Université Pierre et Marie Curie

Licence de Mathématiques

Année 2015–2016

2M270 - Algèbre linéaire 2, espaces affines

par Laurent KOELBLEN et Patrick POLO,

repris et revu par Vincent HUMILIERE

Remerciements:

à Laurent Koelblen et Patrick Polo pour m'avoir généreusement fourni les fichiers du désormais ancien LM270, le présent texte est à 95% le leur;

à Patrick Polo en particulier pour les multiples et éclairantes discussions autour du contenu de ce cours.

Vincent Humilière
Université P. & M. Curie – Paris 6
Institut de Mathématiques de Jussieu (UMR 7586 du CNRS)
URL: http//webusers.imj-prg.fr/~vincent.humiliere
E-mail : vincent.humiliere@imj-prg.fr

TABLE DES MATIÈRES

CHAPITRE 0

RAPPELS : ESPACES VECTORIELS ET APPLICATIONS LINÉAIRES

Ce chapitre 0 est constitué de **rappels**: espaces vectoriels, familles génératrices, familles libres, bases et dimension, applications linéaires et matrices, transposée d'une matrice, théorème du rang, formules de changement de base, matrices équivalentes, matrices semblables. Ces notions ont été introduites en L1 et ne seront pas traitées en cours, mais le lecteur pourra se reporter, si nécessaire, à ce chapitre 0 pour un rappel de ces notions ou résultats. On suppose également connue la théorie du pivot pour la résolution des systèmes linéaires $AX = Y$; ceci équivaut à faire des opérations sur les lignes de la matrice A et on verra dans le chapitre 1 comment faire des opérations sur les *colonnes* (ce qui est plus pratique pour la recherche de vecteurs propres).

On a indiqué par des symboles les définitions et résultats fondamentaux. Par ailleurs, des *compléments de cours*, pour les étudiants intéressés, sont donnés dans un appendice à la fin du chapitre; ces passages n'interviendront pas dans les évaluations.

0.1. Espaces vectoriels: définition et exemples

0.1.1. Trois exemples importants. — (1) Un exemple d'espace vectoriel sur $\mathbb{R}$ est l'espace de dimension 3

$$\mathbb{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbb{R}\}.$$

Dans ce cas, l'espace vectoriel nous est donné comme un ensemble de n-uplets (ici $n = 3$) de "coordonnées".

Deux autres exemples sont les suivants.

(2) Soient $a, b \in \mathbb{R}$; l'ensemble $\mathscr{S}(a, b)$ des suites $(u_n)_{n\in\mathbb{N}}$ de nombres réels vérifiant la relation de récurrence linéaire

$$(*) \qquad u_{n+2} = au_{n+1} + bu_n$$

est un $\mathbb{R}$-espace vectoriel, qui est de dimension 2, car toute suite vérifiant $(*)$ est déterminée par ses termes initiaux u_0 et u_1, qui peuvent être choisis arbitrairement.

(3) Soient $a, b \in \mathbb{R}$ et $t_0 \in \mathbb{R}$; l'ensemble $\mathscr{S}(a, b)$ des fonctions $f : \mathbb{R} \to \mathbb{R}$ de classe C^∞, vérifiant l'équation différentielle linéaire:

$$(E) \qquad \forall t \in \mathbb{R}, \qquad f''(t) = af'(t) + bf(t)$$

est un $\mathbb{R}$-espace vectoriel, qui est de dimension 2, car toute solution f de (E) est déterminée par les "conditions initiales" $f(t_0)$ et $f'(t_0)$, qui peuvent être choisies arbitrairement.

Dans ces deux cas, le choix de "coordonnées" sur l'espace $\mathscr{S}(a, b)$ n'est pas évident ... Une des forces de l'algèbre linéaire est qu'elle permet de décrire simplement tous les éléments de $\mathscr{S}(a, b)$, une fois qu'on a choisi une base appropriée de cet espace ...

Rappelons que la notion d'espace vectoriel sur un corps k est définie pour tout corps k, par exemple, $k = \mathbb{Q}$ ou $\mathbb{C}$, ou un corps fini $\mathbb{F}_q$ à q éléments (q étant une puissance d'un nombre premier p), ou le corps des fractions rationnelles $\mathbb{R}(X)$, etc.

---

[0] version du 13/5/2015

***Définition 0.1.2* (Espaces vectoriels).** — Soit k un corps. Un k-espace vectoriel V est un *groupe abélien* $(V,+)$ (c.-à-d., un ensemble muni d'une loi de groupe + commutative) muni d'une "opération" $(t,v) \mapsto t\cdot v$ de k sur V vérifiant les deux conditions suivantes:

(i) $1\cdot v = v$ et $t\cdot(t'\cdot v) = (tt')\cdot v$

(ii) $(t+t')\cdot v = t\cdot v + t'\cdot v, \qquad t\cdot(v+v') = t\cdot v + t\cdot v'.$

On peut mémoriser la condition (i) en disant que 1 agit par l'identité et que l'opération est "associative", et la condition (ii) en disant que l'action de k sur V est "compatible avec l'addition" (dans k et dans V).

***Remarque 0.1.3* (Vecteur nul).** — Étant un groupe abélien, V est muni d'un élément zéro, qu'on notera provisoirement 0_V ou $\overrightarrow{0}$. Par exemple, dans $V = \mathbb{R}^3$, 0_V est le vecteur nul

$$\overrightarrow{0} = (0,0,0).$$

Notons 0 l'élément zéro du corps k. Alors la condition (ii) entraîne, pour tout $v \in V$,

$$0\cdot v = (0+0)\cdot v = 0\cdot v + 0\cdot v, \qquad \text{d'où } 0\cdot v = 0_V = \overrightarrow{0}.$$

Par conséquent, le vecteur nul $0_V = \overrightarrow{0}$ sera noté simplement (par abus de notation) 0. Ainsi, on note $\{0\}$ l'espace vectoriel nul. Par exemple, dans $\mathbb{R}^2$ l'espace des solutions du système

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

est l'espace vectoriel nul $\{0\} = \{(0,0)\}$.

Terminologie 0.1.4. — On dira que k est (relativement à V) le "*corps des scalaires*", et que les éléments de k sont les scalaires (ceux de V étant les vecteurs).

Exemples 0.1.5. — 1) Pour tout $n \in \mathbb{N}^*$, $k^n = \{(x_1,\dots,x_n) \mid x_i \in k\}$ est un k-espace vectoriel.

2) L'ensemble $M_{m,n}(k)$ des matrices à m lignes et n colonnes à coefficients dans k est un k-espace vectoriel. Lorsque $m=n$, on le note simplement $M_n(k)$.

3) L'anneau de polynômes $k[X]$ est un k-espace vectoriel.

***Définition 0.1.6* (Sous-espaces vectoriels).** — Soit V un k-espace vectoriel. Un *sous-espace vectoriel* (pour abréger, on dira "sev") W de V est un sous-ensemble de V qui est un sous-groupe (en particulier, $0 \in W$) et qui est stable par l'opération de k. Ceci équivaut à dire que $W \neq \varnothing$ et que, pour tous $w, w' \in W$ et $t \in k$, on a $t\cdot w + w' \in W$.

Exemples. — (1) Le plan "horizontal"

$$P = \{(x,y,0) \mid x,y \in \mathbb{R}\},$$

donné par l'équation $z=0$, est un sous-espace vectoriel de $\mathbb{R}^3$.

(2) L'ensemble des matrices $(a_{ij})_{1\le i,j\le 3} \in M_3(k)$ qui sont triangulaires supérieures, i.e. telles que $a_{21} = a_{31} = a_{32} = 0$, est un sous-espace vectoriel de $M_3(k)$.

(3) L'ensemble $k_n[X]$ des polynômes en X de degré $\le n$ est un sous-espace vectoriel de $k[X]$.

Le lemme suivant permet de construire de nombreux espaces vectoriels de fonctions:

Lemme 0.1.7. — *Soient V un k-espace vectoriel et X un ensemble. Alors l'ensemble* $\mathrm{Hom}(X,V)$ *de toutes les applications $X \to V$ est muni d'une structure de k-espace vectoriel, définie comme suit: si $f,g \in \mathrm{Hom}(X,V)$ et $t \in k$, on définit l'application $t\cdot f + g : X \to V$ par:*

$$(t\cdot f+g)(x) = tf(x) + g(x),$$

pour tout $x \in X$.

Démonstration. Soient $f,g \in \mathrm{Hom}(X,V)$ et $t,t' \in k$; il faut voir que les conditions (i) et (ii) de 0.1.2 sont vérifiées. Ceci est facile: pour tout $x \in X$, on a $(1\cdot f)(x) = 1f(x) = f(x)$ donc $1\cdot f = f$, et

$$\big(t\cdot(t'\cdot f)\big)(x) = t\big((t'\cdot f)(x)\big) = t(t'f(x)) = (tt')f(x) = \big((tt')\cdot f\big)(x),$$

donc $t\cdot(t'\cdot f) = (tt')\cdot f$. On vérifie de même la condition (ii).

Exemples 0.1.7.1. — (1) L'ensemble $k^{\mathbb{N}}$ de toutes les suites $(u_n)_{n\in\mathbb{N}}$ d'éléments de k muni de l'addition et de l'opération de k définies "terme à terme", c.-à-d.,

$$(u_n)_{n\in\mathbb{N}} + (v_n)_{n\in\mathbb{N}} = (u_n + v_n)_{n\in\mathbb{N}}, \qquad t \cdot (u_n)_{n\in\mathbb{N}} = (t \cdot u_n)_{n\in\mathbb{N}},$$

est un k-espace vectoriel. En effet, ceci coïncide avec la structure définie plus haut si on regarde une suite $(u_n)_{n\in\mathbb{N}}$ d'éléments de k comme l'application $\mathbb{N} \to k$, $n \mapsto u_n$.

(2) Soit I un intervalle de $\mathbb{R}$. Alors l'ensemble $\mathrm{Fonc}(\mathrm{I}, \mathbb{R})$ de toutes les fonctions $I \to \mathbb{R}$ est un $\mathbb{R}$-espace vectoriel. Pour les applications à l'analyse, il est plus intéressant de constater que si $t \in \mathbb{R}$ et si $f, g : I \to \mathbb{R}$ sont continues, resp. de classe C^1, resp. C^2, ..., resp. C^∞, alors $t \cdot f + g$ est encore continue (i.e. de classe C^0), resp. C^1, ..., resp. C^∞. Par conséquent, pour tout $d \in \mathbb{N} \cup \{\infty\}$, l'ensemble $\mathscr{C}^d(I, \mathbb{R})$ des fonctions $f : I \to \mathbb{R}$ de classe C^d est un $\mathbb{R}$-espace vectoriel.

***Définition 0.1.8* (Applications linéaires et endomorphismes)**. — Soient k un corps, V, W deux k-espaces vectoriels. On dit qu'une application $\phi : V \to W$ est une *application linéaire* (ou "homomorphisme d'espaces vectoriels") si elle préserve l'addition et l'opération des scalaires, c.-à-d., si pour tout $v, v' \in V$ et $t \in k$ on a:

$$\phi(v + v') = \phi(v) + \phi(v'), \qquad \phi(t \cdot v) = t \cdot \phi(v).$$

Notons qu'on peut regrouper ces deux conditions en une seule condition:

$$\text{(AL)} \qquad \phi(t \cdot v + v') = t \cdot \phi(v) + \phi(v')$$

et bien sûr cette condition implique (et est impliquée par) la suivante:

$$\phi(t \cdot v + t' \cdot v') = t \cdot \phi(v) + t' \cdot \phi(v').$$

Si $W = V$, on dit alors que ϕ est un *endomorphisme* de V.

Exemples 0.1.9. — a) Soit $V = \mathbb{R}[X]$; l'application d qui à tout polynôme $P = a_n X^n + \cdots + a_1 X + a_0$ associe le polynôme dérivé $P' = n a_n X^{n-1} + \cdots + a_1$ est une application linéaire de V dans V.

b) Soit $V = \mathscr{C}^0([0, 1], \mathbb{R})$ le $\mathbb{R}$-espace vectoriel des fonctions continues $f : [0, 1] \to \mathbb{R}$. Alors l'application

$$V \to \mathbb{R}, \qquad f \mapsto \int_0^1 f(x)dx$$

est linéaire. Par contre, l'application $f \mapsto \int_0^1 f^2(x)dx$ ne l'est pas.

***Définition 0.1.10* (Isomorphismes)**. — Soient V, W deux k-espaces vectoriels, et $\phi : V \to W$ une application linéaire. Suivant des principes généraux, on dit que ϕ est un *isomorphisme* (d'espaces vectoriels) si elle est *bijective* et si l'application inverse $\psi = \phi^{-1}$ est linéaire.

En fait, la seconde condition est automatiquement vérifiée. En effet, soient $w, w' \in W$ et $t \in k$; comme ϕ est bijective il existe $v, v' \in V$ uniques tels que $\phi(v) = w$ et $\phi(v') = w'$. Alors

$$\phi(t \cdot v + v') = t \cdot \phi(v) + \phi(v') = t \cdot w + w',$$

donc appliquant ψ à cette égalité on obtient

$$\psi(t \cdot w + w') = t \cdot v + v' = t \cdot \psi(v) + \psi(v').$$

Donc: *toute application linéaire bijective est un isomorphisme.*

Exemple 0.1.11. — On rappelle que l'ensemble $k^{\mathbb{N}}$ de toutes les suites d'éléments de k est un k-espace vectoriel (cf. 0.1.7.1). Soient $a, b \in k$ et soit $\mathscr{S}(a, b)$ le sous-ensemble de $k^{\mathbb{N}}$ formé des suites $(u_n)_{n\in\mathbb{N}}$ vérifiant la relation de récurrence linéaire:

$$(*) \qquad u_{n+2} = a u_{n+1} + b u_n.$$

Alors $\mathscr{S}(a, b)$ est un sous-espace vectoriel de $k^{\mathbb{N}}$ (le vérifier!). De plus, l'application $\phi : \mathscr{S}(a, b) \to k^2$ qui à toute suite $(u_n)_{n\in\mathbb{N}}$ associe le couple (u_0, u_1), est linéaire (le vérifier!); elle est surjective (car on peut choisir arbitrairement u_0 et u_1), et injective (car les u_n sont déterminés à partir de u_0 et u_1 par la formule $(*)$). Donc ϕ est bijective, donc c'est un isomorphisme d'espaces vectoriels:

$$\phi : \mathscr{S}(a, b) \xrightarrow{\sim} k^2.$$

Exercices 0.1.12. — 1) Soit $k = \mathbb{R}$. Est-ce que l'ensemble des suites $(u_n)_{n\in\mathbb{N}}$ vérifiant la relation de récurrence
$$u_{n+2} = u_{n+1} + u_n^2$$
est un sous-espace vectoriel de $\mathbb{R}^{\mathbb{N}}$?

2) Soient $k = \mathbb{C}$ et $\mathscr{S}(-1,-1)$ l'espace des suites de nombres complexes $(u_n)_{n\in\mathbb{N}}$ vérifiant la relation de récurrence linéaire
$$u_{n+2} = -u_{n+1} - u_n.$$
Soit $(w_n)_{n\in\mathbb{N}}$ l'élément de $\mathscr{S}(-1,-1)$ défini par $w_0 = 2$ et $w_1 = 1$. Pouvez-vous calculer w_{2015} et w_{2016}?

0.2. Familles génératrices, familles libres, bases et dimension

Définitions 0.2.1 **(Sous-espace engendré. Familles génératrices).** — Soit V un k-espace vectoriel.

(1) Soit S un sous-ensemble (fini ou infini) de V. On note $\boxed{\mathrm{Vect}(S)}$ l'ensemble de toutes les combinaisons linéaires *finies*
$$\sigma = t_1v_1 + \cdots + t_rv_r, \qquad \text{pour } r \in \mathbb{N}^* \text{ (variable)},\ v_i \in S,\ t_i \in k,$$
c'est un sous-espace vectoriel de V, car si $\sigma' = t'_1v'_1 + \cdots + t'_pv'_p$ est une autre combinaison linéaire de ce type et si $\lambda \in k$, alors
$$\lambda\sigma + \sigma' = \lambda t_1v_1 + \cdots + \lambda t_rv_r + t'_1v'_1 + \cdots + t'_pv'_p$$
est encore une combinaison linéaire du même type. De plus, si E est un sous-espace vectoriel de V contenant S, alors il contient toute combinaison linéaire σ comme ci-dessus, i.e. il contient $\mathrm{Vect}(S)$. Donc : $\boxed{\mathrm{Vect}(S) \text{ est le plus petit sous-espace vectoriel de } V \text{ contenant } S.}$ On l'appelle le sous-espace vectoriel *engendré* par S.

Dans la suite, on utilisera principalement ceci dans le cas où S est une famille *finie* de vecteurs $v_1, \dots, v_n$; dans ce cas, on a simplement:
$$\mathrm{Vect}(v_1, \dots, v_n) = \{t_1v_1 + \cdots + t_nv_n \mid t_1, \dots, t_n \in k\}.$$

(2) On dit que la famille $\mathscr{F} = (v_1, \dots, v_n)$ est une *famille génératrice* de V si $\mathrm{Vect}(\mathscr{F}) = V$, c.-à-d., si tout élément de V s'écrit comme combinaison linéaire de $v_1, \dots, v_n$. Dans ce cas, on dit aussi que les vecteurs $v_1, \dots, v_n$ *engendrent* V.

Il résulte de la définition que: *toute famille contenant une famille génératrice est génératrice.*

Exemple 0.2.2. — Pour tout $n \in \mathbb{N}^\times$, l'espace k^n est engendré par les vecteurs
$$e_1 = (1, 0, \dots, 0), \quad e_2 = (0, 1, 0, \dots, 0), \quad \cdots, \quad e_n = (0, \dots, 0, 1).$$

Définition 0.2.3 **(Familles libres ou liées).** — Soient V un k-espace vectoriel et $\mathscr{F} = (v_1, \dots, v_n)$ une famille d'éléments de V. On dit que $v_1, \dots, v_n$ sont *linéairement indépendants*, et que $\mathscr{F}$ est une *famille libre*, s'il n'existe pas de relation linéaire non triviale entre les v_i, c.-à-d., si la condition suivante est vérifiée:

(FL) pour *tous* $t_1, \dots, t_n \in k$, si $t_1v_1 + \cdots + t_nv_n = 0$, *alors* $t_1 = 0 = \cdots = t_n$.

Il résulte de la définition que: *toute sous-famille d'une famille libre est libre.*

Au contraire, on dit que $v_1, \dots, v_n$ sont *linéairement dépendants*, et que $\mathscr{F}$ est une *famille liée*, s'il existe une relation linéaire non triviale entre les v_i, c.-à-d., s'il existe des scalaires non tous nuls $t_1, \dots, t_n$, tels que $t_1v_1 + \cdots + t_nv_n = 0$. Dans ce cas, si par exemple $t_i \neq 0$, on peut exprimer v_i en fonction des v_j, pour $j \neq i$:
$$v_i = -\sum_{j\neq i} \frac{t_j}{t_i} v_j.$$
Il résulte de la définition que: *toute famille contenant une famille liée est liée.*

Exemple 0.2.4. — Dans k^n, la famille $(e_1, \dots, e_n)$ (cf. 0.2.2) est libre. En effet, *pour tous* $t_1, \dots, t_n \in k$ on a
$$t_1e_1 + \cdots + t_ne_n = (t_1, \dots, t_n),$$
donc si la somme de gauche est nulle, alors $t_1 = 0 = \cdots = t_n$.

Définition 0.2.5 **(Bases).** — Soient V un k-espace vectoriel. On dit qu'une famille $\mathscr{B} = (v_1, \dots, v_n)$ est une *base* de V si tout élément v de V *s'écrit de façon unique comme combinaison linéaire des* v_i, c.-à-d., si pour tout $v \in V$, *il existe un unique* n*-uplet* $(t_1, \dots, t_n) \in k^n$ *tel que* $v = t_1v_1 + \cdots + t_nv_n$. Ceci équivaut à dire que la famille $\mathscr{B}$ est à la fois génératrice et libre.

Exemples 0.2.6. — 1) Lorsque $V = k^n$, la famille $(e_1, \ldots, e_n)$ (cf. 0.2.2) est une base, appelée la *base canonique de* k^n.

2) Soit $M_{m,n}(k)$ le k-espace vectoriel des matrices à m lignes et n colonnes à coefficients dans k. On note E_{ij} la "*matrice élémentaire*" dont tous les coefficients sont nuls, sauf celui d'indice (i, j) (c.-à-d., celui situé sur la ligne i et la colonne j), qui vaut 1. Alors, toute matrice $A \in M_{m,n}(k)$ *s'écrit de façon unique* comme combinaison linéaire des E_{ij}:

$$A = \sum_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} a_{ij}\, E_{ij},$$

où a_{ij} est le coefficient d'indice (i, j) de A. Donc la famille $(E_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$ est une *base* de $M_{m,n}(k)$.

3) La famille $(1, X, \ldots, X^d)$ est une base de l'espace vectoriel $k_d[X]$ des polynômes de degré $\leq d$. En effet, tout polynôme $P \in k_d[X]$ s'écrit de façon unique

$$P = a_0 + \cdots + a_d X^d, \qquad \text{avec} \quad a_i \in k.$$

Définition 0.2.7. — Soit V un k-espace vectoriel. Disons provisoirement que V est *finiment engendré* (ou "de type fini", cf. le cours de L1) s'il est engendré par un nombre fini de vecteurs $v_1, \ldots, v_p$.

Remarque : il existe des k-espaces vectoriels qui ne sont pas finiment engendrés, par exemple, l'espace vectoriel $k[X]$, mais dans ce cours on s'intéressera à ceux qui le sont.

Rappelons le résultat suivant, déjà vu en première année :

***Théorème 0.2.8* (Dimension d'un espace vectoriel)**. — *Soit V un k-espace vectoriel finiment engendré.*

(i) *Il existe des bases de V, et toutes ont même cardinal n; cet entier s'appelle la dimension de V sur k et se note* $\dim_k V$ *ou simplement* $\dim V$.

(ii) *De toute famille génératrice $\mathscr{F}$ on peut extraire une base, en particulier $\mathscr{F}$ est de cardinal $\geq n$; de plus si* $\mathrm{card}(\mathscr{F}) = n$ *alors $\mathscr{F}$ est une base de V.*

(iii) *Toute famille libre est de cardinal $\leq n$, et toute famille libre de cardinal n est une base de V.*

(iv) "Théorème de la base incomplète" *: Toute famille libre peut être complétée en une base de V.*

(v) *Tout sous-espace W de V est de dimension finie $\leq \dim_k V$; de plus si $\dim_k W = \dim_k V$, alors $W = V$. En d'autres termes, tout sous-espace vectoriel distinct de V est de dimension $< \dim_k V$.*

Démonstration. Ceci a été vu en L1. Pour être complet, on redonne la démonstration dans un appendice à la fin de ce chapitre, où on introduira aussi les notions de familles génératrices ou libres dans un espace vectoriel arbitraire (i.e. qui n'est pas nécessairement finiment engendré).

Terminologie 0.2.9. — En raison du théorème précédent, on dira désormais "k-espace vectoriel *de dimension finie*" au lieu de "k-espace vectoriel finiment engendré", et si $n = \dim_k V$, on dira que V est de dimension n.

D'après les exemples de 0.2.6, k^n est de dimension n, $M_{m,n}(k)$ de dimension mn, et $k_d[X]$ de dimension $d + 1$.

***Définition 0.2.10* (Coordonnées relativement à une base)**. — Soit V un k-espace vectoriel de dimension n et soit $\mathscr{B} = (v_1, \ldots, v_n)$ une base de V. Alors tout $v \in V$ s'écrit de façon unique

$$v = x_1 v_1 + \cdots + x_n v_n;$$

on dit que $(x_1, \ldots, x_n)$ sont les *coordonnées de v par rapport à la base* $\mathscr{B} = (v_1, \ldots, v_n)$. Donc la donnée de $\mathscr{B}$ fournit un isomorphisme de k-espaces vectoriels

$$\phi_{\mathscr{B}} : \quad k^n \xrightarrow{\sim} V, \qquad (x_1, \ldots, x_n) \mapsto x_1 v_1 + \cdots + x_n v_n.$$

Remarque 0.2.11. — Remarquons qu'une base de V est un n-uplet *ordonné*; par exemple, si $\mathscr{B} = (v_1, v_2)$ est une base de V, alors $\mathscr{C} = (v_2, v_1)$ est une base de V distincte de $\mathscr{B}$: l'image de $(1, 2) \in \mathbb{R}^2$ par $\phi_{\mathscr{B}}$ est le vecteur $v_1 + 2v_2$, tandis que son image par $\phi_{\mathscr{C}}$ est le vecteur $v_2 + 2v_1 \neq v_1 + 2v_2$.

Proposition 0.2.12. — *Soit $f : V \to W$ une application linéaire.*

a) *Si f est injective et si $\mathscr{F} = (v_1, \ldots, v_n)$ est une famille libre de V, alors $f(\mathscr{F})$ est libre.*

b) *Si f est surjective et si $\mathscr{F} = (v_1, \ldots, v_n)$ est une famille génératrice de V, alors $f(\mathscr{F})$ engendre W.*

c) *Si f est bijective et si $\mathscr{B} = (v_1, \ldots, v_n)$ est une base de V, alors $f(\mathscr{B})$ est une base de W, d'où* $\dim W = n = \dim V$.

Démonstration. a) Supposons f injective et soit $\mathscr{F} = (v_1, \dots, v_n)$ une famille libre de V. Supposons qu'il existe une relation linéaire dans W:

$$t_1 f(v_1) + \cdots + t_n f(v_n) = 0,$$

avec $t_i \in k$. Alors $0 = f(t_1v_1 + \cdots + t_nv_n)$ d'où, puisque f est injective, $t_1v_1 + \cdots + t_nv_n = 0$, donc comme $\mathscr{F}$ est libre, $t_i = 0$ pour $i = 1, \dots, n$. Ceci montre que $f(\mathscr{F})$ est libre.

b) Supposons f surjective et soit $\mathscr{F} = (v_1, \dots, v_n)$ une famille génératrice de V. Soit $w \in W$; comme f est surjective, il existe $v \in V$ tel que $f(v) = w$. Comme $\mathscr{F}$ engendre V, il existe $t_1, \dots, t_n \in k$ tels que $v = t_1v_1 + \cdots + t_nv_n$, d'où

$$w = t_1 f(v_1) + \cdots + t_n f(v_n).$$

Ceci montre que $f(\mathscr{F})$ engendre W.

c) Supposons f bijective et soit $\mathscr{B} = (v_1, \dots, v_n)$ une base de V. D'après a) et b), $f(\mathscr{B})$ est une famille libre et génératrice de W, donc une base de W; alors $\dim W = n = \dim V$.

Corollaire 0.2.13. — (i) *Tout k-espace vectoriel V de dimension finie est isomorphe (de façon non canonique) à k^n, pour un unique n, égal à $\dim_k V$.*

(ii) *Deux k-espaces vectoriels de dimension finie V et W sont isomorphes si et seulement si ils ont la même dimension.*

Démonstration. (i) Si V est de dimension n alors, par le choix d'une base, il est isomorphe à k^n. (Cet isomorphisme est "non canonique" car il dépend du choix de la base.) Réciproquement, si $V \simeq k^m$, alors d'après la proposition précédente, on a $\dim V = \dim k^m = m$, donc $m = n$. En particulier, $k^m \not\simeq k^n$ si $m \neq n$.

(ii) Si $V \simeq W$, alors $\dim_k V = \dim_k W$, d'après la proposition précédente. Réciproquement, si $\dim_k V = \dim_k W = n$, alors V et W sont tous deux isomorphes à k^n.

Exemples 0.2.14. — 1) La droite d'équation $x_1 + x_2 = 0$ dans $\mathbb{R}^2$ admet pour base le vecteur $e_1 - e_2$, mais on peut tout aussi bien choisir le vecteur $e_2 - e_1$.

2) Le plan d'équation $x_1 + x_2 + x_3 = 0$ dans $\mathbb{R}^3$ admet pour base $(e_1 - e_2, e_2 - e_3)$, mais on peut aussi choisir $(e_1 - e_3, e_1 - e_3)$, ou $(e_1 - e_2, e_1 + e_2 - 2e_3)$, ou $(e_1 - (e_1 + e_2 + e_3)/3, e_1 + e_2 - 2(e_1 + e_2 + e_3)/3)$, etc.

Exemple 0.2.15. — Reprenons l'espace $\mathscr{S}(a, b)$ des suites $(u_n)_{n\in\mathbb{N}}$ d'éléments de k vérifiant la relation de récurrence linéaire $u_{n+2} = au_{n+1} + bu_n$. On a vu (cf. 0.1.11) qu'il est isomorphe à k^2, par l'application $(u_n)_{n\in\mathbb{N}} \mapsto (u_0, u_1)$, donc est de dimension 2. Supposons que le polynôme $P = X^2 - aX - b$ ait deux racines distinctes $\lambda \neq \mu$ dans k. Considérons les éléments $\mathbf{u} = (u_n)_{n\in\mathbb{N}}$ et $\mathbf{v} = (v_n)_{n\in\mathbb{N}}$ de $\mathscr{S}(a, b)$ définis par

$$u_0 = 1 = v_0, \qquad u_1 = \lambda, \quad v_1 = \mu.$$

Alors la famille $(\mathbf{u}, \mathbf{v})$ est libre (car si $s\mathbf{u} + t\mathbf{v} = 0$, on obtient $s + t = 0 = s\lambda + t\mu$, donc $t = -s$ et $s(\lambda - \mu) = 0$, d'où $s = 0$), donc est une base de $\mathscr{S}(a, b)$. Par conséquent, tout élément $\mathbf{w} = (w_n)_{n\in\mathbb{N}}$ de $\mathscr{S}(a, b)$ s'écrit de façon unique

$$\mathbf{w} = s\mathbf{u} + t\mathbf{v},$$

et s, t sont déterminés par les conditions $s + t = w_0$ et $s\lambda + t\mu = w_1$. Ceci permet-il de calculer w_{2015} et w_{2016} dans l'exercice 0.1.12?

0.3. Noyau, image, et théorème du rang

***Définition 0.3.1* (Noyau, image et rang)**. — Soit $f : V \to W$ une application linéaire. D'une part, on définit son noyau

$$\mathrm{Ker}(f) = \{v \in V \mid f(v) = 0\},$$

c'est un sous-espace vectoriel de V. Noter que f est *injective* si et seulement si $\mathrm{Ker}(f) = \{0\}$: en effet, on a $f(v) = f(v') \iff f(v - v') = 0$.

D'autre part, on définit son image

$$\mathrm{Im}(f) = f(V) = \{f(v) \mid v \in V\} \subseteq W,$$

c'est un sous-espace vectoriel de V. Alors f est *surjective* si et seulement si $\mathrm{Im}(f) = W$.

Lorsque $\mathrm{Im}(f)$ est *de dimension finie* r (ce qui est le cas si W ou V est de dimension finie, cf. ci-dessous), l'entier $r = \dim \mathrm{Im}(f)$ est appelé le **rang** de f et est noté $\mathrm{rg}(f)$ ou $\mathrm{rang}(f)$.

***Théorème 0.3.2* (Théorème du rang).** — *Soit $f : V \to W$ une application linéaire, avec V de dimension finie n. Alors*

$$n = \dim V = \dim \mathrm{Ker}(f) + \mathrm{rg}(f).$$

(En particulier, f est surjective si et seulement si W est de dimension $\mathrm{rg}(f) = n - \dim \mathrm{Ker}(f)$.)

Démonstration. Comme V est de dimension finie n, alors $\mathrm{Ker}(f)$ est de dimension finie $d \leq n$; soit $\mathscr{K} = (e_1, \dots, e_d)$ une base de $\mathrm{Ker}(f)$.

1ère méthode (la plus courte). Complétons $\mathscr{K}$ en une base $\mathscr{B} = (e_1, \dots, e_d, e_{d+1}, \dots, e_n)$ de V. Alors $\mathrm{Im}(f) = f(V)$ est engendré par $f(\mathscr{B})$ (cf. Prop. 0.2.12), donc par $f(e_{d+1}), \dots, f(e_n)$ puisque $f(e_i) = 0$ pour $i \leq d$. Montrons que ces vecteurs sont linéairement indépendants: supposons qu'on ait une relation de dépendance linéaire

$$0 = t_1 f(e_{d+1}) + \cdots + t_{n-d} f(e_n) = f(t_1 e_{d+1} + \cdots + t_{n-d} e_n)$$

alors le vecteur $t_1 e_{d+1} + \cdots + t_{n-d} e_n$ appartient à $\mathrm{Ker}(f)$, donc est combinaison linéaire de $e_1, \dots, e_d$, d'où une égalité

$$t_1 e_{d+1} + \cdots + t_{n-d} e_n - s_1 e_1 - \cdots - s_d e_d = 0.$$

Comme $(e_1, \dots, e_n)$ est une base de V, ceci implique $t_i = 0 = s_j$ pour tous i, j. Ceci montre que les vecteurs $f(e_{d+1}), \dots, f(e_n)$ sont linéairement indépendants, donc forment une base de $f(V)$, d'où $\dim f(V) = n - d$ et donc $\mathrm{rg}(f) = \dim f(V) = n - \dim \mathrm{Ker}(f)$.

2ème méthode. Soit toujours $\mathscr{K} = (e_1, \dots, e_d)$ une base de $\mathrm{Ker}(f)$, donnons une autre démonstration, qui sera utile plus loin (cf. 0.5.10). (1) Comme $\mathrm{Im}(f) = f(V)$ est engendré par n éléments (les images d'une base de V), $\mathrm{Im}(f)$ est de dimension finie $r \leq n$. Soit $(w_1, \dots, w_r)$ une base de $\mathrm{Im}(f)$ et pour $i = 1, \dots, r$, soit v_i un élément de V tel que $f(v_i) = w_i$. Alors la famille

$$\mathscr{B} = (e_1, \dots, e_d, v_1, \dots, v_r)$$

est une *base* de V. En effet, elle est *génératrice*: soit $v \in V$ arbitraire, son image $f(v)$ s'écrit $f(v) = t_1 w_1 + \cdots + t_r w_r$, d'où $f(v - t_1 v_1 - \cdots - t_r v_r) = 0$, donc $v - t_1 v_1 - \cdots - t_r v_r$ appartient à $\mathrm{Ker}(f)$, donc s'écrit $s_1 e_1 + \cdots + s_p e_p$, d'où

$$v = t_1 v_1 + \cdots + t_r v_r + s_1 e_1 + \cdots + s_p e_p.$$

Ceci montre que $\mathscr{B}$ est génératrice. Elle est aussi *libre*: si

$$0 = t_1 v_1 + \cdots + t_r v_r + s_1 e_1 + \cdots + s_p e_p$$

alors $0 = f(0) = t_1 w_1 + \cdots + t_r w_r$, donc chaque t_i est nul (puisque $(w_1, \dots, w_r)$ est libre), d'où $0 = s_1 e_1 + \cdots + s_p e_p$, donc chaque s_j est nul (puisque $(e_1, \dots, e_p)$ est libre). Ceci montre que $\mathscr{B}$ est aussi libre, donc est une base de V. Donc $\dim V = d + r = \dim \mathrm{Ker}(f) + \mathrm{rg}(f)$.

Proposition 0.3.3. — *Soit V un k-espace vectoriel de dimension n et soit $u \in \mathrm{End}_k(V)$ ou, plus généralement, soit $u : V \to W$ une application linéaire, où $\dim W = n = \dim V$. Alors les conditions suivantes sont équivalentes:*

(i) *u est bijectif;*
(ii) *u est injectif;*
(iii) *u est surjectif.*

En effet, il est clair que (i) implique (ii) et (iii). Réciproquement, si u est injectif, i.e. $\mathrm{Ker}(u) = \{0\}$ (resp. surjectif, i.e. $\mathrm{Im}(u) = W$), il résulte du théorème du rang que u est aussi surjectif (resp. injectif), donc bijectif.

0.4. Applications linéaires et matrices

***Définition 0.4.1* (Espaces d'applications linéaires).** — Soient V, W deux k-espaces vectoriels. On note $\mathrm{Hom}_k(V, W)$ ou $\mathscr{L}(V, W)$ l'ensemble des applications linéaires de V dans W. C'est un k-espace vectoriel: plus précisément, c'est un sous-espace vectoriel du k-espace vectoriel $\mathrm{Fonc}(V, W)$ de toutes les fonctions $V \to W$ (cf. 0.1.7). Rappelons que si $t \in k$ et $\phi, \psi \in \mathrm{Fonc}(V, W)$ alors les fonctions $\phi + \psi$ et $t \cdot \phi$ sont définies comme suit: pour tout $v \in V$,

$$(*) \qquad (\phi + \psi)(v) = \phi(v) + \psi(v) \qquad \text{et} \qquad (t \cdot \phi)(v) = t \cdot \phi(v).$$

(1) Cette démonstration a l'avantage suivant. Elle montre que si une base $(w_1, \dots, w_r)$ de $\mathrm{Im}(f)$ est donnée à l'avance et si $v_1, \dots, v_r \in V$ vérifient $f(v_i) = w_i$ pour $i = 1, \dots, r$, alors $(e_1, \dots, e_d, v_1, \dots, v_r)$ est une base de V.

Il s'agit de voir que si $\phi, \psi : V \to W$ sont des applications *linéaires*, alors il en est de même de $\phi + \psi$ et de $t \cdot \phi$. Ceci se vérifie facilement: si $v, v' \in V$ et $s \in k$, alors

$$\begin{aligned}(\phi+\psi)(s\cdot v+v') &= \phi(s\cdot v+v')+\psi(s\cdot v+v') && \text{(par définition)}\\ &= s\cdot\phi(v)+\phi(v')+s\cdot\psi(v)+\psi(v') && \text{(car } \phi,\psi \text{ linéaires)}\\ &= s\cdot(\phi+\psi)(v)+(\phi+\psi)(v') && \text{(par définition)}\end{aligned}$$

et de même

$$(t\cdot\phi)(s\cdot v+v') = t\cdot\phi(s\cdot v+v') = ts\cdot\phi(v)+t\cdot\phi(v') = s\cdot(t\cdot\phi)(v)+(t\cdot\phi)(v').$$

Donc $\mathrm{Hom}_k(V, W) = \mathscr{L}(V, W)$ est bien un k-espace vectoriel; on dit que c'est l'espace des applications linéaires de V dans W.

Supposons V *de dimension finie* n et soit $(e_1, \ldots, e_n)$ une *base* de V (par exemple, $V = k^n$ et $(e_1, \ldots, e_n)$ la base canonique). Soit $\phi \in \mathrm{Hom}_k(V, W)$, posons $w_i = \phi(e_i)$ pour $i = 1, \ldots, n$; alors tout $v \in V$ s'écrit de façon unique $v = x_1e_1 + \cdots + x_ne_n$ et l'on a

$$(*) \qquad \phi(v) = x_1w_1 + \cdots + x_nw_n$$

donc ϕ est déterminée par la donnée des n vecteurs $w_1, \ldots, w_n \in W$. Réciproquement, pour tout n-uplet $(w_1, \ldots, w_n) \in W^n$, l'application $\phi : V \to W$ définie par la formule $(*)$ est linéaire. On a donc obtenu la

Proposition 0.4.2. — *Si $\mathscr{B} = (e_1, \ldots, e_n)$ est une base de V, se donner une application linéaire $\phi : V \to W$* "est la même chose" *que se donner un n-uplet $(w_1, \ldots, w_n) \in W$.*

Supposons de plus que W soit *de dimension finie* m et soit $\mathscr{C} = (f_1, \ldots, f_m)$ une base de W. Alors, chaque $w_j = \phi(e_j)$ s'écrit de façon unique

$$w_j = a_{1j}f_1 + \cdots + a_{mj}f_m,$$

ce qu'on représente par le vecteur *colonne* (d'où le choix de l'indice j pour paramétrer ces colonnes):

$$w_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

et donc ϕ est déterminée par la matrice suivante:

$$\mathrm{Mat}_{\mathscr{C},\mathscr{B}}(\phi) = \mathrm{Mat}_{(f_i),(e_j)}(\phi) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{n2} & \cdots & a_{mn} \end{pmatrix}$$

qui exprime les vecteurs $\phi(e_j)$ (les colonnes) en fonction de $f_1, \ldots, f_m$. Noter que la dimension n de l'espace de *départ* V est le nombre de *colonnes*, et la dimension m de l'espace d'*arrivée* W est le nombre de *lignes*.

Réciproquement, pour toute matrice A comme ci-dessus, ses colonnes définissent de façon unique n vecteurs $w_1, \ldots, w_n$ de W, à savoir

$$w_j = a_{1j}f_1 + \cdots + a_{mj}f_m,$$

et ce n-uplet $(w_1, \ldots, w_n) \in W^n$ définit une application linéaire $\phi : V \to W$ dont la matrice associée est A. On obtient donc une *bijection*:

$$\mathrm{Hom}_k(V, W) \longleftrightarrow M_{m,n}(k).$$

De plus, on voit facilement que si A (resp. B) est la matrice associée à ϕ (resp. ψ), alors $tA + B$ est la matrice associée à $t\phi + \psi$, donc la bijection ci-dessus est un *isomorphisme d'espaces vectoriels.*

***Théorème 0.4.3* (Applications linéaires et matrices)**. — *Soient $\mathscr{B} = (e_1, \ldots, e_n)$ une base de V, et $\mathscr{C} = (f_1, \ldots, f_m)$ une base de W. Si ϕ est une application linéaire $V \to W$, on note* $\mathrm{Mat}_{\mathscr{C},\mathscr{B}}(\phi)$ *sa matrice dans les bases: $\mathscr{B}$ "au départ"et $\mathscr{C}$ "à l'arrivée".* [2]

(1) *L'application $\phi \mapsto \mathrm{Mat}_{\mathscr{C},\mathscr{B}}(\phi)$ est un isomorphisme d'espaces vectoriels:*

$$\mathrm{Hom}_k(V, W) \xrightarrow{\sim} M_{m,n}(k).$$

Donc le "slogan" à retenir est: "Des bases de V et W étant choisies, une application linéaire $V \to W$ est la même chose qu'une matrice à m lignes et n colonnes".

---

(2) L'ordre est choisi afin d'avoir, pour la composition des appplications, la formule qui figure dans le point (2) du théorème.

(2) *Cet isomorphisme transforme la composition des applications linéaires en le produit des matrices: si* U *est un* k*-espace vectoriel de base* $\mathscr{A} = (d_1, \dots, d_p)$ *et si* $\theta \in \mathrm{Hom}_k(U, V)$ *et* $\phi \in \mathrm{Hom}_k(V, W)$*, on peut former la composée* $\phi \circ \theta$ *:*

$$U \xrightarrow{\theta} V \xrightarrow{\phi} W,$$

et l'on a:

$$\boxed{\mathrm{Mat}_{\mathscr{C},\mathscr{A}}(\phi \circ \theta) = \mathrm{Mat}_{\mathscr{C},\mathscr{B}}(\phi) \cdot \mathrm{Mat}_{\mathscr{B},\mathscr{A}}(\theta).}$$

Démonstration. — On a déjà vu l'assertion (1), montrons l'assertion (2). Notons

$$A = \mathrm{Mat}_{\mathscr{C},\mathscr{B}}(\phi) = (a_{ij})_{\substack{j=1,\dots,n \\ i=1,\dots,m}} \qquad B = \mathrm{Mat}_{\mathscr{B},\mathscr{A}}(\theta) = (b_{j\ell})_{\substack{\ell=1,\dots,p \\ j=1,\dots,n}}$$

alors, pour tout $\ell = 1, \dots, p$, on a :

$$(\phi \circ \theta)(d_\ell) = \phi\big(\sum_{j=1}^{n} b_{j\ell}\, e_j\big) = \sum_{j=1}^{n} b_{j\ell}\, \phi(e_j) = \sum_{j=1}^{n}\sum_{i=1}^{m} b_{j\ell}\, a_{ij} f_i = \sum_{i=1}^{m}\big(\sum_{j=1}^{n} a_{ij} b_{j\ell}\big) f_i.$$

Donc le coefficient d'indice (i, ℓ) de $M = \mathrm{Mat}_{(f_i),(d_\ell)}(\phi \circ \theta)$ est $\sum_{j=1}^{n} a_{ij} b_{j\ell}$; ceci montre que $M = AB$. □

Remarque 0.4.4. — Soit $A = (a_{ij}) \in M_{m,n}(k)$ et soient $(e_1, \dots, e_n)$ et $(f_1, \dots, f_m)$ les bases canoniques de k^n et k^m. Alors par l'isomorphisme précédent, A correspond à l'application linéaire $u : k^n \to k^m$ telle que, pour tout $j = 1, \dots, n$,

$$u(e_j) = \sum_{i=1}^{m} a_{ij}\, f_i.$$

On identifiera, chaque fois que ce sera utile, A à l'application linéaire $u : k^n \to k^m$ ainsi définie (i.e. la i-ième colonne de A est l'image du i-ième vecteur de la base canonique de k^n.

Corollaire 0.4.5. — *Soient* $A \in M_{m,n}(k)$*,* $B \in M_{n,p}(k)$ *et* $u : k^n \to k^m$*,* $v : k^p \to k^n$ *les applications linéaires associées. Alors* AB *est la matrice de* $u \circ v : k^p \to k^m$*.*

Remarque 0.4.6. — Soient $B \in M_{m,n}(k)$ et $A \in M_{n,p}(k)$. Si l'on note $A_1, \dots, A_p \in k^n$ les colonnes de A, alors les colonnes de BA sont les vecteurs $BA_1, \dots, BA_p \in k^m$. [3]

En effet, ceci se voit directement sur la formule du produit matriciel: $(BA)_{ij} = \sum_{\ell=1}^{n} b_{i\ell}\, a_{\ell j}$. Ou bien, on peut raisonner comme suit: soit $(e_1, \dots, e_p)$ la base canonique de k^p, alors A correspond à l'application linéaire qui envoie chaque e_j sur le vecteur $Ae_j = A_j \in k^n$, et BA correspond à l'application linéaire qui envoie chaque e_j sur le vecteur $B(Ae_j) = BA_j \in k^n$.

Remarque 0.4.7. — On ne peut effectuer le produit AB de deux matrices $A \in M_{m,n}(k)$ et $B \in M_{q,p}(k)$ que si $n = q$, c.-à-d., si le nombre de colonnes de A égale le nombre de lignes de B.

En raison de son importance, répétons le théorème 0.4.3 et le corollaire 0.4.5 dans le cas particulier où *l'espace de départ est le même que celui d'arrivée*, c.-à-d., le cas où l'on considère des *endomorphismes* d'un espace V de dimension finie n, ou des matrices *carrées de taille* n.

***Proposition 0.4.8* (Endomorphismes de $V \simeq k^n$).** — *Le* k*-espace vectoriel* $M_n(k)$ *est un anneau (non commutatif si* $n \geq 2$*), c.-à-d., la multiplication des matrices carrées est associative:* $A(BC) = (AB)C$*, distributive à gauche et à droite par rapport à l'addition:* $(A + B)C = AC + BC$ *et* $A(B + C) = AB + AC$*, et la matrice identité* I_n *est élément neutre. De plus,* $M_n(k)$ *est une* k-algèbre*, c.-à-d., pour* $A, B \in M_n(k)$ *et* $t \in k$*, on a* $t \cdot (AB) = (t \cdot A)B = A(t \cdot B)$*, où* $\cdot$ *désigne la loi externe.*

De même, si V *est un* k*-espace vectoriel de dimension* n*, l'espace des endomorphismes* $\mathrm{End}_k(V)$ *est une* k*-algèbre (la multiplication étant la composition des endomorphismes). De plus, si l'on choisit une base* $\mathscr{B} = (e_1, \dots, e_n)$ *de* V*, l'application*

$$\mathrm{End}_k(V) \to M_n(k), \qquad u \mapsto \mathrm{Mat}_{\mathscr{B}}(u)$$

est un isomorphisme d'anneaux et de k*-espaces vectoriels, c.-à-d., un isomorphisme de* k*-algèbres.*

---

[3] Cette remarque sera utile pour la construction du déterminant, voir Chap. 3.

***Définition 0.4.9* (Noyau, image et rang d'une matrice).** — Soit $A \in M_{m,n}(k)$, on définit son noyau $\mathrm{Ker}(A)$, son image $\mathrm{Im}(A)$, et son rang, noté $\mathrm{rang}(A)$ ou $\mathrm{rg}(A)$, comme le noyau, l'image et le rang de l'application linéaire $u : k^n \to k^m$ associée. On a $\mathrm{rang}(u) \leq n$ (d'après le théorème du rang), et $\mathrm{rang}(u) \leq m$ (puisque $\mathrm{Im}(u)$ est un sous-espace de k^m), donc $\mathrm{rang}(u) \leq \mathrm{Min}(m, n)$.

Or, l'image de u est le sous-espace de k^m engendré par les vecteurs colonnes $C_1, \ldots, C_n$ de A, donc par définition $\mathrm{rang}(A)$ est le nombre maximum de *colonnes* de A linéairement indépendantes, et l'on a $\mathrm{rang}(A) \leq \mathrm{Min}(m, n)$.

On verra plus bas que $\mathrm{rang}(A)$ est aussi le nombre maximum de *lignes* linéairement indépendantes, et l'on donnera des moyens algorithmiques pour calculer $\mathrm{rang}(A)$.

***Définition 0.4.10* (Transposée d'une matrice).** — Soit dans $M_{m,n}(k)$:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Sa *transposée*, notée ${}^tA$, est la matrice de $M_{n,m}(k)$ suivante:

$${}^tA = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

c.-à-d., la j-ème colonne de A est devenue la j-ème ligne de ${}^tA$, c.-à-d., $({}^tA)_{ji} = a_{ij}$. On a évidemment

$$(\star) \qquad \boxed{{}^t({}^tA) = A.}$$

Proposition 0.4.11. — *L'application $A \mapsto {}^tA$ est linéaire: si $A, A' \in M_{m,n}(k)$ et $s \in k$,*

$${}^t(s \cdot A + A') = s \cdot {}^tA + {}^tA'.$$

De plus, si $B \in M_{n,p}(k)$, alors $AB \in M_{m,p}(k)$ et l'on a dans $M_{p,m}(k)$ l'égalité :

$$(\star\star) \qquad \boxed{{}^t(AB) = {}^tB\,{}^tA.}$$

Démonstration. Écrivons $A = (a_{ij})$ et $A' = (a'_{ij})$. Alors $sA + A' \in M_{m,n}(k)$ est la matrice $(sa_{ij} + a'_{ij})$, et sa transposée est la matrice $C \in M_{n,m}(k)$ telle que, pour tout (j, i),

$$C_{ji} = (sA + A')_{ij} = sa_{ij} + a'_{ij} = s({}^tA)_{ji} + ({}^tA')_{ji},$$

donc $C = s \cdot {}^tA + {}^tA'$. Puis, si l'on pose $B = (b_{j\ell}) \in M_{n,p}(k)$, alors pour tout couple (i, ℓ) on a

$$\big({}^t(AB)\big)_{\ell i} = (AB)_{i\ell} = \sum_{r=1}^{n} a_{ir} b_{r\ell} = \sum_{r=1}^{n} ({}^tB)_{\ell r} \cdot ({}^tA)_{ri} = \big({}^tB\,{}^tA\big)_{\ell i}$$

ce qui montre que ${}^t(AB) = {}^tB\,{}^tA$. □

0.5. Changements de base

***Définition 0.5.1* (Automorphismes et matrices inversibles).** — 1) Soit V un k-espace vectoriel. On dit qu'un endomorphisme f de V est un **automorphisme** s'il possède un inverse dans $\mathrm{End}_k(V)$, i.e. s'il existe un endomorphisme g de V tel que $f \circ g = \mathrm{id}_V = g \circ f$. Ceci équivaut à dire que f est bijectif, car on a vu (cf. 0.1.10) que dans ce cas l'application inverse g est automatiquement linéaire.

2) On note $\mathrm{GL}(V)$ l'ensemble des automorphismes de V; c'est un groupe pour la composition des endomorphismes : en effet, la composition des endomorphismes est associative, l'application identique est élément neutre, et si f, g sont inversibles, alors $f \circ g$ l'est aussi, son inverse étant $g^{-1} \circ f^{-1}$ (puisque $f \circ g \circ g^{-1} \circ f^{-1} = \mathrm{id}_V = g^{-1} \circ f^{-1} \circ f \circ g$). [4]

3) De même, on dit qu'une matrice $A \in M_n(k)$ est **inversible** s'il existe $B \in M_n(k)$ vérifiant $AB = I_n = BA$, où I_n désigne la matrice identité de taille n; dans ce cas B est notée A^{-1}. On note $\mathrm{GL}_n(k)$ l'ensemble des matrices inversibles.

[4] Attention! Noter l'inversion de l'ordre des facteurs: $(f \circ g)^{-1}$ égale $g^{-1} \circ f^{-1}$, tandis que $f^{-1} \circ g^{-1} = (g \circ f)^{-1}$!

Comme la correspondance bijective $\mathrm{End}_k(k^n) \longleftrightarrow M_n(k)$ transforme la composition des endomorphismes en le produit des matrices, on voit qu'une matrice A est inversible si et seulement si l'endomorphisme correspondant u de k^n est bijectif, et dans ce cas A^{-1} est la matrice de u^{-1}; de plus $\mathrm{GL}_n(k)$ est un groupe pour la multiplication des matrices: la matrice I_n est élément neutre, et si A, B sont inversibles, alors AB l'est aussi, son inverse étant $B^{-1}A^{-1}$.

La proposition suivante est importante et très utile:

Proposition 0.5.2. — (i) *Soit V un k-espace vectoriel* de dimension finie *et soient $u, v \in \mathrm{End}_k(V)$ tels que $u \circ v = \mathrm{id}_V$. Alors u et v sont bijectifs et $u = v^{-1}$.*

(i′) *Soient $A, B \in M_n(k)$ telles que $AB = I_n$. Alors on a aussi $BA = I_n$ et donc A et B sont inversibles et inverses l'une de l'autre.*

(ii) *Si A est inversible alors ${}^tA$ est inversible et l'on a $({}^tA)^{-1} = {}^t(A^{-1})$.*

Démonstration. (i) Pour tout $x \in V$, on a $x = u(v(x))$, donc u est surjectif, et v est injectif. Donc, d'après la proposition 0.3.3, u et v sont bijectifs; alors en multipliant l'égalité $u \circ v = \mathrm{id}_V$ à gauche par u^{-1} (ou bien à droite par v^{-1}), on obtient que $v = u^{-1}$.

(i′) Notons u (resp. v) l'endomorphisme de k^n correspondant à A (resp. B). Comme $AB = I_n$ équivaut à $u \circ v = \mathrm{id}_V$ alors, d'après (i), u et v sont bijectifs et inverses l'un de l'autre, donc il en est de même de A et B et l'on a $B = A^{-1}$ et $BA = I_n$.

(ii) Supposons A inversible, alors il existe $B \in M_n(k)$ telle que $AB = I_n = BA$. Prenant la transposée de ces matrices, on obtient, puisque ${}^tI_n = I_n$:

$$ {}^tB\,{}^tA = {}^t(AB) = I_n = {}^t(BA) = {}^tA\,{}^tB. $$

Ceci montre que ${}^tA$ est inversible, d'inverse ${}^tB = {}^t(A^{-1})$.

Remarque 0.5.3. — 1) Soit $V = \mathbb{R}[X]$, soit I l'opérateur "d'intégration", qui envoie chaque monôme X^n sur $X^{n+1}/(n+1)$, et soit D l'opérateur de dérivation, qui envoie chaque polynôme P sur le polynôme dérivé P'. Alors $D \circ I = \mathrm{id}_V$, donc D est surjectif et I injectif, mais D n'est pas injectif car $D(1) = 0$, et I n'est pas surjectif car son image est formée des polynômes de terme constant nul.

2) De même, soit V l'espace des suites $(u_n)_{n\in\mathbb{N}}$ et soient I (resp. D) l'opérateur qui envoie toute suite $(u_n)_{n\in\mathbb{N}}$ sur

$$ (0, u_0, u_1, u_2, \dots) \qquad \text{resp.} \qquad (u_1, u_2, u_3, \dots). $$

Alors $D \circ I = \mathrm{id}_V$, donc D est surjectif et I injectif, mais D n'est pas injectif car D annule la suite telle que $u_0 = 1$ et $u_i = 0$ pour $i \geq 1$, et I n'est pas surjectif car son image est formée des suites de terme u_0 nul.

Ces deux exemples montrent que si V est un espace vectoriel qui n'est pas de dimension finie et si $u, v \in \mathrm{End}_k(V)$ vérifient $u \circ v = \mathrm{id}_V$, alors u et v ne sont pas nécessairement bijectifs.

Lemme 0.5.4. — *Soient f un endomorphisme d'un espace vectoriel V de dimension finie, $\mathscr{B} = (v_1, \dots, v_n)$ une base de V, posons $w_i = f(v_i)$. Si $(w_1, \dots, w_n)$ est une base de V, alors f est bijective, et son inverse est l'endomorphisme g de V défini par $g(w_i) = v_i$ pour $i = 1, \dots, n$.*

Démonstration. On suppose que $\mathscr{C} = (w_1, \dots, w_n)$ est une base de V. Alors f est surjectif, car pour tout $w \in V$ il existe $t_1, \dots, t_n \in k$ tels que

$$ w = t_1w_1 + \cdots + t_nw_n = f(t_1v_1 + \cdots + t_nv_n). $$

Donc, d'après la proposition 0.3.3, f est bijectif. (On peut aussi voir directement que f est injectif: soit $v \in \mathrm{Ker}(f)$, il existe $t_1, \dots, t_n \in k$ tels que $v = x_1v_1 + \cdots + x_nv_n$, alors $0 = f(v) = x_1w_1 + \cdots + x_nw_n$ donc, puisque $(w_1, \dots, w_n)$ est libre, $t_1 = 0 = \cdots = t_n$, d'où $v = 0$.)

Soit g l'endomorphisme de V défini par $g(w_i) = v_i$, pour $i = 1, \dots, n$. Alors on a, d'une part, $(g \circ f)(v_i) = v_i$ pour tout i, d'où $g \circ f = \mathrm{id}_V$, et, d'autre part, $(f \circ g)(w_i) = w_i$ pour tout i, d'où $f \circ g = \mathrm{id}_V$.

***Définition 0.5.5* (Matrice de passage)**. — Soient V un espace vectoriel de dimension finie, $\mathscr{B} = (e_1, \dots, e_n)$ une base de V, et $\mathscr{B}' = (v_1, \dots, v_n)$ une seconde base de V. Soit P la matrice $n \times n$ exprimant la seconde base en fonction de la première, c.-à-d., chaque v_j s'écrit de façon unique

$$ v_j = p_{1j}\,e_1 + \cdots + p_{nj}\,e_n $$

et l'on forme la matrice

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1j} & \cdots & p_{1n} \\ \vdots & & \vdots & & \vdots \\ p_{i1} & \cdots & p_{ij} & \cdots & p_{in} \\ \vdots & & \vdots & & \vdots \\ p_{n1} & \cdots & p_{nj} & \cdots & p_{nn} \end{pmatrix}$$

dont les colonnes sont les vecteurs $v_1, \dots, v_n$ exprimés dans la base $(e_1, \dots, e_n)$. Alors P s'appelle la **matrice de passage** de la base $\mathscr{B} = (e_1, \dots, e_n)$ à la base $\mathscr{B}' = (v_1, \dots, v_n)$ et se note $\boxed{\mathrm{Mat}_{\mathscr{B}}(\mathscr{B}').}$

C'est une matrice **inversible**: son inverse P^{-1} est la matrice exprimant $e_1, \dots, e_n$ dans la base $(v_1, \dots, v_n)$.

Remarquons que P peut être vue comme la matrice de l'application identité id_V, exprimée dans les bases: $\mathscr{B}' = (v_1, \dots, v_n)$ au départ, et $\mathscr{B} = (e_1, \dots, e_n)$ à l'arrivée, c.-à-d.,

$$\boxed{P = \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(\mathrm{id}_V)} \qquad \text{et de même} \qquad \boxed{P^{-1} = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{id}_V).}$$

Conservons les notations précédentes. Tout $v \in V$ s'écrit alors de façon unique

$$v = x_1 e_1 + \cdots + x_n e_n \qquad \text{et} \qquad v = x'_1 v_1 + \cdots + x'_n v_n$$

et les x_i (resp. x'_i) s'appellent les **coordonnées** de v relativement à la base $\mathscr{B}$ (resp. $\mathscr{B}'$), cf. 0.2.10. Relativement à la base $\mathscr{B}$ (resp. $\mathscr{B}'$), on peut représenter v comme le vecteur colonne

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{resp.} \qquad X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

***Proposition 0.5.6* (Changement de coordonnées)**. — *La formule de changement de coordonnées, pour le changement de base $\mathscr{B} \to \mathscr{B}'$ donné par la matrice de passage P, est donnée par:*

$$\boxed{X = PX'}$$

(noter que cette formule exprime les anciennes *coordonnées X en fonction des* nouvelles *X').*

Démonstration. En effet, écrivant $v_j = \sum_{i=1}^n p_{ij} e_i$ puis

$$v = \sum_{j=1}^n x'_j v_j = \sum_{j=1}^n \sum_{i=1}^n x'_j p_{ij} e_i = \sum_{i=1}^n \left(\sum_{j=1}^n p_{ij} x'_j \right) e_i$$

et comparant avec $v = \sum_{i=1}^n x_i e_i$, on obtient qu'on a $x_i = \sum_{j=1}^n p_{ij} x'_j$ pour $i = 1, \dots, n$, d'où $X = PX'$.

Soient $V, \mathscr{B}, \mathscr{B}'$ et P comme plus haut, et considérons maintenant une application linéaire $u : V \to W$. Soient $\mathscr{C} = (f_1, \dots, f_m)$ une base de W, et $A \in M_{m,n}(k)$ la matrice de u dans les bases $\mathscr{B}$ et $\mathscr{C}$. Alors, d'après le théorème 0.4.3, on a:

$$\mathrm{Mat}_{\mathscr{C},\mathscr{B}'}(u) = \mathrm{Mat}_{\mathscr{C},\mathscr{B}}(u) \cdot \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(\mathrm{id}_V)$$

donc la matrice de u dans les bases $\mathscr{B}'$ et $\mathscr{C}$ est AP.

