vmatrix} a & b \\ c' & d' \end{vmatrix}$.

• En particulier, si on multiplie une ligne ou une colonne par un scalaire λ, on multiplie le déterminant par λ :

$$\begin{vmatrix} \lambda a & \lambda b \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ \lambda c & \lambda d \end{vmatrix} = \begin{vmatrix} \lambda a & b \\ \lambda c & d \end{vmatrix} = \begin{vmatrix} a & \lambda b \\ c & \lambda d \end{vmatrix} = \lambda \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

Il en résulte que si A a une ligne ou une colonne formée de 0, alors $\det A$ est nul :

$$\begin{vmatrix} 0 & 0 \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ 0 & 0 \end{vmatrix} = \begin{vmatrix} 0 & b \\ 0 & d \end{vmatrix} = \begin{vmatrix} a & 0 \\ c & 0 \end{vmatrix} = 0.$$

Il en résulte aussi que, si on multiplie la matrice A par un scalaire λ, on multiplie le déterminant par λ^2 :

$$\det(\lambda A) = \begin{vmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{vmatrix} = \lambda^2 \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \lambda^2 \det A.$$

• Si les deux lignes ou les deux colonnes sont égales, ou plus généralement si les deux lignes ou les deux colonnes sont multiples l'une de l'autre par un scalaire, alors le déterminant est nul :

$$\begin{vmatrix} a & b \\ a & b \end{vmatrix} = \begin{vmatrix} a & a \\ c & c \end{vmatrix} = 0. \qquad \begin{vmatrix} a & b \\ \lambda a & \lambda b \end{vmatrix} = \begin{vmatrix} a & \lambda a \\ c & \lambda c \end{vmatrix} = 0.$$

• D'après le deuxième et le quatrième point, on ne change pas la valeur du déterminant si on ajoute à une ligne (ou une colonne) un multiple de l'autre ligne (ou colonne) par un scalaire λ :

$$\begin{vmatrix} a & b \\ c+\lambda a & d+\lambda b \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \lambda \begin{vmatrix} a & b \\ a & b \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \quad \text{et} \quad \begin{vmatrix} a & b+\lambda a \\ c & d+\lambda c \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \lambda \begin{vmatrix} a & a \\ c & c \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

c) Multiplicativité.
Soient $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ et $B = \left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)$ deux matrices de $\mathcal{M}_2(\mathbb{K})$. Alors :

$$\det(AB) = \det A \times \det B.$$

En effet : $AB = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right) = \left(\begin{smallmatrix} aa'+bc' & ab'+bd' \\ ca'+dc' & cb'+dd' \end{smallmatrix}\right)$,
donc $\det(AB) = \left|\begin{smallmatrix} aa'+bc' & ab'+bd' \\ ca'+dc' & cb'+dd' \end{smallmatrix}\right| = (aa'+bc')(cb'+dd') - (ab'+bd')(ca'+dc') =$
$aa'dd' - adb'c' - bca'd' + bcb'c' = (ad-bc)(a'd'-b'c') = \left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right| \times \left|\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right|$.

d) Inversibilité. Une matrice $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ de $\mathcal{M}_2(\mathbb{K})$ est inversible si et seulement si $\det A \neq 0$ dans $\mathbb{K}$, et dans ce cas :

$$A^{-1} = \frac{1}{\det A}\left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right) \quad \text{et} \quad \det(A^{-1}) = \frac{1}{\det A}.$$

En effet : si $ad - bc \neq 0$, on vérifie immédiatement que :

$$\frac{1}{ad-bc}\left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \frac{1}{ad-bc}\left(\begin{smallmatrix} ad-bc & 0 \\ 0 & ad-bc \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) = I_2.$$

Réciproquement, si A est inversible, on a $A^{-1} \times A = I_2$, donc d'après c) :
$\det(A^{-1})\det A = \det I_2 = 1$, d'où $\det A \neq 0$ et $\det(A^{-1}) = \frac{1}{\det A}$.

5.2 Cas général des matrices $n \times n$

5.2.1 Définition du déterminant

On a défini le déterminant pour les matrices carrées d'ordre 2. Par convention, pour une matrice carrée d'ordre 1 (une telle matrice est réduite à un coefficient), c'est-à-dire $A = (a)$ avec $a \in \mathbb{K}$, on pose $\det A = a$. On va maintenant définir le déterminant pour les matrices carrées d'ordre n quelconque, par récurrence sur n, de la façon suivante.

a) Définition. (*Développement par rapport à la première ligne*)
Supposons que l'on ait défini le déterminant des matrices carrées d'ordre $n-1$, pour un certain $n \geq 3$. Fixons dans $\mathcal{M}_n(\mathbb{K})$ une matrice carrée :

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,n-1} & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots & a_{2,n-1} & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \dots & a_{3,n-1} & a_{3,n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \dots & a_{n,n-1} & a_{n,n} \end{pmatrix}. \text{ On veut définir : } \det A = \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,n-1} & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots & a_{2,n-1} & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \dots & a_{3,n-1} & a_{3,n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \dots & a_{n,n-1} & a_{n,n} \end{vmatrix}.$$

Pour tous $1 \leq i, j \leq n$, notons $A_{i,j}$ la matrice extraite de A (au sens de 4.1.4.d) obtenue en supprimant dans A la i-ième ligne et la j-ième colonne ; c'est donc une matrice carrée de $\mathcal{M}_{n-1}(\mathbb{K})$, pour laquelle la notion de déterminant est supposée être définie.
On appelle alors *déterminant* de A le scalaire :

$$\begin{aligned} \det A &= a_{1,1}\det A_{1,1} - a_{1,2}\det A_{1,2} + a_{1,3}\det A_{1,3} + \ldots + (-1)^{1+n}a_{1,n}\det A_{1,n} \\ &= \sum_{k=1}^{n}(-1)^{1+k}a_{1,k}\det A_{1,k}. \end{aligned}$$

Cette formule est dite *développement du déterminant par rapport à la première ligne.*

b) Proposition. (*Permutation des lignes*)
Si on échange deux lignes d'une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$, on multiplie son déterminant par -1.

Preuve : par récurrence sur n (vue en 5.1.b pour $n = 2$) ; cf. ouvrage de référence. □

c) Conséquence. (*Développement par rapport à une ligne quelconque*)

En appliquant le b) ci-dessus, et en reprenant toutes les notations de la définition a), on a pour tout $1 \leq i \leq n$:

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

Cette formule est dite *développement du déterminant par rapport à la i-ième ligne.*

Par exemple, pour $n = 3$, $\begin{vmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{vmatrix}$ est égal à chacun des trois développements :

- $a \begin{vmatrix} b' & b'' \\ c' & c'' \end{vmatrix} - a' \begin{vmatrix} b & b'' \\ c & c'' \end{vmatrix} + a'' \begin{vmatrix} b & b' \\ c & c' \end{vmatrix} = a(b'c'' - b''c') - a'(bc'' - b''c) + a''(bc' - b'c).$
- $-b \begin{vmatrix} a' & a'' \\ c' & c'' \end{vmatrix} + b' \begin{vmatrix} a & a'' \\ c & c'' \end{vmatrix} - b'' \begin{vmatrix} a & a' \\ c & c' \end{vmatrix} = -b(a'c'' - a''c') + b'(ac'' - a''c) - b''(ac' - a'c).$
- $c \begin{vmatrix} a' & a'' \\ b' & b'' \end{vmatrix} - c' \begin{vmatrix} a & a'' \\ b & b'' \end{vmatrix} + c'' \begin{vmatrix} a & a' \\ b & b' \end{vmatrix} = c(a'b'' - a''b') - c'(ab'' - a''b) + c''(ab' - a'b).$

d) Définition. Avec les notations du a), pour tous $1 \leq i, j \leq n$, le scalaire :

$$\Delta_{i,j} = (-1)^{i+j} \det A_{i,j}$$

est appelé le *cofacteur* associé au coefficient $a_{i,j}$ de A.

e) Mises en garde.

(i) Ainsi le développement par rapport à une ligne quelconque d'un déterminant d'ordre n ramène au calcul de n déterminants d'ordre $n-1$. On comprend que ceci aboutit très vite (dès que n est un peu grand...) à des calculs laborieux, voire inextricables. D'où l'importance des méthodes de réduction que l'on va voir plus loin, et qui aboutissent dans la pratique au principe suivant : faire sur le déterminant des transformations faisant apparaître le maximum de zéros, et ne développer que lorsqu'on ne peut plus rien faire d'autre !

(ii) On veillera particulièrement à l'alternance des signes dans l'apparition des cofacteurs (ce signe étant $+$ si la somme de l'indice de ligne et de l'indice de colonne est paire, et $-$ si elle est impaire).

$$\begin{pmatrix} + & - \\ - & + \end{pmatrix} \qquad \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix} \qquad \begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix}.$$

(iii) Enfin, on gardera bien à l'esprit que la notion de déterminant n'est définie que pour des matrices carrées !

5.2.2 Opérations sur les lignes

Proposition *(Opérations sur les lignes)*

(i) *Le déterminant est linéaire par rapport à chaque ligne, ce qui signifie que :*
- *si l'on multiplie par $\lambda \in \mathbb{K}$ une ligne de la matrice A, on multiplie son déterminant par λ ;*
- *si A, A', A'' sont trois matrices carrées d'ordre n telles que, pour un certain $1 \leq i \leq n$, la i-ième ligne de A soit la somme de la i-ième ligne de A' et de la i-ième ligne de A'', et telles que toutes les autres lignes soient les mêmes pour les trois matrices, alors* $\det A = \det A' + \det A''$.

(ii) *Si une matrice a une ligne entière de zéros, alors son déterminant est nul.*

(iii) *Si l'on multiplie par un scalaire λ une matrice A carrée d'ordre n, on multiplie son déterminant par λ^n.*

(iv) *Si une matrice carrée a deux lignes identiques, alors son déterminant est nul.*

(v) *Si dans une matrice carrée, on ajoute à une ligne une combinaison linéaire des autres lignes, on ne change pas la valeur de son déterminant.*

(vi) *Si dans une matrice carrée, une ligne est combinaison linéaire des autres lignes, alors son déterminant est nul.*

Preuve : (i) se démontre sans problème par récurrence.

On déduit immédiatement du premier point de (i) d'une part la propriété (ii) [en multipliant une ligne par $\lambda = 0$] et d'autre part la propriété (iii) [en multipliant par λ chacune des n lignes de A]. Pour montrer (iv), supposons que A ait deux lignes identiques ; si on échange ces deux lignes, on multiplie $\det A$ par -1 d'après la proposition 5.2.1.b, mais la matrice reste la même puisque les deux lignes échangées sont identiques ; donc $\det A = -\det A$, c'est-à-dire $\det A = 0$. La propriété (v) résulte immédiatement de (i) et (iv), et la propriété (vi) découle de (v) et (ii). □

5.2.3 Transposition (Opérations sur les colonnes)

a) Théorème. *Toute matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ a le même déterminant que sa transposée. En d'autres termes :* $\forall\, A \in \mathcal{M}_n(\mathbb{K}),\ \det({}^tA) = \det A$.

Preuve : c'est clair pour $n \leq 3$; pour une preuve générale, cf. ouvrage de référence. □

Avec ce théorème, on déduit immédiatement de 5.2.1.b, 5.2.1.c et 5.2.2 les résultats suivants :

b) Corollaire. *Les règles de calcul démontrées sur les lignes aux propositions 5.2.1.b et 5.2.2 restent vraies en remplaçant "ligne" par "colonne".*

c) Corollaire. *Soit $A = (a_{i,j})_{1\leq i,j\leq n}$ une matrice carrée d'ordre n. Pour tout $1 \leq j \leq n$, on a la formule suivante, dite développement de A par rapport à la j-ième colonne :*

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

où $(-1)^{k+j} \det A_{k,j}$ désigne toujours le cofacteur associé au coefficient $a_{k,j}$.

5.2.4 Exemples de calculs

a) On commence par un résultat général sur les matrices triangulaires (voir 3.1.8.c), qui s'applique donc aussi en particulier aux matrices diagonales (voir 3.1.8.a), et qui est très utile dans la pratique.

Proposition. *Le déterminant d'une matrice triangulaire est égal au produit de ses coefficients diagonaux.*

Preuve : Dans la cas d'une matrice triangulaire supérieure, d'ordre n, on développe le déterminant par rapport à la première colonne, et on réitère n fois :

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n-1} & a_{1,n} \\ 0 & a_{2,2} & a_{2,3} & \cdots & a_{2,n-1} & a_{2,n} \\ 0 & 0 & a_{3,3} & \cdots & a_{3,n-1} & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & 0 & \cdots & 0 & a_{n,n} \end{vmatrix} = a_{1,1} \begin{vmatrix} a_{2,2} & a_{2,3} & \cdots & a_{2,n-1} & a_{2,n} \\ 0 & a_{3,3} & \cdots & a_{3,n-1} & a_{3,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \cdots & 0 & a_{n,n} \end{vmatrix} = \cdots = a_{1,1}a_{2,2}\cdots a_{n,n}.$$

Le cas d'une matrice triangulaire inférieure s'en déduit à l'aide de 5.2.3.a. □

b) En appliquant les règles 5.2.1.b et 5.2.2, ou leurs analogues pour les colonnes, on cherche à faire apparaître le maximum de zéros, et à se ramener à une forme triangulaire pour développer à l'aide de la proposition a) ci-dessus.

Exemple : calculer dans $\mathbb{R}$ les déterminants : $D_1 = \begin{vmatrix} a & 0 & 3 & 5 \\ 0 & 0 & b & 2 \\ 1 & c & 2 & 3 \\ 0 & 0 & 0 & d \end{vmatrix}$ et $D_2 = \begin{vmatrix} a & a & a & a \\ a & b & b & b \\ a & b & c & c \\ a & b & c & d \end{vmatrix}$.

Pour D_1, on échange les lignes l_1 et l_3, puis les lignes l_2 et l_3, puis les colonnes c_1 et c_2 :

$$D_1 = \begin{vmatrix} a & 0 & 3 & 5 \\ 0 & 0 & b & 2 \\ 1 & c & 2 & 3 \\ 0 & 0 & 0 & d \end{vmatrix} = - \begin{vmatrix} 1 & c & 2 & 3 \\ 0 & 0 & b & 2 \\ a & 0 & 3 & 5 \\ 0 & 0 & 0 & d \end{vmatrix} = \begin{vmatrix} 1 & c & 2 & 3 \\ a & 0 & 3 & 5 \\ 0 & 0 & b & 2 \\ 0 & 0 & 0 & d \end{vmatrix} = - \begin{vmatrix} c & 1 & 2 & 3 \\ 0 & a & 3 & 5 \\ 0 & 0 & b & 2 \\ 0 & 0 & 0 & d \end{vmatrix} = -abcd.$$

Pour D_2, on remplace d'abord c_2 par $c_2 - c_1$, c_3 par $c_3 - c_1$ et c_4 par $c_4 - c_1$; ensuite on remplace c_3 par $c_3 - c_2$ et c_4 par $c_4 - c_2$; enfin, on remplace c_4 par $c_4 - c_3$.

$$D_2 = \begin{vmatrix} a & 0 & 0 & 0 \\ a & b-a & b-a & b-a \\ a & b-a & c-a & c-a \\ a & b-a & c-a & d-a \end{vmatrix} = \begin{vmatrix} a & 0 & 0 & 0 \\ a & b-a & 0 & 0 \\ a & b-a & c-b & c-b \\ a & b-a & c-b & d-b \end{vmatrix} = \begin{vmatrix} a & 0 & 0 & 0 \\ a & b-a & 0 & 0 \\ a & b-a & c-b & 0 \\ a & b-a & c-b & d-c \end{vmatrix} = a(b-a)(c-b)(d-c).$$

c) Calculons dans $\mathbb{C}$ le déterminant : $D_3 = \begin{vmatrix} a+i\alpha & \alpha+ia & a+\alpha \\ b+i\beta & \beta+ib & b+\beta \\ c+i\gamma & \gamma+ic & c+\gamma \end{vmatrix}$. On applique la linéarité par rapport à chaque colonne :

$$D_3 = \begin{vmatrix} a & \alpha+ia & a+\alpha \\ b & \beta+ib & b+\beta \\ c & \gamma+ic & c+\gamma \end{vmatrix} + i \begin{vmatrix} \alpha & \alpha+ia & a+\alpha \\ \beta & \beta+ib & b+\beta \\ \gamma & \gamma+ic & c+\gamma \end{vmatrix} = \begin{vmatrix} a & \alpha & a+\alpha \\ b & \beta & b+\beta \\ c & \gamma & c+\gamma \end{vmatrix} + i \begin{vmatrix} a & a & a+\alpha \\ b & b & b+\beta \\ c & c & c+\gamma \end{vmatrix} + i \begin{vmatrix} \alpha & \alpha & a+\alpha \\ \beta & \beta & b+\beta \\ \gamma & \gamma & c+\gamma \end{vmatrix} + i^2 \begin{vmatrix} \alpha & a & a+\alpha \\ \beta & b & b+\beta \\ \gamma & c & c+\gamma \end{vmatrix}.$$

Le 2ième et le 3ième déterminants sont nuls car ils ont deux colonnes identiques ; le 1er et le 4ième déterminants sont nuls car la colonne c_3 est somme des colonnes c_1 et c_2. On conclut $D_3 = 0$.

5.2.5 Multiplicativité et inversibilité

a) Théorème fondamental.

(i) *Pour toutes matrices carrées* $A, B \in \mathcal{M}_n(\mathbb{K})$, *on a :* $\det(AB) = \det A \times \det B$.
En particulier : $\det(AB) = \det(BA)$.

(ii) *Une matrice carrée* $A \in \mathcal{M}_n(\mathbb{K})$ *est inversible si et seulement si* $\det A \neq 0$.
De plus, dans ce cas, $\det(A^{-1}) = (\det A)^{-1}$.

Preuve : on l'a vu en 5.1.c et 5.1.d pour $n = 2$. On pourra à titre d'exercice écrire une preuve élémentaire pour $n = 3$. Pour une preuve générale pour n quelconque, voir ouvrage de référence.□

b) Corollaire. *Deux matrices carrées semblables ont le même déterminant.*

Preuve : Rappelons (cf. 3.1.6.b) que A et B sont semblables signifie que :
$A = PBP^{-1}$avec $P \in GL(n, \mathbb{K})$. On a alors :
$\det A = \det P \times \det B \times \det(P^{-1}) = \det P \times \det B \times (\det P)^{-1}$
$= \det P \times (\det P)^{-1} \times \det B = \det B$. □

On peut préciser le point (ii) du théorème a) par la proposition d) suivante :

c) Définition. Pour toute matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$, on appelle *comatrice* de A, notée $\mathrm{Com}\, A$ la matrice carrée :

$$\mathrm{Com}\, A = (\Delta_{i,j})_{1 \leq i,j \leq n} \ \in \mathcal{M}_n(\mathbb{K})$$

dont le coefficient général est, sur la i-ième ligne et la j-ième colonne, le cofacteur :

$$\Delta_{i,j} = (-1)^{i+j} \det A_{i,j}$$

de la matrice A tel qu'on l'a défini plus haut en 5.2.1.d.

d) Proposition. *Soit* $A \in GL(n, \mathbb{K})$ *une matrice carrée d'ordre* n *inversible ; alors :*

$$A^{-1} = \frac{1}{\det A}\ {}^t(\mathrm{Com}\, A).$$

Preuve : on l'a vu en 5.1.d pour $n = 2$. On pourra à titre d'exercice écrire une preuve élémentaire pour $n = 3$. Pour une preuve générale pour n quelconque, voir ouvrage de référence. □

e) Exemple. Soit $A = \begin{pmatrix} 1 & 4 & 2 \\ 0 & 3 & 1 \\ 2 & 9 & 4 \end{pmatrix}$; on a : $\det A = 1 \times (12 - 9) - 4 \times (0 - 2) + 2 \times (0 - 6) = -1$.
On calcule $\mathrm{Com}\, A = \begin{pmatrix} 3 & 2 & -6 \\ 2 & 0 & -1 \\ -2 & -1 & 3 \end{pmatrix}$; on conclut $A^{-1} = -{}^t(\mathrm{Com}\, A) = \begin{pmatrix} -3 & -2 & 2 \\ -2 & 0 & 1 \\ 6 & 1 & -3 \end{pmatrix}$.

f) Remarque. La proposition d) a un grand intérêt théorique, et on en verra d'autres applications en travaux dirigés. Mais, dans la pratique, elle conduit souvent à des calculs nombreux et complexes (par exemple calculer l'inverse d'une matrice 4×4 revient à calculer 16 déterminants 3×3). C'est pourquoi, sur des exemples concrets de matrices carrées à coefficients numériques explicites, il est souvent beaucoup plus rapide d'utiliser la méthode vue en 3.2.4.

