Algèbre linéaire et géométrie

R. Petit

année académique 2015-2016

Contents

1 Rappels

1.1 Ensembles et fonctions

Définition 1.1. Soient X, Y deux ensembles. Une fonction $f : X \to Y$ 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 $f : X \to Y$ est $f(A) \coloneqq \{f(x) \text{ t.q. } x \in A\}$. L'image de f est $f(X) \subseteq Y$.
- La préimage de $B \subseteq Y$ par f est $f^{-1}(B) \coloneqq \{x \in X \text{ t.q. } f(x) \in B\}$. La préimage de $y \in Y$ est donnée par $f^{-1}(y) \coloneqq f^{-1}(\{y\})$.

Définition 1.3. Soit $f : X \to Y$.

- Une fonction $f : X \to Y$ est injective si $\forall x, x' \in X : f(x) = f(x') \Rightarrow x = x'$.
- Une fonction $f : X \to Y$ 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 $f : X \to Y$ est bijective si elle est injective et surjective.

Définition 1.4. Soit $f : X \to Y$ une fonction bijective. On définit la fonction inverse f^{-1} de f par $f^{-1} : Y \to X : y \mapsto x \text{ t.q. } f(x) = y$.
Définition 1.5. Soient X, Y, Z trois ensembles. Soient $f : X \to Y, g : Y \to Z$ deux fonctions. On définit la composée $g \circ f : X \to Z : x \mapsto g(f(x))$.
Lemme 1.6. *La composée* $f \circ f^{-1} = \mathrm{Id}_Y$ *et la composée* $f^{-1} \circ f = \mathrm{Id}_X$.
Définition 1.7. Soient X, Y deux ensembles. On définit le produit cartésien :

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

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

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

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

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

1.2 Trigonométrie

Définition 1.10. Les fonctions $\cos, \sin : \mathbb{R} \to \mathbb{R}$ 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 $\arccos : [-1, 1] \to [0, \pi]$ et $\arcsin : [-1, 1] \to \left[\frac{\pi}{2}, \frac{3}{2}\pi\right]$.
- On définit également la fonction $\tan :]-\frac{\pi}{2}, \frac{\pi}{2}[\to \mathbb{R} : \theta \mapsto \frac{\sin\theta}{\cos\theta}$ et son inverse $\arctan : \mathbb{R} \to]-\frac{\pi}{2}, \frac{\pi}{2}[$.

1.3 Nombres complexes

Définition 1.15. On définit l'ensemble des complexes $\mathbb{C} \coloneqq \{a + ib \text{ t.q. } (a, b) \in \mathbb{R}^2\}$. 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 $\Re z \coloneqq a$ et $\Im z \coloneqq b$ 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 :

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

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. □

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 $\overline{z} = a - ib = \rho \exp(-i\theta) \in \mathbb{C}$.
Proposition 1.21 (Propriétés du conjugué)**.** *Soient $z, w, \overline{z}, \overline{w} \in \mathbb{C}$ deux complexes et leur conjugué.*

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

Démonstration. Trivial par la définition. □

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

$$z^{-1} \coloneqq \frac{\overline{z}}{|z|^2}.$$

Définition 1.23 (L'identité d'Euler)**.** On définit l'exponentielle complexe par $\exp(i\theta) \coloneqq \cos\theta + i \sin\theta$.
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 $\rho_w = 0$. □

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 :

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

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 $GCD(a, b)$.
Remarque. Pour tout couple (a, b), notons que $0 \leq GCD(a, b) < \max\{|a|, |b|\}$.
Définition 2.2. Soient $(a, b) \in \mathbb{Z} \times \mathbb{Z}_0$. Il existe un unique $q \in \mathbb{Z}$ tel que $r \coloneqq a - qb$ 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 $r_{-1} \coloneqq a$ et $r_0 \coloneqq b$. 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* $\mathrm{GCD}(a, b) \coloneqq |r_k|$

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 $ya + zb = \mathrm{GCD}(a, b)$.*

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 $\mathrm{GCD}(a, b) = r_k = \alpha_k r_0 + \beta_{k+1} r_{-1} = \alpha_k b + \beta_{k+1} a$ 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 :

$$\deg f(x) \coloneqq n.$$

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 $\mathrm{GCD}(f(x), g(x))$ 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 $r_{-1}(x) \coloneqq f(x)$ 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 :*

$$\mathrm{GCD}(f(x), g(x)) \coloneqq \frac{r_k(x)}{[x^{\deg r_k(x)}] r_k(x)}.$$

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. □

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

Démonstration. Preuve similaire à l'algorithme sur des entiers. □

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 :

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

Remarque. Si $f \in \mathbb{R}[x] \subset \mathbb{C}[x]$, alors son conjugué $\overline{f}(x) = f(x)$ car $\forall x \in \mathbb{R} : \overline{x} = x$.
Lemme 2.17. *Soit $f \in \mathbb{C}[x]$. Alors $\overline{f}(\overline{z}) = \overline{f(z)}$.*

Démonstration.

