Chapitre 6

Calcul matriciel

Sommaire

6.1 Matrices et applications linéaires 1
6.2 Produits de matrices 4
6.3 Déterminants 8
6.4 Diagonalisation et vecteurs propres 15
6.5 Polynômes d'endomorphismes 19
6.6 Théorème de Jordan 23

Dans un espace vectoriel de dimension finie, une fois qu'une base est choisie, chaque vecteur est défini de façon unique par ses coordonnées. Les opérations usuelles d'addition et de multiplication par un scalaire peuvent alors se calculer sur les coordonnées. Plus généralement, une fonction entre espaces vectoriels peut alors être représentée par une fonction qui aux coordonnées d'un vecteur associe les coordonnées de son image.

6.1 Matrices et applications linéaires

6.1 **Définition (matrice).** Une *matrice* à coefficients dans un ensemble X est une famille $A = (a_{i,j})_{1\leq i\leq n, 1\leq j\leq p}$ d'éléments de X, où n est un entier appelé *nombre de lignes* de A et p un entier appelé *nombre de colonnes* de A. L'ensemble des matrices à n lignes et p colonnes à coefficients dans X est noté $M_{n,p}(X)$. Une matrice est dite *carrée* si elle a autant de lignes que de colonnes ; l'ensemble des matrices carrées de tailles n peut être noté $M_n(\mathbb{K})$ au lieu de $M_{n,n}(\mathbb{K})$.

Une matrice à coefficients dans X n'est donc rien d'autre qu'un tableau d'éléments de X. C'est une généralisation à deux dimensions de l'idée de n-uplet, qui désigne une famille d'éléments numérotés par un indice : ici on numérote simplement par deux indices. On emploie la plupart du temps la convention (utilisée dans la définition) selon laquelle une matrice est désignée par une lettre majuscule comme A et ses coefficients par la lettre minuscule avec deux indices $a_{i,j}$. On peut parfois se passer de virgule entre les indices et noter a_{ij} si la notation est claire.

Visuellement, on représente une matrice par un tableau entre parenthèses, avec une case de tableau par coefficient. Le premier indice correspond au numéro de la ligne en

partant du haut et le second indice correspond au numéro de la case sur sa ligne en partant de la gauche, en numérotant à partir de 1 dans les deux cas.

6.2 **Proposition (espace des matrices).** *Pour tout corps $\mathbb{K}$ et tous entiers n et p, l'ensemble $M_{n,p}(\mathbb{K})$, muni de l'addition et de la multiplication par un élément de $\mathbb{K}$ coefficient par coefficient, est un espace vectoriel sur $\mathbb{K}$ de dimension np.*

Démonstration. Il est clair que l'addition coefficient par coefficient est une loi de composition interne et qu'elle est associative et commutative, puisque l'addition de $\mathbb{K}$ a ces propriétés. Son élément neutre est la matrice dont tous les coefficients sont nuls, on la notera donc 0. Étant donné un élément λ de $\mathbb{K}$, l'opération qui multiplie chaque coefficient d'une matrice par λ est une action de $\mathbb{K}$ sur $M_{n,p}(\mathbb{K})$ qui vérifie toutes les propriétés voulues par la définition d'un espace vectoriel, du fait des propriétés de la multiplication dans $\mathbb{K}$ (associativité et élément neutre). Ainsi $M_{n,p}(\mathbb{K})$ est un espace vectoriel sur $\mathbb{K}$.

Pour établir qu'il est de dimension finie np, considérons la famille de matrices $(E_{i,j})$ telle que chaque $E_{i,j}$ est la matrice qui a un coefficient 1 en position (i,j) et des coefficients 0 partout ailleurs. Alors il est clair que cette famille est libre et génératrice puisqu'une combinaison linéaire $\sum \lambda_{i,j} E_{i,j}$ définit exactement la matrice dons les $\lambda_{i,j}$ sont les coefficients. C'est donc une base de $M_{n,p}(\mathbb{K})$ et elle a np éléments. □

6.3 **Définition (matrice d'une application linéaire).** Soit $f : E \to F$ une application linéaire entre deux espaces vectoriels de dimensions finies. Soient $\mathcal{B} = (e_1, \dots, e_p)$ une base de E et $\mathcal{C} = (f_1, \dots, f_n)$ une base de F. La matrice de f relative aux bases $\mathcal{B}$ et $\mathcal{C}$ est la matrice à n lignes et p colonnes dont la j-ème colonne, pour chaque $1 \leq j \leq n$, est formée des n coordonnées de $f(e_j)$ dans la base $\mathcal{C}$. On la note $M_{\mathcal{B},\mathcal{C}}(f)$, ou $M(f)$ si le choix des bases $\mathcal{B}$ et $\mathcal{C}$ est clair dans le contexte.

6.4 *Exemple.* Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ la rotation d'angle $\pi/2$ centrée à l'origine. Son expression dans la base canonique $e_1 = (1,0), e_2 = (0,1)$ se détermine ainsi. La première colonne est donnée par les coordonnées dans la base canonique de $f(e_1) = e_2 = (0,1)$, la première colonne est donc $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. La seconde colonne est donnée par les coordonnées dans la base canonique de $f(e_2) = -e_1 = (-1,0)$, c'est donc $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$. La matrice est donc la juxtaposition de ces deux colonnes :

$$M(f) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Plus généralement, si f est une rotation d'angle θ, par le même raisonnement on obtient

$$M(f) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

6.5 *Exemple.* Soit $f : \mathbb{R}^n \to \mathbb{R}$, $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ l'application telle que $f(x_1, \dots, x_n) = a_1 x_1 + \cdots + a_n x_n$, pour une certaine famille $a_1, \dots, a_n \in \mathbb{R}$. La matrice de f dans les bases canoniques est

$$M(f) = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}.$$

6.6 *Exemple.* Soit $f : \mathbb{R} \to \mathbb{R}^n$ l'application telle que $f(t) = (a_1 t, \dots, a_n t)$ pour tout t, pour une certaine $a_1, \dots, a_n \in \mathbb{R}$. La matrice de f dans les bases canoniques est

$$M(f) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

6.7 **Théorème.** *Pour tous espaces vectoriels E et F de dimension finie et toutes bases $\mathcal{B}$ de E et $\mathcal{C}$ de F, la fonction $M_{\mathcal{B},\mathcal{C}}$ est un isomorphisme.*

Démonstration. Pour simplifier les notations, comme les bases $\mathcal{B}$ et $\mathcal{C}$ sont fixées, on se contentera de noter $M(f)$ la matrice d'un application linéaire f donnée, et on notera $m_{i,j}(f)$ le coefficient à la ligne i et à la colonne j de cette matrice. La première chose à montrer est que la fonction M est linéaire, pour cela il suffit de montrer que chaque fonction $m_{i,j}$ est linéaire. Considérons alors deux fonctions $f, g \in \mathcal{L}(E, F)$ et un coefficient $\lambda \in \mathbb{K}$. Pour chaque j on a $(f + \lambda g)(e_j) = f(e_j) + \lambda g(e_j)$ par définition, et considérant la i-ème coordonnée de ce vecteur on obtient $m_{i,j}(f + \lambda g) = m_{i,j}(f) + \lambda m_{i,j}(g)$, ce qui montre bien la linéarité de $m_{i,j}$.

Pour montrer que M est un isomorphisme, on montre que c'est une fonction injective et surjective. Pour l'injectivité, il suffit de montrer que son noyau est réduit à 0, considérons donc une application linéaire f telle que $M(f) = 0$. Alors par définition pour tout j on a $f(e_j) = 0$, donc pour tout vecteur $x \in E$, en notant $x = x_1 e_1 + \cdots + x_n e_n$ on obtient $f(x) = x_1 f(e_1) + \cdots + x_n f(e_n) = x_1 0 + \cdots + x_n 0 = 0$, ainsi f est bien la fonction nulle. Pour la surjectivité, considérons une matrice $A \in M_{n,p}(\mathbb{K})$ quelconque. Par le théorème **??**, on sait qu'il existe une application linéaire f telle que pour chaque j on ait $f(e_j) = a_{1,j} f_1 + \cdots + a_{n,j} f_n$ et par construction cette fonction a bien pour matrice A. □

6.8 **Corollaire.** *Pour tous espaces vectoriels E et F de dimensions finies, $\mathcal{L}(E, F)$ est de dimension finie et* $\dim \mathcal{L}(E, F) = \dim E \times \dim F$.

6.9 Un cas particulier important de matrices est le cas où il n'y a qu'une seule colonne, c'est-à-dire le cas des espaces $M_{n,1}(\mathbb{K})$. On emploie ce genre de matrice en particulier pour représenter des vecteurs : si $\mathcal{B} = (e_1, \dots, e_n)$ est une base d'un espace vectoriel E, alors un vecteur $x = x_1 e_1 + \cdots + x_n e_n$ sera représenté par la matrice

$$X = M_{\mathcal{B}}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Une telle matrice est souvent appelée *vecteur colonne*.

6.10 Soit $f : E \to F$ une application linéaire et soient $\mathcal{B} = (e_1, \dots, e_p)$ et $\mathcal{C} = (f_1, \dots, f_n)$ des bases de E et F respectivement. Posons $A = M_{\mathcal{B},\mathcal{C}}(f)$ et notons $a_{i,j}$ les coefficients de A. Alors l'image d'un vecteur $x = x_1 e_1 + \cdots + x_p e_p$ par f est $x_1 f(e_1) + \cdots + x_p f(e_p)$, on peut en déduire ses coordonnées dans $\mathcal{C}$: la coordonnée de $f(x)$ selon le vecteur f_i est $a_{i,1} x_1 + \cdots + a_{i,p} x_p$. On représente ce calcul sous forme d'une multiplication :

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{1,1} x_1 + \cdots + a_{1,p} x_p \\ \vdots \\ a_{n,1} x_1 + \cdots + a_{n,p} x_p \end{pmatrix}.$$

Cette formule est un cas particulier de la notion plus générale de produit de matrices que l'on verra plus loin.

6.11 *Exemple.* Reprenons la matrice d'une rotation du plan d'angle θ dans la base canonique. On obtient

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos\theta \cdot x_1 - \sin\theta \cdot x_2 \\ \sin\theta \cdot x_1 + \cos\theta \cdot x_2 \end{pmatrix}.$$

6.12 *Exemple.* Pour une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}$ on obtient :

$$\begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_1 x_1 + \cdots + a_n x_n \end{pmatrix}.$$

6.13 Par cette opération, une matrice $A \in M_{n,p}(\mathbb{K})$ peut donc être considérée comme une application linéaire de $M_{p,1}(\mathbb{K})$ (vecteurs colonnes à p coordonnées) dans $M_{n,1}(\mathbb{K})$ (vecteurs colonnes n coordonnées). Les notions qui concernent les applications linéaires concernent donc aussi bien les matrices, ainsi on pourra considérer utilement le noyau d'une matrice, son image, son rang, etc.

6.2 Produits de matrices

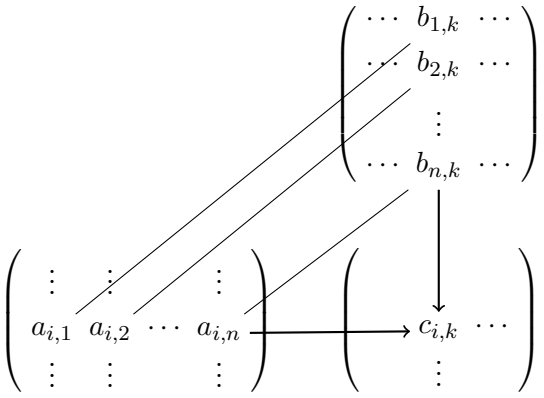
Dans la définition de la structure d'espace vectoriel de $M_{n,p}$, on définit une multiplication par un scalaire qui consiste simplement à multiplier chaque coefficient de la matrice par ce scalaire. On pourrait définir une opération de multiplication entre matrices coefficient par coefficient sur le modèle de l'addition mais il s'avère que la notion de multiplication la plus utile n'est pas celle-là.

