Cours d'algèbre linéaire, 2 ème année d'université.

Gérard Letac[1]

[1]Laboratoire de Statistique et Probabilités, Université Paul Sabatier, 31062, Toulouse, France

Ceci est le cours d'algèbre linéaire enseigné à Toulouse à un bon millier d'étudiants de 1996 à 2002, à raison de 24 heures dans le semestre. Un de ses principes est de n'utiliser des coordonnées ou une structure euclidienne qu'au moment où elles sont nécessaires et s'imposent après analyse. La seconde année d'université est d'une richesse extraordinaire : en maitriser les contenus vous équipe intellectuellement pour le reste de l'existence, et vous rend pratiquement apte à passer l'agrégation. Lisez les démonstrations pour trois raisons :

– Elles vous convaincront de la véracité des énoncés.
– Elles contiennent souvent des idées très originales.
– On est constamment amené à les imiter dans les exercices et les applications.

Ne sautez jamais une ligne, tout est essentiel. Partout où c'est possible, on mentionne des choses élémentaires hors programme : formule de Laplace sur $\det(A+B)$, matrices de Kac, de Hua ou d'Hoffmann, angles d'Euler, l'exponentielle d'un endomorphisme et sa différentielle, semi groupes de matrices stochastiques, le cochonnet monstrueux de l'exercice II 4.10, base de Schmidt du tétraèdre régulier, quaternions, simplicité de $\mathbf{SO(3)}$, ombres d'un cube, algèbres de von Neumann de dimension finie, inégalité de Marčenko Pastur, décomposition de Cholewsky pour les arbres, graphes de Dynkin. J'espère qu'aucun exercice ne laisse le lecteur indifférent. La moitié a été utilisée à l'oral du concours d'entrée à l'Ecole Polytechnique.

G.L.

Table des matières

Chapitre 1

Réduction des endomorphismes

I Représentation matricielle d'un vecteur et d'une application linéaire

On se fixe un corps K. Si $A = (a_{ij})_{1\leq i\leq p, 1\leq j\leq q}$ est une matrice à éléments dans K à p lignes et q colonnes, on note par A^T sa matrice *transposée*, c'est à dire la matrice $(b_{ij})_{1\leq i\leq q, 1\leq j\leq p}$ à q lignes et p colonnes définie par $b_{ij} = a_{ji}$. [1] Ce symbole est souvent utilisé pour écrire une matrice colonne dans un texte sous forme de transposée de matrice ligne, pour économiser le papier.

Si E est un espace vectoriel de dimension finie q sur K, et si $e = \{e_1, \dots, e_q\}$ est une base de E (c'est à dire une partie de E qui soit libre et génératrice), soit x un vecteur de E. Comme e est une base, il existe une suite unique $(x_1, \dots, x_q)$ d'éléments de K telle que

$$x = x_1 e_1 + \cdots + x_q e_q.$$

Nous allons noter par $[x]^e$ la matrice *colonne* $[x_1, \dots, x_n]^T$. On l'appelle la matrice *représentative du vecteur x dans la base e*. Cette notation rappelle que la matrice représentative de x dépend du vecteur x certes, mais aussi de la base dans laquelle on a choisi de le représenter. Le plus simple des espaces vectoriels de dimension q sur K est l'espace K^q des matrices colonnes sur K d'ordre q. Les vecteurs

$$e_1 = [1, 0, \dots, 0]^T,\ e_2 = [0, 1, \dots, 0]^T,\ \dots, e_p = [0, 0, \dots, 1]^T$$

forment la *base canonique* de K^q.

Soit ensuite un autre espace vectoriel F sur K de dimension finie p et soit $f = (f_1, \dots, f_p)$ une base de f. L'ensemble des applications linéaires de E vers F sera noté $L(E, F)$. Soit maintenant $a : E \to F$ un élément de $L(E, F)$. Notons par $A = [a]_e^f$ la matrice $(a_{ij})_{1\leq i\leq p, 1\leq j\leq q}$ *dont les q colonnes sont les* $[a(e_j)]^f$ $j = 1, \dots, q$, c'est à dire les composantes dans la base f de l'espace F d'arrivée des images par a de chacun des vecteurs de la base e de l'espace de départ E. La matrice $[a]_e^f$ est appelée matrice *représentative de l'application a dans la base de départ e et la base d'arrivée f*. Si on se donne une matrice A

[1] D'autres notent ${}^tA$ ou A^t, et les statisticiens notent A'.

arbitraire à coefficients dans K à p lignes et q colonnes (on dira une matrice (p, q) ou $p \times q$, et une matrice carrée d'ordre p pour une matrice (p, p)), on peut toujours l'interpréter comme un $[a]_e^f$ en prenant des espaces E et F sur K arbitraires et munis de bases arbitraires e et f. Cette remarque permet de montrer que $\dim L(E, F) = \dim E \times \dim F$. Le choix le plus simple pour cela serait $E = K^q$ et $F = K^p$ avec les bases e et f canoniques.

Dans le cas particulier où $E = F$ on note $L(E, E) = L(E)$. Les éléments de $L(E)$ s'appellent des *endomorphismes.* Si a est dans E et si on prend les mêmes bases $e = f$, la matrice $[a]_e^e$ est dit représentative de a dans la base e. Le plus célèbre des endomorphismes de E est l'identité, notée id_E et définie par $\mathrm{id}_E(x) = x$ pour tout x de E. Si e est une base quelconque de E on a $[\mathrm{id}_E]_e^e = I_q$ où I_q est la *matrice* identité d'ordre q. Attention, si on représente id_E dans des bases e et e' de E différentes, comme on est amené à le faire en cas de changement de base, alors il est toujours faux que $[\mathrm{id}_E]_e^{e'} = I_q$. Dans ce cas, si e est appelée l'*ancienne* base et si e' est appelée la *nouvelle* base, la matrice carrée $P = [\mathrm{id}_E]_{e'}^e$ est appelée *matrice de changement de base.* Elle est très facile à écrire : ses colonnes sont les composantes de la nouvelle base e' par rapport à l'ancienne base e. On dit parfois que deux matrices carrées A et B d'ordre q sur K sont *semblables* s'il existe une matrice carrée inversible P d'ordre q sur K telle que $B = P^{-1}AP$.

Résumons dans le théorème suivant tout ce qu'il faut savoir sur les notions ci dessus, et qui a été démontré en 1 ère année :

Théorème 1.1. Soient (E, e), (F, f) et (G, g) des espaces de dimension finie sur le corps K équipés de bases.

1. Si $a \in L(E, F)$ et $x \in E$ alors la matrice représentative de l'image de x par a dans la base d'arrivée f est
$$[a(x)]^f = [a]_e^f [x]^e. \tag{1.1}$$
2. Si $a \in L(E, F)$ et si $b \in L(F, G)$ alors la matrice représentative de l'application linéaire composée $b \circ a$ de E dans G avec pour base de départ e et base d'arrivée g est
$$[b \circ a]_e^g = [b]_f^g [a]_e^f. \tag{1.2}$$
3. Si e' et f' sont d'autres bases de E et F, soit $P = [\mathrm{id}_E]_{e'}^e$ et $Q = [\mathrm{id}_F]_{f'}^f$. Alors pour tout $x \in E$ on a
$$[x]^{e'} = [\mathrm{id}]_e^{e'} [\mathrm{id}_E]^e = P^{-1}[x]^e. \tag{1.3}$$
Ensuite, si $a \in L(E, F)$ on a
$$[a]_{e'}^{f'} = [\mathrm{id}_F]_f^{f'} [a]_e^f [\mathrm{id}_E]_{e'}^e = Q^{-1}[a]_e^f P. \tag{1.4}$$
En particulier, si $E = F$ et $e = f$ et $e' = f'$, si $P = [\mathrm{id}_E]_{e'}^e$ et si $a \in L(E)$ alors
$$[a]_{e'}^{e'} = [\mathrm{id}_E]_e^{e'} [a]_e^e [\mathrm{id}_E]_{e'}^e = P^{-1}[a]_e^e P. \tag{1.5}$$

Disons un mot des *matrices par blocs.* Si $p = p_1 + \cdots + p_n$ et $q = q_1 + \cdots + q_m$ on considère parfois les matrices de la forme
$$A = (A_{ij})_{1 \leq i \leq n, 1 \leq j \leq m}$$

où A_{ij} est elle même une matrice sur K à p_i lignes et q_j colonnes. Ceci se prête au calcul numérique des produits de matrices : si $r = r_1 + \cdots + r_l$ et $B = (B_{jk})$ où la matrice B_{jk} a q_j lignes et r_k colonnes, alors $C = AB = (C_{ik})$ où $C_{ik} = \sum_{j=1}^{q} A_{ij}B_{jk}$. Un résultat qui sert souvent donne la condition nécessaire et suffisante pour que une matrice carrée commute avec une matrice diagonale :

Proposition 1.2. Soit $\lambda_1, \ldots, \lambda_p$ des éléments distincts de K et soit

$$D = \mathrm{diag}(\lambda_1 I_{m_1}, \ldots, \lambda_p I_{m_p}).$$

On considère la matrice carrée par blocs $A = (A_{ij})_{1 \leq i,j \leq p}$, où A_{ij} a m_i lignes et m_j colonnes. Alors $DA = AD$ si et seulement si $A_{ij} = 0$ pour $i \neq j$, c'est à dire si A est diagonale par blocs :

$$A = \mathrm{diag}(A_{11}, \ldots, A_{pp}).$$

Démonstration : $DA - AD = ((\lambda_i - \lambda_j)A_{ij})_{1 \leq i,j \leq p} = 0$ entraîne $A_{ij} = 0$ pour $i \neq j$ puisque les λ sont distincts.

A la fin de cette première section, expliquons que nous traiterons dans ce cours les coordonnées des vecteurs non avec répugnance mais avec précaution. Dans la vie pratique, un vecteur est aussi souvent un objet géométrique (une force, une vitesse, où les coordonnées sont artificielles ou inutiles) qu'une suite de nombres (le triplet (poids, hauteur, âge) d'un individu). Beaucoup d'objets mathématiques familiers (polynômes, matrices) sont souvent considérés comme des membres d'un espace vectoriel, comme des vecteurs, donc. Même si ces derniers sont décrits par des nombres (les coefficients du polynôme ou de la matrice) et donc conduisant à une base plus naturelle que les autres (de l'espace des polynômes de degré $\leq n$, des matrices à p lignes et q colonnes) souvent cette base est mal adaptée au problème à traiter (exercice 1.1) voire assez inutilisable et conduisant à des calculs maladroits. Par exemple les matrices carrées symétriques d'ordre q ont une base naturelle à $q(q+1)/2$ éléments dont l'utilisation est très malaisée. Il est pathétique de voir des statisticiens lorsqu'ils ont à manipuler une transformation linéaire de cet espace des matrices symétriques (par exemple $X \mapsto AXA^T$ où A est une matrice carrée d'ordre q) ne pouvoir se la représenter par sa simple définition, de vouloir en expliciter la matrice représentative et pour finir de vouloir absolument écrire la matrice carrée symétrique X comme une matrice colonne de hauteur $q(q+1)/2$. Il est vrai qu'ils ont une excuse informatique : bien des logiciels recoivent la description d'une matrice comme une suite de nombres où les lignes de la matrice sont séparées par des points virgules dans la suite . Bien sûr les coordonnées sont utiles, mais elles le sont bien plus si une étude géométrique préliminaire est faite pour déterminer s'il existe une base plus naturelle que les autres dans laquelle les calculs seront plus simples.

Travailler géométriquement oblige à clarifier les concepts. C'est un gros effort : on se représente mentalement facilement un vecteur membre de l'espace vectoriel réel E, mais l'étape suivante : se représenter une forme linéaire $x \mapsto f(x)$ c'est à dire une application linéaire de E à valeurs dans $\mathbb{R}$ est important et plus difficile (L'ensemble E^* des formes linéaires sur E s'appelle le dual de E). Si on travaille uniquement en coordonnées, on percevra seulement $x \in E$ et $f \in E^*$ comme des suites $(x_1, \dots, x_q)$ et $(f_1, \dots, f_q)$ avec $f(x) = f_1x_1 + \cdots + f_qx_q$ et la différence entre E et son dual sera fortement gommée (C'est toutefois normal pour les espaces euclidiens et hermitiens des chapitres 2 et 3). Pour reprendre l'exemple de l'espace E des polynômes de degré $\leq n$ il n'est pas indispensable quand on considère les formes linéaires $f : P \mapsto P(5)$ ou $g : P \mapsto \int_{-1}^{3} P(x)dx$ de se les représenter par des suites. Faire un calcul où apparaissent simultanément des éléments de E et E^* est grandement simplifié si on garde conscience de l'espace auquel appartient l'objet manipulé, pour ne rien dire du grand nombre d'erreurs de calcul ainsi évitées. De même, il est important de se représenter une transformation linéaire d'un espace vectoriel E vers un autre géométriquement, et non seulement à l'aide d'une matrice. Ainsi équipés peut être pourrons nous mieux comprendre la physique théorique avec son cortège de groupes, d'algèbres, de spineurs, de calcul extérieur et de "calcul tensoriel".

Exercice 1.1. Soit n un entier > 0. Pour k entier tel que $0 \leq k \leq n$ on définit les fonctions sur $\mathbb{R}$ e_k et f_k par

$$e_k(x) = e^{(2k-n)x}, \quad f_k(x) = (\sinh x)^k (\cosh x)^{n-k}.$$

Soit E l'espace vectoriel sur $K = \mathbb{R}$ formé par les fonctions sur $\mathbb{R}$ de la forme $f = \sum_{k=0}^{n} c_k e_k$ avec $c_k \in \mathbb{R}$. Montrer que si $f(x) = 0$ pour tout $x \in \mathbb{R}$ alors $c_k = 0$ pour tout k (Méthode : Montrer qu'alors $P(X) = \sum_{k=0}^{n} c_k X^k$ est un polynôme de degré $\leq n$ qui a plus de n racines). En déduire que $e = (e_0, \dots, e_n)$ est une base de E. Quelle est la dimension de E? Montrer que $f_k \in E$. Montrer que $f = (f_0, \dots, f_n)$ est une base de E (Méthode : montrer que c'est une famille génératrice en utilisant la formule du binôme dans l'expression $e_k(x) = (\cosh x + \sinh x)^k (\cosh x - \sinh x)^{n-k}$). Soit a le procédé qui à tout f de E fait correspondre sa fonction dérivée : $a(f) = f'$. Montrer que a est un endomorphisme de E et calculer directement les matrices $[a]_e^e$ et $[a]_f^f$ (on ne demande pas la matrice de changement de base). Voir aussi l'exercice 3.3.

Exercice 1.2. Il est connu que si a, b, c, d sont dans K avec $a \neq 0$, alors

$$\begin{bmatrix} 1 & 0 \\ \frac{c}{a} & 1 \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & d - \frac{bc}{a} \end{bmatrix} \begin{bmatrix} 1 & \frac{b}{a} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Plus généralement, si A et D sont des matrices carrées sur K d'ordre p et q telles que A soit inversible, trouver les matrices B_1, C_1, D_1 telles que

$$\begin{bmatrix} I_p & 0 \\ C_1 & I_q \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & D_1 \end{bmatrix} \begin{bmatrix} I_p & B_1 \\ 0 & I_q \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Quel est l'inverse de la matrice $\begin{bmatrix} I_p & 0 \\ C_1 & I_q \end{bmatrix}$? Si D_1 est inversible, quel est l'inverse de $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$?

Exercice 1.3. Soit K un corps fixé et $(E, e), (F, f)$ des espaces sur K de dimensions q et p équipés de bases. Soit $a \in L(E, K)$ une forme linéaire sur E et $b \in F$ et $\varphi \in L(E, F)$ défini par $\varphi(x) = ba(x)$. Quel est le rang de φ? Calculer $M = [\varphi]_e^f$ en fonction de $[a]_e^1 = [a_1, \ldots, a_q]$ et $[b]^f = [b_1, \ldots, b_p]^T$. Si $e' = (e'_1, \ldots, e'_q)$ et $f' = (f'_1, \ldots, f'_p)$ sont des bases telles que $a(e'_j) = 0$ pour $j = 2, \ldots, q$ et $b = f'_1$, calculer $M_1 = [\varphi]_{e'}^{f'}$. Voir aussi les exercices 3.5 , 4.5 et 5.4.

Exercice 1.4. Soit $a \in L(E)$ où E est de dimension finie q sur un corps K. Soit $F \subset E$ un sous espace vectoriel de dimension k tel que que $a(F) \subset F$ (on dit aussi que F est *stable* par a). Montrer l'existence d'une base e de E tel que $[a]_e^e$ s'écrive par blocs ainsi :

$$[a]_e^e = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

avec A matrice carrée d'ordre k (Méthode : prendre une base $e_1, \ldots, e_k$ de F et la compléter en une base quelconque e de E).

II Déterminant et trace d'un endomorphisme.

Si A est une matrice *carrée* sur le corps K on note son déterminant par $\det A$ et on appelle la somme de ses éléments diagonaux la *trace* de A, notée $\operatorname{trace} A$.

Proposition 2.1. Si A et B sont des matrices (p, q) et (q, p) sur K alors $\operatorname{trace}(AB) = \operatorname{trace}(BA)$. Si A, B, C sont des matrices $(p, q), (r, p), (q, r)$ alors $\operatorname{trace}(ABC) = \operatorname{trace}(CAB)$.

Démonstration : Soit $A = (a_{ij})$ et $B = (b_{jk})$ alors $M = AB = (m_{ik})_{1 \leq i,k} \leq q$ satisfait

$$m_{ik} = \sum_{j=1}^{p} a_{ij} b_{jk}$$

et donc $\operatorname{trace}(AB) = \sum_{i=1}^{q} \sum_{j=1}^{p} a_{ij} b_{ji}$. Comme i et j sont des variables muettes, si on les échange on trouve que $\operatorname{trace}(AB) = \operatorname{trace}(BA)$. En remplacant dans cette égalité (A, B) par (AB, C) on a le second résultat.

Remarque : Si on a un produit de plus de deux matrices carrées, il est faux qu'on puisse permuter les facteurs sans changer la trace. Par exemple, si

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix},\ B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix},\ C = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

alors $\operatorname{trace}(ABC) = 3$ et $\operatorname{trace}(ACB) = 5$.

Proposition 2.2. Soit E un espace vectoriel de dimension finie sur K et soit e et e' des bases de E. Soit $a \in L(E)$. Alors

1. $\det[a]_e^e = \det[a]_{e'}^{e'}$,
2. $\operatorname{trace}[a]_e^e = \operatorname{trace}[a]_{e'}^{e'}$.

Démonstration : On applique (1.5). Si $A = [a]_e^e$ alors $\det(PAP^{-1}) = \det P \det A \det P^{-1} = \det A$, et par la Proposition 2.1 on a $\mathrm{trace}\,(PAP^{-1}) = \mathrm{trace}\,(P^{-1}PA) = \mathrm{trace}\,A$.

A cause de la Proposition 2.2, on appelle respectivement *déterminant et trace de l'endomorphisme* a de E le déterminant et la trace d'une matrice représentative $A = [a_e]^e$ dans une base *arbitraire*. La proposition a servi à montrer que cette définition est cohérente car ces nombres ne dépendent pas de la base choisie.

On complète cette section par une remarque sur les déterminants des matrices triangulaires par blocs :

Proposition 2.3. Soit A une matrice carrée d'ordre $p = p_1 + \cdots + p_n$ décomposée par blocs $A = (A_{ij})$ où A_{ij} est une matrice à p_i lignes et p_j colonnes. On suppose que A est *triangulaire supérieure* [2] par blocs, c'est à dire que $A_{ij} = 0$ si $i > j$. Alors

$$\det A = \det A_{11} \det A_{22} \dots \det A_{nn}.$$

Démonstration : On le montre d'abord pour $n = 2$. On a alors $A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$. D'autre part si $A = (a_{ij})$, si $\mathcal{S}_p$ est le groupe des permutations σ de p objets et si $\epsilon(\sigma)$ est la signature de σ, on a $\det A = \sum_{\sigma \in \mathcal{S}_p} \epsilon(\sigma) f(\sigma)$ avec

$$f(\sigma) = \prod_{i=1}^{p_1} a_{i\sigma(i)} \prod_{i=p_1+1}^{p} a_{i\sigma(i)}.$$

Si σ est tel que il existe $i > p_1$ avec $\sigma(i) \leq p_1$, alors $f(\sigma) = 0$, puisque $A_{21} = 0$. Donc, si $f(\sigma) \neq 0$ alors $\sigma(i) \leq p_1$ si $i \leq p_1$ et $\sigma(i) > p_1$ si $i > p_1$. C'est dire qu'alors σ est le produit d'une permutation σ_1 de $\{1, \dots, p_1\}$ et d'une permutation σ_2 de $\{p_1 + 1, \dots, p\}$. Comme $\epsilon(\sigma_1\sigma_2) = \epsilon(\sigma_1)\epsilon(\sigma_2)$ on arrive à

$$\det A = (\sum_{\sigma_1} \epsilon(\sigma_1) \prod_{i=1}^{p_1} a_{i\sigma_1(i)})(\sum_{\sigma_2} \epsilon(\sigma_2) \prod_{i=p_1+1}^{p} a_{i\sigma_2(i)}) = \det A_{11} \det A_{22}.$$

L'extension par récurrence au cas n quelconque se fait facilement : si

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ 0 & A_{22} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_{nn} \end{bmatrix}.$$

on applique le cas $n = 2$ à la décomposition en blocs de $A = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix}$ avec

$$B_{11} = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1,n-1} \\ 0 & A_{22} & \dots & A_{2,n-1} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_{n-1,n-1} \end{bmatrix}$$

[2] D'autres disent trigonale supérieure.

et $B_{22} = A_{nn}$. Comme l'hypothèse de récurrence donne $\det B_{11} = \det A_{11} \det A_{22} \dots \det A_{n-1,n-1}$ et que par le cas $n = 2$ on a $\det A = \det B_{11} \det B_{22}$, la récurrence est étendue et la proposition est montrée.

Voici maintenant une formule plus difficile que nous utiliserons à la Proposition 9.4 du chapitre 2. Il ne faut pas la confondre avec la formule de Cauchy-Binet de l'exercice 2.9

Théorème 2.4. * (Formule de Laplace) Soit $A = (a_{ij})$ et $B = (b_{ij})$ deux matrices carrées d'ordre q sur le même corps K. Si $T \subset \{1, \dots, q\}$ on note T' le complémentaire de T et on note A_T la matrice carrée d'ordre $k = |T|$ qui est la restriction de A à T, c'est à dire $A_T = (a_{ij})_{i,j \in T}$. Si T est vide on convient $\det A_T = 1$. Alors $\det(A + B)$ s'exprime comme une somme de 2^q termes ainsi :

$$\det(A + B) = \sum_{T \subset \{1,\dots,q\}} \det A_T \det B_{T'}.$$

Démonstration : Soit $D = \operatorname{diag}(\lambda_1, \dots, \lambda_q)$ et notons $C = DA + B = (c_{ij})$. On a donc $c_{ij} = \lambda_i a_{ij} + b_{ij}$. Par définition d'un déterminant, et en notant $\mathcal{S}_q$ le groupe des permutations σ de $\{1, \dots, q\}$ et $\epsilon(\sigma)$ la signature de σ, on sait que

$$\det(DA + B) = \sum_{\sigma \in \mathcal{S}_q} \epsilon(\sigma) \prod_{i=1}^{q} c_{i\sigma(i)} = \sum_{\sigma \in \mathcal{S}_q} \epsilon(\sigma) \prod_{i=1}^{q} (\lambda_i a_{i\sigma(i)} + b_{i\sigma(i)}). \tag{2.6}$$

On considère alors $\det(DA+B)$ comme un polynôme à q variables $(\lambda_1, \dots, \lambda_q)$. On observe que c'est un polynôme *affine*, c'est à dire que pour tout $j = 1, \dots, q$, si on le considère comme une fonction de λ_j seul, alors il est de la forme $\alpha\lambda_j + \beta$ où α et β sont des fonctions des autres λ : cela se voit avec (2.6). Par conséquent il est de la forme

$$\det(DA + B) = \sum_{T \subset \{1,\dots,q\}} c_T \prod_{j \in T} \lambda_j$$

où les c_T sont dans K. Nous allons montrer que $c_T = \det A_T \det B_{T'}$. En faisant alors $D = I_r$ cela établira la proposition.

Pour calculer c_T, sans perte de généralité on peut supposer $T = \{1, \dots, k\}$ et $\lambda_{k+1} = \dots = \lambda_q = 0$. La formule (2.6) devient

$$\det(DA + B) = \sum_{\sigma \in \mathcal{S}_q} \epsilon(\sigma) \prod_{i=1}^{k} (\lambda_i a_{i\sigma(i)} + b_{i\sigma(i)}) \prod_{i=k+1}^{q} b_{i\sigma(i)}.$$

Notons alors par $\mathcal{S}_{k,q}$ le sous groupe de $\mathcal{S}_q$ formé des σ tels que $\sigma(i) \leq k$ si et seulement si $i \leq k$. Il est clair que $\mathcal{S}_{k,q}$ est isomorphe à $\mathcal{S}_k \times \mathcal{S}_{q-k}$: si $\sigma \in \mathcal{S}_{k,q}$ on écrit $\sigma = (\sigma_1, \sigma_2)$ où σ_1 est dans $\mathcal{S}_k$ et est la restriction de σ à $\{1, \dots, k\}$, et où σ_2 est la restriction de σ à $\{k + 1, \dots, q\}$. Notez aussi qu'on a $\epsilon(\sigma) = \epsilon(\sigma_1)\epsilon(\sigma_2)$. Maintenant, le coefficient de $\lambda_1 \dots \lambda_k$ dans $\prod_{i=1}^{k} (\lambda_i a_{i\sigma(i)} + b_{i\sigma(i)}) \prod_{i=k+1}^{q} b_{i\sigma(i)}$ n'est pas nul si et seulement si σ est dans $\mathcal{S}_{k,q}$. Dans ces conditions il est égal à

$$\prod_{i=1}^{k} a_{i\sigma_1(i)} \prod_{i=k+1}^{k} b_{i\sigma_2(i)}.$$

Multiplions cette expression par $\epsilon(\sigma)=\epsilon(\sigma_1)\epsilon(\sigma_2)$ et sommons les résultats sur tous les $\sigma=(\sigma_1,\sigma_2)\in\mathcal{S}_{k,q}$. Cela donne le coefficient $c_{\{1,\ldots,k\}}$ de $\lambda_1\ldots\lambda_k$ cherché. Il est clair que le résultat est aussi

$$\left(\sum_{\sigma_1\in\mathcal{S}_k}\epsilon(\sigma_1)\prod_{i=1}^{k}a_{i\sigma_1(i)}\right)\left(\sum_{\sigma_2\in\mathcal{S}_{q-k}}\epsilon(\sigma_2)\prod_{i=k+1}^{q}b_{i\sigma_2(i)}\right)=\det A_{\{1,\ldots,k\}}\det B_{\{k+1,\ldots,q\}}$$

et le théorème est démontré.

Exercice 2.1. Si A et D sont carrées avec A inversible, utiliser l'exercice 1.2 et la Proposition 2.3 pour montrer que

$$\det\begin{bmatrix}A & B\\ C & D\end{bmatrix}=\det(A)\det(D-CA^{-1}B).$$

Application : si A et D sont de même ordre q, réelles ou complexes, et si $AC=CA$ montrer que

$$\det\begin{bmatrix}A & B\\ C & D\end{bmatrix}=\det(AD-CB).$$

Méthode : supposer d'abord A inversible, puis sinon montrer qu'il existe une suite de ϵ_n tendant vers 0 tels que $A_n=A+\epsilon_nI_q$ soit inversible et conclure par passage à la limite.

Exercice 2.2. Sur le corps K soit A une matrice (p,q), X et Y des matrices colonnes d'ordre p et q. Montrer que $X^TAY=\mathrm{trace}\,(AYX^T)$.

Exercice 2.3. Si E est de dimension finie, on note $F=L(E)$ et pour $a\in F$ on définit l'élément $\varphi_a\in L(F)$ par $b\mapsto\varphi_a(b)=ab-ba$. Montrer que $\det(\varphi_a)=0$ (Méthode : montrer à l'aide de la Proposition 2.1 que φ_a n'est pas surjectif en considérant la forme linéaire sur F définie par $b\mapsto\mathrm{trace}\,b$).

Exercice 2.4. Soit A et B des matrices carrées réelles d'ordre q. Montrer que

$$\det\begin{bmatrix}A & -B\\ B & A\end{bmatrix}=|\det(A+iB)|^2.$$

Méthode : calculer

$$\begin{bmatrix}I_q & iI_q\\ iI_q & I_q\end{bmatrix}\begin{bmatrix}A & -B\\ B & A\end{bmatrix}\begin{bmatrix}I_q & iI_q\\ iI_q & I_q\end{bmatrix}^{-1}.$$

Exercice 2.5. (Caractérisation de la trace) Soit n un entier ≥ 2 et soit E l'espace vectoriel sur K des matrices carrées à éléments dans K et d'ordre n. Soit $f:E\to K$ une forme linéaire sur E telle que de plus pour tous X et Y dans E on ait $f(XY)=f(YX)$. Dans la suite, pour $k=1,\ldots,n$ $P_k=(p_{ij})\in E$ est défini par $p_{ij}=0$ si $(i,j)\neq(k,k)$ et $p_{kk}=1$.

1. Soit $X=(x_{ij})\in E$. Montrer les affirmations suivantes : $P_1+\ldots+P_n=I_n$, $P_k^2=P_k$, $P_kXP_k=x_{kk}P_k$ et $f(XP_k)=f(P_kXP_k)=x_{kk}f(P_k)$.

2. Dans le cas particulier $n = 2$, trouver $U \in E$ inversible telle que $P_2U = UP_1$. Même question pour n quelconque. Enfin, si $k \geq 2$ trouver $U_k \in E$ inversible telle que $P_k = U_kP_1U_k^{-1}$.
3. On pose $c = f(P_1)$. Montrer à l'aide du 2) que $c = f(P_k)$ pour tout $k = 1, \dots, n$. En écrivant $X = X(P_1 + \dots + P_n)$ montrer à l'aide de la linéarité de f et du 1) que $f(X) = c\,\mathrm{trace}\,X$.

Exercice 2.6. Si J_q est la matrice carrée d'ordre q dont tous les coefficients sont 1 et si a et b sont réels, il existe bien des manières de montrer que $\det(aJ_q + bI_q) = b^q + qab^{q-1}$. Le montrer avec la formule de Laplace appliquée à $A = aJ_q$ et $B = bI_q$ et utilisant $\det A_T = 1, a, 0$ suivant que $|T| = 0, 1, \geq 2$.

Exercice 2.7. Soit $D = \mathrm{diag}(\mu_1, \dots, \mu_q)$ et soit J_q la matrice (q, q) dont tous les coefficients sont égaux à 1. On suppose $\det D \neq 0$. Montrer à l'aide de la formule de Laplace que

$$\det(D - J_q) = \mu_1 \cdots \mu_q \left(1 - \frac{1}{\mu_1} - \dots - \frac{1}{\mu_q}\right).$$

Exercice 2.8. On fixe un corps K. Soit $a = (a_1, \dots, a_q)^T$ et $b = (b_1, \dots, b_q)^T$ des matrices colonnes sur K d'ordre q avec $q \geq 2$. Si $S \subset \{1, \dots, q\}$ on note $a_S = \sum_{s \in S} a_s$. Soit $C = [c_1, \dots, c_q]$ une matrice carrée d'ordre q telle que $c_j \in \{a, b\}$ pour tout j. Soit enfin $S = \{j; c_j = a\}$ et $S' = \{j; c_j = b\}$. Montrer que si $\lambda \in K$ alors

$$\det(\lambda I_q + C) = \lambda^q + (a_S + b_{S'})\lambda^{q-1} + (a_S b_{S'} - a_{S'} b_S)\lambda^{q-2}.$$

Méthode : appliquer la formule de Laplace du Théorème 2.4 au couple $(C, \lambda I_q)$ et utiliser le fait que le rang de C est ≤ 2 pour écrire $\det(\lambda I_q + C) = \lambda^q + \lambda^{q-1}\mathrm{trace}\,C + \lambda^{q-2}\sum_{T; |T|=2} \det C_T$.

Exercice 2.9. (Formule de Cauchy-Binet) Soit $0 < q \leq m$ des entiers, soit $A = (a_{ij})$ une matrice à q lignes et m colonnes et soit $B = (b_{jk})$ une matrice à m lignes et q colonnes. Soit $\mathcal{T}$ l'ensemble des parties T de $\{1, \dots, m\}$ de taille q et, pour $T \in \mathcal{T}$ on considère les matrices carrées d'ordre q

$$A_T = (a_{ij})_{1 \leq i \leq q,\ j \in T}, \quad B_T = (b_{jk})_{j \in T,\ 1 \leq k \leq q}.$$

Montrer que

$$\det(AB) = \sum_{T \in \mathcal{T}} \det A_T \det B_T.$$

Méthode : introduire des variables $(\lambda_1, \dots, \lambda_m)$ et la matrice $D = \mathrm{diag}(\lambda_1, \dots, \lambda_m)$. Montrer que $(\lambda_1, \dots, \lambda_m) \mapsto \det ADB$ est un polynôme affine $\sum_{T \in \mathcal{T}} c_T \prod_{j \in T} \lambda_j$. Montrer que $c_T = \det A_T \det B_T$ en considérant sans perte de généralité le cas particulier $T = \{1, \dots, q\}$ et $\lambda_{q+1} = \dots = \lambda_m = 0$ et en écrivant les matrices A, D, B par blocs. Conclure en prenant $D = I_m$. Remarque : la formule de Cauchy Binet est parfois énoncée avec une apparence de plus grande généralité : on prend des entiers $0 < q \leq n, m, p$ et et des matrices $A = (a_{ij})$ à n

lignes et m colonnes et $B = (b_{jk})$ à m lignes et p colonnes. Soit S et R des parties de taille q de $\{1, n\}$ et $\{1, p\}$ respectivement. Alors, avec des notations évidentes

$$\det(AB)_{S\times R} = \sum_{T\in\mathcal{T}} \det A_{S\times T} \det B_{T\times R}.$$

III Espaces et valeurs propres d'un endomorphisme.

Définition 3.1 : Soit E un espace vectoriel de dimension finie sur K, soit $a \in L(E)$, soit λ dans K et soit $E_\lambda = \ker(a - \lambda \mathrm{id}_E)$. Si $E_\lambda \neq \{0\}$, on dit que λ est une *valeur propre* de a. Dans ce cas, E_λ est appelé l'*espace propre* associé à la valeur propre λ et les éléments *non nuls* de E_λ sont appelés des *vecteurs propres* associés à λ. L'ensemble Λ de toutes les valeurs propres de a est appelé le *spectre* de a.

Remarques : Les définitions précédentes peuvent être reformulées ainsi : λ est une valeur propre de a si et seulement si il existe un vecteur x de E non nul tel que $a(x) = \lambda x$. Dans ce cas un tel x est appelé un vecteur propre associé à λ, et l'espace propre associé à λ est l'ensemble des vecteurs propres associés à λ complété par le vecteur nul. En effet, il est clair que $\ker(a - \lambda \mathrm{id}_E) \neq \{0\} \Leftrightarrow \exists x \in E\ ;\ (a - \lambda \mathrm{id}_E)(x) = 0 \Leftrightarrow \exists x \in E\ ;\ a(x) = \lambda x$. En gros, un vecteur propre est un vecteur de E dont *la direction est conservée* après déformation par a. Une chose surprenante est que le coefficient de proportionalité, à savoir la valeur propre, ait au moins autant d'importance que le vecteur propre lui même.

Avant de donner des exemples et des propriétés des ces nouvelles notions, et d'expliquer en particulier comment les calculer, nous démontrons un théorème important. Rappelons avant que si $F_1, \ldots, F_p$ sont des sous espaces vectoriels de l'espace vectoriel E, on dit que la famille $\{F_1, \ldots, F_p\}$ est *en somme directe* si une suite de vecteurs $(x_1, \ldots, x_p)$ est telle que $x_j \in F_j$ pour $j = 1, \ldots p$ et telle que $x_1 + \cdots + x_p = 0$, alors ceci ne peut arriver *que* si $x_1 = \ldots = x_p = 0$. Dans ce cas le sous espace F de E égal à $F = F_1 + \cdots + F_p$ est appelé *la somme directe* des $F_1, \ldots, F_p$. On écrit alors traditionnellement

$$F = F_1 \oplus \cdots \oplus F_p,$$

Théorème 3.1. Soit E de dimension finie, $a \in L(E)$ et soient $\lambda_1, \ldots, \lambda_p$ des valeurs propres *distinctes* de a. Alors les espaces propres $E_{\lambda_1}, \ldots, E_{\lambda_p}$ sont en somme directe. En particulier, le spectre de a est fini et a au plus $\dim E$ éléments.

Démonstration : Celle ci va nous faire introduire deux objets très intéressants :

- Les polynômes de Lagrange;
- Les polynômes d'endomorphismes.

Si $\lambda_1, \ldots, \lambda_p$ sont des éléments distincts du corps K les p *polynômes de Lagrange* $L_1, \ldots, L_p$ associés sont définis par

$$L_j(X) = \frac{\prod_{i\neq j}(X-\lambda_i)}{\prod_{i\neq j}(\lambda_j-\lambda_i)}.$$

Ils ont trois propriétés remarquables : $\deg L_j = p-1$, $L_j(\lambda_j) = 1$ et, pour $i \neq j$, $L_j(\lambda_j) = 0$. Ils sont utilisés en analyse numérique pour résoudre le problème suivant : étant donné $(a_1, \ldots, a_p) \in K^p$ trouver l'unique polynôme P de degré $\leq p-1$ tel que pour tout j on ait $P(\lambda_j) = a_j$. Réponse : c'est $P = a_1 L_1 + \cdots + a_p L_p$: on laisse le lecteur vérifier que P convient et est unique.

Définissons maintenant les polynômes d'endomorphismes. Si $a \in E$, on définit la suite $(a^k)_{k\geq 0}$ d'éléments de $L(E)$ par la récurrence suivante : $a^0 = \mathrm{id}_E$, $a^{k+1} = a \circ a^k$. Notez qu'à cause de l'associativité de la composition des fonctions on a $a^{k+1} = a^k \circ a$ et plus généralement $a^{j+k} = a^j \circ a^k$. Si e est une base quelconque, alors (1.2) implique que si $A = [a]_e^e$ alors $[a^k]_e^e = A^k$.

Si $P(X) = c_0 + c_1 X + \cdots + c_n X^n$ est un polynôme dont les coefficients c_j sont dans K on définit alors l'endomorphisme $P(a)$ par

$$P(a) = c_0 \mathrm{id}_E + c_1 a + c_2 a^2 + \cdots + c_n a^n.$$

De même donc, si $A = [a]_e^e$ alors

$$[P(a)]_e^e = P(A) = c_0 I_q + c_1 A + c_2 A^2 + \cdots + c_n A^n.$$

Exemple : Matrices circulantes. Soit $P(X) = a_0 + a_1 X + \cdots + a_{q-1} X^{q-1}$ un polynôme sur le corps K. On considère aussi la matrice d'ordre q et ses puissances :

$$R = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \\ 1 & 0 & 0 & \ldots & 0 \end{bmatrix} \quad R^2 = \begin{bmatrix} 0 & 0 & 1 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 0 & 0 & \ldots & 1 \\ 0 & 1 & 0 & \ldots & 0 \end{bmatrix} \ \ldots$$

En d'autres termes si $E = K^q$ est muni de la base canonique $e = (e_1, \ldots, e_q)$ et si on considère $r \in L(E)$ tel que $r(e_j) = e_{j-1}$ si $j = 2, \ldots, q$ et $r(e_1) = e_q$ on a $R = [r]_e^e$. Alors

$$[P(r)]_e^e = P(R) = a_0 I_q + a_1 R + \cdots + a_{q-1} R^{q-1} = \begin{bmatrix} a_0 & a_1 & a_2 & \ldots & a_{q-1} \\ a_{q-1} & a_0 & a_1 & \ldots & a_{q-2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_2 & a_3 & a_4 & \ldots & a_1 \\ a_1 & a_2 & a_3 & \ldots & a_0 \end{bmatrix}. \quad (3.7)$$

$<$

Nous résumons les propriétés de la correspondance $P \mapsto P(a)$ par l'énoncé suivant, dont les démonstrations sont évidentes :

Proposition 3.2. Soit E de dimension finie sur K et a fixé dans $L(E)$. Soit $K[X]$ l'espace des polynômes à coefficients dans K. L'application de $K[X]$ dans $L(E)$ définie par $P \mapsto P(a)$ a les propriétés suivantes : pour tous polynômes P et Q on a

1. $(P+Q)(a) = P(a) + Q(a)$.
2. $(PQ)(a) = P(a) \circ Q(a) = Q(a) \circ P(a)$.

Nous procédons maintenant à la démonstration du Théorème 3.1. Soit pour $j = 1, \ldots, p$, les vecteurs $x_j \in E_{\lambda_j}$. On suppose que $x_1 + \cdots + x_p = 0$ et on veut montrer qu'alors $x_j = 0$ pour tout j. Puisque $a(x_j) = \lambda_j x_j$ on a, en appliquant a aux deux membres de cette égalité :

$$a^2(x_j) = a(a(x_j)) = a(\lambda_j x_j) = \lambda_j a(x_j) = \lambda_j^2 x_j.$$

En fait on a plus généralement $a^k(x_j) = \lambda_j^k x_j$, et plus généralement encore pour tout polynôme $P(a)(x_j) = P(\lambda_j)(x_j)$. Par conséquent :

$$0 = P(a)(0) = P(a)(x_1 + \cdots + x_p) = P(a)(x_1) + \cdots + P(a)(x_p) = P(\lambda_1)(x_1) + \cdots + P(\lambda_p)(x_p).$$

Prenons en particulier P égal au polynôme de Lagrange L_j : l'égalité ci dessus donne $x_j = 0$, le résultat voulu : les E_{λ_j} sont bien en somme directe. Pour finir, considérons le sous espace vectoriel de E suivant

$$F = E_{\lambda_1} \oplus \cdots E_{\lambda_p}.$$

Puisque tous les espaces propres sont de dimension strictement positive, on a

$$\dim E \geq \dim F = \dim E_{\lambda_1} + \cdots + \dim E_{\lambda_p} \geq p.$$

Ceci montre que le spectre est fini et de taille inférieure ou égal à la dimension de E, et la démonstration est complète.

Exercice 3.1. Soit E l'espace des polynômes sur $\mathbb{R}$ de degré $\leq n$. Quelle est sa dimension ? (chercher une base). Soit $a \in L(E)$ défini par $a(P)(X) = XP'(X)$. Trouver valeurs propres et vecteurs propres de a (Méthode : chercher λ pour que les solutions de l'équation différentielle $xy' = \lambda y$ soient des polynômes). Soit $b \in L(E)$ défini par $b(P)(X) = P(X+1) - P(X)$. Montrer que la seule valeur propre de b est 0 avec les polynômes constants pour espace propre (Méthode : observer que si $\deg P \geq 1$ alors $\deg b(P) = (\deg P) - 1$). En déduire les valeurs et espaces propres de $c = b + \mathrm{id}_E$, satisfaisant donc $c(P)(X) = P(X+1)$.

Exercice 3.2. Soit E l'espace sur $\mathbb{C}$ des suites complexes $z = (z_n)_{n \in \mathbf{Z}}$ qui sont de période 3 (c'est à dire $z_{n+3} = z_n$ pour tout $n \in \mathbf{Z}$). de degré $\leq n$. Quelle est sa dimension ? (chercher une base). On définit $a \in L(E)$ par $a(z)_n = z_{n+1}$. Montrer que les trois racines cubiques de l'unité $1, j, j^2$ sont les valeurs propres de a.

Exercice 3.3. Soit n un entier > 0. Utiliser l'exercice 1.1 pour calculer les valeurs propres de la matrice réelle d'ordre $n+1$ suivante :

$$\begin{bmatrix} 0 & n & 0 & 0 & \ldots & 0 & 0 & 0 \\ 1 & 0 & n-1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 2 & 0 & n-2 & \ldots & 0 & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & n-1 & 0 & 1 \\ 0 & 0 & 0 & 0 & \ldots & 0 & n & 0 \end{bmatrix}.$$

Exercice 3.4. Soit $\mathcal{M}$ l'espace vectoriel sur $\mathbb{R}$ des matrices carrées réelles d'ordre q, soit $\mathcal{S}$ le sous espace des matrices symétriques, c'est à dire des matrices A telles que $A = A^T$ et soit $\mathcal{A}$ le sous espace des matrices antisymétriques, c'est à dire des matrices A telles que $A = -A^T$. Quelles sont les dimensions respectives de $\mathcal{M}$, $\mathcal{S}$ et $\mathcal{A}$? Montrer que $\mathcal{S}$ et $\mathcal{A}$ sont en somme directe. Montrer que $\mathcal{M} = \mathcal{S} \oplus \mathcal{A}$ soit par un argument de dimension, soit en écrivant

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T).$$

On considère enfin l'endomorphisme a de $\mathcal{M}$ défini par $a(A) = A^T$. Utiliser ce qui précède pour trouver ses espaces et valeurs propres. Voir aussi l'exercice 5.2.

Exercice 3.5. Dans l'exercice 1.3 on particularise à $E = F$. On suppose $a(b) \neq 0$. Montrer que $a(b)$ et 0 sont les seules valeurs propres de φ, et trouver les espaces propres correspondants en précisant leurs dimensions (Méthode : considérer $\ker a$.) On suppose $a(b) = 0$: trouver valeurs propres et espaces propres. Voir aussi les exercices 4.5 et 5.4.

Exercice 3.6. Soit E et F des espaces vectoriels de dimension finie sur un même corps. Soit $a \in L(E,F)$ et $b \in L(F,E)$. (1) On suppose que $(\mathrm{id}_F - ab)^{-1}$ existe. Montrer que $(\mathrm{id}_E - ba)^{-1}$ existe en montrant $(\mathrm{id}_E - ba)^{-1} = \mathrm{id}_E + b(\mathrm{id}_F - ab)^{-1}a$. (2) En déduire que les valeurs propres non nulles de ab et de ba sont les mêmes (considérer les $\lambda \neq 0$ tels que $(\lambda\mathrm{id}_F - ab)^{-1} = \lambda^{-1}(\mathrm{id}_F - \lambda^{-1}ab)^{-1}$ existe). Voir aussi les exercices 4.2 et 4.3.

IV Cayley Hamilton et polynôme caractéristique

Pour éviter quelques subtilités inintéressantes dans le maniement des polynômes nous supposons désormais que le corps K a un nombre infini d'éléments. Ceci permet d'identifier la *fonction* sur K qui à $\lambda \in K$ fait correspondre l'élément de K égal à $P(\lambda) = c_0 + c_1\lambda + \cdots + c_n\lambda^n$ avec le *polynôme* lui même, c'est à dire à la suite $(c_0, \ldots, c_n)$ de ses coefficients. Si le corps est fini, plusieurs polynômes donnent naissance à la même fonction : par exemple avec le corps $\{0,1\}$ des informaticiens on a $1 + \lambda + \lambda^2 = 1$ pour $\lambda = 0$ et 1. Ceci ne peut arriver avec un corps K infini : $P(\lambda) = Q(\lambda)$ pour tout $\lambda \in K$ entraîne que $P - Q$ a une infinité de racines et est donc le polynôme dont tous les coefficients sont nuls.

Définition 4.1 : Soit E un espace vectoriel de dimension finie sur K et soit $a \in L(E)$. Le *polynôme caractéristique* de a est $P_a(X) = \det(a - X\mathrm{id}_E)$. Si A est une matrice carrée d'ordre q, le polynôme caractéristique de A est $P_A(X) = \det(A - XI_q)$.