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

□

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 $0 = f(z) = \overline{f(z)} = \overline{f}(\overline{z}) = f(\overline{z})$. □

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 :*

$$(x - z_0)(z - \overline{z_0})$$

est un diviseur de $f(x)$.

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

$$f(x) = q(x)(x - z_0)(x - \overline{z_0}) + r(x).$$

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 $(x - z_0)(x - \overline{z_0})$ 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 $z_1, \ldots, z_n$ ses n racines, multiplicité comptée. On définit, pour tout $1 \leq k \leq n$, les fonctions symétriques élémentaires de f par :

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

De plus, on définit $e_0(z_1, \ldots, z_n) \coloneqq 1$.
Lemme 2.22. *Soit un polynôme complexe monique $f \in \mathbb{C}[x]$ de degré n et ses fonctions symétriques élémentaires $e_k : \mathbb{C}^n \to \mathbb{C}$. Alors pour tout $1 \leq k \leq n$, on a:*

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

Démonstration. Cette formule donne deux termes :

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

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. □

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 $e_k : \mathbb{C}^n \to \mathbb{C}$ 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 :

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

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 :

$$v + w \coloneqq (v_1 + w_1, v_2 + w_2, v_3 + w_3),$$
$$\lambda v \coloneqq (\lambda v_1, \lambda v_2, \lambda v_3).$$

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. $v - v \coloneqq v + (-v) = v + (-1)v = 0 = (0, 0, 0)$;
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.
$$P \in \Pi \iff \exists (\lambda, \mu) \in \mathbb{R}^2 \text{ t.q. } \begin{cases} x = \lambda v_1 + \mu v_2 + u_1 \\ y = \lambda v_2 + \mu v_2 + u_2 \\ z = \lambda v_3 + \mu w_3 + u_3 \end{cases}.$$

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:

$$D : \begin{cases} a_1 x + b_1 y + c_1 z = d_1 \\ a_2 x + b_2 y + c_2 z = d_2 \end{cases}.$$

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

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

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. □

Définition 3.9. Soient deux vecteurs $v, w \in \mathbb{R}$. 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 $d \coloneqq \langle v, a \rangle$. 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. □

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. $v_1 = k v_2$, 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. $a_1 = k a_2$).
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 :

$$\|v\| \coloneqq \sqrt{\langle v, v\rangle}.$$

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 :

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

□

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

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

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 $\left|\langle v, w\rangle\right| \leq \|v\|\|w\|$. □

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

Démonstration.

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

□

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 :

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

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 :*

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

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. □

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

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

3. *Soient une droite D de vecteur directeur $v \in \mathbb{R}^3$ et un plan π de vecteur normal $a \in \mathbb{R}^3$. Si $|D \cap \pi| = 1$, alors l'angle entre π et D est donné par :*
$$\alpha \coloneqq \frac{\pi}{2} - \arccos\left(\frac{\left|\langle a, v\rangle\right|}{\|a\|\|v\|}\right) \in \left]0, \frac{\pi}{2}\right].$$

3.6 Produit vectoriel

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

$$v \times w \coloneqq (v_2w_3 - v_3w_2, v_3w_1 - v_1w_3, v_1w_2 - v_2w_1).$$

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

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

Démonstration.

1. Pour la proposition 1, notons que :
$$\begin{aligned}(u + v) \times w &= ((u_2 + v_2)w_3 - w_2(u_3 + v_3), (u_3 + v_3)w_1 - w_3(u_1 + v_1), (u_1 + v_1)w_2 - w_1(u_2 + v_2)) \\ &= (u_2w_3 - w_2u_3 + v_2w_3 - w_2v_3, u_3w_1 - w_3u_1 + v_3w_1 - w_3v_1, u_1w_2 - w_1u_2 + v_1w_2 - w_1v_2) \\ &= (u_2w_3 - w_2u_3, u_3w_1 - w_3u_1, u_1w_2 - w_1u_2) + (v_2w_3 - w_2v_3, v_3w_1 - w_3v_1, v_1w_2 - w_1v_2) \\ &= (u \times w) + (v \times w).\end{aligned}$$

2. Pour la proposition 2, il suffit de mettre λ en évidence :
$$(\lambda v) \times w = (\lambda(v_2w_3 - w_2v_3), \lambda(v_3w_1 - w_3v_1), \lambda(v_1w_2 - w_1v_2)) = \lambda(v \times w).$$

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

4. Pour la proposition 4, on a:
$$v \times (\lambda v) = (v_2\lambda v_3 - v_3\lambda v_2, v_3\lambda v_1 - v_1\lambda v_3, v_1\lambda v_2 - v_2\lambda v_3) = (0, 0, 0) = 0.$$