6.14 **Définition (produit de matrices).** Soient trois entiers m, n, p et un corps $\mathbb{K}$. Pour toutes matrices $A \in M_{m,n}(\mathbb{K})$ et $B \in M_{n,p}(\mathbb{K})$ on définit la matrice $AB \in M_{m,p}(\mathbb{K})$ par

$$AB = (c_{i,k})_{1 \leq i \leq m, 1 \leq k \leq p} \qquad \text{où} \qquad c_{i,k} = \sum_{j=1}^{n} a_{i,j} b_{j,k}.$$

Cette opération de $M_{m,n}(\mathbb{K}) \times M_{n,p}(\mathbb{K})$ dans $M_{m,p}(\mathbb{K})$ est appelée *produit de matrices*.

Pour visualiser cette opération, il est utile de montrer comment on peut poser la multiplication des matrices :



$$c_{i,k} = a_{i,1} b_{1,k} + a_{i,2} b_{2,k} + \cdots + a_{i,n} b_{n,k}$$

Le coefficient en position (i,k) dans le produit est obtenu en prenant la ligne d'indice i dans A et la colonne d'indice k dans B, en multipliant les coefficients dont les indices correspondent, et en additionnant les résultats. Il faut donc que le nombre de colonnes de A soit égal au nombre de lignes de B.

6.15 **Théorème.** *Soient $f : E \to F$ et $g : F \to G$ deux applications linéaires, soient $\mathcal{B}$, $\mathcal{C}$ et $\mathcal{D}$ des bases de E, F et G respectivement. Alors on a $M_{\mathcal{B},\mathcal{D}}(g \circ f) = M_{\mathcal{C},\mathcal{D}}(g) M_{\mathcal{B},\mathcal{C}}(f)$.*

Démonstration. Notons $\mathcal{B} = (e_1, \dots, e_p)$, $\mathcal{C} = (f_1, \dots, f_q)$ et $\mathcal{D} = (g_1, \dots, g_r)$. Notons $a_{i,j}$ le coefficient en ligne i et colonne j de $M_{\mathcal{B},\mathcal{C}}(f)$, $b_{i,j}$ le coefficient en i,j dans $M_{\mathcal{C},\mathcal{D}}(g)$ et $c_{i,j}$ le coefficient en i,j dans $M_{\mathcal{B},\mathcal{D}}(g \circ f)$. Pour chaque $1 \leq j \leq p$ on a $f(e_j) = \sum_{k=1}^{q} a_{kj} f_k$ donc

$$g \circ f(e_j) = \sum_{k=1}^{q} a_{kj} g(f_k) = \sum_{k=1}^{q} a_{kj} \sum_{i=1}^{r} b_{ik} g_i = \sum_{i=1}^{r} \left(\sum_{k=1}^{q} b_{ik} a_{kj} \right) g_i$$

donc $c_{ij} = \sum_{k=1}^{q} b_{ik} a_{kj}$. C'est bien la formule du produit de matrices. □

Cette propriété est la raison d'être de la définition du produit de matrices. Elle indique comment raisonner sur la composition des applications linéaires en raisonnant sur leurs coordonnées.

6.16 *Exemple.* Considérons le corps des nombres complexes $\mathbb{C}$ comme un espace vectoriel de dimension 2 sur le corps $\mathbb{R}$ des réels. Nous choisissons la base $(1, i)$. La multiplication par $a \in \mathbb{C}$ définit toujours une application linéaire m_a. Exprimons le vecteur a dans la base choisie : $a = \mathrm{Re}(a) + i\,\mathrm{Im}(a)$, où $\mathrm{Re}(a)$ est la partie réelle de a et $\mathrm{Im}(a)$ est sa partie imaginaire. La matrice de m_a dans la base $(1, i)$ est alors

$$\begin{pmatrix} \mathrm{Re}(a) & -\mathrm{Im}(a) \\ \mathrm{Im}(a) & \mathrm{Re}(a) \end{pmatrix}.$$

Si le nombre complexe u est de module 1, alors $u = \cos\theta + i \sin\theta$ pour un certain angle θ, et si $\lambda \in \mathbb{R}$, on obtient :

$$M(m_u) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad M(m_\lambda) = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix},$$

$$M(m_\lambda) M(m_u) = \begin{pmatrix} \lambda\cos\theta & -\lambda\sin\theta \\ \lambda\sin\theta & \lambda\cos\theta \end{pmatrix}.$$

Remarquons que $m_\lambda \circ m_u = m_{\lambda u}$. Ainsi nous obtenons l'interprétation géométrique de la multiplication par un nombre complexe. En effet, tout nombre complexe a s'écrit sous la forme $a = \lambda u$ où $\lambda \in \mathbb{R}$ et $u = \cos\theta + i\sin\theta$. La multiplication par $a = \lambda u$ est donc la composition de la rotation de centre 0 et d'angle θ (l'argument de a) avec l'homothétie de centre 0 et de rapport λ (le module de a).

6.17 **Exercice.** La conjugaison complexe $\tau : \mathbb{C} \to \mathbb{C}$, $\tau(z) = \bar{z}$ vérifie $\tau(w + z) = \tau(w) + \tau(z)$, mais n'est pourtant *pas* une application linéaire de $\mathbb{C}$ dans $\mathbb{C}$. Pourquoi ? Par contre si l'on considère $\mathbb{C}$ comme un espace vectoriel de dimension 2 sur $\mathbb{R}$ (et non plus comme un espace vectoriel de dimension 1 sur lui-même), l'application τ est linéaire. Démontrer ceci et trouver la matrice de τ dans la base $(1, i)$.

6.18 **Exercice.** Soit $s : \mathbb{R}^2 \to \mathbb{R}^2$ la symétrie axiale d'axe $0x$. Soit $r : \mathbb{R}^2 \to \mathbb{R}^2$ la rotation d'angle $\pi/2$ autour de zéro. Déterminer dans la base canonique les matrices $M(s \circ r)$ et $M(r \circ s)$ des applications composées $s \circ r$ et $r \circ s$.

6.19 **Exercice.** Soit $r : \mathbb{R}^3 \to \mathbb{R}^3$ la rotation d'angle π autour de l'axe $0x$. Soit $s : \mathbb{R}^3 \to \mathbb{R}^3$ la rotation d'angle π autour de l'axe $0y$. Déterminer dans la base canonique les matrices $M(r)$, $M(s)$ et $M(r \circ s)$ et vérifier que $M(s \circ r) = M(r \circ s)$.

6.20 **Exercice.** Montrer que le produit de matrices est associatif, c'est-à-dire que pour tout corps $\mathbb{K}$ et tous entiers m, n, p, q, pour toutes matrices $A \in M_{m,n}(\mathbb{K})$, $B \in M_{n,p}(\mathbb{K})$ et $C \in M_{p,q}(\mathbb{K})$ on a $(AB)C = A(BC)$.

6.21 En revanche, il est très important de retenir que le produit de matrices *n'est pas commutatif*. Cela signifie que pour deux matrices A et B données, en général AB et BA ne sont pas égales.

Bien entendu, pour se demander si AB et BA sont égales, il faut déjà savoir qu'elles sont définies toutes les deux, ce qui n'est pas forcément le cas selon les tailles de A et B. De plus, si A et B ne sont pas carrées mais que AB et BA sont définies, alors celles-ci sont carrées mais de tailles différentes. La question se pose donc principalement pour des matrices carrées de même tailles. Dans ce cas on peut trouver un exemple simple en posant

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad \text{alors on a} \quad AB = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{et} \quad BA = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

donc $AB \neq BA$.

Il y a bien sûr des cas de matrices A et B pour lesquelles on a effectivement $AB = BA$, ne serait-ce que dans des cas simples comme $A = B$ ou $B = 0$, mais ce n'est pas toujours le cas.

6.22 **Exercice.** Montrer que le produit de matrices est linéaire en des deux côtés, c'est-à-dire que pour tout corps $\mathbb{K}$ et tous entiers m, n, p,

- pour toutes matrices $A, B \in M_{m,n}(\mathbb{K})$ et $C \in M_{n,p}(\mathbb{K})$ on a $(A+B)C = AC + BC$,
- pour toutes matrices $A \in M_{m,n}(\mathbb{K})$ et $B, C \in M_{n,p}(\mathbb{K})$ on a $A(B+C) = AB + AC$,
- pour toutes matrices $A \in M_{m,n}(\mathbb{K})$ et $B \in M_{n,p}(\mathbb{K})$ et tout scalaire $\lambda \in \mathbb{K}$ on a $(\lambda A)B = A(\lambda B) = \lambda(AB)$.

6.23 Dans le cas particulier $m = n = p$, les résultats précédent expriment que le produit est une loi de composition interne sur $M_n(\mathbb{K})$, associative et distributive sur la somme et la multiplication par un scalaire. On dit alors que $M_n(\mathbb{K})$ est une *algèbre* sur $\mathbb{K}$.

6.24 **Définition (diagonale).** La *diagonale* d'une matrice A est l'ensemble des cases de A d'indices (i, i) pour toutes les valeurs possibles de i. On dit qu'un matrice est *diagonale* si tous les coefficients de cette matrice sont nuls en dehors de sa diagonale.

La notion de diagonale est surtout employée à propos des matrices carrées, où il s'agit effectivement des cases situées sur la diagonale allant du coin supérieur gauche au coin inférieur droit. On ne considère quasiment jamais l'autre diagonale parce qu'elle a très peu de propriété remarquables, contrairement à la diagonale principale.

6.25 **Définition (matrice identité).** Soit n un entier. La *matrice identité* de taille n est la matrice carrée de taille n, notée I_n (ou simplement I si n peut être déduit du contexte sans ambiguïté), qui contient des 1 sur la diagonale et des 0 ailleurs.

6.26 Il est facile de constater que la matrice identité est la matrice de l'application linéaire identité. Plus précisément, si E est un espace vectoriel de dimension n et si $\mathrm{id} : E \to E$ est l'application identité (telle que $\mathrm{id}(v) = v$ pour tout vecteur v) alors dans toute base $\mathcal{B}$ on a $M_{\mathcal{B},\mathcal{B}}(\mathrm{id}) = I_n$. Bien entendu, si on choisit deux bases différentes $\mathcal{B}$ et $\mathcal{C}$, alors $M_{\mathcal{B},\mathcal{C}}(\mathrm{id})$ n'est pas la matrice identité.

6.27 **Proposition.** *La matrice identité est* neutre *pour le produit de matrices, c'est-à-dire que*
— *pour tous m, n, pour toute matrice $A \in M_{m,n}(\mathbb{K})$ on a $AI_n = A$,*
— *pour tous n, p, pour toute matrice $B \in M_{n,p}(\mathbb{K})$ on a $I_n B = B$.*

Comme le produit de matrices est une forme de multiplication, il est naturel de se demander s'il y a une notion associée de division, c'est-à-dire d'inverse. On peut le a définir de façon formelle :

6.28 **Définition (matrice inversible).** Une matrice carrée $A \in M_n(\mathbb{K})$ est dite *inversible* s'il existe une matrice B telle que $AB = BA = I_n$. Dans ce cas la matrice B est unique, on l'appelle *inverse* de A et on la note A^{-1}.

6.29 Le fait que l'inverse soit unique peut s'établir simplement en considérant deux matrices B et C qui sont toutes les deux inverses de A. Alors par définition on a $BA = I$ et $AC = I$, alors en appliquant l'associativité du produit on peut calculer le produit BAC de deux façons différentes : d'une part on a $(BA)C = IC = C$ et d'autre part on a $B(AC) = BI = B$, par conséquent $B = C$.

6.30 Si A est la matrice représentative d'un application linéaire $f : E \to E$ dans une base $\mathcal{B}$ donnée et si A est inversible, alors on peut observer que A^{-1} représente la fonction réciproque de f, en effet si $g : E \to E$ est telle que $M_{\mathcal{B}}(g) = A^{-1}$ alors on a $M(f \circ g) = M(f)M(g) = AA^{-1} = I = M(\mathrm{id})$ donc $f \circ g = \mathrm{id}_E$, et de même on a $g \circ f = \mathrm{id}_E$, ce qui correspond bien à la définition du fait que g est réciproque de f.

Par des arguments similaires, on voit que la relation fonctionne aussi dans l'autre sens : si f est une application linéaire inversible, alors sa matrice représentative (quelles que soient les bases choisies) est inversible et son inverse est une matrice de f^{-1}.

6.31 La définition d'inversibilité ne s'applique qu'aux matrices carrées parce que si A n'est pas carrée, alors il ne peut pas exister de matrice B telle que AB et BA soient toutes les deux des matrices identité. Ce fait n'est pas évident si on considère seulement la définition du produit de matrices, mais si on s'intéresse aux applications linéaires associées alors cela peut se justifier par exemple en raisonnant sur les rangs de A et B.