Enfin, soit $\mathscr{C}' = (w_1, \dots, w_m)$ une seconde base de W et soit Q la matrice de passage de $\mathscr{C}$ à $\mathscr{C}'$; alors $Q = \mathrm{Mat}_{\mathscr{C},\mathscr{C}'}(\mathrm{id}_W)$ et $Q^{-1} = \mathrm{Mat}_{\mathscr{C}',\mathscr{C}}(\mathrm{id}_W)$, d'où:

$$\mathrm{Mat}_{\mathscr{C}',\mathscr{B}'}(u) = \mathrm{Mat}_{\mathscr{C}',\mathscr{C}}(\mathrm{id}_W) \cdot \mathrm{Mat}_{\mathscr{C},\mathscr{B}'}(u) = Q^{-1}AP.$$

On a donc obtenu le théorème suivant:

***Théorème 0.5.7* (Changement de bases pour une application linéaire)**

Soit $A \in M_{m,n}(k)$ la matrice d'une application linéaire $u : V \to W$, relativement à des bases $\mathscr{B} = (e_1, \dots, e_n)$ de V et $\mathscr{C} = (f_1, \dots, f_m)$ de W. Soit $\mathscr{B}' = (v_1, \dots, v_n)$ (resp. $\mathscr{C}' = (w_1, \dots, w_m)$) une seconde base de V (resp. de W) et soit P (resp. Q) la matrice de passage correspondante. Alors la matrice de u dans les bases $\mathscr{B}'$ et $\mathscr{C}'$ est :

$$\boxed{\mathrm{Mat}_{\mathscr{C}',\mathscr{B}'}(u) = Q^{-1}AP.}$$

Remarque 0.5.8. — Le théorème précédent est (évidemment) compatible avec la formule de changement de coordonnées 0.5.6: si l'on désigne par X (resp. X') les coordonnées d'un vecteur $v \in V$ relativement à $\mathscr{B}$ (resp. $\mathscr{B}'$), et Y (resp. Y') les coordonnées du vecteur $u(v)$ dans la base $\mathscr{C}$ (resp. $\mathscr{C}'$), alors on a:

$$Y = AX, \qquad X = PX', \qquad Y = QY'$$

d'où $Y' = Q^{-1}Y = Q^{-1}AX = Q^{-1}APX'$. Compte-tenu de cette "compatibilité", on peut utiliser (comme moyen mnémotechnique) l'une de ces formules pour retrouver l'autre...

Remarque 0.5.9. — Même si l'on s'intéresse au départ à une *matrice* $A \in M_{m,n}(k)$, il est souvent utile de considérer A comme une application linéaire $u : k^n \to k^m$ (définie par $u(e_j) = a_{1j}f_1 + \cdots + a_{mj}f_m$, où $(e_1, \ldots, e_n)$, resp. $(f_1, \ldots, f_m)$, est la base canonique de k^n, resp. k^m). Par exemple, le théorème précédent donne alors le corollaire suivant:

Corollaire 0.5.10. — *Soit $A \in M_{m,n}(k)$ et soit $r = \mathrm{rang}(A)$.*

1) *Il existe des matrices inversibles $P \in \mathrm{GL}_n(k)$ et $Q \in \mathrm{GL}_m(k)$ telles que*

$$Q^{-1}AP = \begin{pmatrix} I_r & \mathbf{0}_{r,n-r} \\ \mathbf{0}_{m-r,r} & \mathbf{0}_{m-r,n-r} \end{pmatrix}$$

où I_r est la matrice identité de taille r et où $\mathbf{0}_{p,q}$ désigne la matrice nulle à p lignes et q colonnes.

2) *Réciproquement, s'il existe des matrices inversibles $P \in \mathrm{GL}_n(k)$ et $Q \in \mathrm{GL}_m(k)$ et un entier $s \in \mathbb{N}$ tels que*

$$Q^{-1}AP = \begin{pmatrix} I_s & \mathbf{0}_{s,n-s} \\ \mathbf{0}_{m-s,s} & \mathbf{0}_{m-s,n-s} \end{pmatrix}$$

alors $s = \mathrm{rang}(A)$.

Démonstration. Soient $(e_1, \ldots, e_n)$ et $(f_1, \ldots, f_m)$ les bases canoniques de k^n et k^m et soit u l'application linéaire $k^n \to k^m$ correspondant à A. Par définition, $r = \mathrm{rang}(A)$ est la dimension de $\mathrm{Im}(u)$. Soit donc $(w_1, \ldots, w_r)$ une base de $\mathrm{Im}(u)$, on peut la compléter en une base $\mathscr{C} = (w_1, \ldots, w_m)$ de k^m; notons Q la matrice de passage de $(f_1, \ldots, f_m)$ à $\mathscr{C}$.

Soient $v_1, \ldots, v_r$ des éléments de k^n tels que $u(v_j) = w_j$, pour $j = 1, \ldots, r$, et soit $(e_1, \ldots, e_d)$ une base de $\mathrm{Ker}(u)$. D'après la démonstration (2ème méthode) du théorème du rang 0.3.2, $\mathscr{B} = (v_1, \ldots, v_r, e_1, \ldots, e_d)$ est une base de k^n. Alors, la matrice de u dans les bases $\mathscr{B}$ et $\mathscr{C}$ est

$$\mathrm{Mat}_{\mathscr{C},\mathscr{B}}(u) = \begin{pmatrix} I_r & \mathbf{0}_{r,n-r} \\ \mathbf{0}_{m-r,r} & \mathbf{0}_{m-r,n-r} \end{pmatrix}.$$

Or, si P désigne la matrice de passage de $(e_1, \ldots, e_n)$ à $\mathscr{B}$, on a

$$\mathrm{Mat}_{\mathscr{C},\mathscr{B}}(u) = Q^{-1} \cdot \mathrm{Mat}_{(f_i),(e_j)}(u) \cdot P = Q^{-1}AP,$$

d'où l'assertion 1) du corollaire.

Réciproquement, supposons qu'il existe des matrices inversibles $P \in \mathrm{GL}_n(k)$ et $Q \in \mathrm{GL}_m(k)$ et un entier $s \in \mathbb{N}$ tels que

$$Q^{-1}AP = \begin{pmatrix} I_s & \mathbf{0}_{s,n-s} \\ \mathbf{0}_{m-s,s} & \mathbf{0}_{m-s,n-s} \end{pmatrix}.$$

Ceci signifie qu'il existe des bases $(v_1, \ldots, v_n)$ de k^n et $(w_1, \ldots, w_m)$ de k^m telles que $u(v_i) = w_i$ pour $i = 1, \ldots, s$, et $u(v_j) = 0$ pour $j = s+1, \ldots, n$. Alors $\mathrm{Im}(f) = \mathrm{Vect}(w_1, \ldots, w_s)$ est de dimension s, d'où $s = \mathrm{rang}(A)$.

On déduit du corollaire 0.5.10 la proposition suivante.

Proposition 0.5.11. — *Soit $A \in M_{m,n}(k)$. Alors*

$$\boxed{\mathrm{rang}(A) = \mathrm{rang}({}^tA)}$$

par conséquent $\mathrm{rang}(A)$ *est aussi le nombre maximum de* lignes *de A qui sont linéairement indépendantes.*

Démonstration. D'après ce qui précède, il existe $P \in \mathrm{GL}_n(k)$ et $Q \in \mathrm{GL}_m(k)$ telles que

$$Q^{-1}AP = \begin{pmatrix} I_r & \mathbf{0}_{r,n-r} \\ \mathbf{0}_{m-r,r} & \mathbf{0}_{m-r,n-r} \end{pmatrix}.$$

où $r = \mathrm{rang}(A)$. Alors

$${}^tP\,{}^tA\,{}^tQ^{-1} = \begin{pmatrix} I_r & \mathbf{0}_{r,m-r} \\ \mathbf{0}_{n-r,r} & \mathbf{0}_{n-r,m-r} \end{pmatrix}.$$

Or, d'après la proposition 0.5.2, ${}^tP$ et ${}^tQ^{-1}$ sont inversibles, donc l'égalité ci-dessus entraîne, d'après le corollaire précédent, que $r = \mathrm{rang}({}^tA)$.

***Définition 0.5.12* (Matrices équivalentes).** — Soient $A, B \in M_{m,n}(k)$; on dit que A et B sont *équivalentes* s'il existe des matrices inversibles $P \in \mathrm{GL}_n(k)$ et $Q \in \mathrm{GL}_m(k)$ telles que $Q^{-1}AP = B$. (D'après le corollaire 0.5.10, ceci équivaut à dire que A et B ont même rang).

Cas des endomorphismes. — Le théorème 0.5.7 traite le cas d'une application linéaire $u : V \to W$, où V, W sont *a priori* distincts. Dans ce cas, lorsque qu'on s'autorise des changements de bases arbitraires dans V et dans W, le corollaire 0.5.10 montre que le seul invariant de u est son rang, qui est un entier compris entre 0 et $\mathrm{Min}(\dim V, \dim W)$.

Mais, lorsque $V = W$ et qu'on s'intéresse à la nature géométrique d'un endomorphisme u de V, c.-à-d., lorsqu'on veut comparer $u(x)$ et x, pour x variant dans V, pour pouvoir faire la comparaison on veut exprimer x et $u(x)$ dans la même base, et c'est la raison pour laquelle, dans ce cas, on écrit la matrice de u dans une *même* base $\mathscr{B}$ au départ et à l'arrivée.

(Par exemple, si $V = W = k$, les automorphismes de k comme k-espace vectoriel sont les homothéties $h_\lambda : x \mapsto \lambda x$ avec $\lambda \neq 0$, si on prend $\{1\}$ comme base de départ et $\{\lambda\}$ comme base d'arrivée, la matrice est (1) donc on a "perdu" le rapport λ de l'homothétie; mais si on impose de garder la même base au départ et à l'arrivée, la matrice est (λ)...)

Alors, le théorème 0.5.7 donne dans ce cas:

***Théorème 0.5.13* (Changement de base pour un endomorphisme)**

Soit A la matrice d'un endomorphisme u de V relativement à une base $\mathscr{B}$ de V. Si $\mathscr{B}'$ est une seconde base, et si P est la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$, alors la matrice de u dans la base $\mathscr{B}'$ est:

$$\boxed{\mathrm{Mat}_{\mathscr{B}'}(u) = P^{-1}AP.}$$

***Définition 0.5.14* (Matrices semblables).** — Soit $A, B \in M_n(k)$ des matrices *carrées* de taille n. On dit que A, B sont *semblables* s'il existe une matrice inversible $P \in \mathrm{GL}_n(k)$ telle que $P^{-1}AP = B$. Dans ce cas, on dit que A, B sont dans la même *classe de similitude*.

***Remarques 0.5.15*.** — 1) On notera que cette définition ne fait sens que pour des matrices *carrées*.

2) A et B sont semblables si et seulement si elles représentent, dans des bases différentes, le même endomorphisme de k^n.

3) Si $A, B \in M_n(k)$ sont semblables, elles sont évidemment équivalentes, mais la réciproque est loin d'être vraie: les classes de similitude forment une partition de $M_n(k)$ beaucoup plus raffinée que celle donnée par le rang (cf. le cas $n = 1$, et voir plus loin pour le cas n arbitraire).

0.6. Appendice (†): compléments sur les familles génératrices ou libres

Dans cet appendice, on donne la définition des familles génératrices ou libres éventuellement infinies, ainsi qu'une démonstration du théorème sur la dimension (0.2.8).

***Définitions 0.6.1* (Familles génératrices).** — Soit V un k-espace vectoriel.

1) Soit S un sous-ensemble (fini ou infini) de V. L'ensemble de toutes les combinaisons linéaires *finies*

$$t_1v_1 + \cdots + t_rv_r, \qquad \text{pour } r \in \mathbb{N}^* \text{ (variable)}, \ v_i \in S, \ t_i \in k,$$

forme un sous-espace vectoriel de V, et c'est le plus petit sous-espace vectoriel de V contenant S. On l'appelle le sous-espace *engendré* par S et on le note $\mathrm{Vect}(S)$.

En particulier, si E est un sous-ensemble *fini* $\{v_1, \ldots, v_n\}$, alors $\mathrm{Vect}(E)$ est l'ensemble des combinaisons linéaires

$$t_1v_1 + \cdots + t_nv_n, \qquad \text{où } t_1, \ldots, t_n \in k.$$

2) On dit que S est un ensemble de générateurs (ou une *famille génératrice*) de V si le sous-espace engendré $\mathrm{Vect}(S)$ égale V, c.-à-d., si tout élément de V s'écrit comme combinaison linéaire des éléments de S.

Il résulte de la définition que: *toute famille contenant une famille génératrice est génératrice.*

***Exemple 0.6.2*.** — Les monômes X^n, pour $n \in \mathbb{N}$, engendrent l'espace vectoriel $k[X]$. En effet, tout polynôme $P \in k[X]$ s'écrit comme une combinaison linéaire finie: $P = a_0 + a_1X + \cdots + a_dX^d$.

Exercices 0.6.3. — 1) Soit $\mathbb{R}[[X]]$ l'espace des séries formelles $\sum_{i=0}^{\infty} a_i X^i$, avec $a_i \in \mathbb{R}$. Quel est le sous-espace de $\mathbb{R}[[X]]$ engendré par les monômes X^n, pour $n \in \mathbb{N}$? (Réponse: c'est $\mathbb{R}[X]$.)

2) Pouvez-vous montrer que l'espace vectoriel $\mathbb{R}[X]$ n'est *pas* finiment engendré?

***Définitions 0.6.4* (Familles libres ou liées)**. — Soit V un k-espace vectoriel et soit S un sous-ensemble (fini ou infini) de V. On dit que les éléments de S sont *linéairement indépendants* (ou que S est une *famille libre*) s'il n'existe pas de relation linéaire non triviale entre les éléments de S, c.-à-d., si la condition suivante est vérifiée:

$$\text{(FL)} \qquad \begin{cases} \textit{pour tous } r \in \mathbb{N}^*, \ v_1, \dots, v_r \in S \textit{ deux-à-deux distincts, et } t_1, \dots, t_r \in k, \textit{ si l'on a} \\ \textit{une relation } t_1 v_1 + \cdots + t_r v_r = 0, \ \textit{alors } t_1 = 0 = \cdots = t_r. \end{cases}$$

(Ceci prend une forme plus simple si S est un ensemble *fini* $\{v_1, \dots, v_n\}$; dans ce cas la condition s'écrit plus simplement: pour *tous* $t_1, \dots, t_n \in k$, si $t_1 v_1 + \cdots + t_n v_n = 0$, alors $t_1 = 0 = \cdots = t_n$.)
Il résulte de la définition que: *toute sous-famille d'une famille libre est libre.*

Au contraire, on dit que les éléments de S sont *linéairement dépendants* (ou que S est une *famille liée*) s'il existe une relation linéaire non triviale entre les éléments de S, c.-à-d., s'il existe un entier $r \geq 1$, des éléments $v_1, \dots, v_r \in S$ deux-à-deux distincts, et des scalaires $t_1, \dots, t_r$, non tous nuls, tels que $t_1 v_1 + \cdots + t_r v_r = 0$.
Il résulte de la définition que: *toute famille contenant une famille liée est liée.*

Exemple 0.6.5. — Dans $k[X]$, la famille des monômes $(X^n)_{n\in\mathbb{N}}$ est libre. En effet, soient $i_1 < \cdots < i_r$ dans $\mathbb{N}$ et soient $t_1, \dots, t_r \in k$ non tous nuls. Alors le polynôme

$$t_1 X^{i_1} + \cdots + t_r X^{i_r}$$

est non nul. Ceci montre que $(X^n)_{n\in\mathbb{N}}$ est une famille libre.

Remarque 0.6.6. — Soient V un k-espace vectoriel et $S = (v_i)_{i\in I}$ une famille de vecteurs indexée par un ensemble d'indice I infini. Alors S est libre si et seulement si la condition d'unicité suivante est vérifiée: (U) "si l'on a une égalité

$$\sum_{j\in J} t_j v_j = \sum_{p\in P} s_p v_p \qquad t_j \in k, \quad s_p \in k,$$

où J, P sont deux sous-ensembles finis de I, alors $\{j \in J \mid t_j \neq 0\}$ égale $\{p \in P \mid s_p \neq 0\}$ et, notant L cet ensemble, on a $t_\ell = s_\ell$ pour tout $\ell \in L$".

En effet, supposons cette condition vérifiée; si l'on a une égalité $\sum_{j\in J} t_j v_j = 0$ (le terme de droite correspond à $P = \varnothing$: une somme indexée par $\varnothing$ vaut 0), alors $\{j \in J \mid t_j \neq 0\}$ égale $\varnothing$, i.e. tous les t_j sont nuls; ceci montre que S est une famille libre.

Réciproquement, supposons que S soit une famille libre, et qu'on ait une égalité $\sum_{j\in J} t_j v_j = \sum_{p\in P} s_p v_p$ comme plus haut, alors on a:

$$0 = \sum_{j\in J-P} t_j v_j + \sum_{i\in J\cap P} (t_i - s_i) v_i - \sum_{p\in P-J} s_p v_p$$

et comme S est libre, ceci entraîne que $t_j = 0 = s_p$ pour $j \in J - P$ et $p \in P - J$, et que $t_i = s_i$ pour tout $i \in J \cap P$, donc la condition (U) est vérifiée.

***Définition 0.6.7* (Bases)**. — Soient V un k-espace vectoriel et $(v_i)_{i\in I}$ une famille (finie ou infinie) de vecteurs de V. On dit que $\mathscr{B} = (v_i)_{i\in I}$ est une *base* de V si tout élément v de V *s'écrit de façon unique comme combinaison linéaire des* v_i, c.-à-d., si :

(1) $\mathscr{B}$ est une famille génératrice, i.e. pour tout $v \in V$, il existe un sous-ensemble *fini* J de I (dépendant de v) et des scalaires $t_j \in k$, pour $j \in J$, tels que $v = \sum_{j\in J} t_j v_j$.

(2) $\mathscr{B}$ vérifie la condition (U), i.e. si l'on a un second sous-ensemble fini P de I et des scalaires s_p, pour $p \in P$, tels que

$$v = \sum_{j\in J} t_j v_j = \sum_{p\in P} s_p v_p,$$

alors le sous-ensemble $L = \{j \in J \mid t_j \neq 0\}$ égale $\{p \in P \mid s_p \neq 0\}$ et pour tout $\ell \in L$, on a $t_\ell = s_\ell$.
D'après la remarque précédente, ceci équivaut à dire que $\mathscr{B}$ est une famille *génératrice et libre.*

Exemple 0.6.8. — La famille des monômes $(X^n)_{n\in\mathbb{N}}$ est une base de $k[X]$: en effet, tout polynôme non nul $P \in k[X]$ s'écrit de façon unique

$$P = a_0 + \cdots + a_d X^d, \qquad \text{où } a_i \in k, \ a_d \neq 0, \ d = \deg P.$$

Définitions 0.6.9. — Soient V un k-espace vectoriel, et $\mathscr{F} = (v_i)_{i\in I}$ une famille de vecteurs de V.

(i) On dit que $\mathscr{F}$ est une *famille génératrice minimale* si c'est une famille génératrice, et si "on ne peut pas la rendre plus petite", c.-à-d., si pour tout sous-ensemble $I' \neq I$, la famille $(v_i)_{i\in I'}$ n'est *pas* génératrice.

(ii) On dit que $\mathscr{F}$ est une *famille libre maximale* si c'est une famille libre, et si "on ne peut pas la rendre plus grande", c.-à-d., si pour tout $v \in V - \mathscr{F}$, la famille $\mathscr{F} \cup \{v\}$ est liée.

Proposition 0.6.10. — *Soient V un k-espace vectoriel, et $\mathscr{F} = (v_i)_{i\in I}$ une famille non vide de vecteurs de V. Les conditions suivantes sont équivalentes:*

a) $\mathscr{F}$ *est une* base *de V;*

b) $\mathscr{F}$ *est une* famille génératrice minimale*;*

c) $\mathscr{F}$ *est une* famille libre maximale*.*

Démonstration. Supposons que $(v_i)_{i\in I}$ soit une base de V, c.-à-d., une famile *génératrice et libre.* Alors, pour tout $i \in I$, la famille $(v_j)_{j\in I-\{i\}}$ n'est *pas* génératrice: sinon v_i s'exprimerait comme combinaison linéaire des v_j, pour $j \neq i$, d'où une relation linéaire non triviale $v_i - (t_1 v_{j_1} + \cdots + t_r v_{j_r}) = 0$. Ceci montre que $(v_i)_{i\in I}$ est une famille génératrice minimale.

De plus, tout $v \in V$ s'écrit comme combinaison linéaire (d'un nombre fini) des v_i, donc si $v \notin \mathscr{F}$, la famille strictement plus grande $\mathscr{F} \cup \{v\}$ est liée. Ceci montre que $\mathscr{F}$ est une famille libre maximale. On a donc prouvé que a) implique b) et c).

b) $\Rightarrow$ a) Supposons que $(v_i)_{i\in I}$ soit une famille génératrice minimale et montrons qu'elle est libre. Sinon, on aurait une relation linéaire non triviale

$$t_1 v_{i_1} + \cdots + t_r v_{i_r} = 0$$

avec $r \geq 1$ et $t_p \neq 0$ pour au moins un indice p; alors on aurait $v_{i_p} = -t_p^{-1} \sum_{q\neq p} t_q\, v_{i_q}$ et donc la famille $(v_i)_{i\in I-\{i_p\}}$ serait déjà génératrice, contredisant la minimalité. Ceci prouve b) $\Rightarrow$ a).

c) $\Rightarrow$ a) Supposons que $\mathscr{F}$ soit une famille libre maximale et montrons qu'elle est génératrice. Soit $v \in V - \mathscr{F}$, alors la famille $\mathscr{F} \cup \{v\}$ est liée, donc on a une relation linéaire

$$sv + t_1 v_{i_1} + \cdots + t_r v_{i_r} = 0$$

non triviale (c.-à-d., s et les t_p non tous nuls). On ne peut avoir $s = 0$ car sinon, les v_i étant linéairement indépendants, on aurait $t_1 = 0 = \cdots = t_r$. Donc $s \neq 0$, d'où $v = -s^{-1}(t_1 v_{i_1} + \cdots + t_r v_{i_r})$. Donc $\mathscr{F}$ est une famille libre et génératrice, donc une base de V. Ceci prouve c) $\Rightarrow$ a).

***Corollaire 0.6.11* (Existence de bases).** — *Tout k-espace vectoriel finiment engendré V possède une base. (On convient que l'ensemble vide $\varnothing$ est une base de l'espace nul $\{0\}$.)*

En effet, par hypothèse V est engendré par une famille *finie* $\{v_1, \ldots, v_r\}$. Celle-ci contient une sous-famille génératrice $\mathscr{B}$ de cardinal minimal, donc minimale, et d'après la proposition, $\mathscr{B}$ est une base de V.

On va voir dans un instant (cf. 0.6.13 ci-dessous) que si $(v_1, \ldots, v_n)$ est une base de V, alors *toutes les bases de V ont n éléments.* Commençons par le lemme important suivant.

Lemme 0.6.12. — *Soient V un k-espace vectoriel, $v_1, \ldots, v_n \in V$, et soient $m > n$ et $u_1, \ldots, u_m$ des éléments de V qui sont combinaison linéaire de $v_1, \ldots, v_n$. Alors $u_1, \ldots, u_m$ sont liés.*

Démonstration. On procède par récurrence sur n; le résultat est clair si $n = 1$, supposons donc $n \geq 2$ et le résultat établi pour $n - 1$. Écrivons:

$$\begin{cases} u_1 &= a_{11}v_1 + \cdots + a_{1n}v_n \\ u_2 &= a_{21}v_1 + \cdots + a_{2n}v_n \\ \vdots & \qquad\qquad \vdots \\ u_m &= a_{m1}v_1 + \cdots + a_{mn}v_n. \end{cases}$$

Si tous les u_i sont combinaison linéaire des $n-1$ vecteurs $v_2, \ldots, v_n$, alors les u_i sont liés, d'après l'hypothèse de récurrence. Sinon, quitte à renuméroter les u_i, on peut supposer que $a_{11} \neq 0$. Alors, pour $i = 2, \ldots, m$, les $m - 1$ vecteurs

$$u_i' = u_i - a_{i1}a_{11}^{-1}u_1$$

sont combinaison linéaire des $n-1$ vecteurs $v_2, \dots, v_n$, donc sont liés d'après l'hypothèse de récurrence. Donc il existe des scalaires non tous nuls $t_2, \dots, t_m$ tels que

$$0 = t_2 u_2' + \cdots + t_m u_m' = t_2 u_2 + \cdots + t_m u_m - \big(\sum_{i=2}^{m} t_i a_{i1} a_{11}^{-1}\big) u_1$$

et ceci montre que les u_i sont liés. Le lemme est démontré.

Le lemme précédent a les conséquences très importantes suivantes. Soit V un k-espace vectoriel engendré par un nombre fini d'éléments $x_1, \dots, x_N$. D'après le corollaire 0.6.11, V possède une base $\mathscr{B} = (v_1, \dots, v_n)$ formée de $n \leq N$ éléments. Comme $\mathscr{B}$ est une famille libre de V, il résulte du lemme précédent que:

a) toute famille génératrice de V a au moins n éléments;

b) toute famille génératrice de V ayant n éléments est *minimale*, donc d'après 0.6.10 est une base de V.

D'autre part, comme $\mathscr{B}$ est une famille génératrice de V, il résulte du lemme précédent que:

a$'$) toute famille libre dans V a au plus n éléments;

b$'$) toute famille libre de V ayant n éléments est *maximale*, donc d'après 0.6.10 est une base de V.

Enfin, en combinant a) et b), on voit que: *toute base de V, étant une famille à la fois génératrice et libre, a exactement n éléments.* On obtient donc le théorème fondamental suivant.

***Théorème 0.6.13* (Dimension d'un espace vectoriel).** — *Soit V un k-espace vectoriel finiment engendré.*

(i) *Il existe des bases de V, et toutes ont même cardinal n; cet entier s'appelle la dimension de V sur k et se note* $\dim_k V$ *ou simplement* $\dim V$.

(ii) *De toute famille génératrice $\mathscr{F}$ on peut extraire une base, en particulier $\mathscr{F}$ est de cardinal $\geq n$; de plus si* $\operatorname{card}(\mathscr{F}) = n$ *alors $\mathscr{F}$ est une base de V.*

(iii) *Toute famille libre est de cardinal $\leq n$, et toute famille libre de cardinal n est une base de V.*

(iv) "Théorème de la base incomplète"*: Toute famille libre peut être complétée en une base de V.*

(v) *Tout sous-espace W de V est finiment engendré, et* $\dim_k W \leq \dim_k V$*; de plus si* $\dim_k W = \dim_k V$, *alors $W = V$. En d'autres termes, tout sous-espace vectoriel distinct de V est de dimension* $< \dim_k V$.

Démonstration. On a déjà vu les assertions (i), (ii) et (iii). L'assertion (iv) résulte du fait que toute famille libre peut être agrandie en une famille libre maximale, c.-à-d., en une base de V (cf. la proposition 0.6.10).

Démontrons (v). Soit W un sous-espace vectoriel de V. D'après (iii), toute famille libre d'éléments de W est de cardinal $\leq n = \dim_k V$, donc W possède une famille *libre maximale* $\mathscr{C}$, de cardinal $m \leq n$. Alors $\mathscr{C}$ est une *base* de W, d'après la proposition 0.6.10, donc W est finiment engendré, et de dimension $m \leq n$. Si de plus $m = n$ alors, d'après (iii), $\mathscr{C}$ est une base de V (donc engendre V), d'où $W = V$. Le théorème est démontré.

***Proposition 0.6.14*.** — *Soit $f : V \to W$ une application linéaire.*

a) *Si f est injective et si $\mathscr{F}$ est une famille libre de V, alors $f(\mathscr{F})$ est libre.*

b) *Si f est surjective et si $\mathscr{F}$ est une famille génératrice de V, alors $f(\mathscr{F})$ engendre W.*

c) *Si f est bijective et si $\mathscr{B}$ est une base de V, alors $f(\mathscr{B})$ est une base de W.*

Démonstration. a) Supposons f injective et soit $\mathscr{F}$ une famille libre de V. S'il existe une relation linéaire dans W:

$$t_1 f(x_1) + \cdots + t_n f(x_n) = 0, \qquad t_i \in k, \quad x_i \in \mathscr{F},$$

alors $0 = f(t_1 x_1 + \cdots + t_n x_n)$, donc comme f est injective $t_1 x_1 + \cdots + t_n x_n = 0$, donc comme $\mathscr{F}$ est libre, $t_i = 0$ pour $i = 1, \dots, n$. Ceci montre que $f(\mathscr{F})$ est libre.

b) Supposons f surjective et soit $\mathscr{F}$ une famille génératrice de V. Soit $w \in W$; comme f est surjective, il existe $v \in V$ tel que $f(v) = w$. Comme $\mathscr{F}$ engendre V, il existe $x_1, \dots, x_n \in \mathscr{F}$ et $t_1, \dots, t_n \in k$ tels que $v = t_1 x_1 + \cdots + t_n x_n$, d'où $w = t_1 f(x_1) + \cdots + t_n f(x_n)$. Ceci montre que $f(\mathscr{F})$ engendre W.

c) Supposons f bijective et soit $\mathscr{B}$ une base de V. Alors, d'après a) et b), $f(\mathscr{B})$ est une famille libre et génératrice de W, donc une base de W.

CHAPITRE 1

DUALITÉ

Résumé: Soient V un k-espace vectoriel de dimension n et $\mathscr{B}$ une base de V. Dans ce chapitre, on commence par introduire l'espace dual V^* et la base duale $\mathscr{B}^*$ de $\mathscr{B}$, qui est formée des "fonctions coordonnées"relativement à la base $\mathscr{B}$. Ensuite, on introduit les opérations élémentaires sur les *colonnes* d'une matrice $A \in M_{m,n}(k)$, ce qui est utile pour déterminer une *base* de $\mathrm{Ker}(A)$ et de $\mathrm{Im}(A)$. Puis on revoit les opérations sur les lignes (vues en L1) et l'on fait le lien, d'une part, avec la théorie des systèmes linéaires (vue en L1) et, d'autre part, avec les formes linéaires sur V, i.e. les éléments de V^*, et l'on applique ceci aux équations qui définissent un sous-espace E de V (cf. §1.3).

Au passage, on montre aussi comment les opérations sur les colonnes (ou bien les lignes) permettent de déterminer si une matrice carrée $A \in M_n(k)$ est inversible et de calculer, dans ce cas, son inverse.

1.1. Formes linéaires, espace dual

Définition 1.1.1 **(Espace dual)**. — Un cas particulier très important d'espace $\mathscr{L}(V, W)$ est le cas où $W = k$. Dans ce cas, une application k-linéaire $V \to k$ s'appelle une *forme linéaire* sur V, et $\mathscr{L}(V, k)$ s'appelle *l'espace dual* de V et se note V^*.

Supposons V de dimension finie n et soit $\mathscr{B} = (e_1, \dots, e_n)$ une base de V. On sait, d'après 0.4.2, que se donner une forme linéaire $\phi : V \to k$ équivaut à se donner ses valeurs $\phi(e_j) \in k$ sur les e_j. Pour $i = 1, \dots, n$, notons alors e_i^* la forme linéaire sur V définie par

$$e_i^*(e_i) = 1 \qquad \text{et} \qquad e_i^*(e_j) = 0 \quad \text{pour } j \neq i.$$

Alors, pour tout n-uplet $(t_1, \dots, t_n) \in k^n$, la forme linéaire $t_1 e_1^* + \cdots + t_n e_n^*$ vérifie :

$$(\star) \qquad \forall j = 1, \dots, n, \qquad (t_1 e_1^* + \cdots + t_n e_n^*)(e_j) = \sum_{i=1}^{n} t_i \underbrace{e_i^*(e_j)}_{\substack{=1 \text{ si } i=j \\ =0 \text{ sinon}}} = t_j.$$

Il en résulte que pour tout $\phi \in V^*$, on a $\phi = \phi(e_1)e_1^* + \cdots + \phi(e_n)e_n^*$, puisque ces deux applications linéaires $V \to k$ coïncident sur la base $\mathscr{B}$. De plus, si l'on a une égalité $\phi = t_1 e_1^* + \cdots + t_n e_n^*$, avec les $t_i \in k$, alors d'après $(\star)$ on a nécessairement $t_j = \phi(e_j)$ pour tout j. Ceci montre que tout $\phi \in V^*$ *s'écrit de façon unique* comme combinaison des formes linéaires $e_1^*, \dots, e_n^*$, donc $\mathscr{B}^* = (e_1^*, \dots, e_n^*)$ est une *base* de V^*, appelée la *base duale* de la base $\mathscr{B} = (e_1, \dots, e_n)$. En particulier, V^* *est de dimension* n. On a donc démontré le:

Théorème 1.1.2 **(Base duale)**. — *Soient V un k-espace vectoriel de dimension n et $\mathscr{B} = (e_1, \dots, e_n)$ une base de V. Pour $i = 1, \dots, n$, on note e_i^* la forme linéaire sur V définie par $e_i^*(e_i) = 1$ et $e_i^*(e_j) = 0$ pour $j \neq i$. Alors $\mathscr{B}^* = (e_1^*, \dots, e_n^*)$ est une* base *de V^*, appelée la* base duale *de la base $\mathscr{B}$; en effet toute forme linéaire $\phi \in V^*$ s'écrit de façon unique comme combinaison linéaire des e_i^*, i.e. on a :*

$$\phi = t_1 e_1^* + \cdots + t_n e_n^*, \qquad \textit{avec} \quad t_j = \phi(e_j).$$

En particulier, V^* est de dimension n.

De plus, $e_1^*, \dots, e_n^*$ sont les formes linéaires "coordonnées" *par rapport à la base $(e_1, \dots, e_n)$, c.-à-d., tout $v \in V$ s'écrit de façon unique $v = x_1 e_1 + \cdots + x_n e_n$, et l'on a $e_i^*(v) = x_i$ pour tout i.*

1.1.3. Matrices colonnes et matrices lignes. — Soient V un k-espace vectoriel de dimension n et $\mathscr{B} = (e_1, \dots, e_n)$ une base de V. Alors, tout vecteur $v = x_1 e_1 + \cdots + x_n e_n$ est représenté par le vecteur *colonne*

$$v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in M_{n,1}(k).$$

D'autre part, toute forme linéaire $\phi \in V^*$ est un élément de $\mathscr{L}(V, k)$ donc est représentée par une matrice $L \in M_{1,n}(k)$, i.e. par une matrice *ligne*

$$L = \begin{pmatrix} t_1 & \cdots & t_n \end{pmatrix}, \qquad \text{où} \quad t_j = \phi(e_j).$$

Alors on a $\phi(v) = \sum_{j=1}^n \phi(x_j e_j) = \sum_{j=1}^n x_j \underbrace{\phi(e_j)}_{=t_j} = \sum_{j=1}^n x_j t_j = t_1 x_1 + \cdots + t_n x_n$ i.e. $\phi(v)$ est donné par le produit matriciel suivant : $\phi(v) = L \cdot v = \begin{pmatrix} t_1 & \cdots & t_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Donc, pour des entiers $p, r \geq 1$ arbitraires, si l'on a p formes linéaires $\phi_1, \dots, \phi_p \in V^*$, représentées par des matrices lignes (à n colonnes) $L_1 = \begin{pmatrix} t_{11} & \cdots & t_{1n} \end{pmatrix}, \dots, L_p = \begin{pmatrix} t_{p1} & \cdots & t_{pn} \end{pmatrix}$ et r vecteurs

$$v_1 = \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix} \quad , \quad \dots \quad , \quad v_r = \begin{pmatrix} x_{1r} \\ \vdots \\ x_{nr} \end{pmatrix}$$

alors, notant $\begin{pmatrix} L_1 \\ \vdots \\ L_p \end{pmatrix}$ l'élément de $M_{p,n}(k)$ dont les lignes sont $L_1, \dots, L_p$ et $(v_1 \cdots v_r)$ l'élément de $M_{n,r}(k)$ dont les colonnes sont $v_1, \dots, v_r$, on a l'égalité suivante de matrices à p lignes et r colonnes :

$$(\star) \qquad \boxed{\begin{pmatrix} \phi_1(v_1) & \cdots & \phi_1(v_r) \\ \vdots & & \vdots \\ \phi_p(v_1) & \cdots & \phi_p(v_r) \end{pmatrix} = \begin{pmatrix} L_1 \\ \vdots \\ L_p \end{pmatrix} \begin{pmatrix} v_1 & \cdots & v_r \end{pmatrix}.}$$

On en déduit la proposition suivante:

Proposition 1.1.4 **(Matrice de passage pour les bases duales. Base "préduale")**
Soient V un k-espace vectoriel de dimension n, $\mathscr{B}$ une base de V, et $\mathscr{B}^*$ la base duale de V^*.

(1) *Soient $\mathscr{C}$ une seconde base de V, $\mathscr{C}^*$ la base duale, et $P = \mathrm{Mat}_{\mathscr{B}}(\mathscr{C})$. Alors la matrice de passage $\mathrm{Mat}_{\mathscr{B}^*}(\mathscr{C}^*)$ égale ${}^tP^{-1}$.*

(2) *Pour toute base Δ de V^*, il existe une unique base $\mathscr{D}$ de V telle que $\mathscr{D}^* = \Delta$. On l'appellera la base* "préduale" *de Δ.*

Démonstration. Notons $\mathscr{B} = (e_1, \dots, e_n)$, $\mathscr{B}^* = (e_1^*, \dots, e_n^*)$ et $\mathscr{C} = (v_1, \dots, v_n)$, $\mathscr{C}^* = (v_1^*, \dots, v_n^*)$. Notons $Q = (q_{ij})$ la matrice de passage $\mathrm{Mat}_{\mathscr{B}^*}(\mathscr{C}^*)$, i.e. pour $j = 1, \dots, n$, on a $v_j^* = q_{1j} e_1^* + \cdots + q_{nj} e_n^*$. Donc v_j^* correspond à la matrice ligne

$$L_j = (q_{1j}, \cdots, q_{nj}) = {}^tC_j,$$

où C_j est la j-ième colonne de Q. Donc la matrice $\begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix}$ égale ${}^tQ$. D'autre part, chaque v_j est la j-ième colonne de $P = \mathrm{Mat}_{\mathscr{B}}(\mathscr{C})$. Donc, d'après $(\star)$ on obtient:

$${}^tQP = \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix} \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} = \begin{pmatrix} v_1^*(v_1) & \cdots & v_1^*(v_n) \\ \vdots & & \vdots \\ v_n^*(v_1) & \cdots & v_n^*(v_n) \end{pmatrix} = I_n$$

(la dernière égalité découlant du fait que $v_i^*(v_j) = 1$ si $i = j$ et $= 0$ sinon). Ceci prouve que ${}^tQ = P^{-1}$, i.e. $Q = {}^tP^{-1}$. L'assertion (1) est démontrée.

L'assertion (2) en découle. En effet, notons encore Q la matrice de passage $\mathrm{Mat}_{\mathscr{B}^*}(\Delta)$. S'il existe une base $\mathscr{D}$ de V telle que $\mathscr{D}^* = \Delta$ alors, d'après l'assertion (1), la matrice de passage $P = \mathrm{Mat}_{\mathscr{B}}(\mathscr{D})$ vérifie

${}^tP^{-1} = Q$, i.e. $P = {}^tQ^{-1}$. Donc $\mathscr{D}$ est déterminée (si elle existe) par la condition $\mathrm{Mat}_{\mathscr{B}}(\mathscr{D}) = {}^tQ^{-1}$; or cette condition définit bien une base de V : comme la matrice $M = {}^tQ^{-1}$ est inversible, ses colonnes $v_1, \dots, v_n$ forment une base $\mathscr{D}$ de V, telle que $\mathrm{Mat}_{\mathscr{B}}(\mathscr{D}) = M$. Ceci prouve l'assertion (2).

Exercice 1.1.5. — Etant donnée une base $\mathscr{B}$, calculer la matrice dans $\mathscr{B}$ de chacun des éléments de la base duale $\mathscr{B}^*$. Remarquer que l'on obtient la base canonique de l'espace des matrices lignes.

1.2. Opérations élémentaires sur les colonnes (ou les lignes)

1.2.1. Réductions des colonnes. — Revenons au calcul du rang d'une application linéaire $k^n \to k^m$. Soit $A \in M_{m,n}(k)$, notons $(e_1, \dots, e_n)$ la base canonique de k^n. On peut calculer le rang de A et, plus précisément, une *base* de $\mathrm{Im}(A)$ et de $\mathrm{Ker}(A)$ de la façon suivante.

Observons qu'échanger les colonnes C_i et C_j de A revient à permuter, avant d'appliquer A, les vecteurs e_i et e_j de la base canonique de k^n, ce qui revient aussi à multiplier A *à droite* par la matrice $P(i,j)$ de l'automorphisme de k^n qui échange e_i et e_j, et laisse fixe chaque e_ℓ pour $\ell \neq i, j$. Par exemple, pour $n = 4$ et $(i,j) = (1,4)$ on a $P(1,4) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$.

De même, multiplier une colonne C_j par un scalaire $\lambda_j \neq 0$ revient à remplacer le vecteur e_j par $\lambda_j e_j$, c.-à-d., à multiplier A *à droite* par la matrice diagonale dont tous les termes diagonaux valent 1, sauf celui d'indice j qui vaut λ_j. D'autre part, pour tout $t \in k$ et $i \neq j$, ajouter tC_i à C_j revient à remplacer, avant d'appliquer A, le vecteur e_j par le vecteur $e_j + te_i$, et ceci revient encore à multiplier A *à droite* par la matrice $B_{ij}(t)$ de l'automorphisme de k^n qui laisse fixe e_ℓ pour $\ell \neq j$, et envoie e_j sur $e_j + te_i$, c.-à-d.,

$$B_{ij}(t) = I_n + tE_{ij}.$$

Définition 1.2.2. — On appellera *opérations élémentaires sur les colonnes* les opérations précédentes : échange de colonnes, multiplication d'une colonne par un scalaire $\neq 0$, ou ajout de tC_i à C_j avec $j \neq i$. D'après ce qui précède, on voit qu'effectuer ces opérations sur les colonnes revient à appliquer des automorphismes sur l'espace *de départ* k^n, donc ne change pas l'image de A, ni son rang.

On peut alors calculer $\mathrm{Im}(A)$ de la façon suivante. Soit i_1 l'indice de la première ligne non nulle. Alors en permutant les colonnes, on peut supposer que $a_{i_1,1} \neq 0$ puis, en multipliant la première colonne par $a_{i_1,1}^{-1}$ on se ramène au cas où $a_{i_1,1} = 1$. Ensuite, en soustrayant $a_{i_1,j}C_1$ de C_j, on se ramène au cas où $a_{i_1,j} = 0$ pour $j \geq 2$. Soit alors i_2 le plus petit indice tel qu'il existe $j \geq 2$ tel que $a_{i_2,j} \neq 0$; en procédant comme plus haut, on se ramène au cas où $a_{i_2,2} = 1$ et $a_{i_2,j} = 0$ pour $j \geq 3$. Soit alors i_3 le plus petit indice tel qu'il existe $j \geq 3$ tel que $a_{i_3,j} \neq 0$, en répétant le processus précédent, on obtient une matrice de la forme suivante:

$$(\dagger) \qquad A' = \begin{pmatrix}
0 & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \\
1 & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \\
a_{2,1} & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \\
a_{i_2,1} & 1 & 0 & 0 & \cdots & 0 & \cdots & 0 \\
* & * & 0 & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & * & 0 & 0 & \cdots & 0 & \cdots & 0 \\
a_{i_3,1} & a_{i_3,2} & 1 & 0 & \cdots & 0 & \cdots & 0 \\
* & * & * & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
a_{i_r,1} & a_{i_r,2} & a_{i_r,3} & \cdots & 1 & 0 & \cdots & 0 \\
* & * & * & \cdots & * & 0 & \cdots & 0 \\
* & * & * & \cdots & * & 0 & \cdots & 0
\end{pmatrix}$$

c.-à-d., les colonnes d'indice $> r$ sont nulles et, pour $j = 1, \dots, r$, les colonnes C'_j vérifient

$$C'_j = e_{i_j} + \sum_{\ell > i_j} a_{\ell,j} e_\ell,$$

avec $i_1 < i_2 < \cdots < i_r$.

Il en résulte que les vecteurs $C'_1, \dots, C'_r \in k^m$ sont linéairement indépendants, donc forment une base de $\mathrm{Im}(u)$; en particulier, $r = \mathrm{rang}(A)$.

Le procédé précédent s'appelle "*réduction des colonnes*", la matrice A' ainsi obtenue vérifie

$$A' = AP \tag{1}$$

où P est une matrice inversible correspondant aux opérations élémentaires sur les colonnes qu'on a effectuées. On a vu que $\mathrm{Im}(A') = \mathrm{Im}(A)$; d'autre part, le noyau de A' est le sous-espace V de k^n engendré par les vecteurs $e_{r+1}, \dots, e_n$ de la base canonique. Or, il résulte de (1) que

$$\mathrm{Ker}(A) = P(V) = \mathrm{Vect}(P_{r+1}, \dots, P_n), \tag{2}$$

où $P_{r+1}, \dots, P_n$ désignent les $(n-r)$ dernières colonnes de P. En effet, si $v \in V$, alors $0 = A'v = APv$ donc $Pv \in \mathrm{Ker}(A)$. Réciproquement, comme P est inversible, on a $A = A'P^{-1}$ donc si $x \in \mathrm{Ker}(A)$ alors $0 = Ax = A'P^{-1}x$, d'où $P^{-1}x \in \mathrm{Ker}(A') = V$, donc $x \in P(V)$.

L'égalité (2) permet de déterminer explicitement une *base* de $\mathrm{Ker}(A)$ de la façon suivante. Les opérations sur les colonnes qu'on a faites correspondent à multiplier A à droite par certaines matrices inversibles d'un type particulier, et P est le produit de ces matrices; alors la même suite d'opérations sur les colonnes de la matrice identité I_n conduit à la matrice P.

Donc, en pratique, on écrit en-dessous de A la matrice I_n, où n = nombre de colonnes de A = dimension de l'espace de départ, et à chaque étape on fait les mêmes opérations élémentaires sur les colonnes des deux matrices, on obtient alors à la fin du processus la matrice $A' = AP$ ainsi que la matrice $P = I_nP$. [(1)] Illustrons ceci dans le cas de l'exemple suivant:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 6 \\ 3 & 8 & 15 & 16 \end{pmatrix}.$$

Écrivons I_4 en-dessous de A et, notant $C_1, \dots, C_4$ les colonnes de A, faisons les opérations $C_i \to C_i - iC_1$ pour $i = 2, 3, 4$:

$$\left(\begin{array}{cccc} 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 6 \\ 3 & 8 & 15 & 16 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[C_4 \to C_4 - 4C_1]{\substack{C_2 \to C_2 - 2C_1 \\ C_3 \to C_3 - 3C_1}} \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 1 & 1 & 3 & 2 \\ 3 & 2 & 6 & 4 \\ \hline 1 & -2 & -3 & -4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[C_4 \to C_4 - 2C_2]{C_3 \to C_3 - 3C_2} \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ \hline 1 & -2 & 3 & 0 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

Donc $\mathrm{Im}(A)$ et $\mathrm{Ker}(A)$ ont respectivement pour base les vecteurs

$$\begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \text{ et } \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \qquad \text{resp.} \qquad \begin{pmatrix} 3 \\ -3 \\ 1 \\ 0 \end{pmatrix} \text{ et } \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}.$$

En résumé, la "réduction des colonnes" d'une matrice $A \in M_{m,n}(k)$ fournit des *bases* de $\mathrm{Im}(A)$ et de $\mathrm{Ker}(A)$. Mais, étant donné un vecteur $Y = (y_1, \dots, y_m)$ de k^m (resp. $X = (x_1, \dots, x_n)$ de k^n), il n'est pas forcément immédiat de déterminer si $Y \in \mathrm{Im}(A)$ (resp. si $X \in \mathrm{Ker}(A)$), et il est parfois plus commode de disposer *d'équations* définissant $\mathrm{Im}(A)$ et $\mathrm{Ker}(A)$. On va voir que celles-ci sont obtenues par *réduction des lignes* de A.

1.2.3. Remarque importante. — Dans l'exemple précédent, le premier pivot était en position (1,1), puis le second en position (2,2). Bien sûr, il n'en est pas toujours ainsi; en théorie on peut s'y ramener par des échanges de colonnes, mais avec un peu de pratique on n'a pas besoin de faire ces échanges de colonnes: il suffit de se ramener à une matrice A' dont les colonnes soient échelonnées "à permutation près des colonnes"; alors les colonnes non nulles de A' forment une base de $\mathrm{Im}(A)$, et dans la matrice du dessous

[(1)] Lorsque $m = n$, c.-à-d., lorsqu'on considère des matrices carrées, on peut utiliser ce processus pour déterminer si A est inversible et calculer son inverse; voir 1.2.8 plus bas.

(obtenue à partir de I_n), les colonnes en-dessous des colonnes nulles de A' forment une base de $\mathrm{Ker}(A)$. Illustrons ceci dans le cas suivant:

$$\begin{pmatrix} 3 & 2 & 4 & 1 \\ 6 & 3 & 6 & 1 \\ 15 & 8 & 16 & 3 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[C_1 \to C_1 - 3C_4]{\substack{C_3 \to C_3 - 4C_4 \\ C_2 \to C_2 - 2C_4}} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 3 & 1 & 2 & 1 \\ 6 & 2 & 4 & 3 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -3 & -2 & -4 & 1 \end{pmatrix} \xrightarrow[C_3 \to C_3 - 2C_2]{C_1 \to C_1 - 3C_2} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & 3 \\ \hline 1 & 0 & 0 & 0 \\ -3 & 1 & -2 & 0 \\ 0 & 0 & 1 & 0 \\ 3 & -2 & 0 & 1 \end{pmatrix}.$$

Donc $\mathrm{Im}(A)$ et $\mathrm{Ker}(A)$ ont respectivement pour base les vecteurs

$$\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \text{ et } \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \qquad \text{resp.} \qquad \begin{pmatrix} 1 \\ -3 \\ 0 \\ 3 \end{pmatrix} \text{ et } \begin{pmatrix} 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}.$$

1.2.4. Réduction des lignes. — Soit toujours $A \in M_{m,n}(k)$. On peut calculer le rang de A et, plus précisément, des *équations* de $\mathrm{Ker}(A)$ et de $\mathrm{Im}(A)$ en faisant des opérations sur les *lignes* de A.

Observons que multiplier une ligne L_i de A par un scalaire $\lambda_i \neq 0$ revient à multiplier A *à gauche* par la matrice diagonale inversible dont tous les termes diagonaux valent 1, sauf celui d'indice i qui vaut λ_i. De même, pour tout $t \in k$ et $i \neq j$, ajouter tL_i à L_j revient à multiplier A *à gauche* par la matrice inversible

$$B_{ji}(t) = I_m + tE_{ji}$$

(dont l'inverse est $B_{ji}(-t)$). Enfin, échanger les lignes L_i et L_j de A revient à multiplier A *à gauche* par la matrice de l'automorphisme de k^m qui échange les vecteurs f_i et f_j de la base canonique $(f_1, \dots, f_m)$, et laisse fixe chaque f_ℓ pour $\ell \neq i, j$ (cette matrice est égale à son inverse).

Définition 1.2.5. — On appellera *opérations élémentaires sur les lignes* les opérations précédentes : échange de lignes, multiplication d'une ligne par un scalaire $\neq 0$, ou ajout de tL_i à L_j avec $j \neq i$. D'après ce qui précède, on voit qu'effectuer ces opérations revient à appliquer des automorphismes sur l'espace *d'arrivée* k^m, donc ne change pas le noyau de A, ni son rang.

On peut alors calculer des équations de $\mathrm{Ker}(A)$ et $\mathrm{Im}(A)$ de la façon suivante. Soit j_1 l'indice de la première colonne non nulle. En permutant les lignes, on peut supposer que $a_{1,j_1} \neq 0$ puis, en multipliant la première ligne par a_{1,j_1}^{-1}, on se ramène au cas où $a_{1,j_1} = 1$. Ensuite, en soustrayant $a_{i,j_1}L_1$ de L_i pour tout $i \geq 2$, on se ramène au cas où $a_{i,j_1} = 0$ pour $i \geq 2$. Soit alors j_2 le plus petit indice tel qu'il existe $i \geq 2$ tel que $a_{i,j_2} \neq 0$; en procédant comme plus haut, on se ramène au cas où $a_{2,j_2} = 1$ et $a_{i,j_2} = 0$ pour $i \geq 3$. En répétant ce processus, on obtient une matrice de la forme suivante:

$$(\ddagger) \qquad A'' = \begin{pmatrix} 0 & 1 & * & a_{1,j_2} & * & a_{1,j_3} & \cdots & a_{1,j_r} & * & * \\ 0 & 0 & 0 & 1 & * & a_{2,j_3} & \cdots & a_{2,j_r} & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & a_{3,j_r} & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}$$

On a $A'' = QA$, où $Q \in \mathrm{GL}_m(k)$ est inversible, donc $\mathrm{Ker}(A) = \mathrm{Ker}(A'')$, a fortiori $\mathrm{rg}(A) = \mathrm{rg}(A'') = r$. De plus, A'' donne directement les *équations* de $\mathrm{Ker}(A'')$. En effet, considérant le système

$$A'' \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0$$

on voit qu'on peut choisir arbitrairement x_i pour $i \notin \{j_1, \dots, j_r\}$, et que chaque x_{j_ℓ} s'exprime en fonction des x_i pour $i > j_\ell$.

D'autre part, $\mathrm{Im}(A'')$ est le sous-espace W de k^m engendré par les vecteurs $f_1, \dots, f_r$; on en déduit que $\mathrm{Im}(A) = Q^{-1}(W)$. Mais il n'est pas nécessaire d'inverser la matrice Q pour avoir les *équations* de $\mathrm{Im}(A)$.

En effet, si l'on écrit côte à côte A et un vecteur colonne

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \in k^m$$

et qu'on applique à A et Y les mêmes opérations sur les lignes, on obtient le couple

$$(A'' = QA \mid Y'' = QY)\,;$$

alors $Y \in \mathrm{Im}(A)$ si et seulement si $Y'' \in \mathrm{Im}(A'')$ et, puisque $\mathrm{Im}(A'') = \mathrm{Vect}(f_1, \dots, f_r)$, on a $Y'' \in \mathrm{Im}(A'')$ si et seulement si les $(n-r)$ dernières coordonnées $y''_{r+1}, \dots, y''_n$ de Y'' sont nulles. Donc $\mathrm{Im}(A)$ *est déterminée par les équations* $y''_{r+1} = 0,\ \dots,\ y''_n = 0$.

Illustrons ceci sur le premier exemple considéré au paragraphe précédent, c.-à-d.,

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 6 \\ 3 & 8 & 15 & 16 \end{pmatrix}.$$

Appliquons la *réduction des lignes* de A au couple:

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & y_1 \\ 1 & 3 & 6 & 6 & y_2 \\ 3 & 8 & 15 & 16 & y_3 \end{array}\right).$$

D'abord, $L_2 \to L_2 - L_1$ et $L_3 \to L_3 - 3L_1$ donnent:

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & y_1 \\ 0 & 1 & 3 & 2 & y_2 - y_1 \\ 0 & 2 & 6 & 4 & y_3 - 3y_1 \end{array}\right)$$

puis $L_3 \to L_3 - 2L_2$ donne:

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & y_1 \\ 0 & 1 & 3 & 2 & y_2 - y_1 \\ 0 & 0 & 0 & 0 & y_3 - 3y_1 - 2(y_2 - y_1) = y_3 - y_1 - 2y_2 \end{array}\right).$$

On obtient ainsi que des équations de $\mathrm{Ker}(A)$ sont

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 &= 0 \\ x_2 + 3x_3 + 2x_4 &= 0; \end{cases}$$

ces équations expriment x_2 puis x_1 en fonction de x_3 et x_4, qu'on peut choisir arbitrairement. En choisissant $x_3 = 1$ et $x_4 = 0$ (resp. $x_3 = 0$ et $x_4 = 1$), on obtient les vecteurs

$$u = \begin{pmatrix} 3 \\ -3 \\ 1 \\ 0 \end{pmatrix} \qquad \text{resp.} \qquad v = \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}$$

qui forment donc une base de $\mathrm{Ker}(A)$.

D'autre part, comme Q est inversible, les solutions de l'équation $AX = Y$ sont les mêmes que celles de l'équation $QAX = QY$, c.-à-d., $A''X = Y''$. Or, le système

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 &= y_1 \\ x_2 + 3x_3 + 2x_4 &= y_2 - y_1 \\ 0 &= y_3 - y_1 - 2y_2 \end{cases}$$

admet des solutions si et seulement si

$$(*) \qquad y_3 = y_1 + 2y_2.$$

On voit donc ainsi que $(*)$ est une équation de $\mathrm{Im}(A)$. Par exemple, prenant $y_1 = 1$ et $y_2 = 0$ (resp. $y_1 = 0$ et $y_2 = 1$), on obtient que les vecteurs

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \qquad \text{et} \qquad \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

forment une base de $\mathrm{Im}(A)$.

1.2.6. Remarque importante. — Dans l'exemple précédent, le premier pivot était en position (1,1), puis le second en position (2,2). Bien sûr, il n'en est pas toujours ainsi; en théorie on peut s'y ramener par des échanges de lignes, mais en pratique on n'a pas besoin de faire ces échanges de lignes: il suffit de se ramener à une matrice A' dont les lignes soient échelonnées "à permutation près des lignes"; alors les lignes non nulles de A' donnent des équations de $\mathrm{Ker}(A)$, et les formes linéaires en face des lignes nulles de A' donnent des équations de $\mathrm{Im}(A)$. Illustrons ceci dans le cas suivant:

$$\left(\begin{array}{cccc|c} 3 & 15 & 8 & 16 & y_1 \\ 1 & 6 & 3 & 6 & y_2 \\ 1 & 3 & 2 & 4 & y_3 \end{array}\right) \xrightarrow[L_2\to L_2-L_3]{L_1\to L_1-3L_3} \left(\begin{array}{cccc|c} 0 & 6 & 2 & 4 & y_1-3y_3 \\ 0 & 3 & 1 & 2 & y_2-y_3 \\ 1 & 3 & 2 & 4 & y_3 \end{array}\right)$$

$$\xrightarrow{L_1\to L_1-2L_2} \left(\begin{array}{cccc|c} 0 & 0 & 0 & 0 & y_1-2y_2-y_3 \\ 0 & 3 & 1 & 2 & y_2-y_3 \\ 1 & 3 & 2 & 4 & y_3 \end{array}\right).$$

Donc des équations de $\mathrm{Ker}(A)$ sont

$$\begin{cases} 3x_2+x_3+2x_4 &= 0 \\ x_1+3x_2+2x_3+4x_4 &= 0 \end{cases}$$

et une équation de $\mathrm{Im}(A)$ est $y_1-2y_2-y_3=0$.

1.2.7. Lien avec les systèmes linéaires. — Ce qui précède équivaut à la théorie des systèmes linéaires vue en L1. En effet, notant X le vecteur colonne $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, on associe à la matrice $A\in M_{m,n}(k)$ le système linéaire $AX=0$, c.-à-d., les m équations

$$(L_i) \qquad a_{i1}x_1+\cdots+a_{in}x_n=0$$

données par les lignes $L_1,\dots,L_m$ de A. Chaque L_i peut aussi être vue comme une forme linéaire sur $V=\mathbb{R}^n$, i.e. la forme linéaire qui à tout $v=x_1e_1+\cdots+x_ne_n$ (où $\mathscr{B}=(e_1,\dots,e_n)$ est la base canonique de $\mathbb{R}^n$) associe le scalaire $L_i(v)=a_{i1}x_1+\cdots+a_{in}x_n$.

Soit s le rang de ce système, c.-à-d., le nombre maximum de lignes de A qui sont linéairement indépendantes (donc $s=\mathrm{rang}({}^tA)$). Le paragraphe précédent montre, en faisant des opérations élémentaires sur les lignes de A, que ce système a les mêmes solutions que le système échelonné $A''X=0$, où A'' est la matrice à lignes échelonnées obtenue en (‡) plus haut: si $j_1<\cdots<j_s$ désignent les colonnes où se trouvent les pivots sur les lignes $1,\dots,s$ de A'', on peut choisir arbitrairement x_i pour $i\not\in\{j_1,\dots,j_s\}$, et chaque x_{j_ℓ} s'exprime en fonction des x_i pour $i>j_\ell$. Donc l'espace des solutions de ce système, qui n'est autre que $\mathrm{Ker}(A)$, est de dimension égale à $n-s$. On retrouve ainsi la théorie des systèmes linéaires, vue en L1. D'autre part, comme $\dim\mathrm{Ker}(A)=n-\mathrm{rg}(A)$ d'après le théorème du rang, on obtient le corollaire suivant (dont une autre démonstration a déjà été donnée en 0.5.10):

Corollaire 1.2.7.1. — *Pour tout $A\in M_{m,n}(k)$, on a $\mathrm{rg}({}^tA)=\mathrm{rg}(A)$.*

Remarquons d'autre part que faire des opérations sur les lignes de A revient à faire des opérations sur les colonnes de ${}^tA$ ...

Remarque 1.2.7.2. — Notons $\mathscr{B}=(e_1,\dots,e_n)$ (resp. $\mathscr{C}=(f_1,\dots,f_m)$) la base canonique de $V=\mathbb{R}^n$ (resp. de $W=\mathbb{R}^m$) et $\mathscr{B}^*=(e_1^*,\dots,e_n^*)$ (resp. $\mathscr{C}^*=(f_1^*,\dots,f_m^*)$) la base duale de V^* (resp. de W^*). Alors les formes linéaires $L_1,\dots,L_m$ données par les lignes de A ont les deux interprétations suivantes. Fixons un indice de ligne $i\in\{1,\dots,m\}$. D'une part, pour tout $j\in\{1,\dots,n\}$, on a $L_i(e_j)=a_{ij}$ et donc

$$L_i=a_{i1}e_1^*+\cdots+a_{in}e_n^*.$$

D'autre part, $L_i(e_j)$ est la i-ième coordonnée dans la base $\mathscr{C}$ du vecteur Ae_j, c.-à-d., $L_i(e_j)=f_i^*(Ae_j)=(f_i^*\circ A)(e_j)$. Comme les formes linéaires L_i et $f_i^*\circ A$ coïncident sur la base $\mathscr{B}=(e_1,\dots,e_n)$ de $\mathbb{R}^n$, elles sont égales, et l'on a donc

$$(\star) \qquad \forall i=1,\dots,m, \qquad f_i^*\circ A=L_i=a_{i1}e_1^*+\cdots+a_{in}e_n^*.$$

Si l'on note u l'application linéaire $V=k^n\to k^m=W$ définie par A, alors l'application $W^*\to V^*$ qui à tout $f\in W^*$ associe $f\circ u\in V^*$ est linéaire: on dit que c'est la *transposée* de u et on la note ${}^tu$; il résulte de $(\star)$ ci-dessus que l'on a:

$$\mathrm{Mat}_{\mathscr{B}^*,\mathscr{C}^*}({}^tu)={}^tA$$

(les lecteurs intéressés pourront consulter l'appendice 1.5 à la fin de ce chapitre).