□

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

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

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

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

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 :*

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

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 linéaires de s équations à t inconnues x_j et à coefficients réels $a_{ij}, b_i \in \mathbb{R}$ est :

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

Que l'on peut également noter :

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

Définition 4.2. Soit (S) un système d'équations linéaire. Si $b_i = 0$, alors la $i$eme 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-uples $(x_1, \ldots, x_t) \in \mathbb{R}^t$ 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 linéaires 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 $\mathcal{M}$ 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 $(\mathcal{M}_{\ell r})$ puis, pour toutes ligne i différentes de ℓ, soustraire à la ligne i la ligne ℓ multipliée par le coefficient $\mathcal{M}_{ir}$.

```
Soit M une matrice de dimensions l fois r.
Pour tout 0 < i < l+1
    Si M(i) != (0, 0, ..., 0)
        p = indice du premier coefficient non nul de M(i)
        M(i) /= M(i)(p)
        Pour tout 0 < j < l+1, j != 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 $a \coloneqq [a_{ij}] \in M_{s\times t}(\mathbb{R})$ une matrice quelconque de dimension $s \times t$. a_{ij} représente le coefficient de la $i$eme ligne et de la $j$eme colonne.

Théorème 4.10 (Unicité de la forme échelonée 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 $i_1$eme colonne de c soit vide. Or b est une combinaison linéaire de c et la première ligne de b contient à la $i_1$eme 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 $i_\gamma = j_\gamma$. □

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 4.11. 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 4.12. 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 :

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

Définition 4.13. 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 4.14. Soit $a = [a_{ij}] \in M_{s\times t}(\mathbb{R})$ une matrice. On définit la transformation linéaire qui lui est associée par :

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

Lemme 4.15. *Soit $(S) : ax + b$ un système linéaire d'équations et soit $A : \mathbb{R}^t \to \mathbb{R}^s$ 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 4.16 (Propriétés de la transformation linéaire associée)**.** *Soient $a = [a_{ij}] \in M_{s\times t}(\mathbb{R})$ une matrice et $A : \mathbb{R}^t \to \mathbb{R}^s$ 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 :

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

$$A(\lambda v) = a(\lambda v) = \left[\sum_{k=1}^{t} a_{ik}\lambda v_k\right] = \lambda\left[\sum_{k=1}^{t} a_{ik}v_k\right] = \lambda A(v).$$

□

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

Définition 4.18. Soit $(S) : ax = b$ et (S') son système homogène associé. Alors l'ensemble des solutions de (S') est appelé W_0.

Proposition 4.19. *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 $A(\lambda x + \mu y) = \lambda A(x) + \mu A(y) = \lambda 0 + \mu 0 = 0$. □

Lemme 4.20. *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 4.19, on sait que toutes les combinaisons linéaires de ces vecteurs sont toujours des solutions. □

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 4.21. *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, $A(x - y) = A(x) - A(y) = b - b = 0$. □

Corollaire 4.22. *Soit $\widetilde{x} \in \mathbb{R}^t$ une solution du système linéaire (S). Alors toute solution $y \in W$ de (S) est sous la forme suivante $\widetilde{x} + y$ où $y \in W_0$. Ou encore, l'ensemble des solutions de (S) est l'ensemble :*

$$W = \widetilde{x} + W_0 \coloneqq \{\widetilde{x} + y \ \ t.q. \ \ y \in W_0\}.$$

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

Soit $x \in W$. On sait que $A(x) = A(\widetilde{x}) = b$. Par le lemme 4.21, on sait que le vecteur $x - \widetilde{\in}\mathbb{R}^t$ est une solution du système homogène associé, donc $x - \widetilde{\in} W_0$. Dès lors, on peut réécrire $x = \widetilde{x} + (x - \widetilde{x})$. On peut donc dire $x \in \widetilde{x} + W_0$, ou encore $W \subseteq \widetilde{x} + W_0$. Donc $W = \widetilde{x} + W_0$. □

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

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

où les vecteurs $\widetilde{x}, v_i$ sont donnés par l'algorithme de Gauss (le vecteur $\widetilde{x}$ est le vecteur correspondant à la $(t+1)$eme 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 $\widetilde{x}$. 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. $+ : V \times V \to V : (v, w) \mapsto v + w$;
2. $\cdot : \mathbb{K} \times V \to V : (\lambda, v) \mapsto \lambda v$.

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. $\lambda(v + w) = \lambda v + \lambda w$;

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

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 $f : V \to W$ est une application linéaire si pour tout $v, w \in V, \lambda, \mu \in \mathbb{R}$, on a: $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$.

