

Algèbre linéaire et géométrie

R. Petit

année académique 2015 - 2016

Contents

1	Rappels	1
1.1	Ensembles et fonctions	1
1.2	Trigonométrie	1
1.3	Nombres complexes	2
2	Les polynômes	5
2.1	L'algorithme d'Euclide sur des entiers	5
2.2	L'algorithme d'Euclide sur des polynômes	5
2.3	Racines	6
2.4	Fonctions symétriques	8
3	Géométrie analytique dans \mathbb{R}^3 : rappels	10
3.1	Vecteurs	10
3.2	Droites et plans	10
3.3	Produit scalaire et orthogonalité	11
3.4	Distances dans \mathbb{R}^3	12
3.5	Angles dans \mathbb{R}^3	13
3.6	Produit vectoriel	14
4	Systèmes d'équations linéaires et algèbre matricielle	16
4.1	Définitions	16
4.2	La méthode de Gauss	16
4.3	Calcul matriciel	17
4.4	Structure des solutions d'un système linéaire d'équations	18
5	Espaces vectoriels	20
5.1	Définitions	20
5.2	Sous-espaces	21
5.3	Parties libres et génératrices	22
5.4	Bases et dimension	25
5.4.1	Cardinalité	26
5.4.2	Dimension	27
5.4.3	Dépendance base et système de coordonnées	28
6	Relations entre espaces vectoriels	30
6.1	Isomorphismes	30
6.2	Somme directe	33
6.3	Groupe linéaire	34

7 Algèbre matricielle	35
7.1 Matrice associée à une transformation linéaire	35
7.2 Opérations sur les matrices	35
7.3 Méthode de Gauss pour inverser une matrice	37
7.4 Changement de base	37
7.5 Dualité	39
7.6 Matrices transposées	40
8 Permutations et Déterminants	42
8.1 Définitions	42
8.2 Signe d'une permutation	44
8.3 Déterminants	46
8.4 Matrices et applications linéaires	49
8.5 Cofacteurs, matrices adjointes et inverses	52
9 Algèbre d'ensembles et résultats sur les espaces vectoriels	56
9.1 Définitions	56
9.2 Sous-espaces propres	56
9.3 Équation caractéristique	60
9.3.1 Opérateurs diagonalisables	62
9.3.2 Opérateurs triangulables	63
9.4 Formule canonique de Jordan	64
9.5 Polynôme minimal d'un opérateur et propriétés des polynômes	65
9.6 Sous-espaces généralisés	67
9.7 La forme canonique de Jordan	70
9.7.1 Existence de la forme canonique de Jordan	70
9.7.2 Unicité de la forme canonique de Jordan	72
10 Espaces euclidiens et hermitiens	74
10.1 Définitions	74
10.1.1 Orthogonalité	74
10.1.2 Propriétés de la norme	75
10.1.3 base orthonormale	75
10.1.4 Angles entre vecteurs	76
10.1.5 Métrique	76
10.2 Sous-espaces orthogonaux	77
10.2.1 Droites, hyperplans et projections orthogonales dans \mathbb{R}^n	77
10.3 Isométries	78
10.4 Déterminants et valeurs propres	79
10.5 Le groupe orthogonal O_3	80
11 Matrices symétriques et hermitiennes	81
11.1 Définitions	81
11.2 Diagonalisation de matrices hermitiennes et symétriques	81
11.3 Formes bilinéaires et quadratiques	82
11.3.1 Changement de base pour les formes quadratiques	83
11.3.2 Signature et théorème de Sylvester	84
12 Approche naïve des affinités	87
12.1 Introduction	87
12.2 Lieu des solutions d'une équation polynômiale	87

1 Rappels

1.1 Ensembles et fonctions

Définition 1.1. Soient X, Y deux ensembles. Une fonction $f : X \rightarrow Y$ est une *correspondance* qui associe chaque élément de X à un élément unique de Y que l'on appelle $f(x) \in Y$. X est appelé le domaine de f et Y est le codomaine.

Définition 1.2.

- L'image de $A \subseteq X$ par $f : X \rightarrow Y$ est $f(A) := \{f(x) \text{ t.q. } x \in A\}$. L'image de f est $f(X) \subseteq Y$.
- La préimage de $B \subseteq Y$ par f est $f^{-1}(B) := \{x \in X \text{ t.q. } f(x) \in B\}$. La préimage de $y \in Y$ est donnée par $f^{-1}(y) := f^{-1}(\{y\})$.

Définition 1.3. Soit $f : X \rightarrow Y$.

- Une fonction $f : X \rightarrow Y$ est injective si $\forall x, x' \in X : f(x) = f(x') \Rightarrow x = x'$.
- Une fonction $f : X \rightarrow Y$ est surjective si $\forall y \in Y : \exists x \in X \text{ t.q. } f(x) = y$ (ou encore si $f(X) = Y$).
- Une fonction $f : X \rightarrow Y$ est bijective si elle est injective et surjective.

Définition 1.4. Soit $f : X \rightarrow Y$ une fonction bijective. On définit la fonction inverse f^{-1} de f par $f^{-1} : Y \rightarrow X : y \mapsto x \text{ t.q. } f(x) = y$.

Définition 1.5. Soient X, Y, Z trois ensembles. Soient $f : X \rightarrow Y, g : Y \rightarrow Z$ deux fonctions. On définit la composée $g \circ f : X \rightarrow Z : x \mapsto g(f(x))$.

Lemme 1.6. La composée $f \circ f^{-1} = \text{Id}_Y$ et la composée $f^{-1} \circ f = \text{Id}_X$.

Définition 1.7. Soient X, Y deux ensembles. On définit le produit cartésien :

$$X \times Y := \{(x, y) \text{ t.q. } x \in X, y \in Y\}.$$

Définition 1.8. Soit $\{X_i\}_{i \in I}$ une collection d'ensembles. On définit :

$$\prod_{i \in I} X_i := \{(x_i)_{i \in I} \text{ t.q. } x_i \in X_i\}.$$

Définition 1.9. Soit $f : X \rightarrow Y$ une fonction. On définit le graphe de f par :

$$\Gamma_f := \{(x, f(x)) \text{ t.q. } x \in X\} \subseteq X \times Y.$$

1.2 Trigonométrie

Définition 1.10. Les fonctions $\cos, \sin : \mathbb{R} \rightarrow \mathbb{R}$ représentent respectivement l'abscisse et l'ordonnée d'un point dont les coordonnées polaires sont $(1, \theta)$.

Lemme 1.11. Soit $\theta \in \mathbb{R}$. Il existe un unique $k \in \mathbb{Z}$ tel que $\theta + 2k\pi \in [0, 2\pi[$.

Théorème 1.12. $\forall \theta \in \mathbb{R} : \sin(\theta)^2 + \cos(\theta)^2 = 1$.

Démonstration. Par Pythagore. □

Lemme 1.13. Les fonctions \cos, \sin peuvent être réduites à un domaine afin d'être bijectives (si ce domaine est continu, il doit être de longueur π).

Définition 1.14.

- On définit les fonctions inverses $\arccos : [-1, 1] \rightarrow [0, \pi]$ et $\arcsin : [-1, 1] \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.
- On définit également la fonction $\tan :]-\frac{\pi}{2}, \frac{\pi}{2}[\rightarrow \mathbb{R} : \theta \mapsto \frac{\sin \theta}{\cos \theta}$ et son inverse $\arctan : \mathbb{R} \rightarrow]-\frac{\pi}{2}, \frac{\pi}{2}[$.

1.3 Nombres complexes

Définition 1.15. On définit l'ensemble des complexes $\mathbb{C} := \{a + ib \text{ t.q. } (a, b) \in \mathbb{R}^2\}$. Où i est l'unité imaginaire telle que $i^2 = -1$.

Définition 1.16. Soit $z = a + ib \in \mathbb{C}$ un complexe. On définit $\Re z := a$ et $\Im z := b$ respectivement la partie réelle et la partie imaginaire de z .

Définition 1.17. Soit $z \in \mathbb{C}$ un complexe. z peut s'écrire soit sous sa forme cartésienne $a + ib$ soit sous sa forme polaire $\rho \exp(i\theta)$. ρ est le module de z et θ est son argument. Il existe une relation entre (a, b) et (ρ, θ) : $(a, b) = (\rho \cos \theta, \rho \sin \theta)$.

Remarque. On peut visualiser l'ensemble des réels par $\mathbb{R} \equiv \{z \in \mathbb{C} \text{ t.q. } \Im z = 0\} \subseteq \mathbb{C}$.

Définition 1.18. On définit la somme et le produit de deux complexes $z = a + ib, w = c + id \in \mathbb{C}$ comme suit :

$$\begin{aligned} w + z &:= (a + ib) + (c + id) = (a + c) + (b + d)i && \in \mathbb{C} \\ wz &:= (a + ib)(c + id) = (ac - bd) + (ad + bc)i && \in \mathbb{C} \end{aligned}$$

On définit également le module d'un complexe $z = a + ib = \rho \exp(i\theta) \in \mathbb{C}$ par $|z| = \sqrt{a^2 + b^2} = \rho \in \mathbb{R}^+$.

Proposition 1.19 (Propriétés des opérations complexes). Soient $v, w, z \in \mathbb{C}$. Alors :

1. $z + w = w + z$;
2. $z + (w + v) = (z + w) + v$;
3. $z + 0 = z$;
4. $z + (-1)z = z - z = 0$;
5. $zw = wz$;
6. $z(wv) = (zw)v$;
7. $z(w + v) = zw + zv$;
8. $1z = z$;
9. $|z| \geq 0$ avec $|z| = 0 \iff z = 0$;
10. $|zw| = |z||w|$;
11. $|z + w| \leq |z| + |w|$.

Démonstration. Les points 1 à 8 se démontrent par les mêmes propriétés sur les nombres réels. Le point 9 découle du fait qu'une racine carrée est toujours positive et $\sqrt{a^2 + b^2} = 0 \iff a^2 + b^2 = 0$. Le point 10 se montre par les propriétés de la racine carrée. Le point 11 vient de l'inégalité de Cauchy-Schwartz. \square

Remarque. La somme et le produit peuvent également être exprimés en coordonnées polaires. De plus, le produit est plus instinctif en coordonnées polaires : $\forall z = \rho_z \exp(i\theta_z), w = \rho_w \exp(i\theta_w) \in \mathbb{C} : zw = \rho_z \rho_w \exp(i(\theta_z + \theta_w))$.

Définition 1.20. Soit $z = a + ib = \rho \exp(i\theta) \in \mathbb{C}$ un complexe. On définit son conjugué $\bar{z} = a - ib = \rho \exp(-i\theta) \in \mathbb{C}$.

Proposition 1.21 (Propriétés du conjugué). Soient $z, w, \bar{z}, \bar{w} \in \mathbb{C}$ deux complexes et leur conjugué.

1. $\overline{z + w} = \bar{z} + \bar{w}$;
2. $\overline{zw} = \bar{z} \cdot \bar{w}$;
3. $z\bar{z} = |z|^2$.

Démonstration. Trivial par la définition. \square

Définition 1.22. Par la propriété 3 ci-dessus, on définit pour tout $z \in \mathbb{C} \setminus \{0\}$ un inverse multiplicatif :

$$z^{-1} := \frac{\bar{z}}{|z|^2}.$$

Définition 1.23 (L'identité d'Euler). On définit l'exponentielle complexe par $\exp(i\theta) := \cos \theta + i \sin \theta$.

Proposition 1.24. Soient $z, w \in \mathbb{C}$ deux complexes. Alors $zw = 0 \Rightarrow z = 0 \vee w = 0$.

Démonstration. Soient $z, w \in \mathbb{C}$. On sait que $zw = \rho_z \rho_w \exp(i(\theta_z + \theta_w)) = 0$. Or \exp est toujours strictement positive. Dès lors, $\rho_z = 0$ ou $\rho_w = 0$. \square

Théorème 1.25. Tout polynôme à coefficients complexes de degré n a n racines complexes.

Proposition 1.26. Soit $z = \rho \exp(i\theta) \in \mathbb{C}$. Le polynôme $x^n - z$ admet n racines complexes distinctes données par :

$$w_k = \sqrt[n]{\rho} \exp\left(i \frac{\theta + 2k\pi}{n}\right).$$

Démonstration. On sait par le théorème 1.25 que le polynôme admet n racines. Montrons que w_k est une racine de $x^n - z$. En étendant la définition du produit vu ci-dessus, on sait que :

$$\prod_{j=1}^n z_j = \prod_{j=1}^n (\rho_j) \exp\left(i \sum_{j=1}^n \theta_j\right).$$

On peut étendre cette formule pour la puissance :

$$z^n = \rho^n \exp(ini\theta).$$

Dès lors, en prenant une racine w_k , on a :

$$w_k^n = \left(\sqrt[n]{\rho} \exp\left(i \frac{\theta + 2k\pi}{n}\right) \right)^n = \sqrt[n]{\rho}^n \exp\left(in \frac{\theta + 2k\pi}{n}\right) = \rho \exp(i(\theta + 2k\pi)).$$

On a donc effectivement $w_k^n - z = 0$.

□

2 Les polynômes

2.1 L'algorithme d'Euclide sur des entiers

Définition 2.1. Soient $(a, b) \in \mathbb{Z} \times \mathbb{Z}_0$. Il existe un unique $d \in \mathbb{Z}_0$ tel que d est le plus grand nombre qui divise a et b simultanément. On note ce nombre $\text{GCD}(a, b)$.

Remarque. Pour tout couple (a, b) , notons que $0 \leq \text{GCD}(a, b) < \max\{|a|, |b|\}$.

Définition 2.2. Soient $(a, b) \in \mathbb{Z} \times \mathbb{Z}_0$. Il existe un unique $q \in \mathbb{Z}$ tel que $r := a - qb$ satisfait $0 \leq r < |b|$. q et r sont respectivement le reste et le quotient de la division euclidienne de a par b .

Algorithme 2.3 (Algorithme d'Euclide). L'algorithme d'Euclide est un moyen très efficace pour déterminer le GCD entre deux nombres entiers. En posant $r_{-1} := a$ et $r_0 := b$. Ensuite pour $i > 0$, on trouve les uniques q_i, r_i tels que :

$$r_{i-2} = q_i r_{i-1} + r_i,$$

où $0 \leq r_i < |r_{i-1}|$. On s'arrête lorsque $r_i = 0$.

Théorème 2.4. Soient a, b deux entiers avec b non-nul. Prenons $(r_i)_i, (q_i)_i$ donnés par l'algorithme d'Euclide. Soit $k \geq 0$ le plus petit entier tel que $r_{k+1} = 0$. Alors $\text{GCD}(a, b) := |r_k|$

Démonstration. Montrons d'abord que tout diviseur de a et de b divise r_k et ensuite montrons que r_k divise a et b .

On sait pour tout i que $r_{i-2} = q_i r_{i-1} + r_i$. Donc si d divise r_{i-1} et r_{i-2} , il doit diviser r_i . Dès lors les diviseurs (en particulier le GCD) de a et b doivent diviser r_1 et donc r_2 , etc. jusque r_k .

Montrons que r_k divise a et b . On sait que $r_{k+1} = 0$. Donc $r_{k-1} = q_{k+1} r_k + r_{k+1} = q_{k+1} r_k$. Donc r_k divise r_{k-1} . De plus, $r_{k-2} = q_k r_{k-1} + r_k$. r_k divise donc r_{k-2} , etc. jusque r_{-1} .

On a donc montré que tous les diviseurs de a et de b divisent r_k et que r_k divise a et b . □

Corollaire 2.5. Soient a, b deux entiers. Il existe deux uniques entiers y, z tels que $ya + zb = \text{GCD}(a, b)$.

Démonstration. Par l'algorithme d'Euclide, on sait qu'il existe des uniques $(r_i)_i, (q_i)_i$. Soit $k \geq 0$ le plus petit entier tel que $r_{k+1} = 0$. On sait que $r_k = r_{k-2} - q_k r_{k-1}$. Or r_{k-1} peut s'exprimer en fonction de r_{k-2} et r_{k-3} . Donc $r_k = \alpha_2 r_{k-2} + \beta_3 r_{k-3}$. On répète ce procédé k fois. On a donc $\text{GCD}(a, b) = r_k = \alpha_k r_0 + \beta_{k+1} r_{-1} = \alpha_k b + \beta_{k+1} a$ avec $\alpha, \beta \in \mathbb{Z}$. □

2.2 L'algorithme d'Euclide sur des polynômes

Définition 2.6. On définit l'anneau des polynômes à coefficients réels par $\mathbb{R}[x]$.

Définition 2.7. Soit $f \in \mathbb{R}[x]$. $f(x) = \sum_{i=1}^n a_i x^i$ où $a_i \in \mathbb{R}$ pour tout i et $a_n \neq 0$. n est le degré du polynôme.

On note :

$$\deg f(x) := n.$$

Remarque. Par convention, le degré du polynôme $x \mapsto 0$ est $-\infty$.

Définition 2.8. Soit $f \in \mathbb{R}[x]$ un polynôme de degré n . f est monique (ou unitaire) si le coefficient du monôme de degré n vaut 1 (si $a_n = 1$).

Définition 2.9. Soient $f, g \in \mathbb{R}[x]$. On définit $\text{GCD}(f(x), g(x))$ par l'unique polynôme monique de degré maximal qui divise $f(x)$ et $g(x)$.

Définition 2.10. Soit $f \in \mathbb{R}[x]$ un polynôme de degré n . Pour tout $0 \leq k \leq n$, on définit $[x^k]f(x)$ par le coefficient du monôme de degré k de f .

Théorème 2.11. Soient $f, g \in \mathbb{R}[x]$ tels que $g(x) \neq 0$. Il existe deux uniques polynômes $q, r \in \mathbb{R}[x]$ tels que $f(x) = q(x)g(x) + r(x)$ satisfaisant $\deg r(x) < \deg g(x)$. q est appelé le quotient et r le reste de la division de f par g .

Algorithme 2.12 (Algorithme d'Euclide sur des polynômes). Soient $f, g \in \mathbb{R}[x]$ tels que $g(x) \neq 0$. Comme pour l'algorithme sur des entiers, posons $r_{-1}(x) := f(x)$ et $r_0(x) = g(x)$. Dès lors, pour tout $i > 0$, il existe des uniques q_i, r_i tels que :

$$r_{i-2}(x) = q_i(x)r_{i-1}(x) + r_i(x)$$

où $\deg r_i(x) < \deg r_{i-1}(x)$. On s'arrête lorsque $r_i(x) = 0$.

Théorème 2.13. Soient $f, g \in \mathbb{R}[x]$. Prenons $(r_i)_i, (q_i)_i \subset \mathbb{R}[x]$ donnés par l'algorithme de division de f par g . Soit le plus petit $k \in \mathbb{N}$ tel que $r_{k+1}(x) = 0$. $r_k(x)$ est le GCD non-normalisé de $f(x)$ et $g(x)$. Et on définit :

$$\text{GCD}(f(x), g(x)) := \frac{r_k(x)}{[x^{\deg r_k(x)}]r_k(x)}.$$

Démonstration. Comme pour l'algorithme sur les nombres entiers, si un polynôme divise $f(x)$ et $g(x)$, alors il doit diviser $r_i(x)$ pour tout i . De plus, comme par hypothèse $r_{k+1}(x) = 0$, $r_{k-1}(x) = q_{k+1}(x)r_k(x)$. Donc $r_k(x)$ divise $r_{k-1}(x)$. Et donc $r_k(x)$ divise également $r_{k-2}(x)$, etc. jusque $f(x)$ et $g(x)$. On a donc bien que tout diviseur simultané de f et g divise r_k et r_k divise à la fois f et g . Dès plus, en divisant r_k par le coefficient de son monôme dominant, on en fait un polynôme monique. \square

Corollaire 2.14. Soient $f, g \in \mathbb{R}[x]$. Il existe deux uniques polynômes $y(x), z(x)$ tels que $y(x)f(x) + z(x)g(x) = \text{GCD}(f(x), g(x))$.

Démonstration. Preuve similaire à l'algorithme sur des entiers. \square

2.3 Racines

Définition 2.15. Soit $P \in \mathbb{C}[x]$ un polynôme complexe. Une racine de P est un complexe $z \in \mathbb{C}$ est une complexe tel que $P(z) = 0$.

Remarque. Le théorème fondamental (théorème 1.25) dit qu'un polynôme complexe $f(x) \in \mathbb{C}[x]$ a exactement $\deg f(x)$ racines complexes en comptant la multiplicité. On peut donc réécrire le polynôme f comme :

$$f(x) = c \prod_{i=1}^{\deg f(x)} (x - z_i)$$

où $\{z_i \text{ t.q. } 1 \leq i \leq \deg(f(x))\}$ est l'ensemble des racines (pas obligatoirement distinctes) de f .

Définition 2.16. Soit $f \in \mathbb{C}[x]$ un polynôme complexe définie par :

$$f(x) = \sum_{i=1}^n a_i x^i.$$

On définit son polynôme conjugué par :

$$\bar{f}(x) = \sum_{i=1}^n \bar{a_i} x^i.$$

Remarque. Si $f \in \mathbb{R}[x] \subset \mathbb{C}[x]$, alors son conjugué $\bar{f}(x) = f(x)$ car $\forall x \in \mathbb{R} : \bar{x} = x$.

Lemme 2.17. Soit $f \in \mathbb{C}[x]$. Alors $\bar{f}(\bar{z}) = \overline{f(z)}$.

Démonstration.

$$\bar{f}(\bar{z}) = \sum_{i=1}^n \bar{a_i} \cdot \bar{z} = \sum_{i=1}^n \overline{a_i z} = \overline{\sum_{i=1}^n a_i z} = \overline{f(z)}.$$

□

Proposition 2.18. Soit $f \in \mathbb{R}[x]$ un polynôme à coefficients réels. Le théorème fondamental (théorème 1.25) dit que f a exactement $\deg(f(x))$ racines complexes. Soit z une telle racine. Alors le conjugué de z est également une racine de f .

Démonstration. Soient $f \in \mathbb{R}[x]$ et z une racine de f . Alors $0 = f(z) = \overline{f(z)} = \bar{f}(\bar{z}) = f(\bar{z})$.

□

Proposition 2.19. Soit $f \in \mathbb{R}[x]$ un polynôme complexe. Soit $z_0 \in \mathbb{C}$ une racine complexe de f telle que $\Im z \neq 0$. Alors le polynôme suivant :

$$(x - z_0)(x - \bar{z}_0)$$

est un diviseur de $f(x)$.

Démonstration. Par hypothèse, une telle racine z_0 existe. Dès lors le polynôme $(x - z_0)(x - \bar{z}_0)$ est un polynôme réel. En effet : $(x - z_0)(x - \bar{z}_0) = x^2 - z_0 x - \bar{z}_0 x + z_0 \bar{z}_0 = x^2 - 2\Re z_0 x + |z_0|^2 \in \mathbb{R}$. Dès lors, on peut trouver les polynômes quotient et reste tels que :

$$f(x) = q(x)(x - z_0)(x - \bar{z}_0) + r(x).$$

Ensuite, montrons que $\deg r(x) = 0$. En effet, $0 \leq \deg r(x) < 2$ car 2 est le degré du polynôme diviseur. Supposons par l'absurde que $\deg r(x) = 1$. Alors $r(x) = ax + b$ pour $a, b \in \mathbb{R}$. L'évaluation en z_0 donne donc $0 = f(z_0) = r(z_0) = az_0 + b$. Donc $\Im r(x) = \Im(az_0 + b) = a\Im z_0 \neq 0$. Il y a donc contradiction.

Le reste étant constant (de degré nul), le polynôme $(x - z_0)(x - \bar{z}_0)$ doit diviser $f(x)$.

□

Corollaire 2.20. Tout polynôme $f(x) \in \mathbb{R}[x]$ se factorise en un produit de polynômes de degré 2 à racines purement complexes conjuguées et de polynômes de degré 1 à racines réelles.

Remarque. Si $f(x) \in \mathbb{R}[x]$ est de degré impair, alors il doit avoir une racine réelle.

2.4 Fonctions symétriques

Définition 2.21. Soit $f \in \mathbb{C}[x]$ un polynôme monique de degré n . Soient z_1, \dots, z_n ses n racines, multiplicité comptée. On définit, pour tout $1 \leq k \leq n$, les fonctions symétriques élémentaires de f par :

$$e_k(z_1, \dots, z_n) := \sum_{1 \leq i_1 < \dots < i_k \leq n} \prod_{j=1}^k z_{i_j}.$$

De plus, on définit $e_0(z_1, \dots, z_n) := 1$.

Lemme 2.22. Soit un polynôme complexe monique $f \in \mathbb{C}[x]$ de degré n et ses fonctions symétriques élémentaires $e_k : \mathbb{C}^n \rightarrow \mathbb{C}$. Alors pour tout $1 \leq k \leq n$, on a :

$$e_k(z_1, \dots, z_n) = e_k(z_1, \dots, z_{n-1}) + z_n e_{k-1}(z_1, \dots, z_{n-1}).$$

Démonstration. Cette formule donne deux termes :

$$\sum_{1 \leq i_1 < \dots < i_k \leq n-1} \prod_{j=1}^k z_{i_j} + z_n \sum_{1 \leq i_1 < \dots < i_{k-1} \leq n-1} \prod_{j=1}^{k-1} z_{i_j}.$$

En posant $i_k = n$ dans le terme de droite, on obtient :

$$e_k(z_1, \dots, z_n) = \sum_{1 \leq i_1 < \dots < i_k \leq n-1} \prod_{j=1}^k z_{i_j} + \sum_{1 \leq i_1 < \dots < i_{k-1} \leq n-1} \prod_{j=1}^k z_{i_j}.$$

On a, dans le terme de gauche les $\binom{n-1}{k}$ termes ne contenant pas z_n et dans le terme de droite les $\binom{n-1}{k-1}$ termes contenant z_n . On a donc bien les $\binom{n}{k}$ termes de e_k . \square

Proposition 2.23. Soit un polynôme monique $f \in \mathbb{C}[x]$ de degré n , soient $(z_i)_{1 \leq i \leq n}$ ses racines et soient $e_k : \mathbb{C}^n \rightarrow \mathbb{C}$ ses fonctions symétriques élémentaires. Alors, si f est défini par :

$$f(x) = \sum_{i=1}^n a_i x^i,$$

avec $a_i \in \mathbb{C}$ et $a_n \neq 0$, alors, pour tout $0 \leq i \leq n$:

$$e_{n-i}(z_1, \dots, z_n) = (-1)^{n-i} a_i.$$

Démonstration. Prouvons le par récurrence sur n . Montrons d'abord pour $n = 1$:

$$f(x) = x - z_1 = a_1 x + a_0.$$

De plus :

$$\begin{aligned} (-1)^{1-0} a_0 &= e_{1-0}(z_1) \iff a_0 = -e_1(z_1) = -z_1. \\ (-1)^{1-1} a_1 &= e_{1-1}(z_1) \iff a_1 = 1. \end{aligned}$$

Le cas initial est donc bon. Montrons maintenant le pas de récurrence. Supposons cela vrai pour $n - 1$ et montrons-le pour n :

$$\begin{aligned}
f(x) &= \prod_{i=1}^n (x - z_i) = x \prod_{i=1}^{n-1} (x - z_i) - z_n \prod_{i=1}^{n-1} (x - z_i) = x \sum_{i=0}^{n-1} b_i x^i - z_n \sum_{i=0}^{n-1} b_i x^i \\
&\stackrel{\text{hyp.rec.}}{=} x \sum_{i=0}^{n-1} (-1)^{n-1-i} e_{n-1-i}(z_1, \dots, z_{n-1}) x^i - z_n \sum_{i=0}^{n-1} (-1)^{n-1-i} e_{n-i-1}(z_1, \dots, z_{n-1}) x^i \\
&= \sum_{i=1}^n (-1)^{n-i} e_{n-i}(z_1, \dots, z_{n-1}) x^i + z_n \sum_{i=0}^{n-1} (-1)^{n-i} e_{n-i-1}(z_1, \dots, z_{n-1}) x^i \\
&= (-1)^n e_{n-1}(z_1, \dots, z_{n-1}) z_n + \sum_{i=1}^{n-1} (-1)^{n-i} (e_{n-i}(z_1, \dots, z_{n-1}) + z_n e_{n-i-1}(z_1, \dots, z_{n-1})) x^i + x^n \\
&= \sum_{i=0}^n (-1)^{n-i} e_{n-i}(z_1, \dots, z_n) x^i.
\end{aligned}$$

On a donc bien prouvé l'égalité pour chaque coefficient, donc pour tout i et en faisant la preuve par récurrence sur n . □

3 Géométrie analytique dans \mathbb{R}^3 : rappels

3.1 Vecteurs

Définition 3.1. Soit \mathbb{R}^3 , l'espace à trois dimensions. Pour tout $v \in \mathbb{R}^3$, on définit (naïvement) le vecteur v comme une flèche partant de l'origine $\mathcal{O} = (0, 0, 0)$ au point v . On définit donc les coordonnées de v comme étant $v_1, v_2, v_3 \in \mathbb{R}$ tels que $v = (v_1, v_2, v_3)$.

Définition 3.2. Soient $v, w \in \mathbb{R}^3$, deux vecteurs et $\lambda \in \mathbb{R}$ un scalaire. On définit la somme interne et le produit externe par :

$$\begin{aligned}v + w &:= (v_1 + w_1, v_2 + w_2, v_3 + w_3), \\ \lambda v &:= (\lambda v_1, \lambda v_2, \lambda v_3).\end{aligned}$$

Remarque. Pour tout $\lambda \in \mathbb{R}$ scalaire, le vecteur λv se trouve sur la droite engendrée par le vecteur v (droite passant par \mathcal{O} et v). On parle alors de droite vectorielle $v\mathbb{R}$.

Proposition 3.3. Ces définitions d'opérations ont les propriétés suivantes. Pour tout $u, v, w \in \mathbb{R}^3, \lambda, \mu \in \mathbb{R}$:

1. $v + w = w + v$;
2. $(u + v) + w = u + (v + w)$;
3. $\lambda(\mu v) = (\lambda\mu)v$;
4. $0v = 0$;
5. $0 + v = v$;
6. $1v = v$;
7. $v - v := v + (-v) = v + (-1)v = 0 = (0, 0, 0)$;
8. $\lambda(v + w) = \lambda v + \lambda w$;
9. $(\lambda + \mu)v = \lambda v + \mu v$.

Démonstration. Trivial par ces mêmes opérations sur \mathbb{R} . □

3.2 Droites et plans

Définition 3.4. Soit $v \in \mathbb{R}^3$ un vecteur. Alors $v\mathbb{R}$ est la droite engendrée par le vecteur v . Son équation vectorielle est $D : \lambda v$. Soit $w \in \mathbb{R}^3$ un autre vecteur. L'équation vectorielle de la droite parallèle à D et passant par w est $D_2 : w + \lambda v$. v est appelé le vecteur directeur de ces droites.

Définition 3.5. Soient $v, w \in \mathbb{R}^3$ deux vecteurs non colinéaires. L'équation vectorielle du plan regroupant tous les points de la droite $D : \lambda v$ et ses translations par les vecteurs μw est $\pi : \lambda v + \mu w$. Ce plan passe par l'origine car il contient tous les points de λv . Soit u un troisième vecteur. L'équation vectorielle du plan parallèle à Π et passant par u est $\Pi_u : u + \lambda v + \mu w$.

Définition 3.6. Les deux définitions précédentes concernent les équations dites vectorielles. Une droite ou un plan peuvent également être exprimés en coordonnées paramétriques ou cartésiennes. Soit $P = (x, y, z) \in \mathbb{R}^3$. Soient $D : \lambda v + w, \Pi : \lambda v + \mu w + u$.

1.

$$P \in D \iff \exists \lambda \in \mathbb{R} \text{ t.q. } \begin{cases} x = \lambda v_1 + w_1 \\ y = \lambda v_2 + w_2 \\ z = \lambda v_3 + w_3 \end{cases}$$

2.

$$P \in \Pi \iff \exists (\lambda, \mu) \in \mathbb{R}^2 \text{ t.q. } \begin{cases} x = \lambda v_1 + \mu v_2 + u_1 \\ y = \lambda v_2 + \mu v_2 + u_2 \\ z = \lambda v_3 + \mu v_3 + u_3 \end{cases}.$$

Ces équations sont les équation paramétriques. Pour trouver les équations cartésiennes, il faut résoudre le système (en λ, μ pour le plan et en λ pour la droite). On a donc :

$$\Pi : ax + by + cz = d.$$

Et pour la droite D, on a :

$$D : \begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \end{cases}.$$

Remarque. L'équation cartésienne d'une droite est un couple de coordonnées cartésiennes de plan. Cela peut se comprendre comme suit : l'ensemble des points de la droite sont les points satisfaisant simultanément la première équation et la seconde, ou encore tous les points étant dans l'intersection entre les deux plans. En coordonnées cartésiennes, une droite est définie par deux plans et un point par trois plans (non-parallèles).

3.3 Produit scalaire et orthogonalité

Définition 3.7. Soient $v, w \in \mathbb{R}^3$ deux vecteurs. On définit le produit scalaire entre v et w par

$$\langle v, w \rangle := \sum_{i=1}^3 v_i w_i.$$

Proposition 3.8. Soient $v, w \in \mathbb{R}^3$. Le produit scalaire $\langle v, w \rangle$ satisfait les propriétés suivantes pour tout $\lambda \in \mathbb{R}, u, v, w \in \mathbb{R}^3$:

1. $\langle v, w \rangle = \langle w, v \rangle$;
2. $\langle \lambda v + u, w \rangle = \lambda \langle v, w \rangle + \langle u, w \rangle$;
3. $\langle v, v \rangle \geq 0$ avec égalité $\iff v = 0$.

Démonstration. Le point 1 est trivial, le second se démontre par :

$$\langle \lambda v + u, w \rangle \equiv \sum_{i=1}^3 (\lambda v_i + u_i) w_i = \lambda \sum_{i=1}^3 v_i w_i + \sum_{i=1}^3 u_i w_i = \lambda \langle v, w \rangle + \langle u, w \rangle,$$

et le dernier point découle du fait que $\langle v, v \rangle$ est une somme de carrés ne pouvant donc être nulle que quand tous les termes sont nuls. \square

Définition 3.9. Soient deux vecteurs $v, w \in \mathbb{R}^3$. On dit que v et w sont orthogonaux si $\langle v, w \rangle = 0$.

Remarque. Un plan peut être défini uniquement par un vecteur normal et un point appartenant au plan (au lieu de deux vecteurs directeurs et un point).

Proposition 3.10. Soit un plan $\pi : ax + by + cz = d$. Alors le vecteur $a = (a, b, c)$ est normal à π .

Démonstration. Soit π un plan défini par le point v et le vecteur normal a . Alors pour tout vecteur $w = (x, y, z) \in \pi$, le vecteur $v - w$ est un vecteur directeur du plan. Dès lors, $\langle v - w, a \rangle = 0$, ou encore $\langle v, a \rangle = \langle w, a \rangle$. Posons $d := \langle v, a \rangle$. On a donc $d = \langle v, a \rangle = \langle w, a \rangle = ax + by + cz$. Dès lors, si un plan est de coordonnée cartésienne $\pi : ax + by + cz = d$, alors le vecteur a est normal au plan π . \square

Définition 3.11. Soient deux droites D_1, D_2 de vecteur directeur respectif v_1, v_2 .

1. Si $D_1 \cap D_2 = \emptyset$ et $\exists k \in \mathbb{R}$ t.q. $v_1 = kv_2$, alors D_1 et D_2 sont parallèles.
2. Si $|D_1 \cap D_2| = 1$ et $\langle v_1, v_2 \rangle = 0$, alors D_1 et D_2 sont orthogonales.