6.32 On peut se demander comment exprimer les coefficients de la matrice A^{-1} en fonction de ceux de A. Dans le cas d'une matrice carrée de taille 2, il y a un critère d'inversibilité et une formule relativement simples. On peut les obtenir par exemple en posant l'équation

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} ax + bz & ay + bt \\ cx + dz & cy + dt \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

et en résolvant le système à 4 équations et 4 inconnues que l'on obtient en identifiant les deux dernières matrices (les inconnues sont x, y, z, t, les variables a, b, c, d sont des paramètres). On constate alors que le système a une solution si et seulement si $ad - bc \neq 0$ et que dans ce cas on a

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

Le nombre $ad - bc$ est appelé *déterminant* de la matrice parce qu'il sert à déterminer le statut de la matrice, inversible ou pas. Le déterminant et la formule d'inversion se généralisent en toute dimension mais les formules sont nettement plus compliquées.

6.33 **Proposition.** *Si A et B sont deux matrices inversibles de même taille alors AB est inversible et on a $(AB)^{-1} = B^{-1}A^{-1}$.*

Démonstration. Il suffit d'appliquer l'associativité du produit pour obtenir

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

et de même

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I,$$

ce qui prouve par définition que $B^{-1}A^{-1}$ est inverse de AB. □

Comme le produit des matrices n'est pas commutatif, il est important de retenir ce phénomène : lors du passage à l'inverse, l'ordre des facteurs dans un produit est renversé.

À cause du fait que l'ordre est important, on évite en général d'employer la notation des fractions dans le cas des matrices (et plus généralement dans les produits non commutatifs). On pourrait bien sûr fixer une convention, par exemple que $\frac{A}{B}$ représente toujours le produit AB^{-1}, mais on ne le fait pas parce que certaines règles très naturelles liées à cette notation, comme la règle du produit $\frac{a}{b}\frac{c}{d} = \frac{ac}{bd}$, ne s'appliquent plus.

6.34 **Proposition (formule du binôme).** *Soient A et B deux matrices qui commutent, c'est-à-dire telles que $AB = BA$. Alors pour tout entier naturel n on a*

$$(A+B)^n = \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k} \qquad \text{où } \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Cette formule se démontre exactement comme dans le cas de $(a+b)^n$ où a et b sont des nombres réels ou complexes, en effet seules les règles de calcul générales sont utilisées, dont la commutativité du produit.

6.35 **Exercice.** Trouver deux matrices A et B et un entier n pour lesquels la formule du binôme pour $(A+B)^n$ ne s'applique pas (il faudra donc avoir $AB \neq BA$).

6.3 Déterminants

6.36 **Définition (permutation).** Une *permutation* d'ordre n est une bijection de l'ensemble $\{1, \ldots, n\}$ dans lui-même. L'ensemble des permutation d'ordre n est noté S_n.

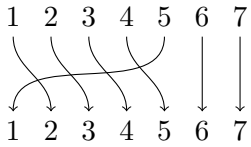
6.37 L'ensemble S_n est fini et son cardinal est $n! = n \times (n-1) \times (n-2) \times \cdots \times 3 \times 2 \times 1$, la factorielle de n.

6.38 La composition des fonctions est une loi de composition interne associative sur S_n pour laquelle la fonction identité est l'élément neutre et pour laquelle tout élément a un inverse qui est sa fonction réciproque. Il forme donc un *groupe*, que l'on appelle *groupe des permutations* d'ordre n.

6.39 **Définition (signature).** Soit σ une permutation d'ordre n. Une *inversion* de σ est un couple (i, j) tel que $i < j$ et $\sigma(j) < \sigma(i)$. On note $I(\sigma)$ le nombre d'inversions de σ. La *signature* de σ et le nombre $\operatorname{sgn}(\sigma) = (-1)^{I(\sigma)}$, c'est-à-dire 1 si $I(\sigma)$ est pair et -1 s'il est impair.

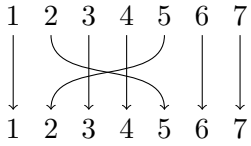
6.40 *Exemple.* La fonction identité est une permutation, par définition elle n'a aucune inversion donc sa signature est 1.

6.41 *Exemple.* Soit n un entier, soit ℓ un entier inférieur ou égal à n. Soit σ la permutation telle que $\sigma(i) = i + 1$ si $1 \leq i < \ell$, $\sigma(\ell) = 1$ et $\sigma(i) = i$ si $l < i \leq n$. Graphiquement, on peut représenter cette permutation par le diagramme suivant :



qui correspond au cas $n = 7$ et $\ell = 5$. Alors les inversions correspondent aux croisements dans le diagramme, il y en a exactement $\ell - 1$ donc la signature de σ est $(-1)^{\ell-1}$.

6.42 *Exemple.* Une *transposition* est une permutation qui fixe tous les éléments de $\{1, \ldots, n\}$ sauf deux éléments i et j distincts qui sont échangés. On note habituellement $(i\ j)$ une telle transposition. En reprenant les conventions de l'exemple précédent, la transposition $(2\ 5)$ dans S_7 est représentée par



Ce diagramme comporte 5 croisements donc la permutation a 5 inversions et la signature de $(2\ 5)$ est -1. Plus généralement, on peut observer que toute transposition est de signature -1.

6.43 **Proposition.** *Toute permutation peut s'écrire comme une composition de transpositions. En d'autres termes, le groupe S_n est* engendré *par les transpositions.*

Démonstration. On procède par récurrence sur n. Si $n = 1$, il n'y a rien a démontrer. Supposons le résultat établi pour un n donné et soit une permutation $\sigma \in S_{n+1}$. Deux cas se présentent :

1. Si $\sigma(1) = 1$, alors la restriction de σ à $\{2, \ldots, n + 1\}$ s'identifie à un élément de S_n puisque l'ensemble $\{2, \ldots, n + 1\}$ est constitué de n éléments. L'hypothèse de récurrence implique dans ce cas que σ est un produit de transpositions.
2. Si $\sigma(1) > 1$, soit $\tau = (1\ \sigma(1))$ la transposition de $n + 1$ qui échange 1 et $\sigma(1)$. Alors $\tau \circ \sigma$ laisse 1 invariant, et par le raisonnement précédent on déduit que $\tau \circ \sigma$ est une composition de transpositions. C'est donc le cas aussi de τ puisqu'on a $\sigma = \tau \circ \tau \circ \sigma$. □

6.44 **Proposition.** *Pour toutes permutations $\sigma, \sigma' \in S_n$ on a* $\operatorname{sgn}(\sigma \circ \sigma') = \operatorname{sgn}(\sigma)\operatorname{sgn}(\sigma')$. *En d'autres termes,* sgn *est un morphisme de groupes de S_n dans $\{-1, 1\}$ muni de la multiplication.*

Démonstration. On commence par remarquer que pour toute signature σ on a l'égalité

$$\operatorname{sgn}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}$$

En effet, pour tout couple (i, j) avec $i < j$, la différence $\sigma(j) - \sigma(i)$ est négative si et seulement si (i, j) est une inversion de σ et la différence $j - i$ est toujours positive, donc le signe du produit est bien celui de $\operatorname{sgn}(\sigma)$. Pour ce qui est de la valeur absolue, on remarque que si (i, j) atteint chaque paire exactement une fois, alors $(\sigma(i), \sigma(j))$ aussi puisque σ est une bijection, donc $\Pi_{1 \leq i < j \leq n} |\sigma(j) - \sigma(i)| = \Pi_{1 \leq i < j \leq n} |j - i|$.

En appliquant cette égalité à $\sigma \circ \sigma'$ on peut écrire

$$\begin{aligned}
\operatorname{sgn}(\sigma \circ \sigma') &= \prod_{1 \leq i < j \leq n} \frac{\sigma(\sigma'(j)) - \sigma(\sigma'(i))}{j - i} = \prod_{1 \leq i < j \leq n} \left(\frac{\sigma(\sigma'(j)) - \sigma(\sigma'(i))}{\sigma'(j) - \sigma'(i)} \times \frac{\sigma'(j) - \sigma'(i)}{j - i} \right) \\
&= \prod_{1 \leq i < j \leq n} \frac{\sigma(\sigma'(j)) - \sigma(\sigma'(i))}{\sigma'(j) - \sigma'(i)} \times \prod_{1 \leq i < j \leq n} \frac{\sigma'(j) - \sigma'(i)}{j - i} \\
&= \prod_{1 \leq i' < j' \leq n} \frac{\sigma(j') - \sigma(i')}{j' - i'} \times \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i} = \operatorname{sgn}(\sigma)\operatorname{sgn}(\sigma')
\end{aligned}$$

en observant que le changement de variable entre (i, j) et (i', j') est valide parce que le quotient $(\sigma'(\sigma(i)) - \sigma'(\sigma(j)))/(\sigma(i) - \sigma(j))$ est préservé si l'on échange i et j, ce qui peut être nécessaire pour obtenir $i' < j'$. □

6.45 **Exercice.** Démontrer que pour toute permutation σ on a $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$.

6.46 **Définition (déterminant d'une matrice).** Soit $A \in M_n(\mathbb{K})$ une matrice carrée. Le *déterminant* de A est

$$\det(A) := \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$

La notation $\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$ désigne le déterminant $\det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$.

Il s'agit donc d'une somme de $n!$ termes qui sont toutes les façons de choisir un coefficient dans chaque ligne en n'en choisissant jamais deux dans u ne même colonne. Pour chaque choix, on fait le produit des coefficients choisis et on compte le terme en positif ou en négatif selon la signature de la permutation considérée.

6.47 Dans le cas d'une matrice de taille 2, on retrouve la formule vue précédemment dans le calcul de l'inverse. En effet, il y a deux permutations dans S_2 : l'identité et la transposition $(1\ 2)$, de signatures respectives 1 et -1. Le résultat est donc $a_{11}a_{22} - a_{12}a_{21}$.

6.48 On peut écrire une formule explicite dans le cas d'une matrice 3×3. Le groupe S_3 a 6 éléments donc la somme a 6 termes :

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} + a_{21}a_{32}a_{13} - a_{21}a_{12}a_{33} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13}.$$

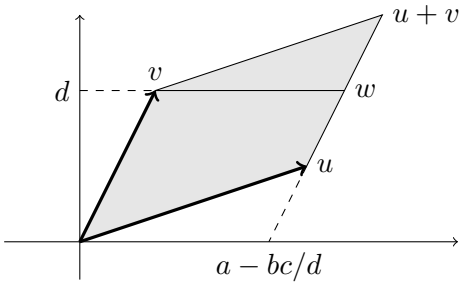
Cependant il y a peut d'intérêt à retenir cette formule. Au delà de la taille 3, il y a d'autant moins d'intérêt à écrire une formule développée dans le cas général. On verra dans la suite qu'il y a d'autres façons plus efficaces de calculer des déterminants.

6.49 Si la matrice A est diagonale, alors $\det(A)$ est le produit des coefficients de la diagonale. En effet dans ce cas tous les termes de la somme qui définit le déterminant contiennent un coefficient $a_{i,\sigma(i)}$ qui est nul sauf dans le cas où σ est la permutation identité. En conséquence on a en particulier $\det(I_n) = 1$ et $\det(\lambda I_n) = \lambda^n$ pour tout λ.

Pour établir les propriétés du déterminant, il est utile d'étendre la définition à d'autres objets, notamment des familles de vecteurs et des applications linéaires.

6.50 **Définition (déterminant d'une famille de vecteurs).** Soit E un espace vectoriel de dimension n et soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E. On définit le déterminant dans $\mathcal{B}$ d'une famille de vecteurs $(v_1, \dots, v_n)$ comme $\det_{\mathcal{B}}(v_1, \dots, v_n) = \det(A)$ où A est la matrice $n \times n$ dont les colonnes sont les coordonnées de $v_1, \dots, v_n$ dans $\mathcal{B}$, c'est-à-dire $v_j = a_{1j}e_1 + \cdots + a_{nj}e_n$ pour tout j.