Pour calculer P_a, le moyen le plus simple est de prendre une base e et de considérer la matrice $[a]_e^e = A = (a_{ij})_{1 \leq i,j \leq q}$. On a alors

$$P_a(X) = \det \begin{bmatrix} a_{11} - X & a_{12} & \ldots & a_{1q} \\ a_{21} & a_{22} - X & \ldots & a_{2q} \\ \ldots & \ldots & \ldots & \ldots \\ a_{q1} & a_{q2} & \ldots & a_{qq} - X \end{bmatrix}, \tag{4.8}$$

ce qui montre que P_a est un polynôme.

Proposition 4.1. Soit E de dimension finie q et a dans $L(E)$. Alors

1. λ est valeur propre de a si et seulement si $P_a(\lambda) = 0$.
2. P_a est de degré q et $P_a(X) = (-1)^q X^q + (-1)^{q-1}\mathrm{trace}\,(a) X^{q-1} + \cdots + \det a$. En particulier, si $q = 2$ on a $P_a(X) = X^2 - X\mathrm{trace}\, a + \det a$.
3. Si m est la multiplicité de la valeur propre λ dans P_a, alors $\dim E_\lambda \leq m$.

Démonstration : (1) La première partie est basée sur le fait suivant vu en première année : si $b \in L(E)$ alors $\det b = 0 \Leftrightarrow \ker b \neq \{0\}$. Il suffit d'appliquer cela à $b = a - \lambda \mathrm{id}_E$.
(2) Pour la seconde on utilise (4.8) et on applique la définition du déterminant à $A - XI_q = B = (b_{ij})$. On obtient si $\mathcal{S}_q$ est le groupe des permutations de q objets et si $\epsilon(\sigma)$ est la signature de σ :

$$P_a(X) = \det B = \sum_{\sigma \in \mathcal{S}_q} \epsilon(\sigma) \prod_{i=1}^{q} b_{i\sigma(i)}.$$

Comme les b_{ij} sont des polynômes de degré ≤ 1 en X ceci montre que $\deg P_a \leq q$. Son terme constant est $P_a(0) = \det a$. Pour les termes de degré q et $q-1$, on observe que si σ n'est pas la permutation identique, elle déplace au moins *deux* objets et donc $\deg \prod_{i=1}^{q} b_{i\sigma(i)} \leq q-2$. La contribution aux termes de degré q et $q-1$ vient donc seulement de $\prod_{i=1}^{q} b_{ii} = \prod_{i=1}^{q}(a_{ii} - X)$ et est donc celle de l'énoncé.
(3) On prend une base e de E telle que $(e_1, \ldots, e_k)$ soit une base de E_λ. Alors la matrice $[a]_e^e$ s'écrit par blocs

$$[a]_e^e = \begin{bmatrix} \lambda I_k & A \\ 0 & B \end{bmatrix}$$

et donc $P_a(X) = (\lambda - X)^k \det(B - XI_{q-k}$, ce qui montre que $k \leq m$. La proposition est montrée.

Voici alors la surprenante *propriété de Cayley Hamilton* de P_a : c'est que $P_a(a) = 0$. Si e est une base de E, si $A = [a]_e^e$ et si $P_a(X) = c_0 + c_1 X + \cdots + (-1)^q X^q$, c'est dire que la matrice carrée d'ordre q définie par $c_0 I_q + c_1 A + \cdots + (-1)^q A^q$ est la matrice nulle. Ainsi si $A = \begin{bmatrix} 5 & 2 \\ -2 & -1 \end{bmatrix}$ alors trace $A = 4$, $\det A = -1$ et donc (d'après la Proposition 4.1 (2)) $P_a(X) = X^2 - 4X - 1$. Comme $A^2 = \begin{bmatrix} 21 & 8 \\ -8 & -3 \end{bmatrix}$ on a bien

$$A^2 - 4A - I_2 = \begin{bmatrix} 21 & 8 \\ -8 & -3 \end{bmatrix} - 4 \begin{bmatrix} 5 & 2 \\ -2 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Notons qu'on en tire pour les matrices d'ordre 2 un moyen rapide de calculer les premières puissances : si $P_a(X) = X^2 - tX + d$, alors

$$A^2 = tA - dI_2,\ A^3 = tA^2 - dA = (t^2 - d)A - tdI_2,\ A^4 = (t^2 - 2dt)A - (t^2 d - d^2)I_4.$$

Ce résultat est vrai pour un corps fini ou infini. Mais la démonstration que nous en donnerons n'est valable que pour un corps infini. Enoncons cette propriété de Cayley

Hamilton dans un vocabulaire légèrement différent qu'on aura à utiliser plus tard : si $a \in L(E)$ on dit que le polynôme P sur K est un *polynôme annulateur de* a si $P(a) = 0$.

Théorème 4.2. (Th. de Cayley-Hamilton) Soit $a \in L(E)$ où E est un espace de dimension finie sur le corps K. Le polynôme caractéristique de a est un polynôme annulateur de a.

Démonstration : La démonstration est basée sur le résultat acquis en première année : si M est une matrice carrée et si $C(M)$ est la matrice de ses cofacteurs, alors $(C(M))^T M = (\det M) I_q$. Soit e une base de E et $A = [a]_e^e$. Soit B la matrice transposée de la matrice des *cofacteurs* de la matrice $M = A - XI_q$. Chacun de ces cofacteurs est un polynôme en λ de degré au plus $q-1$. En regroupant les facteurs de la puissance X^k en une matrice B_k, la matrice B s'écrit

$$B = B_0 + XB_1 + \cdots + B_{q-1}X^{q-1} \tag{4.9}$$

Si $P_a(X) = c_0 + c_1 X + \cdots + c_q X^q$ et puisque $B(A - XI_q) = P_a(X) I_q$ on obtient

$$\begin{aligned} B_0 A + X(B_1 A - B_0) + X^2(B_2 A - B_1) + \cdots + X^{q-1}(B_{q-1}A - B_{q-2}) - X^q B_{q-1} \\ = c_0 I_q + c_1 X I_q + \cdots + c_q X^q I_q. \end{aligned}$$

On identifie alors les coefficients des X^k dans chaque membre (c'est ici qu'on utilise le fait que le corps est infini) et on obtient

$$\begin{array}{rclrcl} B_0 A &=& c_0 I_q & B_0 A &=& c_0 I_q \\ B_1 A - B_0 &=& c_1 I_q & B_1 A^2 - B_0 A &=& c_1 A \\ B_2 A - B_1 &=& c_2 I_q & B_0 A^3 - B_1 A^2 &=& c_0 A^2 \\ \cdots &=& \cdots & \cdots &=& \cdots \\ -B_{q-1} &=& c_q I_q & -B_{q-1} A^q &=& c_q A^q \end{array}$$

La colonne de gauche a été obtenue par identification. La colonne de droite s'obtient en multipliant à droite par A^k la ligne k. Sommons la colonne de droite, on obtient $0 = P_a(A)$ et le théorème est montré.

Exercice 4.1. Quel est le polynôme caractéristique d' une matrice triangulaire supérieure (a_{ij})?

Exercice 4.2. Si A et B sont des matrices carrées de même ordre et si B est inversible, pourquoi A et BAB^{-1} ont ils le même polynôme caractéristique ? pourquoi AB et BA ont ils le même polynôme caractéristique ?

Exercice 4.3. Si A et B sont des matrices carrées de même ordre q (cette fois ci B n'est pas nécessairement inversible), montrer que AB et BA ont même polynôme caractéristique. Méthode : les déterminants des deux matrices par blocs suivantes

$$\begin{bmatrix} A & XI_q \\ I_q & 0 \end{bmatrix} \begin{bmatrix} B & -XI_q \\ -I_q & A \end{bmatrix}, \quad \begin{bmatrix} B & -XI_q \\ -I_q & A \end{bmatrix} \begin{bmatrix} A & XI_q \\ I_q & 0 \end{bmatrix}$$

sont les mêmes. Appliquer alors la Proposition 2.3.

Exercice 4.4. On considère la matrice réelle

$$A = \begin{bmatrix} 3 & -2 & -1 \\ 3 & -1 & 1 \\ 6 & 3 & -2 \end{bmatrix}.$$

Calculer son polynôme caractéristique et son carré. En utilisant Cayley Hamilton, trouver son inverse.

Exercice 4.5. Soit C une matrice carrée d'ordre q. Montrer à l'aide de la formule de Laplace du Théorème 2.4 que son polynôme caractéristique est $\sum_{T\subset\{1,\dots,q\}}(-X)^{|T'|}\det C_T$, où $|T'|$ est le nombre d'éléments du complémentaire T' de T.

Exercice 4.6. Soit $(a_1,\dots,a_q)$ et $(b_1,\dots,b_q)$ dans K^q et la matrice

$$M = \begin{bmatrix} a_1b_1 & a_2b_1 & \dots & a_qb_1 \\ a_1b_2 & a_2b_2 & \dots & a_qb_2 \\ \dots & \dots & \dots & \dots \\ a_1b_q & a_2b_q & \dots & a_qb_q \end{bmatrix}.$$

Trouver sans calculs le polynôme caractéristique de cette matrice

- en l'interprétant comme la matrice représentative $[\varphi]_e^e$ de l'endomorphisme φ de l'exercice 3.5 ;
- en appliquant la formule de Laplace du théorème 2.4 au couple $(M,-XI_q)$ et en remarquant que M est de rang 1.

Exercice 4.7. Soit, pour $j=1,\dots,k\leq q$ les k matrices colonnes $a_j=(a_{1j},\dots,a_{qj})^T$ de K^q. Si $S\subset\{1,\dots,q\}$ on convient $a_j(S)=\sum_{i\in S}a_{ij}$. On considère alors la matrice carrée $C=[c_1,\dots,c_q]$ sur K dont les colonnes c_j sont prises dans l'ensemble de colonnes $\{a_1,\dots,a_k\}$. On note $S_i=\{j\ ;\ c_j=a_i\}$, $m_{ij}=a_i(S_j)$ et $M=(m_{ij})_{1\leq i,j\leq k}$. Montrer que

$$\det(C-XI_q)=(-X)^{q-k}\det(M-XI_k).$$

Méthode : pour $k=2$, c'est une reformulation de l'exercice 2.8.

V La diagonalisation et le polynôme minimal

Voici la partie la plus utile du chapitre. Nous commencons par appliquer le Théorème 3.1 et la Proposition 4.1 à des exemples.

Exemple 5.1 : On prend $K=\mathbb{R}$ l'espace E de dimension 3, une base e et un endomorphisme a de E tel que $A=[a]_e^e$ soit égal à

$$A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}. \tag{5.10}$$

Nous allons rechercher les valeurs propres et espaces propres de a. Ceci repose sur des techniques de première année, en particulier la résolution des systèmes linéaires *homogènes*. On a par calcul $P_a(X) = -X^3 + 5X^2 - 8X + 4$. Une racine évidente est 1. Pour trouver les autres racines on fait la division euclidienne de $-X^3+5X^2-8X+4$ par $X-1$ pour voir que le quotient est $-X^2+4X-4 = -(X-2)^2$. Donc a a deux valeurs propres qui sont 1 et 2. Cherchons une base des espaces propres E_1 et E_2. Pour E_1 c'est donc rechercher les vecteurs v de E avec $v = xe_1 + ye_2 + ze_3$ où $V = (x, y, z)^T \in \mathbb{R}^3$ satisfait au système linéaire $AV = V$ ou encore

$$\begin{array}{rrrcl} 5x & -6y & -6z & = & x \\ -x & +4y & +2z & = & y \\ 3x & -6y & -4z & = & z \end{array}$$

et, de facon équivalente

$$\begin{array}{rrrcl} 4x & -6y & -6z & = & 0 \\ -x & +3y & +2z & = & 0 \\ 3x & -6y & -5z & = & 0 \end{array}$$

On a un système homogène, qui a évidemment une solution non triviale, car le déterminant de $A - I_3$ est nul (puisque 1 est valeur propre). Pour trouver toutes les solutions on détermine d'abord le rang de la matrice $A - I_3$ du système. Les deux premières lignes et colonnes fournissent le déterminant d'ordre 2 non nul $\begin{vmatrix} 4 & -6 \\ -1 & 3 \end{vmatrix} = 18$. Le rang est donc 2, et de plus on peut prendre pour *équations principales* les deux premières, et pour *inconnues principales* les variables x et y. Poursuivant les techniques de première année on fait passer au second membre l'inconnue *non principale* z et on résout le système de Cramer

$$\begin{array}{rrcr} 4x & -6y & = & 6z \\ -x & +3y & = & -2z \end{array}$$

et on obtient $x = z$ et $y = -z/3$. Nous avons ainsi *paramétré l'espace propre par l'inconnue non principale.* On remarque que E_1 est de dimension 1. Une base de E_1 est obtenue en prenant *par exemple* $z = 3$ et la base de E_1 est formée de l'unique vecteur $f_1 = 3e_1 - e_2 + 3e_3$.

Passons à E_2. On recherche donc les vecteurs v de E avec $v = xe_1 + ye_2 + ze_3$ où $V = (x, y, z)^T \in \mathbb{R}^3$ satisfait au système linéaire $AV = 2V$ ou encore $(A - 2I_3)V = 0$, ou encore

$$\begin{array}{rrrcl} 3x & -6y & -6z & = & 0 \\ -x & +2y & +2z & = & 0 \\ 3x & -6y & -6z & = & 0 \end{array}$$

Cette fois ci, le rang de $A-2I_3$ est 1. On peut prendre la première équation pour équation principale, et x pour inconnue principale. Le système de Cramer se réduit à l'unique équation

$$3x = 6y + 6z.$$

L'espace propre E_2 est de dimension 2 et est paramétré par les deux inconnues non principales y et z. C'est un plan d'équation $x = 2(y+z)$. Une base de E_2 est donc obtenue *par exemple* en prenant $(y, z) = (1, 0)$ et $(0, 1)$, fournissant les vecteurs $f_2 = 2e_1 + e_2$ et $f_3 = 2e_1 + e_3$. Il est clair que f_2 et f_3 sont indépendants. Insistons sur le fait que le choix d'une base d'un espace propre a un *caractère arbitraire*.

A ce point rappelons que le Théorème 3.1 garantit que E_1 et E_2 sont en somme directe. Comme E_1 est de dimension 1, E_2 de dimension 2 et E de dimension 3 on a $E = E_1 \oplus E_2$. Donc $f = (f_1, f_2, f_3)$ est une base de E, de matrice de changement de base

$$P = [\mathrm{id}_E]_f^e = \begin{bmatrix} 3 & 2 & 2 \\ -1 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}.$$

Si on décide de représenter a dans cette nouvelle base f, le résultat est immédiat : $[a]_f^f$ doit être diagonale et est égale à

$$[a]_f^f = D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Si on calcule P^{-1} par les techniques standard on obtient

$$P^{-1} = [\mathrm{id}_E]_e^f = \begin{bmatrix} -1 & 2 & 2 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix}.$$

D'après le Théorème 1.1 (3) on a $[a]_e^e = [id_E]_f^e[a]_f^f[\mathrm{id}_E]_e^f = P^{-1}DP$ ou encore

$$A = \begin{bmatrix} 3 & 2 & 2 \\ -1 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} -1 & 2 & 2 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix}.$$

Un des avantages de la représentation dans une base f de vecteurs propres est que le calcul de la représentation matricielle $[a^k]_f^f = D^k$ est immédiat, et donc $[a^k]_e^e = A^k = PD^kP^{-1}$ se calcule moins péniblement.

Exemple 5.2 : On prend $K = \mathbb{R}$ l'espace E de dimension 2, une base e et un endomorphisme a de E tel que $A = [a]_e^e$ soit égal à

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{5.11}$$

Ici, le polynôme caractéristique $P_a(X)$ est X^2, et la seule valeur propre est 0. Des calculs conduits comme à l'exemple 5.1 montrent que l'espace propre E_0 est de dimension 1 et est engendré par e_1. Cet exemple montre que la somme directe des espaces propres n'est pas forcément égale à tout E. Il montre aussi que l'affirmation $B = PAP^{-1} \Rightarrow P_A = P_B$

a une réciproque fausse : Prendre A comme en (5.11) et prendre pour B la matrice nulle d'ordre 2.

Exemple 5.3 : On prend pour K le corps $\mathbb{C}$ des complexes. On se fixe un nombre réel θ. On prend l'espace E de dimension 2, une base e et un endomorphisme a de E tel que $[a]_e^e$ soit égal à
$$R(\theta)=\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix} \tag{5.12}$$
On a $P_a(X)=X^2-2\cos\theta X+1$ et les valeurs propres sont complexes et égales à $\cos\theta\pm i\sin\theta=e^{\pm i\theta}$. Si θ est un multiple de π alors $a=\pm\mathrm{id}_E$ n'a que la valeur propre ±1 et E est l'espace propre. Si $\theta\not\equiv 0 \bmod \pi$, alors l'espace propre $E_{e^{i\theta}}$ est de dimension 1 et est engendré par $f_1=e_1+ie_2$ et l'espace propre $E_{e^{-i\theta}}$ est de dimension 1 et est engendré par e_1-ie_2. Mais par souci de symétrie, nous prenons plutôt un vecteur f_2 obtenu en multipliant le précédent par i, soit $f_2=ie_1+e_2$. Si $f=(f_1,f_2)$ alors $P=[id_E]_f^e=\begin{bmatrix}1 & i\\ i & 1\end{bmatrix}$ et donc $P^{-1}=[id_E]_e^f=\frac{1}{2}\begin{bmatrix}1 & -i\\ -i & 1\end{bmatrix}$. Ici les vecteurs propres forment une base de E et si $[a]_f^f=D=\begin{bmatrix}e^{i\theta} & 0\\ 0 & e^{-i\theta}\end{bmatrix}$ on aura
$$R(\theta)=\frac{1}{2}\begin{bmatrix}1 & i\\ i & 1\end{bmatrix}\begin{bmatrix}e^{i\theta} & 0\\ 0 & e^{-i\theta}\end{bmatrix}\begin{bmatrix}1 & -i\\ -i & 1\end{bmatrix}.$$

Exemple 5.4 : On prend $K=\mathbb{R}$. On se fixe un nombre réel $\theta\not\equiv 0 \bmod \pi$. On prend l'espace E de dimension 2, une base e et un endomorphisme a de E tel que $[a]_e^e=R(\theta)$ défini par (7.12). Puisque qu'alors $P_a(X)=X^2-2\cos\theta X+1$ n'a pas de racines *réelles,* nous avons un exemple d'endomorphisme qui n'a *pas de valeurs propres*, et donc pas de vecteurs propres. Au chapitre des espaces euclidiens, nous verrons que si e est une base orthonormale alors a est une rotation d'angle θ et qu'il n'est pas surprenant de ne pas avoir de vecteur dont la direction serait conservée par a.

Définition 5.2 : Si E est de dimension finie, alors a dans $L(E)$ est dit *diagonalisable* si il existe une base f de E telle que $[a]_f^f$ soit une matrice diagonale. On dit qu'une matrice carrée A d'ordre q sur le corps K est diagonalisable si elle représente un endomorphisme diagonalisable, c'est à dire si il existe E, $a\in L(E)$ et une base e tels que $A=[a]_e^e$, avec a diagonalisable.

Remarques : Il est évident que si $[a]_f^f$ est diagonale, alors f_j est un vecteur propre de a, et qu'il est associé à la valeur propre d_j, le j ième élément de la diagonale. Les exemples 5.1 et 5.3 sont diagonalisables. Comme on l'a vu, les endomorphismes diagonalisables sont très faciles à manier une fois écrits dans *une* base de diagonalisation (celle ci n'est pas unique). Nous allons donner dans la suite des conditions nécessaires et des conditions suffisantes de diagonalisabilité. Par exemple, on se doute que les exemples 5.2 et 5.4 ne sont pas diagonalisables.

Définitions 5.3 : On dit qu'un polynôme P sur le corps K est *scindé*[3] si il est le produit de polynômes de degré 1. En particulier, si $K = \mathbb{R}$ si $\deg P = q$, alors P est scindé si P a toutes ses racines réelles. On dit qu'un polynôme P non nul est *monique* [4] si le coefficient du terme de plus haut degré est 1.

Proposition 5.1. Soit E de dimension finie sur K et $a \in L(E)$. Alors

1. a est diagonalisable $\Leftrightarrow$ E a une base de formée de vecteurs propres de a $\Leftrightarrow$ E est la somme directe des espaces propres de E.
2. Si $\dim E = q$ et si P_a a q racines dans K *distinctes*, alors a est diagonalisable. De plus, tous les espaces propres sont de dimension 1.
3. Si a est diagonalisable de spectre $\{\lambda_1, \ldots, \lambda_p\}$, la dimension de l'espace propre E_{λ_j} est égale à la multiplicité m_j de la racine λ_j de P_a. De plus $\det a = \lambda_1^{m_1} \ldots \lambda_p^{m_p}$.
4. Si a est diagonalisable, alors P_a est scindé.

Démonstration : (1) Très facile. Le travail est fait par le Théorème 3.1.
(2) S'il y a q valeurs propres λ_j, et puisque tout espace propre est de dimension ≥ 1, alors

$$q = \dim E \geq \dim(E_{\lambda_1} + \cdots + \dim E_{\lambda_q}) \geq q.$$

Ceci entraîne que ces inégalités deviennent des égalités. Donc $E = E_{\lambda_1} \oplus \cdots E_{\lambda_q}$ et $\dim E_{\lambda_j} = 1$ pour tout j.
(3) et (4) Si f est une base de diagonalisation de a, alors $D = [a]_f^f$ s'écrit par blocs :

$$D = \begin{bmatrix} \lambda_1 I_{m_1} & \ldots & 0 \\ 0 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & \lambda_p I_{m_p}. \end{bmatrix}. \tag{5.13}$$

Pour (4) il suffit d'écrire que $P_a(X) = \det([a]_f^f - XI_q)$.

La fin de cette section et les sections suivantes sont relativement difficiles (et considérées en partie comme hors du programme d'une classe de Mathématiques Spéciales au lycée) et d'une utilité pratique plus modeste. Elles donnent une meilleure maitrise de l'algèbre linéaire, et nous suivrons une tradition universitaire vénérable en les exposant. Le Théorème 4.2 fournit une condition nécessaire et suffisante de diagonalisabilité de a en termes du polynôme minimal de a. Nous allons réactiver la notion de polynômes annulateurs d'un endomorphisme $a \in L(E)$ introduite avant le Théorème de Cayley Hamilton 3.2.

Théorème 5.2. Soit E de dimension finie sur K et $a \in L(E)$. Soit $\mathcal{I}(a)$ l'ensemble des polynômes sur K tels que $P(a) = 0$.

[3] D'autres disent dissocié.
[4] D'autres disent unitaire.

1. Il existe un polynôme monique unique m_a (appelé *polynôme minimal* de a) dans $\mathcal{I}(a)$ tel que tout P de $\mathcal{I}(a)$ soit un multiple de m_a.
2. Toute valeur propre de a est racine de m_a.
3. a est diagonalisable si et seulement si m_a est scindé et n'a que des racines simples.

Remarques : Le polynôme m_a est dit minimal car c'est le polynôme monique de plus bas degré de $\mathcal{I}(a)$. Le polynôme caractéristique P_a est dans $\mathcal{I}(a)$ d'après la propriété de Cayley Hamilton et donc m_a divise P_a. Le fait que $\mathcal{I}(a)$ ne soit pas réduit à 0 se déduit de Cayley Hamilton certes, mais peut être montré d'une manière plus élémentaire : puisque $L(E)$ est un espace vectoriel de dimension finie $n = (\dim E)^2$, alors les $n+1$ vecteurs de $L(E)$ formés par $\mathrm{id}_E, a, a^2, \ldots, a^n$ sont dépendants, c'est à dire qu'on peut trouver des éléments $c_0, \ldots, c_n$ dans K non tous nuls tels que

$$c_0\mathrm{id}_E + c_1 a + \cdots + c_n a^n = 0,$$

c'est à dire encore que $P(X) = c_0 + c_1X + \cdots + c_nX^n$ définit un élément non nul de $\mathcal{I}(a)$.

Il n'y a pas d'algorithme pour calculer m_a en dehors de sa définition : il faut rechercher le diviseur m_a non nul de P_a de plus bas degré tel que $m_a(a) = 0$. Toutefois la recherche de m_a est plus ou moins simple suivant la base dans laquelle on représente a. L'exemple 4.1 est instructif. On sait que $P_a(X) = -(X-1)(X-2)^2$ et le théorème précédent dit que $m_a(X) = (X-1)(X-2)$ car a est diagonalisable (et donc les racines de m_a sont simples) et toutes les racines de P_a doivent être présentes. Enfin, avec les bases e et f de l'exemple 4.1, il est beaucoup plus facile de vérifier que $([a]_f^f - I_3)(([a]_f^f - 2I_3) = 0$ que $([a]_e^e - I_3)(([a]_e^e - 2I_3) = 0$.

Une utilisation fréquente du Théorème 5.2 est la suivante : si a est tel qu'il existe un polynôme scindé à racines simples satisfait $P(a) = 0$, alors a est diagonalisable. En effet $P \in \mathcal{I}(a)$ et donc m_a divise P. Comme P est scindé et à racines simples, il en va de même pour m_a et le théorème s'applique.

Un *exemple important* est celui des *projections* : Rappelons qu'en première année on a dit que le sous espace vectoriel F' de E est un supplémentaire du sous espace vectoriel F de E si $F + F' = E$ et si $F \cap F' = \{0\}$, autrement dit $E = F \oplus F'$. Si $x \in F$ et $x' \in F'$ le procédé a qui envoie $x + x'$ sur x est appelé projection de E sur F *parallèlement à* F'. Il est clair que $a^2 = a$. Inversement soit $a \in L(E)$ tel que $a^2 = a$. Alors a est une projection. En effet, d'après le Théorème 5.2 si $a^2 = a$ alors $P(a) = 0$ avec $P(X) = X^2 - X = X(X-1)$: les valeurs propres de a sont dans $\{0, 1\}$ et a est diagonalisable. On prend alors pour F l'espace propre de la valeur propre 1 et pour F' l'espace propre de la valeur propre 0. Le cas où 0 n'est pas valeur propre est $a = \mathrm{id}_E$ et celui où 1 n'est pas valeur propre est $a = 0$.

Exemple 5.5 : On reprend les notations de l'exemple des matrices circulantes de la section 3. Il est facile de voir que $R^q = I_q$. Ce ci montre déjà que le polynôme $X^q - 1$ est dans l'idéal $\mathcal{I}(r)$. Pour en déduire que le polynôme minimal de r est $m_r(X) = X^q - 1$ observons que si m_r est de degré plus petit alors la matrice $m_q(R)$ est nulle ce qui contredit 3.7. Donc polynômes minimal et caractéristique coïncident ici :

$$P_r(X) = m_r(X) = X^q - 1.$$

On en déduit si K est le corps de complexes $\mathbb{C}$ que les valeurs propres de r et R sont les q racines de l'unité $e^{\frac{2ik\pi}{q}}$ avec $k = 0, \dots, q-1$.

Une conséquence très utile est que les valeurs propres de la matrice circulante $P(R)$ de 3.7 sont les q nombres complexes $P(e^{\frac{2ik\pi}{q}})$ avec $k = 0, \dots, q-1$.

Démonstration du Théorème 5.2 : (1) On observe que $\mathcal{I}(a)$ est un *idéal* de l'algèbre $K[X]$ des polynômes sur K, ce qui signifie que $\mathcal{I}(a)$ est un sous espace vectoriel de $K[X]$ tel que de plus $QP \in \mathcal{I}(a)$ pour tout polynôme $Q \in K[X]$ et tout polynôme $P \in \mathcal{I}(a)$. Il a été vu en première année qu'alors il existe m_a dans $\mathcal{I}(a)$ unique avec les propriétés annoncées.
(2) Si λ est une valeur propre de a et si $m_a(\lambda) \neq 0$, alors en écrivant

$$m_a(X) = b_0 + b_1(X-\lambda) + b_2(X-\lambda)^2 + \dots + b_p(X-\lambda)^p$$

on a $b_0 \neq 0$. Comme $m_a(a) = 0$ cela entraîne

$$0 = b_0\mathrm{id}_E + b_1(a-\lambda\mathrm{id}_E) + b_2(a-\lambda\mathrm{id}_E)^2 + \dots + b_p(a-\lambda\mathrm{id}_E)^p,$$

$$(a-\lambda\mathrm{id}_E)[-\frac{b_1}{b_0}\mathrm{id}_E - \frac{b_2}{b_0}(a-\lambda\mathrm{id}_E) + \dots - \frac{b_p}{b_0}(a-\lambda\mathrm{id}_E)^{p-1}] = \mathrm{id}_E.$$

Donc $a-\lambda\mathrm{id}_E$ est inversible et de noyau réduit à 0, ce qui contredit le fait que λ soit une valeur propre.
(3) $\Rightarrow$ Si a est diagonalisable, soit f une base de diagonalisation, soit $\lambda_1, \dots, \lambda_p$ les valeurs propres distinctes de a et soit $m_1, \dots, m_p$ les dimensions respectives des espaces propres (c'est à dire les multiplicités des racines du polynôme caractéristique) On a $q = \dim E = m_1 + \dots + m_p$. Alors $D = [a]_f^f$ s'écrit par blocs comme en (5.13). Alors on en déduit

$$D - \lambda_1 I_q = \begin{bmatrix} (\lambda_1-\lambda_1)I_{m_1} & \dots & 0 \\ 0 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & (\lambda_p-\lambda_1)I_{m_p} \end{bmatrix},$$

et on voit que $(D-\lambda_1 I_q)\dots(D-\lambda_1 I_q) = 0$. C'est dire que $(X-\lambda_1)\dots(X-\lambda_p)$ est un polynôme de $\mathcal{I}(a)$. D'après le (2), c'est m_a.
(3) $\Leftarrow$ Cette partie est plus coriace et utilise le résultat suivant, dit *lemme des noyaux,* et qui resservira :

Théorème 5.3. Soit $Q_1, \dots, Q_p$ des polynômes sur K deux à deux premiers entre eux, $P = Q_1 \dots Q_p$ et $a \in L(E)$. Alors les $\ker Q_j(a)$ sont en somme directe et

$$\ker P(a) = \ker Q_1(a) \oplus \dots \oplus \ker Q_p(a).$$

Acceptons ce théorème quelques instants pour achever la démonstration du Théorème 5.2 : si $m_a(X) = (X-\lambda_1)\dots(X-\lambda_p)$ on applique le Théorème 5.3 aux $Q_j(X) = X - \lambda_j$. Comme $\ker m_a(a) = E$ et que $\ker Q_j(a) = E_{\lambda_j}$ on a $E = E_{\lambda_1} \oplus \dots \oplus E_{\lambda_p}$ et E est somme directe des espaces propres de a : c'est dire que a est diagonalisable.

Démonstration du Théorème 5.3. On suppose d'abord $p = 2$. On applique le *lemme de Bézout* (voir cours de 1 ère année) qui dit que puisque Q_1 et Q_2 sont premiers entre eux alors il existe deux polynômes P_1 et P_2 tels que

$$\begin{aligned} 1 &= P_1(X)Q_1(X) + P_2(X)Q_2(X), \\ \mathrm{id}_E &= P_1(a)Q_1(a) + P_2(a)Q_2(a), \\ x &= P_1(a)Q_1(a)(x) + P_2(a)Q_2(a)(x). \end{aligned} \tag{5.14}$$

Dans (5.14) prenons $x \in \ker Q_1(a) \cap \ker Q_2(a)$: (5.14) entraîne qu'alors $x = 0$ et donc $\ker Q_1(a)$ et $\ker Q_2(a)$ sont en somme directe. Ensuite, si $x \in \ker Q_1(a)Q_2(a) = \ker(Q_1Q_2)(a)$, notons $x_j = P_j(a)Q_j(a)(x)$. Alors $x_1 \in \ker Q_2(a)$, puisque $Q_2(a)(x_1) = Q_2(a)P_1(a)Q_1(a)(x) = P_1(a)Q_1(a)Q_2(a)(x) = 0$ par définition de x. De même $x_2 \in \ker Q_1(a)$. Enfin $x = x_1 + x_2$ par (5.14). Donc :

$$\ker(Q_1Q_2)(a) \subset \ker(Q_1)(a) \oplus \ker(Q_2)(a).$$

Pour montrer que cette inclusion est en fait une égalité, prenons $x_j = \ker Q_j(a)$ avec $j = 1, 2$ et appliquons $Q_1(a)Q_2(a)$ à $x_1 + x_2$:

$$Q_1(a)Q_2(a)(x_1 + x_2) = Q_2(a)Q_1(a)(x_1) + Q_1(a)Q_2(a)(x_2) = 0 + 0.$$

Donc

$$\ker(Q_1Q_2)(a) \supset \ker(Q_1)(a) \oplus \ker(Q_2)(a),$$

et le théorème est montré pour $p = 2$.

Pour terminer, on procède par récurrence sur p. Rappelons d'abord le *lemme de Gauss :* si P, Q, R sont des polynômes tels que R divise PQ et R soit premier avec Q, alors R divise P (démonstration : par Bézout il existe Q_1 et R_1 tels que $QQ_1 + RR_1 = 1$; comme $PQ = RS$ alors $QS_1 = S$ avec $S_1 = SQ_1 + PR_1$; d'où $PQ = RQS_1$ et $P = RS_1$). Ensuite, on observe que si $k \leq p$ si R divise $Q_1 \dots Q_k$ et si R est premier avec chaque Q_j, $j = 1, \dots, k$, alors R est constant. En effet, en appliquant le lemme de Gauss au couple $(Q_1 \dots Q_{k-1}, Q_k)$, R divise $Q_1 \dots Q_{k-1}$. De même R divise $Q_1 \dots Q_{k-2}$, etc jusqu'à Q_1, d'où le résultat. Ce résultat entraîne que $Q_1 \dots Q_{k-1}$ et Q_k *sont premiers entre eux.*

Supposons alors le théorème vrai pour $p-1$ et montrons le pour p. Puisque $Q_1 \dots Q_{p-1}$ et Q_p sont premiers entre eux, le cas $p = 2$ déjà traité entraîne

$$\ker(Q_1 \dots Q_p)(a) = \ker(Q_1 \dots Q_{p-1})(a) \oplus \ker Q_p(a).$$

Il reste à appliquer l'hypothèse de récurrence à $\ker(Q_1 \dots Q_{p-1})(a)$ et la démonstration du lemme des noyaux est achevée.

Exercice 5.1. Quel est le polynôme minimal de l' endomorphisme nul ? De $a = \lambda \mathrm{id}_E$?

Exercice 5.2. Montrer que l'endomorphisme $A \mapsto A^T$ de l'exercice 3.4 est diagonalisable, et calculer son déterminant en fonction de q au moyen de la Proposition 5.1, partie 3.

Exercice 5.3. Soit E et F des espaces de dimension finie sur K. Soit $a \in L(E)$ diagonalisable et φ_a l'application linéaire de $L(E, F)$ dans lui même définie par $b \mapsto \varphi_a(b) = b \circ a$. Montrer que

φ_a est a les mêmes valeurs propres que a et en déduire que $\det \varphi_a = (\det a)^{\dim F}$ (Méthode : prendre une base e de diagonalisation de a et une base quelconque f de F, noter $D = [a]_e^e$ et $B = [b]_e^f$ et examiner l'endomorphisme de l'espace des matrices (p, q) défini par $B \mapsto BD$).

Exercice 5.4. Montrer que φ défini à l'exercice 3.5 est diagonalisable si et seulement si $a(b) \neq 0$.

Exercice 5.5. Soit E un espace vectoriel complexe de dimension q et soit a et b dans $L(E)$ de spectres disjoints, c'est à dire que les polynômes caractéristiques P_a et P_b sont premiers entre eux. Montrer que $P_b(a)$ est inversible (appliquer le lemme des noyaux à a et à P_aP_b). On pose ensuite $F = L(E)$ et on considère $\varphi \in L(F)$ défini pour tout $x \in F$ par $\varphi(x) = ax - bx$. Montrer que $\ker \varphi = \{0\}$ (Méthode : soit $x_0 \in F$ tel que $ax_0 = x_0b$; montrer que pour tout polynôme Q on a $Q(a)x_0 = x_0Q(b)$ et appliquer cela à $Q = P_b$). Soit enfin A et B des matrices carrées complexes d'ordre q ayant une valeur propre commune λ et soit U et V des matrices colonnes non nulles d'ordre q telles que $AU = \lambda U$ et $B^TV = \lambda V$. Si $X = UV^T$ calculer $AX - XB$.. En déduire le *théorème de Liapounov* : pour tout $c \in L(E)$ il existe un unique $x \in L(E)$ tel que $ax - xb = c$ si et seulement si a et b ont des spectres disjoints.

Exercice 5.6. Soit M l'espace vectoriel des matrices (n, n) à coefficients dans le corps de complexes $\mathbb{C}$. Si $1 \leq k, l \leq n$, la matrice E_{kl} de M est la matrice (x_{ij}) telle que $x_{ij} = 0$ si $(i, j) \neq (k, l)$ et $x_{kl} = 1$. Donc $e = (E_{11}, E_{12}, E_{13}, \ldots, E_{nn})$ forme une base de M. Si A est fixé dans M et si $X \in M$ on pose $\varphi_A(X) = AX - XA$. Ceci définit donc un endomorphisme φ_A de M. Le but de l'exercice est de montrer que φ_A est diagonalisable si et seulement si A est diagonalisable. On se fixe une matrice diagonale $D = \mathrm{Diag}(c_1, c_2, \ldots, c_n)$ de M. Calculer $\varphi_D(E_{kl})$ pour $1 \leq k, l \leq n$. En déduire les valeurs propres de φ_D en fonction de $c_1, c_2, \ldots, c_n$, et en déduire que φ_D admet une base de diagonalisation. On se fixe de plus $P \in M$ inversible. Calculer $\varphi_{PDP^{-1}}(PE_{kl}P^{-1})$ pour $1 \leq k, l \leq 3$. En déduire les valeurs propres de $\varphi_{PDP^{-1}}$ et en déduire que $\varphi_{PDP^{-1}}$ admet une base de diagonalisation. On suppose ensuite que φ_A admet une base de diagonalisation $(F_1, \ldots, F_{n^2})$ correspondant aux valeurs propres $(\lambda_1, \ldots, \lambda_{n^2})$ satisfaisant $\varphi_A(F_k) = \lambda_k F_k$ pour $k = 1, \ldots, n^2$. Pourquoi existe t-il $V \in \mathbb{C}^n$ non nul et un complexe λ tel que $AV = \lambda V$? Calculer $\varphi_A(F_k)V$ et en déduire que $F_kV = (\lambda_k + \lambda)V$. Montrer que le fait que les $(F_1, \ldots, F_{n^2})$ forment une famille génératrice de M entraîne que les $(F_1V, \ldots, F_{n^2}V)$ forment une famile génératrice de $\mathbb{C}^n$. En déduire que A admet une base de diagonalisation.

Exercice 5.7. [5]Soit n un entier ≥ 2, soit J la matrice (n, n) dont tous les coefficients sont 1 et soit I_n la matrice identité d'ordre n. On considère une matrice (n, n) symétrique réelle A dont les coefficients sont 0 ou 1, telle que trace $A = 0$ et telle qu'il existe un entier $d > 0$ avec

$$A^2 + A - (d-1)I_n = J. \quad (*)$$

Montrer que les éléments de la diagonale de A^2 sont tous égaux à d. Montrer en prenant la trace de $(*)$ que $n = d^2+1$. Montrer que si $\mathbf{1} = (1, \ldots, 1)^T$ alors $A\mathbf{1} = d\mathbf{1}$, que $(A - dI_n)J = 0$

[5]Source : A.J. Hoffmann "Eigenvalues of graphs" *Studies in Graph Theory II*, 225-245, D. Fulkerson ed. Mathematical Association of America (1975).

et que $(X-d)(X^2+X-(d-1))$ est polynôme annulateur de A. Si α et β sont les racines de $X^2+X-(d-1)$ montrer qu'il existe une matrice inversible P et des entiers positifs ou nuls (n_d, n_α, n_β) tels que si

$$D = \mathrm{diag}(dI_{n_d}, \alpha I_{n_\alpha}, \beta I_{n_\beta})$$

alors $A = PDP^{-1}$. Montrer que (n_d, n_α, n_β) satisfait au systême linéaire

$$n_d + n_\alpha + n_\beta = d^2+1, \quad dn_d + \alpha n_\alpha + \beta n_\beta = 0, \quad d^2 n_d + \alpha^2 n_\alpha + \beta^2 n_\beta = d(d^2+1)$$

(Méthode : comparer trace A et trace A^2 à trace D et trace D^2). En déduire $n_d = 1$ et $n_\alpha = \frac{1}{\beta-\alpha} d(d\beta+1)$. Montrer enfin que $d \in \{1, 2, 5, 7, 57\}$ (Méthode : observer que n_α est entier, que

$$(2n_\alpha - d^2)\sqrt{4d-3} = d(2-d)$$

et que $s = \sqrt{4d-3}$ doit être entier si $d \neq 2$, et en déduire que dans ce cas s divise 15. Préciser dans les 5 cas les polynômes caractéristique et minimal de A. On ne sait pas si A existe pour d=57 et n=3250.

VI Diagonalisation simultanée *

Cette section contient un résultat très utile[6].

Théorème 6.1. Soit E de dimension finie et soit $\mathcal{A}$ un sous ensemble de $L(E)$ formé d'endomorphismes tous diagonalisables et qui commutent deux à deux entre eux. Alors il existe une base e de E telle que pour tout a de $\mathcal{A}$ la matrice $[a]_e^e$ soit diagonale.

Remarques : Il est essentiel de supposer que les éléments de $\mathcal{A}$ sont diagonalisables : des endomorphismes non diagonalisables peuvent commuter. Par exemple pour $E = \mathbb{R}^2$ les $A_t = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$ satisfont $A_{t+s} = A_t A_s$. Pourtant A_t n'est pas diagonalisable pour $t \neq 0$: son polynôme minimal serait $X-1$, ce qui n'est pas, puisque $A_t - I_2 \neq 0$. Très souvent on applique le théorème au cas où $\mathcal{A}$ est une *algèbre*, c'est à dire un sous espace vectoriel de $L(E)$ qui est de plus fermé pour la composition des endomorphismes. Nous n'utiliserons ce résultat qu'une fois dans la suite du cours, et seulement pour une famille $\mathcal{A}$ à deux éléments, pour montrer l'unicité de la décomposition de Dunford au Théorème 8.2.

Démonstration : Elle se fait en deux parties : le cas où $\mathcal{A}$ est fini ; le cas où $\mathcal{A}$ est infini.

Si $\mathcal{A} = \{a_1, \ldots, a_N\}$ soit Λ_j le spectre de a_j et soit

$$(E^{(j)}_{\lambda_j})_{\lambda_j \in \Lambda_j}$$

la famille des espaces propres de a_j. Pour k fixé dans $\{1, \ldots, N\}$ on note $\Lambda^k = \Lambda_1 \times \ldots \times \Lambda_k$, dont un élément typique est noté $\lambda^k = (\lambda_1, \ldots, \lambda_k)$. Enfin on note

$$F_{\lambda^k} = \bigcap_{j=1}^{k} E^{(j)}_{\lambda_j}$$

[6]Une section ou un énoncé marqués d'une étoile * ne sont pas traités en cours et sont là pour la culture.

(F_{λ^k} peut être réduit à zéro) et on note par B_k la réunion de tous les espaces F_{λ^k}. Nous montrons alors par récurrence sur k le résultat suivant

1. Il existe une base e contenue dans B_k;
2. Pour toute base e contenue dans B_k, alors $[a_1]_e^e, \dots, [a_k]_e^e$ sont diagonales.

L'application à $k = N$ donnera le résultat voulu dans le cas où $\mathcal{A}$ est fini. L'hypothèse de récurrence est vérifiée pour $k = 1$, d'après la Proposition 5.1. Supposons la vérifée à un ordre $k < N$ fixé et montrons qu'elle est vraie à l'ordre $k + 1$.

Pour simplifier la notation, fixons $\lambda^k = (\lambda_1, \dots, \lambda_k)$ dans Λ^k et notons $F = F_{\lambda^k}$. Alors

– F est *stable* par a_{k+1}, c'est à dire que $a_{k+1}(F) \subset F$. En effet, si $x \in F$, alors pour $j \leq k$ on a, puisque a_j et a_{k+1} commutent :

$$a_j \circ a_{k+1}(x) = a_{k+1} \circ a_j(x) = \lambda_j a_{k+1}(x),$$

ce qui est dire que $a_{k+1}(x)$ est un vecteur propre de a_j pour la valeur propre λ_j. Donc $a_{k+1}(x) \in E_{\lambda_j}^{(j)}$ pour tout $j \leq k$, donc $a_{k+1}(x) \in F$.

– Soit b la *restriction* de a_{k+1} à F et soit $m = m_{a_{k+1}}$ le polynôme minimal de a_{k+1} (relativement à E, non au sous espace stable F). Si $x \in F$, la définition du polynôme minimal entraîne

$$0 = m(a_{k+1}),\ 0 = m(a_{k+1})(x) = m_b(x)$$

et *donc* $m(b) = 0$. On en déduit que le polynôme minimal m_b *divise* m. Or m est scindé et n'a que des racines simples (Théorème 5.2 appliqué à a_{k+1} diagonalisable). Donc b est diagonalisable sur F par le même théorème. Enfin, $x \in F$ est un vecteur propre de b si et seulement si x est dans

$$F \bigcap \bigcup_{\lambda_{k+1} \in \Lambda_{k+1}} = F \cap B_{k+1}.$$

F admet donc une base contenue dans B_{k+1}, qui diagonalise b.

Revenons aux F_{λ^k} en général. D'après l'hypothèse de récurrence, partie 1), on a

$$E = \bigoplus_{\lambda^k \in \Lambda^k} F_{\lambda^k}.$$

Comme chaque F_{λ^k} avait une base contenue dans B_{k+1}, E a donc une base contenue dans B_{k+1}.

On a vu d'autre part que toute base contenue dans B_{k+1} diagonalise a_{k+1}. Enfin, $B_k \supset B_{k+1}$ par définition, et donc d'après l'hypothèse de récurrence partie 2), toute base contenue dans B_{k+1} diagonalise $a_1, a_2, \dots, a_{k+1}$, ce qui achève la démonstration quand $\mathcal{A}$ est fini.

Si $\mathcal{A}$ est infini, soit $\mathcal{A}' = \{a_1, \dots, a_N\} \subset \mathcal{A}$ tel que $\mathcal{A}'$ soit une base du sous espace de $L(E)$ engendré par $\mathcal{A}$. Appliquons la première partie à $\mathcal{A}'$: si e est une base qui diagonalise les éléments de $\mathcal{A}'$, elle diagonalise les éléments de $\mathcal{A}$, qui sont des combinaisons linéaires de $\mathcal{A}'$, et la démonstration est complète.

VII Triangularisation et nilpotence

Il y a des des endomorphismes qu'il est impossible de diagonaliser (voir exemples 5.2 et 5.4). La prochaine catégorie d'endomorphismes à considérer sont ceux pour lesquels il existe une base dans laquelle la matrice représentative est triangulaire supérieure : les puissances successives d'une matrice triangulaire sont un peu plus difficiles à calculer que celles des matrices diagonales et beaucoup moins que celles d'une matrice quelconque. La proposition suivante en donne la caractérisation attendue : il faut et il suffit que le polynôme caractéristique soit scindé : par exemple tout endomorphisme sur un espace de dimension finie sur le corps $K = \mathbb{C}$ est triangularisable. La proposition recèle aussi un algorithme relativement rapide pour construire une base de triangularisation.

Proposition 7.1. Si E est de dimension finie et si $a \in L(E)$, alors il existe une base e telle que $[a]_e^e$ soit triangulaire supérieure si et seulement si le polynôme caractéristique P_a est scindé.