Définition 3.12. Soient deux plans π_1, π_2 de vecteur orthogonal respectif a_1, a_2 .

1. Si $\pi_1 \cap \pi_2 = \emptyset$, alors π_1 et π_2 sont parallèles (et donc $\exists k \in \mathbb{R}$ t.q. $a_1 = ka_2$).
2. Si $\langle a_1, a_2 \rangle = 0$, alors π_1 et π_2 sont orthogonaux.

Définition 3.13. Soient D une droite de vecteur directeur v et π un plan de vecteur normal a .

1. Si $\langle a, v \rangle = 0$, alors le vecteur normal du plan et le vecteur directeur de la droite et le vecteur normal du plan sont orthogonaux, donc la droite et le plan sont parallèles.
2. Si $\exists k \in \mathbb{R}$ t.q. $a = kv$, alors le vecteur directeurs de la droite et le vecteur normal du plan sont colinéaires, donc la droite est orthogonale au plan.

3.4 Distances dans \mathbb{R}^3

Définition 3.14. Soit $v \in \mathbb{R}^3$ un vecteur. On définit sa norme (sa longueur) par :

$$\|v\| := \sqrt{\langle v, v \rangle}.$$

Remarque. Comme $\|v\|$ représente la longueur de v , c'est également la distance entre v et l'origine \mathcal{O} . Donc la distance entre deux points p et q est donnée par $\|p - q\|$.

Proposition 3.15 (Propriétés de la norme). Soit $v \in \mathbb{R}^3$ un vecteur. Alors :

1. $\|v\| \geq 0$ et $\|v\| = 0 \iff v = 0$;
2. $\|\lambda v\| = |\lambda| \|v\|$.

Démonstration. La proposition 1 est directe par la définition de la norme et la proposition 2 se montre par :

$$\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v \rangle} = \sqrt{\lambda^2 \langle v, v \rangle} = |\lambda| \sqrt{\langle v, v \rangle} = |\lambda| \|v\|.$$

\square

Théorème 3.16 (Inégalité de Cauchy-Schwartz). Soient deux vecteurs $v, w \in \mathbb{R}^3$. Alors :

$$|\langle v, w \rangle| \leq \|v\| \|w\|,$$

avec égalité si et seulement si v et w sont colinéaires.

Démonstration. Si un des deux vecteurs vaut 0, alors la preuve est évidente. S'ils sont tous les deux différents de 0, alors pour tout $\lambda \in \mathbb{R}$, on a :

$$0 \leq \|v - \lambda w\|^2 = \langle v - \lambda w, v - \lambda w \rangle = \langle v, v \rangle - 2\lambda \langle v, w \rangle + \lambda^2 \langle w, w \rangle = \|v\|^2 - 2\lambda \langle v, w \rangle + \lambda^2 \|w\|^2.$$

Cette équation de degré 2 en λ a pour discriminant $(2 \langle v, w \rangle)^2 - 4 \|w\|^2 \|v\|^2$. Et comme cette équation doit être positive, le discriminant doit être négatif (pas de racine) ou nul (une unique racine λ_0 telle que $\|v - \lambda_0 w\| = 0$). Il faut donc $4 \langle v, w \rangle^2 - 4 \|v\|^2 \|w\|^2 \leq 0$, ou encore $|\langle v, w \rangle| \leq \|v\| \|w\|$. \square

Proposition 3.17. Soient $v, w \in \mathbb{R}^3$. Alors $\|v + w\| \leq \|v\| + \|w\|$.

Démonstration.

$$\|v + w\|^2 = \langle v + w, v + w \rangle = \langle v, v \rangle + 2 \langle v, w \rangle + \langle w, w \rangle = \|v\|^2 + 2 \langle v, w \rangle + \|w\|^2 \leq \|v\|^2 + 2 \|v\| \|w\| + \|w\|^2 = (\|v\| + \|w\|)^2.$$

\square

3.5 Angles dans \mathbb{R}^3

Définition 3.18. Soient $v, w \in \mathbb{R}^3$. Alors On définit l'angle entre v et w par :

$$\alpha := \arccos \left(\frac{\langle v, w \rangle}{\|v\| \|w\|} \right).$$

Remarque. Cette définition est bonne car deux vecteurs orthogonaux ont un produit scalaire nul, dès lors $\alpha = \arccos(0) = \frac{\pi}{2}$. De même, deux vecteurs colinéaires ont un produit scalaire valant $\|v\| \|w\|$ (ou $-\|v\| \|w\|$ selon l'orientation des vecteurs). Dès lors, $\alpha = \arccos(\pm 1) = 0$ ou π .

Remarque. On remarque que cette définition implique que $\langle v, w \rangle = \|v\| \|w\| \cos \alpha$.

Lemme 3.19. Soient deux vecteurs $v, w \in \mathbb{R}^3$. Alors :

$$\beta := \arccos \left(\frac{|\langle v, w \rangle|}{\|v\| \|w\|} \right) \in \left[0, \frac{\pi}{2} \right]$$

Démonstration. Par la définition de la fonction arccos, pour tout $x \in \left[0, \frac{\pi}{2} \right]$, on a $\arccos(x) \geq 0$ et c'est une bijection. \square

Proposition 3.20 (Parallélisme et orthogonalité entre droites et plans).

1. Soient deux droites D_1, D_2 de vecteurs directeurs $v_1, v_2 \in \mathbb{R}^3$. Si $|D_1 \cap D_2| = 1$, alors l'angle entre ces deux droites est donné par :

$$\alpha := \arccos \left(\frac{|\langle v_1, v_2 \rangle|}{\|v_1\| \|v_2\|} \right) \in \left] 0, \frac{\pi}{2} \right[.$$

2. Soient deux plans π_1, π_2 de vecteurs normaux $a_1, a_2 \in \mathbb{R}^3$. Si $\pi_1 \cap \pi_2$ est une droite, alors l'angle entre π_1 et π_2 est donné par :

$$\alpha := \arccos \left(\frac{|\langle a_1, a_2 \rangle|}{\|a_1\| \|a_2\|} \right) \in \left] 0, \frac{\pi}{2} \right[.$$

3. Soient une droite D de vecteur directeur $\mathbf{v} \in \mathbb{R}^3$ et un plan π de vecteur normal $\mathbf{a} \in \mathbb{R}^3$. Si $|D \cap \pi| = 1$, alors l'angle entre π et D est donné par :

$$\alpha := \frac{\pi}{2} - \arccos \left(\frac{|\langle \mathbf{a}, \mathbf{v} \rangle|}{\|\mathbf{a}\| \|\mathbf{v}\|} \right) \in \left] 0, \frac{\pi}{2} \right].$$

3.6 Produit vectoriel

Définition 3.21. Soient $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ deux vecteurs. On définit le produit vectoriel entre \mathbf{v} et \mathbf{w} par :

$$\mathbf{v} \times \mathbf{w} := (v_2 w_3 - v_3 w_2, v_3 w_1 - v_1 w_3, v_1 w_2 - v_2 w_1).$$

Proposition 3.22 (Propriétés du produit vectoriel). Soient $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ trois vecteurs et $\lambda \in \mathbb{R}$ un scalaire.

1. $(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = (\mathbf{u} \times \mathbf{w}) + (\mathbf{v} \times \mathbf{w})$;
2. $(\lambda \mathbf{v}) \times \mathbf{w} = \lambda(\mathbf{v} \times \mathbf{w})$;
3. $\mathbf{v} \times \mathbf{w} = -(\mathbf{w} \times \mathbf{v})$;
4. $\mathbf{v} \times (\lambda \mathbf{v}) = 0$.

Démonstration.

1. Pour la proposition 1, notons que :

$$\begin{aligned} (\mathbf{u} + \mathbf{v}) \times \mathbf{w} &= ((u_2 + v_2)w_3 - w_2(u_3 + v_3), (u_3 + v_3)w_1 - w_3(u_1 + v_1), (u_1 + v_1)w_2 - w_1(u_2 + v_2)) \\ &= (u_2 w_3 - w_2 u_3 + v_2 w_3 - w_2 v_3, u_3 w_1 - w_3 u_1 + v_3 w_1 - w_3 v_1, u_1 w_2 - w_1 u_2 + v_1 w_2 - w_1 v_2) \\ &= (u_2 w_3 - w_2 u_3, u_3 w_1 - w_3 u_1, u_1 w_2 - w_1 u_2) + (v_2 w_3 - w_2 v_3, v_3 w_1 - w_3 v_1, v_1 w_2 - w_1 v_2) \\ &= (\mathbf{u} \times \mathbf{w}) + (\mathbf{v} \times \mathbf{w}). \end{aligned}$$

2. Pour la proposition 2, il suffit de mettre λ en évidence :

$$(\lambda \mathbf{v}) \times \mathbf{w} = (\lambda(v_2 w_3 - w_2 v_3), \lambda(v_3 w_1 - w_3 v_1), \lambda(v_1 w_2 - w_1 v_2)) = \lambda(\mathbf{v} \times \mathbf{w}).$$

3. Pour la proposition 3, il suffit de regarder les indices pour s'en convaincre.
4. Pour la proposition 4, on a :

$$\mathbf{v} \times (\lambda \mathbf{v}) = (v_2 \lambda v_3 - v_3 \lambda v_2, v_3 \lambda v_1 - v_1 \lambda v_3, v_1 \lambda v_2 - v_2 \lambda v_3) = (0, 0, 0) = 0.$$

□

Proposition 3.23. Soient $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ deux vecteurs. Alors leur produit vectoriel est orthogonal à \mathbf{v} et à \mathbf{w} .

Démonstration. Trivial par la définition (expansion de la formule du produit scalaire et arriver à 0). □

Lemme 3.24. Soient $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ deux vecteurs. Alors :

$$\|\mathbf{v}\|^2 \|\mathbf{w}\|^2 - (\langle \mathbf{v}, \mathbf{w} \rangle)^2 = \|\mathbf{v} \times \mathbf{w}\|^2.$$

Démonstration. Trivial par expansion de la formule. □

Corollaire 3.25 (du lemme précédent). Soient $v, w \in \mathbb{R}^3$. Alors $\|v \times w\| = \|v\|^2 \|w\|^2 \sin(\alpha)^2$.

Démonstration. Par le lemme 3.24, on a:

$$\|v \times w\|^2 = (\|v\| \|w\|)^2 - (\langle v, w \rangle)^2 = (\|v\| \|w\|)^2 \left(1 - \left(\frac{\langle v, w \rangle}{\|v\| \|w\|} \right)^2 \right) = (\|v\| \|w\|)^2 (1 - \cos(\alpha)^2) = \|v\|^2 \|w\|^2 \sin(\alpha)^2.$$

□

Proposition 3.26. Soient $v, w \in \mathbb{R}^3$ deux vecteurs. La projection $p \in \mathbb{R}^3$ de v sur w est donnée par :

$$p := \frac{\langle v, w \rangle}{\langle w, w \rangle} w.$$

Démonstration. On veut trouver λ tel que $p = \lambda w$. Géométriquement, il faut que $v - p$ et w soient orthogonaux :

$$0 = \langle v - p, w \rangle = \langle v - \lambda w, w \rangle = \langle v, w \rangle - \lambda \langle w, w \rangle.$$

On a donc $\lambda = \frac{\langle v, w \rangle}{\langle w, w \rangle}$.

□

4 Systèmes d'équations linéaires et algèbre matricielle

4.1 Définitions

Définition 4.1. Soient $s, t \in \mathbb{N}_0$. Un système linéaire de s équations à t inconnues x_j et à coefficients réels $a_{ij}, b_i \in \mathbb{R}$ est :

$$(S) : \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1t}x_t = b_1 \\ \dots \\ a_{s1}x_1 + a_{s2}x_2 + \dots + a_{st}x_t = b_t \end{cases}$$

Que l'on peut également noter :

$$(S) : \sum_{j=1}^t a_{ij}x_j = b_i. \quad i \in [s]$$

Définition 4.2. Soit (S) un système d'équations linéaire. Si $b_i = 0$, alors la i ème équation est dite homogène. Si toutes les équations sont homogènes ($b = 0$), alors le système (S) est dit homogène.

Définition 4.3. L'ensemble W des solutions de (S) est l'ensemble des t -uplets $(x_1, \dots, x_t) \in \mathbb{R}^t$ qui satisfont toutes les équations de (S) .

Définition 4.4. Deux systèmes linéaires sont dits équivalents s'ils ont le même ensemble de solutions

4.2 La méthode de Gauss

Proposition 4.5. Soit (S) un système linéaire d'équations. Il existe trois opérations fondamentales applicables sur (S) qui donnent un système équivalent :

1. échanger deux équations du système ;
2. remplacer une équation par une combinaison linéaire d'équations du système ;
3. multiplier une équation par une constante $\lambda \neq 0$.

Remarque. Les trois opérations fondamentales ci-dessus sont toutes inversibles et leur inverse restent fondamentales.

Définition 4.6. la matrice associée à un système linéaire d'équations (S) est la matrice de dimensions $s \times (t + 1)$ contenant tous les coefficients du système :

$$\begin{bmatrix} a_{11} & \dots & a_{1t} & -b_1 \\ a_{21} & \dots & a_{2t} & -b_2 \\ \dots & \dots & \dots & \dots \\ a_{s1} & \dots & a_{st} & -b_t \end{bmatrix}.$$

Définition 4.7. Soit \mathcal{M} la matrice associée à un système linéaire d'équations (S) . Si :

1. dans chaque ligne de \mathcal{M} , si le premier coefficient non-nul (s'il existe) est 1 (on l'appelle le pivot de la ligne) ;
2. dans la colonne de chaque pivot, tous les coefficients n'étant pas le pivot valent 0 ;

3. Le pivot d'une ligne est à gauche des pivots des lignes suivantes (s'ils existent), alors on dit que M est sous sa forme échelonnée réduite.

Algorithme 4.8 (Méthode de Gauss). L'algorithme de Gauss permet de transformer une matrice en sa forme échelonnée réduite en utilisant les opérations fondamentales de la proposition 4.5. Le principe de cet algorithme est pour chaque ligne ℓ , s'il existe un coefficient non-nul dans la ligne, notons r l'indice de ce coefficient. Ensuite, il faut diviser la ligne ℓ par ce même coefficient ($M_{\ell r}$) puis, pour toutes ligne i différentes de ℓ , soustraire à la ligne i la ligne ℓ multipliée par le coefficient M_{ir} .

Soit M une matrice de dimensions l fois r .

Pour tout $0 < i < l+1$

Si $M(i) \neq (0, 0, \dots, 0)$

p = indice du premier coefficient non nul de $M(i)$

$M(i) /= M(i)(p)$

Pour tout $0 < j < l+1, j \neq i$

$M(j) -= M(i) * M(j)(p)$

Définition 4.9. L'ensemble des matrices réelles de dimensions $s \times t$ est donné par $M_{s \times t}(\mathbb{R})$. On note $a := [a_{ij}] \in M_{s \times t}(\mathbb{R})$ une matrice quelconque de dimension $s \times t$. a_{ij} représente le coefficient de la i ème ligne et de la j ème colonne.

Définition 4.10. Soit $a \in M_{n \times n}(\mathbb{R})$ une matrice. a est une matrice triangulaire si pour tout $i > j : a_{ij} = 0$.

Théorème 4.11 (Unicité de la forme échelonnée réduite). Soient trois matrices $a, b, c \in M_{s \times t}(\mathbb{R})$. Si les matrices b et c sont toutes deux obtenues par une suite d'opérations fondamentales sur a et si b et c sont toutes deux sous leur forme échelonnée réduite, alors $b = c$.

Démonstration. Les matrices b et c sont obtenues par combinaisons linéaires des lignes de la matrice a . Dès lors, les lignes de b sont également des combinaisons linéaires des lignes de c . Dès lors, les indices $1 \leq i_1 < i_2 < \dots < i_k \leq t$ des pivots de la matrice b doivent être les mêmes que les indices des pivots de la matrice c . En effet : soient $1 \leq j_1 < j_2 < \dots < j_k \leq t$ les indices des pivots de la matrice c .

Prouvons-le par récurrence.

Cas de base : supposons $i_1 < j_1$. Alors, il faudrait que la i_1 ème colonne de c soit vide. Or b est une combinaison linéaire de c et la première ligne de b contient à la i_1 ème colonne le coefficient 1 ce qui est une contradiction avec le fait que b et c soient combinaisons linéaires l'un de l'autre. En supposant $i_1 > j_1$, on obtient la même contradiction. Il faut donc $i_1 = j_1$.

Pas de récurrence : soit $\gamma \in \{1, \dots, t\}$. Supposons $i_\gamma < j_\gamma$. Par le même argument, on montre que $i_\gamma = j_\gamma$. \square

4.3 Calcul matriciel

Remarque. Soit $x \in \mathbb{R}^t$ un vecteur. $x \in M_{t \times 1}(\mathbb{R})$ donc x est une matrice colonne.

Définition 4.12. Soient $v_1, \dots, v_k \in \mathbb{R}^t$ k vecteurs et $\lambda_1, \dots, \lambda_k$ k scalaires. Le vecteur

$$w = \sum_{i=1}^k \lambda_i v_i$$

est une combinaison linéaire des vecteurs v_i .

Définition 4.13. Soient $a \in M_{s \times t}(\mathbb{R})$ une matrice et $b \in \mathbb{R}^t$ un vecteur (une matrice colonne). On définit le produit entre a et b par :

$$ab := \left[\sum_{k=1}^s a_{ik} b_k \right] \in \mathbb{R}^s \simeq M_{s \times 1}(\mathbb{R}).$$

Définition 4.14. À l'aide de cette notation, on peut maintenant noter le système d'équations linéaire $(S) : ax = b$ où $a \in M_{s \times t}(\mathbb{R})$, $x \in \mathbb{R}^t$ et $b \in \mathbb{R}^s$.

Définition 4.15. Soit $a = [a_{ij}] \in M_{s \times t}(\mathbb{R})$ une matrice. On définit la transformation linéaire qui lui est associée par :

$$A : \mathbb{R}^t \rightarrow \mathbb{R}^s : v \mapsto av.$$

Lemme 4.16. Soit $(S) : ax = b$ un système linéaire d'équations et soit $A : \mathbb{R}^t \rightarrow \mathbb{R}^s$ la transformation linéaire associée à a . Alors l'ensemble W des solutions du système correspond à la préimage du vecteur $b \in \mathbb{R}^s$.

Démonstration. On peut constater :

$$A^{-1}(b) = \{x \in \mathbb{R}^t \text{ t.q. } A(x) = b\} = \{x \in \mathbb{R}^t \text{ t.q. } ax = b\} = W.$$

□

Proposition 4.17 (Propriétés de la transformation linéaire associée). Soient $a = [a_{ij}] \in M_{s \times t}(\mathbb{R})$ une matrice et $A : \mathbb{R}^t \rightarrow \mathbb{R}^s$ sa transformation linéaire associée (TLA). Soient $v = [v_i], w = [w_i] \in \mathbb{R}^t, \lambda \in \mathbb{R}$. On définit les propriétés de linéarité de A par :

1. $A(v + w) = A(v) + A(w)$;
2. $A(\lambda v) = \lambda A(v)$.

Démonstration. En repartant de la définition de la TLA, on voit que :

$$\begin{aligned} A(v + w) &= a(v + w) = \left[\sum_{k=1}^t a_{ik}(v_k + w_k) \right] = \left[\sum_{k=1}^t a_{ik}v_k + \sum_{k=1}^t a_{ik}w_k \right] = A(v) + A(w), \\ A(\lambda v) &= a(\lambda v) = \left[\sum_{k=1}^t a_{ik}\lambda v_k \right] = \lambda \left[\sum_{k=1}^t a_{ik}v_k \right] = \lambda A(v). \end{aligned}$$

□

Remarque. On remarque que $A(0) = A(v + (-1)v) = A(v) - A(v) = 0 \in \mathbb{R}^s$.

4.4 Structure des solutions d'un système linéaire d'équations

Définition 4.18. Soit $(S) : ax = b$ un système linéaire d'équations. On définit le système homogène associé par $(S') : ax = 0$.

Définition 4.19. Soit $(S) : ax = b$ et (S') son système homogène associé. Alors l'ensemble des solutions de (S') est appelé W_0 .

Proposition 4.20. Soient (S) et (S') un système linéaire et son système homogène associé. Soient $x, y \in W_0$ deux solutions de (S') . Alors toute combinaison linéaire de ces deux solutions est également une solution de (S') .

Démonstration. Effectivement, si x et y sont deux solutions de (S') , on a $A(\lambda x + \mu y) = \lambda A(x) + \mu A(y) = \lambda 0 + \mu 0 = 0$. \square

Lemme 4.21. *L'ensemble des solutions de (S') , le système homogène associé à $(S) : ax = b$ est :*

$$W_0 = \left\{ \sum_{i=1}^k \lambda_i v_i \text{ t.q. } \lambda_i \in \mathbb{R} \right\},$$

où les vecteurs v_i sont les vecteurs donnés par l'algorithme de Gauss.

Démonstration. On sait que les vecteurs v_i sont les solutions de (S') et donc sont dans W_0 . Or, par la proposition 4.20, on sait que toutes les combinaisons linéaires de ces vecteurs sont toujours des solutions. \square

Remarque. Géométriquement, selon le degré de liberté k du système, l'ensemble des solutions est un espace de dimension k passant par l'origine. En effet, si $k = 0$, alors l'ensemble des solutions W_0 est l'ensemble contenant uniquement l'origine, si $k = 1$, l'ensemble est la droite contenant tous les multiples λv de v , solution donnée par Gauss, si $k = 3$, W_0 est un plan ayant pour vecteurs directeurs v_1, v_2 donnés par Gauss et passant par l'origine. Etc.

Lemme 4.22. *Soit (S) un système linéaire d'équations et (S') son système homogène associé. Si $x, y \in \mathbb{R}^t$ sont des solutions de (S) , alors $x - y$ est une solution de (S') .*

Démonstration. Soit a la matrice associée à (S) et A la TLA à a . Si $x, y \in W$, alors $A(x) = A(y) = b$. Dès lors, $A(x - y) = A(x) - A(y) = b - b = 0$. \square

Corollaire 4.23. *Soit $\tilde{x} \in \mathbb{R}^t$ une solution du système linéaire (S) . Alors toute solution $y \in W$ de (S) est sous la forme suivante $\tilde{x} + y$ où $y \in W_0$. Ou encore, l'ensemble des solutions de (S) est l'ensemble :*

$$W = \tilde{x} + W_0 := \{\tilde{x} + y \text{ t.q. } y \in W_0\}.$$

Démonstration. Montrons d'abord que $\tilde{x} + W_0 \subseteq W$ et ensuite montrons que $W \subseteq \tilde{x} + W_0$. Soit $\tilde{x} + y \in \tilde{x} + W_0$. On sait que $A(\tilde{x} + y) = A(\tilde{x}) + A(y) = b + 0 = b$. Dès lors, $\tilde{x} + y$ est une solution de (S) et donc $\tilde{x} + y \in W$.

Soit $x \in W$. On sait que $A(x) = A(\tilde{x}) = b$. Par le lemme 4.22, on sait que le vecteur $x - \tilde{x} \in \mathbb{R}^t$ est une solution du système homogène associé, donc $x - \tilde{x} \in W_0$. Dès lors, on peut réécrire $x = \tilde{x} + (x - \tilde{x})$. On peut donc dire $x \in \tilde{x} + W_0$, ou encore $W \subseteq \tilde{x} + W_0$. Donc $W = \tilde{x} + W_0$. \square

Remarque. L'ensemble des solutions de (S) est donc l'ensemble

$$W := \left\{ \sum_{i=1}^k \lambda_i v_i \text{ t.q. } \lambda_i \in \mathbb{R} \right\} + \tilde{x}$$

où les vecteurs \tilde{x}, v_i sont donnés par l'algorithme de Gauss (le vecteur \tilde{x} est le vecteur correspondant à la $(t + 1)$ ème colonne de la matrice sous sa forme échelonnée réduite). Géométriquement, cela correspond à une translation de l'espace des solutions de (S') par le vecteur \tilde{x} . C'est donc un point, une droite, un plan, etc. comme pour (S') mais ne devant pas contenir l'origine.

5 Espaces vectoriels

5.1 Définitions

Définition 5.1. Soit \mathbb{K} un corps. Un espace vectoriel V sur \mathbb{K}^1 est un sous-ensemble de vecteurs muni d'une addition interne et d'un produit externe définis par :

1. $+: V \times V \rightarrow V : (v, w) \mapsto v + w$;
2. $\cdot: \mathbb{K} \times V \rightarrow V : (\lambda, v) \mapsto \lambda v$.

telles que pour tout $u, v, w \in V, \lambda, \mu \in \mathbb{K}$:

1. $u + (v + w) = (u + v) + w$;
2. $\exists 0 \in V$ t.q. $v + 0 = v$;
3. $\exists (-v) \in V$ t.q. $v + (-v) = 0$;
4. $v + w = w + v$;
5. $\lambda(\mu v) = (\lambda\mu)v$;
6. $\lambda(v + w) = \lambda v + \lambda w$;
7. $(\lambda + \mu)v = \lambda v + \mu v$;
8. $1v = v$.

Définition 5.2. Soit $W \subseteq V$ un sous-ensemble de V . W est un sous-espace vectoriel de V si W est également un espace vectoriel muni des mêmes opérations de somme et de produit.

Proposition 5.3. Soit V un espace vectoriel réel. $W \subseteq V$ est un sous-espace vectoriel de V si et seulement si pour tout $v, w \in W, \lambda \in \mathbb{R}$ on a :

1. $0 \in W$;
2. $v + w \in W$;
3. $\lambda v \in W$.

Démonstration. Montrons d'abord que ces trois hypothèses font de W un sous-espace vectoriel. Montrons ensuite que si W est un sous-espace, alors ces trois propositions sont vérifiées.

Soit $W \subseteq V$ tel que $\forall v, w \in W, \lambda \in \mathbb{R} : (v + w \in W) \wedge (\lambda v \in W) \wedge (0 \in W)$. Les propriétés 1, 4, 5, 6, 7, 8 viennent de la définition des opérations et des hypothèses 2 et 3. La propriété 2 vient de l'hypothèse 0 et la propriété 3 vient de l'hypothèse 3 pour $\lambda = -1$.

Soit $W \subseteq V$ un sous-espace de V . Les hypothèses sont vérifiées de manière triviale. □

Définition 5.4. Soient V, W deux espaces vectoriels. La fonction $f : V \rightarrow W$ est une application linéaire si pour tout $v, w \in V, \lambda, \mu \in \mathbb{R}$, on a : $f(\lambda v + \mu w) = \lambda f(v) + \mu f(w)$.

Proposition 5.5. Soit V un espace vectoriel réel. Soient $v, v_1, \dots, v_k \in V, \lambda, \lambda_1, \lambda_k \in \mathbb{R}$. Alors :

1. $\lambda 0 = 0$;
2. $0v = 0$;

¹Il sera question ici presque exclusivement d'espaces vectoriels sur \mathbb{R} .

$$3. \lambda v = 0 \Rightarrow (\lambda = 0) \vee (v = 0) ;$$

$$4. (-\lambda)v = \lambda(-v) = -\lambda v ;$$

$$5. v \sum_{i=1}^k \lambda_i = \sum_{i=1}^k (\lambda_i v) ;$$

$$6. \lambda \sum_{i=1}^k v_i = \sum_{i=1}^k (\lambda v_i)$$

Démonstration.

1. $0 + \lambda 0 = \lambda 0 = \lambda(0 + 0) = \lambda 0 + \lambda 0$. Dès lors, par l'existence de l'inverse additif, on a $0 = \lambda 0$;
2. $0 + 1v = 1v = v = v(1 + 0) = 1v + 0v$. Dès lors, par l'existence d'un inverse additif, on a $0 = 0v$;
3. Supposons $\lambda v = 0$ et $\lambda \neq 0$. Alors $0 = \lambda^{-1}0 = \lambda^{-1}(\lambda v) = 1v = v$;
4. $-\lambda v = (-1 \cdot \lambda)v = (-\lambda)v = (\lambda \cdot -1)v = \lambda(-v)$;
5. par les propriétés 6 et 7 d'un espace vectoriels, les propriétés 5 et 6 de la proposition sont vérifiées.

□

Définition 5.6. Soient V, W deux espaces vectoriels. Soit $A : V \rightarrow W$ une application linéaire. On définit :

1. $\text{Ker}(A) := \{v \in V \text{ t.q. } A(v) = 0 \in W\} \subseteq V = A^{-1}(0)$;
2. $\text{Im}(A) := \{A(v) \text{ t.q. } v \in V\} \subseteq W = A(V)$.

5.2 Sous-espaces

Théorème 5.7. Soient V, W deux espaces vectoriels et $A : V \rightarrow W$ une application linéaire. Alors :

1. si $Y \subseteq W$ est un sous-espace vectoriel de W , alors $A^{-1}(Y)$ est un sous-espace vectoriel de V ;
2. si $Z \subseteq V$ est un sous-espace vectoriel de V , alors $A(Z)$ est un sous-espace vectoriel de W .

Démonstration. Pour montrer qu'un sous-ensemble est un sous-espace vectoriel, on peut utiliser la proposition 5.3. Dès lors, montrons que $A(Z)$ est un sous-espace de W . Soient $y, z \in Z$. Notons $v = A(y), w = A(z)$. Alors, comme Z est un sous-espace, on sait que $y + \lambda z \in Z$. Donc $A(Z) \ni A(y + \lambda z) = A(y) + \lambda A(z) = v + \lambda w$.

Montrons maintenant que $A^{-1}(Y)$ est un sous-espace de V . Soient $x, y \in Y$. Comme Y est un sous-espace, $x + \lambda y \in Y$. Prenons v, w tels que $A(v) = x$ et $A(w) = y$. Donc $A^{-1}(Y) \ni A^{-1}(x + \lambda y) = A^{-1}(A(v) + \lambda A(w)) = A^{-1}(A(v + \lambda w)) = v + \lambda w$. □

Remarque. La preuve ci-dessus utilise le fait que 0_V est envoyé sur 0_W par une application linéaire, ce qui garantit la présence de 0_W dans $A(Z)$ et la présence de 0_V dans $A^{-1}(Y)$.

Corollaire 5.8. Soient V, W deux espaces vectoriels et $A : V \rightarrow W$ une application linéaire. Alors $\text{Ker}(A)$ est un sous-espace de W et $\text{Im}(A)$ est un sous-espace de V .

Démonstration. V est un sous-espace de lui-même et $\{0\}$ est un sous-espace de W . Donc $\text{Ker}(A) = A^{-1}(\{0\})$ est un sous-espace de V et $\text{Im}(A) = A(V)$ est un sous-espace de W par le théorème 5.7. □

Définition 5.9. Soient V un e.v. et W_1, W_2 deux sous-espaces de V . On définit la somme de ces sous-espaces par :

$$W = W_1 + W_2 := \{w_1 + w_2 \text{ t.q. } w_1 \in W_1, w_2 \in W_2\} \subseteq V.$$

Définition 5.10. Soit V un e.v. et W_1, W_2 deux sous-espaces. La somme $W_1 + W_2$ est dite directe si $W_1 \cap W_2 = \emptyset$.

Proposition 5.11. Soit V un e.v. réel et W_1, W_2 deux sous-espaces. La somme $W_1 + W_2$ est un sous-espace de V .

Démonstration. Les deux sous-espaces contiennent 0_V donc $0_V \in W_1 + W_2$. De plus, soient $x_1, x_2 \in W_1, \lambda \in \mathbb{R}, w_1, w_2 \in W_2$. On sait que $v_1 = x_1 + w_1, v_2 = x_2 + w_2 \in W_1 + W_2$. De plus, $v_1 + \lambda v_2 = (x_1 + \lambda x_2) \in W_1 + (w_1 + \lambda w_2) \in W_2$. On a donc bien $v_1 + \lambda v_2 \in W_1 + W_2$. \square

Définition 5.12. Soit V un e.v. Soit $\{W_i\}_{i \in I}$ une famille de sous-espaces. On définit l'intersection de ces sous-espaces par :

$$\bigcap_{i \in I} W_i := \{v \in V \text{ t.q. } \forall i \in I : v \in W_i\}.$$

Proposition 5.13. Soit V un e.v. et $\{W_i\}_{i \in I}$ une famille de sous-espaces. Alors l'espace intersection $\bigcap_{i \in I} W_i$ est un sous-espace de V .

Démonstration. Soient V un e.v. et $\{W_i\}_{i \in I}$ une famille de sous-espaces. Soient $w_1, w_2 \in \bigcap_{i \in I} W_i$ deux vecteurs dans l'intersection. On sait dès lors que $w_1, w_2 \in W_i$ pour tout $i \in I$. S'ils le sont, alors $w_1 + \lambda w_2 \in W_i$ pour tout $i \in I$ également. Dès lors, $w_1 + \lambda w_2 \in \bigcap_{i \in I} W_i$. \square

5.3 Parties libres et génératrices

Définition 5.14. Soient V un espace vectoriel et $X \subseteq V$ un sous-ensemble de V . On définit le sous-espace engendré par X par :

$$\langle X \rangle := \left\{ \sum_{i=1}^k \lambda_i v_i \text{ t.q. } k \in \mathbb{N}, \lambda_i \in \mathbb{R}, v_i \in X \right\}.$$

Si $X = \emptyset$, on pose $\langle X \rangle := \{0\}$.

Proposition 5.15. Soient V un e.v. et $X \subseteq V$ un sous-ensemble de V . Alors $\langle X \rangle$ est un sous-espace de V .

Démonstration. Si X est l'ensemble vide, alors $\langle X \rangle = \{0\}$ et est un sous-espace de V . Si $|X| = n > 0$, alors prenons deux vecteurs v_1, v_2 dans $\langle X \rangle$. On sait que v_1 et v_2 sont des combinaisons linéaires des vecteurs x_i de X . On sait :

$$v_1 = \sum_{i=1}^n \lambda_i x_i,$$

$$v_2 = \sum_{i=1}^n \mu_i x_i.$$

Dès lors, on trouve que :

$$v_1 + \beta v_2 = \sum_{i=1}^n \lambda_i x_i + \beta \sum_{i=1}^n \mu_i x_i = \sum_{i=1}^n (\lambda_i + \beta \mu_i) x_i \in \langle X \rangle.$$

Effectivement, $v_1 + \beta v_2$ est une combinaison linéaire des vecteurs de X , il fait donc partie de $\langle X \rangle$. \square

Lemme 5.16. Soit $X \subseteq V$ un sous-ensemble d'un espace vectoriel. Alors $X \subseteq \langle X \rangle$

Démonstration. En prenant les n vecteurs x_i de X , on sait que pour tout $x_i \in X$, on a :

$$x_i = \sum_{j=1}^n \delta_{ij} x_j \in \langle X \rangle.$$

□

Proposition 5.17. Soit un espace vectoriel V , un sous-ensemble $X \subseteq V$ et W l'intersection des sous-espaces de V contenant X . Alors :

$$W = \langle X \rangle.$$

Démonstration. Si $X = \emptyset$, alors tous les sous-espaces contiennent X . Donc l'intersection de tous ces sous-espaces est $\{0\}$. Si X est non-vide, montrons que $\langle X \rangle \subseteq W$ et puis montrons que $W \subseteq \langle X \rangle$. Soit $v = \sum_{i=1}^k \lambda_i v_i \in \langle X \rangle$. Puisque les $v_i \in X$, v est dans tout sous-espace contenant X et est donc dans l'intersection de ceux-ci. On a donc $\langle X \rangle \subseteq W$.

Par le lemme 5.16, on sait que $\langle X \rangle$ est un sous-espace de V contenant X . Dès lors, $W \subseteq \langle X \rangle$. On a alors montré l'égalité. □

Définition 5.18. Soient V un e.v. et $X \subseteq V$. On appelle X une partie génératrice si $\langle X \rangle = V$.