5.3 Applications

5.3.1 Déterminant d'un endomorphisme ; application à la bijectivité

a) Lemme. *Soient $E = \mathbb{K}^n$ et f un endomorphisme de E. Pour toute base $\mathcal{B}$ de E, le déterminant de la matrice de f par rapport à la base $\mathcal{B}$ garde la même valeur.*

Preuve : Prenons $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E, et P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$.
D'après 3.3.4.c, en notant $A = M_{\mathcal{B}}(f)$ et $A' = M_{\mathcal{B}'}(f)$, on a $A = PA'P^{-1}$. Donc $\det A = \det A'$ d'après 5.2.5.b. □

b) Définition. Soient $E = \mathbb{K}^n$ et f un endomorphisme de E. La valeur commune du déterminant de la matrice de f par rapport à une base quelconque de E est appelé le déterminant de f. On le note : $\det f$.

c) Proposition. *Soient $E = \mathbb{K}^n$ et f un endomorphisme de E. Alors f est un automorphisme (c'est-à-dire f est bijectif) si et seulement si* $\det f \neq 0$.

Preuve : résulte immédiatement de 3.2.3.b et 5.2.5.a.(ii). □

5.3.2 Application à l'indépendance linéaire d'une famille de n vecteurs

a) Proposition. *Soient $E = \mathbb{K}^n$ et $\mathcal{C}$ une famille de n vecteurs de E. Alors $\mathcal{C}$ est une base de E si et seulement si la matrice de $\mathcal{C}$ dans une base quelconque de E a un déterminant non-nul.*

Preuve : notons A la matrice de la famille $\mathcal{C}$ dans une base $\mathcal{B}$ au sens de 3.3.1. Donc A est carrée d'ordre n. D'après 3.3.2.a, $\mathcal{C}$ est une base de E si et seulement si A est inversible. D'après 5.2.5.a.(ii), cela équivaut à $\det A \neq 0$. □

b) Remarque de synthèse. On a donné en 4.1.6.a différentes façons de traduire (en termes de famille de vecteurs, d'endomorphisme, de système linéaire homogène) l'inversibilité d'une matrice carrée. Aux cinq assertions (i) à (v) de ce théorème, on peut donc maintenant ajouter en (vi) la propriété : $\det A \neq 0$.

5.3.3 Application à certains systèmes linéaires : formules de Cramer

a) Données. Considérons un système linéaire (S) de n équations à n inconnues dans $\mathbb{K}$, et notons (S_0) le système homogène associé.

$$(S) : \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = b_2 \\ \cdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n = b_n \end{cases}, \qquad (S_0) : \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = 0 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = 0 \\ \cdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n = 0 \end{cases},$$

que l'on peut noter matriciellement :

$$(S) : \ A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}, \qquad (S_0) : \ A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

en notant $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{K})$ la matrice carrée des coefficients.
Un n-uplet $(x_1, \ldots, x_n)$ de $\mathbb{K}^n$ qui satisfait dans $\mathbb{K}$ chacune des n équations de (S) s'appelle une *solution* de (S). On sait (cf. 2.3.3 et 4.1.3) que l'ensemble des solutions de (S_0) est un ss-e.v. de

$\mathbb{K}^n$, ce qui n'est pas le cas de l'ensemble des solutions de (S) lorsque le second membre (la matrice colonne des b_i) est non-nul. On appelle *rang* de (S), noté $\text{rg}(S)$, le rang de (S_0) au sens de 4.1.3, c'est-à-dire le rang de la matrice A.

b) Définition. Avec les notations précédentes, on dit que (S) est un *système de Cramer* lorsqu'il est de rang maximum, c'est-à-dire de rang n. D'après 4.1.6.a, cela est équivalent à dire que A est inversible, ou encore d'après 5.2.5.a.(ii), à dire que $\det A \neq 0$.

c) Théorème. *Avec les notations précédentes, (S) est un système de Cramer si et seulement si (S) admet dans $\mathbb{K}^n$ une unique solution. De plus, cette unique solution $(x_1, \dots, x_n)$ est donnée par les formules de Cramer :*

$$x_i = \frac{\det(A|B)_i}{\det A} \qquad \textit{pour tout } 1 \leq i \leq n,$$

où $(A|B)_i$ désigne la matrice carrée d'ordre n obtenue en remplaçant dans A la i-ième colonne par la colonne du second membre de (S).

Preuve : l'unicité est claire, car la différence de deux solutions de (S) est une solution de (S_0), et ce système homogène n'admet que la solution nulle lorsque A est inversible (cf. 4.1.6.a). Pour la preuve des formules de Cramer, cf. ouvrage de référence. □

d) Exemple. Soit (S) : $\left\{\begin{array}{lllll} (2a+1)x - & ay & + (a+1)z & = a-1 & (1) \\ (a-2)x + & (a-1)y & + (a-2)z & = a & (2) \\ (2a-1)x + & (a-1)y & + (2a-1)z & = a & (3) \end{array}\right.$ $\qquad$ ($a \in \mathbb{R}$ paramètre fixé).

Pour la matrice A de (S), on calcule $\det A = a(a-1)(a+1)$. D'où quatre cas :

Premier cas : si $a \neq 0, a \neq 1, a \neq -1$; alors $\det A \neq 0$ donc (S) est de Cramer et son unique solution (x_0, y_0, z_0) dans $\mathbb{R}$ est donnée par :

$$x_0 = \frac{1}{\det A}\begin{vmatrix} a-1 & -a & a+1 \\ a & a-1 & a-2 \\ a & a-1 & 2a-1 \end{vmatrix}, \quad y_0 = \frac{1}{\det A}\begin{vmatrix} 2a+1 & a-1 & a+1 \\ a-2 & a & a-2 \\ 2a-1 & a & 2a-1 \end{vmatrix}, \quad z_0 = \frac{1}{\det A}\begin{vmatrix} 2a+1 & -a & a-1 \\ a-2 & a-1 & a \\ 2a-1 & a-1 & a \end{vmatrix},$$

que l'on simplifie après calculs en : $x_0 = \frac{2a^2-2a+1}{a(a-1)}$, $y_0 = \frac{a}{a-1}$, $z_0 = -\frac{2a^2-2a+1}{a(a-1)}$.

Deuxième cas : si $a = 0$, les équations (1), (2) et (3) sont incompatibles ; (S) n'a pas de solution.
Troisième cas : si $a = 1$, idem.
Quatrième cas : si $a = -1$, les équations (2) et (3) sont équivalentes ; (S) est de rang 2. L'ensemble des solutions de (S) est $\{(-\frac{3}{5}z + 1, -\frac{3}{5}z - 1, z) \,;\, z \in \mathbb{R}\}$. On dit qu'il y a indétermination d'ordre 1.

d) Remarque. On verra en TD d'autres applications des déterminants pour résoudre des systèmes d'équations linéaires dont la matrice n'est pas nécessairement carrée (c'est-à-dire que le nombre d'équations n'est pas nécessairement égal au nombre d'inconnues).

5.4 Exercices sur le chapitre 5

Exercice 1. Calculer dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ les déterminants suivants :

$$D_1 = \begin{vmatrix} 1 & 0 & -3 \\ 4 & 1 & -1 \\ 2 & -1 & 1 \end{vmatrix}, \quad D_2 = \begin{vmatrix} -3 & 1 & 0 \\ 1 & 0 & 1 \\ -2 & -2 & 5 \end{vmatrix}, \quad D_3 = \begin{vmatrix} 1+i & 0 & -1 \\ 1-i & i & i \\ -2 & 2i & -i \end{vmatrix}, \quad D_4 = \begin{vmatrix} 3 & -2 & -5 & 4 \\ -5 & 2 & 8 & -5 \\ -2 & 4 & 7 & -3 \\ 2 & -3 & -5 & 8 \end{vmatrix}.$$

Exercice 2. Calculer dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ les déterminants suivants :

$$\Delta_1 = \begin{vmatrix} x & y & x+y \\ y & x+y & x \\ x+y & x & y \end{vmatrix}, \quad \Delta_2 = \begin{vmatrix} a & b & b & b \\ b & a & b & b \\ b & b & a & b \\ b & b & b & a \end{vmatrix}, \quad \Delta_3 = \begin{vmatrix} a & a & a & a \\ a & b & b & b \\ a & b & c & c \\ a & b & c & d \end{vmatrix}, \quad \Delta_4 = \begin{vmatrix} a & b & c & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{vmatrix}, \quad \Delta_5 = \begin{vmatrix} a & b & c & d \\ -a & b & 1 & 1 \\ -a & -b & c & 1 \\ -a & -b & -c & d \end{vmatrix},$$

$$\Delta_6 = \begin{vmatrix} x & 0 & -1 & 1 & 0 \\ 1 & x & -1 & 1 & 6 \\ 1 & 0 & x-1 & 0 & 6 \\ 0 & 1 & -1 & x & 1 \\ 0 & 1 & -1 & 0 & x \end{vmatrix}, \quad \Delta_7 = \begin{vmatrix} a-b-c & 2a & 2a \\ 2b & -a+b-c & 2b \\ 2c & 2c & -a-b+c \end{vmatrix}, \quad \Delta_8 = \begin{vmatrix} (b+c)^2 & a^2 & a^2 \\ b^2 & (a+c)^2 & b^2 \\ c^2 & c^2 & (a+b)^2 \end{vmatrix}.$$

Exercice 3. a) On considère dans $\mathcal{M}_4(\mathbb{R})$ la matrice $A = \begin{pmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{pmatrix}$. Calculer le produit de A par sa transposée, et en déduire $\det A$.

b) On considère dans $\mathcal{M}_3(\mathbb{C})$ les matrices $M = \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}$ et $J = \begin{pmatrix} 1 & 1 & 1 \\ 1 & j & j^2 \\ 1 & j^2 & j \end{pmatrix}$.

[Rappelons que $j = -\frac{1}{2} + i\frac{\sqrt{3}}{2} = \exp(\frac{2i\pi}{3})$, et qu'il vérifie $j^3 = 1$ et $1 + j + j^2 = 0$].
Calculer MJ et $\det J$; en déduire $\det M$.

Exercice 4. On considère dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ les déterminants suivants :

$$\Lambda_n = \begin{vmatrix} x & a_1 & a_2 & a_3 & \dots & a_{n-1} & 1 \\ a_1 & x & a_2 & a_3 & \dots & a_{n-1} & 1 \\ a_1 & a_2 & x & a_3 & \dots & a_{n-1} & 1 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_1 & a_2 & a_3 & a_4 & \dots & x & 1 \\ a_1 & a_2 & a_3 & a_4 & \dots & a_n & 1 \end{vmatrix}, \quad \Gamma_n = \begin{vmatrix} 1 & a & a^2 & a^3 & \dots & a^n \\ a & 1 & a & a^2 & \dots & a^{n-1} \\ a^2 & a & 1 & a & \dots & a^{n-2} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ a^n & a^{n-1} & a^{n-2} & a^{n-3} & \dots & 1 \end{vmatrix}, \quad \Delta_n = \begin{vmatrix} a & c & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ b & a & c & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & b & a & c & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & b & a & c & & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & a & c & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & b & a & c \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & b & a \end{vmatrix}.$$

Calculer Λ_n et Γ_n. Ecrire une relation de récurrence liant Δ_n à Δ_{n-1} et Δ_{n-2}. En déduire la valeur de Δ_n lorsque $a = x + y, b = 1, c = xy$; puis lorsque $a = 1 + x^2, b = c = x$.

Exercice 5. Montrer qu'une matrice carrée d'ordre impair antisymétrique à coefficients dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ n'est jamais inversible.

Exercice 6. Montrer que les matrices suivantes sont inversibles, et calculer leur inverse par la méthode des cofacteurs.

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}, \quad D = \begin{pmatrix} -1+i & 1-i & -1 \\ 1-2i & -1+i & -i \\ 2-i & 0 & 1-i \end{pmatrix}.$$

Exercice 7. On considère dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ le déterminant : $\Delta = \begin{vmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{vmatrix}$.

a) En développant par rapport à la première ligne, montrer qu'il existe un polynôme $P(x)$ de degré ≤ 3 à coefficients dans $\mathbb{K}$ tel que $\Delta = P(a)$.
b) Calculer $P(b), P(c)$ et $P(d)$. En déduire Δ.
c) Généraliser à l'ordre n quelconque.

Exercice 8. Résoudre dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ les systèmes suivants :

$$\begin{cases} x + y + z = 1 \\ x + 2y - z = -1 \\ 2x + y + 4z = 4 \end{cases}, \quad \begin{cases} 3x + 2y + z = 3 \\ x + 3y + 2z = 0 \\ 2x + y + 3z = -3 \end{cases}, \quad \begin{cases} x + 2y - z = 1 \\ 2x + y + 2z = 2 \\ x - 4y + 7z = 3 \end{cases}, \quad \begin{cases} x + y + z + t = 0 \\ x + y + z - t = 4 \\ x + y - z + t = -4 \\ x - y + z + t = 2 \end{cases}.$$

Exercice 9. Résoudre en discutant suivant les valeurs des paramètres les systèmes :

$$\left\{\begin{array}{l} mx + y + z = 1 \\ x + my + z = 1 \\ x + y + mz = 1 \end{array}\right., \qquad \left\{\begin{array}{l} x + y + mz = 2 \\ 3x + 4y + 2z = m \\ 2x + 3y - z = 1 \end{array}\right., \qquad \left\{\begin{array}{l} (-3m+5)x + (2m+1)y + (4m-5) = 1 \\ (-2m+6)x + 2my + (4m-6) = m \\ (-m+7)x + (4m-1)y + (6m-7) = -1 \end{array}\right.,$$

$$\left\{\begin{array}{l} px + (p-3)y + qz = 2p-7 \\ 3x + (2p-7)y + qz = 1 \\ 2x + (p-3)y + qz = 1 \end{array}\right., \qquad \left\{\begin{array}{l} ax + y + z + t = 1 \\ x + ay + z + t = a \\ x + y + az + t = a^2 \\ x + y + z + at = a^3 \end{array}\right., \qquad \left\{\begin{array}{l} x + z + t = 1 \\ y + 2z + t = 1 \\ 2x - y + t = 1 \\ 2px + pt = q \end{array}\right..$$

Exercice 10. Déterminer l'unique polynôme $P(x)$ de degré 3 à coefficients réels tel que :

$$P(0) = 0, \quad P(1) = 1, \quad P(2) = 16, \quad P(3) = 81.$$

Exercice 11. Soit n un entier ≥ 2. On considère le système de n équations à n inconnues $x_1, x_2, \ldots, x_n$ défini par : $x_i + x_{i+1} = 2$ pour tout $1 \leq i \leq n-1$ et $x_n + x_1 = 2$. Quel est son rang ? Dans quel cas est-il de Cramer ?

Application : Soient dans le plan rapporté à un repère les quatre points $A(1,0)$, $B(0,0)$, $C(0,1)$ et $D(1,1)$. Montrer qu'il existe un unique triangle dont les milieux des côtés sont A, B, C, mais qu'il existe une infinité de quadrilatères dont les milieux des côtés sont A, B, C, D.

Chapitre 6

Réduction des endomorphismes (1)

6.1 Notions générales sur les valeurs propres

6.1.1 Remarques préliminaires et notations

$\mathbb{K}$ désigne toujours soit le corps $\mathbb{R}$ des réels, soit le corps $\mathbb{C}$ des complexes.

a) On a déjà vu dans les chapitres précédents que les calculs sur les matrices triangulaires, et parmi elles les matrices diagonales, sont particulièrement simples. Rappelons que :
• le produit de deux matrices triangulaires est triangulaire (cf. 3.1.8.c) ; et en particulier (cf. 3.1.8.a) pour deux matrices diagonales :

$$\begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & a_n \end{pmatrix} \begin{pmatrix} b_1 & 0 & \dots & 0 \\ 0 & b_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & b_n \end{pmatrix} = \begin{pmatrix} b_1 & 0 & \dots & 0 \\ 0 & b_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & b_n \end{pmatrix} \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & a_n \end{pmatrix} = \begin{pmatrix} a_1b_1 & 0 & \dots & 0 \\ 0 & a_2b_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & a_nb_n \end{pmatrix}$$

• le déterminant d'une matrice triangulaire (et donc en particulier d'une matrice diagonale) est égal au produit de ses coefficients diagonaux (cf. 5.2.4.a). Il en résulte avec 5.2.5.a qu'une matrice triangulaire (et donc en particulier une matrice diagonale) est inversible si et seulement si tous ses coefficients diagonaux sont non-nuls.

b) On comprend donc que, pour représenter un endomorphisme par sa matrice A dans une base $\mathcal{B}$, on a tout intérêt à choisir $\mathcal{B}$ telle que A soit diagonale, ou à défaut triangulaire. Encore faut-il que cela soit possible, c'est-à-dire qu'une telle base existe. C'est ce problème (pour un endomorphisme f donné, déterminer une base dans laquelle la matrice de f soit la plus simple possible : diagonale, triangulaire, ... ou d'autres formes standard que l'on verra plus loin) que l'on va aborder dans ce chapitre. Les applications (en analyse et en géométrie en particulier) sont extrêmement importantes. Nous en verrons quelques-unes en cours ou en T.D.

c) On a déjà introduit dans les chapitres précédents divers *invariants* pour les matrices carrées : il s'agit d'objets mathématiques qui ne changent pas quand on remplace $A \in \mathcal{M}_n(\mathbb{K})$ par une matrice semblable (au sens de 3.1.6.b), c'est-à-dire une matrice carrée $A' = P^{-1}AP$, avec $P \in GL(n, \mathbb{K})$. D'après 3.3.4.c, cela signifie que ces objets sont en fait attachés à l'endomorphisme f de E (où E est un $\mathbb{K}$-e.v. de dimension n, par exemple $\mathbb{K}^n$), dont A ou A' sont les matrices dans deux bases de E. Citons :
• le rang, qui est un entier (cf. chap. 4) : $\mathrm{rg}\, A = \mathrm{rg}(P^{-1}AP) = \mathrm{rg}\, f$;
• la trace, qui est un scalaire (cf. 3.1.9) : $\mathrm{tr}\, A = \mathrm{tr}(P^{-1}AP) = \mathrm{tr}\, f$;
• le déterminant, qui est aussi un scalaire (cf. chap. 5) : $\det A = \det(P^{-1}AP) = \det f$.

On va dans ce chapitre définir un autre invariant, qui est cette fois un polynôme (et qui en un certain sens, que l'on précisera plus loin, "contient" les précédents).

$$\boxed{\text{Dans toute la suite, } n \text{ est un entier fixé } \geq 1, \text{ et } E \text{ désigne le } \mathbb{K}\text{-e.v. } \mathbb{K}^n.}$$

6.1.2 Valeurs propres, vecteurs propres, sous-espaces propres

Soit f un endomorphisme de E.

a) Définition.
• Une *valeur propre* de f est un scalaire $\lambda \in \mathbb{K}$ pour lequel il existe un vecteur $v \in E$ non-nul tel que $f(v) = \lambda.v$. On utilisera l'abréviation v.p. pour valeur propre.
• Si λ est une v.p. de f, on appelle *vecteur propre* de f associé à λ tout vecteur $v \in E$ non-nul tel que $f(v) = \lambda.v$.
• On appelle *spectre* de f, noté $\mathrm{spec} f$, l'ensemble des v.p. de f.

b) Remarque. Si v est un vecteur propre associé à une valeur propre λ, alors v ne peut pas être vecteur propre pour une autre valeur propre $\mu \neq \lambda$. En effet, $f(v) = \lambda.v = \mu.v$ implique $(\lambda - \mu).v = 0_E$, ce qui, puisque $v \neq 0_E$, conduit nécessairement à $\lambda = \mu$.

c) Remarque. Soit λ un scalaire. Notons $E_\lambda = \mathrm{Ker}(f - \lambda.\,\mathrm{id}_E)$.
Par définition, E_λ est l'ensemble des vecteurs $v \in E$ tels que $f(v) = \lambda.v$. Donc, dire que λ est une v.p. de f signifie que le ss-e.v. E_λ n'est pas réduit à $\{0_E\}$. En d'autres termes, λ est une v.p. de f si et seulement si l'endomorphisme $(f - \lambda.\,\mathrm{id}_E)$ n'est pas injectif.

d) Définition. Pour toute v.p. λ de f, on appelle *sous-espace propre* de f associé à la v.p. λ le sous-espace vectoriel $E_\lambda = \mathrm{Ker}(f - \lambda.\,\mathrm{id}_E)$ de E.
Ce ss-e.v. E_λ est donc l'ensemble des vecteurs propres associés à λ (qui par définition sont tous non-nuls), plus le vecteur nul (qui n'est pas un vecteur propre, mais qui appartient quand même au sous-espace propre). Par définition, on a : $1 \leq \dim E_\lambda \leq n$.