Démonstration : $\Rightarrow$ Si $A = [a]_e^e = (a_{ij})$ est triangulaire supérieure alors il est clair (voir la Proposition 2.3) que

$$P_a(X) = \det(A - XI_q) = (a_{11} - X) \dots (a_{qq} - X).$$

$\Leftarrow$ On procède par récurrence sur la dimension q de E. C'est trivial pour $q = 1$. Supposons le résultat vrai pour $q - 1$. Puisque P_a est scindé, a a au moins une valeur propre qu'on note a_{11} associée à un vecteur noté e_1. Complétons e_1 en une base $f = (e_1, f_2 \dots, f_q)$. Alors $[a]_f^f$ s'écrit par blocs

$$[a]_f^f = \begin{bmatrix} a_{11} & B \\ 0 & A \end{bmatrix}$$

avec A matrice carrée d'ordre $q - 1$ et B matrice ligne d'ordre $q - 1$. L'interprétation géométrique de A et B est la suivante : soit F le sous espace vectoriel de E engendré par $f' = (f_2, \dots, f_q)$, soit p la projection de E sur F parallèlement à Ke_q et soit $a_1 : F \to E$ la restriction de a à F. Alors $[a_1]_{f'}^e = \begin{bmatrix} B \\ A \end{bmatrix}$ et $[p \circ a_1]_{f'}^{f'} = A$. D'après la Proposition 2.3, $P_a(X) = (a_{qq} - X)\det(A - XI_{q-1})$. Donc

$$P_{p \circ a_1}(X) = \det(A - XI_{q-1})$$

est scindé lui aussi et on peut appliquer l'hypothèse de récurrence. Il existe donc une base e' de F telle que $[p \circ a_1]_{e'}^{e'} = T$ soit triangulaire. Notons $P = [\mathrm{id}_F]_{e'}^{f'}$. On a donc $P^{-1}AP = T$. La base $e = \{e_1\} \cup e'$ satisfait alors (Théorème 1.1) :

$$[a]_e^e = [\mathrm{id}_E]_f^e [a]_f^f [\mathrm{id}_E]_e^f = \begin{bmatrix} 1 & 0 \\ 0 & P^{-1} \end{bmatrix} \begin{bmatrix} a_{11} & B \\ 0 & A \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & P \end{bmatrix} = \begin{bmatrix} a_{11} & BP \\ 0 & T \end{bmatrix}$$

et la proposition est montrée.

Définition 7.1. Un endomorphisme n de l'espace de dimension finie E tel que il existe un entier $k > 0$ avec $n^k = 0$ est dit *nilpotent.*

Remarques : Un endomorphisme nilpotent n d'un espace de dimension q a un polynôme annulateur de la forme X^k. Son polynôme minimal est donc de la forme X^p et l'entier p est appelé l'*indice de nilpotence* de n. On a donc $n^{p-1} \neq 0$ et $n^p = 0$. On a aussi $p \leq q$ et donc toujours $n^q = 0$.

Proposition 7.2. Soit n un endomorphisme d'un espace vectoriel E de dimension finie. Alors n est nilpotent si et seulement si il existe une base f de E telle que $[n]_f^f$ soit triangulaire supérieure et telle que la diagonale de $[n]_f^f$ soit nulle.

Démonstration : $\Rightarrow$ Le polynôme minimal de m_n est donc une puissance de X. Donc le polynôme caractéristique de n est $P_n(X) = (-1)^q X^q$, il est scindé et on peut appliquer la proposition précédente. $\Leftarrow$ Pour $0 \leq k$ soit T_k^q l'espace des matrices $t = (t_{ij})_{1 \leq i,j \leq q}$ triangulaires supérieures sur K telles que $t_{ij} \neq 0$ entraîne $i - j = k$. Par exemple on obtient les matrices diagonales si $k = 0$ et la matrice nulle si $k \geq q$. Si $A \in T_k^q$ et $A' \in T_{k'}^q$ on voit que $AA' \in T_{k+k'}^q$ Si alors $N = [n]_f^f$ est triangulaire supérieure avec diagonale nulle, on écrit $N = \sum_{k=1}^{q-1} N_k$ avec $N_k \in T_k^q$. Alors $N^q = (\sum_{k=1}^{q-1} N_k)^q$ est somme de termes de la forme $N_{k_1} \ldots N_{k_q}$. Si $s = k_1 + \ldots + k_q$ alors $N_{k_1} \ldots N_{k_q}$ est dans T_s^q. Comme $k_j \geq 1$ on a $s \geq q$, et on a vu que le seul élément de T_s^q pour $s \geq q$ est 0. Donc $N^q = 0$ et n est nilpotent.

Pour conclure cette section, mentionnons une application de la triangulation à l'analyse :

Proposition 7.3. Si E est un espace complexe de dimension finie et soit $a \in L(E)$. Alors il existe une suite $a_n \in L(E)$ telle que $\lim_n a_n = a$ et a_n est diagonalisable. De plus, si a est inversible, on peut choisir les a_n inversibles.

Démonstration :

Si a est diagonalisable, on prend $a_n = a$. Sinon, soit f une base de triangularisation de b (qui existe toujours, puisque $K = \mathbb{C}$ et donc P_b est scindé). Soit $A = D + N = [a]_f^f$ avec D diagonale et N triangulaire supérieure avec diagonale nulle. Soit $J = \mathrm{diag}(1, 2, \ldots, q)$ et $A_n = A + \frac{1}{n}J$. Alors les valeurs propres de A_n sont distinctes à partir d'un certain rang et a_n défini par $A_n = [a_n]_f^f$ est diagonalisable. Il est clair que a_n tend vers a. De plus, la construction montre que si a est inversible, alors les a_n ci dessus sont inversibles à partir d'un certain rang.

Exercice 7.1. Utiliser la méthode de la démonstration de la Proposition 7.1 pour triangulariser un endomorphisme a dont une matrice représentative est $\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$.

Exercice 7.2. Soit $b \in E$ et $a \in L(E, K) = E^*$ tels que $a(b) = 0$. Calculer l'indice de nilpotence de $n \in L(E)$ défini par $n(x) = ba(x)$.

Exercice 7.3. Soit $a \in L(E)$ et un entier $k \geq 1$. Montrer que $\ker a^{k-1} \subset \ker a^k$. Montrer que si $\ker a^{k-1} = \ker a^k$ alors $\ker a^k = \ker a^{k+1}$ (Méthode : L'hypothèse entraîne que tout $y \in \ker a^k$ satisfait $a^{k-1}(y) = 0$. Appliquer cela à $y = a(z)$ si $z \in \ker a^{k+1}$). Si n est nilpotent d'indice de nilpotence p et $F_k = \ker n^k$, montrer que les $(F_k)_{k=0}^p$ sont tous distincts.

Exercice 7.4. 1) Utiliser la méthode de la démonstration de la Proposition 7.1 pour trianguler un endomorphisme a dont une matrice représentative est $A = \begin{bmatrix} 4 & -1 \\ 4 & 0 \end{bmatrix}$. Une réponse possible est

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

Posant $T = \begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix}$, calculer $T^n = \begin{bmatrix} 2^n & -a_n \\ 0 & 2^n \end{bmatrix}$. Méthode : montrer que $a_{n+1} = 2a_n + 2^n$ pour tout n avec $a_0 = 0$. Pour en déduire a_n, faire le changement de suite inconnue $a_n = 2^n b_n$. et déterminer facilement b_n. Calculer alors A^n à l'aide de tout ce qui précède. 2) Si $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ est inversible on considère l'application h_M de $\mathbb{R} \cup \{\infty\}$ dans lui même définie par $h_M(x) = \frac{ax+b}{cx+d}$ avec les conventions $h(M)(\infty) = a/c$ et $h_M(-d/c) = \infty$ si $c \neq 0$, et avec $h(M)(\infty) = \infty$ si $c = 0$. Vérifier que $h_M(h_{M_1}(x)) = h_{MM_1}(x)$ si M et M_1 sont inversibles, et en déduire que si $x_{n+1} = h_M(x_n)$ alors $x_n = h_{M^n}(x_0)$. 3) Si $x_{n+1} = 1 - \frac{1}{4x_n}$ calculer x_n en fonction de x_0. Méthode : appliquer le 2) à $M = A$ et utiliser 1). Voir aussi la fin de l'exercice 4.5 du chapitre 2.

VIII Espaces caractéristiques et décomposition de Dunford*

Nous allons maintenant perfectionner la Proposition 7.1, qui était la triangularisation du pauvre. Les bases de triangularisation sont loin d'être uniques. La décomposition de Dunford que nous allons exposer maintenant met en évidence une triangularisation plus intéressante. Voici d'abord une proposition technique :

Proposition 8.1. Soit E de dimension finie, a dans $L(E)$ tel que P_a soit scindé, soit $n_1, \ldots, n_p$ les multiplicités respectives des valeurs propres $\lambda_1, \ldots, \lambda_p$ dans le polynôme minimal m_a, et soit $m_1, \ldots, m_p$ les multiplicités respectives des valeurs propres $\lambda_1, \ldots, \lambda_p$ dans le polynôme caractéristique P_a. On considère les polynômes Q_j de degré $< n_j$ définis par

$$\frac{1}{m_a(X)} = \sum_{j=1}^{p} \frac{Q_j(X)}{(X - \lambda_j)^{n_j}} \tag{8.15}$$

et on pose $P_j(X) = Q_j(X)(X - \lambda_j)^{-n_j} m_a(X)$. Alors

1. Si $\delta_{ij} = 0$ pour $i \neq j$ et $\delta_{ii} = 1$ on a
$$\sum_{j=1}^{p} P_j(a) = \mathrm{id}_E, \quad P_i(a)P_j(a) = \delta_{ij}\mathrm{id}_E, \quad (a - \lambda_j \mathrm{id}_E)^{n_j} P_j(a) = 0. \tag{8.16}$$
2. Les $F_{\lambda_j} = \ker(a - \lambda_j \mathrm{id}_E)^{n_j}$, sont en somme directe et
$$E = F_{\lambda_1} \oplus \cdots \oplus F_{\lambda_j}.$$
De plus, $P_j(a)$ est la projection sur F_{λ_j} parallèlement à $\bigoplus_{i \neq j} F_{\lambda_i}$.
3. $F_{\lambda_j} = \ker(a - \lambda_j \mathrm{id}_E)^{m_j}$ et $\dim F_{\lambda_j} = m_j$.

Remarques : L'existence des Q_j est héritée de la *décomposition en éléments simples* de la fraction rationnelle $1/m_a$ vue en première année. En fait celle ci affirme l'existence de polynômes E_j de degrés n_j et de valuation > 0 tels que
$$\frac{1}{m_a(X)} = \sum_{j=1}^{p} E_j(\frac{1}{X - \lambda_j}).$$
On passe par réduction au même dénominateur de cette expression à (8.15).

Notons que $m_j \geq n_j$ par Cayley Hamilton et la définition de m_a. Notons que $n_j > 0$ par le Théorème 5.2 partie 2. Le sous espace $F_{\lambda_j} = \ker(a - \lambda_j \mathrm{id}_E)^{n_j}$ est appelé l'*espace caractéristique* associé à la valeur propre λ_j. Si on ne connait que le polynôme caractéristique de a et non son polynôme minimal, la partie 3 de la Proposition *permet de calculer* F_{λ_j}.

Démonstration : 1) En multipliant les deux côtés de (8.15) par m_a on obtient $\sum_{j=1}^{p} P_j(X) = 1$ et donc $\sum_{j=1}^{p} P_j(a) = \mathrm{id}_E$. Si $i \neq j$ alors $P_i(X)P_j(X)$ est divisible par m_a et donc $P_i(a)P_j(a) = 0$. Puisque $P_i(a) = \sum_{j=1}^{p} P_i(a)P_j(a)$ on en déduit $P_i(a) = P_i(a)^2$. Enfin $(X - \lambda_j)^{n_j} P_j(X)$ est un multiple de m_a et cela entraîne la troisième égalité de (8.16).
2) On applique le lemme des noyaux (Théorème 5.3) à $m_a(X) = (X - \lambda_1)^{n_1} \ldots (X - \lambda_p)^{n_p}$. Ensuite si $i \neq j$, alors $P_j(a)(F_{\lambda_i}) = \{0\}$ par définition de $P_j(a)$ et de F_{λ_j}. Enfin si $x \in F_{\lambda_j}$ alors en appliquant $\mathrm{id}_E = \sum_i P_i(a)$ à x on obtient $P_j(a)(x) = x$, ce qui montre que $P_j(a)$ est la projection indiquée.
3) On observe d'abord que F_{λ_j} est stable par a, c'est à dire que $a(F_{\lambda_j}) \subset F_{\lambda_j}$. En effet si $x \in F_{\lambda_j}$ alors $(a - \lambda_j \mathrm{id}_E)^{n_j}(x) = 0$. Appliquons a aux deux membres de cette égalité, et utilisons le fait que a et $(a - \lambda_j \mathrm{id}_E)^{n_j}$ commutent : on obtient $(a - \lambda_j \mathrm{id}_E)^{n_j}(a(x)) = 0$. C'est dire que $a(x) \in F_{\lambda_j}$. Appelons a_j la restriction a_j de a à F_{λ_j}. Choisissons alors une base e_j pour F_{λ_j} et notons $A_j = [a_j]_{e_j}^{e_j}$. Si $e = e_1 \cup \ldots \cup e_p$ alors $[a]_e^e$ est la matrice diagonale par blocs suivante :
$$[a]_e^e = \begin{bmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & A_p \end{bmatrix}. \tag{8.17}$$

Notons aussi par d_j la dimension de F_{λ_j} (rappelons qu'on cherche à montrer que $d_j = m_j$). Posons $B_j = A_j - \lambda_i I_{d_j}$. La définition de F_{λ_j} entraîne que $B_j^{n_j} = 0$ et donc que le polynôme minimal de a_j est une puissance de $X - \lambda_j$. Donc (Théorème 5.2, 2)) $P_{a_j}(X) = (X - \lambda_j)^{d_j}$. Puisque d'après (8.17) on a $P_{a_1}(X) \dots P_{a_p}(X) = P_a(X)$ on en tire

$$P_a(X) = (X - \lambda_1)^{d_1} \dots (X - \lambda_p)^{d_p} = (X - \lambda_1)^{m_1} \dots (X - \lambda_p)^{m_p},$$

et on a bien $d_j = m_j$. Le fait que $F_{\lambda_j} = \ker(a - \lambda_j \mathrm{id}_E)^{m_j}$ vient de $B_j^{m_j} = 0$, qui vient de $B_j^{n_j} = 0$ puisque $n_j \leq m_j$. La proposition est montrée.

Théorème 8.2. Soit E de dimension finie et a dans $L(E)$ tel que P_a soit scindé. Alors il existe un couple unique (d, n) d'endomorphismes de E tels que d soit diagonalisable, n soit nilpotent, $dn = nd$ et enfin $a = d + n$. Dans ces conditions d et n sont des polynômes d'endomorphismes de a, et il existe une base e de E telle que $[d]_e^e$ soit diagonale et $[n]_e^e$ soit triangulaire supérieure.

Remarques : Cette décomposition unique $a = d + n$ en diagonalisable plus nilpotent qui commutent est de nos jours appelée la *décomposition de Dunford* de a. En fait elle est plutôt due à Camille Jordan, né 100 ans plus tôt. Mais celui ci a encore beaucoup raffiné son théorème en fournissant une description très précise des différentes formes que peuvent prendre les endomorphismes nilpotents, conduisant à ce qu'on appelle la décomposition de Jordan. Toutefois, certains appellent Jordan ce que nous appelons Dunford ici. Finalement, nous appelerons *base de Dunford* de a une base telle que $[d]_e^e$ soit diagonale et $[n]_e^e$ soit triangulaire supérieure.

Si $[a]_f^f$ est triangulaire supérieure, cela n'entraîne pas que f soit une base de Dunford. En effet, supposons que D soit une matrice diagonale, que N soit une matrice triangulaire supérieure et à diagonale nulle. La condition nécessaire et suffisante pour que $DN = ND$ est facile à obtenir. Sans perte de généralité, on peut supposer que D s'écrit par blocs, avec des λ_j dans K tous distincts :

$$D = \begin{bmatrix} \lambda_1 I_{m_1} & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \lambda_p I_{m_p} \end{bmatrix}.$$

Ecrivons aussi $N = (N_{ij})_{1 \leq i,j \leq p}$ par blocs, tels que N_{ij} a m_i lignes et m_j colonnes. Alors $DN - ND = (((\lambda_i - \lambda_j)N_{ij})_{1 \leq i,j \leq p})$. Par conséquent $DN - ND = 0$ si et seulement si $N_{ij} = 0$ pour $i \neq j$. Par exemple, si un endomorphisme a d'un espace réel de dimension 4 est représenté par

$$[a]_e^e = A = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = D + N, \tag{8.18}$$

alors e n'est pas une base de Dunford : avec les notations précédentes on a $p = 2$, $\lambda_1 = 2$, $\lambda_2 = 1$, $m_1 = 3$, $m_2 = 1$ et $N_{12} = [0, 1, 0]^T \neq 0$, et donc D et N ne commutent pas.

Démonstration : 1) *Existence de la décomposition.* On garde les notations de la Proposition 8.1. On prend

$$d=\sum_{j=1}^{p}\lambda_j P_j(a)$$

et $n=a-d$. Il est clair que d et n étant des polynômes en a, ils commutent. Pour voir que d est diagonalisable, on prend une base e_j pour F_{λ_j} et $e=e_1\cup\ldots\cup e_p$. Alors $[d]_e^e$ est la matrice diagonale par blocs suivante :

$$[d]_e^e=\left[\begin{array}{cccc}\lambda_1 I_{m_1} & 0 & \ldots & 0\\ 0 & \lambda_2 I_{m_2} & \ldots & 0\\ \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & \ldots & \lambda_p I_{m_p}\end{array}\right]. \tag{8.19}$$

Pour voir que n est nilpotent on écrit grâce à la Proposition 8.1 partie 1 :

$$n=\sum_{j=1}^{p}(a-\lambda_j\mathrm{id}_E)P_j(a),\ldots,n^k=\sum_{j=1}^{p}(a-\lambda_j\mathrm{id}_E)^k P_j(a).$$

Comme (8.15) entraîne que si $k\geq m_j$ pour tout j on a $n^k=0$, la nilpotence de n est montrée.
2) *Unicité de la décomposition.* Si $a=d+n$, avec les d et n définis ci dessus, et si $a=d'+n'$ avec d' diagonalisable, n' nilpotent et $d'n'-n'd'=0$, montrons qu'alors $d=d'$ (et donc $n=n'$). D'abord, puisque d' et n' commutent, ils commutent avec leur somme a, et donc avec tout polynôme en a, donc avec d et n, qui sont des polynômes en a. Puisque n et n' commutent, et si $q=\dim E$, on applique la formule du binôme :

$$(d'-d)^{2q}=(n-n')^{2q}=\sum_{k=0}^{2q}(-1)^k C_{2q}^k n^{2q-k}n'^k=0;$$

la raison pour laquelle cette somme est nulle est amusante : si $0\leq k\leq 2q$, l'un des deux entiers k et $2q-k$ est $\geq q$ et comme $n^q=n'^q=0$ on a bien $(d'-d)^{2q}=0$. Invoquons maintenant le Théorème 6.1 : d et d' sont des endomorphismes diagonalisables qui commutent, il existe donc une base e telle que $D=[d]_e^e$ et $D'=[d']_e^e$ soient diagonales. Comme $D'-D$ est diagonale et que $(D'-D)^{2q}=0$ cela entraîne que $D'-D=0$ et l'unicité s'ensuit.
3) *Existence d'une base de Dunford.* Il est clair que F_{λ_j} est stable par n. Notons par n_j la restriction de n à F_{λ_j}. Il ne suffit pas que e_j soit une base de l'espace F_{λ_j} pour que $[n]_{e_j}^{e_j}$ soit triangulaire supérieure. Mais il suffit pour terminer d'appliquer la Proposition 7.2 à n_j : on crée ainsi une base f_j de F_{λ_j} telle que $N_j=[n_j]_{f_j}^{f_j}$ soit triangulaire supérieure. Puisque la restriction d_j de d à F_{λ_j} est $\lambda_j\mathrm{id}_{F_{\lambda_j}}$ on a $[d_j]_{f_j}^{f_j}=\lambda_j I_{m_j}$. Finalement $f=f_1\cup\ldots\cup f_p$ est une base de Dunford, puisque $[d]_f^f=[d]_e^e$ comme en (8.19) et que

$$[n]_f^f=\left[\begin{array}{cccc}N_1 & 0 & \ldots & 0\\ 0 & N_2 & \ldots & 0\\ \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & \ldots & N_p\end{array}\right].$$

Exemple 8.1. Traitons en détail la recherche des espaces caractéristiques, du polynôme minimal et d'une base de Dunford f pour l'exemple (8.18). Ces trois problèmes sont liés et *se traitent simultanément.* Il est clair que $P_a(X) = (2-X)^3(1-X)$. L'espace caractéristique F_2 associé à la valeur propre 2 est de dimension 3. Pour en trouver une base, il faut chercher successivement des bases de

$$\ker(a - 2\mathrm{id}_E) \subset \ker(a - 2\mathrm{id}_E)^2 \subset \ker(a - 2\mathrm{id}_E)^3,$$

et donc résoudre les systèmes linéaires homogènes gouvernés par les matrices $A - 2I_4$, $(A-2I_4)^2$, $(A-2I_4)^3$. On a

$$A - 2I_4 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad (A-2I_4)^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

En examinant le système linéaire homogène $(A - 2I_3)V = 0$, avec $V = [x, y, z, t]^T$, on voit que celui ci se réduit à $z = t = 0$. C'est dire qu'il est de rang 2 , que z et t sont les inconnues principales et x et y sont les inconnues non principales. L'espace des solutions est donc égal au nombre d'inconnues non principales et est de dimension 2. Au passage nous avons montré que *l'espace propre* E_2 est de dimension 2. Une base en est (e_1, e_2).

En examinant le système linéaire homogène $(A-2I_4)^2V = 0$, avec $V = [x, y, z, t]^T$, on voit que celui ci se réduit à $t = 0$. Il est de rang 1, t est l'inconnue principale et x, y, z sont les inconnues non principales. L'espace des solutions est donc égal au nombre d'inconnues non principales et est de dimension 3. Une base en est (e_1, e_2, e_3).

Comme on vient de voir que $\dim\ker(a - 2\mathrm{id}_E) = 2$ et que $\dim\ker(a - 2\mathrm{id}_E)^2 = 3$ qui est la multiplicité m_2 de la valeur propre 2, on en déduit que l'espace caractéristique $F_2 = \ker(a-2\mathrm{id}_E)^2$ et que la multiplicité de 2 dans le polynôme minimal est $n_2 = 2$. Il est en particulier inutile de calculer $(A-2I_4)^3$, puisque a priori $\ker(a-2\mathrm{id}_E)^2 = \ker(a-2\mathrm{id}_E)^3$.

On procède de même avec la valeur propre 1. C'est plus facile, car sa multiplicité dans P_a est 1, et donc $m_1 = n_1$: l'espace propre E_1 coincide avec l'espace caractéristique F_1. L'examen du système linéaire homogène $(A - I_4)V = 0$ montre que $F_1 = E_1$ est de dimension 1 (on s'en doutait) et on peut prendre comme base $f_4 = -e_2 + e_4$. Le polynôme minimal est $m_a(X) = (X-2)^2(X-1)$.

Si on note $f_1 = e_1, f_2 = e_2, f_3 = e_3$ alors $f = (f_1, f_2, f_3, f_4)$ est une base de Dunford. On a

$$P = [\mathrm{id}_E]^e_f = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \; [a]^f_f = P^{-1}AP = [d]^f_f + [n]^f_f = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Exercice 8.1. L'endomorphisme a d'un espace réel de dimension 3 est tel que dans une base e on ait

$$[a]_e^e = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Calculer le polynôme caractéristique P_a puis, en fonction de e, donner des bases des espaces propres et des espaces caractéristiques de a. En déduire le polynôme minimal et une base de Dunford de a.

Exercice 8.2. Soit E un espace complexe de dimension q et soit $\mathbf{Z}$ l'ensemble des entiers relatifs. On dit qu'une suite $(b_k)_{k\in\mathbf{Z}}$ d'éléments de $L(E)$ est *bornée* si pour toute base e de E il existe une constante $M > 0$ telle que pour tout $k \in \mathbf{Z}$, les modules des coefficients de la matrice $[b_k]_e^e$ sont $\leq M$. Soit $a \in L(E)$ inversible.

1. Si a est diagonalisable et si ses valeurs propres sont de module 1, montrer que la suite $(a^k)_{k\in\mathbf{Z}}$ est bornée.
2. Si a n'a qu'une valeur propre λ, et si la suite $(a^k)_{k\in\mathbf{Z}}$ est bornée montrer que $a = \lambda \mathrm{id}_E$ et que λ est de module 1 (Méthode : si f est une base de Dunford pour a, montrer que les coefficients $P_{ij}(k)$ de la matrice $[\lambda^{-k}a^k]_f^f$ sont des polynômes en k. Utiliser le fait que $(a^k)_{k\in\mathbf{Z}}$ est bornée pour montrer que les $P_{ij}(k)$ sont constants).
3. Si la suite $(a^k)_{k\in\mathbf{Z}}$ est bornée, montrer que a est diagonalisable et que ses valeurs propres sont de module 1 (Méthode : introduire une base de Dunford pour a et appliquer le 2).
4. Dans le cas particulier où $q = 2$, où $\det a = 1$ et où il existe une base e telle que $[a]_e^e$ soit réelle, montrer le théorème de Liapounov : $(a^k)_{k\in\mathbf{Z}}$ est bornée si et seulement si ou bien $a = \pm\mathrm{id}_E$ ou bien $\mathrm{trace}\, a \in]-2, 2[$.

IX Exponentielle d'un endomorphisme*

Cette section suppose que le cours de base K est $\mathbb{C}$ ou $\mathbb{R}$ et utilise l'analyse. On se donne un espace E de dimension finie sur K et on munit l'algèbre $L(E)$ sur K d'une *norme sous multiplicative*, c'est à dire d'une application $a \mapsto \|a\|$ de $L(E)$ dans $[0, \infty[$ telle que

1. Pour tous a et b dans $L(E)$ on a $\|a + b\| \leq \|a\| + \|b\|$ (*sous additivité*).
2. Pour $a \in L(E)$ et $\lambda \in K$ on a $\|\lambda a\| = |\lambda|\|a\|$ (*homogénéïté positive*).
3. $\|a\| = 0$ si et seulement si $a = 0$ (*séparation*).
4. Pour tous a et b dans $L(E)$ on a $\|ab\| \leq \|a\|\|b\|$ (*sous multiplicativité*).

Les trois premiers axiomes sont ceux d'un espace normé ordinaire, le quatrième n'a de sens que sur une algèbre. De telles normes existent. Pour le voir, par exemple, on munit E d'une norme p quelconque (son existence étant renvoyée au cours d'analyse) et on définit $\|a\| = \sup_{x\in E\setminus\{0\}} p(a(x))/p(x)$. C'est un exercice facile que de vérifier que les 4 axiomes sont vérifiés.

Théorème 9.1. Soit E un espace de dimension finie réel ou complexe et $a \in E$. Soit

$$Q_n(X) = 1 + X + \frac{1}{2}X^2 + \ldots + \frac{1}{n!}X^n. \tag{9.20}$$

Alors $\lim_{n\to\infty} Q_n(a) = \exp(a)$ existe dans $L(E)$ (On notera aussi $e^a = \exp a$). De plus si a et b commutent on a $\exp(a+b) = \exp a \exp b$. Enfin $\lim_{n\to\infty}(\mathrm{id}_E + \frac{1}{n}a)^n = \exp a$.

Démonstration : On munit $L(E)$ d'une norme sous multiplicative. Alors la suite $(Q_n(a))_{n\geq 0}$ a la propriété de Cauchy car si $n \leq n'$ on a

$$\begin{aligned}
\|Q_{n'}(a) - Q_n(a)\| &= \|\sum_{k=n+1}^{n'} \frac{1}{k!}a^k\| \\
&\leq \sum_{k=n+1}^{n'} \frac{1}{k!}\|a^k\| \\
&\leq \sum_{k=n+1}^{n'} \frac{1}{k!}\|a\|^k = Q_{n'}(\|a\|) - Q_n(\|a\|)
\end{aligned}$$

Comme la série $\sum_{k=0}^{\infty} \frac{1}{k!}\|a\|^k$ converge par le critère de D'Alembert, $(Q_n(\|a\|))_{n\geq 0}$ a la propriété de Cauchy et donc $(Q_n(a))_{n\geq 0}$ aussi. Comme un espace normé de dimension finie est toujours complet, la suite $(Q_n(a))_{n\geq 0}$ est donc convergente. Si alors a et b commutent on peut écrire

$$\begin{aligned}
Q_n(a+b) &= \sum_{i+j\leq n, 0\leq i, 0\leq j} \frac{a^i}{i!}\frac{b^j}{j!}, \\
Q_{2n}(a+b) - Q_n(a)Q_n(b) &= \sum_{i+j\leq 2n, n+1\leq i, n+1\leq j} \frac{a^i}{i!}\frac{b^j}{j!}, \\
Q_{2n}(a+b) - Q_n(a+b) &= \sum_{n+1\leq i+j\leq 2n, n+1\leq i, n+1\leq j} \frac{a^i}{i!}\frac{b^j}{j!}.
\end{aligned}$$

Par conséquent

$$\begin{aligned}
\|Q_{2n}(a+b) - Q_n(a)Q_n(b)\| &\leq Q_{2n}(\|a\|+\|b\|) - Q_n(\|a\|)Q_n(\|b\|) \\
&\leq Q_{2n}(\|a\|+\|b\|) - Q_n(\|a\|+\|b\|).
\end{aligned}$$

Comme pour $p = \|a\|+\|b\|$ on a $\lim_{n\to\infty} Q_{2n}(p) - Q_n(p) = 0$, on en déduit que $\exp(a+b) = \exp a \exp b$ si a et b commutent.

Finalement, si $R_n(X) = (1+\frac{X}{n})^n$, on constate que $Q_n(X) - R_n(X)$ est à coefficients positifs : celui de X^k est

$$\frac{1}{k!}\left(1 - (1-\frac{1}{n})(1-\frac{2}{n})\ldots(1-\frac{k-1}{n})\right).$$

Donc, en employant la sous multiplicativité de la norme on a

$$\|Q_n(a) - R_n(a)\| \leq Q_n(\|a\|) - R_n(\|a\|).$$

Comme on sait déjà que pour $p \geq 0$ on a $\lim_{n\to\infty} Q_n(p) - R_n(p) = e^p - e^p = 0$ on en déduit que $\lim_{n\to\infty} Q_n(a) - R_n(a) = 0$. Comme $\lim_{n\to\infty} Q_n(a) = e^a$, on en déduit que la limite de $R_n(a) = Q_n(a) - (Q_n(a) - R_n(a))$ existe et est $e^a - 0 = e^a$.

Corollaire 9.2. Si $a \in L(E)$ où E est un espace de dimension finie sur $K = \mathbb{C}$ ou $\mathbb{R}$, alors

1. $(e^a)^{-1} = e^{-a}$.
2. Si $[a]_e^e = \mathrm{diag}(\mu_1, \ldots, \mu_q)$ est diagonale, alors $\exp a = \mathrm{diag}(e^{\mu_1}, \ldots, e^{\mu_q})$.
3. Si a est nilpotent d'indice de nilpotence p alors
$$\exp a = \mathrm{id}_E + a + \frac{1}{2}a^2 + \ldots + \frac{1}{(p-1)!}a^{p-1}.$$
4. Si (d, n) est la décomposition de Dunford de a alors
$$e^a = e^d(\mathrm{id}_E + n + \frac{1}{2}n^2 + \ldots + \frac{1}{(p-1)!}n^{p-1}).$$
et la décomposition de Dunford de e^a est $(e^d,\ e^d(e^n - \mathrm{id}_E)$. De plus, si f est une base de Dunford pour a, c'est une base de Dunford pour e^a.
5. $\det e^a = \exp(\mathrm{trace}\, a)$.
6. Si A et P sont des matrices carrées d'ordre q et si P est inversible, alors
$$\exp(PAP^{-1}) = P(\exp A)P^{-1}.$$

Démonstration : Elle est très simple. 1) vient du fait que a et $-a$ commutent. 2) et 3) sont évidents. 4) vient du fait que d et n commutent, ce qui entraîne en particulier que e^d et $N = e^d(e^n - \mathrm{id}_E)$ commutent. Mais e^d est diagonalisable par le 2) et N est nilpotent car si $q = \dim E$ alors $(e^n - \mathrm{id}_E)^q = 0$ et donc $N^q = e^{qd}(e^n - \mathrm{id}_E)^q = 0$. Enfin 5) se voit si $K = \mathbb{C}$ à l'aide de la décomposition de Dunford, qui existe toujours puisqu'alors P_a est scindé : si (d, n) est la décomposition de Dunford de a il est clair que $\det a = \det d$ et $\mathrm{trace}\, a = \mathrm{trace}\, d$ (prendre une base de Dunford de a pour s'en convaincre). Appliquons ce principe à e^a et le 4) : cela entraîne que $\det e^a = \det e^d$ qui est lui même $\exp(\mathrm{trace}\, d)$ d'après le 2) et donc $\exp(\mathrm{trace}\, a)$. Pour traiter le cas $K = \mathbb{R}$, on utilise l'artifice suivant : on prend une base quelconque e de E et on considère la matrice réelle $A = [a]_e^e$ comme la matrice d'un endomorphisme de $\mathbb{C}^q$ dans la base canonique. Le résultat dans les complexes déjà obtenu entraîne $\det e^A = \exp \mathrm{trace}\, A$ et la démonstration du 5) est complète. Le 6) vient du fait que $(PAP^{-1})^k = PA^kP^{-1}$, donc que avec la notation (9.20) on a $Q_n(PAP^{-1}) = PQ_n(A)P^{-1}$ d'où le résultat par passage à la limite.

Exemple 9.1. Soit $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Alors $A^2 = -I_2$, et donc $A^3 = -A$ et $A^4 = I_2$. Soit $\theta \in \mathbb{R}$. Si Q_n est défini par (9.20) on voit que

$$Q_{4n-1}(\theta A) = (\sum_{k=0}^{n-1} \frac{(-1)^k}{(2k)!}\theta^{2k})I_2 + (\sum_{k=0}^{n-1} \frac{(-1)^k}{(2k+1)!}\theta^{2k+1})A \to_{n\to\infty} I_2 \cos\theta + A\sin\theta.$$

D'où le résultat dont nous reparlerons au moment du groupe orthogonal (Chapitre 2, section 7)

$$\exp \begin{bmatrix} 0 & -\theta \\ \theta & 0 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = R(\theta).$$

Exemple 9.2. Soit $a, b \in (0,1)$. Il existe $\alpha, \beta \geq 0$ tels que

$$\begin{bmatrix} a & 1-a \\ 1-b & b \end{bmatrix} = \exp \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix}$$

si et seulement si $a + b > 1$. En effet on écrit

$$S = \begin{bmatrix} a & 1-a \\ 1-b & b \end{bmatrix}, \quad A = \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix}$$

et on voit que les valeurs propres de A sont 0 et $-\alpha - \beta$. Si $\alpha + \beta = 0$ alors $A = 0$ et $S = I_2$ a des coefficients nuls. Donc on peut supposer $\alpha + \beta > 0$. Notons $\gamma = e^{-\alpha-\beta}$ et

$$P = \begin{bmatrix} 1 & -\alpha \\ 1 & \beta \end{bmatrix}, \ P = \frac{1}{\alpha+\beta}\begin{bmatrix} \beta & \alpha \\ -1 & 1 \end{bmatrix}, \ D = \begin{bmatrix} 0 & 0 \\ 0 & -\alpha-\beta \end{bmatrix}, \ \exp D = \begin{bmatrix} 1 & 0 \\ 0 & \gamma \end{bmatrix}.$$

Nous obtenons $A = PDP^{-1}$, $S = \exp A = P(\exp D)P^{-1}$ et donc

$$S = \begin{bmatrix} a & 1-a \\ 1-b & b \end{bmatrix} = \frac{1}{\alpha+\beta}\begin{bmatrix} \beta+\alpha\gamma & \alpha-\alpha\gamma \\ \beta-\beta\gamma & \alpha+\beta\gamma \end{bmatrix}$$

Par conséquent

$$a = \frac{\beta+\alpha\gamma}{\alpha+\beta}, \ b = \frac{\beta+\alpha\gamma}{\alpha+\beta}, \ a+b = 1+\gamma > 1.$$

Inversement, étant donnés a et b dans $(0,1)$ tels que $\gamma = a+b-1 \in (0,1)$ on obtient

$$\alpha = \frac{-(1-a)\log\gamma}{1-\gamma}, \ \beta = \frac{-(1-b)\log\gamma}{1-\gamma}.$$

Si E est complexe, on peut montrer que l'image de $L(E)$ par $a \mapsto e^a$ est le groupe linéaire $GL(E)$ des automorphismes de E. Nous omettrons la démonstration, assez délicate, et l'étude de $e^a = b$ comme équation en a. Cette application n'est certes pas injective : voici une agréable application de Dunford qui étudie le cas $b = \mathrm{id}_E$.

Proposition 9.3. Si E est un espace complexe de dimension finie, alors $e^a = \mathrm{id}_E$ si et seulement si a est diagonalisable et si ses valeurs propres sont dans $2i\pi\mathbf{Z}$.

Démonstration : Si $e^a = \mathrm{id}_E$, soit (d,n) la décomposition de Dunford de a. D'après le Corollaire 9.3 la décomposition de Dunford de e^a est $(e^d, e^d(e^n - \mathrm{id}_E))$. Comme elle est unique cela entraîne $e^d = \mathrm{id}_E$ et $e^d(e^n - \mathrm{id}_E) = 0$, ou $e^n - \mathrm{id}_E = 0$. Mais le Corollaire 9.3 entraîne que si $n \neq 0$, cad si son indice de nilpotence est $p \geq 2$ alors $e^n - \mathrm{id}_E = nm = 0$ avec
$$m = \mathrm{id}_E + \cdots + \frac{1}{(p-2)!}n^{p-2}.$$
Mais la décomposition de Dunford de m est $(\mathrm{id}_E, m - \mathrm{id}_E)$ donc m est inversible et $nm = 0$ entraîne $n = 0$ et a est diagonalisable. Soit e une base de diagonalisation de a, soit $[a]_e^e = D = \mathrm{diag}(z_1, \ldots, z_q)$. Alors $e^D = I_q$. Les seules solutions de l'équation dans les complexes $e^z = 1$ sont les z de $2i\pi\mathbf{Z}$: pour le voir on écrit $z = x + i\theta$ avec x et θ réels. Comme $1 = |e^z|^2 = e^{x+i\theta}e^{x-i\theta} = e^{2x}$, on a $x = 1$. D'après la formule de Moivre $e^z = \cos\theta + i\sin\theta = 1$, et donc $\cos\theta = 1$, ce qui conduit au résultat.

Mentionnons pour terminer le rôle joué par l'exponentielle des endomorphismes dans la solution des systèmes différentiels à coefficients constants :

Proposition 9.4. Soit E un espace réel ou complexe de dimension finie et $a \in L(E)$. Alors la dérivée de l'application $t \mapsto e^{at}$ de $\mathbb{R}$ dans $L(E)$ est $ae^{at} = e^{at}a$. Si I est un intervalle contenant 0 et $Y : I \to E$ est dérivable, alors $Y'(t) = a(Y(t))$ sur I si et seulement si $Y(t) = e^{at}(Y(0))$ sur I.

Démonstration : $\frac{d}{dt}e^{at}ae^{at} = e^{at}a$ résulte de la dérivation de la série entière en t à valeurs dans $L(E)$ définie par
$$\sum_{k=0}^{\infty}\frac{a^k}{k!}t^k.$$
Pour la suite, $\Rightarrow$ se montre en considérant $Y_1(t) = e^{-at}(Y(t))$ et en vérifiant que sa dérivée est nulle sur I :
$$Y_1'(t) = -e^{-at}a(Y(t)) + e^{-at}(Y'(t)) = -e^{-at}a(Y(t)) + e^{-at}(a(Y(t))) = 0.$$
Comme I est un intervalle, $Y_1(t) = Y_1(0) = Y(0)$. $\Leftarrow$ se voit directement car si $Y(t) = e^{at}(Y(0))$ alors $Y'(t) = ae^{at}(Y(0)) = a(Y(t))$.

L'analogie avec l'exponentielle réelle suggérée par l'énoncé précédent ne s'étend pas très loin : la proposition suivante montre que la dérivée de $t \mapsto e^{a_t}$ n'est certes pas $e^{a_t}(a_t)'$.

Proposition 9.5. Soit E un espace réel ou complexe de dimension finie, I un intervalle ouvert et $t \mapsto a_t$ une fonction continûment dérivable de I à valeurs dans $L(E)$. Alors
$$\frac{d}{dt}e^{a_t} = \int_0^1 e^{(1-u)a_t}(\frac{da_t}{dt})e^{ua_t}du. \tag{9.21}$$

Démonstration : Commencons par observer que si a et $h \in L(E)$ alors on a pour tout s réel

$$e^{-as}e^{(a+h)s} - \mathrm{id}_E = \int_0^s e^{-au}he^{(a+h)u}du. \tag{9.22}$$

Pour le voir, il suffit de dériver chaque membre par rapport à s. En application de la Proposition 9.4, la dérivée du premier membre est

$$-e^{-as}ae^{(a+h)s} + e^{-as}(a+h)e^{(a+h)s} = e^{-as}he^{(a+h)s}.$$

En application du théorème fondamental du calcul intégral, la dérivée du second membre est aussi $e^{-as}he^{(a+h)s}$. Les deux membres ont la même dérivée dans $\mathbb{R}$ et coincident en $s = 0$. Il y a donc égalité pour tout s dans (9.22). En faisant $s = 1$ dans (9.22) et en multipliant à gauche par e^a on obtient

$$e^{a+h} - e^a = \int_0^1 e^{a(1-u)}he^{(a+h)u}du.$$

Soit maintenant t et $t+k$ dans I et appliquons la dernière égalité à $a = a_t$ et $h = a_{t+k} - a_t$. Après division par k on obtient

$$\frac{1}{k}(\exp a_{t+k} - \exp a_t) = \int_0^1 e^{(1-u)a_t}\frac{1}{k}(a_{t+k} - a_t)e^{ua_{t+k}}du.$$

Faisons enfin tendre k vers 0. Le second membre tend vers le second membre de (9.21) car la fonction $(u,k) \mapsto e^{(1-u)a_t}\frac{1}{k}(a_{t+k} - a_t)e^{ua_{t+k}}$ est continue sur $[0,1] \times (I-t)$. La proposition est donc montrée.

Exercice 9.1. Calculer pour x et θ réels la matrice $\exp\begin{bmatrix} x & -\theta \\ \theta & x \end{bmatrix}$ (Méthode : écrire la matrice sous la forme $xI_2 + \theta A$ où A est comme à l'exemple 9.1).

Exercice 9.2. Si a et b dans $L(E)$ sont tels que pour tout t réel on a $e^{(a+b)t} = e^{at}e^{bt}$ montrer que a et b commutent (Méthode : à l'aide de la Proposition 9.4, dériver deux fois et faire $t = 0$).

Exercice 9.3. Si $a \in L(E)$ montrer que

$$\lim_{h \to 0} \frac{1}{h}\det(\mathrm{id}_E + ha) = \mathsf{trace}\, a.$$

A partir de $\lim_{n\to\infty}(\mathrm{id}_E + \frac{1}{n}a)^n = \exp a$, en déduire une autre démonstration de $\det e^a = \exp \mathsf{trace}\, a$.

Exercice 9.4. Si $a \in L(E)$ soit $y_1 : \mathbb{R} \to L(E)$ et $y_2 : \mathbb{R} \to L(E)$ dérivables telles que $y_1'(t) = ay_1(t)$ et $y_2'(t) = y_2(t)a$ pour tout $t \in \mathbb{R}$. Montrer que $y_1(t) = e^{at}y_1(0)$ et que $y_2(t) = y_2(0)e^{at}$.

Exercice 9.5. (Matrices stochastiques et exponentielle) On note $\mathcal{S}_n$ l'ensemble des matrices carrées d'ordre n réelles $S = (s_{ij})$ telles que $s_{ij} \geq 0$ pour tous $i, j = 1, \ldots, n$ et telles que

$\sum_{j=1}^n s_{ij} = 1$ pour tout $i = 1, \ldots, n$. (Ces matrices servent en Calcul des Probabilités et sont dites stochastiques). On note $\mathcal{A}_n$ l'ensemble des matrices carrées réelles $A = (a_{ij})$ telles que $a_{ij} \geq 0$ pour tous $i \neq j = 1, \ldots, n$ et telles que $\sum_{j=1}^n a_{ij} = 0$ pour tout $i = 1, \ldots, n$. On note $\mathcal{P}_n$ les matrices e^A avec $A \in \mathcal{A}_n$.

1. Si $\mathbf{1}_n$ est le vecteur colonne d'ordre n formé de 1 montrer que $S\mathbf{1}_n = \mathbf{1}_n$ et $A\mathbf{1}_n = 0$ pour $S \in \mathcal{S}_n$ et $A \in \mathcal{A}_n$.
2. Si S et S' sont dans $\mathcal{S}_n$ montrer à l'aide du 1 que $SS' \in \mathcal{S}_n$.
3. Montrer que $\mathcal{P}_n \subset \mathcal{S}_n$. (Méthode : si $A \in \mathcal{A}_n$ prendre $\lambda \geq \max_i(-a_{ii})$ et montrer que $e^{\lambda I_n + A}$ et $e^{-\lambda I_n} e^{\lambda I_n + A}$ sont à coefficients positifs ou nuls).
4. Soit $S \in \mathcal{S}_2$. Montrer à l'aide de l'exemple 9.2 que $S \in \mathcal{P}_2$ si et seulement si trace $S > 1$ (on ne connait pas de caractérisation simple de $\mathcal{P}_n$ pour $n \geq 3$).
5. Soit $S_t \in \mathcal{S}_n$ pour tout $t \geq 0$ de la forme $S = e^{tA}$ pour une matrice carrée A. Montrer qu'alors $A \in \mathcal{A}_n$ (Méthode : examiner $A = \lim_{t \to 0} \frac{1}{t}(S_t - I_n)$).

Exercice 9.6. (Semi groupes de matrices stochastiques) On garde les notations de l'exercice précédent. Soit l'application $t \mapsto S_t$ de $[0, \infty[$ dans $\mathcal{S}_n$ telle que $S_t S_s = S_{t+s}$ pour tous t et $s \geq 0$, telle que $t \mapsto S_t$ soit continue et telle que $S_0 = I_n$. On veut montrer qu'alors il existe $A \in \mathcal{A}_n$ telle que $S_t = e^{tA}$ pour tout $t \geq 0$ (voir aussi les exercices 9.8 (3) et 9.7). On considère pour cela pour tout $\lambda > 0$ la matrice à coefficients ≥ 0

$$R_\lambda = \int_0^\infty e^{-t\lambda} S_t dt.$$

1. Pourquoi l'intégrale définissant R_λ est elle convergente ?
2. Montrer que $R_\lambda - R_\mu = (\mu - \lambda) R_\lambda R_\mu$ (Méthode : faire le changement de variable $(t, s) \mapsto (t, u) = (t, t + s)$ dans l'intégrale double qui représente $(\mu - \lambda) R_\lambda R_\mu$).
3. Notons $E_\lambda \subset \mathbb{R}^n$ l'espace image de R_λ. En écrivant $R_\lambda = (I_n + (\mu - \lambda) R_\lambda) R_\mu$ montrer que $E_\lambda \subset E_\mu$. Comme symétriquement $E_\mu \subset E_\lambda$ en déduire que $E = E_\lambda$ ne dépend pas de λ.
4. Montrer que $\lim_{\lambda \to \infty} \lambda R_\lambda = I_n$. Pour cela on prend pour norme sur l'espace vectoriel des matrices carrées d'ordre n le nombre $\|M\| = \max_{ij} m_{ij}$ et un nombre arbitraire t_0 pour écrire
$$\begin{aligned} \|I_n - \lambda R_\lambda\| &= \|\int_0^\infty \lambda e^{-t\lambda}(I_n - S_t) dt\| \\ &\leq \int_0^{t_0} \lambda e^{-t\lambda} \|I_n - S_t\| dt + \int_{t_0}^\infty \lambda e^{-t\lambda} dt. \end{aligned}$$
En déduire que $E = \mathbb{R}^n$ et donc que R_λ^{-1} existe.
5. Soit $\lambda > 0$ fixé. Soit $y \in \mathbb{R}^n$ et $x = R_\lambda(y)$. Montrer que
$$\lim_{t \to 0} \frac{1}{t}(S_t - I_n) x = \lambda x - y.$$

Ecrire pour cela

$$\begin{aligned}\frac{1}{t}(S_t - I_n)R_\lambda(y) &= \frac{1}{t}\int_0^\infty e^{-s\lambda}S_{t+s}(y)ds - \frac{1}{t}\int_0^\infty e^{-s\lambda}S_s(y)ds \\ &= \frac{e^{t\lambda}}{t}\int_t^\infty e^{-s\lambda}S_{t+s}(y)ds - \frac{1}{t}\int_0^\infty e^{-s\lambda}S_s(y)ds \\ &= \frac{e^{t\lambda}-1}{t}\int_t^\infty e^{-s\lambda}S_{t+s}(y)ds - \frac{1}{t}\int_0^t e^{-s\lambda}S_s(y)ds.\end{aligned}$$

En déduire que $\lim_{t\to 0}\frac{1}{t}(S_t - I_n) = A$ avec $A = \lambda I_n - R_\lambda^{-1}$. Pourquoi A ne dépend il pas de λ?