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

1. $\lambda 0 = 0$;
2. $0v = 0$;
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 $A : V \to W$ une application linéaire. On définit :

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

5.2 Sous-espaces

Théorème 5.7. *Soient V, W deux espaces vectoriels et $A : V \to W$ 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 $A : V \to W$ une application linéaire. Alors* $\mathrm{Ker}(A)$ *est un sous-espace de W et* $\mathrm{Im}(A)$ *est un sous-espace de V.*

Démonstration. V est un sous-espace de lui-même et $\{0\}$ est un sous-espace de W. Donc $\mathrm{Ker}(A) = A^{-1}(\{0\})$ est un sous-espace de V et $\mathrm{Im}(A) = A(V)$ 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 :

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

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 $v_1 + \lambda v_2 \in W_1 + W_2$. □

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 :

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

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

Démonstration. Soient V un e.v. et $\{W_i\}_{i \in I}$ une famille de sous-espaces. Soient $w_1, w_2 \in \cap_{i \in I} W_i$ 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, $w_1 + \lambda w_2 \in \cap_{i \in I} W_i$. □

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 :

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

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

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 $\langle X\rangle$. □

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 $i \in \{1, \ldots, k\}$ les vecteurs de Y. On sait par hypothèse que pour $i \in \{1, \ldots, n\}$, 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 $i \in \{1, \ldots, k-n\}$, 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 $\big\langle \langle X\rangle \big\rangle = \langle X\rangle$.*

Démonstration. On sait que

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

En posant :

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

on peut réécrire :

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

□

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 $x \in \left\langle X \setminus \{x\} \right\rangle$, 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 $x \in X : x \notin \left\langle X \setminus \{x\} \right\rangle$.*

Démonstration. Montrons que si X est une partie libre, alors $\forall x \in X : x \notin \left\langle X \setminus \{x\} \right\rangle$. Soit X une partie libre de V. Supposons par l'absurde qu'il existe $x \in X$ tel que $x \notin \left\langle X \setminus \{x\} \right\rangle$. Dès lors, il existe $v_1, \ldots, v_k \in X \setminus \{x\}$ et $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ tels que $x = \lambda_1 v_1 + \ldots + \lambda_k v_k$. La combinaison linéaire $x - \lambda_1 v_1 - \ldots - \lambda_k v_k = 0$ 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 $\lambda_1 v_1 + \ldots + \lambda_k v_k = 0$ 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 $\lambda_1 = \lambda_2 = \ldots = 0$. 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 $t \in \{n+1, \ldots, k\}$. 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 $v_1, \ldots, v_k \in X, \lambda, \lambda_1, \ldots, \lambda_k \in \mathbb{R}$ 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 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 $\mathcal{S} \coloneqq \{Z \subseteq V \text{ t.q. } X \subseteq Z \subseteq Y, Z \text{ partie libre de } V\}$. 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 $Z \coloneqq \cup_{i\in I} Z_i$. 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 $X \subseteq B \subseteq Y$. □

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 $\langle V \rangle = V$. □

Lemme 5.32. *Soi 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 :*

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

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 $Z = \langle W \cup \{v\} \rangle$. □

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 $f : X \to Y$. Et on note $|X| = |Y|$. S'il existe une fonction injective $f : X \to Y$, 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 $f : X \to Y, g : Y \to X$ deux fonctions injectives. Si g est bijective, la preuve est finie. Posons $A_0 \coloneqq X \setminus g(Y)$. On sait A_0 non-vide car g n'est pas bijective. On pose ensuite :

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

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

On sait :

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

Définissons ensuite :

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

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 on a 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 $B \coloneqq g(Y)$. Donc $g^{-1} : g(Y) \to Y$ 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 $\mathcal{F}_X \coloneqq \{Y \subseteq X \text{ t.q. } Y \text{ est une partie finie}\}$. 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 $|E| = \left|E'\right|$.*

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

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