Proposition 5.19. Soient V un e.v. et $X \subseteq Y \subseteq V$ deux sous-ensembles de V . Alors $\langle X \rangle \subseteq \langle Y \rangle$.

Démonstration. Supposons $|X| = n$ et $|Y| = k$ avec $k \geq n$. Soit $v \in \langle X \rangle$. On sait que

$$v = \sum_{i=1}^n \alpha_i x_i,$$

où $x_i \in X$. Soient y_i pour $i \in \{1, \dots, k\}$ les vecteurs de Y . On sait par hypothèse que pour $i \in \{1, \dots, n\}$, on a $x_i = y_i$. Donc un vecteur de $\langle Y \rangle$ s'écrit sous la forme

$$w = \sum_{i=1}^k \lambda_i y_i = \sum_{i=1}^n \lambda_i x_i + \sum_{i=1}^{k-n} \mu_i y_{i+n}.$$

En prenant $\mu_i = 0$ pour tout $i \in \{1, \dots, k-n\}$, on sait écrire

$$v = \sum_{i=1}^n \alpha_i x_i + \sum_{i=1}^{k-n} \mu_i y_{i+n} \in \langle Y \rangle.$$

On a donc bien $\langle X \rangle \subseteq \langle Y \rangle$. □

Proposition 5.20. Soit V un e.v. réel et $X \subseteq V$. Alors $\langle \langle X \rangle \rangle = \langle X \rangle$.

Démonstration. On sait que

$$\begin{aligned} \langle \langle X \rangle \rangle &= \left\{ \sum_{i=1}^k \lambda_i v_i \text{ t.q. } k \in \mathbb{N}, \lambda_i \in \mathbb{R}, v_i \in \langle X \rangle \right\} = \left\{ \sum_{i=1}^k \left(\lambda_i \sum_{j=1}^{\ell} \mu_j w_j \right) \text{ t.q. } k, \ell \in \mathbb{N}, w_j \in X, \lambda_i, \mu_j \in \mathbb{R} \right\} \\ &= \left\{ \sum_{j=1}^{\ell} \left(\mu_j \sum_{i=1}^k \lambda_i \right) w_j \text{ t.q. } k, \ell \in \mathbb{N}, w_j \in X, \lambda_i, \mu_j \in \mathbb{R} \right\}. \end{aligned}$$

En posant :

$$\alpha_j := \mu_j \sum_{i=1}^k \lambda_i,$$

on peut réécrire :

$$\langle \langle X \rangle \rangle = \left\{ \sum_{j=1}^{\ell} \alpha_j w_j \text{ t.q. } \ell \in \mathbb{N}, w_j \in X, \alpha_j \in \mathbb{R} \right\} = \langle X \rangle.$$

□

Lemme 5.21. Soit V un e.v. et $X \subseteq Y \subseteq V$. Si X est une partie génératrice de V , alors Y est également une partie génératrice de V .

Démonstration. Par la proposition 5.19, on sait que $\langle X \rangle \subseteq \langle Y \rangle$. Or $\langle X \rangle = V$. Donc $V \subseteq \langle Y \rangle$. Donc Y est une partie génératrice de V . □

Proposition 5.22. Soit V un e.v. et X une partie génératrice de V . S'il existe $x \in X \setminus \{x\}$ tel que $x \in \langle X \setminus \{x\} \rangle$, alors $X \setminus \{x\}$ est une partie génératrice de V .

Démonstration. Supposons qu'il existe un tel x . On peut donc réécrire

$$x = \sum_{i=1}^k \lambda_i v_i$$

où $v_i \in X \setminus \{x\}$. Un vecteur quelconque $w \in V$ peut s'écrire comme une combinaison linéaire de vecteurs de X (car X est génératrice). On a donc :

$$v = \sum_{i=1}^k \mu_i v_i + \mu x$$

où $v_i \in X \setminus \{x\}$. Or, comme x est également une combinaison linéaire des vecteurs v_i , on peut réécrire :

$$v = \sum_{i=1}^k \mu_i v_i + \sum_{i=1}^k \lambda_i v_i = \sum_{i=1}^k (\lambda_i + \mu_i) v_i.$$

Et comme $v_i \in X \setminus \{x\}$, l'ensemble $X \setminus \{x\}$ est une partie génératrice de V . □

Définition 5.23. Soit V un e.v. et $X \subseteq V$ un sous-ensemble de V . X est appelé partie libre si pour tout $v \in \langle X \rangle$, on a $v = 0 \Rightarrow \lambda_i = 0 \forall i$.

Proposition 5.24. Soit X un sous-ensemble de l'e.v. V . X est une partie libre de V si et seulement si pour tout $x \in X : x \notin \langle X \setminus \{x\} \rangle$.

Démonstration. Montrons que si X est une partie libre, alors $\forall x \in X : x \notin \langle X \setminus \{x\} \rangle$. Soit X une partie libre de V . Supposons par l'absurde qu'il existe $x \in X$ tel que $x \in \langle X \setminus \{x\} \rangle$. Dès lors, il existe $v_1, \dots, v_k \in X \setminus \{x\}$ et $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ tels que $x = \lambda_1 v_1 + \dots + \lambda_k v_k$. La combinaison linéaire $x - \lambda_1 v_1 - \dots - \lambda_k v_k = 0$ est donc une combinaison nulle non-triviale, ce qui contredit le fait que X est libre.

Montrons maintenant que si $\forall x \in X : x \notin \langle X \setminus \{x\} \rangle$, alors X est libre. Supposons par l'absurde que X n'est pas libre. Alors il existe une combinaison linéaire $\lambda_1 v_1 + \dots + \lambda_k v_k = 0$ avec au moins un coefficient non-nul. Appelons-le λ_i . On a donc :

$$v_i = -\lambda_i^{-1} \left(\sum_{j=1, j \neq i}^k \lambda_j v_j \right) \in X \setminus \{v_i\},$$

ce qui contredit l'hypothèse. □

Proposition 5.25. Soit V un e.v. et $X \subseteq Y \subseteq V$. Si Y est une partie libre, alors X est également une partie libre.

Démonstration. Soit une combinaison linéaire des vecteurs de Y :

$$y = \sum_{i=1}^k \lambda_i v_i$$

où $v_i \in Y$. Comme Y est libre, la seule combinaison linéaire nulle possible est $\lambda_1 = \lambda_2 = \dots = 0$. Soit une combinaison linéaire des vecteurs de X :

$$x = \sum_{i=1}^n \mu_i v_i$$

où $n \leq k$ et $v_i \in X$. Si $x = 0$, on a déjà $\mu_t = 0$ pour $t \in \{n+1, \dots, k\}$. Par hypothèse, il faut donc que les autres λ_i valent 0. □

Proposition 5.26. Soit V un e.v. et X une partie libre de V . S'il existe $v \in V \setminus \langle X \rangle$, alors $X \cup \{v\}$ est également une partie libre.

Démonstration. Supposons que $X \cup \{v\}$ ne soit pas libre. Alors il existe $v_1, \dots, v_k \in X, \lambda, \lambda_1, \dots, \lambda_k \in \mathbb{R}$ tels que :

$$\lambda v + \sum_{i=1}^k \lambda_i v_i = 0.$$

Si $\lambda = 0$, il y a contradiction car on a supposé $X \cup \{v\}$ non libre. Si $\lambda \neq 0$, alors v est une combinaison linéaire des vecteurs v_i de X . Or par hypothèse, $v \in V \setminus \langle X \rangle$. Il y a donc à nouveau contradiction. Un tel λ n'existe pas, et donc $X \cup \{v\}$ est libre. □

5.4 Bases et dimension

Théorème 5.27. Soient V un e.v. et $X \subseteq V$. Alors les trois propositions suivantes sont équivalentes :

1. X est une partie génératrice minimale (il n'existe pas de $x \in X$ tel que $x \in \langle X \setminus \{x\} \rangle$) ;
2. X est une partie libre maximale (il n'existe pas de $v \in V$ tel que $X \cup \{v\}$ soit libre) ;
3. X est une partie libre et génératrice.

Démonstration. Montrons que (1) implique (3) : par la proposition 5.24, s'il n'existe pas de $x \in X$ tel que $x \in \langle X \setminus \{x\} \rangle$, alors X est libre.

Montrons que (3) implique (2) : X est libre et génératrice. Dès lors, si $v \in V$, alors $v \in \langle X \rangle$. Donc $X \cup \{v\}$ ne peut être libre.

Montrons que (2) implique (1) : X est libre maximale. Donc il n'existe pas de $v \in V$ tel que $v \notin \langle X \rangle$. Donc $\langle X \rangle = V$ ou encore X est génératrice. Mais X est génératrice minimale car s'il existait $x \in X$ tel que $x \in \langle X \setminus \{x\} \rangle$, alors cela voudrait dire que X n'est pas libre maximale. □

Définition 5.28. Soit un espace vectoriel V et $X \subseteq V$. Si X est une partie libre et génératrice, on appelle X une base de V .

Lemme 5.29 (lemme de Zorn). Soit (S, \leq) un ensemble partiellement ordonné. Si tout sous-ensemble totalement ordonné $X \subseteq S$ a une borne supérieure dans S (un $M \in S$ tel que $\forall x \in X : x \leq M$), alors il existe un élément maximal $m \in S$.

Théorème 5.30. Soient V un espace vectoriel et $X \subseteq Y \subseteq V$ tels que X est libre. Alors il existe une base B de $\langle Y \rangle$ telle que $X \subseteq B \subseteq \langle Y \rangle$.

Démonstration. Soit $\mathcal{S} := \{Z \subseteq V \text{ t.q. } X \subseteq Z \subseteq Y, Z \text{ partie libre de } V\}$. Associons une relation d'ordre partiel à \mathcal{S} par l'inclusion. On sait \mathcal{S} non vide car $X \in \mathcal{S}$.

Soient $\{Z_i\}_{i \in I}$ une famille totalement ordonnée de \mathcal{S} et $Z := \bigcup_{i \in I} Z_i$. On sait que $Z \in \mathcal{S}$ car $X \subseteq Z \subseteq Y$ (en effet : comme Z est défini par l'union, $X \subseteq Z$ et de plus, il n'existe aucun $z \in Z$ tel qu'il n'existe pas de i pour $z \in Z_i$ dès lors $Z \subseteq Y$) et Z est libre car défini par l'inclusion de parties libres **incluses**. De plus Z est une borne supérieure de $\{Z_i\}_{i \in I}$. Dès lors, par le lemme de Zorn, on sait qu'il existe un élément maximal $B \in \mathcal{S}$.

On remarque que B est une base de $\langle Y \rangle$ car sinon il existerait $v \in V \setminus \langle Y \rangle$ tel que $B \cup \{v\}$ serait dans \mathcal{S} , or B est l'élément maximal. Donc $\langle B \rangle = \langle Y \rangle$, ou encore B est une base de $\langle Y \rangle$ telle que $X \subseteq B \subseteq Y$. \square

Corollaire 5.31. Tout espace vectoriel V admet une base.

Démonstration. En posant $X = \emptyset$ et $Y = V$, par le théorème 5.30, il existe une base de $\langle V \rangle = V$. \square

Lemme 5.32. Soit V un e.v. et W un sous-espace de V . Soit $v \in V \setminus W$. Si $Z \subseteq V$ est un sous-espace de V tel que :

$$W \subseteq Z \subseteq \langle W \cup \{v\} \rangle,$$

alors, $W = Z$ ou $Z = \langle W \cup \{v\} \rangle$.

Démonstration. Soit B une base de W . Comme $v \notin W$, alors $B \cup \{v\}$ est une base de $\langle W \cup \{v\} \rangle$ car $B \cup \{v\}$ est une partie libre de $\langle W \cup \{v\} \rangle$ et est génératrice.

Si $Z = W$, alors la démonstration est faite. Supposons donc $Z \neq W$. Prenons $z \in Z \setminus W$. Puisque z n'est pas dans W , on sait que $B \cup \{z\}$ est une partie libre de $\langle W \cup \{v\} \rangle$. Alors il faut $v \in \langle B \cup \{z\} \rangle$ sinon $B \cup \{z, v\}$ serait une partie libre de $\langle W \cup \{v\} \rangle$ ce qui contredirait le fait que $B \cup \{v\}$ en soit une base. On a donc :

$$\langle W \cup \{v\} \rangle = \langle B \cup \{v\} \rangle \subseteq \langle B \cup \{z\} \rangle \subseteq Z.$$

Or par hypothèse, on avait $Z \subseteq \langle W \cup \{v\} \rangle$. Donc on sait $Z = \langle W \cup \{v\} \rangle$. \square

5.4.1 Cardinalité

Définition 5.33. Soient deux ensembles X et Y . On dit qu'ils sont de même cardinalité s'il existe une bijection $f : X \rightarrow Y$. Et on note $|X| = |Y|$. S'il existe une fonction injective $f : X \rightarrow Y$, on écrit $|X| \leq |Y|$.

Théorème 5.34 (Théorème de Cantor-Bernstein). Soient deux ensembles X et Y . Si $|X| \leq |Y|$ et $|Y| \leq |X|$, alors $|X| = |Y|$.

Démonstration. Soient $f : X \rightarrow Y, g : Y \rightarrow X$ deux fonctions injectives. Si g est bijective, la preuve est finie. Posons $A_0 := X \setminus g(Y)$. On sait A_0 non-vide car g n'est pas bijective. On pose ensuite :

$$\forall i \in \mathbb{N} \setminus \{0\} : A_i := (g \circ f)(A_{i-1}).$$

On définit $A := \bigcup_{i \in \mathbb{N}} A_i$ et $A' := X \setminus A$. L'ensemble A est non-vide.

On sait :

$$(g \circ f)(A) = \bigcup_{i \in I} (g \circ f)(A_i) = \bigcup_{i \in \mathbb{N}} A_{i+1} = X \setminus X_0 \subset A.$$

Définissons ensuite :

$$\varphi : X \rightarrow Y : x \mapsto \begin{cases} f(x) & \text{si } x \in A, \\ g^{-1}(x) & \text{si } x \in A' . \end{cases}$$

On sait que $g^{-1}(x)$ existe car $x \in A'$ et $A' \subseteq g(Y)$.

Montrons que φ est bijective :

1. montrons que $\varphi(A) \cap \varphi(A') = \emptyset$. Supposons qu'il existe $x \in \varphi(A) \cap \varphi(A')$. Puisque $x \in \varphi(A)$, on sait que $x \in f(A) \subseteq Y$, donc il existe $y \in A$ tel que $f(y) = x$ donc $g(x) = g(f(y)) = (g \circ f)(y)$. On a donc montré que l'image de $(g \circ f)$ est incluse dans A . On a donc $g(x) \in A$.

Mais puisque $x \in \varphi(A')$, on sait qu'il existe $y' \in A'$ tel que $g(x) = y'$. Donc $g(x) \in A'$. Ce qui est une contradiction avec le fait que $g(x) \in A$ car A et A' sont disjoints.

2. Montrons que φ est injective. Soient $x, x' \in X$ tels que $\varphi(x) = \varphi(x')$. Les combinaisons possibles sont :

- (a) $x \in A$ et $x \in A'$ ou inversement ce qui est impossible car $\varphi(x) \in \varphi(A)$ et $\varphi(x') \in \varphi(A')$ qui sont disjoints ;
- (b) $x, x' \in A$. Dès lors, $f(x) = \varphi(x) = \varphi(x') = f(x')$, et par injectivité de f , il faut $x = x'$;
- (c) $x, x' \in A'$. On sait que g est une bijection entre Y et $B := g(Y)$. Donc $g^{-1} : g(Y) \rightarrow Y$ est également bijective. Dès lors, si $g^{-1}(x) = g^{-1}(x')$, alors $x = x'$.

On a montré que dans tous les cas, si $\varphi(x) = \varphi(x')$, il faut $x = x'$. φ est donc injective.

3. Montrons que φ est surjective. Soit $y \in Y$. On sait que $g(y) \in A \iff g(y) \notin g(A')$.

- (a) Si $g(y) \in A$, alors on sait que $g(y) \notin A_0$ car $A_0 \equiv X \setminus g(Y)$. Donc $g(y) \in X \setminus A_0 = (g \circ f)(A) \subset X$. Donc il existe $x \in X$ tel que $g(y) = (g \circ f)(x)$ ou encore tel que $y = f(x) = \varphi(x)$ (par l'injectivité de g).
- (b) Si $g(y) \in A'$, alors il existe $x' \in A'$ tel que $g(y) = x'$ ou encore $y = g^{-1}(x') = \varphi(x')$.

On a montré que pour tout $y \in Y$, il existe un $x \in X$ tel que $\varphi(x) = y$. φ est donc surjective.

□

5.4.2 Dimension

Remarque. Les deux théorèmes suivants ne sont pas démontrés et sont considérés comme admis.

Théorème 5.35. Soit X un ensemble infini. Soit $\mathcal{F}_X := \{Y \subseteq X \text{ t.q. } Y \text{ est une partie finie}\}$. Alors $|\mathcal{F}_X| = |X|$.

Théorème 5.36. Soient I un ensemble d'indices infini et $\{X_i\}_{i \in I}$ une famille d'ensembles finis indicés par I . Alors :

$$\left| \bigcup_{i \in I} X_i \right| \leq |I|.$$

Théorème 5.37. Soient V un espace vectoriel et E, E' deux bases quelconques de V . Alors $|E| = |E'|$.

Démonstration. Supposons d'abord $|E| = n \in \mathbb{N}$ donc E fini. Posons $k := |E \cap E'|$. Montrons par récurrence sur $n - k$ que $|E| = |E'|$.

cas de base : si $n - k = 0$, alors $|E \cap E'| = |E|$ et il faut donc $E \subseteq E'$. Or E et E' sont tous deux des familles génératrices minimales donc il faut $E = E'$.

pas de récurrence : si $n - k = i > 0$, on sait que $E \cap E' \subseteq E$ et donc si $E = \{e_1, \dots, e_n\}$, on a $E \cap E' = \{e_1, \dots, e_k\}$. De plus, on sait que E est une partie génératrice minimale. Donc $E \setminus \{e_{k+1}\}$ n'est plus génératrice. Dès lors, on sait qu'il existe $y \in E' \setminus \langle E \setminus \{e_{k+1}\} \rangle$. Soit $E'' := \{y\} \cup E \setminus \{e_{k+1}\}$. On sait donc dire :

$$\langle E \setminus \{e_{k+1}\} \rangle \subset \langle E'' \rangle \subseteq \langle E \rangle = V.$$

Par le lemme 5.32, on sait $\langle E'' \rangle = V$. De plus, E'' est libre. Donc E'' est une base de V . On a donc $|E''| = |E| = n$. De plus, puisque $E' \cap E'' = \{e_1, \dots, e_{k+1}\}$, peut dire : $|E''| - |E' \cap E''| = n - (k + 1) = n - k - 1$. Alors Par hypothèse de récurrence, $|E'| = |E''|$.

Supposons ensuite que E est de cardinal infini (et donc E' également). Posons $\mathcal{F}_E := \{X \subseteq E \text{ t.q. } X \text{ partie finie}\}$ et $\mathcal{F}_{E'} := \{Y \subseteq E' \text{ t.q. } Y \text{ partie finie}\}$. Par le théorème 5.35 (admis), on sait que $|\mathcal{F}_E| = |E|$ et $|\mathcal{F}_{E'}| = |E'|$.

Soit $Y \in \mathcal{F}_{E'}$. Alors, par définition de Y , on sait que $E' \cap \langle Y \rangle$ est fini. Dès lors, on sait qu'il existe une base B de $\langle Y \rangle$ telle que $E' \cap \langle Y \rangle \subseteq B \subseteq \langle Y \rangle$. Or Y est également une base de $\langle Y \rangle$. Donc $|Y| = |B|$ et donc Y est fini.

Posons $f : \mathcal{F}_E \rightarrow \mathcal{F}_{E'} : Y \mapsto E' \cap \langle Y \rangle$. Soit $b \in E'$. On sait qu'il existe $Y \in \mathcal{F}_E$ tel que $b \in \langle Y \rangle$. Donc $b \in f(Y) = E' \cap \langle Y \rangle$. On peut déduire que E' est égal à l'union de tous les éléments de \mathcal{F}_E :

$$E' = \bigcup_{Y \in \mathcal{F}_E} Y.$$

Dès lors, par le théorème 5.36, on sait que $|E'| \leq |f(\mathcal{F}_E)|$. De plus, on sait que $f(\mathcal{F}_E) \subseteq \mathcal{F}_{E'}$. Donc $|E'| \leq |f(\mathcal{F}_E)| \leq |\mathcal{F}_{E'}| = |E'|$. Il faut donc nécessairement $|f(\mathcal{F}_E)| = |\mathcal{F}_{E'}| = |E'|$.

Et comme, de plus, on a $|f(\mathcal{F}_E)| \leq |\mathcal{F}_E|$. On sait donc dire $|E'| = |f(\mathcal{F}_E)| \leq |\mathcal{F}_E| = |E|$, d'où $|E'| \leq |E|$.

Par symétrie, on trouve la même chose pour E : à savoir $|E| \leq |E'|$. Donc par le théorème de Cantor-Bernstein (théorème 5.34), on sait que $|E| = |E'|$. \square

Définition 5.38. Soit V un e.v. On définit la dimension réelle de V par le cardinal de ses bases : $\dim_{\mathbb{R}} V = |E|$ avec E , base de V .

5.4.3 Dépendance base et système de coordonnées

Définition 5.39. Soit V un espace vectoriel réel. Tout vecteur $v \in V$ peut s'écrire comme une combinaison linéaire de $\lambda_i e_i$ où e_i sont les vecteurs de la base E de V . On appelle le n -uplet de λ_i suivant $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$ les coordonnées de v dans la base E .

Proposition 5.40. Soit V un espace vectoriel et E une base de V . Alors chaque vecteur $v \in V$ s'écrit de manière unique comme une combinaison linéaire des vecteurs de la base E .

Démonstration. Par hypothèse, E est génératrice. Soit $v \in V$. Alors :

$$v = \sum_{i=1}^n \lambda_i e_i = \sum_{i=1}^n \mu_i e_i.$$

Alors, on peut écrire $v - v = 0$, d'où :

$$0 = v - v = \sum_{i=1}^n \lambda_i e_i - \sum_{i=1}^n \mu_i e_i = \sum_{i=1}^n (\lambda_i - \mu_i) e_i.$$

Comme E est libre, la seule combinaison nulle possible est la combinaison linéaire nulle triviale avec $\lambda_i - \mu_i = 0$ pour tout i , ou encore $\lambda_i = \mu_i$. \square

Proposition 5.41. Soient V un e.v. et W un sous-espace de V . Soient E_V et E_W des bases respectivement de V et W . Alors on peut :

1. compléter E_W pour en faire une base de V ;
2. réduire E_V pour en faire une base de W .

Démonstration. Montrons d'abord comment étendre E_W . Si $V = W$, E_W est une base de V . Supposons $W \neq V$. Alors il existe $v_1 \in V \setminus W = V \setminus \langle E_W \rangle$. De plus, $E_W \cup \{v_1\}$ est toujours libre. Si $\langle E_W \cup \{v_1\} \rangle = V$, alors c'est bon, sinon, il doit exister $v_2 \in V \setminus \langle E_W \cup \{v_1\} \rangle$. Mais $E_W \cup \{v_1, v_2\}$ est toujours libre. Si $\langle E_W \cup \{v_1, v_2\} \rangle = V$, il existe v_3 etc. Une fois $k = \dim_{\mathbb{R}} V - \dim_{\mathbb{R}} W$ vecteurs ajoutés, on obtient $\langle E_W \cup \{v_1, \dots, v_k\} \rangle = V$ où $E_W \cup \{v_1, \dots, v_k\}$ est libre. Donc on a complété E_W pour en faire une base de V .

Montrons maintenant comment réduire E_V pour en faire une base de W . Si $V = W$, alors E_V est une base de W . Supposons $W \neq V$. Alors il existe $v_1 \in \langle E_V \rangle \setminus W$ tel que v_1 est une combinaison linéaire des des vecteurs de $E_V \setminus \{v_1\}$. $E_V \setminus \{v_1\}$ est toujours génératrice. Si $E_V \setminus \{v_1\}$ n'est pas libre, c'est qu'il existe un vecteur $v_2 \in \langle E_V \setminus \{v_1\} \rangle \setminus W$ qui est une combinaison linéaire des vecteurs de $E_V \setminus \{v_1\}$. $E_V \setminus \{v_1, v_2\}$ est toujours génératrice. Et si elle n'est pas libre, alors il existe v_3 etc. Une fois $k = \dim_{\mathbb{R}} V - \dim_{\mathbb{R}} W$ vecteurs retirés, on obtient $E_V \setminus \{v_1, \dots, v_k\}$ est libre dans W et est restée génératrice. Donc on a réduit E_V pour en faire une base de W . \square

Corollaire 5.42. Soit W un sous-espace de l'e.v. V . Alors $\dim_{\mathbb{R}} W \leq \dim_{\mathbb{R}} V$.

Démonstration. Par la proposition 5.41, on sait qu'il faut ajouter (ou retirer) un nombre positif de vecteurs pour passer d'une base du sous-espace (ou de l'espace) à une base de l'espace (ou du sous-espace). \square

Proposition 5.43. Soit V un e.v. réel de dimension finie d . Si X est une partie libre de V , alors $|X| \leq d$ et si $|X| = d$, alors X est une base. Si X est une partie génératrice, alors $|X| \geq d$ et si $|X| = d$, alors X est une base de V .

Démonstration. Soit E une base de V . Supposons $X = \{x_1, \dots, x_k\}$ une partie libre. Si $k > d$, alors x_i peut être écrit comme une combinaison linéaire des vecteurs de E pour $1 \leq i \leq d$. Alors x_{d+1} peut s'écrire comme une combinaison linéaire des x_i pour $1 \leq i \leq d$, ce qui contredit le fait que X est libre. Il faut donc $|X| \leq d$. Si $|X| = d$, X est une partie libre maximale, donc une base par le théorème 5.27.

Supposons $X = \{x_1, \dots, x_k\}$ une partie génératrice. Si $k < d$, alors Il existe $v \in V \setminus \langle X \rangle$ ce qui contredit le fait que X est génératrice. Si $|X| = d$, alors X est une partie génératrice minimale, donc une base par le théorème 5.27. \square

6 Relations entre espaces vectoriels

6.1 Isomorphismes

Définition 6.1. Soient V, W deux e.v. Soit $A : V \rightarrow W$ une application linéaire bijective. Alors A est un isomorphisme.

Définition 6.2. Deux espaces vectoriels V et W sont dits isomorphes s'il existe un isomorphisme $I : V \rightarrow W$.

Lemme 6.3. Soit $A : V \rightarrow W$ un isomorphisme entre deux espaces vectoriels. Alors l'inverse A^{-1} de A est également un isomorphisme.

Démonstration. Si A est bijective, alors A^{-1} est également bijective. Soient $\lambda \in \mathbb{R}, v_1, v_2 \in V$ et $w_1, w_2 \in W$ tels que $A(v_1) = w_1$ et $A(v_2) = w_2$. Alors, par la linéarité de A :

$$A^{-1}(w_1 + \lambda w_2) = A^{-1}(A(v_1) + \lambda A(v_2)) = A^{-1}(A(v_1 + \lambda v_2)) = v_1 + \lambda v_2 = A^{-1}(w_1) + \lambda A^{-1}(w_2).$$

On a bien $A^{-1} : W \rightarrow V$ linéaire et bijective. □

Proposition 6.4. Soit V un espace vectoriel réel de dimension d . Alors V est isomorphe à \mathbb{R}^d .

Démonstration. Soient E la base canonique de \mathbb{R}^d et F une base de V . Soit $\phi : V \rightarrow \mathbb{R}^d$ telle que si $v = \sum_{i=1}^d \lambda_i f_i$, alors on définit :

$$f(v) := (\lambda_i)_{i \in \{1, \dots, d\}}.$$

La fonction f est linéaire car si $v = \sum_{i=1}^n \lambda_i f_i$ et $w = \sum_{i=1}^n \mu_i f_i$, alors :

$$f(v + \alpha w) = f\left(\sum_{i=1}^n (\lambda_i + \alpha \mu_i) f_i\right) = (\lambda_i + \alpha \mu_i)_{i \in \{1, \dots, d\}}.$$

De plus, f est bijective car elle est injective :

$$(\lambda_1, \dots, \lambda_d) = (\mu_1, \dots, \mu_d) \Rightarrow \sum_{i=1}^d \lambda_i f_i = \sum_{i=1}^d \mu_i f_i,$$

et elle est surjective :

$$\forall (\lambda_i)_{i \in \{1, \dots, d\}} : \exists v \in V \text{ t.q. } f(v) = (\lambda_i)_{i \in \{1, \dots, d\}}.$$

Il suffit de prendre $v = \sum_{i=1}^n \lambda_i f_i$. □

Proposition 6.5. Tous les espaces vectoriels réels de dimension d sont isomorphes entre eux.

Démonstration. Soient V et W deux espaces vectoriels de dimension d . Par la proposition 6.4, on sait qu'il existe $f : V \rightarrow \mathbb{R}^d$ et $g : W \rightarrow \mathbb{R}^d$ deux isomorphismes. Puisque g^{-1} est également un isomorphisme (par le lemme 6.3), et puisque la composée de bijection reste une bijection ainsi que la composée de transformations linéaires reste une transformation linéaire, la fonction $(g^{-1} \circ f) : V \rightarrow W$ est un isomorphisme entre V et W . □

Proposition 6.6. Soient V, W deux espaces vectoriels réels de dimension finie. Alors :

$$\dim_{\mathbb{R}}(V + W) + \dim_{\mathbb{R}}(V \cap W) = \dim_{\mathbb{R}} V + \dim_{\mathbb{R}} W.$$

Démonstration. Soient $B_{V \cap W} = \{e_1, \dots, e_r\}$ une base de $V \cap W$, $B_V = \{e_1, \dots, e_r, x_1, \dots, x_n\}$ une base de V , $B_W = \{e_1, \dots, e_r, y_1, \dots, y_t\}$ une base de W . Soit $B := B_V \cup B_W = \{e_1, \dots, e_r, x_1, \dots, x_n, y_1, \dots, y_t\}$. B est une partie génératrice de $V + W$. Montrons que B est également une partie libre de $V + W$. Supposons :

$$\sum_{i=1}^r \alpha_i e_i + \sum_{i=1}^n \beta_i x_i + \sum_{i=1}^t \gamma_i y_i = 0.$$

Posons :

$$v := \sum_{i=1}^r \alpha_i e_i + \sum_{i=1}^n \beta_i x_i \in V = - \sum_{i=1}^t \gamma_i y_i \in W.$$

Si $v \in V$ et $v \in W$, alors $v \in V \cap W$. Donc il existe $(\lambda_1, \dots, \lambda_r) \in \mathbb{R}^r$ tels que :

$$v = \sum_{i=1}^r \lambda_i e_i.$$

On a donc :

$$v := \sum_{i=1}^r \alpha_i e_i + \sum_{i=1}^n \beta_i x_i = - \sum_{i=1}^t \gamma_i y_i = \sum_{i=1}^r \lambda_i e_i.$$

Donc :

$$\sum_{i=1}^r \lambda_i e_i + \sum_{i=1}^t \gamma_i y_i = 0,$$

ce qui implique $\lambda_i = \gamma_j = 0$ pour tout i, j puisque B_W est libre. Dès lors, $v = 0$, ou encore :

$$v = \sum_{i=1}^r \alpha_i e_i + \sum_{i=1}^n \beta_i x_i = 0.$$

Mais puisque B_V est libre, il faut $\alpha_i = \beta_j = 0$ pour tout i, j . La seule combinaison linéaire nulle possible est donc la combinaison linéaire nulle triviale. Alors B est libre.

Dès lors, on sait que $\dim_{\mathbb{R}}(V + W) + \dim_{\mathbb{R}}(V \cap W) = r + n + t + r = (r + n) + (r + t) = \dim_{\mathbb{R}} V + \dim_{\mathbb{R}} W$. \square

Proposition 6.7. Soient V, W deux espaces vectoriels réels. Soit $A : V \rightarrow W$ une application linéaire. Alors $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} \text{Ker}(A) + \dim_{\mathbb{R}} \text{Im}(A)$.

Démonstration. Soient $B_{\text{Ker}} = \{e_1, \dots, e_m\}$ une base du $\text{Ker}(A)$ et $B_{\text{Im}} = \{f_1, \dots, f_n\}$ une base de $\text{Im}(A)$. Soient $v_1, \dots, v_n \in V$ tels que $A(v_i) = f_i$ pour tout $1 \leq i \leq n$. Montrons que $B := B_{\text{Ker}} \cup \{v_1, \dots, v_n\}$ est une base de V . Montrons d'abord que B est génératrice.

Soit $v \in V$. Alors on a :

$$A(v) = \sum_{i=1}^n \lambda_i f_i = A \left(\sum_{i=1}^n \lambda_i v_i \right).$$

Dès lors, on sait construire $0 \in V$ par :

$$0 = A(v) - A \left(\sum_{i=1}^n \lambda_i v_i \right) = A \left(v - \sum_{i=1}^n \lambda_i v_i \right).$$

Il en découle directement que le vecteur $v - \sum_{i=1}^n \lambda_i v_i \in \text{Ker}(A)$. Or, tout vecteur de $\text{Ker}(A)$ peut s'exprimer comme :

$$\sum_{i=1}^m \mu_i e_i.$$

On a donc :

$$v - \sum_{i=1}^n \lambda_i v_i = \sum_{i=1}^m \mu_i e_i \iff v = \sum_{i=1}^n \lambda_i v_i + \sum_{i=1}^m \mu_i e_i \in \langle B \rangle.$$

Maintenant, montrons que B est libre. Soient $\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_n$ tels que :

$$\sum_{i=1}^m \alpha_i e_i + \sum_{i=1}^n \beta_i v_i = 0.$$

Alors, on peut écrire :

$$0 = A(0) = A\left(\sum_{i=1}^m \alpha_i e_i + \sum_{i=1}^n \beta_i v_i\right) = A\left(\sum_{i=1}^m \alpha_i e_i\right) + \sum_{i=1}^n \beta_i A(v_i) = A\left(\sum_{i=1}^m \alpha_i e_i\right) + \sum_{i=1}^n \beta_i f_i = 0 + \sum_{i=1}^n \beta_i f_i.$$

Or B_{Im} est libre, donc il faut $\beta_i = 0$ pour tout i . Il reste donc :

$$\sum_{i=1}^m \alpha_i e_i = 0,$$

et comme B_{Ker} est libre, il faut $\alpha_i = 0$ pour tout i . On a donc B libre. Puisque B est libre et génératrice de V , alors B est une base de V . Donc $\dim_{\mathbb{R}} V = |B| = n + m = \dim_{\mathbb{R}} \text{Im}(A) + \dim_{\mathbb{R}} \text{Ker}(A)$. \square

Proposition 6.8. Soient V, W deux espaces vectoriels et $A : V \rightarrow W$ une transformation linéaire. Alors :

1. A est injective si et seulement si $\dim_{\mathbb{R}} \text{Ker}(A) = 0$;
2. A est surjective si et seulement si $\dim_{\mathbb{R}} \text{Im}(A) = \dim_{\mathbb{R}} W$;
3. A est un isomorphisme si et seulement si $\dim_{\mathbb{R}} \text{Ker}(A) = 0$ et $\dim_{\mathbb{R}} \text{Im}(A) = \dim_{\mathbb{R}} W$;
4. A est un isomorphisme si et seulement si $\dim_{\mathbb{R}} \text{Ker}(A) = 0$ et $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} W$;
5. A est un isomorphisme si et seulement si $\dim_{\mathbb{R}} \text{Im}(A) = \dim_{\mathbb{R}} W$ et $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} W$.