6. En écrivant
$$\frac{1}{s}(S_{s+t} - S_t) = \frac{1}{s}(S_s - I_n)S_t = S_t\frac{1}{s}(S_s - I_n),$$
montrer que $AS_t = S_tA$ pour tout t. A l'aide de la question 4 montrer que $t \mapsto S_t$ est dérivable et que $\frac{d}{dt}S_t = AP_t = P_tA$. En déduire l'existence d'une matrice C telle que $S_t = Ce^{tA}$. Utiliser $S_0 = I_n$ pour montrer $C = I_n$. Utiliser l'exercice précédent pour montrer que $A \in \mathcal{A}_n$.

Exercice 9.7. On pose $S(a,b) = \begin{bmatrix} a & 1-a \\ 1-b & b \end{bmatrix}$. Soit $\alpha, \beta \geq 0$ tels que $\alpha + \beta > 0$. On pose $\gamma = e^{-\alpha-\beta}$ et

$$S_t = S(\frac{\beta + \alpha\gamma^t}{\alpha+\beta},\ \frac{\alpha+\beta\gamma^t}{\alpha+\beta}).$$

Montrer de deux manières que $S_tS_s = S_{t+s}$ pour tous $s, t \geq 0$ (calcul direct ou exemple 9.2).

Exercice 9.8. Soit E un espace vectoriel réel ou complexe de dimension finie. Soit $b \in L(E)$. On considère l'intégrale

$$I(b) = \int_0^\infty e^{-tb}dt = \lim_{t_0\to\infty}\int_0^{t_0} e^{-tb}dt$$

à valeurs dans $L(E)$. Si cette limite dans $L(E)$ n'existe pas, on dit que l'intégrale diverge.

1. Montrer que si b n'est pas inversible, l'intégrale diverge (considérer $v \in E \setminus \{0\}$ tel que $b(v) = 0$ et montrer que $e^{-tb}(v) = v$).
2. Si b^{-1} existe montrer avec la proposition 9.4 que $\int_0^{t_0} e^{-tb}dt = b^{-1} - b^{-1}e^{-t_0b}$. Avec le corollaire 9.2 (4) montrer que $\lim_{t_0\to\infty} e^{-t_0b}$ existe dans E si et seulement si toutes les valeurs propres de b ont une partie réelle strictement positive. Quelle est alors la valeur de $I(b)$?
3. Avec les notations des exercices 9.5 et 9.6, déduire du (2) que si $S_t = e^{tA}$ avec $A \in \mathcal{A}_n$ alors $R_\lambda = (\lambda I_n - A)^{-1}$ et que toutes les valeurs propres de A ont une partie réelle ≤ 0.
4. Plus généralement, pour $p > 0$ et pour $b \in L(E)$ dont toutes les valeurs propres ont une partie réelle strictement positive on considère

$$I_p(b) = \frac{1}{\Gamma(p)}\int_0^\infty e^{-tb}t^{p-1}dt$$

avec $\Gamma(p) = \int_0^\infty e^{-t}t^{p-1}dt$. Montrer que $I_p(b)$ converge. Si de plus p est entier montrer que $I_p(b) = b^{-p}$. Si p et $q > 0$ ne sont pas nécessairement entiers montrer en imitant le (2) de l'exercice 9.6 que $I_p(b)I_q(b) = I_{p+q}(b)$ (Si p n'est pas entier on peut prendre $I_p(b)$ comme définition de "puissance fractionnaire" de l'endomorphisme b).

Chapitre 2

Espaces euclidiens

I Particularités des espaces réels.

Dans le chapitre 1, nous avons obtenu des résultats d'une belle généralité, valables pour des espaces vectoriels de dimension finie sur un corps quelconque. *Jamais* pourtant nous n'y avons fait allusion à des objets d'intuition géométrique courante : distance entre deux points, perpendicularité, angles : l'axiomatique était trop pauvre. Le présent chapitre est beaucoup plus proche de l'espace physique. La présente section est un prélude qui met en évidence trois particularités des espaces réels de dimension finie : un endomorphisme d'un espace vectoriel sur $\mathbb{R}$ de dimension *impaire* a toujours un vecteur propre ; un endomorphisme d'espace réel a toujours un espace stable de dimension 1 ou 2 ; un espace réel peut être orienté.

Proposition 1.1. Soit E un espace réel de dimension finie q et $a \in L(E)$. Alors a a au moins une valeur propre si q est impair, et au moins un espace stable de dimension 2 si q est pair.

Démonstration : Si q est impair, le polynôme caractéristique P_a tend vers $+\infty$ si X tend vers $-\infty$ et P_a tend vers $-\infty$ si X tend vers ∞ : il a donc au moins une racine réelle. Si q est pair, soit e une base quelconque de E et $A = [a]_e^e$. Considérons A comme la matrice représentative d'un endomorphisme b de l'espace complexe $\mathbb{C}^q$ exprimé dans la base canonique f. Alors $P_a = P_A = P_b$. Comme P_b est scindé dans $\mathbb{C}$, l'endomorphisme b admet au moins une valeur propre $\lambda = \alpha + i\beta$ avec α et β réels, à laquelle est associé un vecteur propre v. Si $V = X + iY = [v]_f^f$ avec X et Y matrices colonnes réelles d'ordre q l'égalité $b(v) = \lambda v$ se traduit, en identifiant parties réelle et imaginaire

$$AV = \lambda V, \;\; A(X + iY) = (\alpha + i\beta)(X + iY), \; AX = \alpha X - \beta Y, \; AY = \beta X + \alpha Y.$$

Soit enfin x et y dans E définis par $[x]_e^e = X$ et $[y]_e^e = y$. Les égalités ci dessus deviennent $a(x) = \alpha x - \beta y, \; a(y) = \beta x + \alpha y$, ce qui est dire que le sous espace F de E engendré par x et y est stable par a. Il ne peut être de dimension 0, car $X = Y = 0$ entraînerait $V = 0$, ce qui n'est pas, car un vecteur propre est toujours non nul. Si x et y sont indépendants alors F est de dimension 2 comme annoncé. Si x et y sont colinéaires, supposons $x \neq 0$. C'est dire que x est un vecteur propre de a et que $\beta = 0$. On prend une base $g = (g_1, \ldots, g_q)$ de E telle que $g_1 = x$. Alors, avec L matrice ligne d'ordre $q-1$ et A_1 matrice carrée d'ordre $q-1$ on écrit par blocs

$$[a]_g^g = \begin{bmatrix} \alpha & L \\ 0 & A_1 \end{bmatrix}.$$

Comme A_1 est d'ordre impair, elle a au moins une valeur propre réelle, et un vecteur propre X_1 correspondant. Si $X_1 = [t_2, \ldots, t_q]$ alors $x_1 = t_2 g_2 + \cdots + t_q g_q$ est un vecteur propre de a indépendant de x. Le raisonnement quand $x = 0$ et $y \neq 0$ est similaire. La proposition est montrée.

Définition 1.1. Soit E un espace *réel* de dimension finie, et soit $\mathcal{B}$ l'ensemble de toutes ses *bases-ordonnées*. On dit que e et e' dans $\mathcal{B}$ ont *même orientation* , ce qu'on note $e \sim e'$ si le déterminant de la matrice de changement de base $P = [\mathrm{id}_E]_{e'}^{e}$ est *positif.*

Remarques : L'idée de base-ordonnée d'un espace vectoriel E de dimension finie mérite un rappel. Une base (tout court) $e = \{e_1, \ldots, e_q\}$ de E est une *partie* de E, libre et génératrice. Une base-ordonnée $e = (e_1, \ldots, e_q)$ est une *suite* d'éléments de E distincts telle que l'ensemble associé $\{e_1, \ldots, e_q\}$ est une base. L'ordre compte donc : si $q = 3$ la base-ordonnée $e = (e_1, e_2, e_3)$ est distincte de (e_2, e_1, e_3) alors que $\{e_1, e_2, e_3\} = \{e_2, e_1, e_3\}$. Il faut remarquer que si on a besoin d'écrire une matrice représentative, on utilise implicitement plutôt une base-ordonnée.

Proposition 1.2. La relation $\sim$ sur $\mathcal{B}$ est une relation d'équivalence, et l'ensemble quotient $\mathcal{B}/\sim$ a deux éléments si $\dim E > 0$.

Démonstration : *Réflexivité :* $e \sim e$ trivialement car $[\mathrm{id}_E]_e^e = I_q$ et $\det I_q = 1 > 0$. *Symétrie :* si $e \sim e'$, alors $\det[\mathrm{id}_E]_{e'}^{e} > 0$ par définition. Mais $[\mathrm{id}_E]_e^{e'} = ([\mathrm{id}_E]_{e'}^{e})^{-1}$ et donc $\det[\mathrm{id}_E]_e^{e'} = (\det[\mathrm{id}_E]_{e'}^{e})^{-1} > 0$, d'où $e' \sim e$. *Transitivité :* Si $e \sim e'$ et $e' \sim e''$ alors d'après le Théorème 1.1

$$\det[\mathrm{id}_E]_{e''}^{e} = \det[\mathrm{id}_E]_{e'}^{e} \det[\mathrm{id}_E]_{e''}^{e'} > 0 \tag{1.1}$$

et donc $e \sim e''$.

Pour voir qu'il y a au plus deux classes d'équivalence, on suppose que e et e' sont dans deux classes différentes. Donc $\det[\mathrm{id}_E]_{e'}^{e} < 0$ Si alors $e'' \not\sim e'$ alors $\det[\mathrm{id}_E]_{e''}^{e'} < 0$ aussi et (1.1) est encore vrai, ce qui montre $e'' \sim e$: il n'y a pas plus de deux classes. Enfin il y en a toujours deux, car $(-e_1, e_2, \ldots, e_q) \not\sim e$ si $q > 0$.

Définition 1.2. Soit E un espace *réel* de dimension finie. On dit qu'il est *orienté* si on a sélectionné une des deux classes de $\mathcal{B}$. Les éléments de cette classe sont appelés bases-ordonnées *directes*, les autres sont les bases-ordonnées *indirectes.*

Remarques : L'orientation d'un espace a un caractère arbitraire. On a réalisé que la communication avec des civilisations extra terrestres ne permettrait pas d'expliquer nos notions de droite et de gauche, qui sont culturelles et transmises par la tradition : elles concernent l'orientation d'un espace réel de dimension 1. Considérons ensuite le cas $q = 2$ et soit donc une base ordonnée (e_1, e_2) du plan E. La droite $\mathbb{R}e_1$ partage E en deux parties. Plus précisément, si a est une forme linéaire non nulle sur E telle que $a(e_1) = 0$, appelons E_+ le demi plan des $x \in E$ tels que $a(x) > 0$ et E_- le demi plan des $x \in E$ tels que $a(x) < 0$. C'est un exercice de voir que $(e_1, e_2) \sim (e_1, e'_2)$ si et seulement si $a(e_2)$ et $a(e'_2)$ ont même signe, autrement dit si e_2 et e'_2 sont tous deux dans E_+ ou tous deux dans E_-. *Traditionnellement,* on regarde le plan de sorte que e_1 soit horizontal et soit dirigé vers la droite et on place e_2 dans le demi plan supérieur pour avoir une base directe. Toutefois, il faut réaliser que ces choix utilisent des notions non définies mathématiquement : "regarder", "horizontal", "la droite". Dans le cas $q = 3$, l'orientation traditionnelle est fournie par la règle du *bonhomme d'Ampère* : adossé à e_3, qui lui rentre par les pieds et ressort par la tête, le bonhomme voit e_1 à sa droite et e_2 à sa gauche si (e_1, e_2, e_3) est directe. D'autres règles comme celle *des trois doigts* (pouce $= e_1$, index $=e_2$, et majeur $=e_3$ de la main *droite* fournissent un trièdre positif) ou celle du *tirebouchon de Maxwell* fournissent la même orientation.

Exercice 1.1. Soit $e = (e_1, \ldots, e_q)$ une base-ordonnée de E, soit $\sigma \in \mathcal{S}_q$ une permutation de q objets et soit $e_\sigma = (e_{\sigma(1)}, \ldots, e_{\sigma(q)})$. Montrer que $e \sim e_\sigma$ si et seulement si la signature $\epsilon(\sigma)$ est 1 (Méthode : Utiliser la définition du déterminant par signatures).

Exercice 1.2. Soit la matrice réelle $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ telle que $bc > 0$. Montrer que l'endomorphisme de $\mathbb{R}^2$ associé est diagonalisable.

II Produit scalaire, polarisation, parallélogramme

Pour quelques pages (sections 2,3,4), nous n'allons plus nous limiter aux les espaces réels de dimension finie, pour considérer aussi des espaces vectoriels de dimension infinie. Les notions de produit scalaire et de projection orthogonale qu'on va introduire sont très utiles même en dimension infinie, par exemple dans l'étude des séries de Fourier. Il y a tellement de définitions au début qu'il vaut mieux les donner toutes et les commenter ensuite.

Définitions 2.1. Soit E et F des espaces vectoriels sur un corps K. Une *forme bilinéaire* B sur $E \times F$ est une application de $E \times F$ dans K :

$$(x, y) \mapsto B(x, y),$$

telle que pour tout $x \in E$ fixé l'application $y \mapsto B(x, y)$ soit une forme linéaire sur F, et telle que pour tout $y \in F$ fixé l'application $x \mapsto B(x, y)$ soit une forme linéaire sur E. Si de plus $E = F$, la forme bilinéaire est dite *symétrique* si pour tous x et y de E on a $B(x, y) = B(y, x)$. Par abus de langage, on parle alors de forme bilinéaire symétrique sur E plutôt que sur $E \times E$. Dans ces conditions, la fonction Q_B sur E définie par $x \mapsto Q_B(x) = B(x, x)$ est appelée la *forme quadratique* associée à la forme bilinéaire symétrique B. Si de plus $K = \mathbb{R}$, la forme bilinéaire symétrique B est dite *positive* si pour tout $x \in E$ on a $Q_B(x) \geq 0$. Elle est dite *définie positive* si pour tout $x \in E\setminus$ on a $Q_B(x) > 0$. Une forme bilinéaire symétrique définie positive sur $E \times E$ est dite plus brièvement *produit scalaire*[1] sur l'espace réel E. Un espace vectoriel réel dans lequel on a fixé un produit scalaire est appelé un espace *préhilbertien réel.* Si B est ce produit scalaire, le nombre $\|x\| = \sqrt{Q_B(x)}$ s'appelle la *norme* du vecteur x. Le produit scalaire $B(x,y)$ vecteurs x et y est plus souvent noté $\langle x, y\rangle$. Un espace préhilbertien réel de dimension finie s'appelle un espace *euclidien.*

Exemples 2.1. Si $E = K^2$ et $F = K^3$, alors $B_1(x,y) = B_1(x_1,x_2;y_1,y_2,y_3) = 3x_1y_1 - 5x_2y_3$ est bilinéaire sur $E\times F$. Si $E = F = K^2$ alors $B_2(x,y) = 3x_1y_1 - 5x_2y_1$ est bilinéaire sur $E \times E$ mais non symétrique car $B(y,x) = 3x_1y_1 - 5y_2x_1 \neq 3x_1y_1 - 5x_2y_1$. Soit maintenant $E = \mathbb{R}^2$. Alors $B_3(x,y) = 3x_1y_1 - 5x_2y_1 - 5x_1y_2$ est bilinéaire symétrique sur $E\times E$ et la forme quadratique associée est $Q_{B_3}(x) = Q_{B_3}(x_1,x_2) = 3x_1^2 - 10x_1x_2$. Toutefois B_3 n'est pas positive car $Q_{B_3}(1,4) = -37 < 0$. La forme $B_4(x,y) = x_1y_1 - x_2y_1 - x_1y_2 + x_2y_2$ est une forme bilinéaire symétrique qui est positive car $Q_{B_4}(x_1,x_2) = x_1^2 - 2x_1x_2 + x_2^2 = (x_1 - x_2)^2 \geq 0$. Toutefois, B_4 n'est pas définie positive, car $Q_{B_4}(1,1) = 0$ alors que $(1,1) \neq (0,0)$. Enfin $B_5(x,y) = x_1y_1 - x_2y_1 - x_1y_2 + 5x_2y_2$ est une forme bilinéaire symétrique qui est définie positive car $Q_{B_4}(x_1,x_2) = (x_1-x_2)^2 + 4x_2^2$ est toujours ≥ 0 et ne peut être nul que si $x_1 - x_2 = 2x_2 = 0$, c'est à dire si $x = 0$.

Exemple 2.2. On prend $E = \mathbb{R}^q$ et la forme bilinéaire $B(x,y) = \langle x,y\rangle = x_1y_1 + \cdots + x_qy_q$. Alors la forme quadratique associée est $Q_B(x) = x_1^2 + \cdots + x_q^2$. Il est clair que B est symétrique définie positive. La forme B définit donc un produit scalaire sur $\mathbb{R}^q$, qui est ainsi muni de ce qu'on appelle le produit scalaire *canonique* [2] (ou encore de la *structure euclidienne canonique*) parce que c'est le plus simple et le plus naturel des produits scalaires possibles sur $\mathbb{R}^q$. Il faut noter qu'alors $\|x\| = (x_1^2 + \cdots + x_q^2)^{1/2}$.

Il y a un cas particulier intéressant dans cet exemple : si p et q sont des entiers positifs, soit E l'espace des matrices réelles à p lignes et q colonnes. Alors la forme bilinéaire sur $E \times E$ définie par $(A,B) \mapsto \mathrm{trace}\,(AB^T)$ est symétrique, d'après la Proposition 2.1 du Chapitre 1. Si $A = (a_{ij})$ on aura

$$\mathrm{trace}\,(AA^T) = \sum_{i=1}^{p}\sum_{j=1}^{q} a_{ij}^2 \geq 0.$$

On voit que cette forme définit un produit scalaire sur E. Il est clair que E s'identifie à $\mathbb{R}^{pq}$ euclidien canonique.

Exemple 2.3. Soit E l'espace des polynômes trigonométriques réels, c'est à dire l'espace de dimension infinie des fonctions f définies sur $\mathbb{R}$ telles qu'il existe un entier $n \geq 0$ et des nombres réels $a_0, \ldots, a_n, b_1, \ldots, b_n$ tels que

$$f(x) = a_0 + \sum_{k=1}^{n}(a_k \cos kx + b_k \sin kx).$$

[1] D'autres disent produit *intérieur.*

[2] Dans la littérature nord américaine, qui a parfois du mal à concevoir un espace vectoriel sans recours à des coordonnées, *Euclidean space* signifie $\mathbb{R}^q$ avec son produit scalaire canonique. Notre espace euclidien s'appelle alors *finite dimensional inner product space.*

Alors c'est un espace préhilbertien réel pour le produit scalaire

$$(f,g) \mapsto \frac{1}{2\pi}\int_0^{2\pi} f(x)g(x)dx.$$

La bilinéarité, la symétrie et la positivité sont évidentes. La définie positivité découle du fait que si $\int_0^{2\pi} f^2(x)dx = 0$ alors $f(x) = 0$ pour tout $x \in [0, 2\pi]$ car f est continue (voir cours de première année) et est donc nulle sur tout $\mathbb{R}$ car de période 2π : nous avons bien un produit scalaire.

Terminons cette section en expliquant que la connaissance de la forme quadratique associée Q_B à la forme bilinéaire symétrique B sur un espace E sur le corps K donne la connaissance de B lui même, au moyen des *égalités de polarisation* ci dessous.

Proposition 2.1. Soit K un corps tel que $x + x = 0$ entraîne $x = 0$. Soit E un espace vectoriel sur K et soit B et B' deux formes bilinéaires symétriques sur E telles que les formes quadratiques associées soit les mêmes. Alors $B = B'$. De plus pour tous $x, y \in E$ on a

$$\begin{aligned} B(x,y) &= \frac{1}{2}(Q_B(x+y) - Q_B(x) - Q_B(y)) \qquad (2.2)\\ &= \frac{1}{4}(Q_B(x+y) - Q_B(x-y))\\ &= \frac{1}{2}(Q_B(x) + Q_B(y) - Q_B(x-y)). \end{aligned}$$

On a aussi l'*égalité du parallélogramme* :

$$Q_B(x+y) + Q(x-y) = 2Q(x) + 2Q(y).$$

Démonstration : Contentons nous de montrer (2.2), les autres sont semblables. Par définition, le second membre de (2.2) est, en utilisant lentement la bilinéarité :

$$\begin{aligned} &\frac{1}{2}(B(x+y,x+y) - B(x,x) - B(y,y))\\ &= \frac{1}{2}(B(x,x+y) + B(y,x+y) - B(x,x) - B(y,y))\\ &= \frac{1}{2}(B(x,x) + B(x,y) + B(y,x) + B(y,y) - B(x,x) - B(y,y)) = B(x,y). \end{aligned}$$

Donc Q_B détermine B.

Remarques : Cette étrange condition sur le corps K est faite pour avoir le droit de diviser par 2 dans le corps : on ne peut le faire par exemple dans le corps à deux éléments. On dit aussi que K n'est pas de caractéristique 2. Enfin on peut définir une *forme quadratique* sur l'espace vectoriel E sur un corps K, (non de caractéristique 2) comme une fonction Q sur E telle que $B_Q(x,y) = \frac{1}{2}(Q(x+y) - Q(x) - Q(y))$ soit une forme bilinéaire. B_Q est alors appelée la *forme polaire* ou polarisée de la forme quadratique Q. Nous étudierons les formes quadratiques au chapitre 4.

Finalement, le lecteur avec des objectifs à court terme peut se contenter du résumé suivant :

Résumé. Un espace euclidien E est un espace vectoriel de dimension finie sur le corps des réels muni d'une fonction réelle sur $E \times E$ notée $(x, y) \mapsto \langle x, y\rangle$, appelée produit scalaire et satisfaisant aux axiomes suivants

- Pour tous x, y dans E on a $\langle x, y\rangle = \langle y, z\rangle$ (Symétrie).
- Pour tous λ et μ réels et pour tous x, y, z dans E on a $\langle \lambda x+\mu z, y\rangle = \lambda\langle x, y\rangle+\mu\langle z, y\rangle$ (Bilinéarité).
- Pour tout $x \neq 0$ on a $\langle x, x\rangle > 0$ (Définie-positivité).

Dans ces conditions, le nombre positif $\|x\|$ défini par $\|x\|^2 = \langle x, x\rangle$ est appelé norme de x et on a les identités de polarisation et du parallélogramme

$$\langle x, y\rangle = \frac{1}{2}(\|x + y\|^2 - \|x\|^2 - \|y\|^2) = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2),$$

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Il est *conseillé* quand on manipule les normes dans un espace euclidien de *les élever au carré*, les propriétés algébriques du carré de la norme euclidienne étant nombreuses.

Exercice 2.1. Soit E un espace euclidien et $a \in E$. Montrer que la fonction sur E définie par $x \mapsto \|x\|^2 - 2\langle a, x\rangle$ est minimum en a. Application : si $a_1, \ldots, a_n$ sont dans E trouver le minimum de

$$x \mapsto \|x - a_1\|^2 + \cdots + \|x - a_n\|^2.$$

Exercice 2.2. Soit E un espace euclidien. Montrer que pour tous x et y dans E et pour tout $\lambda \in [0, 1]$ on a

$$\lambda\|x\|^2 + (1 - \lambda)\|y\|^2 - \|\lambda x + (1 - \lambda)y\|^2 = \lambda(1 - \lambda)\|x - y\|^2$$

et en déduire que la fonction $x \mapsto \|x\|^2$ est convexe.

III Inégalités de Schwarz et du triangle

Proposition 3.1. Soit E un espace vectoriel réel et B une forme linéaire sur E symétrique et positive et soit Q_B la forme quadratique associée. Alors pour tous x et y de E :

1. $B(x, y)^2 \leq Q_B(x)Q_B(y)$ (Inégalité de *Schwarz*).
2. $\sqrt{Q_B(x + y)} \leq \sqrt{Q_B(x)} + \sqrt{Q_B(y)}$ (Inégalité *du triangle*)
3. $F = \{x \in E\ ;\ Q_B(x) = 0\}$ est un sous espace vectoriel de E.

4. L'inégalité de Schwarz est une égalité si et seulement si il existe des réels $(\lambda, \mu) \neq (0,0)$ tels que $\lambda x - \mu y \in F$. Dans ces conditions, $B(x,y)$ est du signe [3] de $\lambda\mu$.
5. L'inégalité du triangle est une égalité si et seulement si si il existe des réels $(\lambda, \mu) \neq (0,0)$ tels que $\lambda x - \mu y \in F$ avec $\lambda\mu \geq 0$.

Démonstration : 1) Soit $\lambda \in \mathbb{R}$. Alors

$$Q_B(\lambda x + y) = \lambda^2 Q_B(x) + 2\lambda B(x,y) + Q_B(y) \tag{3.3}$$

en utilisant la bilinéarité de B. Comme B est positive alors le polynôme de degré ≤ 2 défini par (3.3) est positif pour tout λ réel.

- Si $Q_B(x) = 0$ le polynôme affine (3.3) en λ ne peut être positif sur tout $\mathbb{R}$ sans être constant, cad qu'on a $B(x,y) = 0$. Dans ce cas l'inégalité de Schwarz est vérifiée et est même une égalité, et les conditions du 4) sont remplies avec $(\lambda, \mu) = (1,0)$. La reciproque du 4) est triviale dans ce cas $Q_B(x) = 0$.
- Si $Q_B(x) > 0$ le trinôme du second degré (3.3) en λ ne peut être positif pour tout λ réel que si il n'a pas de racines réelles distinctes, c'est à dire si son discriminant simplifié $\Delta' = B(x,y)^2 - Q_B(x)Q_B(y)$ est ≤ 0, ce qui est une reformulation de l'inégalité de Schwarz. Si $Q_B(x) > 0$ toujours, l'inégalité est une égalité si et seulement si $\Delta' = 0$, cad si le trinôme a une racine double, disons λ_0. Dans ce cas $Q_B(\lambda_0 x + y) = 0$, cad que $\lambda_0 x + y \in F$. Les conditions du 4) sont remplies avec $(\lambda, \mu) = (\lambda_0, -1)$. L'égalité $-2\lambda_0 B(x,y) = \lambda_0^2 Q_B(x) + Q_B(y) > 0$ montre que $B(x,y)$ est du signe de $-\lambda_0$. Quant à la réciproque du 4) si $(\lambda, \mu) \neq (0,0)$ est tel que $Q_B(\lambda x - \mu y) = 0$ on obtient

  $$\lambda^2 Q_B(x) - 2\lambda\mu B(x,y) + \mu^2 Q_B(y) = 0.$$

  Si $\mu = 0$ alors $\lambda \neq 0$ et donc $Q_B(x) = 0$: contradiction. Si $\mu \neq 0$ alors $\lambda_0 = -\lambda/\mu$ est racine d'un trinôme du second degré toujours positif ou nul, donc une racine double. Donc le trinôme a un discriminant simplifié Δ' qui est nul : l'inégalité de Schwarz est une égalité. L'égalité $2\lambda\mu B(x,y) = \lambda^2 Q_B(x) + \mu^2 Q_B(y)$ montre que $B(x,y)$ a le signe de $\lambda\mu$.

Pour l'inégalité du triangle on a par Schwarz :

$$\left[\sqrt{Q_B(x)} + \sqrt{Q_B(y)}\right]^2 - Q_B(x+y) = 2\left[\sqrt{Q_B(x)}\sqrt{Q_B(y)} - B(x,y)\right] \geq 0. \tag{3.4}$$

Il est clair qu'il y a égalité si et seulement si il y a égalité dans l'inégalité de Schwarz avec $B(x,y) \geq 0$. D'après 4), ceci est équivalent à l'affirmation du 5).

Reste à montrer le 3). Pour voir que F est un sous espace vectoriel, on constate aisément que si λ est réel et si x est dans F alors $Q_B(\lambda x) = \lambda^2 Q_B(x) = 0$ et donc λx est dans F. De même si x et y sont dans F, l'inégalité du triangle entraîne que $x + y$ est dans F, qui est donc un sous espace de E.

Proposition 3.2. Si E est un espace préhilbertien réel alors

[3] Signe $a = 1$ si $a > 0$, Signe $a = -1$ si $a < 0$, Signe $0 = 0$.

1. Pour tous x et y de E on a
$$|\langle x, y\rangle| \leq \|x\|\|y\|.$$
De plus $\langle x, y\rangle = \epsilon\|x\|\|y\|$ avec $\epsilon = \pm 1$ si et seulement si il existe un couple (λ, μ) non nul de réels ≥ 0 tels que $\lambda x = \epsilon \mu y$. Enfin, si x et y sont non nuls le nombre de $[0, \pi]$ égal à $\theta = \arccos(\langle x, y\rangle/(\|x\|\|y\|)$ est bien défini (et est appelé *angle* de x et de y).
2. Pour tous $x_1, \ldots, x_n$ de E on a $\|x_1 + \cdots + x_n\| \leq \|x_1\| + \ldots + \|x_n\|$, avec égalité si et seulement si il existe un vecteur $u \neq 0$ et des réels $\lambda_j \geq 0$ tels que $x_j = \lambda_j u$ pour tout $j = 1, \ldots, n$.

Démonstration : La partie 1) est une application immédiate de la Proposition 3.1 parties 1) et 4) à la forme bilinéaire définie-positive $B(x, y) = \langle x, y\rangle$. L'existence de l'angle vient du fait que $\langle x, y\rangle/(\|x\|\|y\|$ est dans $[-1, 1]$ c'est à dire le domaine de définition d' $\arccos$. Pour la partie 2) l'inégalité se déduit de l'inégalité du triangle de la Proposition 3.1 par une récurrence facile qu'on ne va pas faire. En revanche l'étude du cas d'égalité se fait par une récurrence plus délicate que nous détaillons.

Par la Proposition 3.1 partie 5) le résultat est vrai pour $n = 2$: en effet puisque B est définie positive, ici $F = \{0\}$. Supposons le résultat du 2) vrai pour $n \geq 2$ et montrons le pour $n + 1$. Notons $y = x_n + x_{n+1}$. Alors

$$\|x_1\|+\ldots+\|x_{n-1}\|+\|x_n\|+\|x_{n+1}\| = \|x_1+\ldots+x_{n-1}+y\| \leq \|x_1\|+\ldots+\|x_{n-1}\|+\|y\|. \quad (3.5)$$

Donc $\|x_n\| + \|x_{n+1}\| \leq \|y\|$ et donc $\|x_n\| + \|x_{n+1}\| = \|x_n + x_{n+1}\|$. D'après le cas $n = 2$ il existe un vecteur $v \neq 0$ et $\lambda_n, \lambda_{n+1} \geq 0$ tels que $x_n = \lambda_n v$ et $x_{n+1} = \lambda_{n+1} v$. Donc l'inégalité dans (3.5) est devenue une égalité, à laquelle on applique l'hypothèse de récurrence. Il existe donc un vecteur $u \neq 0$ est des scalaires λ_j et λ'_n positifs ou nuls, et tels que $x_j = \lambda_j$ avec $j = 1, \ldots, n-1$, et $y = \lambda'_n u$. Si $\lambda'_n = 0$, alors $y = 0$ et donc $\lambda_n = \lambda_{n+1} = 0$. Si $\lambda'_n > 0$, on peut prendre évidemment $u = v$.

Corollaire 3.3. Inégalité de Cauchy-Schwarz. Soit $(a_1, \ldots, a_q)$ et $(b_1, \ldots, b_q)$ deux suites de nombres réels de longueur q. Alors
$$(\sum_{i=1}^{q} a_i b_i)^2 \leq (\sum_{i=1}^{q} a_i^2)(\sum_{i=1}^{q} b_i^2).$$

Démonstration : Il suffit d'appliquer la proposition 3.2 partie 1 à l'espace $\mathbb{R}^q$ muni de sa structure euclidienne canonique et aux vecteurs $x = (a_1, \ldots, a_q)$ et $y = (b_1, \ldots, b_q)$. Pour la discussion des cas d'égalité il faut se reporter à la proposition.

Exercice 3.1. Si a, b et c sont réels, montrer que $|6a + 3b + 2c| \leq 7\sqrt{a^2 + b^2 + c^2}$. Cas d'égalité ?

Exercice 3.2. Donner une autre démonstration de l'inégalité de Schwarz dans un espace euclidien basée sur l'inégalité

$$\left\| \frac{x}{\|x\|} - \frac{y}{\|y\|} \right\|^2 \geq 0.$$

Exercice 3.3. Soient $p_1, q_1, \dots, p_n, q_n$ des nombres > 0 tels que $\sum_{i=1}^n p_i = \sum_{i=1}^n q_i = 1$. Montrer

$$1 \leq \sum_{i=1}^n \frac{p_i^2}{q_i}.$$

Méthode : appliquer Cauchy- Schwarz à $a_i = p_i/\sqrt{q_i}$ et $b_i = \sqrt{q_i}$.

IV Pythagore, Schmidt, Bessel

Définition 4.1. Soit E un espace préhilbertien réel et x et y dans E. On dit que x et y sont *orthogonaux* si $\langle x, y\rangle = 0$. Une partie V de E est dite *orthogonale* si tout couple x, y de vecteurs distincts de V est orthogonal. Enfin U est dite *orthonormale* si *de plus* tous les vecteurs de V sont de norme 1.

Proposition 4.1. Soit V une famille orthogonale de E ne comprenant pas 0. Alors

1. La famille V est *libre*.
2. Si x est dans l'espace vectoriel engendré par la famille V, avec $x = \lambda_1 v_1 + \dots \lambda_k v_k$ et $\{v_1, \dots, v_k\} \subset V$, alors
$$\lambda_j = \frac{\langle x, v_j\rangle}{\|v_j\|^2}, \quad \|x\|^2 = \frac{(\langle x, v_1\rangle)^2}{\|v_1\|^2} + \dots + \frac{(\langle x, v_k\rangle)^2}{\|v_k\|^2} \tag{4.6}$$
3. En particulier, pour la famille orthogonale $V = \{v_1, \dots, v_k\}$ on a le *Théorème de Pythagore* :
$$\|v_1 + \dots + v_k\|^2 = \|v_1\|^2 + \dots + \|v_k\|^2.$$
4. Si E est de dimension finie, si $e = (e_1, \dots, e_q)$ est une base orthonormale, si $[x]^e = [x_1, \dots, x_q]^T$ et si $[y]^e = [y_1, \dots, y_q^T$ alors $x_j = \langle x, e_j\rangle$ et $y_j = \langle y, e_j\rangle$ et
$$\langle x, y\rangle = x_1 y_1 + x_2 y_2 + \dots + x_q y_q.$$

Démonstration : Montrons simultanément 1) et 2). Soit $\{v_1, \dots, v_k\} \subset V$, soit des nombres λ_j réels, $j = 1, \dots k$ et soit $x = \lambda_1 v_1 + \dots \lambda_k v_k$. Formons le produit scalaire $\langle x, v_j\rangle$. Comme les $\{v_1, \dots, v_k\}$ sont orthogonaux on a donc $\langle x, v_j\rangle = \lambda_j \|v_j\|^2$. Comme V ne comprend pas le vecteur nul, $\|v_j\|^2 \neq 0$. Dans ces conditions, si $x = 0$ alors $\lambda_j = 0$ pour tout j : ceci montre le 1). Pour terminer le 2) on forme

$$\|x\|^2 = \langle x, x\rangle = \langle x, \lambda_1 v_1 + \dots \lambda_k v_k\rangle$$

et on conclut facilement avec $\langle x, v_j\rangle = \lambda_j\|v_j\|^2$. Le 3) est le cas particulier $\lambda_j = 1$ pour tout j. Le 4) est évident.

Théorème 4.2. (Procédé d'orthonormalisation de Schmidt). Soit E un espace préhilbertien réel et $f = (f_k)_{k=1}^q$ une suite finie de q vecteurs de E *indépendants*. Soit F_k le sous espace vectoriel de dimension k engendré par $\{f_1, f_2, \dots, f_k\}$. Alors il existe une suite (appelée *base de Schmidt* associée à f) *orthonormale unique* $e_f = (e_1, \dots, e_q)$ telle que

- Pour tout k, la suite $(e_1, \dots, e_k)$ est une base de F_k.
- Le nombre $\langle e_k, f_k\rangle$ est strictement positif.

De plus, si E est euclidien de dimension q, de bases f et e où e est orthonormale, alors e est la base de Schmidt associée à f si et seulement si $[id_E]_f^e$ est triangulaire supérieure avec diagonale formée d'éléments > 0.

Démonstration : On procède par récurrence. L'indépendance de f entraîne que $f_1 \neq 0$. On prend $e_1 = f_1/\|f_1\|$ et il est clair que c'est le seul choix possible pour satisfaire 1) et 2) avec $k = 1$. Supposons $(e_1, \dots, e_{n-1})$ déterminés de sorte que 1) et 2) sont satisfaits avec $k = 1, \dots, n-1$ et supposons qu'on ait montré que cette propriété entraînait l'unicité de $(e_1, \dots, e_{n-1})$. Nous étendons cela à l'ordre n ainsi : soit

$$g_n = f_n - \sum_{k=1}^{n-1}\langle f_n, e_k\rangle e_k. \tag{4.7}$$

Existence. Il est impossible que $g_n = 0$ car alors $f_n \in F_{n-1}$, ce qui contredit l'indépendance de $(f_1, \dots, f_n)$. On remarque alors que g_n est orthogonal à F_{n-1}. En effet pour $j = 1, \dots, n-1$ on a

$$\langle g_n, e_j\rangle = \langle f_n, e_j\rangle - \sum_{k=1}^{n-1}\langle f_n, e_k\rangle\langle e_k, e_j\rangle = 0,$$

car $(e_1, \dots, e_{n-1})$ est orthonormale et donc $\langle e_k, e_j\rangle = 0$ si $j \neq k$ et $\langle e_j, e_j\rangle = 1$. Puisque g_n est orthogonal à chacun des vecteurs e_j et que ceux ci forment une base de F_{n-1}, g_n est donc orthogonal à chacun des vecteurs de F_{n-1}. De plus

$$\langle g_n, f_n\rangle = \langle g_n, g_n + \sum_{k=1}^{n-1}\langle f_n, e_k\rangle e_k\rangle = \|g_n\|^2 > 0.$$

En prenant $e_n = g_n/\|g_n\|$ on a montré l'existence.
Unicité. Pour voir l'unicité de e_n, on considère

$$V = \{x \in F_n;\ \langle x, y\rangle = 0\ \forall y \in F_{n-1}\}.$$

Il est clair que V est un sous espace de F_n et que $V \cap F_{n-1} = \{0\}$, car un vecteur de cette intersection est orthogonal à lui même. Donc V et F_{n-1} sont en somme directe et $F_{n-1} \oplus V \subset F_n$. Toutefois $\dim V \geq 1$ puisque V contient e_n. Donc comme $\dim F_k = k$ on a $\dim V = 1$. Or si un autre e'_n convenait, il serait dans V, de norme 1. Si $e'_n \neq e_n$ alors $e'_n = -e_n$, mais puisque $\langle f_n, e'_n\rangle = -\langle f_n, e_n\rangle < 0$ la condition 2 n'est pas remplie. Donc $e'n = e_n$ et l'unicité est montrée.

Changement de base. Si $e = e_f$ alors avec les notations ci dessus on a pour $n = 1, \dots q$:

$$f_n = \langle f_n, e_1 \rangle e_1 + \dots + \langle f_n, e_{n-1} \rangle e_{n-1} + \|g_n\| e_n,$$

$$[\mathrm{id}_E]_f^e = \begin{bmatrix} \|g_1\| & \langle f_2, e_1 \rangle & \dots & \langle f_q, e_1 \rangle \\ 0 & \|g_2\| & \dots & \langle f_q, e_2 \rangle \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \|g_q\| \end{bmatrix}.$$

La réciproque est facile.

Remarques : La démonstration précédente contient un algorithme pour calculer la base de Schmidt. En pratique on calcule à l'étape n le vecteur g_n apparaissant en (4.7) puis sa norme par la formule

$$\|g_n\|^2 = \langle g_n, g_n \rangle = \langle f_n, f_n \rangle - \sum_{k=1}^{n-1} (\langle f_n, e_k \rangle)^2. \tag{4.8}$$

Notez que si $e = e_f$, alors $[\mathrm{id}_E]_e^f$ est aussi triangulaire supérieure avec diagonale positive, comme inverse d'une matrice de ce type.

La première conséquence du théorème est que tout espace euclidien a une base orthonormale (qu'*on abréviera* désormais par *bon*) : si E est un tel espace, en tant qu'espace de dimension finie il a une base f, et la base de Schmidt e_f est orthonormale. Ensuite, ce théorème a plusieurs versions voisines. L'une d'entre elles part de f et fabrique g ortho*gonale* telle que pour tout k $(g_1, \dots, g_k)$ est une base de F_k et telle que $g_k - f_k$ soit dans F_{k-1}. L'autre suppose que E est de dimension infinie et part d'une suite $f = (f_k)_{k \geq 1}$ infinie de vecteurs indépendants de E. . On fabrique alors une suite $e = (f_k)_{k \geq 1}$ avec les propriétés 1) et 2) du théorème. Ceci est particulièrement utilisé quand E est l'espace des polynômes à coefficients réels, qu'on le munit d'un produit scalaire et qu'on applique le théorème à la suite $f = (f_k)_{k \geq 0}$ définie par $f_0(X) = 1,\ f_1(X) = X, \dots, f_k(X) = X^k$. Alors e_k est un polynôme de degré k. Ceci conduit à la riche théorie des *polynômes orthogonaux*. Les e_k dépendent du produit scalaire choisi. Un exemple fréquent est

$$\langle P, Q \rangle = \int_{-\infty}^{\infty} P(x)Q(x)f(x)dx$$

où f est une fonction positive, mais il y en a d'autres très différents. Les polynômes orthogonaux interviennent dans beaucoup de domaines, des mathématiques pures à la physique.

Voici maintenant l'important théorème de *projection orthogonale* sur un sous espace de dimension finie :

Théorème 4.3. Soit E un espace préhilbertien réel, F un sous espace de E de *dimension finie*, et soit $F^{\perp}$ l'ensemble de tous les vecteurs de E orthogonaux à tous les éléments de F. Alors

1. $F^\perp$ est un sous espace, F et $F^\perp$ sont en somme directe et $E = F \oplus F^\perp$.
2. Soit p_F la projection de E sur F parallèlement à $F^\perp$. Si $x \in E$ et $y_0 \in F$ alors il y a équivalence entre les trois propriétés suivantes : (a) $y_0 = p_F(x)$; (b) $\|x - y_0\| \leq \|x - y\|$ pour tout $y \in F$; (c) $\langle x - y_0, y\rangle = 0$ pour tout $y \in F$.

Remarques : Donc la projection orthogonale $p_F(x)$ est aussi caractérisée par le fait que c'est le vecteur de F qui *minimise* la distance entre x et F, et par le fait que $x - p_F(x)$ est orthogonal à F. L'endomorphisme de E défini par $x \mapsto 2p_F(x) - x$ est appelé la *symétrie orthogonale* par rapport à F.

Démonstration : 1) $0 \in F^\perp$, qui est donc non vide. Si x et x_1 sont dans $F^\perp$ et si λ et λ_1 sont dans $\mathbb{R}$, alors pour tout y de F on a

$$\langle \lambda x + \lambda_1 x_1, y\rangle = \lambda\langle x, y\rangle + \lambda_1\langle x_1, y\rangle = 0$$

et donc $F^\perp$ est bien un sous espace vectoriel. Ensuite, y dans $F \cap F^\perp$ est orthogonal à lui même et est donc nul. Enfin soit $f = (f_1, \ldots, f_n)$ une bon de F. Notons alors

$$p_F(x) = \sum_{j=1}^{n} \langle f_j, x\rangle f_j \in F.$$

Alors $x - p_F(x) \in F^\perp$, car $\langle f_j, x - p_F(x)\rangle = 0$ pour tout $j = 1, \ldots, n$. Donc pour tout x de E on a

$$x = (x - p_F(x)) + p_F(x) \in F^\perp + F$$

et donc $F^\perp$ est un supplémentaire de F. De plus $x \mapsto p_F(x)$ est la projection de x sur F parallèlement à $F^\perp$.
2) $(a) \Rightarrow (c)$: par définition de p_F. $(c) \Rightarrow (a)$: notons $z = p_F(x) - y_0 \in F$. Alors

$$0 = \langle x - y_0, z\rangle = \langle x - p_F(x) + z, z\rangle = \langle x - p_F(x), z\rangle + \|z\|^2.$$

Donc $z = 0$. $(a) \Rightarrow (b)$: notons $z = p_F(x) - y$. Alors, d'après Pythagore

$$\|x - y\|^2 = \|x - p_F(x) + z\|^2 = \|x - p_F(x)\|^2 + \|z\|^2 \geq \|x - p_F(x)\|^2.$$

$(b) \Rightarrow (a)$: c'est la partie la plus délicate. Posons $X = x - y_0$ et $Y = x - p_F(x)$. D'après $(a) \Rightarrow (b)$ on a $\|X\|^2 = \|Y\|^2$. Ensuite, si nous notons $y_1 = \frac{1}{2}(y_0 + p_F(x))$, par hypothèse on a

$$\|X\|^2 = \|Y\|^2 \leq \|x - y_1\|^2 = \|\frac{1}{2}(X + Y)\|^2.$$

Appliquons alors l'égalité du parallélogramme à (X, Y). On obtient

$$\|X - Y\|^2 = 2\|X\|^2 + 2\|Y\|^2 - \|X + Y\|^2 = 4(\|X\|^2 - \|\frac{1}{2}(X + Y)\|^2) \leq 0.$$

On en déduit que $X = Y$ et donc que $y_0 = P_F(x)$. Le théorème est montré.

Corollaire 4.4. (Inégalité de *Bessel*) Soit $e = (e_1, \dots, e_n)$ une suite orthonormale finie d'un espace préhilbertien réel E, soit F le sous espace engendré par e et soit x dans E. Alors

$$\|x\|^2 \geq \sum_{j=1}^{n} (\langle x, e_j \rangle)^2,$$

avec égalité si et seulement si x est dans F.