6.1.3 Polynôme caractéristique

a) Définition. Soit f un endomorphisme de E. On appelle *polynôme caractéristique de* f le polynôme $P_f(x) = \det(f - x.\,\mathrm{id}_E)$.

b) Théorème. *Soit f un endomorphisme de E. Les valeurs propres de f dans $\mathbb{K}$ sont exactement les zéros dans $\mathbb{K}$ de son polynôme caractéristique $P_f(x)$.*

Preuve. Soit $\lambda \in \mathbb{K}$. On a vu en 6.1.2.c que λ est une v.p. de f si et seulement si l'endomorphisme $(f - \lambda.\,\mathrm{id}_E)$ n'est pas injectif. Comme $(f - \lambda.\,\mathrm{id}_E)$ est un endomorphisme de E qui est de dimension finie, cela est équivalent d'après 2.1.5.f à dire que $(f - \lambda.\,\mathrm{id}_E)$ n'est pas bijectif, et donc, d'après 5.3.1.c, à $\det(f - \lambda.\,\mathrm{id}_E) = 0$. En résumé, les v.p. de f sont exactement les scalaires λ tels que $P_f(\lambda) = 0$. □

c) Remarque. Ainsi, déterminer les v.p. de f revient, après avoir calculé le polynôme $P_f(x)$, à chercher ses zéros. Lorsque $\mathbb{K} = \mathbb{C}$, il y a forcément des v.p. car tout polynôme à coefficients complexes admet des zéros dans $\mathbb{C}$. Ce n'est plus le cas si $\mathbb{K} = \mathbb{R}$ (par exemple un polynôme à coefficients réels de degré 2 de discriminant < 0 n'a pas de zéro réel). De plus, comme il est clair que $P_f(x)$ est de degré n, il admet forcément au plus n zéros dans $\mathbb{K}$. Donc f admet au plus n valeurs propres dans $\mathbb{K}$.

d) Dans la pratique : pour calculer $P_f(x)$, on choisit une base quelconque $\mathcal{B}$ de E. En notant A la matrice de f par rapport à la base $\mathcal{B}$, on sait (d'après 3.2.2.b) que la matrice par rapport à la base $\mathcal{B}$ de l'endomorphisme $(f - x.\,\mathrm{id}_E)$ est $(A - x.I_n)$. On a alors $\det(f - x.\,\mathrm{id}_E) = \det(A - x.I_n)$, et ceci (comme on l'a montré en 5.3.1.a) quelle que soit la base $\mathcal{B}$ choisie.
Réciproquement, toute matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ peut être considérée comme la matrice d'un certain endomorphisme f de $E = \mathbb{K}^n$ par rapport à la base canonique (cf. 3.2.2.a). On appelle *polynôme caractéristique* de la matrice A le polynôme caractéristique de f, c'est-à-dire le polynôme $P_A(x) = \det(A - x.I_n)$. Ses zéros dans $\mathbb{K}$ seront appelés les *valeurs propres* de la matrice A dans $\mathbb{K}$; ce sont bien sûr les valeurs propres de l'endomorphisme f.

e) Proposition. *Soit f un endomorphisme de E. Son polynôme caractéristique $P_f(x) = \det(f - x.\mathrm{id}_E)$ est un polynôme de degré n, à coefficients dans $\mathbb{K}$, tel que :*
- *le coefficient de x^n est $(-1)^n$;*
- *le coefficient de x^{n-1} est $(-1)^{n-1}\operatorname{tr} f$;*
- *le coefficient constant est $\det f$.*

On a donc :
$$P_f(x) = (-1)^n x^n + (-1)^{n-1}(\operatorname{tr} f)x^{n-1} + \cdots + \det f.$$

En d'autres termes, pour toute matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$, on a :
$$P_A(x) = (-1)^n x^n + (-1)^{n-1}(\operatorname{tr} A)x^{n-1} + \cdots + \det A.$$

Preuve : pour $n = 2$, on a bien $P_A(x) = \begin{vmatrix} a-x & b \\ c & d-x \end{vmatrix} = x^2 - x(a+d) + (ad - bc)$.

Pour $n \geq 3$, développons par rapport à la première ligne :
$$\begin{vmatrix} a_{1,1}-x & a_{1,2} & a_{1,3} & \dots & a_{1,n-1} & a_{1,n} \\ a_{2,1} & a_{2,2}-x & a_{2,3} & \dots & a_{2,n-1} & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3}-x & \dots & a_{3,n-1} & a_{3,n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \dots & a_{n,n-1} & a_{n,n}-x \end{vmatrix} = (a_{1,1} - x) \begin{vmatrix} a_{2,2}-x & a_{2,3} & \dots & a_{2,n-1} & a_{2,n} \\ a_{3,2} & a_{3,3}-x & \dots & a_{3,n-1} & a_{3,n} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n,2} & a_{n,3} & \dots & a_{n,n-1} & a_{n,n}-x \end{vmatrix} + \cdots$$

où le reste désigné par les points de suspension est de degré $< n-1$. En d'autres termes, $P_A(x) = (a_{1,1} - x)P_{A'}(x) + \cdots$, où A' est la matrice carrée d'ordre $n-1$ extraite de A obtenue en supprimant la première ligne et la première colonne.

Par récurrence, supposons que : $P_{A'}(x) = (-1)^{n-1}x^{n-1} + (-1)^n(\operatorname{tr} A')x^{n-2} + \cdots$.

Donc : $$\begin{aligned} P_A(x) &= (-x + a_{1,1})[(-1)^{n-1}x^{n-1} + (-1)^n(\operatorname{tr} A')x^{n-2} + \cdots] + \cdots \\ &= (-1)^n x^n + [a_{1,1}(-1)^{n-1} - (-1)^n(\operatorname{tr} A')]x^{n-1} + \cdots \\ &= (-1)^n x^n + (-1)^{n-1}\underbrace{(a_{1,1} + \operatorname{tr} A')}_{\operatorname{tr} A}x^{n-1} + \cdots \end{aligned}$$

ce qui prouve les deux premiers points de la proposition. Le troisième est clair en évaluant le polynôme $P_A(x) = \det(A - x.I_n)$ en $x = 0$. □

6.1.4 Multiplicité d'une valeur propre

a) Rappel sur les polynômes. On note $\mathbb{K}[x]$ la $\mathbb{K}$-algèbre des polynômes à coefficients dans $\mathbb{K}$. Soit $P(x) \in \mathbb{K}[x]$ de degré n. Soit α un zéro de $P(x)$ dans $\mathbb{K}$, c'est-à-dire un nombre $\alpha \in \mathbb{K}$ tel que $P(\alpha) = 0$. On peut alors mettre $(x - \alpha)$ en facteur dans $P(x)$. Il existe donc un unique entier $1 \leq q \leq n$, appelé la *multiplicité* de α, tel que :
$$P(x) = (x - \alpha)^q Q(x)$$
avec $Q(x) \in \mathbb{K}[x]$ de degré $n - q$ vérifiant $Q(\alpha) \neq 0$.

b) Définition. Soit λ une v.p. d'un endomorphisme f de $E = \mathbb{K}^n$ (ou d'une matrice A carrée d'ordre n à coefficients dans $\mathbb{K}$). On appelle *multiplicité algébrique* de λ sa multiplicité en tant que zéro du polynôme caractéristique de f (ou de A).

c) Proposition. *Avec les données et notations ci-dessus, et en notant E_λ le sous-epace propre associé à la v.p. λ, on a :*
$$1 \;\leq\; \dim E_\lambda \;\leq\; (\text{multiplicité algébrique de } \lambda) \;\leq\; n.$$

Preuve : Notons $p = \dim E_\lambda$ et q la multiplicité algébrique de λ.

Comme E_λ est un ss-e.v. de E, on a $p \leq n$ (cf. 1.5.2.c). Soit $\mathcal{C} = (u_1, \ldots, u_p)$ une base de E_λ. Comme c'est une famille libre, on la complète (d'après 1.5.4.b) en une base $\mathcal{B} = (u_1, \ldots, u_p, v_{p+1}, \ldots, v_n)$ de E. Notons M la matrice de f par rapport à cette base $\mathcal{B}$. Pour tout $1 \leq i \leq p$, le vecteur u_i est un vecteur propre de f associé à λ, donc $f(u_i) = \lambda.u_i$. Donc, par définition même de la matrice d'un endomorphisme par rapport à une base, les p premières colonnes de la matrice M sont :

$$M=\begin{pmatrix} \lambda & 0 & 0 & \dots & 0 & \star & \dots & \star \\ 0 & \lambda & 0 & \dots & 0 & \star & \dots & \star \\ 0 & 0 & \lambda & & 0 & \star & \dots & \star \\ \vdots & \vdots & & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \lambda & \star & \dots & \star \\ 0 & 0 & 0 & \dots & 0 & \star & \dots & \star \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & \star & \dots & \star \end{pmatrix} \quad \text{donc} \quad \det(M-x.I_n)=\begin{vmatrix} \lambda-x & 0 & 0 & \dots & 0 & \star & \dots & \star \\ 0 & \lambda-x & 0 & \dots & 0 & \star & \dots & \star \\ 0 & 0 & \lambda-x & & 0 & \star & \dots & \star \\ \vdots & \vdots & & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \lambda-x & \star & \dots & \star \\ 0 & 0 & 0 & \dots & 0 & \star & \dots & \star \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & \star & \dots & \star \end{vmatrix}.$$

En développant ce déterminant par rapport aux p premières colonnes, on déduit que $P_f(x)$ est de la forme $P_f(x)=(\lambda-x)^pQ(x)$, avec $Q(x)\in\mathbb{K}[x]$ de degré $n-p$. Par définition de la multiplicité algébrique q de $P_f(x)$, on a donc $p\leq q$. □

d) Terminologie. La dimension du sous-espace propre E_λ est parfois appelée la multiplicité géométrique de la v.p. λ. Elle est donc toujours $\leq$ à sa multiplicité algébrique.
Une v.p. est dite *simple* si sa multiplicité algébrique est 1 (notons que sa multiplicité géométrique est alors forcément 1 aussi). Une v.p. est dite *double* si sa multiplicité algébrique est 2 (sa multiplicité géométrique est soit 1 soit 2). Une v.p. est dite *triple* si sa multiplicité algébrique est 3 (sa multiplicité géométrique est alors soit 1 soit 2 soit 3).

6.1.5 Quelques exemples de calculs

Exemple (a) Soit $A=\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$. On calcule :

$P_A(x)=\begin{vmatrix} 1-x & 2 & 0 \\ 0 & 1-x & 1 \\ 1 & 1 & -1-x \end{vmatrix}=(1-x)(x^2-2)+2=-x^3+x^2+2x=-x(x-2)(x+1).$
Donc A admet trois valeurs propres simples qui sont 0, 2 et -1.
Déterminons les sous-espaces propres correspondants.

• $(x,y,z)\in E_2$ ssi $\begin{pmatrix} 1-2 & 2 & 0 \\ 0 & 1-2 & 1 \\ 1 & 1 & -1-2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, donc ssi $\begin{cases} -x+2y=0 \\ -y+z=0 \\ x+y-3z=0 \end{cases}$.
Donc $E_2=\{(2y,y,y)\in\mathbb{R}^3\,;\,y\in\mathbb{R}\}$ est la droite de base (u_2), où $u_2=(2,1,1)$.

• $(x,y,z)\in E_{-1}$ ssi $\begin{pmatrix} 1-(-1) & 2 & 0 \\ 0 & 1-(-1) & 1 \\ 1 & 1 & -1-(-1) \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, donc ssi $\begin{cases} 2x+2y=0 \\ 2y+z=0 \\ x+y=0 \end{cases}$.
Donc $E_{-1}=\{(-y,y,-2y)\in\mathbb{R}^3\,;\,y\in\mathbb{R}\}$ est la droite de base (u_{-1}), où $u_{-1}=(1,-1,2)$.

• $(x,y,z)\in E_0$ ssi $\begin{pmatrix} 1-0 & 2 & 0 \\ 0 & 1-0 & 1 \\ 1 & 1 & -1-0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, donc ssi $\begin{cases} x+2y=0 \\ y+z=0 \\ x+y-z=0 \end{cases}$.
Donc $E_0=\{(-2y,y,-y)\in\mathbb{R}^3\,;\,y\in\mathbb{R}\}$ est la droite de base (u_0), où $u_0=(2,-1,1)$.

Exemple (b) Soit $A=\begin{pmatrix} 8 & 2 & -2 \\ 2 & 5 & 4 \\ -2 & 4 & 5 \end{pmatrix}$. On calcule :

$$P_A(x)=\begin{vmatrix} 8-x & 2 & -2 \\ 2 & 5-x & 4 \\ -2 & 4 & 5-x \end{vmatrix}=\begin{vmatrix} 8-x & 2 & -2 \\ 2 & 5-x & 4 \\ 0 & 9-x & 9-x \end{vmatrix}=(9-x)\begin{vmatrix} 8-x & 2 & -2 \\ 2 & 5-x & 4 \\ 0 & 1 & 1 \end{vmatrix}=(9-x)\begin{vmatrix} 8-x & 2 & -4 \\ 2 & 5-x & x-1 \\ 0 & 1 & 0 \end{vmatrix}$$
$$=(9-x)(-1)\begin{vmatrix} 8-x & -4 \\ 2 & x-1 \end{vmatrix}=(9-x)(x^2-9x)=-x(x-9)^2.$$
Donc 0 est v.p. simple et 9 est v.p. double. On sait qu'alors E_0 est forcément une droite, mais E_9 peut être soit une droite, soit un plan. Déterminons-les.

• $(x,y,z)\in E_0$ ssi $\begin{pmatrix} 8-0 & 2 & -2 \\ 2 & 5-0 & 4 \\ -2 & 4 & 5-0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, donc ssi $\begin{cases} 8x+2y-2z=0 \\ 2x+5y+4z=0 \\ -2x+4y+5z=0 \end{cases}$.
Après calculs, on trouve que E_0 est la droite de base (u_0), où $u_0=(1,-2,2)$.

• $(x,y,z)\in E_9$ ssi $\begin{pmatrix} 8-9 & 2 & -2 \\ 2 & 5-9 & 4 \\ -2 & 4 & 5-9 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, donc ssi $\begin{cases} -x+2y-2z=0 \\ 2x-4y+4z=0 \\ -2x+4y-4z=0 \end{cases}$.

Donc E_9 est le plan d'équation $x - 2y + 2z = 0$, dont une base est (u_9, v_9), où $u_9 = (2,1,0)$ et $v_9 = (2,0,-1)$. Donc, sur cet exemple, la multiplicité de la v.p. double 9 est égale à la dimension du sous-espace propre correspondant.

Exemple (c) Soit $A = \begin{pmatrix} 3 & 2 & 4 \\ -1 & 3 & -1 \\ -2 & -1 & -3 \end{pmatrix}$. On calcule : $P_A(x) = \begin{vmatrix} 3-x & 2 & 4 \\ -1 & 3-x & -1 \\ -2 & -1 & -3-x \end{vmatrix} = \begin{vmatrix} -1-x & 2 & 4 \\ 0 & 3-x & -1 \\ x+1 & -1 & -3-x \end{vmatrix}$
$= (x+1) \begin{vmatrix} -1 & 2 & 4 \\ 0 & 3-x & -1 \\ 1 & -1 & -3-x \end{vmatrix} = (x+1) \begin{vmatrix} -1 & 2 & 4 \\ 0 & 3-x & -1 \\ 0 & 1 & 1-x \end{vmatrix} = -(x+1)(x^2 - 4x + 4) = -(x+1)(x-2)^2$.

Donc -1 est v.p. simple et 2 est v.p. double. On sait qu'alors E_{-1} est forcément une droite, mais E_2 peut être soit une droite, soit un plan. Déterminons-les.

• $(x,y,z) \in E_{-1}$ ssi $\begin{pmatrix} 3-(-1) & 2 & 4 \\ -1 & 3-(-1) & -1 \\ -2 & -1 & -3-(-1) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, donc ssi $\left\{ \begin{array}{l} 4x + 2y + 4z = 0 \\ -x + 4y - z = 0 \\ -2x - y - 2z = 0 \end{array} \right.$.

Après calculs, on trouve que E_{-1} est la droite de base (u_{-1}), où $u_{-1} = (1,0,-1)$.

• $(x,y,z) \in E_2$ ssi $\begin{pmatrix} 3-2 & 2 & 4 \\ -1 & 3-2 & -1 \\ -2 & -1 & -3-2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, donc ssi $\left\{ \begin{array}{l} x + 2y + 4z = 0 \\ -x + y - z = 0 \\ -2x - y - 5z = 0 \end{array} \right.$.

Après calculs, on trouve que E_2 est la droite de base (u_2), où $u_2 = (2,1,-1)$. Donc, sur cet exemple, la multiplicité de la v.p. double 2 est strictement supérieure à la dimension du sous-espace propre correspondant.

Exemple (d) Soit $A = \begin{pmatrix} -1 & 1 & 1 & 0 \\ -1 & 2 & 1 & -1 \\ 5 & -3 & -2 & 5 \\ 4 & -2 & -2 & 3 \end{pmatrix}$. Après calculs, $P_A(x) = (x+1)(x-1)^3$.

• -1 est v.p. simple. On sait qu'alors E_{-1} est forcément une droite. Après calculs, on trouve que E_{-1} est la droite de base (u_{-1}), où $u_{-1} = (1,0,0,-1)$.

• 1 est v.p. triple. $(x,y,z) \in E_1$ ssi $\left\{ \begin{array}{l} -2x + y + z = 0 \\ -x + y + z - t = 0 \\ 5x - 3y - 3z + 5t = 0 \\ 4x - 2y - 2z + 2t = 0 \end{array} \right.$.

On vérifie que ce système équivaut à $x = t = 0$ et $y + z = 0$. Donc E_1 est la droite de base (u_1), où $u_1 = (0,1,-1,0)$. Sur cet exemple, le sous-espace propre associé à la v.p. triple 1 est seulement de dimension 1.

Exemple (e) Soit $A = \begin{pmatrix} 1 & -1 \\ 3 & -2 \end{pmatrix}$. Après calculs, $P_A(x) = x^2 + x + 1$.
Donc A n'admet pas de v.p. dans $\mathbb{R}$.
En revanche, dans $\mathbb{C}$, elle admet deux v.p. simples distinctes qui sont j et j^2. On vérifie aisément (par la même méthode que sur les exemples précédents) que E_j est la droite de base (u_j), où $u_j = (1, 1-j)$ et E_{j^2} est la droite de base (u_{j^2}), où $u_{j^2} = (1, 1-j^2)$.

6.2 Diagonalisation et trigonalisation

6.2.1 Endomorphisme diagonalisable, matrice diagonalisable

a) Lemme préliminaire. *Soit f un endomorphisme de E. On suppose que f admet s valeurs propres distinctes $\lambda_1, \lambda_2, \ldots, \lambda_s$ dans $\mathbb{K}$. Le sous-espace somme $F = E_{\lambda_1} + E_{\lambda_2} + \cdots + E_{\lambda_s}$ des sous-espaces propres de f est une somme directe : $F = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_s}$.*

Preuve : D'après 1.5.3.e, il s'agit de montrer que tout vecteur de F se décompose de façon unique en une somme $u_1 + u_2 + \cdots + u_s$ avec $u_1 \in E_{\lambda_1}, u_2 \in E_{\lambda_2}, \ldots, u_s \in E_{\lambda_s}$. Par définition de F, l'existence d'une telle décomposition est claire. Le problème est seulement l'unicité. Quitte à faire la différence membre à membre de deux telles décompositions, on est ramené à montrer simplement que :
si $u_1 + u_2 + \cdots + u_s = 0_E$ avec $u_1 \in E_{\lambda_1}, u_2 \in E_{\lambda_2}, \ldots, u_s \in E_{\lambda_s}$,
alors $u_1 = u_2 = \cdots = u_s = 0_E$.