Démonstration.

1. Montrons que si $\dim_{\mathbb{R}} \text{Ker}(A) = 0$, alors A est injective. Si $\dim_{\mathbb{R}} \text{Ker}(A) = 0$, alors $\text{Ker}(A) = \{0_V\}$. Donc soient $v_1, v_2 \in V$. Supposons que $f(v_1) = f(v_2)$. Dès lors, $0 = f(v_1) - f(v_2) = f(v_1 - v_2) = f(0_V)$. Comme 0_V est le seul élément du noyau, il faut $v_1 - v_2 = 0$, ou encore $v_1 = v_2$.

Montrons maintenant que si A est injective, alors $\dim_{\mathbb{R}} \text{Ker}(A) = 0$. Soient $v_1, v_2 \in V$. On sait que $f(v_1) = f(v_2) \Rightarrow v_1 = v_2$. On sait également que $f(0_V) = 0_W$. Donc pour tout $v \in V$, $f(v) = f(0) = 0$, alors $v = 0$. Autrement dit, seul 0_V est envoyé sur 0_W . On a bien $\dim_{\mathbb{R}} \text{Ker}(A) = 0$.

2. Montrons que si $\dim_{\mathbb{R}} \text{Im}(A) = \dim_{\mathbb{R}} W$, alors A est surjective. Soit $E = \{e_1, \dots, e_m\}$ une base de W . Soit $F = \{f_1, \dots, f_n\}$ une base de $\text{Im}(A)$. Soient v_1, \dots, v_m tels que $A(v_i) = f_i$. Par la proposition 6.5, on sait que $\text{Im}(A)$ et W sont isomorphes. Donc soit $w \in W$, on peut écrire w comme :

$$w = \sum_{i=1}^m \lambda_i e_i = \sum_{i=1}^m \mu_i f_i = \sum_{i=1}^m \mu_i A(v_i) = A\left(\sum_{i=1}^m \mu_i v_i\right).$$

Il existe donc un vecteur $v \in V$ tel que $A(v) = w$.

Montrons maintenant que si A est surjective, alors $\dim_{\mathbb{R}} \text{Im}(A) = \dim_{\mathbb{R}} W$. Soit $w \in W$. On sait qu'il existe $v \in V$ tel que $A(v) = w$. Prenons $E = \{e_1, \dots, e_m\}$ une base de W . Prenons $v_1, \dots, v_m \in V$ tels

que $A(v_i) = e_i$. On a $\{A(v_1), \dots, A(v_m)\}$ est une base de $\text{Im}(A)$ car pour tout vecteur $w \in W$, il existe $v \in V$ tel que :

$$w = \sum_{i=1}^m \lambda_i e_i = \sum_{i=1}^m \lambda_i A(v_i) = A\left(\sum_{i=1}^m \lambda_i v_i\right) = A(v).$$

3. Par les propositions 1 et 2.

4. Montrons que si A est un isomorphisme et $\dim_{\mathbb{R}} \text{Ker}(A) = 0$ (par l'injectivité de A), alors $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} W$. Par la proposition 6.7, on sait $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} \text{Ker}(A) + \dim_{\mathbb{R}} \text{Im}(A)$. Donc si $\dim_{\mathbb{R}} \text{Ker}(A) = 0$, on a $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} W$.

Montrons maintenant que si $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} W$ et $\dim_{\mathbb{R}} \text{Ker}(A) = 0$, alors A est un isomorphisme. Soient $E_V = \{e_1, \dots, e_n\}$ une base de V et $E_W = \{f_1, \dots, f_n\}$ une base de W . En définissant A par $A(e_i) = f_i$, on a bien un isomorphisme tel que $\dim_{\mathbb{R}} \text{Ker}(A) = 0$ car seul 0_V est envoyé sur 0_W .

5. Montrons d'abord que si A est un isomorphisme, alors $\dim_{\mathbb{R}} \text{Im}(A) = \dim_{\mathbb{R}} W$ et $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} W$. À nouveau, en partant de la proposition 6.7, on sait que $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} \text{Im}(A) + \dim_{\mathbb{R}} \text{Ker}(A)$. Or on sait A surjectif. Donc on sait que $\dim_{\mathbb{R}} W = \dim_{\mathbb{R}} \text{Im}(A)$. Et puisque A est injectif, on a $\dim_{\mathbb{R}} \text{Ker}(A) = 0$ donc $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} \text{Im}(A) = \dim_{\mathbb{R}} W$.

Si A n'est pas un isomorphisme, alors $\dim_{\mathbb{R}} \text{Im}(A) \neq \dim_{\mathbb{R}} W$.

□

6.2 Somme directe

Définition 6.9. Soient V et W deux espaces vectoriels. On définit la somme directe de V et W par $V \oplus W := V \times W$ muni de l'addition interne et de la multiplication externe triviale membre à membre.

Proposition 6.10. la somme directe de n espaces vectoriels est un espace vectoriel.

Démonstration. Les 8 axiomes se prouvent grâce à leur homologue des espaces vectoriels pères. □

Proposition 6.11. Soit V un espace vectoriel réel et soient W_1, W_2 deux sous-espaces de V . Si $W_1 \cap W_2 = \{0\}$, alors $W_1 + W_2$ est isomorphe à $W_1 \oplus W_2$.

Démonstration. Soit $A : W_1 \oplus W_2 \rightarrow W_1 + W_2 : (v, w) \mapsto v + w$. A est linéaire car pour tout $v_1, w_1 \in W_1, w_2, v_2 \in W_2, \lambda \in \mathbb{R}$, on a : $A((w_1, w_2) + \lambda(v_1, v_2)) = A(w_1 + \lambda v_1, w_2 + \lambda v_2) = (w_1 + \lambda v_1) + (w_2 + \lambda v_2) = (w_1 + w_2) + \lambda(v_1 + v_2) = A(w_1, w_2) + \lambda A(v_1, v_2)$.

De plus, $\text{Ker}(A) = \{(w_1, w_2) \in W_1 \oplus W_2 \text{ t.q. } w_1 + w_2 = 0\}$. Or, par hypothèse, $W_1 \cap W_2 = \{0\}$. Donc la seule possibilité pour avoir $w_1 + w_2 = 0$ est d'avoir $w_1 = w_2 = 0$. Donc $\dim_{\mathbb{R}} \text{Ker}(A) = 0$. Soit $x \in W_1 + W_2$. On sait qu'il existe $w_1 \in W_1, w_2 \in W_2$ tels que $x = w_1 + w_2 = A(w_1, w_2)$. Donc A est surjective. Dès lors, par la proposition 6.8, A est un isomorphisme. □

Proposition 6.12. Soit V un e.v. et $E = \{e_1, \dots, e_d\}$ une base de V . Soit W un autre e.v. Soit $F = \{f_1, \dots, f_d\} \subset W$. Alors il existe une unique transformation linéaire $A : V \rightarrow W$ telle que $A(e_i) = f_i$ pour tout $1 \leq i \leq d$.

Démonstration. Soit $g : E \rightarrow F : e_i \mapsto f_i$. On définit $A_g : V \rightarrow W : v = \sum_{i=1}^d \lambda_i e_i \mapsto w = \sum_{i=1}^d \lambda_i g(e_i)$.

Puisque g est bijective, A_g est également bijective. De plus, A_g est linéaire :

$$A_g \left(\sum_{i=1}^d \lambda_i e_i + \alpha \sum_{i=1}^d \mu_i e_i \right) = \sum_{i=1}^d (\lambda_i + \alpha \mu_i) g(e_i) = A_g \left(\sum_{i=1}^d \lambda_i e_i \right) + \alpha A_g \left(\sum_{i=1}^d \mu_i e_i \right).$$

□

Définition 6.13. Soient V, W deux e.v. et $A : V \rightarrow W$ une application linéaire. Alors on appelle $\dim_{\mathbb{R}} \text{Im}(A)$ le rang de A et on le note $\text{rang } A$.

6.3 Groupe linéaire

Définition 6.14. Soit V un espace vectoriel réel. Le groupe linéaire (général) est l'ensemble des isomorphismes $\{A : V \rightarrow V \text{ t.q. } A \text{ est un automorphisme}\}$. On le note $\text{GL}(V)$. Si $V = \mathbb{R}^d$, on note $\text{GL}_d(\mathbb{R})$ ou $\text{GL}(d, \mathbb{R})$.

Remarque. La composée de deux automorphismes est toujours un automorphisme car la composée de bijections reste bijective et si $f : V \rightarrow V$ et $g : V \rightarrow V$, alors $(f \circ g) : V \rightarrow V$ également.

Proposition 6.15. Soient deux espaces vectoriels V et W . Soit $E = \{e_1, \dots, e_d\}$. L'application linéaire $A : V \rightarrow W$ est un isomorphisme si et seulement si $A(E) = \{A(e_1), A(e_2), \dots, A(e_d)\}$ est une base de W .

Démonstration. Montrons que si A est un isomorphisme, alors $A(E)$ est une base de W . Les vecteurs de $A(E)$ sont des combinaisons linéaires des vecteurs de E . Donc $A(E)$ est libre. De plus, soit $w \in W$, il existe $v \in V$ tel que $w = A(v) = \sum_{i=1}^d \lambda_i A(e_i)$. Donc $A(E)$ est une partie génératrice de W .

Montrons maintenant que si $A(E)$ est une base de W , alors A est un isomorphisme. Puisque $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} W$ et $\text{Im}(A) = W$, par la proposition 6.8, A est un isomorphisme. □

7 Algèbre matricielle

7.1 Matrice associée à une transformation linéaire

Définition 7.1. Soient deux espaces vectoriels V et W . On définit $\text{Hom}(V, W) := \{A : V \rightarrow W \text{ t.q. } A \text{ transformation linéaire}\}$.

Proposition 7.2. Soient V, W deux e.v. En munissant $\text{Hom}(V, W)$ de l'addition interne

$$(A + B)(v) := A(v) + B(v),$$

et de la multiplication externe

$$(\lambda A)(v) := \lambda(A(v)),$$

$\text{Hom}(V, W)$ est une espace vectoriel.

Démonstration. Les 8 propriétés sont montrées par les propriétés identiques de V et W . □

Définition 7.3. Soit $W = \text{Hom}(V, V)$ un espace vectoriel. Les éléments de V sont appelés opérateurs linéaires (ou endomorphismes) de V .

Définition 7.4. Soient V, W deux espaces vectoriels de dimension finie, $A \in \text{Hom}(V, W)$. Si $E = \{e_1, \dots, e_t\}$, $F = \{f_1, \dots, f_s\}$ sont une base de V et W respectivement, alors on définit $m_{F,E}(A) := [A_{ij}] \in M_{s \times t}(\mathbb{R})$ telle que :

$$A(e_j) = \sum_{i=1}^s A_{ij} f_i.$$

7.2 Opérations sur les matrices

Proposition 7.5. Soient V, W deux espaces vectoriels avec une base respective E, F , soient $A, B \in \text{Hom}(V, W)$, et $m_{F,E}(A), m_{F,E}(B)$ les matrices associées à A et B . Alors :

$$\begin{aligned} m_{F,E}(\lambda A) &= \lambda m_{F,E}(A) \\ m_{F,E}(A + B) &= m_{F,E}(A) + m_{F,E}(B). \end{aligned}$$

Démonstration. On définit $m_{F,E}(A + \lambda B) = [(A + \lambda B)_{ij}]$ comme étant la matrice associée la transformation linéaire $A + \lambda B$. Donc elle est telle que :

$$(A + \lambda B)(e_j) = \sum_{i=1}^s (A + \lambda B)_{ij} f_i.$$

Cependant, on sait que $(A + \lambda B)(v) = A(v) + \lambda B(v)$. Donc on sait :

$$\sum_{i=1}^s (A + \lambda B)_{ij} f_i = (A + \lambda B)(e_j) = A(e_j) + \lambda B(e_j) = \sum_{i=1}^s A_{ij} f_i + \lambda \sum_{i=1}^s B_{ij} f_i = \sum_{i=1}^s (A_{ij} + \lambda B_{ij}) f_i.$$

On a donc pour tout i, j : $(A + \lambda B)_{ij} = A_{ij} + \lambda B_{ij}$. □

Remarque. Une fois les bases E et F fixées, $m_{F,E}$ est une fonction telle que $m_{F,E} : \text{Hom}(V, W) \rightarrow M_{\dim_{\mathbb{R}} W \times \dim_{\mathbb{R}} V}(\mathbb{R}) : A \mapsto m_{F,E}(A)$. Et la proposition 7.5 montre que cette fonction est linéaire.

Définition 7.6. Soient trois espaces vectoriels U, V, W de base respective $D = \{d_1, \dots, d_s\}, E = \{e_1, \dots, e_t\}, F = \{f_1, \dots, f_u\}$. Soient $A \in \text{Hom}(U, V)$ et $B \in \text{Hom}(V, W)$. La composée $(B \circ A) \in \text{Hom}(U, W)$. Soient les matrices $m_{E,D}(A) = [A_{ij}] \in M_{t \times s}(\mathbb{R})$ et $m_{F,E}(B) = [B_{ij}] \in M_{u \times t}(\mathbb{R})$. On définit le produit de ces deux matrices par $m_{F,E}(B \circ A) = [(BA)_{ij}] \in M_{u \times s}(\mathbb{R})$. Il faut que :

$$(B \circ A)(d_j) = \sum_{i=1}^u (BA)_{ij} f_i.$$

Or, par la définition de la composée, on sait :

$$(B \circ A)(d_j) = B(A(d_j)) = B\left(\sum_{i=1}^t A_{ij} e_i\right) = \sum_{i=1}^t A_{ij} B(e_i) = \sum_{i=1}^t A_{ij} \sum_{k=1}^u B_{ki} f_k = \sum_{k=1}^u \left(\sum_{i=1}^t B_{ki} A_{ij}\right) f_k.$$

On peut donc conclure pour tout i, j :

$$(BA)_{ij} = \sum_{k=1}^t B_{ki} A_{kj}.$$

Remarque. Selon cette définition, pour $A = [A_{ij}] \in M_{s \times t}(\mathbb{R}), B = [B_{ij}] \in M_{t \times u}(\mathbb{R})$, on a $(AB) = [(AB)_{ij}] = \left[\sum_{k=1}^t A_{ik} B_{kj}\right] \in M_{s \times u}(\mathbb{R})$.

Donc pour utiliser les notations des matrices d'applications, on peut écrire $m_{F,D}(B \circ A) = m_{F,E}(B) m_{E,D}(A)$.

Proposition 7.7. Si V et W sont deux espaces vectoriels. Soit $A : V \rightarrow W$ une transformation linéaire. Alors si $x \in V : A(x) = m_{F,E}(A)x$.

Démonstration. Soient $x \in V$ et $E = \{e_1, \dots, e_t\}$ une base de V . Soit $F = \{f_1, \dots, f_u\}$ une base de W . Dès lors $x = \sum_{i=1}^t x_i e_i$. On peut donc écrire :

$$A(x) = \sum_{i=1}^t x_i A(e_i) = \sum_{i=1}^t x_i \sum_{k=1}^u A_{ki} f_k = \sum_{k=1}^u \left(\sum_{i=1}^t A_{ki} x_i\right) f_k = m_{F,E}(A)x.$$

En posant $x \in M_{t \times 1}(\mathbb{R})$ le vecteur colonne contenant les coordonnées de x dans la base E . □

Proposition 7.8. Soient V, W deux espaces vectoriels de dimension respective t et s avec une base E, F . La transformation $m_{F,E} : \text{Hom}(V, W) \rightarrow M_{s \times t}(\mathbb{R}) : A \mapsto m_{F,E}(A)$ est un isomorphisme.

Démonstration. La proposition 7.5 montre la linéarité de cette transformation. Montrons maintenant qu'elle est bijective.

$m_{F,E}$ est injective car les matrices déterminent totalement les applications. Dès lors, deux matrices identiques donnent deux applications identiques également. Et $m_{F,E}$ est surjective car pour toute matrice $[A_{ij}] \in M_{s \times t}(\mathbb{R})$, on définit $A \in \text{Hom}(V, W)$ telle que $A(x) := \sum_{i=1}^s \sum_{j=1}^t A_{ij} x_j$. □

Définition 7.9. Soit V un espace vectoriel de dimension finie n et une base E de V . Soit $A \in \text{Hom}(V, V)$. Alors il existe $A^{-1} \in \text{Hom}(V, V)$ l'inverse de A . On définit la matrice inverse de $m_{E,E}(A)$ par $m_{E,E}(A)^{-1} := m_{E,E}(A^{-1})$.

7.3 Méthode de Gauss pour inverser une matrice

Définition 7.10. On définit la matrice $\Lambda_{k\ell} = [\lambda_{ij}]$ telle que $\lambda_{ij} = 1$ si $(i, j) = (k, \ell)$ et $\lambda_{ij} = 0$ sinon.

Remarque. En partant de l'algorithme de Gauss (algorithme 4.8), et en se concentrant sur les définitions des opérations matricielles, on remarque que les opérations fondamentales de l'algorithme correspondent à des opérations algébriques sur les matrices.

Changer la i ème ligne de la matrice $A = [A_{ij}]$ par $L_i + \mu L_j$ revient à faire :

$$(I + \mu \Lambda_{ij})A.$$

De plus, multiplier à gauche par $(I + \mu \Lambda_{ii})$ correspond à multiplier la i ème ligne par $\mu + 1$. Donc multiplier une ligne par un facteur α revient à faire :

$$(I + (\alpha - 1)\Lambda_{ii})A.$$

Finalement, remplacer la i ème ligne par la j ème revient à faire :

$$(I - (\Lambda_{ii} + \Lambda_{jj}) + \Lambda_{ij} + \Lambda_{ji})A.$$

Lemme 7.11. Soit une matrice $A = [A_{ij}] \in M_{n \times n}(\mathbb{R})$. En appliquant l'algorithme de Gauss, on obtient une suite de matrices b_k, \dots, b_1 telles que :

$$I = b_k \dots b_1 A.$$

Alors $b_k \dots b_1$ est l'inverse à gauche A^{-1} de A .

Démonstration. $A^{-1} := b_k \dots b_1$ est effectivement l'inverse à gauche car $A^{-1}A = b_k \dots b_1 A$ qui, par hypothèse, vaut I . \square

Proposition 7.12. Soit $a \in M_{n \times n}(\mathbb{R})$ une matrice inversible à gauche. Si a^{-1} est l'inverse à gauche de a , alors a^{-1} est également l'inverse à droite (et donc l'inverse) de a .

Démonstration. Soit E une base de V . Soient $a \in M_{n \times n}(\mathbb{R})$ une matrice et a^{-1} son inverse à gauche. a représente une application linéaire γ . γ est injective puisque son inverse sur $\gamma(\mathbb{R}^n)$ existe. Donc $\dim_{\mathbb{R}} \text{Ker}(\gamma) = 0$. De plus, $\gamma : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Son domaine et son codomaine sont donc de même dimension. Alors par la proposition 6.8, γ est un isomorphisme. Pareillement pour a^{-1} qui est la matrice d'une application linéaire γ^{-1} qui est bijective pour les mêmes raisons. Dès lors on peut dire $(\gamma \circ \gamma^{-1}) = \text{Id}_{\mathbb{R}^n} = (\gamma^{-1} \circ \gamma)$. On peut donc écrire :

$$aa^{-1} = m_{E,E}(\gamma)m_{E,E}(\gamma^{-1}) = m_{E,E}(\gamma \circ \gamma^{-1}) = m_{E,E}(\text{Id}_{\mathbb{R}^n}) = I_n.$$

On vient donc de montrer que si a^{-1} est l'inverse à gauche de a , alors a^{-1} est l'inverse à droite de a . \square

Remarque. L'existence d'un inverse à gauche (respectivement droite) n'implique pas toujours l'existence d'un inverse à droite (respectivement à gauche). Le cas des matrices ou celui des réels est un exemple mais pas une généralité : soit $f : X \rightarrow Y$ une fonction injective mais pas surjective. On peut trouver $g : f(X) \rightarrow X$ telle que $(g \circ f)$ soit l'identité sur X mais il n'existe pas d'inverse à droite pour f car elle n'est pas surjective.

7.4 Changement de base

Définition 7.13. Soit V un e.v. et $E = \{e_1, \dots, e_n\}, F = \{f_1, \dots, f_n\}$ deux bases de V et soit $A \in \text{Hom}(V, V)$ un automorphisme de V . On définit la matrice de changement de base de E vers F exprimée avec les coordonnées de E par $b := [B_{ij}] \in M_{n \times n}(\mathbb{R})$, matrice associée à l'application linéaire $B \in \text{Hom}(V, V)$ où $B(e_j) = \sum_{i=1}^n B_{ij} f_i$.

Lemme 7.14. Soient V un e.v. avec deux bases E et F . Soit b la matrice de changement de base de E vers F exprimée dans les coordonnées de E . Alors si $\lambda_1, \dots, \lambda_n$ représente les coordonnées de $x \in V$ dans la base E et μ_1, \dots, μ_n représente ses coordonnées dans la base F , on a :

$$\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} = b \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}.$$

Démonstration. On sait :

$$\sum_{j=1}^n \mu_j f_j = B(x) = \sum_{i=1}^n \lambda_i B(e_i) = \sum_{i=1}^n \lambda_i \sum_{j=1}^n B_{ji} f_j = \sum_{j=1}^n \left(\sum_{i=1}^n B_{ji} \lambda_i \right) f_j.$$

Il faut donc avoir, pour tout j :

$$\mu_j = \sum_{i=1}^n B_{ji} \lambda_i,$$

ou encore :

$$\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} = b \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}.$$

□

Proposition 7.15. Soit V un e.v., E, F deux bases de V et b la matrice de changement de base de E vers F de V . Soit $A \in \text{Hom}(V, V)$ un automorphisme de V . Soient $m_{E,E}(A) = [A_{ij}]$ est la matrice de l'application A dans les coordonnées de la base E et $m_{F,F}(A) = [A'_{ij}]$ la matrice de A dans la base F . Alors :

$$b \cdot m_{F,F}(A) = m_{E,E}(A) \cdot b.$$

Démonstration. Par définition des matrices $[A_{ij}]$ et $[A'_{ij}]$, on sait que :

$$A(e_j) = \sum_{k=1}^n A_{kj} e_k \quad \text{et} \quad A(f_j) = \sum_{k=1}^n A'_{kj} f_k.$$

Donc, en partant de $A(f_j)$, on obtient les égalités suivantes :

$$\begin{aligned} A(f_j) &= \sum_{k=1}^n A'_{kj} f_k = \sum_{k=1}^n A'_{kj} B(e_k) = \sum_{k=1}^n A'_{kj} \sum_{i=1}^n B_{ik} e_i = \sum_{i=1}^n \left(\sum_{k=1}^n B_{ik} A'_{kj} \right) e_i, \\ A(f_j) &= A(B(e_j)) = A \left(\sum_{k=1}^n B_{kj} e_k \right) = \sum_{k=1}^n B_{kj} A(e_k) = \sum_{k=1}^n B_{kj} \sum_{i=1}^n A_{ik} e_i = \sum_{i=1}^n \left(\sum_{k=1}^n A_{ik} B_{kj} \right) e_i. \end{aligned}$$

On sait donc que pour tout i, j , l'égalité suivante est vérifiée :

$$\sum_{k=1}^n B_{ik} A'_{kj} = \sum_{k=1}^n A_{ik} B_{kj},$$

ce qui revient à faire l'égalité suivante :

$$b \cdot m_{F,F}(A) = [B_{ij}][A'_{ij}] = [A_{ij}][B_{ij}] = m_{E,E}(A) \cdot b.$$

□

7.5 Dualité

Définition 7.16 (Définition naïve de la dualité). Soient deux objets X et Y et une application A . Si $A(X) = Y$ et $A(Y) = X$, on dit que X et Y sont duals.

Remarque. La géométrie projective comporte des exemples de dualité. On sait que deux points distincts désignent une droite et que deux droites distinctes (non-parallèles) désignent un point.

Définition 7.17. Soit V un espace vectoriel réel. On définit V^* l'espace dual de V par $V^* := \text{Hom}(V, \mathbb{R})$. On appelle les éléments de V^* des formes linéaires de V .

Proposition 7.18. Soit V un espace vectoriel de dimension finie n . Alors l'espace V^* est isomorphe à \mathbb{R}^n .

Démonstration. Soit $E = \{e_1, \dots, e_n\}$ une base de V et soit $F = \{1\}$ la base canonique de \mathbb{R} . Soit $f \in V^*$ une forme linéaire sur V . La forme f est caractérisée par la matrice $m_{F,E}(f) = [f(e_i)]_i$. La transformation linéaire $m_{F,E} : V^* \rightarrow M_{1 \times n}(\mathbb{R}) \sim \mathbb{R}^n$ est donc un isomorphisme. \square

Corollaire 7.19. Un espace vectoriel de dimension finie est isomorphe avec son dual.

Démonstration. Soient V et V^* un espace vectoriel et son dual. Si V est de dimension finie n , alors par la proposition 7.18, on sait que V^* est isomorphe à \mathbb{R}^n . Par la proposition 6.4, on sait que V est isomorphe à \mathbb{R}^n . Donc par composition des isomorphismes, il existe un isomorphisme entre V et V^* . \square

Définition 7.20. Soient un espace vectoriel V et une base $E = \{e_1, \dots, e_n\}$ de V ainsi que V^* le dual de V . On définit E^* la base duale de E par $E^* = \{e_1^*, \dots, e_n^*\}$ où pour tout k, j , on définit :

$$e_k^*(e_j) := \delta_{kj}.$$

Proposition 7.21. Soient V, V^* et E un espace vectoriel et son dual avec une base de E . Alors la base duale E^* de E est une base de V^* .

Démonstration. E^* est libre car :

$$0 = \left(\sum_{i=1}^n \lambda_i e_i^* \right) (e_j) = \sum_{i=1}^n \lambda_i e_i^*(e_j) = \sum_{i=1}^n \lambda_i \delta_{ij} = \lambda_j.$$

Montrons alors que E^* est génératrice. Soit $f \in V^*$. En effet, $f = \sum_{i=1}^n f(e_i) e_i^*$:

$$\left(\sum_{i=1}^n f(e_i) e_i^* \right) (e_j) = \sum_{i=1}^n f(e_i) e_i^*(e_j) = \sum_{i=1}^n f(e_i) \delta_{ij} = f(e_j).$$

\square

Remarque. Par rapport au corollaire 7.19, on a un isomorphisme $\phi : V \rightarrow V^*$ tel que $\phi(e_i) := e_i^*$.

Lemme 7.22. Soient V un espace vectoriel de dimension finie n et V^* son dual. Alors pour $f \in V^*$, si f n'est pas la fonction constante nulle, alors $\dim_{\mathbb{R}} \text{Ker}(f) = n - 1$

Démonstration. Si f n'est pas la fonction constante nulle, on sait que $\dim_{\mathbb{R}} \text{Im}(f) \geq 1$. De plus, $\text{Im}(f) \subseteq \mathbb{R}$. Donc $\dim_{\mathbb{R}} \text{Im}(f) \leq \dim_{\mathbb{R}} \mathbb{R} = 1$. Donc par la proposition 6.7, on sait que $\dim_{\mathbb{R}} \text{Ker}(f) = \dim_{\mathbb{R}} V - \dim_{\mathbb{R}} \text{Im}(f) = n - 1$. \square

Remarque. Le noyau de f est donc un hyperplan passant par l'origine et si $\lambda \in \mathbb{R}$, $f^{-1}(\lambda)$ est un hyperplan parallèle à $\text{Ker}(f)$.

Remarque. On sait que $f^{-1}(\lambda) = \{x \in V \text{ t.q. } f(x) = \lambda\}$. Soit $v \in V \setminus \text{Ker}(f)$. Tout vecteur $w \in V$ peut s'écrire comme : $v = u + \mu v$ avec $u \in \text{Ker}(f)$ et $\mu \in \mathbb{R}$. Dès lors $f(w) = f(u) + \mu f(v) = \mu f(v)$. On sait donc déterminer f par son noyau et un vecteur dans $V \setminus \text{Ker}(f)$.

De même, avec un espace V , un hyperplan W et $v \in V \setminus W$, il existe une unique forme linéaire $f \in V^*$ telle que $f(W) = \{0\}$ et $f(v) = \xi$ pour ξ fixé.

Définition 7.23. Soit V un espace vectoriel. On définit l'espace bidual de V par $V^{**} := (V^*)^* = \text{Hom}(V^*, \mathbb{R})$.

Proposition 7.24. Si V est un espace vectoriel de dimension finie, V^{**} est isomorphe à V .

Démonstration. On définit une fonction $\alpha : V \rightarrow V^{**}$ telle que pour $v \in V$, $\alpha(v) \in \text{Hom}(V^*, \mathbb{R})$ est définie par $\alpha(v)(f) = f(v)$. Il est évident que α est linéaire :

$$\alpha(v + \lambda w)(f) = f(v + \lambda w) = f(v) + \lambda f(w) = \alpha(v)(f) + \lambda \alpha(w)(f).$$

De plus, si $E = \{e_1, \dots, e_n\}$ est une base de V , on définit la base biduale de E par $E^{**} = \{\alpha(e_1), \dots, \alpha(e_n)\}$. On sait que E^{**} est une base de V^{**} car E^{**} est la base duale de E^* . En effet :

$$\alpha(e_i)(e_j^*) = e_j^*(e_i) = \delta_{ij}.$$

□

7.6 Matrices transposées

Définition 7.25. Soit $A : V \rightarrow W$ une application linéaire. On définit la transposée de A par $A^T : W^* \rightarrow V^* : f \mapsto (f \circ A)$.

Proposition 7.26. L'application transposée est linéaire.

Démonstration. Montrons la linéarité :

$$A^T(f + \lambda g)(v) = ((f + \lambda g) \circ A)(v) = (f \circ A)(v) + \lambda(g \circ A)(v) = A^T(f)(v) + \lambda A^T(g)(v).$$

□

Définition 7.27. Soient V, W deux e.v. avec des bases E, F . On définit la matrice $A^T := m_{E^*, F^*}(A^T) = [A_{ij}^T]$ comme la matrice transposée de A .

Proposition 7.28. Soit $A \in \text{Hom}(V, W)$ une application linéaire et $A = [A_{ij}]$ sa matrice associée. Alors la matrice transposée A^T est définie telle que pour tout i, j , on a $A_{ij}^T = A_{ji}$.

Démonstration. Si $A^T = m_{E^*, F^*}(A^T)$ est la matrice de l'application transposée, alors :

$$A^T(f_j^*)(e_i) = (f_j^* \circ A)(e_i) = f_j^*(A(e_i)) = f_j \left(\sum_{k=1}^n A_{ki} f_k \right) = \sum_{k=1}^n A_{ki} f_j^*(f_k) = \sum_{k=1}^n A_{ki} \delta_{jk} = A_{ji}.$$

Mais par définition de la matrice, on a également :

$$A^T(f_j^*)(e_i) = \left(\sum_{k=1}^n A_{kj}^T e_k^* \right) (e_i) = \sum_{k=1}^n A_{kj}^T e_k^*(e_i) = \sum_{k=1}^n A_{kj}^T \delta_{ki} = A_{ij}^T.$$

Il faut donc avoir $A_{ij}^T = A_{ji}$. □

Remarque. On peut donc noter $m_{E^*, F^*}(A^T) = m_{F, E}(A)^T$.

Définition 7.29. Soit $a \in M_{n \times n}(\mathbb{R})$ une matrice. Si $a = a^T$, on dit que a est symétrique. Si $a = -a^T$, on dit que a est antisymétrique.

Définition 7.30. Soit $a = [a_{ij}] \in M_{s \times t}(\mathbb{R})$ une matrice. Soit $A : V \rightarrow W : v \mapsto av$ la transformation linéaire associée à a . On définit le rang de la matrice a comme le rang de l'application A . Donc $\text{rang}(a) = \text{rang}(A) = \dim_{\mathbb{R}} \text{Im}(A)$.

Proposition 7.31. Soit $A : V \rightarrow W$ une transformation linéaire. Alors $\text{rang}(A) = \text{rang}(A^T)$.

Démonstration. Soit $E = \{e_1, \dots, e_n\}$ une base de V . Soit $A(E) = \{A(e_1), \dots, A(e_r)\}$ une base de l'image de A . Soit $F = \{A(e_1), \dots, A(e_r), f_{r+1}, \dots, f_s\}$ une base de W . Dès lors, la matrice $m_{F, E}$ correspond à la matrice identité sur $[A_{ij}]_{1 \leq i, j \leq r}$. De plus, la matrice est nulle sur $[A_{ij}]_{r+1 \leq i \leq s, 1 \leq j \leq r}$. Donc forcément, le rang de la transposée doit être au moins égal au rang de A . On a donc $\text{rang}(A) \leq \text{rang}(A^T)$. Par un raisonnement similaire, on trouve $\text{rang}(A^T) \leq \text{rang}((A^T)^T)$. Or on sait que $(A^T)^T = A$. Donc on a $\text{rang}(A) \leq \text{rang}(A^T) \leq \text{rang}(A)$. Il faut alors $\text{rang}(A) = \text{rang}(A^T)$. □

8 Permutations et Déterminants

8.1 Définitions

Définition 8.1. Soit $n \in \mathbb{N}^*$. On note l'ensemble des permutations de $\{1, 2, \dots, n\}$:

$$\mathfrak{S}_n := \{f : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\} \text{ t.q. } f \text{ est bijective}\}.$$

On appelle cet ensemble le **groupe symétrique de degré n** .

Définition 8.2. Soient n un naturel non nul et \mathfrak{S}_n , le groupe symétrique de degré n . Soit $\sigma \in \mathfrak{S}_n$ une permutation de $\{1, \dots, n\}$. Il existe quatre notations pour σ :

1. le tableau double :

$$\sigma = \begin{bmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{bmatrix} ;$$

2. la chaîne :

$$\sigma = \sigma(1)\sigma(2)\dots\sigma(n) ;$$

3. la tresse : deux lignes comportant les points $\{1, 2, \dots, n\}$ reliés par des arêtes $(i, \sigma(i))$;

4. le produit de cycles :

$$\sigma = (\sigma(k_{1,1})\sigma(k_{1,2})\dots\sigma(k_{1,r})) (\sigma(k_{2,1})\sigma(k_{2,2})\dots\sigma(k_{2,s})) \dots (\sigma(k_{t,1})\sigma(k_{t,2})\dots\sigma(k_{t,n})) .$$

Remarque. Quand le degré du groupe symétrique \mathfrak{S}_n n'est pas ambigu, lors d'une notation par produit de cycles, les points fixes sont omis. De plus, l'ordre d'expression des cycles n'est pas important.

Définition 8.3. Soit $\sigma \in \mathfrak{S}_n$ une permutation. Si σ ne comporte qu'un seul cycle de longueur $k \geq 2$, alors σ est dit **cycle de longueur k** . Un cycle de longueur deux est appelé **transposition**.

Théorème 8.4. Le nombre de permutations possibles d'ordre n est $n!$.

Démonstration. Soit \mathfrak{S}_n , le groupe symétrique de degré n . Soit $\sigma \in \mathfrak{S}_n$. Déterminons $\sigma(i)$ pour $i \in \{1, 2, \dots, n\}$:

- il y a n choix pour $\sigma(1)$ (toute valeur dans $\{1, 2, 3, \dots, n\}$) ;
- il y a $(n - 1)$ choix pour $\sigma(2)$ (toute valeur dans $\{1, 2, \dots, n\} \setminus \{\sigma(1)\}$) ;
- il y a $(n - 2)$ choix pour $\sigma(3)$ (toute valeur dans $\{1, 2, \dots, n\} \setminus \{\sigma(1), \sigma(2)\}$) ;
- ... ;
- il y a 1 choix pour $\sigma(n)$.