1.2.8. Calcul de l'inverse d'une matrice carrée. — Soit $A \in M_n(k)$. On peut utiliser l'algorithme de réduction des colonnes pour déterminer si A est inversible et calculer son inverse, de la façon suivante. Notons $(e_1, \ldots, e_n)$ la base canonique de k^n.

Si la première ligne de A est nulle, alors évidemment A n'est pas inversible, car alors l'image de A est contenue dans le sous-espace engendré par $e_2, \ldots, e_n$. On peut donc supposer que la première ligne de A est non nulle; alors la première étape de l'algorithme de réduction des colonnes nous fournit une matrice $A_1 = AP_1$ dont la première ligne est:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}$$

Soit $t \in \{2, \ldots, n\}$ et supposons qu'après $t-1$ étapes on ait obtenu une matrice

$$A_{t-1} = AP_1 \cdots P_{t-1}$$

dont les $t-1$ premières lignes sont de la forme

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ * & 1 & 0 & 0 & \cdots & 0 \\ * & * & 1 & 0 & \cdots & 0 \end{pmatrix};$$

si la t-ème ligne a tous ses coefficients d'indice $\geq t$ nuls, alors $\mathrm{Im}(A)$ est contenue dans le sous-espace engendré par les colonnes $A_1, \ldots, A_{t-1}$ et par les vecteurs $e_{t+1}, \ldots, e_n$ de la base canonique, donc A n'est pas inversible.

On obtient donc l'alternative suivante: ou bien au cours du processus on obtient une matrice de la forme

$$A_t = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & \cdots & 0 \\ * & 1 & 0 & 0 & \cdots & 0 \\ * & * & 1 & 0 & \cdots & 0 \\ \hline * & * & * & 0 & \cdots & 0 \\ * & * & * & * & \cdots & * \\ * & * & * & * & \cdots & * \end{array}\right)$$

auquel cas A n'est pas inversible (d'après ce qui précède), ou bien on arrive à une matrice triangulaire inférieure avec des 1 sur la diagonale:

$$A_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ * & 1 & 0 & \cdots & 0 \\ * & * & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & 1 \end{pmatrix};$$

alors soustrayant $a_{nj}C_n$ à la j-ème colonne, pour $j = 1, \ldots, n-1$, on obtient une matrice

$$A_{n+1} = A_nP_{n+1} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ * & 1 & 0 & \cdots & 0 \\ * & * & 1 & \ddots & 0 \\ a_{n-1,1} & a_{n-1,2} & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

retranchant alors $a_{n-1,j}C_{n-1}$ à la j-ème colonne, pour $j = 1, \ldots, n-2$, on obtient une matrice

$$A_{n+2} = A_nP_{n+1}P_{n+2} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ * & 1 & 0 & \cdots & 0 \\ * & * & 1 & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

etc. En continuant ainsi, on arrive à une forme réduite $A' = AP$ qui égale la matrice identité I_n; alors, d'après la proposition 0.5.2, on a $P = A^{-1}$. De plus, comme indiqué dans le paragraphe 1.2.1, la matrice P est obtenue en écrivant au début du processus la matrice I_n en-dessous de A, et en effectuant à chaque étape les *mêmes* opérations élémentaires sur les colonnes des deux matrices; on obtient ainsi à la fin du processus, en haut la matrice réduite $A' = AP = I_n$, et en bas la matrice $I_nP = P = A^{-1}$.

Illustrons ceci par un exemple: soit $A = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 3 \\ 1 & 2 & 0 \end{pmatrix} \in M_3(\mathbb{Q})$; on a:

$$\left(\begin{array}{ccc} 2 & -1 & 1 \\ 1 & 0 & 3 \\ 1 & 2 & 0 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \xrightarrow[C_3 \to C_3 - C_1/2]{\substack{C_1 \to C_1/2 \\ C_2 \to C_2 + C_1/2}} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 1/2 & 1/2 & 5/2 \\ 1/2 & 5/2 & -1/2 \\ \hline 1/2 & 1/2 & -1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \xrightarrow[C_3 \to C_3 - 5C_2]{C_2 \to 2C_2} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/2 & 5 & -13 \\ \hline 1/2 & 1 & -3 \\ 0 & 2 & -5 \\ 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[C_3 \to -C_3/13]{\substack{C_1 \to C_1 + C_3/26 \\ C_2 \to C_2 + 5C_3/13}} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 5/13 & -2/13 & 3/13 \\ -5/26 & 1/13 & 5/13 \\ 1/26 & 5/13 & -1/13 \end{array}\right) \xrightarrow{C_1 \to C_1 - C_2/2} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 6/13 & -2/13 & 3/13 \\ -3/13 & 1/13 & 5/13 \\ -2/13 & 5/13 & -1/13 \end{array}\right).$$

Notons que, au lieu de la suite d'opérations sur les colonnes utilisée plus haut, on peut choisir n'importe quelle suite d'opérations sur les colonnes de façon à obtenir des calculs les plus simples possibles. Ainsi, dans l'exemple considéré, il apparaît plus avantageux de procéder comme suit:

$$\left(\begin{array}{ccc} 2 & -1 & 1 \\ 1 & 0 & 3 \\ 1 & 2 & 0 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \xrightarrow[C_3 \to C_3 + C_2]{\substack{C_1 \to -C_2 \\ C_2 \to C_1 + 2C_2}} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 3 \\ -2 & 5 & 2 \\ \hline 0 & 1 & 0 \\ -1 & 2 & 1 \\ 0 & 0 & 1 \end{array}\right) \xrightarrow{C_3 \to C_3 - 3C_2} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 5 & -13 \\ \hline 0 & 1 & -3 \\ -1 & 2 & -5 \\ 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[C_3 \to -C_3/13]{\substack{C_1 \to C_1 - 2C_3/13 \\ C_2 \to C_2 + 5C_3/13}} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 6/13 & -2/13 & 3/13 \\ -3/13 & 1/13 & 5/13 \\ -2/13 & 5/13 & -1/13 \end{array}\right).$$

Bien sûr, on peut aussi procéder par réduction des lignes: partons de $A \in M_n(k)$, si la première colonne de A est nulle, alors évidemment A n'est pas inversible. On peut donc supposer que la première colonne de A est non nulle; alors la première étape de l'algorithme de réduction des lignes nous fournit une matrice $A_1 = Q_1 A$ dont la première colonne est:

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Soit $t \in \{2, \ldots, n\}$ et supposons qu'après $t-1$ étapes on ait obtenu une matrice

$$A_{t-1} = Q_{t-1} \cdots Q_1 A$$

dont les $t-1$ premières colonnes sont de la forme

$$\begin{pmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \\ \vdots & \vdots & 0 \\ 0 & 0 & 0 \end{pmatrix};$$

si la t-ième colonne a tous ses coefficients d'indice $\geq t$ nuls, alors $\mathrm{Im}(A)$ est contenue dans le sous-espace engendré par les vecteurs $e_1, \ldots, e_{t-1}$ de la base canonique et par les colonnes $A_{t+1}, \ldots, A_n$, donc A n'est pas inversible.

On obtient donc l'alternative suivante: ou bien au cours du processus on obtient une matrice de la forme

$$A_t = \left(\begin{array}{ccc|ccc} 1 & * & * & * & \cdots & * \\ 0 & 1 & * & * & \cdots & * \\ 0 & 0 & 1 & * & \cdots & * \\ \hline 0 & 0 & 0 & 0 & * & * \\ \vdots & \vdots & \vdots & 0 & \vdots & * \\ 0 & 0 & 0 & 0 & * & * \end{array}\right)$$

auquel cas A n'est pas inversible (d'après ce qui précède), ou bien on arrive à une matrice triangulaire supérieure avec des 1 sur la diagonale:

$$A_n = \begin{pmatrix} 1 & * & * & \cdots & a_{1n} \\ 0 & 1 & * & \cdots & a_{2n} \\ 0 & 0 & 1 & * & * \\ \vdots & \ddots & \ddots & \ddots & a_{n-1,n} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix};$$

alors soustrayant $a_{in}L_n$ à la i-ième ligne, pour $i = 1, \dots, n-1$, on obtient une matrice

$$A_{n+1} = Q_{n+1}A_n = \begin{pmatrix} 1 & * & * & a_{1,n-1} & 0 \\ 0 & 1 & * & a_{2,n-1} & 0 \\ 0 & 0 & 1 & \vdots & 0 \\ 0 & \ddots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix};$$

retranchant alors $a_{i,n-1}L_{n-1}$ à la i-ième ligne, pour $i = 1, \dots, n-2$, on obtient une matrice

$$A_{n+2} = Q_{n+2}Q_{n+1}A_n = \begin{pmatrix} 1 & * & * & 0 & 0 \\ 0 & 1 & * & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \ddots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

etc. En continuant ainsi, on arrive à une forme réduite $A' = QA$ qui égale la matrice identité I_n; alors, d'après la proposition 0.5.2, on a $Q = A^{-1}$. De plus, la matrice Q s'obtient en écrivant au début du processus les matrices A et I_n côte à côte, et en effectuant à chaque étape les *mêmes* opérations élémentaires sur les lignes des deux matrices; on obtient ainsi à la fin du processus, d'un côté la matrice réduite $A' = QA = I_n$, et de l'autre la matrice $QI_n = Q = A^{-1}$.

Illustrons ceci en reprenant l'exemple précédent:

$$\left(\begin{array}{ccc|ccc} 2 & -1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[L_3 \to L_3 - L_1/2]{\substack{L_1 \to L_1/2 \\ L_2 \to L_2 - L_1/2}} \left(\begin{array}{ccc|ccc} 1 & -1/2 & 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 5/2 & -1/2 & 1 & 0 \\ 0 & 5/2 & -1/2 & -1/2 & 0 & 1 \end{array}\right) \xrightarrow[L_3 \to L_3 - 5L_2]{L_2 \to 2L_2}$$

$$\left(\begin{array}{ccc|ccc} 1 & -1/2 & 1/2 & 1/2 & 0 & 0 \\ 0 & 1 & 5 & -1 & 2 & 0 \\ 0 & 0 & -13 & 2 & -5 & 1 \end{array}\right) \xrightarrow[L_3 \to -L_3/13]{\substack{L_1 \to L_1 + L_3/26 \\ L_2 \to L_2 + 5L_3/13}} \left(\begin{array}{ccc|ccc} 1 & -1/2 & 0 & 15/26 & -5/26 & 1/26 \\ 0 & 1 & 0 & -3/13 & 1/13 & 5/13 \\ 0 & 0 & 1 & -2/13 & 5/13 & -1/13 \end{array}\right)$$

$$\xrightarrow{L_1 \to L_1 + L_2/2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 6/13 & -2/13 & 3/13 \\ 0 & 1 & 0 & -3/13 & 1/13 & 5/13 \\ 0 & 0 & 1 & -2/13 & 5/13 & -1/13 \end{array}\right).$$

Remarquons à nouveau que, au lieu de la suite d'opérations sur les lignes utilisée plus haut, on peut choisir n'importe quelle suite d'opérations sur les lignes de façon à obtenir des calculs les plus simples possibles. Ainsi, dans cet exemple, il apparaît plus avantageux de procéder comme suit:

$$\left(\begin{array}{ccc|ccc} 2 & -1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[L_3 \to L_3 - L_2]{\substack{L_1 \to L_2 \\ L_2 \to L_1 - 2L_2}} \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & -1 & -5 & 1 & -2 & 0 \\ 0 & 2 & -3 & 0 & -1 & 1 \end{array}\right) \xrightarrow[L_3 \to L_3 + 2L_2]{L_2 \to -L_2}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 5 & -1 & 2 & 0 \\ 0 & 0 & -13 & 2 & -5 & 1 \end{array}\right) \xrightarrow[L_3 \to -L_3/13]{\substack{L_1 \to L_1 + 3L_3/13 \\ L_2 \to L_2 + 5L_3/13}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 6/13 & -2/13 & 3/13 \\ 0 & 1 & 0 & -3/13 & 1/13 & 5/13 \\ 0 & 0 & 1 & -2/13 & 5/13 & -1/13 \end{array}\right).$$

Remarque 1.2.8.1. — Attention, dans cet algorithme pour calculer A^{-1} il faut choisir de faire des opérations sur les colonnes *ou bien* sur les lignes, mais il ne faut *pas mélanger les deux*!

En effet, si l'on fait à la fois des opérations sur les lignes et sur les colonnes de A, et qu'on fait les mêmes opérations sur la matrice I_n, on arrive au bout du processus à une paire de matrices:

$$(QAP = I_n, \quad QI_nP = QP)\,;$$

alors la première égalité donne $A = Q^{-1}P^{-1}$ d'où $A^{-1} = PQ$, mais de l'autre côté c'est QP que l'on a calculé...

1.3. Codimension et équations d'un sous-espace

Dans ce paragraphe, nous introduisons la notion de codimension d'un sous-espace vectoriel, puis nous justifions l'idée que *la codimension d'un sous-espace est le nombre d'équations nécessaire pour le décrire.*

***Définition 1.3.1* (Codimension)**. — Soient V un k-espace vectoriel de dimension n et E un sous-espace vectoriel de V, on appelle *codimension* de E dans V et l'on note $\operatorname{codim}_V(E)$ l'entier $\dim V - \dim E$.

Soit $\mathscr{B} = (e_1, \dots, e_n)$ une base de V. Notons $(x_1, \dots, x_n)$ les "*fonctions coordonnées*" par rapport à $\mathscr{B}$, i.e. tout $v \in V$ s'écrit de façon unique

$$v = x_1(v)e_1 + \cdots + x_n(v)e_n.$$

Si $\mathscr{B}^* = (e_1^*, \dots, e_n^*)$ est la base duale de V^*, alors pour tout $v \in V$ on a $e_j^*(v) = x_j(v)$, donc la fonction $x_j : V \to k$, $v \mapsto x_j(v)$ n'est autre que la forme linéaire e_j^*. Considérons maintenant des "équations linéaires"

$$(L_i) \qquad a_{i1}x_1 + \cdots + a_{in}x_n = 0$$

pour $i = 1, \dots, m$ et soit E le sous-espace de V défini par ces équations, i.e. E est l'ensemble des vecteurs $v = x_1e_1 + \cdots + x_ne_n$ dont les coordonnées $(x_1, \dots, x_n)$ vérifient les équations précédentes. Alors E est le noyau de la matrice $A \in M_{m,n}(k)$ dont les lignes sont $L_1, \dots, L_m$, donc $\dim E = n - s$ où s est le rang du système, i.e. le nombre maximum d'équations linéairement indépendantes.

Comme dans le paragraphe 1.2.7, ces équations peuvent aussi être vues comme des formes linéaires sur V, i.e. chaque L_i est la forme linéaire $a_{i1}e_1^* + \cdots + a_{in}e_n^*$; alors E est l'ensemble des $v \in V$ tels que $L_i(v) = 0$ pour $i = 1, \dots, m$. Quitte à renuméroter les L_i, on peut supposer que $L_1, \dots, L_s$ sont linéairement indépendantes; alors pour $i > s$ chaque L_i est combinaison linéaire de $L_1, \dots, L_s$ donc l'égalité $L_i(v) = 0$ est conséquence des égalités $L_t(v) = 0$ pour $t = 1, \dots, s$, et donc E est le sous-espace de V défini par les équations $L_t(v) = 0$ pour $t = 1, \dots, s$, et $\dim E = n - s$. Pour résumer, on a obtenu la

Proposition 1.3.2. — *Soit V un espace vectoriel de dimension n et soient $L_1, \dots, L_m \in V^*$. Le sous-espace E de V défini par ces équations (i.e. $E = \{v \in V \mid L_t(v) = 0,\ \forall t = 1, \dots, m\}$) est de dimension $n - s$, où $s = \dim \operatorname{Vect}(L_1, \dots, L_m)$. En particulier, lorsque $L_1, \dots, L_m$ sont linéairement indépendantes, on a $\dim E = n - m$.*

Remarque 1.3.3. — Remarquons que E ne dépend que du sous-espace $F = \operatorname{Vect}(L_1, \dots, L_m)$ de V^* engendré par les L_t. En effet, d'une part, tout $\phi \in F$ s'annule sur E (car ϕ est combinaison linéaire de $L_1, \dots, L_m$). D'autre part, si l'on pose

$$F^\circ = \{v \in V \mid \phi(v) = 0, \quad \forall \phi \in F\}$$

on a $E \subseteq F^\circ$ d'après ce qui précède, et $F^\circ \subseteq \{v \in V \mid L_t(v) = 0,\ \forall t = 1, \dots, m\} = E$, d'où $E = F^\circ$. On appelle F° *l'orthogonal* de F dans V; alors la proposition précédente peut se récrire en disant que $\dim F^\circ = n - \dim F$.

Réciproquement, si E est un sous-espace vectoriel de V de dimension r, il peut être défini par $n - r$ équations linéairement indépendantes. En effet, soit $(e_1, \dots, e_r)$ une base de E; complétons-la en une base $\mathscr{B} = (e_1, \dots, e_n)$ de V et soit $\mathscr{B}^* = (e_1^*, \dots, e_n^*)$ la base duale de V^*. Alors E est défini par les formes linéaires $e_{r+1}^*, \dots, e_n^*$, c.-à-d., par les équations $x_i = 0$ pour $i = r + 1, \dots, n$. De plus, une forme linéaire $\phi = t_1e_1^* + \cdots + t_ne_n^*$ s'annule sur E si et seulement si $t_1 = 0 = \cdots = t_r$, donc, si l'on définit *l'orthogonal* $E^\perp$ de E dans V^* par:

$$E^\perp = \{\phi \in V^* \mid \phi(x) = 0 \quad \text{pour tout } x \in E\},$$

alors $E^\perp$ égale $\operatorname{Vect}(e_{r+1}^*, \dots, e_n^*)$ donc est de dimension $n-r$. Donc, si $L_1, \dots, L_m$ sont des formes linéaires qui définissent E, c.-à-d., telles que

$$E = \{v \in V \mid L_i(v) = 0, \quad \forall i = 1, \dots, m\}$$

alors $L_1, \ldots, L_m \in E^\perp$, et d'après la proposition précédente, on a:

$$n - \dim \mathrm{Vect}(L_1, \ldots, L_m) = \dim E = r$$

et donc $\dim \mathrm{Vect}(L_1, \ldots, L_m) = n - r = \dim E^\perp$, donc $(L_1, \ldots, L_m)$ est une famille génératrice de $E^\perp$ et l'on peut en extraire un système de $n - r$ équations qui définissent E. Pour résumer, on a obtenu la

Proposition 1.3.4. — *Soient V un espace vectoriel de dimension n et E un sous-espace vectoriel de dimension r. Alors E peut être défini par $n - r$ équations linéairement indépendantes; de plus de tout système d'équations $L_1, \ldots, L_m$ définissant E, on peut extraire $n-r$ équations linéairement indépendantes définissant E (en particulier, on a $m \geq n - r$).*

Exemples 1.3.5. — 1) Dans $\mathbb{R}^2$, la droite D engendrée par le vecteur $\binom{a}{b} \neq 0$ est de codimension $2-1 = 1$. L'orthogonal $D^\perp \subseteq (\mathbb{R}^2)^*$ est engendré par la forme linéaire $\phi = be_1^* - ae_2^* \neq 0$, i.e. D est définie par l'équation $bx - ay = 0$. Cette équation est unique à un scalaire près, i.e. toute autre équation de D est de la forme $\lambda(bx - ay) = 0$, avec $\lambda \in \mathbb{R}^*$.

2) Dans $\mathbb{R}^3$, le plan P engendré par les vecteurs $e_1 - e_2$ et $e_2 - e_3$ est de codimension $3 - 2 = 1$. Son orthogonal $P^\perp \subseteq (\mathbb{R}^3)^*$ est engendré par la forme linéaire $\phi = e_1^* + e_2^* + e_3^* \neq 0$, i.e. P est défini par l'équation $x + y + z = 0$. Cette équation est unique à un scalaire près, i.e. toute autre équation de P est de la forme $\lambda(x + y + z) = 0$, avec $\lambda \in \mathbb{R}^*$.

3) Dans $\mathbb{R}^3$, la droite D engendrée par le vecteur $ae_1 + be_2 - e_3$, avec $(a, b) \neq (0, 0)$, est de codimension $3 - 1 = 2$, donc son orthogonal $D^\perp \subseteq (\mathbb{R}^3)^*$ est de dimension 2, il admet pour base, par exemple, $(e_1^* + ae_3^*, e_2^* + be_3^*)$, i.e. D est définie par les équations $x + az = 0$ et $y + bz = 0$. Une autre base de $D^\perp$ est, par exemple, $(be_1^* - ae_2^*, e_1^* + e_2^* + (a + b)e_3^*)$, donc D est aussi définie par les équations $bx - ay = 0$ et $x + y + (a + b)z = 0$.

Remarques 1.3.6. — Soient V un espace vectoriel de dimension n et V^* son dual.

(1) On voit facilement que pour tout sous-espace E de V (resp. F de V^*), on a $E \subseteq (E^\perp)^\circ$ (resp. $F \subseteq (F^\circ)^\perp$), et il résulte des deux propositions précédentes que $\dim(E^\perp)^\circ = n - \dim E^\perp = \dim E$ et $\dim(F^\circ)^\perp = n - F^\circ = \dim F$. On a donc les égalités: $E = (E^\perp)^\circ$ et $F = (F^\circ)^\perp$.

(2) On a introduit temporairement la notation F° pour l'orthogonal de F dans V, pour le distinguer a priori de son orthogonal $F^\perp$ dans le dual $V^{**} = (V^*)^*$ de V^*. Mais en fait on a un isomorphisme canonique $V = V^{**}$, i.e. V est le dual de son dual (voir l'appendice 1.6), et les deux notions d'orthogonalité coïncident, donc pour un sous-espace F de V^* on peut aussi noter $F^\perp$ son orthogonal dans V...

1.4. Conclusion: notion de dualité

La notion de "dualité" est apparue vers 1820 dans les travaux de Gergonne et Poncelet en géométrie projective[(2)]. Elle est aujourd'hui omniprésente en mathématiques et dans les sciences en générale, en particulier en physique. Elle exprime l'idée que certains objets mathématiques admettent des objets "duaux", sortes d'images miroir dont les propriétés se déduisent de celles des objets initiaux. Ceci présente un intéret certain pour le mathématicien: on peut produire de nouveaux théorèmes à partir d'anciens simplement en "passant au dual"; mais aussi pour l'étudiant en mathématiques: il suffit de retenir la moitié des énoncés puisque la deuxième moitié se déduit de la première par dualité!

Dans ce chapitre, nous avons illustré plusieurs aspects de la dualité dans les espaces vectoriels, parfois appelée "dualité linéaire". Représentons dans un tableau les objets, propriétés et notions dont nous avons identifié les duaux dans ce chapitre. Le lecteur pourra compléter au fur et à mesure que sa culture mathématique s'enrichira.

Espace vectoriel V	Espace dual V^*
Vecteur	Forme linéaire
Base $\mathscr{B}$	Base duale $\mathscr{B}^*$
Matrice A	Matrice transposée ${}^tA$
Colonnes d'une matrice	Lignes d'une matrice
Opérations sur les lignes	Opérations sur les colonnes
Base d'un sous-espace	Equations d'un sous-espace
Sous-espace vectoriel E	Orthogonal $E^\perp$
Somme de sous-espaces	intersection de sous-espaces

(2) Ceux-ci ont par exemple remarqué l'analogie qu'il y a entre les propriétés des points alignés dans le plan et celles des droites concourrantes. Les propriétés des droites concourantes sont "duales" de celles des points alignés.

1.5. Appendice (†): Transposée d'une application linéaire

Soit $A = (a_{ij}) \in M_{m,n}(k)$ et soit u l'endomorphisme $k^n \to k^m$ associé à A. Notant $\mathscr{B} = (e_1, \dots, e_n)$ (resp. $\mathscr{C} = (f_1, \dots, f_m)$) la base canonique de k^n (resp. k^m) et $\mathscr{B}^* = (e_1^*, \dots, e_n^*)$ (resp. $\mathscr{C}^* = (f_1^*, \dots, f_m^*)$) la base duale, u est défini par $u(e_j) = \sum_{i=1} a_{ij} f_i$ pour $j = 1, \dots, n$, donc aussi par:

$$f_i^*(u(e_j)) = a_{ij} \qquad \forall j = 1, \dots, n, \quad \forall i = 1, \dots, m.$$

Donc, pour $i = 1, \dots, m$, la forme linéaire $f_i^* \circ u : k^n \to k$ s'écrit:

$$(\dagger) \qquad f_i^* \circ u = \sum_{j=1}^{n} a_{ij}\, e_j^*.$$

Considérons alors l'application $\theta : (k^m)^* \to (k^n)^*$ qui à toute forme linéaire $\phi : k^m \to k$ associe la forme linéaire $\phi \circ u : k^n \to k$. Alors θ est linéaire, car

$$\theta(t \cdot \phi + \psi) = (t \cdot \phi + \psi) \circ u = t \cdot \phi \circ u + \psi \circ u = t \cdot \theta(\phi) + \theta(\psi),$$

et (†) montre que $\theta(f_i^*) = \sum_{j=1}^{n} a_{ij}\, e_j^*$, d'où : $\mathrm{Mat}_{\mathscr{B}^*,\mathscr{C}^*}(\theta) = {}^t A$. Ceci conduit à la définition et au théorème suivants.

Définition 1.5.1 **(Transposée d'une application linéaire)**. — Soient E, F deux k-espaces vectoriels et $u : E \to F$ une application linéaire. On appelle *transposée de* u et l'on note ${}^t u$ l'application linéaire $F^* \to E^*$ définie par ${}^t u(\phi) = \phi \circ u$, pour tout $\phi \in F^*$. On peut visualiser ceci par le diagramme:

$$\begin{tikzcd} E \arrow[r, "u"] \arrow[dr, dashed, "{}^t u(\phi)=\phi\circ u"'] & F \arrow[d, "\phi"] \\ & k. \end{tikzcd}$$

Remarquons que si v est une application linéaire $F \to G$, alors $\boxed{{}^t(v \circ u) = {}^t u \circ {}^t v.}$ En effet, pour tout $\psi \in G^*$ on a

$${}^t(v \circ u)(\psi) = \psi \circ v \circ u = {}^t u(\psi \circ v) = {}^t u({}^t v(\psi)) = ({}^t u \circ {}^t v)(\psi).$$

Lemme 1.5.2. — *Soit $u : E \to F$ une application linéaire. Alors* $\mathrm{Ker}({}^t u) = \mathrm{Im}(u)^\perp = \{\phi \in F^* \mid \phi(u(E)) = 0\}$.

Démonstration. On a $\mathrm{Ker}({}^t u) = \{\phi \in F^* \mid \phi \circ u = 0\} = \{\phi \in F^* \mid \phi(u(E)) = 0\} = \mathrm{Im}(u)^\perp$.

Théorème 1.5.3. — *Soient E, F deux k-espaces vectoriels de dimension finie, $u : E \to F$ une application linéaire, $\mathscr{B}$ (resp. $\mathscr{C}$) une base de E (resp. de F) et $A = \mathrm{Mat}_{\mathscr{C},\mathscr{B}}(u)$.*

(i) *On a* $\mathrm{Mat}_{\mathscr{B}^*,\mathscr{C}^*}({}^t u) = {}^t A$.

(ii) *On a:* $\mathrm{rang}({}^t u) = \mathrm{rang}(u)^\perp$ *et* $\mathrm{rang}(A) = \mathrm{rang}({}^t A)$. *Par conséquent,* $\mathrm{rang}(A)$ *est aussi le nombre maximum de lignes de A linéairement indépendantes.*

Démonstration. On a déjà vu le point (i) au début de ce paragraphe. D'après le lemme, on a $\mathrm{Ker}({}^t u) = \dim \mathrm{Im}(u)^\perp$ donc, d'après le théorème du rang, appliqué à ${}^t u$, et la proposition 1.3.4, on a

$$n - \mathrm{rg}({}^t u) = \dim \mathrm{Ker}({}^t u) = \dim \mathrm{Im}(u)^\perp = n - \mathrm{rg}(u),$$

d'où $\mathrm{rg}({}^t u) = \mathrm{rg}(u)$. Le point (ii) en découle.

1.6. Appendice (†) : Bidual

Soient V un k-espace vectoriel et V^* son dual. Puisque V^* est un k-espace vectoriel, on peut aussi former son dual $(V^*)^*$, qu'on note V^{**} et qu'on appelle le *bidual* de V. Remarquons que tout vecteur $v \in V$ définit une forme linéaire ε_v sur V^*, l'application "d'évaluation en v":

$$\varepsilon_v : V^* \to k, \qquad f \mapsto f(v).$$

On a bien $\varepsilon_v(t \cdot f + f') = (t \cdot f + f')(v) = t \cdot f(v) + f'(v) = t \cdot \varepsilon_v(f) + \varepsilon_v(f')$, donc ε_v est un élément du bidual V^{**}.

De plus, l'application $v \mapsto \varepsilon_v$ est linéaire. En effet, pour tout $f \in V^*$, on a

$$\varepsilon_{t\cdot v+v'}(f) = f(t \cdot v + v') = t \cdot f(v) + f(v') = t \cdot \varepsilon_v(f) + \varepsilon_{v'}(f),$$

ce qui montre que $\varepsilon_{t\cdot v+v'} = t \cdot \varepsilon_v + \varepsilon_{v'}$. On obtient donc une application linéaire "canonique" (c.-à-d., qui ne dépend d'aucun choix):

$$\varepsilon : V \to V^{**}, \qquad v \mapsto \varepsilon_v.$$

Proposition 1.6.1. — *Supposons V de dimension finie n. Alors l'application canonique $\varepsilon : V \to V^{**}$ est un isomorphisme, donc permet d'identifier canoniquement V à son bidual V^{**}.*

Démonstration. D'abord, on a $\dim_k(V^{**}) = \dim_k(V^*) = \dim_k(V)$. D'autre part, soit $v_1 \in V$ non nul, on peut compléter la famille libre $\{v_1\}$ en une base $(v_1, \dots, v_n)$ de V, et si v_1^* est la forme linéaire définie par $v_1^*(v_1) = 1$ et $v_1^*(v_i) = 0$ pour $i = 2, \dots, n$, alors $\varepsilon_{v_1}(v_1^*) = 1$, donc $\varepsilon_{v_1} \neq 0$. Ceci montre que ε est injective; comme $\dim_k(V^{**}) = \dim_k(V)$ alors ε est un isomorphisme.

Remarque 1.6.1.1. — Attention! Si V n'est pas de dimension finie, on peut montrer que $\varepsilon : V \to V^{**}$ est injective mais pas surjective. (C'est facile dès qu'on dispose de l'existence de bases de V et de V^*). (On peut même montrer (mais c'est plus difficile) qu'il n'existe aucun isomorphisme de V sur V^{**}.)

Remarque 1.6.2. — Soient V un k-espace vectoriel de dimension n. L'isomorphisme $\varepsilon : V \xrightarrow{\sim} V^{**}$ permet, d'une part, de démontrer le point (2) de la proposition 1.1.4 comme suit. Soit $\Delta = (f_1, \dots, f_n)$ une base de V^*, et soit $\Delta^* = (f_1^*, \dots, f_n^*)$ la base duale de V^{**}. Via l'isomorphisme $\varepsilon : V \xrightarrow{\sim} V^{**}$, Δ^* est l'image d'une base $\mathscr{B} = (e_1, \dots, e_n)$ de V; alors, pour tous i, j, on a:

$$f_i(e_j) = \varepsilon_{e_j}(f_i) = f_j^*(f_i) = \begin{cases} 1 & 1 \text{ si } i = j, \\ 0 & \text{sinon,} \end{cases}$$

et donc $\Delta = (f_1, \dots, f_n)$ est la base duale de $\mathscr{B} = (e_1, \dots, e_n)$.

D'autre part, soient F un sous-espace vectoriel de V^*, et $F^\perp$ son orthogonal dans V^{**}. Alors, via l'isomorphisme $\varepsilon : V \xrightarrow{\sim} V^{**}$, $F^\perp$ s'identifie à

$$\{x \in V \mid \forall f \in F, \quad 0 = \varepsilon_x(f) = f(x)\} = F^\circ$$

(cf. la remarque 1.3.6).

CHAPITRE 2

ESPACES AFFINES

Résumé: Dans ce chapitre, on introduit la notion importante d'espace affine. De façon imagée, un espace affine $\mathscr{E}$ est "la même chose" qu'un espace vectoriel E dont on a oublié l'origine; c'est le cadre naturel pour faire de la géométrie. On introduit les notions de repère, de changement de repère, d'application affine, de barycentres, et enfin de sous-espace affine.

2.1. Espaces affines réels

***Définition 2.1.1* (Espaces affines réels).** — Soit E un $\mathbb{R}$-espace vectoriel. Un **espace affine $\mathscr{E}$ de direction E** est un ensemble non-vide $\mathscr{E}$ muni d'une application

$$\mathscr{E} \times \mathscr{E} \to E, \qquad (x, y) \mapsto \overrightarrow{xy}$$

vérifiant les deux propriétés suivantes :

$$(1) \qquad \boxed{\text{Relation de Chasles :} \quad \overrightarrow{xz} = \overrightarrow{xy} + \overrightarrow{yz}} \qquad \forall x, y, z \in \mathscr{E}.$$

$$(2) \qquad \text{L'application} \quad \mathscr{E} \times \mathscr{E} \to \mathscr{E} \times E, \quad (x, y) \mapsto (x, \overrightarrow{xy}) \quad \text{est } \textbf{bijective}.$$

On notera $(x, \overrightarrow{u}) \mapsto (x, x + \overrightarrow{u})$ la **bijection inverse**, c.-à-d., pour tout $x \in \mathscr{E}$ et $\overrightarrow{u} \in E$, $x + \overrightarrow{u}$ désigne l'unique $y \in \mathscr{E}$ tel que $\overrightarrow{xy} = \overrightarrow{u}$.

Vocabulaire: les éléments de $\mathscr{E}$ sont appelés "points", ceux de E sont appelés "vecteurs". Si E est de dimension finie n, ce que nous supposerons par la suite, on pose $\dim \mathscr{E} = \dim E$.

Remarque 2.1.2. — Dans la définition précédente, on peut remplacer $\mathbb{R}$ par un corps arbitraire k, on obtient ainsi la notion d'espace affine sur k. Comme nous ne considérerons que des espaces affines sur $\mathbb{R}$, nous dirons simplement "espaces affines", sans préciser "réels". De plus, pour abréger, on écrira souvent: "soit $(\mathscr{E}, E)$ un espace affine", ou même: "soit $\mathscr{E}$ un espace affine (de dimension n)", sans préciser l'espace vectoriel E (appelé la "direction" de $\mathscr{E}$).

***Exemple fondamental 2.1.3* (Sous-espaces affines d'un espace vectoriel)**

Soit V un espace vectoriel, E un sous-espace vectoriel de V et $\mathscr{E}$ une partie de V de la forme

$$\mathscr{E} = x_0 + E = \{x_0 + y \mid y \in E\}.$$

Alors $\mathscr{E}$ est un espace affine de direction E pour l'application $\mathscr{E} \times \mathscr{E} \to E$, $(x, y) \mapsto \overrightarrow{xy} = y - x$ (le vérifier!) On dit alors que $\mathscr{E}$ est un sous-espace affine de V.

L'exemple ci-dessus s'applique à de nombreuses situations courantes, et donne ainsi une longue liste d'exemple d'espaces affines. En voici quelques exemples.

Exemples 2.1.4. — Dans chacun des exemples ci-dessous, on a $\overrightarrow{xy} = y - x$ pour tout couple de points (x, y).

(1) Toute droite D du plan $\mathbb{R}^2$ est un espace affine dont la direction est la droite parallèle à D et passant par l'origine.

(2) Tout plan P de $\mathbb{R}^3$ est espace affine dont la direction est le plan parallèle à P et passant par l'origine.

(3) Si E est un $\mathbb{R}$-espace vectoriel, E est également un espace affine, de direction lui-même.

[(0)]version du 11/6/2015

D'une manière générale, l'ensemble des solutions d'une équation linéaire avec second membre est un espace affine, dont la direction est l'espace des solutions de l'équation homogène (c.-à-d.sans second membre) associée. Ceci est illustré par les exemples suvants. Ici encore, on a $\overrightarrow{xy} = y - x$ pour tout couple de points (x, y).

Exemples 2.1.5. — (1) Soient $A \in M_{m,n}(\mathbb{R})$ une matrice et $B \in M_{m,1}(\mathbb{R})$ un vecteur colonne; l'ensemble $\mathscr{S}$ des vecteurs colonnes $X \in M_{n,1}$ vérifiant $AX = B$ est un espace affine de direction $\ker(A)$.

(2) Soient $a, b, c \in \mathbb{R}$; l'ensemble $\mathscr{S}(a, b; c)$ des suites $(u_n)_{n\in\mathbb{N}}$ de nombres réels vérifiant la relation

$$u_{n+2} = au_{n+1} + bu_n + c$$

est un espace affine dont la direction est l'espace vectoriel $\mathscr{S}(a, b)$ des suites vérifiant la relation

$$u_{n+2} = au_{n+1} + bu_n.$$

(3) Soient $a, b, c \in \mathbb{R}$; l'ensemble $\mathscr{S}(a, b; c)$ des fonctions $f : \mathbb{R} \to \mathbb{R}$ de classe C^∞ vérifiant l'équation

$$\forall t \in \mathbb{R}, \quad f''(t) = af'(t) + bf(t) + c$$

est un espace affine dont la direction est l'espace vectoriel $\mathscr{S}(a, b)$ des fonctions vérifiant la relation

$$\forall t \in \mathbb{R}, \quad f''(t) = af'(t) + bf(t).$$

Définition 2.1.6 **(Repères cartésiens d'un espace affine de dimension finie)**
Soit $(\mathscr{E}, E)$ un espace affine, avec $\dim E = n$. Un **repère** (cartésien) $\mathscr{R}$ de $\mathscr{E}$ est un couple $(O, \mathscr{B})$, où O est un point de $\mathscr{E}$ et $\mathscr{B} = (e_1, \dots, e_n)$ est une base de E; on dira aussi que le $(n+1)$-uplet $(O, e_1, \dots, e_n)$ est un repère de $\mathscr{E}$. Dans ce cas, pour tout point P de $\mathscr{E}$, il existe un unique n-uplet $(x_1, \dots, x_n) \in \mathbb{R}^n$ tel que

$$\overrightarrow{OP} = x_1e_1 + \cdots + x_ne_n$$

et l'on dit que $(x_1, \dots, x_n)$ sont les **coordonnées** de P dans le repère $\mathscr{R}$.

Exemple 2.1.7. — Soit $\mathscr{B} = (e_1, \dots, e_n)$ la base canonique de l'espace vectoriel $E = \mathbb{R}^n$, posons

$$\mathscr{E} = \mathbb{R}^n = \{(x_1, \dots, x_n) \mid \forall i = 1, \dots, n, \quad x_i \in \mathbb{R}\}$$

considéré comme espace affine de direction $E = \mathbb{R}^n$, c.-à-d., muni de l'application $\mathscr{E} \times \mathscr{E} \to E$, $(x, y) \mapsto \overrightarrow{xy} = y - x = \sum_{i=1}^n (y_i - x_i)e_i$. Notons O le point $(0, \dots, 0)$ de $\mathscr{E}$ et $\mathscr{R}$ le repère $(O, \mathscr{B})$. Alors pour tout point $P = (x_1, \dots, x_n)$ de $\mathscr{E}$, ses coordonnées dans le repère $\mathscr{R}$ sont $(x_1, \dots, x_n)$. Si on fixe un point arbitraire $A = (a_1, \dots, a_n)$ de $\mathscr{E}$ et qu'on considère le repère $\mathscr{R}' = (A, \mathscr{B})$, alors les coordonnées $(x'_1, \dots, x'_n)$ de P dans $\mathscr{R}'$ sont données par l'égalité $\overrightarrow{AP} = x'_1e_1 + \cdots x'_ne_n$, et comme par définition $\overrightarrow{AP} = \sum_{i=1}^n (x_i - a_i)e_i$, on obtient $x'_i = x_i - a_i$ pour tout i.

Ce qui précède est un cas particulier du théorème suivant:

Théorème 2.1.8 **(Changement de repère).** — *Soit $(\mathscr{E}, E)$ un espace affine de dimension n, $\mathscr{B} = (e_1, \dots, e_n)$ et $\mathscr{C} = (f_1, \dots, f_n)$ deux bases de E, soit $P = \mathrm{Mat}_{\mathscr{B}}(\mathscr{C})$ la matrice de passage, et soient $O, O' \in \mathscr{E}$. Posons $\overrightarrow{OO'} = t_1e_1 + \cdots + t_ne_n$. Pour $M \in \mathscr{E}$ arbitraire, notons $(x_1, \dots, x_n)$ (resp. $(x'_1, \dots, x'_n)$) ses coordonnées dans le repère $\mathscr{R} = (O, \mathscr{B})$ (resp. $\mathscr{R}' = (O', \mathscr{C})$). Alors on a*

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} + \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix}, \qquad \textit{et donc} \qquad \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = P^{-1} \begin{pmatrix} x_1 - t_1 \\ \vdots \\ x_n - t_n \end{pmatrix}.$$

De façon abrégée, si l'on note X, X' et T les vecteurs colonnes ci-dessus, on a

$$\boxed{X = PX' + T} \qquad \boxed{X' = P^{-1}(X - T)}$$

Démonstration. — Soient X, X' comme ci-dessus et notons Y le vecteur colonne PX'. Alors, d'après la formule de changement de base dans E, on a:

$$\overrightarrow{O'M} = x'_1f_1 + \cdots + x'_nf_n = y_1e_1 + \cdots + y_ne_n,$$

et comme $\overrightarrow{OM} = \overrightarrow{OO'} + \overrightarrow{O'M}$ on obtient

$$X = \overrightarrow{OM} = \sum_{i=1}^n (t_i + y_i)e_i = T + PX'$$

et l'on a donc aussi $X' = P^{-1}(X - T)$, ce que l'on peut aussi retrouver en écrivant que

$$\sum_{i=1}^{n} x'_i f_i = \overrightarrow{O'M} = \overrightarrow{OM} - \overrightarrow{OO'} = \sum_{i=1}^{n} (x_i - t_i) e_i$$

d'où $X' = P^{-1}(X - T)$. □

2.2. Applications affines

Définition et proposition 2.2.1 **(Applications affines).** — *Soient $(\mathscr{E}, E)$ et $(\mathscr{E}', E')$ deux espaces affines. Une application $f : \mathscr{E} \to \mathscr{E}'$ est une* **application affine** *s'il existe un point $O \in \mathscr{E}$ tel que l'application $\phi : E \to E'$ définie par:*

$$\forall P \in \mathscr{E}, \qquad \phi(\overrightarrow{OP}) = \overrightarrow{f(O)f(P)}$$

soit **linéaire**. *Dans ce cas, on note $\phi = \overrightarrow{f}$, et les égalités ci-dessus sont vraies pour tout couple de points (I, P), c.-à-d., on a* **pour tout** *$I \in \mathscr{E}$:*

$$\boxed{\forall P \in \mathscr{E}, \quad \overrightarrow{f(I)f(P)} = \overrightarrow{f}(\overrightarrow{IP})} \qquad \text{et} \qquad \boxed{\forall \overrightarrow{u} \in E, \quad f(I + \overrightarrow{u}) = f(I) + \overrightarrow{f}(\overrightarrow{u}).}$$

On dit alors que $\overrightarrow{f}$ est la **partie linéaire** *de f.*

Démonstration. — Fixons un point $O \in \mathscr{E}$. Comme l'application $\theta : E \to \mathscr{E}$, $\overrightarrow{u} \mapsto O + \overrightarrow{u}$ est une bijection, d'inverse $P \mapsto \overrightarrow{OP}$ alors l'application $\theta^{-1} \circ f \circ \theta$ est l'unique application $\phi : E \to E$ telle que

$$\forall \overrightarrow{u} \in E, \qquad f(O + \overrightarrow{u}) = f(O) + \phi(\overrightarrow{u}) \qquad \text{(i.e. } \overrightarrow{f(O)f(P)} = \phi(\overrightarrow{OP}) \text{, pour tout } P \in \mathscr{E}).$$

On suppose que, pour ce choix de O, l'application $\phi : E \to E$ est **linéaire**. Soit $I \in \mathscr{E}$ arbitraire. D'après la relation de Chasles, et la définition de ϕ, on a, pour tout $P \in \mathscr{E}$:

$$\overrightarrow{f(I)f(P)} = \overrightarrow{f(O)f(P)} - \overrightarrow{f(O)f(I)} = \phi(\overrightarrow{OP}) - \phi(\overrightarrow{OI})$$

et comme ϕ est supposée linéaire, ceci égale $\phi(\overrightarrow{OP} - \overrightarrow{OI})$. D'après la relation de Chasles, à nouveau, on a $\overrightarrow{OP} - \overrightarrow{OI} = \overrightarrow{IP}$. Posant $\overrightarrow{f} = \phi$, on a donc obtenu:

$$\forall I, P \in \mathscr{E}, \qquad \overrightarrow{f(I)f(P)} = \overrightarrow{f}(\overrightarrow{IP}).$$

La proposition est démontrée. □

Remarque 2.2.2. — Pour $f : \mathscr{E} \to \mathscr{E}$ arbitraire, et $O \in \mathscr{E}$, l'application $\phi : E \to E$ définie plus haut n'a aucune raison d'être linéaire: prendre par exemple $\mathscr{E} = E = \mathbb{R}$, $O = 0$ et $f(t) = t^2$ pour tout $t \in \mathbb{R}$, alors $\phi(t) = f(t)$ n'est pas linéaire.

Proposition 2.2.3 **(Écriture d'une application affine dans des repères cartésiens)**
Soient $(\mathscr{E}, E)$ et $(\mathscr{E}', E')$ deux espaces affines, avec $\dim E = m$ et $\dim E' = n$, et soit $f : \mathscr{E} \to \mathscr{E}'$ une application affine. Soient $\mathscr{R} = (O, \mathscr{B})$ et $\mathscr{R}' = (O', \mathscr{C})$ des repères de $\mathscr{E}$ et $\mathscr{E}'$, soit $A = \mathrm{Mat}_{\mathscr{C},\mathscr{B}}(\overrightarrow{f}) \in M_{n,m}(\mathbb{R})$ et soient $(b'_1, \dots, b'_n)$ les coordonnées de $f(O)$ dans $\mathscr{R}'$.

Alors, pour tout $P \in \mathscr{P}$, de coordonnées $(x_1, \dots, x_m)$ dans $\mathscr{R}$, les coordonnées $(y_1, \dots, y_n)$ de $f(P)$ dans $\mathscr{R}'$ sont données par :

$$\boxed{\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} + \begin{pmatrix} b'_1 \\ \vdots \\ b'_n \end{pmatrix}.}$$

Démonstration. — En effet, on a $\begin{pmatrix} y_1 - b'_1 \\ \vdots \\ y_n - b'_n \end{pmatrix} = \overrightarrow{f(O)f(P)} = \overrightarrow{f}(\overrightarrow{OP}) = A \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$. □

Définition et proposition 2.2.4 **(Translations de $\mathscr{E}$).** — *Pour tout $\overrightarrow{u} \in E$, on appelle* **translation de vecteur** *$\overrightarrow{u}$ l'application $t_{\overrightarrow{u}} : \mathscr{E} \to \mathscr{E}$, qui à tout $x \in \mathscr{E}$ associe le point $x + \overrightarrow{u}$. C'est une application* **affine**, *dont la partie linéaire est id_E.*

Démonstration. — On fixe $O \in \mathscr{E}$ et pour tout $P \in \mathscr{E}$ on pose $P' = P + \overrightarrow{u}$. Alors

$$\overrightarrow{O'P'} = \underbrace{\overrightarrow{O'O}}_{=-\overrightarrow{u}} + \overrightarrow{OP} + \underbrace{\overrightarrow{PP'}}_{=\overrightarrow{u}} = \overrightarrow{OP}.$$

Avec les notations de 2.2.1, ceci montre que l'application ϕ égale id_E, qui est linéaire. Donc $t_{\overrightarrow{u}}$ est une application affine, de partie linéaire id_E. $\square$

Lemme 2.2.5. — *Pour tout $\overrightarrow{u}, \overrightarrow{v} \in E$, on a $\boxed{t_{\overrightarrow{u}} \circ t_{\overrightarrow{v}} = t_{\overrightarrow{u}+\overrightarrow{v}}}$. Donc l'ensemble des translations est un groupe commutatif; en particulier, toute translation $t_{\overrightarrow{u}}$ est* **bijective**, **d'inverse** $t_{-\overrightarrow{u}}$.

Démonstration. — Pour tout $P \in \mathscr{E}$, on a $(t_{\overrightarrow{u}} \circ t_{\overrightarrow{v}})(P) = t_{\overrightarrow{u}}(P + \overrightarrow{v}) = P + \overrightarrow{v} + \overrightarrow{u} = t_{\overrightarrow{u}+\overrightarrow{v}}(P)$. Ceci prouve la première assertion, et la seconde en découle facilement. $\square$

Proposition 2.2.6 **(Composée d'applications affines).** — *Soient $\mathscr{E} \xrightarrow{f} \mathscr{E}' \xrightarrow{g} \mathscr{E}''$ deux applications affines, alors la composée $g \circ f : \mathscr{E} \to \mathscr{E}''$ est affine, et sa partie linéaire est* $\boxed{\overrightarrow{(g \circ f)} = \overrightarrow{g} \circ \overrightarrow{f}.}$

Démonstration. — Fixons $O \in \mathscr{E}$ et posons $O' = f(O)$. Pour tout $M \in \mathscr{E}$ et $P \in \mathscr{E}'$, on a:

$$\overrightarrow{f(O)f(M)} = \overrightarrow{f}(\overrightarrow{OM}) \qquad \text{et} \qquad \overrightarrow{g(O')g(P)} = \overrightarrow{g}(\overrightarrow{O'P}).$$

En particulier, appliquant la deuxième égalité au cas où $P = f(M)$, on obtient:

$$\overrightarrow{(g \circ f)(O)(g \circ f)(M)} = \overrightarrow{g(f(O))g(f(M))} = \overrightarrow{g}(\overrightarrow{f(O)f(M)}) = \overrightarrow{g}\Big(\overrightarrow{f}(\overrightarrow{OM})\Big) = (\overrightarrow{g} \circ \overrightarrow{f})(\overrightarrow{OM}).$$

D'après 2.2.1, ceci prouve que $g \circ f$ est affine, de partie linéaire $\overrightarrow{g} \circ \overrightarrow{f}$. $\square$

Théorème et définition 2.2.7 **(Transformations affines et groupe** $\mathrm{GA}(\mathscr{E})$**)**

Soit $(\mathscr{E}, E)$ un espace affine.

(1) *On note $\mathrm{TAff}(\mathscr{E})$ l'ensemble des transformations affines de $\mathscr{E}$, i.e. applications affines $\mathscr{E} \to \mathscr{E}$; il est stable par composition, c.-à-d., si $f, g \in \mathrm{TAff}(\mathscr{E})$ alors $g \circ f \in \mathrm{TAff}(\mathscr{E})$.*

(2) *Pour tout $O \in \mathscr{E}$, tout $f \in \mathrm{TAff}(\mathscr{E})$ s'écrit $f = t_{\overrightarrow{Of(O)}} \circ \phi_O$, où ϕ_O est l'application affine définie par $\phi_O(P) = O + \overrightarrow{f}(\overrightarrow{OP})$.*

(3) *Soit $f \in \mathrm{TAff}(\mathscr{E})$, alors* $\boxed{f \text{ bijective} \iff \overrightarrow{f} \text{ bijective.}}$

(4) *On note $\mathrm{GA}(\mathscr{E}) = \{f \in \mathrm{TAff}(\mathscr{E}) \mid f \text{ bijective}\}$; c'est un* **groupe**, *et l'application $\mathrm{GA}(\mathscr{E}) \to \mathrm{GL}(E)$, $f \mapsto \overrightarrow{f}$ est un morphisme de groupes.*

(5) *Pour tout $f \in \mathrm{TAff}(\mathscr{E})$ et $\overrightarrow{u} \in E$, on a $f \circ t_{\overrightarrow{u}} = t_{\overrightarrow{f}(\overrightarrow{u})} \circ f$; en particulier si $f \in \mathrm{GA}(\mathscr{E})$ alors $f \circ t_{\overrightarrow{u}} \circ f^{-1} = t_{\overrightarrow{f}(\overrightarrow{u})}$.*

Démonstration. — (1) découle de 2.2.6. Prouvons (2). Soient $O \in \mathscr{E}$ et $f \in \mathrm{TAff}(\mathscr{E})$, posons $\overrightarrow{u} = \overrightarrow{Of(O)}$ et $g = t_{-\overrightarrow{u}} \circ f \in \mathrm{TAff}(\mathscr{E})$. Alors g est affine, sa partie linéaire est $\mathrm{id}_E \circ \overrightarrow{f} = \overrightarrow{f}$, et l'on a $g(O) = O$. Donc, pour tout $P \in \mathscr{E}$, on a $\overrightarrow{Og(P)} = \overrightarrow{f}(\overrightarrow{OP})$ donc g égale ϕ_O, et l'on a $f = t_{\overrightarrow{u}} \circ \phi_O$, ce qui prouve (2). De plus, comme $t_{\overrightarrow{u}}$ est bijective et $f = t_{\overrightarrow{u}} \circ \phi_O$, on a:

$$f \text{ bijective} \iff \phi_O \text{ bijective} \iff \overrightarrow{f} \text{ bijective}$$

ce qui prouve (3).

Prouvons (4). Soit $f \in \mathrm{GA}(\mathscr{E})$ et soient $O, P \in \mathscr{E}$. Posons $O' = f^{-1}(O)$ et $P' = f^{-1}(P)$, comme f est affine et $f(O') = O$, $f(P') = P$, on a:

$$\overrightarrow{f}(\overrightarrow{O'P'}) = \overrightarrow{f(O')f(P')} = \overrightarrow{OP}$$

et comme $\overrightarrow{f}$ est bijective, d'après (3), on obtient que

$$\overrightarrow{f^{-1}(O)f^{-1}(P)} = \overrightarrow{O'P'} = (\overrightarrow{f})^{-1}(\overrightarrow{OP})$$

et ceci prouve que f^{-1} est affine, de partie linéaire $(\overrightarrow{f})^{-1}$. Le reste du point (4) découle alors de 2.2.6.

Prouvons (5). Pour tout $P \in \mathscr{E}$, on a

$$(f \circ t_{\overrightarrow{u}})(P) = f(P + \overrightarrow{u}) = f(P) + \overrightarrow{f}(\overrightarrow{u}) = (t_{\overrightarrow{f}(\overrightarrow{u})} \circ f)(P)$$

ce qui prouve le premier point de (5), et le second en découle. $\square$

Remarque 2.2.8. — Soient $(\mathscr{E}, E)$ et $(\mathscr{E}', E')$ deux espaces affines de dimension finie, et soit $f : \mathscr{E} \to \mathscr{E}'$ une application affine bijective. Alors $\overrightarrow{f} : E \to E'$ est une application linéaire bijective, donc E et E' sont isomorphes; en particulier ils ont même dimension, donc $\dim \mathscr{E} = \dim \mathscr{E}'$.

2.3. Barycentres

***Théorème et définition 2.3.1* (Barycentres).** — *Soit $\mathscr{E}$ un espace affine, soient $P_1, \ldots, P_n \in \mathscr{E}$ et $t_1, \ldots, t_n$ des réels* **tels que** $\boxed{t_1 + \cdots + t_n = 1}$. *Alors il existe un unique point $G \in \mathscr{E}$ tel que:*

$$(*) \qquad \boxed{\forall I \in \mathscr{E}, \qquad \overrightarrow{IG} = t_1\overrightarrow{IP_1} + \cdots + t_n\overrightarrow{IP_n}}$$

et G s'appelle le **barycentre** *des points pondérés (P_i, t_i) (i.e. chaque point P_i est affecté du "poids" t_i). On notera:*

$$(**) \qquad \boxed{G = t_1P_1 + \cdots + t_nP_n.}$$

Lorsque $t_1 = \cdots = t_n$, d'où $t_1 = \cdots = t_n = 1/n$, on dit que G est le **centre de gravité** *ou* **isobarycentre** *des points $P_1, \ldots, P_n$.*

Démonstration. — Fixons une origine $O \in \mathscr{E}$. Si G existe, il vérifie nécessairement

$$(\dagger) \qquad \overrightarrow{OG} = t_1\overrightarrow{OP_1} + \cdots + t_n\overrightarrow{OP_n}$$

d'où l'unicité de G (s'il existe). Réciproquement, définissons G par l'égalité ci-dessus et montrons que G vérifie $(*)$ pour tout $I \in \mathscr{E}$. D'après la relation de Chasles et $(\dagger)$, on a

$$\sum_{i=1}^{n} t_i\overrightarrow{IP_i} = \sum_{i=1}^{n} t_i(\overrightarrow{IO} + \overrightarrow{OP_i}) = \underbrace{\Big(\sum_{i=1}^{n} t_i\Big)}_{=1}\overrightarrow{IO} + \sum_{i=1}^{n} t_i\overrightarrow{OP_i} = \overrightarrow{IO} + \overrightarrow{OG} = \overrightarrow{IG}.$$

Ceci prouve le théorème. □

Remarque 2.3.1.1. — La démonstration du théorème montre que si un point G' vérifie l'égalité $(*)$ pour *un* point I, alors G' égale G et vérifie l'égalité $(*)$ pour *tout* point I.

***Définition 2.3.2* (Segments).** — Soient $A, B \in \mathscr{E}$, on définit le **segment** $[A, B]$ comme l'ensemble des points qui sont "entre" A et B, i.e. l'ensemble des points P de la forme

$$P = A + t\overrightarrow{AB} = B + (1-t)\overrightarrow{BA}, \qquad \text{avec } t \in [0,1],$$

c'est donc aussi l'ensemble des barycentres $P = (1-t)A + tB$, avec $t \in [0,1]$. En particulier, le *centre de gravité* de A, B (qui correspond à $t = 1/2$) est le **milieu** du segment $[A, B]$.

***Proposition 2.3.3* (Associativité des barycentres).** — *Soient $A_1, \ldots, A_n \in \mathscr{E}$, $t_1, \ldots, t_n \in \mathbb{R}$ tels que $t_1 + \cdots + t_n = 1$, et G le barycentre des points pondérés (A_1, t_1), ..., (A_n, t_n).*

(1) *On suppose que $\mu = t_1 + \cdots + t_{n-1} = 1 - t_n$ est non nul. Soit H le barycentre des points pondérés $(A_1, t_1/\mu)$, ..., $(A_{n-1}, t_{n-1}/\mu)$, alors* $\boxed{G \text{ est le barycentre de } (H, \mu) \text{ et de } (A_n, t_n).}$

(2) *En particulier, si $t_1 = \cdots = t_n = 1/n$, soit G (resp. H) le centre de gravité G de $A_1, \ldots, A_n$ (resp. de $A_1, \ldots, A_{n-1}$), alors G est le barycentre de $(H, (n-1)/n)$ et de $(A_n, 1/n)$.*

Démonstration. — Soit $O \in \mathscr{E}$. Alors H est défini par $\overrightarrow{OH} = \dfrac{1}{\mu}\sum_{i=1}^{n-1} t_i\overrightarrow{OA_i}$ et le barycentre G' de (H, μ) et de (A_n, t_n) est défini par

$$\overrightarrow{OG'} = \mu\overrightarrow{OH} + t_n\overrightarrow{OA_n} = \sum_{i=1}^{n-1} t_i\overrightarrow{OA_i} + t_n\overrightarrow{OA_n} = \overrightarrow{OG}$$

d'où $G' = G$. □

***Exemple 2.3.4* (Centre de gravité d'un triangle).** — Dans le plan affine $\mathscr{P}$, soient A, B, C trois points non alignés, soit G le centre de gravité des points A, B, C, et soit $O \in \mathscr{P}$ arbitraire. Notons I le milieu du segment $[BC]$, alors la droite (AI) s'appelle la **médiane** du triangle **issue de** A. Appliquons la proposition précédente à $A_1 = B$, $A_2 = C$, $A_3 = A$, alors on a $\mu = 2/3$ et $t_i/\mu = 1/2$ pour $i = 1, 2$, donc $H = I$ et l'on a:

$$\overrightarrow{OG} = \frac{2}{3}\overrightarrow{OI} + \frac{1}{3}\overrightarrow{OA}.$$

Prenant $O = G$, on obtient $\overrightarrow{GA} = -2\overrightarrow{GI}$: ceci montre que G est situé sur le segment $[AI]$, aux deux-tiers du segment en partant du sommet A. On a le même résultat si l'on introduit les milieux J et K de $[BC]$ et $[CA]$, donc on obtient le résultat suivant: "*les médianes d'un triangle se coupent aux deux-tiers de leur longueur (en partant des sommets), et le point d'intersection est le centre de gravité du triangle*".