Supposons donc $u_1 + u_2 + \cdots + u_s = 0_E$ avec $u_1 \in E_{\lambda_1}, u_2 \in E_{\lambda_2}, \ldots, u_s \in E_{\lambda_s}$.
En appliquant f, on a :
$\lambda_1.u_1 + \lambda_2.u_2 + \cdots + \lambda_s.u_s = f(u_1) + f(u_2) + \cdots + f(u_s) = f(u_1 + u_2 + \cdots + u_s) = f(0_E) = 0_E$.
Par ailleurs, en multipliant par λ_1 : $\lambda_1 u_1 + \lambda_1 u_2 + \cdots + \lambda_1 u_s = 0_E$.
D'où par différence membre à membre : $\sum_{i=2}^{s} (\lambda_i - \lambda_1).u_i = 0_E$.
On réitère. En appliquant f, il vient :
$\sum_{i=2}^{s} (\lambda_i - \lambda_1)\lambda_i.u_i = \sum_{i=2}^{s} (\lambda_i - \lambda_1).f(u_i) = f(\sum_{i=2}^{s} (\lambda_i - \lambda_1).u_i) = f(0_E) = 0_E$.
Par ailleurs, en multipliant par λ_2 : $\sum_{i=2}^{s} (\lambda_i - \lambda_1)\lambda_2.u_i = 0_E$.
D'où par différence membre à membre : $\sum_{i=3}^{s} (\lambda_i - \lambda_1)(\lambda_i - \lambda_2).u_i = 0_E$.
De proche en proche, on parvient ainsi à : $(\lambda_s - \lambda_1)(\lambda_s - \lambda_2) \ldots (\lambda_s - \lambda_{s-1}).u_s = 0_E$. Comme les λ_i sont deux à deux distincts, on conclut $u_s = 0_E$. La somme de départ $u_1 + \cdots + u_s = 0_E$ se réduit donc à $u_1 + \cdots + u_{s-1} = 0_E$. En réitérant le même raisonnement, on obtient successivement $u_{s-1} = \cdots = u_2 = u_1 = 0_E$. Ce qui achève la preuve. □

b) Définition. Un endomorphisme f de E est dit *diagonalisable* lorsqu'il existe une base $\mathcal{C}$ de E telle que la matrice de f dans la base $\mathcal{C}$ soit une matrice diagonale D. Une matrice carrée d'ordre n est dite *diagonalisable* lorsqu'elle est semblable (au sens de 3.1.6.b) à une matrice diagonale. Si A désigne la matrice de f dans une base quelconque $\mathcal{B}$ de E, il résulte donc de 3.3.4.c que :

(l'endomorphisme f est diagonalisable) $\Leftrightarrow$ (la matrice A est diagonalisable).

Dans ce cas, $A = PDP^{-1}$ avec $A = M_{\mathcal{B}}(f)$, $D = M_{\mathcal{C}}(f)$ diagonale, $P = \mathrm{Pass}_{\mathcal{B} \to \mathcal{C}} \in GL(n, \mathbb{K})$.

Remarquons qu'alors les coefficients diagonaux de la matrice D sont exactement les v.p. de f, chacune apparaissant un nombre de fois égal à sa multiplicité algébrique [il suffit pour le voir de développer $\det(D - x.I_n) = P_f(x)$].

c) Théorème. *Soit f un endomorphisme de E. Les conditions suivantes sont équivalentes :*

(i) *f est diagonalisable ;*

(ii) *il existe une base $\mathcal{C}$ de E qui est formée de vecteurs propres de f ;*

(iii) *E est la somme directe des sous-espaces propres de f ;*

(iv) *le polynôme caractéristique $P_f(x)$ admet n zéros (comptés avec leur multiplicité), et la multiplicité algébrique de chaque valeur propre λ de f est égale à la dimension du sous-espace propre E_λ correspondant ;*

(v) *le polynôme caractéristique $P_f(x)$ est produit de n facteurs de degré 1, et la multiplicité algébrique de chaque valeur propre λ de f est égale à la dimension du sous-espace propre E_λ correspondant.*

Preuve. Supposons que l'on a (i). Il existe donc une base $\mathcal{C}$ de E telle que la matrice D de f par rapport à $\mathcal{C}$ est diagonale. Certains de ses coefficients peuvent être égaux : on note $\lambda_1, \lambda_2, \ldots, \lambda_s$ les valeurs distinctes de ces coefficients diagonaux (donc $1 \leq s \leq n$). Quitte à permuter les vecteurs de $\mathcal{C}$, on peut sans restriction supposer qu'ils apparaissent dans l'ordre :

$$D = M_{\mathcal{C}}(f) = \begin{pmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_1 \end{bmatrix} & & & \\ & \begin{bmatrix} \lambda_2 & & \\ & \ddots & \\ & & \lambda_2 \end{bmatrix} & & \\ & & \ddots & \\ & & & \begin{bmatrix} \lambda_s & & \\ & \ddots & \\ & & \lambda_s \end{bmatrix} \end{pmatrix}$$

(tous les autres coefficients hors de la diagonale étant des zéros).

Pour tout $1 \leq i \leq s$, notons q_i le nombre de fois où apparaît le coefficient λ_i sur la diagonale. Donc la base $\mathcal{C}$ de E est la réunion $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_s$ où chaque $\mathcal{C}_i$ est formé de q_i vecteurs de $\mathcal{C}$ vérifiant $f(u) = \lambda_i.u$. Avec la remarque 6.1.2.b, chaque λ_i est donc une v.p. de f, et la dimension du sous-espace propre associé E_{λ_i} est q_i. Cet entier q_i est aussi la multiplicité algébrique de la v.p. λ_i, puisque qu'il suffit de développer le déterminant $\det(D - x.I_n)$ pour obtenir $P_f(x) = \det(D - x.I_n) = (-1)^n(x-\lambda_1)^{q_1}(x-\lambda_2)^{q_2}\ldots(x-\lambda_s)^{q_s}$. Donc (v) est vérifiée.

• L'implication (v) $\Rightarrow$ (iv) est claire.

• Supposons (iv). Cela signifie que l'on a :
d'une part : $P_f(x) = (-1)^n(x-\lambda_1)^{q_1}(x-\lambda_2)^{q_2}\ldots(x-\lambda_s)^{q_s}$,
où $\lambda_1, \lambda_2, \ldots, \lambda_s$ sont (d'après 6.1.3.b) les v.p. distinctes de f,
d'autre part : $\dim E_{\lambda_i} = q_i$ pour tout $1 \leq i \leq s$.

Considérons le ss-e.v. somme $F = E_{\lambda_1} + E_{\lambda_2} + \cdots + E_{\lambda_s}$. D'après le lemme a), on a en fait $F = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_s}$, donc $\dim F = \dim E_{\lambda_1} + \dim E_{\lambda_2} + \cdots + \dim E_{\lambda_s} = q_1 + q_2 + \cdots + q_n = \deg P_f(x) = n = \dim E$. Donc $E = F = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_s}$, ce qui prouve (iii).

• Il résulte de (iii) que l'on peut former une base $\mathcal{C}$ de E en prenant la réunion disjointe : $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2 \cup \cdots \cup \mathcal{C}_s$, où $\mathcal{C}_i$ est une base de E_{λ_i} (donc formée de q_i vecteurs propres associés à la v.p. λ_i) pour tout $1 \leq i \leq s$. Ceci prouve (ii).

• Supposons (ii). Avec les notations précédentes, pour tout vecteur $u \in \mathcal{C}$, il existe un unique indice $1 \leq i \leq s$ tel que $u \in \mathcal{C}_i$, et l'on a alors $f(u) = \lambda_i.u$. La matrice de f par rapport à la base $\mathcal{C}$ est donc par construction diagonale.
On a ainsi montré que : (i) $\Rightarrow$ (v) $\Rightarrow$ (iv) $\Rightarrow$ (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i), ce qui prouve l'équivalence des cinq assertions. □

d) Corollaire. *Soit f un endomorphisme de E. Si f admet n v.p. distinctes (chacune étant nécessairement une v.p. simple), alors f est diagonalisable.*

Preuve. On applique tout ce qui précède au cas $n = s$, d'où $q_i = 1$ pour tout $1 \leq i \leq n$.
Dans ce cas, chaque sous-espace propre est une droite. □

ATTENTION : ce corollaire ne donne qu'une condition suffisante pour qu'un endomorphisme soit diagonalisable ; elle n'est nullement nécessaire, comme l'exprime le théorème général .c) et comme le montre le deuxième des exemples ci-dessous.

e) Exemples. Reprenons les exemples traités en 6.1.5.

• Dans l'exemple (a), la matrice A considérée est diagonalisable, d'après le corollaire ci-dessus. Une base de vecteurs propres est $\mathcal{C} = (u_2, u_{-1}, u_0)$, donc :

$$A = PDP^{-1}, \text{ avec } P = \begin{pmatrix} 2 & 1 & 2 \\ 1 & -1 & -1 \\ 1 & 2 & 1 \end{pmatrix} \text{ et } D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

• Dans l'exemple (b), la matrice A considérée est diagonalisable, d'après le (iv) du théorème c) ci-dessus. Une base de vecteurs propres est $\mathcal{C} = (u_0, u_9, v_9)$, donc :

$$A = PDP^{-1}, \text{ avec } P = \begin{pmatrix} 1 & 2 & 2 \\ -2 & 1 & 0 \\ 2 & 0 & -1 \end{pmatrix} \text{ et } D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix}$$

• Dans l'exemple (c), la matrice A considérée n'est pas diagonalisable, car le sous-espace propre associé à la valeur propre double 2 est seulement de dimension 1.

• De même dans l'exemple (d), la matrice A considérée n'est pas diagonalisable, car le sous-espace propre associé à la valeur propre triple 1 est seulement de dimension 1.

• Enfin, dans l'exemple (e), A n'est pas diagonalisable dans $\mathbb{R}$, mais est diagonalisable dans $\mathbb{C}$ (car admet deux v.p. simples distinctes complexes). Une base de vecteurs propres est $\mathcal{C} = (u_j, u_{j^2})$, donc : $A = PDP^{-1}$, avec $P = \begin{pmatrix} 1 & 1 \\ 1-j & 1-j^2 \end{pmatrix}$ et $D = \begin{pmatrix} j & 0 \\ 0 & j^2 \end{pmatrix}$.

6.2.2 Endomorphisme trigonalisable, matrice trigonalisable

a) Remarque préliminaire. On a traité au paragraphe précédent la situation la plus favorable : celle où la matrice carrée A considérée est semblable à une matrice diagonale. Bien sûr, on l'a vu, ce n'est pas toujours le cas ! On traite donc dans ce paragraphe une situation moins favorable, mais utile quand même pour les applications pratiques, celle où A est seulement semblable à une matrice triangulaire.

b) Définition. Un endomorphisme f de E est dit *trigonalisable* (on dit aussi triangularisable) sur $\mathbb{K}$ lorsqu'il existe une base $\mathcal{C}$ de E telle que la matrice de f dans la base $\mathcal{C}$ soit une matrice triangulaire (supérieure) T. Une matrice carrée d'ordre n à coefficients dans $\mathbb{K}$ est dite *trigonalisable* sur $\mathbb{K}$ lorsqu'elle est semblable (au sens de 3.1.6.b) à une matrice triangulaire (supérieure). Si A désigne la matrice de f dans une base quelconque $\mathcal{B}$ de E, il résulte donc de 3.3.4.c que :

$$\text{(l'endomorphisme } f \text{ est trigonalisable)} \Leftrightarrow \text{(la matrice } A \text{ est trigonalisable).}$$

Alors : $A = PTP^{-1}$ avec $A = M_{\mathcal{B}}(f)$, $T = M_{\mathcal{C}}(f)$ triangulaire, $P = \mathrm{Pass}_{\mathcal{B}\to\mathcal{C}} \in GL(n,\mathbb{K})$.

c) Théorème.

Soit f un endomorphisme de E. Les conditions suivantes sont équivalentes :

(i) *f est trigonalisable sur $\mathbb{K}$;*

(ii) *le polynôme caractéristique $P_f(x)$ est produit de n facteurs de degré 1 sur $\mathbb{K}$.*

Preuve. Supposons que l'on a (i). Il existe donc une base $\mathcal{C}$ de E telle que la matrice T de f par rapport à $\mathcal{C}$ est triangulaire supérieure. Donc $P_f(x) = P_T(x)$. Comme T est triangulaire, il est clair que $P_T(x) = (-1)^n(x-\alpha_1)(x-\alpha_2)\dots(x-\alpha_n)$. Donc (ii) est vérifié (et les coefficients diagonaux de T sont les v.p. de f).

• Pour la réciproque, on raisonne par récurrence sur n. Le résultat est trivial si $n = 1$. Supposons-le vrai pour tout endomorphisme de tout $\mathbb{K}$-e.v. de dimension $n-1$. Prenons f un endomorphisme d'un $\mathbb{K}$-e.v. de dimension n tel que $P_f(x)$ est produit de n facteurs de degré 1 sur $\mathbb{K}$. Il admet donc au moins un zéro dans $\mathbb{K}$. Soit donc λ un zéro de $P_f(x)$ dans $\mathbb{K}$, c'est-à-dire une v.p. de f. Soit u un vecteur propre de f associé à λ. On a $u \neq 0$, donc d'après 1.5.4.b, on peut compléter u en une base $\mathcal{B} = (u, v_2, v_3, \dots, v_n)$ de E. Notons $\mathcal{C} = (v_2, v_3, \dots, v_n)$, qui est une famille libre (car sous-famille de $\mathcal{B}$). Notons $F = \mathrm{Vect}\,\mathcal{C}$. D'après 1.5.2.b, $\mathcal{C}$ est une base de F, de sorte que $\dim F = n-1$. D'après 1.5.3.b, F est un supplémentaire dans E de la droite vectorielle Δ de base (u).

Comme $f(u) = \lambda.u$, la matrice de f par rapport à $\mathcal{B}$ est de la forme :

$$A = \begin{pmatrix} \lambda & \alpha_2 \; \alpha_3 \dots \; \alpha_n \\ 0 & \\ 0 & \left[\quad B \quad \right] \\ \vdots & \\ 0 & \end{pmatrix}$$

où $B \in \mathcal{M}_{n-1}(\mathbb{K})$ et $\alpha_j \in \mathbb{K}$ pour tout $2 \leq j \leq n$.

Appelons g l'endomorphisme de F dont B est la matrice par rapport à la base $\mathcal{C}$ (attention, ce n'est pas la restriction de f à F, car F n'est a priori pas stable par f). On a donc :

$$f(v_j) = \alpha_j.u + g(v_j) \quad \text{pour tout } 2 \leq j \leq n.$$

On calcule alors à partir de A : $P_A(f) = \det(A - x.I_n) = (\lambda - x)\det(B - x.I_{n-1})$, c'est-à-dire : $P_f(x) = (\lambda - x)P_g(x)$. Comme on a supposé que $P_f(x)$ est produit de n facteurs de degré 1, il en résulte que $P_g(x)$ est produit de $n-1$ facteurs de degré 1. On peut appliquer à g l'hypothèse de récurrence, il existe une base $\mathcal{C}' = (w_2, w_3, \dots, w_n)$ de F telle que la matrice T de g par rapport à la base $\mathcal{C}'$ est triangulaire supérieure. Comme $E = \Delta \oplus F$, il résulte de 1.5.3.b que $\mathcal{B}' = (u, w_2, w_3, \dots, w_n)$ est une base de E. Chaque w_i (pour $2 \leq i \leq n$) est c.l. de $v_2, \dots, v_n$, donc :

$$f(w_i) = \beta_i.u + g(w_i) \quad \text{pour tout } 2 \leq i \leq n,$$

où le scalaire β_i est c.l. des α_j. Donc la matrice de f par rapport à $\mathcal{C}'$ est :

$$A' = \begin{pmatrix} \lambda & \beta_2 \ \beta_3 \dots \ \beta_n \\ 0 & \\ 0 & \\ \vdots & \left[\quad T \quad \right] \\ 0 & \end{pmatrix},$$

qui est triangulaire ; ce qui achève la preuve. □

d) Corollaire. (Cas où $\mathbb{K} = \mathbb{C}$)
Tout endomorphisme de $\mathbb{C}^n$ est trigonalisable sur $\mathbb{C}$.
Toute matrice carrée à coefficients complexes est trigonalisable sur $\mathbb{C}$.

Preuve. On sait que tout polynôme à coefficients dans $\mathbb{C}$ se décompose en un produit de facteurs de degré 1 sur $\mathbb{C}$; on peut donc appliquer le théorème précédent. □

e) Remarque très importante. Soit A une matrice carrée d'ordre n à coefficients dans $\mathbb{R}$. On peut toujours la considérer comme un élément de $\mathcal{M}_n(\mathbb{C})$. Elle est donc toujours trigonalisable sur $\mathbb{C}$, mais cela ne signifie bien sûr pas qu'elle est trigonalisable sur $\mathbb{R}$.
Si par exemple $P_A(x) = -(x-2)(x^2+1)$, A n'est pas trigonalisable sur $\mathbb{R}$ mais l'est sur $\mathbb{C}$ puisque $P_A(x) = -(x-2)(x-i)(x+i)$.

f) Exemples. Reprenons les exemples du paragraphe 6.1.5.

On a déjà vu en 6.2.1.e que les exemples (a) et (b) correspondent à des matrices diagonalisables sur $\mathbb{R}$, et que l'exemple e) est diagonalisable sur $\mathbb{C}$ (il n'est sur $\mathbb{R}$ ni diagonalisable ni trigonalisable puisque qu'il n'admet pas de v.p. réelles.)

Dans les exemples (c) et (d), les matrices considérées ne sont pas diagonalisables, mais elles sont trigonalisables sur $\mathbb{R}$ puisque le polynôme caractéristique se décompose en produit de facteurs de degré 1. Détaillons :

• Reprenons la matrice A de l'exemple 6.1.5.c). Soit f l'endomorphisme de $\mathbb{R}^3$ dont A est la matrice par rapport à la base canonique. Le sous-espace propre associé à la valeur simple -1 est la droite de base (u_{-1}) avec $u_{-1} = (1, 0, -1)$. Le sous-espace propre associé à la valeur double 2 est la droite de base (u_2) avec $u_2 = (2, 1, -1)$. Quelle que soit la façon de compléter la famille libre (u_{-1}, u_2) en une base $\mathcal{C}$ par l'adjonction d'un vecteur quelconque qui n'est pas combinaison linéaire de u_{-1} et u_2, la matrice de f par rapport à la base $\mathcal{C}$ est triangulaire. Prenons par exemple $\mathcal{C} = (u_{-1}, u_2, e_1)$ avec $e_1 = (1, 0, 0)$. Notons P la matrice de passage de la base canonique à la base $\mathcal{C}$. Donc

$$P = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ -1 & -1 & 0 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 0 & -1 & -1 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}, \quad T = \mathrm{Mat}_{\mathcal{C}}(f) = P^{-1}AP = \begin{pmatrix} -1 & 0 & 3 \\ 0 & 2 & -1 \\ 0 & 0 & 2 \end{pmatrix}.$$

• Reprenons la matrice A de l'exemple 6.1.5.d). Soit f l'endomorphisme de $\mathbb{R}^4$ dont A est la matrice par rapport à la base canonique. Le sous-espace propre associé à la valeur simple -1 est la droite de base (u_{-1}) avec $u_{-1} = (1, 0, 0, -1)$. Le sous-espace propre associé à la valeur triple 1 est la droite de base (u_1) avec $u_1 = (0, 1, -1, 0)$. On sait que l'on peut compléter (u_{-1}, u_1) en une base $\mathcal{C} = (u_{-1}, u_1, v, w)$ de $\mathbb{R}^4$ telle que la matrice de f par rapport à la base $\mathcal{C}$ est triangulaire. Mais ici, le choix des deux vecteurs v, w n'est pas indifférent ; on verra plus loin une méthode générale pour les déterminer. Contentons-nous de noter ici que, pour $v = (1, 0, 2, 0)$ et $w = (0, 0, 1, 1)$, on a :

$$P = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 2 & 1 \\ -1 & 0 & 0 & 1 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 2 & -1 & -1 & 1 \\ 0 & 1 & 0 & 0 \\ -1 & 1 & 1 & -1 \\ 2 & -1 & -1 & 2 \end{pmatrix}, \quad T = \mathrm{Mat}_{\mathcal{C}}(f) = P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

6.2.3 Application à la résolution de systèmes différentiels linéaires

a) Premier exemple (cas diagonalisable avec v.p. simples).