6.51 On peut donner une interprétation instructive du déterminant d'une famille de vecteurs en géométrie. En dimension 2, si on se place dans un repère orthonormé et si on considère deux vecteurs $u = (a, c)$ et $v = (b, d)$, alors on a la situation suivante :



Calculons l'aire du parallélogramme porté par les vecteurs u et v, dont les sommets sont donc $0, u, u + v, v$. Si on déplace le côté $(u, u + v)$ sur la même droite de sorte qu'il parte de l'axe des abscisses et arrive en un point w alors l'aire est préservée, de plus w a la même ordonnée que v c'est-à-dire d, et le point sur l'axe des abscisses doit avoir pour coordonnées $(a - bc/d, 0)$ puisqu'il est à l'intersection de l'axe d'équation $y = 0$ et de la droite qui contient $u = (a, c)$ et $u + v = (a + b, c + d)$, d'équation $b(y - c) = d(x - a)$. Le parallélogramme obtenu a pour base $a - bc/d$ et pour hauteur d donc son aire est $ad - bc$, le déterminant de (u, v) dans la base considérée.

Si on échange u et v dans le raisonnement précédent, on obtient une aire $bc - ad$ opposée de la précédente : il s'agit en fait d'une aire *algébrique*, positive si les sommets $0, u, u + v, v$ sont parcourus dans le sens direct (trigonométrique) et négative s'ils sont parcourus dans le sens indirect (horaire). Le déterminant est nul si et seulement si le parallélogramme est aplati, c'est-à-dire si les deux vecteurs sont colinéaires.

En dimension 3, le déterminant de trois vecteurs correspondra de la même façon au volume (algébrique) d'un parallélépipède porté par les trois vecteurs considérés. En dimension supérieure, l'interprétation est évidemment plus abstraite.

6.52 **Définition.** Soient E un espace vectoriel sur un corps $\mathbb{K}$ et d un entier naturel. Une *forme multilinéaire alternée* de degré d sur V est une application $f : V^d \to \mathbb{K}$ telle que

— f est linéaire en chacune de ses d variables séparément ;

— pour toute famille $(v_1, \dots, v_d) \in V^d$ et toute permutation $\sigma \in S_d$, on a l'égalité $f(v_{\sigma(1)}, \dots, v_{\sigma(d)}) = \operatorname{sgn}(\sigma) f(v_1, \dots, v_d)$.

Au lieu de « multilinéaire de degré d » on emploie souvent l'expression « d-linéaire » qui généralise l'emploi de préfixes comme dans « bilinéaire », « trilinéaire » etc.

6.53 Il est facile de vérifier, en appliquant la formule du déterminant, que la fonction $\det_{\mathcal{B}}$ est n-linéaire alternée.

6.54 Par linéarité, la seconde condition est équivalente au fait que pour toute famille $(v_1, \dots, v_d) \in V^d$, s'il existe i et j distincts tels que $v_i = v_j$ alors $f(v_1, \dots, v_d) = 0$. En effet, une forme multilinéaire alternée a forcément cette propriété (puisqu'échanger v_i et v_j change le signe alors que la famille reste inchangée).

Réciproquement, si une forme multilinéaire f a cette propriété, alors pour toute transposition τ et toute famille $(v_1, \dots, v_d)$ on a $f(v_{\tau(1)}, \dots, v_{\tau(n)}) = -f(v_1, \dots, v_n)$. En effet, dans le cas de la transposition (1 2), on a

$$\begin{aligned}
&f(v_1, v_2, v_3, \dots, v_d) + f(v_2, v_2, v_3, \dots, v_d) \\
&= f(v_1, v_1, v_3, \dots, v_d) + f(v_1, v_2, v_3, \dots, v_d) + f(v_2, v_1, v_3, \dots, v_d) + f(v_2, v_2, v_3, \dots, v_d) \\
&= f(v_1, v_1 + v_2, v_3, \dots, v_d) + f(v_2, v_1 + v_2, v_3, \dots, v_d) = f(v_1 + v_2, v_1 + v_2, v_3, \dots, v_d) = 0
\end{aligned}$$

et le raisonnement est le même pour toute transposition. On en déduit que pour toute permutation σ on doit avoir $f(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = \operatorname{sgn}(\sigma) f(v_1, \dots, v_n)$, en écrivant σ comme une composition de transpositions.

6.55 **Lemme.** *Pour toute forme n-linéaire alternée $f : E^n \to \mathbb{K}$, toute base $\mathcal{B}$ de E et toute famille $(v_1, \dots, v_n) \in E^n$ on a $f(v_1, \dots, v_n) = f(\mathcal{B}) \det_{\mathcal{B}}(v_1, \dots, v_n)$.*

Démonstration. Soit $(v_1, \dots, v_n)$ une famille de n vecteurs de E. En notant $\mathcal{B} = (e_1, \dots, e_n)$ et en appelant a_{ij} les coordonnées des vecteurs v_j dans $\mathcal{B}$ on a

$$\begin{aligned}
f(v_1, \dots, v_n) &= f\Big(\sum_{i_1=1}^{n} a_{i_1 1} e_{i_1}, \dots, \sum_{i_n=1}^{n} a_{i_n n} e_{i_n}\Big) && \text{par définition} \\
&= \sum_{i_1=1}^{n} a_{i_1 1} \cdots \sum_{i_n=1}^{n} a_{i_n n} f(e_{i_1}, \dots, e_{i_n}) && \text{par multilinéarité} \\
&= \sum_{\sigma \in S_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n} f(e_{\sigma(1)}, \dots, e_{\sigma(n)}) && \text{par alternance} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} f(e_1, \dots, e_n) && \text{par alternance} \\
&= \det_{\mathcal{B}}(v_1, \dots, v_n) f(\mathcal{B}) && \text{par définition}
\end{aligned}$$

où le premier argument par alternance vient du fait que dans tous les cas où deux indices i_p et i_q sont égaux, le déterminant de la famille $(e_{i_1}, \dots, e_{i_n})$ est nul et qu'il suffit donc de considérer les familles $(i_1, \dots, i_n)$ qui sont des bijections. □

Ce résultat signifie donc que toutes les formes n-linéaires alternées d'un espace de dimension n sont égales à une constante multiplicative près. On peut en particulier en déduire une formule de changement de base pour le déterminant. En effet, pour deux bases $\mathcal{B}$ et $\mathcal{C}$, en choisissant pour f la fonction $\det_{\mathcal{C}}$ on obtient

$$\det_{\mathcal{C}}(v_1, \dots, v_n) = \det_{\mathcal{C}}(\mathcal{B}) \det_{\mathcal{B}}(v_1, \dots, v_n).$$

6.56 En appliquant la formule on peut facilement constater que pour toute base $\mathcal{B}$ on a $\det_{\mathcal{B}}(\mathcal{B}) = 1$ parce que ce déterminant est celui de la matrice identité.

Du fait du lemme 6.55, on peut donner une définition alternative de la fonction $\det_{\mathcal{B}}$: c'est l'unique forme n-linéaire alternée qui envoie $\mathcal{B}$ sur 1.

6.57 **Théorème.** *Le déterminant d'une famille de vecteurs est nul et si et seulement si cette famille est liée.*

Démonstration. On se place dans une base quelconque $\mathcal{B}$. Si une famille $(v_1, \dots, v_n)$ est liée alors l'un des vecteurs est combinaison linéaire des autres. Quitte à changer l'ordre, on suppose qu'il s'agit de v_n qui s'écrit $a_1 v_1 + \cdots + a_{n-1} v_{n-1}$. Par linéarité on a donc $\det_{\mathcal{B}}(v_1, \dots, v_n) = \sum_{i=1}^{n-1} a_i \det_{\mathcal{B}}(v_1, \dots, v_{n-1}, v_i)$ or les déterminants dans cette formule sont tous nuls puisqu'ils portent sur des familles dont deux vecteurs sont égaux. Au contraire, si $(v_1, \dots, v_n)$ est libre alors c'est une base $\mathcal{C}$ de E et on a $\det_{\mathcal{C}}(\mathcal{B}) \det_{\mathcal{B}}(\mathcal{C}) = \det_{\mathcal{C}}(\mathcal{C}) = 1$, donc $\det_{\mathcal{B}}(\mathcal{C}) \neq 0$. □

6.58 **Proposition (déterminant d'un endomorphisme).** *Soit E un espace vectoriel de dimension finie et $f : E \to E$ une application linéaire. Pour toute base $\mathcal{B} = (e_1, \dots, e_n)$ de E on pose $\det_{\mathcal{B}}(f) = \det_{\mathcal{B}}(f(e_1), \dots, f(e_n))$. Alors $\det_{\mathcal{B}}(f)$ ne dépend pas de $\mathcal{B}$, on l'appelle* déterminant *de f et on le note* $\det(f)$.

Démonstration. Considérons deux bases $\mathcal{B}$ et $\mathcal{C}$ de E et une application linéaire $f : E \to E$ et montrons que $\det_{\mathcal{B}}(f) = \det_{\mathcal{C}}(f)$. On a

$$\begin{aligned}
\det_{\mathcal{C}}(f) &= \det_{\mathcal{C}}(f(\mathcal{C})) = \det_{\mathcal{C}}(\mathcal{B}) \det_{\mathcal{B}}(f(\mathcal{C})) && \text{par changement de base} \\
&= \det_{\mathcal{C}}(\mathcal{B}) \det_{\mathcal{B}}(f(\mathcal{B})) \det_{\mathcal{B}}(\mathcal{C}) && \text{par le lemme 6.55 appliqué à } \det_{\mathcal{B}} \circ f \\
&= \det_{\mathcal{C}}(\mathcal{B}) \det_{\mathcal{B}}(f) \det_{\mathcal{B}}(\mathcal{C}) && \text{par définition} \\
&= \det_{\mathcal{C}}(\mathcal{B}) \det_{\mathcal{B}}(\mathcal{C}) \det_{\mathcal{B}}(f) = \det_{\mathcal{B}}(f).
\end{aligned}$$

du fait que $\det_{\mathcal{B}}(\mathcal{C})$ et $\det_{\mathcal{C}}(\mathcal{B})$ sont inverses l'un de l'autre. □

6.59 **Théorème.** *Un endomorphisme $f : E \to E$ d'un espace vectoriel de dimension finie est inversible si et seulement si son déterminant n'est pas nul.*

Démonstration. Soit $\mathcal{B}$ une base de E. Un endomorphisme f est inversible si et seulement si la famille $f(\mathcal{B})$ est libre, et par le théorème 6.57 ceci équivaut au fait que $\det_{\mathcal{B}}(f(\mathcal{B})) \neq 0$, c'est-à-dire $\det(f) \neq 0$. □

6.60 **Théorème.** *Pour tous endomorphismes $f, g : E \to E$ d'un espace vectoriel de dimension finie, on a* $\det(f \circ g) = \det(f)\det(g)$.

Démonstration. Si g est inversible, alors pour tout base $\mathcal{B}$ la famille $g(\mathcal{B})$ est une base et on peut écrire $\det(f \circ g) = \det_{\mathcal{B}}(f(g(\mathcal{B}))) = \det_{\mathcal{B}}(g(\mathcal{B}))\det_{g(\mathcal{B})}(f(g(\mathcal{B}))) = \det(g)\det(f)$. Si g n'est pas inversible alors $f \circ g$ non plus et $\det(f \circ g) = \det(f)\det(g) = 0$ par le théorème précédent. □

6.61 Pour tout espace vectoriel E de dimension finie et toute base $\mathcal{B}$ de E, pour tout endomorphisme $f : E \to E$ on a $\det(f) = \det(M_{\mathcal{B}}(f))$ par définition. On peut donc transposer les propriétés du déterminant des endomorphismes au cas des matrices :

— Pour toutes matrices A et B, on a $\det(AB) = \det(A)\det(B)$.
— Une matrice A est inversible si et seulement si $\det(A) \neq 0$.
— Le déterminant commute à l'inversion : $\det(A^{-1}) = \det(A)^{-1}$.

On va maintenant voir comment généraliser la formule de l'inversion des matrices à toute dimension.

6.62 **Exercice.** Démontrer que le déterminant d'une matrice ne change pas lorsqu'on ajoute à une de ses colonnes une combinaisons linéaire de ses colonnes.