Théorème 2.3.5* (Applications affines et barycentres).** — *Soit $f : \mathscr{E} \to \mathscr{E}'$ une application affine. Alors f* **préserve les barycentres: si $G \in \mathscr{E}$ est le barycentre des points (A_i, t_i) alors $f(G) \in \mathscr{E}'$ est le barycentre des points $(f(A_i), t_i)$.*

Démonstration. — Fixons $O \in \mathscr{E}$ et pour tout $P \in \mathscr{E}$ notons $P' = f(P)$. Par définition, on a $\overrightarrow{OG} = \sum_{i=1}^n t_i \overrightarrow{OA_i}$ d'où, puisque $\overrightarrow{f}$ est linéaire:

$$\overrightarrow{O'G'} = \overrightarrow{f}\Big(\sum_{i=1}^n t_i \overrightarrow{OA_i}\Big) = \sum_{i=1}^n t_i \overrightarrow{f}(\overrightarrow{OA_i}) = \sum_{i=1}^n t_i \overrightarrow{O'A'_i}.$$

Ceci montre que G' est le barycentre des points (A'_i, t_i). □

2.4. Sous-espaces affines

Définition et proposition 2.4.1* (Sous-espaces affines).** — *Soient $(\mathscr{E}, E)$ un espace affine et $\mathscr{F}$ un sous-ensemble* **non-vide. Les conditions suivantes sont équivalentes:*

(1) *Pour tous $M_0, M_1, \dots, M_p \in \mathscr{F}$ et tous $t_0, t_1, \dots, t_p \in \mathbb{R}$ tels que $t_0 + \cdots + t_p = 1$, le barycentre $G = t_0 M_0 + \cdots + t_p M_p$ appartient à $\mathscr{F}$.*

(2) *Pour tous $M_0, M_1, \dots, M_p \in \mathscr{F}$ et tous $\lambda_1, \dots, \lambda_p \in \mathbb{R}$, le point $M_0 + \lambda_1 \overrightarrow{M_0M_1} + \cdots + \lambda_p \overrightarrow{M_0M_p}$ appartient à $\mathscr{F}$.*

(3) *Il existe $M_0 \in \mathscr{F}$ tel que $F = \{\overrightarrow{M_0M} \mid M \in \mathscr{F}\}$ soit un sous-espace vectoriel de E.*

(4) *Pour tout $M_0 \in \mathscr{F}$, $F = \{\overrightarrow{M_0M} \mid M \in \mathscr{F}\}$ est un sous-espace vectoriel de E, indépendant du choix de $M_0 \in \mathscr{F}$.*

Lorsque ces conditions sont vérifiées, on dit que $\mathscr{F}$ est un **sous-espace affine** *(en abrégé:* **sea***),* **de direction** *F, et l'on pose $\dim \mathscr{F} = \dim F$.*

Démonstration. — Montrons que (1) ⇔ (2): en prenant M_0 comme origine, G est défini par l'égalité

$$\overrightarrow{M_0G} = \sum_{i=1}^p t_i \overrightarrow{M_0M_i} \qquad \text{i.e.} \qquad G = M_0 + \sum_{i=1}^p t_i \overrightarrow{M_0M_i}$$

et l'équivalence (1) ⇔ (2) en découle (si $\lambda_1, \dots, \lambda_p$ sont donnés, on prend $t_i = \lambda_i$ pour $i = 1, \dots, p$ et $t_0 = 1 - \sum_{i=1}^p \lambda_i$).

Il est clair que (2) ⇒ (3). Supposons (3) vérifié pour un point M_0 fixé et soit $M'_0 \in \mathscr{F}$, notons provisoirement

$$F' = \{\overrightarrow{M'_0M} \mid M \in \mathscr{F}\}.$$

Posons $\overrightarrow{u} = \overrightarrow{M_0M'_0}$. Comme $M'_0 \in \mathscr{F}$, on a $\overrightarrow{u} \in F$, et comme F est un sous-espace vectoriel de E, l'application $\overrightarrow{v} \mapsto \overrightarrow{v} - \overrightarrow{u}$ est une bijection de F sur lui-même (d'inverse $\overrightarrow{w} \mapsto \overrightarrow{w} + \overrightarrow{u}$). D'après la relation de Chasles, pour tout $M \in \mathscr{F}$ on a

$$\overrightarrow{M'_0M} = \overrightarrow{M_0M} - \overrightarrow{M_0M'_0} = \overrightarrow{M_0M} - \overrightarrow{u}$$

et donc

$$F' = \{\overrightarrow{v} - \overrightarrow{u} \mid \overrightarrow{v} \in F\} = F.$$

Ceci prouve que (3) ⇒ (4). Enfin, il est clair que (4) ⇒ (2). La proposition est démontrée. □

Remarque 2.4.2. — Soient $(\mathscr{F}, F)$ comme dans la proposition précédente, alors $\mathscr{F}$ est un espace affine de direction F, ce qui justifie la terminologie "sous-espace affine de $\mathscr{E}$ de direction F". En effet, l'application

$$\mathscr{F} \times \mathscr{F} \to F, \qquad (M_0, M) \mapsto \overrightarrow{M_0M}$$

est la restriction à $\mathscr{F} \times \mathscr{F}$ de l'application $\mathscr{E} \times \mathscr{E} \to E$, donc vérifie la relation de Chasles. D'autre part, d'après le point (4) de la proposition, pour tout $M_0 \in \mathscr{F}$, l'application $\mathscr{F} \to F$, $M \mapsto \overrightarrow{M_0M}$ est bijective. Ceci prouve que $\mathscr{F}$ est bien un espace affine de direction F (cf. définition 2.1.1).

***Définition 2.4.3* (Sea de direction F passant par un point A).** — Soient $(\mathscr{E}, E)$ un espace affine et F un sev de E. Pour tout $A \in \mathscr{E}$, il existe un unique sea de direction F passant par A, c'est:

$$\mathscr{F} = \{P \in \mathscr{E} \mid \overrightarrow{AP} \in F\} = \{A + \overrightarrow{u} \mid \overrightarrow{u} \in F\}$$

on le note $A + F$. De plus, pour tout point $B \in \mathscr{F}$, on a aussi $\mathscr{F} = B + F$.

Exemples 2.4.4. — (1) Pour tout $P \in \mathscr{E}$, le singleton $\{P\}$ est un sea de $\mathscr{E}$ de dimension 0, et réciproquement.

(2) Soit $\mathscr{D}$ un sea de $\mathscr{E}$ de dimension 1, et soit $D = \mathbb{R}\overrightarrow{u}$ sa direction (une droite vectorielle dans E). Pour tout A fixé dans $\mathscr{D}$, on a:

$$\mathscr{D} = A + \mathbb{R}\overrightarrow{u} = \{A + t\overrightarrow{u} \mid t \in \mathbb{R}\} = \{(1-t)A + tB \mid t \in \mathbb{R}\}$$

où l'on a posé $B = A + \overrightarrow{u}$. Réciproquement, pour tout $A \neq B$ dans $\mathscr{E}$, l'ensemble des barycentres:

$$\{(1-t)A + tB \mid t \in \mathbb{R}\}$$

est un sea de $\mathscr{E}$ de direction $\mathbb{R}\overrightarrow{AB}$, qu'on appelle la "**droite affine**" (AB).

***Proposition 2.4.5* (Sous-espaces affines définis par des équations)**

Soit $(\mathscr{E}, E)$ un espace affine. On fixe $O \in \mathscr{E}$. Soient $f_1, \dots, f_p \in E^*$ des formes linéaires sur E et $c_1, \dots, c_p \in \mathbb{R}$. On pose

$$\mathscr{F} = \{P \in \mathscr{E} \mid f_i(\overrightarrow{OP}) = c_i, \quad \forall i = 1, \dots, p\}.$$

Alors, si $\mathscr{F} \neq \varnothing$, c'est un sea de $\mathscr{E}$, de direction l'espace vectoriel

$$F = \{\overrightarrow{u} \in E \mid f_i(\overrightarrow{u}) = 0, \quad \forall i = 1, \dots, p\} = \bigcap_{i=1}^{p} \operatorname{Ker}(f_i).$$

Démonstration. — Supposons $\mathscr{F} \neq \varnothing$ et soit $M_0 \in \mathscr{F}$. Alors un point arbitraire $M \in \mathscr{E}$ appartient à $\mathscr{F}$ si et seulement si on a:

$$\forall i = 1, \dots, p, \qquad f_i(\overrightarrow{OM}) = c_i = f_i(\overrightarrow{OM_0})$$

ce qui équivaut à:

$$\forall i = 1, \dots, p, \qquad f_i(\overrightarrow{M_0M}) = f_i(\overrightarrow{OM} - \overrightarrow{OM_0}) = 0$$

et encore à: $\overrightarrow{M_0M} \in F = \bigcap_{i=1}^{p} \operatorname{Ker}(f_i)$. D'après la proposition 2.4.1, ceci montre que $\mathscr{F}$, s'il est non-vide, est un sea de direction F. $\square$

Exemples 2.4.6. — Soit $(\mathscr{E}, E)$ un espace affine de dimension n. Fixons un repère $(O, e_1, \dots, e_n)$ de $\mathscr{E}$, d'où des coordonnées $(x_1, \dots, x_n)$. Alors toute forme linéaire sur E est de la forme $x \mapsto a_1x_1 + \cdots + a_nx_n$ et donc se donner des équations $f_i(\overrightarrow{OM}) = c_i$ pour $i = 1, \dots, p$ équivaut à se donner un système:

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n &= c_1 \\ \cdots &= \cdots \\ a_{p1}x_1 + \cdots + a_{pn}x_n &= c_p \end{cases}$$

et la proposition précédente signifie la chose suivante: **si l'ensemble $\mathscr{F}$ des solutions de ce système est non-vide, alors c'est un sea de $\mathscr{E}$ de direction l'espace vectoriel F formé des solutions du système homogène:**

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ \cdots &= \cdots \\ a_{p1}x_1 + \cdots + a_{pn}x_n &= 0. \end{cases}$$

Illustrons ceci par les deux exemples suivants:

(1) $\mathscr{F} = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 1, \quad x_1 + 2x_2 = 2\} \neq \varnothing$ (il contient, par exemple, le point $(0, 1, 0)$), c'est un sea de $\mathbb{R}^3$ de direction la droite vectorielle

$$F = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0 = x_1 + 2x_2\}.$$

(2) Par contre,

$$\mathscr{F} = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 + x_2 = 1, \quad 2x_1 + 2x_2 = 3\}$$

est vide !

De plus, dans un espace affine $\mathscr{E}$ de dimension n, tout sous-espace affine $\mathscr{F}$ de dimension r peut être défini par exactement $n - r$ équations:

Proposition 2.4.7. — *Soient* $(\mathscr{E}, E)$ *un espace affine de dimension* n *et* $\mathscr{F}$ *un sea de direction* F, $\dim F = r$. *Soit* $O \in \mathscr{E}$ *arbitraire et* $(e_1, \dots, e_r)$ *une base de* F, *complétons-la en une base* $\mathscr{B} = (e_1, \dots, e_n)$ *de* E *et notons* $(x_1, \dots, x_n)$ *les coordonnées dans le repère* $\mathscr{R} = (O, \mathscr{B})$. *Alors il existe* $c_{r+1}, \dots, c_n \in \mathbb{R}$ *tels que*

$$\mathscr{F} = \{M(x_1, \dots, x_n) \in \mathscr{E} \mid x_i = c_i, \quad \forall i = r+1, \dots, n\}.$$

Démonstration. — Soit $M_0 \in \mathscr{F}$, posons $c_i = x_i(M_0)$ pour $i = r+1, \dots, n$. Pour tout $M \in \mathscr{F}$, on a

$$\overrightarrow{M_0M} \in F = \mathrm{Vect}(e_1, \dots, e_r)$$

donc $\overrightarrow{M_0M}$ a toutes ses coordonnées d'indice $i > r$ nulles, et donc $M = M_0 + \overrightarrow{M_0M}$ vérifie $x_i(M) = x_i(M_0) = c_i$ pour $i > r$.

Réciproquement, si $M \in \mathscr{E}$ vérifie $x_i(M) = x_i(M_0) = c_i$ pour $i > r$, alors le vecteur $\overrightarrow{M_0M}$ a toutes ses coordonnées d'indice $i > r$ nulles, donc appartient à F, et donc $M = M_0 + \overrightarrow{M_0M}$ appartient à $\mathscr{F}$. Ceci prouve la proposition. □

Définition et proposition 2.4.8 **(Sous-espace affine engendré par une partie non-vide)**
Soit $(\mathscr{E}, E)$ *un espace affine,* X *une partie non vide de* $\mathscr{E}$. *Alors l'ensemble* $\mathscr{F}$ *de tous les barycentres*

$$G = t_0A_0 + \cdots + t_pA_p, \qquad \text{où} \quad A_0, \dots, A_p \in X, \quad t_0, \dots, t_p \in \mathbb{R}, \quad t_0 + \cdots + t_p = 1$$

est un sous-espace affine de $\mathscr{E}$, *et c'est le* **plus petit sea de** $\mathscr{E}$ **contenant** X. *On l'appelle le sea* **engendré par** X *et on le note* $\mathrm{Aff}\langle X\rangle$. *De plus, pour tout choix d'un point* $A_0 \in X$, *on a:*

$$\begin{aligned}\mathrm{Aff}\langle X\rangle &= \{A_0 + \lambda_1\overrightarrow{A_0A_1} + \cdots + \lambda_p\overrightarrow{A_0A_p} \mid A_1, \dots, A_p \in X,\ \lambda_1, \dots, \lambda_p \in \mathbb{R}\}\\ &= A_0 + \mathrm{Vect}\{\overrightarrow{A_0B} \mid B \in X\}\end{aligned}$$

Démonstration. — On définit $\mathscr{F}$ par:

$$\mathscr{F} = \{G = t_0A_0 + \cdots + t_pA_p \mid A_0, \dots, A_p \in X, \quad t_0, \dots, t_p \in \mathbb{R}, \quad t_0 + \cdots + t_p = 1\},$$

cette définition ne dépend d'aucun choix et tous les points de X y jouent le même rôle. On pourrait montrer directement, avec cette définition, que $\mathscr{F}$ vérifie la propriété (1) de la proposition 2.4.1, mais il est plus commode de procéder comme suit. Fixons un point $A_0 \in X$, alors, comme dans la démonstration de (1) $\Leftrightarrow$ (2) dans 2.4.1, on voit que

$$\begin{aligned}\mathrm{Aff}\langle X\rangle &= \{A_0 + \lambda_1\overrightarrow{A_0A_1} + \cdots + \lambda_p\overrightarrow{A_0A_p} \mid A_1, \dots, A_p \in X,\ \lambda_1, \dots, \lambda_p \in \mathbb{R}\}\\ &= A_0 + \mathrm{Vect}\{\overrightarrow{A_0B} \mid B \in X\}.\end{aligned}$$

Ceci montre que $\mathscr{F}$ est un sea de direction $F = \mathrm{Vect}\{\overrightarrow{A_0B} \mid B \in X\}$ (qui ne dépend pas du choix de A_0, cf. 2.4.1). De plus, $\mathscr{F}$ contient X, et si $\mathscr{F}'$ est un sea contenant X, alors sa direction F' contient tous les vecteurs $\overrightarrow{A_0B}$, pour $B \in \mathscr{F}$, donc F' contient F, et $\mathscr{F}' = A_0 + F'$ contient $\mathscr{F}$. Ceci montre que $\mathscr{F}$ est le **plus petit** sea de $\mathscr{E}$ contenant X, ce qui justifie la notation $\mathrm{Aff}\langle X\rangle$. □

Corollaire 2.4.9 **(Sea engendré par des points** $A_0, \dots, A_p$**)**. — *Soit* $(\mathscr{E}, E)$ *un espace affine, le sous-espace affine engendré par des points* $A_0, \dots, A_p \in \mathscr{E}$ *est*

$$\begin{aligned}\mathrm{Aff}\langle A_0, \dots, A_p\rangle &= \{G = t_0A_0 + \cdots + t_pA_p \mid t_0, \dots, t_p \in \mathbb{R}, \quad t_0 + \cdots + t_p = 1\}\\ &= A_0 + \mathrm{Vect}\{\overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_p}\}\end{aligned}$$

sa dimension est celle de l'espace vectoriel $F = \mathrm{Vect}\{\overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_p}\}$; *en particulier,* $\dim \mathrm{Aff}\langle A_0, \dots, A_p\rangle = p \Leftrightarrow$ *les vecteurs* $\overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_p}$ *sont linéairement indépendants.*

Remarque 2.4.10. — Dans le corollaire précédent, on a choisi le point A_0 comme "origine", mais bien sûr le même résultat est valable pour tout point A_i, i.e. on a aussi $F = \mathrm{Vect}\{\overrightarrow{A_iA_j} \mid j \neq i\}$ et $\mathrm{Aff}\langle A_0, \dots, A_p\rangle = A_i + F$.

Définition 2.4.11 **(Points affinement indépendants ou liés)**. — Soit $(\mathscr{E}, E)$ un espace affine et soit $p \in \mathbb{N}^*$. On dit que $(p+1)$ points $A_0, \dots, A_p \in \mathscr{E}$ sont **affinement indépendants** si le sea de $\mathscr{E}$ qu'ils engendrent (i.e. le plus petit sea de $\mathscr{E}$ qui les contient) est de dimension p, c.-à-d., si les vecteurs $\overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_p}$ sont linéairement indépendants (ceci équivaut à dire que, pour i fixé, les p vecteurs $\overrightarrow{A_iA_j}$, pour $j \neq i$, sont linéairement indépendants). Dans le cas contraire, on dit que $A_0, \dots, A_p \in \mathscr{E}$ sont **affinement liés**.

Si $p = 2$, les trois points A_0, A_1, A_2 sont affinement liés $\Longleftrightarrow$ ils sont alignés. Donc A_0, A_1, A_2 sont affinement indépendants $\Longleftrightarrow$ ils sont **non alignés**.

On dit que des points $A_0, \dots, A_p \in \mathscr{E}$ sont **coplanaires** s'ils sont contenus dans un sea $\mathscr{P}$ de dimension 2 (un plan affine), i.e. si $\dim \mathrm{Aff}\langle A_0, \dots, A_p\rangle \leq 2$. Donc quatre points $A_0, \dots, A_3$ sont affinement indépendants $\Longleftrightarrow$ ils sont **non coplanaires**.

***Définition 2.4.12* (Retour sur les repères).** — Soit $(\mathscr{E}, E)$ un espace affine de dimension n. Se donner un repère $\mathscr{R} = (A_0, \mathscr{B})$ de $\mathscr{E}$ (où $\mathscr{B}$ est une base de E) équivaut à se donner un $(n+1)$-uplet de points $A_0, A_1, \dots, A_n$ **affinement indépendants**: dans ce cas les n vecteurs $\overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_n}$ sont linéairement indépendants donc forment une base $\mathscr{B}$ de E, et la donnée de $(A_0, \mathscr{B})$ équivaut bien sûr à celle du $(n+1)$-uplet $(A_0, A_1, \dots, A_n)$. On dira aussi qu'un tel $(n+1)$-uplet de points affinement indépendants forme un repère de $\mathscr{E}$.

Attention, cette fois l'ordre des points est important: A_0 est l'origine et la base $\mathscr{B} = (\overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_n})$ est ordonnée, i.e. les coordonnées $(x_1, \dots, x_n)$ dans ce repère correspondent au point

$$A_0 + x_1\overrightarrow{A_0A_1} + \cdots + x_n\overrightarrow{A_0A_n} = \big(1 - \sum_{i=1}^{n} x_i\big)A_0 + x_1A_1 + \cdots + x_nA_n.$$

***Définitions 2.4.13* (Sea parallèles).** — Soit $(\mathscr{E}, E)$ un espace affine et $\mathscr{F}_1, \mathscr{F}_2$ deux sea, de directions respectives F_1 et F_2.

(1) On dit que $\mathscr{F}_1$ est *faiblement parallèle* à $\mathscr{F}_2$ si l'on a $F_1 \subseteq F_2$. Dans ce cas, si $A_0 \in \mathscr{F}_1 \cap \mathscr{F}_2$, alors $\mathscr{F}_1 = A_0 + F_1 \subseteq A_0 + F_2 = \mathscr{F}_2$, donc on a l'alternative suivante: ou bien

$$\mathscr{F}_1 \subseteq \mathscr{F}_2 \qquad \text{ou bien} \qquad \mathscr{F}_1 \cap \mathscr{F}_2 = \varnothing.$$

(2) On dit que $\mathscr{F}_1$ et $\mathscr{F}_2$ sont **parallèles** si l'on a $F_1 = F_2$. Dans ce cas, on a l'alternative suivante: ou bien $\mathscr{F}_1 = \mathscr{F}_2$ ou bien $\mathscr{F}_1 \cap \mathscr{F}_2 = \varnothing$.

Remarque 2.4.13.1. — Attention, deux sea $\mathscr{F}_1, \mathscr{F}_2$ qui vérifient $\mathscr{F}_1 \cap \mathscr{F}_2 = \varnothing$ ne sont **pas nécessairement** parallèles! Par exemple, dans $\mathbb{R}^3$ les droites affines $\mathscr{D}_1$ et $\mathscr{D}_2$ d'équations respectives $x = y = 0$ et $x - 1 = 0 = z$ vérifient $\mathscr{D}_1 \cap \mathscr{D}_2 = \varnothing$ mais ne sont pas parallèles: la direction D_1 de $\mathscr{D}_1$ (resp. D_2 de $\mathscr{D}_2$) est la droite vectorielle d'équation $x = 0 = y$ (resp. $x = 0 = z$) et l'on a $D_1 \neq D_2$.

***Proposition 2.4.14* (Sea engendré par deux sea $\mathscr{F}$ et $\mathscr{F}'$).** — *Soit $(\mathscr{E}, E)$ un espace affine, soient $(\mathscr{F}, F)$, $(\mathscr{F}', F')$ deux sous-espaces affines, $P \in \mathscr{F}$ et $P' \in \mathscr{F}'$.*

(1) *Le sous-espace vectoriel $V = F + F' + \mathbb{R}\overrightarrow{PP'}$ est indépendant du choix de $P \in \mathscr{F}$ et $P' \in \mathscr{F}'$.*
(2) *On a $V = F + F' \Longleftrightarrow \mathscr{F} \cap \mathscr{F}' \neq \varnothing$.*
(3) *Le sous-espace affine engendré par $\mathscr{F} \cup \mathscr{F}'$ est $P + V = P' + V$, on le notera $\mathscr{F} + \mathscr{F}'$.* [(1)]
(4) *Par conséquent, si $\mathscr{F}$ et $\mathscr{F}'$ sont de dimension finie (par exemple, si $\mathscr{E}$ l'est), on a:*

$$\dim(\mathscr{F} + \mathscr{F}') = \begin{cases} \dim(F + F') & si\ \mathscr{F} \cap \mathscr{F}' \neq \varnothing, \\ \dim(F + F') + 1 & si\ \mathscr{F} \cap \mathscr{F}' = \varnothing. \end{cases}$$

Démonstration. — (1) Si $Q \in \mathscr{F}$ et $Q' \in \mathscr{F}'$, alors $\overrightarrow{QQ'} = \overbrace{\overrightarrow{QP}}^{\in F} + \overrightarrow{PP'} + \overbrace{\overrightarrow{P'Q'}}^{\in F'}$ et donc $F + F' + \overrightarrow{QQ'} = F + F' + \overrightarrow{PP'}$.

Prouvons (2). L'implication $\Leftarrow$ est évidente, car si $Q \in \mathscr{F} \cap \mathscr{F}'$, on peut prendre $P = Q = P'$, d'où $V = F + F'$. Réciproquement, supposons que $\overrightarrow{PP'} = \overrightarrow{u} + \overrightarrow{u}'$ avec $\overrightarrow{u} \in F$ et $\overrightarrow{u}' \in F'$. Alors le point $Q = P + \overrightarrow{u}$ appartient à $\mathscr{F}$, et l'on a aussi

$$\overrightarrow{P'Q} = \overrightarrow{PQ} - \overrightarrow{PP'} = \overrightarrow{u} - (\overrightarrow{u} + \overrightarrow{u}') = -\overrightarrow{u}' \in F'$$

donc $Q \in \mathscr{F}' \cap \mathscr{F}$. Ceci prouve (2).

Prouvons (3). Par définition, $\mathrm{Aff}\langle \mathscr{F} \cup \mathscr{F}'\rangle = P + W$, où W est le sous-espace vectoriel engendré par les vecteurs $\overrightarrow{PM}$ et $\overrightarrow{PN} = \overrightarrow{PP'} + \overrightarrow{P'N}$, où M (resp. N) parcourt $\mathscr{F}$ (resp. $\mathscr{F}'$). Donc W est le sev engendré par F, F' et $\overrightarrow{PP'}$, i.e. $W = V$. Ceci prouve (3). Enfin, (4) découle de (2), car $\dim V \leq \dim(F + F') + 1$, avec égalité si et seulement si $\overrightarrow{PP'} \notin F + F'$. $\square$

Remarque 2.4.15. — Dans la proposition précédente, si $\dim \mathscr{F} = p$, $\dim \mathscr{F}' = r$ et si $(P, A_1, \dots, A_p)$ et $(P', B_1, \dots, B_r)$ sont des repères de $\mathscr{F}$ et $\mathscr{F}'$, alors $\mathscr{F} + \mathscr{F}' = \mathrm{Aff}\langle P, A_1, \dots, A_p, P', B_1, \dots, B_r\rangle$.

[(1)] Attention, cette notation n'est peut-être pas standard, peut-être les puristes s'en tiennent-ils à la notation $\mathrm{Aff}\langle \mathscr{F} \cup \mathscr{F}'\rangle$...

***Proposition 2.4.16* (Intersection de deux sea $\mathscr{F}$ et $\mathscr{F}'$).** — *Soit $(\mathscr{E}, E)$ un espace affine et soient $(\mathscr{F}, F)$, $(\mathscr{F}', F')$ deux sous-espaces affines. Si $\mathscr{F} \cap \mathscr{F}'$ est non vide, c'est un sea de direction $F \cap F'$.*

Démonstration. — Supposons que $A \in \mathscr{F} \cap \mathscr{F}'$. Alors un point arbitraire $P \in \mathscr{E}$ appartient à $\mathscr{F} \cap \mathscr{F}'$ si et seulement si $\overrightarrow{AP}$ appartient à F et à F', i.e. à $F \cap F'$. D'après la proposition 2.4.1, ceci montre que $\mathscr{F} \cap \mathscr{F}'$ est un sea de direction $F \cap F'$. □

***Proposition 2.4.17* (Applications affines et sous-espaces).** — *Soit $f : \mathscr{E} \to \mathscr{E}'$ une application affine, notons ϕ sa partie linéaire.*

(1) *Soit $(\mathscr{F}, F)$ un sea de $\mathscr{E}$, alors $f(\mathscr{F})$ est un sea de $\mathscr{E}'$; plus précisément, si $P \in \mathscr{F}$ alors $f(\mathscr{F}) = f(P) + \phi(F)$.*

(2) *Soit $(\mathscr{F}', F')$ un sea de $\mathscr{E}'$. On a $f^{-1}(\mathscr{F}') = \varnothing$ si $f(\mathscr{E}) \cap \mathscr{F}' = \varnothing$, sinon $f^{-1}(\mathscr{F}')$ est un sea de $\mathscr{E}$ de direction $\phi^{-1}(F')$; plus précisément, si $P \in \mathscr{F}$ est tel que $f(P) \in \mathscr{F}'$ alors*

$$\begin{aligned} f^{-1}(\mathscr{F}') &= \{Q = P + \overrightarrow{u} \in \mathscr{E} \mid f(Q) = f(P) + \phi(\overrightarrow{u}) \in \mathscr{F}'\} \\ &= \{Q = P + \overrightarrow{u} \in \mathscr{E} \mid \overrightarrow{f(P)f(Q)} = \phi(\overrightarrow{u}) \in F'\} = P + \phi^{-1}(F'). \end{aligned}$$

Démonstration. — Laissée au lecteur comme exercice. □

2.5. Projections, symétries, points fixes

Proposition 2.5.1. — *Soit $(\mathscr{E}, E)$ un espace affine et soient $(\mathscr{F}, F)$, $(\mathscr{F}', F')$ deux sous-espaces affines tels que $F + F' = E$. Alors $\mathscr{F} \cap \mathscr{F}'$ est non vide et est un sous-espace affine de direction $F \cap F'$.*

Démonstration. — Ceci résulte des propositions 2.4.14 et 2.4.16. En effet, soient $P \in \mathscr{F}$ et $P' \in \mathscr{F}'$. Comme $F+F' = E$, on a nécessairement $\overrightarrow{PP'} \in F+F'$ donc, d'après la proposition 2.4.14, on a $\mathscr{F} \cap \mathscr{F}' \neq \varnothing$. Donc, d'après la proposition 2.4.16, $\mathscr{F} \cap \mathscr{F}'$ est un sea de direction $F \cap F'$. □

***Corollaire 2.5.2* (Sous-espaces supplémentaires).** — *Si $E = F \oplus F'$, i.e. si F et F' sont supplémentaires, alors pour tous sea $\mathscr{F}$ de direction F et $\mathscr{F}'$ de direction F', l'intersection $\mathscr{F} \cap \mathscr{F}'$ est un singleton $\{P\}$.*

Démonstration. — D'après la proposition précédente, $\mathscr{F} \cap \mathscr{F}'$ est non vide et est un sea de direction $F \cap F' = \{0\}$, donc c'est un singleton $\{P\}$. □

Définition 2.5.3. — Soient $(\mathscr{E}, E)$ un espace affine et $(\mathscr{F}, F)$ et $(\mathscr{F}', F')$ deux sea tels que $E = F \oplus F'$. D'après le corollaire précédent, $\mathscr{F} \cap \mathscr{F}'$ est un singleton $\{O\}$ et, comme $E = F \oplus F'$, on a:

$$\forall M \in \mathscr{E}, \quad \exists!(P, P') \in \mathscr{F} \times \mathscr{F}' \quad \text{tel que} \quad \overrightarrow{OM} = \overrightarrow{OP} + \overrightarrow{OP'}.$$

On définit alors:

(1) la **projection** $\pi = \pi_{\mathscr{F},\mathscr{F}'}$ **sur $\mathscr{F}$ parallèlement à $\mathscr{F}'$** par $\pi(M) = P$,

(2) la **symétrie** $s = s_{\mathscr{F},\mathscr{F}'}$ **par rapport à $\mathscr{F}$ parallèlement à $\mathscr{F}'$** par $s(M) = M + 2\overrightarrow{MP} = M - 2\overrightarrow{OP'}$, i.e. $\overrightarrow{Ms(M)} = 2\overrightarrow{MP} = 2\overrightarrow{OP'}$.

(Faire un dessin dans le plan affine $\mathscr{P}$, pour deux droites sécantes $\mathscr{F}$ et $\mathscr{F}'$, non nécessairement orthogonales, par exemple $\mathscr{F}$ d'équation $y = 1$ et $\mathscr{F}'$ d'équation $y - x = 1$.)

Si de plus $\dim \mathscr{F} = \dim \mathscr{E} - 1$ (de sorte que $\dim \mathscr{F}' = 1$, i.e. $\mathscr{F}'$ est une droite affine), on dit que $\mathscr{F}$ est un **hyperplan** affine de $\mathscr{E}$ et que $s_{\mathscr{F},\mathscr{F}'}$ est la **réflexion par rapport à l'hyperplan $\mathscr{F}$ parallèlement à la droite $\mathscr{F}'$**.

***Définition 2.5.4* (Points fixes).** — Si X est un ensemble et f une application $X \to X$, on dit que $x \in X$ est un **point fixe** de f si $f(x) = x$. On notera $\mathrm{Fix}(f)$ l'ensemble des points fixes de f; il peut être vide.

Soit $(\mathscr{E}, E)$ un espace affine et $f : \mathscr{E} \to \mathscr{E}$ une transformation affine. Si f possède un point fixe O, on a vu que, via la bijection $\mathscr{E} \xrightarrow{\sim} E$, $P \mapsto \overrightarrow{OP}$, f s'identifie à sa partie linéaire $\overrightarrow{f}$ (cf. point (2) du théorème 2.2.7). Donc, lorsque f possède un point fixe, l'étude de la transformation f se ramène à l'étude de l'endomorphisme $\overrightarrow{f}$ de E, et l'on peut appliquer les résultats connus sur les endomorphismes. Ceci explique l'intérêt d'étudier l'ensemble des points fixes de f. On a la proposition suivante.

***Proposition 2.5.5* (Points fixes de f).** — *Soient $(\mathscr{E}, E)$ un espace affine de dimension finie, $f : \mathscr{E} \to \mathscr{E}$ une application affine, et $\overrightarrow{f}$ sa partie linéaire.*

(1) *Si* $\mathrm{Fix}(f)$ *est non vide, c'est un sea de direction* $\mathrm{Fix}(\overrightarrow{f}) = \mathrm{Ker}(\overrightarrow{f} - \mathrm{id}_E)$.

(2) *Si* $\mathrm{Ker}(\overrightarrow{f} - \mathrm{id}_E) = \{0\}$*, alors* f *a un point fixe unique* I.

Démonstration. — (1) Si $\mathrm{Fix}(f) \neq \varnothing$, soit $I \in \mathrm{Fix}(f)$. Alors, pour un point arbitraire $P \in \mathscr{E}$ on a:

$$P \in \mathrm{Fix}(f) \iff f(P) = P \iff \overrightarrow{IP} = \overrightarrow{If(P)} = \overrightarrow{f(I)f(P)} = \overrightarrow{f}(\overrightarrow{IP}) \iff \overrightarrow{IP} \in \mathrm{Fix}(\overrightarrow{f}).$$

Ceci montre qu'alors $\mathrm{Fix}(f)$ est un sea de direction $\mathrm{Fix}(\overrightarrow{f}) = \mathrm{Ker}(\overrightarrow{f} - \mathrm{id}_E)$.

(2) Supposons $\mathrm{Ker}(\overrightarrow{f} - \mathrm{id}_E) = \{0\}$ et fixons une origine $O \in \mathscr{E}$ arbitraire. Pour tout point $P \in \mathscr{E}$, on a

$$\overrightarrow{Of(P)} = \overrightarrow{Of(O)} + \overrightarrow{f(O)f(P)} = \overrightarrow{Of(O)} + \overrightarrow{f}(\overrightarrow{OP})$$

et donc on a:

$$P = f(P) \iff \overrightarrow{OP} = \overrightarrow{Of(P)} \iff (\overrightarrow{f} - \mathrm{id}_E)(\overrightarrow{OP}) = -\overrightarrow{Of(O)}.$$

Or, d'après hypothèse, $\overrightarrow{f} - \mathrm{id}_E$ est injective, donc **bijective** puisque E est de dimension finie. Donc il existe un unique vecteur $\overrightarrow{u} \in E$ tel que $(\overrightarrow{f} - \mathrm{id}_E)(\overrightarrow{u}) = -\overrightarrow{Of(O)}$, et il existe donc un unique point $I \in \mathscr{E}$ tel que $(\overrightarrow{f} - \mathrm{id}_E)(\overrightarrow{OI}) = -\overrightarrow{Of(O)}$, i.e. I est l'unique point fixe de f. La proposition est démontrée. $\square$

CHAPITRE 3

APPLICATIONS MULTILINÉAIRES, GROUPE SYMÉTRIQUE, DÉTERMINANT

Résumé: On introduit les formes multilinéaires, puis on étudie le groupe symétrique, c'est à dire l'ensemble des permutations de $\{1, \ldots, n\}$. En particulier on définit la signature d'une permutation ce qui nous permet d'étudier précisément les formes n-linéaires antisymétriques sur un espace de dimension n et ainsi définir le déterminant d'une famille de vecteurs. On introduit ensuite le déterminant d'une matrice, et démontrons ses propriétés. Enfin, nous terminons le chapitre en définissant pour un endomorphisme, le déterminant, le polynome caractéristique, la trace, les valeurs et vecteurs propres.

3.1. Applications multilinéaires

***Définition 3.1.1* (Espace des applications à valeurs dans un espace vectoriel)**
Soient X un ensemble et F un k-espace vectoriel, on note $\mathrm{Applic}(X, F)$ l'ensemble de toutes les applications $X \to F$. Si $f, g \in \mathrm{Applic}(X, F)$ et $\lambda \in k$, on définit l'application $\lambda \cdot f + g$ par

$$(\lambda \cdot f + g)(x) = \lambda \cdot f(x) + g(x)$$

pour tout $x \in X$; on vérifie facilement que ceci munit $\mathrm{Applic}(X, F)$ d'une structure de k-espace vectoriel.

***Définition 3.1.2* (Applications et formes p-linéaires)**. — Soit p un entier ≥ 2.

(1) Soient $E_1, \ldots, E_p$ et F des k-espaces vectoriels; une application

$$f : E_1 \times \cdots \times E_p \to F$$

est dite *p-linéaire* ou *multilinéaire* (si $p = 2$, on dit *bilinéaire*) si elle est *linéaire par rapport à chacune des variables* (les autres variables étant fixées), c.-à-d., si pour tout $i = 1, \ldots, p$, on a :

$$(*_i) \qquad f(x_1, \ldots, t_i \cdot x_i + x'_i, \ldots, x_p) = t_i \cdot f(x_1, \ldots, x_i, \ldots, x_p) + f(x_1, \ldots, x'_i, \ldots, x_p)$$

où chaque x_j est dans E_j (et $x'_i \in E_i$), et $t_i \in k$. On notera $\mathscr{L}(E_1, \ldots, E_p; F)$ l'ensemble de ces applications, et si tous les E_i sont égaux à un même espace E, on le notera simplement $\boxed{\mathscr{L}_p(E, F).}$

Si $f, g \in \mathscr{L}(E_1, \ldots, E_p; F)$ et $\lambda \in k$, on voit aisément que l'application $\lambda \cdot f + g : E_1 \times \cdots \times E_p \to F$ vérifie $(*_i)$ pour tout i; ceci montre que $\mathscr{L}(E_1, \ldots, E_p; F)$ est un k-espace vectoriel (c'est un sous-espace vectoriel de $\mathrm{Applic}(E_1 \times \cdots \times E_p, F)$).

(2) Lorsque l'espace d'arrivée F est le corps k, un élément de $\mathscr{L}_p(E, k)$, i.e. une application p-linéaire $\underbrace{E \times \cdots \times E}_{p \text{ facteurs}} \to k$, s'appelle une *forme p-linéaire* sur E (et *forme bilinéaire* si $p = 2$).

Remarque 3.1.3. — Il résulte de la définition que $f(x_1, \ldots, x_n) = 0$ s'il existe un indice i tel que $x_i = 0$ (puisque f est linéaire par rapport à la i-ème variable).

Remarque 3.1.4. — Attention, si l'on considère, pour $p \geq 2$, l'espace vectoriel

$$E_1 \times \cdots \times E_p = E_1 \oplus \cdots \oplus E_p$$

une application multilinéaire $E_1 \times \cdots \times E_p \to F$ **n'est pas** la même chose qu'une application linéaire $E_1 \oplus \cdots \oplus E_p \to F$. Par exemple, si $t \in k$ on a:

$$f(tx_1, \ldots, tx_p) = \begin{cases} t\, f(x_1, \ldots, x_p) & \text{si } f \text{ est linéaire;} \\ t^p\, f(x_1, \ldots, x_p) & \text{si } f \text{ est multilinéaire;} \end{cases}$$

[(0)] version du 12/06/2015

d'autre part, prenant $p = 2$ pour simplifier, $f(x_1 + y_1, x_2 + y_2)$ égale :

$$\begin{cases} f(x_1, x_2) + f(y_1, y_2) & \text{si } f \text{ est linéaire;} \\ f(x_1, x_2) + f(x_1, y_2) + f(y_1, x_2) + f(y_1, y_2) & \text{si } f \text{ est multilinéaire.} \end{cases}$$

***Exemple 3.1.5* (Produit scalaire sur $\mathbb{R}^3$).** — Soit $\phi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ le produit scalaire "euclidien" usuel:

$$\phi(X, Y) = x_1y_1 + x_2y_2 + x_3y_3 \qquad \text{si} \qquad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

alors $\phi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ est une application bilinéaire, c.-à-d.,

$$\phi(X + sX', Y + tY') = \phi(X, Y) + s\phi(X', Y) + t\phi(X, Y') + st\phi(X', Y').$$

Comme ϕ est à valeurs dans le corps des scalaires (ici $\mathbb{R}$), on dit que ϕ est une <u>forme</u> bilinéaire sur $\mathbb{R}^3$. De plus, on dit que ϕ est <u>symétrique</u> car

$$\phi(X, Y) = \phi(Y, X), \qquad \forall X, Y \in \mathbb{R}^3.$$

Plus généralement, pour tout $n \geq 1$, le produit scalaire "euclidien" sur $\mathbb{R}^n$ défini par

$$\phi(X, Y) = x_1y_1 + \cdots + x_ny_n \qquad \text{si} \qquad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

est une forme bilinéaire symétrique sur $\mathbb{R}^n$.

***Exemple 3.1.6* (Produit vectoriel dans $\mathbb{R}^3$ "euclidien").** — Le produit vectoriel $\wedge : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ défini par

$$(*) \qquad X \wedge Y = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix} \qquad \text{si} \qquad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

est une application bilinéaire $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ alternée, c.-à-d., $X \wedge X = 0$ pour tout X. Pour $X, Y \in \mathbb{R}^3$ arbitraires, ceci entraîne: $0 = (X + Y) \wedge (X + Y) = X \wedge X + Y \wedge Y + X \wedge Y + Y \wedge X$ d'où

$$X \wedge Y = -Y \wedge X, \qquad \forall X, Y \in \mathbb{R}^3$$

donc le produit vectoriel est antisymétrique (ce qu'on voit aussi directement d'après la définition $(*)$).

***Exemple 3.1.7* (Couplage $V^* \times V \to k$).** — Soient V un k-espace vectoriel, V^* l'espace dual. Alors l'application $V^* \times V \to k$, $(f, v) \mapsto f(v)$ est bilinéaire.

3.2. Groupes symétriques: quelques propriétés

***Définition 3.2.1* (Groupe symétrique S_n).** — Soit S_n le groupe des permutations de $\{1, \ldots, n\}$, i.e. des bijections de $\{1, \ldots, n\}$ sur lui-même [1]. C'est un groupe pour la composition des applications, et il est de cardinal $n!$, car une permutation σ est déterminée par la donnée de $\sigma(1)$, pour lequel il y a n choix, puis de $\sigma(2)$, pour lequel il reste $n - 1$ choix, etc.

Notation 3.2.2. — On représente en général un élément σ de S_n par son écriture "à deux lignes": sur la première ligne, on écrit $1, 2, 3, \ldots, n$, dans cet ordre, et sur la seconde on écrit les nombres $\sigma(1), \sigma(2), \sigma(3), \ldots, \sigma(n)$. Ainsi, par exemple,

$$\sigma = \begin{pmatrix} 123456 \\ 356124 \end{pmatrix} \qquad \text{et} \qquad \tau = \begin{pmatrix} 123456 \\ 615234 \end{pmatrix}$$

désignent les éléments de S_6 tels que $\sigma(1) = 3$, $\sigma(2) = 5$, $\sigma(3) = 6$, $\sigma(4) = 1$, $\sigma(5) = 2$, $\sigma(6) = 4$ et, respectivement, $\tau(1) = 6$, $\tau(2) = 1$, $\tau(3) = 5$, $\tau(4) = 2$, $\tau(5) = 3$, $\tau(6) = 4$. Par définition, le produit $\sigma\tau$ est la bijection composée $\sigma \circ \tau$, i.e. on fait d'abord τ puis σ, c.-à-d., $\sigma\tau(1) = \sigma(\tau(1)) = \sigma(6) = 4$, etc. On trouve alors:

$$\sigma\tau = \begin{pmatrix} 123456 \\ 432561 \end{pmatrix} \qquad \text{et} \qquad \tau\sigma = \begin{pmatrix} 123456 \\ 534612 \end{pmatrix}.$$

On notera que le groupe S_n n'est pas commutatif en général (il l'est seulement pour $n = 1$ ou 2).

---

[1] Il est parfois noté Σ_n (= "Sigma"n)

Pour certaines permutations, on utilise une écriture plus condensée, introduite ci-dessous.

***Définition 3.2.3* (Transpositions et cycles).** — Pour $i \neq j$, on note (ij) la permutation qui échange i et j et laisse les autres éléments inchangés; on dit que (ij) est une **transposition**. Remarquons que si τ est une transposition, alors $\tau^2 = \mathrm{id}$ et donc $\tau = \tau^{-1}$.

Plus généralement, pour $r \geq 2$, on dit que $c \in S_n$ est un r**-cycle** s'il existe $i_1, \ldots, i_r$, deux à deux distincts, tels que $c(j) = j$ pour $j \notin \{i_1, \ldots, i_r\}$ et

$$c(i_1) = i_2, \quad c(i_2) = i_3, \quad \cdots, \quad c(i_{r-1}) = i_r, \quad c(i_r) = i_1.$$

Dans ce cas, on note $c = (i_1 i_2 \cdots i_r)$, et l'on dit que l'ensemble $\{i_1, \ldots, i_r\}$ est le **support** du cycle c.

Par exemple, dans S_6, les cycles $c_1 = (253)$ et $c_2 = (1635)$ désignent, respectivement, les permutations suivantes:

$$\begin{pmatrix} 123456 \\ 152436 \end{pmatrix} \qquad \text{et} \qquad \begin{pmatrix} 123456 \\ 625413 \end{pmatrix}.$$

Le terme de "cycle" s'explique si l'on représente graphiquement l'effet du cycle c sur les éléments $1, \ldots, n$: dans l'exemple précédent, pour c_1 et c_2 on a respectivement :

(c_1)
$$\begin{tikzcd} & 2 \arrow[dr] & & & & \\ 3 \arrow[ur] & & 5 \arrow[ll] & 1 & 4 & 6 \end{tikzcd}$$

(c_2)
$$\begin{tikzcd} 1 \arrow[r] & 6 \arrow[d] & & \\ 5 \arrow[u] & 3 \arrow[l] & 2 & 4 \end{tikzcd}$$

i.e. les éléments hors du support de c sont inchangés, et c permute "cycliquement" les éléments de son support.

On voit ainsi que l'inverse de c est le r-cycle de même support, que l'on parcourt en sens inverse: par exemple, $c_1^{-1} = (235)$ et $c_2^{-1} = (1536)$. On voit aussi que c est un élément *d'ordre* r, c.-à-d., $c^r = \mathrm{id}$ mais $c^s \neq \mathrm{id}$ pour $s = 1, \ldots, r-1$. En particulier, on a donc $c^{-1} = c^{r-1}$.

Remarque 3.2.4. — Soit X un ensemble, *réunion disjointe* de sous-ensembles $A_1, \ldots, A_d$, et pour chaque $i = 1, \ldots, d$, soit c_i une permutation de A_i, i.e. une bijection de A_i dans lui-même; on étend c_i en une bijection σ_i de X en prolongeant c_i par l'identité hors de A_i, i.e. on pose $\sigma_i(x) = c_i(x)$ si $x \in A_i$ et $\sigma_i(x) = x$ si $x \notin A_i$. Alors on voit facilement que la composée $\sigma = \sigma_1 \circ \cdots \circ \sigma_d$ *ne dépend pas de l'ordre des facteurs*, c.-à-d., $\sigma = \sigma_d \circ \cdots \circ \sigma_1$, etc. En effet, quel que soit l'ordre des facteurs, le produit est la permutation σ de X qui coïncide avec c_i sur chaque A_i, i.e. $\sigma(x) = c_i(x)$ si $x \in A_i$ (ceci est bien défini, puisque les A_i sont disjoints et leur réunion est X). On a donc obtenu le résultat suivant:

$(\star)$ $$\boxed{\text{Dans } S_n\text{, des cycles de supports disjoints commutent.}}$$

Par exemple, dans S_7, si $\sigma = (257)$ et $\tau = (1364)$ alors

$$\sigma\tau = \tau\sigma = \quad \begin{tikzcd} & 2 \arrow[dr] & & & 1 \arrow[r] & 3 \arrow[d] \\ 7 \arrow[ur] & & 5 \arrow[ll] & & 4 \arrow[u] & 6 \arrow[l] \end{tikzcd}$$

***Théorème 3.2.5* (Décomposition en produit de cycles de supports disjoints)**
Soit $\sigma \in S_n$. Alors il existe des cycles $c_1, \ldots, c_d$ de supports disjoints, uniquement déterminés, tels que $\sigma = c_1 \cdots c_n$. (Et donc l'écriture précédente est unique, à l'ordre des facteurs près).

On admettra ce résultat (la démonstration n'est pas difficile, mais un peu pénible à écrire). Illustrons ceci par un exemple: soit

$$\sigma = \begin{pmatrix} 123456789 \\ 356184927 \end{pmatrix} \in S_9.$$

Choisissons un élément, par exemple 1, et regardons la suite de ses transformés par σ: $1 \to 3 \to 6 \to 4 \to 1$, ceci donne le cycle (1364) auquel appartient l'élément 1. Puis choisissons un élément parmi les éléments restants, par exemple 2, et regardons à nouveau la suite de ses transformés par σ: $2 \to 5 \to 8 \to 2$, ceci donne le cycle (258) auquel appartient l'élément 2. Puis on recommence ... Ici, la dernière étape consiste à choisir, disons, l'élément 7, et l'on obtient la transposition, i.e. le 2-cycle, (79), d'où la décomposition

$$\sigma = (1364)(258)(79) = \begin{tikzcd} & 2 \arrow[dr] & & 1 \arrow[r] & 3 \arrow[d] & 7 \arrow[d, leftrightarrow] \\ 8 \arrow[ur] & & 5 \arrow[ll] & 4 \arrow[u] & 6 \arrow[l] & 9 \end{tikzcd}$$

Théorème 3.2.6 **(S_n est engendré par les transpositions)**. — *Les transpositions $s_i = (i, i+1)$, pour $i = 1, \ldots, n-1$, engendrent S_n.*

Démonstration. — On procède par récurrence sur n. C'est clair pour $n = 2$ car $S_2 = \{\mathrm{id}, (12)\}$. Supposons $n \geq 3$ et le résultat établi pour $n-1$. On identifie S_{n-1} au sous-groupe de S_n formé des permutations τ telles que $\tau(n) = n$. Posons $s_i = (i, i+1)$, pour $i = 1, \ldots, n-1$, et notons H le sous-groupe de S_n engendré par les s_i.

Soit $\sigma \in S_n$. Si $\sigma(n) = n$, alors $\sigma \in S_{n-1}$ et donc, par hypothèse de récurrence, σ appartient au sous-groupe engendré par les s_i, pour $i \leq n-2$, donc a fortiori $\sigma \in H$. On peut donc supposer que $\sigma(n) = i < n$. Mais alors, $s_i\sigma(n) = i+1$, et si $i+1 < n$ alors $s_{i+1}s_i\sigma(n) = i+2$, etc. En continuant ainsi, on arrive à l'égalité $s_{n-1} \cdots s_i\sigma(n) = n$. Alors, d'après ce qui précède, $\tau = s_{n-1} \cdots s_i\sigma$ appartient à H.

Multipliant l'égalité $\tau = s_{n-1} \cdots s_i\sigma$ à gauche par $s_{n-1}^{-1} = s_{n-1}$, puis par $s_{n-2}^{-1} = s_{n-2}$, etc., on arrive à l'égalité $\sigma = s_i \cdots s_{n-1}\tau$. Comme τ et les s_j appartiennent à H, il en est de même de σ. Ceci prouve que $H = S_n$, d'où le théorème. $\square$

Lemme 3.2.7. — *Soit k un corps. Tout élément $\sigma \in S_n$ induit un automorphisme ϕ_σ de la k-algèbre $k[X_1, \ldots, X_n]$, défini par*

$$(*) \qquad \phi_\sigma(X_i) = X_{\sigma(i)}, \quad \forall i = 1, \ldots, n.$$

$$(**) \qquad \phi_\sigma(P)(X_1, \ldots, X_n) = P(X_{\sigma(1)}, \ldots, X_{\sigma(n)}), \qquad \forall P \in k[X_1, \ldots, X_n].$$

De plus, l'application $\sigma \mapsto \phi_\sigma$ est un morphisme de groupes.

Démonstration. — D'après la propriété universelle de $A = k[X_1, \ldots, X_n]$, il existe, pour tout $\sigma \in S_n$, un unique morphisme de k-algèbres $\phi_\sigma : A \to A$ vérifiant $(*)$ et $(**)$. De plus, il résulte de $(*)$ que $\phi_{\mathrm{id}} = \mathrm{id}_A$ et que $\phi_\sigma \circ \phi_\tau = \phi_{\sigma\tau}$.

Ceci entraîne, d'une part, que chaque ϕ_σ est un automorphisme de A, d'inverse $\phi_{\sigma^{-1}}$, et, d'autre part, que l'application $\sigma \mapsto \phi_\sigma$ est un morphisme de groupes, de S_n dans le groupe des automorphismes de k-algèbre de A. $\square$

Notation. Pour tout $P \in k[X_1, \ldots, X_n]$, on écrira simplement $\sigma(P)$ au lieu de $\phi_\sigma(P)$.

Théorème et définition 3.2.8 **(Signature d'une permutation)**. — *Soit n un entier ≥ 2.*

(i) *Il existe un unique morphisme de groupes $\varepsilon : S_n \to \{\pm 1\}$ tel que $\varepsilon(\tau) = -1$ pour toute transposition $\tau = (ij)$. On l'appelle la* **signature**.

(ii) *Par conséquent, si l'on écrit un élément $\sigma \in S_n$ comme produit de transpositions de deux façons différentes : $\sigma = s_1 \cdots s_p = t_1 \cdots t_q$ où les s_i et t_j sont des transpositions, alors p et q ont la même parité, puisque $\varepsilon(\sigma) = (-1)^p = (-1)^q$.*

(iii) *Enfin, pour tout $\sigma \in S_n$, on a $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)$.*

Démonstration. — Remarquons d'abord que ε, s'il existe, est nécessairement unique, puisque les transpositions engendrent S_n, d'après le théorème 3.2.6. Il s'agit donc de montrer l'existence. (Ceci équivaut à montrer que, si $s_1, \ldots, s_p$ et $t_1, \ldots, t_q$ sont des transpositions et si $s_1 \cdots s_p = t_1 \cdots t_q$, alors p et q ont la même parité, ce qui n'est pas évident *a priori*.)

D'après le lemme précédent, S_n opère par automorphismes d'algèbre sur $A = \mathbb{Q}[X_1, \ldots, X_n]$. Considérons le polynôme

$$V_n = \prod_{1 \leq i < j \leq n} (X_i - X_j).$$

Soit $\sigma \in S_n$. Alors $\sigma(V_n) = \prod_{1 \leq i < j \leq n} (X_{\sigma(i)} - X_{\sigma(j)})$ et, pour tout $i < j$,

$$(1) \qquad \sigma(X_i - X_j) = \begin{cases} X_{\sigma(i)} - X_{\sigma(j)}, & \text{si } \sigma(i) < \sigma(j); \\ -(X_{\sigma(j)} - X_{\sigma(i)}), & \text{si } \sigma(i) > \sigma(j). \end{cases}$$

On dit qu'un couple (i, j) avec $i < j$ est une *inversion* de σ si $\sigma(i) > \sigma(j)$; on introduit le **nombre d'inversions** de σ:

$$\boxed{\ell(\sigma) = |\{i < j \mid \sigma(i) > \sigma(j)\}|} \tag{2}$$

et l'on définit la signature de σ par:

$$\boxed{\varepsilon(\sigma) = (-1)^{\ell(\sigma)}.} \tag{3}$$

On déduit alors de (1) que, pour tout $\sigma \in S_n$,

$$\sigma(V_n) = (-1)^{\ell(\sigma)} V_n. \tag{4}$$

Ceci entraîne que $\varepsilon : S_n \to \{\pm 1\}$ est un morphisme de groupes: en effet, pour $\sigma, \tau \in S_n$ on a

$$\varepsilon(\sigma\tau)V_n = (\sigma\tau)(V_n) = \sigma(\varepsilon(\tau)V_n) = \varepsilon(\tau)\varepsilon(\sigma)V_n,$$

d'où $\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau)$. En particulier, pour tout σ on a $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)^{-1}$, et comme $\varepsilon(\sigma) \in \{\pm 1\}$, on a $\varepsilon(\sigma)^{-1} = \varepsilon(\sigma)$. Ceci prouve déjà l'assertion (iii).

Enfin, pour $i < j$ notons τ_{ij} la transposition qui échange i et j. On vérifie facilement que les inversions de τ_{ij} sont les couples (i, j) et (i, k), (k, j) pour $i < k < j$; leur nombre est $1 + 2(j - i - 1)$, d'où $\varepsilon(\tau_{ij}) = -1$. Ceci prouve l'assertion (i) et, comme noté dans le théorème, l'assertion (ii) en découle aussitôt. Le théorème est démontré. □

Définitions 3.2.9. — 1) On dit qu'une permutation $\sigma \in S_n$ est **paire**, resp. **impaire**, si $\varepsilon(\sigma) = 1$, resp. -1. Ceci équivaut à dire que σ s'écrit comme produit d'un nombre pair (resp. impair) de transpositions.

2) $\mathrm{Ker}(\varepsilon) = \{\sigma \in S_n \mid \varepsilon(\sigma) = 1\}$ est appelé **groupe alterné** d'ordre n, et noté A_n. Il est formé des permutations paires, et est de cardinal $n!/2$. (En effet, si τ est une transposition (par exemple, $\tau = \tau_{12}$), alors S_n est la réunion disjointe de A_n et de $\tau A_n = \{\tau\sigma \mid \sigma \in A_n\}$.)

***Proposition 3.2.10* (Signature d'un cycle)**. — *Tout r-cycle est de signature* $(-1)^{r-1}$.

Démonstration. — Soit $c = (i_1 \cdots i_r)$ un r-cycle. On voit facilement que le produit ci-dessous de $(r-1)$ transpositions :

$$(i_1 i_2)(i_2 i_3) \cdots (i_{r-1} i_r)$$

envoie i_1 sur i_2, i_2 sur i_3, …, i_{r-1} sur i_r, et i_r sur i_1, donc égale c. Par conséquent, $\varepsilon(c) = (-1)^{r-1}$. □

***Définition 3.2.11* (Diagramme d'une permutation)**. — [2] Soit $\sigma \in S_n$; on a défini l'ensemble de ses inversions:

$$I(\sigma) = \{(i, i') \in \{1, \dots, n\}^2 \mid i < i' \text{ et } \sigma(i) > \sigma(i')\}$$

et noté $\ell(\sigma)$ le cardinal de $I(\sigma)$. On définit le diagramme $D(\sigma)$ de σ par

$$D(\sigma) = (\mathrm{id} \times \sigma) I(\sigma) = \{(i, \sigma(i')) \in \{1, \dots, n\}^2 \mid i < i' \text{ et } \sigma(i) > \sigma(i')\}.$$

Alors, d'une part, $D(\sigma)$ a même cardinal $\ell(\sigma)$ que $I(\sigma)$. D'autre part, en posant $j = \sigma(i')$, de sorte que $i' = \sigma^{-1}(j) > i$, on voit que

$$D(\sigma) = \{(i, j) \in \{1, \dots, n\}^2 \mid j < \sigma(i) \text{ et } \sigma^{-1}(j) > i\}.$$

Si l'on dessine le quadrillage du carré $[0, n] \times [0, n]$ donné par les droites $x = 0, 1, \dots, n$ et $y = 0, 1, \dots, n$ et qu'on représente le graphe de σ en mettant une croix au centre de la case (carré de côté 1):

$$\{(x, y) \mid i - 1 \le x \le i, \quad \sigma(i) - 1 \le y \le \sigma(i)\},$$

alors on voit que $D(\sigma)$ est formé des cases qui restent lorsqu'on a barré toutes les cases qui sont à droite ou bien au-dessus d'un point du graphe. Ceci donne un moyen graphique commode pour compter $\ell(\sigma) =$ le nombre d'inversions de σ. En particulier, si σ est la transposition (i, i') avec $i < i'$, on voit que le diagramme de σ est une "équerre", formée de la bande horizontale (à $i' - i$ cases):

$$\{(x, y) \mid i - 1 \le x \le i' - 1, \quad i - 1 \le y \le i'\}$$

et de la bande verticale (à $i' - i$ cases):

$$\{(x, y) \mid i - 1 \le x \le i, \quad i - 1 \le y \le i' - 1\},$$

dont l'intersection est la case $i - 1 \le x, y \le i$. On retrouve ainsi que $\ell((i, i')) = 2(i' - i) - 1$.

[2] Cette définition est donnée pour les lecteurs intéressés; elle n'interviendra pas dans les évaluations.

3.3. Applications multilinéaires antisymétriques et déterminant d'une famille de vecteurs

Désormais, on fixe un corps k.

Définition et proposition 3.3.1 **(Applications p-linéaires antisymétriques)**
Soient E, F deux k-espaces vectoriels et $f \in \mathscr{L}_p(E,F)$ (i.e. f est une application p-linéaire $E^p \to F$). On dit que f est **antisymétrique** *si elle vérifie les conditions équivalentes suivantes:*

$$(\dagger) \qquad \forall i \neq j, \qquad f(x_1,\dots,x_j,\dots,x_i,\dots,x_p) = -f(x_1,\dots,x_i,\dots,x_j,\dots,x_p)$$

$$(\dagger') \qquad \forall \sigma \in S_p, \qquad f(x_{\sigma(1)},\dots,x_{\sigma(p)}) = \varepsilon(\sigma) f(x_1,\dots,x_p)$$

Démonstration. — L'équivalence de $(\dagger)$ et $(\dagger')$ résulte du théorème 3.2.8. □

Proposition 3.3.2. — *On suppose que k n'est pas de caractéristique 2 (c'est à dire $2 \neq 0$ dans k). Soit $f : E^p \to F$ une application p-linéaire antisymétrique. Alors pour toute famille* ***liée*** $(v_1,\dots,v_p) \in E^p$*, on a*

$$f(v_1,\dots,v_p) = 0.$$

Démonstration. — Remarquons d'abord que sous nos hypothèses, si deux vecteurs v_i et v_j d'une famille $(v_1,\dots,v_p)$ sont égaux, alors $f(v_1,\dots,v_p) = 0$. En effet,

$$-f(v_1,\dots,v_i,\dots,v_j,\dots,v_p) = f(v_1,\dots,v_j,\dots,v_i,\dots,v_p) = f(v_1,\dots,v_i,\dots,v_j,\dots,v_p)$$

(la seconde égalité puisque $v_i = v_j$), d'où $2f(v_1,\dots,v_p) = 0$. Donc, si k est de caractéristique $\neq 2$, $f(v_1,\dots,v_p) = 0$.

Soit maintenant $(v_1,\dots,v_p)$ une famille liée quelconque. L'un de ses éléments peut s'écrire comme combinaison linéaire des autres. Supposons que cet élément est v_1 et écrivons donc $v_1 = a_2v_2 + \ldots + a_pv_p$ avec $a_2,\dots,a_p \in k$. Par linéarité,

$$f(v_1,\dots,v_p) = a_2 f(v_2,v_2,\dots,v_p) + \cdots + a_p f(v_p,v_2,\dots,v_p).$$

Or, la remarque ci-dessus implique que chacun des termes du membre de droite s'annule. D'où $f(v_1,\dots,v_p) = 0$. □

Examinons de plus près le cas d'une forme α que l'on suppose n-linéaire antisymétrique sur un k-espace vectoriel E de dimension n. Nous allons voir que si $\mathscr{B} = (e_1,\dots,e_n)$ est une base de E, alors α est entièrement déterminée par la valeur du scalaire $\alpha(e_1,\dots,e_n)$.