On cherche à déterminer tous les triplets de fonctions réelles $x(t), y(t), z(t)$ dérivables sur $\mathbb{R}$ solutions du système différentiel :

$$(S)\ \begin{cases} x'(t) = & -2x(t) - 2y(t) + 2z(t) \\ y'(t) = & 3x(t) + 5y(t) \\ z'(t) = & x(t) + 2y(t) + 3z(t) \end{cases}.$$

On note matriciellement ce système sous la forme :

$$X'(t) = AX(t), \text{ avec } X(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix},\ X'(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix},\ A = \begin{pmatrix} -2 & -2 & 2 \\ 3 & 5 & 0 \\ 1 & 2 & 3 \end{pmatrix}.$$

Suivant les méthodes vues ci-dessus, on montre que A est diagonalisable et on réalise la diagonalisation :

$$P_A(x) = -(x+1)(x-2)(x-5), \text{ et } A = PDP^{-1} \text{ où } D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix} \text{ et } P = \begin{pmatrix} 2 & 1 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

On pose : $U(t) = \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix} = P^{-1}X(t)$; d'où : $U'(t) = \begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = P^{-1}X'(t)$.

On a alors :

$(S) \Leftrightarrow X'(t) = AX(t) \Leftrightarrow PU'(t) = APU(t) \Leftrightarrow U'(t) = P^{-1}APU(t) \Leftrightarrow U'(t) = DU(t)$.

On est donc ramené à résoudre le système (beaucoup plus simple) :

$$\begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix}, \quad \text{c'est-à-dire : } \begin{cases} u'(t) = -u(t) \\ v'(t) = 2v(t) \\ w'(t) = 5w(t) \end{cases}.$$

D'après le cours d'analyse, on sait que les solutions de ce dernier système sont :

$$u(t) = \alpha e^{-t},\ v(t) = \beta e^{2t},\ w(t) = \gamma e^{5t}, \text{ pour } \alpha, \beta, \gamma \text{ décrivant } \mathbb{R}.$$

En revenant à $X(t) = PU(t) = \begin{pmatrix} 2 & 1 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \alpha e^{-t} \\ \beta e^{2t} \\ \gamma e^{5t} \end{pmatrix}$, on conclut que les solutions du système différentiel (S) sont :

$$\begin{cases} x(t) = 2\alpha e^{-t} + \beta e^{2t} \\ y(t) = -\alpha e^{-t} - \beta e^{2t} + \gamma e^{5t} \\ z(t) = \beta e^{2t} + \gamma e^{5t} \end{cases}, \qquad \text{avec } \alpha, \beta, \gamma \in \mathbb{R}.$$

Remarque : algébriquement, les trois fonctions vectorielles $\mathbb{R} \to \mathbb{R}^3$ définies par $t \mapsto \begin{pmatrix} 2e^{-t} \\ -e^{-t} \\ 0 \end{pmatrix}$, $t \mapsto \begin{pmatrix} e^{2t} \\ -e^{2t} \\ e^{2t} \end{pmatrix}$, $t \mapsto \begin{pmatrix} 0 \\ e^{5t} \\ e^{5t} \end{pmatrix}$, forment une base de l'espace vectoriel des solutions de (S).

b) Deuxième exemple (cas diagonalisable avec v.p. multiples).

On cherche à déterminer tous les triplets de fonctions réelles $x(t), y(t), z(t)$ dérivables sur $\mathbb{R}$ solutions du système différentiel :

$$(S)\ \begin{cases} x'(t) = & 2x(t) + 3y(t) - 6z(t) \\ y'(t) = & -6x(t) - 7y(t) + 12z(t) \\ z'(t) = & -3x(t) - 3y(t) + 5z(t) \end{cases}.$$

On note matriciellement ce système sous la forme :

$$X'(t) = AX(t), \text{ avec } X(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix},\ X'(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix},\ A = \begin{pmatrix} 2 & 3 & -6 \\ -6 & -7 & 12 \\ -3 & -3 & 5 \end{pmatrix}.$$

Suivant les méthodes vues ci-dessus, on montre que A est diagonalisable et on réalise la diagonalisation :

$$P_A(x) = -(x+1)^2(x-2), \text{ et } A = PDP^{-1} \text{ où } D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \text{ et } P = \begin{pmatrix} -1 & 1 & 1 \\ 2 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

On pose : $U(t) = \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix} = P^{-1}X(t)$; d'où : $U'(t) = \begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = P^{-1}X'(t)$.

On a alors :

$(S) \Leftrightarrow X'(t) = AX(t) \Leftrightarrow PU'(t) = APU(t) \Leftrightarrow U'(t) = P^{-1}APU(t) \Leftrightarrow U'(t) = DU(t)$.

On est donc ramené à résoudre le système (beaucoup plus simple) :

$$\begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix}, \quad \text{c'est-à-dire :} \quad \begin{cases} u'(t) = 2u(t) \\ v'(t) = -v(t) \\ w'(t) = -w(t) \end{cases}.$$

D'après le cours d'analyse, on sait que les solutions de ce dernier système sont :

$$u(t) = \alpha e^{2t},\ v(t) = \beta e^{-t},\ w(t) = \gamma e^{-t}, \text{ pour } \alpha, \beta, \gamma \text{ décrivant } \mathbb{R}.$$

En revenant à $X(t) = PU(t) = \begin{pmatrix} -1 & 1 & 1 \\ 2 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha e^{2t} \\ \beta e^{-t} \\ \gamma e^{-t} \end{pmatrix}$, on conclut que les solutions du système différentiel (S) sont :

$$\begin{cases} x(t) = -\alpha e^{2t} + (\beta+\gamma)e^{-t} \\ y(t) = 2\alpha e^{2t} + (-\beta+\gamma)e^{-t} \\ z(t) = \alpha e^{2t} + \gamma e^{-t} \end{cases}, \qquad \text{avec } \alpha, \beta, \gamma \in \mathbb{R}.$$

Remarque : algébriquement, les trois fonctions vectorielles $\mathbb{R} \to \mathbb{R}^3$ définies par $t \mapsto \begin{pmatrix} -e^{2t} \\ 2e^{2t} \\ e^{2t} \end{pmatrix}$, $t \mapsto \begin{pmatrix} e^{-t} \\ -e^{-t} \\ 0 \end{pmatrix}$, $t \mapsto \begin{pmatrix} e^{-t} \\ e^{-t} \\ e^{-t} \end{pmatrix}$, forment une base de l'espace vectoriel des solutions de (S).

c) Troisième exemple (cas diagonalisable sur $\mathbb{C}$ mais pas sur $\mathbb{R}$).

On cherche à déterminer tous les triplets de fonctions réelles $x(t), y(t), z(t)$ dérivables sur $\mathbb{R}$ solutions du système différentiel :

$$(S) \quad \begin{cases} x'(t) = 2y(t) - 2z(t) \\ y'(t) = -2x(t) + z(t) \\ z'(t) = 2x(t) - y(t) \end{cases}.$$

On note matriciellement ce système sous la forme :

$$X'(t) = AX(t), \text{ avec } X(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix},\ X'(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix},\ A = \begin{pmatrix} 0 & 2 & -2 \\ -2 & 0 & 1 \\ 2 & -1 & 0 \end{pmatrix}.$$

On calcule : $P_A(x) = -x(x^2 + 9)$.

• *On résoud d'abord sur* $\mathbb{C}$.

$P_A(x) = -x(x - 3i)(x + 3i)$, donc A est diagonalisable. On a :

$$A = PDP^{-1} \text{ où } D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3i & 0 \\ 0 & 0 & -3i \end{pmatrix} \text{ et } P = \begin{pmatrix} 1 & 4 & 4 \\ 2 & -1+3i & -1-3i \\ 2 & -1-3i & -1+3i \end{pmatrix}.$$

On raisonnant comme sur les 2 exemples a) et b) précédents, on montre que les solutions du système différentiel (S) sont :

$$\begin{cases} x(t) = \alpha + 4\beta e^{3it} + 4\gamma e^{-3it} \\ y(t) = 2\alpha + (-1+3i)\beta e^{3it} + (-1-3i)\gamma e^{-3it} \\ z(t) = 2\alpha + (-1-3i)\beta e^{3it} + (-1+3i)\gamma e^{-3it} \end{cases}, \qquad \text{avec } \alpha, \beta, \gamma \in \mathbb{C}.$$

• *On cherche ensuite les solutions réelles.*

D'après ce qui précède, les 3 fonctions vectorielles $\mathbb{R} \to \mathbb{C}^3$ définies par :

$$u_0 : t \mapsto \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad v : t \mapsto \begin{pmatrix} 4e^{3it} \\ (-1+3i)e^{3it} \\ (-1-3i)e^{3it} \end{pmatrix}, \quad \overline{v} : t \mapsto \begin{pmatrix} 4e^{-3it} \\ (-1-3i)e^{-3it} \\ (-1+3i)e^{-3it} \end{pmatrix},$$

forment une base de l'espace vectoriel des solutions complexes de (S).

On en déduit qu'une base de l'espace vectoriel des solutions réelles de (S) est formée des 3 fonctions vectorielles $\mathbb{R} \to \mathbb{C}^3$ définies par : u_0, $w_1 = \frac{1}{2}(v + \overline{v})$, $w_2 = \frac{1}{2i}(v - \overline{v})$. En d'autres termes :

$$w_1 : t \mapsto \begin{pmatrix} 4\cos 3t \\ -\cos 3t - 3\sin 3t \\ -\cos 3t + 3\sin 3t \end{pmatrix}, \qquad w_2 : t \mapsto \begin{pmatrix} 4\sin 3t \\ -\sin 3t + 3\cos 3t \\ -\sin 3t - 3\cos 3t \end{pmatrix}.$$

On conclut que les solutions réelles du système différentiel (S) sont :

$$\begin{cases} x(t) = \alpha + 4\beta\cos 3t + 4\gamma\sin 3t \\ y(t) = 2\alpha + (-\beta+3\gamma)\cos 3t + (-3\beta-\gamma)\sin 3t \\ z(t) = 2\alpha + (-\beta-3\gamma)\cos 3t + (3\beta-\gamma)\sin 3t \end{cases}, \qquad \text{avec } \alpha, \beta, \gamma \in \mathbb{R}.$$

d) Quatrième exemple (cas triangularisable non diagonalisable).

On cherche à déterminer tous les triplets de fonctions réelles $x(t), y(t), z(t)$ dérivables sur $\mathbb{R}$ solutions du système différentiel :

$$(S) \begin{cases} x'(t) = -y(t) - 2z(t) \\ y'(t) = x(t) + z(t) \\ 5z'(t) = -6x(t) + 8y(t) \end{cases}.$$

On note matriciellement ce système sous la forme :

$$X'(t) = AX(t), \quad \text{avec} \quad X(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}, \quad X'(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix}, \ A = \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & 1 \\ -\frac{6}{5} & \frac{8}{5} & 0 \end{pmatrix}.$$

On calcule : $P_A(x) = -(x-1)^2(x+2)$.

$\lambda = -2$ est v.p. simple ; le sous-espace propre associé est la droite dirigée par $v_0 = \begin{pmatrix} 3 \\ -4 \\ 5 \end{pmatrix}$.

$\lambda = 1$ est v.p. double ; le sous-espace propre associé est la droite dirigée par $v_1 = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}$.

Donc A n'est pas diagonalisable. On sait que si l'on complète la famille libre (v_0, v_1) en une base en lui adjoignant un vecteur v_2, on a :

$$A = PTP^{-1} \text{ avec } T = \begin{pmatrix} -2 & 0 & * \\ 0 & 1 & * \\ 0 & 0 & 1 \end{pmatrix} \text{ et } P = \begin{pmatrix} 3 & 3 & * \\ -4 & 1 & * \\ 5 & -2 & * \end{pmatrix}.$$

On verra au chapitre suivant une raison théorique pour laquelle on peut toujours choisir v_2 de telle sorte que $T = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Pour l'instant, contentons-nous de vérifier par le calcul que, si l'on choisit $v_2 = \begin{pmatrix} -1 \\ -2 \\ 0 \end{pmatrix}$, on a $\mathcal{C} = (v_0, v_1, v_2)$ base de $\mathbb{R}^3$, et :

$$A = PTP^{-1} \text{ avec } T = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \text{ et } P = \mathrm{Pass}_{\mathcal{B}\to\mathcal{C}} = \begin{pmatrix} 3 & 3 & -1 \\ -4 & 1 & -2 \\ 5 & -2 & 0 \end{pmatrix}.$$

On pose : $U(t) = \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix} = P^{-1}X(t)$; d'où : $U'(t) = \begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = P^{-1}X'(t)$. Donc :

$$(S) \Leftrightarrow X'(t) = AX(t) \Leftrightarrow PU'(t) = APU(t) \Leftrightarrow U'(t) = P^{-1}APU(t) \Leftrightarrow U'(t) = TU(t).$$

On est ainsi ramené à résoudre le système (beaucoup plus simple) :

$$\begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix}, \qquad \text{c'est-à-dire : } \begin{cases} u'(t) = -2u(t) \\ v'(t) = v(t) + w(t) \\ w'(t) = w(t) \end{cases}.$$

D'après le cours d'analyse, on sait que les solutions des équations (1) et (3) sont $u(t) = \alpha e^{-2t}$ et $w(t) = \gamma e^t$ pour α, γ décrivant $\mathbb{R}$. La seconde équation devient $v'(t) = v(t) + \gamma e^t$, dont la solution générale est $v(t) = (\gamma t + \beta)e^t$.

En revenant à $X(t) = PU(t) = \begin{pmatrix} 3 & 3 & -1 \\ -4 & 1 & -2 \\ 5 & -2 & 0 \end{pmatrix} \begin{pmatrix} \alpha e^{-2t} \\ (\gamma t+\beta)e^t \\ \gamma e^t \end{pmatrix}$, on conclut que les solutions du système différentiel (S) sont :

$$\begin{cases} x(t) = 3\alpha e^{-2t} + (3\gamma t + 3\beta - \gamma)e^t \\ y(t) = -4\alpha e^{-2t} + (\gamma t + \beta - 2\gamma)e^t \\ z(t) = 5\alpha e^{-2t} + (-2\gamma t - 2\beta)e^t \end{cases}, \qquad \text{avec } \alpha, \beta, \gamma \in \mathbb{R}.$$

6.2.4 Application au calcul des puissances d'une matrice

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Supposons que A est diagonalisable. On a donc $A = PDP^{-1}$ avec D diagonale et $P \in GL(n, \mathbb{K})$. Comme on l'a vu en 3.1.6.d, on a alors :

$$A^n = PDP^{-1}PDP^{-1} \cdots PDP^{-1} = PD^nP^{-1} \text{ pour tout entier } n \geq 1.$$

Mais D^n est elle-même diagonale (voir 6.1.1.a), ce qui permet de calculer aisément A^n.

On utilise ce type de calculs en particulier pour la détermination du terme général de suites définies par certains types de relations de récurrence.

Des exercices sur cette méthode seront vus en TD (voir aussi ci-dessous ex. 7 et 10).

6.3 Exercices sur le chapitre 6

Exercice 1. Déterminer les valeurs propres, avec leur multiplicité, de chacune des matrices suivantes à cœfficients réels. Lesquelles sont diagonalisables ?

$$A = \begin{pmatrix} 0 & 5 & 8 \\ 5 & 0 & 8 \\ 8 & 5 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} -4 & 1 & 0 & 1 \\ -2 & -1 & 0 & 1 \\ -12 & 6 & 3 & 1 \\ -2 & 1 & 0 & -1 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & -1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & -1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 2 & -3 \end{pmatrix}.$$

Exercice 2. Diagonaliser $A = \begin{pmatrix} -4 & -6 & 0 \\ 3 & 5 & 0 \\ 3 & 6 & 5 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$ et calculer A^n pour tout $n \in \mathbb{N}$.

Exercice 3. Pour quelles valeurs des paramètres réels les matrices suivantes sont-elles diagonalisables ?

$$A = \begin{pmatrix} 2 & 3 & -6 \\ m-6 & m-7 & -m+12 \\ m-3 & m-3 & -m+5 \end{pmatrix}, \qquad B = \begin{pmatrix} a & 0 & b \\ 0 & a+b & 0 \\ b & 0 & a \end{pmatrix}, \qquad C = \begin{pmatrix} a^2 & ab & ab & b^2 \\ ab & a^2 & b^2 & ab \\ ab & b^2 & a^2 & ab \\ b^2 & ab & ab & a^2 \end{pmatrix}.$$

Exercice 4. a) Montrer que, si λ est une valeur propre d'un endomorphisme f d'un $\mathbb{K}$-espace vectoriel E, alors λ^n est valeur propre de l'endomorphisme f^n pour tout $n \in \mathbb{N}^*$; que peut-on dire des sous-espaces propres correspondants ?

b) Pour tout $n \in \mathbb{N}^*$, déterminer les valeurs propres et sous-espaces propres dans $\mathbb{R}^4$ de la matrice A^n, où : $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 2 & 0 & -1 & 0 \\ 0 & 7 & 0 & 6 \\ 0 & 0 & 3 & 0 \end{pmatrix}$. (Utiliser la question **a**.)

Exercice 5. Soient $a, b \in \mathbb{C}$. On considère dans $\mathcal{M}_n(\mathbb{C})$ la matrice $A = (a_{i,j})_{1 \leq i,j \leq n}$ définie par : $a_{i,j} = b$ si $i \neq j$ et $a_{i,i} = a$. Montrer que A est diagonalisable.

Exercice 6. a) Un endomorphisme d'un $\mathbb{K}$-espace vectoriel E non-nul admet-il nécessairement des valeurs propres ?
b) Deux matrices semblables dans $\mathcal{M}_n(\mathbb{K})$ ont le même polynôme caractéristique (rappeler la preuve) ; a-t-on la réciproque ?
c) Soit $A \in \mathcal{M}_n(\mathbb{K})$. Montrer que la somme des valeurs propres de A dans $\mathbb{C}$ (chacune répétée autant de fois que sa multiplicité) est égale à la trace de A.
d) Soient $A \in \mathcal{M}_n(\mathbb{K})$ et $P(x)$ son polynôme caractéristique. Quelle relation a-t-on entre $P(x)$ et le polynôme caractéristique de la transposée ${}^tA$. On suppose de plus que $A \in GL(n, \mathbb{K})$, et l'on note $Q(x)$ est le polynôme caractéristique de A^{-1} ; quelle relation a-t-on pour tout $\lambda \in \mathbb{K}^*$ entre $Q(\lambda)$ et $P(\lambda^{-1})$?

Exercice 7. On considère deux suites de nombres réels $(u_n)_{n \geq 0}$ et $(v_n)_{n \geq 0}$ telles que : $u_{n+1} = -10u_n - 28v_n$ et $v_{n+1} = 6u_n + 16v_n$ pour tout $n \in \mathbb{N}$. En diagonalisant une matrice de $\mathcal{M}_2(\mathbb{R})$ adaptée, calculer les termes généraux u_n et v_n en fonction de u_0, v_0 et n.

Exercice 8. Soit (E) la courbe du plan d'équation : $x^2 - xy + y^2 - \frac{1}{2} = 0$, par rapport à un repère orthonormé. En diagonalisant la matrice $\left(\begin{smallmatrix} 2 & -1 \\ -1 & 2 \end{smallmatrix}\right)$, montrer que (E) est une ellipse centrée en l'origine.

Exercice 9. Déterminer les fonctions réelles $x(t), y(t), z(t)$ dérivables sur $\mathbb{R}$ solutions des systèmes différentiels :

$$\left\{\begin{array}{l} x'(t) = 3x(t) - 5y(t) \\ y'(t) = \quad\quad - 2y(t) \\ z'(t) = 5x(t) - 5y(t) - 2z(t) \end{array}\right. \qquad \left\{\begin{array}{l} x'(t) = 3x(t) - 2y(t) \\ y'(t) = \quad\quad y(t) \\ z'(t) = 2x(t) - 2y(t) + z(t) \end{array}\right. \qquad \left\{\begin{array}{l} x'(t) = 9x(t) + 5y(t) + 5z(t) \\ y'(t) = -5x(t) - y(t) - 5z(t) \\ z'(t) = -5x(t) - 5y(t) - z(t) \end{array}\right.$$

Exercice 10. Soient $a \in \mathbb{R}^*$ et A la matrice $\begin{pmatrix} 0 & a & a^2 \\ a^{-1} & 0 & a \\ a^{-2} & a^{-1} & 0 \end{pmatrix}$. Diagonaliser A. En déduire, pour tout $n \in \mathbb{N}$, le calcul des réels α_n et β_n tels que $A^n = \alpha_n A + \beta_n I_3$. Montrer que $A \in GL(3, \mathbb{R})$, et calculer A^{-1}.