Le développement d'un déterminant par rapport à une colonne (ou à une ligne) est une méthode permettant le calcul des déterminants. Soit

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

une matrice $n \times n$ à coefficients a_{ij} dans un corps $\mathbb{K}$. Pour chaque couple $1 \leq i, j \leq n$, notons

6.63 **Proposition (développement du déterminant).** *Soit $A = (a_{ij})_{1 \leq i,j \leq n}$ une matrice $n \times n$ à coefficients a_{ij} dans un corps $\mathbb{K}$. Pour tout couple (i, j) on note A_{ij} la matrice $(n-1) \times (n-1)$ obtenue à partir de A en supprimant la i-ème ligne et la j-ème colonne. Alors pour tout $i \in \{1, \ldots, n\}$ on a*

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

Démonstration. On peut procéder simplement en appliquant la formule du déterminant. □

6.64 **Exercice.** Démontrer que le déterminant d'une matrice carrée qui possède au moins deux lignes égales est nul.

6.65 **Théorème.** *Soit $A = (a_{ij})_{1 \leq i,j \leq n}$ une matrice $n \times n$ à coefficients dans un corps $\mathbb{K}$. Soit $B = (b_{ij})_{1 \leq i,j \leq n}$, la matrice définie par*

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

Alors on a

$$AB = \det(A) I_n$$

En particulier, si $\det(A) \neq 0$, *alors on a* $A^{-1} = (1/\det(A))B$.

Démonstration. Soit c_{ij} le i-ème terme de la j-ème colonne du produit AB. On a donc :

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

Si $i = j$, on obtient $c_{ii} = \det(A)$ d'après la proposition 6.63. Sinon, soit A' la matrice obtenue à partir de A en remplaçant la j-ème ligne par la i-ème. On a donc $\det(A') = 0$ puisque les lignes i et j de A' sont égales (voir l'exercice 6.64). En développant $\det(A')$ selon la j-ème ligne, on obtient

$$\det(A') = \sum_{k=1}^{n} (-1)^{j+k} a'_{jk} \det(A'_{jk}) = \sum_{k=1}^{n} (-1)^{j+k} a_{ik} \det(A_{jk}) = c_{ij}.$$

puisque $a'_{jk} = a_{ik}$ et $A'_{jk} = A_{jk}$ par construction, d'où $c_{ij} = 0$. □

6.4 Diagonalisation et vecteurs propres

La représentation matricielle d'une application linéaire dépend du choix des bases. Afin de trouver les bases qui produisent les matrices les plus simples, on étudie les propriétés caractéristiques de l'application considérée, en particulier ses vecteur propres, que l'on définit dans la suite. Dans un premier temps on étudie comment se transforme la matrice d'une application linéaire lors d'un changement de base.

Soit $f : E \to E$ une application linéaire d'un espace vectoriel E de dimension n. Soit $\mathcal{B} = (b_1, \dots, b_n)$ une base de E. Par définition, si on note $a_{ij}^{\mathcal{B}}$ le coefficient en i-ème ligne et j-ème colonne de $M_{\mathcal{B}}(f)$, alors on a pour tout $1 \leq j \leq n$ l'égalité

$$f(b_j) = a_{1j}^{\mathcal{B}} b_1 + \cdots + a_{nj}^{\mathcal{B}} b_n.$$

En choisissant une autre base $\mathcal{C} = (c_1, \dots, c_n)$ de E et en notant $a_{ij}^{\mathcal{C}}$ le coefficient en i-ème ligne et j-ème colonne de $M_{\mathcal{C}}(f)$, on a pour tout $1 \leq j \leq n$ l'égalité

$$f(c_j) = a_{1j}^{\mathcal{C}} c_1 + \cdots + a_{nj}^{\mathcal{C}} c_n.$$

Pour savoir comment exprimer les $a_{ij}^{\mathcal{C}}$ et fonction des $a_{ij}^{\mathcal{B}}$, on peut appliquer le théorème 6.15 qui lie la composition d'applications linéaires et le produit des matrices, en utilisant l'application identité sur E. En effet, on a

$$M_{\mathcal{C}}(f) = M_{\mathcal{C}}(\mathrm{id}_E \circ f \circ \mathrm{id}_E) = M_{\mathcal{B},\mathcal{C}}(\mathrm{id}_E)\, M_{\mathcal{B}}(f)\, M_{\mathcal{C},\mathcal{B}}(\mathrm{id}_E)$$

de plus en appliquant le même théorème on constate que $M_{\mathcal{B},\mathcal{C}}(\mathrm{id}_E)$ et $M_{\mathcal{C},\mathcal{B}}(\mathrm{id}_E)$ sont inverses l'une de l'autre.

6.66 **Définition (matrice de passage).** Soit E un espace vectoriel et soient $\mathcal{B}$ et $\mathcal{C}$ deux bases de E. La matrice $M_{\mathcal{C},\mathcal{B}}(\mathrm{id}_E)$ est appelée *matrice de passage* ou *matrice de changement de base* de $\mathcal{B}$ dans $\mathcal{C}$.

6.67 Une autre façon d'exprimer la matrice de passage de $\mathcal{B}$ vers $\mathcal{C}$ est d'observer que c'est la matrice dont les colonnes représentent les coordonnées des vecteurs de $\mathcal{C}$ dans la base $\mathcal{B}$. Ainsi, si P est la matrice de passage de $\mathcal{B}$ dans $\mathcal{C}$ et si X est le vecteur colonne représentant un vecteur x dans $\mathcal{C}$, alors PX est le vecteur colonne représentant x dans $\mathcal{B}$, et à l'inverse P^{-1} transforme des coordonnées dans $\mathcal{C}$ en coordonnées dans $\mathcal{B}$. Par conséquent, si A est la matrice de f dans une $\mathcal{B}$ et P est la matrice de passage de $\mathcal{B}$ dans $\mathcal{C}$, la matrice de f dans $\mathcal{C}$ est donc $P^{-1}AP$.

6.68 *Exemple.* On se place dans $\mathbb{R}^2$ muni de la base canonique $\mathcal{B} = ((1,0),(0,1))$ et on considère l'application linéaire $f : \mathbb{R}^2 \to \mathbb{R}^2$ telle que $f(1,0) = (1,2)$ et $f(0,1) = (3,4)$. Alors la matrice A de f dans $\mathcal{B}$ est $\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$. Si on considère maintenant la base $\mathcal{C}$ obtenue en échangeant les deux vecteurs de $\mathcal{B}$, c'est-à-dire $\mathcal{C} = ((0,1),(1,0))$, comme $(3,4) = 4(0,1) + 3(1,0)$ et $(1,2) = 2(0,1) + 1(1,0)$ la matrice de f dans $\mathcal{C}$ est $B = \begin{pmatrix} 4 & 2 \\ 3 & 1 \end{pmatrix}$. La matrice de changement de base de $\mathcal{B}$ dans $\mathcal{C}$ est $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ et on vérifie facilement que $B = P^{-1}AP$, en fait dans ce cas on a même $P = P^{-1}$.

6.69 **Exercice.** Soit E un espace vectoriel de dimension n. Soit f une application linéaire de E dans E. Soit $\mathcal{B} = (b_1, \ldots, b_n)$ une base de E et soit $M_{\mathcal{B}}(f)$ la matrice de f dans cette base. Montrer que la somme

$$\mathrm{Tr}(M_{\mathcal{B}}(f)) = \sum_{i=1}^{n} m_{ii}$$

des coefficients diagonaux de la matrice $M_{\mathcal{B}}(f)$ ne dépend pas du choix de la base, c'est-à-dire si $\mathcal{C} = (c_1, \ldots, c_n)$ est une base de E et si $M_{\mathcal{C}}(f)$ est la matrice de f dans $\mathcal{C}$, alors $\mathrm{Tr}(M_{\mathcal{B}}(f)) = \mathrm{Tr}(M_{\mathcal{C}}(f))$. Cette somme, qui ne dépend donc que de f, est appelée *la trace* de l'endomorphisme f. Indication : prouver que si A et B sont deux matrices carrées alors $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, puis appliquer la formule du changement de base.

6.70 **Définition (vecteur propre, valeur propre).** Soient E un espace vectoriel et $f : E \to E$ une application linéaire. Un *vecteur propre* de f est un vecteur $v \in E$ non nul tel que $f(v) = \lambda v$ pour un certain scalaire λ. Ce scalaire est la *valeur propre* de f associée au vecteur propre v.

6.71 Remarquons que contrairement à v qui, par définition d'un vecteur propre, n'est pas nul, le scalaire λ peut tout à fait être égal à zéro. C'est le cas si et seulement si v est dans le noyau de f.

6.72 *Exemple.* Si $f : E \to E$ est l'identité alors tout vecteur non nul de E est un vecteur propre dont la valeur propre associée est 1.

6.73 *Exemple.* Considérons dans le plan $\mathbb{R}^2$ une droite D passant par l'origine et soit s la symétrie d'axe D. Alors tous les vecteurs de D sont invariants par s donc ils sont tous vecteurs propres associés à la valeur propre 1. D'autre part, pour tout vecteur v perpendiculaire à D on a $s(v) = -v$ donc v est vecteur propre associé à la valeur propre -1.

6.74 *Exemple.* Supposons que dans une base $(b_1, \ldots, b_n)$ de E, la matrice d'une application

linéaire f soit diagonale, c'est-à-dire de la forme

$$M(f) = \begin{pmatrix} a_{11} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{nn} \end{pmatrix}.$$

Dans ce cas, chaque vecteur de base b_i est un vecteur propre pour f de valeur propre a_{ii}. Si de plus tous les éléments diagonaux sont égaux à un même scalaire λ, c'est-à-dire $\lambda = a_{11} = \cdots = a_{nn}$, alors f est une homothétie de rapport λ centrée à l'origine.

6.75 **Exercice (sous-espace propre).** Soit $f : E \to E$ une application linéaire. Soit λ une valeur propre de f. Montrer que l'ensemble $E_\lambda = \{v \in E \mid f(v) = \lambda v\}$ des vecteurs propres de f de valeur propre λ est un sous-espace vectoriel de E. Montrer que E_λ est stable par f, c'est-à-dire que $f(E_\lambda) \subseteq E_\lambda$. Montrer que si μ est une valeur propre de A qui est distincte λ alors $E_\lambda \cap E_\mu = \{0\}$. Le sous-espace E_λ est appelé le *sous-espace propre* de f associé à la valeur propre λ.

6.76 **Exercice (matrice par blocs).** Soit E un espace vectoriel et soient U et V des sous-espaces vectoriels de E tels que $E = U \oplus V$. Soit $f : E \to E$ une application linéaire. Supposons que f préserve U et V, c'est-à-dire que $f(U) \subseteq U$ et $f(V) \subseteq V$. Notons $f_U : U \to U$ et $f_V : V \to V$ les restrictions de f à U et à V. Montrer qu'il existe une base de E dans laquelle la matrice de A est de la forme

$$M(f) = \begin{pmatrix} M(f_U) & 0 \\ 0 & M(f_V) \end{pmatrix},$$

où $M(f_U)$ est la matrice de f_U dans une base de U et $M(f_V)$ est la matrice de f_V dans une base de V. La matrice ci-dessus est une matrice constituée de quatre blocs. Si $\dim(U) = m$ et $\dim(V) = n$, alors le bloc $M(f_U)$ est une matrice $m \times m$, $M(f_V)$ est une matrice $n \times n$, le symbole 0 qui apparaît en haut à droite représente une matrice à m lignes et n colonnes dont tous les coefficients sont nuls, le symbole 0 qui apparaît en bas à gauche représente une matrice à n lignes et m colonnes dont tous les coefficients sont nuls.

6.77 **Exercice (matrice à un bloc).** Soit E un espace vectoriel et soit U un sous-espace vectoriel de E. Soit $f : E \to E$ une application linéaire. Supposons que f préserve U. Notons $f_U : U \to U$ la restriction de f à U. Montrer qu'il existe une base de E dans laquelle la matrice de f est de la forme

$$M(A) = \begin{pmatrix} M(f_U) & * \\ 0 & * \end{pmatrix},$$

où $M(f_U)$ est la matrice de f_U dans une base de U. Comme dans l'exercice 6.76, la matrice est constituée de quatre blocs, ceux de la diagonale étant carrés. Le symbole $*$ dans la colonne de droite représente des matrices quelconques de taille convenable.