Le nombre total de choix possibles est donc $n(n - 1)(n - 2) \dots 1 = n!$. □

Remarque. Les permutations sont des fonctions, on peut donc les composer. La composition $\sigma \circ \tau$ se note également $\sigma\tau$. La composition k fois de la permutation σ avec elle-même se note σ^k .

Définition 8.5. Soit $\sigma \in \mathfrak{S}_n$. On appelle le **degré** de σ le plus petit nombre naturel non nul k tel que $\sigma^k = \text{Id}_{\mathfrak{S}_n}$.

Définition 8.6. Soit $\sigma \in \mathfrak{S}_n$. Comme σ est une bijection, elle admet un inverse noté σ^{-1} tel que $\sigma \circ \sigma^{-1} = \text{Id}_{\mathfrak{S}_n} = \sigma^{-1} \circ \sigma$.

Lemme 8.7. Soit $\tau \in \mathfrak{S}_n$ une transposition. L'inverse τ^{-1} de τ est τ .

Démonstration. Soit $\tau = (i, j) \in \mathfrak{S}_n$ où $1 \leq i < j \leq n$, montrons que $\tau^{-1} = \tau$. En effet, $\tau^2 = \tau\tau = \tau \circ \tau$ revient à permuter deux fois les positions i et j de la permutation, et donc de remettre i et j à leur place initiale :

$$(\tau\tau)(n) = \tau \left(\begin{cases} j & \text{si } n = i \\ i & \text{si } n = j \\ n & \text{sinon} \end{cases} \right) = \begin{cases} i & \text{si } n = i \\ j & \text{si } n = j \\ n & \text{sinon} \end{cases} = \text{Id}_{\mathfrak{S}_n}$$

□

Lemme 8.8. Soient $\tau_1, \tau_2, \dots, \tau_n$, n transpositions et $\sigma \in \mathfrak{S}_n$ une permutation telle que $\sigma = \tau_1\tau_2 \dots \tau_n$. Alors $\sigma^{-1} = \tau_n\tau_{n-1} \dots \tau_1$.

Démonstration. Soient $\sigma \in \mathfrak{S}_n$ et τ_1, \dots, τ_n des transpositions telles que $\sigma = \tau_1 \dots \tau_n$. Montrons que $\sigma\sigma^{-1} = \sigma^{-1}\sigma = \text{Id}_{\mathfrak{S}_n}$.

On sait par le lemme 8.7 que $\tau_i\tau_i = \text{Id}_{\mathfrak{S}_n}$ pour tout i , et on calcule :

$$\begin{aligned} \tau_1\tau_2 \dots \tau_{n-1}\tau_n\tau_n\tau_{n-1} \dots \tau_2\tau_1 &= \tau_1\tau_2 \dots \tau_{n-1}(\tau_n\tau_n)\tau_{n-1} \dots \tau_2\tau_1 = \tau_1\tau_2 \dots \tau_{n-1} \text{Id}_{\mathfrak{S}_n} \tau_{n-1} \dots \tau_2\tau_1 \\ &= \tau_1\tau_2 \dots (\tau_{n-1}\tau_{n-1}) \dots \tau_2\tau_1 = \tau_1\tau_2 \dots \text{Id}_{\mathfrak{S}_n} \dots \tau_2\tau_1 = \dots = \tau_1\tau_2\tau_2\tau_1 = \tau_1 \text{Id}_{\mathfrak{S}_n} \tau_1 \\ &= \text{Id}_{\mathfrak{S}_n} . \end{aligned}$$

Par un raisonnement similaire, on montre que :

$$\prod_{j=1}^n \tau_{n+1-j} \prod_{j=1}^n \tau_j = \text{Id}_{\mathfrak{S}_n} .$$

Et puisque $\sigma = \tau_1 \dots \tau_n$, on sait que $\tau_n \dots \tau_1$ est son inverse, σ^{-1} .

□

Proposition 8.9. Soit $\sigma \in \mathfrak{S}_n$ un cycle de longueur k . Alors $\forall m \in \mathbb{N}^* : \sigma^{mk} = \text{Id}_{\mathfrak{S}_n}$.

Démonstration. Montrons que $\sigma^k = \text{Id}_{\mathfrak{S}_n}$. Par définition, on sait qu'il existe k nombres deux à deux distincts $n_1, \dots, n_k \in \{1, \dots, n\}$ tels que $\forall i \in \{1, \dots, n\} : n_i = \sigma^{i-1}(n_1)$ et $n - k$ nombres restants n_{k+i} (avec $i \in \{1, \dots, n - k\}$) tels que $\sigma(n_{k+i}) = n_{k+i}$. Dès lors, pour tout $j \in \{1, \dots, k\}$, on a $\sigma^k(n_j) = \sigma^k(\sigma^{j-1}(n_1)) = \sigma^{k+j-1}(n_1) = \sigma^{j-1}(\sigma^k(n_1)) = \sigma^{j-1}(\sigma(n_k)) = \sigma^{j-1}(n_1) =: n_j$. On a donc bien $\sigma^k = \text{Id}_{\mathfrak{S}_n}$.

Maintenant montrons que $\forall m \geq 1 : \sigma^{mk} = \sigma^{(m-1)k}$. En effet, $\sigma^{mk}(n) = \sigma^{(m-1)k}(\sigma^k(n)) = \sigma^{(m-1)k}(n)$.

Dès lors, on sait que $\sigma^{mk}(n) = \sigma^{(m-1)k}(n) = \sigma^{(m-2)k}(n) = \dots = \sigma^{2k}(n) = \sigma^k(n) = n$.

□

Remarque. Le lemme 8.7 est un cas particulier de cette proposition pour un cycle de longueur $k = 2$.

Corollaire 8.10. Soit $\sigma \in \mathfrak{S}_n$ une permutation admettant n cycles de longueur respective k_1, k_2, \dots, k_n . Le degré de σ est le plus petit commun multiple des longueurs des cycles. Donc $\deg \sigma = \text{LCM}(k_1, k_2, \dots, k_n)$.

Démonstration. On sait par la proposition 8.9 qu'un cycle mis à la puissance de sa longueur donne l'identité. Soit $K := \text{LCM}(k_1, \dots, k_n)$, le plus petit commun multiple des longueurs des cycles de la permutation σ . Par définition, K est un multiple de tous les nombres k_i pour $i \in \{1, \dots, n\}$, donc pour chaque cycle individuel, la puissance K donne l'identité. On a donc $\sigma^K = \text{Id}_{\mathfrak{S}_n}$ et on sait par définition que K est le plus petit nombre naturel ≥ 1 satisfaisant cette propriété. K est donc le degré de la permutation σ . \square

Définition 8.11. Soient $\sigma, \tau \in \mathfrak{S}_n$ deux permutations. La **permutation conjuguée** de σ par τ^{-1} est la bijection donnée par :

$$\tau\sigma\tau^{-1} := \tau \circ \sigma \circ \tau^{-1}.$$

Théorème 8.12. La permutation conjuguée conserve les cycles telle que si $\sigma = (a_1, a_2, \dots, a_n) \dots (b_1, b_2, \dots, b_m)$, alors :

$$\tau\sigma\tau^{-1} = (\tau(a_1), \tau(a_2), \dots, \tau(a_n)) \dots (\tau(b_1), \tau(b_2), \dots, \tau(b_m)).$$

Démonstration. Montrons que $(\tau(a_1), \dots, \tau(a_m))$ est une permutation de $\gamma := \tau\sigma\tau^{-1}$. Soit $i \in \{1, \dots, m\}$. On sait $\gamma(\tau(a_i)) = (\gamma\tau)(a_i) = (\tau\sigma\tau^{-1}\tau)(a_i) = (\tau\sigma)(a_i) = \tau(\sigma(a_i))$. Si $i < m$, alors $\gamma(\tau(a_i)) = \tau(a_{i+1})$, et si $i = m$, alors $\gamma(\tau(a_i)) = \tau(\sigma(a_m)) = \tau(a_1)$. En appliquant le même raisonnement à tous les cycles de σ , on montre que les cycles c sont conservés et qu'ils sont donnés par $\tau(c)$. \square

8.2 Signe d'une permutation

Définition 8.13. Soit $\sigma \in \mathfrak{S}_n$, une permutation. Une inversion dans σ est un couple $(i, j) \in \{1, 2, \dots, n\}^2$ tel que $i < j$ et $\sigma(i) > \sigma(j)$.

Définition 8.14. Soit $\sigma \in \mathfrak{S}_n$ une permutation. On définit $\text{sign}(\sigma)$ le signe de σ par :

$$\text{sign}(\sigma) = (-1)^{N(\sigma)},$$

où $N(\sigma)$ est le plus petit nombre d'inversions possible pour transformer $\text{Id}_{\mathfrak{S}_n}$ en σ . Si $\text{sign}(\sigma) = -1$, on dit que σ est impaire, sinon σ est paire.

Proposition 8.15. Soit $\sigma \in \mathfrak{S}_n$ une permutation. Supposons qu'il existe $m+n$ transpositions $\tau_1, \dots, \tau_m, \tau'_1, \dots, \tau'_n$ telles que :

$$\sigma = \prod_{i=1}^m \tau_i = \prod_{j=1}^n \tau'_j.$$

Alors m et n sont de même parité ($m \bmod 2 = n \bmod 2 = p \in \{0, 1\}$).

Démonstration. Toute inversion peut s'écrire comme un produit (une composition) de transpositions adjacentes : soient $1 \leq i < j \leq n$. L'inversion (i, j) peut s'écrire comme suit :

$$\tau = (j, j-1)(j-1, j-2) \dots (i+1, i+2)(i, i+1)(i+1, i+2) \dots (j-2, j-1)(j-1, j).$$

Le nombre de transpositions adjacentes est donc $2(j-i-2)+1$ qui est impair. Dès lors, prenons $\hat{\tau}_1, \dots, \hat{\tau}_m$, l'ensemble des transpositions adjacentes permettant de réécrire $\tau_1 \dots \tau_m$ et $\hat{\tau}'_1, \dots, \hat{\tau}'_n$, l'ensemble des transpositions adjacentes permettant de réécrire $\tau'_1 \dots \tau'_n$. On a donc les égalités suivantes :

$$\begin{aligned} \sigma &= \tau_1 \dots \tau_m &= \hat{\tau}'_1 \dots \hat{\tau}'_n \\ &= \hat{\tau}_1 \dots \hat{\tau}_m &= \hat{\tau}'_1 \dots \hat{\tau}'_n. \end{aligned}$$

Soient ν_1, \dots, ν_m et ν'_1, \dots, ν'_n le nombre de transpositions adjacentes de τ_i et τ'_i . On sait que $\sum_i \nu_i = m'$ et $\sum_i \nu'_i = n'$. Dès lors, on sait calculer :

$$m' - m = \sum_{i=1}^m \nu_i - m = \sum_{i=1}^m (\nu_i - 1).$$

De plus, étant donné que ν_i est impair pour tout i , la valeur $(\nu_i - 1)$ est paire. La quantité $m' - m$ est donc une somme de valeurs paires et est donc paire également. Par un raisonnement similaire, on obtient $n' - n$ pair. Par le lemme 8.8, on sait que $\sigma^{-1} = \tau_m \dots \tau_1$. En composant σ et son inverse σ^{-1} , on obtient l'identité $\text{Id}_{\mathfrak{S}_n}$.

Soit $N(\sigma)$ le nombre d'inversions de σ . En évaluant $N(\sigma) - m' \pmod{2} \equiv N(\sigma \tau_k \dots \tau_1) \pmod{2} = N(\text{Id}_{\mathfrak{S}_n}) \pmod{2} = 0$, on trouve que $N(\sigma) - m'$ est pair et donc $N(\sigma) - m' + (m' - m) = N(\sigma) - m$ est pair également car une somme de nombres pairs est paire. De manière similaire, on trouve $N(\sigma) - n$ pair également. La quantité $(N(\sigma) - n) - (N(\sigma) - m) = m - n$ est une différence de valeurs paires et est donc paire également. Or m et n sont de même parité si et seulement si $m - n$ est pair. \square

Proposition 8.16. Soient $\sigma, \tau \in \mathfrak{S}_n$. L'opérateur sign respecte les propriétés suivantes :

1. $\text{sign}(\sigma\tau) = \text{sign}(\sigma) \text{sign}(\tau)$;
2. $\text{sign}(\sigma^{-1}) = \text{sign}(\sigma)$.

Démonstration. Soient $\sigma, \tau \in \mathfrak{S}_n$.

1. Supposons $\sigma = \tau_1 \dots \tau_n$ et $\tau = \tau'_1 \dots \tau'_m$. On sait $\text{sign}(\sigma\tau) = (-1)^{m+n} = (-1)^m (-1)^n = \text{sign}(\tau) \text{sign}(\sigma)$.
2. Supposons $\sigma = \tau_1 \dots \tau_n$. On sait $\sigma^{-1} = \tau_n \dots \tau_1$. σ et son inverse σ^{-1} ont donc la même parité, et alors le même signe.

\square

Remarque. Si σ est une permutation, alors $\text{sign}(\sigma) = 1 \iff \sigma$ est paire.

Lemme 8.17. Soit $\sigma \in \mathfrak{S}_n$ un cycle de longueur k . Si $\sigma = (n_1, n_2, \dots, n_k)$ et $\tau_i = (n_i, n_{i+1})$ pour $1 \leq i < k$, alors $\sigma = \tau_1 \tau_2 \dots \tau_{k-1}$.

Démonstration. Soient $k - 1$ transpositions $\tau_i = (n_i, n_{i+1})$. La composition s'évaluant de droite à gauche, on a :

$$(\tau_1 \dots \tau_{k-1})(n_i) = (\tau_1 \dots \tau_{i-1}) \left(\tau_i [(\tau_{i+1} \dots \tau_{k-1})(n_i)] \right) = (\tau_1 \dots \tau_{i-1}) (\tau_i[n_i]) = (\tau_1 \dots \tau_{i-1})(n_{i+1}) = n_{i+1}.$$

En prenant $m \notin \{n_1, \dots, n_{k-1}\}$, alors aucune des $k - 1$ transpositions ne modifie m , donc $(\tau_1 \dots \tau_{k-1})$ agit comme l'identité $\text{Id}_{\mathfrak{S}_n}$. On a donc bien $\sigma = \tau_1 \dots \tau_{k-1}$. \square

Corollaire 8.18. Soit $\sigma \in \mathfrak{S}_n$ un cycle de longueur k . Alors σ est pair si et seulement si k est impair.

Démonstration. On sait par le lemme 8.17 qu'il existe $k - 1$ transpositions τ_i telles que $\sigma = \tau_1 \dots \tau_{k-1}$. Dès lors, $\text{sign}(\sigma) = (-1)^{k-1}$. Si k est pair, alors $\text{sign}(\sigma) = -1$ et donc σ est impair, et si k est impair, alors $\text{sign}(\sigma) = 1$ et donc σ est pair. \square

8.3 Déterminants

Définition 8.19. Soient $x, y, z \in \mathbb{R}^2$ des points du plan, $\lambda \in \mathbb{R}$. On définit l'**aire orientée** $A : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ par les propriétés suivantes :

1. $A^{\text{or}}(x, y + z) = A^{\text{or}}(x, y) + A^{\text{or}}(x, z) ;$
2. $A^{\text{or}}(x + y, z) = A^{\text{or}}(x, z) + A^{\text{or}}(y, z) ;$
3. $A^{\text{or}}(\lambda x, y) = \lambda A^{\text{or}}(x, y) = A^{\text{or}}(x, \lambda y) ;$
4. $A^{\text{or}}(x, x) = 0.$

Si $\{e_1, e_2\}$ est une base de \mathbb{R}^2 , alors on définit l'unité d'aire par $A(e_1, e_2) = 1$.

Remarque. L'aire est dite *orientée* car $A(x, y) = -A(y, x)$. En effet :

$$0 = A^{\text{or}}(x + y, x + y) = A^{\text{or}}(x, x) + A^{\text{or}}(x, y) + A^{\text{or}}(y, x) + A^{\text{or}}(y, y) = A^{\text{or}}(x, y) + A^{\text{or}}(y, x).$$

Définition 8.20. Soit V un espace réel. Soit $f : V^n \rightarrow \mathbb{R}$ une fonction à valeurs réelles. On dit que f est une **forme multilinéaire** si f est linéaire pour toute composante v_i de $(v_1, \dots, v_n) \in V^n$. Si lorsque $a_i = a_j$ pour $i \neq j$, alors on dit que f est une **forme multilinéaire alternée**.

Proposition 8.21. Soient V un espace réel, $f : V^n \rightarrow \mathbb{R}$ une forme multilinéaire variée et $1 \leq i < j \leq n$. Alors $f(v_1, \dots, v_i, \dots, v_j, \dots, v_n) = -f(v_1, \dots, v_j, \dots, v_i, \dots, v_n)$.

Démonstration. Par définition d'une forme multilinéaire variée, si deux paramètres v_i et v_j sont égaux pour $i \neq j$, alors la fonction s'annule en ce point. Dès lors :

$$\begin{aligned} 0 &= f(v_1, \dots, v_i + v_j, \dots, v_i + v_j, \dots, v_n) = f(v_1, \dots, v_i, \dots, v_j, \dots, v_n) + f(v_1, \dots, v_i, \dots, v_i, \dots, v_n) \\ &\quad + f(v_1, \dots, v_j, \dots, v_j, \dots, v_n) + f(v_1, \dots, v_j, \dots, v_i, \dots, v_n) \\ &= f(v_1, \dots, v_i, \dots, v_j, \dots, v_n) + 0 + 0 + f(v_1, \dots, v_j, \dots, v_i, \dots, v_n). \end{aligned}$$

□

Proposition 8.22. Soient V un espace vectoriel réel de dimension n et $f : V^n \rightarrow \mathbb{R}$ une forme multilinéaire alternée. Soient $v_1, \dots, v_n \in V$ tels qu'il existe $1 \leq i \leq n$ tel que $v_i = 0$. Alors

$$f(v_1, \dots, v_n) = 0.$$

Démonstration. Soit un tel i et soit $j \neq i$. On pose $v_i := v_j - v_j$, ce qui est admissible car $v_i = 0$. Dès lors, par linéarité, on a :

$$\begin{aligned} f(v_1, \dots, v_i, \dots, v_j, \dots, v_n) &= f(v_1, \dots, v_j - v_j, \dots, v_j, \dots, v_n) \\ &= f(v_1, \dots, v_j, \dots, v_j, \dots, v_n) - f(v_1, \dots, v_j, \dots, v_j, \dots, v_n) = 0. \end{aligned}$$

□

Proposition 8.23. Soient V un espace vectoriel réel de dimension n et $E = \{e_1, \dots, e_n\}$ une base de V . Alors il existe une et une seule forme multilinéaire alternée $f : V^n \rightarrow \mathbb{R}$ telle que :

$$f(e_1, \dots, e_n) = 1.$$

Démonstration. Supposons qu'une telle fonction existe, et montrons que :

$$f(v_1, \dots, v_n) = \sum_{\sigma \in \mathfrak{S}_n} \left(\prod_{j=1}^n a_{j \sigma(j)} \right) f(e_{1 \sigma(1)}, \dots, e_{n \sigma(n)}).$$

En effet, en posant pour tout $1 \leq i \leq n$:

$$v_i = \sum_{j=1}^n v_{ij} e_j,$$

on peut déterminer, par (multi)linéarité de f :

$$\begin{aligned} f(v_1, \dots, v_n) &= f\left(\sum_{j=1}^n v_{1j} e_j, \dots, \sum_{j=1}^n v_{nj} e_j\right) = \sum_{j=1}^n v_{1j} f\left(e_i, \dots, \sum_{k=1}^n v_{nk} e_k\right) \\ &= \sum_{k_1=1}^n \sum_{k_2=1}^n \dots \sum_{k_n=1}^n v_{1k_1} v_{2k_2} \dots v_{nk_n} f(e_{k_1}, \dots, e_{k_n}). \end{aligned}$$

Or, si deux indices k_α sont égaux, la fonction s'annule. Dès lors, k_2 doit être différent de k_1 , k_3 doit être différent de k_1 et de k_2 , etc. Dès lors, l'ensemble des k_i forme l'ensemble $\{1, 2, \dots, n\}$ et alors $i \mapsto k_i$ est une bijection. Comme une bijection d'un ensemble discret est une permutation, on peut sommer sur l'ensemble des permutations :

$$f(v_1, \dots, v_n) = \sum_{\sigma \in \mathfrak{S}_n} v_{1 \sigma(1)} \dots v_{n \sigma(n)} f(e_{\sigma(1)}, \dots, e_{\sigma(n)}).$$

Par la proposition 8.21, on sait que $f(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = \text{sign}(\sigma) f(e_1, \dots, e_n)$. Dès lors, on a bien :

$$f(v_1, \dots, v_n) = \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \left(\prod_{i=1}^n v_{i \sigma(i)} \right) f(e_1, \dots, e_n).$$

De plus, en imposant $f(e_1, \dots, e_n) = 1$, on peut poser :

$$\det(v_1, \dots, v_n) := f(v_1, \dots, v_n) = \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \left(\prod_{i=1}^n v_{i \sigma(i)} \right).$$

Maintenant, montrons que \det est une forme multilinéaire alternée. Soient $\lambda \in \mathbb{R}$, $a_1, \dots, a_{i-1}, a_i^{(1)}, a_i^{(2)}, a_{i+1}, \dots, a_n \in V$.

$$\begin{aligned} \det(a_1, \dots, a_{i-1}, a_i^{(1)} + \lambda a_i^{(2)}, a_{i+1}, \dots, a_n) &:= \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \left(\prod_{k=1, k \neq i}^n a_{k \sigma(k)} \right) (a_{i \sigma(i)}^{(1)} + a_{i \sigma(i)}^{(2)}) \\ &= \sum_{\sigma \in \mathfrak{S}_n} \left[\text{sign}(\sigma) \left(\prod_{k=1, k \neq i}^n a_{k \sigma(k)} \right) a_i^{(1)} + \text{sign}(\sigma) \left(\prod_{k=1, k \neq i}^n a_{k \sigma(k)} \right) \lambda a_i^{(2)} \right]. \end{aligned}$$

En posant :

$$\begin{aligned} a_i &:= a_i^{(1)} + \lambda a_i^{(2)}, \\ b_\alpha &:= \begin{cases} a_\alpha & \text{si } \alpha \neq i \\ a_i^{(1)} & \text{si } \alpha = i \end{cases}, \\ c_\beta &:= \begin{cases} a_\beta & \text{si } \beta \neq i \\ a_i^{(2)} & \text{si } \beta = i \end{cases}. \end{aligned}$$

on obtient :

$$\begin{aligned} \det(a_1, \dots, a_n) &= \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \prod_{k=1}^n b_{k \sigma(k)} + \lambda \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \prod_{k=1}^n c_{k \sigma(k)} \\ &= \det(a_1, \dots, a_{i-1}, a_i^{(1)}, a_{i+1}, \dots, a_n) + \lambda \det(a_1, \dots, a_{i-1}, a_i^{(2)}, a_{i+1}, \dots, a_n). \end{aligned}$$

\det est donc bien une forme multilinéaire car linéaire en chacun de ses paramètres. Il reste à montrer que \det est une forme multilinéaire *alternée*. Pour ce faire, supposons qu'il existe $1 \leq i < j \leq n$ tels que $a_i = a_j$. Dès lors, on sait :

$$\det(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = \det(a_1, \dots, a_j, \dots, a_i, \dots, a_n).$$

De plus, $\forall t \in \{1, 2, \dots, n\} : a_{it} = a_{jt}$. Soit τ_{ij} la transposition (i, j) . Intervertir les coordonnées i et j n'a pas d'influence sur le résultat car $a_{\tau_{ij}(j)} = a_i = a_j = a_{\tau_{ij}(i)}$. Dès lors :

$$\begin{aligned} \det(a_1, \dots, a_i, \dots, a_j, \dots, a_n) &= \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma \tau_{ij}) \left(\prod_{k=1}^n a_{k (\sigma \tau_{ij})(k)} \right) \\ &= \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma \tau_{ij}) \left(\prod_{k=1, k \notin \{i, j\}}^n a_{k (\sigma \tau_{ij})(k)} \right) a_{i (\sigma \tau_{ij})(i)} a_{j (\sigma \tau_{ij})(j)} \\ &= \sum_{\sigma \in \mathfrak{S}_n} -\text{sign}(\sigma) \left(\prod_{k=1, k \notin \{i, j\}}^n a_{k \sigma(k)} \right) a_{i \sigma(j)} a_{j \sigma(i)} \\ &= - \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \left(\prod_{k=1, k \notin \{i, j\}}^n a_{k \sigma(k)} \right) a_{j \sigma(j)} a_{i \sigma(j)} \\ &= - \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \left(\prod_{k=1}^n a_{k \sigma(k)} \right) \\ &= -\det(a_1, \dots, a_i, \dots, a_j, \dots, a_n). \end{aligned}$$

Dès lors, on a $\det(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = -\det(a_1, \dots, a_i, \dots, a_j, \dots, a_n)$, ce qui implique :

$$\det(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = 0.$$

□

8.4 Matrices et applications linéaires

Définition 8.24. Soit $a \in M_{n \times n}(\mathbb{R})$ une matrice carrée à valeurs réelles. On appelle le **déterminant de la matrice** a la valeur :

$$\det(a) = \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \prod_{k=1}^n a_{k \sigma(k)}.$$

Proposition 8.25. Soit $a \in M_{n \times n}(\mathbb{R})$ une matrice. Alors $\det(a) = \det(a^T)$.

Démonstration. Soit $a = [a_{ij}] \in M_{n \times n}(\mathbb{R})$. On sait que $a^T = [a_{ji}] \in M_{n \times n}(\mathbb{R})$. Dès lors, le déterminant de la matrice transposée vaut :

$$\begin{aligned} \det(a^T) &= \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \prod_{k=1}^n (a^T)_{k \sigma(k)} = \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \prod_{k=1}^n a_{\sigma(k) k} \\ &= \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \prod_{k=1}^n a_{(\sigma^{-1} \sigma)(k) \sigma^{-1}(k)}. \end{aligned}$$

On peut permuter tous les indices car σ^{-1} est également une permutation et donc une bijection. Les couples d'indices sont donc conservés. Dès lors :

$$\det(a^T) = \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \prod_{k=1}^n a_{k \sigma^{-1}(k)}.$$

Or, par la proposition 8.16, on sait que $\text{sign}(\sigma) = \text{sign}(\sigma^{-1})$. Dès lors :

$$\det(a^T) = \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma^{-1}) \prod_{k=1}^n a_{k \sigma^{-1}(k)}.$$

De plus, \mathfrak{S}_n est un groupe, donc toute permutations $\sigma \in \mathfrak{S}_n$ admet une permutation inverse $\sigma^{-1} \in \mathfrak{S}_n$. On peut donc conclure par :

$$\det(a^T) = \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \prod_{k=1}^n a_{k \sigma(k)} = \det(a).$$

□

Remarque. L'application \det étant une forme multilinéaire, l'inversion de deux lignes (ou colonnes) d'une matrice change le signe de son déterminant (par la proposition 8.21).

De plus, on sait que si une matrice a deux lignes (ou colonnes) égales, alors le déterminant s'annule car c'est une forme multilinéaire **variée**.

Proposition 8.26. Soit $a \in M_{n \times n}(\mathbb{R})$ une matrice triangulaire. Alors :

$$\det(a) = \prod_{i=1}^n a_{ii}.$$

Démonstration. Supposons que $a \in M_{n \times n}(\mathbb{R})$ soit triangulaire. On sait donc que si $i < j$, alors $a_{ij} = 0$. Le déterminant est défini par la somme des produits des *diagonales* $a_{i \sigma(i)}$. Or, si $\sigma(i) > i$, on sait que $a_{i \sigma(i)} = 0$. Dès lors, la seule permutation telle que $\forall i : i \geq \sigma(i)$ est $\text{Id}_{\mathfrak{S}_n}$. Cela peut se construire comme suit : il faut

que $1 \leq \sigma(1)$, donc on impose $\sigma(1) = 1$. Il faut également $\sigma(2) \leq 2$, ce qui impose que $\sigma(2) \in \{1, 2\}$, or σ doit être une bijection, et 1 est déjà l'image de $\sigma(1)$. En réitérant jusque n , on construit la permutation identité qui est la seule à respecter cette propriété. On a donc :

$$\begin{aligned} \det(a) &= \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \prod_{k=1}^n a_{k \sigma(k)} \\ &= \text{sign}(\text{Id}_{\mathfrak{S}_n}) \prod_{k=1}^n a_{k \text{Id}_{\mathfrak{S}_n}(k)} + \sum_{\sigma \in \mathfrak{S}_n \setminus \{\text{Id}_{\mathfrak{S}_n}\}} \text{sign}(\sigma) \prod_{k=1}^n a_{k \sigma(k)} \\ &= \text{sign}(\text{Id}_{\mathfrak{S}_n}) \prod_{k=1}^n a_{k \text{Id}_{\mathfrak{S}_n}(k)} + 0 = \text{sign}(\text{Id}_{\mathfrak{S}_n}) \prod_{k=1}^n a_{k \text{Id}_{\mathfrak{S}_n}(k)} = 1 \times \prod_{k=1}^n a_{kk}. \end{aligned}$$

□

Remarque. Afin de calculer le déterminant de la matrice, on peut utiliser la méthode de Gauss pour trianguler la matrice. De là, il ne reste plus qu'à évaluer le produit de la diagonale ; les opérations élémentaires de Gauss ne modifiant pas le déterminant.

Proposition 8.27. Soient $a, b \in M_{n \times n}(\mathbb{R})$ deux matrices carrées. Alors :

$$\det(ab) = \det(a) \det(b).$$

Démonstration. Soient $a, b \in M_{n \times n}(\mathbb{R})$ et $A, B \in \text{Hom}(V, V)$ les fonctions linéaires associées aux matrices. On sait que :

$$\det(B(A(e_1)), \dots, B(A(e_n))) = \det(a) \det(B(e_1), \dots, B(e_n)) = \det(a) \det(b) \det(e_1, \dots, e_n) = \det(a) \det(b).$$

Cependant, on sait également que :

$$\det(B(A(e_1)), \dots, B(A(e_n))) = \det((B \circ A)(e_1), \dots, (B \circ A)(e_n)) = \det(ba) \det(e_1, \dots, e_n) = \det(ba).$$

Dès lors, on a bien :

$$\det(ab) = \det(a) \det(b),$$

ce qui veut dire que le déterminant d'un produit est égal au produit des déterminants. □

Proposition 8.28. Soit $b \in M_{n \times n}(\mathbb{R})$, une matrice carrée inversible. Alors $\det(b^{-1}) = \det(b)^{-1}$.

Démonstration. On sait que l'identité est une matrice triangulaire. Dès lors, par la proposition 8.26, on sait que son déterminant est égal au produit des éléments de sa diagonale. Donc $\det(\text{Id}_{M_{n \times n}(\mathbb{R})}) = 1$.

Cependant, par hypothèse, b est inversible. Dès lors, $\text{Id}_{M_{n \times n}(\mathbb{R})} = b^{-1}b = bb^{-1}$. Et par la proposition 8.27, on sait que $\det(bb^{-1}) = \det(b) \det(b^{-1})$. On a alors :

$$1 = \det(\text{Id}_{M_{n \times n}(\mathbb{R})}) = \det(b^{-1}b) = \det(b^{-1}) \det(b).$$

D'où on déduit $\det(b^{-1}) = \det(b)^{-1}$. □

Remarque. En ayant supposé b inversible, on a obtenu $\det(b) \neq 0$. Dès lors, par contraposée, on a également $\det(b) = 0 \Rightarrow b$ pas inversible.

Corollaire 8.29. Soient $a, b \in M_{n \times n}(\mathbb{R})$ deux matrices carrées inversibles. $\det(bab^{-1}) = \det(a)$.

Démonstration. On sait que $\det(bab^{-1}) = \det(b) \det(a) \det(b^{-1})$, et donc on a :

$$\det(bab^{-1}) = \det(b) \det(a) \det(b)^{-1} = \det(a).$$

□

Remarque. Ce corollaire veut dire que si V est un espace vectoriel de dimension n , et $A \in \text{Hom}(V, V)$, alors le déterminant de la matrice associée à A ne dépend pas de la base choisie.

Soient E et F , deux bases de V . Si $m_{F,E}$ est la matrice de changement de base de E vers F , alors :

$$\det(m_{E,E}(A)) = \det(m_{E,F}(A) m_{F,F}(A) m_{F,E}(A)) = \det(m_{F,F}(A)).$$

Définition 8.30. On peut définir le **déterminant de l'opérateur linéaire** $A \in \text{Hom}(V, V)$ par :

$$\det(A) := \det(m_{E,E}(A))$$

où E est une base quelconque de V . La valeur est bien définie car le déterminant ne dépend pas de la base.

Théorème 8.31. Soient V un espace vectoriel réel de dimension n et $A \in \text{Hom}(V, V)$ un opérateur linéaire. Alors $\det(A) \neq 0$ si et seulement si A est inversible.

Démonstration. En supposant A inversible, supposons également, par l'absurde, $\det(A) = 0$. On a alors :

$$1 = \det(\text{Id}_{\text{Hom}(V,V)}) = \det(A \circ A^{-1}) = \det(A) \det(A^{-1}) = 0,$$

ce qui est une contradiction. L'hypothèse est donc fausse et $\det(A) \neq 0$.

Supposons maintenant A non inversible et montrons que $\det(A) = 0$. Soit E une base de V . Si A n'est pas inversible, alors il existe i tel que $A(e_i) = 0$ où $e_i \neq 0$ car c'est un vecteur de la base. Soit f , la seule forme multilinéaire alternée telle que $f(e_1, \dots, e_n) = 1$. On a donc :

$$\det(A) = \det(A)f(e_1, \dots, e_i, \dots, e_n) = f(A(e_1), \dots, A(e_i), \dots, A(e_n)) = f(A(e_1), \dots, 0, \dots, A(e_n)) = 0,$$

par la proposition 8.22. □

Remarque. Soient $A \in \text{Hom}(\mathbb{R}^n, \mathbb{R}^n)$ et $a := m_{E,E}(A) \in M_{n \times n}(\mathbb{R})$ sa matrice associée dans la base canonique de \mathbb{R}^n . L'application A envoie le cube unitaire de dimension n sur le parallélogramme de dimension n dont les coordonnées des points $A(e_i)$ sont les colonnes de a . Le volume orienté de l'image de $P \subset \mathbb{R}^n$ par A est donné par :

$$V^{\text{or}}(A(P)) = \det(A) V^{\text{or}}(P).$$

Dès lors, si P est le cube unitaire de dimension n et Q est un parallélogramme de dimension n , alors en construisant la matrice $a \in M_{n \times n}(\mathbb{R})$ en y mettant colonne par colonne les coordonnées des points de Q , alors $V^{\text{or}}(Q) = \det(a)$.

Si le déterminant est négatif, c'est que le parallélogramme a été *retourné* et donc son volume orienté change de signe. De là, on voit bien que $\det(A) = 0$ si $\text{rang}(A) := \dim_{\mathbb{R}} \mathcal{J}(A) < n$ car $A(P)$ n'est plus un parallélogramme à n dimensions, et donc son volume en dimension n est nul également.