<u>pas de récurrence :</u> si $n - k = i > 0$, on sait que $E \cap E' \subseteq E$ et donc si $E = \{e_1, \ldots, e_n\}$, on a $E \cap E' = \{e_1, \ldots, e_k\}$. 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 $y \in E' \setminus \left\langle E \setminus \{e_{k+1}\} \right\rangle$. Soit $E'' \coloneqq \{y\} \cup E \setminus \{e_{k+1}\}$. On sait donc dire :

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

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

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

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 $f : \mathcal{F}_E \to \mathcal{F}_{E'} : Y \mapsto E' \cap \langle Y \rangle$. 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 $\mathcal{F}_E$:

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

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

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

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

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-uple de λ_i suivant $(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$ 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 $\lambda_i = \mu_i$. □

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 $\langle E_W \cup \{v_1, v_2\} \rangle$, il existe v_3 etc. Une fois $k = \dim_{\mathbb{R}} V - \dim_{\mathbb{R}} W$ vecteurs ajoutés, on obtient $\langle E_W \cup \{v_1, \ldots, v_k\} \rangle = V$ où $E_W \cup \{v_1, \ldots, v_k\}$ 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 $E_V \setminus \{v_1, \ldots, v_k\}$ est libre dans W et est restée génératrice. Donc on a réduit E_V pour en faire une base de W. □

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). □

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 $X = \{x_1, \ldots, x_k\}$ 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} sait 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 $X = \{x_1, \ldots, x_k\}$ 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. □

6 Relations entre espaces vectoriels

6.1 Isomorphismes

Définition 6.1. Soient V, W deux e.v. Soit $A : V \to W$ 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 $I : V \to W$.
Lemme 6.3. *Soit $A : V \to W$ 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 $A^{-1} : W \to V$ 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 $\phi : V \to \mathbb{R}^d$ telle que si $v = \sum_{i=1}^{d} \lambda_i f_i$, alors on définit :

$$f(v) \coloneqq (\lambda_i)_{i \in \{1,\ldots,d\}}.$$

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 :

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

De plus, f est bijective car elle est injective :

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

et elle est surjective :

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

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 $f : V \to \mathbb{R}^d$ et $g : W \to \mathbb{R}^d$ 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 $(g^{-1} \circ f) : V \to W$ 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 $B \coloneqq B_V \cup B_W = \{e_1, \dots, e_r, x_1, \dots, x_n, y_1, \dots, y_t\}$. 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 :

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

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 :

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

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 $\dim_{\mathbb{R}}(V+W) + \dim_{\mathbb{R}}(V \cap W) = r + n + t + r = (r+n) + (r+t) = \dim_{\mathbb{R}} V + \dim_{\mathbb{R}} W$. □

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

Démonstration. Soient $B_{\mathrm{Ker}}\{e_1, \dots, e_m\}$ une base du $\operatorname{Ker}(A)$ et $B_{\mathrm{Im}} = \{f_1, \dots, f_n\}$ une base de $\operatorname{Im}(A)$. Soient $v_1, \dots, v_n \in V$ tels que $A(v_i) = f_i$ pour tout $1 \leq i \leq n$. Montrons que $B \coloneqq B_{\mathrm{Ker}} \cup \{v, \dots, v_n\}$ 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 $v - \sum_{i=1} \mu_i v_i \in \operatorname{Ker}(A)$. Or, tout vecteur de $\operatorname{Ker}(A)$ 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 $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n$ tels que :

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

Alors, on peut écrire :

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

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 $\dim_{\mathbb{R}} V = |B| = n + m = \dim_{\mathbb{R}} \mathrm{Im}(A) + \dim_{\mathbb{R}} \ker(A)$. □

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

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

Démonstration.

1. Montrons que si $\dim_{\mathbb{R}} \mathrm{Ker}(A) = 0$, alors A est injective. Si $\dim_{\mathbb{R}} \mathrm{Ker}(A) = 0$, alors $\mathrm{Ker}(A) = \{0_V\}$. 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 $\dim_{\mathbb{R}} \mathrm{Ker}(A) = 0$. 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 si $v \in V$, $f(v) = f(0) = 0$, alors $v = 0$. Autrement dit, seul 0_V est envoyé sur 0_W. On a bien $\dim_{\mathbb{R}} \mathrm{Ker}(A) = 0$.