Exercice 11. Pour tout $\alpha \in \mathbb{R}$, on considère dans $M_3(\mathbb{R})$ la matrice :

$$A_\alpha = \frac{1}{4} \begin{pmatrix} 6\alpha+1 & 2\alpha-1 & -\sqrt{2}(2\alpha-1) \\ 2\alpha-1 & 6\alpha+1 & \sqrt{2}(2\alpha-1) \\ -\sqrt{2}(2\alpha-1) & \sqrt{2}(2\alpha-1) & 4\alpha+2 \end{pmatrix}.$$

On note ϕ_α l'endomorphisme du $\mathbb{R}$-espace vectoriel $\mathbb{R}^3$ dont la matrice par rapport à la base canonique $\mathcal{B} = (e_1, e_2, e_3)$ de $\mathbb{R}^3$ est A_α.
a) Déterminer une base $\mathcal{B}'$ de $\mathbb{R}^3$ formée de vecteurs propres de A_0.
b) Montrer que, pour tout $\alpha \in \mathbb{R}$, la matrice de ϕ_α dans la base $\mathcal{B}'$ est une matrice diagonale D_α.
c) En déduire, par un calcul simple, que $A_\alpha A_\beta = A_{2\alpha\beta}$ pour tous $\alpha, \beta \in \mathbb{R}$, et que $A_{1/2} = I_3$. Pour quelles valeurs de α la matrice A_α est-elle inversible ?
d) Que peut-on en déduire pour l'ensemble $G = \{A_\alpha \,;\, \alpha \in \mathbb{R}^*\}$?

Chapitre 7

Réduction des endomorphismes (2)

7.1 Polynôme minimal

7.1.1 Théorème de Cayley-Hamilton

a) Données et notations.
Dans tout le chapitre, on note $E = \mathbb{K}^n$, avec n fixé ≥ 1.
On désigne par $\mathcal{B} = (e_1, \ldots, e_n)$ la base canonique de E.
Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée d'ordre n à coefficients dans $\mathbb{K}$.
Soit f l'endomorphisme de E dont A est la matrice par rapport à la base canonique : $A = \mathrm{Mat}_{\mathcal{B}}(f)$.
Soit $P_A(x)$ le polynôme caractéristique de A (ou encore de f).
On sait que $P_A(x)$ est de la forme :

$$P_A(x) = \alpha_n x^n + \alpha_{n-1}x^{n-1} + \cdots + \alpha_1 x + \alpha_0,$$

où les coefficients α_i (pour $0 \leq i \leq n$) appartiennent à $\mathbb{K}$. Comme on l'a vu au chapitre 6, certains sont connus : $\alpha_n = (-1)^n$, $\alpha_{n-1} = (-1)^{n-1} \operatorname{tr} A$, $\alpha_0 = \det A$.

On peut alors former la matrice :

$$P_A(A) = \alpha_n A^n + \alpha_{n-1}A^{n-1} + \cdots + \alpha_1 A + \alpha_0 I_n.$$

Cette écriture a bien un sens dans l'algèbre $\mathcal{M}_n(\mathbb{K})$: c'est une combinaison linéaire (avec les coefficients α_i) des puissances A^i de A pour i allant de n à 0.
La notation $P_A(A)$ est naturelle dans la mesure où cette matrice est obtenue en "appliquant" le polynôme caractéristique $P_A(x)$ à A.
Le théorème suivant (qui est à la base de tout ce que l'on va faire dans ce chapitre) montre qu'en fait, la matrice $P_A(A)$ que l'on obtient est toujours égale à la matrice nulle.

b) Théorème (de Cayley-Hamilton).
Toute matrice carrée annule son polynôme caractéristique. En d'autres termes :

$$\textit{pour toute } A \in \mathcal{M}_n(\mathbb{K}), \textit{ on a : } P_A(A) = O_n.$$

Preuve. Admis. □

7.1.2 Notion de polynôme minimal

a) Données et notations.
On reprend les données ci-dessus.
Pour tout polynôme $Q(x) = \lambda_m x^m + \cdots + \lambda_1 x + \lambda_0$, on peut former de la même façon qu'au paragraphe précédent la matrice $Q(A) = \lambda_m A^m + \cdots + \lambda_1 A + \lambda_0 I_n$. On dit que la matrice A annule le polynôme $Q(x)$ lorsque la matrice $Q(A)$ est égale à la matrice nulle O_n.

Le théorème de Cayley-Hamilton affirme que A annule le polynôme caractéristique $P_A(x)$; la proposition suivante montre que l'on peut trouver un polynôme plus simple que $P_A(x)$ (en un sens que l'on va préciser) qui est aussi annulé par A.

b) Proposition.

Pour toute matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$, il existe un unique polynôme $\pi_A(X)$ vérifiant :

(i) *$P_A(x)$ est un multiple de $\pi_A(x)$,*
[ce qui signifie qu'il existe un polynôme $Q(x)$ tel que $P_A(x) = \pi_A(x) \times Q(x)$,
et ce qui implique en particulier que $\deg \pi_A(x) \leq \deg P_A(x)$] ;

(ii) *A annule $\pi_A(x)$ [c'est-à-dire $\pi_A(A) = O_n$] ;*

(iii) *tout polynôme annulé par A est un multiple de $\pi_A(x)$;*

(iv) *$\pi_A(x)$ est unitaire,*
[ce qui signifie que le coefficient du terme de plus haut degré dans $\pi_A(x)$ est 1].

Principe de la preuve. Elle repose sur les propriétés arithmétiques élémentaires de l'algèbre des polynômes $\mathbb{K}[x]$ à coefficients dans $\mathbb{K}$ (le fait que tout idéal est principal, et donc en particulier l'idéal d'annulation de la matrice A). Dans le cadre de cet enseignement, on admet ce résultat. □

c) Définition. Pour toute matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$, le polynôme $\pi_A(X)$ déterminé à la proposition précédente est appelé le *polynôme minimal* de A.

d) Proposition.

Pour toute matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$, les zéros du polynôme minimal $\pi_A(X)$ sont exactement les valeurs propres de A.

En d'autres termes :

$P_A(x)$ et $\pi_A(x)$ ont exactement les mêmes zéros, qui sont les valeurs propres de A.

En outre, comme $P_A(x)$ est un multiple de $\pi_A(x)$, il est clair que pour toute valeur propre λ de A :

$$\begin{pmatrix} \text{la multiplicité de } \lambda \text{ en} \\ \text{tant que zéro de } \pi_A(x) \end{pmatrix} \leq \begin{pmatrix} \text{la multiplicité de } \lambda \text{ en} \\ \text{tant que zéro de } P_A(x) \end{pmatrix}.$$

Preuve. Soit λ un zéro de $\pi_A(x)$. Donc $\pi_A(\lambda) = 0$. Mais d'après le point (i) de la proposition 7.1.2.b, on sait que $P_A(x)$ est de la forme $\pi_A(x)Q(x)$ pour un certain polynôme $Q(x)$. Il est clair alors que $P_A(\lambda) = 0$.
Réciproquement, soit λ un zéro de $P_A(x)$. Donc λ est une v.p. de A (c'est-à-dire de l'endomorphisme f dont A est la matrice par rapport à $\mathcal{B}$). Il existe donc un vecteur non-nul v de E tel que $f(v) = \lambda.v$, ou encore une matrice colonne non-nulle $X \in \mathcal{M}_{n,1}(\mathbb{K})$ (celle des composantes de v dans la base $\mathcal{B}$), telle que $AX = \lambda.X$. En multipliant à gauche par A, on en déduit que : $A^2X = \lambda^2.X$, puis $A^3X = \lambda^3.X$ et finalement $A^jX = \lambda^j.X$ pour tout $j \geq 0$.
Posons $\pi_A(x) = x^m + \beta_{m-1}x^{m-1} + \cdots + \beta_1 x + \beta_0$, où $m = \deg \pi_A(x) \leq \deg P_A(x) = n$ et où les β_i appartiennent à $\mathbb{K}$. D'après le point (ii) de la proposition 7.1.2.b, on a $\pi_A(A) = O_n$. En notant ici 0 la matrice colonne nulle, on a donc :

$$\begin{aligned}
\pi_A(A)X = 0 \;&\Rightarrow\; \left(A^m + \beta_{m-1}A^{m-1} + \cdots + \beta_1 A + \beta_0 I_n\right) X = 0, \\
&\Rightarrow\; A^m X + \beta_{m-1}A^{m-1}X + \cdots + \beta_1 AX + \beta_0 X = 0, \\
&\Rightarrow\; \lambda^m X + \beta_{m-1}\lambda^{m-1}X + \cdots + \beta_1 \lambda X + \beta_0 X = 0, \\
&\Rightarrow\; \pi_A(\lambda)X = 0,
\end{aligned}$$

ce qui, puisque $X \neq 0$ par hypothèse, prouve que $\pi_A(\lambda)$ est nul dans $\mathbb{K}$. □

e) En résumé.

Dans toute la suite, on supposera que $P_A(x)$ se décompose en produit de facteurs de degré 1 (c'est-à-dire que A est trigonalisable) ce qui est toujours possible dans $\mathbb{C}$ (voir chapitre 6).
On notera $\lambda_1, \lambda_2, \ldots, \lambda_s$ les valeurs propres *distinctes* de A, et :

$$P_A(x) = (-1)^n(x-\lambda_1)^{q_1}(x-\lambda_2)^{q_2}\cdots(x-\lambda_s)^{q_s}, \quad n = \deg P_A(x) = q_1+q_2+\cdots+q_s,$$
$$\pi_A(x) = (x-\lambda_1)^{p_1}(x-\lambda_2)^{p_2}\cdots(x-\lambda_s)^{p_s}, \quad m = \deg \pi_A(x) = p_1+p_2+\cdots+p_s,$$

avec $1 \leq p_i \leq q_i \leq n$ pour tout $1 \leq i \leq s$.

f) Exemple.
Supposons que l'on ait une matrice $A \in \mathcal{M}_5(\mathbb{K})$ telle que $P_A(x) = -(x-2)^3(x+1)^2$.
Alors, a priori, $\pi_A(x)$ peut valoir :

$$\begin{array}{lllll} (x-2)(x+1), & \text{ou} & (x-2)^2(x+1), & \text{ou} & (x-2)^3(x+1), \\ (x-2)(x+1)^2, & \text{ou} & (x-2)^2(x+1)^2, & \text{ou} & (x-2)^3(x+1)^2. \end{array}$$

Pour déterminer ce que vaut effectivement $\pi_A(x)$:
on calcule $(A-2I_5)(A+I_5)$.
Si ce produit est nul, c'est fini : $\pi_A(x) = (x-2)(x+1)$.
Sinon, on calcule $(A-2I_5)^2(A+I_5)$.
Si ce produit est nul, c'est fini : $\pi_A(x) = (x-2)^2(x+1)$.
Sinon, on calcule $(A-2I_5)^3(A+I_5)$.
Si ce produit est nul, c'est fini : $\pi_A(x) = (x-2)^3(x+1)$.
Sinon, on calcule $(A-2I_5)(A+I_5)^2$.
Si ce produit est nul, c'est fini : $\pi_A(x) = (x-2)(x+1)^2$.
Sinon, on calcule $(A-2I_5)^2(A+I_5)^2$.
Si ce produit est nul, c'est fini : $\pi_A(x) = (x-2)^2(x+1)^2$.
Sinon, on sait d'après le théorème de Cayley-Hamilton que $(A-2I_5)^3(A+I_5)^2 = O_5$.
Donc dans ce dernier cas, $\pi_A(x) = (x-2)^3(x+1)^2 = -P_A(x)$.

On conçoit que de tels calculs directs sont vite fastidieux, voire inextricables à la main pour des matrices un peu grandes. D'où l'importance des arguments théoriques plus généraux que l'on va maintenant développer.

7.2 Sous-espaces caractéristiques

7.2.1 Notion de sous-espace caractéristique

a) Rappel. Avec les données et notations précédentes, rappelons que, pour toute v.p. λ_i de A, on appelle sous-espace propre associé à λ_i le noyau :

$$E_i = \operatorname{Ker}(f - \lambda_i \operatorname{id}_E),$$

et que sa dimension, que l'on notera r_i, vérifie $r_i \leq q_i$.

b) Définition. Pour toute v.p. λ_i de A, on appelle sous-espace caractéristique associé à λ_i le noyau :

$$F_i = \operatorname{Ker}(f - \lambda_i \operatorname{id}_E)^{q_i}.$$

c) Remarque. Précisons que la notation avec une puissance est relative à la loi $\circ$ dans $\operatorname{End} E$, c'est-à-dire que : $(f-\lambda_i \operatorname{id}_E)^{q_i} = (f-\lambda_i \operatorname{id}_E)\circ(f-\lambda_i \operatorname{id}_E)\circ\ldots\circ(f-\lambda_i \operatorname{id}_E)$. En particulier, il est clair que $E_i \subseteq F_i$ pour tout $1 \leq i \leq s$. Le théorème suivant précise cette inclusion.

d) Théorème. *On reprend toutes les données et notations précédentes.*

(i) *Pour toute valeur propre λ_i de A, on a :* $F_i = \operatorname{Ker}(f-\lambda_i \operatorname{id}_E)^{p_i}$, $E_i \subseteq F_i$, $\dim F_i = q_i$.

(ii) *E est somme directe des sous-espaces caractéristiques : $E = F_1 \oplus F_2 \oplus \cdots \oplus F_s$.*

Commentaire. On admet ce résultat. Remarquons simplement que l'égalité $F_i = \operatorname{Ker}(f - \lambda_i \operatorname{id}_E)^{p_i}$ est meilleure que la définition $F_i = \operatorname{Ker}(f - \lambda_i \operatorname{id}_E)^{q_i}$ puisqu'a priori $p_i \leq q_i$.

e) Corollaire. *Les conditions suivantes sont équivalentes :*
(i) *A diagonalisable ;*
(ii) *pour toute valeur propre λ_i de A, on a $F_i = E_i$;*
(iii) *pour toute valeur propre λ_i de A, on a $p_i = 1$;*
(iv) *le polynôme minimal de A n'a que des termes de degré 1 : $\pi_A(x) = (x - \lambda_1) \cdots (x - \lambda_s)$.*

Commentaire. Il est clair que (iii) $\Leftrightarrow$ (iv). Comme par définition A diagonalisable $\Leftrightarrow E = E_1 \oplus E_2 \oplus \cdots \oplus E_s$, l'équivalence (i) $\Leftrightarrow$ (ii) résulte du point (ii) du théorème précédent. L'implication (iii) $\Rightarrow$ (ii) découlant du point (i) du théorème précédent, seule la réciproque (ii) $\Rightarrow$ (iii) est non-triviale. Sa preuve repose sur le lemme dit "des noyaux" que l'on va voir au paragraphe suivant.

7.2.2 Suite des noyaux

a) Définition. Pour toute valeur propre λ_i de A, on appelle suite des noyaux associée à λ_i la suite croissante des sous-espaces vectoriels :

$$\underbrace{E_i}_{\dim = r_i} = \operatorname{Ker}(f - \lambda_i \operatorname{id}_E) \subseteq \operatorname{Ker}(f - \lambda_i \operatorname{id}_E)^2 \subseteq \cdots \subseteq \operatorname{Ker}(f - \lambda_i \operatorname{id}_E)^{q_i} = \underbrace{F_i}_{\dim = q_i}.$$

Cette suite est donc formée de q_i sous-espaces, mais d'après le point (i) du théorème 7.2.1.d, elle est stationnaire à partir de $\operatorname{Ker}(f - \lambda_i \operatorname{id}_E)^{p_i} = \cdots = \operatorname{Ker}(f - \lambda_i \operatorname{id}_E)^{q_i} = F_i$.

b) Lemme des noyaux. *Pour toute valeur propre λ_i de A, choisissons un entier quelconque a_i tel que $1 \leq a_i \leq q_i$. On a :*
(i) $E_i \subseteq \operatorname{Ker}(f - \lambda_i \operatorname{id}_E)^{a_i} \subseteq F_i$ *pour tout $1 \leq i \leq s$;*
(ii)
$$\operatorname{Ker}\big[(f - \lambda_1 \operatorname{id}_E)^{a_1} \circ (f - \lambda_2 \operatorname{id}_E)^{a_2} \circ \cdots \circ (f - \lambda_s \operatorname{id}_E)^{a_s}\big]$$
$$= \operatorname{Ker}\big[(f - \lambda_1 \operatorname{id}_E)^{a_1}\big] \oplus \operatorname{Ker}\big[(f - \lambda_2 \operatorname{id}_E)^{a_2}\big] \oplus \cdots \oplus \operatorname{Ker}\big[(f - \lambda_s \operatorname{id}_E)^{a_s}\big].$$

Preuve. Elle repose sur des propriétés arithmétiques des polynômes. On admet ici le résultat. □

c) Commentaire. Appliquons le point (ii) du lemme avec $a_i = q_i$ pour tout $1 \leq i \leq s$, et donc $\operatorname{Ker}\big[(f - \lambda_i \operatorname{id}_E)^{a_i}\big] = F_i$; on déduit que $F_1 \oplus \cdots \oplus F_s$ est égal au noyau de l'endomorphisme $h = (f - \lambda_1 \operatorname{id}_E)^{q_1} \circ (f - \lambda_2 \operatorname{id}_E)^{q_2} \circ \cdots \circ (f - \lambda_s \operatorname{id}_E)^{q_s}$. Or h n'est autre que $P_f(f)$. Ce dernier est nul d'après le théorème de Cayley-Hamilton, d'où $\operatorname{Ker} h = E$, et donc $F_1 \oplus \cdots \oplus F_s = E$. C'est le point (ii) du théorème 7.2.1.d.

7.2.3 Exemples

a) Premier exemple. Soit $A = \begin{pmatrix} -4 & 1 & 0 & 1 \\ -2 & -1 & 0 & 1 \\ -12 & 6 & 3 & 1 \\ -2 & 1 & 0 & -1 \end{pmatrix}$. On calcule $P_A(x) = (x-3)(x+2)^3$.

Par les méthodes habituelles, on détermine $E_1 = \operatorname{Ker}(f - 3 \operatorname{id}_E)$ (on sait qu'il est de dimension 1) et $E_2 = \operatorname{Ker}(f + 2 \operatorname{id}_E)$. On trouve que $\dim E_2 = 1$ ce qui, comme -2 est v.p. triple, prouve que A n'est pas diagonalisable. A priori, le polynôme minimal de A peut valoir : $(x-3)(x+2)^3$, ou $(x-3)(x+2)^2$, ou $(x-3)(x+2)$. Mais ce dernier cas est exclu puisque n'est pas A diagonalisable (voir 7.2.1.e). Donc $(A - 3I_4)(A + 2I_4)$ est non-nulle. Comme par ailleurs on sait que $(A - 3I_4)(A + 2I_4)^3$ est nulle d'après le théorème de Cayley-Hamilton, c'est le calcul de $(A - 3I_4)(A + 2I_4)^2$ qui permet de trancher. On fait le calcul de ce produit matriciel :

$$\begin{pmatrix} -7 & 1 & 0 & 1 \\ -2 & -4 & 0 & 1 \\ -12 & 6 & 0 & 1 \\ -2 & 1 & 0 & -4 \end{pmatrix} \begin{pmatrix} -2 & 1 & 0 & 1 \\ -2 & 1 & 0 & 1 \\ -12 & 6 & 5 & 1 \\ -2 & 1 & 0 & 1 \end{pmatrix}^2 = \begin{pmatrix} -7 & 1 & 0 & 1 \\ -2 & -4 & 0 & 1 \\ -12 & 6 & 0 & 1 \\ -2 & 1 & 0 & -4 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -50 & 25 & 25 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = O_4.$$

On trouve $(A - 3I_4)(A + 2I_4)^2 = O_4$; on conclut que $\pi_A(x) = (x-3)(x+2)^2$.

b) Second exemple. Soit $A = \begin{pmatrix} 1 & -1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & -1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 2 & -3 \end{pmatrix}$. On calcule $P_A(x) = -(x-1)^3(x+1)^2$.