8.5 Cofacteurs, matrices adjointes et inverses

Définition 8.32. Soit $a := [a_{ij}] \in M_{n \times n}(\mathbb{R})$ une matrice. Le cofacteur A_i^j est défini par :

$$A_i^j := (-1)^{i+j} \det_{(n-1) \times (n-1)} \begin{bmatrix} a_{11} & \dots & a_{1j-1} & a_{1j+1} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-11} & \dots & a_{i-1j-1} & a_{i-1j+1} & \dots & a_{i-1n} \\ a_{i+11} & \dots & a_{i+1j-1} & a_{i+1j+1} & \dots & a_{i+1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nj-1} & a_{nj+1} & \dots & a_{nn} \end{bmatrix}$$

Théorème 8.33 (Théorème de Laplace). Soit $a \in M_{n \times n}(\mathbb{R})$. Pour $1 \leq i, j \leq n$, on a :

$$\det(a) = \sum_{k=1}^n a_{ik} A_i^k = \sum_{k=1}^n a_{kj} A_k^j.$$

Démonstration. Prouvons d'abord la première égalité pour $i = n$:

$$\begin{aligned} \det(a) &= \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \prod_{k=1}^n a_{k\sigma(k)} = \sum_{j=1}^n \sum_{\sigma \in \mathfrak{S}_n, \sigma(n)=j} \text{sign}(\sigma) \prod_{k=1}^n a_{k\sigma(k)} \\ &= \sum_{j=1}^n a_{nj} \sum_{\sigma \in \mathfrak{S}_n, \sigma(n)=j} \text{sign}(\sigma) \prod_{k=1}^{n-1} a_{k\sigma(k)} \\ &= \sum_{j=1}^n a_{nj} \det \begin{bmatrix} a_{11} & \dots & a_{1j-1} & a_{1j} & a_{1j+1} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n-11} & \dots & a_{n-1j-1} & a_{n-1j} & a_{n-1j+1} & \dots & a_{n-1n} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \\ &= \sum_{j=1}^n a_{nj} (-1)^{j-1} \det \begin{bmatrix} a_{1j} & a_{11} & \dots & a_{1j-1} & a_{1j+1} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n-1j} & a_{n-11} & \dots & a_{n-1j-1} & a_{n-1j+1} & \dots & a_{n-1n} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ &= \sum_{j=1}^n a_{nj} (-1)^{j-1} (-1)^{n-1} \det \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{1j} & a_{11} & \dots & a_{1j-1} & a_{1j+1} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n-1j} & a_{n-11} & \dots & a_{n-1j-1} & a_{n-1j+1} & \dots & a_{n-1n} \end{bmatrix} \\ &= \sum_{j=1}^n a_{nj} (-1)^{j+n} \sum_{\gamma=1}^n b_{1\gamma} \sum_{\sigma \in \mathfrak{S}_n, \sigma(1)=\gamma} \text{sign}(\sigma) \prod_{k=2}^n b_{k\sigma(k)} \\ &= \sum_{j=1}^n a_{nj} (-1)^{j+n} \sum_{\sigma \in \mathfrak{S}_n, \sigma(1)=1} \text{sign}(\sigma) \prod_{k=2}^n b_{k\sigma(k)}. \end{aligned}$$

Ici, on a posé :

$$b := [b_{ij}] := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{1j} & a_{11} & \dots & a_{1j-1} & a_{1j+1} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n-1j} & a_{n-11} & \dots & a_{n-1j-1} & a_{n-1j+1} & \dots & a_{n-1n} \end{bmatrix} \in M_{n \times n}(\mathbb{R}).$$

De plus, en imposant $\sigma(1) = 1$, pour tout $n \geq i > 1$, on a $\sigma(i) > 1$ car σ est une bijection. Donc avec cette contrainte, et en itérant sur $k \in \{2, 3, \dots, n\}$, les éléments de la première ligne ou de la première colonne ne sont pas accédés. En posant :

$$c := [c_{ij}] := \begin{bmatrix} a_{11} & \dots & a_{1j-1} & a_{1j+1} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n-11} & \dots & a_{n-1j-1} & a_{n-1j+1} & \dots & a_{n-1n} \end{bmatrix} \in M_{(n-1) \times (n-1)}(\mathbb{R}),$$

à savoir c est la matrice b sans la première ligne et la première colonne, on a alors :

$$\begin{aligned} \det(a) &= \sum_{j=1}^n a_{nj} (-1)^{j+n} \sum_{\sigma \in \mathfrak{S}_n, \sigma(1)=1} \text{sign}(\sigma) \prod_{k=2}^n b_{k\sigma(k)} \\ &= \sum_{j=1}^n a_{nj} (-1)^{j+n} \sum_{\sigma \in \mathfrak{S}_n, \sigma(1)=1} \text{sign}(\sigma) \prod_{k=2}^n c_{k-1\sigma(k-1)} \\ &= \sum_{j=1}^n a_{nj} (-1)^{j+n} \sum_{\sigma \in \mathfrak{S}_{n-1}} \text{sign}(\sigma) \prod_{k=1}^{n-1} c_{k\sigma(k)} \\ &= \sum_{j=1}^n a_{nj} \mathcal{A}_n^j. \end{aligned}$$

Afin de prouver le théorème pour tout i , il suffit de passer la proposition 8.21 en échangeant les colonnes, ce qui change le signe (à cause de l'échange de deux colonnes), mais qui change également les signes des cofacteurs. Afin de prouver également le théorème, on pose $\gamma \in \{1, \dots, n\}$, et on part de la proposition 8.25 qui dit que $\det(a) = \det(a^T)$ et on obtient :

$$\det(a) = \det(A^T) = \sum_{j=1}^n (a^T)_{\gamma j} (A^T)_{\gamma}^j = \sum_{j=1}^n a_{j\gamma} \mathcal{A}_j^{\gamma}.$$

□

Définition 8.34. Soit $a := [a_{ij}] \in M_{n \times n}(\mathbb{R})$ une matrice carrée. On définit sa matrice adjacente par :

$$\text{adj } a = [A_j^i] \in M_{n \times n}(\mathbb{R}).$$

Proposition 8.35. Soit $a = [a_{ij}] \in M_{n \times n}(\mathbb{R})$. Le produit matriciel $a \text{adj}(a)$ donne une matrice diagonale telle que $(a \text{adj}(a))_{ii} = \det(a)$ pour tout i .

Démonstration. En procédant au produit de a et $\text{adj}(a)$, on obtient :

$$(a \text{adj}(a))_{ij} := \sum_{k=1}^n a_{ik} (\text{adj}(a))_{kj} = \sum_{k=1}^n a_{ik} \mathcal{A}_j^k = \det(a) \delta_{ij}.$$

Le facteur δ_{ij} impose que tout élément $(a \operatorname{adj}(a))_{ij}$ où $i \neq j$ soit nul. Il ne reste dès lors que la diagonale où chaque $(a \operatorname{adj}(a))_{ii}$ vaut $\det(a)$. \square

Corollaire 8.36. Soit $a \in M_{n \times n}(\mathbb{R})$ une matrice carrée inversible. Son inverse a^{-1} vaut :

$$a^{-1} = \frac{1}{\det(a)} \operatorname{adj}(a)$$

Démonstration. On sait que $a \operatorname{adj}(a) = \det(a) \operatorname{Id}_{M_{n \times n}(\mathbb{R})}$. Dès lors, en divisant de part et d'autre par $\det(a)$ (qui est non nul car a est inversible), et en multipliant par a^{-1} à gauche, on obtient :

$$\begin{aligned} a^{-1} a \operatorname{adj}(a) \det(a)^{-1} &= a^{-1} \det(a) \operatorname{Id}_{M_{n \times n}(\mathbb{R})} \det(a)^{-1} \\ \operatorname{adj}(a) \frac{1}{\det(a)} &= a^{-1}. \end{aligned}$$

\square

Définition 8.37. Soient $a = [a_{ij}] \in M_{n \times n}(\mathbb{R})$ et $b = [b_i], x = [x_i] \in M_{n \times 1}(\mathbb{R})$ tels que :

$$ax = b.$$

On pose :

$$S_j(a, b) := \begin{bmatrix} a_{11} & \dots & a_{1j-1} & b_1 & a_{1j+1} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nj-1} & b_n & a_{nj+1} & \dots & a_{nn} \end{bmatrix}$$

La valeur $S_j(a, b)$ est donc la matrice a où la j ème colonne a été remplacée par b .

Théorème 8.38 (Règle de Cramer). Soit $(S) : ax = b$ un système linéaire d'équations où $a \in M_{n \times n}(\mathbb{R}), x, b \in M_{n \times 1}(\mathbb{R})$. Le système (S) admet une et une seule solution si et seulement si $\det(a) \neq 0$. De plus, cette solution est donnée par :

$$x_j = \frac{\det(S_j(a, b))}{\det(a)}.$$

Démonstration. Soient $a \in M_{n \times n}(\mathbb{R}), b, x \in M_{n \times 1}(\mathbb{R})$ tels que $ax = b$. On peut déterminer :

$$\begin{aligned} \det(S_j(a, b)) &= \det \begin{bmatrix} a_{11} & \dots & a_{1j-1} & b_1 & a_{1j+1} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nj-1} & b_n & a_{nj+1} & \dots & a_{nn} \end{bmatrix} \\ &= \det \begin{bmatrix} a_{11} & \dots & a_{1j-1} & \sum_{k=1}^n a_{1k} x_k & a_{1j+1} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nj-1} & \sum_{k=1}^n a_{nk} x_k & a_{nj+1} & \dots & a_{nn} \end{bmatrix} \\ &= \sum_{k=1}^n \det \begin{bmatrix} a_{11} & \dots & a_{1j-1} & a_{1k} x_k & a_{1j+1} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nj-1} & a_{nk} x_k & a_{nj+1} & \dots & a_{nn} \end{bmatrix} \\ &= \sum_{k=1}^n x_k \det \begin{bmatrix} a_{11} & \dots & a_{1j-1} & a_{1k} & a_{1j+1} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nj-1} & a_{nk} & a_{nj+1} & \dots & a_{nn} \end{bmatrix} \end{aligned}$$

Ici, pour toute valeur de k différente de j , la matrice aura deux colonnes identiques (à savoir la colonne j et la colonne k). Dès lors, le déterminant s'annule, et il ne reste plus que :

$$\det(S_j(a, b)) = x_j \det(a).$$

On en déduit que :

$$x_j = \frac{\det(S_j(a, b))}{\det(a)}.$$

□

9 Algèbre d'ensembles et résultats sur les espaces vectoriels

9.1 Définitions

Définition 9.1. Soit \mathbb{K} un ensemble et soit $*$ une opération interne telle que $*$: $\mathbb{K} \times \mathbb{K} : (a, b) \mapsto a * b =: ab$. Si \mathbb{K} respecte les trois propriétés suivantes :

1. $\exists 0 \in \mathbb{K}$ tel que $\forall a \in \mathbb{K} : a * 0 = 0 * a = a$;
2. $\forall a, b, c \in \mathbb{K} : a * (b * c) = (a * b) * c$;
3. $\forall a \in \mathbb{K} : \exists b \in \mathbb{K}$ t.q. $a + b = b + a = 0$ (on note $b = -a$ et on l'appelle **l'inverse de a**),

on appelle $(\mathbb{K}, *)$ un **groupe**.

Définition 9.2. Soit $(\mathbb{K}, *)$ un groupe. Si $*$ est une opération commutative ($\forall a, b \in \mathbb{K} : a * b = b * a$), on appelle $(\mathbb{K}, *)$ un **groupe abélien**.

Remarque. L'opération $*$ d'un groupe abélien est souvent notée $+$ pour montrer son caractère commutatif.

Définition 9.3. Soit $(\mathbb{K}, +)$ un groupe abélien. On définit une opération \cdot : $\mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K} : (a, b) \mapsto a \cdot b =: ab$. Si $(\mathbb{K}, +, \cdot)$ respecte les propriétés suivantes :

1. $\exists 1 \in \mathbb{K}$ t.q. $\forall a \in \mathbb{K} : a \cdot 1 = 1 \cdot a = a$;
2. $\forall a, b, c \in \mathbb{K} : a \cdot (b \cdot c) = (a \cdot b) \cdot c$;
3. $\forall a, b, c \in \mathbb{K} : (a \cdot (b + c) = a \cdot b + a \cdot c) \wedge ((a + b) \cdot c = a \cdot c + b \cdot c)$,

on appelle $(\mathbb{K}, +, \cdot)$ un **anneau**.

Définition 9.4. Soit $(\mathbb{K}, +, \cdot)$ un anneau. Si $\forall a \neq 0 \in \mathbb{K} : \exists b \in \mathbb{K}$ t.q. $a \cdot b = b \cdot a = 1$ (où b est noté a^{-1}), alors on appelle $(\mathbb{K}, +, \cdot)$ un **corps**.

Définition 9.5. Soit $(\mathbb{K}, +, \cdot)$ un corps. Si l'opération \cdot est commutative, alors on appelle $(\mathbb{K}, +, \cdot)$ un **corps commutatif**.

9.2 Sous-espaces propres

Définition 9.6. Soient \mathbb{K} un corps commutatif, V un espace vectoriel sur \mathbb{K} , et $A \in \text{Hom}(V, V)$ un opérateur linéaire.

- On appelle **vecteur propre** de A tout vecteur $v \in V$ t.q. $\exists \lambda \in \mathbb{K}$ t.q. $A(v) = \lambda v$;
- on appelle **valeur propre** de A toute valeur $\lambda \in \mathbb{K}$ t.q. $\exists v \neq 0 \in V$ et v est une valeur propre de A ;
- l'ensemble des valeurs propres de A est appelé le **spectre** de A .

Définition 9.7. Soit V un espace vectoriel sur un corps \mathbb{K} . Soit $A \in \text{Hom}(V, V)$ un opérateur linéaire. Soit $\lambda \in \mathbb{K}$ une valeur propre de A . On appelle **sous-espace propre** relatif à λ l'ensemble :

$$V_\lambda := \{v \in V \text{ t.q. } A(v) = \lambda v\}.$$

Proposition 9.8. Soit V un espace vectoriel sur un corps \mathbb{K} et V_λ un sous-espace propre relatif à λ , une valeur propre de $A \in \text{Hom}(V, V)$. Alors V_λ muni des opérateurs $+_V$ et \cdot_V est un sous-espace vectoriel de V .

Démonstration. Pour montrer que V_λ est un sous-espace de V , il faut montrer que :

1. $0_V \in V_\lambda$;
2. $\forall w_1, w_2 \in V_\lambda : w_1 + w_2 \in V_\lambda$;
3. $\forall \mu \in \mathbb{K}, w \in V_\lambda : \mu w \in V_\lambda$.

Montrons donc chacune de ces trois propriétés :

1. On sait que $A(0) = 0 = \lambda 0$ pour tout $\lambda \in \mathbb{K}$. Dès lors $0 \in V_\lambda$.
2. Soient $w_1, w_2 \in V_\lambda$. Alors $A(w_1 + w_2) = A(w_1) + A(w_2) = \lambda v_1 + \lambda v_2 = \lambda(v_1 + v_2)$. Dès lors $w_1 + w_2 \in V_\lambda$.
3. Soient $\mu \in \mathbb{K}, w \in V_\lambda$. Alors $A(\mu w) = \mu A(w) = \mu(\lambda v) = \lambda(\mu v)$. Dès lors $\mu w \in V_\lambda$.

On a donc bien V_λ est un sous-espace de V par la proposition 5.3. □

Lemme 9.9. Soit \mathbb{K} un corps commutatif et V un espace vectoriel sur \mathbb{K} . Si $A \in \text{Hom}(V, V)$, alors $\text{Ker}(A) = V_0$.

Démonstration. On sait que $\text{Ker}(A) := \{v \in V \text{ t.q. } A(v) = 0\}$. Également, $V_0 := \{v \in V \text{ t.q. } A(v) = 0 \cdot v = 0\} =: \text{Ker}(A)$. □

Proposition 9.10. Soient $a := [a_{ij}] \in M_{n \times n}(\mathbb{K})$ une matrice diagonale et $A \in \text{Hom}(\mathbb{K}^n, \mathbb{K}^n)$ son opérateur linéaire associé. Soit $E := \{e_1, \dots, e_n\}$ la base canonique de \mathbb{K}^n . Alors :

$$V_\lambda = \langle \{e_i \in E \text{ t.q. } a_{ii} = \lambda\} \rangle.$$

Démonstration. Montrons que $V_\lambda \subseteq \langle \{e_i \in E \text{ t.q. } a_{ii} = \lambda\} \rangle \subseteq V_\lambda$.

Soit $v \in V_\lambda$. On sait que $A(v) = \lambda v$. Or, on sait également que $A(v) = av$. On a donc :

$$A(v) = \left[\sum_{k=1}^n a_{ik} v_k \right]_i = \sum_{i=1}^n e_i \sum_{k=1}^n a_{ik} v_k.$$

Cependant, a est une matrice diagonale. Donc $\forall i \neq k : a_{ik} = 0$. On trouve donc :

$$A(v) = \sum_{i=1}^n e_i a_{ii} v_i.$$

En reprenant le fait que $A(v) = \lambda v$, on obtient :

$$\sum_{i=1}^n \lambda v_i e_i = \lambda v = A(v) = \sum_{i=1}^n a_{ii} v_i e_i.$$

Que l'on peut transformer en :

$$0 = \sum_{i=1}^n a_{ii} v_i e_i - \lambda v_i e_i = \sum_{i=1}^n (\lambda - a_{ii}) v_i e_i.$$

Or, e_i est un vecteur de la base. On sait donc que $\forall 1 \leq i \leq n : e_i \neq 0$. Ce qui laisse donc soit $v_i = 0$, soit

$a_{ii} = \lambda$. On a alors :

$$\begin{aligned} A(v) &= \sum_{i=1}^n a_{ii} v_i e_i = \sum_{i \text{ t.q. } a_{ii} = \lambda} a_{ii} v_i e_i + \sum_{i \text{ t.q. } a_{ii} \neq \lambda} a_{ii} v_i e_i \\ &= \sum_{i \text{ t.q. } a_{ii} = \lambda} a_{ii} v_i e_i + \sum_{i \text{ t.q. } a_{ii} \neq \lambda} a_{ii} \cdot 0 \cdot e_i \\ &= \sum_{i \text{ t.q. } a_{ii} = \lambda} a_{ii} v_i e_i \in \langle \{e_i \in E \text{ t.q. } a_{ii} = \lambda\} \rangle. \end{aligned}$$

Soit $v \in \langle \{e_i \in E \text{ t.q. } a_{ii} = \lambda\} \rangle$. On sait donc que :

$$v = \sum_{i \text{ t.q. } a_{ii} = \lambda} \mu_i e_i.$$

De plus, en évaluant l'image de v par A , on obtient :

$$\begin{aligned} A(v) &= A\left(\sum_{i \text{ t.q. } a_{ii} = \lambda} \mu_i e_i\right) = \sum_{i \text{ t.q. } a_{ii} = \lambda} \mu_i A(e_i) = \sum_{i \text{ t.q. } a_{ii} = \lambda} \mu_i a e_i \\ &= \sum_{i \text{ t.q. } a_{ii} = \lambda} \mu_i \sum_{j=1}^n e_j \sum_{k=1}^n a_{jk} (e_i)_k. \end{aligned}$$

Or, e_i est un vecteur de la base canonique. Ce qui veut dire que $(e_i)_k = \delta_{ik}$. De plus, a est une matrice diagonale, et donc pour tout i, j , on a : $a_{ij} = a_{ii} \delta_{ij}$. Dès lors :

$$\begin{aligned} A(v) &= \sum_{i \text{ t.q. } a_{ii} = \lambda} \mu_i \sum_{j=1}^n e_j \sum_{k=1}^n (a_{jk} \delta_{jk}) ((e_i)_k \delta_{ik}) \\ &= \sum_{i \text{ t.q. } a_{ii} = \lambda} \mu_i \sum_{j=1}^n e_j a_{jj} (e_i)_j \delta_{ij} = \sum_{i \text{ t.q. } a_{ii} = \lambda} \mu_i e_i a_{ii} \\ &= \sum_{i \text{ t.q. } a_{ii} = \lambda} \mu_i e_i \lambda = \lambda \sum_{i \text{ t.q. } a_{ii} = \lambda} \mu_i e_i = \lambda v \in V_\lambda. \end{aligned}$$

□

Lemme 9.11. Soient V un espace vectoriel sur un corps commutatif \mathbb{K} , $A \in \text{Hom}(V, V)$ un opérateur linéaire, $\lambda_1, \dots, \lambda_n$ des valeurs propres de A deux à deux distinctes et $(v_1, \dots, v_n) \in V_{\lambda_1} \times V_{\lambda_2} \times \dots \times V_{\lambda_n}$ tels que $\forall i \leq n : v_i \neq 0$. Alors $\{v_1, \dots, v_n\}$ est une partie libre.

Démonstration. Montrons cela par récurrence sur n .

Cas de base : $n = 1$. Soit $W := \{v_1\} \in V_{\lambda_1}$. W est libre car $v_1 \neq 0$ donc la seule combinaison linéaire nulle possible est la combinaison linéaire nulle triviale.

Pas de récurrence : supposons que $\{v_1, \dots, v_n\}$ est une partie libre et montrons que $\{v_1, \dots, v_{n+1}\}$ est une partie libre, où $v_{n+1} \in V_{\lambda_{n+1}}$. Soient $\mu_1, \dots, \mu_{n+1} \in \mathbb{K}$ tels que :

$$\omega := \sum_{i=1}^{n+1} \mu_i v_i = 0.$$

On sait :

$$0 = A(0) = A \left(\sum_{i=1}^{n+1} \mu_i v_i \right) = \sum_{i=1}^{n+1} \mu_i A(v_i) = \sum_{i=1}^{n+1} \mu_i \lambda_i v_i.$$

De plus, on peut ajouter ω sans changer la valeur car $\omega = 0$. Dès lors :

$$0 = \sum_{i=1}^{n+1} \mu_i \lambda_i v_i - \lambda_{n+1} \omega = \sum_{i=1}^{n+1} \mu_i v_i (\lambda_i - \lambda_{n+1}).$$

Pour $1 \leq i \leq n$, $(\lambda_i - \lambda_{n+1}) \neq 0$. De plus, par hypothèse de récurrence, $\{v_1, \dots, v_n\}$ forme une partie libre. La seule combinaison linéaire nulle est formée par $\mu_i = 0 \forall 1 \leq i \leq n$. On a donc :

$$0 = \sum_{i=1}^n \mu_i v_i + \mu_{n+1} v_{n+1} = \mu_{n+1} v_{n+1}.$$

Or, par hypothèse, on sait $v_{n+1} \neq 0$. On a donc $\mu_{n+1} = 0$. Dès lors, $\{v_1, \dots, v_{n+1}\}$ est une partie libre. \square

Proposition 9.12. Soit V un espace vectoriel sur un corps commutatif \mathbb{K} . Soient $A \in \text{Hom}(V, V)$ un opérateur linéaire et $\lambda_1, \dots, \lambda_n$ des valeurs propres de A deux à deux distinctes. Pour $1 \leq i \leq n$, on définit $B_i \subseteq V_{\lambda_i}$ des parties libres. Alors, l'ensemble :

$$W := \bigcup_{i=1}^n B_i$$

est également une partie libre.

Démonstration. Montrons à nouveau cela par récurrence sur n .

Cas de base : $n = 1$. Soit $W := B_1 \subseteq V_{\lambda_1}$. W est une partie libre car B_1 est libre par hypothèse.

Pas de récurrence : supposons que $B_1 \cup \dots \cup B_n$ soit une partie libre et montrons que $B_1 \cup \dots \cup B_{n+1}$ est également une partie libre. Pour tout $1 \leq i \leq n+1$, on définit $B_i := \{x_1^{(i)}, \dots, x_{n_i}^{(i)}\} \subseteq V_{\lambda_i}$.

Prenons ω défini par :

$$\omega := \sum_{i=1}^{n+1} \sum_{j=1}^{n_i} \mu_j^{(i)} x_j^{(i)} = 0.$$

On veut ensuite montrer que $\forall 1 \leq i \leq n : \forall 1 \leq j \leq n_i : \mu_j^{(i)} = 0$. On définit également :

$$\forall 1 \leq i \leq n : v_i := \sum_{j=1}^{n_i} \mu_j^{(i)} x_j^{(i)}.$$

On sait que $v_i \in V_{\lambda_i}$. Dès lors, on sait :

$$\sum_{i=1}^n v_i = 0.$$

Par le lemme 9.11, on sait que :

$$0 = v_{n+1} = \sum_{j=1}^{n_{n+1}} \mu_j^{(n+1)} x_j^{(n+1)}.$$

Cependant, on sait que B_{n+1} est une partie libre. Dès lors, la seule combinaison linéaire nulle est $\mu_j^{(n+1)} = 0$ pour tout j . Il reste :

$$0 = \sum_{i=1}^n v_i.$$

Par hypothèse de récurrence, on sait que $B_1 \cup \dots \cup B_n$ est une partie libre. Dès lors, la seule combinaison linéaire nulle possible pour $0 = \sum_i v_i$ est $\mu_j^{(i)} = 0$ pour tout $1 \leq i \leq n$ et $1 \leq j \leq n_i$.

La seule combinaison linéaire possible au final pour :

$$0 = \sum_{i=1}^{n+1} \sum_{j=1}^{n_i} \mu_j^{(i)} x_j^{(i)}$$

est la combinaison linéaire nulle triviale. L'ensemble $W := B_1 \cup \dots \cup B_{n+1}$ est donc bien une partie libre. \square

9.3 Équation caractéristique

Définition 9.13. Soit A un opérateur linéaire sur V , un \mathbb{K} -espace vectoriel où \mathbb{K} est un corps commutatif. Alors :

$$\pi_A(\lambda) := \det(A - \lambda \text{Id}_{\text{Hom}(V, V)})$$

est appelé le **polynôme caractéristique** de A .

Lemme 9.14. Soit A un opérateur linéaire sur un espace vectoriel V sur un corps commutatif \mathbb{K} . Alors $\lambda \in \mathbb{K}$ est une valeur propre de A si et seulement si $V_\lambda \neq \{0\}$.

Démonstration. Supposons que $V_\lambda = \{0\}$. Alors seul 0 est tel que $A(0) = \lambda 0 = 0$. Dès lors, λ n'est pas une valeur propre.

Supposons maintenant que λ est une valeur propre de A . Alors, il existe $v \in V$ tel que $A(v) = \lambda v$. Donc $v \in V_\lambda$. On a alors bien $V_\lambda \neq \{0\}$. \square

Théorème 9.15. Soit V un espace vectoriel de dimension finie n sur un corps commutatif \mathbb{K} et soit $A \in \text{Hom}(V, V)$. Alors :

1. $\lambda \in \mathbb{K}$ est une valeur propre de A si et seulement si $\det(A - \lambda \text{Id}_{\text{Hom}(V, V)}) = 0$;
2. $V_\lambda = \text{Ker}(A - \lambda \text{Id}_{\text{Hom}(V, V)})$;
3. $\pi_A(\lambda)$ est un polynôme de degré n en λ ;
4. l'opérateur A a au plus n valeurs propres distinctes qui sont les racines de l'équation caractéristique.

Démonstration. Soient \mathbb{K} un corps commutatif, V un espace vectoriel sur \mathbb{K} , et $A \in \text{Hom}(V, V)$, un opérateur linéaire sur V .

1. Par définition, $\lambda \in \mathbb{K}$ est une valeur propre de A si et seulement si il existe $v \neq 0 \in V$ t.q. $A(v) = \lambda v$, donc $A(v) - \lambda v = 0$, ou encore $(A - \lambda \text{Id}_{\text{Hom}(V, V)})(v) = 0$.
2. Par le point 1, on peut réécrire $\text{Ker}(A - \lambda \text{Id}_{\text{Hom}(V, V)}) \neq \{0\} \iff \det(A - \lambda \text{Id}_{\text{Hom}(V, V)}) = 0$, ce qui a déjà été montré par le théorème 8.31.
3. Montrons cela par récurrence sur n .

Cas de base : $n = 1$. On sait que $\det\left(\begin{bmatrix} a_{11} \end{bmatrix} - \lambda \begin{bmatrix} 1 \end{bmatrix}\right) = a_{11} - \lambda$, qui est effectivement un polynôme de degré 1 en λ .

Pas de récurrence : supposons que pour V de dimension n , $\det(a - \lambda \text{Id}_{\text{Hom}(V, V)})$ est un polynôme de degré n en λ , et montrons que pour V de dimension $n + 1$, le déterminant est un polynôme de degré $n + 1$.

Soient $a \in M_{(n+1) \times (n+1)}(\mathbb{K})$ et $\gamma \in \mathbb{N}$ tel que $1 \leq \gamma \leq (n+1)$. Alors, par le théorème de Laplace :

$$\det(a - \lambda \text{Id}_{M_{(n+1) \times (n+1)}(\mathbb{K})}) = \sum_{i=1}^{n+1} (a_{i\gamma} - \lambda \delta_{i\gamma}) A_i^\gamma = \sum_{i=1}^{n+1} a_{i\gamma} A_i^\gamma - \lambda A_\gamma^\gamma.$$

Or, A_i^γ est, par hypothèse de récurrence, un polynôme de degré n . On a donc la somme de n polynômes de degré n et un polynôme de degré $n+1$. Le degré étant la plus grande puissance à laquelle est montée une inconnue, le degré final du déterminant est $n+1$.

4. La démonstration du point 4 dépasse l'objectif de ce cours. La démonstration est donc omise.

□

Lemme 9.16. Soient \mathbb{K} un corps commutatif, V un espace vectoriel de dimension finie n sur \mathbb{K} et $A \in \text{Hom}(V, V)$. Alors $(-1)^n \pi_A(\lambda) \in \mathbb{K}[\lambda]$ est un polynôme normalisé. À savoir, si :

$$\pi_A(\lambda) = \sum_{i=0}^n \lambda^i \pi_i,$$

alors $\pi_n = (-1)^n$.

Démonstration. Soit $a := m_{E,E}(A)$, la matrice de A dans une base E de V . Alors, $\det(a - \lambda \text{Id}_{M_{n \times n}(\mathbb{K})})$ ne peut avoir comme seul monôme de degré n celui compris dans, en prenant $\sigma = \text{Id}_{\mathfrak{S}_n}$:

$$\text{sign}(\sigma) \prod_{i=1}^n (a_{i\sigma(i)} - \lambda) = \prod_{i=1}^n (a_{ii} - \lambda),$$

à savoir :

$$\prod_{i=1}^n (-\lambda) = (-1)^n \lambda^n.$$

□

Théorème 9.17 (Théorème de Cayley-Hamilton). Soit V un espace vectoriel de dimension finie n sur un corps commutatif \mathbb{K} . Soit $A \in \text{Hom}(V, V)$. Alors $\pi_A(A) = 0$.

Démonstration. Soit $\pi_A(\lambda) = \det(a - \lambda \text{Id}_{M_{n \times n}(\mathbb{K})})$, le polynôme caractéristique de l'opérateur linéaire A . On sait que π_A est un polynôme de degré n en λ . Dès lors :

$$\pi_A = \sum_{i=0}^n \pi_i \lambda^i,$$

où $\pi_i \in \mathbb{K}$ pour $i \in \{1, \dots, n\}$. Prenons E une base de V , et $a := m_{E,E}(A) \in M_{n \times n}(\mathbb{K})$ la matrice associée à A .

Posons $b := \text{adj}(a - \lambda \text{Id}_{M_{n \times n}(\mathbb{K})})$. Soient $b_0, \dots, b_{n-1} \in M_{n \times n}(\mathbb{K})$ tels que :

$$b = \sum_{i=0}^{n-1} \lambda^i b_i.$$

Alors :

$$\begin{aligned}
\pi_A(\lambda) \text{Id}_{M_{n \times n}(\mathbb{K})} &= \det(a - \lambda \text{Id}_{M_{n \times n}(\mathbb{K})}) \text{Id}_{M_{n \times n}(\mathbb{K})} = (a - \lambda \text{Id}_{M_{n \times n}(\mathbb{K})}) \text{adj}(a - \lambda \text{Id}_{M_{n \times n}(\mathbb{K})}) \\
&= (a - \lambda \text{Id}_{M_{n \times n}(\mathbb{K})})b = (a - \lambda \text{Id}_{M_{n \times n}(\mathbb{K})}) \sum_{i=0}^{n-1} \lambda^i b_i \\
&= \sum_{i=0}^{n-1} a \lambda^i b_i - \sum_{i=0}^n \lambda \text{Id}_{M_{n \times n}(\mathbb{K})} \lambda^i b_i = \sum_{i=0}^{n-1} a \lambda^i b_i - \sum_{i=0}^{n-1} \lambda^{i+1} b_i \\
&= ab_0 + \sum_{i=1}^{n-1} \lambda^i (ab_i - b_{i-1}) - \lambda^n b_{n-1}
\end{aligned}$$

On sait du lemme 9.16 que :

$$\pi_A(\lambda) = (-1)^n \lambda^n + \sum_{i=0}^{n-1} c_i \lambda^i.$$

Dès lors, on doit avoir $b_{n-1} = (-1)^n \text{Id}_{M_{n \times n}(\mathbb{K})}$, $ab_0 = c_0$, et $ab_i - b_{i-1} = c_i \text{Id}_{M_{n \times n}(\mathbb{K})}$ pour $1 \leq i \leq n-1$.
Dès lors, on a :

$$\begin{aligned}
\pi_A(a) &= ab_0 + \sum_{i=1}^{n-1} a^i (ab_i - b_{i-1}) - a^n b_{n-1} = ab_0 + \sum_{i=1}^{n-1} a^{i+1} b_i - \sum_{i=1}^{n-1} a^i b_{i-1} \\
&= ab_0 + \sum_{i=1}^{n-2} a^{i+1} b_i + a^n b_{n-1} - \sum_{i=2}^{n-1} a^i b_{i-1} - a^1 b_0 - a^n b_{n-1} = 0.
\end{aligned}$$

□

9.3.1 Opérateurs diagonalisables

Définition 9.18. Soient V un espace vectoriel de dimension n sur un corps commutatif \mathbb{K} et $A \in \text{Hom}(V, V)$. On dit que A est diagonalisable s'il existe une base E de V telle que $m_{E,E}(A) = [\delta_{ij} \lambda_i]$. Donc A est diagonalisable si et seulement s'il existe une base E de V composée de vecteurs propres de A .

Proposition 9.19. Soient V un espace vectoriel de dimension finie n sur un corps commutatif \mathbb{K} et $A \in \text{Hom}(V, V)$. Si A possède n valeurs propres distinctes, alors A est diagonalisable.

Démonstration. Soient $\lambda_1, \dots, \lambda_n$ les valeurs propres de A telles que pour tout $i \neq j$, on a $\lambda_i \neq \lambda_j$. Soient $e_i \in V_{\lambda_i}$ pour $1 \leq i \leq n$. Par la proposition 9.12, on sait que $E := \{e_1, \dots, e_n\}$ est une base de V . On a donc :

$$\forall 1 \leq i \leq n : A(e_i) = \lambda_i e_i.$$

Dès lors :

$$m_{E,E}(A) = [\lambda_i \delta_{ij}] \in M_{n \times n}(\mathbb{K}).$$

□

Proposition 9.20. Soient V un espace vectoriel de dimension finie n sur un corps commutatif \mathbb{K} et $A \in \text{Hom}(V, V)$. Si A est diagonalisable, alors $\pi_A(\lambda)$ peut s'exprimer comme un produit de n polynômes linéaires à coefficients dans \mathbb{K} .

Démonstration. Le polynôme caractéristique est un polynôme de degré n en λ ayant pour zéros les valeurs propres de A . Dès lors :

$$P(\lambda) := (-1)^n \prod_{i=1}^n (\lambda - \lambda_i) = \prod_{i=1}^n (\lambda_i - \lambda)$$

est le polynôme caractéristique de A dans une base de vecteurs propres. \square

Remarque. Il y a deux raisons pour lesquelles un opérateur linéaire A ne serait pas diagonalisable :

1. un problème algébrique : les racines de π_A ne sont pas dans \mathbb{K} ;
2. un problème géométrique : il n'y a pas suffisamment de vecteurs propres pour former une base de V .