2. Montons que si $\dim_{\mathbb{R}} \mathrm{Im}(A) = \dim_{\mathbb{R}} W$, alors A est surjective. Soit $E\{e_1, \ldots, e_m\}$ une base de W. Soit $F = \{f_1, \ldots, f_n\}$ une base de $\mathrm{Im}(A)$. Soient $v_1, \ldots, v_m$ tels que $A(v_i) = f_i$. Par la proposition 6.5, on sait que $\mathrm{Im}(A)$ 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 $\dim_{\mathbb{R}} \mathrm{Im}(A) = \dim_{\mathbb{R}} W$. Soit $w \in W$. On sait qu'il existe $v \in V$ tel que $A(v) = w$. Prenons $E = \{e_1, \ldots, e_m\}$ une base de W. Prenons $v_1, \ldots, v_m \in V$ tels que $A(v_i) = e_i$. On a $\{A(v_1), \ldots, A(v_m)\}$ est une base de $\mathrm{Im}(A)$ 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 $\dim_{\mathbb{R}} \operatorname{Ker}(A) = 0$ (par l'injectivité de A), alors $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} W$. Par la proposition 6.7, on sait $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} \operatorname{Ker}(A) + \dim_{\mathbb{R}} \operatorname{Im}(A)$. Donc si $\dim_{\mathbb{R}} \operatorname{Ker}(A) = 0$, on a $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} W$.

   Montrons maintenant que si $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} W$ et $\dim_{\mathbb{R}} \operatorname{Ker}(A) = 0$, 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 $\dim_{\mathbb{R}} \operatorname{Ker}(A) = 0$ car seul 0_V est envoyé sur 0_W.

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

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

□

6.2 Somme directe

Définition 6.9. Soient V et W deux espaces vectoriels. On définit la somme directe de V et W par $V \oplus W \coloneqq V \times W$ muni de l'addition interne et de la multiplication externe triviales 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 $A : W_1 \oplus W_2 \to W_1 + W_2 : (v, w) \mapsto v + w$. 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, $\operatorname{Ker}(A) = \{(w_1, w_2) \in W_1 \oplus W_2 \text{ t.q. } w_1 + w_2 = 0\}$. 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 $\dim_{\mathbb{R}} \operatorname{Ker}(A) = 0$. 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 $A : V \to W$ telle que $A(e_i) = f_i$ pour tout $1 \leq i \leq d$.*

Démonstration. Soit $g : E \to F : e_i \mapsto f_i$. On définit $A_g : V \to W : v = \sum_{i=1}^{d} \lambda_i e_i \mapsto w = \sum_{i=1}^{d} \lambda_i g(e_i)$. 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 $A : V \to W$ une application linéaire. Alors on appelle $\dim_{\mathbb{R}} \operatorname{Im}(A)$ le rang de A et on le note $\operatorname{rang} A$.

6.3 Groupe linéaire

Définition 6.14. Soit V un espace vectoriel réel. Le groupe linéaire (général) est l'ensemble des isomorphismes $\{A : V \to V \text{ t.q. } A \text{ est un automorphisme}\}$. On le note $GL(V)$. Si $V = \mathbb{R}^d$, on note $GL_d(\mathbb{R})$ ou $GL(d, \mathbb{R})$.
Remarque. La composée de deux automorphismes est toujours un automorphisme car la composée de bijections reste bijective et si $f : V \to V$ et $g : V \to V$, alors $(f \circ g) : V \to V$ également.
Proposition 6.15. *Soient deux espaces vectoriels V et W. Soit $E = \{e_1, \dots, e_d\}$. L'application linéaire $A : V \to W$ 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 $v \in W$, il existe $w \in W$ 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,a lors A est un isomorphisme. Puisque $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} W$ et $\mathrm{Im}(A) = W$, 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 $\mathrm{Hom}(V, W) \coloneqq \{A : V \to W \text{ t.q. } A \text{ transformation linéaire}\}$.
Proposition 7.2. *Soient V, W deux e.v. En munissant $\mathrm{Hom}(V, W)$ de l'addition interne*

$$(A + B)(v) \coloneqq A(v) + B(v),$$

et de la multiplication externe

$$(\lambda A)(v) \coloneqq \lambda(A(v)),$$

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

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

Définition 7.3. Soit $W = \mathrm{Hom}(V, V)$ 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, $A \in \mathrm{Hom}(V, W)$. Si $E = \{e_1, \ldots, e_t\}, F = \{f_1, \ldots, f_s\}$ sont une base de V et W respectivement, alors on définit $m_{F,E}(A) \coloneqq [A_{ij}] \in M_{s\times t}(\mathbb{R})$ 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 $A, B \in \mathrm{Hom}(V, W)$, et $m_{F,E}(A), m_{F,E}(B)$ les matrices associées à A et B. Alors :*

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

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 $m_{F,E} : \mathrm{Hom}(V, W) \to M_{\dim_{\mathbb{R}} W \times \dim_{\mathbb{R}} V}(\mathbb{R})$: $A \mapsto m_{F,E}(A)$. Et la proposition 7.5 montre que cette fonction est linéaire.

Définition 7.6. Soient trois espaces vectoriels U, V, W de base respective $D = \{d_1, \ldots, d_s\}, E = \{e_1, \ldots, e_t\}, F = \{f_1, \ldots, f_u\}$. Soient $A \in \mathrm{Hom}(U, V)$ et $B \in \mathrm{Hom}(V, W)$. La composée $(B \circ A) \in \mathrm{Hom}(U, W)$. 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 :

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

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 $A : V \to W$ une transformation linéaire. Alors si $x \in V : A(x) = m_{F,E}(A)x$.*

Démonstration. Soient $x \in V$ et $E = \{e_1, \ldots, e_t\}$ une base de V. Soit $F = \{f_1, \ldots, f_u\}$ 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 $m_{F,E} : \mathrm{Hom}(V, W) \to M_{s\times t}(\mathbb{R}) : A \mapsto m_{F,E}(A)$ 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 $A \in \mathrm{Hom}(V, W)$ telle que $A(x) \coloneqq \sum_{i=1}^{s} \sum_{j=1}^{t} A_{ij} x_j$. □