Soient $v_1,\dots,v_n \in V$. Pour chaque $i,j = 1,\dots,n$ on note v_{ij} le j-ième coefficient de v_i décomposé dans la base $\mathscr{B}$:

$$v_i = \sum_{j=1}^n v_{ij} e_j.$$

Alors, en utilisant successivement la linéarité par rapport à chacune des variables, on peut écrire:

$$\begin{aligned}
\alpha(v_1,\dots,v_n) &= \sum_{j=1}^n v_{1j}\alpha(e_j,v_2,\dots,v_n) \\
&= \sum_{j_1,j_2=1}^n v_{1j_1}v_{2j_2}\alpha(e_{j_1},e_{j_2},v_3,\dots,v_n) \\
&\;\vdots \\
&= \sum_{j_1,j_2,\dots,j_n=1}^n v_{1j_1}v_{2j_2}\cdots v_{nj_n}\alpha(e_{j_1},e_{j_2},\dots,e_{j_n})
\end{aligned}$$

Remarquons maintenant que si deux indices $j_i, j_{i'}$ avec $i \neq i'$ coincident, alors par antisymétrie, $\alpha(e_{j_1},e_{j_2},\dots,e_{j_n}) = 0$. Donc dans la somme ci-dessus, seuls restent les termes pour lesquels l'application $\{1,\dots,n\} \to \{1,\dots,n\}, i \mapsto j_i$ est injective, et donc bijective. Nous pouvons donc écrire

$$\alpha(v_1,\dots,v_n) = \sum_{\sigma\in S_n} v_{1\sigma(1)}\cdots v_{n\sigma(n)}\alpha(e_{\sigma(1)},\dots,e_{\sigma(n)}),$$

et par antisymétrie

$$(*) \qquad \alpha(v_1,\dots,v_n) = \alpha(e_1,\dots,e_n)\sum_{\sigma\in S_n}\varepsilon(\sigma)v_{1\sigma(1)}\cdots v_{n\sigma(n)}.$$

Ce qui précède nous mène à la définition suivante.

Définition 3.3.3. — Soient E un k-espace vectoriel de dimension n, $\mathscr{B}$ une base de E et $(v_1, \ldots, v_n)$ une famille de n vecteurs de E. Pour chaques $i, j = 1, \ldots, n$ on note v_{ij} le j-ième coefficient de v_i décomposé dans la base $\mathscr{B}$. On appelle **déterminant** de $(v_1, \ldots, v_n)$ dans la base $\mathscr{B}$, le scalaire

$$\mathrm{d\acute{e}t}_{\mathscr{B}}(v_1, \ldots, v_n) = \sum_{\sigma \in S_n} \varepsilon(\sigma) v_{1\sigma(1)} \cdots v_{n\sigma(n)}.$$

Proposition 3.3.4. — *Soient E un k-espace vectoriel de dimension n, et $\mathscr{B} = (e_1, \ldots, e_n)$ une base de E. Alors,*

(1) *L'application* $\mathrm{d\acute{e}t}_{\mathscr{B}} : E^n \to k$ *est n-linéaire antisymétrique.*
(2) *Toute forme n-linéaire antisymétrique α sur E est un multiple de* $\mathrm{d\acute{e}t}_{\mathscr{B}}$*, plus précisément:*

$$\forall (v_1, \ldots, v_n) \in E^n, \quad \alpha(v_1, \ldots, v_n) = \alpha(e_1, \ldots, e_n)\, \mathrm{d\acute{e}t}_{\mathscr{B}}(v_1, \ldots, v_n).$$

Démonstration. — (1) se démontre comme suit. La n-linéarité étant évidente, il ne reste qu'à démontrer l'antisymétrie. Soit $\tau \in S_n$ une permutation quelconque. Alors pour toute famille de vecteurs $(v_1, \ldots, v_n)$, on a par définition

$$\mathrm{d\acute{e}t}_{\mathscr{B}}(v_{\tau(1)}, \ldots, v_{\tau(n)}) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} v_{\tau(i)\sigma(i)}.$$

Comme τ est une bijection, on peut réindicer le produit en posant $j = \tau(i)$, ce qui donne

$$\prod_{i=1}^{n} v_{\tau(i)\sigma(i)} = \prod_{j=1}^{n} v_{j\sigma(\tau^{-1}(j))}.$$

En utilisant ensuite la propriété de la signature $\varepsilon(\sigma) = \varepsilon(\tau)\varepsilon(\sigma \circ \tau^{-1})$, nous obtenons:

$$\mathrm{d\acute{e}t}_{\mathscr{B}}(v_{\tau(1)}, \ldots, v_{\tau(n)}) = \varepsilon(\tau) \sum_{\sigma \in S_n} \varepsilon(\sigma \circ \tau^{-1}) \prod_{j=1}^{n} v_{j\sigma(\tau^{-1}(j))}.$$

Enfin, comme $\sigma \mapsto \sigma \circ \tau^{-1}$, $S_n \to S_n$ est une bijection (en effet, on vérifie aisément qu'elle admet une réciproque, $\sigma \mapsto \sigma \circ \tau$), on peut réindicer la somme en posant $\rho = \sigma \circ \tau^{-1}$. On obtient alors

$$\begin{aligned} \mathrm{d\acute{e}t}_{\mathscr{B}}(v_{\tau(1)}, \ldots, v_{\tau(n)}) &= \varepsilon(\tau) \sum_{\rho \in S_n} \varepsilon(\rho) \prod_{j=1}^{n} v_{j\rho(j)} \\ &= \varepsilon(\tau)\, \mathrm{d\acute{e}t}_{\mathscr{B}}(v_1, \ldots, v_n). \end{aligned}$$

Nous avons bien prouvé que $\mathrm{d\acute{e}t}_{\mathscr{B}}$ est antisymétrique.

(2) résulte immédiatement de la formule (*) établie précédemment.

□

***Proposition 3.3.5* (Déterminant et changement de base).** — *Soient $\mathscr{B}$ et $\mathscr{B}'$ deux bases d'un même espace vectoriel E de dimension n. Alors pour toute famille $(v_1, \ldots, v_n) \in E^n$,*

$$\mathrm{d\acute{e}t}_{\mathscr{B}'}(v_1, \ldots, v_n) = \mathrm{d\acute{e}t}_{\mathscr{B}'} \mathscr{B} \cdot \mathrm{d\acute{e}t}_{\mathscr{B}}(v_1, \ldots, v_n).$$

Démonstration. — D'après le (1) de la proposition 3.3.4, $\mathrm{d\acute{e}t}_{\mathscr{B}'}$ est une application multilinéaire antisymétrique. On peut donc lui appliquer le (2) de la proposition 3.3.4 et obtenir $\mathrm{d\acute{e}t}_{\mathscr{B}'} = \mathrm{d\acute{e}t}_{\mathscr{B}'} \mathscr{B} \cdot \mathrm{d\acute{e}t}_{\mathscr{B}}$, comme voulu. □

***Remarque 3.3.6* (Interprétation géométrique du déterminant: aire et volume)**

Etant donnés deux vecteurs u et v de $\mathbb{R}^2$, on note $\mathscr{A}$ l'aire orienté du parallélogramme engendré par u et v (dans les notations du chapitre 2, il s'agit d'un parallélogramme ABCD du plan affine tel que $\overrightarrow{AB} = u$ et $\overrightarrow{AD} = v$; un tel parallélogramme est unique à translation près). Ici "orientée" signifie que l'on compte l'aire positivement si l'on passe de u à v dans le sens positif (comme sur le dessin, et négativement sinon. On se convainc facilement sur un dessin que le parallélogramme engendré par $u_1 + u_2$ et v peut se voir comme la réunion d'un parallélogramme engendré par u_1 et v et d'un autre engendré par u_2 et v, d'où l'on tire l'égalité

$$\mathscr{A}(u_1 + u_2, v) = \mathscr{A}(u_1, v) + \mathscr{A}(u_2, v).$$

On se convainc aussi assez facilement que lorsque l'on étire le parallélogramme dans la direction de u d'un facteur λ, alors l'aire est multipliée par λ:

$$\mathscr{A}(\lambda u, v) = \lambda \mathscr{A}(u, v).$$

En résumé, $\mathscr{A}$ est linéaire en sa première variable, et bien sûr, un raisonnement analogue sur la deuxième variable montrerait que $\mathscr{A}$ est bilinéaire. Enfin, échanger u et v change l'orientation donc le signe de l'aire du parallélogramme. L'application $\mathscr{A}$ est donc bilinéaire antisymétrique sur $\mathbb{R}^2$.

Appliquons le point (2) de la proposition 3.3.4 à $\mathscr{A}$. Si l'on note (e_1, e_2) la base canonique de $\mathbb{R}^2$, on a pour tous $u, v \in \mathbb{R}^2$,

$$\mathscr{A}(u,v) = \mathscr{A}(e_1,e_2)\,\text{dét}_{(e_1,e_2)}(u,v).$$

Or $\mathscr{A}(e_1, e_2) = 1$ puisque c'est l'aire d'un carré de côté 1. D'où $\mathscr{A}(u,v) = \text{dét}_{(e_1,e_2)}(u,v)$. En conclusion:

Le déterminant dans la base canonique de deux vecteurs u, v de $\mathbb{R}^2$
est l'aire orientée du parallélogramme engendré par u et v.

Un raisonnement analogue justifierait que le déterminant dans la base canonique de trois vecteurs u, v, w de $\mathbb{R}^3$ est le volume orienté du parallélépipède engendré par u, v et w; et de même en dimension supérieure. Ceci explique par exemple qu'il y ait un déterminant apparaissant dans la formule de changement de variable dans les intégrales multiples.

3.4. Déterminant des matrices

On fixe un corps k.

Définition 3.4.1. — Le **déterminant** d'une matrice $A = (a_{ij})_{i,j=1,\dots,n} \in M_n(k)$ est le scalaire

$$\text{dét}\, A = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}.$$

Exemple 3.4.2. — Pour une matrice de 2×2, on a

$$\text{dét} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{21}a_{12}.$$

Exercice 3.4.3. — Exprimer comme dans l'exemple ci-dessus le déterminant d'une matrice 3×3.

Proposition 3.4.4. — *(Propriétés du déterminant)*

(1) *Pour toute matrice A, on a* $\text{dét}(A) = \text{dét}({}^tA)$.

(2) *Déterminant de la matrice identité:*

$$\text{dét}(I_n) = 1.$$

(3) *Le déterminant est multiplicatif:*

$$\forall A, B \in M_n(k), \quad \text{dét}(AB) = \text{dét}(A)\,\text{dét}(B).$$

(4) *Une matrice A est inversible si et seulement si* $\text{dét}(A) \neq 0$ *et dans ce cas* $\text{dét}(A^{-1}) = \frac{1}{\text{dét}(A)}$.

(5) *Deux matrices semblables ont même déterminant:*

$$\forall A \in M_n(k), \forall P \in Gl_n(k), \quad \text{dét}(P^{-1}AP) = \text{dét}(A).$$

Remarque 3.4.5. — Une conséquence immédiate des définitions 3.3.3 et 3.4.1 est que le *déterminant d'une matrice $A \in M_n(k)$ est égal au déterminant de ses vecteurs lignes dans la base canonique de k^n*. Il résulte également du (1) de la proposition 3.4.4 que le *déterminant d'une matrice $A \in M_n(k)$ est égal au déterminant de ses vecteurs colonnes dans la base canonique de k^n*.

Démonstration de la proposition 3.4.4. — (1) Soit $A \in M_n(k)$. Par définition,

$$\text{dét}({}^tA) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}.$$

Comme σ est une bijection, on peut réindicer le produit en posant $j = \sigma(i)$, ce qui donne

$$\text{dét}({}^tA) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} a_{j\sigma^{-1}(j)}.$$

L'application $\sigma \mapsto \sigma^{-1}$, $S_n \to S_n$ est une bijection, et l'on peut donc aussi réindicer la somme en posant $\rho = \sigma^{-1}$. On obtient

$$\text{dét}({}^tA) = \sum_{\rho \in S_n} \varepsilon(\rho^{-1}) \prod_{j=1}^{n} a_{j\rho(j)}.$$

Comme $\varepsilon(\rho^{-1}) = \varepsilon(\rho)$ pour tout $\rho \in S_n$, on en déduit bien $\text{dét}({}^tA) = \text{dét}(A)$.

(2) Notons δ_{ij} les coefficients de la matrice identité I_n. Soit $\sigma \in S_n$ une permutation. Si $\sigma \neq \mathrm{id}$, alors, il existe $i \in \{1, \dots, n\}$ tel que $\sigma(i) \neq i$. Pour un tel i, $\delta_{i\sigma(i)} = 0$. Le produit $\varepsilon(\sigma) \prod_{i=1}^n \delta_{i\sigma(i)}$ s'annule donc. Maintenant si $\sigma = \mathrm{id}$, alors on a $\delta_{i\sigma(i)} = 1$ pour tout i et $\varepsilon(\mathrm{id}) = 1$, donc le produit $\varepsilon(\sigma) \prod_{i=1}^n \delta_{i\sigma(i)}$ vaut 1. D'où

$$\mathrm{dét}(I_n) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^n \delta_{i\sigma(i)} = 1.$$

(3) Soit $A, B \in M_n(k)$. Notons $\mathscr{B} = (e_1, \dots, e_n)$ la base canonique de k^n et considérons l'application

$$\alpha : (k^n)^n \to k, \quad (v_1, \dots, v_n) \mapsto \mathrm{dét}_{\mathscr{B}}(Av_1, \dots, Av_n).$$

Il s'agit d'une application n-linéaire antisymétrique et d'après la Proposition 3.3.4, on a

$$\alpha(v_1, \dots, v_n) = \alpha(e_1, \dots, e_n) \, \mathrm{dét}_{\mathscr{B}}(v_1, \dots, v_n),$$

et donc

$$\mathrm{dét}_{\mathscr{B}}(Av_1, \dots, Av_n) = \mathrm{dét}_{\mathscr{B}}(Ae_1, \dots, Ae_n) \, \mathrm{dét}_{\mathscr{B}}(v_1, \dots, v_n)$$

pour tous vecteurs $v_1, \dots, v_n$. En particulier, si $v_1, \dots, v_n$ sont les colonnes de B, $Av_1, \dots, Av_n$ sont les colonnes de AB et $Ae_1, \dots, Ae_n$ sont les colonnes de A. D'après la remarque 3.4.5, on en déduit l'égalité souhaitée

$$\mathrm{dét}(AB) = \mathrm{dét}(A) \, \mathrm{dét}(B).$$

(4) Soit A une matrice. Si A n'est pas inversible, alors les lignes de A sont liées et d'après la proposition 3.3.2, le déterminant des lignes de A dans la base canonique, c'est à dire le déterminant de A est nul. Réciproquement, si A est inversible, alors, d'après (3) et le (2) ci-dessus

$$\mathrm{dét}(A) \, \mathrm{dét}(A^{-1}) = \mathrm{dét}(AA^{-1}) = \mathrm{dét}(I_n) = 1;$$

d'où l'on tire immédiatement $\mathrm{dét}(A) \neq 0$ et $\mathrm{dét}(A^{-1}) = \frac{1}{\mathrm{dét}(A)}$.

(5) C'est une conséquence facile de (3) et (4).

□

Pour calculer un déterminant explicitement, il y a essentiellement 3 méthodes à connaître: les opérations élémentaires sur les lignes ou les colonnes, le cas des matrices triangulaires supérieures (éventuellement par bloc), le développement par rapport à une ligne ou une colonne.

***Proposition 3.4.6* (Opérations élémentaires et déterminant).** — *Soient A, A' deux matrices et $\lambda \in k$.*

(1) *Si A' est obtenue à partir de A en échangeant deux lignes (ou deux colonnes) de A, alors*

$$\mathrm{dét}(A') = -\mathrm{dét}(A).$$

(2) *Si A' est obtenue à partir de A en multipliant une ligne (ou une colonne) de A par λ, alors*

$$\mathrm{dét}(A') = \lambda \, \mathrm{dét}(A).$$

(3) *Si A' est obtenue à partir de A en ajoutant à l'une des lignes (resp. des colonnes) de A un multiple d'une autre ligne (resp. colonne) de A, alors*

$$\mathrm{dét}(A') = \mathrm{dét}(A).$$

Démonstration. — C'est une conséquence immédiate de la remarque 3.4.5 et du fait que le déterminant est n-linéaire et antisymétrique. □

Remarque 3.4.7. — Attention, une conséquence du point (2) de la proposition 3.4.6 est

$$\forall A \in M_n(k), \forall \lambda \in k, \quad \mathrm{dét}(\lambda A) = \lambda^n \, \mathrm{dét}(A).$$

(et non $\lambda \, \mathrm{dét}(A)$!)

***Proposition 3.4.8* (Déterminant de matrices triangulaires).** — (1) *Si A est une matrice triangulaire de termes diagonaux $\lambda_1, \dots, \lambda_n$, alors*

$$\mathrm{dét}(A) = \lambda_1 \cdots \lambda_n.$$

(2) *Plus généralement, si A est une matrice triangulaire par blocs, c.-à-d., de la forme*

$$A = \left(\begin{array}{c|c|c|c} A_1 & * & \cdots & * \\ \hline 0 & A_2 & \ddots & * \\ \hline 0 & \ddots & \ddots & * \\ \hline 0 & \cdots & 0 & A_n \end{array}\right)$$

(où les $*$ désignent des coefficients arbitraires), alors

$$\text{dét}(A) = \text{dét}(A_1) \cdots \text{dét}(A_n).$$

Nous allons commencer par démontrer un lemme qui est un cas particulier de la proposition précédente.

Lemme 3.4.9. — *Soit $A \in M_n(k)$ une matrice qui admet l'écriture par bloc suivante:*

$$A = \left(\begin{array}{c|ccc} 1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array}\right),$$

avec $A' \in M_{n-1}(k)$.

Démonstration. — D'après la proposition 3.4.4,

$$\text{dét}(A) = \text{dét}({}^tA) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}.$$

L'hypothèse faite sur A implique que $a_{\sigma(1)1}$ vaut 1 si $\sigma(1) = 1$ et 0 sinon. La somme ci-dessus ne porte donc que sur les permutations σ qui fixent 1, ce qui revient à sommer sur les permutations de $\{2, \ldots, n\}$:

$$\text{dét}(A) = \sum_{\sigma \text{ permutation de } \{2,\ldots,n\}} \varepsilon(\sigma) a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

Quitte à réindicer, on reconnaît dans le membre de droite le déterminant de A'. □

Nous pouvons maintenant démontrer la proposition 3.4.8.

Démonstration. — 1) Supposons que A soit une matrice triangulaire supérieure:

$$A = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

alors la linéarité par rapport à la première colonne donne

$$\text{dét}(A) = \lambda_1 \,\text{dét} \begin{pmatrix} 1 & * & \cdots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix};$$

puis par le lemme 3.4.9,

$$\text{dét}(A) = \lambda_1 \,\text{dét} \begin{pmatrix} \lambda_2 & * & \cdots & * \\ 0 & \lambda_3 & \ddots & \vdots \\ 0 & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix};$$

le résultat en découle, en répétant l'opération ou en procédant par récurrence sur n.

2) En procédant par récurrence sur n, il suffit de montrer que si

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

où B (resp. D) est une matrice carrée de taille p (resp. q), C est une matrice à p lignes et q colonnes, et 0 désigne la matrice nulle à q lignes et p colonnes, alors

$$\text{dét}(A) = \text{dét}(B) \cdot \text{dét}(D).$$

Pour cela on remarque

$$\begin{pmatrix} B & C \\ 0 & D \end{pmatrix} = \begin{pmatrix} I_p & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} B & C \\ 0 & I_q \end{pmatrix},$$

En appliquant p fois le lemme 3.4.9, on obtient

$$\text{dét} \begin{pmatrix} I_p & 0 \\ 0 & D \end{pmatrix} = \text{dét}(D),$$

et de même

$$\text{dét} \begin{pmatrix} B & C \\ 0 & I_q \end{pmatrix} = \text{dét}(B).$$

Et l'on obtient bien

$$\text{dét} \begin{pmatrix} B & C \\ 0 & D \end{pmatrix} = \text{dét} \begin{pmatrix} I_p & 0 \\ 0 & D \end{pmatrix} \text{dét} \begin{pmatrix} B & C \\ 0 & I_q \end{pmatrix} = \text{dét}(B)\,\text{dét}(D).$$

□

Remarque 3.4.9.1. — Attention! Si on décompose une matrice M sous la forme

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

où A, B, C, D sont toutes des matrices carrées (nécessairement de même taille p, de sorte que M est de taille $2p$), il n'est pas vrai en général que $\text{dét}(M)$ égale $\text{dét}(A)\,\text{dét}(D) - \text{dét}(B)\,\text{dét}(C)$. Par exemple si

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

alors $\text{dét}(A) = 0 = \text{dét}(B) = \text{dét}(C) = \text{dét}(D)$ mais $\text{dét}(M) = -1$.

***Proposition 3.4.10* (Développement par rapport à une ligne ou une colonne)**

Soit $A \in M_n(k)$ une matrice. Pour chaques indices $1 \leq i, j \leq n$, on note Δ_{ij} le déterminant de la matrice $A - L_i - C_j$, c'est à dire la matrice obtenue à partir de A en supprimant la i-ème ligne et la j-ème colonne. Alors,

(1) *(développement par rapport à la j-ème colonne)*

$$\forall j \in \{1, \dots, n\}, \quad \text{dét}(A) = \sum_{i=1}^{n} a_{ij}(-1)^{i+j}\Delta_{ij}.$$

(2) *(développement par rapport à la i-ème ligne)*

$$\forall i \in \{1, \dots, n\}, \quad \text{dét}(A) = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}\Delta_{ij}.$$

Démonstration. — Nous allons montrer le point (1), la démonstration du point (2) serait en tout point analogue. C'est une conséquence de la n-linéarité du déterminant. Soit $j \in \{1, \dots, n\}$ et $A \in M_n(k)$. La linéarité par rapport à la j-ème colonne donne:

$$\text{dét}(A) = \sum_{i=1}^{n} a_{ij}\, \text{dét} \begin{pmatrix} a_{11} & \dots & a_{1(j-1)} & 0 & a_{1(j+1)} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{(i-1)1} & \dots & a_{(i-1)(j-1)} & 0 & a_{(i-1)(j+1)} & \dots & a_{(i-1)n} \\ a_{i1} & \dots & a_{i(j-1)} & 1 & a_{i(j+1)} & \dots & a_{in} \\ a_{(i-1)1} & \dots & a_{(i-1)(j-1)} & 0 & a_{(i-1)(j+1)} & \dots & a_{(i-1)n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n(j-1)} & 0 & a_{n(j+1)} & \dots & a_{nn}. \end{pmatrix}$$

Dans chacun des termes ci-dessus, on applique sur la matrice les opérations consistant à déplacer la j-ème colonne vers la première colonne et à décaller les $j-1$ premières colonnes d'une cran vers la droite. Cela revient à faire un cycle de longueur j sur les colonnes. On sait (voir proposition 3.2.10) qu'un tel cycle peut être réalisé à l'aide de $j-1$ échanges de deux colonnes. Cette opération multiplie donc le déterminant par $(-1)^{j-1}$. On fait ensuite une opération analogue sur la matrice apparaissant dans chaque terme. Pour le i-ème terme, on déplace la i-ème ligne à la première et on décalle d'un cran vers le bas les $i-1$ premières lignes. Cela revient à faire un cycle de longueur i sur les lignes. Comme précédemment, cette opération multiplie donc le déterminant par $(-1)^{i-1}$. On obtient donc:

$$\text{dét}(A) = \sum_{i=1}^{n} a_{ij}(-1)^{i-1+j-1}\,\text{dét}\begin{pmatrix} 1 & a_{i1} & \cdots & a_{i(j-1)} & a_{i(j+1)} & \cdots & a_{in} \\ 0 & a_{11} & \cdots & a_{1(j-1)} & a_{1(j+1)} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{(i-1)1} & \cdots & a_{(i-1)(j-1)} & a_{(i-1)(j+1)} & \cdots & a_{(i-1)n} \\ 0 & a_{(i-1)1} & \cdots & a_{(i-1)(j-1)} & a_{(i-1)(j+1)} & \cdots & a_{(i-1)n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{n1} & \cdots & a_{n(j-1)} & a_{n(j+1)} & \cdots & a_{nn} \end{pmatrix}.$$

Comme $(-1)^{i-1+j-1} = (-1)^{i+j}$, le lemme 3.4.9 donne bien

$$\text{dét}(A) = \sum_{i=1}^{n} a_{ij}(-1)^{i+j}\Delta_{ij}.$$

□

Dans la pratique, pour calculer un déterminant, on ne procède pas véritablement en développant suivant les lignes ou les colonnes (ce qui serait trop fastidieux, et couteux en temps de calcul), mais on essaie de faire des opérations élémentaires sur les lignes ou les colonnes, pour rendre la matrice triangulaire.

Illustrons ceci sur un exemple.

Exemple 3.4.10.1. — Calculons le déterminant suivant[3] (où $n \in \mathbb{Z}$):

$$D = \begin{vmatrix} 2 & 3 & 4 & 5 \\ 7 & 6 & 9 & 8 \\ 10 & 11 & 12 & 13 \\ 14 & 15 & 16 & n \end{vmatrix}$$

Mettant $a_{11} = 2$ en facteur dans la première colonne et remplaçant C_2, C_3 et C_4 par $C_2 - (3/2)C_1$, $C_3 - 2C_1$ et $C_4 - (5/2)C_1$ on obtient:

$$D = 2\begin{vmatrix} 1 & 0 & 0 & 0 \\ 7/2 & -9/2 & -5 & -19/2 \\ 5 & -4 & -8 & -12 \\ 7 & -6 & -12 & n-35 \end{vmatrix} = 2\begin{vmatrix} -9/2 & -5 & -19/2 \\ -4 & -8 & -12 \\ -6 & -12 & n-35 \end{vmatrix}$$

puis remplaçant chaque colonne par son opposé (ce qui introduit un signe $(-1)^3$), mettant 4 en facteur dans la 2ème ligne, puis échangeant L_1 et L_2 (ce qui introduit un signe -1), on obtient :

$$D = 2\cdot 4\cdot(-1)^4\begin{vmatrix} 1 & 2 & 3 \\ 9/2 & 5 & 19/2 \\ 6 & 12 & 35-n \end{vmatrix} = 8\begin{vmatrix} 1 & 2 & 3 \\ 9/2 & 5 & 19/2 \\ 6 & 12 & 35-n \end{vmatrix}$$

puis remplaçant C_2 et C_3 par $C_2 - 2C_1$ et $C_2 - 3C_1$ on obtient :

$$D = 8\begin{vmatrix} 1 & 0 & 0 \\ 9/2 & -4 & -4 \\ 6 & 0 & 17-n \end{vmatrix} = 8\begin{vmatrix} -4 & -4 \\ 0 & 17-n \end{vmatrix} = 32(n-17).$$

***Définition 3.4.11* (Matrice des cofacteurs)**. — Pour $i, j \in \{1, \ldots, n\}$, $(-1)^{i+j}\Delta_{ij}(A)$ est appelé le *cofacteur* de A d'indice (i, j). On appelle *matrice des cofacteurs* de A la matrice $\widetilde{A}$ dont le coefficient d'indice (i, j) est $\widetilde{A}_{ij} = (-1)^{i+j}\Delta_{ij}(A)$.

Proposition 3.4.12. — *Soit $A \in M_n(k)$ et $\widetilde{A}$ la matrice des cofacteurs de A. Alors,*

$${}^t\widetilde{A}\,A = A\,{}^t\widetilde{A} = \text{dét}(A)\cdot I_n.$$

En particulier, si A est inversible,

$$A^{-1} = \frac{1}{\text{dét}(A)}\,{}^t\widetilde{A}.$$

[3] lorsqu'une matrice est notée entre barre et non entre parenthèse, cela signifie que c'est le déterminant de la matrice que l'on considère.

Démonstration. — Pour tout $i, \ell \in \{1, \ldots, n\}$, on a:

$$(A\,{}^t\widetilde{A})_{i\ell} = \sum_{j=1}^{n} a_{ij}\,({}^t\widetilde{A})_{j\ell} = \sum_{j=1}^{n} (-1)^{j+\ell} a_{ij}\,\Delta_{\ell j}(A)$$

et l'on reconnaît là le développement suivant la ligne ℓ du déterminant de la matrice $B(\ell, i)$ déduite de A en remplaçant la ligne d'indice ℓ par celle d'indice i. Donc $(A\,{}^t\widetilde{A})_{i\ell} = 0$ si $\ell \neq i$ (car $B(\ell, i)$ a alors deux lignes égales), et $(A\,{}^t\widetilde{A})_{ii} = \text{dét}(A)$ si $\ell = i$. Ceci montre que

$$A\,{}^t\widetilde{A} = \text{dét}(A) \cdot I_n.$$

De même,

$$({}^t\widetilde{A}\,A)_{i\ell} = \sum_{j=1}^{n} (-1)^{i+j} \Delta_{ji}(A)\, a_{j\ell}$$

et l'on reconnaît là le développement suivant la colonne i du déterminant de la matrice $B'(i, \ell)$ déduite de A en remplaçant la colonne d'indice i par celle d'indice ℓ. Donc, à nouveau, $({}^t\widetilde{A}\,A)_{i\ell} = 0$ si $\ell \neq i$, et $= \text{dét}(A)$ si $\ell = i$, d'où

$${}^t\widetilde{A}\,A = \text{dét}(A) \cdot I_n.$$

□

Exemple 3.4.13. — Dans le cas d'une matrice inversible de taille 2×2, la formule de la proposition 3.4.12 s'écrit

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

Remarque 3.4.14. — La formule de la proposition 3.4.12 n'est pas utilisable en pratique pour calculer une inverse de matrice, sauf pour une matrice de taille 2×2 ou 3×3. Elle demande un temps de calcul beaucoup plus grand que la méthode basée sur le pivot de Gauss, vue au chapitre 1. Cette formula a en revanche un intérêt théorique. Par exemple, elle implique que l'application $A \mapsto A^{-1}$ est continue, différentiable et même infiniment différentiable sur $GL_n(\mathbb{R})$ (voir un ouvrage ou un cours traitant du calcul différentiel en plusieurs variables).

3.5. Déterminant d'un endomorphisme, polynome caractéristique, trace

Soit V un espace vectoriel de dimension n. En raison de son importance, répétons encore ici le théorème de changement de base pour un endomorphisme u de V (étant entendu qu'on exprime la matrice de u dans une même base au départ et à l'arrivée) (cf. 0.5.13) :

Théorème 3.5.1 **(Changement de base pour un endomorphisme)**
Soit A la matrice d'un endomorphisme u de V relativement à une base $\mathscr{B}$ de V. Si $\mathscr{B}'$ est une seconde base, et si P est la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$, alors la matrice de u dans la base B' est:

$$\text{Mat}_{\mathscr{B}'}(u) = P^{-1}AP.$$

Ceci permet de définir le déterminant et le polynôme caractéristique de u comme suit.

Théorème et définition 3.5.2 **(Déterminant, polynôme caractéristique et trace d'un endomorphisme)**
Soit $u \in \text{End}_k(V)$ et soit A sa matrice dans une base $\mathscr{B}$ de V. On définit le déterminant et le polynôme caractéristique de u par:

$$\text{dét}(u) = \text{dét}(A), \qquad P_u(X) = \text{dét}(A - XI_n) \in k[X];$$

ceci ne dépend pas du choix de la base $\mathscr{B}$. De plus, $P_u(X)$ est de degré $n = \dim_k(V)$ et est de la forme

$$P_u(X) = (-1)^n X^n + (-1)^{n-1}\,\text{Tr}(u) X^{n-1} + \cdots + \text{dét}(u).$$

Le coefficient $\text{Tr}(u)$ *s'appelle la* trace *de u, il est égal à la somme des coefficients diagonaux a_{ii} de A (et ceci ne dépend pas de la base choisie).*

Démonstration. — En effet, soient $\mathscr{B}'$ une autre base, P la matrice de passage, et A' la matrice de u dans la base $\mathscr{B}'$. Alors on a $A' = P^{-1}AP$ et donc aussi

$$A' - XI_n = P^{-1}(A - XI_n)P$$

(égalité dans l'anneau $M_n(k[X])$ des matrices à coefficients dans $k[X]$). Donc, d'après la multiplicativité du déterminant (cf. proposition 3.4.4), on a

$$\text{dét}(A') = \text{dét}(A), \qquad \text{dét}(A' - XI_n) = \text{dét}(A - XI_n)$$

ce qui prouve que $\text{dét}(u)$ et $P_u(X)$ sont bien définis. De plus, notons b_{ij} les coefficients de la matrice $A - XI_n$, i.e. $b_{ij} = a_{ij}$ si $i \neq j$ et $b_{ii} = a_{ii} - X$. Par définition, on a

$$P_u(X) = \sum_{\sigma \in S_n} \varepsilon(\sigma)\, b_{1\sigma(1)} \cdots b_{n\sigma(n)}$$

c.-à-d., c'est la somme, avec certains signes $+$ ou $-$, de tous les produits de n coefficients de $B = A - XI_n$, en ne prenant dans chaque produit qu'un seul coefficient dans chaque ligne et chaque colonne, et de plus on a $\varepsilon(\text{id}) = 1$, i.e. le terme $b_{11}b_{22}\cdots b_{nn}$ apparaît avec le signe $+$.

Donc, le terme de degré maximal en X est $(-X)^n$, qui s'obtient en développant le produit des termes diagonaux:

$$b_{11}b_{22}\cdots b_{nn} = (a_{11} - X)\cdots(a_{nn} - X).$$

Pour avoir un terme en X^{n-1}, il faut prendre $n-1$ fois $(-X)$ sur la diagonale, mais alors, comme chaque produit de n coefficients n'a qu'un seul coefficient par ligne et par colonne, le dernier terme du produit est le coefficient diagonal restant, dans lequel on prend le terme a_{ii}. Ceci montre que le coefficient de $(-X)^{n-1}$ est

$$a_{11} + \cdots + a_{nn},$$

qu'on appelle la trace de A. Comme $P_u(X)$ ne dépend pas de la base choisie, ce coefficient n'en dépend pas non plus; on l'appelle la trace de u et on le note $\text{Tr}(u)$. □

Remarque 3.5.3. — On peut aussi montrer directement que, pour tout $A, B \in M_n(k)$, on a $\text{Tr}(AB) = \text{Tr}(BA)$, d'où $\text{Tr}(P^{-1}AP) = \text{Tr}(APP^{-1}) = \text{Tr}(A)$.

D'après la proposition 3.4.4, on peut alors compléter la proposition 0.3.3 comme suit:

Proposition 3.5.4. — *Soit $u \in \text{End}_k(V)$ $(\dim_k(V) = n)$. Les conditions suivante sont équivalentes:*
(i) *u est injectif;*
(ii) *u est surjectif;*
(iii) *u est bijectif;*
(iv) $\text{dét}(u) \neq 0$.

Définition 3.5.5 **(Valeurs, vecteurs et sous-espaces propres).** — 1) Soit $u \in \text{End}_k(V)$. On dit que $\lambda \in k$ est une *valeur propre* de u si $\text{Ker}(u - \lambda\,\text{id}_V) \neq 0$ (i.e. s'il existe $v \in V - \{0\}$ tel que $u(v) = \lambda v$). Dans ce cas, $V_\lambda = \text{Ker}(u - \lambda\,\text{id}_V)$ est appelé *l'espace propre* associé à λ, et tout vecteur $v \in V_\lambda - \{0\}$ est appelé un *vecteur propre* de u, associé à la valeur propre λ.

2) Pour $A \in M_n(k)$, on définit ses valeurs, vecteurs et sous-espaces propres comme étant ceux de l'endomorphisme u de k^n défini par A, c.-à-d., λ est valeur propre de A si et seulement si $V_\lambda = \text{Ker}(A - \lambda I_n)$ est non nul.

Proposition 3.5.6. — *Soient $u \in \text{End}_k(V)$ et $\lambda \in k$, alors λ est une valeur propre de $u \Leftrightarrow P_u(\lambda) = 0$.*

Démonstration. — D'après la proposition 3.5.4, λ est valeur propre de u si et seulement si $\text{dét}(u - \lambda\,\text{id}_V) = 0$. □

3.6. Appendice (†) : Bases et dimension de $\mathscr{L}(E_1, \ldots, E_p; F)$

Soit p un entier ≥ 2 et soient $E_1, \ldots, E_p$ et F des k-espaces vectoriels de dimension finie, soit $\mathscr{C} = (f_1, \ldots, f_r)$ une base de F et, pour $q = 1, \ldots, p$, soient $d_q = \dim E_q$ et

$$\mathscr{B}_q = (e_1^q, \ldots, e_{d_q}^q)$$

une base de E_i. Notons P l'ensemble produit $P = \{1, \ldots, d_1\} \times \cdots \times \{1, \ldots, d_p\}$, il est de cardinal $D = d_1 \cdots d_p$.

Théorème 3.6.1 **(Bases et dimension de $\mathscr{L}(E_1, \ldots, E_p; F)$).** — *Avec les notations précédentes, on a :*

(i) *Tout $\phi \in \mathscr{L}(E_1,\dots,E_p;F)$ est entièrement déterminé par ses valeurs $u_{(i_1,\dots,i_p)} = \phi(e^1_{i_1},\dots,e^p_{i_p}) \in F$ pour $(i_1,\dots,i_p) \in P$, et ces éléments peuvent être choisis arbitrairement.*

(ii) *Par conséquent, les éléments $\phi_s^{a_1,\dots,a_p}$, pour $(a_1,\dots,a_p) \in P$ et $s = 1,\dots,r$, définis par:*

$$\forall (i_1,\dots,i_p), \qquad \phi_s^{a_1,\dots,a_p}(e_{i_1},\dots,e_{i_p}) = \begin{cases} f_s \in F & \textit{si } i_1 = a_1,\dots,i_p = a_p \\ 0 & \textit{sinon} \end{cases}$$

forment une base de $\mathscr{L}(E_1,\dots,E_p;F)$; celui-ci est donc de dimension $rD = rd_1\cdots d_p$.

Démonstration. — Soit $\phi \in \mathscr{L}(E_1,\dots,E_p;F)$. Pour tout $(v_1,\dots,v_p) \in E_1\times\cdots\times E_p$, chaque v_q s'écrit de façon unique

$$(0) \qquad v_q = t_{q,1}\, e^q_1 + \cdots + t_{q,d}\, e^q_{d_q}, \qquad \text{avec} \quad t_{q,i} \in k,$$

donc, en appliquant les égalités $(*_1)$, …, $(*_p)$ de 3.1.2, on obtient la formule :

$$(1) \quad \phi(v_1,\dots,v_p) = \phi\left(t_{1,1}\, e^1_1 + \cdots + t_{1,d}\, e^1_{d_1},\dots,t_{p,1}\, e^p_1 + \cdots + t_{p,d}\, e^p_{d_p}\right) = \sum_{(i_1,\dots,i_p)\in P} t_{1,i_1}\cdots t_{p,i_p}\, \phi(e^1_{i_1},\dots,e^p_{i_p}).$$

Ceci montre que ϕ est entièrement déterminée par les D éléments $\phi(e_{i_1},\dots,e_{i_p}) \in F$; en d'autres termes, l'application "d'évaluation sur les p-uplets $(e_{i_1},\dots,e_{i_p})$":

$$\mathrm{ev} : \mathscr{L}(E_1,\dots,E_p;k) \to F^D \simeq k^{rD}, \qquad \phi \mapsto \Big(\phi(e_{i_1},\dots,e_{i_p})\Big)_{(i_1,\dots,i_p)\in P}$$

est *injective*. D'autre part, cette application est *linéaire*: si $\mu \in k$ et $\phi,\psi \in \mathscr{L}(E_1,\dots,E_p;k)$, on a

$$(\mu\phi+\psi)(v_1,\dots,v_p) = \mu\phi(v_1,\dots,v_p) + \psi(v_1,\dots,v_p)$$

pour *tout* $(v_1,\dots v_p) \in E_1\times\cdots\times E_p$, donc *a fortiori* pour tout p-uplet $(e_{i_1},\dots,e_{i_p})$.

Réciproquement, pour tout $(a_1,\dots,a_p) \in P$, notons $\phi_s^{a_1,\dots,a_p}$ l'application $E_1\times\cdots\times E_p \to F$ définie comme suit: pour tout p-uplet $(v_1,\dots,v_p)$ comme en (0) ci-dessus,

$$(2) \qquad \phi_s^{a_1,\dots,a_p}(v_1,\dots,v_p) = (\underbrace{t_{1,a_1}\cdots t_{p,a_p}}_{\in k})\cdot f_s \in F.$$

Fixons un indice $q \in \{1,\dots,p\}$, et montrons que $\phi_s^{a_1,\dots,a_p}$ est linéaire par rapport à la q-ème variable, i.e. que, pour tout $\lambda \in k$ et $w \in E_q$, on a

$$(\star_q) \qquad \phi_s^{a_1,\dots,a_p}(v_1,\dots,\lambda v_q + w,\dots,v_p) = \lambda\phi_s^{a_1,\dots,a_p}(v_1,\dots,v_q,\dots,v_p) + \phi_s^{a_1,\dots,a_p}(v_1,\dots,w,\dots,v_p).$$

Écrivons $w = s_1 e^q_1 + \cdots + s_{d_q} e^q_{d_q}$, avec $s_i \in k$. Alors $\lambda v_q + w = t'_{q,1}\, e^q_1 + \cdots + t'_{q,d}\, e^q_d$, avec $t'_{q,i} = \lambda t_{q,i} + s_i$, et donc la formule (2) donne :

$$\begin{aligned} \phi_s^{a_1,\dots,a_p}(v_1,\dots,\lambda v_q + w,\dots,v_p) &= \big(t_{1,a_1}\cdots t'_{q,a_q}\cdots t_{p,a_p}\big)\cdot f_s = \big(t_{1,a_1}\cdots(\lambda t_{q,a_q} + s_{a_q})\cdots t_{p,a_p}\big)\cdot f_s \\ &= \lambda t_{1,a_1}\cdots t_{q,a_q}\cdots t_{p,a_p} + t_{1,a_1}\cdots s_{a_q}\cdots t_{p,a_p} \\ &= \lambda\phi_s^{a_1,\dots,a_p}(v_1,\dots,v_q,\dots,v_p) + \phi_s^{a_1,\dots,a_p}(v_1,\dots,w,\dots,v_p) \end{aligned}$$

ce qui montre que $(\star_q)$ est vérifiée, d'où $\phi_s^{a_1,\dots,a_p} \in \mathscr{L}(E_1,\dots,E_p;k)$. Remarquons de plus que, pour tout $(i_1,\dots,i_p) \in P$, l'élément $(v_1,\dots,v_p) = (e_{i_1},\dots,e_{i_p})$ de $E_1\times\cdots\times E_p$ correspond dans l'écriture (0) à:

$$\forall q = 1,\dots,p, \qquad t_{q,j_q} = \begin{cases} 1 & \text{si } j_q = i_q \\ 0 & \text{sinon} \end{cases}$$

donc (2) donne, en particulier:

$$(2') \qquad \forall (i_1,\dots,i_p) \in P, \qquad \phi_s^{a_1,\dots,a_p}(e_{i_1},\dots,e_{i_p}) = \begin{cases} f_s & \text{si } i_1 = a_1,\dots,i_p = a_p \\ 0 & \text{sinon.} \end{cases}$$

Maintenant, pour tout élément $U = \big(\sum_{s=1}^r \lambda^s_{(j_1,\dots,j_p)} f_s\big)_{(j_1,\dots,j_p)\in P} \in F^D$, l'élément

$$\phi_U = \sum_{(j_1,\dots,j_p)\in P} \sum_{s=1}^r \lambda^s_{(j_1,\dots,j_p)}\, \phi_s^{(j_1,\dots,j_p)}$$

appartient à $\mathscr{L}(E_1,\dots,E_p;k)$ et, d'après (2′) il vérifie, pour tout $(i_1,\dots,i_p)\in P$:

$$(3)\qquad \phi_U(e_{i_1},\dots,e_{i_p}) = \sum_{(j_1,\dots,j_p)\in P}\sum_{s=1}^{r}\lambda^s_{(j_1,\dots,j_p)}\underbrace{\phi_s^{(j_1,\dots,j_p)}(e_{i_1},\dots,e_{i_p})}_{\substack{=f_s \text{ si } (j_1,\dots,j_p)=(i_1,\dots,i_p)\\ =0 \text{ sinon}}} = \sum_{s=1}^{r}\lambda_{(i_1,\dots,i_p)}\, f_s.$$

Ceci prouve que l'application linéaire injective ev : $\mathscr{L}(E_1,\dots,E_p;F)\to F^D$ est aussi *surjective*, donc est un isomomorphisme; en particulier, $\dim\mathscr{L}(E_1,\dots,E_p;k) = rD = rd_1\cdots d_p$. Ceci prouve déjà l'assertion (i).

De plus, (3) implique que si l'on a une relation $\sum_{(j_1,\dots,j_p)\in P}\sum_{s=1}^{r}\lambda^s_{(j_1,\dots,j_p)}\,\phi_s^{(j_1,\dots,j_p)} = 0$, alors pour tout $(i_1,\dots,i_p)\in P$ on a:

$$\sum_{s=1}^{r}\lambda_{(i_1,\dots,i_p)}\, f_s = 0$$

donc chaque $\lambda^s_{(i_1,\dots,i_p)}$ est nul (puisque $(f_1,\dots,f_r)$ est une base de F). La famille

$$\mathscr{F} = \left(\phi_s^{(a_1,\dots,a_p)}\right)_{\substack{(a_1,\dots,a_p)\in P\\ s=1,\dots,r}}$$

est donc *libre*, et comme elle est de cardinal $rD = rd_1\cdots d_p$, c'est une base de $\mathscr{L}(E_1,\dots,E_p;k)$. Ceci achève la preuve de l'assertion (ii). Le théorème est démontré. □

Dans le cas particulier où $E_1 = \ldots = E_p = E$ où $F = k$, c'est à dire où l'on considère l'espace des formes p-linéaires sur E, le théorème 3.6.1 se traduit par le corollaire suivant.

Corollaire 3.6.2 **(Bases et dimension de $\mathscr{L}_p(E,k)$).** — *Soient p un entier ≥ 2 et E un k-espace vectoriel de dimension d. Notons P l'ensemble produit $\{1,\dots,d\}\times\cdots\times\{1,\dots,d\}$ (p facteurs), et soit $\mathscr{B}=(e_1,\dots,e_d)$ une base de E.*

(i) *Tout élément $\phi\in\mathscr{L}_p(E,k)$ est entièrement déterminé par ses valeurs $\lambda_{(i_1,\dots,i_p)} = \phi(e_{i_1},\dots,e_{i_p})\in k$, et ces éléments peuvent être choisis arbitrairement.*

(ii) *Par conséquent, les éléments $\phi^{a_1,\dots,a_p}$, pour $(a_1,\dots,a_p)\in P$, définis par:*

$$\forall(i_1,\dots,i_p),\qquad \phi^{a_1,\dots,a_p}(e_{i_1},\dots,e_{i_p}) = \begin{cases}1 & si\ i_1=a_1,\dots,i_p=a_p\\ 0 & sinon\end{cases}$$

forment une base de $\mathscr{L}_p(E,k)$; on a donc $\boxed{\dim\mathscr{L}_p(E,k) = d^p = (\dim E)^p.}$

3.7. Appendice (†) : Produit tensoriel

Soient $E_1,\dots,E_p$ et F des k-espaces vectoriels, chaque E_q étant de dimension finie d_q. Pour $q=1,\dots,p$, soit $\mathscr{B}_q = (e^q_1,\dots,e^q_{d_q})$ une base de E_q. D'après le théorème 3.6.1, on voit que l'espace des applications p-linéaires $E_1\times\cdots\times E_p\to F$ est la même chose que l'espace des applications *linéaires* de "$E_1\otimes\cdots\otimes E_p$" vers F, où "$E_1\otimes\cdots\otimes E_p$" désigne un espace vectoriel de base les p-uplets

$$(e^1_{j_1},\dots,e^p_{j_p})\in\mathscr{B}_1\times\cdots\times\mathscr{B}_p.$$

On pourrait *définir* de la sorte l'espace "$E_1\otimes\cdots\otimes E_p$", mais il faudrait alors démontrer qu'il ne dépend pas des bases choisies. En fait, on définit le *produit tensoriel*

$$E_1\otimes\cdots\otimes E_p$$

comme suit. On considère d'abord le k-espace vectoriel V ayant pour base $\mathscr{B}$ *l'ensemble* $E_1\times\cdots\times E_p$, c.-à-d., V admet pour base des vecteurs

$$b_{x_1,\dots,x_p}$$

indexés par $E_1\times\cdots\times E_p$. Alors, se donner une application linéaire $\phi: V\to F$ est la même chose que se donner pour tout vecteur de base $b_{x_1,\dots,x_p}$ un élément $f(x_1,\dots,x_p)$ de F; on a donc une bijection naturelle

$$\begin{array}{ccc}\mathscr{L}(V,F) & \xrightarrow{\sim} & \mathrm{Applic}(E_1\times\cdots\times E_p, F)\\ \phi & \longmapsto & \Big((x_1,\dots,x_p)\mapsto\phi(b_{x_1,\dots,x_p})\Big)\end{array}$$

où dans le terme de droite de la première ligne, $E_1\times\cdots\times E_p$ est regardé juste comme un ensemble.

Considérons maintenant le sous-espace W de V engendré par tous les éléments de l'une des formes suivantes, pour $q = 1, \dots, p$:

$$(*_q) \qquad \begin{cases} b_{x_1,\dots,x_q+y_q,\dots,x_p} - b_{x_1,\dots,x_q,\dots,x_p} - b_{x_1,\dots,y_q,\dots,x_p} \\ b_{x_1,\dots,\lambda\cdot x_q,\dots,x_p} - \lambda \cdot b_{x_1,\dots,x_q,\dots,x_p} \end{cases}$$

on pose

$$E_1 \otimes \cdots \otimes E_p = V/W$$

et pour $x_1 \in E_1, \dots, x_p \in E_p$, on note $x_1 \otimes \cdots \otimes x_p$ l'image de $b_{x_1,\dots,x_p}$ dans $V/W = E_1 \otimes \cdots \otimes E_p$.

Alors, d'après le théorème 5.5.3, une application linéaire $E_1 \otimes \cdots \otimes E_p \to F$ est la même chose qu'une application linéaire $\phi : V \to F$ qui s'annule sur les générateurs de W; si l'on note f l'application d'ensembles $E_1 \times \cdots \times E_p \to F$ correspondant à ϕ (c.-à-d., $f(x_1, \dots, x_p) = \phi(b_{x_1,\dots,x_p})$), la condition que ϕ s'annule sur les générateurs $(*_q)$ de W équivaut alors à dire que f est linéaire par rapport à la q-ème variable. On obtient ainsi que l'on a un isomorphisme d'espaces vectoriels :

$$(1) \qquad \begin{array}{ccc} \mathscr{L}(E_1 \otimes \cdots \otimes E_p, F) & \xrightarrow{\sim} & \mathscr{L}(E_1, \dots, E_p; F) \\ \phi & \longmapsto & \Big((x_1, \dots, x_p) \mapsto \phi(b_{x_1,\dots,x_p})\Big). \end{array}$$

D'autre part, tout $x_1 \in E_1$ s'écrit de façon unique $x_1 = \sum_{j_1=1}^{d_1} t_{1,j_1} e^1_{j_1}$ et donc, notant π la projection $V \to E_1 \otimes \cdots \otimes E_p$, on a, d'après $(*_1)$:

$$x_1 \otimes \cdots \otimes x_p = \pi(b_{x_1 \otimes \cdots \otimes x_p}) = \sum_{j_1=1}^{d_1} t_{1,j_1} \pi(b_{e_{j_1}, x_2, \dots, x_p}) = \sum_{j_1=1}^{d_1} t_{1 j_1} e^1_{j_1} \otimes x_2 \otimes \cdots \otimes x_p$$

donc écrivant $x_q = \sum_{j_i q=1}^{d_q} t_{q,j_q} e^q_{j_q}$ pour $q = 2, \dots, p$ et appliquant $(*_2), \dots, (*_p)$, on obtient que:

$$x_1 \otimes \cdots \otimes x_p = \sum_{j_1,\dots,j_p} t_{1,j_1} \cdots t_{p,j_p} \, e^1_{j_1} \otimes \cdots \otimes e^p_{j_p},$$

ce qui montre que $E = E_1 \otimes \cdots \otimes E_p$ est engendré par les $d_1 \cdots d_p$ vecteurs

$$e^1_{j_1} \otimes \cdots \otimes e^p_{j_p}$$

où $j_q \in \{1, \dots, d_q\}$ pour $q = 1, \dots, p$, donc E est de dimension finie $n \leq d_1 \cdots d_p$. Alors n est aussi la dimension de l'espace dual $E^* = \mathscr{L}(E, k)$; or d'après (1) appliqué à $F = k$, on a un isomorphisme de k-espaces vectoriels

$$\mathscr{L}(E, k) \simeq \mathscr{L}(E_1, \dots, E_p; k)$$

et l'on a vu en 3.6.1 que $\dim \mathscr{L}(E_1, \dots, E_p; k) = d_1 \cdots d_p$. On obtient donc que

$$\dim(E_1 \otimes \cdots \otimes E_p) = d_1 \cdots d_p$$

et donc les $d_1 \cdots d_p$ vecteurs $e^1_{j_1} \otimes \cdots \otimes e^p_{j_p}$ forment une *base* de $E_1 \otimes \cdots \otimes E_p$. On a donc obtenu le :

Théorème 3.7.1. — *Soient $E_1, \dots, E_p$ des k-espaces vectoriels de dimension finie et, pour $q = 1, \dots, p$, soit $(e^q_1, \dots, e^q_{d_q})$ une base de E_q. Alors, pour tout k-espace vectoriel F, on a un isomorphisme de k-espaces vectoriels*

$$\boxed{\mathscr{L}(E_1 \otimes \cdots \otimes E_p, F) \simeq \mathscr{L}(E_1, \dots, E_p; F).}$$

De plus, $E_1 \otimes \cdots \otimes E_p$ est de dimension $d_1 \cdots d_p$ et les vecteurs $e^1_{j_1} \otimes \cdots \otimes e^p_{j_p}$, où chaque j_q parcourt $\{1, \dots, d_q\}$, en forment une base.

CHAPITRE 4

RÉDUCTION DES ENDOMORPHISMES

Résumé: Dans ce chapitre, on commence par des rappels sur les espaces propres et des critères de diagonalisabilité (4.1.8, 4.1.11 et 4.1.12). Puis, prenant $\mathbb{C}$ comme corps de base, on démontre les théorèmes de trigonalisation, de Cayley-Hamilton, et de décomposition en espaces caractéristiques.

Par ailleurs, des *compléments de cours*, pour les étudiants intéressés, sont donnés dans des appendices à la fin du chapitre. (on y démontre le théorème de Bézout et le fait que le corps $\mathbb{C}$ est algébriquement clos).

4.1. Espaces propres et critères de diagonalisabilité

Définition 4.1.1 **(Sous-espaces en somme directe).** — Soient V un k-espace vectoriel, $E_1, \dots, E_n$ des sous-espaces de V. (Ni V ni les E_i ne sont supposés de dimension finie.)

(1) D'abord, on note $E_1 + \cdots + E_n$ (ou $\sum_{i=1}^n E_i$) le sous-espace de V engendré par $E_1 \cup \cdots \cup E_n$; c'est l'ensemble de toutes les sommes

$$(*) \qquad x_1 + \cdots + x_n, \qquad \text{avec } x_i \in E_i.$$

(2) On dit que les E_i *sont en somme directe* si pour tous $x_1 \in E_1, \dots, x_n \in E_n$, l'égalité $x_1+\cdots+x_n = 0$ entraîne $x_1 = 0 = \cdots = x_n$. Ceci équivaut à dire que tout élément x de $E_1 + \cdots + E_n$ s'écrit *de façon unique* $x = x_1 + \cdots + x_n$ avec $x_i \in E_i$. Dans ce cas, $E_1 + \cdots + E_n$ est noté $E_1 \oplus \cdots \oplus E_n$ ou $\bigoplus_{i=1}^n E_i$.

(3) Si chaque E_i est de dimension finie d_i, et si $\mathscr{B}_1 = (e_1, \dots, e_{d_1})$ est une base de E_1, et $\mathscr{B}_2 = (e_{d_1+1}, \dots, e_{d_1+d_2})$ une base de E_2, ... puis $\mathscr{B}_n = (e_{d_1+\cdots+d_{n-1}+1}, \dots, e_{d_1+\cdots+d_n})$ une base de E_n, la condition précédente équivaut à dire que la famille $\mathscr{F} = (e_1, \dots, e_{d_1+\cdots+d_n})$ est une base de $E_1 + \cdots + E_n$, et comme $\mathscr{F}$ engendre de toute façon $E_1 + \cdots + E_n$, ceci équivaut aussi à dire que

$$(\star) \qquad \dim(E_1 + \cdots + E_n) = \dim(E_1) + \cdots + \dim(E_n).$$

Terminologie 4.1.2. — Si $E_1, \dots, E_n$ sont en somme directe et si de plus $E_1 \oplus \cdots \oplus E_n$ égale V, alors on dit que V est la somme directe des E_i.

Remarques 4.1.3. — (1) Il résulte de la définition que $E_1, \dots, E_n$ sont en somme directe si et seulement si, pour tout $i = 1, \dots, n$, on a: $E_i \cap \sum_{j \neq i} E_j = 0$.

(2) En particulier, si $n = 2$, alors E_1 et E_2 sont en somme directe si et seulement si $E_1 \cap E_2 = (0)$.

(3) **Attention!** Si des sous-espaces sont en somme directe, leur somme n'est pas nécessairement égale à l'espace tout entier: par exemple si E_1, E_2 sont deux droites distinctes dans $\mathbb{R}^3$, leur somme est directe, et c'est un plan de $\mathbb{R}^3$, et non $\mathbb{R}^3$ tout entier!

(4) **Attention!** Si $n \geq 3$, la condition $E_i \cap E_j = \{0\}$ pour $i \neq j$ *n'entraîne pas* que la somme des E_i soit directe : par exemple si E_1, E_2, E_3 sont trois droites distinctes dans $\mathbb{R}^2$, elles vérifient $E_i \cap E_j = \{0\}$ pour $i \neq j$, mais leur somme n'est pas directe (car $E_1 + E_2$ égale $\mathbb{R}^2$ donc contient E_3).

Définition 4.1.4 **(Sous-espaces supplémentaires).** — Soient V un espace vectoriel, E, F deux sous-espaces de V. On dit que E et F sont des sous-espaces *supplémentaires* si $V = E \oplus F$, c.-à-d., si $E \cap F = (0)$ et $E + F = V$.

Si V est de dimension finie, ceci équivaut à dire que $E \cap F = (0)$ et $\dim(E) + \dim(F) = \dim(V)$.

Proposition 4.1.5. — *Soit V un k-espace vectoriel de dimension finie n. Tout sous-espace E de V admet un supplémentaire.*

[(0)]version du 15/06/2015

Démonstration. Soit $(e_1, \dots, e_r)$ une base de E, complétons-la en une base $(e_1, \dots, e_n)$ de V et soit F le sous-espace de V engendré par $e_{r+1}, \dots, e_n$. Alors $E \cap F = \{0\}$ et $E + F = V$, donc $V = E \oplus F$.

Exercice 4.1.6. — Soient V un espace vectoriel, E (resp. F) un sous-espace de dimension finie m (resp. n). Soit $(v_1, \dots, v_r)$ une base de $E \cap F$, complétons-la en une base $(e_1, \dots, e_{m-r}, v_1, \dots, v_r)$ de E (resp. $(v_1, \dots, v_r, f_1, \dots, f_{n-r})$ de F). Montrer que $(e_1, \dots, e_{m-r}, v_1, \dots, v_r, f_1, \dots, f_{n-r})$ est une base de $E + F$. En déduire l'égalité $\dim(E + F) = \dim(E) + \dim(F) - \dim(E \cap F)$.

Remarque 4.1.7. — Soient V un k-espace vectoriel et f une forme linéaire sur V (cf. 1.1.1). Supposons $f \neq 0$ et notons $H = \mathrm{Ker}(f)$. Comme $f \neq 0$, il existe $v \in V$ tel que $f(v) \neq 0$, et remplaçant v par $f(v)^{-1}\, v$, on peut supposer $f(v) = 1$. Alors, pour tout $w \in V$, on a $f(w - f(w)v) = 0$, donc $w - f(w)v \in \mathrm{Ker}(f) = H$ et donc w s'écrit

$$w = h + f(w)v \qquad \text{avec} \quad h = w - f(w)v \in H,$$

d'où $V = H + kv$. D'autre part, si $tv \in H$ alors $0 = f(tv) = t$; on a donc $H \cap kv = (0)$ et donc $V = H \oplus kv$, i.e. H et la droite kv sont supplémentaires. On dit que H est un *hyperplan* de V; si V est de dimension finie n, alors H est de dimension $n - 1$.

Un exemple très important de sous-espaces en somme directe est celui des sous-espaces propres d'un endomorphisme de V; rappelons-le ici.

Théorème 4.1.8. — *Soient V un k-espace vectoriel (pas nécessairement de dimension finie), u un k-endomorphisme de V, et $\lambda_1, \dots, \lambda_r$ des valeurs propres, deux à deux distinctes, de u. Pour $i = 1, \dots, r$, on note*

$$E_i = V_{\lambda_i} = \{v \in V \mid u(v) = \lambda_i v\}$$

le sous-espace propre associé. Alors les V_{λ_i} sont en somme directe. (Mais bien sûr, leur somme $\bigoplus_{i=1}^r V_{\lambda_i}$ n'est pas nécessairement égale à V; si $\dim(V) < \infty$, c'est le cas si et seulement si u est diagonalisable.)