Démonstration : Ecrivons $x = y + z$, avec $y \in F$ et $z \in F^\perp$. Ensuite, $\langle x, e_j \rangle = \langle y, e_j \rangle + \langle z, e_j \rangle = \langle y, e_j \rangle$ et d'après la Proposition 4.1 :

$$\|y\|^2 = \sum_{j=1}^{n} (\langle x, e_j \rangle)^2 \leq \|y\|^2 + \|z\|^2 = \|x\|^2,$$

avec égalité si et seulement si $z = 0$, c'est à dire si $x \in F$.

Exercice 4.1. Soit $e_0, e_1, \dots, e_n$ la base canonique de l'espace $\mathbb{R}^{n+1}$ muni de sa structure euclidienne canonique. On pose

$$h = \frac{1}{n}(e_1 + \cdots + e_n), \ g = \frac{1}{n+1}(e_0 + e_1 + \cdots + e_n).$$

1. Calculer $\|e_0 - g\|^2$ et $\|e_0 - h\|^2$.
2. Montrer que $e_0 - h$ est orthogonal au sous espace vectoriel de $\mathbb{R}^{n+1}$ engendré par les n vecteurs $e_i - h$, $i = 1, 2, \dots, n$.
3. Les vecteurs $e_0, e_1, \dots, e_n$ forment les sommets d'un tétraèdre régulier situé dans l'hyperplan d'équation $x_0 + x_1 + \ldots + x_n = 1$. Quelle est la longueur de l'arête de ce tétraèdre ? Quelle est sa hauteur ? Quel est le rayon de la sphère de ce plan qui passe par tous les sommets ? Si $n = 3$ et si un tétraèdre régulier est d'arête a, quelle est sa hauteur en fonction de a? Quel est le rayon de sa sphère circonscrite en fonction de a?

Exercice 4.2. Dans $\mathbb{R}^3$ euclidien canonique, on considère le plan vectoriel

$$F = \{(x, y, z); \ 6x + 3y - 2z = 0\}.$$

Donner une base orthonormale de $F^\perp$. Calculer le projeté orthogonal du vecteur $\vec{v} = (7, 7, 7)$ sur $F^\perp$, puis en déduire le projeté orthogonal de $\vec{v}$ sur F.

Exercice 4.3. L'espace $\mathbb{R}^5$ est muni de sa structure euclidienne canonique. Soit E le sous espace de $\mathbb{R}^5$ des $(x_1, \cdots, x_5)$ tels que $x_1 + \cdots + x_5 = 0$. E est donc également un espace euclidien, dont les vecteurs

$$\vec{f_1} = (1, -1, 0, 0, 0), \ \vec{f_2} = (1, 0, -1, 0, 0), \ \vec{f_3} = (1, 0, 0, -1, 0), \ \vec{f_4} = (1, 0, 0, 0, -1)$$

forment une base f ordonnée. Calculer la base orthonormale de E associée à f par le procédé de Schmidt.

Exercice 4.4. Dans $\mathbb{R}^2$, soient n points $P_i = (x_i, y_i)$, où les x_i ne sont pas tous égaux. et soit D la droite d'équation $y = ax + b$, avec $(a, b) \neq (0, 0)$. On veut trouver D tel que $\sum_{i=1}^n (y_i - ax_i - b)^2$ soit minimum (droite des *moindres carrés*). Pour cela on considère l'espace euclidien $E = \mathbb{R}^n$ muni de son produit scalaire canonique et les trois vecteurs colonnes $\mathbf{1}$ X et Y de E définis par $\mathbf{1}^T = (1, \ldots, 1)$, $X^T = (x_1, \ldots, x_n)$, $Y^T = (y_1, \ldots, y_n)$, et le sous espace vectoriel F de E engendré par $\mathbf{1}$ et X. Quelle est la dimension de F? Calculer en fonction de X et Y les nombres a et b tels que $Z = aX + b\mathbf{1}$ soit la projection orthogonale de Y sur F. Dans le cas particulier où $n = 4$ et où $P_1 = (1, 3)$, $P_2 = (2, 4)$, $P_3 = (4, 3)$, $P_4 = (5, 6)$, donner l'équation de la droite des moindres carrés.

Exercice 4.5. (Base de Schmidt du tétraèdre régulier). Une suite de vecteurs de l'espace euclidien E $(v_k)_{k=0}^n$ est dite régulière si $v_0 = 0$ et si, pour $0 \leq i < j \leq n$ on a $||v_i - v_j||^2 = 1$. Montrer qu'alors $(v_1, \cdots, v_n)$ est une suite indépendante (Méthode : calculer $\langle v_i, v_j \rangle$). Montrer qu'il existe des suites régulières. (Par exemple, dans un espace euclidien F de dimension $n+1$ muni d'une base orthonormale $f = (\vec{f_0}, \cdots, \vec{f_n})$, considérer $v_k = \frac{1}{\sqrt{2}}(f_k - f_0)$). Soit $(v_k)_{k=0}^n$ une suite régulière de l'espace euclidien E et soit $e = (e_1, \cdots, e_n)$ la base de Schmidt associée à $(v_1, \cdots, v_n)$. Montrer qu'il existe des nombres réels $(a_1, \cdots, a_n)$ et $(b_1, \cdots, b_{n-1})$ tels que pour tout $k = 1, \cdots, n$ on ait

$$v_k = a_k e_k + \sum_{i<k} b_i e_i.$$

(Méthode : procéder par récurrence sur k. Si c'est vrai pour k, soit

$$v_{k+1} = c_1 e_1 + \cdots + c_k e_k + a_{k+1} e_{k+1}.$$

Pour montrer que $b_1 = c_1, \ldots, b_{k-1} = c_{k-1}$, montrer que $v_{k+1} - v_k$ est orthogonal à F_{k-1}, le sous espace engendré par $v_1, \cdots, v_{k-1}$). Calculer les $(a_1, \cdots, a_n)$ et $(b_1, \cdots, b_{n-1})$. Méthode : utiliser $\|v_k\|^2$, $\langle v_k, v_{k+1} \rangle$ et $\|v_{k+1}\|^2$ pour montrer que

$$a_{k+1}^2 = 1 - \frac{1}{4a_k^2}.$$

Exercice 4.6. Soit E un espace euclidien de dimension $d+1$. Soit H un sous espace vectoriel de dimension d et soit $(v_0, \ldots, v_d)$ des vecteurs de H tels que $\langle v_i, v_j \rangle = -1$ si $i \neq j$.

1. Montrer que $\sum_{j=0}^d \frac{1}{1+\|v_j\|^2} = 1$. Méthode : soit u un vecteur de norme 1 orthogonal à H. Montrer que $w_j = v_j + u$ définit une suite orthogonale de E. Si $a_j = 1/\|w_j\|$ montrer que $e_j = a_j w_j$ définit une suite orthonormale de E. Exprimer alors les composantes de u dans la bon $e = (e_0, e_1, \ldots, e_d)$ et conclure.
2. Montrer que $\sum_{j=0}^d \frac{v_j}{1+\|v_j\|^2} = 0$. Méthode : multiplier scalairement les deux membres par v_i avec $i = 0, \ldots, d$.
3. Si a et b sont des vecteurs de E l'endomorphisme de E défini par $x \mapsto a\langle b, x \rangle$ est noté $a \otimes b$. Montrer que $\sum_{j=0}^d \frac{v_j \otimes v_j}{1+\|v_j\|^2} = \mathrm{id}_H$. Méthode : montrer au préalable que $\mathrm{id}_H = \mathrm{id}_E - u \otimes u$.

Exercice 4.7. (Réciproque de l'exercice 4.6) Soit H un espace euclidien de dimension d, un ensemble $V = \{v_0, \dots, v_d\}$ de vecteurs de H et une suite $(p_0, \dots, p_d)$ de nombres strictement positifs. On suppose que

$$\sum_{j=0}^{d} p_j = 1, \quad \sum_{j=0}^{d} p_j v_j = 0, \quad \sum_{j=0}^{d} p_j v_j \otimes v_j = \mathrm{id}_H.$$

Montrer qu'alors $\langle v_i, v_j \rangle = -1$ si $i \neq j$ et que $p_j = 1/(1 + \|v_j\|^2)$. Méthode :

1. Montrer que la famille $V \setminus \{v_0\}$ est indépendante (sinon les $v_1, \dots, v_d$ sont orthogonaux à un même vecteur non nul h de H, et par la condition 2, v_0 aussi. Appliquer alors l'endomorphisme $\sum_{j=0}^{d} p_j v_j \otimes v_j$ au vecteur h pour obtenir une contradiction).
2. Par la condition 2, calculer les composantes de v_0 dans la base $(v_1, \dots, v_d)$ en fonction des $(p_0, \dots, p_d)$.
3. Appliquer l'endomorphisme $\sum_{j=0}^{d} p_j v_j \otimes v_j$ au vecteur v_0, en déduire une autre expression des composantes de v_0 dans la base $(v_1, \dots, v_d)$ et comparer les résultats.
4. Répéter la procédure en remplaçant v_0 par v_i et en déduire l'existence d'un nombre λ tel que $\langle v_i, v_j \rangle = -\lambda$ si $i \neq j$.
5. Utiliser la condition 2 pour montrer que $\lambda > 0$. Appliquer l'exercice 4.6 à la suite $v'_j = \lambda^{-1/2} v_j$ et la condition 1 pour voir que $\lambda = 1$.

Exercice 4.8. On garde les notations et les hypothèses de l'exercice 4.7. Montrer que le carré de la distance de 0 à l'hyperplan affine engendré par $v_1, \dots, v_d$ est égale à $p_0/(1 - p_0)$. (Méthode : tout point x de cet hyperplan affine étant de la forme $x = \sum_{i=1}^{d} \lambda_i v_i$ avec $\lambda_1 + \cdots + \lambda_d = 1$, montrer à l'aide de $\langle v_i, v_j \rangle = -1$ puis de Cauchy-Schwarz que

$$\|x\|^2 = -1 + \sum_{i=1}^{d} \frac{\lambda_i^2}{p_i} \geq -1 + \frac{1}{1 - p_0}$$

et que l'égalité est atteinte en x_0 défini par $\lambda_i = p_i/(1 - p_0)$ pour $i = 1, \dots, d$). Montrer aussi que x_0 et v_0 sont colinéaires. En déduire par symétrie que les hauteurs du tétraèdre de sommets $v_0, \dots, v_d$ passent toutes par 0. Autrement dit, fait exceptionnel, le tétraèdre a un orthocentre, qui est ici 0.

Exercice 4.9. Calculer $B^T B$ si

$$B = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{6} \\ 0 & 0 & \frac{\sqrt{6}}{3} \end{bmatrix}.$$

Soit $(f_i)_{i=1}^{3}$ une base ordonnée de l'espace euclidien de dimension 3 telle que $\|f_i\|^2 = 1$ pour tout i et telle que si $i \neq j$ on ait $\langle f_i, f_j \rangle \geq 0$. Soit $(e_i)_{i=1}^{3}$ la base de Schmidt engendrée, et soit $f_j = \sum_{i=1}^{3} b_{ij} e_i$. Est il vrai que b_{ij} soit toujours positif ?

Exercice 4.10. LE COCHONNET MONSTRUEUX. Soit $e = (e_1, \ldots, e_q)$ une bon de l'espace euclidien E de dimension q. Si $a \in E$ et $r > 0$ le cube $C(a, r)$ est l'ensemble des $x = a + \lambda_1 e_1 + \ldots + \lambda_q e_q$ avec les λ_j dans $[-1, 1]$, et la boule $B(a, r)$ est l'ensemble des x tels que $\|x - a\|^2 \leq r^2$.

1. Montrer que $B(a, r) \subset C(a, r)$.
2. Si $\varepsilon = \varepsilon_1, \ldots, \varepsilon_q) \in \{-1, 1\}^q$ notons $a(\varepsilon) = \varepsilon_1 e_1 + \cdots, \varepsilon_q e_q$. Dans le cube $C(0, 2)$ on place les 2^q boules $B(a(\varepsilon), 1)$. Par exemple si $q = 3$ le cube $C(0, 2)$est la boite qui contient huit boules de pétanque de diamètre 2. Il reste de la place au centre pour y loger le cochonnet, c'est à dire la plus grosse boule possible $B(0, r)$ telle que $B(0, r)$ soit tangente aux 2^q boules $B(a(\varepsilon), 1)$. Calculer r pour les différentes valeurs de q.
3. Si $q = 9$ montrer que $r = 2$ c'est dire que le cochonnet est tangent aux parois de la boîte $C(0, 2)$. Si $q \geq 10$ montrer que le cochonnet sort de la boîte $C(0, 2)$.

Exercice 4.11. Si p et q sont des entiers strictement positifs, on note par $J_{p,q}$ la matrice à p lignes et q colonnes dont tous les coefficients sont égaux à 1. Soit a et b des réels. Le but de l'exercice est de trouver le polynôme caractéristique de la matrice carrée d'ordre $p + q$

$$M = \begin{bmatrix} aJ_{p,p} & bJ_{p,q} \\ bJ_{q,p} & aJ_{q,q} \end{bmatrix}.$$

par choix d'une base adaptée. Soit E et F deux sous espaces vectoriels orthogonaux de l'espace euclidien G tels que $G = E \oplus F$. Soit $e = (e_1, \ldots, e_p)$ et $f = (f_1, \ldots, f_q)$ des bon de E et F et soit $g = e \cup f$. On note $s = e_1 + \cdots + e_p$ et $t = f_1 + \cdots + f_q$ et on définit l'endomorphisme m de G par $m(e_i) = as + bt$ pour $i = 1, \ldots, p$ et $m(f_j) = bs + at$ pour $j = 1, \ldots, q$.

1. Montrer que $M = [m]_g^g$.
2. Si $x \in E$ et $y \in F$ montrer que
$$m(x + y) = \langle x, s\rangle(as + bt) + \langle y, t\rangle(bs + at).$$
3. Soit $e'_1 = \frac{1}{\sqrt{p}} s$ et $f'_1 = \frac{1}{\sqrt{q}} t$ et soit $e' = (e'_1, \ldots, e'_p)$ et $f' = (f'_1, \ldots, f'_q)$ des bon de E et F. Calculer les $m(e'_i)$ et $m(f'_j)$ à l'aide du 2.
4. Soit la bon $g' = (e'_1, f'_1, e'_2, \ldots, e'_p, f'_2, \ldots, f'_q)$ de G. Si $S = \begin{bmatrix} ap & b\sqrt{pq} \\ b\sqrt{pq} & aq \end{bmatrix}$ montrer à l'aide du 3 que par blocs 2×2 et $(p + q - 2) \times (p + q - 2)$ on a $[m]_{g'}^{g'} = \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}$. Calculer le polynôme caractéristique de S. En déduire le polynôme caractéristique de m et de M.

Exercice 4.11. Soit U et V des sous espaces vectoriels de l'espace euclidien E, pas nécessairement orthogonaux mais en somme directe et tels que $E = U + V$. Soit p la projection de E sur U parallèlement à V (c'est à dire que si $x = u + v$ avec $u \in U$ et $v \in V$ alors $p(x) = u$). Montrer que U et V sont orthogonaux si et seulement si pour tout $x \in E$ on a $\|p(x)\|^2 \leq \|x\|^2$. Méthode pour $\Leftarrow$: si il existait $(u, v) \in U \times V$ tel que $0 \neq \langle u, v\rangle$ considérons pour tout $t \in \mathbb{R}$ le vecteur de E $x(t) = u + tv$. Explicitez le polynôme du second degré P défini par

$$P(t) = \|x(t)\|^2 - \|p(x(t))\|^2 = \|u + tv\|^2 - \|u\|^2$$

et utilisez le fait que son discriminant doit être ≤ 0 puisque $P(t) \geq 0$ pour tout $t \in \mathbb{R}$.

V Dual et adjoints

Nous ne considérons plus les espaces préhilbertiens réels en général, mais seulement les espaces euclidiens. Si E est un espace de dimension finie q sur un corps K, l'espace $E^* = L(E, K)$ des applications linéaires de E dans le corps de base est appelé le *dual* de E et ses éléments sont appelés des *formes linéaires.* Puisque $\dim L(E, F) = \dim E \times \dim F$, on a donc $\dim E^* = \dim E$. Toutefois les deux espaces E et E^* sont isomorphes seulement au sens où deux espaces de même dimension sur K le sont, sans que parmi les nombreux isomorphismes possibles -paramétrés par les matrices carrées d'ordre q inversibles- l'un d'entre se distingue et se montre plus utile[4]. Quand $K = \mathbb{R}$ et que E est euclidien il en va différemment et il y a un isomorphisme *canonique*, c'est à dire plus naturel que les autres, entre un espace euclidien et son dual :

Proposition 5.1. Soit E un espace euclidien et soit E^* son dual. Si $y \in E$, on note par F_y l'élément de E^* défini par $F_y(x) = \langle x, y\rangle$. Alors $\varphi : \ y \mapsto F_y = \varphi(y)$ est un isomorphisme entre E et E^*.

Démonstration : Soient y et y_1 dans E, λ et λ_1 dans $\mathbb{R}$. Le fait que pour tout x de E on ait

$$\langle x, \lambda y + \lambda_1 y_1\rangle = \lambda\langle x, y\rangle + \lambda_1\langle x, y_1\rangle$$

se traduit par $F_{\lambda y + \lambda_1 y_1} = \lambda F_y + \lambda_1 F_{y_1}$, et donc φ est bien une application linéaire de E dans E^*. Soit ensuite une bon $e = (e_1, \ldots, e_q)$ de E. Si $y \in E$ est tel que F_y est la forme linéaire nulle, cela entraîne $\langle x, y\rangle = 0$ pour tout $x \in E$, et en particulier $\langle e_j, y\rangle = 0$ pour tout $j = 1, \ldots, q$, et donc $y = \sum_{j=1}^q \langle y, e_j\rangle e_j = 0$ (Proposition 4.1). Donc le noyau de φ est $\{0\}$ et φ est injective. Comme E et E^* ont même dimension, on en déduit que φ est surjective et que c'est un isomorphisme.

Exemples. Si une bon e de E est donnée, et si $f \in E^*$ est donné pour $[x]^e = [x_1, \ldots, x_q]^T$ par

$$f(x) = a_1 x_1 + \cdots + a_q x_q,$$

alors le $y \in E$ tel que $f = F_y$, c'est à dire $y = \varphi^{-1}(f)$, est donné par $[y]^e = [a_1, \ldots, a_q]^T$. Ce peut être plus pénible si une bon n'a pas été explicitée. Ainsi considérons le cas où E est l'espace des polynômes réels de degré ≤ 2 muni du produit scalaire

$$\langle P, Q\rangle = \int_0^1 P(x)Q(x)dx,$$

[4] L'étude du dual sur un espace quelconque sera reprise au chapitre 4.

et considérons la forme linéaire f sur E définie par $f(P) = P(2)$. Pour trouver le polynôme $Q = \varphi^{-1}(f)$ tel que pour tout p de E on ait $P(2) = \int_0^1 P(x)Q(x)dx$, on peut utiliser la base non orthogonale définie par $e_0(x) = 1,\ e_1(x) = x,\ e_2(x) = x^2$, noter $Q(x) = q_0+q_1x+q_2x^2$, et faire successivement $P = e_0, e_1, e_2$ pour obtenir un système linéaire en (q_0, q_1, q_2) : la proposition précédente garantit que c'est un système de Cramer.

$$1 = q_0 + \frac{1}{2}q_1 + \frac{1}{3}q_2,\ 2 = \frac{1}{2}q_0 + \frac{1}{3}q_1 + \frac{1}{4}q_2,\ \ 4 = \frac{1}{3}q_0 + \frac{1}{4}q_1 + \frac{1}{5}q_2.$$

Son déterminant est $1/2160$. La solution est $q_0 = 57,\ q_1 = -372,\ q_2 = 390$.

Nous introduisons maintenant la notion fondamentale d'adjoint d'une application linéaire a entre deux espaces euclidiens[5] E et F.

Proposition 5.2. Soit E et F deux espaces euclidiens. Alors pour tout $a \in L(E, F)$ il existe un unique $a^* \in L(F, E)$ appelé *adjoint* de a tel que pour tout $x \in E$ et tout $y \in F$ on ait

$$\langle a(x), y\rangle = \langle x, a^*(y)\rangle. \tag{5.9}$$

Dans ces conditions, $a \mapsto a^*$ est un isomorphisme entre les espaces vectoriels $L(E, F)$ et $L(F, E)$ et $a^{**} = a$. De plus, si e et f sont des *bon* de E et F, alors $[a^*]^e_f = ([a]^f_e)^T$. Ensuite, si $\dim E = \dim F$ et si a^{-1} existe, alors $(a^{-1})^* = (a^*)^{-1}$. Si G est euclidien, si $a \in L(E, F)$, si $b \in L(F, G)$ alors $(b \circ a)^* = a^* \circ b^*$. Enfin si $a \in L(E)$ alors $\det a^* = \det a$ et $(\exp a)^* = \exp(a^*)$.

Remarques : L'énoncé ci dessus est compliqué par le fait qu'on ne suppose pas nécessairement $E = F$. Par exemple, dans (2.3), il faut réaliser que le produit scalaire de gauche est celui de F et celui de droite est celui de E. En principe, il n'y a pas d'ambiguité, mais certains auteurs préfèrent écrire $\langle a(x), y\rangle_F = \langle x, a^*(y)\rangle_E$. Le cas le plus important est celui où $E = F$ et où on parle d'adjoint d'un endomorphisme d'espace euclidien. Notez que $[a^*]^e_e = ([a]^e_e)^T$ n'est pas vrai en général si e n'est pas une bon. Notez que l'adjoint d'un produit d'endomorphismes n'est pas le produit des adjoints, mais le produit renversé.

Démonstration : *Existence de l'adjoint.* Soit $a \in L(E, F)$. Fixons des bon e et f aux espaces euclidiens E et F et soit

$$A = [a]^f_e = (a_{ij})_{1\leq i\leq p, 1\leq j\leq q}.$$

Définissons $a^* \in L(F, E)$ par sa matrice représentative $[a^*]^e_f = A^T$. Alors si $x \in E$ et $y \in F$ ont pour matrices représentatives $[x]^e = [x_1, \dots, x_q]$ et $[y]^f = [y_1, \dots, y_p]$ on a puisque les bases sont orthonormales

$$\langle a(x), y\rangle = \sum_{i=1}^p y_i(\sum_{j=1}^q a_{ij}x_j) = \sum_{j=1}^q x_j(\sum_{i=1}^p a_{ij}y_i) = \langle x, a^*(y)\rangle.$$

[5]Une notion d'adjoint pour des espaces quelconques qui généralise le cas euclidien est faite au chapitre 4. Elle est plus abstraite.

Unicité de l'adjoint. Si a' satisfait aussi $\langle a(x), y\rangle = \langle x, a'(y)\rangle$, alors
$$0 = \langle x, a^*(y)\rangle - \langle x, a'(y)\rangle = \langle x, (a^* - a')(y)\rangle.$$
En prenant $x = (a^* - a')(y)$, cela entraîne $\|(a^* - a')(y)\|^2 = 0$ pour tout $y \in F$ et donc $a^* - a' = 0$.
Autres propriétés. Elles découlent de l'isomorphisme entre $L(E, F)$ et l'espace $\mathcal{M}_{pq}(\mathbb{R})$ des matrices réelles à p lignes et q colonnes.

Exercice 5.1. L'espace vectoriel réel E des polynômes réels de degré $\leq n$ est muni de la structure euclidienne $\langle P, Q\rangle = \int_0^1 P(t)Q(t)dt$. Soit f la forme linéaire sur E définie par $f(P) = P(1/3)$. (a) Dans le cas où $n = 1$, trouver Q dans E tel que pour tout P de E on ait $f(P) = \langle P, Q\rangle$. (b) Dans le cas d'un n quelconque, montrer l'existence et l'unicité d'un Q dans E tel que pour tout P de E on ait $f(P) = \langle P, Q\rangle$.

Exercice 5.2. Soit $a = (a_0, \ldots, a_n) \in \mathbb{R}^{n+1}$ tel que $a_0 \neq 0$. Soit E le sous espace de l'espace $\mathbb{R}^{n+1}$ euclidien canonique formé par les suites $x = (x_0, x_1, \ldots, x_n)$ de réels tels que $a_0x_0 + a_1x_1 + \cdots + a_nx_n = 0$. Quelle est la dimension de E? On considère la forme linéaire g sur E définie par $g(x) = x_0$. Trouver l'unique $y = (y_0, y_1, \ldots, y_n)$ dans E tel que pour tout $x \in E$ on ait $g(x) = \langle x, y\rangle$ (Voir Proposition 5.1. Méthode : Trouver le réel v_0 tel que $v = (v_0, a_1, \ldots, a_n)$ soit dans E et chercher ensuite y sous la forme $y = \lambda v$). Application : montrer que
$$\max_{x \in E \setminus \{0\}} \frac{x_0^2}{\|x\|^2} = 1 - \frac{a_0^2}{\|a\|^2}.$$
(Méthode : appliquer l'inégalité de Schwarz à $x_0^2 = \langle y, x\rangle^2$).

Exercice 5.3. Soit y dans E euclidien et $F_y(x) = \langle y, x\rangle$. Donc $(F_y)^*$ est une application linéaire de $\mathbb{R}$ dans E qu'on demande de préciser en termes de y.

Exercice 5.4. Soit a un endomorphisme de l'espace euclidien E et $b = aa^* - a^*a$. Montrer que $b = 0$ (c'est à dire que a et a^* commutent) si et seulement si pour tout $x \in E$ on a $\|a(x)\|^2 = \|a^*(x)\|^2$. Méthode pour $\Leftarrow$: en utilisant une identité de polarisation (Proposition 2.1) montrer que pour tous x et $y \in E$ on a
$$\langle a(x), a(y)\rangle = \langle a^*(x), a^*(y)\rangle, \quad \langle x, b(y)\rangle = 0$$
puis choisir $x = b(y)$. Autre méthode quand on a lu le Théorème spectral 9.1 plus bas : sans polarisation, montrer que pour tout $x \in E$ on a $\langle x, b(x)\rangle = 0$ et utiliser le fait que $b = b^*$ pour voir que b est diagonalisable, puis en déduire $b = 0$ (Comparer cet exercice avec le Th. 3.1 du chapitre 3).

Exercice 5.5. Soit U et V des sous espaces vectoriels de l'espace euclidien E, pas nécessairement orthogonaux mais en somme directe et tels que $E = U + V$. Soit p la projection de E sur U parallèlement à V. On veut décrire p^*. Montrer que $U^\perp \cap V^\perp = \{0\}$ et, par un argument de dimension, que $U^\perp + V^\perp = E$. Montrer que pour tout $u \in U$, $v \in V$ et $x \in E$ on a $\langle u + v, p^*(x)\rangle = \langle u, x\rangle$ et donc
$$\langle u, x - p^*(x)\rangle = \langle v, p^*(x)\rangle.$$

En déduire que les deux membres sont nuls. De l'égalité $x = p^*(x) + (x - p^*(x)$ déduire que p^* est la projection sur $V^\perp$ parallèlement à $U^\perp$. Montrer que $p = p^*$ si et seulement si U et V sont orthogonaux.

VI Le groupe orthogonal, les matrices orthogonales

Théorème 6.1. Soit E un espace euclidien et soit a un endomorphisme de E. Alors les propriétés suivantes sont équivalentes

1. Pour tous x et y de E on a $\langle a(x), a(y)\rangle = \langle x, y\rangle$ (*Conservation du produit scalaire*).
2. Pour tout $x \in E$ on a $\|a(x)\|^2 = \|x\|^2$ (*Conservation de la norme*).
3. a^{-1} existe et $a^{-1} = a^*$ (*L'adjoint égale l'inverse*).
4. Il existe une bon $e = (e_1, \dots, e_q)$ de E telle que $(a(e_1), \dots, a(e_q))$ soit aussi une bon.
5. *Pour toute* bon $e = (e_1, \dots, e_q)$ de E alors $(a(e_1), \dots, a(e_q))$ est aussi une bon (*Conservation des bon).*

Remarques : Un endomorphisme a satisfaisant une des propriétés équivalentes ci dessus est dit *orthogonal*. On l'appelle aussi parfois une *isométrie vectorielle*. On prend en général le 1), la préservation du produit scalaire, comme définition. Les autres sont alors appelées propriétés caractéristiques. Certains auteurs prennent toutefois $a^{-1} = a^*$ comme définition (rappel : $a^{-1} = a^* \Leftrightarrow aa^* = \mathrm{id}_E \Leftrightarrow a^*a = \mathrm{id}_E$). L' ensemble des endomorphismes orthogonaux de l'espace euclidien E est noté $\mathbb{O}(E)$. Un exemple important d'endomorphisme orthogonal est la *symétrie orthogonale* $x \mapsto 2p_F(x) - x$ par rapport à un sous espace F de E (rappelons que $p_F(x)$ est la projection orthogonale de x sur F et que donc $x - p_F(x)$ est orthogonal à $p_F(x)$). En effet, cet endomorphisme conserve la norme puisque d'après Pythagore on a

$$\|2p_F(x) - x\|^2 = \|p_F(x)\|^2 + \|p_F(x) - x\|^2 = \|p_F(x)\|^2 + \|x - p_F(x)\|^2 = \|x\|^2.$$

Démonstration : 1) $\Rightarrow$ 2) est trivial en faisant $x = y$ dans 1). Ensuite, 2) $\Rightarrow$ 1) est clair par polarisation. 3) $\Rightarrow$ 1) vient de

$$\langle a(x), a(y)\rangle = \langle x, a^*a(y)\rangle = \langle x, a^{-1}a(y)\rangle = \langle x, y\rangle.$$

1) $\Rightarrow$ 3) : d'après le 1) on a que pour tous x et y de E : $\langle x, (a^*a - \mathrm{id}_E)(y)\rangle = 0$. Appliquons cela à $x = (a^*a - \mathrm{id}_E)(y)$ pour obtenir $\|(a^*a - \mathrm{id}_E)(y)\|^2 = 0$ pour tout y et donc $a^*a - \mathrm{id}_E = 0$. Comme E est de dimension finie, cela entraîne aussi $aa^* = \mathrm{id}_E$ et 3) est vrai.

5) $\Rightarrow$ 4) est évident. 1) $\Rightarrow$ 5) : en effet $\langle a(e_j), a(e_j)\rangle = \langle e_j, e_j\rangle = 1$, et pour $i \neq j$ $\langle a(e_i), a(e_j)\rangle = \langle e_i, e_j\rangle = 0$. Donc $a(e) = (a(e_1), \dots, a(e_q))$ est aussi une bon. Montrons enfin 4) $\Rightarrow$ 1).

Avec $[x]^e = [x_1, \ldots, x_q]$ et $[y]^e = [y_1, \ldots, y_q]$ on a puisque les bases sont orthonormales

$$\langle a(x), a(y)\rangle = \langle a(\sum_{i=1}^q x_i e_i), a(\sum_{j=1}^q y_j e_j)\rangle = \sum_{i=1}^q \sum_{j=1}^q x_i y_j \langle a(e_i), a(e_j)\rangle.$$

Comme la base $a(e)$ est orthogonale alors $(\langle a(e_i), a(e_j)\rangle) = I_q$ et on en tire $\langle a(x), a(y)\rangle = \sum_{i=1}^q x_i y_i = \langle x, y\rangle$. Le théorème est démontré.

Corollaire 6.2. *Si E est euclidien, alors $\mathbb{O}(E)$ est un sous groupe (appelé groupe orthogonal de E) pour la composition du groupe $\mathbf{GL}(E)$ des automorphismes de E (appelé groupe linéaire de E). De plus, si $a \in \mathbb{O}(E)$ alors $\det a = \pm 1$ et les seules valeurs propres de a sont ± 1. Enfin, si $\mathbb{O}_+(E)$ est l'ensemble des $a \in \mathbb{O}(E)$ tels que $\det a = 1$ (appelés rotations) et si $\mathbb{O}_-(E)$ est l'ensemble des $a \in \mathbb{O}(E)$ tels que $\det a = -1$ alors $\mathbb{O}_+(E)$ est un sous groupe de $\mathbb{O}(E)$ (appelé groupe spécial orthogonal de E et aussi noté $\mathbb{SO}(E)$), et pour a fixé dans $\mathbb{O}_-(E)$ l'application de $\mathbb{O}_-(E)$ dans $\mathbb{O}_+(E)$ définie par $b \mapsto ab$ est bijective.*

Démonstration : a et b dans $\mathbb{O}(E)$ implique $(ab)^* = b^* a^* = b^{-1}a^{-1} = (ab)^{-1}$ et donc $ab \in \mathbb{O}(E)$. Ensuite, $\mathrm{id}_E \in \mathbb{O}(E)$ trivialement. Enfin $a \in \mathbb{O}(E)$ implique $(a^{-1})^* = (a^*)^* = a = (a^*)^{-1}$ et donc $a^{-1} \in \mathbb{O}(E)$. Donc $\mathbb{O}(E)$ est bien un sous groupe de $\mathbf{GL}(E)$. Ensuite, si $a \in \mathbb{O}(E)$, alors $aa^* = \mathrm{id}_E$ entraîne $\det a \det a^* = 1$ et donc $(\det a)^2 = 1$. Si $a \in \mathbb{O}(E)$ toujours, et si λ est une valeur propre de a, soit $v \neq 0$ un vecteur propre associé. Puisque $a(v) = \lambda v$ et que a conserve la norme on a $\lambda^2 \|v\|^2 = \|a(v)\|^2 = \|v\|^2$ et donc en simplifiant par le nombre *non nul* $\|v\|^2$ on a $\lambda^2 = 1$ et $\lambda = \pm 1$. Il est clair que $\mathbb{O}_+(E)$ est un sous groupe de $\mathbb{O}(E)$ (et que $\mathbb{O}_-(E)$ n'est pas un). Enfin, si $a \in \mathbb{O}_-(E)$ et $b \in \mathbb{O}_-(E)$ il est clair que $ab \in \mathbb{O}_+(E)$. Cette application $b \mapsto ab$ de $\mathbb{O}_-(E)$ dans $\mathbb{O}_+(E)$ est injective puisque $ab = ab'$ entraîne $b = b'$ par multiplication à gauche par a^{-1}. Elle est surjective, car si $c \in \mathbb{O}_+(E)$ alors $b = a^{-1}c$ est dans $\mathbb{O}_-(E)$ et donc $ab = c$.

Définition 6.1. Une matrice carrée réelle A est dite *orthogonale* si elle est inversible et si $A^{-1} = A^T$. L'ensemble des matrices orthogonales d'ordre q est noté $\mathbb{O}(q)$, l'ensemble des matrices orthogonales d'ordre q à déterminant positif est noté $\mathbb{O}_+(q)$ ou $\mathbb{SO}(q)$ et l'ensemble des matrices orthogonales d'ordre q à déterminant négatif est noté $\mathbb{O}_-(q)$.

Commentons cette définition. D'après le cours de première année, la définition $A^{-1} = A^T$ des matrices orthogonales est équivalente à *chacune* des propriétés suivantes

$$AA^T = I_q, \quad A^T A = I_q.$$

Si $A = (a_{ij})_{1 \leq i,j \leq q}$, alors $AA^T = I_q$ est équivalent à

$$\begin{aligned} \sum_{k=1}^q a_{ik} a_{jk} &= 0 \text{ si } i \neq j \\ &= 1 \text{ si } i = j. \end{aligned} \tag{6.10}$$

C'est dire que si on considère les *lignes* de la matrice A comme des vecteurs de $\mathbb{R}^q$ muni de sa structure euclidienne et de sa bon canonique, alors ces lignes forment une autre bon de $\mathbb{R}^q$. De la même manière, $A^T A^T = I_q$ est équivalent à

$$\begin{aligned}\sum_{k=1}^{q} a_{ki}a_{kj} &= 0 \text{ si } i \neq j \\ &= 1 \text{ si } i = j. \end{aligned} \tag{6.11}$$

C'est dire que si on considère les *colonnes* de la matrice A comme des vecteurs de $\mathbb{R}^q$, alors ces colonnes forment une autre bon de $\mathbb{R}^q$.

Théorème 6.3. Soit E un espace euclidien de dimension q et soit A une matrice carrée réelle d'ordre q. Alors les propriétés suivantes sont équivalentes :

1. $A \in \mathbb{O}(q)$.
2. Il existe une bon e de E et $a \in \mathbb{O}(E)$ tels que $[a]_e^e = A$.
3. Pour toute bon e de E il existe $a \in \mathbb{O}(E)$ tel que $[a]_e^e = A$.
4. Il existe deux bon e et f de E telles que $[\mathrm{id}_E]_f^e = A$.
5. Pour toute bon e de E il existe une bon f de E telle que $[\mathrm{id}_E]_f^e = A$.
6. Pour toute bon f de E il existe une bon e de E telle que $[\mathrm{id}_E]_f^e = A$.

Remarque : La philosophie de ce théorème est que les matrices orthogonales ont deux interprétations bien distinctes :

- Elles sont les matrices représentatives des endomorphismes orthogonaux dans une base orthonormale ; de cela il découle du Corollaire 6.2 que $\mathbb{O}(q)$ et $\mathbb{SO}(q)$ sont des groupes, et qui sont isomorphes à $\mathbb{O}(E)$ et $\mathbb{SO}(E)$ si $\dim E = q$.
- Elles sont les matrices de passage d'une bon vers une autre.

Démonstration : 3) $\Rightarrow$ 2) est évident. 2) $\Rightarrow$ 1), car $A^T = ([a]_e^e)^T = [a^*]_e^e$ et $A^{-1} = ([a]_e^e)^{-1} = [a^{-1}]_e^e = [a^*]_e^e$ et donc $A^{-1} = A^T$. Enfin 1) $\Rightarrow$ 3), car

$$[a^* - a^{-1}]_e^e = [a^*]_e^e - [a^{-1}]_e^e = A^T - A^{-1} = 0.$$

5) $\Rightarrow$ 4) est évident. 4) $\Rightarrow$ 1) : définissons $a \in L(E)$ par $a(e_j) = f_j$ pour tout $j = 1, \ldots, q$. Alors $[a]_e^e = [\mathrm{id}_E]_f^e = A$. D'après le Théorème 6.1 on a $a \in \mathbb{O}(E)$ et d'après la partie 1) $\Rightarrow$ 3) on a le résultat. 1) $\Rightarrow$ 5) : définissons $a \in L(E)$ par $[a]_e^e = A$. D'après 1) $\Leftrightarrow$ 3) on a $a \in \mathbb{O}(E)$. Donc $f = a(e)$ est une bon d'après le Théorème 6.1. Enfin $[a]_e^e = [\mathrm{id}_E]_f^e$. L'équivalence 1) $\Leftrightarrow$ 6) si montre comme 1) $\Leftrightarrow$ 5) et le théorème est montré.

On complète cette section par quelques résultats utiles.

Corollaire 6.4. Soit M une matrice carrée inversible d'ordre q. Alors il existe un couple unique (A, T) avec $A \in \mathbb{O}(q)$ et T matrice triangulaire supérieure à diagonale positive tel que $M = AT$. Mêmes affirmations avec $M = TA$.

Remarques : C'est une proposition fort utile en analyse numérique, qui relie Schmidt et $\mathbb{O}(q)$. En prenant des transposées dans le résultat précédent on obtient un résultat analogue avec les triangulaires inférieures. En écrivant dans le résultat précédent $T = DN$

avec D diagonale positive et N triangulaire supérieure avec des 1 sur la diagonale, on obtient encore des décompositions uniques : $M = ADN$, ou NDA, DNA, AND.

Démonstration : Soit E un espace euclidien de dimension q et e une bon de E. Puisque M est inversible, il existe une base unique f, non nécessairement orthonormale, telle que $M = [\mathrm{id}_E]^e_f$. Soit alors $e' = e_f$ la bon de Schmidt associée à f. Alors

$$M = [\mathrm{id}_E]^e_f = [\mathrm{id}_E]^e_{e'}[\mathrm{id}_E]^{e'}_f = AT$$

avec $A = [\mathrm{id}_E]^e_{e'} \in \mathbb{O}(q)$ par le Théorème 6.3 et $T = [\mathrm{id}_E]^{e'}_f$ triangulaire supérieure à diagonale positive par le Théorème 4.2. De même, pour avoir plutôt $M = TA$ on part d'un f tel que

$$M = [\mathrm{id}_E]^f_e = [\mathrm{id}_E]^f_{e'}[\mathrm{id}_E]^{e'}_e = TA.$$

Si $T = [\mathrm{id}_E]^f_{e'}$ alors $T^{-1} = [\mathrm{id}_E]^{e'}_f$ est triangulaire supérieure à diagonale positive par le Théorème 4.2 et il en est donc de même pour T. L'unicité dans les deux cas est facile à montrer, à partir de l'unicité du procédé de Schmidt.

Proposition 6.5. Soit $a \in \mathbb{O}(E)$ et F un sous espace de E stable par a, c'est à dire tel que $a(F) \subset F$. Alors $a(F) = F$ et $a(F^\perp) = F^\perp$.

Démonstration : La restriction a_F de a est un endomorphisme de F dont le noyau $\ker a_F = \ker a \cap F$ est réduit à $\{0\}$ puisque a est inversible. Donc a_F est surjectif. Ensuite, soit $y \in F^\perp$. On veut montrer que $a(y) \in F^\perp$. Pour cela on utilise la surjectivité de a_F : tout élément de F est de la forme $a(x)$ pour un $x \in F$. Donc

$$\langle a(x), a(y)\rangle = \langle x, y\rangle = 0$$

montre que $F^\perp$ est stable par a, et par la première partie $a(F^\perp) = F^\perp$.

Proposition 6.6. (*Déterminant d'un système de vecteurs dans un espace euclidien orienté).* Soit E un espace euclidien *orienté* de dimension q. Soit $v = (v_1, \ldots, v_q)$ une suite de q vecteurs et soit e une bon *directe*. Alors le nombre

$$\det([v_1]^e, \ldots, [v_q]^e)$$

ne dépend pas de la bon directe e choisie.

Remarques : Dans un espace de dimension finie quelconque, même sur le corps $K = \mathbb{R}$, la notion de déterminant de système de vecteurs n'a pas de sens intrinsèque car le nombre $\det([v_1]^e, \ldots, [v_q]^e)$ varie avec la base e. Dans le cas d'un espace euclidien, si on se limite aux bases orthonormales et directes, on a une invariance. Si on considère le parallélépipède engendré par v :

$$P(v) = \{x = \lambda_1 v_1 + \cdots + \lambda_q v_q;\ 0 \leq \lambda_j \leq 1,\ j = 1, \ldots, q\},$$

alors le nombre $|\det([v_1]^e, \ldots, [v_q]^e)|$ est appelé le *volume* de $P(v)$. Si $q = 2$ on voit facilement en partitionnant le parallélogramme $P(v)$ en un triangle plus un trapèze que ceci est l'aire d'un rectangle convenable. Si $q = 3$ le fait que $|\det([v_1]^e, [v_2]^e, [v_3]^e)|$ soit le volume au sens intuitif de la physique est en fait plus difficile à justifier rigoureusement et relève de la troisième année d'université. Le nombre relatif $\det([v_1]^e, \ldots, [v_q]^e)$ est appelé *volume algébrique* de $P(v)$, une notion que nous n'allons pas approfondir puisqu'elle nécessite de définir une orientation de $P(v)$.

Démonstration : Si f est une autre bon directe, soit $P = [\mathrm{id}_E]_f^e$ la matrice de changement de base. Alors $\det P > 0$ car e et f ont même orientation, $P \in \mathbb{O}(q)$ car e et f sont des bon et le Théorème 6.3 s'applique. Donc $\det P = 1$. Ensuite pour tout x de E on a $[x]^e = [\mathrm{id}_E]_f^e[x]^f = P[x]^f$. Par conséquent, en écrivant la matrice carrée des vecteurs colonnes $[v_j]^e$ on a

$$[[v_1]^e, \ldots, [v_q]^e] = [P[v_1]^f, \ldots, P[v_q]^f] = P[[v_1]^f, \ldots, [v_q]^f],$$
$$\det([v_1]^e, \ldots, [v_q]^e) = \det P \det([v_1]^f, \ldots, [v_q]^f) = \det([v_1]^f, \ldots, [v_q]^f).$$

Voici pour terminer une curiosité sans grand intérêt qu'adorent les jurys de concours : la caractérisation des endomorphismes orthogonaux *sans hypothèse de linéarité* :

Proposition 6.7. Soit E un espace euclidien et a une application de E dans E telle que $a(0) = 0$ et telle que pour tous x et y de E on ait

$$\|a(x) - a(y)\|^2 = \|x - y\|^2.$$

Alors $a \in \mathbb{O}(E)$.

Démonstration : Par polarisation on a

$$\langle a(x), a(y)\rangle = \frac{1}{2}(\|a(x)\|^2 + \|a(y)\|^2 - \|a(x) - a(y)\|^2).$$

De plus, comme $a(0) = 0$ on a $\|a(x)\|^2 = \|x\|^2$. Donc $\langle a(x), a(y)\rangle = \langle x, y\rangle$. Reste donc à montrer que a est linéaire. Soit λ et μ des réels. On veut donc montrer que

$$z = a(\lambda x + \mu y) - \lambda a(x) - \mu a(y)$$

est nul. On écrit pour cela pour tout $v \in E$:

$$\begin{aligned} \langle z, a(v)\rangle &= \langle a(\lambda x + \mu y), a(v)\rangle - \lambda\langle a(x), a(v)\rangle - \mu\langle a(y), a(v)\rangle \\ &= \langle \lambda x + \mu y, v\rangle - \lambda\langle x, v\rangle - \mu\langle y, v\rangle = 0 \end{aligned}$$

Appliquant alors successivement cette égalité à $v = \lambda x + \mu y$, à $v = x$ et à $v = y$, on en tire $\langle z, z\rangle = 0$ et donc $z = 0$, ce qui achève la démonstration.

Exercice 6.1. Trouver les nombres réels a, b, c, pour que la matrice suivante

$$U = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & a \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & b \\ \frac{1}{\sqrt{3}} & 0 & c \end{bmatrix}$$

soit dans $\mathbb{SO}(3)$, c'est à dire telle que U soit orthogonale de déterminant positif.

Exercice 6.2. (Suites stationnaires dans un espace euclidien) Soit E un espace euclidien de dimension q et soit $(b_n)_{n\in\mathbb{Z}}$ une suite de E indexée par l'ensemble $\mathbb{Z}$ des entiers relatifs non concentrée sur un sous espace affine de E, c'est à dire qu' il existe $n_0 < n_1 < n_2 < \ldots < n_q$ tels que $(b_{n_1} - b_{n_0}, \ldots, b_{n_q} - b_{n_0})$ est une base de E. On suppose qu'il existe $u \in \mathbb{O}(E)$ et $t \in E$ tels que si f est la transformation affine de E définie par $f(x) = u(x) + t$ alors pour tout $n \in \mathbb{Z}$ on a $b_n = f^n(b_0)$. Montrer que c'est équivalent au fait que pour tout $k \in \mathbb{Z}$ le nombre $\psi(k) = \|b_{n+k} - b_n\|^2$ ne dépend pas de $n \in \mathbb{Z}$. Montrer dans ces conditions qu'il existe des nombres $p_0, p_1, \ldots, p_r \geq 0$ et des nombres réels $\theta_1, \ldots, \theta_r$ tels que

$$\psi(k) = k^2 p_0 + \sum_{j=1}^{r} (1 - \cos k\theta_j) p_j.$$

Exercice 6.3. Soit F un sous espace vectoriel d'un espace euclidien E et soit s_F la symétrie orthogonale par rapport à F. Si e est une bon contenue dans $F \cup F^\perp$ calculer $[s_F]_e^e$ et montrer que $-s_F = s_{F^\perp}$.

Exercice 6.4. Soit a un endomorphisme de l'espace euclidien tel que $\|x\| \leq \|y\|$ entraine $\|a(x)\| \leq \|a(y)\|$. Montrer qu'il existe $u \in \mathbb{O}(E)$ et $\lambda \in [0, 1]$ tels que $a = \lambda u$. Méthode : montrer que $\|x\| = \|y\|$ entraine $\|a(x)\| = \|a(y)\|$ et que $\lambda = \|a(x)\|/\|x\|$ ne dépend pas de x. Si $\lambda > 0$ montrer que $x \mapsto a(x)/\lambda$ est orthogonal.