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

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 $i$eme 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 $i$eme ligne par $\mu + 1$. Donc multiplier une ligne par un facteur α revient à faire :

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

Finalement, remplacer la $i$eme ligne par la $j$eme 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. $A^{-1} \coloneqq b_k \dots b_1$ est effectivement l'inverse à gauche car $A^{-1}A = b_k \dots b_1 A$ qui, par hypothèse, vaut I. □

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 $\dim_{\mathbb{R}} \operatorname{Ker}(\gamma) = 0$. De plus, $\gamma : \mathbb{R}^n \to \mathbb{R}^n$. 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 $(\gamma \circ \gamma^{-1}) = Id_{\mathbb{R}^n} = (\gamma^{-1} \circ \gamma)$. On peut donc écrire :

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

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. □

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 $f : X \to Y$ une fonction injective mais pas surjective. On peut trouver $g : f(X) \to X$ 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 $A \in \operatorname{Hom}(V, V)$ un automorphisme de V. On définit la matrice de changement de base de E vers F exprimée avec les coordonnées de E par $b \coloneqq [B_{ij}] \in M_{n\times n}(\mathbb{R})$, matrice associée à l'application linéaire $B \in \operatorname{Hom}(V, V)$ 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 $\mu_1, \dots, \mu_n$ 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 $A \in \mathrm{Hom}(V,V)$ un automorphisme de V. Soient $m_{E,E}(A) = [A_{ij}]$ est la matrice e 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 :

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

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

$$A(f_j) = \sum_{k=1}^{n} A'_{kj} f_k = \sum_{k=1}^{n} A'_{kj} B(e_k) = \sum_{k=1}^{n} A'_{kj} \sum_{i=1}^{n} B_{ik} e_i = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} B_{ik} A'_{kj} \right) e_i,$$

$$A(f_j) = A(B(e_j)) = A\left(\sum_{k=1}^{n} B_{kj} e_k \right) = \sum_{k=1}^{n} B_{kj} A(e_k) = \sum_{k=1}^{n} B_{kj} \sum_{i=1}^{n} A_{ik} e_i = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} A_{ik} B_{kj} \right) e_i.$$

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 $V^* \coloneqq \mathrm{Hom}(V, \mathbb{R})$. 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 $E = \{e_1, \ldots, e_n\}$ 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 $m_{F,E} : V^* \to M_{1\times n}(\mathbb{R}) \sim \mathbb{R}^n$ est donc un isomorphisme. □

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^*. □

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 :

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

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, $f = \sum_{i=1}^{n} f(e_i) e_i^*$:

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

□

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

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* $\dim_{\mathbb{R}} \operatorname{Ker}(f) = n - 1$

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

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 $\operatorname{Ker}(f)$.

Remarque. On sait que $f^{-1}(\lambda) = \{x \in V \text{ t.q. } f(x) = \lambda\}$. Soit $v \in V \setminus \operatorname{Ker}(f)$. Tout vecteur $w \in V$ peut s'écrire comme : $v = u + \mu v$ avec $u \in \operatorname{Ker}(f)$ 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 $V \setminus \operatorname{Ker}(f)$.

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 $V^{**} \coloneqq (V^*)^* = \operatorname{Hom}(V^*, \mathbb{R})$.

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

Démonstration. On définit une fonction $\alpha : V \to V^{**}$ telle que pour $v \in V$, $\alpha(v) \in \operatorname{Hom}(V^*, \mathbb{R})$ 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 $A : V \to W$ une application linéaire. On définit la transposée de A par $A^T : W^* \to V^* : f \mapsto (f \circ A)$.

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 $A^T \coloneqq m_{E^*,F^*}(A^T) = [A^T_{ij}]$ comme la matrice transposée de A.
Proposition 7.28. *Soit $A \in \mathrm{Hom}(V,W)$ 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 $A^T_{ij} = A_{ji}$.*

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 :

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

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

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 $A : V \to W : v \mapsto av$ la transformation linéaire associée à a. On définit le rang de la matrice a comme le rang de l'application A. Donc $\mathrm{rang}(a) = \mathrm{rang}(A) = \dim_{\mathbb{R}} \mathrm{Im}(A)$.
Proposition 7.31. *Soit $A : V \to W$ une transformation linéaire. Alors* $\mathrm{rang}(A) = \mathrm{rang}(A^T)$.