Démonstration. On va montrer par récurrence sur r l'assertion: $(\star_r)$ si l'on a une égalité $x_1 + \cdots + x_r = 0$, avec $x_i \in V_{\lambda_i}$, alors $x_1 = 0 = \cdots = x_r$. C'est évident si $r = 1$, donc on peut supposer $r \geq 2$ et l'assertion établie pour $r - 1$. Supposons qu'on ait une égalité $x_1 + \cdots + x_r = 0$, avec $x_i \in V_{\lambda_i}$. En appliquant l'endomorphisme u, d'une part, et en multipliant par λ_r, d'autre part, on obtient les égalités :

$$\begin{cases} \lambda_1 x_1 + \cdots + \lambda_{r-1} x_{r-1} + \lambda_r x_r = 0 \\ \lambda_r x_1 + \cdots + \lambda_r x_{r-1} + \lambda_r x_r = 0 \end{cases}$$

d'où par soustraction l'égalité

$$(*) \qquad (\lambda_1 - \lambda_r) x_1 + \cdots + (\lambda_{r-1} - \lambda_r) x_{r-1} = 0.$$

Chaque vecteur $y_i = (\lambda_i - \lambda_r) x_i$ appartient à V_{λ_i} donc, d'après l'hypothèse de récurrence $(\star_{r-1})$, l'égalité $(*)$ entraîne $y_i = 0$ pour tout $i = 1, \dots, r - 1$, et comme $\lambda_i - \lambda_r \neq 0$ ceci entraîne $x_i = 0$ pour tout $i = 1, \dots, r - 1$. Enfin, reportant ceci dans l'égalité initiale $x_1 + \cdots + x_r = 0$, on obtient $x_r = 0$. Ceci montre que $(\star_r)$ est vérifiée, et la proposition est démontrée.

Exemples 4.1.9. — (1) Soit $V = \mathscr{C}^\infty(\mathbb{R}, \mathbb{R})$ le $\mathbb{R}$-espace vectoriel des fonctions $f : \mathbb{R} \to \mathbb{R}$ de classe C^∞. Alors les fonctions $f_\lambda : t \mapsto \exp(\lambda t)$, pour $\lambda \in \mathbb{R}$ sont linéairement indépendantes, c.-à-d., quelques soient $n \in \mathbb{N}^*$ et $\lambda_1, \dots, \lambda_n$ des réels deux à deux distincts, les fonctions $f_{\lambda_1}, \dots, f_{\lambda_n}$ sont linéairement indépendantes. En effet, l'opérateur de dérivation $d : f \mapsto f'$ est un endomorphisme de V (car f' est C^∞ si f l'est), et chaque f_λ est un vecteur propre de d pour la valeur propre λ.

(2) Soit $V = \mathbb{R}^{\mathbb{N}}$ le $\mathbb{R}$-espace vectoriel des suites réelles $(u_n)_{n \in \mathbb{N}}$. Alors, pour $\lambda \in \mathbb{R}$, les suites géométrique $u(\lambda)$, définies par $u(\lambda)_n = \lambda^n$, sont linéairement indépendantes, c.-à-d., quelques soient $n \in \mathbb{N}^*$ et $\lambda_1, \dots, \lambda_n$ des réels deux à deux distincts, les suites $u(\lambda_1), \dots, u(\lambda_n)$ sont linéairement indépendantes. En effet, soit $D : V \to V$ l'opérateur de décalage, défini par $(D(u))_n = u_{n+1}$ (i.e. l'image par D de la suite $(u_0, u_1, u_2, \dots)$ est la suite $(u_1, u_2, u_3, \dots)$); alors chaque $u(\lambda)$ est un vecteur propre de D pour la valeur propre λ.

***Définition 4.1.10* (Endomorphismes diagonalisables).** — Soient V un k-espace vectoriel de dimension finie et $u \in \mathrm{End}_k(V)$. Les conditions suivantes sont équivalentes:

(i) V admet une base formée de vecteurs propres de u;
(ii) les vecteurs propres de u engendrent V;
(iii) la somme des espaces propres de u égale V;
(iv) V est la somme directe des espaces propres de u.

Si ces conditions sont vérifiées, on dit que u est *diagonalisable*.

Démonstration. — En effet, il est clair que (iv) $\Rightarrow$ (iii) $\Leftrightarrow$ (ii) $\Leftarrow$ (i), et (ii) $\Rightarrow$ (i) car d'un système de générateurs on peut extraire une base. Enfin (iii) $\Rightarrow$ (iv) d'après le théorème précédent. $\square$

Une condition *suffisante* de diagonalisabilité est donnée par la proposition ci-dessous. (Bien entendu, cette condition n'est **pas nécessaire**: par exemple la matrice identité I_n (≥ 2) est diagonale et a toutes ses valeurs propres égales à 1!)

***Proposition 4.1.11* (Valeurs propres distinctes).** — *Soit $u \in \operatorname{End}_k(V)$ ($\dim_k(V) = n$). Si $P_u(X)$ a n valeurs propres distinctes, alors u est diagonalisable.*

Démonstration. — En effet, u possède alors n espaces propres distincts $V_1, \dots, V_n$, qui sont en somme directe d'après le théorème précédent. Alors le sous-espace $E = V_1 \oplus \cdots \oplus V_n$ de V est de dimension

$$\sum_{i=1}^{n} \dim V_i \geq n = \dim V$$

donc égale V (et de plus chaque V_i est de dimension 1). $\square$

Enfin, une CNS (condition nécessaire et suffisante) de diagonalisabilité est donnée par la :

***Définition et proposition 4.1.12* (Multiplicités algébrique et géométrique d'une valeur propre)** *Soient V un $\mathbb{C}$-espace vectoriel de dimension n, $u \in \operatorname{End}_{\mathbb{C}}(V)$, $\lambda_1, \dots, \lambda_r$ les racines (deux à deux distinctes) de $P_u(X)$ dans $\mathbb{C}$. D'une part, $P_u(X)$ se factorise*

$$P_u(X) = (-1)^n (X - \lambda_1)^{m_1} \cdots (X - \lambda_r)^{m_r}$$

où m_i est la multiplicité de λ_i comme racine de $P_u(X)$. D'autre part, d'après 3.5.6, *$\lambda_1, \dots, \lambda_r$ sont les valeurs propres de u.*

(1) *On appelle multiplicité* algébrique *(resp.* géométrique*) de la valeur propre λ_i sa multiplicité m_i comme racine de $P_u(X)$ (resp. la dimension n_i de l'espace propre V_{λ_i}).*

(2) *On a $\dim V_{\lambda_i} \leq m_i$ pour tout i.*

(3) *u est diagonalisable $\Longleftrightarrow \dim V_{\lambda_i} = m_i$ pour tout i.*

Démonstration. — (2) Pour tout i, soit $\mathscr{C}^i$ une base de V_{λ_i}. Comme les espaces propres sont en somme directe, la famille $\mathscr{C} = \mathscr{C}^1 \cup \cdots \cup \mathscr{C}^r$ est une famille libre, donc on peut la compléter en une base $\mathscr{B}$ de V. Alors $A = \operatorname{Mat}_{\mathscr{B}}(u)$ est de la forme suivante:

$$A = \left(\begin{array}{c|c|c|c|c} \lambda_1 I_{n_1} & 0 & \cdots & 0 & * \\ \hline 0 & \lambda_2 I_{n_2} & \ddots & \vdots & * \\ \hline 0 & \ddots & \ddots & 0 & * \\ \hline \vdots & \ddots & 0 & \lambda_r I_{n_r} & * \\ \hline 0 & \cdots & 0 & 0 & B \end{array}\right)$$

où B est une matrice carrée de taille $p = n - (n_1 + \cdots + n_r)$. En particulier, A est triangulaire par blocs. Donc, d'après 3.4.8, on a

$$P_u(X) = \operatorname{dét}(A - XI_n) = P_B(X) \prod_{i=1}^{r} (\lambda_i - X)^{n_i}.$$

Donc $\prod_{i=1}^{r} (\lambda_i - X)^{n_i}$ divise $P_u(X)$, d'où $n_i \leq m_i$ pour tout i, ce qui prouve (2).

Si $n_i = m_i$ pour tout i, alors le sous-espace $E = \bigoplus_{i=1}^{r} V_{\lambda_i}$ est de dimension $\sum_{i=1}^{r} m_i = n$, donc égale V, donc u est diagonalisable. Réciproquement, si u est diagonalisable, il existe une base $\mathscr{B}$ de V telle que

$$A = \operatorname{Mat}_{\mathscr{B}}(u) = \left(\begin{array}{c|c|c|c} \lambda_1 I_{n_1} & 0 & \cdots & 0 \\ \hline 0 & \lambda_2 I_{n_2} & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & \vdots \\ \hline 0 & \cdots & 0 & \lambda_r I_{n_r} \end{array}\right)$$

alors $P_u(X) = \operatorname{dét}(A - XI_n) = \prod_{i=1}^{r} (\lambda_i - X)^{n_i} = (-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{n_i}$, d'où $n_i = m_i$ pour tout i. $\square$

Donnons de plus une propriété remarquable des endomorphismes diagonalisables, qui sera utile par la suite. Commençons par une définition:

***Définition 4.1.13* (Restriction de u à un sous-espace stable).** — Soit $u \in \mathrm{End}_k(V)$. On dit qu'un sous-espace E de V est stable par u si $u(E) \subseteq E$. Dans ce cas, la restriction de u à E induit un endomorphisme de E, que l'on notera u_E.

***Théorème 4.1.14* (Restriction d'un endomorphisme diagonalisable)**
Soient V un k-espace vectoriel de dimension finie, u un endomorphisme diagonalisable de V, et E un sous-espace de V stable par u. Alors E admet une base formée de vecteurs propres de u, i.e. la restriction u_E de u à E est diagonalisable.

Démonstration. — D'après la proposition précédente, il suffit de montrer que E est engendré par des vecteurs propres de u. Comme u est diagonalisable, tout $x \in E$ s'écrit dans V comme une somme de vecteurs propres :

$$(\dagger) \qquad x = x_1 + \cdots + x_r, \qquad \text{avec} \quad x_i \in V_{\mu_i} \quad \text{et} \quad \mu_i \neq \mu_j \text{ si } i \neq j.$$

Montrons par récurrence sur r que pour tout $x \in E$ et toute écriture $(\dagger)$ comme ci-dessus, chaque x_i appartient à E (ce qui prouvera le théorème). C'est OK pour $r = 1$, donc on peut supposer $r \geq 2$ et le résultat démontré pour $r - 1$. Appliquant $u - \mu_r \,\mathrm{id}_V$ à $(\dagger)$ on obtient

$$x' = (u - \mu_r \,\mathrm{id}_V)(x) = \sum_{i=1}^{r-1} (\mu_i - \mu_r) x_i$$

et $x' \in E$ puisque E est stable par u. Donc par hypothèse de récurrence, $(\mu_i - \mu_r)x_i$ appartient à E pour $i = 1, \ldots, r-1$, donc x_i y appartient aussi (puisque $\mu_i - \mu_r \neq 0$), et reportant ceci dans $(\dagger)$ on obtient aussi $x_r \in E$. Ceci prouve le théorème. □

Pour terminer ce paragraphe, donnons encore l'exemple ci-dessous d'endomorphismes diagonalisables.

Rappels 4.1.15. — Soit k un corps. Si $n \cdot 1_k = 1_k + \cdots + 1_k$ (n termes) est $\neq 0$ pour tout entier $n > 0$, on dit que k est de caractéristique 0; c'est le cas par exemple pour $\mathbb{Q}, \mathbb{R}, \mathbb{C}$. Sinon, le plus petit entier $p > 0$ tel que $p \cdot 1_k = 0$ est nécessairement un nombre premier (car si $p = rs$ avec $r, s \geq 1$, l'égalité $0 = (r \cdot 1_k)(s \cdot 1_k)$ entraîne que $r \cdot 1_k = 0$ ou $s \cdot 1_k = 0$, disons $r \cdot 1_k = 0$, mais alors la minimalité de p entraîne que $r = p$); dans ce cas on dit que k est de caractéristique p. D'autre part, si V est un k-espace vectoriel et $p \in \mathrm{End}_k(V)$, rappelons qu'on dit que p est un **projecteur** si $p^2 = p \circ p$ est égal à p.

***Proposition 4.1.16* (Symétries).** — *Soient k un corps de caractéristique $\neq 2$ (par exemple, $k = \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$), V un k-espace vectoriel de dimension n, et $s \in \mathrm{End}_k(V)$ tel que $s^2 = \mathrm{id}_V$. Alors s est diagonalisable; plus précisément, soient*

$$p_+ = \frac{\mathrm{id}_V + s}{2}, \qquad p_- = \frac{\mathrm{id}_V - s}{2}, \qquad V_\pm = \mathrm{Im}(p_\pm).$$

Alors p_+ et p_- sont des projecteurs et l'on a:

$$V = V_+ \oplus V_- \qquad \textit{et} \qquad \forall x \in V_\pm, \quad s(x) = \pm x.$$

Donc, si $s \neq \pm \mathrm{id}_V$, alors V_+ et V_- sont non nuls, et $V_\pm$ est l'espace propre associé à la valeur propre ± 1; dans ce cas, s est la symétrie par rapport à V_+ parallèlement à V_-.

Démonstration. — On a $p_+^2 = p_+$, $p_-^2 = p_-$ et $p_- = \mathrm{id}_V - p_+$ d'où $p_+ p_- = 0 = p_- p_+$, donc p_+ et p_- sont des projecteurs et l'on a $V = V_+ \oplus V_-$. De plus, si $x \in V_\pm$, on voit aussitôt que $s(x) = \pm x$, d'où la proposition. □

Remarque 4.1.16.1. — Attention, si k est de caractéristique 2, c.-à-d., si $2 = 0$ dans k (par exemple, si k est le corps à deux éléments $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$), la matrice $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in M_2(k)$ vérifie $A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = I_2$ mais A n'est pas diagonalisable: en effet sa seule valeur propre est 1, donc si A était diagonalisable on aurait $A = I_2$, ce qui n'est pas le cas.

4.2. Trigonalisation

***Définition 4.2.1* (Endomorphismes trigonalisables).** — Soit $u \in \mathrm{End}_k(V)$ ($\dim_k(V) = n$). On dit que u est *trigonalisable* s'il existe une base $\mathscr{B}$ de V dans laquelle la matrice $A = \mathrm{Mat}_{\mathscr{B}}(u)$ est triangulaire, disons supérieure. [(1)]

Dans ce cas, soient $\lambda_1, \ldots, \lambda_n$ les coefficients diagonaux, et soit X une indéterminée. Alors, d'après 3.4.8, on a

$$P_u(X) = \text{dét}(A - XI_n) = \prod_{i=1}^{n} (\lambda_i - X)$$

donc $\lambda_1, \ldots, \lambda_n$ sont les n racines (comptées avec multiplicités) du polynôme caractéristique $P_u(X)$. On voit donc qu'une condition *nécessaire* pour que u soit trigonalisable est que $P_u(X)$ ait toutes ses racines dans k. Ceci conduit à la définition suivante:

***Définition 4.2.2* (Polynômes scindés, corps algébriquement clos)**

Soient k un corps et $P \in k[X]$ un polynôme de degré $n \geq 1$.

(1) On dit que P est *scindé* dans $k[X]$ s'il admet n racines $\lambda_1, \ldots, \lambda_n$ dans k (comptées avec multiplicités), c.-à-d., si P se factorise dans $k[X]$ en produit de facteurs de degré 1 :

$$P = a(X - \lambda_1) \cdots (X - \lambda_n)$$

(où a est le coefficient dominant de P).

(2) On dit que k est *algébriquement clos* si tout polynôme $P \in k[X]$ de degré ≥ 1 est scindé. Par exemple, on sait que $\mathbb{C}$ est algébriquement clos (une démonstration est donnée dans un appendice à ce chapitre).

Pour simplifier l'exposition, on suppose dans la suite de ce paragraphe 4.2 que $k = \mathbb{C}$. (Mais tous les résultats, à l'exception du corollaire 4.3.11, sont valables sur un corps algébriquement clos arbitraire [et le corollaire 4.3.11 est valable pour tout corps algébriquement clos de caractéristique 0]).

***Théorème 4.2.3* (Trigonalisation).** — (1) *Soient V un $\mathbb{C}$-espace vectoriel de dimension n et $u \in \mathrm{End}_{\mathbb{C}}(V)$. Il existe une base $\mathscr{C}$ de V telle que $\mathrm{Mat}_{\mathscr{C}}(u)$ soit triangulaire supérieure (les coefficients diagonaux $\lambda_1, \ldots, \lambda_n$ sont alors les racines, comptées avec multiplicité, de $P_u(X)$ dans $\mathbb{C}$).*

En termes matriciels: toute matrice $A \in M_n(\mathbb{C})$ est semblable à une matrice triangulaire, i.e. il existe $P \in \mathrm{GL}_n(\mathbb{C})$ telle que $P^{-1}AP$ soit triangulaire.

(2) *Plus généralement, soient k un corps, V un k-espace vectoriel de dimension n et $u \in \mathrm{End}_k(V)$; on suppose que $P_u(X)$ est scindé, i.e. qu'il a n racines $\lambda_1, \ldots, \lambda_n$ (pas nécessairement distinctes) dans k. Alors il existe une base $\mathscr{B}$ de V dans laquelle la matrice de u est triangulaire supérieure, les coefficients diagonaux étant alors $\lambda_1, \ldots, \lambda_n$.*

En termes matriciels: si le polynôme caractéristique de $A \in M_n(k)$ a toutes ses racines dans k, alors A est semblable à une matrice triangulaire, i.e. il existe $P \in \mathrm{GL}_n(k)$ telle que $P^{-1}AP$ soit triangulaire.

Démonstration. — (1) On procède par récurrence sur $n = \dim(V)$. Il n'y a rien à montrer si $n = 1$, donc on peut supposer $n \geq 2$ et le théorème démontré pour $n-1$. Soient $\mathscr{B}$ une base arbitraire de V et $A = \mathrm{Mat}_{\mathscr{B}}(u)$. Considérons la matrice $B = {}^tA$. Comme $\mathbb{C}$ est algébriquement clos, le polynôme caractéristique de B a au moins une racine λ dans $\mathbb{C}$ donc, d'après 3.5.6, il existe un vecteur $T = \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} \neq 0$ tel que ${}^tAT = BT = \lambda T$.

Prenant la transposée de cette égalité, on obtient:

$$(\star) \qquad \lambda(t_1, \ldots, t_n) = (t_1, \ldots, t_n)\,{}^t({}^tA) = (t_1, \ldots, t_n)A.$$

D'après 1.3.2, le sous-espace H de V défini par l'équation $t_1x_1 + \cdots + t_nx_n = 0$, i.e. le sous-espace de V formé des vecteurs colonnes $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ tels que $(t_1, \ldots, t_n)X = (t_1, \ldots, t_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0$, est de dimension $n-1$. Pour tout $X \in H$, on a, d'après $(\star)$:

$$(t_1, \ldots, t_n)AX = \lambda \underbrace{(t_1, \ldots, t_n)X}_{=0} = 0,$$

[(1)] Si la matrice de u dans une base $(v_1, \ldots, v_n)$ est triangulaire supérieure, alors la matrice dans la base $(v_n, \ldots, v_1)$ est triangulaire inférieure, et vice-versa, donc on pourrait dans la définition remplacer le mot "supérieure" par "inférieure".

donc $AX \in H$. Comme AX représente dans la base $\mathscr{B}$ le vecteur $u(X)$, ceci montre que H est stable par u. Comme $\dim H = n-1$, alors par hypothèse de récurrence il existe une base $\mathscr{C}'$ de H telle $\mathrm{Mat}_{\mathscr{C}'}(u_H)$ soit triangulaire supérieure. Choisissons arbitrairement $w \in V - H$, alors la famille $\mathscr{C} = \mathscr{C}' \cup \{w\}$ est libre, donc est une base de V, et $\mathrm{Mat}_{\mathscr{C}}(u)$ est de la forme :

$$A' = \left(\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,n-1} & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n-1} & a_{2,n} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & a_{n-1,n-1} & a_{n-1,n} \\ \hline 0 & \cdots & 0 & 0 & a_{n,n} \end{array}\right) \qquad \text{où} \qquad \mathrm{Mat}_{\mathscr{C}'}(u_H) = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n-1} \\ 0 & a_{2,2} & \cdots & a_{2,n-1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n-1,n-1} \end{pmatrix}$$

donc triangulaire supérieure. Enfin, on a déjà noté en 4.2.1 que, dans ce cas, les coefficients diagonaux sont les racines (comptées avec multiplicité) de $P_u(X)$ dans $\mathbb{C}$. Ceci prouve (1).

Prouvons l'assertion plus générale (2), i.e. k est un corps arbitraire et l'on suppose que $P = P_u(X)$ a n racines $\lambda_1, \dots, \lambda_n$ (pas nécessairement distinctes) dans k. On procède par récurrence sur $n = \dim(V)$. Il n'y a rien à montrer si $n = 1$, donc on peut supposer $n \geq 2$ et le théorème démontré pour $n-1$. Soit $\mathscr{B}$ une base arbitraire de V et soient $A = \mathrm{Mat}_{\mathscr{B}}(u)$ et $B = {}^tA$. Comme

$${}^t(A - XI_n) = B - XI_n$$

alors $P_B(X) = P_A(X) = P_u(X) = P$, donc $\lambda = \lambda_1$ est racine de $P_B(X)$, donc valeur propre de B (cf. 3.5.6). Il existe donc un vecteur $T = \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} \neq 0$ tel que ${}^tAT = BT = \lambda T$. Prenant la transposée de cette égalité, on obtient:

$$(\star) \qquad \lambda(t_1, \dots, t_n) = (t_1, \dots, t_n)\,{}^t({}^tA) = (t_1, \dots, t_n)A.$$

D'après 1.3.2, le sous-espace H de V défini par l'équation $t_1x_1 + \cdots + t_nx_n = 0$, i.e. le sous-espace de V formé des vecteurs colonnes $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ tels que $(t_1, \dots, t_n)X = (t_1, \dots, t_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0$, est de dimension $n-1$. Pour tout $X \in H$, on a, d'après $(\star)$:

$$(t_1, \dots, t_n)AX = \lambda \underbrace{(t_1, \dots, t_n)X}_{=0} = 0,$$

donc $AX \in H$. Comme AX représente dans la base $\mathscr{B}$ le vecteur $u(X)$, ceci montre que H est stable par u. Soient $\mathscr{B}_H$ une base de H et $w \in V - H$, alors la famille $\widetilde{\mathscr{B}} = \mathscr{B}_H \cup \{w\}$ est libre, donc est une base de V. Soient u_H la restriction de u à H et $A' = \mathrm{Mat}_{\mathscr{B}_H}(u_H) \in M_{n-1}(k)$; alors $\mathrm{Mat}_{\widetilde{\mathscr{B}}}(u)$ est de la forme:

$$(*) \qquad \widetilde{A} = \left(\begin{array}{c|c} A' & Y \\ \hline \mathbf{0}_{1,n-1} & \mu \end{array}\right)$$

où $\mu \in k$, $Y \in M_{n-1,1}(k)$ est une matrice colonne, et $\mathbf{0}_{1,n-1}$ est la matrice ligne $(0, \dots, 0)$.

Ceci entraîne que $P_{A'}(X)(\mu - X) = P_{\widetilde{A}}(X) = P_u(X)$. Il en résulte que μ est l'une des racines de P, disons λ_i, et que $P_{A'}(X)$ est le produit, pour $j \neq i$, des $(\lambda_j - X)$; en particulier, $P_{A'}(X)$ a toutes ses racines dans k. (En fait, on peut montrer que $\mu = \lambda = \lambda_1$, voir plus bas). Comme $\dim H = n-1$, alors par hypothèse de récurrence il existe une base $\mathscr{C}'$ de H telle $\mathrm{Mat}_{\mathscr{C}'}(u_H)$ soit triangulaire supérieure (avec les λ_j, pour $j \neq i$, comme termes diagonaux). Alors $\mathscr{C} = \mathscr{C}' \cup \{w\}$ est une base de V, et $\mathrm{Mat}_{\mathscr{C}}(u)$ est de la forme :

$$A' = \left(\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,n-1} & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n-1} & a_{2,n} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & a_{n-1,n-1} & a_{n-1,n} \\ \hline 0 & \cdots & 0 & 0 & \mu \end{array}\right) \qquad \text{où} \qquad \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n-1} \\ 0 & a_{2,2} & \cdots & a_{2,n-1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n-1,n-1} \end{pmatrix} = \mathrm{Mat}_{\mathscr{C}'}(u_H),$$

donc triangulaire supérieure. Enfin, on a déjà noté en 4.2.1 que, dans ce cas, les coefficients diagonaux sont les racines (comptées avec multiplicité) de $P_u(X)$ dans k. Le théorème est démontré.

Remarquons enfin que $\mu = \lambda$. En effet, ${}^tA$ est la matrice de l'endomorphisme ${}^tu : V^* \to V^*$, $\phi \mapsto \phi \circ u$, et le vecteur T obtenu au début de la démonstration correspond donc à une forme linéaire $f \in V^*$ telle que

$$f \circ u = ({}^tu)(f) = \lambda f.$$

De plus, on a $H = \mathrm{Ker}(f)$ et puisque $w \notin H$ alors le scalaire $z = f(w)$ est $\neq 0$. On a $u(w) = \mu w + y$, avec $y \in H$ et donc

$$f(u(w)) = f(\mu w + y) = \mu f(w). \tag{†}$$

D'autre part,

$$f(u(w)) = (f \circ u)(w) = (\lambda f)(w) = \lambda f(w), \tag{‡}$$

et comme $f(w) \neq 0$ la comparaison de (†) et (‡) donne $\mu = \lambda$. □

Corollaire 4.2.4 **(Déterminant (resp. trace) = produit (resp. somme) des valeurs propres)**
Soient $A \in M_n(\mathbb{C})$ et $\lambda_1, \dots, \lambda_n$ les n racines dans $\mathbb{C}$ (comptées avec multiplicité) du polynôme caractéristique $P_A(X)$. Alors $\boxed{\mathrm{dét}(A) = \lambda_1 \cdots \lambda_n}$ *et* $\boxed{\mathrm{Tr}(A) = \lambda_1 + \cdots + \lambda_n.}$

Démonstration. — D'après le théorème 4.2.3, il existe $P \in \mathrm{GL}_n(\mathbb{C})$ telle que $A' = P^{-1}AP$ soit triangulaire; notons $\lambda_1, \dots, \lambda_n$ ses coefficients diagonaux, alors

$$P_{A'}(X) = \prod_{i=1}^{n}(\lambda_i - X), \qquad \mathrm{dét}(A') = \lambda_1 \cdots \lambda_n, \qquad \mathrm{Tr}(A') = \lambda_1 + \cdots + \lambda_n.$$

D'autre part, d'après le théorème 3.5.2, A' et A ont même polynôme caractéristique, même déterminant et même trace. Alors $\lambda_1, \dots, \lambda_n$ sont les racines de $P_{A'}(X) = P_A(X)$ et l'on a $\mathrm{dét}(A) = \mathrm{dét}(A') = \lambda_1 \cdots \lambda_n$ et $\mathrm{Tr}(A) = \mathrm{Tr}(A') = \lambda_1 + \cdots + \lambda_n$. □

4.3. Polynomes d'endomorphismes et applications

Définition 4.3.1 **(Polynômes d'endomorphismes ou de matrices)**. — Soit $Q = a_0 + a_1X + \cdots + a_dX^d$ un polynôme à coefficients dans k.

(1) Soient V un k-espace vectoriel et $u \in \mathrm{End}_k(V)$. On pose

$$Q(u) = a_0 \,\mathrm{id}_V + a_1 u + \cdots + a_d u^d$$

où, bien sûr, u^i désigne $u \circ \cdots \circ u$ (i fois); alors $Q(u)$ est un élément de $\mathrm{End}_k(V)$, qu'on appelle un "polynôme en u". On vérifie aussitôt que si $R \in k[X]$ est un second polynôme, alors

$$Q(u) \circ R(u) = (QR)(u) = (RQ)(u) = R(u) \circ Q(u);$$

en particulier, "les polynômes en u commutent entre eux". Donc, l'application $\phi : k[X] \to \mathrm{End}_k(V)$, $Q \mapsto Q(u)$, est un *homomorphisme de k-algèbres*, c.-à-d., ϕ est linéaire, envoie $1 \in k[X]$ sur id_V et vérifie $\phi(QR) = \phi(Q)\phi(R)$.

(2) De même, si $A \in M_n(k)$, on note $Q(A)$ la matrice $a_0I_n + a_1A + \cdots + a_dA^d$. Si $R \in k[X]$ est un second polynôme, alors

$$Q(A)R(A) = (QR)(A) = (RQ)(A) = R(A)Q(A);$$

en particulier, "les polynômes en A commutent entre eux". De plus, si $\dim(V) = n$ et si A est la matrice de u dans une base $\mathscr{B}$, alors la matrice dans $\mathscr{B}$ de $Q(u)$ est $Q(A)$.

Remarque 4.3.1.1. — Soit $u \in \mathrm{End}_k(V)$. Si $v \in V$ est un vecteur propre de u pour la valeur propre λ, alors

$$u^2(v) = u(u(v)) = u(\lambda v) = \lambda u(v) = \lambda^2 v$$

et l'on montre ainsi, par récurrence sur i, que $u^i(v) = \lambda^i v$ pour tout $i \in \mathbb{N}^*$ (et aussi pour $i = 0$, avec la convention $u^0 = \mathrm{id}_V$ et $\lambda^0 = 1$). On en déduit le lemme suivant.

Lemme 4.3.2. — *Soient $u \in \mathrm{End}_k(V)$ et $v \in V$ un vecteur propre pour une valeur propre $\lambda \in k$. Alors, pour tout $Q \in k[X]$ on a $Q(u)(v) = Q(\lambda)\, v$.*

En particulier, si $Q(u) = 0$, alors λ est une racine de Q.

Théorème 4.3.3 **(Théorème de Cayley-Hamilton)**. — *Soit V un $\mathbb{C}$-espace vectoriel de dimension n, et soient $u \in \mathrm{End}_{\mathbb{C}}(V)$ et $P_u(X)$ son polynôme caractéristique. Alors l'endomorphisme $P_u(u)$ est nul, c.-à-d.: "u est annulé par son polynôme caractéristique".*

Démonstration. — On a $P_u(X) = (-1)^n \prod_{i=1}^n (X - \lambda_i)$, où $\lambda_1, \ldots, \lambda_n$ sont les n racines (comptées avec multiplicité) de $P_u(X)$ dans $\mathbb{C}$. D'après le théorème 4.2.3, il existe une base $\mathscr{B} = (f_1, \ldots, f_n)$ de V dans laquelle la matrice de u est triangulaire supérieure, avec $\lambda_1, \ldots, \lambda_n$ sur la diagonale. Ceci équivaut à dire que, pour tout $i = 1, \ldots, n$, le sous-espace F_i de V engendré par $f_1, \ldots, f_i$ est stable par u et, plus précisément, que l'on a, pour $i = 1, \ldots, n$:

$$(u - \lambda_i \,\mathrm{id}_V)(F_i) \subseteq F_{i-1},$$

avec la convention $F_0 = \{0\}$. Comme $F_n = V$, on déduit des inclusions ci-dessus que $(u - \lambda_n \,\mathrm{id}_V)(V) \subseteq F_{n-1}$, puis que $(u - \lambda_{n-1} \,\mathrm{id}_V)(u - \lambda_n \,\mathrm{id}_V)(V) \subseteq F_{n-2}$, etc., d'où finalement :

$$(u - \lambda_1 \,\mathrm{id}_V) \cdots (u - \lambda_n \,\mathrm{id}_V)(V) \subseteq F_0 = \{0\}.$$

Ceci montre que $(-1)^n P_u(u) = 0$, d'où $P_u(u) = 0$. □

***Corollaire 4.3.4* (Cayley-Hamilton pour $k \subseteq \mathbb{C}$).** — *Soient k un sous-corps de $\mathbb{C}$ (par exemple $k = \mathbb{R}$) et V un k-espace vectoriel de dimension n, et soient $u \in \mathrm{End}_k(V)$ et $P_u(X)$ son polynôme caractéristique. Alors $P_u(u) = 0$*

Démonstration. — Soient $\mathscr{B}$ une base de V et $A = \mathrm{Mat}_{\mathscr{B}}(u) \in M_n(k)$. D'une part, $P_u(X) = P_A(X)$, notons P ce polynôme. D'autre part, comme $k \subseteq \mathbb{C}$, on peut considérer A comme élément de $M_n(\mathbb{C})$, donc d'après Cayley-Hamilton on a $P(A) = 0$. Or $P(A)$ est la matrice dans la base $\mathscr{B}$ de $P(u)$, d'où $P(u) = 0$. □

***Définition 4.3.5* (Espaces caractéristiques).** — Soient V un $\mathbb{C}$-espace vectoriel de dimension n, $u \in \mathrm{End}_{\mathbb{C}}(V)$, $\lambda \in \mathbb{C}$ une valeur propre de u, et m sa multiplicité algébrique (i.e. sa multiplicité comme racine de $P_u(X)$). On pose

$$V_{(\mu)} = \mathrm{Ker}\left((u - \lambda \,\mathrm{id}_V)^m\right)$$

et on l'appelle l'espace caractéristique associé à μ; il est stable par u (car il est stable par $u - \lambda \,\mathrm{id}_V$ donc aussi par $u = (u - \lambda \,\mathrm{id}_V) + \lambda \,\mathrm{id}_V$). Notons aussi que $V_{(\mu)}$ contient l'espace propre $V_\mu = \mathrm{Ker}(u - \lambda \,\mathrm{id}_V)$.

Exemple 4.3.6. — Si u est l'endomorphisme de $\mathbb{C}^2$ de matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ (i.e. , $u(e_2) = e_1$ et $u(e_1) = 0$), alors $P_u(X) = X^2$ a 0 comme unique racine. D'après Cayley-Hamilton (ou par un calcul direct), on a $u^2 = 0$, donc $V_{(0)} = \mathbb{C}^2$ tandis que $V_0 = \mathrm{Ker}(u) = \mathbb{C}e_1$.

Afin de démontrer plus bas un théorème sur les espaces caractéristiques, on aura besoin du théorème ci-dessous, que les étudiants de 2M220 ont déjà vu (une démonstration est donnée en appendice à la fin de ce chapitre).

***Théorème 4.3.7* (Théorème de Bézout).** — *Soient $P_1, \ldots, P_r \in \mathbb{C}[X]$ des polynômes sans racine commune. Alors il existe des polynômes $S_1, \ldots, S_r \in \mathbb{C}[X]$ tels que $P_1 S_1 + \cdots + P_r S_r = 1$.*

***Théorème 4.3.8* (Décomposition en espaces caractéristiques).** — *Soient V un $\mathbb{C}$-espace vectoriel de dimension n et $u \in \mathrm{End}_{\mathbb{C}}(V)$. Écrivons $P_u(X) = (-1)^n \prod_{i=1}^r (X - \lambda_i)^{m_i}$, où $\lambda_1, \ldots, \lambda_r$ sont les valeurs propres, deux à deux distinctes, de u et m_i est la multiplicité algébrique de λ_i. Alors :*

(1) *V est la somme directe des espaces caractéristiques $V_{(\lambda_1)}, \ldots, V_{(\lambda_r)}$.*

(2) *On a $\dim V_{(\lambda_i)} = m_i$ pour tout i.*

Démonstration. — Pour tout i, posons

$$P_i = (-1)^n \prod_{j \neq i} (X - \lambda_j)^{m_j}.$$

Alors $P_1, \ldots, P_r \in \mathbb{C}[X]$ sont sans racines communes, donc d'après le théorème de Bézout il existe des polynômes $S_1, \ldots, S_r \in \mathbb{C}[X]$ tels que $P_1 S_1 + \cdots + P_r S_r = 1$. Prenant les polynômes en u correspondants, on obtient :

$$P_1(u) S_1(u) + \cdots + P_r(u) S_r(u) = \mathrm{id}_V$$

Donc, pour tout $v \in V$, on a :

$$(\star) \qquad v = P_1(u)(x_1) + \cdots + P_r(u)(x_r), \qquad \text{où} \quad x_i = S_i(u)(v).$$

D'autre part, d'après Cayley-Hamilton, $P_u(u) = 0$; comme de plus $P_u(X) = (X - \lambda_i)^{m_i} P_i(X)$ alors

$$(u - \lambda_i \,\mathrm{id}_V)^{m_i} \left(P_i(u)(x_i)\right) = P_u(u)(x_i) = 0, \qquad \text{pour tout } i.$$

Ceci montre que chaque $v_i = P_i(u)(x_i)$ appartient à l'espace caractéristique $V_{(\lambda_i)} = \operatorname{Ker}\big((u - \lambda_i \operatorname{id}_V)^{m_i}\big)$; par conséquent l'égalité $(\star)$ montre déjà que $V = \sum_{i=1}^r V_{(\lambda_i)}$.

Reste à montrer que la somme des $V_{(\lambda_i)}$ est directe: supposons qu'on ait une égalité

$$v_1 + \cdots + v_r = 0, \qquad \text{avec} \quad v_i \in V_{(\lambda_i)}$$

et montrons que chaque v_i est nul. Fixons un indice i, alors $v_i = -\sum_{j\neq i} v_j$ appartient à $V_{(\lambda_i)} \cap \sum_{j\neq i} V_{(\lambda_j)}$ donc est annulé par $(u - \lambda_i \operatorname{id}_V)^{m_i}$ et par $P_i(u) = (-1)^n \prod_{j\neq i}(u - \lambda_j \operatorname{id}_V)^{m_j}$. Or, les polynômes $Q_i(X) = (X - \lambda_i)^{m_i}$ et $P_i(X)$ n'ont pas de racines communes donc, d'après Bézout, il existe $R, T \in \mathbb{C}[X]$ tels que $RP_i + TQ_i = 1$. Alors $\operatorname{id}_V = R(u)P_i(u) + T(u)Q_i(u)$, d'où

$$v_i = \operatorname{id}_V(v_i) = R(u)\underbrace{P_i(u)(v_i)}_{=0} + T(u)\underbrace{Q_i(u)(v_i)}_{=0} = 0.$$

Ceci montre que la somme des $V_{(\lambda_i)}$ est directe, et achève la démonstration de l'assertion (1).

Enfin, pour tout i, soit $d_i = \dim V_{(\lambda_i)}$, soit u_i la restriction de u à $V_{(\lambda_i)}$ et soit $\mathscr{B}_i$ une base de $\dim V_{(\lambda_i)}$. Alors $\mathscr{B} = \mathscr{B}_1 \cup \cdots \cup \mathscr{B}_r$ est une base de V et, notant $A_i = \operatorname{Mat}_{\mathscr{B}_i}(u_i)$, on a :

$$(\dagger) \qquad \operatorname{Mat}_{\mathscr{B}}(u) = \left(\begin{array}{c|c|c|c} A_1 & 0 & \cdots & 0 \\ \hline 0 & A_2 & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & 0 \\ \hline 0 & \cdots & 0 & A_r \end{array}\right) \qquad \text{d'où} \quad P_u(X) = \prod_{i=1}^r P_{u_i}(X).$$

D'autre part, si μ est une valeur propre de u_i et $x \neq 0$ un vecteur propre associé, on a $(u_i - \lambda_i \operatorname{id}_V)(x) = (\mu - \lambda_i)x$, d'où $(u_i - \lambda_i \operatorname{id}_V)^p(x) = (\mu - \lambda_i)^p x$ pour tout $p \in \mathbb{N}^*$. Or $(u_i - \lambda_i \operatorname{id}_V)^{m_i} = 0$, d'où $(\mu - \lambda_i)^{m_i} x = 0$ et donc $\mu = \lambda_i$. Ceci montre que λ_i est la seule valeur propre de u_i, d'où $P_{u_i}(X) = (-1)^{d_i}(X - \lambda_i)^{d_i}$. D'après $(\dagger)$, on obtient donc que

$$P_u(X) = (-1)^n \prod_{i=1}^r (X - \lambda_i)^{d_i}$$

d'où l'égalité $d_i = m_i$ pour tout i. Le théorème est démontré. □

Remarque 4.3.9. — On conserve les notations précédentes. En fait, chaque espace caractéristique $V_{(\lambda_i)}$ est l'ensemble de tous les $v \in V$ qui sont annulés par une *certaine puissance* de $(u - \lambda_i \operatorname{id}_V)$, i.e. on a l'égalité

$$V_{(\lambda_i)} = \{v \in V \mid \exists p \in \mathbb{N}^* \text{ tel que } (u - \lambda_i \operatorname{id}_V)^p(v) = 0\}.$$

En effet, l'inclusion $\subseteq$ est claire, prouvons la réciproque. Quitte à renuméroter les λ_i, il suffit de le faire pour λ_1. Supposons que $(u - \lambda_1 \operatorname{id}_V)^p(v) = 0$. On peut écrire $v = v_1 + v_2 + \cdots + v_r$, avec $v_i \in V_{(\lambda_i)}$. Comme chaque $V_i = V_{(\lambda_i)}$ est stable par u donc aussi par $u - \lambda_1 \operatorname{id}_V$, l'égalité

$$0 = (u - \lambda_1 \operatorname{id}_V)^p(v) = (u - \lambda_1 \operatorname{id}_V)^p(v_1) + \cdots + (u - \lambda_1 \operatorname{id}_V)^p(v_r)$$

jointe au fait que les V_i sont en somme directe, entraîne que $(u - \lambda_i \operatorname{id}_V)^p(v_i) = 0$ pour tout i. Or on a vu que la restriction u_i de u à V_i a pour polynôme caractéristique $(\lambda_i - X)^{m_i}$, donc $\det(u_i - \lambda_1 \operatorname{id}_{V_i}) = (\lambda_i - \lambda_1)^{m_i}$ est non nul lorsque $i \neq 1$, donc l'égalité $(u - \lambda_i \operatorname{id}_V)^p(v_i) = 0$ entraîne $v_i = 0$ pour $i \neq 1$. Donc v égale v_1 et appartient donc à $V_{(\lambda_1)}$.

Remarquons aussi qu'une démonstration analogue (et en fait, plus simple) donne la :

***Proposition 4.3.10* (Polynômes sans racines multiples).** — *Soient V un $\mathbb{C}$-espace vectoriel de dimension n et $u \in \operatorname{End}_{\mathbb{C}}(V)$. On suppose que $Q(u) = 0$, où $Q \in \mathbb{C}[X]$ un polynôme de degré $d \geq 1$ ayant d racines* distinctes *dans* $\mathbb{C}$. *Alors u est diagonalisable.*

Démonstration. — Par hypothèse, on a $Q(X) = c\prod_{i=1}^d (X - \alpha_i)$, où $c \in \mathbb{C}^*$ est le coefficient dominant de Q, et $\alpha_1, \ldots, \alpha_d \in \mathbb{C}$ sont deux à deux distincts. Pour tout i, posons

$$P_i = c\prod_{j\neq i}(X - \alpha_j).$$

Alors $P_1, \ldots, P_d \in \mathbb{C}[X]$ sont sans racines communes donc, d'après le théorème de Bézout il existe des polynômes $S_1, \ldots, S_d \in \mathbb{C}[X]$ tels que $P_1S_1 + \cdots + P_dS_d = 1$. Prenant les polynômes en u correspondants, on obtient :

$$P_1(u)S_1(u) + \cdots + P_d(u)S_d(u) = \operatorname{id}_V$$

Donc, pour tout $v \in V$, on a :

$$(\star) \qquad v = P_1(u)(x_1) + \cdots + P_d(u)(x_d), \qquad \text{où} \quad x_i = S_i(u)(v).$$

Par hypothèse, $Q(u)$ est l'endomorphisme nul; comme de plus $Q(X) = (X - \alpha_i)P_i(X)$ alors

$$(u - \alpha_i \,\mathrm{id}_V)\big(P_i(u)(x_i)\big) = Q(u)(x_i) = 0,$$

pour tout i. Ceci montre que chaque $v_i = P_i(u)(x_i)$ appartient à $\mathrm{Ker}(u - \alpha_i \,\mathrm{id}_V)$; par conséquent l'égalité $(\star)$ montre que V est la somme des espaces $\mathrm{Ker}(u - \alpha_i \,\mathrm{id}_V)$.

Or, ceux de ces espaces qui sont non nuls sont des espaces propres de u, donc sont en somme directe (cf. 4.1.8). Donc, notant I le sous-ensemble de $\{1, \dots, d\}$ formé des i tels que $\mathrm{Ker}(u - \alpha_i \,\mathrm{id}_V) \neq \{0\}$, on obtient que $V = \bigoplus_{i \in I} V_{\alpha_i}$, ce qui montre que u est diagonalisable. □

Corollaire 4.3.11 **(Automorphismes d'ordre fini de $\mathbb{C}^n$).** — *Soient V un $\mathbb{C}$-espace vectoriel de dimension n et $u \in \mathrm{End}_{\mathbb{C}}(V)$ tel que $u^d = \mathrm{id}_V$ pour un entier $d \geq 1$ (dans ce cas, on dit que u est un automorphisme d'ordre fini). Alors u est diagonalisable (et ses valeurs propres sont des racines d-èmes de l'unité).*

Démonstration. — En effet, le polynôme $X^d - 1$ a d racines distinctes dans $\mathbb{C}$, à savoir $\exp(\dfrac{2ir\pi}{d}) = \cos(\dfrac{2ir\pi}{d}) + i\sin(\dfrac{2ir\pi}{d})$ pour $r = 0, 1, \dots d-1$. □

Exercice 4.3.12. — Soit p un projecteur, c'est à dire un endomorphisme vérifiant $p^2 = p$. Trouver un polynome annulé par p. En déduire que les valeurs propres de P sont 0 et 1 et montrer que p est diagonalisable. Vérifier que l'espace propre V_1 est l'image de p.

On dit que p est le *projecteur* sur V_1 parallèlement à V_0.

4.4. Appendice (†) : somme directe externe d'espaces vectoriels

Définition 4.4.1 **(Sommes directes).** — Soient $V_1, \dots, V_n$ des k-espaces vectoriels. L'ensemble produit

$$V_1 \times \cdots \times V_n = \{(v_1, \dots, v_n) \mid v_i \in V_i\}$$

est muni d'une structure d'espace vectoriel définie "composante par composante, c.-à-d., $t \cdot (v_1, \dots, v_n) = (t \cdot v_1, \dots, t \cdot v_n)$ et $(v_1, \dots, v_n) + (v_1', \dots, v_n') = (v_1 + v_1', \dots, v_n + v_n')$.

On l'appelle la *somme directe* (externe) des V_i et on le note

$$V_1 \oplus \cdots \oplus V_n \qquad \text{ou} \qquad \bigoplus_{i=1}^{n} V_i.$$

De même, un n-uplet $(v_1, \dots, v_n)$ (avec $v_i \in V_i$) est aussi noté $v_1 + \cdots + v_n$ ou $\sum_{i=1}^n v_i$, c.-à-d., on identifie l'élément $v_i \in V_i$ au n-uplet $(0, \dots, 0, v_i, 0, \dots, 0)$ (où v_i est à la i-ème place).

Supposons que $\mathscr{B}_1 = (e_1, \dots, e_{d_1})$ soit une base de V_1, puis $\mathscr{B}_2 = (e_{d_1+1}, \dots, e_{d_1+d_2})$ une base de V_2, …puis $\mathscr{B}_n = (e_{d_1+\cdots+d_{n-1}+1}, \dots, e_{d_1+\cdots+d_n})$ une base de V_n. Alors tout élément v de $\bigoplus_{i=1}^n V_i$ s'écrit de façon unique

$$v = \underbrace{t_1 e_1 + \cdots + t_{d_1} e_{d_1}}_{v_1} + \cdots + \underbrace{t_{d_1+\cdots+d_{n-1}+1} e_{d_1+\cdots+d_{n-1}+1} + \cdots + t_{d_1+\cdots+d_n} e_{d_1+\cdots+d_n}}_{v_n}$$

avec les t_i dans k, d'où l'on déduit que la réunion disjointe des $\mathscr{B}_i$ est une base de $\bigoplus_{i=1}^n V_i$. Par conséquent, on a la formule:

$$(*) \qquad \boxed{\dim(\textstyle\bigoplus_{i=1}^n V_i) = d_1 + \cdots + d_n = \sum_{i=1}^n \dim(V_i).}$$

Remarque 4.4.2. — Supposons maintenant que $E_1, \dots, E_n$ soient des sous-espaces d'un k-espace vectoriel V. D'une part, on note $E_1 + \cdots + E_n$ le sous-espace de V engendré par $E_1 \cup \cdots \cup E_n$; c'est l'ensemble de toutes les sommes

$$(*) \qquad x_1 + \cdots + x_n, \qquad \text{avec } x_i \in E_i.$$

D'autre part, on peut former, la somme directe externe $S = E_1 \oplus \cdots \oplus E_n$ des E_i; ce n'est pas un sous-espace de V, mais on a une application linéaire naturelle

$$\sigma : E_1 \oplus \cdots \oplus E_n \to V, \qquad (x_1, \dots, x_n) \mapsto x_1 + \cdots + x_n$$

dont l'image est le sous-espace $E_1 + \cdots + E_n$ de V, et le noyau est le sous-espace de S formé des n-uplets $(x_1, \dots, x_n)$ tels que $x_1 + \cdots + x_n = 0$.

On voit donc que $\mathrm{Ker}(\sigma) = (0)$ si et seulement si les sous-espaces $E_1, \dots, E_n$ sont en somme directe dans V, et dans ce cas σ est un isomorphisme de la somme directe externe S sur le sous-espace de V noté $E_1 \oplus \cdots \oplus E_n$ en 4.1.1. Ceci justifie l'usage de la notation $E_1 \oplus \cdots \oplus E_n$ dans les deux cas. D'autre part, pour des espaces vectoriels arbitraires $E_1, \dots, E_n$, la somme directe "externe"$E_1 \oplus \cdots \oplus E_n$ sera appelée simplement "somme directe".

4.5. Appendice (†) : division euclidienne dans $\mathbb{C}[X]$ et théorème de Bézout

***Théorème 4.5.1* (Division euclidienne dans $k[X]$).** — *Soit k un corps et soit $P \in k[X]$ un polynôme de degré $d \geq 1$. Pour tout $F \in k[X]$, il existe un unique couple (Q, R) d'éléments de $k[X]$ tel que*

$$F = PQ + R, \qquad \textit{et} \quad R = 0 \ \textit{ou bien} \ \deg(R) < \deg(P).$$

On appelle Q (resp. R) le quotient (resp. le reste) de la division euclidienne de F par P.

Démonstration. — Montrons l'existence de (Q, R) en procédant par récurrence sur $\deg(F)$. Si $F = 0$ ou si $\deg(F) < d = \deg(P)$, on prend $Q = 0$ et $R = F$. Soit $n \geq d$ et supposons l'existence établie pour les degrés $< n$. Soit F de degré n. Notons a le coefficient dominant de F et c celui de P. Alors $ac^{-1}X^{n-d}P$ est de degré n et de coefficient dominant a, donc $F - ac^{-1}X^{n-d}P$ est de degré $< n$. Par hypothèse de récurrence, il existe $Q, R \in k[X]$ tels que

$$F - ac^{-1}X^{n-d}P = PQ + R, \qquad \text{et} \quad R = 0 \text{ ou bien } \deg(R) < \deg(P).$$

Alors $F = P(Q + ac^{-1}X^{n-d}) + R$, ce qui prouve le résultat d'existence.

Montrons l'unicité: si Q_1, R_1 vérifie les mêmes conditions, les égalités $PQ + R = F = PQ_1 + R_1$ donnent

$$P(Q - Q_1) = R_1 - R.$$

Si $Q - Q_1$ était $\neq 0$ alors $P(Q - Q_1)$ serait de degré $d + \deg(Q - Q_1) \geq d$. Or, $R_1 - R$ est nul ou de degré $< d$. Donc nécessairement $Q - Q_1 = 0$, d'où $R_1 - R = 0$, d'où $Q = Q_1$ et $R = R_1$. Ceci prouve l'unicité. □

Définitions 4.5.2. — 1) Soit I un sous-ensemble de $k[X]$. On dit que I est un *idéal* de $k[X]$ si c'est un sous-espace vectoriel et si, pour tout $P \in I$ et $S \in k[X]$, on a $SP \in I$.

2) Soient $P_1, \dots, P_r \in k[X]$. Alors l'ensemble I des sommes $S_1P_1 + \cdots + S_rP_r$ est un idéal de $k[X]$, et c'est le plus petit idéal de $k[X]$ contenant $P_1, \dots, P_r$. En effet, toute combinaison linéaire de telles sommes est encore une somme de ce type, donc I est un sous-espace vectoriel de $k[X]$; de plus, pour tout $S \in k[X]$, $S(S_1P_1 + \cdots + S_rP_r) = SS_1P_1 + \cdots + SS_rP_r$ est encore une somme de ce type, donc I est un idéal. Réciproquement, tout idéal J contenant $P_1, \dots, P_r$ contient toute les sommes $S_1P_1 + \cdots + S_rP_r$, donc contient I.

On dit que I est l'idéal *engendré par* $P_1, \dots, P_r$ et on le note $(P_1, \dots, P_r)$.

3) On dit qu'un idéal I de $k[X]$ est principal s'il peut être engendré par un seul élément, c.-à-d., s'il existe $P \in I$ tel que $I = \{SP \mid S \in k[X]\} = (P)$.

Théorème 4.5.3. — *Soit k un corps. Tout idéal I de $k[X]$ est principal. Plus précisément, si I est un idéal non nul de $k[X]$, il existe un unique polynôme unitaire $P \in I$ tel que $I = (P)$.*

Démonstration. — Si I est l'idéal nul $\{0\}$, il est engendré par le polynôme nul 0. Donc on peut supposer $I \neq \{0\}$. Dans ce cas, l'ensemble $\{\deg(Q) \mid Q \in I - \{0\}\}$ est un sous-ensemble non-vide de $\mathbb{N}$, donc admet un plus petit élément d. Soit $P \in I - \{0\}$ tel que $\deg(P) = d$; quitte à remplacer P par $a^{-1}P$, où a est le coefficient dominant de P, on peut supposer P unitaire.

Soit F un élément arbitraire de I, d'après le théorème 4.5.1, on peut écrire $F = PQ + R$, avec $R = 0$ ou bien $\deg(R) < \deg(P) = d$. Comme I est un idéal, alors $PQ \in I$ et donc $R = F - PQ$ appartient à I. Si on avait $R \neq 0$, ce serait un élément non nul de I de degré $< d$, contredisant la minimalité de d. Donc $R = 0$ et donc $F = PQ$. Il en résulte que $I = \{PQ \mid Q \in k[X]\} = (P)$, i.e. I est principal, engendré par le polynôme unitaire P. De plus, P est unique. En effet, si P_1 est un second polynôme unitaire tel que $I = (P_1)$, alors il existe $Q, Q_1 \in k[X]$ tels que $P_1 = PQ$ et $P = P_1Q_1$. Il en résulte que Q et Q_1 sont de degré zéro, donc des éléments de k, et comme P et P_1 sont unitaires, l'égalité $P_1 = PQ$ entraîne $Q = 1$ d'où $P_1 = P$. □

***Théorème 4.5.4* (Théorème de Bézout).** — *Soient $P_1, \dots, P_r \in \mathbb{C}[X]$ des polynômes non nuls, sans racine commune. Alors il existe $S_1, \dots, S_r \in \mathbb{C}[X]$ tels que $S_1P_1 + \cdots + S_rP_r = 1$.*

Démonstration. — Soit I l'idéal de $\mathbb{C}[X]$ engendré par $P_1, \dots, P_r$. D'une part, c'est l'ensemble des sommes $S_1P_1 + \cdots + S_rP_r = 1$, avec $S_1, \dots S_r \in \mathbb{C}[X]$. D'autre part, on sait que c'est un idéal principal non nul, engendré par un certain polynôme $D \neq 0$.

D'une part, comme $D \in I = (P_1, \dots, P_r)$, il existe $T_1, \dots, T_r \in \mathbb{C}[X]$ tels que $T_1P_1 + \cdots + T_rP_r = D$. D'autre part, comme $P_i \in I = (D)$, alors chaque P_i égale DQ_i, pour un certain $Q_i \in k[X]$. Or, $\mathbb{C}$ est algébriquement clos (cf. paragraphe suivant), donc si D était non constant il aurait une racine $\alpha \in \mathbb{C}$, et alors α serait une racine commune à tous les P_i (puisque $P_i = DQ_i$), contredisant l'hypothèse. Donc D est un polynôme constant non nul, i.e. un élément $z \in \mathbb{C}^*$. Alors remplaçant T_i par $S_i = z^{-1}T_i$, on obtient l'égalité voulue $S_1P_1 + \cdots + S_rP_r = 1$. $\square$

***Définition 4.5.5* (Polynôme minimal d'un endomorphisme).** — Soient k un sous-corps de $\mathbb{C}$ (par exemple $k = \mathbb{R}$), V un k-espace vectoriel de dimension finie et $u \in \mathrm{End}_k(V)$. Alors $I_u = \{Q \in k[X] \mid Q(u) = 0\}$ est un idéal de $k[X]$, non nul puisque $P_u(X) \in I_u$ d'après le théorème de Cayley-Hamilton. Donc, d'après le théorème 4.5.3, il existe un unique polynôme unitaire M_u tel que $I_u = (M_u)$, c.-à-d., M_u annule u et tout polynôme Q annulant u est un multiple de M_u. On dit que M_u est le *polynôme minimal* de u.

Proposition 4.5.6. — *Soit V un $\mathbb{C}$-espace vectoriel de dimension finie et soit $u \in \mathrm{End}_{\mathbb{C}}(V)$. Alors u est diagonalisable $\Leftrightarrow$ M_u a des racines simples.*

Démonstration. — $\Rightarrow$ Si u est diagonalisable, alors V est la somme directe des espaces propres $V_i = V_{\lambda_i}$, pour $i = 1, \dots, r$. Comme $u - \lambda_i \,\mathrm{id}_V$ est nul sur V_i, alors $\prod_{i=1}^r (u - \lambda_i \,\mathrm{id}_V)$ est nul, donc M_u divise le polynôme $P = \prod_{i=1}^r (X - \lambda_i \,\mathrm{id}_V)$ qui n'a que des racines simples, donc il en est de même de M_u. (En fait, on a $M_u = P$ car chaque produit partiel $\prod_{j \neq i} (u - \lambda_j \,\mathrm{id}_V)$ est non nul sur V_i.)

$\Leftarrow$ Si M_u a des racines simples, alors u est diagonalisable d'après la proposition 4.3.10 (qui utilise le théorème de Bézout). $\square$

4.6. Appendice (†) : $\mathbb{C}$ est algébriquement clos

Théorème 4.6.1. — *$\mathbb{C}$ est algébriquement clos, c.-à-d., tout polynôme $P \in \mathbb{C}[X]$ de degré $n \geq 1$ se factorise en produit de facteurs de degré 1, i.e. $P = a(X - \lambda_1) \dots (X - \lambda_n)$, où a est le coefficient dominant de P.*

Démonstration. — Remarquons d'abord qu'il suffit de montrer l'assertion suivante:

$$(\star) \qquad \text{Tout } P \in \mathbb{C}[X] \text{ non constant admet une racine dans } \mathbb{C}.$$

En effet, supposons $(\star)$ établie et montrons par récurrence sur n que tout $P \in \mathbb{C}[X]$ de degré $n \geq 1$ se factorise comme indiqué dans le théorème. C'est évident pour $n = 1$, donc on peut supposer $n \geq 2$ et le résultat établi pour $n - 1$. Soit P de degré n et de coefficient dominant a. D'après $(\star)$, P a au moins une racine λ_1 dans $\mathbb{C}$. Faisant la division euclidienne de P par $X - \lambda_1$, on peut écrire

$$P = (X - \lambda_1)P_1 + R, \qquad \text{avec} \qquad R = 0 \text{ ou bien } \deg(R) < 1.$$

Donc $R = 0$ ou bien R est une constante $c \neq 0$. Or, évaluant l'égalité ci-dessus en $X = \lambda_1$, on trouve $R(\lambda_1) = P(\lambda_1) = 0$, donc nécessairement $R = 0$. Donc $P = (X - \lambda_1)P_1$, avec P_1 non nul, de degré $n - 1$ et de coefficient dominant a. Par hypothèse de récurrence, P_1 se factorise en $P_1 = a(X - \lambda_2) \dots (X - \lambda_n)$, et donc $P = (X - \lambda_1)P_1$ égale $a(X - \lambda_1) \dots (X - \lambda_n)$. Ceci montre que le théorème découle de l'assertion $(\star)$.

Démontrons maintenant l'assertion $(\star)$. Soit $P \in \mathbb{C}[X]$ un polynôme de degré $n \geq 1$. Sans perte de généralité, on peut supposer P *unitaire*, i.e. de coefficient dominant égal à 1. Écrivons

$$P = X^n + a_1 X^{n-1} + \cdots + a_n.$$

Raisonnons par l'absurde et supposons que P ne s'annule pas sur $\mathbb{C}$. Alors, en particulier, $a_n \neq 0$. Notons $|\cdot|$ la norme usuelle sur $\mathbb{C}$, c.-à-d., si $z = x + iy$ alors $|z| = \sqrt{z\overline{z}} = \sqrt{x^2 + y^2}$.

Comme $\lim_{|z| \to +\infty} |P(z)| = +\infty$, il existe $R > 0$ tel que

$$(1) \qquad |z| \geq R \Longrightarrow |P(z)| \geq |a_n|.$$

Explicitement, on peut prendre $R = R_0 = \max\{1, 2na\}$, où $a = \max_{i=1}^n |a_i|$. En effet, pour $|z| \geq R_0$ et $d = 1, \dots, n$, on a $|z^d| \geq |z| \geq 2na$ d'où

$$\left| \sum_{d=1}^n \frac{a_d}{z^d} \right| \leq \sum_{d=1}^n \frac{|a_d|}{2na} \leq \frac{1}{2}.$$

Comme $|u+v| \geq |u| - |v|$, on obtient que, pour $|z| \geq R_0$, on a

$$|P(z)| = |z^n| \cdot \left|1 + \sum_{d=1}^{n} \frac{a_d}{z^d}\right| \geq 2na(1 - \frac{1}{2}) = na \geq n|a_n|.$$

Comme le disque D de centre 0 et de rayon R est compact, la fonction continue $f : z \mapsto |P(z)|$ y atteint son minimum r_0, et $r_0 > 0$ puisqu'on a supposé que P ne s'annule pas. Comme de plus

$$(2) \qquad \forall z \notin D, \qquad f(z) = |P(z)| \geq |a_n| = |P(0)| \geq r_0$$

alors r_0 est le minimum de f sur $\mathbb{C}$ tout entier.