6.78 **Théorème.** *Soient E un espace vectoriel de dimension finie et $f : E \to E$ une application linéaire. Les conditions suivantes sont équivalentes.*

1. *Il existe une base de E dans laquelle la matrice de f est diagonale.*
2. *Il existe une base de E qui est formée de vecteurs propres de f.*

3. L'espace E est somme directe des sous-espaces propres de f.

Démonstration. Comme on l'a vu dans l'exemple 6.74, s'il existe un base dans laquelle la matrice de f est diagonale, alors chaque vecteur de cette base est un vecteur propre de f.

Supposons maintenant que $\mathcal{B} = (b_1, \dots, b_n)$ est une base de E formée de vecteurs propres de f. Pour chaque valeur propre λ correspondant à un vecteur de cette base (il se peut bien entendu que plusieurs vecteurs de cette base soient vecteurs propres de f de même valeur propre), notons F_λ le sous-espace propre de f associé à λ et E_λ le plus petit sous-espace vectoriel de E contenant tous les vecteurs de $\mathcal{B}$ dont la valeur propre est λ. On a alors $E_\lambda \subseteq F_\lambda$ donc l'intersection $E_\lambda \cap E_\mu$ est réduite à $\{0\}$ dès que $\lambda \neq \mu$. Comme la réunion des E_λ contient la base $\mathcal{B}$, on obtient E comme somme directe $E = \bigoplus_\lambda E_\lambda \subset \bigoplus_\lambda F_\lambda \subset E$, ce qui montre que $E = \bigoplus_\lambda F_\lambda$.

Pour finir, supposons que E est somme directe des sous-espaces propres de f et choisissons pour chaque sous-espace propre F_λ de f une base $\mathcal{B}_\lambda$. Comme la restriction de f à F_λ est une homothétie de rapport λ, la matrice de la restriction de f à F_λ est diagonale, quelle que soit la base $\mathcal{B}_\lambda$. La réunion $\mathcal{B} = \bigcup_\lambda \mathcal{B}_\lambda$ est donc une base de E dans la quelle la matrice de f est diagonale. □

Une application linéaire qui satisfait les conditions du théorème est dite *diagonalisable* puisque les conditions entraînent qu'il existe une base dans laquelle la matrice de f est diagonale. Par extension, on dit qu'une matrice A est diagonalisable s'il existe une matrice inversible P telle que $P^{-1}AP$ est diagonale, ce qui revient à dire que l'on peut changer de base (par la matrice de passage P) pour obtenir une représentation diagonale de l'application linéaire représentée par A.

6.79 *Exemple.* Considérons la matrice $A = \begin{pmatrix} -1 & -2 \\ 3 & 4 \end{pmatrix}$. Alors on peut observer que

$$A\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \qquad \text{et} \qquad A\begin{pmatrix} -2 \\ 3 \end{pmatrix} = \begin{pmatrix} -4 \\ 6 \end{pmatrix}$$

donc $V = (1, -1)$ est vecteur propre associé à la valeur propre 1 et $W = (-2, 3)$ est vecteur propre associé à la valeur propre 2. Ces deux vecteurs sont linéairement indépendants donc ils forment une base de $\mathbb{R}^2$, ainsi A est diagonalisable et il existe une matrice P telle que

$$P^{-1}AP = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

Pour trouver cette matrice P il suffit de remarquer que puisque l'on a $AV = V$ et $AW = 2W$, il suffit que P appliquée à $(1, 0)$ donne V et que P appliqué à $(0, 1)$ donne W. Ainsi on obtient P en rassemblant les colonnes V et W :

$$P = \begin{pmatrix} V & W \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix}$$

et on peut obtenir l'inverse avec la formule pour la dimension 2 (voir paragraphe 7) :

$$P^{-1} = \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix}^{-1} = \frac{1}{1 \times 3 - (-2) \times (-1)} \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}.$$

6.80 Un avantage qu'il y a à diagonaliser une matrice est que cela simplifie le calcul de son inverse et de ses puissances. En effet, pour toute matrice inversible P, l'application de « changement de base » $A \mapsto P^{-1}AP$ préserve l'addition, la multiplication et l'inversion des matrices, ainsi on a $(P^{-1}AP)^n = P^{-1}A^nP$ pour tout n. Dans le cas des matrices diagonales, ces puissances sont immédiates à calculer puisqu'il suffit de calculer la puissance correspondante de chaque coefficient de la diagonale.

Pour reprendre l'exemple précédent, après diagonalisation on obtient

$$P^{-1}AP = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \qquad \text{donc} \qquad A = P \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} P^{-1}$$

et pour tout $n \in \mathbb{Z}$ on a donc les égalités

$$\begin{aligned} A^n = P \begin{pmatrix} 1 & 0 \\ 0 & 2^n \end{pmatrix} P^{-1} &= \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2^n \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -2^{n+1} \\ -1 & 3 \cdot 2^n \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 3 - 2^{n+2} & 1 - 2^{n+1} \\ 3(2^{n+1} - 1) & 3 \cdot 2^n - 1 \end{pmatrix}. \end{aligned}$$

Le calcul direct de la dernière formule n'est pas facile et nécessiterait au mieux de deviner la formule puis de la démontrer par récurrence sur n.

De même l'écriture diagonale permet de déduire immédiatement que les valeurs propres de A^n sont 1 et 2^n et que les sous-espaces propres associés sont ceux de A.

6.81 *Exemple.* Il est important de noter que toute matrice n'est pas diagonalisable. L'exemple le plus simple est donné par les rotation du plan $\mathbb{R}^2$: on a vu dans l'exemple 6.4 que la rotation d'angle θ autour de l'origine a pour matrice

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

dans la base canonique. Géométriquement, il est clair que cette transformation n'a pas de vecteur propre dès lors que θ n'est pas multiple de π, en effet la transformation préserve la norme des vecteurs donc un vecteur propre devrait être envoyé soit sur lui-même soit sur son opposé, mais ces cas ne peuvent survenir dans une rotation que si l'angle est 0 (c'est la fonction identité) ou π (c'est la symétrie centrale). Ainsi une rotation de $\mathbb{R}^2$ d'angle non multiple de π n'est pas diagonalisable.

6.5 Polynômes d'endomorphismes

Pour diagonaliser une matrice, on en cherche une base de vecteurs propres de cette matrice, s'il en existe. Pour cela on cherche à résoudre l'équation $AX = \lambda X$ dont les inconnues sont λ et X, avec la contrainte que X ne doit pas être le vecteur nul (puisque celui-ci est toujours solution de l'équation mais qu'un vecteur propre n'est pas nul par définition). Cette équation est équivalente à $(A - \lambda I)X = 0$, ainsi λ est valeur propre de A si et seulement s'il existe un $X \neq 0$ tel que $(A - \lambda I)X = 0$, c'est-à-dire si $\mathrm{Ker}(A - \lambda I) \neq \{0\}$, c'est-à-dire encore si $A - \lambda I$ n'est pas inversible. L'ensemble des valeurs propres d'une matrice A est donc l'ensemble des solutions de l'équation $\det(A - \lambda I) = 0$.

6.82 **Proposition (polynôme caractéristique).** *Soit A une matrice $n \times n$ à coefficients dans un corps $\mathbb{K}$. La fonction $\lambda \mapsto \det(A - \lambda I)$ est une fonction polynomiale de degré n. Le polynôme correspondant, élément de $\mathbb{K}[X]$, est appelé* polynôme caractéristique *de A et est noté P_A ou $P_A(X)$.*

Ce polynôme est invariant par changement de base. Pour un endomorphisme f d'un espace vectoriel de dimension finie, on appelle polynôme caractéristique *de f et on note P_f ou $P_f(X)$ le polynôme caractéristique de la matrice de f dans une base quelconque.*

Démonstration. On commence par considérer le cas d'une matrice carrée A de taille n à coefficients dans $\mathbb{K}$. En appliquant la formule du déterminant on observe immédiatement que $\det(A-\lambda I)$ est un polynôme en λ (puisqu'il est défini uniquement avec des sommes est des produits), de plus son terme dominant est $(-\lambda)^n$ et il est issu du terme correspondant à la permutation identité. Si A est la matrice de f dans une base $\mathcal{B}$ donnée, alors $A - \lambda I$ est bien la matrice de $f - \lambda \,\mathrm{id}$ dans cette base. De plus on a vu que le déterminant d'un endomorphisme ne dépend pas de la base choisie, donc la définition est valide. □

6.83 Pour être très formel, on pourrait considérer la matrice $A - XI$ à coefficients dans $\mathbb{K}[X]$ et faire le calcul du déterminant de cette matrice dans $\mathbb{K}[X]$. Cependant $\mathbb{K}[X]$ n'est pas un corps donc la construction précise nécessiterait des complications techniques dont l'intérêt est limité dans le cadre de ce cours.

6.84 *Exemple.* Reprenons la matrice A de l'exemple 6.79. On a alors

$$\det(A - \lambda I) = \det \begin{pmatrix} -1-\lambda & -2 \\ 3 & 4-\lambda \end{pmatrix} = (-1-\lambda)(4-\lambda) - (-2)3 = \lambda^2 - 3\lambda + 2$$

donc $P_A = X^2 - 3X + 2$. On constate que ce polynôme peut se factoriser en $P_A = (X-1)(X-2)$ donc ses racines sont 1 et 2, ce qui permet de retrouver le fait que les valeurs propres de A sont 1 et 2.

6.85 *Exemple.* Dans le cas de la rotation du plan d'angle θ, notée r_θ, on a

$$\det(r_\theta - \lambda \,\mathrm{id}) = \det \begin{pmatrix} \cos\theta - \lambda & -\sin\theta \\ \sin\theta & \cos\theta - \lambda \end{pmatrix} = (\cos\theta - \lambda)^2 + \sin^2\theta$$

donc $P_{r_\theta} = (\cos\theta - X)^2 + \sin^2\theta = X^2 - 2\cos\theta\lambda + 1$. Le discriminant de ce polynôme du second degré est $4\cos^2\theta - 4 = -4\sin^2\theta$. Ce polynôme n'a de racine réelle que si $\sin\theta = 0$ c'est-à-dire si θ est multiple entier de π, et dans ce cas l'unique racine est $\cos\theta$, ce qui correspond aux observations précédentes sur les valeurs et vecteurs propres de r_θ.

6.86 **Proposition.** *Soit E un espace vectoriel de dimension finie. Soit f un endomorphisme de E. Un élément $\lambda \in \mathbb{K}$ est valeur propre de f si et seulement si λ est racine du polynôme caractéristique de f.*

Démonstration. Soit $\lambda \in \mathbb{K}$. On a $P_f(\lambda) = \det(f - \lambda \,\mathrm{id}) = 0$ si et seulement si $f - \lambda \,\mathrm{id}$ n'est pas inversible, ce qui équivaut à $\mathrm{Ker}(f - \lambda \,\mathrm{id}) \neq \{0\}$ sont à l'existence d'un vecteur $v \in E$ non nul tel que $f(v) = \lambda v$. □

On va voir dans la suite que l'application de polynôme aux endomorphismes et aux matrices donne des techniques très puissantes pour étudier les propriétés de ces objets.

On commence par définit formellement ce que signifie appliquer un polynôme à un endomorphisme ; il s'agit simplement d'un cas particulier d'évaluation d'un polynôme dans une algèbre.

6.87 **Définition (polynôme d'endomorphisme).** Soit f un endomorphisme d'un espace vectoriel E sur un corps $\mathbb{K}$. Pour chaque polynôme $P = a_d X^d + \cdots + a_1 X + a_0$ de $\mathbb{K}[X]$, on note $P(f)$ l'endomorphisme $a_d f^d + \cdots + a_1 f + a_0$, où f^d désigne le morphisme $f \circ \cdots \circ f$ obtenu en composant d exemplaires de f.

6.88 Pour f fixé, l'application de l'algèbre des polynômes $\mathbb{K}[X]$ dans l'algèbre des endomorphismes $\mathcal{L}(E)$ qui à chaque P associe $P(f)$ est un *morphisme d'algèbres*, c'est-à-dire que pour tous $P, Q \in \mathbb{K}[X]$ et $\lambda \in \mathbb{K}$ on a les égalités

$$(P+Q)(f) = P(f) + Q(f), \qquad (\lambda P)(f) = \lambda P(f), \quad \text{et} \quad (PQ)(f) = P(f) \circ Q(f).$$

Les deux premières égalités sont immédiates, la troisième peut surprendre mais se vérifie simplement en écrivant $P = a_m X^m + \cdots + a_1 X + a_0$ et $Q = b_n X^n + \cdots + b_1 X + b_0$ et en développant le produit.