On détermine les sous-espaces propres ; on obtient :

$\lambda_1 = 1, \quad E_1 = \mathrm{Ker}(f - \mathrm{id}_E)$, de dimension $r_1 = 2$ [une base est $(e_1, e_2 + e_3)$] ;
$\lambda_2 = -1, \quad E_2 = \mathrm{Ker}(f + \mathrm{id}_E)$, de dimension $r_2 = 1$ [une base est $(e_1 + e_2 + e_3 - 2e_4 - 2e_5)$].

Donc A n'est pas diagonalisable. En particulier, $\pi_A(x) \neq (x+1)(x-1)$.
On forme la suite des noyaux :

$E_1 = \mathrm{Ker}(f - \mathrm{id}_E) \subseteq \mathrm{Ker}(f - \mathrm{id}_E)^2 \subseteq \mathrm{Ker}(f - \mathrm{id}_E)^3 = F_1$, avec $\dim E_1 = r_1 = 2$ et $\dim F_1 = q_1 = 3$,
$E_2 = \mathrm{Ker}(f + \mathrm{id}_E) \subseteq \mathrm{Ker}(f + \mathrm{id}_E)^2 = F_2$, avec $\dim E_2 = r_2 = 1$ et $\dim F_2 = q_2 = 2$.

Le seul noyau à déterminer est $\mathrm{Ker}(f - \mathrm{id}_E)^2$ qui, au vu des dimensions, est égal à E_1 ou à F_1.

On calcule pour cela $(A - I_5)^2 = \begin{pmatrix} 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & -4 & 8 \\ 0 & 0 & 0 & -8 & 12 \end{pmatrix}$; on en déduit qu'une base de $\mathrm{Ker}(f - \mathrm{id}_E)^2$ est (e_1, e_2, e_3), d'où $\dim \mathrm{Ker}(f - \mathrm{id}_E)^2 = 3$ et donc $\mathrm{Ker}(f - \mathrm{id}_E)^2 = F_1$. En résumé :

$$\underbrace{E_1}_{r_1=2} = \mathrm{Ker}(f - \mathrm{id}_E) \subsetneq \mathrm{Ker}(f - \mathrm{id}_E)^2 = \mathrm{Ker}(f - \mathrm{id}_E)^3 = \underbrace{F_1}_{q_1=3},$$
$$\underbrace{E_2}_{r_2=1} = \mathrm{Ker}(f + \mathrm{id}_E) \subsetneq \mathrm{Ker}(f + \mathrm{id}_E)^2 = \underbrace{F_2}_{q_2=2}.$$

D'après le lemme des noyaux, $\mathrm{Ker}\big[(f - \mathrm{id}_E) \circ (f + \mathrm{id}_E)^2\big] = \mathrm{Ker}(f - \mathrm{id}_E) \oplus \mathrm{Ker}(f + \mathrm{id}_E)^2 = E_1 \oplus F_2$.
Ce noyau est donc de dimension $r_1 + q_2 = 4 < 5$, de sorte que l'endomorphisme $(f - \mathrm{id}_E) \circ (f + \mathrm{id}_E)^2$ n'est pas nul, ou encore $(A - I_5)(A + I_5)^2 \neq O_5$.
De même, $\mathrm{Ker}\big[(f - \mathrm{id}_E)^2 \circ (f + \mathrm{id}_E)\big] = \mathrm{Ker}(f - \mathrm{id}_E)^2 \oplus \mathrm{Ker}(f + \mathrm{id}_E) = F_1 \oplus E_2$ est de dimension $q_1 + r_2 = 4 < 5$, de sorte que l'endomorphisme $(f - \mathrm{id}_E)^2 \circ (f + \mathrm{id}_E)$ n'est pas nul, ou encore $(A - I_5)^2(A + I_5) \neq O_5$.
En revanche, $\mathrm{Ker}\big[(f - \mathrm{id}_E)^2 \circ (f + \mathrm{id}_E)^2\big] = \mathrm{Ker}(f - \mathrm{id}_E)^2 \oplus \mathrm{Ker}(f + \mathrm{id}_E)^2 = F_1 \oplus F_2 = \mathbb{R}^5$, de sorte que l'endomorphisme $(f - \mathrm{id}_E)^2 \circ (f + \mathrm{id}_E)^2$ est nul, c'est-à-dire $(A - I_5)^2(A + I_5)^2 = O_5$.
On conclut que le polynôme minimal est $\pi_A(x) = (x-1)^2(x+1)^2$.

7.3 Réduction de Jordan

7.3.1 Synthèse sur la méthode de réduction de Jordan

a) Données et résumé des résultats précédents.

On fixe une matrice A carrée d'ordre n à coefficients dans $\mathbb{K}$, et l'on note :

- f l'endomorphisme de E tel que $\mathrm{Mat}_{\mathcal{B}}(f) = A$,
- $\lambda_1, \lambda_2, \ldots, \lambda_s$ les valeurs propres distinctes de A (ou de f).

Pour toute valeur propre λ_i, on note :

- $P_A(x) = (-1)^n (x - \lambda_1)^{q_1} (x - \lambda_2)^{q_2} \cdots (x - \lambda_s)^{q_s}$ le polynôme caractéristique de A,
- $\pi_A(x) = (x - \lambda_1)^{p_1} (x - \lambda_2)^{p_2} \cdots (x - \lambda_s)^{p_s}$ le polynôme minimal de A,

avec donc :

- $n = \deg P_A(x) = q_1 + q_2 + \cdots + q_s$,
- $m = \deg \pi_A(x) = p_1 + p_2 + \cdots + p_s$,
- $1 \leq p_i \leq q_i \leq n$ pour tout $1 \leq i \leq s$.

On introduit également pour toute valeur propre λ_i :

- le sous-espace propre $E_i = \mathrm{Ker}(f - \lambda_i \,\mathrm{id}_E)$,
- le sous-espace caractéristique $F_i = \mathrm{Ker}(f - \lambda_i \,\mathrm{id}_E)^{q_i} = \mathrm{Ker}(f - \lambda_i \,\mathrm{id}_E)^{p_i}$.

Rappelons que $E_i \subseteq F_i$, que F_i est de dimension q_i, et que si l'on note r_i la dimension de E_i, on a :

$$1 \;\leq\; \underbrace{r_i = \dim E_i}_{\substack{\text{multiplicité}\\ \text{géométrique}}} \;\leq\; \underbrace{q_i = \dim F_i}_{\substack{\text{multiplicité}\\ \text{algébrique}}}.$$

On a vu que $E = F_1 \oplus F_2 \oplus \cdots \oplus F_s$; le théorème suivant affirme qu'il existe une base de E, formée en réunissant des bases des sous-espaces F_i, dans laquelle la matrice de f a une forme triangulaire standard d'un type particulier, dite FORME RÉDUITE DE JORDAN.

b) Le théorème principal.

Il existe une base $\mathcal{C}$ de E telle que

$$\mathrm{Mat}_{\mathcal{C}}(f) = \begin{pmatrix} \boxed{\boxed{T_1}} & \bigcirc & \bigcirc & \cdots & \bigcirc \\ \bigcirc & \boxed{\boxed{T_2}} & \bigcirc & \cdots & \bigcirc \\ \bigcirc & \bigcirc & \boxed{\boxed{T_3}} & \cdots & \bigcirc \\ \vdots & \vdots & & \ddots & \vdots \\ \bigcirc & \bigcirc & \bigcirc & \cdots & \boxed{\boxed{T_s}} \end{pmatrix},$$

où chaque matrice T_i est d'ordre q_i, de la forme

$$T_i = \begin{pmatrix} \boxed{J_1} & \bigcirc & \bigcirc & \cdots & \bigcirc \\ \bigcirc & \boxed{J_2} & \bigcirc & \cdots & \bigcirc \\ \bigcirc & \bigcirc & \boxed{J_3} & \cdots & \bigcirc \\ \vdots & \vdots & & \ddots & \vdots \\ \bigcirc & \bigcirc & \bigcirc & \cdots & \boxed{J_{r_i}} \end{pmatrix},$$

avec $J_1, J_2, \ldots, J_{r_i}$ des blocs de Jordan, c'est-à-dire des matrices de la forme :

$$J_k = \begin{pmatrix} \lambda_i & 1 & 0 & 0 & \cdots & 0 & 0 \\ & \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ & & \lambda_i & 1 & \cdots & 0 & 0 \\ & & & \ddots & \ddots & \vdots & \vdots \\ & \bigcirc & & & \ddots & 1 & 0 \\ & & & & & \lambda_i & 1 \\ & & & & & & \lambda_i \end{pmatrix}$$

dont le nombre et la taille permettent de déterminer explicitement T_i pour tout $1 \leq i \leq s$, et vérifient :

- *le nombre r_i de blocs de Jordan dans T_i = la dimension du sous-espace propre E_i,*
- *la taille du plus grand bloc de Jordan dans T_i = l'exposant p_i de $(x - \lambda_i)$ dans $\pi_A(x)$.*

Cette écriture est unique à l'ordre près des blocs (on les ordonne généralement par taille décroissante).

c) Remarque et exemple.

Un bloc de Jordan d'ordre 3 est de la forme $\begin{pmatrix} \lambda_i & 1 & 0 \\ 0 & \lambda_i & 1 \\ 0 & 0 & \lambda_i \end{pmatrix}$, un bloc de Jordan d'ordre 2 est de la forme $\begin{pmatrix} \lambda_i & 1 \\ 0 & \lambda_i \end{pmatrix}$, un bloc de Jordan d'ordre 1 est de la forme (λ_i). Ainsi, si dans le théorème ci-dessus la matrice T_i relative à la v.p. λ_i est d'ordre 4, elle peut être de l'une des cinq formes suivantes :

$$\begin{pmatrix} \lambda_i & 1 & 0 & 0 \\ 0 & \lambda_i & 1 & 0 \\ 0 & 0 & \lambda_i & 1 \\ 0 & 0 & 0 & \lambda_i \end{pmatrix}, \quad \begin{pmatrix} \boxed{\begin{matrix} \lambda_i & 1 & 0 \\ 0 & \lambda_i & 1 \\ 0 & 0 & \lambda_i \end{matrix}} & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ 0\ 0\ 0 & \boxed{\lambda_i} \end{pmatrix}, \quad \begin{pmatrix} \boxed{\begin{matrix} \lambda_i & 1 \\ 0 & \lambda_i \end{matrix}} & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & \boxed{\begin{matrix} \lambda_i & 1 \\ 0 & \lambda_i \end{matrix}} \end{pmatrix}, \quad \begin{pmatrix} \boxed{\begin{matrix} \lambda_i & 1 \\ 0 & \lambda_i \end{matrix}} & \begin{matrix} 0 \\ 0 \end{matrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ 0\ 0 & \boxed{\lambda_i} & 0 \\ 0\ 0 & 0 & \boxed{\lambda_i} \end{pmatrix}, \quad \begin{pmatrix} \boxed{\lambda_i} & 0 & 0 & 0 \\ 0 & \boxed{\lambda_i} & 0 & 0 \\ 0 & 0 & \boxed{\lambda_i} & 0 \\ 0 & 0 & 0 & \boxed{\lambda_i} \end{pmatrix}.$$

cas $\begin{cases} p_i=4 \\ r_i=1 \end{cases}$ cas $\begin{cases} p_i=3 \\ r_i=2 \end{cases}$ cas $\begin{cases} p_i=2 \\ r_i=2 \end{cases}$ cas $\begin{cases} p_i=2 \\ r_i=3 \end{cases}$ cas $\begin{cases} p_i=1 \\ r_i=4 \end{cases}$

d) Remarques.

• Concrètement, sur les exemples numériques que l'on aura à traiter, on détermine pour chaque v.p. λ_i de A le sous-espace propre associé (sa dimension donne le nombre r_i de blocs de Jordan dans T_i), puis (si nécessaire) la taille p_i du plus grand bloc en calculant le polynôme minimal.

• On détermine la base $\mathcal{C}$ du théorème en utilisant la suite des noyaux.

• En particulier,

(A diagonalisable) $\Leftrightarrow$ (pour tout $1 \leq i \leq s$, chaque bloc de Jordan J_k dans T_i est de taille 1),
$\Leftrightarrow$ (pour tout $1 \leq i \leq s$, T_i a autant de blocs de Jordan J_k que sa taille q_i).

On retrouve donc les conditions de diagonalisabilité vues précédemment :

(A diagonalisable) $\Leftrightarrow$ (pour tout $1 \leq i \leq s$, $p_i = 1$) $\Leftrightarrow$ ($\pi_A(x) = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_s)$),
(A diagonalisable) $\Leftrightarrow$ (pour tout $1 \leq i \leq s$, $r_i = q_i$) $\Leftrightarrow$ (pour tout $1 \leq i \leq s$, $E_i = F_i$).

7.3.2 Exemples

1) Soient a et b deux réels, et $T \in \mathcal{M}_8(\mathbb{R})$ la matrice (réduite à la forme de Jordan) :

$$T = \begin{pmatrix} a&1&0&0&0&0&0&0 \\ 0&a&1&0&0&0&0&0 \\ 0&0&a&0&0&0&0&0 \\ 0&0&0&a&0&0&0&0 \\ 0&0&0&0&b&1&0&0 \\ 0&0&0&0&0&b&0&0 \\ 0&0&0&0&0&0&b&1 \\ 0&0&0&0&0&0&0&b \end{pmatrix}.$$

• *Supposons d'abord que* $a \neq b$. Le nombre r_a de blocs de Jordan relatifs à la v.p. a est égal à la dimension du sous-espace propre E_a, alors que la taille p_a du plus grand bloc de Jordan relatif à la v.p. a est égal à l'exposant de $(x-a)$ dans le polynôme minimal de A. Ici, $r_a = 2$ et $p_a = 3$. De même, $r_b = 2 = p_b$. Donc $\dim E_a = \dim E_b = 2$ et $\pi_A(x) = (x-a)^3(x-b)^2$.
Par ailleurs, l'exposant q_a de $(x-a)$ dans le polynôme caractéristique de A est égal à la dimension du sous-espace caractéristique F_a ; comme $P_A(x) = (x-a)^4(x-b)^4$, on déduit $\dim F_a = \dim F_b = 4$.
• *Supposons maintenant que* $a = b$. On obtient de même : $r_a = 4$, $p_a = 3$ et $q_a = 8$, donc $\dim F_a = 8$, $\dim E_a = 4$, $P_A(x) = (x-a)$ et $\pi_A(x) = (x-a)^3$.

2) Soit $A \in \mathcal{M}_8(\mathbb{R})$ telle que le polynôme caractéristique $P_A(x)$ et le polynôme minimal $\pi_A(x)$ de A soient donnés par : $P_A(x) = (x-3)^4(x-5)^4$ et $\pi_A(x) = (x-3)^2(x-5)^3$.
Dans la réduite de Jordan T de A, la taille du plus grand bloc de Jordan relatif à la v.p. 5 est trois, donc le bloc complémentaire est nécessairement de taille un. La taille du plus grand bloc de Jordan relatif à la v.p. 3 est deux, donc le complément est formé soit d'un unique bloc d'ordre 2, soit de deux blocs d'ordre un. A l'ordre près des blocs (et sans noter les zéros), on a donc :

$$T = \begin{pmatrix} 5&1&&&&&& \\ &5&1&&&&& \\ &&5&&&&& \\ &&&5&&&& \\ &&&&3&1&& \\ &&&&&3&& \\ &&&&&&3&1 \\ &&&&&&&3 \end{pmatrix} \quad \text{ou} \quad T = \begin{pmatrix} 5&1&&&&&& \\ &5&1&&&&& \\ &&5&&&&& \\ &&&5&&&& \\ &&&&3&1&& \\ &&&&&3&& \\ &&&&&&3& \\ &&&&&&&3 \end{pmatrix}$$

En considérant le nombre de blocs de Jordan, on déduit que les dimensions des sous-espaces propres sont $\dim E_5 = \dim E_3 = 2$ dans le premier cas, et $\dim E_5 = 2$ et $\dim E_3 = 3$ dans le second cas.

7.3.3 Exemples

Exemple (a) (cas de deux valeurs propres doubles) Soit $A = \begin{pmatrix} 1 & -1 & 2 & -2 \\ 0 & 0 & 1 & -1 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & 1 & 0 \end{pmatrix} \in M_4(\mathbb{R})$.

On note f l'endomorphisme de $\mathbb{R}^4$ dont A est la matrice dans la base canonique $\mathcal{B}$.
On calcule $P_f(x) = P_A(x) = x^2(x-1)^2$. On a deux valeurs propres doubles $\lambda_1 = 0$ et $\lambda_2 = 1$.

La réduite de Jordan T de A est donc a priori de la forme $T = \begin{pmatrix} \boxed{T_1} & \begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix} \\ \begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix} & \boxed{T_2} \end{pmatrix}$, avec

$$\Big[\, T_1 = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right) \text{ (lorsque } r_1 = 2 \text{ et } p_1 = 1) \,\Big] \text{ ou } \Big[\, T_1 = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) \text{ (lorsque } r_1 = 1 \text{ et } p_1 = 2) \,\Big]$$

et

$$\Big[\, T_2 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \text{ (lorsque } r_2 = 2 \text{ et } p_2 = 1) \,\Big] \text{ ou } \Big[\, T_1 = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \text{ (lorsque } r_2 = 1 \text{ et } p_2 = 2) \,\Big]$$

• *On étudie d'abord la valeur propre $\lambda_1 = 0$.*

Un calcul simple à partir de $(A - 0.I_4) = A = \begin{pmatrix} 1 & -1 & 2 & -2 \\ 0 & 0 & 1 & -1 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & 1 & 0 \end{pmatrix}$ donne :

$$N_{1,1} = E_1 = \mathrm{Ker}(f - 0.\,\mathrm{id}_{\mathbb{R}^4}) = \mathrm{Ker}\, f = \{(x, x, 0, 0)\,;\ x \in \mathbb{R}\}, \text{ de dimension } 1,$$

donc $r_1 = 1$ et $T_1 = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$.

Il s'agit de trouver ensuite une base du sous-espace caractéristique F_1 associé à $\lambda_1 = 0$ par rapport laquelle la matrice de la restriction de f à F_1 est précisément T_1. Pour celà, un calcul simple à partir de $(A - 0.\,\mathrm{id}_n)^2 = A^2 = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 2 & -2 & 2 & -1 \\ 2 & -2 & 2 & -1 \end{pmatrix}$ donne :

$$N_{1,2} = F_1 = \mathrm{Ker}(f - 0.\,\mathrm{id}_{\mathbb{R}^4})^2 = \{(x, x+z, z, 0)\,;\ x, z \in \mathbb{R}\}, \text{ de dimension } 2.$$

Le sous-espace propre $N_{1,1} = E_1$ est strictement inclus dans le sous-espace caractéristique $N_{1,2} = F_1$. Choisissons un vecteur $v_2 \in N_{1,2}$ tel que $v_2 \notin N_{1,1}$; par exemple $v_2 = (0, 1, 1, 0)$. Posons $v_1 = f(v_2)$, qui vérifie $v_1 \in N_{1,1}$ c'est-à-dire $f(v_1) = 0$; on calcule $v_1 = f(0, 1, 1, 0) = (1, 1, 0, 0)$. Ainsi $f(v_1) = 0$ et $f(v_2) = v_1$, donc la matrice dans la base (v_1, v_2) de F_1 de la restriction de f à F_1 est bien $T_1 = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$.

• *On étudie ensuite la valeur propre $\lambda_2 = 1$.*

Un calcul simple à partir de $A - 1.I_4 = \begin{pmatrix} 0 & -1 & 2 & -2 \\ 0 & -1 & 1 & -1 \\ 1 & -1 & 0 & 0 \\ 1 & -1 & 1 & -1 \end{pmatrix}$ donne :

$$N_{2,1} = E_2 = \mathrm{Ker}(f - 1.\,\mathrm{id}_{\mathbb{R}^4}) = \{(0, 0, z, z)\,;\ z \in \mathbb{R}\}, \text{ de dimension } 1, \text{ donc } r_2 = 1 \text{ et } T_2 = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right).$$

Il s'agit de trouver ensuite une base du sous-espace caractéristique F_2 associé à $\lambda_2 = 1$ par rapport laquelle la matrice de la restriction de f à F_2 est précisément T_2. Pour celà, un calcul simple à partir de $(A - 1.I_4)^2 = \begin{pmatrix} 0 & 1 & -3 & 3 \\ 0 & 0 & -2 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ donne :

$$N_{2,2} = F_2 = \mathrm{Ker}(f - 1.\,\mathrm{id}_{\mathbb{R}^4})^2 = \{(x, 0, z, z)\,;\ x, z \in \mathbb{R}\}, \text{ de dimension } 2.$$

Le sous-espace propre $N_{2,1} = E_2$ est strictement inclus dans le sous-espace caractéristique $N_{2,2} = F_2$. Choisissons un vecteur $w_2 \in N_{2,2}$ tel que $w_2 \notin N_{2,1}$; par exemple $w_2 = (1, 0, 0, 0)$. Posons $w_1 = (f - \mathrm{id}_{\mathbb{R}^4})(w_2)$, qui vérifie $w_1 \in N_{2,1}$ c'est-à-dire $f(w_1) = w_1$; on calcule $w_1 = (f - \mathrm{id}_{\mathbb{R}^4})(1, 0, 0, 0) = (0, 0, 1, 1)$. Ainsi $f(w_1) = 1.w_1 + 0.w_2$ et $f(w_2) = 1.w_1 + 1.w_2$, donc la matrice dans la base (w_1, w_2) de F_2 de la restriction de f à F_2 est bien $T_2 = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$.