Soit $z_0 \in D$ tel que $f(z_0) = r_0$. En remplaçant z par $z + z_0$ et $P(z)$ par $Q(z) := P(z_0)^{-1}P(z+z_0)$, on se ramène au cas où $z_0 = 0$ et où $Q(0) = 1$ est le minimum de $g = |Q|$ sur $\mathbb{C}$.

Observons que Q est, comme P, de degré n. Notons k l'ordre d'annulation en 0 de $Q - 1$. On peut alors écrire

$$Q(X) = 1 + b_k X^k + \cdots + b_n X^n.$$

avec b_k et b_n tous deux $\neq 0$. Écrivons $b_k = re^{i\theta}$, avec $r > 0$ et $\theta \in [0, 2\pi[$ et posons $z_\varepsilon = \varepsilon e^{i(\pi-\theta)/k}$, pour tout $\varepsilon \in \mathbb{R}_+^*$. Comme $z_\varepsilon^k = \varepsilon^k\, e^{i(\pi-\theta)} = -\varepsilon^k\, e^{-i\theta}$, alors

$$Q(z_\varepsilon) = 1 - r\varepsilon^k + \varepsilon^k h(\varepsilon), \qquad \text{où } h(\varepsilon) = \sum_{j=1}^{n-k} b_{k+j} z_\varepsilon^j.$$

Comme $\lim_{\varepsilon\to 0} \varepsilon^k = 0$ et $\lim_{\varepsilon\to 0} h(\varepsilon) = 0$, il existe $\varepsilon_0 \in]0,1[$ tel que

$$\forall \varepsilon \leq \varepsilon_0, \qquad r\varepsilon^k < 1 \quad \text{et} \quad |h(\varepsilon)| \leq \frac{r}{2}.$$

On a alors

$$|Q(z_{\varepsilon_0})| = |1 - r\varepsilon_0^k + \varepsilon_0^k h(\varepsilon_0)| \leq |1 - r\varepsilon_0^k| + \frac{r}{2}\varepsilon_0^k = 1 - r\varepsilon_0^k + \frac{r}{2}\varepsilon_0^k = 1 - \frac{r}{2}\varepsilon_0^k < 1.$$

Ceci contredit l'hypothèse que $1 = Q(0)$ était le minimum de $g = |Q|$ sur $\mathbb{C}$. Cette contradiction montre que l'hypothèse que P ne s'annule pas sur $\mathbb{C}$ est impossible. Ceci achève la démonstration du théorème 4.6.1. □

CHAPITRE 5

DÉCOMPOSITION DE JORDAN, DÉCOMPOSITION DE DUNFORD, EXPONENTIELLES DE MATRICES

Résumé: Dans les sections 1 et 2 de ce chapitre, on raffine la décomposition en espaces caractéristiques en étudiant la suite des noyaux et en introduisant la forme normale de Jordan (ceci donne lieu à une jolie interprétation "graphique" en termes de *partitions*). Puis, dans les sections 3 et 4, on se place sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et on introduit les exponentielles de matrices et leur utilisation pour l'étude des équations différentielles linéaires (à coefficients constants).

5.1. Endomorphismes nilpotents, partitions et formes normales de Jordan

Revenons pour un moment à un corps k arbitraire.

Définition 5.1.1 **(Endomorphismes nilpotents).** — Soit V un k-espace vectoriel. On dit que $u \in \mathrm{End}_k(V)$ est *nilpotent* s'il existe un entier $r \geq 1$ tel que $u^r = 0$. Dans ce cas, le plus petit entier r ayant cette propriété s'appelle l'*indice de nilpotence* de u. Par exemple, si $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ alors $A^2 = 0$ donc A est d'indice de nilpotence 2; si $B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, on a $B^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ et $B^3 = 0$, donc B est d'indice de nilpotence 3.

Un exemple très important d'endomorphismes nilpotents est fourni par les matrices triangulaires strictes, i.e. les matrices triangulaires avec des 0 sur la diagonale:

Proposition 5.1.2 **(Matrices triangulaires strictes).** — *Soit $A \in M_n(k)$ une matrice triangulaire stricte. Alors $A^n = 0$, donc A est nilpotente.*

Démonstration. — Comme A est triangulaire stricte, on a $P_A(X) = (-1)^n X^n$, d'où aussitôt le résultat si $k = \mathbb{C}$, d'après Cayley-Hamilton. Pour un corps arbitraire k, reprenons la démonstration du théorème de Cayley-Hamilton. Notons $\mathscr{B} = (e_1, \dots, e_n)$ la base canonique de k^n et F_i le sous-espace de k^n engendré par $e_1, \dots, e_i$. Si A est triangulaire supérieure stricte on a, pour $i = 1, \dots, n$:

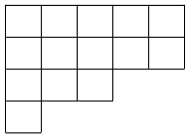
$$A(F_i) \subseteq F_{i-1},$$

avec la convention $F_0 = \{0\}$. Comme $F_n = k^n$, on déduit des inclusions ci-dessus que $A(k^n) \subseteq F_{n-1}$, puis $A^2(k^n) \subseteq F_{n-2}$, etc., d'où finalement $A^n(k^n) \subseteq F_0 = \{0\}$, donc $A^n = 0$. Enfin, si A est triangulaire inférieure stricte, en appliquant ce qui précède à $B = {}^tA$, on obtient que $B^n = {}^t(A^n)$ est nulle, d'où aussi $A^n = 0$. $\square$

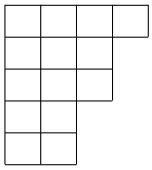
Définition 5.1.3 **(Partitions d'un entier).** — Soit n un entier ≥ 1. On appelle *partition* de n une suite décroissante $\mathbf{p} = (p_1, \dots, p_r)$ d'entiers ≥ 1 tels que $p_1 + \cdots + p_r = n$. Les p_i s'appellent les *parts* de la partition, r le nombre de parts, et p_1 la plus grande part. On peut représenter une telle partition par un diagramme formé de "boîtes", chaque boîte étant un carré de côté 1, la première ligne contenant p_1 boîtes,

[(0)] version du 15/06/2015

la seconde p_2 boîtes, etc., toutes les lignes étant alignées à gauche; par exemple, la partition $\mathbf{p} = (5,5,3,1)$ de $5+5+3+1 = 14$ est représentée par le diagramme :



On voit alors que l'ensemble des partitions de n est muni d'une involution $\mathbf{p} \mapsto \widetilde{\mathbf{p}}$, où le diagramme de $\widetilde{\mathbf{p}}$ est obtenu en "transposant" celui de $\mathbf{p}$, i.e. les lignes de $\widetilde{\mathbf{p}}$ sont les colonnes de $\mathbf{p}$, et vice-versa (on a donc $\widetilde{\widetilde{\mathbf{p}}} = \mathbf{p}$). Ainsi, dans l'exemple précédent, le diagramme de $\widetilde{\mathbf{p}}$ est :



i.e. $\widetilde{\mathbf{p}} = (4,3,3,2,2)$.

***Proposition 5.1.4* (La suite des noyaux).** — *Soient V un k-espace vectoriel de dimension n, u un endomorphisme nilpotent de V, et d l'indice de nilpotence de u.*

(1) *On a une suite croissante:*

$$(*) \qquad \{0\} = \operatorname{Ker}(\mathrm{id}_V) \subseteq \operatorname{Ker}(u) \subseteq \operatorname{Ker}(u^2) \subseteq \cdots \subseteq \operatorname{Ker}(u^d) = V.$$

(2) *Pour tout $i = 1, \dots, d$, posons* $\boxed{K_i = \dim \operatorname{Ker}(u^i)}$ *et* $\boxed{q_i = K_i - K_{i-1}.}$ *(On a $u^0 = \mathrm{id}_V$ et $K_0 = 0$.) Alors on a* $\boxed{q_{i-1} \geq q_i}$ *pour $i = 2, \dots, d$.*

(3) *La suite $(*)$ ci-dessus est strictement croissante, i.e.* $\operatorname{Ker}(u^{i-1}) \neq \operatorname{Ker}(u^i)$ *pour tout $i = 1, \dots, d$. En particulier, on a $d \leq n$.*

(4) *La suite $\mathbf{q} = (q_1, \dots, q_d)$ est une* partition *de n, et la suite des dimensions des noyaux, $\mathbf{K} = (K_1, K_2, \dots, K_d)$, est la suite $\sigma(\mathbf{q})$ des sommes partielles de $\mathbf{q}$.*

Démonstration. — L'assertion (1) est immédiate: si $x \in V$ vérifie $u^i(x) = 0$ alors *a fortiori* $u^{i+1}(x) = u(u^i(x)) = 0$, d'où $\operatorname{Ker}(u^i) \subseteq \operatorname{Ker}(u^{i+1})$ pour $i = 0, \dots, d-1$. De plus, $\operatorname{Ker}(u^d) = V$ puisque $u^d = 0$.

Pour montrer (2), fixons un indice $i \geq 2$ et posons $q = q_i$. Soit $\mathscr{B} = (e_1, \dots, e_q)$ une base d'un supplémentaire de $\operatorname{Ker}(u^{i-1})$ dans $\operatorname{Ker}(u^i)$. Supposons qu'on ait une égalité

$$(\dagger) \qquad x + t_1 u(e_1) + \cdots + t_q u(e_q) = 0, \qquad \text{avec} \quad x \in \operatorname{Ker}(u^{i-2}), \quad t_1, \dots, t_q \in k$$

alors, appliquant u^{i-2} à cette égalité, on obtient $u^{i-1}(t_1 e_1 + \cdots + t_q e_q) = 0$, d'où $t_1 e_1 + \cdots + t_q e_q \in \operatorname{Ker}(u^{i-1})$. Comme $\mathscr{B}$ est une base d'un supplémentaire de $\operatorname{Ker}(u^{i-1})$ dans $\operatorname{Ker}(u^i)$, ceci entraîne $t_1 = 0 = \cdots = t_q$, puis reportant ceci dans $(\dagger)$ on obtient aussi que $x = 0$. Ceci montre que les sous-espaces $\operatorname{Ker}(u^{i-2})$ et $ku(e_1)$, …, $ku(e_q)$ de $\operatorname{Ker}(u^{i-1})$ sont en somme directe, d'où

$$K_{i-2} + q = \dim\Big(\operatorname{Ker}(u^{i-2}) \oplus \underbrace{ku(e_1) \oplus \cdots \oplus ku(e_q)}_{=\operatorname{Vect}(u(\mathscr{B}))}\Big) \leq K_{i-1}$$

donc $q_i = q \leq K_{i-1} - K_{i-2} = q_{i-1}$, ce qui prouve l'assertion (2). Signalons aussi le point suivant, dont on aura besoin plus loin : si $\mathscr{B}' = (f_1, \dots, f_t)$ est une base d'un supplémentaire de $\operatorname{Ker}(u^{i-2}) \oplus \operatorname{Vect}(u(\mathscr{B}))$ dans $\operatorname{Ker}(u^{i-1})$, alors $t = K_{i-1} - (K_{i-2} + q_i) = q_{i-1} - q_i$, et $u(\mathscr{B}) \cup \mathscr{B}'$ est une base d'un supplémentaire de $\operatorname{Ker}(u^{i-2})$ dans $\operatorname{Ker}(u^{i-1})$. On a donc obtenu le résultat suivant:

$$(\star) \quad \left\{ \begin{array}{l} \text{si } \mathscr{B}_i \text{ est une base d'un supplémentaire de } \operatorname{Ker}(u^{i-1}) \text{ dans } \operatorname{Ker}(u^i), \text{ alors la famille } u(\mathscr{B}_i) \text{ est} \\ \text{libre et peut se compléter en une base d'un supplémentaire de } \operatorname{Ker}(u^{i-2}) \text{ dans } \operatorname{Ker}(u^{i-1}). \end{array} \right.$$

L'assertion (3) en découle. En effet, si pour un certain $i \leq d$ on avait $\operatorname{Ker}(u^{i-1}) = \operatorname{Ker}(u^i)$, i.e. $q_i = 0$, on aurait $q_{i+1} = 0 = \cdots = q_d$, d'où $\operatorname{Ker}(u^{i-1}) = \operatorname{Ker}(u^d) = V$, donc $u^{i-1} = 0$, contredisant le fait que l'indice de nilpotence est d. Donc la suite $(*)$ est strictement croissante. Comme la dimension croît d'au moins 1 à chaque cran, on a donc $\dim \operatorname{Ker}(u^i) \geq i$ pour $i = 1, \dots, d$, et comme d'autre part $\dim V = n$, on conclut que $d \leq n$.

Prouvons l'assertion (4). Par définition des q_i, on a:

$$K_1 = q_1, \quad K_2 = q_2 + K_1 = q_2 + q_1, \quad \dots \quad K_i = q_i + K_{i-1} = q_i + \cdots + q_1, \quad \dots$$
$$K_d = q_d + K_{d-1} = q_d + \cdots + q_1.$$

D'autre part, d'après (2), on a $q_i \geq q_{i+1}$. Il en résulte que $\mathbf{q} = (q_1, \dots, q_d)$ est une partition de $K_d = n$, et que la suite $\mathbf{K} = (K_1, K_2, \dots, K_d)$ est la suite $\sigma(\mathbf{q})$ des sommes partielles de $\mathbf{q}$. □

Définitions 5.1.5 **(Blocs de Jordan nilpotents et matrices de Jordan nilpotentes)**
1) Pour tout entier $n \geq 1$, on appelle "bloc de Jordan nilpotent" de taille n, et l'on note J_n, la matrice carrée de taille n ayant des 1 sur la diagonale juste au-dessus de la diagonale principale et des 0 partout ailleurs. Ainsi, J_1 est la matrice nulle (0), et

$$J_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad J_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \text{etc.}$$

Notant $\mathscr{B} = (e_1, \dots, e_n)$ la base canonique de k^n, on a donc $J_n e_i = e_{i-1}$ pour $i = n, \dots, 2$ et $J_n e_1 = 0$. Il en résulte que $\mathscr{B}$ égale $(u^{n-1}(e_n), \dots, u(e_n), e_n)$ et que $\mathrm{Ker}(J_n)$ est la droite $ke_1 = ku^{n-1}(e_n)$.

2) On appelle "matrice de Jordan nilpotente" de taille n toute matrice $A \in M_n(k)$ diagonale par blocs, dont les blocs diagonaux $A_1, \dots, A_d$ sont des blocs de Jordan nilpotents $J_{p_1}, \dots J_{p_d}$, rangés par ordre de taille décroissante, i.e. toute matrice $A \in M_n(k)$ de la forme suivante:

$$A = \left(\begin{array}{c|c|c|c} J_{p_1} & 0 & \cdots & 0 \\ \hline 0 & J_{p_2} & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & 0 \\ \hline 0 & \cdots & 0 & J_{p_d} \end{array}\right) \qquad \text{où} \quad p_1 \geq \cdots \geq p_d \geq 1.$$

Comme la somme des tailles des blocs diagonaux égale n, on a $p_1 + \cdots + p_d = n$, i.e. $\mathbf{p} = (p_1, \dots, p_d)$ est une partition de n, et l'on notera $J_{\mathbf{p}}$ la matrice ci-dessus. On voit ainsi que les matrices de Jordan nilpotentes de taille n sont en bijection avec l'ensemble $\mathscr{P}(n)$ des partitions de n.

Théorème et définition 5.1.6 **(Forme normale de Jordan d'un endomorphisme nilpotent)**
Soient V un k-espace vectoriel de dimension n, et u un endomorphisme nilpotent de V, d'indice de nilpotence d.

(1) *Il existe une base $\mathscr{B}$ de V telle que* $\mathrm{Mat}_{\mathscr{B}}(u)$ *soit une matrice de Jordan nilpotente $J_{\mathbf{p}}$, pour une certaine partition $\mathbf{p}$ de n.*

(2) *De plus, $\mathbf{p}$ est uniquement déterminée par u: en effet, $\widetilde{\mathbf{p}}$ est la partition $\mathbf{q}$ associée à la suite des noyaux* $\mathrm{Ker}(u) \subseteq \cdots \subseteq \mathrm{Ker}(u^d)$. *En particulier, d est la plus grande part p_1 de $\mathbf{p}$, et le nombre de parts de $\mathbf{p}$ est $q_1 = \dim \mathrm{Ker}(u)$.*

(3) *En termes matriciels: toute matrice nilpotente $A \in M_n(k)$ est semblable à une unique matrice de Jordan nilpotente de taille n, et donc l'ensemble des classes de similitude de matrices nilpotentes $A \in M_n(k)$ est en bijection avec l'ensemble $\mathscr{P}(n)$ des partitions de n.*

(4) *On dit que la matrice $J_{\mathbf{p}}$ obtenue est la* **forme normale de Jordan** *ou la* **réduction de Jordan** *de u (ou de A).*

Démonstration. — Soit $\mathbf{q} = (q_1, \dots, q_d)$ la partition $\mathbf{q}$ associée à la suite des noyaux de u, i.e. $q_i = K_i - K_{i-1}$, où $K_i = \dim \mathrm{Ker}(u^i)$, pour $i = 1, \dots, d$ (et $K_0 = 0$). Montrons l'existence d'une base $\mathscr{B}$ de V telle que $\mathrm{Mat}_{\mathscr{B}} = J_{\widetilde{\mathbf{q}}}$. Rappelons le résultat suivant, qu'on a obtenu dans la démonstration du point (2) de 5.1.4:

$$(\star) \quad \left\{ \begin{array}{l} \text{si } \mathscr{C}_i \text{ est une base d'un supplémentaire de } \mathrm{Ker}(u^{i-1}) \text{ dans } \mathrm{Ker}(u^i), \text{ alors la famille } u(\mathscr{C}_i) \text{ est} \\ \text{libre et peut se compléter en une base d'un supplémentaire de } \mathrm{Ker}(u^{i-2}) \text{ dans } \mathrm{Ker}(u^{i-1}) \end{array} \right.$$

Soit $\mathscr{B}_d = (e_1, \dots, e_{q_d})$ une base d'un supplémentaire de $\mathrm{Ker}(u^{d-1})$ dans $V = \mathrm{Ker}(u^d)$. D'après $(\star)$, la famille $u(\mathscr{B}_d)$ est libre et se complète, par ajout d'une famille libre $\mathscr{B}_{d-1} = (e_{q_d+1}, \dots, e_{q_{d-1}})$, en une base $\mathscr{L}_{d-1} = u(\mathscr{B}_d) \cup \mathscr{B}_{d-1}$ d'un supplémentaire de $\mathrm{Ker}(u^{d-2})$ dans $\mathrm{Ker}(u^{d-1})$. (Et par conséquent, $\mathscr{B}_d \cup \mathscr{L}_{d-1}$ est une base d'un supplémentaire de $\mathrm{Ker}(u^{d-2})$ dans V.)

Pour aider le lecteur, indiquons que la démonstration peut se "visualiser" comme suit: on dessine le diagramme de la partition $\mathbf{q}$, disons pour $\mathbf{q} = (5, 5, 3, 1)$, et l'on écrit les vecteurs $e_1, \dots, e_{q_d}$ dans les dernières cases des colonnes 1 à q_d (i.e. dans la ligne d), puis on écrit au-dessus les vecteurs $u(e_1), \dots, u(e_{q_d})$ et l'on complète la ligne $d-1$ en écrivant les vecteurs $e_{q_d+1}, \dots, e_{q_{d-1}}$, dans les dernières cases des colonnes

$q_d + 1$ à q_{d-1}:

base d'un supplém. de $\mathrm{Ker}(u^{d-2})$ dans $\mathrm{Ker}(u^{d-1})$:	$u(e_1)$	e_2	e_3		
base d'un supplém. de $\mathrm{Ker}(u^{d-1})$ dans $\mathrm{Ker}(u^{d})$:	e_1				

D'après ce qui précède, les vecteurs qu'on a écrits forment une base d'un supplémentaire de $\mathrm{Ker}(u^{d-2})$ dans V.

Puis, d'après $(\star)$ à nouveau, il existe une famille libre $\mathscr{B}_{d-2} = (e_{q_{d-1}+1}, \dots, e_{q_{d-2}})$ dans $\mathrm{Ker}(u^{d-2})$ telle que

$$\mathscr{L}_{d-2} = u(\mathscr{L}_{d-1}) \cup \mathscr{B}_{d-2} = u^2(\mathscr{B}_d) \cup u(\mathscr{B}_{d-1}) \cup \mathscr{B}_{d-2}$$

soit une base d'un supplémentaire de $\mathrm{Ker}(u^{d-3})$ dans $\mathrm{Ker}(u^{d-2})$; alors $\mathscr{B}_d \cup \mathscr{L}_{d-1} \cup \mathscr{L}_{d-2}$ est une base d'un supplémentaire de $\mathrm{Ker}(u^{d-3})$ dans V. Dans l'exemple précédent, ceci donne:

base d'un supplém. de $\mathrm{Ker}(u)$ dans $\mathrm{Ker}(u^2)$:	$u^2(e_1)$	$u(e_2)$	$u(e_3)$	e_4	e_5
base d'un supplém. de $\mathrm{Ker}(u^2)$ dans $\mathrm{Ker}(u^3)$:	$u(e_1)$	e_2	e_3		
base d'un supplém. de $\mathrm{Ker}(u^3)$ dans $\mathrm{Ker}(u^4)$:	e_1				

Ainsi, après $p \leq d$ étapes, on a rempli les p lignes du bas du diagramme par des vecteurs qui forment une base d'un supplémentaire de $\mathrm{Ker}(u^{d-p})$ dans V (plus précisément, les vecteurs de la i-ème ligne en partant du bas forment une base d'un supplémentaire de $\mathrm{Ker}(u^{d-i})$ dans $\mathrm{Ker}(u^{d-i+1})$). Continuant ainsi, on remplit le diagramme de la partition $\mathbf{q}$ par des vecteurs qui forment une base $\mathscr{L}$ de V:

base de $\mathrm{Ker}(u)$:	$u^3(e_1)$	$u^2(e_2)$	$u^2(e_3)$	$u(e_4)$	$u(e_5)$
base d'un supplém. de $\mathrm{Ker}(u)$ dans $\mathrm{Ker}(u^2)$:	$u^2(e_1)$	$u(e_2)$	$u(e_3)$	e_4	e_5
base d'un supplém. de $\mathrm{Ker}(u^2)$ dans $\mathrm{Ker}(u^3)$:	$u(e_1)$	e_2	e_3		
base d'un supplém. de $\mathrm{Ker}(u^3)$ dans $\mathrm{Ker}(u^4)$:	e_1				

(Noter que les vecteurs $e_1, \dots, e_{q_1}$ qui ont été introduits à chaque étape du processus sont situés dans la dernière case de chaque colonne.)

Lisons maintenant le diagramme colonne par colonne, de haut en bas: la 1ère colonne est formée des vecteurs $v_1 = u^{d-1}(e_1)$, $v_2 = u^{d-2}(e_1)$, ..., $v_d = e_1$; ils forment une base $\mathscr{C}_1$ d'un sous-espace E_1 stable par u, dans laquelle la restriction u_{E_1} de u à E_1 a pour matrice le bloc de Jordan nilpotent J_d. De même, pour chaque $j = 1, \dots, q_d$, la j-ème colonne est formée des vecteurs $u^{d-i}(e_j)$, pour $i = 1, \dots, d$, qui forment une base $\mathscr{C}_j$ d'un sous-espace E_j de dimension d, telle que $\mathrm{Mat}_{\mathscr{C}_j}(u_{E_j}) = J_d$. Ensuite, pour j variant de $q_d + 1$ à q_{d-1}, la j-ème colonne est formée des vecteurs $u^{d-1-i}(e_j)$, pour $i = 1, \dots, d-1$, qui forment une base $\mathscr{C}_j$ d'un sous-espace E_j stable par u, de dimension $d-1$, telle que $\mathrm{Mat}_{\mathscr{C}_j}(u_{E_j}) = J_{d-1}$, etc.

Donc chaque *colonne* C_j du diagramme correspond à une base $\mathscr{C}_j$ d'un sous-espace E_j stable par u, de dimension la *hauteur* h_j de la colonne C_j, telle que $\mathrm{Mat}_{\mathscr{C}_j}(u_{E_j})$ soit le bloc de Jordan nilpotent J_{h_j}. On voit donc que les blocs de Jordan nilpotents qui apparaissent correspondent aux colonnes du diagramme de $\mathbf{q}$, c.-à-d., aux lignes du diagramme de $\widetilde{\mathbf{q}}$. On obtient donc que dans la base $\mathscr{C} = \mathscr{C}_1 \cup \cdots \cup \mathscr{C}_{q_1}$ (ses éléments sont les mêmes que ceux de $\mathscr{L}$, mais dans un ordre différent), la matrice de u est la matrice de Jordan nilpotente $J_{\widetilde{\mathbf{q}}}$. Par exemple, pour le diagramme précédent, on obtient la matrice

$$J_{(4,3,3,2,2)} = \left(\begin{array}{c|c|c|c|c} J_4 & 0 & 0 & 0 & 0 \\ \hline 0 & J_3 & 0 & 0 & 0 \\ \hline 0 & 0 & J_3 & 0 & 0 \\ \hline 0 & 0 & 0 & J_2 & 0 \\ \hline 0 & 0 & 0 & 0 & J_2 \end{array}\right)$$

Il reste à montrer l'assertion d'unicité: supposons que dans une certaine base $\mathscr{B} = (v_1, \dots, v_n)$ de V, u ait pour matrice la matrice de Jordan nilpotente $J_{\mathbf{p}}$, pour une certaine partition $\mathbf{p} = (p_1, \dots, p_s)$. Considérons alors le diagramme de $\mathbf{p}$ et écrivons les vecteurs $v_1, \dots, v_n$ de la base $\mathscr{B}$ dans ce diagramme, de gauche à droite dans chaque ligne, en commençant par la ligne du haut. Alors, par définition de la matrice $J_{\mathbf{p}}$, si l'on note e_j le vecteur qui est dans la dernière case de la j-ème ligne en partant du haut, alors les vecteurs de cette ligne, lus de gauche à droite, sont:

$$(*_j) \qquad u^{p_j-1}(e_j), \quad u^{p_j-2}(e_j), \quad \dots \quad u(e_j), \quad e_j.$$

Donc, en transposant ce diagramme, on obtient le diagramme de $\widetilde{\mathbf{p}}$, contenant dans sa j-ème colonne les vecteurs $(*_j)$ précédents :

$u^{p_1-1}(e_1)$	$u^{p_2-1}(e_2)$	$\cdots$	$u^{p_s-1}(e_s)$
$u^{p_1-2}(e_1)$	$u^{p_2-2}(e_2)$	$\cdots$	$u^{p_s-2}(e_s)$
$\vdots$	$\vdots$	$\dots$	$\vdots$
$\vdots$	$\vdots$	$\dots$	e_s
$\vdots$	$\vdots$	$\dots$	
$u(e_1)$	e_2		
e_1			

Comme ces vecteurs forment une base de V, on voit facilement que les vecteurs de la ligne 1 du haut forment une base de $\operatorname{Ker}(u)$, ceux des lignes 1 et 2 du haut forment une base de $\operatorname{Ker}(u^2)$, etc. On obtient donc que $\widetilde{\mathbf{p}}$ est la partition associée à la suite des noyaux de u, d'où $\widetilde{\mathbf{p}} = \mathbf{q}$ et donc $\mathbf{p} = \widetilde{\mathbf{q}}$. Ceci achève la démonstration des assertions (1) et (2) du théorème 5.1.6. Enfin, l'assertion (3) est une conséquence immédiate des assertions (1) et (2). Le théorème est démontré. $\square$

Remarque 5.1.7. — Dans la démonstration précédente, on a utilisé de façon répétée *l'existence* d'une base d'un supplémentaire de $\operatorname{Ker}(u^{i-1})$ dans $\operatorname{Ker}(u^i)$. Pour construire *explicitement* une base de Jordan, on peut utiliser la méthode de réduction des colonnes pour construire successivement une base de $\operatorname{Ker}(u)$, puis d'un supplémentaire de $\operatorname{Ker}(u)$ dans $\operatorname{Ker}(u^2)$, puis d'un supplémentaire de $\operatorname{Ker}(u^2)$ dans $\operatorname{Ker}(u^3)$, etc. Voir le paragraphe 5.1.14 plus bas.

Exercice 5.1.7.1. — Quel est le nombre de classes de similitude de matrices nilpotentes dans $M_6(k)$? Donner un représentant de chaque classe.

Revenons maintenant au cas où $k = \mathbb{C}$ et combinons les résultats obtenus dans les théorèmes 4.3.8 et 5.1.6. Commençons par la définition suivante:

***Définitions 5.1.8* (Blocs de Jordan et matrices de Jordan)**. — (1) Pour $\lambda \in \mathbb{C}$ et $n \in \mathbb{N}^*$, on appelle “bloc de Jordan” de taille n associé à λ, et l'on note $J_n(\lambda)$, la matrice carrée de taille n égale à $\lambda I_n + J_n$. (Donc $J_n(0) = J_n$.) Ainsi $J_1(\lambda) = (\lambda)$, et

$$J_2(\lambda) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad J_3(\lambda) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}, \quad \text{etc.}$$

(2) Pour $\lambda \in \mathbb{C}$, $n \in \mathbb{N}^*$, et $\mathbf{p}$ une partition de n, on note $J_{\mathbf{p}}(\lambda) = \lambda I_n + J_{\mathbf{p}}$ et on dit que c'est une “matrice de Jordan” de taille n associé à λ.

(3) Enfin, étant donnés des r-uplets $(\lambda_1, \dots, \lambda_r) \in \mathbb{C}^r$ et $(m_1, \dots, m_r) \in (\mathbb{N}^*)^r$, ainsi qu'une partition $\mathbf{p}_i$ de chaque m_i, on note $J_{\mathbf{p}_1,\dots,\mathbf{p}_r}(\lambda_1, \dots, \lambda_r)$ la matrice carrée de taille $n = m_1 + \cdots + m_r$, diagonale par blocs, dont les blocs diagonaux sont $J_{\mathbf{p}_1}(\lambda_1), \dots J_{\mathbf{p}_r}(\lambda_r)$. On dira que c'est une “matrice de Jordan associée au r-uplet $\big((\lambda_1, m_1), \dots, (\lambda_r, m_r)\big)$”.

Remarque 5.1.9. — Soit $J_{\mathbf{p}_1,\dots,\mathbf{p}_r}(\lambda_1, \dots, \lambda_r)$ comme ci-dessus; pour $i = 1, \dots, r$, notons $\mathscr{C}_i$ l'ensemble des vecteurs de la base canonique de $\mathbb{C}^n$ qui correspondent au bloc diagonal $J_{\mathbf{p}_i}(\lambda_i)$, alors chaque $\mathscr{C}_i$ est une base d'un sous-espace W_i de V stable par u, telle que $\operatorname{Mat}_{\mathscr{C}_i}(u_{W_i}) = J_{\mathbf{p}_i}(\lambda_i)$, et $J_{\mathbf{p}_1,\dots,\mathbf{p}_r}(\lambda_1, \dots, \lambda_r)$ est la matrice de u dans la base $\mathscr{C} = \mathscr{C}_1 \cup \cdots \cup \mathscr{C}_r$.

Remarquons tout de suite que si l'on permute les $\mathscr{C}_i$ entre elles, i.e. si σ est une permutation de $\{1, \dots, r\}$ [(1)] et qu'on considère la base $\mathscr{C}_\sigma = \mathscr{C}_{\sigma(1)} \cup \cdots \cup \mathscr{C}_{\sigma(r)}$, alors la matrice de u dans la base $\mathscr{C}_\sigma$ est $J_{\mathbf{p}_{\sigma(1)},\dots,\mathbf{p}_{\sigma(r)}}(\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(r)})$. Pour toute permutation σ de $\{1, \dots, r\}$, les matrices $J_{\mathbf{p}_1,\dots,\mathbf{p}_r}(\lambda_1, \dots, \lambda_r)$ et $J_{\mathbf{p}_{\sigma(1)},\dots,\mathbf{p}_{\sigma(r)}}(\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(r)})$ sont donc semblables.

***Théorème 5.1.10* (Forme normale de Jordan d'un endomorphisme de $\mathbb{C}^n$)**

Soient V un $\mathbb{C}$-espace vectoriel de dimension n, $u \in \operatorname{End}_{\mathbb{C}}(V)$, $\lambda_1, \dots, \lambda_r$ les valeurs propres, deux à deux distinctes, de u et pour tout i, soient u_i la restriction de u à l'espace caractéristique $V_{(\lambda_i)}$ et $\mathbf{q}_i$ la partition de $m_i = \dim V_{(\lambda_i)}$ associée à la suite des noyaux de u_i.

[(1)] c.-à-d., une bijection de l'ensemble $\{1, \dots, r\}$ dans lui-même

(1) *Il existe une base* $\mathscr{B}$ *de* V *telle que* $\mathrm{Mat}_{\mathscr{B}}(u)$ *soit la matrice de Jordan* $J_{\mathbf{p}_1,\dots,\mathbf{p}_r}(\lambda_1,\dots,\lambda_r)$, *où* $\mathbf{p}_i = \widetilde{\mathbf{q}}_i$.

(2) *Le* r*-uplet* $\big((\lambda_1,\mathbf{p}_1),\dots,(\lambda_r,\mathbf{p}_r)\big)$ *est uniquement déterminé, à l'ordre près, c.-à-d., si pour une base* $\mathscr{B}'$ *on a* $\mathrm{Mat}_{\mathscr{B}'}(u) = J_{\mathbf{p}'_1,\dots,\mathbf{p}'_s}(\mu_1,\dots,\mu_s)$, *avec* $\mu_i \neq \mu_j$ *si* $i \neq j$, *alors* $s = r$ *et il existe une permutation* σ *de* $\{1,\dots,r\}$ *telle que* $\mu_i = \lambda_{\sigma(i)}$ *et* $\mathbf{p}'_i = \mathbf{p}_{\sigma(i)}$.

(3) *En termes matriciels : toute* $A \in M_n(\mathbb{C})$ *est semblable à une matrice de Jordan* $J_{\mathbf{p}_1,\dots,\mathbf{p}_r}(\lambda_1,\dots,\lambda_r)$, *qui est unique à permutation près des blocs diagonaux entre eux.*

(4) *On dit que la matrice* $J_{\mathbf{p}_1,\dots,\mathbf{p}_r}(\lambda_1,\dots,\lambda_r)$ *est la* **forme normale de Jordan** *ou la* **réduction de Jordan** *de* u *(ou de* A*).*

Démonstration. — L'existence découle immédiatement des théorèmes 4.3.8 et 5.1.6. En effet, pour tout i, notons $N_i = V_{(\lambda_i)}$. Comme chaque $u_i - \lambda_i \,\mathrm{id}_{N_i}$ est nilpotent, alors, d'après le théorème 5.1.6, il existe une base $\mathscr{B}_i$ de N_i telle que $\mathrm{Mat}_{\mathscr{B}_i}(u_i - \lambda_i \,\mathrm{id}_{N_i}) = J_{\mathbf{p}_i}$, où $\mathbf{p}_i = \widetilde{\mathbf{q}}_i$, d'où $\mathrm{Mat}_{\mathscr{B}_i}(u_i) = J_{\mathbf{p}_i} + \lambda I_{m_i} = J_{\mathbf{p}_i}(\lambda_i)$.

D'autre part, d'après le théorème 4.3.8, on a

$$(*) \qquad V = N_1 \oplus \cdots \oplus N_r \qquad \text{et} \qquad P_u(X) = (-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{m_i}.$$

Donc $\mathscr{B} = \mathscr{B}_1 \cup \cdots \cup \mathscr{B}_r$ est une base de V, et dans cette base la matrice de u est $J_{\mathbf{p}_1,\dots,\mathbf{p}_r}(\lambda_1,\dots,\lambda_r)$. Ceci prouve l'existence.

Montrons l'unicité. Supposons que pour une base $\mathscr{B}'$ on ait $\mathrm{Mat}_{\mathscr{B}'}(u) = J_{\mathbf{p}'_1,\dots,\mathbf{p}'_s}(\mu_1,\dots,\mu_s)$, avec $\mu_i \neq \mu_j$ si $i \neq j$. Comme cette matrice est triangulaire supérieure, on obtient que

$$P_u(X) = (-1)^n \prod_{i=1}^{s} (X - \mu_i)^{d_i}$$

où chaque d_i est la somme des parts de $\mathbf{p}'_i$ (i.e. $\mathbf{p}'_i$ est une partition de d_i). Comme $\mu_i \neq \mu_j$ si $i \neq j$ alors, $\mu_1,\dots,\mu_s$ sont les racines, deux à deux distinctes, de P_u et donc, comparant avec $(*)$ plus haut, on obtient déjà que $s = r$ et qu'il existe une permutation σ de $\{1,\dots,r\}$ telle que $\mu_i = \lambda_{\sigma(i)}$ et $d_i = m_{\sigma(i)}$ pour tout $i = 1,\dots,r$.

Reste à montrer que $\mathbf{p}'_i = \mathbf{p}_{\sigma(i)}$. En renumérotant les μ_i, on se ramène d'abord au cas où $\sigma = \mathrm{id}$ (i.e. $\mu_i = \lambda_i$ et $d_i = m_i$ pour tout i), et il faut montrer que $\mathbf{p}'_i = \mathbf{p}_i$. Pour tout i, notons $\mathscr{B}'_i$ l'ensemble des vecteurs de la base $\mathscr{B}'$ qui correspondent au bloc diagonal $J_{\mathbf{p}'_i}(\lambda_i)$, alors chaque $\mathscr{B}'_i$ est une base d'un sous-espace W_i de dimension m_i de V stable par u, et l'on a $\mathrm{Mat}_{\mathscr{B}'_i}(u_{W_i}) = J_{\mathbf{p}'_i}(\lambda_i)$. Comme $J_{\mathbf{p}'_i}$ est triangulaire stricte, de taille m_i, on a $0 = J_{\mathbf{p}'_i}^{m_i} = (J_{\mathbf{p}'_i}(\lambda_i) - \lambda \,\mathrm{id}_{W_i})^{m_i}$ et donc $(u_{W_i} - \lambda_i \,\mathrm{id}_{W_i})^{m_i} = 0$. Ceci montre que W_i est contenu dans $\mathrm{Ker}\big((u - \lambda_i)^{m_i}\big) = V_{(\lambda_i)} = N_i$, et comme tous deux sont de dimension m_i, on obtient que $W_i = N_i$ pour tout i.

Donc, $\mathscr{B}'_i$ et $\mathscr{B}_i$ sont deux bases du *même espace*, N_i, telles que la matrice de l'endomorphisme nilpotent $u_i - \lambda_i \,\mathrm{id}_{N_i}$ de N_i dans $\mathscr{B}'_i$ (resp. $\mathscr{B}_i$) est $J_{\mathbf{p}'_i}$ (resp. $J_{\mathbf{p}'_i}$). D'après le résultat d'unicité dans le cas nilpotent, on conclut que $\mathbf{p}'_i = \mathbf{p}_i$. Ceci achève la preuve des assertions (1) et (2) du théorème 5.1.10. Enfin, l'assertion (3) est une conséquence immédiate des assertions (1) et (2). Le théorème est démontré. □

Corollaire 5.1.11. — *Soient* $A, A' \in M_n(\mathbb{C})$. *Les conditions suivantes sont équivalentes:*

(1) A *et* A' *sont semblables (i.e. sont dans la même classe de similitude).*

(2) A *et* A' *ont même polynôme caractéristique, disons* $\prod_{i=1}^{r}(\lambda_i - X)^{m_i}$ *avec* $\lambda_i \neq \lambda_j$ *si* $i \neq j$, **et** *pour tout* $i = 1,\dots,r$, $A - \lambda_i I_n$ *et* $A' - \lambda_i I_n$ *ont la même partition associée à la suite des noyaux.*

(3) A *et* A' *ont la même forme normale de Jordan.*

Remarque 5.1.12. — **En résumé**, pour déterminer la forme normale de Jordan d'une matrice $A \in M_n(\mathbb{C})$, on calcule d'abord son polynôme caractéristique $P_A(X) = (-1)^n \prod_{i=1}^{r}(X-\lambda_i)^{m_i}$ puis, pour chaque i, on calcule la dimension $K_{i,s} = \dim \mathrm{Ker}\big((A - \lambda_i I_n)^s\big)$ (en déterminant le rang de $(A - \lambda_i I_n)^s$), pour $s = 1, 2, \dots$, en s'arrêtant au cran t tel que $K_{i,t} = m_i$. Posant $q_{i,s} = K_{i,s} - K_{i,s-1}$, on obtient la partition $\mathbf{q}_i = (q_{i,1},\dots,q_{i,t})$ associée à la suite des noyaux de $A - \lambda_i I_n$; notant $\mathbf{p}_i$ sa transposée $\widetilde{\mathbf{q}}_i$, on obtient alors la matrice de Jordan $J_{\mathbf{p}_i}(\lambda_i)$, dont le nombre de blocs de Jordan est $q_{i,1} = K_{i,1} = \dim \mathrm{Ker}(A - \lambda_i I_n)$. En particulier, si $K_{i,1} = 1$, alors $J_{\mathbf{p}_i}(\lambda_i)$ est le bloc de Jordan $J_{m_i}(\lambda_i)$.

Exercice 5.1.13. — Soient $\lambda_1,\dots,\lambda_r \in \mathbb{C}$, deux à deux distincts, $m_1,\dots,m_r$ des entiers ≥ 1, et soient $n = m_1 + \cdots + m_r$ et P le polynôme $(-1)^n \prod_{i=1}^{r}(X - \lambda_i)^{m_i}$. Alors, d'après le théorème précédent, le nombre de classes de similitude de matrices $A \in M_n(\mathbb{C})$ dont le polynôme caractéristique est égal à P, est égal à $p(m_1)\cdots p(m_r)$, où $p(m_i)$ désigne le nombre de partitions de m_i.

(1) Quel est le nombre de classes de similitudes de matrices dans $M_7(\mathbb{C})$ dont le polynôme caractéristique est $-X^7+3X^6-3X^5+X^4$?

(2) À quelle condition sur $P=(-1)^n\prod_{i=1}^r(X-\lambda_i)^{m_i}\in\mathbb{C}[X]$, l'ensemble $\mathscr{C}(P)=\{A\in M_n(\mathbb{C})\mid P_A(X)=P\}$ est-il formé d'une seule classe de similitude?

5.1.14. Opérations sur les colonnes et bases de Jordan. — Étant donné $A\in M_n(\mathbb{C})$ ou, plus généralement $A\in M_n(k)$ avec $P_A(X)$ scindé, la méthode de réduction des colonnes s'applique très bien au calcul explicite d'une "base de Jordan" de l'endomorphisme u défini par A, i.e. d'une base $\mathscr{C}$ dans laquelle la matrice de u est la forme normale de Jordan J_A de A. Illustrons ceci par deux exemples.

Exemple 1. Soit $A=\begin{pmatrix}0&1&-1&0\\-3&-1&3&1\\-2&1&1&1\\0&0&0&2\end{pmatrix}\in M_4(\mathbb{C})$ et soit u l'endomorphisme de $\mathbb{C}^4$ défini par A. Calculer le polynôme caractéristique $P_A(X)$ et déterminer ses racines et leur multiplicité. Puis, pour chaque racine λ, déterminer une base de $\operatorname{Ker}(A-\lambda I_4)$, puis de $\operatorname{Ker}\big((A-\lambda I_4)^2\big)$, etc. jusqu'à obtenir l'espace caractéristique $V_{(\lambda)}$. Enfin, donner une base $\mathscr{C}$ de $\mathbb{C}^4$ telle que $\mathrm{Mat}_{\mathscr{C}}(u)$ soit la forme normale de Jordan J_A de A.

En développant par rapport à la dernière ligne, on obtient que $P_A(X)$ égale:

$$(2-X)\begin{vmatrix}-X&1&-1\\-3&-1-X&3\\-2&1&1-X\end{vmatrix}\xrightarrow{C_3\to C_3+C_2}(2-X)\begin{vmatrix}-X&1&0\\-3&-1-X&2-X\\-2&1&2-X\end{vmatrix}$$

$$=(2-X)^2\begin{vmatrix}-X&1&0\\-3&-1-X&1\\-2&1&1\end{vmatrix}\xrightarrow{L_2\to L_2-L_3}(2-X)^2\begin{vmatrix}-X&1&0\\-1&-2-X&0\\-2&1&1\end{vmatrix}=(2-X)^2\,(X^2+2X+1)$$

$$=(X-2)^2\,(X+1)^2.$$

Pour la valeur propre $\lambda=-1$, posons $B=A+I_4$ et faisons des opérations sur les colonnes de $B=A+I_4$:

$$\left(\begin{array}{cccc}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\\hline 1&1&-1&0\\-3&0&3&1\\-2&1&2&1\\0&0&0&3\end{array}\right)\xrightarrow[C_1\to C_1-C_2]{C_3\to C_3+C_1}\left(\begin{array}{cccc}1&0&1&0\\-1&1&0&0\\0&0&1&0\\0&0&0&1\\\hline 0&1&0&0\\-3&0&0&1\\-3&1&0&1\\0&0&0&3\end{array}\right)\xrightarrow{C_1\to C_1+3C_4}\left(\begin{array}{cccc}1&0&1&0\\-1&1&0&0\\0&0&1&0\\3&0&0&1\\\hline 0&1&0&0\\0&0&0&1\\0&1&0&1\\9&0&0&3\end{array}\right)=\left(\begin{array}{c}P\\\hline B'\end{array}\right),$$

où $B'=BP$. Donc, comme les colonnes non nulles de B' sont échelonnées, $\operatorname{Ker}(B)$ est engendré par le vecteur $v_1=e_1+e_3$. Pour la suite, on a deux méthodes.

<u>1ère méthode</u>. Calculons $B^2=(A+I_4)^2$:

$$\begin{pmatrix}1&1&-1&0\\-3&0&3&1\\-2&1&2&1\\0&0&0&3\end{pmatrix}\begin{pmatrix}1&1&-1&0\\-3&0&3&1\\-2&1&2&1\\0&0&0&3\end{pmatrix}=\begin{pmatrix}0&0&0&0\\-9&0&9&6\\-9&0&9&6\\0&0&0&9\end{pmatrix}$$

Donc on voit que $B^2=(A+I_4)^2$ est de rang 2, et $\operatorname{Ker}(B^2)$ contient les vecteurs e_2 et e_1+e_3, qui forment donc une base de $\operatorname{Ker}(B^2)$. De plus, comme $\operatorname{Ker}(B)$ est engendré par e_1+e_3, alors e_2 engendre un supplémentaire de $\operatorname{Ker}(B)$ dans $\operatorname{Ker}(B^2)$.

<u>2ème méthode</u>. On a aussi la méthode plus courte qui suit, suggérée par A. Moussaoui (en réponse à une question de T. de La Rochefoucauld). Les colonnes de la matrice inversible P forment une base (f_1,f_2,f_3,f_4) de $\mathbb{C}^4$, et les colonnes de $B'=BP$ sont les images par B des f_i. Il suffit donc de multiplier B' par B pour avoir la matrice $BB'=B^2P$ dont les colonnes donnent les images par B^2 des f_i. Comme ici $Bf_3=0$, on a bien sûr $B^2f_3=0$ et il suffira de faire les mêmes opérations sur BB' et P pour créer une autre colonne nulle de BB'. Dans le cas présent, le calcul est particulièrement simple (voir plus bas pour

un exemple plus élaboré). On a :

$$BB' = \begin{pmatrix} 1 & 1 & -1 & 0 \\ -3 & 0 & 3 & 1 \\ -2 & 1 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 9 & 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 9 & 0 & 0 & 6 \\ 9 & 0 & 0 & 6 \\ 27 & 0 & 0 & 9 \end{pmatrix}$$

d'où

$$\left(\frac{P}{BB'}\right) = \left(\begin{array}{cccc} 1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 \\ 9 & 0 & 0 & 6 \\ 9 & 0 & 0 & 6 \\ 27 & 0 & 0 & 9 \end{array}\right)$$

donc on voit que, outre le vecteur $e_1 + e_3$ qui appartenait déjà à $\text{Ker}(B)$, $\text{Ker}(B^2)$ contient le vecteur e_2, qui engendre donc un supplémentaire de $\text{Ker}(B)$ dans $\text{Ker}(B^2)$.

Finalement, utilisant l'une ou l'autre méthode, comme la dimension de l'espace caractéristique $V_{(-1)}$ est la multiplicité algébrique de la valeur propre -1, à savoir 2, on conclut que

$$V_{(-1)} = \text{Ker}\left((A + I_4)^2\right) = \text{Vect}(e_1 + e_3, e_2).$$

Considérons maintenant la valeur propre $\lambda = 2$ et faisons des opérations sur les colonnes de la matrice $C = A - 2I_4$:

$$\left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline -2 & 1 & -1 & 0 \\ -3 & -3 & 3 & 1 \\ -2 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow[C_1 \to C_1 + 2C_2]{C_3 \to C_3 + C_2} \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 2 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 0 & 1 & 0 & 0 \\ -9 & -3 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow{C_1 \to C_1 + 9C_4} \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 2 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 9 & 0 & 0 & 1 \\ \hline 0 & 1 & 0 & 0 \\ 0 & -3 & 0 & 1 \\ 9 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right) = \left(\frac{Q}{C'}\right),$$

où $C' = CQ$. Donc, comme les colonnes non nulles de C' sont échelonnées, $\text{Ker}(C)$ est engendré par le vecteur $v_1 = e_2 + e_3$. Pour la suite, on a comme avant deux méthodes.

<u>1ère méthode</u>. Calculons $C^2 = (A - 2I_4)^2$:

$$\begin{pmatrix} -2 & 1 & -1 & 0 \\ -3 & -3 & 3 & 1 \\ -2 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -2 & 1 & -1 & 0 \\ -3 & -3 & 3 & 1 \\ -2 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & -6 & 6 & 0 \\ 9 & 9 & -9 & 0 \\ 3 & -6 & 6 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

On voit donc que $C^2 = (A - 2I_4)^2$ est de rang 2 (car les deux premières colonnes sont linéairement indépendantes), et que $\text{Ker}(C^2)$ contient les vecteurs e_4 et $e_2 + e_3$, qui forment donc une base de $\text{Ker}(C^2)$. De plus, comme $\text{Ker}(C)$ est engendré par $e_2 + e_3$, alors e_4 engendre un supplémentaire de $\text{Ker}(C)$ dans $\text{Ker}(C^2)$.

<u>2ème méthode</u>. On a:

$$CC' = \begin{pmatrix} -2 & 1 & -1 & 0 \\ -3 & -3 & 3 & 1 \\ -2 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -3 & 0 & 1 \\ 9 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -9 & -6 & 0 & 0 \\ 27 & 9 & 0 & 0 \\ -9 & -6 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

d'où

$$\left(\frac{Q}{CC'}\right) = \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 2 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 9 & 0 & 0 & 1 \\ \hline -9 & -6 & 0 & 0 \\ 27 & 9 & 0 & 0 \\ -9 & -6 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

donc on voit que, outre le vecteur $e_2 + e_3$ qui appartenait déjà à $\mathrm{Ker}(C)$, $\mathrm{Ker}(C^2)$ contient le vecteur e_4, qui engendre donc un supplémentaire de $\mathrm{Ker}(C)$ dans $\mathrm{Ker}(C^2)$.

Finalement, utilisant l'une ou l'autre méthode, comme la dimension de l'espace caractéristique $V_{(2)}$ est la multiplicité algébrique de la valeur propre 2, à savoir 2, on conclut que

$$V_{(2)} = \mathrm{Ker}\left((A - 2I_4)^2\right) = \mathrm{Vect}(e_2 + e_3, e_4).$$

Il résulte de ce qui précède que la forme normale de Jordan de A est la matrice

$$J_A = \left(\begin{array}{rr|rr} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{array}\right).$$

Plus précisément, soit u l'endomorphisme de $\mathbb{C}^4$ défini par A. Comme $(u + \mathrm{id})(e_2) = e_1 + e_3 \in V_{-1}$ et $(u - 2\,\mathrm{id})(e_4) = e_2 + e_3 \in V_2$, alors $\mathscr{C} = (e_1 + e_3, e_2, e_2 + e_3, e_4)$ est une base de $V = \mathbb{C}^4$ dans laquelle la matrice de u est la matrice J_A ci-dessus.

Exemple 2. Soit $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & -1 & 3 & -1 \\ -1 & 1 & 1 & 3 \end{pmatrix} \in M_4(\mathbb{R})$ et soit u l'endomorphisme de $\mathbb{R}^4$ défini par A. Calculer $P_A(X)$ et déterminer ses racines. Pour chaque racine λ, déterminer une base de chaque $\mathrm{Ker}\big((A - \lambda I_4)^i\big)$, pour $i = 1, 2, \ldots$ et en déduire la forme normale de Jordan J de A, ainsi qu'une base $\mathscr{C}$ de $\mathbb{R}^4$ telle que $\mathrm{Mat}_{\mathscr{C}}(u) = J$.

On a :

$$\begin{vmatrix} 1-X & 1 & 1 & 1 \\ -1 & 1-X & 1 & 1 \\ 1 & -1 & 3-X & -1 \\ -1 & 1 & 1 & 3-X \end{vmatrix} \xrightarrow{C_1 \to C_1 + C_4} \begin{vmatrix} 2-X & 1 & 1 & 1 \\ 0 & 1-X & 1 & 1 \\ 0 & -1 & 3-X & -1 \\ 2-X & 1 & 1 & 3-X \end{vmatrix} =$$

$$(2-X)\begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1-X & 1 & 1 \\ 0 & -1 & 3-X & -1 \\ 1 & 1 & 1 & 3-X \end{vmatrix} \xrightarrow{L_4 \to L_4 - L_1} (2-X)\begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1-X & 1 & 1 \\ 0 & -1 & 3-X & -1 \\ 0 & 0 & 0 & 2-X \end{vmatrix} =$$

$$(2-X)^2 \begin{vmatrix} 1-X & 1 \\ -1 & 3-X \end{vmatrix} = (2-X)^2\,(X^2 - 4X + 4) = (X-2)^4$$

donc $\lambda = 2$ est la seule valeur propre de A. Posons $B = A - 2I_4$ et déterminons une base de $\mathrm{Ker}(B^i)$, pour $i = 1, 2, \ldots$ en faisant des opérations sur les colonnes.

Notons (e_1, e_2, e_3, e_4) la base canonique de $\mathbb{R}^4$. Le calcul précédent montre déjà que $e_1 + e_4 \in \mathrm{Ker}(B)$:

$$\left(\begin{array}{rrrr} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline -1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & 1 \end{array}\right) \xrightarrow[\substack{C_3 \to C_3 - C_2 \\ C_4 \to C_4 - C_2}]{C_1 \to C_1 + C_4} \left(\begin{array}{rrrr} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ \hline 0 & 1 & 0 & 0 \\ 0 & -1 & 2 & 2 \\ 0 & -1 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right) = \left(\frac{P}{B'}\right), \quad \text{où } B' = BP.$$

Comme les colonnes non nulles de B' sont échelonnées, on obtient que $\mathrm{Ker}(B)$ est de dimension 1, engendré par $e_1 + e_4$. Comme le nombre de blocs de Jordan pour la valeur propre $\lambda = 2$ est la dimension de $\mathrm{Ker}(A - 2I_4) = \mathrm{Ker}(B)$, on sait donc déjà que la forme normale de Jordan de A est

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

Calculons maintenant:

$$BB' = \begin{pmatrix} -1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -1 & 2 & 2 \\ 0 & -1 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -2 & 4 & 2 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & -2 \\ 0 & -2 & 4 & 2 \end{pmatrix}$$

d'où

$$\left(\frac{P}{BB'}\right) = \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ \hline 0 & -2 & 4 & 2 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & -2 \\ 0 & -2 & 4 & 2 \end{array}\right) \xrightarrow[C_4 \to C_4 + C_2]{C_3 \to C_3 + 2C_2} \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ \hline 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & -2 \\ 0 & -2 & 0 & 0 \end{array}\right) = \left(\frac{PQ}{B''}\right), \quad \text{où } B'' = B^2PQ.$$

Comme les colonnes non nulles de B'' sont échelonnées, on obtient que $\mathrm{Ker}(B^2)$ est de dimension 2, et que $e_2 + e_3$ engendre un supplémentaire de $\mathrm{Ker}(B) = \mathbb{R}(e_1 + e_4)$ dans $\mathrm{Ker}(B^2)$. Calculons maintenant:

$$BB'' = \begin{pmatrix} -1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & -2 \\ 0 & -2 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4 \end{pmatrix}$$

d'où

$$\left(\frac{PQ}{BB''}\right) = \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4 \end{array}\right) = \left(\frac{PQ}{B'''}\right), \quad \text{où } B''' = B^3PQ$$

donc $\mathrm{Ker}(B^3)$ est de dimension 3, et e_2 engendre un supplémentaire de $\mathrm{Ker}(B^2)$ dans $\mathrm{Ker}(B^3)$.

Enfin, on voit que $B^3(e_4) = -4(e_1 + e_4) \neq 0$, donc e_4 engendre un supplémentaire de $\mathrm{Ker}(B^3)$ dans $\mathrm{Ker}(B^4) = \mathbb{R}^4$.

Donc les vecteurs $v_4 = e_4$, $v_3 = (u - 2\,\mathrm{id})(e_4) = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$,

$$v_2 = (u - 2\,\mathrm{id})(v_3) = \begin{pmatrix} -1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ -2 \\ 0 \end{pmatrix} = B^2 e_4$$

et

$$v_1 = (u - 2\,\mathrm{id})(v_2) = B^3 e_4 = \begin{pmatrix} -4 \\ 0 \\ 0 \\ -4 \end{pmatrix}$$

forment une base $\mathscr{C}$ de $\mathbb{R}^4$ telle que

$$\mathrm{Mat}_{\mathscr{C}}(u) = J = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

<u>Remarque</u>. Les matrices échelonnées B', B'' et B''' montrent que

$$\mathrm{Im}(u) = \mathrm{Vect}(e_1 + e_4, e_2, e_3), \qquad \mathrm{Im}(u^2) = \mathrm{Vect}(e_1 + e_4, e_2 + e_3), \qquad \mathrm{Im}(u^3) = \mathrm{Vect}(e_1 + e_4).$$

5.2. Décomposition de Dunford

Soient k un corps et V un k-espace vectoriel de dimension finie.

***Définition 5.2.1* (Endomorphismes qui commutent).** — On dit que deux endomorphismes u, v de V *commutent* si $u \circ v = v \circ u$. Dans ce cas, pour tout entier $n \geq 1$ on peut calculer $(u+v)^n$ par la formule du binôme:

$$(u+v)^n = \sum_{i=0}^{n} \binom{n}{i} u^{n-i} v^i, \qquad \text{où} \quad \binom{n}{i} = \frac{n!}{i!\,(n-i)!} = \frac{n(n-1)\cdots(n-i+1}{i!}.$$

(Le coefficient binomial $\binom{n}{i}$ est aussi noté C_n^i.)

Attention, cette formule est fausse si u et v ne commutent pas! Par exemple, on a $(u+v)^2 = u^2 + uv + vu + v^2$ et ceci est $\neq u^2 + 2uv + v^2$ si $vu \neq uv$, par exemple si $u = E_{12}$ et $v = E_{21}$ dans $M_2(k)$.

Lemme 5.2.2. — *Soient u, v deux endomorphismes de V qui commutent.*

(1) *Soient λ une valeur propre de u, m sa multiplicité algébrique, et $V_\lambda = \mathrm{Ker}(u - \lambda)$ (resp. $V_{(\lambda)} = \mathrm{Ker}\big((u - \lambda\,\mathrm{id}_V)^m\big)$) l'espace propre (resp. caractéristique) associé. Alors V_λ et $V_{(\lambda)}$ sont stables par v.*

(2) *Si u et v sont diagonalisables, alors V possède une base formée de vecteurs propres communs à u et v. Par conséquent, $u + v$ et uv sont diagonalisables.*

(3) *Si u et v sont nilpotents, il en est de même de uv et de $u + v$.*

Démonstration. — (1) Soit $x \in V_\lambda$, alors $u(v(x)) = v(u(x)) = \lambda v(x)$, donc $v(x) \in V_\lambda$. De même, comme v commute à u, il commute aussi à $U = (u - \lambda\,\mathrm{id}_V)^m$, donc si $x \in V_{(\lambda)}$, alors $U(v(x)) = v(U(x)) = 0$, donc $v(x) \in \mathrm{Ker}(U) = V_{(\lambda)}$. Ceci montre que V_λ et $V_{(\lambda)}$ sont stables par v.

(2) Soit $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}$ la décomposition de V en espaces propres de u. Fixons un indice i. D'après (1), V_{λ_i} est stable par v et d'après le théorème 4.1.14, V_{λ_i} admet une base $\mathscr{B}_i$ formée de vecteurs propres de v, qui sont tous des vecteurs propres de u pour la valeur propre λ_i. Alors $\mathscr{B} = \mathscr{B}_1 \cup \cdots \cup \mathscr{B}_r$ est une base de V formée de vecteurs propres communs à u et v; et dans cette base les matrices de $u + v$ et uv sont diagonales.

(3) Supposons $u^r = 0 = v^s$. Comme u et v commutent, on a $(uv)^n = u^n v^n = 0$ si $n \geq \max(r, s)$. D'autre part, pour tout $n \in \mathbb{N}$ on a

$$(u+v)^n = \sum_{i=0}^{n} \binom{n}{i} u^i v^{n-i}\,;$$

le terme u^i (resp. v^{n-i}) est nul si $i \geq r$ (resp. $n - i \geq s$), donc pour que le terme $u^i\, v^{n-i}$ soit $\neq 0$, il faut que $i \leq r - 1$ et $n - i \leq s - 1$, d'où $n \leq r + s - 2$: ceci montre que $(u+v)^{r+s-1} = 0$. $\square$

Remarques 5.2.3. — 1) Attention, si u et v sont diagonalisables mais ne commutent pas, alors en général $u + v$ et uv ne sont pas diagonalisables! Par exemple, les matrices $A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$ dans $M_2(\mathbb{R})$ sont diagonalisables, mais ni $A + B$ ni AB ne l'est. (Exercice: vérifier ces assertions.)

2) Attention, si u et v sont des endomorphismes nilpotents, alors en général ni uv ni $u + v$ ne sont nilpotents! Par exemple, dans $M_2(k)$ les matrices élémentaires E_{12} et E_{21} sont de carré nul, mais $E_{12}E_{21} = E_{11}$ n'est pas nilpotente, et $S = E_{21} + E_{12}$ non plus (car $S^2 = I_2$).