Du fait que la multiplication des polynômes est commutative, cette égalité implique que deux morphismes $P(f)$ et $Q(f)$ (des polynômes du même morphisme f) commutent toujours, même si l'algèbre des endomorphismes n'est pas commutative. En d'autres termes, la *sous-algèbre* des polynômes de f est commutative.

6.89 **Théorème (lemme des noyaux).** *Soit E un espace vectoriel et soit f un endomorphisme de E. Soient $P, Q \in \mathbb{K}[X]$ deux polynômes premiers entre eux. Alors*

$$\operatorname{Ker}((PQ)(f)) = \operatorname{Ker}(P(f)) \oplus \operatorname{Ker}(Q(f)).$$

Démonstration. Comme P et Q sont premiers entre eux, il existe deux polynômes U et V dans $\mathbb{K}[X]$ tels que $UP + VQ = 1$, d'après le théorème de Bézout appliqué aux polynômes. On peut écrire l'endomorphisme identité $\mathrm{id} : E \to E$ comme $\mathrm{id} = (UP + VQ)(f)$, d'où on déduit que pour tout $v \in E$ on a $v = (UP)(f)(v) + (VQ)(f)(v)$.

Soit $v \in \operatorname{Ker}(P(f)) \cap \operatorname{Ker}(Q(f))$. Par définition on a donc $P(f)(v) = Q(f)(v) = 0$, d'où $(UP)(f)(v) = 0$ et $(VQ)(f)(v) = 0$, et d'après l'égalité précédente on en déduit $v = 0$. Ainsi $\operatorname{Ker}(P(f)) \cap \operatorname{Ker}(Q(f)) = \{0\}$ donc $\operatorname{Ker}(P(f))$ et $\operatorname{Ker}(Q(f))$ sont en somme directe.

Montrons maintenant l'inclusion $\operatorname{Ker}(P(f)) \oplus \operatorname{Ker}(Q(f)) \subseteq \operatorname{Ker}((PQ)(f))$. Pour tout vecteur $v \in \operatorname{Ker}(P(f))$, on a $(QP)(f)(v) = Q(f)(P(f)(v)) = Q(f)(0) = 0$ donc $v \in \operatorname{Ker}((QP)(f))$ et $\operatorname{Ker}(P(f)) \subseteq \operatorname{Ker}((QP)(f))$, or $(QP)(f) = (PQ)(f)$ donc $\operatorname{Ker}(P(f)) \subseteq \operatorname{Ker}((PQ)(f))$. On prouve de même que $\operatorname{Ker}(Q(f)) \subseteq \operatorname{Ker}((PQ)(f))$ et donc le sous-espace vectoriel de $\operatorname{Ker}((PQ)(f))$ engendré par $\operatorname{Ker}(P(f))$ et $\operatorname{Ker}(Q(f))$, qui est la somme $\operatorname{Ker}(P(f)) \oplus \operatorname{Ker}(Q(f))$, est lui aussi inclus dans $\operatorname{Ker}((PQ)(f))$.

Pour l'inclusion réciproque, considérons un vecteur $v \in \operatorname{Ker}((PQ)(f))$. Comme on a $(PQ)(f)(v) = 0$ on déduit $(QUP)(f)(v) = (UPQ)(f)(v) = 0$ donc $(UP)(f)(v) \in \operatorname{Ker}(Q(f))$. Par un argument similaire on déduit $(VQ)(f)(v) \in \operatorname{Ker}(P(f))$, en additionnant les deux on on obtient $(UP)(f)(v) + (VQ)(f)(v) \in \operatorname{Ker}(P(f)) \oplus \operatorname{Ker}(Q(f))$ d'où $v \in \operatorname{Ker}(P(f)) \oplus \operatorname{Ker}(Q(f))$ par définition de U et V, ainsi on a $\operatorname{Ker}((PQ)(f)) \subseteq \operatorname{Ker}(P(f)) \oplus \operatorname{Ker}(Q(f))$. □

On rappelle qu'un polynôme $P \in \mathbb{K}[X]$ est *scindé* $\mathbb{K}$ s'il se décompose dans $\mathbb{K}[X]$ en facteurs linéaires : $P = \lambda(X - \alpha_1)(X - \alpha_2) \cdots (X - \alpha_n)$ où les α_i sont éléments de $\mathbb{K}$. On

sait que la liste $\alpha_1, \dots, \alpha_n$ est unique à permutation près. La *multiplicité* d'une racine α est le nombre de fois où elle apparaît dans cette liste ; une racine est *simple* si sa multiplicité est 1, elle est *double* si sa multiplicité est 2, etc.

6.90 **Théorème.** *Soit E un espace vectoriel de dimension finie. Un endomorphisme f de E est diagonalisable si et seulement s'il existe un polynôme $P \in \mathbb{K}[X]$ scindé, à racines simples, tel que $P(f) = 0$.*

Démonstration. Supposons que f soit diagonalisable. Par définition il existe donc une base $\mathcal{B}$ de E formée de vecteurs propres de f. Soient $\{\lambda_1, \dots, \lambda_k\}$ l'ensemble des valeurs propres de f (avec $\lambda_i \neq \lambda_j$ pour $i \neq j$), posons $P = (X - \lambda_1) \cdots (X - \lambda_k)$. Pour tout vecteur b dans la base $\mathcal{B}$ il existe un i tel que $f(b) = \lambda_i b$ donc $(f - \lambda_i \operatorname{id})(b) = 0$ et donc $P(f)(b) = 0$, par conséquent $P(f)$ envoie chaque vecteur de la base $\mathcal{B}$ sur 0 donc $P(f) = 0$.

Réciproquement, supposons qu'il existe $P \in \mathbb{K}[X]$, scindé à racines simples, tel que $P(f) = 0$. Sans perte de généralité, on peut supposer que le coefficient dominant de P est 1 et que l'on a donc $P = (X - \alpha_1) \cdots (X - \alpha_k)$ pour une famille $\alpha_1, \dots, \alpha_k$ d'éléments deux à deux distincts de $\mathbb{K}$. En appliquant le lemme des noyaux, on obtient

$$E = \operatorname{Ker}(P(f)) = \operatorname{Ker}(f - \alpha_1 \operatorname{id}) \oplus \cdots \oplus \operatorname{Ker}(f - \alpha_k \operatorname{id}).$$

Comme pour chaque i, les endomorphismes f et $f - \alpha_i \operatorname{id}$ commutent (puisque ce sont des polynômes en f), on déduit que chaque sous-espace $\operatorname{Ker}(f - \alpha_i \operatorname{id})$ est préservé par f. Ainsi en choisissant pour chaque i une base de $\operatorname{Ker}(f - \alpha_i \operatorname{id})$ et en considérant la base de E constituée de la réunion des bases choisies des sous-espaces $\operatorname{Ker}(f - \alpha_i \operatorname{id})$, on obtient une matrice de f constituée de blocs, chacun des blocs étant une matrice de la restriction de f à un sous-espace $\operatorname{Ker}(f - \alpha_i \operatorname{id})$. Comme la restriction de f à un tel sous-espace est une homothétie de rapport α_i, sa matrice est diagonale quelle que soit la base choisie. Ainsi la matrice construite pour f est diagonale. □

6.91 *Exemple.* Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ l'endomorphisme du plan défini dans la base canonique par la matrice

$$M(f) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Le polynôme caractéristique de f est

$$P_f = \det \begin{pmatrix} 2 - X & 1 \\ 1 & 1 - X \end{pmatrix} = X^2 - 3X + 1 = \left(X - \frac{3 + \sqrt{5}}{2} \right) \left(X - \frac{3 - \sqrt{5}}{2} \right).$$

On vérifie que $f^2 - 3f + \operatorname{id} = 0$ (en fait, pour tout endomorphisme f on a $P_f(f) = 0$, c'est un résultat connu sous le nom de *théorème de Cayley-Hamilton*). Le théorème précédant implique donc que f est diagonalisable. Les valeurs propres sont les racines

$$\lambda = \frac{3 + \sqrt{5}}{2}, \qquad \mu = \frac{3 - \sqrt{5}}{2}$$

de P_f, remarquons que $\lambda\mu = \det(f) = 1$ et que $\lambda + \mu = \operatorname{Tr}(f) = 3$. Pour trouver une base de vecteurs propres il faut donc trouver $v, w \in \mathbb{R}^2$ tels que $f(v) = \lambda v$ et $f(w) = \mu w$. En posant le calcul on trouve $v = (\mu, 1 - 2\mu)$ (tout multiple de v est aussi une solution) et

$w = (\lambda, 1-2\lambda)$ (tout multiple de w est aussi une solution). Dans la base (v, w), la matrice de f est

$$M_{(v,w)}(f) = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$$

et la matrice de changement de base de la base canonique dans (v, w) est

$$P = \begin{pmatrix} \mu & \lambda \\ 1-2\mu & 1-2\lambda \end{pmatrix}.$$

6.92 **Proposition (polynôme minimal).** *Soit A une matrice carrée (ou un endomorphisme d'un espace vectoriel de dimension finie). Parmi l'ensemble des polynômes non nuls P tels que $P(A) = 0$, il y en a exactement un qui est de degré minimal et unitaire. Ce polynôme est appelé* polynôme minimal *de A.*

Démonstration. Soit n la taille de la matrice A. L'espace $M_n(\mathbb{K})$ est de dimension n^2 donc la famille de matrices $(I, A, A^2, \dots, A^{n^2})$, qui a n^2+1 éléments, est liée, ce qui implique qu'il existe une combinaison linéaire $a_0 I + a_1 A + a_2 A^2 + \cdots + a_{n^2} A^{n^2}$ qui est nulle mais dont certains coefficients ne sont pas nuls. Cela signifie que le polynôme $a_0 + a_1 X + a_2 X^2 + \cdots + a_{n^2} X^{n^2}$ n'est pas nul mais annule A. De plus, si un polynôme non nul P de coefficient dominant a annule A, alors $a^{-1}P$ est unitaire et annule A. Ceci entraîne qu'il existe un polynôme unitaire qui annule A, il en existe donc un de degré minimal.

Pour montrer l'unicité, considérons deux polynômes P et Q, unitaires et de degré minimal d, qui annulent A, c'est-à-dire que $P(A) = Q(A) = 0$. Alors on a $(P-Q)(A) = 0$ donc $P - Q$ annule A. Comme P et Q sont unitaires et de même degré d, on a $\deg(P - Q) < d$. Si $P - Q$ n'était pas nul, alors en le divisant par son coefficient dominant on obtiendrait un polynôme unitaire de degré strictement inférieur à d qui annulerait A, ce qui est impossible par définition de d. Par conséquent $P - Q = 0$ donc $P = Q$, ce qui prouve l'unicité. □

Une façon plus abstraite de formuler cette propriété est d'observer que l'ensemble des polynômes qui annulent A est un idéal de l'anneau $\mathbb{K}[X]$, or tout idéal de $\mathbb{K}[X]$ est principal, donc engendré par un seul polynôme (nécessairement de degré minimal), et le polynôme minimal de A est l'unique polynôme unitaire qui engendre cet idéal.

Dans la démonstration, on trouve que le degré du polynôme minimal est au plus n^2 par un simple argument sur la dimension de $M_n(\mathbb{K})$, mais on peut en fait montrer que son degré est au plus n, grâce qu théorème de Cayley-Hamilton qui prouve que le polynôme caractéristique de A, de degré n, annule A.

On étudiera les propriétés du polynôme minimal dans les exercices en fin de chapitre.

6.6 Théorème de Jordan

Le théorème 6.90 donne une caractérisation des matrices et endomorphismes diagonalisables par le fait d'être annulé par un polynôme scindé à racine simples. Dans le cas où le polynôme n'est pas à racines simples, le fait d'être diagonalisable n'est pas garanti.