Exercice 6.5. Une similitude vectorielle de l'espace euclidien E est un endomorphisme f de E de la forme $x \mapsto \lambda u(x)$ avec $\lambda > 0$ et $u \in \mathbb{O}(E)$ (le nombre λ est appelé rapport de similitude). Soit alors $f \in GL(E)$. On suppose que pour tous x et y de E orthogonaux alors $f(x)$ et $f(y)$ sont orthogonaux. Montrer qu'alors f est une similitude vectorielle. Méthode : on prend une bon $e = (e_1, \ldots, e_q)$ de E quelconque on montre que pour $i \neq 1$ alors $f(e_1 + e_i)$ et $f(e_1 - e_i)$ sont orthogonaux. On en déduit que $\lambda = \|f(e_i)\|$ ne dépend pas de i. Pour terminer on démontre que la matrice $\frac{1}{\lambda}[f]_e^e$ est orthogonale en utilisant la définition d'une matrice orthogonale.

Exercice 6.6. Soit H le sous espace de l'espace $\mathbb{R}^3$ euclidien canonique défini par $H = \{x = (x_1, x_2, x_3) \ ; \ x_1 + x_2 + x_3 = 0\}$ et soit f l'endomorphisme de H défini par

$$f(x) = f(x_1, x_2, x_3) = \frac{1}{\sqrt{3}}(x_1 - x_2, x_2 - x_3, x_3 - x_1).$$

Montrer que $\|f(x)\|^2 = \|x\|^2$ (méthode : expliciter $(x_1 + x_2 + x_3)^2$).

Exercice 6.7. Soit $U = (u_{ij})_{1\leq i,j\leq q} \in \mathbb{O}(q)$. Montrer que

$$|\sum_{i,j} u_{ij}| \leq q.$$

Méthode : dans $E=\mathbb{R}^q$ euclidien canonique considérer la bon canonique $e=(\vec{e}_1,\dots,\vec{e}_q)$ et la suite de vecteurs $s=(\vec{u}_1,\dots,\vec{u}_q)$ avec $\vec{u}_i=(u_{i1},\dots,u_{iq})$. Montrer à l'aide de (6.11) que s est une bon de E. Former ensuite $\vec{f}=\vec{e}_1+\cdots+\vec{e}_q$ et $\vec{g}=\vec{u}_1+\cdots+\vec{u}_q$, montrer que $\sum_{i,j}u_{ij}=\langle\vec{f},\vec{g}\rangle$ et calculer les normes de $\vec{f}$ et $\vec{g}$. Conclure avec l'inégalité de Schwarz. Autre méthode, utilisant le Théorème spectral 9.1 ci dessous : Si J est la matrice carrée d'ordre q dont tous les coefficients sont 1 soit u et j les endomorphismes de E tels que $[u]_e^e=U$ et $[j]_e^e=J$. Observer que trace $(uj)=\sum_{i,j}u_{ij}$ et montrer qu'il existe une bon f de E telle que $[j]_f^f=\mathrm{diag}(q,0,\dots,0)$ (c'est à dire que la matrice J de rang 1 a q pour unique valeur propre non nulle). Notant $[u]_f^f=(v_{ij})_{1\leq i,j\leq q}\in\mathbb{O}(q)$ montrer que trace $(uj)=qv_{11}$ et conclure.

VII Le groupe orthogonal du plan ; dimensions supérieures.

Théorème 7.1. Soit E euclidien de dimension 2 et a dans $\mathbb{O}(E)$. Alors

- ou bien $a\in\mathbb{O}_-(E)$. Alors il existe une bon $e=(e_1,e_2)$ telle que $[a]_e^e=\begin{bmatrix}1&0\\0&-1\end{bmatrix}$. Dans ce cas, a est une symétrie orthogonale par rapport à la droite du plan $\mathbb{R}e_1$.
- ou bien $a\in\mathbb{O}_+(E)$. Si de plus E est orienté, alors il existe un nombre θ tel que *pour toute* bon directe e on ait
$$[a]_e^e=R(\theta)=\begin{bmatrix}\cos\theta&-\sin\theta\\ \sin\theta&\cos\theta\end{bmatrix},\tag{7.12}$$
et pour toute bon indirecte on ait $[a]_e^e=R(-\theta)$.

Remarques : Le nombre θ apparaissant quand $a\in\mathbb{SO}(E)$ est appelé *angle orienté* de la rotation a. Il n'est pas unique en ce sens qu'on peut le remplacer par $\theta+2k\pi$ pour $k\in\mathbb{Z}$. Il est remarquable qu'il ne dépende pas de la bon directe e. Il dépend de l'orientation : changer l'orientation du plan euclidien changerait θ en $-\theta$. L'angle orienté est relié à la notion d'angle introduite à la Proposition 3.2 ainsi : si $a\in\mathbb{SO}(E)$ et si $[a]_e^e=R(\theta)$ dans une bon directe alors pour $x\neq 0$ on a $\langle a(x),x\rangle=\|x\|^2\cos\theta$. Bien que θ ne soit pas nécessairement dans $[0,\pi]$, il existe certainement un $\theta_1\in[0,\pi]$ tel que $\cos\theta=\cos\theta_1$. Ce θ_1 est l'angle des deux vecteurs $a(x)$ et x. Il n'est pas orienté, il ne change pas si on échange $a(x)$ et x, il ne dépend pas de l'orientation du plan euclidien. Cette notion d'angle orienté est spéciale au plan euclidien et ne généralise pas aux dimensions supérieures.

Démonstration : Si $\det a=-1$ alors le polynôme caractéristique de a est $P_a(X)=X^2-\mathrm{trace}\,aX-1$. Il a donc deux racines réelles distinctes, qui sont donc -1 et 1 d'après le Corollaire 6.2. a est donc diagonalisable car $q=2$ et il y a 2 valeurs propres distinctes. Soit e_1 un vecteur propre associé à la valeur propre 1 et e_2 un vecteur propre associé à la valeur propre -1. On les prend de norme 1. Ils sont orthogonaux car
$$\langle e_1,e_2\rangle=\langle a(e_1),a(e_2)\rangle=\langle e_1,-e_2\rangle=-\langle e_1,e_2\rangle.$$

Le résultat s'ensuit.

Si maintenant $\det a = 1$, soit e une bon directe et $[a]_e^e = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$. Comme c'est une matrice orthogonale on a $\alpha^2 + \beta^2 = \gamma^2 + \delta^2 = 1$, et le cours de première année entraîne l'existence de nombres θ et θ_1 tels que

$$\alpha = \cos\theta,\ \beta = -\sin\theta,\ \gamma = \sin\theta_1,\ \delta = \cos\theta_1.$$

De plus

$$1 = \det a = \alpha\delta - \beta\gamma = \cos(\theta - \theta_1),$$

et donc $\theta = \theta_1 \mod 2\pi$: donc $[a]_e^e = R(\theta)$. Au passage, nous venons de montrer que si P est une matrice orthogonale de $\mathbb{SO}(2)$ alors il existe un nombre t tel que $P = R(t)$.

Observons ensuite que $R(\theta)R(\theta') = R(\theta + \theta')$ par les formules de trigonométrie. Si alors f est une autre bon directe, alors $P = [\mathrm{id}_E]_f^e$ est dans $\mathbb{SO}(2)$ et il existe un réel t tel que $P = R(t)$. Donc $P^{-1} = R(-t)$ et

$$[a]_f^f = [\mathrm{id}_E]_e^f[a]_e^e[\mathrm{id}_E]_f^e = R(-t)R(\theta)R(t) = R(-t + \theta + t) = R(\theta).$$

Finalement, si $f = (f_1, f_2)$ est une bon indirecte alors $e = (f_1, -f_2)$ est une bon directe et donc

$$[a]_f^f = [\mathrm{id}_E]_e^f[a]_e^e[\mathrm{id}_E]_f^e = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} R(\theta) \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = R(-\theta).$$

Théorème 7.2. Soit E un espace euclidien de dimension q et soit $a \in \mathbb{O}(E)$. Alors il existe une bon e, des entiers positifs ou nuls r, p et n tels que $p + n + 2r = q$ et des nombres réels $\theta_1, \dots, \theta_r$ dans tels que $[a]_e^e$ s'écrive par blocs :

$$[a]_e^e = \mathrm{diag}(R(\theta_1), \dots, R(\theta_r), I_p, -I_n).$$

De plus, si $a \in \mathbb{SO}(E)$ on peut prendre $n = 0$. Finalement, pour $a \in L(E)$ alors $a \in \mathbb{SO}(E)$ si et seulement si il existe $b \in L(E)$ tel que $b + b^* = 0$ et $a = e^b$.

Remarques : Il n'y pas tout à fait unicité de (p, n, r) avec l'énoncé précédent puisque $I_2 = R(0)$ et $-I_2 = R(\pi)$. On pourrait arriver à cette unicité en imposant que les θ_j soient dans $]0, \pi[$ mais nous ignorons ce raffinement, laissé en exercice. L'énoncé montre que $\mathbb{SO}(E)$ est *connexe par arcs,* un terme défini dans le cours d'analyse : si $a \in \mathbb{SO}(E)$ soit $b \in L(E)$ tel que $b^* = -b$ (on dit que b est *antisymétrique*) et tel que $a = e^b$. Soit $a_t = e^{bt}$. Alors $t \mapsto a_t$ est une application continue de $[0, 1]$ dans $\mathbb{SO}(E)$, ce qui conduit facilement à la connexité annoncée.

Démonstration : On procède par récurrence sur q. C'est évident pour $q = 1$ et c'est un conséquence du Théorème 7.1 si $q = 2$. Si c'est vrai pour tous les entiers inférieurs à q on utilise le Théorème 1.2 : puisque E est un espace réel, il y a un sous espace F de E de dimension 1 ou 2 qui est stable par a. On utilise alors la Proposition 6.5 qui dit que $F^\perp$ est stable par a. Comme $\dim F^\perp < q$ on peut trouver une bon e' de $F^\perp$ telle que la restriction $a_{F^\perp}$ de a à $F^\perp$ ait la forme voulue. De même il y a une bon f de F pour que a_F ait aussi la forme voulue, et la bon $e = e' \cup f$ convient : la récurrence est étendue. Si $a \in \mathbb{SO}(E)$ il est clair qu'alors $n = 2m$ doit être pair. En écrivant $-I_2 = R(\pi)$ on a la forme annoncée.

Enfin si $a \in \mathbb{SO}(E)$, pour voir qu'il existe $b \in L(E)$ tel que $b + b^* = 0$ et $a = e^b$ on utilise l'exemple 9.1 du chapitre 1 qui dit que si $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ alors $\exp(\theta A) = R(\theta)$. Donc pour $a \in \mathbb{SO}(E)$ on peut écrire dans une certaine bon e

$$[a]_e^e = \mathrm{diag}(R(\theta_1), \ldots, R(\theta_r), I_p) = \exp B$$

avec $B = \mathrm{diag}(\theta_1 A, \ldots, \theta_r A, 0_p)$. Définissons $b \in L(E)$ par $[b]_e^e = B$. On a bien $a = \exp b$. Puisque e est une bon, alors $[b^*]_e^e = B^T$. Comme $A^T = -A$ on a donc $B^T = -B$ et $b^* = -b$. Inversement, $a = \exp b$ satisfait $a^* = \exp b^*$ par définition de l'exponentielle. Si alors $b^* = -b$ alors $a^* = a^{-1}$ et donc $a \in \mathbb{O}(E)$. Mais $\det \exp b > 0$ (ou bien parce que $\det \exp b = \exp \mathrm{trace}\, b$, ou bien plus simplement parce que $\exp b = \exp(b/2)\exp(b/2)$ et donc $\det \exp b = (\det \exp(b/2))^2$). Donc $a \in \mathbb{SO}(E)$.

Corollaire 7.3. Soit E un espace euclidien orienté de dimension 3 et $a \in \mathbb{SO}(E)$. Alors 1 est valeur propre, a a une droite $\mathbb{R}e_3$ de vecteurs fixes (appelée *axe de rotation*) et il existe un nombre θ tel que pour toute bon directe $e = (e_1, e_2, e_3)$ on ait

$$[a]_e^e = \begin{bmatrix} R(\theta) & 0 \\ 0 & 1 \end{bmatrix}.$$

En particulier, $\cos\theta = \frac{1}{2}(\mathrm{trace}\, a - 1)$. Si $a \in \mathbb{O}_-(E)$ alors -1 est valeur propre. Si e_3 est tel que $a(e_3) = -e_3$ alors il existe un nombre θ tel que pour toute bon directe $e = (e_1, e_2, e_3)$ on ait

$$[a]_e^e = \begin{bmatrix} R(\theta) & 0 \\ 0 & -1 \end{bmatrix}.$$

Remarques : On pourrait énoncer un résultat voisin sans orienter E, mais la présente version est plus utile. Notez qu'on ne peut pas parler d'angle algébrique de rotation en 3 dimensions, même à 2π près. En effet si $a \in \mathbb{SO}(E)$ le choix de e_3 a une ambiguité de signe qui se reporte sur θ. On peut quand même visualiser un élément de $\mathbb{SO}(E)$ comme une rotation autour d'un axe, et un élément de $\mathbb{O}_-(E)$ comme la donnée d'un plan vectoriel F par rapport auquel on fait une symétrie orthogonale suivie d'une rotation autour de l'axe perpendiculaire au plan F. Si $a \in \mathbb{SO}(E)$, on parlera de sa représentation $a = e^b$ à la Proposition 8.2 plus loin.

Démonstration : Il suffit d'appliquer le Théorème 7.2 : si $a \in \mathbb{SO}(E)$ alors il existe une bon e telle que $[a]_e^e = \mathrm{diag}(R(\theta), 1)$ ou $[a]_e^e = I_3$, puisque $2r + p = 3$ n'a que les solutions $(r, p) = (1, 1)$ ou $(0, 3)$. Si $a \in \mathbb{O}_-(E)$ alors il faut n impair, et donc $n = 1$ correspond à $[a]_e^e = \mathrm{diag}(R(\theta), -1)$ ou $[a]_e^e = \mathrm{diag}(I_2, -1)$, et $n = 3$ correspond à $[a]_e^e = -I_3$.

Proposition 7.4. Si $\theta \in \mathbb{R}$ notons

$$P(\theta) = \begin{bmatrix} R(\theta) & 0 \\ & 0 & 1 \end{bmatrix}, \quad Q(\theta) = \begin{bmatrix} 1 & & 0 \\ 0 & & R(\theta) \end{bmatrix}.$$

Alors pour tout $A \in \mathbb{SO}(3)$ il existe trois nombres (ψ, θ, φ) tels que $A = P(\varphi)Q(\theta)P(\psi)$.

Démonstration : L'espace $\mathbb{R}^3$ est muni de sa structure euclidienne et de sa bon ordonnée $e = (e_1, e_2, e_3)$ canoniques et il est orienté pour que e soit directe. Soit $a \in \mathbb{SO}(\mathbb{R}^3)$ tel que $A = [a]_e^e$. Soit $f = (f_1, f_2, f_3)$ la bon directe définie par $a(e_j) = f_j$. Le résultat est trivial si $e_3 = f_3$: il suffit de prendre alors $\theta = \psi = 0$. Si $f_3 = -e_3$ il suffit de prendre $\theta = \pi$ et $\psi = 0$. Si $e_3 \neq \pm f_3$, soit D la droite d'intersection des plans (e_1, e_2) et (f_1, f_2) et soit u celui des deux vecteurs de D de norme 1 tel que la base (u, e_3, f_3) soit directe. On considère alors a_1, a_2 et a_3 dans $\mathbb{SO}(\mathbb{R}^3)$ définis par $a_1(e_1) = u$ et $a_1(e_3) = e_3$ qui satisfait donc $[a_1]_e^e = P(\psi)$ pour un ψ convenable ; $a_2(u) = u$ et $a_2(e_3) = f_3$, qui satisfait donc $[a_2 \circ a_1]_e^e = Q(\theta)P(\psi)$ pour un θ convenable ; $a_3(f_3) = f_3$ $a_3(u) = f_1$ qui satisfait donc $[a_3 \circ a_2 \circ a_1]_e^e = P(\varphi)Q(\theta)P(\psi)$ pour un ψ convenable. Reste à vérifier $a = a_3 \circ a_2 \circ a_1$. Pour cela on calcule $a_3 \circ a_2 \circ a_1$ sur e_1 et e_3 et on trouve f_1 et f_3. Comme $a_3 \circ a_2 \circ a_1$ est dans $\mathbb{SO}(\mathbb{R}^3)$ qui respecte les bon directes cela entraîne que $a_3 \circ a_2 \circ a_1(e_2) = f_2$ et le résultat est montré.

Remarques : L'angle ψ est dit de *précession*, l'angle θ est dit de *nutation* et l'angle φ est dit de *rotation propre*. Ce sont les trois *angles d'Euler* d'une rotation, qui fournissent le paramétrage de $\mathbb{SO}(E)$ qu'utilisent les mécaniciens. La construction ci dessus montre que ces angles ne sont pas tout à fait uniques, mais on peut les rendre uniques si on reste pour A dans un voisinage de I_3 assez petit. Ce paramétrage en fait ne renseigne pas directement sur l'axe de rotation et l'angle de rotation.

Exercice 7.1. Soit D et D' deux droites vectorielles d'un plan euclidien orienté E. Soit $\vec{u}$ et $\vec{u'}$ des vecteurs de norme 1 qui engendrent D et D', et on suppose que $(\vec{u}, \vec{u'})$ est une base directe. On note $\langle \vec{u}, \vec{u'} \rangle = \cos\alpha$, avec $0 < \alpha < \pi$. Soit s_D et $s_{D'}$ les symétries orthogonales par rapport à D et D'. (a) Calculer l'angle des rotations $s_D \circ s_{D'}$ et $s_{D'} \circ s_D$ par rapport à α. (b) Soit p_D et $p_{D'}$ les projections orthogonales sur D et D'. Soit e_1 un vecteur de norme 1 proportionnel à $u + u'$. Soit e_2 orthogonal à e_1 de norme 1, et soit la bon $e = (e_1, e_2)$. Exprimer $[p_D]_e^e$ et $[p_{D'}]_e^e$. Montrer que $\pm(p_D - p_{D'})/\sin\alpha$ sont des symétries orthogonales par rapport aux droites $\mathbb{R}(e_1 \pm e_2)$, les bissectrices des axes engendrés par e_1 et e_2.

Exercice 7.2. Soit E l'espace vectoriel sur $\mathbb{R}$ des matrices $x = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ réelles $(2, 2)$. On note par x^* la transposée de x. On note par P_+ le sous espace vectoriel de E formé des x tels que $d = a$ et $c = -b$, et par P_- le sous espace vectoriel de E formé des x tels que $d = -a$ et $c = b$. On note par $\mathbb{O}_+$ le sous groupe du groupe orthogonal $\mathbb{O}(2)$ des x tels que $\det x = 1$. On note par $\mathbb{O}_-$ le sous ensemble du groupe orthogonal $\mathbb{O}(2)$ des x tels que $\det x = -1$. Montrer que $\langle x, x' \rangle = \text{trace}\,(x^* x')$ est un produit scalaire sur E, qu'on considère désormais comme un espace euclidien de dimension 4. Montrer que P_+ et P_- sont orthogonaux et que

$E = P_+ \oplus P_-$. Soit S la sphère de E centrée en 0 et de rayon $\sqrt{2}$. Montrer que $\mathbb{O}_+ = S \cap P_+$ et que $\mathbb{O}_- = S \cap P_-$.

Exercice 7.3. Soit a un endomorphisme antisymétrique de l'espace euclidien E. Montrer que si $y : \mathbb{R} \to L(E)$ est dérivable et satisfait $y' = ay$, alors $y(t) \in \mathbb{SO}(E)$ pour tout t (Méthode : utiliser l'exercice 9.4 du chapitre 1).

Exercice 7.4. Soit a un endomorphisme orthogonal de l'espace euclidien E de dimension n. Montrer à l'aide du Théorème 7.2 et de l'exercice 7.1 a) que a est le produit d'au plus n symétries orthogonales par rapport à des hyperplans (c'est à dire des sous espaces vectoriels de dimension $n-1$).

Exercice 7.5. L'endomorphisme orthogonal du plan H de l'exercice 6.6 est-il une symétrie orthogonale ? (méthode : chercher s'il y a des vecteurs propres).

VIII Produit vectoriel en dimension 3 et quaternions

Nous donnons d'abord une définition intrinsèque du produit vectoriel, c'est à dire libre d'un système de coordonnées. Rappelons que le déterminant d'un système de q vecteurs dans un espace euclidien orienté de dimension q a été défini à la Proposition 6.6.

Définition 8.1. Soit E un espace euclidien de dimension 3 orienté. Le *produit vectoriel* $u \wedge v$ des vecteurs u et v de E est l'unique vecteur[6] de E tel que pour tout $w \in E$ on ait $\langle u \wedge v, w\rangle = \det(u, v, w)$. En d'autres termes, $u \wedge v$ est le vecteur associé (voir section 5) à la forme linéaire sur E définie par $w \mapsto \det(u, v, w)$. Le nombre $\det(u, v, w)$ est appelé *produit mixte* des vecteurs u, v, w.

Proposition 8.1. Soit E euclidien orienté de dimension 3. On a les propriétés suivantes :

1. $(u, v) \mapsto u \wedge v$ est bilinéaire.
2. $(u, v) \mapsto u \wedge v$ est antisymétrique : $u \wedge v = -v \wedge u$.
3. Soit $e = (i, j, k)$ une bon directe de E, $[u]^e = [a, b, c]^T$ et $[v]^e = [x, y, z]^T$. Alors
$$[u \wedge v]^e = [bz - cy, cx - az, ay - bx]^T.$$
En particulier
$$i \wedge j = k,\ j \wedge k = i,\ k \wedge i = j. \tag{8.13}$$
4. $\langle u \wedge v, u\rangle = 0$ et $\langle u \wedge v, v\rangle = 0$.
5. Si $u \neq 0$ et $v \neq 0$, soit θ l'angle entre u et v, c'est à dire le nombre de $[0, \pi]$ défini par $\theta = \arccos(\langle u, v\rangle/(\|u\|\|v\|))$. Alors $\|u \wedge v\|^2 = \|u\|^2\|v\|^2 \sin^2\theta$.
6. *Double produit vectoriel* : $(u \wedge v) \wedge w = \langle u, w\rangle v - \langle v, w\rangle u$.
7. Si u et v sont indépendants, alors $(u, v, u \wedge v)$ est une base *directe*.

[6] Noté $u \times v$ et appelé *cross product* dans les ouvrages anglo-saxons.

Démonstration : 1) et 2) découlent du fait que le déterminant est une forme multilinéaire alternée. 3) s'obtient par exemple avec la règle de Sarrus : si $[w]^e = [X, Y, Z]^T$ alors

$$\det \begin{bmatrix} a & x & X \\ b & y & Y \\ c & z & Z \end{bmatrix} = (bz - cy)X + (cx - az)Y + (ay - bx)Z.$$

4) découle du fait que $\det(u, v, u) = \det(u, v, v) = 0$. 5) peut se montrer par la méthode peu élégante suivante : si e est une bon directe et $[u]^e = [a, b, c]^T$ et $[v]^e = [x, y, z]^T$ alors [7]

$$\|u \wedge v\|^2 - \|u\|^2\|v\|^2(1 - \cos^2\theta) = \|u \wedge v\|^2 - \|u\|^2\|v\|^2 + \langle u, v\rangle^2 =$$
$$(bz - cy)^2 + (cx - az)^2 + (ay - bx)^2 - (a^2 + b^2 + c^2)(x^2 + y^2 + z^2) + (ax + by + cz)^2 = 0.$$

Une autre méthode plus rapide, mais moins symétrique consiste à choisir d'abord la bon directe $e = (e_1, e_2, e_3)$ en fonction de u et v en posant $u = ae_1$ et $v = xe_1 + ye_2$ comme dans le procédé de Schmidt. Alors $[u \wedge v]^e = [0, 0, ay]^T$ et le calcul ci dessus est très facile. C'est cette méthode que nous suivons pour le 6) : Si $u = ae_1$, $v = xe_1 + ye_2$ et $w = Xe_1 + Ye_2 + Ze_3$. Alors

$$[u \wedge v]^e = [0, 0, ay]^T,\ [(u \wedge v) \wedge w]^e = [-ayY, ayX, 0]^T,$$

$$[\langle u, w\rangle v - \langle v, w\rangle u]^e = [aXx, aXy, 0]^T - [(xX + yY)a, 0, 0]^T = [-ayY, ayX, 0]^T.$$

De même pour le 7), avec $u = ae_1$, $v = xe_1 + ye_2$ alors $ay \neq 0$ si u et v sont indépendants, et il est clair que $f = (u, v, u \wedge v)$ est une base puisque, d'après le 4) le vecteur $u \wedge v$ est orthogonal au plan engendré par u et v. Comme $[\mathrm{id}_E]^e_f$ est triangulaire supérieure de diagonale (a, y, ay) son déterminant a^2y^2 est > 0 et f est directe : la proposition est montrée.

Proposition 8.2. Soit E euclidien orienté de dimension 3 et soit $\mathcal{A}$ le sous espace vectoriel de $L(E)$ formé des endomorphismes antisymétriques. Soit $u \in E$ et $\varphi_u(x) = u \wedge x$. Alors φ_u est dans $\mathcal{A}$ et l'application $u \mapsto \varphi_u$ est un isomorphisme entre E et $\mathcal{A}$. Si e est une bon directe et si $[u]^e = [a, b, c]^T$, alors

$$[\varphi_u]^e_e = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix}, \tag{8.14}$$

Enfin, si $\theta = \|u\| > 0$, si $f = (f_1, f_2, f_3)$ est une bon directe telle que $f_3 = u/\theta$, alors on a

$$[\exp \varphi_u] = \begin{bmatrix} R(\theta) & 0 \\ 0 & 1 \end{bmatrix},$$

c'est à dire que $\exp \varphi_u$ est une rotation d'axe u et d'angle θ.

[7] On a ici une nouvelle démonstration de l'inégalité de Schwarz en dimension 3, avec représentation explicite de la différence par $\|u \wedge v\|^2 = \|u\|^2\|v\|^2 - \langle u, v\rangle^2$.

Démonstration : D'après la Proposition 8.1, partie 1), φ_u est bien linéaire. Calculons son adjoint :

$$\langle v, \varphi_u^*(w)\rangle = \langle \varphi_u(v), w\rangle = \det(u, v, w) = -\det(u, w, v) = -\langle \varphi_u(w), v\rangle = -\langle v, \varphi_u(w)\rangle.$$

Donc $\varphi_u^* = -\varphi_u$ et φ_u est bien antisymétrique. La Proposition 8.1 partie 3) montre que $[\varphi_u]_e^e$ a bien la forme indiquée en (8.14). Il est clair que $u \mapsto \varphi_u$ est linéaire de E dans $\mathcal{A}$ et que toute matrice rélle $(3,3)$ antisymétrique a la forme du second membre de (8.14). Donc $\dim \mathcal{A} = 3$ et $u \mapsto \varphi_u$ est un isomorphisme.

Ensuite, si la bon f est comme indiquée, alors $[a, b, c]^T = [0, 0, \theta]^T$, avec $\theta = \|u\|$. Si $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ alors (8.14) devient $[\varphi_u]_e^e = \mathrm{diag}[\theta A, 1]$ et avec l'exemple 9.1 du chapitre 1 qui affirme $\exp(\theta A) = R(\theta)$, on a le résultat annoncé.

Corollaire 8.3. Soit E euclidien orienté de dimension 3 et $u \in E$ tel que $\theta = \|u\| > 0$. Soit $a = \exp \varphi_u$ et $v \in E$ défini par $\varphi_v = a - a^*$. Alors $v = \frac{2\sin\theta}{\theta} u$. En particulier si $\theta \not\equiv \pi/2 \bmod \pi$, u est proportionnel à v.

Démonstration : L'existence de v découle de la proposition et du fait que $b = a - a^*$ est antisymétrique. Soit $f = (f_1, f_2, f_3)$ une bon directe telle que $f_3 = u/\theta$. Alors d'après la Proposition 8.2

$$[\varphi_v]_f^f = [a - a^*]_f^f = \begin{bmatrix} 0 & -2\sin\theta & 0 \\ 2\sin\theta & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{8.15}$$

et donc d'après (8.14) on a $[v]^f = [0, 0, 2\sin\theta]^T$. Comme $[u]^f = [0, 0, \theta]^T$, on a le résultat.

Remarque. L'intérêt du corollaire est qu'il permet de calculer rapidement l'axe de rotation d'un élément $a \in \mathbb{SO}(E)$ quand celui ci est connu par sa matrice représentative $A = [a]_e^e$ dans une bon directe quelconque, et d'avoir avec $\|v\| = 2|\sin\theta|$ un renseignement partiel sur l'angle de rotation. En effet le vecteur v dans la base e est donné par $[v]^e = [a, b, c]^T$ avec

$$A - A^T = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix}.$$

Expliquons maintenant le lien qui existe entre ces questions et l'*axe instantané de rotation* des mécaniciens. Quand on considère le mouvement d'un solide gardant un point fixe O dans l'espace euclidien E à trois dimensions, on considère en fait une application continuement dérivable $t \mapsto a_t$ d'un intervalle $[0, t_1]$ dans $\mathbb{SO}(E)$ telle que $a_0 = \mathrm{id}_E$. Si on veut, si s est une bon directe attachée au solide, soit e_t la position de s à l'instant t. C'est encore une bon directe, le mouvement du solide se faisant sans déformation et de facon continue. L'endomorphisme orthogonal a_t est défini par $a_t(e_0) = e_t$.

Proposition 8.4. Soit E un espace euclidien de dimension quelconque, I un intervalle ouvert et $t \mapsto a_t$ une application continûment dérivable de I dans $\mathbb{SO}(E)$ et soit $(a_t)' \in$

$L(E)$ la dérivée de a_t au point t. Alors

$$b_t = (a_t)' a_t^* = \lim_{h \to 0} \frac{1}{h}(a_{t+h} a_t^{-1} - \mathrm{id}_E)$$

est antisymétrique. En particulier, si $\dim E = 3$, il existe un vecteur v_t (et $\mathbb{R}v_t$ est alors appelé axe instantané de rotation) tel que pour tout $x \in E$ on a

$$\lim_{h \to 0} \frac{1}{h}(a_{t+h} a_t^{-1} - \mathrm{id}_E)(x) = v_t \wedge x.$$

Démonstration : Dérivons par rapport à t l'égalité $a_t a_t^* = \mathrm{id}_E$. On obtient $a_t(a_t^*)' + (a_t)' a_t^* = 0$. Notons que $(a_t^*)'$ est l'adjoint de $(a_t)'$. Donc $(b_t)^* + b_t = 0$. Dans le cas $\dim E = 3$ on applique alors la Proposition 8.2 à $b_t = \varphi_{v_t}$.

Pour terminer cette section sur les liens entre le produit vectoriel et le groupe des rotations en dimension 3, nous présentons *l'algèbre des quaternions*, qui sert en particulier à paramétrer $\mathbb{SO}(3)$ et $\mathbb{SO}(4)$ comme on va le voir au Théorème 8.6. Nous définissons l'algèbre $\mathbf{H}$ des quaternions comme un espace euclidien orienté de dimension 4 dans lequel on a sélectionné un vecteur de norme 1 appelé *unité* et noté $\mathbf{1}$. On note alors E l'espace de dimension 3 orthogonal à $\mathbf{1}$. On oriente enfin E de sorte que $(\mathbf{1}, e)$ soit une base directe de $\mathbf{H}$ si e est une base directe de E. On identifiera $\mathbf{H} = \mathbb{R}\mathbf{1} \oplus E$ avec $(\mathbb{R}, E)$ et nous noterons alors les éléments x de $\mathbf{H}$ par $x = (x_0, \vec{x})$ avec $\Re x = x_0 \in \mathbb{R}$ et $\vec{x} \in E$. Cette notation est plus commode que $x = x_0\mathbf{1} + \vec{x}$ pour les calculs. Le produit scalaire dans E sera noté par $\vec{x} \cdot \vec{y}$ si bien que le produit scalaire dans $\mathbf{H}$ est

$$\langle (x_0, \vec{x}), (y_0, \vec{y}) \rangle = x_0 y_0 + \vec{x} \cdot \vec{y}.$$

Les quaternions de la forme $(x_0, 0)$, c'est à dire les multiples de $\mathbf{1}$ sont dits *réels*, les quaternions de la forme $(0, \vec{x})$ sont dits *purs*. Le quaternion $\overline{x} = (x_0, -\vec{x})$ est dit *conjugué* de $x = (x_0, \vec{x})$. On définit enfin le *produit* des quaternions $x = (x_0, \vec{x})$ et $y = (y_0, \vec{y})$ par

$$xy = (x_0 y_0 - \vec{x} \cdot \vec{y}, x_0 \vec{y} + y_0 \vec{x} + \vec{x} \wedge \vec{y}). \tag{8.16}$$

On voit immédiatement que $xy = yx$ si et seulement si $\vec{x} \wedge \vec{y} = \vec{0}$, ce qui n'est pas le cas général. La proposition suivante centralise les propriétés algébriques du produit des quaternions.

Proposition 8.5. Soit x, y, z des quaternions. Alors

1. $(xy)z = x(yz)$.
2. $\|xy\|^2 = \|x\|^2 \|y\|^2$.
3. $x\overline{x} = \overline{x}x = \|x\|^2 \mathbf{1}$. En particulier si $x \neq 0$ et $x^{-1} = \overline{x}/\|x\|^2$ alors $xx^{-1} = x^{-1}x = \mathbf{1}$. De plus $\langle xy, z \rangle = \langle y, \overline{x}z \rangle$, $\langle yx, z \rangle = \langle y, z\overline{x} \rangle$, $\overline{xy} = \overline{y}\,\overline{x}$ et $\langle x, y \rangle = \Re(x\overline{y})$.
4. Si $(\vec{i}, \vec{j}, \vec{k})$ est une bon directe de l'espace E des quaternions purs alors

$$\vec{i}\vec{j} = -\vec{j}\vec{i} = \vec{k},\ \vec{j}\vec{k} = -\vec{k}\vec{j} = \vec{i},\ \vec{k}\vec{i} = -\vec{i}\vec{k} = \vec{j},\ \vec{i}^2 = \vec{j}^2 = \vec{k}^2 = \vec{i}\vec{j}\vec{k} = -\mathbf{1}.$$

De plus, si

$$\mathcal{A} = \left\{ \begin{bmatrix} a & b \\ -\overline{b} & \overline{a} \end{bmatrix} ;\ a, b \in \mathbb{C} \right\}$$

alors l'application $x \mapsto A(x)$ de $\mathbf{H}$ dans $\mathcal{A}$ définie par

$$A(x_0\mathbf{1} + x_1\vec{i} + x_2\vec{j} + x_3\vec{k}) = \begin{bmatrix} x_0 + ix_1 & i(x_2 + ix_3) \\ i(x_2 - ix_3) & x_0 - ix_1 \end{bmatrix} \tag{8.17}$$

est un isomorphisme d'algèbres, c'est à dire que A est linéaire bijective et $A(xy) = A(x)A(y)$.

Remarques. Le 1) montre que l'application bilinéaire $(x, y) \mapsto xy$ de $\mathbf{H} \times \mathbf{H}$ dans $\mathbf{H}$ fait de $\mathbf{H}$ une algèbre *associative.* Le 2) montre que si $\|x\|^2 = 1$ alors $y \mapsto xy$ et $y \mapsto yx$ sont des endomorphismes orthogonaux de l'espace euclidien $\mathbf{H}$. Le 3), qui dit que tout élément non nul de $\mathbf{H}$ a un inverse, montre que $\mathbf{H}$ est un *corps non commutatif* (ou corps gauche). Le 3) montre aussi que les adjoints de $y \mapsto xy$ et $y \mapsto yx$ sont $z \mapsto \overline{x}z$ et $z \mapsto z\overline{x}$. Le 4) a une importance historique car c'est sous cette forme avec coordonnées que William Hamilton a découvert les quaternions. Comme dans le reste de ce cours, nous avons préféré une exposition avec un minimum de coordonnées lorsque celles ci ne s'imposent pas géométriquement. Le 4) montre que $\mathbf{H}$ peut être paramétré par $\mathbb{C}^2$. Le 4) implique aussi que si $p \in E$ est de norme 1 alors $\mathbb{R}\mathbf{1} + \mathbb{R}p$ est une sous algèbre de $\mathbf{H}$ de dimension 2 isomorphe au corps des nombres complexes qu'Hamilton cherchait à généraliser.

Démonstration : Pour le 1) on le vérifie d'abord en supposant que (x, y, z) sont des quaternions purs. Dans ces conditions on a

$$\begin{aligned} (xy)z - x(yz) &= (\vec{x} \cdot \vec{y}, \vec{x} \wedge \vec{y})z - x(\vec{y} \cdot \vec{z}, \vec{y} \wedge \vec{z}) \\ &= ((\vec{x} \wedge \vec{y}) \cdot \vec{z}, (\vec{x} \cdot \vec{y})\vec{z} + (\vec{x} \wedge \vec{y}) \wedge \vec{y}) \wedge \vec{z} \\ &\quad -(\vec{x} \cdot (\vec{y} \wedge \vec{z}), \vec{x} \wedge (\vec{y} \cdot \vec{z}) + \vec{x} \wedge (\vec{y} \wedge \vec{z}) = (0, \vec{0}), \end{aligned}$$

la dernière égalité étant héritée des propriétés du produit mixte (pour la composante réelle) et de la formule du double produit vectoriel de la Proposition 8.1 (pour la composante quaternions purs). Pour passer au cas général, on observe que 1) est trivial si un des (x, y, z) est un quaternion réel, ce qui permet de constater facilement que $(xy)z - x(yz) = (\vec{x}\vec{y})\vec{z} - \vec{x}(\vec{y}\vec{z}) = 0$.

Pour le 2) rappelons que d'après la Proposition 8.2

$$\|\vec{x} \wedge \vec{y}\|^2 = \|\vec{x}\|^2\|\vec{y}\|^2 - (\vec{x} \cdot \vec{y})^2.$$

Donc

$$\begin{aligned} \|xy\|^2 &= (x_0y_0 - \vec{x} \cdot \vec{y})^2 + x_0^2|\vec{y}\|^2 + y_0^2|\vec{x}\|^2 + 2x_0y_0\vec{x} \cdot \vec{y} + \|\vec{x} \wedge (\vec{y}\|^2 \\ &= (x_0^2 + |\vec{x}\|^2)(y_0^2 + |\vec{y}\|^2) = \|x\|^2\|y\|^2. \end{aligned}$$

Le 3) et le 4) sont des conséquences immédiates de la définition (8.16) du produit des quaternions. Pour l'isomorphisme avec l'algèbre $\mathcal{A}$, la linéarité et la bijectivité sont évidentes. La vérification de $A(xy) = A(x)A(y)$ est un peu laborieuse, mais élémentaire.

Voici maintenant deux extraordinaires paramétrisations des groupes $\mathbb{SO}(3)$ et $\mathbb{SO}(4)$ par la sphère unité des quaternions.

Théorème 8.6. Soit S la sphère unité de l'algèbre $\mathbf{H}$ des quaternions et soit E l'ensemble des quaternions purs. Alors

1. S est un groupe pour la multiplication.
2. Si $s \in S$ alors E est stable pour $x \mapsto sxs^{-1}$. Si φ_s est la restriction de $x \mapsto sxs^{-1}$ à E alors $s \mapsto \varphi_s$ est un homomorphisme surjectif du groupe S sur le groupe $\mathbb{SO}(E)$ de noyau $\pm\mathbf{1}$.
3. Si s et t sont dans S et si $\psi_{s,t}(x) = sxt^{-1}$ alors $(s,t) \mapsto \psi_{s,t}$ est un homomorphisme surjectif du groupe[8] $S \times S$ sur le groupe $\mathbb{SO}(\mathbf{H})$ de noyau $\pm(\mathbf{1},\mathbf{1})$.

Démonstration : Le 1) est évident avec la Proposition 8.5. Pour le 2), on observe que $x \mapsto sxs^{-1}$ est dans $\mathbb{O}(\mathbf{H})$ puisque linéaire et puisque préservant la norme (d'après le 2) de la Proposition 8.5). De plus $s\mathbf{1}s^{-1} = \mathbf{1}$ et donc $\mathbb{R}\mathbf{1}$ est stable. Donc d'après la Proposition 6.5 son orthogonal E est stable. La restriction φ_s à E est donc un élément de $\mathbb{O}(E)$. Explicitement :

$$\varphi_s(\vec{x}) = (s_0^2 - \|\vec{s}\|^2)\vec{x} + 2s_0(\vec{s} \wedge \vec{x}) + 2(\vec{s} \cdot \vec{x})\vec{s}. \tag{8.18}$$

Si $\vec{s} = \vec{0}$ alors $s = \pm\mathbf{1}$ et $\varphi_s = \mathrm{id}_E$. Si $\vec{s} \neq \vec{0}$ et si $e = (i,j,k)$ est une bon directe de E telle que $k = \vec{s}/\|\vec{s}\|$, alors

$$[\varphi_s]_e^2 = \mathrm{diag}(R(\theta), 1) \tag{8.19}$$

avec $\cos\theta = s_0^2 - \|\vec{s}\|^2$ et $\sin\theta = 2s_0\|\vec{s}\|$, et $\det\varphi_s = 1$ est clair. En fait, (8.18) et (8.19) permettent de voir que tout élément a de $\mathbb{SO}(E)$ est de la forme φ_s pour un $s \in S$ convenable. En effet si $a \in \mathbb{SO}(E)$ alors il existe une bon directe $e = (i,j,k)$ telle que $[a]_e^e = \mathrm{diag}(R(\theta), 1)$ pour un $\theta \in [0, 2\pi[$. On prend alors $\alpha = \theta/2$ et $s = (\cos\alpha, k\sin\alpha)$. Enfin $s \mapsto \varphi_s$ est un homomorphisme car

$$\varphi_t \circ \varphi_s(x) = \varphi_t(sxs^{-1}) = tsxs^{-1}t^{-1} = (ts)x(ts)^{-1} = \varphi_{ts}(x).$$

On a vu qu'il est surjectif. Son noyau est l'ensemble des $s \in S$ tels que $s\vec{x}s^{-1} = \vec{x}$ pour tout $\vec{x} \in E$ c'est à dire tels que $\vec{s} \wedge \vec{x} = \vec{0}$ pour tout $\vec{x} \in E$. Ceci n'est possible que si $\vec{s} = \vec{0}$, c'est à dire si $s = \pm\mathbf{1}$. La partie 2) est donc montrée.

Pour le 3) il est clair que $\psi_{s,t}$ est dans $\mathbb{O}(\mathbf{H})$ par le 2) de la Proposition 8.7. Pour voir que $\det\psi_{s,t} = 1$ on peut écrire $\psi_{s,t} = \psi_{s,\mathbf{1}} \circ \psi_{\mathbf{1},t}$, calculer la matrice représentative de $\psi_{s,\mathbf{1}}$ dans une bon $(\mathbf{1}, i, j, k)$ telle que $\vec{s}$ est proportionnelle à k et calculer le déterminant de cette matrice d'ordre 4 pas trop compliquée. On procède ensuite de même pour $\psi_{\mathbf{1},t}$. Toutefois une méthode moins élémentaire mais plus rapide utilise la connexité : il est clair que $s \mapsto \det\psi_{s,\mathbf{1}}$ est une application polynomiale, donc continue sur S. La sphère unité d'un espace euclidien étant une partie connexe par arc, l'ensemble des valeurs prises sur

[8] muni du produit $(s,t)(s_1,t_1) = (ss_1, tt_1)$.

S par $s \mapsto \det \psi_{s,\mathbf{1}}$ est connexe. Cet ensemble est une partie de $\{-1, 1\}$, et il contient 1 comme on le voit en faisant $s = \mathbf{1}$. Il est donc réduit au singleton $\{1\}$ et donc $\psi_{s,\mathbf{1}} \in \mathbb{SO}(\mathbf{H})$.

Le point plus ingénieux de la démonstration est de montrer que tout élément a de $\mathbb{SO}(\mathbf{H})$ est de la forme $\psi_{s,t}$ pour quelque (s,t) de $S \times S$. Pour cela on note $r = a(\mathbf{1})$, qui est de norme 1. On introduit ensuite $b = \psi_{r,\mathbf{1}} \in \mathbb{SO}(\mathbf{H})$. Alors $b^{-1}a$ est dans $\mathbb{SO}(\mathbf{H})$ *et préserve* $\mathbf{1}$. Donc sa restriction à E est un élément de $\mathbb{SO}(\mathbf{E})$. D'après le 2) il existe donc $t \in S$ tel que $b^{-1}a(x) = txt^{-1}$, ce qui montre que $a = \psi_{rt,t}$. Vérifier que $(s,t) \mapsto \psi_{s,t}$ est bien un homomorphisme de $S \times S$ dans $\mathbb{SO}(\mathbf{H})$ est facile. Pour voir que $\psi_{s,t} = \mathrm{id}_{\mathbf{H}}$ si et seulement si $(s,t) = \pm(\mathbf{1},\mathbf{1})$, on exploite $sx = xt$ pour tout $x \in \mathbf{H}$ en faisant $x = s^{-1}$ ce qui montre $t = s$. Le 2) permet de conclure.

Exercice 8.1. Si $c = 1/\sqrt{2}$ soit la matrice de $\mathbb{SO}(3)$ suivante

$$A = \begin{bmatrix} -c & c^2 & c^2 \\ c & c^2 & c^2 \\ 0 & c & c \end{bmatrix}.$$

A l'aide du corollaire 8.3 donner l'axe de rotation dans $\mathbb{R}^3$ correspondant. A l'aide du corollaire 7.3 donner le cosinus de l'angle de rotation correspondant.

Exercice 8.2. Calculer $(u \wedge v) \wedge (w \wedge x)$.

Exercice 8.3. Soit $(\vec{i}, \vec{j}, \vec{k})$ une bon directe de l'espace des quaternions purs et $e = (\mathbf{1}, \vec{i}, \vec{j}, \vec{k})$. Si $[x]^e = (x_0, x_1, x_2, x_3)^T$ calculer les matrices carrées d'ordre 4 représentatives dans la base e de l'espace $\mathbf{H}$ des quaternions des endomorphismes $y \mapsto xy$ et $y \mapsto yx$.

Exercice 8.4. Soit $(\vec{i}, \vec{j}, \vec{k})$ une bon directe de l'espace des quaternions purs. Montrer que pour tout quaternion x on a

$$\overline{x} = -\frac{1}{2}(x + \vec{i}x\vec{i} + \vec{j}x\vec{j} + \vec{k}x\vec{k})$$

Exercice 8.5. Trouver tous les quaternions x tels que $x^2 + 1 = 0$.

Exercice 8.6. Soit $(\vec{i}, \vec{j}, \vec{k})$ une bon directe de l'espace des quaternions purs. Montrer que les 24 éléments de la sphère S unité des quaternions

$$\pm 1, \quad \pm\vec{i}, \quad \pm\vec{j}, \quad \pm\vec{k}, \quad \frac{1}{2}(\pm 1 \pm \vec{i} \pm \vec{j} \pm \vec{k})$$

forment un sous groupe G du groupe multiplicatif S. Ecrire les 12 matrices des rotations de l'espace E des quaternions purs dans la base $(\vec{i}, \vec{j}, \vec{k})$ qui sont l'image de G par l'application $s \mapsto \varphi_s$ du théorème 8.6 partie 2. Soit G_1 l'image de G . Montrer que si on note

$$s = \vec{i} + \vec{j} + \vec{k}, \ s_1 = 2\vec{i} - s, \ s_2 = 2\vec{j} - s, \ s_3 = 2\vec{k} - s$$

alors $T = \{s, s_1, s_2, s_3\}$ forme un tétraèdre régulier et que $g(T) = T$ pour toute rotation g de G_1.