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 $[A_{ij}]_{1\le i,j\le r}$. De plus, la matrice est nulle sur $[A_{ij}]_{r+1\le i\le s, 1\le j\le r}$. Donc forcément, le range de la transposée doit être au moins égal au rang de A. On a donc $\mathrm{rang}(A) \le \mathrm{rang}(A^T)$. Par un raisonnement similaire, on trouve $\mathrm{rang}(A^T) \le \mathrm{rang}\left(\left(A^T\right)^T\right)$. Or on sait que $\left(A^T\right)^T = A$. Donc on a $\mathrm{rang}(A) \le \mathrm{rang}(A^T) \le \mathrm{rang}(A)$. Il faut alors $\mathrm{rang}(A) = \mathrm{rang}(A^T)$. □

8 Permutations

8.1 Définitions

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

$$\mathfrak{S}_n \coloneqq \left\{f : \{1,2,\dots,n\} \to \{1,2,\dots,n\} \text{ t.q. } f \text{ est bijective}\right\}.$$

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

4. le produit de cycles :

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

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 $\{1, 2, \dots, n\} \setminus \{\sigma(1), \sigma(2)\}$) :
- $\dots$;
- 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 $\sigma^k = \mathrm{Id}_{\mathfrak{S}_n}$.

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

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 :

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

□

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 $\tau_1, \dots, \tau_n$ des transpositions telles que $\sigma = \tau_1 \dots \tau_n$. Montrons que $\sigma\sigma^{-1} = \sigma^{-1}\sigma = \mathrm{Id}_{\mathfrak{S}_n}$.

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

$$\begin{aligned} \tau_1\tau_2\dots\tau_{n-1}\tau_n\tau_n\tau_{n-1}\dots\tau_2\tau_1 &= \tau_1\tau_2\dots\tau_{n-1}(\tau_n\tau_n)\tau_{n-1}\dots\tau_2\tau_1 = \tau_1\tau_2\dots\tau_{n-1}\,\mathrm{Id}_{\mathfrak{S}_n}\,\tau_{n-1}\dots\tau_2\tau_1 \\ &= \tau_1\tau_2\dots(\tau_{n-1}\tau_{n-1})\dots\tau 2\tau_1 = \tau_1\tau_2\dots\mathrm{Id}_{\mathfrak{S}_n}\dots\tau_2\tau_1 = \dots = \tau_1\tau_2\tau_2\tau_1 = \tau_1\,\mathrm{Id}_{\mathfrak{S}_n}\,\tau_1 \\ &= \mathrm{Id}_{\mathfrak{S}_n}\,. \end{aligned}$$

Par un raisonnement similaire, on montre que :

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

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 $\forall m \in \mathbb{N}^* : \sigma^{mk} = \mathrm{Id}_{\mathfrak{S}_n}$.*

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

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

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

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 $k_1, k_2, \ldots, k_n$. Le degré de σ est le plus petit commun multiple des longueurs des cycles. Donc* $\deg \sigma = LCM(k_1, k_2, \ldots, k_n)$.

Démonstration. On sait par la proposition 8.9 qu'un cycle mis à la puissance de sa longueur donne l'identité. Soit $K \coloneqq LCM(k_1, \ldots, k_n)$, 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 $i \in \{1, \ldots, n\}$, donc pour chaque cycle individuel, la puissance K donne l'identité. On a donc $\sigma^K = \mathrm{Id}_{\mathfrak{S}_n}$ 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. □

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 :

$$\tau\sigma\tau^{-1} \coloneqq \tau \circ \sigma \circ \tau - 1.$$

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

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

Démonstration. Montrons que $(\tau(a_1), \ldots, \tau(a_m))$ est une permutation de $\gamma \coloneqq \tau\sigma\tau^{-1}$. Soit $i \in \{1, \ldots, m\}$. 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 $\tau(c)$. □

8.2 Signe d'une permutation

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

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

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

où $N(\sigma)$ est le plus petit nombre d'inversions possible pour transformer $\mathrm{Id}_{\mathfrak{S}_n}$ en σ. Si $\mathrm{sign}(\sigma) = -1$, 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 $\tau_1, \ldots, \tau_m, \tau'_1, \ldots, \tau'_n$ telles que :*

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

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

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 :

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

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

$$\begin{aligned} \sigma = \tau_1 \ldots \tau_m \quad &= \tau'_1 \ldots \tau'_n \\ = \hat{\tau_1} \ldots \hat{\tau_{m'}} \quad &= \hat{\tau'_1} \ldots \hat{\tau'_{n'}}. \end{aligned}$$

Soient $\nu_1, \ldots, \nu_m$ et $\nu'_1, \ldots, \nu'_n$ le nombre de transpositions adjacentes de $t\hat{a}u_i$ et $\hat{\tau}'_i$. 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 $\sigma^{-1} = \hat{\tau_m} \ldots \hat{\tau_1}$. En composant σ et son inverse σ^{-1}, on obtient l'identité $\mathrm{Id}_{\mathfrak{S}_n}$.

Soit $N(\sigma)$ le nombre d'inversions de σ. En évaluant $N(\sigma) - m' \mod 2 \equiv N(\sigma\hat{\tau_k} \ldots \hat{\tau_1}) \mod 2 = N(\mathrm{Id}_{\mathfrak{S}_n}) \mod 2 = 0 \mod 2 = 0$, 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 $\big(N(\sigma) - n\big) - \big(N(\sigma) - m\big) = m - n$ 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. □

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

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

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

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

□