6.93 *Exemple.* Considérons la matrice

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

Le polynôme caractéristique de A est $P_A = (2 - X)^2$, il est donc scindé mais son unique racine 2 est double. On peut facilement constater que l'on a bien $P_A(A) = 0$. Or A n'est pas diagonalisable, en effet son seul sous-espace propre est est E_2, associé à l'unique valeur propre 2, or $(x, y) \in E_2$ si et seulement si

$$A \begin{pmatrix} x \\ y \end{pmatrix} = 2 \begin{pmatrix} x \\ y \end{pmatrix} \Leftrightarrow \begin{cases} 2x + y = 2x \\ 2y = 2y \end{cases} \quad \Leftrightarrow y = 0$$

donc $E_2 = \mathrm{Vect}((1, 0))$ et $\dim(E_2) = 1$. Ainsi $\mathbb{R}^2$ n'est pas somme directe des sous-espaces propres de A, ce qui contredit la caractérisation du théorème 6.78.

Le théorème de Jordan, qui conclut ce chapitre, étend le théorème précédent au cas de polynômes qui ne sont pas forcément à racines simples. La matrice de l'exemple précédent donne la forme générale : le mieux que l'on puisse faire est de décomposer en une matrice qui a les valeurs propres sur la diagonale et quelques 1 juste au dessus de la diagonale, là où des valeurs propres apparaissent plusieurs fois.

6.94 **Définition (bloc de Jordan).** Un bloc de Jordan $J_n(\lambda)$ est une matrice carrée dont les seuls coefficients non nuls sont d'une part les coefficients diagonaux, qui sont tous égaux au scalaire λ, et d'autre part les coefficients sur diagonaux, qui sont tous égaux à 1.

6.95 *Exemple.* On a donc

$$J_2(\lambda) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \qquad J_3(\lambda) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}, \qquad J_4(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}.$$

De façon générale $J_n(\lambda)$ est la matrice (a_{ij}) telle que $a_{ii} = \lambda$ pour chaque i, $a_{i,i+1} = 1$ pour tout $i < n$ et $a_{ij} = 0$ si $j \notin \{i, i + 1\}$.

6.96 **Définition (nilpotence).** Soit E un espace vectoriel. Un endomorphisme $f : E \to E$ est *nilpotent* s'il existe un entier positif d tel que $f^d = 0$.

6.97 **Proposition.** *Soient E un espace vectoriel, f un endomorphisme nilpotent de E et $v \in E$ un vecteur non nul. Soit c le plus petit entier tel que $f^{c+1}(v) = 0$. Notons $Z(v)$ le sous-espace vectoriel de E engendré par les vecteurs $v, f(v), \ldots, f^c(v)$. Alors $Z(v)$ est stable par f, la famille $(f^c(v), \ldots, f(v), v)$ est une base de $Z(v)$ et la matrice de la restriction de f à $Z(v)$ dans cette base est $J_{c+1}(0)$. Le sous-espace $Z(v)$ est appelé le* sous-espace cyclique *de f engendré par v.*

Démonstration. Par définition, $x \in Z(v)$ si et seulement s'il existe $a_0, \ldots, a_c \in \mathbb{K}$ tels que $x = a_0 v + \cdots + a_c f^c(v)$. Comme $f(f^c(v)) = 0$, en voit qu'en appliquant f au vecteur x, on retrouve un élément de $Z(v)$. Ainsi $Z(v)$ est préservé par f. La famille $(v, f(v), \ldots, f^c(v))$ engendre $Z(v)$ par définition. Afin de montrer qu'elle est libre supposons que l'on ait $a_0 v + \cdots + a_c f^c(v) = 0$ et montrons que $a_i = 0$ pour tout i.

On procède par récurrence forte : soit $i \leq c$, supposons que pour tout $j < i$ on ait $a_j = 0$. On a donc $a_i f^i(v) + a_{i+1} f^{i+1}(v) + \cdots + a_c f^c(v) = 0$. En appliquant f^{c-i} à cette égalité on obtient $a_i f^c(v) + a_{i+1} f^{c+1}(v) + \cdots + a_c f^{2c-i}(v) = 0$, or tous les termes de cette égalité s'annulent sauf le premier donc on a $a_i f^c(v) = 0$ d'où $a_i = 0$ et l'étape de récurrence est validée.

Ainsi $(f^c(v), \ldots, v)$ est bien une base de $Z(v)$. Comme l'image du premier vecteur $f^c(v)$ de cette base par f est 0 et que tout autre vecteur de cette base est appliqué par f sur son prédécesseur, la matrice de la restriction de f a $Z(v)$ dans cette base est $J_{c+1}(0)$. □

6.98 **Lemme.** *Soient E un espace vectoriel et f un endomorphisme nilpotent de E. Soient $x, y \in \mathrm{Im}(f) \setminus \{0\}$ tels que $Z(x) \cap Z(y) = \{0\}$. Pour tous $v, w \in E$ tels que $f(v) = x$ et $f(w) = y$ on a $Z(v) \cap Z(w) = \{0\}$.*

Démonstration. Pour chaque vecteur x non nul, notons $c(x)$ le plus petit entier tel que $f^{c(x)+1}(x) = 0$ Quitte à échanger les rôles de x et y, on peut supposer que $c(x) \leq c(y)$. Par définition de v et w on a $c(v) = c(x) + 1$ et $c(w) = c(y) + 1$. Soit $z \in Z(v) \cap Z(w)$. Il existe donc $a_0, \ldots, a_{c(v)}, b_0, \ldots, b_{c(w)} \in \mathbb{K}$ tels que

$$z = a_0 v + \cdots + a_{c(v)} f^{c(v)}(v) = b_0 w + \cdots + b_{c(w)} f^{c(w)}(w).$$

En appliquant $f^{c(v)}$ à cette égalité, on obtient

$$a_0 f^{c(v)}(v) = b_0 f^{c(v)}(w) + \cdots + b_{c(w)} f^{c(v)+c(w)}(w).$$

Comme $c(v) \geq 1$, on en déduit que $a_0 f^{c(v)}(v) \in Z(x) \cap Z(y)$. Ainsi $a_0 f^{c(v)}(v) = 0$ or $f^{c(v)}(v) \neq 0$ donc $a_0 = 0$ et

$$0 = b_0 f^{c(v)}(w) + \cdots + b_{c(w)} f^{c(v)+c(w)}(w).$$

On a $c(v) \leq c(w)$ par hypothèse, les vecteurs $f^{c(v)}(w), \ldots, f^{c(w)}(w)$ sont linéairement indépendants et $f^{c(w)+1}(w) = \cdots = f^{c(v)+c(w)}(w) = 0$ donc on en déduit $b_0 = 0$. Ceci entraîne $z \in Z(x) \cap Z(y)$ et donc $z = 0$. □

6.99 **Proposition.** *Soient E un espace vectoriel de dimension finie et f un endomorphisme nilpotent de E. Alors E est somme directe d'un nombre fini de sous-espaces cycliques : il existe une famille finie $(x_1, \ldots, x_k)$ telle que*

$$E = Z(x_1) \oplus \cdots \oplus Z(x_k).$$

Démonstration. On procède par récurrence sur la dimension de E. Si $E = \{0\}$, le résultat est satisfait avec $k = 0$. Sinon, le noyau de f n'est pas réduit à 0 puisque f est nilpotent, donc $\dim(\mathrm{Im}(f)) < \dim(E)$. Par hypothèse de récurrence, il existe donc une famille finie $x_1, \ldots, x_r \in \mathrm{Im}(f) \setminus \{0\}$, tels que $\mathrm{Im}(f) = Z(x_1) \oplus \cdots \oplus Z(x_k)$. Pour chaque x_i, choisissons $v_i \in E$ tel que $f(v_i) = x_i$. Le lemme précédent montre que l'on a une somme directe $W = Z(v_1) \oplus \cdots \oplus Z(v_k)$.

Montrons que $E = \mathrm{Ker}(f) + W$. Pour tout $y \in E$, comme $f(y) \in Z(x_1) \oplus \cdots \oplus Z(x_k)$ on peut écrire $f(y) = y_1 + \cdots + y_k$ avec $y_i \in Z(x_i)$ pour chaque i, ainsi chaque y_i est combinaison linéaire de vecteurs de la forme $f^j(x_i)$ donc il existe un polynôme $P_i \in \mathbb{K}[X]$

tel que $y_i = P_i(f)(x_i)$, et on a finalement $f(y) = \sum_{i=1}^r P_i(f)(x_i)$. Posons alors $w = \sum_{i=1}^r P_i(f)(v_i)$. On a $w \in W$ et

$$f(y-w) = f(y) - f(w) = \sum_{i=1}^r P_i(f)(x_i) - \sum_{i=1}^r P_i(f)(f(v_i)) = \sum_{i=1}^r P_i(f)(x_i - f(v_i)) = 0$$

puisque $f(v_i) = x_i$. On a donc $y - w \in \mathrm{Ker}(f)$ or $w \in W$ donc $y \in \mathrm{Ker}(f) + W$, d'où on conclut $E = \mathrm{Ker}(f) + W$.

Soit Y un supplémentaire de $\mathrm{Ker}(f) \cap W$ dans $\mathrm{Ker}(f)$. Comme $E = \mathrm{Ker}(f) + W$ on a donc $E = Y \oplus W$. Soit $(y_1, \dots, y_m)$ une base de Y, pour chaque i on a $f(y_i) = 0$ donc $Z(y_i) = \mathrm{Vect}(y_i)$ et donc $Y = Z(y_1) \oplus \cdots \oplus Z(y_m)$. On en déduit $E = Y \oplus W = Z(y_1) \oplus \cdots \oplus Z(y_m) \oplus Z(v_1) \oplus \cdots \oplus Z(v_k)$. □

6.100 **Théorème (Jordan).** *Soient E un espace vectoriel de dimension finie sur un corps $\mathbb{K}$ et f un endomorphisme de E. Il existe une base de E dans laquelle la matrice de f est formée de blocs de Jordan si et seulement s'il existe un polynôme $P \in \mathbb{K}[X]$ scindé tel que $P(f) = 0$.*

Démonstration. Supposons qu'il existe une base $\mathcal{B}$ dans laquelle la matrice de f soit formée de blocs de Jordan :

$$M(f) = \begin{pmatrix} J_{n_1}(\lambda_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J_{n_r}(\lambda_r) \end{pmatrix}.$$

Le polynôme caractéristique d'un bloc de Jordan $J_n(\lambda)$ est $(\lambda - X)^n$ donc on obtient

$$P_f = (\lambda_1 - X)^{n_1} \cdots (\lambda_r - X)^{n_r}.$$

donc P_f est scindé. Si v est un vecteur de la base $\mathcal{B}$, il existe un i tel que $v \in \mathrm{Ker}(f - \lambda_i \,\mathrm{id})^{n_i}$ donc $P_f(f)(v) = 0$, par conséquent $P_f(f) = 0$.

Réciproquement, supposons qu'il existe $P \in \mathbb{K}[X]$ scindé tel que $P(f) = 0$. Quitte à multiplier P par une constante non nulle, on peut donc écrire

$$P = (\lambda_1 - X)^{n_1} \cdots (\lambda_r - X)^{n_r}$$

pour une certaine famille $(\lambda_1, \dots, \lambda_r)$ d'éléments de $\mathbb{K}$ deux à deux distincts et une certaine famille $(n_1, \dots, n_r)$ d'entiers naturels non nuls. Les polynômes $(\lambda_i - X)^{n_i}$ sont premiers entre eux donc le lemme des noyaux implique que

$$E = \mathrm{Ker}(P(f)) = \mathrm{Ker}(\lambda_1 - f)^{n_1} \oplus \cdots \oplus \mathrm{Ker}(\lambda_r - f)^{n_r}.$$

Pour chaque i, comme f et $(\lambda_i - f)^{n_i}$ commutent, le sous-espace $\mathrm{Ker}(\lambda_i - f)^{n_i}$ est préservé par f. La restriction de $\lambda_i - f$ à $\mathrm{Ker}(\lambda_i - f)^{n_i}$ est nilpotente, il existe donc une base $\mathcal{B}_i$ de $\mathrm{Ker}(\lambda_i - f)^{n_i}$ dans laquelle la matrice de la restriction de $\lambda_i - f$ à $\mathrm{Ker}(\lambda_i - f)^{n_i}$ est constituée de blocs de Jordan de valeur propre 0. Dans $\mathcal{B}_i$, la restriction de f à $\mathrm{Ker}(\lambda_i - f)^{n_i}$ est donc constituée de blocs de Jordan de valeur propre λ_i. Soit $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_r)$ la base de E obtenue en rassemblant les $\mathcal{B}_i$, alors dans $\mathcal{B}$ la matrice de f est constituée de blocs de Jordan. □