Exercice 8.7. Imiter l'exercice 8.6 quand le tétraèdre est remplacé par le cube C dont les huit sommets sont $\pm\vec{i}\pm\vec{j}\pm\vec{k}$. On admet que le groupe $I(C)$ des isométries g de E qui préserve C est ici formé des 48 g dont la matrice représentative dans la base $(\vec{i}, \vec{j}, \vec{k})$ est de la forme PD où P est une des 6 matrices de permutation de 3 éléments et $D = \mathrm{diag}(\pm 1, \pm 1, \pm 1)$. Le sous groupe G_1 formé des seules rotations a donc 24 éléments. Trouver les 48 éléments du sous groupe G de S tel que G_1 soit l'image de G par l'homomorphisme $s \mapsto \varphi_s$. Méthode : utiliser la formule 8.18 pour trouver les deux s de S tels que $\varphi_s = g$.

Exercice 8.8. Soit $s = (s_0, \overrightarrow{s})$ dans la sphère unité S des quaternions tel que $\overrightarrow{s} \neq 0$. Soit θ tel que $s_0 = \cos\frac{\theta}{2}$ et $\|\overrightarrow{s}\| = \sin\frac{\theta}{2}$. Montrer que $x \mapsto sxs^{-1}$ est une rotation de l'espace E des quaternions purs d'axe $\overrightarrow{s}/\|\overrightarrow{s}\|$ et d'angle θ. Méthode : analyser la démonstration du théorème 8.6 partie 2.

IX Endomorphismes symétriques, positifs et définis positifs

Si A est une matrice carrée réelle, elle est dite symétrique si $A = A^T$. Si E est euclidien et si $a \in L(E)$ on dit que a est *symétrique* si $a = a^*$. Puisque si e est une bon on a $[a^*]_e^e = ([a]_e^e)^T$, il est clair que les trois propriétés suivantes sont équivalentes :

- L'endomorphisme a de l'espace euclidien E est symétrique.
- Il existe une bon e de E telle que $A = [a]_e^e$ soit symétrique.
- Pour toute bon e de E, alors $A = [a]_e^e$ est symétrique

Il est évident aussi qu'une matrice diagonale est symétrique : la propriété 2) ci dessus entraîne donc que si a dans $L(E)$ est diagonalisable *en base orthonormale* alors a est symétrique. Le *théorème le plus important de l'algèbre linéaire* dit que la réciproque est vraie. C'est le 2) du théorème ci dessous. On l'appelle aussi souvent le *théorème spectral.*

Théorème 9.1. Soit $a \in L(E)$ avec E euclidien tel que a soit symétrique. Alors

1. Si F est un sous espace vectoriel de E tel que $a(F) \subset F$, alors $a(F^\perp) \subset F^\perp$. En particulier, les espaces propres sont deux à deux orthogonaux.
2. Il existe une bon e telle que $[a]_e^e$ soit diagonale.
3. On a $\langle a(x), x\rangle \geq 0$ pour tout $x \in E$ si et seulement si les valeurs propres de a sont ≥ 0.
4. On a $\langle a(x), x\rangle > 0$ pour tout $x \in E \setminus \{0\}$ si et seulement si les valeurs propres de a sont > 0.

Démonstration : Pour le 1) soit $y \in F^\perp$. Donc pour tout $x \in F$ on a $\langle a(y), x\rangle = \langle y, a(x)\rangle = 0$ car $a = a^*$ et car F est stable par a. Pour le 2), on procède par récurrence sur $q = \dim_E$. C'est trivial pour $q = 1$. Pour $q = 2$, on prend une bon e quelconque et on

écrit $[a]_e^e = \begin{bmatrix} r & s \\ s & t \end{bmatrix}$. Alors le polynôme caractéristique $P_a(X) = X^2 - (r+t)X + (rt - s^2)$ a pour discriminant $\Delta = (r-t)^2 + 4s^2 \geq 0$ et donc a a au moins un vecteur propre e_1, avec e_1 de norme 1. Appliquons le 1) à $F = \mathbb{R}e_1$. Donc avec e_2 de norme 1, $F^\perp = \mathbb{R}e_2$ est aussi un espace stable et $e = (e_1, e_2)$ est une base de diagonalisation. Supposons enfin le résultat montré pour tout entier inférieur à $q > 2$. On sait d'après la Proposition 1.1 que tout endomorphisme d'un espace réel de dimension finie a au moins un espace stable F de dimension 1 ou 2. Appliquons ceci à a. D'après le 1), $F^\perp$ est stable par a, et notons a_1 et a_2 les restrictions respectives de a à F et $F^\perp$. Il est clair que a_1 et a_2 sont symétriques. Appliquons l'hypothèse de récurrence et soit e_1 et e_2 des bon de diagonalisation de a_1 et a_2. Alors $e = e_1 \cup e_2$ est une bon de diagonalisation de a et la récurrence est étendue.
3) et 4) : Appliquons le 2) et soit e une bon de diagonalisation de a. Notons $[a]_e^e = \mathrm{diag}(\mu_1, \ldots, \mu_q)$. La suite $(\mu_1, \ldots, \mu_q)$ est donc celle des valeurs propres de a répétées en fonction de leurs multiplicités. Alors $\langle a(e_j), e_j \rangle = \mu_j$. Si $\langle a(x), x \rangle \geq 0$ pour tout x comme dans 3) alors $\mu_j \geq 0$. Si $\langle a(x), x \rangle > 0$ pour tout $x \neq 0$ comme dans 4) alors $\mu_j > 0$. Inversement si $[x]^e = [x_1, \ldots, x_q]$ alors

$$\langle a(x), x \rangle = \mu_1 x_1^2 + \mu_2 x_2^2 + \cdots + \mu_q x_q^2. \tag{9.20}$$

Si $\mu_j \geq 0$ pour tout $j = 1, \ldots, q$ comme dans 3), alors pour tout $(x_1, \ldots, x_q)$ le second membre de (9.20) est ≥ 0. Si $\mu_j > 0$ pour tout $j = 1, \ldots, q$ comme dans 4), alors pour tout $(x_1, \ldots, x_q) \neq (0, \ldots, 0)$ le second membre de (9.20) est > 0 et le théorème est montré.

Corollaire 9.2. Soit A une matrice réelle symétrique d'ordre q. Alors il existe des matrices diagonale D et orthogonale P d'ordre q telles que $A = PDP^{-1} = PDP^T$.

Démonstration : Considérons A comme la matrice représentative d'un endomorphisme a de $\mathbb{R}^q$ espace euclidien canonique dans la base e canonique. Celle ci étant orthonormale, alors a est symétrique et il existe une bon f telle que $D = [a]_f^f$ soit diagonale. Puisque $P = [\mathrm{id}_E]_f^e$ est la matrice de changement de la bon e vers la bon f, alors $P \in \mathbb{O}(q)$. L'égalité $[a]_e^e = [\mathrm{id}_E]_f^e [a]_f^f [\mathrm{id}_E]_f^e$ se traduit donc $A = PDP^{-1}$, ou $A = PDP^T$ puisque $PP^T = I_q$.

Exemple 9.1. Si $t \in \mathbb{R}$ et soit $A = \begin{bmatrix} 0 & 1 \\ 1 & e^t - e^{-t} \end{bmatrix}$. Pour l'écrire $A = PDP^T$ avec D diagonale et P orthogonale, on calcule les valeurs propres de A, qui vont donner D. On obtient facilement que les valeurs propres sont e^t et $-e^{-t}$ et donc $D = \begin{bmatrix} e^t & 0 \\ 0 & -e^{-t} \end{bmatrix}$. Dans la base canonique e de $E = \mathbb{R}^2$, un vecteur propre pour e^t est $[v_1]^e = [1, e^t]^T$ et un vecteur propre pour $-e^{-t}$ est $[v_2]^e = [-e^t, 1]^T$. Ces vecteurs propres sont orthogonaux, mais ne sont pas de norme 1 et la matrice de changement de base $[\mathrm{id}_E]_v^e$ n'est pas orthogonale. Pour fabriquer une bon f de vecteurs propres il faut encore normaliser v_1 et v_2. On obtient en posant $c = (1 + e^{2t})^{-1/2}$ les vecteurs propres normalisés $f_1 = cv_1$ et $f_2 = cv_2$. Alors $P = [\mathrm{id}_E]_f^e = c \begin{bmatrix} 1 & e^t \\ -e^t & 1 \end{bmatrix}$. L'égalité $A = [a]_e^e = [\mathrm{id}_E]_f^e [a]_f^f [\mathrm{id}_E]_e^f = PAP^{-1} = PAP^T$ se traduit par

$$\begin{bmatrix} 0 & 1 \\ 1 & e^t - e^{-t} \end{bmatrix} = \begin{bmatrix} c & -ce^t \\ ce^t & c \end{bmatrix} \begin{bmatrix} e^t & 0 \\ 0 & -e^{-t} \end{bmatrix} \begin{bmatrix} c & ce^t \\ -ce^t & c \end{bmatrix}.$$

Définition 9.1. Un endomorphisme *symétrique* a de l'espace euclidien est dit *positif* s'il satisfait la propriété suivante : pour tout $x \in E$ on a $\langle a(x), x\rangle \geq 0$. Il est dit *défini-positif* s'il satisfait la propriété suivante : pour tout $x \in E \setminus \{0\}$ on a $\langle a(x), x\rangle > 0$. Une matrice symétrique A d'ordre q est dite *positive*[9] si elle satisfait la propriété suivante : pour tout $X \in \mathbb{R}^q$ on a $X^TAX \geq 0$. Elle est dite *définie-positive* si elle satisfait la propriété suivante : pour tout $X \in \mathbb{R}^q \setminus \{0\}$ on a $X^TAX > 0$.

Remarques. Il est clair que si a est symétrique et si e est une bon, alors $A = [a]_e^e$ est positive ou définie positive en même temps que a. Par le Théorème 9.1 parties 3) et 4), l'endomorphisme symétrique a (ou la matrice $A = [a]_e^e$) est positif si et seulement si ses valeurs propres sont toutes ≥ 0, et il est défini-positif si et seulement si ses valeurs propres sont toutes > 0. Ces propriétés caractéristiques sont parfois prises comme définition par certains auteurs. Une matrice positive n'a pas nécessairement ses coefficients ≥ 0 : exemple $\begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}$. Une matrice symétrique à coefficients positifs n'est pas nécessairement positive : exemple $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$.

Voici un ensemble de propriétés simples des endomorphismes positifs et définis positifs.

Proposition 9.3. Soit E et F euclidiens. Alors

1. Si $a \in L(E)$ est symétrique alors la forme bilinéaire sur $E \times E$ définie par $B_a(x, y) = \langle a(x), y\rangle$ est symétrique. Inversement si B est une forme biliéaire symétrique E il existe un unique $a \in L(E)$ symétrique tel que $B = B_a$. Elle est positive ou définie-positive si et seulement si a l'est.
2. Si $\lambda > 0$ et $\mu > 0$ et si a et b dans $L(E)$ sont symétriques positifs (ou définis positifs), alors $\lambda a + \mu b$ sont positifs (ou définis positifs).
3. Si $a \in L(E)$ est symétrique positif, alors $\det a \geq 0$. Si a est symétrique défini-positif, alors $\det a > 0$. Si a est positif, alors a est défini-positif si et seulement si $\det a \neq 0$.
4. Si $a \in L(E)$ est symétrique, alors a est défini-positif si et seulement si il existe $b \in L(E)$ tel que $a = \exp b$.

[9]En parlant d'endomorphismes positifs et de matrices positives, nous nous conformons aux termes utilisés dans les programmes des classes de mathématiques supérieures et spéciales des lycées, très soigneusement rédigés. Cependant, il faut reconnaître que les utilisateurs ont souvent d'autres noms et parlent de matrices symétriques "semi-définies positives" ou "définies non-négatives". Et les mêmes utilisateurs réservent alors le nom de matrices positives à des matrices réelles non nécessairement symétriques ni même carrées, mais dont tous les coefficients sont ≥ 0. Leur théorie est certainement utile, mais leur étude est loin de l'esprit de ce cours, car la géométrie des applications linéaires correspondantes est alors liée à un système de coordonnées particulier. Il est sage que le lecteur s'assure du contexte quand il rencontrera le terme "matrice positive" dans la littérature.

5. Si $a \in L(E,F)$ alors aa^* dans $L(F)$ et a^*a dans $L(E)$ sont symétriques positives. De même si A est une matrice réelle alors AA^T et A^TA sont positives.
6. Si $a \in \mathbf{GL}(E)$ alors aa^* et a^*a sont symétriques définis positifs. De même si A est carrée inversible, AA^T et A^TA sont symétriques définies positives.

Démonstration : Le 1) est clair, sauf pour l'existence de a tel que $B = B_a$. Fixons $y \in E$ et considérons la forme linéaire $x \mapsto B(x,y)$. D'après la Proposition 5.1 il existe un vecteur $a(x) \in E$ unique tel que $B(x,y) = \langle a(x), y\rangle$. Puisque $x \mapsto B(x,y)$ est linéaire alors $x \mapsto a(x)$ est linéaire. Enfin $a = a^*$ vient de $B(x,y) = B(y,x)$. Le 2) et le 5) découlent des définitions. Le 3) découle du Théorème 9.1 parties 3) et 4). Montrons le 4) $\Leftarrow$: si e est une bon de diagonalisation de b alors

$$[b]_e^e = \mathrm{diag}(\mu_1, \dots, \mu_q), \quad [a]_e^e = \mathrm{diag}(e^{\mu_1}, \dots, e^{\mu_q})$$

et un tel a est donc bien symétrique (car diagonalisable en bon) et défini positif car à valeurs propres >0. Le 4) $\Rightarrow$ est analogue. En fait b est unique : voir exercice 10.6. Le 6) découle du 3) et du 5).

Voici des caractérisations des matrices positives et des matrices définies-positives.

Théorème 9.4. Soit $A = (a_{ij})_{1\leq i,j\leq q}$ et B des matrices carrées symétriques réelles d'ordre q. On pose $A_k = (a_{ij})_{1\leq i,j\leq k}$ pour $k = 1, \dots q$ et $A_S = (a_{ij})_{i,j\in S}$ pour $S \subset \{1, \dots, q\}$. On suppose qu'il existe une matrice inversible P d'ordre q telle que $A = P^TBP$. Alors

1. On suppose que A est positive. Alors A_S est positive pour tout S. En particulier $a_{ii} \geq 0$, et si $a_{ii} = 0$ alors $a_{ij} = a_{ji} = 0$ pour tout j. De plus B est positive.
2. Si A est définie- positive alors A_S est définie-positive pour tout S. En particulier les éléments de la diagonale de A sont > 0. De plus B est définie-positive.
3. A est définie-positive si et seulement si $\det A_k > 0$ pour tout $k = 1, \dots q$.
4. A est positive si et seulement si $\det A_S \geq 0$ pour tout $S \subset \{1, \dots, q\}$. Le rang de A est alors le maximum des S tels que $\det A_S > 0$.
5. A est positive si et seulement si le polynôme

$$\det(A + XI_q) = X^q + c_1X^{q-1} + \cdots + c_q$$

   satisfait $c_j \geq 0$ pour tout $j = 1, \dots, q$.
6. Avec cette notation, il y a équivalence entre
   (a) A est définie positive,
   (b) $c_j \geq 0$ pour tout $j = 1, \dots, q$ et $c_q > 0$,
   (c) $c_j > 0$ pour tout $j = 1, \dots, q$ et $c_q > 0$.

Démonstration : 1) prenons $X \in \mathbb{R}^q$ avec $X = [x_1, \dots, x_q]^T$ tel que $x_j = 0$ si $j \notin S$, et notons X_S la restriction de X à S. Alors $0 \leq X^TAX = X_S^TA_SX_S$ et le résultat est clair.

Si on applique cela à $S=\{i\}$ on obtient $a_{ii}\geq 0$. Si on applique cela à $S=\{i,j\}$ quand $a_{ii}=0$, alors $\det A_S=-a_{ij}^2\geq 0$ et donc $a_{ij}=0$. Pour voir que B est positive, on observe que $Y^TBY=X^TAX\geq 0$ en prenant $Y=PX$. On procède de même pour 2). Pour le 3) $\Rightarrow$, cela découle du 2) et de la Proposition 9.3 partie 3).

Le 3) $\Leftarrow$ est plus délicat. Observons qu'en général on a si A et C sont carrées symétriques d'ordre p et n avec A inversible

$$\left[\begin{array}{cc} A & B \\ B^T & C \end{array}\right]=\left[\begin{array}{cc} I_p & 0 \\ B^TA^{-1} & I_n \end{array}\right]\left[\begin{array}{cc} A & 0 \\ 0 & C-B^TA^{-1}B \end{array}\right]\left[\begin{array}{cc} I_p & A^{-1}B \\ 0 & I_n \end{array}\right]. \tag{9.21}$$

D'après les 1) et 2) appliqués à $P=\left[\begin{array}{cc} I_p & A^{-1}B \\ 0 & I_n \end{array}\right]$, la matrice $\left[\begin{array}{cc} A & B \\ B^T & C \end{array}\right]$ est positive (respectivement définie positive) si et seulement si la matrice $\left[\begin{array}{cc} A & 0 \\ 0 & C-B^TA^{-1}B \end{array}\right]$ est positive (respectivement définie positive).

Nous montrons le résultat 3) $\Leftarrow$ par récurrence sur q. C'est trivial pour $q=1$. Si c'est vrai pour $q-1$, alors A_{q-1} est inversible et (9.21) est applicable, avec $p=q-1$ et $n=1$:

$$A=\left[\begin{array}{cc} A_{q-1} & b \\ b^T & c \end{array}\right]=\left[\begin{array}{cc} I_{q-1} & 0 \\ b^TA_{q-1}^{-1} & I_q \end{array}\right]\left[\begin{array}{cc} A_{q-1} & 0 \\ 0 & c-b^TA_{q-1}^{-1}b \end{array}\right]\left[\begin{array}{cc} I_p & A_{q-1}^{-1}b \\ 0 & I_q \end{array}\right]. \tag{9.22}$$

En prenant le déterminant des deux membres on constate que si $c'=c-b^TA_{q-1}^{-1}b$ alors $\det A=c'\det A_{q-1}$ et l'hypothèse entraîne $c'>0$.

On écrit ensuite $X\in\mathbb{R}^q$ en blocs $X^T=[X_{q-1}^T,x_q]$, puis $Y^T=[X_{q-1}^T+x_qb^TA_{q-1}^{-1},x_q]$. Ceci conduit à

$$X^TAX=Y_{q-1}^TA_qY_{q-1}+x_q^2c'.$$

Si $X\neq 0$ alors ou bien $x_q\neq 0$ et $X^TAX\geq x_q^2c'>0$. Ou bien $x_q=0$, alors $Y_{q-1}=X_{q-1}\neq 0$ et, comme par l'hypothèse de récurrence A_{q-1} est définie positive on a bien $X^TAX=Y_{q-1}^TA_{q-1}Y_{q-1}>0$.

Le 4) $\Rightarrow$ est facile. Montrons le 4) $\Leftarrow$. D'après le 2) si $a_{ii}=0$ quel que soit i alors $A=0$ et A est donc positive. Supposons donc qu'il existe i tel que $a_{ii}>0$. Sans perte de généralité on suppose $a_{11}>0$. La démonstration du 4) $\Leftarrow$ est alors basée sur la formule de Laplace (Théorème 2.4 du chapitre 1) appliquée au couple $(A,B)=(A,XI_q)$: si $\det(A+XI_q)=\sum_{j=0}^q c_jX^{q-j}$ alors

$$c_j=\sum_{S;|S|=j}\det A_S, \tag{9.23}$$

qui généralise les cas déjà connus $j=1$ et $j=q$. Appliquons la au cas $X>0$. Donc, puisque par hypothèse $\det A_S\geq 0$ on en déduit $c_j\geq 0$. Donc comme $a_{11}>0$ on a $\det(A+XI_q)>0$. De la même manière on démontre que $\det(A_k+XI_k)>0$. La partie 3 entraîne donc que $A+XI_q$ est définie positive. Donc $A=\lim_{X\to 0}(A+XI_q)$ est positive.

Pour terminer, soit r le rang de A et k la taille maximum des S tels que $\det A_S>0$. D'après le résultat de première année qui dit que le rang est la taille du plus grand determinant non nul extrait on a $k\leq r$. Or c_j défini par 9.23 est nul si $j>k$ et donc le polynôme en X égal à $\det(A + XI_q)$ a zéro pour racine d'ordre au moins égal à $q - k$. En diagonalisant dans une bon l'endomorphisme symétrique a de $\mathbb{R}^q$ dont A est la matrice représentative dans la base canonique, on voit donc que le rang r de a est $\leq k$ ce qui montre que $r = k$ et achève la démonstration du 4).

5) Soit $(\lambda_1, \ldots, \lambda_q)$ une énumération des valeurs propres de la matrice symétrique réelle. Alors, par définition du polynôme caractéristique de A on a

$$(A + XI_q) = (X + \lambda_1) \ldots (X + \lambda_q) = X^q + c_1 X^{q-1} + \cdots + c_q \tag{9.24}$$

Si A est positive alors $\lambda_j \geq 0$ pour tout j et les c_j sont trivialement positifs ou nuls. Inversement, si une valeur propre était strictement négative (disons $\lambda_1 = -a$ avec $a > 0$) alors en faisant $X = a$ on obtient la contradiction

$$0 = a^q + c_1 a^{q-1} + \cdots + c_q \geq a^q.$$

La réciproque est donc montrée.

6) $(c) \Rightarrow (b)$ est évident. $(b) \Rightarrow (a)$ vient du fait que A est positive par le 5). De plus $c_q = \det A > 0$ et donc A est définie positive, par la Proposition 9.3, partie 3. Finalement $(a) \Rightarrow (c)$ vient de 9.24 qui montre que $c_j \geq \lambda_1 \ldots \lambda_j > 0$.

Remarque. Les nombres $c_j = c_j(A)$ apparaissant au 5) du théorème précédent sont parfois appelés les *invariants* de la matrice symétrique A. En effet d'après 9.24 ils ne dépendent que des valeurs propres de A, donc de l'endomorphisme symétrique associé et de sa représentation particulière par une base. C'est familier pour deux d'entre eux : $c_1 = \lambda_1 + \cdots + \lambda_q = \operatorname{trace} A$ et $c_q = \lambda_1 \cdots \lambda_q = \det A$. Pour les autres, ce sont les fonctions symétriques élémentaires des $(\lambda_1, \ldots, \lambda_q)$, à savoir

$$c_j = \sum_{T\ ;|T|=j} \prod_{i \in T} \lambda_i.$$

où la somme est prise pour tous les sous ensembles T de $\{1, \ldots, q\}$ de taille $|T| = j$. Par exemple si $q = 4$ alors

$$c_3 = \lambda_1\lambda_2\lambda_3 + \lambda_1\lambda_2\lambda_4 + \lambda_1\lambda_3\lambda_4 + \lambda_2\lambda_3\lambda_4.$$

On peut facilement montrer à partir de la définition de c_j que si $A_T = (a_{ij})_{i,j \in T}$ alors

$$c_j = \sum_{T\ ;|T|=j} \det A_T. \tag{9.25}$$

Attention toutefois, il ne faut pas croire que $\det A_T = \prod_{i \in T} \lambda_i$.

Finalement, si on normalise les $c_j = c_j(A)$ en les divisant par le nombre de sous ensembles $T \subset \{1, \ldots, q\}$ de taille j on obtient les nombres $p_j = \frac{j!(q-j)!}{q!} c_j$ qui satisfont aux *inégalités de Maclaurin* : pour $k = 1, \ldots, q-1$ on a

$$p_k^2 \geq p_{k-1} p_{k+1}$$

ou $k(q-k)c_k^2 \geq (k+1)(q-k+1)c_{k-1}c_{k+1}$. La démonstration se fait en considérant le polynôme de deux variables

$$Q(x,y) = (x+\lambda_1 y)\dots(x+\lambda_q y) = \sum_{j=0}^{q} \frac{q!}{j!(q-j)!} p_j x^j y^{q-j}$$

et en observant que

$$\frac{\partial^{q-2} Q}{\partial x^{k-1}\partial y^{q-k-1}}(x,y) = \frac{q!}{2}(p_{k+1}x^2 + 2p_k xy + p_{k-1}y^2).$$

On applique alors le principe suivant : soit le polynôme réel $X \mapsto P(X)$ de degré n à une variable ayant toutes ses racines réelles distinctes et non nulles, soit $Q(x,y) = y^n P(x/y)$ soit $Q_1(x,y) = \frac{\partial Q}{\partial x}(x,y)$, $Q_2(x,y) = \frac{\partial Q}{\partial y}(x,y)$ et soit $P_1(X) = Q_1(X,1)$ et $P_2(X) = Q_2(X,1)$. Alors d'après le théorème de Rolle P_1 et P_2 ont toutes leurs racines réelles distinctes et non nulles. En itérant ce principe on voit que le trinôme du second degré $p_{k+1}X^2 + 2p_k X + p_{k-1}$ a ses racines réelles et distinctes. Il a donc un discriminant positif ce qui entraîne le résultat désiré si tous les λ_j sont distincts non nuls. Le cas où certains sont nuls ou non distincts se traite par un passage à la limite. (Attention la réciproque est fausse : si un polynôme à coefficients réels satisfait les inégalités de Maclaurin, cela n'entraîne pas que ses racines soient réelles : ainsi $(X+2)(X^2+2X+1+\epsilon)$ avec $0 < \epsilon$ assez petit satisfait $p_2^2 > p_1 p_3$ et $p_1^2 > p_2 p_0$).

Corollaire 9.5. Soit la matrice symétrique réelle d'ordre 2 $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$. Alors

1. A est positive si et seulement si $\operatorname{trace} A \geq 0$ et $\det A \geq 0$.
2. A est définie-positive si et seulement si $\operatorname{trace} A > 0$ et $\det A > 0$.

Démonstration : Pour la première partie en appliquant le Théorème 9.1 partie 5, puisqu'alors $c_1 = \operatorname{trace} A$ et $c_2 = \det A$. Pour la seconde partie, appliquer la partie 6.

Un moyen populaire de fabriquer des matrices positives ou définies-positives est de considérer la *matrice de Gram* d'une suite finie de vecteurs d'un espace préhilbertien réel :

Proposition 9.6. Soit $(v_1, \dots, v_p)$ une suite de vecteurs de l'espace préhilbertien réel E. Alors la matrice $A = (\langle v_i, v_j \rangle)_{1 \leq i,j \leq p}$ est symétrique positive. De plus, A est définie positive si et seulement si les $(v_1, \dots, v_p)$ sont indépendants.

Démonstration : Si $X = [x_1, \dots, x_p]^T$ alors

$$X^T A X = \sum_{i=1}^{p} \sum_{j=1}^{p} x_i x_j \langle v_i, v_j \rangle = \| \sum_{i=1}^{p} x_i v_i \|^2 \geq 0.$$

De plus, donc, $X^T A X = 0$ si et seulement si $\sum_{i=1}^{p} x_i v_i = 0$. Si les v_i sont indépendants ceci ne peut arriver que si $x_1 = \dots = x_p = 0$, et donc A est alors définie-positive. La réciproque est du même genre.

Les matrices de Gram ont de nombreuses applications, comme on le verra au chapitre 5. Elles fournissent déjà un moyen de trouver les projections orthogonales d'une bon :

Proposition 9.7. Soit E un espace euclidien de dimension q, soit $(v_1, \ldots, v_q)$ une suite de vecteurs de E, soit F le sous espace vectoriel de E qu'ils engendrent et soit $p \in L(E)$ la projection orthogonale de E sur F. Alors il existe une bon $e = (e_1, \ldots, e_q)$ de E telle que $p(e_i) = v_i$ pour $i = 1, \ldots, q$ si et seulement si la matrice de Gram $A = (\langle v_i, v_j \rangle)_{1 \leq i,j \leq}$ a ses valeurs propres dans $\{0, 1\}$. De plus, la multiplicité de 1 est la dimension de F.

Démonstration : $\Longrightarrow$. Pour tous x et y dans E on a $([x]^e)^T A [y]^e = \langle p(x), p(y) \rangle$. C'est dire que $A = [p^* p]]_e^e$. Donc A est la matrice représentative dans e de l'endomorphisme $p^* p$ de E. Soit alors $f = (f_1, \ldots, f_q)$ une bon de E telle que $(f_1, \ldots, f_k)$ soit une base de F, avec $k = \dim F$. Alors
$$[p]_f^f = [p^*]_f^f = \operatorname{diag}(I_k, 0) = [p^* p]_f^f.$$
Donc les valeurs propres de A sont 0 et 1 et k est la multiplicité de 1.

$\Longleftarrow$. D'après le théorème spectral, il existe une matrice orthogonale U d'ordre q telle que si $Q = \operatorname{diag}(I_k, 0)$ alors $A = U^T\, Q\, U$, où k est la multiplicité de 1. Si $X = (x_1, \ldots, x_q)^T$ et $Y = (y_1, \ldots, y_q)^T$ sont dans $\mathbb{R}^q$ alors

$$I = \langle \sum_{i=1}^{q} x_i v_i, \sum_{j=1}^{q} y_j v_j \rangle = X^T A Y = X^T U^T\ Q\ U Y = (X')^T\ Q\ Y'$$

avec $X' = UX = (x'_1, \ldots, x'_q)^T$ et $Y' = UY = (y'_1, \ldots, y'_q)^T$. Définissons alors $e' = (e'_1, \ldots, e'_q) = (v_1, \ldots, v_q) U^T$. Notons que par définition tous les e'_i sont dans F et que $(e'_1, \ldots, e'_q)$ engendre F. Alors I s'écrit

$$I = \langle \sum_{i=1}^{q} x'_i e_i, \sum_{j=1}^{q} y'_j e'_j \rangle = x'_1 y'_1 + \cdots + x'_k y'_k$$

Puisque c'est vrai pour tous X' et Y' ceci montre que la matrice de Gram de e' est Q. Donc $(e'_1, \ldots, e'_k)$ forme une bon de F et $e'_j = 0$ si $j > k$. Complétons $(e'_1, \ldots, e'_k)$ en une bon e'' de E et définissons enfin $e = (e_1, \ldots, e_q) = (e''_1, \ldots, e''_k, e''_{k+1}, \ldots, e''_q) U$. C'est une bon comme e'' car U est orthogonale. On a alors

$$\begin{aligned}
(v_1, \ldots, v_q) &= (e'_1, \ldots, e'_q) U = (e''_1, \ldots, e''_k, 0, \ldots, 0) U \\
&= (e''_1, \ldots, e''_k, e''_{k+1}, \ldots, e''_q) Q U = (e''_1, \ldots, e''_k, e''_{k+1}, \ldots, e''_q) U U^T Q U \\
&= (e_1, \ldots, e_q) U^T Q U = (e_1, \ldots, e_q) A.
\end{aligned}$$

Définissons alors l'endomorphisme p de E par $p(e_i) = v_i$. L'égalité précédente entraîne que $[p]_e^e = A$, ce qui montre que p est symétrique comme A, et est une projection orthogonale puisque ses valeurs propres sont 0 et 1 comme celles de A. Son image est tout F, par définition de F comme espace engendré par les (v_i). La proposition est montrée.

Ombres d'un cube. Si C est un cube d'arête unité dans l'espace euclidien de dimension 3, on l'écrit à l'aide d'une bon $e = (e_1, e_2, e_3)$ par

$$C = \{x \in E; 0 \leq x_1, x_2, x_3 \leq 1\}.$$

Si on projette orthogonalement C sur un plan, on obtient un hexagone symétrique par rapport à un point ou un parallélogramme. Inversement, soit un hexagone symétrique par rapport à un point, ou un parallélogramme. Quand peut on dire que c'est l'ombre d'un cube unité ? Notons par v_1, v_2, v_3 les vecteurs formés par trois cotés consécutifs de l'hexagone orientés arbitrairement. Si on a un parallélogramme on prend par exemple $v_3 = 0$. La proposition 9.7 entraîne que l'hexagone est l'ombre d'un cube unité si et seulement si la matrice de Gram

$$A = \begin{bmatrix} \|v_1\|^2 & \langle v_1, v_2\rangle & \langle v_1, v_3\rangle \\ \langle v_2, v_1\rangle & \|v_2\|^2 & \langle v_2, v_3\rangle \\ \langle v_3, v_1\rangle & \langle v_3, v_2\rangle & \|v_3\|^2 \end{bmatrix} = \begin{bmatrix} a_1 & b_3 & b_2 \\ b_3 & a_2 & b_1 \\ b_2 & b_1 & a_3 \end{bmatrix}$$

a pour valeurs propres $0, 1, 1$ ou que son polynôme caractéristique est

$$P_A(X) = X(X-1)^2 = X^3 - 2X^2 + X.$$

Comme d'après la proposition 9.6 le déterminant de A est toujours nul puisque (v_1, v_2, v_3) sont coplanaires, une condition nécessaire et suffisante est donc le couple d'égalités

$$a_1 + a_2 + a_3 = 2, \quad (a_1a_2 - b_3^2) + (a_2a_3 - b_1^2) + (a_3a_1 - b_2^2) = 1.$$

Si on cherche l'ombre d'un cube pas nécessairement unité, par homénéité la condition sur la matrice de Gram devient

$$\frac{(a_1a_2 - b_3^2) + (a_2a_3 - b_1^2) + (a_3a_1 - b_2^2)}{(a_1 + a_2 + a_3)^2} = \frac{1}{4}.$$

Voici enfin quelques considérations géométriques sur la place des endomorphismes positifs ou définis-positifs dans l'espace des endomorphismes symétriques :

Proposition 9.8. Soit E un espace euclidien, V l'espace des endomorphismes symétriques sur E, V_+ l'ensemble des endomorphismes symétriques définis-positifs de E et $\overline{V_+}$ l'ensemble des endomorphismes symétriques positifs de E. Alors

- V_+ est un cône convexe ouvert.
- $\overline{V_+}$ est un cône convexe fermé qui est la fermeture de V_+.

Démonstration : On en donne trois démonstrations, toutes instructives.

Première méthode. Cette méthode s'appuie seulement sur la définition des endomorphismes positifs et définis positifs. Le fait que V_+ soit un cône convexe vient du fait que si a et b sont dans V_+ alors pour tout $x \in E \setminus \{0\}$ on a $\langle a(x), x\rangle > 0$, $\langle b(x), x\rangle > 0$ et donc $\langle (a+b)(x), x\rangle > 0$, ce qui donne $a + b \in V_+$. De plus si $\lambda > 0$ et $a \in V_+$ alors $\langle a(x), x\rangle > 0$

implique $\langle \lambda a(x), x\rangle > 0$ et donc $\lambda a \in V_+$. La démonstration du fait que $\overline{V_+}$ est un cône convexe est semblable. Montrer que V_+ est ouvert est plus créatif :

Observons que l'application $(a, x) \mapsto \langle a(x), x\rangle$ de $V \times E$ dans $\mathbb{R}$ est continue. En effet pour x fixé elle est linéaire en a, pour a fixé elle est quadratique en x. Elle est donc polynomiale, donc continue d'après le cours d'analyse.

Montrons que le complémentaire de V_+ par rapport à V est un fermé. Pour cela on considère une suite $(a_n)_{n\geq 0}$ d'éléments de $V \setminus V_+$ qui converge vers un élément a de V et on a à montrer que $a \notin V_+$. Dire que a_n n'est pas dans V_+ est dire qu'il existe un vecteur x non nul tel que $\langle a_n(x), x\rangle \leq 0$. En fait, ce vecteur x dépend de n, et on le note x_n. De plus, puisque $x_n \neq 0$, sans perte de généralité on peut le prendre dans la sphère unité S de E. Cette sphère S est bornée par définition, et est fermée, car c'est l'image inverse du fermé $\{1\}$ de $\mathbb{R}$ par l'application continue $x \mapsto \|x\|^2$ de E dans $\mathbb{R}$ (continue, parce qu'une norme est toujours continue, ou plus simplement ici parce que $x \mapsto \|x\|^2$ est polynomiale). Donc S est compacte, et d'après un important théorème du cours d'analyse on peut trouver une suite strictement croissante d'entiers $(n_k)_{k\geq 0}$ et un $x \in S$ tels que la suite extraite $(x_{n_k})_{k\geq 0}$ converge vers x. Comme

$$\langle a_{n_k}(x_{n_k}), x_{n_k}\rangle \leq 0,$$

et que $(a, x) \mapsto \langle a(x), x\rangle$ est continue on en déduit que $\langle a(x), x\rangle \leq 0$ et donc V_+ est ouvert.

Montrons que $\overline{V_+}$ est la fermeture de V_+ : si a est symétrique positif, ses valeurs propres sont positives ou nulles, et celles de $a_n = a + \frac{1}{n}\mathrm{id}_E$ sont supérieures ou égales à $1/n$. Donc a_n est dans V_+ et converge vers a. Donc $\overline{V_+}$ est contenu dans la fermeture de V_+. Inversement si a_n est une suite convergente de V_+, soit a sa limite. Alors pour tout $x \in E \setminus \{0\}$ on a $\langle a_n(x), x\rangle > 0$ et donc $\langle a(x), x\rangle \geq 0$. Donc a est dans $\overline{V_+}$, qui contient donc la fermeture de V_+. La démonstration est complète.

Seconde méthode. Elle s'appuie sur les parties 3) et 4) du Théorème 9.1. Les matrices positives A sont caractérisées dans l'espace V des matrices symétriques par le fait que $\det A_T \geq 0$ pour tout $T \subset \{1, \ldots, q\}$. Soit $F_T = \{A \in V ; \det A_T \geq 0\}$. Comme $A \mapsto \det A_T$ est polynomiale, elle est continue et donc F_T est fermé dans V. Donc $\overline{V}_+ = \cap_T F_T$ est fermé, comme intersection de fermés. Les matrices définies-positives A sont caractérisées dans l'espace V des matrices symétriques par le fait que $\det A_k > 0$ pour tout $k \in \{1, \ldots, q\}$. Soit $U_k = \{A \in V ; \det A_k > 0\}$. Alors de même U_k est ouvert dans V. Donc $V_+ = \cap_k U_k$ est ouvert comme intersection d'un nombre *fini* d'ouverts.

Troisième méthode. Elle s'appuie sur les parties 5) et 6) du Théorème 9.1. et sur le fait que d'après 9.25 les $c_j = c_j(A)$ sont des fonctions polynomiales de A donc continues. V_+ est ouvert comme intersection finie des ouverts $c_j(A) > 0$, $\overline{V_+}$ est fermé comme intersection des fermés $c_j(A) \geq 0$.

Exercice 9.1. Soit $a > 0$. Montrer que la matrice symétrique d'ordre $n + 1$

$$A = ((a + i + j)^{-1})_{0\leq i,j\leq n}$$

est définie-positive (Méthode : appliquer la Proposition 9.6 à l'espace préhilbertien des polynômes sur $\mathbb{R}$ muni du produit scalaire $\langle P, Q\rangle = \int_0^1 x^{a-1}P(x)Q(x)dx$ et à une suite $(P_0, \ldots, P_n)$ convenable.

Exercice 9.2. Soit f une fonction continue positive sur $[0,\infty[$ telle que l'intégrale $\int_0^\infty f(x)dx$ converge et soit, pour $s\geq 0$ le nombre $L(s)=\int_0^\infty e^{-sx}f(x)dx$. Soit $(s_1,\dots,s_q)$ une suite de $[0,\infty[$. Montrer que la matrice $A=(L(s_i+s_j))_{1\leq i,j\leq q}$ est symétrique positive. Méthode : si $X=(x_1,\dots,x_q)^T\in\mathbb{R}^q$ exprimer X^TAX comme une intégrale utilisant la fonction $x\mapsto (\sum_{j=1}^q x_je^{-s_jx})^2$.

Exercice 9.3. Soit $\theta\in]0,\pi[$. On considère les deux matrices d'ordre n :

$$A=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0\\ 1 & 0 & 1 & \cdots & 0 & 0\\ 0 & 1 & 0 & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & 0 & 1\\ 0 & 0 & 0 & \cdots & 1 & 0\end{bmatrix}, B=\begin{bmatrix} 2\cos\theta & 1 & 0 & \cdots & 0 & 0\\ 1 & 2\cos\theta & 1 & \cdots & 0 & 0\\ 0 & 1 & 2\cos\theta & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & 2\cos\theta & 1\\ 0 & 0 & 0 & \cdots & 1 & 2\cos\theta\end{bmatrix}$$

Montrer par récurrence que $\det B=\frac{\sin(n+1)\theta}{\sin\theta}$ (Méthode : développer par rapport à la dernière ligne). Montrer que $\det B$ s'annule pour n valeurs distinctes de θ de $]0,\pi[$, et les déterminer. Si P_A est le polynôme caractéristique de A, calculer $P_A(-2\cos\theta)$ et déduire de ce qui précède les valeurs propres de A. Montrer que les matrices $2I_n+A$ et $2I_n-A$ sont définies positives.

Exercice 9.4. Soit $\theta\in]0,\pi[$. On considère les deux matrices d'ordre $n\geq 1$:

$$A_n=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0\\ 1 & 0 & 1 & \cdots & 0 & 0\\ 0 & 1 & 0 & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & 0 & 1\\ 0 & 0 & 0 & \cdots & 1 & 1\end{bmatrix}, B_n=\begin{bmatrix} 2\cos\theta & 1 & 0 & \cdots & 0 & 0\\ 1 & 2\cos\theta & 1 & \cdots & 0 & 0\\ 0 & 1 & 2\cos\theta & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & 2\cos\theta & 1\\ 0 & 0 & 0 & \cdots & 1 & 1+2\cos\theta\end{bmatrix}$$

(en convenant que pour $n=1$ on a $A_1=[1]$ et $B_1=[1+2\cos\theta]$). Montrer que pour $n\geq 2$ on a $\det B_{n+1}=2\cos\theta\det B_n-\det B_{n-1}$ (Méthode : développer par rapport à la première ligne : la récurrence n'est pas nécessaire). Montrer par récurrence que

$$\det B_n=\frac{\sin(n+1)\theta}{\sin\theta}+\frac{\sin n\theta}{\sin\theta}=\frac{\sin(n+\frac{1}{2})\theta}{\sin\frac{\theta}{2}}$$

(Voir Szegö "Orthogonal polynomials" page 29). Montrer que $\det B_n$ s'annule pour n valeurs distinctes de θ de $]0,\pi[$, et les déterminer. Si P_{A_n} est le polynôme caractéristique de A_n, calculer $P_{A_n}(-2\cos\theta)$ et déduire de ce qui précède les valeurs propres de A_n. La matrice A_n est elle diagonalisable ? Montrer que $2I_n-A_n$ est définie positive.

Exercice 9.5. Ombres d'un parallélépipède : Montrer que tout hexagone symétrique par rapport à un point est l'ombre d'un parallélépipède. Méthode : considérer les vecteurs w_1,w_2,w_3 engendrés par trois cotés consécutifs et montrer qu'on peut trouver trois nombres $\lambda_1,\lambda_2,\lambda_3$ tels que si $v_i=\lambda_iw_i$ alors la matrice de Gram de (v_1,v_2,v_3) a pour valeurs propres $(0,1,1)$.

Exercice 9.6. Matrice de Hua. Si n est un entier ≥ 1 soit J_n la matrice (n,n) dont tous les coefficients sont 1. Soit $\mathbf{1}_n = (1, 1, \ldots, 1)$. Soit $p \geq 0$. Montrer que la matrice symétrique d'ordre $n+1$ définie par blocs

$$H = \left[\begin{array}{cc} \frac{n}{n+p} & \mathbf{1}_n \\ \mathbf{1}_n^T & pI_n + J_n \end{array} \right]$$

est positive. Première méthode : interpréter H comme la matrice représentative $[h]_e^e$ d'un endomorphisme symétrique h de l'espace euclidien canonique $E = \mathbb{R}^{n+1}$ dans sa base canonique $e = (e_0, e_1, \ldots, e_n)$ et calculer $[h]_f^f$ si f est une bon de E telle que $f_0 = e_0$ et $f_1 = \frac{1}{\sqrt{n}}(e_1 + \cdots + e_n)$. On trouve

$$[h]_f^f = \left[\begin{array}{ccc} \frac{n}{n+p} & \sqrt{n} & 0 \\ \sqrt{n} & n+p & 0 \\ 0 & 0 & pI_{n-1} \end{array} \right],$$

ce qui permet de calculer par exemple les valeurs propres de h. Plus généralement si $a \in \mathbb{R}$ donner par cette méthode les valeurs propres de $A + pI_{n+1}$ avec

$$A = \left[\begin{array}{cc} 1+a & \mathbf{1}_n \\ \mathbf{1}_n^T & J_n \end{array} \right]$$

et dire pour quels (a, p) cette matrice est positive ou définie-positive. Deuxième méthode : appliquer la formule 9.21 à $A = n/(n+p)I_1$ et $C = pI_n + J_n$ et montrer que C est définie positive. Troisième méthode : appliquer la formule 9.21 à $C = n/(n+p)I_1$ et $A = pI_n + J_n$. Pour ces deux dernières méthodes, il faut observer que $J_n^2 = nJ_n$ et que $(aI_n + bJ_n)^{-1} = a_1 I_n + b_1 J_n$ pour des a_1 et b_1 convenables si l'inverse existe.

Exercice 9.7. Soit $D = \mathrm{diag}(\mu_1, \ldots, \mu_q)$ et soit J_q la matrice (q,q) dont tous les coefficients sont égaux à 1. Montrer que $D - J_q$ est positive si et seulement si $\mu_j \geq 1$ pour tout j et si $1 - \frac{1}{\mu_1} - \cdots - \frac{1}{\mu_q} \geq 0$. Méthode : utiliser le théorème 9.4 partie 4 et le résultat de l'exercice 2.7 du chapitre 1. Donner de même une condition nécessaire et suffisante de définie positivité. Application : soit $(v_1, \ldots, v_q)$ des vecteurs de l'espace euclidien tels que $\langle v_i, v_j \rangle = -1$ si $i \neq j$. Alors

$$\sum_{j=1}^{q} \frac{1}{1 + \|v_j\|^2} \leq 1$$

(voir exercice 4.6). Méthode : introduire $\mu_j = 1 + \|v_j\|^2$. Discuter le cas d'égalité.

Exercice 9.8. Soit $a_1, \ldots, a_n$ des endomorphismes d'un espace E de dimension q tels que $a_1 + \cdots + a_n = \mathrm{id}_E$ et tels que $\sum_{i=1}^n \mathrm{rang}(a_i) = q$. Montrer que $a_i = a_i^2$ et que $a_i a_j = 0$ pour tous $i \neq j$. Méthode : notons $V = L(E)$, de dimension q^2 et soit $A_i \subset V$ l'image de l'endomorphisme de V défini par $x \mapsto a_i x$. Montrer que $\dim A_i = q \, \mathrm{rang}(a_i)$ et déduire de l'hypothèse sur les a_i que V est la somme directe des A_i. De l'égalité

$$a_i = a_1 a_i + a_2 a_i + \ldots + a_n a_i$$

qui donne la décomposition dans cette somme directe de a_i de deux manières, déduire alors le résultat. Application : THÉORÈME DE COCHRAN : Soit $a_1, \dots, a_n$ des endomorphismes symétriques d'un espace euclidien E de dimension q tels que $a_1 + \dots + a_n = \mathrm{id}_E$ et tels que $\sum_{i=1}^n \mathrm{rang}(a_i) = q$. Montrer qu'il existe une décomposition en somme directe orthogonale $E = E_1 \oplus \dots \oplus E_n$ telle que a_i soit la projection orthogonale sur E_i. Méthode : a_i est symétrique avec $a_i^2 = a_i$ et est donc la projection orthogonale sur quelque espace E_i. Utiliser $a_i a_j = 0$ pour voir que E_i et E_j sont orthogonaux et montrer $E = E_1 \oplus \dots \oplus E_n$ à l'aide de $a_1 + \dots + a_n = \mathrm{id}_E$.