Lemme 5.2.4. — *Soit $u \in \mathrm{End}_k(V)$.*

(1) *Si u est nilpotent, 0 est valeur propre de u, et c'est la seule valeur propre.*
(2) *En particulier, si u est diagonalisable et nilpotent, alors $u = 0$.*

Démonstration. — (1) Supposons u nilpotent et soit r son indice de nilpotence, c.-à-d., $u^r = 0$ mais $u^{r-1} \neq 0$. Soit $x \in V$ tel que $u^{r-1}(x) \neq 0$, alors $u^{r-1}(x)$ appartient à $\mathrm{Ker}(u)$ donc est vecteur propre pour la valeur propre 0.

Réciproquement, si μ est une valeur propre de u et $x \neq 0$ un vecteur propre associé, alors $u(x) = \mu x$ entraîne $0 = u^r(x) = \mu^r x$, d'où $\mu = 0$. Ceci prouve (1).

(2) en découle, car si V admet une base $(e_1, \dots, e_n)$ formée de vecteurs propres de u, chacun associé à la valeur propre 0, alors $u(e_i) = 0$ pour tout i, donc $u = 0$. $\square$

***Théorème et définition 5.2.5* (Décomposition de Dunford).** — *Soient V un $\mathbb{C}$-espace vectoriel de dimension finie et $u \in \mathrm{End}_{\mathbb{C}}(V)$. Alors u se décompose de façon unique sous la forme*

$$(\dagger) \qquad u = s + n, \qquad \textit{avec} \begin{cases} s \textit{ diagonalisable et } n \textit{ nilpotent,} \\ s \textit{ et } n \textit{ qui commutent, i.e. } sn = ns. \end{cases}$$

On dit que s (resp. u) est la partie semi-simple *(resp.* partie nilpotente*) de u.*

Démonstration. — Écrivons $P_u(X) = (-1)^n \prod_{i=1}^r (X - \lambda_i)^{m_i}$, où $\lambda_1, \ldots, \lambda_r$ sont les valeurs propres, deux à deux distinctes, de u. Pour tout i, soit u_i la restriction de u à l'espace caractéristique $N_i = V_{(\lambda_i)}$ et soit $\mathscr{B}_i$ une base de N_i telle que $\mathrm{Mat}_{\mathscr{B}_i}(u_i)$ soit une matrice de Jordan $J_{\mathbf{p}_i}(\lambda_i)$ (alors, nécessairement, $\widetilde{\mathbf{p}}_i$ correspond à la suite des noyaux de u_i).

Soit s l'endomorphisme de V qui égale $\lambda_i \,\mathrm{id}_{N_i}$ sur chaque N_i. Alors, pour tout $x \in V$, on a :

$$(s \circ u)(x) = (u \circ s)(x).$$

En effet, comme les deux membres sont linéaires en x, il suffit de vérifier cette égalité lorsque $x \in N_i$; dans ce cas les deux membres égalent $\lambda_i u(x)$. Donc s et u commutent.

D'autre part, la matrice de $n = u - s$ est triangulaire stricte, donc nilpotente (cf. 5.1.2). On a donc décomposé u sous la forme: $u = s + n$ avec s et n qui commutent, s diagonalisable et n nilpotent. Ceci prouve l'existence.

Montrons l'unicité. Soit $u = s' + n'$ une autre décomposition ayant les mêmes propriétés. Alors on a :

$$(*) \qquad s - s' = n' - n.$$

D'autre part, comme s' et n' commutent, ils commutent avec leur somme $s' + n' = u$, donc, d'après le lemme 5.2.2, ils préservent chaque espace caractéristique N_i de u, donc commutent avec s qui est une homothétie sur chacun de ces espaces. Ils commutent donc aussi avec $n = u - s$.

Alors, comme s et s' (resp. n et n') commutent et sont diagonalisables (resp. nilpotents), $s - s'$ est diagonalisable et $n' - n$ est nilpotent, d'après 5.2.2. Donc, $s - s' = n' - n$ est à la fois diagonalisable et nilpotent, donc nul d'après 5.2.4, d'où $s = s'$ et $n = n'$. Ceci prouve l'unicité de la décomposition ($\dagger$). ☐

Terminologie 5.2.5.1. — La décomposition de Dunford est aussi appelé *décomposition de Jordan*. On a évité cette terminologie, pour éviter une confusion avec la "réduction à la forme normale de Jordan".

Remarque 5.2.6. — Attention! Si A est une matrice triangulaire supérieure, et si l'on note D la "partie diagonale" de A et T la partie de A "au-dessus de la diagonale" alors l'écriture

$$A = D + T$$

n'est pas en général la décomposition de Dunford, car D et T ne commutent pas nécessairement! (Le théorème précédent dit juste qu'on peut faire un changement de base, c.-à-d., remplacer A par une matrice conjuguée $A' = P^{-1}AP$, de sorte que $A' = D' + T'$ soit la décomposition de Dunford de l'endomorphisme u défini par A.) Par exemple, dans $M_2(\mathbb{C})$,

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

n'est pas la décomposition de Dunford-Jordan de la matrice A de gauche: celle-ci est diagonalisable ($Ae_1 = e_1$ et $A(e_2 - e_1) = -(e_2 + e_1)$) donc égale à sa partie semi-simple! (Et bien sûr, si l'on note u l'endomorphisme de $\mathbb{R}^2$ défini par A, la matrice de u dans la base $\mathscr{B}' = (e_1, e_2 - e_1)$ est $A' = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$.)

5.3. Exponentielles de matrices

Dans cette section, on désigne par $\mathbb{K}$ le corps $\mathbb{R}$ des réels ou le corps $\mathbb{C}$ des nombres complexes et l'on note $|\cdot|$ la valeur absolue usuelle sur $\mathbb{K}$, c.-à-d., si $x \in \mathbb{R}$, $|x| = \sqrt{x^2}$ et si $z = a + ib \in \mathbb{C}$ (où $a, b \in \mathbb{R}$ et $i^2 = -1$), $|z| = \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}$.

***Définition 5.3.1* (Normes).** — Soit E un $\mathbb{K}$-espace vectoriel. Une **norme** $\|\cdot\|$ sur E est une application $E \to \mathbb{R}_+$, $x \mapsto \|x\|$ vérifiant les trois propriétés suivantes:

(1) $\|x\| = 0 \Leftrightarrow x = 0$.
(2) Pour tout $t \in \mathbb{K}$, $x \in E$, on a $\|tx\| = |t| \cdot \|x\|$ (où $|t|$ est la valeur absolue de t).
(3) $\|u + v\| \leq \|u\| + \|v\|$, pour tout $u, v \in E$.

Dans ce cas, on dit que E est un $\mathbb{K}$-*espace vectoriel normé* (en abrégé: evn).

***Définitions 5.3.2* (Suites de Cauchy).** — Soit $(E, \|\cdot\|)$ un $\mathbb{K}$-espace vectoriel normé.

(1) On dit qu'une suite $(u_n)_{n\in\mathbb{N}}$ d'éléments de E est une suite *de Cauchy* si la propriété suivante est vérifiée:

$$\text{(Cauchy)} \qquad \forall \varepsilon > 0, \quad \exists n_0 \quad \text{tel que} \quad \forall m, n \geq n_0, \quad \|u_m - u_n\| < \varepsilon.$$

(On peut mémoriser ceci en disant que la "suite des différences $u_m - u_n$"tend vers 0 quand $m, n \to +\infty$.)

(2) On dit que $(E, \|\cdot\|)$ est *complet* si toute suite de Cauchy $(u_n)_{n\in\mathbb{N}}$ est convergente (i.e. il existe $\ell \in E$ tel que $\lim_{n\to+\infty} \|u_n - \ell\| = 0$, un tel ℓ étant alors unique).

On admet la proposition ci-dessous (une démonstration est donnée dans un appendice à la fin de ce chapitre).

Proposition 5.3.3. — *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie. Alors toutes les normes sur E sont* équivalentes, *c.-à-d., si $\|\cdot\|$ et $\|\cdot\|'$ sont deux normes sur E, il existe des constantes $c, C \in \mathbb{R}_+^*$ telles que:*

$$(\dagger) \qquad \forall x \in E, \qquad c\cdot\|x\| \leq \|x\|' \leq C\cdot\|x\|.$$

On en déduit les deux théorèmes suivants.

Théorème 5.3.4. — *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, muni d'une norme $\|\cdot\|$.*

(i) *La notion de suite de Cauchy dans E ne dépend pas de la norme choisie. C'est-à-dire, si $(u_n)_{n\in\mathbb{N}}$ est une suite d'éléments de E et si $\|\cdot\|'$ est une seconde norme sur E, alors $(u_n)_{n\in\mathbb{N}}$ est de Cauchy pour $\|\cdot\|$ si et seulement si elle l'est pour $\|\cdot\|'$.*

(ii) *E est complet pour la norme $\|\cdot\|$.*

Démonstration. — Le point (i) est une conséquence immédiate de la proposition précédente. Prouvons le point (ii). Soit $\mathscr{B} = (e_1, \ldots, e_d)$ une base de E. D'après le point (i), il suffit de montrer que E est complet pour la norme $\|\cdot\| = \|\cdot\|_{\mathscr{B},\infty}$ définie par:

$$\forall (x_1, \ldots, x_d) \in \mathbb{K}^d, \qquad \|x_1e_1 + \cdots + x_de_d\| = \mathrm{Max}(|x_1|, \ldots, |x_d|).$$

Soit alors $(u_n)_{n\in\mathbb{N}}$ une suite de Cauchy pour cette norme. Pour tout $i = 1, \ldots, d$, notons $(u_n^i)_{n\in\mathbb{N}}$ la suite à valeurs dans $\mathbb{K}$ formée par les i-ièmes coordonnées des vecteurs u_n (c.-à-d., $u_n = u_n^1e_1 + \cdots + u_n^de_d$). Alors pour tout i et tous m, n on a

$$|u_m^i - u_n^i| \leq \|u_m - u_n\|$$

et donc la suite $(u_n^i)_{n\in\mathbb{N}}$ est de Cauchy, donc converge vers une limite ℓ_i (puisque $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ est complet). Alors on voit facilement que la suite $(u_n)_{n\in\mathbb{N}}$ converge vers l'élément $\ell = \ell_1e_1 + \cdots + \ell_de_d$ de E. Ceci prouve le théorème. □

Théorème 5.3.5. — *Soient E, F deux $\mathbb{K}$-espaces vectoriels de dimension finie, munis de normes $\|\cdot\|_E$ et $\|\cdot\|_F$, et soit $f : E \to F$ une application linéaire. Alors il existe une constante $k \in \mathbb{R}_+^*$ telle que:*

$$(\ddagger) \qquad \forall x \in E, \qquad \|f(x)\|_F \leq k\cdot\|x\|_E\ .$$

En particulier, f est continue.

Démonstration. — Soit $\mathscr{B} = (e_1, \ldots, e_d)$ une base de E et soit $M = \sum_{i=1}^d \|f(e_i)\|_F$. Notons $\|\cdot\|_{\mathscr{B},\infty}$ la norme sur E définie plus haut. D'après la proposition 5.3.3, il existe $C \in \mathbb{R}_+^*$ tel que $\|v\|_\infty \leq C\cdot\|v\|_E$, pour tout $v \in E$. Écrivant $v = x_1e_1 + \cdots + x_de_d$, on a alors

$$\|f(v)\|_F \leq \sum_{i=1}^d |x_i|\cdot\|f(e_i)\|_F \leq \|v\|_\infty\cdot M \leq CM\cdot\|v\|_E,$$

d'où le théorème. □

Dans la suite, on munit $\mathbb{K}^d$ de la norme $\|\cdot\|_\infty$, définie par

$$\|(x_1, \ldots, x_d)\|_\infty = \mathrm{Max}(|x_1|, \ldots, |x_d|), \qquad \forall (x_1, \ldots, x_d) \in \mathbb{K}^d. \quad ^{(2)}$$

Alors la sphère unité $S^{d-1} = \{x \in \mathbb{K}^d \mid \|x\|_\infty = 1\}$ est compacte (étant un fermé borné de $\mathbb{K}^d$).

---

(2) D'autre choix possibles sont les normes $\|(x_1, \ldots, x_d)\|_1 = |x_1| + \cdots + |x_d|$ ou $\|(x_1, \ldots, x_d)\|_2 = \sqrt{|x_1|^2 + \cdots + |x_d|^2}$; tous ces choix sont équivalents, d'après la proposition 5.3.3.

Définition 5.3.6 **(Normes matricielles sur $M_d(\mathbb{K})$).** — Pour tout $A \in M_d(\mathbb{K})$, l'application $x \mapsto \|Ax\|_\infty$ est continue, donc est bornée sur le compact S^{d-1}. On pose alors

$$|||A||| = \mathrm{Max}_{x\in S^{d-1}} \|Ax\|_\infty.$$

On vérifie facilement que $|||\cdot|||$ est une norme sur $M_d(\mathbb{K})$, appelée *norme matricielle* associée à la norme $\|\cdot\|_\infty$ donnée sur $\mathbb{K}^d$.

Remarque 5.3.7. — Pour tout $x \in \mathbb{K}^d - \{0\}$, $x' = \dfrac{1}{\|x\|_\infty} x$ appartient à S^{d-1} et comme $x = \|x\|_\infty \cdot x'$ on a:

$$\|Ax\|_\infty = \|x\|_\infty \cdot \|Ax'\| \leq |||A||| \cdot \|x\|_\infty,$$

inégalité qui est aussi vérifiée pour $x = 0$. Pour tout $A, B \in M_d(\mathbb{K})$, et $x \in S^{d-1}$, on a donc

$$\|ABx\|_\infty \leq |||A||| \cdot \|Bx\|_\infty \leq |||A||| \cdot |||B|||$$

d'où

$$(\star) \qquad |||AB||| \leq |||A||| \cdot |||B|||.$$

Par conséquent, pour tout $n \in \mathbb{N}^*$, on a

$$(*) \qquad |||A^n||| \leq |||A|||^n.$$

(Par ailleurs, la norme matricielle de $A^0 = I_d$ est égale à 1.)

Proposition et définition 5.3.8. — *Pour tout $A \in M_d(\mathbb{K})$, la suite de matrices*

$$S_n = \sum_{i=0}^{n} \frac{A^i}{i!} = I_d + A + \frac{A^2}{2} + \cdots + \frac{A^n}{n!}$$

est de Cauchy donc converge. Sa limite est notée $\exp(A)$*, et l'on a donc* $\exp(A) = \sum_{i=0}^{\infty} \dfrac{A^i}{i!}$.

Démonstration. — En effet, d'après le théorème 5.3.4, le $\mathbb{K}$-espace vectoriel $M_d(\mathbb{K}) \simeq \mathbb{K}^{d^2}$, muni de la norme matricielle, est complet. D'autre part, pour tout $q > p$, on a

$$|||S_q - S_p||| = |||\sum_{i=p+1}^{q} \frac{A^i}{i!}||| \leq \sum_{i=p+1}^{q} \frac{|||A^i|||}{i!} \leq \sum_{i=p+1}^{q} \frac{|||A|||^i}{i!}$$

et comme la suite réelle $\sum_{i=0}^{n} \dfrac{|||A|||^i}{i!}$ converge vers $\exp(|||A|||)$, alors pour tout $\varepsilon > 0$ il existe p_0 tel que pour tous $q > p \geq p_0$ on ait $\sum_{i=p+1}^{q} \dfrac{|||A|||^i}{i!} < \varepsilon$, et donc la suite (S_n) est de Cauchy, donc converge vers une matrice qu'on note $\exp(A)$. □

Remarque 5.3.9. — Si A est *nilpotente*, i.e. s'il existe $n \in \mathbb{N}^*$ tel que $A^n = 0$, alors $\exp(A)$ égale la somme finie $I_d + A + \cdots + A^{n-1}/(n-1)!$; plus généralement, pour tout $t \in \mathbb{K}$ on a alors:

$$\exp(tA) = I_d + tA + \frac{t^2}{2} A^2 + \cdots + \frac{t^{n-1}}{(n-1)!} A^{n-1} \qquad (\text{si } A^n = 0).$$

Proposition 5.3.10. — *Soit $A \in M_d(\mathbb{K})$ une matrice triangulaire, de termes diagonaux $\lambda_1, \ldots, \lambda_d$. Alors* $\exp(A)$ *est une matrice triangulaire, de termes diagonaux* $\exp(\lambda_1), \ldots, \exp(\lambda_d)$.

Démonstration. — Pour tout i, j, la forme linéaire $\phi_{i,j} : M_d(\mathbb{K}) \to \mathbb{K}$, qui à toute matrice $A = (a_{ij})$ associe son coefficient a_{ij}, est continue. En effet, c'est un cas particulier du théorème 5.3.5 en prenant $F = \mathbb{K}$ muni de la norme définie par la valeur absolue mais, plus simplement, ceci se voit directement comme suit. Comme a_{ij} est la i-ième coordonnée du vecteur Ae_j alors:

$$|a_{ij}| \leq \|Ae_j\|_\infty \leq |||A||| \cdot \underbrace{\|e_j\|_\infty}_{=1} = |||A|||.$$

Il en résulte que pour toute suite de matrices $(B_n)_{n\in\mathbb{N}}$ convergeant vers une matrice $B \in M_d(\mathbb{K})$, on a $\phi_{ij}(B) = \lim_{n\to+\infty} \phi_{ij}(B_n)$. Appliquons ceci à la suite $B_n = \sum_{p=0}^{n} A^p/p!$ des sommes partielles de $\exp(A)$. Comme A est supposée triangulaire, disons supérieure, alors chaque produit $A^p/p!$ est une matrice triangulaire, de termes diagonaux les $\lambda_i^p/p!$, pour $i = 1, \dots, d$, et donc B_n est aussi triangulaire, de termes diagonaux les sommes partielles $\sum_{p=0}^n \lambda_i^p/p!$, pour $i = 1, \dots, d$. Donc, pour tout $i, j = 1, \dots, d$, on a

$$\phi_{ij}(\exp(A)) = \lim_{n\to+\infty} \phi_{ij}(B_n) = \begin{cases} 0 & \text{si} \quad i > j \\ \exp(\lambda_i) & \text{si} \quad i = j \end{cases}$$

d'où la proposition. □

Proposition 5.3.11. — *Soient* $A \in M_d(\mathbb{K})$ *et* $P \in \mathrm{GL}_d(\mathbb{K})$. *Alors* $\exp({}^tA) = {}^t\exp(A)$ *et* $\exp(P^{-1}AP) = P^{-1}\exp(A)P$.

Démonstration. — Les applications $A \mapsto {}^tA$ et $A \mapsto P^{-1}AP$ sont linéaires, donc continues d'après le théorème 5.3.5. Par conséquent, pour toute suite de matrices $(B_n)_{n\in\mathbb{N}}$ convergeant vers une matrice $B \in M_d(\mathbb{K})$, on a

$${}^tB = \lim_{n\to+\infty} {}^tB_n \qquad \text{et} \qquad P^{-1}BP = \lim_{n\to+\infty} P^{-1}B_nP.$$

Appliquons ceci à la suite $B_n = \sum_{p=0}^n A^p/p!$ des sommes partielles de $\exp(A)$. Comme ${}^t(A^n) = ({}^tA)^n$ et $P^{-1}A^nP = (P^{-1}AP)^n$, on obtient que

$${}^t\exp(A) = \lim_{n\to+\infty} \sum_{i=0}^n \frac{({}^tA)^n}{n!} = \exp({}^tA) \qquad \text{et} \qquad P^{-1}BP = \lim_{n\to+\infty} \sum_{i=0}^n \frac{(P^{-1}AP)^n}{n!} = \exp(P^{-1}AP).$$

□

Corollaire 5.3.12. — *Pour tout* $A \in M_d(\mathbb{K})$ *on a* $\boxed{\text{dét}(\exp(A)) = \exp(\mathrm{Tr}(A)).}$

Démonstration. — Comme $M_d(\mathbb{R}) \subset M_d(\mathbb{C})$, il suffit d'établir la formule lorsque $\mathbb{K} = \mathbb{C}$. Dans ce cas, d'après le théorème de trigonalisation 4.2.3, il existe $P \in \mathrm{GL}_d(\mathbb{C})$ telle que $T = P^{-1}AP$ soit une matrice triangulaire supérieure. Notons $\lambda_1, \dots, \lambda_d$ ses termes diagonaux. Alors, d'une part,

$$\lambda_1 + \cdots + \lambda_d = \mathrm{Tr}(T) = \mathrm{Tr}(A).$$

D'autre part, comme $P^{-1}\exp(A)P = \exp(T)$, on a $\text{dét}(\exp(A)) = \text{dét}(\exp(T))$. Or, d'après la proposition 5.3.10, $\exp(T)$ est une matrice triangulaire, de termes diagonaux les $\exp(\lambda_i)$. On a donc

$$\text{dét}(\exp(T)) = \prod_{i=1}^d \exp(\lambda_i) = \exp(\lambda_1 + \cdots + \lambda_d) = \exp(\mathrm{Tr}(A)).$$

□

Proposition 5.3.13. — *Si* $A, B \in M_d(\mathbb{K})$ commutent *(i.e. vérifient* $AB = BA$*), alors* $\exp(A+B) = \exp(A)\exp(B)$. *En particulier, pour tout* $t, t' \in \mathbb{K}$, *on a*

$$(\star) \qquad \exp((t+t')A) = \exp(tA)\exp(t'A).$$

En particulier, on a $\exp(-A)\exp(A) = \exp(-A+A) = \exp(0) = I_d$ *(où* 0 *désigne la matrice nulle de* $M_d(\mathbb{K})$*): ceci montre que* $\exp(A)$ *est inversible, d'inverse* $\exp(-A)$.

Démonstration. — Comme A et B commutent, on a pour tout $n \in \mathbb{N}$ la formule du binôme:

$$(A+B)^n = \sum_{p=0}^n \binom{n}{p} A^p B^{n-p} = \sum_{\substack{p,q\geq 0 \\ p+q=n}} \frac{n!}{p!q!} A^p B^q.$$

Donc (renvoyant à l'appendice en fin de chapitre pour la justification de l'égalité $(*)$ ci-dessous) on a :

$$\begin{aligned} \exp(A)\exp(B) = \Big(\sum_{p\geq 0} \frac{A^p}{p!}\Big)\Big(\sum_{q\geq 0} \frac{B^q}{q!}\Big) &\overset{(*)}{=} \sum_{n\geq 0} \Big(\sum_{\substack{p,q\geq 0 \\ p+q=n}} \frac{A^p}{p!}\frac{B^q}{q!}\Big) \\ &= \sum_{n\geq 0} \frac{1}{n!}\Big(\sum_{\substack{p,q\geq 0 \\ p+q=n}} \frac{n!}{p!q!} A^p B^q\Big) = \sum_{n\geq 0} \frac{(A+B)^n}{n!} = \exp(A+B). \end{aligned}$$

La formule $(\star)$ en découle, puisque tA et $t'A$ commutent. □

Remarque 5.3.14. — Attention! Si $AB \neq BA$, alors $\exp(A+B) \neq \exp(A)\exp(B)$ en général. Par exemple, soient $A = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Alors $AB = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \neq BA = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$; d'autre part, comme $A^2 = 0 = B^2$ on a $\exp(A) = I_2 + A$ et $\exp(B) = I_2 + B$ et donc

$$\exp(A)\exp(B) = I_2 + (A+B) + AB = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \neq \exp(B)\exp(A) = I_2 + (A+B) + BA = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}.$$

Donc $\exp(A)\exp(B)$ et $\exp(B)\exp(A)$ ne peuvent pas tous les deux être égaux à $\exp(A+B)$. En fait on peut montrer (cf. TE3a 2010-11) que $\exp(A+B)$ égale $\begin{pmatrix} \cos(1) & -\sin(1) \\ \sin(1) & \cos(1) \end{pmatrix}$, donc est différent de $\exp(A)\exp(B)$ et de $\exp(B)\exp(A)$.

Proposition 5.3.15. — *Soit J le bloc de Jordan nilpotent $J_d(0)$. Pour tout bloc de Jordan $J_d(\lambda) = \lambda I_d + J$ et tout $t \in \mathbb{K}$, on a*

$$\exp(tJ_d(\lambda)) = e^{\lambda t}\big(I_d + tJ + \frac{t^2}{2}J^2 + \cdots + \frac{t^{d-1}}{(d-1)!}J^{d-1}\big) = e^{\lambda t} \begin{pmatrix} 1 & t & \dfrac{t^2}{2} & \cdots & \dfrac{t^{d-1}}{(d-1)!} \\ 0 & 1 & t & \ddots & \vdots \\ 0 & 0 & 1 & \ddots & \dfrac{t^2}{2} \\ \vdots & \ddots & \ddots & \ddots & t \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}$$

et pour toute matrice de Jordan $A = \left(\begin{array}{c|c|c|c} J_{d_1}(\lambda_1) & 0 & \cdots & 0 \\ \hline 0 & J_{d_2}(\lambda_2) & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & 0 \\ \hline 0 & \cdots & 0 & J_{d_r}(\lambda_r) \end{array}\right)$, *on a*

$$\exp(tA) = \left(\begin{array}{c|c|c|c} \exp(tJ_{d_1}(\lambda_1)) & 0 & \cdots & 0 \\ \hline 0 & \exp(tJ_{d_2}(\lambda_2)) & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & 0 \\ \hline 0 & \cdots & 0 & \exp(tJ_{d_r}(\lambda_r)) \end{array}\right).$$

Démonstration. — En effet, soit $\mathscr{B} = (e_1, \dots, e_d)$ la base canonique de $\mathbb{K}^d$. Alors $Je_1 = 0$ et $Je_i = e_{i-1}$ pour $i = 2, \dots, d$, et l'on en déduit, par récurrence sur $i = 1, \dots, d$, que l'on a

$$J^i e_j = \begin{cases} 0 & \text{si } j = 1, \dots, i, \\ e_{j-i} & \text{si } j = i+1, \dots, d, \end{cases}$$

i.e. la matrice de J^i a tous ses coefficients nuls sauf ceux de la i-ième diagonale au-dessus de la diagonale principale, sur laquelle les coefficients valent 1. Ceci donne la forme explicite de $\exp(tJ)$ donnée dans la proposition. De plus, comme tJ et $\lambda t I_d$ commutent, on a

$$\exp(\lambda t I_d + tJ) = \exp(\lambda t I_d) \cdot \exp(tJ) = \exp(\lambda t) I_d \cdot \exp(tJ) = e^{\lambda t} \exp(tJ),$$

ce qui prouve la première égalité.

Lorsque A est une matrice de Jordan ayant r blocs, notons $\mathscr{B}_i$ le sous-ensemble de $\mathscr{B}$ correspondant au bloc $J_{d_i}(\lambda_i)$ et E_i le sous-espace vectoriel de $\mathbb{K}^d$ de base $\mathscr{B}_i$, et soit u_i l'endomorphisme de $\mathbb{K}^d$ qui est nul sur E_j pour $j \neq i$ et tel que $\mathrm{Mat}_{\mathscr{B}_i}(u_i) = J_{d_i}(\lambda_i)$. Alors on obtient comme plus haut que $\mathrm{Mat}_{\mathscr{B}}(\exp(tu_i))$ est la matrice diagonale par blocs dont tous les blocs diagonaux sont l'identité, sauf le i-ième qui est $\exp(tJ_{d_i}(\lambda_i))$. De plus, comme $A = u_1 + \cdots + u_r$ et que les u_i commutent deux à deux, on déduit de 5.3.13 par récurrence sur r que

$$\exp(tA) = \exp(tu_1)\exp(tu_2)\cdots\exp(tu_r)$$

et en effectuant le produit de ces matrices diagonales par blocs, on obtient bien la matrice $\exp(tA)$ indiquée dans la proposition. $\square$

Théorème 5.3.16. — *Soit $A \in M_d(\mathbb{K})$. La fonction $\mathbb{R} \to M_d(\mathbb{K})$, $t \mapsto \exp(tA)$ est dérivable, de dérivée la fonction $t \mapsto A\exp(tA) = \exp(tA)A$. Par conséquent, la fonction $t \mapsto \exp(tA)$ est de classe C^∞, sa dérivée n-ième étant la fonction $t \mapsto A^n \exp(tA) = \exp(tA)A^n$.*

Démonstration. — Pour tout $t \in \mathbb{R}$, notons $S_n(t)$ la somme partielle $\sum_{i=0}^n t^i \dfrac{A^i}{i!}$. Soit $t \in \mathbb{R}$ tel que $|t| < 1$, alors $|t|^i \leq t^2$ pour tout $i \geq 2$ et donc

$$|||S_n(t) - I_d - tA||| \leq \sum_{i=2}^n |t|^i \cdot \frac{|||A|||^i}{i!} \leq t^2 \sum_{i=2}^n \frac{|||A|||^i}{i!} \leq t^2 \exp(|||A|||).$$

Ceci étant vrai pour tout n, on en déduit que

$$(\dagger) \qquad |||\exp(tA) - I_d - tA||| \leq t^2 \exp(|||A|||).$$

Ceci montre que la fonction $t \mapsto \exp(tA)$ est dérivable (donc a fortiori continue) en $t = 0$, de dérivée A.

Puis, pour $t_0 \in \mathbb{R}$ arbitraire et t comme ci-dessus (i.e. $|t| < 1$), on a

$$\exp((t_0 + t)A) - \exp(t_0A) - tA\exp(t_0A) = \exp(t_0A)(\exp(tA) - I_d - tA)$$

et donc

$$(\ddagger) \qquad |||\exp((t_0 + t)A) - \exp(t_0A) - tA\exp(t_0A)||| \leq t^2\, |||\exp(t_0A)||| \cdot \exp(|||A|||),$$

et ceci montre que $t \mapsto \exp(tA)$ est dérivable (donc a fortiori continue) en $t = t_0$, de dérivée $A\exp(t_0A)$. On a donc montré que la fonction $t \mapsto \exp(tA)$ est dérivable, de dérivée la fonction $t \mapsto A\exp(tA)$.

De plus, pour tout $B \in M_d(\mathbb{K})$ on obtient, en reprenant la démonstration précédente, que la fonction $t \mapsto B\exp(tA)$ est dérivable, de dérivée $t \mapsto BA\exp(tA)$; elle est donc indéfiniment dérivable, sa dérivée n-ième étant la fonction $t \mapsto BA^n \exp(tA)$. Le théorème est démontré. □

5.4. Exponentielles de matrices et équations différentielles linéaires

Théorème 5.4.1. — *Soit toujours $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et soient $A \in M_n(\mathbb{K})$ et $X : \mathbb{R} \to \mathbb{K}^n$, $t \mapsto X(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$ une fonction de classe C^∞ vérifiant l'équation différentielle*

$$X'(t) = A \cdot X(t), \qquad \textit{c.-à-d.,} \qquad \begin{pmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{pmatrix} = A \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}.$$

Alors on a $X(t) = \exp(tA) \cdot X(0)$ pour tout $t \in \mathbb{R}$.

Démonstration. — Pour démontrer ce théorème, on a besoin du:

Lemme 5.4.2. — *Soient $X : \mathbb{R} \to \mathbb{K}^n$, $t \mapsto X(t)$ et $B : \mathbb{R} \to M_n(\mathbb{K})$, $t \mapsto B(t)$ des fonctions dérivables. Alors la fonction $F : \mathbb{R} \to \mathbb{K}^n$, $t \mapsto F(t) = B(t) \cdot X(t)$ est dérivable et pour tout $t \in \mathbb{R}$ on a*

$$(\dagger) \qquad F'(t) = B'(t) \cdot X(t) + B(t) \cdot X'(t).$$

Démonstration du lemme. — En effet, on a $F(t) = \begin{pmatrix} F_1(t) \\ \vdots \\ F_n(t) \end{pmatrix}$, où pour tout $i = 1, \dots, n$,

$$F_i(t) = \sum_{j=1}^n B(t)_{ij} X_j(t) \in \mathbb{K}.$$

Alors chaque F_i est dérivable, de dérivée $F_i'(t) = \sum_{j=1}^n \big(B'(t)_{ij} X_j(t) + B(t)_{ij} X_j'(t)\big)$ et il en résulte que F est dérivable et qu'on a bien l'égalité de vecteurs colonnes:

$$F'(t) = B'(t) \cdot X(t) + B(t) \cdot X'(t).$$

□

La démonstration du théorème est maintenant facile: considérons la fonction $Z : \mathbb{R} \to \mathbb{K}^n$ définie par $Z(t) = \exp(-tA) \cdot X(t)$. D'après le lemme précédent et le théorème 5.3.16, Z est dérivable et pour tout $t \in \mathbb{R}$ on a

$$Z'(t) = -\exp(-tA)A \cdot X(t) + \exp(-tA) \cdot X'(t) = \exp(-tA) \cdot \big(X'(t) - A \cdot X(t)\big) = 0.$$

Donc Z est constante, d'où $Z(t) = Z(0) = X(0)$, et donc $X(t) = \exp(tA) \cdot X(0)$ pour tout $t \in \mathbb{R}$. □

Corollaire 5.4.3. — *Soit $A \in M_n(\mathbb{K})$. Alors l'ensemble E des fonctions $X : \mathbb{R} \to \mathbb{K}^n$, de classe C^∞, qui sont solutions de l'équation différentielle $X' = A \cdot X$, est un $\mathbb{K}$-espace vectoriel de dimension n, isomorphe à $\mathbb{K}^n$ par l'application $E \to \mathbb{K}^n$, $X \mapsto X(0)$.*

Démonstration. — D'abord, il est clair que si $X_1, X_2 \in E$ et $\lambda \in \mathbb{K}$, alors $\lambda X_1 + X_2$ est encore solution de l'équation différentielle, i.e. appartient à E. Ceci montre que E est un $\mathbb{K}$-espace vectoriel. Il est clair que l'application $\phi : E \to \mathbb{K}^n$, $X \mapsto X(0)$ est linéaire, et elle est injective puisque X est déterminé par la "condition initiale"$X(0)$ d'après le théorème précédent.

Réciproquement, soit $Y \in \mathbb{K}^n$. D'après le lemme précédent et le théorème 5.3.16, la fonction $X : t \mapsto \exp(tA) \cdot Y$ est dérivable, de dérivée $X'(t) = A\exp(tA) \cdot Y = A \cdot X(t)$, donc X appartient à E et vérifie $X(0) = Y$. Ceci montre que ϕ est aussi surjectif; c'est donc un isomorphisme, d'où le corollaire. □

Considérons maintenant une équation différentielle linéaire:

$$(\star) \qquad f^{(n)} = a_{n-1}f^{(n-1)} + \cdots + a_2 f'' + a_1 f' + a_0 f,$$

à coefficients constants $a_i \in \mathbb{K}$ et notons E le $\mathbb{K}$-espace vectoriel des fonctions $f : \mathbb{R} \to \mathbb{K}$, de classe C^∞, qui sont solutions de $(\star)$. (Lorsque les a_i sont dans $\mathbb{R}$, même si l'on ne s'intéresse qu'aux solutions $f : \mathbb{R} \to \mathbb{R}$, il peut être utile de considérer aussi les solutions à valeurs dans $\mathbb{C}$.)

Ceci devient un cas particulier du cas étudié plus haut, en posant pour tout $f \in E$:

$$X(t) = \begin{pmatrix} f(t) \\ f'(t) \\ \vdots \\ f^{(n-1)}(t) \end{pmatrix}.$$

Alors $X : \mathbb{R} \to \mathbb{K}^n$ est de classe C^∞, et vérifie l'équation différentielle:

$$X' = \begin{pmatrix} f'(t) \\ f''(t) \\ \vdots \\ f^{(n)}(t) \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & 1 \\ a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \end{pmatrix}}_{=A} \begin{pmatrix} f(t) \\ f'(t) \\ \vdots \\ f^{(n-1)}(t) \end{pmatrix}.$$

Réciproquement, si $X(t) = \begin{pmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_{n-1}(t) \end{pmatrix}$ est une fonction $\mathbb{R} \to \mathbb{K}^n$, de classe C^∞, solution de l'équation différentielle $X' = AX$, alors on a

$$\begin{pmatrix} x_0'(t) \\ x_1'(t) \\ \vdots \\ x_{n-2}'(t) \\ x_{n-1}'(t) \end{pmatrix} = A \begin{pmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_{n-1}(t) \\ x_n(t) \end{pmatrix} = \begin{pmatrix} x_1(t) \\ x_2 \\ \vdots \\ x_{n-1}(t) \\ a_0x_0(t) + \cdots + a_{n-1}x_{n-1}(t) \end{pmatrix}$$

d'où $x_1 = x_0'$, puis $x_2 = x_1' = x_0''$, etc., et donc $x_i = x_0^{(i)}$ pour $i = 1, \ldots, n-1$, et

$$x_0^{(n)} = x_{n-1}' = a_{n-1}x_{n-1} + \cdots + a_0x_0 = a_{n-1}x_0^{(n-1)} + \cdots + a_0x_0$$

donc $f = x_0$ est solution de $(\star)$ et $X = \begin{pmatrix} f(t) \\ f'(t) \\ \vdots \\ f^{(n-1)}(t) \end{pmatrix}$. On a ainsi obtenu que l'application linéaire

$$\{\text{solutions } f \text{ de } (\star)\} \to \{\text{solutions } X \text{ de } X' = AX\}, \qquad f \mapsto \begin{pmatrix} f \\ f' \\ \vdots \\ f^{(n-1)} \end{pmatrix}$$

est bijective, donc un isomorphisme de $\mathbb{K}$-espaces vectoriels. Combinant ceci avec le théorème précédent, on obtient le:

Théorème 5.4.4. — *Soient $a_0, \dots, a_{n-1} \in \mathbb{K}$. Alors le $\mathbb{K}$-espace vectoriel E des fonctions $f : \mathbb{R} \to \mathbb{K}$, de classe C^∞, vérifiant l'équation différentielle linéaire $(\star)$ est de dimension n. Plus précisément, l'application $E \to \mathbb{K}^n$, $f \mapsto \begin{pmatrix} f(0) \\ f'(0) \\ \vdots \\ f^{(n-1)}(0) \end{pmatrix}$ est un isomorphisme de $\mathbb{K}$-espaces vectoriels. En d'autres termes: toute solution f de $(\star)$ est entièrement déterminée par les “données initiales”$(f(0), f'(0), \dots, f^{(n-1)}(0))$, et celles-ci peuvent être choisies arbitrairement.*

Exemple 5.4.5. — Soit $\omega \in \mathbb{R}_+^*$ et soit E l'ensemble des fonctions $f : \mathbb{R} \to \mathbb{R}$, de classe C^∞, vérifiant l'équation différentielle

$$(\dagger) \qquad f'' + \omega^2 f = 0.$$

D'après le théorème précédent, E est un $\mathbb{R}$-espace vectoriel de dimension 2. D'autre part, les fonctions $f_1 : t \mapsto \cos(\omega t)$ et $f_2 : t \mapsto \sin(\omega t)$ appartiennent à E (car $f_1'(t) = -\omega \sin(\omega t)$ et $f_2'(t) = \omega \cos(\omega t)$); de plus elles sont linéairement indépendantes car $f_1(0) = 1 = f_2'(0)$ et $f_1'(0) = 0 = f_2(0)$, donc elles forment une base de E. Donc toute solution f de $(\dagger)$ s'écrit de façon unique $f(t) = a\cos(\omega t) + b\sin(\omega t)$, avec $a, b \in \mathbb{R}$. De plus, posons $A = \sqrt{a^2 + b^2}$; si $(a, b) \neq (0, 0)$, il existe $\varphi \in \mathbb{R}$, unique modulo 2π, tel que $\sin(\varphi) = a/A$ et $\cos(\varphi) = b/A$, et l'on a donc $f(t) = A\sin(\omega t + \varphi)$.

5.5. Appendice (†): Espaces quotients

5.5.0. Introduction. — Soit $f : V \to W$ une application linéaire *surjective*, et soit $K = \operatorname{Ker}(f)$. Alors, pour tout $w \in W$, il existe $v \in V$ tel que $f(v) = w$, et si v' a la même propriété, on a $v' - v \in K$, c.-à-d., $v' \in v + K$. On peut donc identifier les éléments de W aux classes :

$$v + K = \{v + x \mid x \in K\} = \{v' \in V \mid v' - v \in K\}$$

Cette construction se généralise: pour tout sous-espace vectoriel E de V, on peut construire de façon canonique un espace vectoriel noté V/E et appelé “quotient de V par E”, et une application linéaire surjective $\pi : V \to V/E$ dont le noyau est E.

Cette construction a d'innombrables applications, dont certaines sont indiquées plus bas. En termes imagés, on peut dire que travailler dans V/E permet de “négliger” les éléments de E, et donc dans certains cas de rendre la situation plus simple. (Une analogie est qu'il est souvent plus facile de calculer “modulo n” que dans $\mathbb{Z}$; par exemple on voit facilement que $3^{2010} \equiv 1 \mod 7$ sans avoir à calculer (heureusement!) 3^{2010}.)

Soit A un groupe *abélien* (c.-à-d., commutatif) et E un sous-groupe. On considère sur A la relation d'équivalence $\sim_E$ définie par E, c.-à-d., telle que :

$$x \sim_E y \Leftrightarrow x - y \in E,$$

pour tout $a \in A$ on note

$$a + E = \{b \in A \mid b - a \in E\} = \{a + x \mid x \in E\}$$

sa classe d'équivalence, on note A/E l'ensemble de ces classes (c.-à-d., l'ensemble *quotient de A par la relation d'équivalence $\sim_E$*), et π la projection $A \to A/E$ qui à tout $a \in A$ associe sa classe $a + E$.

On va munir A/E d'une loi de groupe abélien, telle que l'application $\pi : A \to A/E$ soit un morphisme de groupes abéliens. Cette condition impose que l'on ait:

$$(*) \qquad (a+E)+(b+E) = \pi(a)+\pi(b) = \pi(a+b) = a+b+E.$$

On voudrait donc prendre ceci comme définition de l'addition dans A/E, mais il faut vérifier que ceci a un sens. En effet, la classe $a+E$ ne détermine pas uniquement l'élément a; c'est aussi la classe $a'+E$ pour tout $a' \sim_E a$. Il faut donc vérifier que pour tous $a' \sim_E a$ et $b' \sim_E b$, le "résultat de l'addition"

$$a'+b'+E$$

est le même. Or, il existe $x, y \in E$ tels que $a' = a+x$ et $b' = b+y$, et comme l'addition dans A est commutative, on a:

$$a'+b' = a+x+b+y = a+b+(x+y),$$

et $x+y \in E$, ce qui montre que $a'+b'+E = a+b+E$. On peut donc *définir* une addition dans A/E par la formule $(*)$, et d'après cette formule, il est clair que l'addition est associative et commutative, et admet pour élément neutre la classe $0+E$. Enfin, pour tout $a \in A$, la classe $-a+E$ est l'opposée de $a+E$, puisque

$$(a+E)+(-a+E) = a-a+E = 0+E.$$

On a donc obtenu la

***Proposition 5.5.1* (Groupes abéliens quotients)**. — *Soient A un groupe abélien et E un sous-groupe de A. Il existe sur l'ensemble A/E une unique loi de groupe abélien telle que la projection $\pi : A \to A/E$ soit un morphisme de groupes abéliens; si on note $\overline{a} = \pi(a) = a+E$, cette loi est définie par*

$$\overline{a}+\overline{b} = \overline{a+b}.$$

Remarque 5.5.1.1. — C'est ainsi qu'on construit les "entiers modulo n" $\mathbb{Z}/n\mathbb{Z}$, pour tout $n \in \mathbb{Z}$.

Soient maintenant k un corps, V un k-espace vectoriel, E un sous-espace vectoriel. D'après ce qui précède, l'ensemble quotient V/E est muni d'une structure de groupe abélien. On va le munir d'une structure de k-espace vectoriel, telle que l'application $\pi : V \to V/E$ soit linéaire. Cette condition impose que l'on ait, pour tous $t \in k$ et $v \in V$:

$$(**) \qquad t \cdot (v+E) = t \cdot \pi(v) = \pi(t \cdot v) = t \cdot v + E.$$

On voudrait donc prendre ceci comme définition de la loi externe, mais à nouveau il faut vérifier que ceci a un sens, puisque la classe $v+E$ ne détermine pas uniquement l'élément v; c'est aussi la classe $v'+E$ pour tout $v' \sim_E v$. Il faut donc vérifier que pour tout $v' \sim_E v$, on a

$$t \cdot (v'+E) = t \cdot (v+E).$$

Or il existe $x \in E$ tel que $v' = v+x$, d'où $t \cdot v' = t \cdot v + t \cdot x$, et $t \cdot x \in E$ puisque E est un sous-espace vectoriel de V. On a donc obtenu l'assertion (1) du théorème suivant.

***Théorème 5.5.2* (Espaces vectoriels quotients)**. — *Soient k un corps, V un k-espace vectoriel et E un sous-espace vectoriel de V.*

(1) *Il existe sur le groupe abélien V/E une unique structure d'espace vectoriel telle que la projection $\pi : V \to V/E$ soit linéaire: pour tout $t \in k$ et $v \in V$, on a $t \cdot \overline{v} = \overline{t \cdot v}$. On dit que V/E est l'espace vectoriel quotient de V par E.*

(2) *Si V est de dimension finie,* $\boxed{\dim(V) = \dim(E) + \dim(V/E)}$ *i.e.* $\boxed{\dim(V/E) = \dim(V) - \dim(E).}$

(3) *Supposons V de dimension finie, et soit F un supplémentaire de E dans V* (cf. 4.1.5). *Alors π induit un* isomorphisme $F \xrightarrow{\sim} V/E$. *En particulier, si $\mathscr{C} = (f_1, \dots, f_d)$ est une base de F, où $d =$* $\dim(V) - \dim(E)$, *alors $\pi(\mathscr{C}) = (\pi(f_1), \dots, \pi(f_d))$ est une base de V/E.*

Démonstration. — On a déjà vu l'assertion (1), démontrons (2). L'application linéaire $\pi : V \to V/E$ est surjective, et son noyau est E. Donc, si V est de dimension finie on a, d'après le théorème du rang :

$$\dim(V) = \dim(E) + \dim(V/E).$$

(3) Notons $\pi_F : F \to V/E$ la restriction de π à F. D'abord, $\mathrm{Ker}(\pi_F) = F \cap E = \{0\}$ donc π_F est injectif. D'autre part, tout $x \in V/E$ égale $\pi(v)$, pour un certain $v \in V$. Comme $V = E \oplus F$, on peut écrire $x = e+f$ avec $e \in E$ et $f \in F$, et comme $\pi(e) = 0$ on obtient que $x = \pi(f) = \pi_F(f)$. Ceci montre que $\pi_F : F \to V/E$ est aussi surjectif, donc c'est un isomorphisme. Par conséquent, l'image par π de toute base de F est une base de V/E. $\square$

Le quotient V/E a la “propriété universelle” suivante, relativement aux applications linéaires $f : V \to W$ qui s'annulent sur E, c.-à-d., pour une telle application f, on peut remplacer V par l'espace “plus simple” (car de dimension plus petite) V/E:

***Théorème 5.5.3* (Passage au quotient d'un homomorphisme).** — *Soient V, W des espaces vectoriels, E un sous-espace vectoriel de V, π la projection $V \to V/E$, et $f : V \to W$ une application linéaire telle que $E \subseteq \mathrm{Ker}(f)$. Alors l'application*

$$V/E \to W, \qquad \pi(v) \mapsto f(v)$$

est bien définie et est l'unique application linéaire $\overline{f} : V/E \to W$ telle que $\overline{f} \circ \pi = f$.

Démonstration. — Si l'on a une application linéaire $\overline{f} : V/E \to W$ telle que $\overline{f} \circ \pi = f$ alors pour tout $v \in V$ on a nécessairement

$$(*) \qquad \overline{f}(\pi(v)) = f(v).$$

Réciproquement, montrons que cette formule définit bien une application de V/E dans W. Il faut montrer que si $\pi(v) = \pi(v')$, c.-à-d., si $v - v' \in E$, alors $f(v) = f(v')$. Mais ceci est clair puisque $E \subseteq \mathrm{Ker}(f)$. Donc $\overline{f}$ est bien définie. Alors, pour tout $v, v' \in V$ et $t \in k$, on a:

$$\overline{f}\big(t \cdot \pi(v) + \pi(v')\big) = \overline{f}\big(\pi(t \cdot v + v')\big) = f(t \cdot v + v') = t \cdot f(v) + f(v') = t \cdot \overline{f}(v) + \overline{f}(v')$$

ce qui montre que $\overline{f}$ est linéaire. □

***Théorème 5.5.4* (Théorème d'isomorphisme de Noether).** — *Soient V, W des espaces vectoriels, $f : V \to W$ une application linéaire. Alors f induit un isomorphisme d'espaces vectoriels*

$$\overline{f} : V/\mathrm{Ker}(f) \xrightarrow{\sim} \mathrm{Im}(f).$$

Démonstration. — On peut considérer f comme un homomorphisme $f : V \to \mathrm{Im}(f)$, qui est alors surjectif. D'après le théorème 5.5.3, il existe donc un (unique) homomorphisme $\overline{f} : V/\mathrm{Ker}(f) \to \mathrm{Im}(f)$ tel que $\overline{f}(\overline{v}) = f(v)$ pour tout $v \in V$, et $\overline{f}$ est, comme f, surjectif.

De plus, $\overline{f}$ est injectif: en effet, soit $x \in \mathrm{Ker}(\overline{f})$ et soit $v \in V$ tel que $\overline{v} = x$, alors $0 = \overline{f}(\overline{v}) = f(v)$, donc $v \in \mathrm{Ker}(f)$, d'où $x = 0$. Donc $\overline{f} : V/\mathrm{Ker}(f) \to \mathrm{Im}(f)$ est injectif et surjectif, donc c'est un isomorphisme. □

Pour illustrer l'usage des espaces quotients, terminons ce paragraphe 5.5 en donnant une autre démonstration [(3)] du théorème de trigonalisation; on y remplace aussi l'hypothèse que le corps de base soit $\mathbb{C}$ par l'hypothèse que le polynôme caractéristique $P_u(X)$ soit scindé. Commençons par la proposition suivante, qui établit un lien entre espaces quotients et matrices triangulaires par blocs.

***Proposition 5.5.5* (Quotients et matrices triangulaires par blocs)**

Soient V un k-espace vectoriel de dimension n, $u \in \mathrm{End}_k(V)$, et E un sous-espace de V stable par u. Soit $u_E : E \to E$ la restriction de u à E et soit π la projection $V \to V/E$.

(1) *u induit un endomorphisme $\overline{u} = u_{V/E}$ de V/E, tel que $u_{V/E}(\pi(v)) = \pi(u(v))$ pour tout $v \in V$.*

(2) *Soit $\mathscr{B} = (e_1, \dots, e_n)$ une base de V telle que $\mathscr{C} = (e_1, \dots, e_r)$ soit une base de E. Alors $\mathrm{Mat}_{\mathscr{B}}(u)$ est triangulaire par blocs:*

$$\mathrm{Mat}_{\mathscr{B}}(u) = \left(\begin{array}{c|c} A & B \\ \hline \mathbf{0}_{n-r,r} & D \end{array}\right) \qquad \textit{avec} \qquad \begin{cases} A = \mathrm{Mat}_{\mathscr{C}}(u_E) \\ B \in M_{r,n-r}(k), \quad D \in M_{n-r}(k). \end{cases}$$

De plus, D est la matrice de $u_{V/E}$ dans la base $\mathscr{D} = (\pi(e_{r+1}), \dots, \pi(e_n))$ de V/E.

(3) *Par conséquent, on a*

$$(*) \qquad P_u(X) = P_{u_E}(X) \cdot P_{u_{V/E}}(X).$$

[(3)] en quelque sorte “duale” de celle de 4.2.3

Démonstration. — (1) Comme $u(E) \subseteq E$, on a $(\pi \circ u)(E) = \{0\}$ donc l'existence de $\overline{u} = u_{V/E}$ résulte du théorème 5.5.3 appliqué à $f = \pi \circ u : V \to V/E$.

(2) Comme $E = \mathrm{Vect}(e_1, \dots, e_r)$ est stable par u, il est clair que $M = \mathrm{Mat}_{\mathscr{B}}(u)$ a la forme indiquée. De plus, écrivant $B = (b_{ij})$ et $D = (d_{\ell j})$ pour $i = 1, \dots, r$ et $j, \ell = 1, \dots, n-r$, on a:

$$u(e_{r+j}) = \sum_{i=1}^{r} b_{ij}\, e_i + \sum_{\ell=1}^{r} d_{\ell j}\, e_{r+\ell}$$

donc $\overline{u}(\pi(e_{r+j})) = \pi(u(e_{r+j}))$ égale $\sum_{\ell=1}^{r} d_{\ell j}\pi(e_{r+\ell})$ pour $j = 1, \dots, n-r$, ce qui prouve que D est la matrice de $\overline{u}$ dans la base $\mathscr{D} = (\pi(e_{r+1}), \dots, \pi(e_n))$ de V/E (cf. point (3) de 5.5.2). Alors, comme

$$M - XI_n = \left(\begin{array}{c|c} A - XI_r & B \\ \hline \mathbf{0}_{n-r,r} & D - XI_{n-r} \end{array}\right)$$

on déduit de 3.4.8 que $P_u(X) = \mathrm{dét}(A - XI_r) \cdot \mathrm{dét}(D - XI_{n-r})$ d'où $P_u(X) = P_{u_E}(X) \cdot P_{u_{V/E}}(X)$. La proposition est démontrée. □

Théorème 5.5.6 **(Trigonalisation lorsque $P_u(X)$ est scindé).** — *Soient V un k-espace vectoriel de dimension n et $u \in \mathrm{End}_k(V)$, on suppose que $P_u(X)$ est scindé, i.e. qu'il a n racines $\lambda_1, \dots, \lambda_n$ (pas nécessairement distinctes) dans k. Alors il existe une base $\mathscr{B}$ de V dans laquelle la matrice de u est triangulaire supérieure, les coefficients diagonaux étant alors $\lambda_1, \dots, \lambda_n$.*

En termes matriciels: si le polynôme caractéristique de $A \in M_n(k)$ a toutes ses racines dans k, alors A est semblable à une matrice triangulaire, i.e. il existe $P \in \mathrm{GL}_n(k)$ telle que $P^{-1}AP$ soit triangulaire.

Démonstration. — On procède par récurrence sur n, il n'y a rien à montrer si $n = 1$. Supposons donc $n \geq 2$ et le théorème établi pour $n-1$. Puisque, par hypothèse, $P_u(X)$ est scindé, alors u possède au moins un vecteur propre e_1 associé à la valeur propre λ_1. La droite ke_1 est stable par u, donc d'après le corollaire 5.5.5, u induit un endomorphisme $\overline{u}$ du quotient $\overline{V} = V/ke_1$, qui est de dimension $n-1$. D'après la formule $(*)$ de 5.5.5, on a

$$P_u(X) = (\lambda_1 - X)P_{\overline{u}}(X)$$

et donc $P_{\overline{u}}(X)$ égale $(-1)^{n-1}\prod_{i=2}^{n}(X - \lambda_i)$, donc est scindé.

Par hypothèse de récurrence, il existe une base $\mathscr{D} = (v_2, \dots, v_n)$ de $\overline{V}$ telle que $D = \mathrm{Mat}_{\mathscr{D}}(\overline{u})$ soit triangulaire supérieure. Soient $e_2, \dots, e_n \in V$ dont les images dans V/ke_1 sont $v_2, \dots, v_n$. Alors $\mathscr{B} = (e_1, e_2, \dots, e_n)$ est une base de V (cf. la preuve du théorème du rang 0.3.2) et, d'après 5.5.5, la matrice M de u dans $\mathscr{B}$ est de la forme:

$$\left(\begin{array}{c|c} \lambda_1 & B \\ \hline \mathbf{0}_{n-1,1} & D \end{array}\right)$$

avec $B \in M_{1,n-1}(k)$, donc est triangulaire supérieure. Le théorème est démontré. □

5.6. Appendice (†): Normes sur $\mathbb{K}^n$ et produits de séries absolument convergentes

Soit $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Commençons par démontrer la proposition 5.3.3 :

Proposition 5.6.1. — *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie. Alors toutes les normes sur E sont* équivalentes, *c.-à-d., si $\|\cdot\|$ et $\|\cdot\|'$ sont deux normes sur E, il existe des constantes $c, C \in \mathbb{R}_+^*$ telles que:*

$$(\dagger) \qquad \forall v \in E, \qquad c \cdot \|v\| \leq \|v\|' \leq C \cdot \|v\|.$$

Démonstration. — Soit $\mathscr{B} = (e_1, \dots, e_d)$ une base de E et soit $\|\cdot\|_\infty$ la norme sur E définie par $\|v\|_\infty = \mathrm{Max}_i\, |x_i|$ si $v = \sum_i x_i e_i$. Il suffit de montrer qu'il existe des constantes $c, C \in \mathbb{R}_+^*$ telles que:

$$(\star) \qquad \forall v \in E, \qquad c \cdot \|v\|_\infty \leq \|v\| \leq C \cdot \|v\|_\infty.$$

En effet, on aura de même $c' \cdot \|v\|_\infty \leq \|v\|' \leq C' \cdot \|v\|_\infty$ pour des constantes $c', C' > 0$, et donc:

$$\forall v \in E, \qquad \frac{c'}{C} \cdot \|v\| \leq \|v\|' \leq \frac{C'}{c} \cdot \|v\|.$$

La seconde inégalité de $(\star)$ s'obtient facilement: posant $C = \sum_{i=1}^{d} \|e_i\|$, on a

$$(*) \qquad \|\sum_{i=1}^{d} x_i e_i\| \leq \sum_{i=1}^{d} |x_i| \cdot \|e_i\| \leq C \cdot \|v\|_\infty.$$

Montrons maintenant qu'il existe $m > 0$ tel que, pour tout $v \in E$, on ait

$$(\ddagger) \qquad \|v\| = 1 \quad \Rightarrow \quad \|v\|_\infty \leq m.$$

Dans le cas contraire, il existerait une suite $(v_p)_{p\in\mathbb{N}}$ de vecteurs tels que $\|v_p\| = 1$ et que $t_p = \|v_p\|_\infty$ tende vers $+\infty$. Quitte à remplacer $(v_p)_{p\in\mathbb{N}}$ par une sous-suite et à permuter les coordonnées $(x_1, \ldots, x_d)$, on peut supposer que

$$\frac{1}{t_p} v_p = (1, x_{p,2}, \ldots, x_{p,d})$$

avec $|x_{p,i}| \leq 1$. Comme le "cube unité" de E pour la norme $\|\cdot\|_\infty$ est compact, on peut supposer, quitte à remplacer $(v_p)_{p\in\mathbb{N}}$ par une sous-suite, que la suite $(v_p)_{p\in\mathbb{N}}$ converge pour la norme $\|\cdot\|_\infty$ vers une limite ℓ. Comme la fonction "1ère coordonnée" est continue pour la norme $\|\cdot\|_\infty$ (c'est clair), et comme $\|\cdot\|$ l'est aussi (d'après $(*)$), alors la 1ère coordonnée de ℓ est 1 et d'autre part on a $\|\ell\| = \lim_{p\to+\infty} \dfrac{1}{t_p} = 0$, donc $\ell = 0$, d'où une contradiction. Ceci prouve $(\ddagger)$. Alors, pour tout $v \neq 0$, appliquant $(\ddagger)$ à $v' = \dfrac{1}{\|v\|} v$, on obtient que $\|v\|_\infty \leq m \cdot \|v\|$, inégalité qui est encore valable pour $v = 0$. On a donc:

$$\forall v \in E, \qquad \frac{1}{m} \cdot \|v\|_\infty \leq \|v\|,$$

ce qui achève la démonstration. □

Démontrons maintenant la proposition suivante, utilisée dans la démonstration de la proposition 5.3.13.

Proposition 5.6.2. — *Soient $\sum_{i=0}^\infty A_i$ et $\sum_{j=0}^\infty B_j$ deux séries convergentes dans $M_d(\mathbb{K})$. Supposons de plus que $\sum_{i=0}^\infty A_i$ soit* absolument convergente, *i.e. que la série $\sum_{i=0}^\infty |||A_i|||$ converge.*

Alors, posant $C_p = \sum_{i=0}^p A_i B_{p-i}$, la série $\sum_{p=0}^\infty C_p$ converge vers le produit $\big(\sum_{i=0}^\infty A_i\big)\big(\sum_{j=0}^\infty B_j\big)$.

Démonstration. — Notons (S_n) (resp. (T_n)) la suite des sommes partielles $\sum_{i=0}^n A_i$ (resp. $\sum_{j=0}^n B_j$) et S (resp. T) sa limite. Comme

$$|||ST - S_nT||| \leq |||S - S_n||| \cdot |||T|||,$$

la suite S_nT converge vers ST. Posant

$$U_n = \sum_{p=0}^{2n} C_p = \sum_{p=0}^{2n} \sum_{i=0}^{p} A_i B_{p-i} = \sum_{i=0}^{2n} A_i \big(\sum_{j=0}^{2n-i} B_j \big),$$

il suffit donc de montrer que la suite $S_nT - U_n$ tend vers 0. Or on a

$$\begin{aligned} |||S_nT - U_n||| &\leq \sum_{i=0}^{n} |||A_i\big(T - \sum_{j=0}^{2n-i} B_j\big)||| + \sum_{i=n+1}^{2n} |||A_i\big(\sum_{j=0}^{2n-i} B_j\big)||| \\ &\leq \underbrace{\sum_{i=0}^{n} |||A_i||| \cdot |||T - \sum_{j=0}^{2n-i} B_j|||}_{M_n} + \underbrace{\sum_{i=n+1}^{2n} |||A_i||| \cdot |||\sum_{j=0}^{2n-i} B_j|||}_{M'_n}. \end{aligned}$$

Posons $a = \sum_{i=0}^\infty |||A_i|||$ et remarquons que comme la suite T_n converge vers T, la suite des normes $|||T_n|||$ est bornée par un réel $C > 0$. Alors, pour tout $\varepsilon > 0$, il existe n_0 tel que $|||T - \sum_{j=0}^n B_j||| < \dfrac{\varepsilon}{2a}$ et $\sum_{i=n+1}^{2n} |||A_i||| < \dfrac{\varepsilon}{2C}$ pour tout $n \geq n_0$, d'où $M_n + M'_n < \varepsilon$ pour tout $n \geq n_0$. Ceci prouve que $|||S_nT - U_n|||$ tend vers 0, d'où la proposition. □

INDEX