9.3.2 Opérateurs triangulaires

Définition 9.21. Soient V un espace vectoriel de dimension finie n sur un corps commutatif \mathbb{K} et $A \in \text{Hom}(V, V)$. On dit que A est triangulable s'il existe une base E de V telle que $m_{E,E}(A)$ est triangulaire.

Définition 9.22. Soient V un \mathbb{K} -espace vectoriel où \mathbb{K} est un corps commutatif et W_i t.q. $0 \leq i \leq n$ un ensemble de sous-espaces de V . Si les W_i sont tels que :

$$\{0\} = W_0 \subset W_1 \subset \dots \subset W_n = V,$$

et pour tout i , on a $A(W_i) \subseteq W_i$, alors on dit que W_0, \dots, W_n forment un **drapeau maximal invariant par A** .

Proposition 9.23. Soient V un espace vectoriel sur un corps commutatif \mathbb{K} et $A \in \text{Hom}(V, V)$. Alors A est triangulable si et seulement si A possède un drapeau maximal invariant.

Démonstration. Soit $\{W_i\}_{0 \leq i \leq n}$ un drapeau maximal invariant par A . Montrons que A est triangulable. Pour tout $i \in \{1, \dots, n\}$, on prend $e_i \in W_i \setminus W_{i-1}$. Posons $E := \{e_1, \dots, e_n\}$. Montrons que E est une base de V et que $m_{E,E}(A)$ est une matrice triangulaire.

Montrons que E est libre. Soit $1 \leq i \leq n$. Posons $E_i := \{e_1, \dots, e_i\}$. Supposons que E_i n'est pas libre. Alors il existe $\mu_1, \dots, \mu_{i-1} \in \mathbb{K}$ tels que $\mu_j \neq 0$ pour tout $j < i$ et :

$$e_i = \sum_{j=1}^{i-1} \mu_j e_j.$$

Puisque W_i est un espace vectoriel pour tout i , on a $e_i \in W_{i-1}$. Or par définition, $e_i \in W_i \setminus W_{i-1}$. Il y a donc contradiction et E_i est libre. Dès lors, $E = E_n$ est libre.

Montrons que E est génératrice. Pour cela, montrons par récurrence que E_i est génératrice de W_i . Le cas de base est $i = 1$ et $\{e_1\}$ est une base de W_1 .

Le pas de récurrence suppose que E_{i-1} est génératrice de W_{i-1} . Montrons que E_i est génératrice de W_i . Soit $v_i \in W_i$. On distingue deux cas : soit $v_i \in W_{i-1}$, soit $v_i \in W_i \setminus W_{i-1}$. Dans le second cas, par hypothèse de récurrence, on prend $\mu_1, \dots, \mu_{i-1} \in \mathbb{K}$ tels que $\mu_1 e_1 + \dots + \mu_{i-1} e_{i-1} = v_i$ et prenons $\mu_i = 0$.

Dans le premier cas, prenons $F = \{f_1, \dots, f_{i-1}\}$ une base de W_{i-1} , $F \cup \{v_i\}$ une base de W_i . On a alors $\mu_j = 0$ pour $j \in \{1, \dots, i-1\}$ et $\mu_i = 1$.

On a donc bien E_i est génératrice de W_i . En particulier E_n est génératrice de $W_n = V$. Donc E est une base de V .

Montrons maintenant que $m_{E,E}(A)$ est triangulaire.

Soit $\alpha := [\alpha_{ij}] = m_{E,E}(A) \in M_{n \times n}(\mathbb{K})$. Soit $w_i = \sum_j \alpha_j e_j \in W_i$. On sait que $A(w_i) \in W_i$ et donc :

$$W_i \ni \sum_{j=1}^i \alpha_j A(e_j).$$

Ce qui implique $A(e_j) \in W_i$ pour tout $j \leq i$. On peut donc écrire :

$$W_i \ni A(e_j) = \sum_{k=1}^i \alpha_{kj} e_k.$$

Or, on sait également que $A(e_j) \in V$. Donc :

$$V \ni A(e_j) = \sum_{k=1}^n \alpha_{kj} e_k.$$

On a alors l'égalité :

$$\begin{aligned} \sum_{k=1}^i \alpha_{kj} e_k &= \sum_{k=1}^n \alpha_{kj} e_k \\ \sum_{k=1}^i \alpha_{kj} e_k &= \sum_{k=1}^i \alpha_{kj} e_k + \sum_{k=i+1}^n \alpha_{kj} e_k \\ 0 &= \sum_{k=i+1}^n \alpha_{kj} e_k. \end{aligned}$$

De plus, on sait que $E = \{e_1, \dots, e_n\}$ est libre, donc $\{e_{i+1}, e_{i+2}, \dots, e_n\}$ est libre également. Alors la seule combinaison linéaire nulle possible est $\alpha_{kj} = 0$ pour tout $j \in \{i+1, i+2, \dots, n\}$. Dès lors, on sait que $m_{E,E}(A)_{ij} = 0$ pour tout $j > i$. La matrice est bien triangulaire.

Maintenant soit $E = \{e_1, \dots, e_n\}$ une base de V telle que $m_{E,E}(A) =: [\alpha_{ij}]$ est une matrice triangulaire. Montrons que $\{0\} \subset \langle e_1 \rangle \subset \langle \{e_1, e_2\} \rangle \subset \dots \subset \langle \{e_1, \dots, e_n\} \rangle = V$ est un drapeau maximal invariant par A .

En posant $W_i := \langle \{e_1, \dots, e_i\} \rangle$, on voit bien que $W_n = V$, et que $W_i \subset W_{i+1}$ pour tout $0 \leq i < n$. Et on sait que W_i est un espace vectoriel par définition de l'espace engendré $\langle X \rangle$. \square

Proposition 9.24. Soit V un espace vectoriel de dimension finie n sur \mathbb{C} . Soit $A \in \text{Hom}(V, V)$. Alors A est triangulable.

Remarque. La preuve est omise.

9.4 Formule canonique de Jordan

Théorème 9.25 (Théorème de Jordan). Soit V un espace vectoriel de dimension finie n sur \mathbb{C} . Soit $A \in \text{Hom}(V, V)$. Il existe E , une base de V telle que $\alpha := m_{E,E}(A)$ soit donnée par :

$$\alpha = \begin{bmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_k \end{bmatrix},$$

avec $k \in \mathbb{N}^*$, et où, pour $\gamma \in \{1, 2, \dots, k\}$, on a :

$$J_\gamma = \begin{bmatrix} \lambda_\gamma & & \\ & \ddots & \\ & & \lambda_\gamma \\ 1 & & & \\ & \ddots & & \\ & & 0 & \end{bmatrix} \in M_{n_\gamma \times n_\gamma}(\mathbb{C}),$$

où $\lambda_\gamma \in \mathbb{C}$ est une valeur propre de A et $n_\gamma \in \mathbb{N}^*$.

Remarque. Les blocs J_i d'une telle matrice sont appelés ses *blocs de Jordan*. Lorsque la matrice $m_{E,E}(A)$ est exprimée dans une telle base, on dit qu'elle est dans sa *forme (canonique) de Jordan*.

Remarque (bis). Le théorème de Jordan dit également que s'il existe une autre base $F \neq E$ respectant cette propriété, alors les blocs de Jordan de $m_{F,F}(A)$ sont une permutation des blocs de Jordan de $m_{E,E}(A)$.

Corollaire 9.26. Soit V , un espace vectoriel de dimension finie n sur \mathbb{C} . Soit $A \in \text{Hom}(V, V)$. La matrice A est inversible si et seulement si sa forme de Jordan ne contient que des blocs de Jordan de taille 1.

Définition 9.27. Soit V un espace vectoriel de dimension finie n sur un corps commutatif \mathbb{K} . Soit $A \in \text{Hom}(V, V)$. Une valeur propre λ de A est dite de **multiplicité algébrique** m si $(x - \lambda)^m | \pi_A(x) = \det(A - x \text{Id}_V)$ et $(x - \lambda)^{m+1} \nmid \pi_A(x)$ et est dite de **multiplicité géométrique** m si $\dim_{\mathbb{K}} V_\lambda = m$.

Proposition 9.28. Soient V un espace vectoriel de dimension n sur \mathbb{K} un corps commutatif et $A \in \text{Hom}(V, V)$. Si $\lambda \in \mathbb{K}$ est une valeur propre de A , alors la multiplicité géométrique de λ est inférieure ou égale à sa multiplicité algébrique.

Démonstration. Soit m , la dimension de V_λ , et soit E une base de V telle que les m premiers vecteurs sont dans V_λ . On sait alors que les m premières colonnes de $m_{E,E}(A)$ correspondent à λe_i pour $1 \leq i \leq m$ et qu'il existe deux matrices B, C de dimensions respectives $m \times (n - m)$ et $(n - m) \times (n - m)$. On a alors $\pi_A(x) = (x - \lambda)^m \pi_C(x)$. On sait que m est la multiplicité géométrique de λ par A . Soit β la multiplicité arithmétique de λ par C . On sait $\beta \in \mathbb{N}^*$. Donc on a bien que $(x - \lambda)^{m+\beta} | \pi_A(x)$ mais $(x - \lambda)^{m+\beta+1} \nmid \pi_A(x)$. On a bien $m \leq m + \beta$. Or $m + \beta$ représente la multiplicité arithmétique de λ par A . \square

9.5 Polynôme minimal d'un opérateur et propriétés des polynômes

Définition 9.29. Soit V un espace vectoriel de dimension n sur un corps commutatif \mathbb{K} . Soit $A \in \text{Hom}(V, V)$. Le **polynôme minimal** de A est l'unique polynôme monique de degré minimal $\mu_A(x) \in \mathbb{K}[x]$ tel que $\mu_A(A) = 0_{V^V} \in \text{Hom}(V, V)$.

Proposition 9.30. Tout endomorphisme sur un espace vectoriel admet un polynôme minimal.

Démonstration. Par le théorème de Cayley-Hamilton (théorème 9.17), on sait que $\pi_A(A) = 0 \in \text{Hom}(V, V)$. On en déduit donc l'existence de $\mu_A(x)$. Son unicité se montre en supposant $m_1(x), m_2(x) \in \mathbb{K}[x]$ deux polynômes moniques distincts de degré minimal tels que $m_1(A) = m_2(A) = 0_{V^V}$. On a alors $\deg((m_1 - m_2)(x)) < \deg(m_1(x)) = \deg(m_2(x))$. On a également :

$$m_1(A) - m_2(A) = 0_{V^V} - 0_{V^V} = 0_{V^V}.$$

Donc $(m_1 - m_2)(x)$ est un polynôme minimal de A . Or m_1 et m_2 sont supposés de degré minimal. Il y a donc contradiction et $m_1(x) = m_2(x)$. \square

Proposition 9.31. Le polynôme minimal d'un endomorphisme divise son polynôme caractéristique.

Démonstration. Soit A , un endomorphisme sur un espace vectoriel V de dimension n sur un corps commutatif \mathbb{K} . Par Euclide, on sait qu'il existe $q(x), r(x) \in \mathbb{K}[x]$ tels que $\pi_A(x) = q(x)\mu_A(x) + r(x)$. Or, on sait que :

$$r(A) = \pi_A(A) - q(A)\mu_A(A) = 0_{V^V} - q(A)0_{V^V} = 0_{V^V}.$$

On a donc le fait que le reste est nul, et donc $\mu_A(x) | \pi_A(x)$. □

Lemme 9.32. Les valeurs propres de $A \in \text{Hom}(V, V)$ pour V un espace vectoriel complexe de dimension n sont les racines de son polynôme minimal $\mu_A(x) \in \mathbb{K}[x]$.

Démonstration. Supposons par l'absurde qu'il existe λ une valeur propre de A telle que $\mu_A(\lambda) \neq 0$. On prend $v \neq 0 \in V_\lambda$. On a alors :

$$0 \neq \mu_A(\lambda) = \sum_{i=0}^m \mu_i \lambda^i,$$

et donc également :

$$\mu_A(A)(v) = \sum_{i=0}^m \mu_i A^i(v) = \sum_{i=0}^m \lambda^i v = \mu_A(\lambda)v \neq 0,$$

où $A^i(v)$ représente la composée de A avec lui-même i fois. Et comme v est un vecteur propre de A dans le sous-espace propre V_λ , on sait que $A^i(v) = A^{i-1}(\lambda v) = A^{i-2}(\lambda^2 v) = \dots = \lambda^i v$. On a également :

$$\mu_A(A) = 0_{V^V},$$

et donc :

$$\mu_A(A)(v) = 0_{V^V}(v) = 0.$$

Donc on a une contradiction $0 \neq 0$. Il n'y a donc pas de tel λ . □

Proposition 9.33. Soit $A \in \text{Hom}(V, V)$ un endomorphisme sur un espace vectoriel V complexe de dimension n . Alors le polynôme minimal $m_A(x)$ est donné par :

$$m_A(x) = \prod_{i=1}^k (x - \lambda_i)^{m_i},$$

où k est le nombre de valeurs propres de A , et pour $i \in \{1, \dots, k\}$, on a $m_i \leq n_i$, la multiplicité algébrique de la valeur propre λ_i dans μ_A telle que :

$$\sum_{i=1}^k m_i = \deg(\mu_A(x)).$$

Démonstration. On sait par le lemme 9.32 que toutes les valeurs propres de A sont des racines de μ_A . Si on note $m_i \geq 1$ la multiplicité algébrique de la valeur propre λ_i dans μ_A , on a bien :

$$\mu_A(x) = \prod_{i=1}^k (x - \lambda_i)^{m_i}.$$

Il faut également $m_i \leq n_i$ si n_i est la multiplicité de λ_i dans π_A car il faut que μ_A divise π_A . Pour cette même raison, on ne peut avoir une racine de μ_A n'étant pas une racine de π . Et par définition, les racines de π_A sont les valeurs propres. On a donc bien $m_1 + m_2 + \dots + m_k = \deg(\mu_A(x))$. □

Définition 9.34. Soit \mathbb{K} un corps commutatif. On dit que $f(x) \in \mathbb{K}[x]$ est un polynôme irréductible si f n'est pas constant et si $\exists a(x), b(x) \in \mathbb{K}[x] \Rightarrow a(x)$ est constant ou $b(x)$ est constant.

Lemme 9.35. Soient $p(x), q(x) \in \mathbb{K}[x]$ deux polynômes non nuls. Soient $k \in \mathbb{N}^*, w(x), f(x) \in \mathbb{K}$ tels que $q(x) = w(x)f(x)^k$ et $\text{GCD}(w(x), f(x)) = 1$. Alors il existe $s(x), t(x) \in \mathbb{K}[x]$ uniques tels que $\deg(s(x)) < \deg(f(x))$ et :

$$\frac{p(x)}{q(x)} = \frac{s(x)}{f(x)^k} + \frac{t(x)}{f(x)^{k-1}}.$$

Démonstration. En multipliant par $q(x)$ de part et d'autre, on obtient :

$$p(x) = s(x)w(x) + t(x)f(x). \quad (1)$$

Par Euclide, on sait qu'il existe $y(x), z(x) \in \mathbb{K}[x]$ tels que :

$$1 = \text{GCD}(w(x), f(x)) = y(x)w(x) + z(x)f(x). \quad (2)$$

En multipliant (2) par $p(x)$, on obtient :

$$p(x) = (p(x)y(x))w(x) + (p(x)z(x))f(x). \quad (3)$$

Soient $Q(x), R(x) \in \mathbb{K}[x]$ avec $\deg(r(x)) < \deg(f(x))$ tels que :

$$p(x)y(x) = Q(x)f(x) + R(x). \quad (4)$$

En réinsufflant (4) dans (3), on obtient :

$$\begin{aligned} p(x) &= (p(x)y(x))w(x) + (p(x)z(x))f(x) \\ &= (Q(x)f(x) + R(x))w(x) + (p(x)z(x))f(x) \\ &= (Q(x)w(x))f(x) + R(x)w(x) + (p(x)z(x))f(x) \\ &= R(x)w(x) + (Q(x)w(x) + p(x)z(x))f(x), \end{aligned}$$

ce qui correspond bien à (1).

La démonstration d'unicité se fait par l'absurde en supposant qu'il existe $s_1(x), s_2(x), t_1(x), t_2(x) \in \mathbb{K}[x]$ et avec le fait que $\text{GCD}(f(x), w(x)) = 1$. \square

Proposition 9.36 (Décomposition en fractions partielles). Soit \mathbb{K} , un corps commutatif. Soient $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ et $p(x) \in \mathbb{K}[x]$ tels que :

$$p(x) = \prod_{j=1}^k (x - \lambda_j)^{m_j}$$

pour $m_j \in \mathbb{N}^*$. Alors il existe $a \in M_{n \times m_i}(\mathbb{K})$ tel que :

$$\frac{1}{p(x)} = \sum_{i=1}^k \sum_{j=1}^{m_i} \frac{a_{ij}}{(x - \lambda_i)^j}.$$

Remarque. La démonstration est omise.

9.6 Sous-espaces généralisés

Proposition 9.37. Soit V un espace vectoriel de dimension n sur un corps \mathbb{K} et soit $A \in \text{Hom}(V, V)$. Soient $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ les valeurs propres de μ_A , le polynôme minimal de A et $a_i \in \mathbb{K}^{m_i}$ pour $i \in \{1, \dots, k\}$, les coefficients de sa décomposition en fractions partielles. En posant :

$$h_i(x) := \sum_{j=1}^{m_i} a_{ij} (x - \lambda_i)^{m_i-j},$$

et :

$$\psi_i(x) := \prod_{j \neq i} (x - \lambda_j)^{m_j},$$

on a :

$$\frac{1}{\mu_A(x)} = \sum_{i=1}^k \frac{h_i(x) \psi_i(x)}{\mu_A(x)}. \quad (5)$$

Démonstration. En développant le calcul, on a :

$$\begin{aligned} \frac{1}{\mu_A(x)} &= \sum_{i=1}^k \sum_{j=1}^{m_i} \frac{a_{ij}}{(x - \lambda_i)^j} \\ &= \sum_{i=1}^k \sum_{j=1}^{m_i} \frac{a_{ij} (x - \lambda_i)^{m_i}}{(x - \lambda_i)^j (x - \lambda_i)^{m_i}} \\ &= \sum_{i=1}^k \frac{1}{(x - \lambda_i)^{m_i}} \sum_{j=1}^{m_i} a_{ij} (x - \lambda_i)^{m_i - j} \\ &= \sum_{i=1}^k \frac{1}{(x - \lambda_i)^{m_i}} h_i(x) \\ &= \sum_{i=1}^k \frac{1}{(x - \lambda_i)^{m_i}} h_i(x) \frac{\prod_{j \neq i} (x - \lambda_j)^{m_j}}{\prod_{j \neq i} (x - \lambda_j)^{m_j}} \\ &= \sum_{i=1}^k \frac{h_i(x) \psi_i(x)}{\mu_A(x)}. \end{aligned}$$

□

Remarque. En multipliant (5) par $\mu_A(x)$ de part et d'autre, on obtient :

$$1 = \sum_{i=1}^k h_i(x) \psi_i(x). \quad (6)$$

Corollaire 9.38. En définissant $\mathbb{C}[x] \ni E_i := h_i(x) \psi_i(x)$, on obtient :

$$\sum_{i=1}^k E_i(A) = \text{Id}_{\mathbb{C}[x]}. \quad (7)$$

Démonstration. On remarque :

$$0 = h_i(A) 0_{\mathbb{C}[x]}(v) = h_i(A) \mu_A(A)(v) = h_i(A) \left(A - \lambda_i \text{Id}_{\mathbb{C}[x]} \right)^{m_i} \psi_i(A)(v) = \left(A - \lambda_i \text{Id}_{\mathbb{C}[x]} \right)^{m_i} E_i(A)(v).$$

Et donc, puisque $\left(A - \lambda_i \text{Id}_{\mathbb{C}[x]} \right)^{m_i}$ divise $\psi_j(A)$ pour tout $j \neq i$, on sait également que $\left(A - \lambda_i \text{Id}_{\mathbb{C}[x]} \right)^{m_i}$ divise $E_j(A) := h_j(A) \psi_j(A)$. Dès lors, pour tout $i \neq j$, on a :

$$E_i(A) E_j(A)(v) = 0. \quad (8)$$

Dès lors, on sait, par (6) :

$$E_i(A)(v) = E_i(A)(\text{Id}_{\mathbb{C}[x]}(v)) = E_i(A) \sum_{j=1}^k E_j(A)(v) = E_i(A)^2(v).$$

On a donc bien :

$$\text{Id}_{\mathbb{C}[x]}(v) = \sum_{j=1}^k E_j(A)(v).$$

□

Proposition 9.39. Soient V un \mathbb{C} -espace vectoriel de dimension finie n et $A \in \text{Hom}(V, V)$. Alors :

$$V = \bigoplus_{i=1}^k E_i(A)(V). \quad (9)$$

Démonstration. On remarque que :

$$V = \sum_{i=1}^k E_i(A)(V).$$

Prenons $v \in E_i(A)(V) \cap \langle \bigcup_{j \neq i} E_j(A)(V) \rangle$. Alors pour $j \in \{1, \dots, k\}$, il existe $w_j \in V$ tels que :

$$v = E_i(A)(w_i) = \sum_{j \neq i} E_j(A)(w_j).$$

Dès lors, pour $h \neq i$, on trouve :

$$0 = E_h(A)E_i(A)(v) = E_h(A) \sum_{j \neq i} E_j(A)(w_j) = E_h(A)^2(w_h) = E_h(A)(w_h).$$

Par conséquent, on trouve également $v = E_i(A)(w_i) = 0$.

□

Proposition 9.40. Si V est un \mathbb{C} -espace vectoriel de dimension finie n et $A \in \text{Hom}(V, V)$, alors pour tout $i \in \{1, \dots, k\}$: $E_i(A)(V) \in \text{Ker} \left((A - \lambda_i \text{Id}_{\mathbb{C}[x]})^{m_i} \right)$.

Démonstration. On prend $v \in V$. On calcule ensuite :

$$0 = \mu_A(A)(v) = h_i \mu_A(A)(v) = h_i (A - \lambda_i)^{m_i} \psi_i(A)(v) = (A - \lambda_i)^{m_i} E_i(A)(v).$$

On a donc $E_i(A)(v) \in \text{Ker}(A - \lambda_i)^{m_i}$.

Prenons maintenant $v \in \text{Ker}(A - \lambda_i)^{m_i}$. On sait que $(x - \lambda_i)^{m_i}$ divise $E_j(x)$ pour $j \neq i$. Et donc, on trouve :

$$v = \sum_{j=1}^k E_j(A)(v) = E_i(A)(v).$$

On trouve donc bien $E_i(A)(V) = \text{Ker}(A - \lambda_i \text{Id}_{\mathbb{C}[x]})^{m_i}$.

□

Remarque. On observe alors que l'écriture (9) est équivalente à :

$$V = \bigoplus_{i=1}^k \text{Ker}(A - \lambda_i \text{Id}_{\mathbb{C}[x]})^{m_i}.$$

Définition 9.41. Soit V un espace vectoriel complexe de dimension finie n . Soit $A \in \text{Hom}(V, V)$ et $\lambda_1, \dots, \lambda_k$, les valeurs propres distinctes de A . On appelle le **sous-espace propre généralisé de A associé à la valeur propre λ_i** la valeur :

$$V(\lambda_i) := \text{Ker}(A - \lambda_i \text{Id}_{\mathbb{C}[X]})^{m_i}.$$

Remarque. On remarque que $V(\lambda)$ est un sous-espace de V qui contient le sous-espace propre V_λ de A .

Proposition 9.42. Soit $A \in \text{Hom}(V, V)$ un opérateur linéaire sur un espace vectoriel complexe V de dimension finie n . Soient $\{\lambda_i\}_{1 \leq i \leq k}$ les valeurs propres distinctes de A . Alors :

$$V = \bigoplus_{i=1}^k V(\lambda_i), \quad (10)$$

et pour tout $i \in \{1, \dots, k\}$:

$$f|_{V(\lambda_i)} \in \text{Hom}(V(\lambda_i), V(\lambda_i)).$$

9.7 La forme canonique de Jordan

9.7.1 Existence de la forme canonique de Jordan

Lemme 9.43. Soit $B \in \text{Hom}(V, V)$ un opérateur sur un espace vectoriel V complexe de dimension finie. S'il existe $h \in \mathbb{N}$ tel que $B^h = 0 \in \text{Hom}(V, V)$, alors il existe $v_1, \dots, v_k \in V$ et $m_1, \dots, m_k \in \mathbb{N}$ tels que :

$$\bigcup_{\ell=1}^k \bigcup_{\gamma=1}^{m_\ell} \{B^\gamma(v_\ell)\}$$

est une base de V et :

$$\{B^{m_\ell}(v_\ell) \text{ t.q. } 1 \leq \ell \leq k\}$$

est une base de $\ker(B)$.

Démonstration. Procédons par récurrence sur n , la dimension de V .

Si $n = 1$, il est nécessaire que $B = 0$, et donc le lemme est vrai.

Soit $n > 1$ et soit $h \in \mathbb{N}$ tel que $B^h = 0$. Dès lors, on sait que $\text{Ker } B \neq \{0\}$ et $\text{Im } B \neq V$. On considère la restriction de B à $\text{Im } B$, ce qui donne $(B^h)|_{\text{Im}(B)} = 0 \in \text{Hom}(\text{Im}(B), \text{Im}(B))$. Par hypothèse de récurrence, il existe $u_1, \dots, u_s \in \text{Im}(B)$ et $n_1, \dots, n_s \in \mathbb{N}$ tels que :

$$E' := \bigcup_{\alpha=1}^s \bigcup_{\beta=1}^{n_\alpha} \{B^\beta(u_\alpha)\}$$

est une base de $\text{Im}(B)$ et $\{B^{n_i}(v_i) \text{ t.q. } 1 \leq i \leq s\}$ est une base de $\text{Ker}(B|_{\text{Im}(B)}) = \ker(B) \cap \text{Im}(B)$. On sait que $u_i \in \text{Im}(B)$, donc il existe $v_i \in V$ tel que $(v_i) = u_i$.

Si $\text{Ker}(B) \cap \mathcal{J}(B) \subseteq \text{Ker}(B)$, on peut compléter une base de $\text{Ker}(B|_{\text{Im}(B)}) = \text{Ker}(B) \cap \text{Im}(B)$ en une base de $\text{Ker}(B)$. Soit donc $\{B^{n_1}(v_1), \dots, B^{n_s}(v_s), v_{s+1}, v_{s+2}, \dots, v_k\}$ une base de $\text{Ker}(B)$.

En posant $m_i := n_i + 1$ pour tout $1 \leq i \leq s$, on remarque que :

$$\{B^{m_1}(v_1), \dots, B^{m_s}(v_s), v_{s+1}, v_{s+2}, \dots, v_k\}$$

est une base de $\text{Ker}(B)$ par construction. De plus, montrons que :

$$E := \{v_1, \dots, B^{m_1}(v_1), v_2, \dots, B^{m_2}(v_2), \dots, v_s, \dots, B^{m_s}(v_s), v_{s+1}, v_{s+2}, \dots, v_k\}$$

est une base de V . Pour ce faire, on sait que $|E| = \dim \text{Ker}(B) + \dim \text{Im}(B) = \dim V$, il faut uniquement prouver que E est libre.

Soit la combinaison linéaire nulle suivante :

$$0 = \sum_{\alpha=1}^k \sum_{\beta=0}^{m_\alpha} a_{\alpha\beta} B^\beta(v_\alpha),$$

où pour tout $s < \alpha \leq k$, on pose $m_\alpha = 0$ et où pour tout $1 \leq \alpha \leq k, 1 \leq \beta \leq m_\alpha, a_{\alpha\beta} \in \mathbb{C}$. On sait que $B^{m_i}(v_i) \in \text{Ker}(B)$ pour tout i , et donc $B(B^{m_i}(v_i)) = 0$. En appliquant B de part et d'autre, on trouve :

$$0 = B(0) = \sum_{\alpha=1}^k \sum_{\beta=0}^{n_\alpha} a_{\alpha\beta} B^\beta(u_\alpha).$$

Et de plus, E' est libre et donc, pour tout $1 \leq \alpha \leq s$ et $1 \leq \beta \leq n_\alpha$, on trouve $a_{\alpha\beta} = 0$. Il reste :

$$\sum_{\gamma=1}^s a_{\gamma m_\gamma} B^{m_\gamma}(v_\gamma) + \sum_{\gamma=s+1}^k a_{\gamma 0} = 0.$$

Or, on sait que $\{B^{m_1}(v_1), B^{m_2}(v_2), \dots, B^{m_s}(v_s), v_{s+1}, v_{s+2}, \dots, v_k\}$ est une base de $\text{ker}(B)$ et est donc libre. Il en résulte que $a_{\alpha\beta} = 0$ pour tout α, β . \square

Remarque. Soit $A \in \text{Hom}(V, V)$ un opérateur sur un espace vectoriel complexe V de dimension finie n . On s'intéresse à $\lambda \in \mathbb{C}$ une valeur propre de A et $V(\lambda) = \text{Ker}(A - \lambda I)^m$, le sous-espace propre généralisé de V par λ .

On considère la restriction $B := (A - \lambda I)|_{V(\lambda)} : V(\lambda) \rightarrow V(\lambda)$. Puisque $B^m = 0 \in \text{Hom}(V(\lambda), V(\lambda))$, on sait par le lemme précédent qu'il existe une base de V donnée par :

$$\{v_1, B(v_1), \dots, B^{m_1}(v_1), \dots, v_k, B(v_k), \dots, B^{m_k}(v_k)\}.$$

On remarque que la matrice $m_{E,E}(B) \in M_{m_\lambda \times m_\lambda}(\mathbb{C})$ est donnée par :

$$m_{E,E}(B) = \begin{bmatrix} \tilde{J}_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \tilde{J}_k \end{bmatrix}$$

où les \tilde{J}_i sont sous la forme :

$$\begin{bmatrix} 0 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 \\ 0 & \dots & 0 & 0 \end{bmatrix}$$

Or, on peut exprimer la matrice $m_{E,E}$ par rapport à A :

$$m_{E,E}(B) = m_{E,E} \left((A - \lambda I) \Big|_{V(\lambda)} \right) = m_{E,E}(A \Big|_{V(\lambda)}) - m_{E,E}(\text{Id}_{V(\lambda)}) = m_{E,E} \left(A \Big|_{V(\lambda)} \right) - \lambda I,$$

et donc en réarrangeant, on trouve effectivement la matrice $m_{E,E}(A)$ sous sa forme canonique de Jordan.

Remarque. La preuve du lemme ci-dessus donne un algorithme pour mettre une matrice sous sa forme canonique de Jordan. Il faut :

1. déterminer les valeurs propres de A (les racines du polynôme caractéristique π_A) ;
2. déterminer leur multiplicité algébrique ;
3. pour chacune d'entre elles, résoudre le système $(a - \lambda I)^m x = 0$ où $1 \leq m \leq m_\lambda$;
4. mettre ces vecteurs dans les colonnes de b .

Alors la matrice $b^{-1}ab$ est sous sa forme canonique de Jordan.

9.7.2 Unicité de la forme canonique de Jordan

Théorème 9.44 (Unicité de la forme de Jordan). *Soit V un espace vectoriel complexe de dimension finie n et soit $A \in \text{Hom}(V, V)$. Soit E une base de V telle que $m_{E,E}(A) \in M_{n \times n}(\mathbb{C})$ est sous sa forme de Jordan. S'il existe $F \neq E$ une base de V telle que $m_{F,F}(A)$ est sous sa forme de Jordan, alors les blocs de Jordan de $m_{F,F}(A)$ sont une permutation des blocs de Jordan de $m_{E,E}(A)$.*

Démonstration. On nomme $\lambda_1, \dots, \lambda_k$ les valeurs propres distinctes de A . Soit $\mu_A(x)$ le polynôme minimal de A . On prend i et v un vecteur propre de λ_i (donc $v \in \text{Ker}(A - \lambda_i I)$, ou encore $Av = \lambda_i v$). Soit $j \neq i$ et soit $h \in \mathbb{N}^*$. On remarque que :

$$\begin{aligned} (A - \lambda_j I)^h(v) &= \left(\sum_{k=0}^h \binom{h}{k} A^k (-\lambda_j)^{h-k} I \right) v = \sum_{k=0}^h \binom{h}{k} A^k (-\lambda_j)^{h-k} v \\ &= \sum_{k=0}^h \binom{h}{k} (-\lambda_j)^{h-k} A^k v = \sum_{k=0}^h \binom{h}{k} (-\lambda_j)^{h-k} (\lambda_i)^k v \\ &= \left(\sum_{k=0}^h \binom{h}{k} (\lambda_i)^k (-\lambda_j)^{h-k} \right) v = (\lambda_i - \lambda_j)^h v \neq 0. \end{aligned}$$

On remarque de plus que :

$$\text{Ker}(A - \lambda_i I)^h \subseteq \text{Ker}(A - \lambda_i I)^{h+1},$$

qui est assez évident. De plus, pour tout $h \in \mathbb{N}$, on trouve :

$$V(\lambda_i) = \text{Ker}(A - \lambda_i I)^{m_i} = \text{Ker}(A - \lambda_i I)^{m_i+h}.$$

Supposons qu'il existe $h \in \llbracket 1, m_i - 1 \rrbracket$ tel que :

$$\text{Ker}(A - \lambda_i I)^{m_i} = \text{Ker}(A - \lambda_i I)^{m_i-h}.$$

On peut alors définir $\tilde{\mu}(x)$ par :

$$\tilde{\mu}(x) = \frac{\mu_A(x)}{(x - \lambda_i)^h} = (x - \lambda_i)^{m_i-h} \prod_{j \neq i} (x - \lambda_j)^{m_j}.$$

Or $\deg(\tilde{\mu}(x)) < \deg(\mu_A(x))$ par construction, ce qui est une contradiction avec la définition du polynôme minimal. m_i est donc le plus petit naturel tel que $\text{Ker}(A - \lambda_i I)^{m_i} = V(\lambda_i)$.

On pose alors :

$$\tilde{V}(\lambda_i) := \left\{ v \in V \text{ t.q. } \exists h \in \mathbb{N} \text{ t.q. } v \in \text{Ker}(A - \lambda_i I)^h \right\}.$$

On constate donc effectivement $V(\lambda_i) = \tilde{V}(\lambda_i)$.

Par le lemme précédent, on observe que si $B = A - \lambda_i I$ et $j \geq 1$, alors :

$$\dim_{\mathbb{C}} \text{Ker}(B^j) - \dim_{\mathbb{C}} \text{Ker}(B^{j-1}) = \beta,$$

où β est le nombre de blocs de Jordan associés à λ_i et de taille $\geq j$. La décomposition en blocs de Jordan est donc unique. \square

Corollaire 9.45. *La dimension $\dim_{\mathbb{C}} V(\lambda_i)$ du sous-espace propre généralisé de $A \in \text{Hom}(V, V)$ associé à λ_i une valeur propre de A est égale à la multiplicité algébrique de λ_i .*

Démonstration. On a montré que $V(\lambda_i) = \text{Ker}(A - \lambda_i I)^{m_i}$, et donc la dimension de cet espace est m_i , la multiplicité algébrique de λ_i . \square

10 Espaces euclidiens et hermitiens

10.1 Définitions

Définition 10.1. On définit le *produit hermitien* par une application $V \subseteq \mathbb{C}^n \times V \rightarrow \mathbb{C}$ noté \cdot tel que pour tout $x, x', y, y' \in V, \lambda \in \mathbb{C}$:

- (i) $x(y + y') = xy + xy'$ et $(x + x')y = xy + x'y$;
- (ii) $(\lambda x)y = \lambda(xy)$ et $x(\lambda y) = \bar{\lambda}(xy)$;
- (iii) $xy = \overline{yx}$;
- (iv) $xx \geq 0$ si $x \neq 0$ et $xx = 0$ si $x = 0$.