• *Synthèse et conclusion.* Comme on sait que $\mathbb{R}^4 = F_1 \oplus F_2$, on peut considérer la base $\mathcal{C} = (v_1, v_2, w_1, w_2)$ de $\mathbb{R}^4$. La matrice de f par rapport à $\mathcal{C}$ est la rduite de Jordan T de A. En notant P la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$, on a :

$$A = PTP^{-1}, \quad \text{avec} \quad P = \begin{pmatrix} 1&0&0&1\\ 1&1&0&0\\ 0&1&1&0\\ 0&0&1&0 \end{pmatrix} \quad \text{et} \quad T = \begin{pmatrix} 0&1&0&0\\ 0&0&0&0\\ 0&0&1&1\\ 0&0&0&1 \end{pmatrix}.$$

On a $p_1 = q_1 = p_2 = q_2 = 2$ et $\pi_A(x) = P_A(x) = x^2(x-1)^2$.

Exemple (b) (cas d'une v.p. simple et une v.p. triple) Soit $A = \begin{pmatrix} -1&1&1&0\\ -1&2&1&-1\\ 5&-3&-2&5\\ 4&-2&-2&3 \end{pmatrix} \in M_4(\mathbb{R})$.

On note f l'endomorphisme de $\mathbb{R}^4$ dont A est la matrice dans la base canonique $\mathcal{B}$.
On calcule $P_f(x) = P_A(x) = (x+1)(x-1)^3$. On a deux valeurs propres doubles $\lambda_1 = -1$ et $\lambda_2 = 1$.

La réduite de Jordan T de A est donc a priori de la forme $T = \begin{pmatrix} -1 & 0\,0\,0 \\ \begin{matrix}0\\0\\0\end{matrix} & \boxed{T_2} \end{pmatrix}$, avec

$$\underbrace{T_2 = \begin{pmatrix} 1&0&0\\ 0&1&0\\ 0&0&1 \end{pmatrix}}_{(\text{lorsque } r_2 = 3 \text{ et } p_2 = 1)} \quad \text{ou} \quad \underbrace{T_2 = \begin{pmatrix} 1&0&0\\ 0&1&1\\ 0&0&1 \end{pmatrix}}_{(\text{lorsque } r_2 = 2 \text{ et } p_2 = 2)} \quad \text{ou} \quad \underbrace{T_2 = \begin{pmatrix} 1&1&0\\ 0&1&1\\ 0&0&1 \end{pmatrix}}_{(\text{lorsque } r_2 = 1 \text{ et } p_2 = 3)}$$

• *On étudie d'abord la valeur simple propre $\lambda_1 = -1$.*

On détermine de façon évidente $E_1 = F_1 = \mathrm{Ker}(f + \mathrm{id}_{\mathbb{R}^4})$.
On obtient la droite $E_1 = \{(x, 0, 0, -x)\,;\, x \in \mathbb{R}\}$, de base $\{u_1\}$ où $u_1 = (1, 0, 0, -1)$.

• *On étudie ensuite la valeur propre triple $\lambda_2 = 1$.*

Le calcul de :

$$A - 1.I_4 = \begin{pmatrix} -2&1&1&0\\ -1&1&1&-1\\ 5&-3&-3&5\\ 4&-2&-2&2 \end{pmatrix}, \quad (A - 1.I_4)^2 = \begin{pmatrix} 8&-4&-4&4\\ 2&-1&-1&2\\ -2&1&1&-2\\ -8&4&4&-4 \end{pmatrix}, \quad (A - 1.I_4)^3 = \begin{pmatrix} -16&8&8&-8\\ 0&0&0&0\\ 0&0&0&0\\ 16&-8&-8&8 \end{pmatrix}$$

permet de déterminer les noyaux :

$N_{2,1} = \mathrm{Ker}(f - \mathrm{id}_{\mathbb{R}^4})$ d'équations $x = t = 0 = y + z = 0$, de dimension 1,
$N_{2,2} = \mathrm{Ker}(f - \mathrm{id}_{\mathbb{R}^4})^2$ d'équations $2x - y - z = t = 0$, de dimension 2,
$N_{2,3} = \mathrm{Ker}(f - \mathrm{id}_{\mathbb{R}^4})^3$ d'équation $2x - y - z - t = 0$, de dimension 3.

On est donc dans le cas où $r_2 = 1$ et $p_2 = q_2 = 3$, donc $T_2 = \begin{pmatrix} 1&1&0\\ 0&1&1\\ 0&0&1 \end{pmatrix}$ et l'on a les inclusions strictes :

$$E_2 N_{2,1} \subset N_{2,2} \subset N_{2,3} = F_2.$$

Il s'agit de trouver ensuite une base de F_2 par rapport laquelle la matrice de la restriction de f à F_2 est précisément T_2.
On choisit un vecteur $v_3 \in N_{2,3}$ tel que $v_3 \notin N_{2,2}$. On pose $v_2 = (f - \mathrm{id}_{\mathbb{R}^4})(v_3)$, qui vérifie $v_2 \in N_{2,2}$ et $v_2 \notin N_{2,1}$; par définition de v_2, on a $f(v_3) = v_2 + v_3$. On pose $v_1 = (f - \mathrm{id}_{\mathbb{R}^4})(v_2)$, qui vérifie $v_1 \in N_{2,1}$, donc $f(v_1) = v_1$; de plus par définition de v_1, on a $f(v_2) = v_1 + v_2$. Par construction, (v_1, v_2, v_3) est une base de F_2.
Si l'on part par exemple de $v_3 = (0, 0, 1, 1)$, on obtient $v_2 = (1, 0, 2, 0)$ et $v_1 = (0, 1, -1, 0)$.

• *Synthèse et conclusion.* On considère la base $\mathcal{C} = (u_1, v_1, v_2, v_3)$ de $\mathbb{R}^4$. La matrice de f par rapport à $\mathcal{C}$ est la rduite de Jordan T de A. En notant P la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$, on a :

$$A = PTP^{-1}, \quad \text{avec} \quad P = \begin{pmatrix} 1&0&1&0\\ 0&1&0&0\\ 0&-1&2&1\\ -1&0&0&1 \end{pmatrix} \quad \text{et} \quad T = \begin{pmatrix} -1&0&0&0\\ 0&1&1&0\\ 0&0&1&1\\ 0&0&0&1 \end{pmatrix}.$$

Exemple (c) (cas d'une unique valeur propre d'ordre cinq) Soit $A = \begin{pmatrix} 3&1&1&-1&0\\ 0&3&0&0&1\\ 1&-1&2&1&0\\ 1&-1&-1&4&1\\ 0&0&0&0&3 \end{pmatrix}$

On note f l'endomorphisme de $\mathbb{R}^5$ dont A est la matrice dans la base canonique $\mathcal{B}$.
On calcule $P_f(x) = P_A(x) = -(x-3)^5$. On a une unique valeur propre $\lambda = 3$ d'ordre 5.
Le calcul de :

$$A-3I_4=\begin{pmatrix}0&1&1&-1&0\\0&0&0&0&1\\1&-1&-1&1&0\\1&-1&-1&1&1\\0&0&0&0&0\end{pmatrix},\quad (A-3I_4)^2=\begin{pmatrix}0&0&0&0&0\\0&0&0&0&0\\0&1&1&-1&0\\0&1&1&-1&0\\0&0&0&0&0\end{pmatrix},\quad (A-3I_4)^3=\begin{pmatrix}0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}=O_5$$

permet de déterminer les noyaux :

$N_1=\mathrm{Ker}(f-3\,\mathrm{id}_{\mathbb{R}^5})=\{(0,y,z,z+y,0)\,;\,y,z\in\mathbb{R}\}$ de dimension 2,
$N_2=\mathrm{Ker}(f-3\,\mathrm{id}_{\mathbb{R}^5})^2=\{(x,y,z,y+z,s)\,;\,x,y,z,s\in\mathbb{R}\}$ de dimension 4,
$N_3=\mathrm{Ker}(f-3\,\mathrm{id}_{\mathbb{R}^5})^3=\mathbb{R}^5$ de dimension 5.

Le polynôme minimal est $\pi_A(x)=(x-3)^3$, donc $p=3$, qui est la taille du plus grand bloc de Jordan dans T. Le dimension du sous-espace propre N_1 est $r=2$, qui est le nombre de blocs de Jordan dans T. Donc :

$$T=\begin{pmatrix}\boxed{\begin{matrix}3&1&0\\0&3&1\\0&0&3\end{matrix}}&\begin{matrix}0&0\\0&0\\0&0\end{matrix}\\\begin{matrix}0&0&0\\0&0&0\end{matrix}&\boxed{\begin{matrix}3&1\\0&3\end{matrix}}\end{pmatrix}.$$

Il s'agit de trouver ensuite une base de $\mathbb{R}^5$ par rapport laquelle la matrice de f est précisément T. On choisit un vecteur $v_3\in\mathbb{R}^5$ quelconque mais tel que $v_3\notin N_2$. On pose $v_2=(f-3\,\mathrm{id}_{\mathbb{R}^5})(v_3)$, qui vérifie $v_2\in N_2$ et $v_2\notin N_1$; par définition de v_2, on a $f(v_3)=v_2+3v_3$. On pose $v_1=(f-3\,\mathrm{id}_{\mathbb{R}^5})(v_2)$, qui vérifie $v_1\in N_1$, donc $f(v_1)=3v_1$; de plus par définition de v_1, on a $f(v_2)=v_1+3v_2$. La famille libre (v_1,v_2,v_3) correspond donc au premier bloc de Jordan de T.

Dans cete construction, la droite vectorielle S_2 engendrée par v_3 est un supplémentaire dans $\mathbb{R}^5$ de N_2. Or N_2 de dimension 4 contient N_1 de dimension 2. On peut donc choisir un vecteur $w_2\in N_2$ tel que $w_2\notin N_1$ et tel que la famille (v_2,w_2) soit libre et engendre un supplémentaire S_1 de N_1 dans N_2. On pose $w_1=(f-3\,\mathrm{id}_{\mathbb{R}^5})(w_2)$, qui vérifie $w_1\in N_1$, donc $f(w_1)=3w_1$; de plus par définition de w_1, on a $f(w_2)=w_1+3w_2$. La famille libre (w_1,w_2) correspond donc au second bloc de Jordan de T. En résumé :

$$\mathbb{R}^5=N_2\oplus S_2=N_1\oplus S_1\oplus S_2,\ \text{ avec }\begin{cases}(v_1,w_1)\text{ base de } N_1,\\(v_2,w_2)\text{ base de } S_1,\\v_3\text{ base de } S_2,\end{cases}\ \text{ et }\begin{cases}f(v_1)=3v_1,\\f(v_2)=v_1+3v_2,\\f(v_3)=v_2+3f(v_3),\\f(w_1)=3w_1,\\f(w_2)=w_1+3w_2,\end{cases}$$

On en déduit que $\mathcal{C}=(v_1,v_2,v_3,w_1,w_2)$ est une base de $\mathbb{R}^5$, et que la matrice de f dans $\mathcal{C}$ est T. Numériquement, on peut choisir par exemple $v_3=(0,1,0,0,0)$, en déduire que $v_2=(1,0,-1,-1,0)$ et $v_1=(0,0,1,1,0)$; puis choisir $w_2=(0,0,0,0,1)$, d'où $w_1=(0,1,0,1,0)$. En notant P la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$, on conclut que :

$$A=PTP^{-1},\quad\text{avec}\quad P=\begin{pmatrix}0&1&0&0&0\\0&0&1&1&0\\1&-1&0&0&0\\1&-1&0&1&0\\0&0&0&0&1\end{pmatrix}\quad\text{et}\quad T=\begin{pmatrix}3&1&0&0&0\\0&3&1&0&0\\0&0&3&0&0\\0&0&0&3&1\\0&0&0&0&3\end{pmatrix}.$$

7.4 Exercices sur le chapitre 7

Exercice 1. Calculer le polynôme minimal des matrices suivantes :

$$\begin{pmatrix}0&0&0\\0&0&0\\1&0&0\end{pmatrix},\quad\begin{pmatrix}1&0&0\\0&1&0\\1&0&1\end{pmatrix},\quad\begin{pmatrix}0&0&0\\1&0&0\\0&1&0\end{pmatrix},\quad\begin{pmatrix}-4&1&0&1\\-2&-1&0&1\\-12&6&3&1\\-2&1&0&-1\end{pmatrix},\quad\begin{pmatrix}1&-1&1&0&1\\0&0&1&1&0\\0&-1&2&0&1\\0&0&0&1&-2\\0&0&0&2&-3\end{pmatrix}.$$

Exercice 2. Soit $A=\left(\begin{smallmatrix}3&1\\-1&1\end{smallmatrix}\right)\in\mathcal{M}_2(\mathbb{C})$. Montrer que A n'est pas diagonalisable. Utiliser le théorème de Cayley et Hamilton pour montrer que, pour tout $p\in\mathbb{N}$, on a :

$$A^{2p}=2^{2p}(pA-(2p-1)I_2)\qquad\text{et}\qquad A^{2p+1}=2^{2p}((2p+1)A-4pI_2).$$

Exercice 3. Soient $a\in\mathbb{R}$ et $A=\left(\begin{smallmatrix}0&0&1\\1&0&-1\\-1&1&a\end{smallmatrix}\right)\in\mathcal{M}_3(\mathbb{R})$.

Calculer le polynôme caractéristique $P_A(x)$. En utilisant le théorème de Cayley et Hamilton, en déduire que $A \in GL(3, \mathbb{R})$ et calculer A^{-1}.

Exercice 4. On considère la matrice : $A = \begin{pmatrix} 3 & 1 & 1 & -1 & 0 \\ 0 & 3 & 0 & 0 & 1 \\ 1 & -1 & 2 & 1 & 0 \\ 1 & -1 & -1 & 4 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}$.

Calculer son polynôme caractéristique $P_A(x)$. Montrer qu'elle admet une unique valeur propre. Déterminer le sous-espace propre associé. Quelles sont a priori les formes possibles de la réduite de Jordan de A et du polynôme minimal $\pi_A(x)$? Expliciter la suite des noyaux et déterminer la réduite de Jordan de A, en précisant une base adaptée et en donnant la matrice de passage correspondante. Calculer $\pi_A(x)$.

Exercice 5. On considère la matrice : $A = \begin{pmatrix} 1 & -1 & 2 & -2 \\ 0 & 0 & 1 & -1 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & 1 & 0 \end{pmatrix}$.

Calculer son polynôme caractéristique $P_A(x)$. Montrer qu'elle admet deux valeurs propres. Quelles sont a priori les formes possibles de la réduite de Jordan de A et du polynôme minimal $\pi_A(x)$? Déterminer pour chacune des valeurs propres le sous-espace propre associé, la suite des noyaux et le sous-espace caractéristique. Déterminer la réduite de Jordan de A, en précisant une base adaptée et en donnant la matrice de passage correspondante. Calculer $\pi_A(x)$.

Exercice 6. Même exercice que l'exercice 5 pour la matrice : $A = \begin{pmatrix} -1 & 1 & 1 & 0 \\ -1 & 2 & 1 & -1 \\ 5 & -3 & -2 & 5 \\ 4 & -2 & -2 & 3 \end{pmatrix}$.

En déduire le calcul de A^n pour tout $n \in \mathbb{N}$.

Exercice 7. Même exercice que l'exercice 5 pour la matrice : $A = \begin{pmatrix} 1 & 1 & -1 & 2 & -1 \\ 2 & 0 & 1 & -4 & -1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & -3 & 3 & -1 \end{pmatrix}$.

Exercice 8. Soit $A \in \mathcal{M}_5(\mathbb{R})$ de polynôme caractéristique $P_A(x) = -(x-a)^3(x-b)^2$ avec $a, b \in \mathbb{R},\ a \neq b$. Exprimer tous les cas possibles pour la forme réduite de Jordan de A, en précisant dans chaque cas le polynôme minimal $\pi_A(x)$, et les dimensions des sous-espaces propres.

Exercice 9. Dans chacun des cas suivants, $P_A(x)$ et $\pi_A(x)$ désignant respectivement le polynôme caractéristique et le polynôme minimal d'une matrice $A \in \mathcal{M}_n(\mathbb{R})$, déterminer tous les cas possibles pour la forme réduite de Jordan de A :

$$\begin{array}{llll} n = 6, & P_A(x) = (x-2)^4(x-3)^2, & \pi_A(x) = (x-2)^2(x-3)^2; & (1) \\ n = 5, & P_A(x) = -(x-7)^5 \quad , & \pi_A(x) = (x-7)^2; & (2) \\ n = 7, & P_A(x) = -(x-2)^7 \quad , & \pi_A(x) = (x-2)^3; & (3) \\ n = 8, & P_A(x) = (x-3)^4(x-5)^4, & \pi_A(x) = (x-3)^2(x-5)^2. & (4) \end{array}$$

Exercice 10. On considère la suite de réels $(u_n)_{n \geq 0}$ définie par : $u_0 = 1$, $u_1 = -2$, $u_2 = 3$, et $u_{n+3} = -u_{n+2} + 8u_{n+1} + 12u_n$ pour tout $n \in \mathbb{N}$. Utiliser la réduction de Jordan d'une matrice appropriée pour calculer le terme général u_n en fonction de n.

Exercice 11. Utiliser une réduction de Jordan pour déterminer les fonctions réelles $x(t), y(t), z(t)$ de la variable réelle t, de classe C^1 sur $\mathbb{R}$, solutions du système différentiel :

$$\begin{cases} x'(t) = 8x(t) - y(t) - 5z(t) \\ y'(t) = -2x(t) + 3y(t) + z(t) \\ z'(t) = 4x(t) - y(t) - z(t) \end{cases}$$

Exercice 12. Soient $\alpha, \beta \in \mathbb{R}$ non-nuls tels que $\alpha \neq \beta$, et $f_{\alpha,\beta}$ l'endomorphisme de $\mathbb{R}^4$ dont la matrice par rapport à la base canonique $\mathcal{B}$ de $\mathbb{R}^4$ est :

$$A_{\alpha,\beta} = \begin{pmatrix} \alpha+1 & \alpha & -1 & 1-\alpha \\ \alpha-\beta & 2\alpha & \beta-\alpha & -\alpha \\ \alpha-\beta+1 & \alpha & \beta-1 & 1-\alpha \\ \alpha-\beta & 0 & \beta-\alpha & \alpha \end{pmatrix}.$$

a) Déterminer le polynôme caractéristique $P_{\alpha,\beta}(x)$ de $f_{\alpha,\beta}$. Déterminer, en discutant suivant les valeurs de α et β, les valeurs propres de $f_{\alpha,\beta}$ et les sous-espaces propres associés ; en déduire la forme réduite de Jordan $T_{\alpha,\beta}$ de $A_{\alpha,\beta}$.
b) On fixe $\alpha = 2$ et $\beta = 1$. Déterminer $T_{2,1}$; construire une base $\mathcal{C}$ de $\mathbb{R}^4$ dans laquelle $T_{2,1}$ représente $f_{2,1}$; calculer la matrice de passage P de la base $\mathcal{B}$ à la base $\mathcal{C}$, et son inverse P^{-1}.
c) *Application :* déterminer les fonctions réelles $x(t), y(t), z(t), s(t)$ de la variable réelle t, de classe C^1 sur $\mathbb{R}$, solutions du système différentiel : $\begin{cases} x'(t) = 3x(t) + 2y(t) - z(t) - s(t) \\ y'(t) = x(t) + 4y(t) - z(t) - 2s(t) \\ z'(t) = 2x(t) + 2y(t) - s(t) \\ s'(t) = x(t) - z(t) + 2s(t) \end{cases}$