Définition 10.2. Le produit hermitien usuel est :

$$\forall x, y \in \mathbb{C}^n : xy = \sum_{i=1}^n x_i \overline{y_i}.$$

Définition 10.3. Un espace euclidien est un espace vectoriel V sur \mathbb{R} muni d'un produit scalaire.

Définition 10.4. Un espace hermitien est un espace vectoriel V sur \mathbb{C} muni d'un produit hermitien.

10.1.1 Orthogonalité

Définition 10.5. Dans un espace euclidien ou hermitien, on dit que x, y sont orthogonaux si $xy = 0$, ce que l'on note $x \perp y$.

Proposition 10.6. Soient $x, y \in V$ où V est un espace euclidien ou hermitien. La projection orthogonale de x sur y est donnée par :

$$p = \frac{xy}{yy} y.$$

Démonstration. On veut en effet que $p = \lambda y$ et $x - p$ orthogonal à y . Dès lors, on sait :

$$(x - \lambda y)y = 0.$$

En réarrangeant, on obtient :

$$\lambda yy = xy,$$

et en isolant λ , on obtient effectivement :

$$\lambda = \frac{xy}{yy}.$$

Dès lors, $p = \frac{xy}{yy} y$. □

Définition 10.7. On appelle $\frac{xy}{yy}$ le *coefficient de Fourier* de x par rapport à y .

Proposition 10.8. Dans un espace euclidien ou hermitien, tout ensemble de vecteurs non nuls et deux à deux orthogonaux est une partie libre.

Démonstration. Soit la combinaison linéaire nulle $\sum_{i \in I} \lambda_i e_i = 0$. Soit $j \in I$. On trouve alors :

$$0 = 0e_j = \sum_{i \in I} \lambda_i e_i e_j = \lambda_j (e_j e_j).$$

Il faut donc $\lambda_j = 0$ pour tout $j \in I$, et donc $\{e_i \text{ t.q. } i \in I\}$ est libre. \square

Proposition 10.9 (Orthogonalisation de Gram-Schmidt). *Soit V un espace euclidien ou hermitien de dimension finie n . Si $\{e_1, \dots, e_i\}$ (avec $i \leq n$) est une partie libre orthogonale de V , alors il existe $(n - i)$ vecteurs e_{i+1}, \dots, e_n tels que $\{e_i \text{ t.q. } 1 \leq i \leq n\}$ est une base orthogonale de V .*

Démonstration. On sait que toute partie libre de V peut être complétée en une base. Doit donc $\{e_1, \dots, e_i, f_{i+1}, \dots, f_n\}$ une base de V . On pose pour tout $1 \leq k \leq (n - i)$:

$$e_{i+k} := f_{i+k} - \sum_{j=1}^{i+k-1} \frac{f_{i+k} e_j}{e_j e_j} e_j.$$

Soit $1 \leq k \leq n - i$ et soit $\ell < i + k$. On trouve :

$$e_{i+k} e_\ell = f_{i+k} e_\ell - \sum_{j=1}^{i+k-1} \frac{f_{i+k} e_j}{e_j e_j} e_j e_\ell = f_{i+k} e_\ell - \frac{f_{i+k} e_\ell}{e_\ell e_\ell} e_\ell e_\ell = 0.$$

Dès lors, $E := \{e_i \text{ t.q. } 1 \leq i \leq n\}$ est une partie libre par la proposition 10.8. Et comme $|E| = n$ on sait que E est une base de V . \square

10.1.2 Propriétés de la norme

Définition 10.10. Soit V un espace euclidien ou hermitien et $x \in V$. On définit la norme de x par :

$$\|x\| := \sqrt{x x}.$$

Proposition 10.11 (Propriétés de la norme). *Soit V un espace euclidien ou hermitien. Pour tout $x, y \in V$, on a :*

1. $\|x\| \geq 0$ si $x \neq 0$ et $\|x\| = 0$ si $x = 0$;
2. $\|\lambda x\| = |\lambda| \|x\|$;
3. $|x y| \leq \|x\| \|y\|$ (inégalité de Cauchy-Schwartz) ;
4. $\|x + y\| \leq \|x\| + \|y\|$ (inégalité triangulaire).

10.1.3 base orthonormale

Définition 10.12. Une base orthonormale d'un espace euclidien ou hermitien V est une base orthogonale E de V telle que :

$$\forall e \in E : \|e\| = 1.$$

Proposition 10.13. Un espace euclidien ou hermitien admet toujours une base orthonormale.

Démonstration. Par Gram-Schmidt, on obtient $E = \{e_i \text{ t.q. } 1 \leq i \leq n\}$, une base orthogonale. On pose alors :

$$E' := \left\{ \frac{e_i}{\|e_i\|} \text{ t.q. } 1 \leq i \leq n \right\},$$

qui est une base orthonormale. □

Théorème 10.14. *Tout espace euclidien de dimension réelle n est isomorphe à \mathbb{R}^n avec le produit scalaire usuel et tout espace hermitien de dimension m est isomorphe à \mathbb{C}^m avec le produit hermitien usuel.*

Démonstration. Corollaire direct de la proposition 6.4. □

10.1.4 Angles entre vecteurs

Définition 10.15. Soient $x, y \in V$, un espace euclidien ou hermitien. On appelle $\theta \in [0, \pi]$ tel que $\cos \theta = \frac{xy}{\|x\|\|y\|}$ l'angle entre x et y .

Remarque. Par l'inégalité de Cauchy-Schwartz, on sait que $-\|x\|\|y\| \leq xy \leq \|x\|\|y\|$ et donc :

$$-1 \leq \frac{xy}{\|x\|\|y\|} \leq 1.$$

Dès lors, on sait qu'il existe un tel θ .

Remarque. On remarque également que cette définition permet de dire que deux vecteurs sont orthogonaux si et seulement si $\theta = \frac{\pi}{2}$.

10.1.5 Métrique

Définition 10.16. Soit V un espace euclidien ou hermitien et soient $x, y \in V$. On définit la *distance* entre x et y par :

$$|xy| := \|x - y\|.$$

Définition 10.17. Un espace euclidien ou hermitien muni d'une distance est appelé un *espace métrique*, c'est-à-dire, pour tout $x, y, z \in V$:

1. $|xy| > 0$ si $x \neq y$ et $|xy| = 0$ si $x = y$;
2. $|xy| = |yx|$;
3. $|xy| \leq |xz| + |zx|$.

Théorème 10.18 (Théorème de Pythagore). *Soit V un espace euclidien ou hermitien et soient $x, y, z \in V$. Si $(x - y) \perp (x - z)$, alors :*

$$|yz|^2 = |xy|^2 + |xz|^2.$$

Démonstration. On sait que si $a, b \in V$ sont orthogonaux, alors :

$$|ab| = \|a - b\|^2 = (a - b)(a - b) = \|a\|^2 - ab - ba + \|b\|^2 = \|a\|^2 + \|b\|^2.$$

En posant $a = x - y$ et $b = x - z$, on obtient :

$$|yz| = \|z - y\|^2 = \|(x - y) - (x - z)\|^2 = \|x - y\|^2 + \|x - z\|^2 = |xy|^2 + |xz|^2.$$

□

Théorème 10.19 (Loi des cosinus). Soit V un espace euclidien et soient $x, y, z \in V$. Si θ est l'angle entre les vecteurs $(x - y)$ et $(x - z)$, alors :

$$|yz|^2 = |xy|^2 + |xz|^2 - 2|xy||xz| \cos \theta.$$

Démonstration. Si $a, b \in V$, on sait que :

$$\|a - b\|^2 = \|a\|^2 - ab - ba + \|b\|^2 = \|a\|^2 + \|b\|^2 - 2ab = \|a\|^2 + \|b\|^2 - 2\|a\|\|b\| \cos \theta.$$

En posant $a = x - y$ et $b = x - z$, on obtient bien la loi des cosinus. \square

10.2 Sous-espaces orthogonaux

Définition 10.20. Soit V un espace euclidien ou hermitien. Soit $S \subseteq V$ un sous-ensemble non-vide de V . On définit le sous-espace orthogonal à S par :

$$S^\perp := \{x \in V \text{ t.q. } \forall s \in S : xs = 0\}.$$

Proposition 10.21. Si S est un sous-ensemble non-vide d'un espace euclidien ou hermitien V , alors S^\perp est un sous-espace de V tel que :

$$S^\perp \cap S \in \{\emptyset, \{0\}\}.$$

Démonstration. Par linéarité du produit scalaire ou hermitien, on sait que S^\perp est un sous-espace de V . De plus, si on suppose $S^\perp \cap S \neq \emptyset$, alors il faut obligatoirement $S \cap S^\perp = \{0\}$ car si $x \in S^\perp \cap S$, alors $xx = 0$, et donc $x = 0$. \square

10.2.1 Droites, hyperplans et projections orthogonales dans \mathbb{R}^n

Définition 10.22. Soient $a, b \in \mathbb{R}^n$. On dit que l'ensemble $D := \{ta + b \text{ t.q. } t \in \mathbb{R}\}$ où $a \neq 0 \in \mathbb{R}^n$ est une droite dans \mathbb{R}^n . a est appelé le vecteur directeur de D .

Définition 10.23. Soit $a \in \mathbb{R}^n$ et soit $S = \{a\}$. On dit que le sous-espace S^\perp est un hyperplan de \mathbb{R}^n passant par l'origine.

Définition 10.24. Soient V un espace euclidien ou hermitien de dimension finie n et W un sous-espace de V de dimension $m < n$. Soit $E = \{e_i \text{ t.q. } 1 \leq i \leq m\}$ une base de W . On dit que la projection orthogonale de $v \in V$ dans W est le vecteur $w \in W$ tel que :

$$w := \sum_{i=1}^m \frac{ve_i}{e_i e_i} e_i.$$

Théorème 10.25 (Caractérisation de la projection orthogonale dans un sous-espace). Soit V un espace euclidien ou hermitien de dimension finie n . Soit W un sous-espace de V de dimension $m < n$. Soit $\{e_i \text{ t.q. } 1 \leq i \leq m\}$ une base de W . Soient $v \in V$ et $w \in W$ tels que :

$$w = \sum_{i=1}^m \frac{ve_i}{e_i e_i} e_i.$$

Alors :

1. $v - w \in W^\perp$;

2. si $w' \in W$ tel que $v - w' \in W^\perp$, alors $w' = w$, en particulier, w ne dépend pas de la base orthogonale choisie ;
3. pour tout $w' \in W$, on a $\|v - w\| \leq \|v - w'\|$, donc w est une approximation optimale de v dans W ;
4. si $\omega \in W$ et si pour tout $w' \in W : \|v - \omega\| \leq \|v - w'\|$, alors $w = \omega$, en particulier l'approximation optimale est unique.

Théorème 10.26. *Si V est un espace euclidien ou hermitien de dimension finie n et W est un sous-espace de V , alors :*

$$V = W \oplus W^\perp,$$

en particulier :

$$\dim_{\mathbb{K}} W + \dim_{\mathbb{K}} W^{\perp} = \dim_{\mathbb{K}} V$$

où $\mathbb{K} = \mathbb{R}$ si V est euclidien et $\mathbb{K} = \mathbb{C}$ si V est hermitien.

Démonstration. Soit $v \in V$ et soit $w \in W$ la projection orthogonale de v dans W . On sait $v - w \in W^\perp$, or $v = (v - w) + w$ où $w \in W$. Donc $V = W + W^\perp$. On a également vu que $W \cap W^\perp = \{0\}$, ce qui implique $W \oplus W^\perp = V$. \square

10.3 Isométries

Définition 10.27. Une isométrie dans un espace euclidien ou hermitien V est un opérateur α inversible tel que :

$$\forall x, y \in V : \alpha(x)\alpha(y) = xy.$$

Remarque. Une isométrie conserve le produit scalaire (ou hermitien) et donc conserve les distance (d'où son nom *iso-métrie*) et les angles.

De plus, on remarque qu'une isométrie α envoie une base orthonormale sur une base orthonormale étant donné que les angles et distances sont conservés.

Théorème 10.28. Soit V un espace euclidien ou hermitien et soient $E = \{e_i \text{ t.q. } 1 \leq i \leq n\}$ et $F = \{f_i \text{ t.q. } 1 \leq i \leq n\}$ deux bases orthonormales de V . Alors il existe une unique isométrie α telle que $\forall 1 \leq i \leq n : \alpha(e_i) = f_i$.

Démonstration. Il existe un unique opérateur α (isomorphisme) déterminé par $\alpha(e_i) = f_i$. Montrons maintenant qu'il est isométrique.

Soient $x, y \in V$ tels que :

$$x = \sum_{i=1}^n x_i e_i, \quad \text{et} \quad y = \sum_{i=1}^n y_i e_i.$$

On remarque que :

$$xy = \left(\sum_{i=1}^n x_i e_i \right) \left(\sum_{j=1}^n y_j e_j \right) = \sum -i = 1^n \sum_{j=1}^n x_i \overline{y_j} (e_i e_j) = \sum_{i=1}^n x_i \overline{y_i}.$$

De plus :

$$\alpha(x)\alpha(y) = \sum_{i=1}^n \sum_{j=1}^n x_i \overline{y_j} \alpha(e_i) \alpha(e_j) = \sum_{i=1}^n \sum_{j=1}^n x_i \overline{y_j} (f_i f_j) = \sum_{i=1}^n x_i \overline{y_i} = xy.$$

α est donc bien une isométrie.

Proposition 10.29. Soit V un espace euclidien ou hermitien de dimension finie n et soit E une base de V . Une matrice $[a_{ij}] \in M_{n \times n}(\mathbb{K})$ est la matrice d'une isométrie α dans la base E si et seulement si :

$$a^{-1} = \overline{a}^T.$$

Démonstration. Soit α une isométrie et soit $[a_{ij}] = m_{E,E}(\alpha)$. Or on sait que les vecteurs $\alpha(e_i)$ sont de norme 1 et sont deux à deux orthogonaux. Donc pour tout $1 \leq i, j \leq n$:

$$\delta_{ij} = \alpha(e_i) \alpha(e_j) = \left(\sum_{k=1}^n a_{ki} e_k \right) \left(\sum_{k=1}^n a_{kj} e_k \right) = \sum_{k=1}^n a_{ki} \overline{a_{kj}}.$$

On en déduit alors :

$$a^T \overline{a} = I = \overline{a}^T a = a \overline{a}^T.$$

Et donc :

$$a^{-1} = \overline{a}^T.$$

Réciproquement, si $[a_{ij}]$ est une telle matrice, alors cette matrice définit un opérateur linéaire α inversible qui envoie la base orthonormale E sur une base orthonormale. α est donc une isométrie. \square

Définition 10.30. Soit $a \in M_{n \times n}(\mathbb{R})$. On dit que a est *orthogonale* si $a^{-1} = a^T$.

Soit $u \in M_{n \times n}(\mathbb{C})$. On dit que u est *unitaire* si $u^{-1} = \overline{u}^T$.

Définition 10.31. On définit $O_n(\mathbb{R})$ comme l'ensemble des matrices carrées $n \times n$ orthogonales.

On définit $U_n(\mathbb{C})$ comme l'ensemble des matrices carrées $n \times n$ unitaires.

Remarque. Les ensembles U_n et O_n forment des groupes multiplicatifs. De plus, ces groupes sont isomorphes aux isométries de \mathbb{R}^n (respectivement \mathbb{C}^n).

Proposition 10.32. $U_n(\mathbb{C}) \cap M_{n \times n}(\mathbb{R}) = O_n(\mathbb{R})$.

Remarque. On remarque que dans O_2 , les seules matrices (donc isométries) existantes sont les rotations qui fixent l'origine, les symétries par rapport à une droite passant par l'origine, ainsi que leurs composées.

10.4 Déterminants et valeurs propres

Proposition 10.33. La norme du déterminant d'une matrice orthogonale ou unitaire vaut 1.

Démonstration. Soit $a \in U_n$. On calcule :

$$1 = \det(I) = \det(a \overline{a}^T) = \det(a) \det(\overline{a}^T) = \det(a) \overline{\det(a)} = |\det(a)|^2.$$

\square

Proposition 10.34. Si λ est une valeur propre de a une matrice orthogonale ou unitaire, alors $|\lambda| = 1$.

Démonstration. Soit V un espace euclidien ou hermitien. On appelle α l'isométrie associée à a . On prend x un vecteur propre de a associé à la valeur propre λ . On calcule :

$$xx = \alpha(x) \alpha(x) = (ax)(ax) = (\lambda x)(\lambda x) = \lambda \overline{\lambda} xx = |\lambda|^2 xx.$$

Il faut donc $|\lambda| = 1$. \square

Remarque. Ce résultat est assez intuitif du fait qu'une isométrie conserve les distances. Il faut donc que les valeurs propres soient de norme 1.

10.5 Le groupe orthogonal O_3

Si α est une isométrie $\in O_3$, alors son polynôme caractéristique est de degré 3 et donc admet au moins une racine réelle. On sait que cette racine λ est une valeur propre et donc $\lambda = \pm 1$. Par Gram-Schmidt, on peut construire E une base orthonormale de \mathbb{R}^3 telle que :

$$m_{E,E}(\alpha) = \begin{bmatrix} \pm 1 & f & g \\ 0 & a & b \\ 0 & c & d \end{bmatrix}.$$

Or les colonnes de $m_{E,E}(\alpha)$ forment une base orthonormale. Donc il est impératif d'avoir $f = g = 0$. De plus, il faut que la matrice

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

soit une isométrie de O_2 . L'ensemble des isométries de O_3 sont donc les rotations qui fixent l'origine, les symétries par rapport à un plan qui passe par l'origine et leur composées.

Remarque. La forme générale des matrices de O_4 est :

$$\begin{bmatrix} \cos \theta & -\sin \theta & 0 & 0 \\ \sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & \cos \varphi & -\sin \varphi \\ 0 & 0 & \sin \varphi & \cos \varphi \end{bmatrix} \in O_4(\mathbb{R}), \theta, \varphi \in [0, 2\pi).$$

Elles représentent donc des rotations qui ne sont pas des rotations *simples* mais des rotations doubles simultanées.

11 Matrices symétriques et hermitiennes

11.1 Définitions

Définition 11.1. La matrice $a \in M_{n \times n}(\mathbb{R})$ est dite symétrique si :

$$a = a^T.$$

Définition 11.2. La matrice $b \in M_{n \times n}(\mathbb{C})$ est dite hermitienne si :

$$a = \bar{a}^T.$$

Lemme 11.3. Soient $x, y \in \mathbb{C}^n$. Soit $a \in M_{n \times n}(\mathbb{C})$. Alors :

$$(ax)y = x(\bar{a}^T y).$$

Démonstration. On calcule :

$$(ax)y = \sum_{i=1}^n \sum_{j=1}^n (a_{ij}x_j) \bar{y}_i = \sum_{j=1}^n x_j \sum_{i=1}^n a_{ij} \bar{y}_i = \sum_{j=1}^n x_j (\bar{a}^T y)_j = x(\bar{a}^T y).$$

□

Proposition 11.4. Soit a une matrice symétrique ou hermitienne. Alors toute valeur propre complexe de a est réelle.

Démonstration. Soit λ une valeur propre de a et x un vecteur propre de a associé à λ . On calcule :

$$(ax)x = x(\bar{a}^T x) = x(ax) = \overline{(ax)x}.$$

Donc $(ax)x \in \mathbb{R}$. Mais également :

$$\mathbb{R} \ni (ax)x = (\lambda x)x = \lambda xx,$$

où $xx \in \mathbb{R}_0$, et donc $\lambda \in \mathbb{R}$.

□

Proposition 11.5. Deux vecteurs propres d'une matrice a symétrique ou hermitienne associés à deux valeurs propres distinctes sont orthogonaux.

Démonstration. Soient λ, λ' deux valeurs propres de a et soient x, x' deux vecteurs propres associées respectivement à λ et λ' . On calcule :

$$\lambda(xx') = (\lambda x)x' = (ax)x' = x(\bar{a}^T x') = x(ax') = x(\lambda' x) = \lambda'(xx').$$

Donc on peut exprimer :

$$0 = \lambda(xx') - \lambda'(xx') = (\lambda - \lambda')(xx').$$

Or on sait $\lambda \neq \lambda'$ et donc il faut $xx' = 0$.

□

11.2 Diagonalisation de matrices hermitiennes et symétriques

Théorème 11.6. Si a est une matrice hermitienne (respectivement symétrique), alors il existe une base orthogonale de \mathbb{C}^n (respectivement \mathbb{R}^n) dans laquelle l'opérateur associé à a est diagonal.

Remarque. Il est équivalent de dire que pour toute matrice hermitienne (ou symétrique), il existe une matrice b unitaire (orthogonale) telle que la matrice $b^{-1}ab$ est diagonale.

Démonstration. la démonstration se fait ici sur les matrices symétriques, le cas hermitien se démontre de la même manière. Procédons par récurrence sur n , la dimension de a . Pour $n = 1$, a est diagonale.

Supposons le théorème vrai pour $n \in \mathbb{N}^*$ et montrons-le pour $n + 1$. Soit $a \in M_{(n+1) \times (n+1)}(\mathbb{R})$ symétrique avec $n \geq 1$. Il existe $\lambda \in \mathbb{R}$ une valeur propre de a . Donc il existe $v \neq 0$ un vecteur propre de a associé à λ et il existe doc une base orthonormale de \mathbb{R}^n qui *commence* par $\frac{v}{\|v\|}$.

On pose $c \in M_{(n+1) \times (n+1)}(\mathbb{R})$ la matrice dont les colonnes sont les vecteurs de cette base orthonormale. La matrice c est une matrice orthogonale. Dès lors, on sait que :

$$c^{-1}ac = \begin{bmatrix} \lambda & *_1 & \dots & *_2 \\ 0 & & & \\ \vdots & & a_{(1)} & \\ 0 & & & \end{bmatrix}.$$

On trouve également que la matrice $c^{-1}ac$ est symétrique :

$$(c^{-1}ac)^T = c^T a^T (c^{-1})^T = c^{-1}ac.$$

On peut donc trouver la première ligne de $c^{-1}ac$ qui est la première colonne transposée. On peut donc écrire :

$$c^{-1}ac = \begin{bmatrix} \lambda & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & a_{(1)} & \\ 0 & & & \end{bmatrix}.$$

Dès lors, $a_{(1)}$ est également symétrique mais de dimension $n \times n$. Par hypothèse de récurrence, on sait qu'il existe $d_{(1)}$ orthogonale telle que :

$$(d_{(1)})^{-1}(a_{(1)})d_{(1)}$$

est diagonale.

En posant :

$$d := \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & d_{(1)} & \\ 0 & & & \end{bmatrix},$$

on peut dire que $d^{-1}(c^{-1}ac)d$ est diagonale. En posant $b := cd$, le théorème est prouvé. □

11.3 Formes bilinéaires et quadratiques

Définition 11.7. Soit V un espace euclidien. Une fonction $f : V \times V \rightarrow \mathbb{R}$ est une forme bilinéaire symétrique si f est linéaire dans ses deux composantes et si $\forall v, w : f(v, w) = f(w, v)$.

Définition 11.8. La forme quadratique associée à la forme bilinéaire $f : V \times V \rightarrow \mathbb{R}$ est l'application $q : V \rightarrow \mathbb{R}$ définie par :

$$\forall v \in V : q(v) := f(v, v).$$

Remarque. On remarque que la forme quadratique est déterminée par l'application bilinéaire f , mais réciproquement, la forme quadratique est également entièrement définie par la forme bilinéaire f . En effet, pour tout $v, w \in V$:

$$q(v + w) = f(v + w, v + w) = f(v, v) + f(w, v) + f(v, w) + f(w, w) = q(v) + q(w) + 2f(v, w),$$

et donc :

$$f(v, w) = \frac{1}{2} (q(v + w) - q(v) - q(w))$$

Soit V un espace euclidien et soit $E = \{e_1, \dots, e_n\}$ une base de V . Si $f : V \times V \rightarrow \mathbb{R}$ est une forme bilinéaire sur V , on sait :

$$f(v, w) = \sum_{i=1}^n \sum_{j=1}^n v_i w_j f(e_i, e_j).$$

On pose $g_{ij} := f(e_i, e_j)$.

Définition 11.9. On appelle $g = [g_{ij}] \in M_{n \times n}(\mathbb{R})$ la *matrice associée à la forme bilinéaire f dans la base E* .

Remarque. Étant donné que la forme quadratique q est déterminée par la forme bilinéaire f , on peut également dire que g est la matrice de la forme quadratique q . En effet :

$$q(v) = f(v, v) = \sum_{i=1}^n \sum_{j=1}^n v_i v_j g_{ij} = \sum_{i=1}^n v_i^2 g_{ii} + \sum_{1 \leq i < j \leq n} (g_{ij} + g_{ji}) v_i v_j = \sum_{i=1}^n g_{ii} x_i^2 + \sum_{1 \leq i < j \leq n} 2g_{ij} v_i v_j.$$

On peut donc remplir la matrice g en connaissant q : il faut mettre les coefficients des termes en x_i^2 sur la diagonale, et ensuite mettre dans les autres g_{ij} en divisant le coefficient de $x_i x_j$ par deux.

11.3.1 Changement de base pour les formes quadratiques

Proposition 11.10. la forme matricielle d'une forme bilinéaire symétrique $f : V \times V \rightarrow \mathbb{R}$ est :

$$f(v, w) = v^T g w.$$

Proposition 11.11. Soient $E = \{e_1, \dots, e_n\}$ et $F = \{f_1, \dots, f_n\}$ deux bases distinctes de V , un espace euclidien de dimension n . Soit $f : V \times V \rightarrow \mathbb{R}$ une forme bilinéaire symétrique sur V . Si $g = [g_{ij}]$ est la matrice de f dans la base E , on pose $b \in M_{n \times n}(\mathbb{R})$ la matrice de changement de base de F vers E . Donc :

$$f_j = \sum_{i=1}^n b_{ij} e_i.$$

Alors $g' := b^T g b$ est la matrice de f dans la base F .

Démonstration. On prend $v, w \in V$ que l'on peut exprimer dans les bases E et F par :

$$\begin{aligned} v &= \sum_{i=1}^n v_i e_i = \sum_{i=1}^n v'_i f_i, \\ w &= \sum_{i=1}^n w_i e_i = \sum_{i=1}^n w'_i f_i. \end{aligned}$$

Puisque $bv' = v$ et $bw' = w$, on peut écrire :

$$f(v, w) = v^T g w = (bv')^T g (bw') = v'^T (b^T g b) w'.$$

on a bien $g' = b^T g b$ la matrice de f dans la base F . □

Définition 11.12. Soit q une forme bilinéaire sur un espace euclidien V . On dit que :

- q est *définie positive* si pour tout $v \in V_0$: $q(v) > 0$ et $q(0_V) = 0_V$;
- q est *définie négative* si pour tout $v \in V_0$: $q(v) < 0$ et $q(0_V) = 0_V$;
- q est *semi-définie positive* si pour tout $v \in V$: $q(v) \geq 0$;
- q est *semi-définie négative* si pour tout $v \in V$: $q(v) \leq 0$;
- q est *indéfinie* sinon.

Proposition 11.13. Une forme quadratique q sur un espace euclidien V est définie positive (respectivement négative) si et seulement si toute valeur propre de la matrice g de q dans une base E de V est positive (respectivement négative).

Démonstration. Par le théorème 11.6, on sait qu'il existe $a \in M_{n \times n}(\mathbb{R})$ une matrice orthogonale telle que :

$$a^{-1} g a = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} =: d$$

est diagonale et λ_i est une valeur propre de g . Par orthogonalité de a , on sait que $(a a^{-1}) = (a a^T) = I$. Dès lors :

$$q(v) = v^T g v = v^T I g I v = (v^T a) a^{-1} g a (a^{-1} v) = (a^{-1} v)^T (a^{-1} g a) (a^{-1} v) = (a^T v)^T d (a^T v).$$

En posant $w = a^T v$, on peut exprimer :

$$q(v) = w^T d w = \sum_{i=1}^n \lambda_i w_i^2.$$

On remarque que si $\forall 1 \leq i \leq n$: $\lambda_i < 0$, alors $q(v) < 0$ (et si $\lambda_i > 0$, alors $q(v) > 0$). Inversement, en prenant le vecteur e_i on remarque que si $q(e_i) > 0$, alors $\lambda_i > 0$ (et si $q(e_i) < 0$, alors $\lambda_i < 0$). Donc si $q(v)$ est définie positive (ou négative), il faut $\lambda_i > 0$ (ou $\lambda_i < 0$) pour tout i . □

Proposition 11.14. Soit une forme quadratique $q : V \times V \rightarrow \mathbb{R}$ et $g \in M_{n \times n}(\mathbb{R})$ sa matrice dans une base E de V . Alors :

1. q est définie positive si et seulement si les coefficients du polynôme caractéristique de g alternent en signe ;
2. q est définie négative si et seulement si les coefficients du polynôme caractéristique de g ont tous le même signe.

11.3.2 Signature et théorème de Sylvester

Définition 11.15. Soit $a \in M_{n \times n}(\mathbb{R})$ une matrice symétrique. On définit la *signature* de a par le triplet (n_+, n_-, n_0) où :

- n_0 est le nombre de valeurs propres nulles de a ;
- n_+ est le nombre de valeurs propres strictement positives de a ;

- n_- est le nombre de valeurs propres strictement négatives de a .

Remarque. Les valeurs propres sont comptées avec leur multiplicité dans la signature de la forme quadratique (ou de la matrice g).

Théorème 11.16 (Loi d'intertie de Sylvester). *Soit $g \in M_{n \times n}(\mathbb{R})$ et soient $a, b \in M_{n \times n}(\mathbb{R})$ inversibles. Alors la signature de $(a^T g a)$ et la signature de $(b^T g b)$ sont équivalentes.*

Remarque. Cela veut dire que la signature d'une forme quadratique est un invariant.

Démonstration. On remarque que $(a^T g a)$ est une matrice symétrique et $(b^T g b)$ également. Soient (m_+, m_-, m_0) et (n_+, n_-, n_0) la signature de $(a^T g a)$ et $(b^T g b)$. Soit $r := \text{rang}(g)$. Puisque les matrices a et b sont inversibles, on remarque que $\text{rang}(a^T g a) = r = \text{rang}(b^T g b)$. Or $r = \dim(\text{Im}(a^T g a)) = m_+ + m_-$. Similairement, $r = n_+ + n_-$. On en déduit donc $m_0 = n_0$.

On sait également qu'il existe $\tilde{a}, \tilde{b} \in M_{n \times n}(\mathbb{R})$ telles que :

$$\left\{ \begin{array}{l} \tilde{a}^T a^T g a \tilde{a} = \begin{bmatrix} \lambda_1 & & & & 0 \\ & \ddots & & & \\ & & \lambda_r & & \\ & & & 0 & \\ & & & & \ddots \\ 0 & & & & & 0 \end{bmatrix}, \\ \tilde{b}^T b^T g b \tilde{b} = \begin{bmatrix} \mu_1 & & & & 0 \\ & \ddots & & & \\ & & \mu_r & & \\ & & & 0 & \\ & & & & \ddots \\ 0 & & & & & 0 \end{bmatrix}, \end{array} \right.$$

où $\forall 1 \leq i \leq m_+ : \lambda_i > 0, \forall m_+ + 1 \leq i \leq r : \lambda_i < 0$ et $\forall 1 \leq i \leq n_+ : \mu_i > 0, \forall n_+ + 1 \leq i \leq r : \mu_i < 0$.

Soient $E = \{e_i \text{ t.q. } 1 \leq i \leq n\}, F = \{f_i \text{ t.q. } 1 \leq i \leq n\}$ deux bases de V telles que $m_{E,E}(q) = g$ et $m_{F,F}(q) = F$.

On remarque que $\Delta = \{e_1, \dots, e_{m_+}, f_{n_++1}, \dots, f_n\}$. Dès lors :

$$m_+ + (n - n_+) \leq n \quad \Longleftrightarrow \quad m_+ \leq n_+.$$

Similairement, on trouve $n_+ \leq m_+$, ce qui implique $m_+ = n_+$. On en conclut $m_- = n_-$.

Pour montrer que δ est libre, on procède à une combinaison linéaire :

$$0 = \alpha_1 e_1 + \dots + \alpha_{m_+} e_{m_+} + \beta_{n_++1} f_{n_++1} + \dots + \beta_n f_n = 0.$$

On pose :

$$v = \alpha_1 e_1 + \dots + \alpha_{m_+} e_{m_+} = -(\beta_{n_++1} f_{n_++1} + \dots + \beta_n f_n).$$

On évalue :

$$q(v) = \sum_{i=1}^{m_+} \lambda_i \alpha_i^2 \geq 0,$$

car les λ_i sont > 0 par définition de n_+ . On évalue également :

$$q(v) = \sum_{i=n_++1}^n \beta_i^2 \mu_i = \sum_{i=1}^r \beta_i^2 \mu_i \leq 0,$$

car les μ_i sont < 0 par définition de n_+ . Il reste alors uniquement les f_{r+1}, \dots, f_n qui sont tous dans F , qui est une famille libre. Dès lors, il faut $0 = f_{r+1} = f_{r+2} = \dots = f_n$. On a donc bien $0 = \alpha_i = \beta_j$. \square

Remarque. La signature d'une forme quadratique est invariante, mais ses valeurs propres peuvent changer lors d'une transformation $g \mapsto a^T g a$.

12 Approche naïve des affinités

12.1 Introduction

Définition 12.1. On pose $\mathcal{A}(\mathbb{R}^n)$ l'espace \mathbb{R}^n en tant qu'espace affine (espace linéaire en considérant les translations).

Définition 12.2. On appelle *affinités* les transformations de la forme :

$$L : \mathcal{A}(\mathbb{R}^n) \rightarrow \mathcal{A}(\mathbb{R}^n) : x \mapsto ax + b,$$

où $a \in M_{n \times n}(\mathbb{R})$ et $b \in \mathbb{R}^n$.

Proposition 12.3. *La composée de deux affinités est une affinité.*

Démonstration. Soient L_1, L_2 deux affinités de $\mathcal{A}(\mathbb{R}^n)$ dans $\mathcal{A}(\mathbb{R}^n)$. Alors :

$$(L_1 \circ L_2)(x) = L_1(a_2x + b_2) = a_1(a_2x + b_2) + b_1 = (a_1a_2)x + (a_1b_2 + b_1).$$

□

Définition 12.4. Une isométrie affine est une isométrie tradlatée par $b \in \mathbb{R}^n$. C'est donc une affinité $L : x \mapsto ax + b$ où a est orthogonale.

12.2 Lieu des solutions d'une équation polynômiale

Définition 12.5. Soient $f : \mathcal{A}(\mathbb{R}^n) \rightarrow \mathbb{R}$ une fonction et $S := \{v \in S \text{ t.q. } f(v) = 0\}$. On dit que S est le *lieu des solutions de l'équation* $f(x) = 0$.

Remarque. Si L est un affinité, $L(S)$ est S transformé puis tradlaté. Donc $v \in L(S) \iff L^{-1}(v) \in S$. $L(S)$ est donc le lieu des solutions de l'équation $f(L^{-1}(x)) = 0$.

Proposition 12.6. *Une affinité envoie un sous-espace affine sur un sous-espace affine.*

13 Coniques et quadriques