Exercice 9.9. Soit a et b des endomorphismes symétriques positifs d'un espace euclidien. Montrer que trace $(ab) \geq 0$ et discuter le cas d'égalité. Méthode : choisir une bon e de diagonalisation de a et écrire explicitement trace (ab) en fonction des matrices $A = [a]_e^e = \mathrm{diag}(a_1, \dots, a_q)$ et $B = [b]_e^e = (b_{ij})_{1 \leq i,j \leq q}$. Utiliser le fait que $b_{ii} \geq 0$, dû au Th. 9.4 partie 4 appliquée à $S = \{i\}$, ainsi que le fait qu'un endomorphisme positif a ses valeurs propres positives ou nulles. Pour étudier le cas trace $(ab) = 0$ observer que $b_{ii} = 0$ entraîne b_{ij} si la matrice B est positive. Voir aussi l'exercice 10.9 pour une autre méthode.

Exercice 9.10. On considère la matrice réelle

$$M(a, b, c) = \begin{bmatrix} 1 & c & b \\ c & 1 & a \\ b & a & 1 \end{bmatrix}$$

et on la suppose positive.

1. Est il toujours vrai que $M(a, b, -c)$ est aussi positive ? Même question avec $M(a, -b, -c)$ et $M(-a, -b, -c)$ (Méthode : utiliser le théorème 9.4 partie 3).
2. On considère l'unique matrice B telle que $M(a, b, c) = B^T B$ et telle que B soit triangulaire supérieure à coefficients ≥ 0. On suppose que a, b, c sont ≥ 0. Montrer que B est à coefficients positifs si et seulement si $a \geq bc$.

Exercice 9.11. Soit a un endomorphisme symétrique de l'espace euclidien E tel que $\mathrm{id}_E - a$ soit défini positif (et donc $(\mathrm{id}_E - a)^{-1}$ existe).

1. Montrer par récurrence sur n que
$$(\mathrm{id}_E - a)^{-1} = \mathrm{id}_E + a + a^2 + \dots + a^{2n-1} + a^n (\mathrm{id}_E - a)^{-1} a^n$$
et montrer que l'endomorphisme $a^n (\mathrm{id}_E - a)^{-1} a^n$ est symétrique positif (il est expéditif de se placer dans une bon de diagonalisation de a).
2. Déduire de 1 que la suite $n \mapsto \sum_{k=0}^{2n-1} \mathrm{trace}\, a^k$ est bornée supérieurement.
3. On suppose de plus que trace $a^k \geq 0$ pour tout entier $k \geq 0$. Déduire du 2 que la série $\sum_{k=0}^{\infty} \mathrm{trace}\, a^k$ converge, et que $\lim_{n\to\infty} \mathrm{trace}\, a^n = 0$. Pourquoi les valeurs propres de a sont elles dans $]-1, 1[$? Pourquoi alors a-t-on $\lim_{n\to\infty} a^n = 0$? (pensez à considérer trace $a^{2n} =$ trace $a^n (a^n)^T$). A l'aide du 1 montrer que
$$(\mathrm{id}_E - a)^{-1} = \sum_{k=0}^{\infty} a^k.$$

4. Application : LEMME DE STIELTJES. Soit A une matrice symétrique réelle d'ordre q à coefficients ≥ 0 telle que $I_q - A$ soit définie positive. Montrer à l'aide du 3 que $(I_q - A)^{-1}$ est à coefficients ≥ 0.

Exercice 9.12. Soit $(f_i)_{i=1}^q$ une base ordonnée de l'espace euclidien telle que $\|f_i\|^2 = 1$ pour tout i et telle que si $i \neq j$ on ait $\langle f_i, f_j \rangle \leq 0$. Soit $(e_i)_{i=1}^q$ la base de Schmidt engendrée, et soit

$$e_j = \sum_{i=1}^{j} c_{ij} f_i$$

Montrer que $c_{ij} \geq 0$ pour tous $1 \leq i \leq j \leq q$ (Comparer avec l'exercice 4.8). Méthode : Si $\alpha_j = \langle e_j, f_j \rangle$, qui est positif d'après la définition de la base de Schmidt et si $I_j - A_j = (\langle f_i, f_k \rangle)_{1 \leq i,k \leq j}$, observer que

$$(c_{1j}, \ldots, c_{jj})(I_j - A_j) = (0, \ldots, 0, \alpha_j)$$

Appliquer alors le lemme de Stieltjes de l'exercice 9.11 à la matrice $I_j - A_j$ pour conclure que $c_{1j} \geq 0, \ldots, c_{jj} \geq 0$. Cette application du lemme de Stieltjes est due à Wilson, Proc.A.M.S. 1971.

Exercice 9.13. Si a et b sont des endomorphismes positifs de l'espace euclidien E, montrer qu'il est faux en général que l'endomorphisme symétrique $ab + ba$ soit encore positif. Méthode : prendre $E = \mathbb{R}^2$ euclidien canonique, prendre une matrice (2,2) positive A arbitraire non diagonale and $B = \mathrm{diag}(\lambda, 1)$. Vérifier qu'alors $\det(AB + BA) < 0$ si λ est assez grand.

Exercice 9.14. (Entrelacement des valeurs propres). Soit A la matrice symétrique réelle d'ordre $q \geq 2$ écrite par blocs

$$A = \begin{bmatrix} A_{q-1} & b \\ b^T & c \end{bmatrix}$$

avec A_{q-1} d'ordre $q-1$. Soit $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_q$ et $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_{q-1}$ les valeurs propres de A et de A_{q-1}. En appliquant la formule 9.22 à $A - \lambda I_q$ montrer que

$$\lambda_1 \leq \mu_1 \leq \lambda_2 \leq \mu_2 \leq \ldots \leq \mu_{q-1} \leq \lambda_q.$$

Méthode : écrire $A_{q-1} = U^T D U$ avec $U \in \mathbb{O}(q-1)$ et $D = \mathrm{diag}(\mu_1, \mu_2, \ldots, \mu_{q-1})$ et tracer le graphe de la fonction

$$\lambda \mapsto c - \lambda - b^T (A_{q-1} - \lambda I_{q-1})^{-1} b = c - \lambda - \sum_{i=1}^{q-1} \frac{p_i}{\mu_i - \lambda}$$

avec $p_i = [(Ub)_i]^2 \geq 0$. La place des zéros de cette fraction rationnelle par rapport à ses pôles permet de répondre immédiatement quand les valeurs propres de A sont distinctes, et un peu de réflexion permet de passer au cas général.

Exercice 9.15. Soit $A = (a_{ij})_{1 \leq i,j \leq q}$ une matrice symétrique réelle d'ordre q. Si $S \subset \{1, \ldots, q\}$ avec $S \neq \emptyset$ on note $A_S = (a_{ij})_{i,j \in S}$. Montrer alors l'équivalence des deux propriétés suivantes

1. Pour tout $S \subset \{1, \dots, q\}$ avec $S \neq \emptyset$ on a $\det A_S < 0$.
2. Pour tout $S \subset \{1, \dots, q\}$ avec $S \neq \emptyset$ la matrice A_S a une valeur propre strictement négative et toutes les autres strictement positives.

Méthode : $2 \Rightarrow 1$ est facile. Pour $1 \Rightarrow 2$ procéder par récurrence sur q en utilisant l'exercice précédent 9.14 : on appliquera l'hypothèse de récurrence à A_{q-1}.

Exercice 9.16.[10] Si $S = (s_{ij})_{1 \leq i,j \leq n}$ est une matrice réelle symétrique d'ordre n on note $\lambda(S)$ sa plus petite valeur propre. Soit $-b < a < b$. Le but de l'exercice est de trouver une borne inférieure à $\lambda(S)$ sous la contrainte $a \leq s_{ij} \leq b$ pour tous $i, j = 1, \dots, n$, ce qu'on suppose dans toute la suite.

1. Soit $M = M_p$ définie à l'exercice 4.11 avec $p+q = n$. A l'aide de l'exercice 4.11 montrer que $\lambda(M_p)$ est la plus petite racine de $X^2 - naX + pq(a^2 - b^2)$.
2. Calculer explicitement le nombre $\min_{p=0,\dots,n} \lambda(M_p) = m_n(a,b)$ (on montrera $m_n(a,b) = \lambda(M_r)$ quand $n = 2r$ ou $2r+1$ avec r entier).
3. Soit $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ un vecteur propre de S pour la valeur propre $\lambda(S)$ tel que $\|x\|^2 = x_1^2 + \dots + x_n^2 = 1$. On note par p le nombre de $i = 1, \dots, n$ tels que $x_i \geq 0$, et on va montrer que $\lambda(M_p) \leq \lambda(S)$. Sans perte de généralité on suppose que $x_i \geq 0$ si $i = 1, \dots, p$ et $x_i < 0$ si $i = p+1, \dots, n$. Montrer que
$$x^T M_p x \leq x^T S x = \lambda(S).$$
Méthode : montrer que $x_i m_{ij} x_j \leq x_i s_{ij} x_j$ pour chaque $1 \leq i, j \leq n$, en notant $M_p = (m_{ij})_{1 \leq i,j \leq n}$ pour simplifier.
4. Montrer que $\lambda(M_p) \leq x^T M_p x$ (Méthode : si f est une bon de diagonalisation de M_p, si $([x]^f)^T = (c_1, \dots, c_n)$ et si $\lambda_1 \geq \dots \geq \lambda_n = \lambda(M_p)$ sont les valeurs propres de M_p alors $c_1^2 + \dots + c_n^2 = 1$ et $x^T M_p x = c_1^2 \lambda_1 + \dots + c_n^2 \lambda_n$).
5. A l'aide du 3 et du 4 montrer que $\lambda(S) \geq m_n(a,b)$. Pour quelles valeurs de S a- t- on $\lambda(S) = m_n(a,b)$?

Remarque : si $a < b$ mais avec $|a| \geq b$ (et donc $a < 0$) on montre de la même manière que $\lambda(S) \geq na$ si $a \leq s_{ij} \leq b$ pour tous $i, j = 1, \dots, n$. Cela vient du fait qu'au contraire de 2 on a $m_n(a,b) = \lambda(M_n)$. Ceci permet de donner aussi pour tout intervalle $[a,b] \subset \mathbb{R}$ une borne supérieure à la plus grande valeur propre $\mu(S) = -\lambda(-S)$, à savoir $-m_n(-b,-a)$

Exercice 9.16. Soit a un endomorphisme positif de l'espace euclidien E et soit x dans E. Montrer que $\langle x, a(x) \rangle = 0$ si et seulement si $a(x) = 0$. Méthode : prendre une bon de a et utiliser $\lambda_i x_i^2 = 0 \Leftrightarrow \lambda_i x_i = 0$.

Exercice 9.17. On considère la matrice réelle symétrique
$$S = \begin{bmatrix} a & c & b \\ c & b & a \\ b & a & c \end{bmatrix}.$$

[10]Source : X. Zhan "Extremal values of real symmetric matrices with entries in an interval" (2006) : http ://www.siam.org/journals/simax/2è-3/62781.html

Trouver ses trois valeurs propres. Méthode : montrer que la trace de S est une valeur propre de S et en déduire que les deux autres sont opposées. Si on les note $\pm r$ montrer que

$$r^2 = -\det S/\text{trace}\, S = a^2 + b^2 + c^2 - ab - bc - cd.$$

Exercice 9.18.[11] On munit $E = \mathbb{R}^n$ de sa structure euclidienne canonique. Soit $A = (a_{ij})_{1\leq i,j\leq n}$ une matrice symétrique à coefficients 0 ou 1. On considère la matrice colonne $\mathbf{1} = (1,\ldots,1)^T$ d'ordre n et on suppose qu'il existe d tel que $A\mathbf{1} = d\mathbf{1}$. On note a l'endomorphisme associé. 1. Montrer que d est un entier tel que $0 \leq d \leq n$. 2. Si λ est une valeur propre de A montrer que $|\lambda| \leq d$ (Méthode : si $x = (x_1,\ldots,x_n)^T$ est le vecteur propre correspondant, soit $i_0 \in \{1,\ldots,n\}$ tel que $|x_{i_0}| \geq x_i$ pour tout $i = 1,\ldots$. Soit S l'ensemble des d entiers $j \in \{1,\ldots,n\}$ tels que $a_{i_0,j} = 1$. Utiliser $\lambda x_{i_0} = \sum_{j\in S} x_j$). 3. Montrer que d est une valeur propre simple de A (Méthode : si $\lambda = d$ montrer que $0 \leq \sum_{j\in S}(|x_j| - |x_{i_0}|) \leq 0$ pour voir que l'espace propre est formé des matrices colonnes proportionnelles à $\mathbf{1}$). 4. Si $P \subset \{1,\ldots,n\}$ soit $v_P = (v_1,\ldots,v_n)^T$ avec $v_i = 1$ si $i \in P$ et $v_i = 0$ sinon. Calculer $\|v_P\|$ en fonction de la cardinalité $|P|$ de P. 5. Soit une bon de diagonalisation $f = (f_1,\ldots,f_n)$ de a telle que $f_1 = \mathbf{1}/\sqrt{n}$, soit $[a]_f^f = \text{diag}(d,\lambda_2,\ldots,\lambda_n)$, soit $[v_P]^f = (p_1,\ldots,p_n)^T$ soit $Q \subset \{1,\ldots,n\}$ et soit $[v_Q]^f = (q_1,\ldots,q_n)^T$. Montrer que $p_1 = \langle v_P, f_1\rangle = |P|/\sqrt{n}$, que $q_1 = |Q|/\sqrt{n}$ et que

$$\langle v_Q, a(v_P)\rangle - \frac{d|P||Q|}{n} = \sum_{j=2}^{n} \lambda_i p_i q_i.$$

6. Si $m = \max\{|\lambda_2|,\ldots,|\lambda_n|\}$ montrer à l'aide du 5, de l'inégalité de Schwarz et du 4 que

$$\left|\langle v_Q, a(v_P)\rangle - \frac{d|P||Q|}{n}\right| \leq m\sqrt{|P||Q|}.$$

7. Montrer que

$$nd \leq \text{trace}\, A^2 \leq d^2 + m^2(n-1).$$

Commentaires : On interprète A comme la matrice d'un graphe non orienté de sommets $\{1,\ldots,n\}$ avec une arête entre i et j si $a_{ij} = 1$. Le nombre $\langle v_Q, a(v_P)\rangle$ est le nombre d'arêtes reliant un élément de P à un élément de Q.

X Racine carrée, décomposition polaire et valeurs singulières .

Théorème 10.1. Soit a un endomorphisme symétrique positif de l'espace euclidien E. Alors il existe un unique endomorphisme symétrique positif b, appelé *racine carrée* de a, tel que $b^2 = a$.

[11] Source : N. Alon et F. R. K. Chung. (1989) 'Explicit construction of linear sized tolerant networks, *Discrete Mathematics* **72**, 15-19

Démonstration : Soit e une bon de diagonalisation de a et

$$D = [a]_e^e = \mathrm{diag}(\lambda_1 I_{m_1}, \ldots, \lambda_p I_{m_p})$$

avec $0 \leq \lambda_1 < \ldots < \lambda_d$. Il est clair que b défini par

$$[b]_e^e = \mathrm{diag}(\sqrt{\lambda_1} I_{m_1}, \ldots, \sqrt{\lambda_p} I_{m_p})$$

convient. Pour montrer l'unicité, soit b_1 symétrique et positif tel que $b_1^2 = a$. Soit μ une valeur propre de b_1 et soit F_μ l'espace propre de b_1 correspondant. Si $x \in F_\mu$ alors $b_1(x) = \mu x$ et $a(x) = b_1^2(x) = \mu^2 x$. Donc μ^2 est une valeur propre de a et si E_λ est l'espace propre associé à la valeur propre λ pour a on a $F_\mu \subset E_{\mu^2}$. Or E est égal à la somme directe de tous les espaces propres de b_1. Cela entraîne que les valeurs propres de b_1 sont exactement les $\mu_j = \sqrt{\lambda_j}$, que les espaces propres sont $F_{\sqrt{\lambda_j}} = E_{\lambda_j}$ et donc que b_1 égale b.

Théorème 10.2. Soit a un endomorphisme inversible de l'espace euclidien E. Alors il existe un couple unique (p, u) d'endomorphismes appelé *décomposition polaire* de a, tel que $a = pu$, p est symétrique défini-positif et u est orthogonal.

Démonstration : *Existence* : Puisque d'après la Proposition 9.3 aa^* est défini-positif, soit p sa racine carrée. Comme $(\det p)^2 = \det p^2 = \det(aa^*) = (\det a)^2 \neq 0$, alors p^{-1} existe. Définissons u par $u = p^{-1}a$. Alors $u^* = a^*p^{-1}$ et donc $uu^* = p^{-1}aa^*p^{-1} = p^{-1}p^2p^{-1} = \mathrm{id}_E$. Donc u est un endomorphisme orthogonal et $a = pu$.
Unicité : si $a = p_1u_1$ est une autre décomposition, alors $a^* = u_1^*p_1$ et donc $aa^* = p_1u_1u_1^*p_1 = p_1^2$. Donc $p_1 = p$ par l'unicité de la racine carrée. Et donc $u = u_1$.

Corollaire 10.3. Soit A une matrice réelle inversible d'ordre q. Alors il existe un triplet (U, D, V) tel que $A = VDU$ avec D diagonale et U et V dans $\mathbb{O}(q)$.

Démonstration : Appliquons le théorème précédent à $E = \mathbb{R}^q$ euclidien canonique, muni de sa base e canonique, et à $a \in \mathbf{GL}(E)$ défini par $[a]_e^e = A$. Soit f une bon de diagonalisation de p. Alors $a = pu$ s'écrit

$$A = [a]_e^e = [p]_e^e[u]_e^e = [\mathrm{id}_E]_f^e[p]_f^f[\mathrm{id}_E]_e^f[u]_e^e = VDU$$

avec $U = [\mathrm{id}_E]_e^f[u]_e^e$, $D = [p]_f^f$ et $V = [\mathrm{id}_E]_f$. Comme e et f sont des bon, U et V sont bien orthogonales.

Remarque : Ceci est appelé décomposition polaire par analogie avec l'écriture $z = re^{i\theta}$ des nombres complexes non nuls. Si en particulier $\det a > 0$ alors $a = pe^b$ où p est défini-positif et b antisymétrique. Il y a de nombreuses versions voisines, dans lesquelles on décompose en up plutôt qu'en pu, ou qui n'exigent pas a inversible ; p est alors positif seulement et est toujours unique, $u \in \mathbb{O}(E)$ tel que $a = pu$ existe mais n'est plus unique.

Définition 10.1. Soit a un endomorphisme de l'espace euclidien E. Les valeurs propres de la racine carrée de a^*a sont appelées les *valeurs singulières* de a (L'exercice 4.3 du chapitre 1 montre que aa^* et a^*a ont les mêmes valeurs propres.)

Proposition 10.4. Soit a un endomorphisme de l'espace euclidien E. Alors la *norme* induite de a par la norme euclidienne de E, c'est à dire le nombre

$$\|a\| = \sup_{x \in E \setminus \{0\}} \frac{\|a(x)\|}{\|x\|}$$

est égal à la plus grande valeur singulière de a. Si a est symétrique, $\|a\|$ est en particulier le maximum des $|\lambda_j|$ où les λ_j sont les valeurs propres. Si a est symétrique positif $\|a\|$ est la plus grande valeur propre.

Démonstration : Soit e une bon de diagonalisation de la racine carrée p de a^*a. Notons $[p]_e^e = \mathrm{diag}[\mu_1, \dots, \mu_q]$, avec $\mu_1 \geq \dots \geq \mu_q \geq 0$. Donc

$$[a^*a]_e^e = \mathrm{diag}[\mu_1^2, \dots, \mu_q^2].$$

Notons $S(E)$ la sphère unité de l'espace euclidien E. On a alors

$$\begin{aligned} \|a\|^2 &= \sup\{\langle a(x), a(x)\rangle;\ x \in S(E)\} \\ &= \sup\{\langle a^*a(x), x\rangle;\ x \in S(E)\} \\ &= \sup\{\sum_{k=1}^{q} \mu_k^2 x_k^2;\ x_1^2 + \dots + x_q^2 = 1\}. \end{aligned}$$

Ce sup est μ_1^2 et est atteint pour $x_1 = \pm 1$, ce qui donne le résultat. Dans le cas symétrique, les valeurs singulières sont les valeurs absolues des valeurs propres. Dans le cas positif, les valeurs singulières sont les valeurs propres.

Exercice 10.1. Dans $\mathbb{R}^2$ euclidien canonique, on considère l'endomorphisme symétrique a dont la matrice représentative dans la base canonique $e = (e_1, e_2)$ est

$$A = \begin{bmatrix} 73 & -36 \\ -36 & 52 \end{bmatrix},$$

de trace 125 et de déterminant 2500. Trouver une base orthonormale $f = (f_1, f_2)$ telle que la matrice représentative de a dans cette base soit diagonale (on précisera donc les coordonnées de f_1 et f_2 dans la base e). En déduire une matrice orthogonale U et une matrice diagonale D telles que $A = UDU^{-1}$. À l'aide des résultats précédents, montrer que A est une matrice définie positive et calculer $B = \sqrt{A}$, c'est à dire la matrice symétrique définie positive telle que $B^2 = A$.

Exercice 10.2. (Inégalité de Minkowski) Soit a et b des endomorphismes symétriques définis positifs de l'espace euclidien E de dimension q. Montrer que

$$(\det a)^{1/q} + (\det b)^{1/q} \leq (\det(a+b))^{1/q},$$

et préciser les cas d'égalité (Méthode : le démontrer d'abord pour $a = \mathrm{id}_E$ en se placant dans une bon de diagonalisation de b. Puis se ramener à ce cas en considérant $b' = a^{-1/2} b a^{-1/2}$).

Exercice 10.3. Soit a et b des endomorphismes symétriques de l'espace euclidien E. On suppose que a est défini positif et que $c = b - a$ est positif. Montrer que b est défini positif. Montrer que $a^{-1} - b^{-1}$ est positif (Méthode : le démontrer d'abord pour $a = \mathrm{id}_E$ en se placant dans une bon de diagonalisation de b. Puis se ramener à ce cas en considérant $c' = a^{-1/2}ca^{-1/2}$).

Exercice 10.4. Soit a un endomorphisme positif du plan euclidien E et soit

$$b = \frac{a + (\det a)^{1/2}\mathrm{id}_E}{\mathsf{trace}\, a + 2(\det a)^{1/2}}.$$

Montrer que b est symétrique et positif et, à l'aide de Cayley Hamilton montrer que $b^2 = a$.

Exercice 10.5. Si E est euclidien et si $L(E)$ est muni de la structure euclidienne définie par $\langle a, b\rangle = \mathsf{trace}\,(a^*b)$, montrer que la norme euclidienne sur $L(E)$ qui en découle est sous multiplicative. Méthode : écrire $\|ab\|^2 = \mathsf{trace}\,(a^*abb^*)$ et introduire une bon de diagonalisation e de l'endomorphisme symétrique positif a^*a. Montrer que si a est positif non nul et b défini positif alors $\langle a, b\rangle > 0$ (observer qu'alors $b^{1/2}ab^{1/2}$ est positif non nul et donc de trace strictement positive).

Exercice 10.6. Imiter la démonstration de l'unicité dans le Théorème 10.1 pour montrer que si b et b_1 sont des endomorphismes symétriques d'un espace euclidien tels que $\exp b = \exp b_1$, alors $b = b_1$.

Exercice 10.7. On considère la matrice d'ordre $n \geq 2$:

$$N_n = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Soit a l'endomorphisme de $\mathbb{R}^n$ euclidien canonique dont la matrice représentative est N_n dans la bon canonique. Calculer la norme de a (Méthode : utiliser la Proposition 10.4 et l'exercice 9.4).

Exercice 10.8. Soit $\overline{E}_+$ l'ensemble des endomorphismes symétriques positifs de l'espace euclidien de dimension q. On veut montrer que $x \mapsto y = x^2 + x\,\mathsf{trace}\, x$ est une bijection de $\overline{E}_+$ sur lui même. 1) Montrer que

$$x + \frac{\mathsf{trace}\, x}{2}\mathrm{id}_E = \left(y + \frac{(\mathsf{trace}\, x)^2}{4}\mathrm{id}_E\right)^{1/2}.$$

2) En prenant la trace des deux côtés de 1) montrer que la fonction h sur les réels définie par $h(t) = -t(1 + \frac{q}{2}) + \mathsf{trace}\,(y + \frac{t^2}{4}\mathrm{id}_E)^{1/2}$ s'annule en $t = \mathsf{trace}\, x$. 3) Montrer que la fonction h

est strictement croissante, que trace x est son seul zéro et en déduire que $t(y) =$ trace x est une fonction de y. 4) Déduire du 1) et du 3) le résultat annoncé.

Exercice 10.9. Soit a et b des endomorphismes symétriques positifs d'un espace euclidien. Montrer que trace $(ab) \geq 0$. Méthode : utiliser trace $(ab) =$ trace $(a^{1/2}ba^{1/2})$.

Exercice 10.10. Soient a et b des endomorphismes positifs de l'espace euclidien E. Montrer l'inégalité de Marčenko-Pastur :

$$\text{trace}\,((\mathrm{id}_E + a)^{-1} - (\mathrm{id}_E + a + b)^{-1}) \leq \mathrm{rang}(b).$$

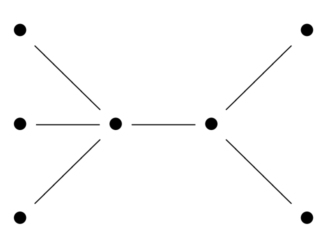
Méthode :

1. Poser $c = (\mathrm{id}_E + a)^{-1/2}$ et dire pourquoi les valeurs propres de c^2 sont dans $]0, 1]$.
2. Poser $d = cbc$ et dire pourquoi $\mathrm{rang}(b) = \mathrm{rang}(d)$.
3. Montrer que
$$(\mathrm{id}_E + a)^{-1} - (\mathrm{id}_E + a + b)^{-1} = cd(\mathrm{id}_E + d)^{-1}c.$$
4. Montrer que
$$\text{trace}\,(cd(\mathrm{id}_E + d)^{-1}c) = \text{trace}\,(c^2 d(\mathrm{id}_E + d)^{-1}) \overset{*}{\leq} \text{trace}\,(d(\mathrm{id}_E + d)^{-1}) \overset{**}{\leq} \mathrm{rang}(d).$$
Pour $*$ utiliser l'exercice 10.9 ou 9.9 appliqué à $\mathrm{id}_E - c^2$ et $d(\mathrm{id}_E + d)^{-1}$. Pour $**$ écrire d dans une bon de diagonalisation.

XI Cholesky et les arbres à racines.*

Si P est une matrice symétrique définie positive d'ordre q la décomposition de Cholesky consiste écrire $P = T^*T$ où T est une matrice triangulaire supérieure et T^* est la transposée de T. Une telle décomposition est unique. Nous avons vu ce résultat comme une conséquence du théorème d'orthonormalisation de Schmidt. Nous allons le retrouver comme cas particulier d'un théorème général qui considère un certain ordre partiel sur $\{1, \dots, q\}$. Pour cela, expliquons avant ce qu'est un arbre et un arbre à racine.

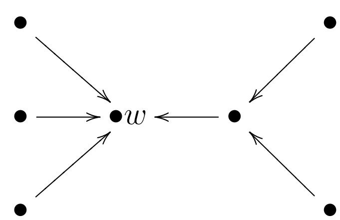
Un arbre est un ensemble fini A muni d'un ensemble $\mathcal{E}$ de parties de A à deux éléments avec la propriété suivante : pour tous x et y dans A il existe un unique entier $n \geq 0$ et une unique suite $(x_1, \dots, x_n)$ de points distincts de A tels que, en convenant $x_0 = x$ on ait $\{x_{i-1}, x_i\} \in \mathcal{E}$ pour tout $i = 1, \dots, n$ et $x_n = y$. Si $x = y$ notez que cela entraîne $n = 0$. On appelle parfois les éléments de A les sommets et les éléments de $\mathcal{E}$ les arêtes de l'arbre, bien que ce vocabulaire de théorie générale des graphes soit peu imagé dans le cas des arbres. La suite $(x_0, \dots, x_n)$ est appelée le chemin de x à y. Exemple :



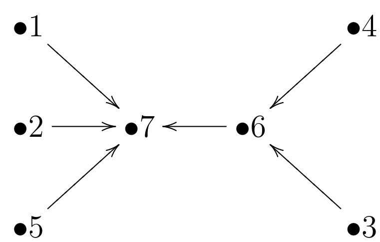
Si on sélectionne maintenant un sommet quelconque w de l'arbre A et qu'on l'appelle racine, le couple (A, w) est un arbre à racine. Le choix d'une racine équipe alors naturellement l'ensemble A des sommets de la structure d'ordre partiel $\preceq$ suivante : on écrit $x \preceq y$ si l'unique chemin de x à w contient y. Il est clair que cette relation binaire entre les éléments de A satisfait aux deux propriétés

– Si $x \preceq y$ et $y \preceq x$ alors $x = y$.
– Si $x \preceq y$ et $y \preceq z$ alors $x \preceq z$.

Notez ce n'est pas nécessairement un ordre total : il peut exister des couples tels que à la fois $x \preceq y$ et $y \preceq x$ soient faux. Voici l'exemple précédent où on a choisi une racine. On a $x \preceq y$ si on peut voyager de x à y en suivant les flèches.



Notons que si A a q sommets, il est possible de numéroter ceux ci $(1, \ldots, q)$ de sorte que $i \preceq j$ implique $i \leq j$. Cela implique évidemment $q = w$. Voici une numérotation acceptable parmi d'autres pour l'exemple précédent :



Voici le théorème concernant la décomposition de Cholesky :

Théorème 11.1. Soit (A, w) un arbre à racine tel que $A = \{1, \ldots, q\}$. Soit $\mathcal{G}$ l'ensemble des matrices $T = (t(x, y))_{x,y \in A}$ telles que si $t(x, y) \neq 0$ alors $x \preceq y$ et telles que $t(x, x) > 0$ pour tout $x \in A$. Soit $P = (p(x, y))_{x,y \in A}$ une matrice symétrique réelle d'ordre q. Alors les deux propriétés suivantes sont équivalentes

1. P est définie positive et est telle que si $p(x, y) \neq 0$ alors ou bien $x \preceq y$ ou bien $y \preceq x$.
2. Il existe une matrice T dans $\mathcal{G}$ telle que $T^*T = P$.

Dans ces conditions, la matrice T est unique. De plus $\mathcal{G}$ est un groupe pour la multiplication des matrices. Enfin si $w_1 \in A$ soit

$$A_1 = \{x \in A;\ x \preceq w_1\},$$

et soit P_1 et T_1 les restrictions respectives de P et T à $A_1 \times A_1$. Alors $P_1 = T_1^* T_1$.

Démonstration. 2 $\Rightarrow$ 1. Il est d'abord clair que $P = T^*T$ est une matrice positive, puisque pour tout vecteur colonne X on a $X^*PX = (TX)^*(TX) \geq 0$. Pour voir qu'elle est définie positive on numérote A tel que $x \preceq y$ implique $x \leq y$. Avec un tel choix la matrice $T = (t(x,y))_{x,y \in A}$ est triangulaire supérieure car $t(x,y) \neq 0 \Rightarrow x \preceq y \Rightarrow x \leq y$. Cela entraîne $\det T = \prod_{x \in A} t(x,x) > 0$, donc $\det P = (\det T)^2 > 0$ et donc la définie positivité de P.

Pour tous x et y de A, par définition du produit des matrices et par définition de la matrice transposée T^* on a $p(x,y) = \sum_{z \in A} t(z,x)t(z,y)$. Par définition de T cela entraîne

$$p(x,y) = \sum_{z \preceq x;\ z \preceq y} t(z,x)t(z,y). \tag{11.26}$$

Si il existait un couple (x,y) tel que $p(x,y) \neq 0$ et tel que $x \not\preceq y$ et $y \not\preceq x$ l'égalité 11.26 entraînerait l'existence d'un $z \in A$ tel que $t(z,x)t(z,y) \neq 0$ donc tel que $z \preceq x$ et $z \preceq y$. Le chemin de z à la racine w étant unique, il passerait par x et par y et donc x et y seraient comparables : contradiction.

1 $\Rightarrow$ 2. Nous procédons par récurrence sur la taille q de l'arbre. C'est évident pour $q = 1$. Supposons le résultat acquis pour tout arbre de taille $\leq q-1$. Numérotons A tel que $x \preceq y$ implique $x \leq y$. Alors il n'y a aucun $x \neq 1$ tel que $x \preceq 1$, c'est à dire que 1 est minimal. Notons $A' = A \setminus \{1\}$: c'est un arbre de racine w. Ecrivons alors la matrice P par blocs

$$P = \begin{bmatrix} a & b \\ b^* & c \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ b^*a^{-1} & I_{q-1} \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & c - b^*a^{-1}b \end{bmatrix} \begin{bmatrix} 1 & a^{-1}b \\ 0 & I_{q-1} \end{bmatrix}$$

où a est un nombre et où c est carrée d'ordre $q-1$. Quant à b c'est un vecteur ligne $b = (b(y))_{y \in A'}$ qui par définition de P est tel que $b(y) \neq 0$ seulement si $1 \preceq y$. Le fait que P soit définie positive entraîne que $a > 0$ et que la matrice symétrique $c - b^*a^{-1}b$ est définie positive. La remarque importante est maintenant que si l'entrée (x,y) de la matrice carrée $b^*a^{-1}b$ est non nulle, alors $b(x)b(y) \neq 0$ et donc $1 \preceq x$ et $1 \preceq y$. Comme A est un arbre à racine, cela entraine que ou bien $x \preceq y$ ou bien $y \preceq x$. Par conséquent on peut appliquer l'hypothèse de récurrence à l'arbre à racine A' et à la matrice définie positive $c - b^*a^{-1}b$. On écrit celle ci $(T')^*T'$ où $T' = (t(x,y))_{x,y \in A'}$ satisfait $t(x,x) \neq 0$. On définit enfin

$$T = \begin{bmatrix} a^{1/2} & a^{-1/2}b \\ 0 & T' \end{bmatrix}$$

qui satisfait $P = T^*T$ par un calcul immédiat ainsi que les autres propriétes demandées.

Il y a ensuite à montrer l'unicité de T. Montrons la d'abord dans le cas particulier $P = I_q$. Alors $T^*T = I_q$ implique que T est orthogonale et triangulaire supérieure (si on a numéroté pour que $x \preceq y$ implique $x \leq y$). Une telle matrice T ne peut être que diagonale avec des entrées ± 1 sur la diagonales. Cependant $t(x,x) > 0$ pour tout x entraine que T est la matrice identité I_q et l'unicité est montrée dans ce cas $P = I_q$.

Pour passer au cas général nous considérons l'ensemble $\mathcal{G}$ de l'énoncé. Nous montrons que $\mathcal{G}$ est un groupe, c'est à dire que $TS \in \mathcal{G}$ et que $T^{-1} \in \mathcal{G}$ si T et S sont dans $\mathcal{G}$. Ceci permettra de conclure, car $P = S^*S = T^*T$ implique $(ST^{-1})^*ST^{-1} = I_q$ et donc $ST^{-1} = I_q$ par la remarque précédente.

On a avec des notations évidentes

$$(TS)(x,y) = \sum_{z\in A} T(x,z)S(z,y) = \sum_{x\preceq z\preceq y} T(x,z)S(z,y)$$

et $(TS)(x,x) = T(x,x)S(x,x)$ ce qui montre que TS est dans $\mathcal{G}$. Ensuite $S = T^{-1}$ existe puisque nous avons vu que $\det T \neq 0$. Mais vérifier que S est dans $\mathcal{G}$ n'est nullement évident. En fait on le voit par récurrence en écrivant si 1 est minimal alors

$$T = \begin{bmatrix} a & b \\ 0 & T' \end{bmatrix}, \quad S = T^{-1} = \begin{bmatrix} a^{-1} & a^{-1}b \\ 0 & (T')^{-1} \end{bmatrix}.$$

Ceci permet de faire la récurrence et achève la démonstration de l'unicité de T en même temps que le fait que $\mathcal{G}$ est un groupe.

Finalement si w_1, A_1, P_1 et T_1 sont comme dans l'énoncé, écrivons par blocs correspondants à A_1 et $A \setminus A_1$ les matrices P et T :

$$P = \begin{bmatrix} P_1 & P_{12} \\ P_{21} & P_2 \end{bmatrix}, \quad T = \begin{bmatrix} T_1 & T_{12} \\ T_{21} & T_2 \end{bmatrix}.$$

On constate que si $x \notin A_1$ et $y \in A_1$ alors $t(x,y) = 0$ par définition de T et de A_1. C'est dire que $T_{21} = 0$. Donc

$$P = T^*T = \begin{bmatrix} T_1^* & 0 \\ T_{12}^* & T_2^* \end{bmatrix} \begin{bmatrix} T_1 & T_{12} \\ 0 & T_2 \end{bmatrix} = \begin{bmatrix} T_1^*T_1 & T_1^*T_{12} \\ T_{12}^*T_1 & T_{12}^*T_{12} + T_2^*T_2 \end{bmatrix}$$

ce qui démontre $P_1 = T_1^*T_1$ et achève la démonstration du théorème.

Remarques. On trouve le cas classique de la décomposition de Cholesky : si P est une matrice définie positive d'ordre q alors il existe une matrice triangulaire supérieure T à diagonale positive et telle que $P = T^*T$. Il suffit d'appliquer le théorème précédent à l'arbre à racine particulier

$$\bullet 1 - \bullet 2 - \ldots - \bullet q = w$$

On peut se demander si dans le théorème 11.1 l'arbre à racine ne pourrait pas être remplacé par un ensemble partiellement ordonné quelconque. Il n'en est rien : si cet ensemble ordonné est par exemple

$$\begin{tikzcd} \bullet 3 \arrow[r] & \bullet 4 \\ \bullet 1 \arrow[u] \arrow[r] \arrow[ur] & \bullet 2 \arrow[u] \end{tikzcd}$$

l'ensemble $\mathcal{G}$ est alors celui des matrices

$$T = \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ 0 & t_{22} & 0 & t_{24} \\ 0 & 0 & t_{33} & t_{34} \\ 0 & 0 & 0 & t_{44} \end{bmatrix}$$

telles que $t_{ii} > 0$ pour $i = 1, 2, 3, 4$. C'est toujours un groupe pour la multiplication des matrices, mais l'entrée (2,3) de T^*T est $t_{12}t_{13}$ et n'est en général pas nulle.

Exercice 11.1. Soit l'arbre à racine de sommets $A = \{1, 2, \ldots, q\}$, d'arêtes $\{i, q\}$ pour $i = 1, \ldots, q-1$ et de racine q. Soit $P = (P(x, y))_{x,y \in A}$ une matrice symétrique réelle d'ordre q telle que si $p(x, y) \neq 0$ alors ou bien $x \preceq y$ ou bien $y \preceq x$. Si P est définie positive calculer l'unique $T = (t(x, y))_{x,y \in A}$ telle que si $t(x, y) \neq 0$ alors $x \preceq y$, telle que $t(x, x) > 0$ pour tout $x \in A$ et telle que $T^*T = P$. En déduire que P est définie positive si et seulement si $\det P > 0$ et si $P(x, x)P(q, q) - P(x, q)^2 > 0$ pour tout $x = 1, \ldots, q-1$.

Chapitre 3

Espaces hermitiens.

I Produit hermitien, Schwarz.

Soit H un espace vectoriel complexe, pas nécessairement de dimension finie. Un *produit scalaire-hermitien* sur H est la donnée d'une application de $H \times H$ à valeurs dans $\mathbb{C}$, notée $(x, y \mapsto \langle x, y\rangle$ et satisfaisant aux axiomes suivants :

1. Pour $x \in H$ fixé, l'application $y \mapsto \langle x, y\rangle$ est une forme linéaire sur H.
2. Pour $y \in H$ fixé l'application $x \mapsto \langle x, y\rangle$ satisfait la propriété de *semi linéarité*, c'est à dire que $\langle x + x_1, y\rangle = \langle x, y\rangle + \langle x_1, y\rangle$ et pour $\lambda \in \mathbb{C}$, que $\langle \lambda x, y\rangle = \overline{\lambda}\langle x, y\rangle$.
3. $\overline{\langle y, x\rangle} = \langle x, y\rangle$ (*symétrie hermitienne*).
4. Pour tout $x \in H \setminus \{0\}$ on a $\langle x, x\rangle > 0$ (définie positivité).

Un espace complexe H muni d'un produit scalaire-hermitien fixé est appelé un espace *préhilbertien complexe.*[1] Si de plus H est de dimension finie, il est dit *hermitien* [2]. Le nombre $\|x\| = (\langle x, x\rangle)^{1/2}$ est appelé la norme de x. Il y a un certain arbitraire à décider que le produit scalaire hermitien $\langle x, y\rangle$ sera semi linéaire par rapport à x plutôt qu'à y, mais il faut faire un choix. C'est celui du programme officiel des classes de spéciales et c'est celui dont on a besoin pour les représentations de groupes évoquées ci dessous. Toutefois, dans un texte concernant les espaces hermitiens, il est prudent de s'assurer que l'auteur n'a pas fait le choix inverse.

Exemples : Si $H = \mathbb{C}^q$ on définit pour $x = (x_1, \ldots, x_q)$ et $y = (y_1, \ldots, y_q)$ dans H le nombre complexe

$$\langle x, y\rangle = \overline{x}_1 y_1 + \overline{x}_2 y_2 + \cdots + \overline{x}_q y_q.$$

Il est clair que H est hermitien. Avec cette structure, $\mathbb{C}^q$ s'appelle l'espace hermitien *canonique* de dimension q.

[1] D'autres disent espace unitaire.

[2] Une fonction de $H \times H$ dans $\mathbb{C}$ satisfaisant les deux premiers axiomes est dite *sesquilinéaire*, le préfixe sesqui signifiant une fois et demie. Si elle satisfait les trois premiers axiomes, elle est souvent appelée forme hermitienne dans la littérature.

Si H est l'espace des polynômes trigonométriques, c'est à dire des combinaisons linéaires complexes des fonctions $\theta \mapsto e^{in\theta}$ avec $n \in \mathbb{Z}$, définissons pour P et Q dans H le nombre

$$\langle P, Q\rangle = \frac{1}{2\pi}\int_0^{2\pi} \overline{P(\theta)}Q(\theta)d\theta, \tag{1.1}$$

qui se calcule facilement à l'aide de la formule essentielle

$$\begin{aligned} \frac{1}{2\pi}\int_0^{2\pi} \overline{e^{ik\theta}}e^{in\theta}d\theta &= 1 \text{ si } n = k \\ &= 0 \text{ si } n \neq k. \end{aligned} \tag{1.2}$$

Il est à peu près évident que les axiomes du produit scalaire-hermitien sont vérifiés, y compris le fait que $\langle P, P\rangle = 0$ ne peut arriver que si $P = 0$, car P est continue et de période 2π. Nous avons donc un espace préhilbertien complexe.

Remarques. Le mot "hermitien" vient du francais Charles Hermite. L'espace hermitien est indispensable en physique, pas seulement quantique. C'est également l'outil essentiel en mathématiques pour étudier les groupes finis et les groupes compacts. Il peut être utile au lecteur de savoir que les endomorphismes des espaces hermitiens sont aussi appelés des *opérateurs*, et pas seulement par les physiciens.

L'étude de l'espace hermitien est très analogue à celle de l'espace euclidien, mais comprend des pièges en ce qui concerne l'intuition géométrique. Pour cette raison, nous n'insistons pas autant sur l'utilité qu'il y a à travailler sans coordonnées. Il est souvent utile de considérer aussi un espace complexe de dimension finie q comme un espace réel de dimension $2q$. Appliquant cela à l'espace hermitien, on voit que l'espace hermitien H de dimension q peut être canoniquement considéré comme un espace euclidien H_R de dimension $2q$ pour la norme $x \mapsto \|x\|$. On laisse en exercice le fait de montrer que le produit scalaire de H_R est alors $\langle x, y\rangle_R = \Re\langle x, y\rangle$. Attention donc : si les normes coincident, ce n'est pas vrai des produits scalaires : *les identités de polarisation ne sont pas les mêmes* dans le cas euclidien et le cas hermitien : voir la Proposition 1.1. Si $a \in L(\mathbb{C}^q)$ on lui associe naturellement un élément $\varphi_a \in L(\mathbb{R}^{2q})$. Considérons par exemple le cas minimal $q = 1$. Si $s + it$ et $z = x + iy$ sont des complexes, soit $a(z) = (s + it)(x + iy)$ (tout endomorphisme de $\mathbb{C}$ est de cette forme). Alors $\varphi_a(x, y) = (sx - ty, tx + sy)$ est l'endomorphisme de $\mathbb{R}^2$ associé, de matrice dans la base canonique

$$\begin{bmatrix} s & -t \\ t & s \end{bmatrix}.$$

Ceci montre que l'image de $L(\mathbb{C}^q)$ dans $L(\mathbb{R}^{2q})$ de $a \mapsto \varphi_a$ n'est pas surjective. La dimension complexe de $L(\mathbb{C}^q)$ est q^2 et donc sa dimension comme espace réel est $2q^2$, alors que $\dim L(\mathbb{R}^{2q}) = 4q^2$.

Beaucoup de démonstrations sont proches du cas réel, et on se contentera de signaler les points plus délicats. Attention aux démonstrations des cas d'égalité dans les inégalités de Schwarz et du triangle ci dessous, auxquelles les jurys de concours sont très attentifs.

Proposition 1.1. (*Polarisation*) Si x et y sont dans un espace préhilbertien complexe, alors

$$\langle x, y\rangle = \frac{1}{4}\sum_{k=0}^{3} i^{-k}\|x + i^k y\|^2.$$

Démonstration. On écrit d'abord

$$\|x + i^k y\|^2 = \|x\|^2 + \|y\|^2 + i^k\langle x, y\rangle + i^{-k}\langle y, x\rangle.$$

On observe ensuite que $\sum_{k=0}^{3} i^{-k} = 0$ et que $\sum_{k=0}^{3} i^{-2k} = 0$. En rassemblant le tout on a le résultat.

Proposition 1.2. Si x et y sont dans l'espace préhilbertien complexe H alors on a

1. l'inégalité de *Schwarz* :
$$|\langle x, y\rangle| \leq \|x\|\|y\|,$$
avec égalité si et seulement si un des deux vecteurs x ou y est nul, ou bien si les vecteurs x et y sont proportionnels ;
2. l'inégalité du *triangle* :
$$\|x + y\| \leq \|x\| + \|y\|$$
avec égalité si et seulement si un des deux vecteurs x ou y est nul, ou bien si les vecteurs x et y sont tels que $y = \lambda x$ avec $\lambda > 0$.

Démonstration : Si x ou y est nul, ou si $\langle x, y\rangle = 0$, c'est évident. Si x, y et $\langle x, y\rangle$ sont non nuls, soit $\theta \in \mathbb{R}$. Alors

$$\begin{aligned} 0 &\leq \left\|\frac{x}{\|x\|} - e^{-i\theta}\frac{y}{\|y\|}\right\|^2 = \langle \frac{x}{\|x\|} - e^{-i\theta}\frac{y}{\|y\|}, \frac{x}{\|x\|} - e^{-i\theta}\frac{y}{\|y\|}\rangle \\ &= 2 - \frac{1}{\|x\|\|y\|}(\langle x, e^{-i\theta}y\rangle + \langle e^{-i\theta}y, x\rangle) = 2 - \frac{2}{\|x\|\|y\|}\Re\langle x, e^{-i\theta}y\rangle, \end{aligned}$$

la dernière égalité venant du fait que si $z = \langle x, e^{-i\theta}y\rangle$, alors $z + \overline{z} = 2\Re z$. Choisissons maintenant θ tel que $\Re z$ soit égal au module ρ de $\langle x, y\rangle$. Cela est possible, car si $\langle x, y\rangle = \rho e^{i\alpha}$ alors $z = e^{-i\theta}\rho e^{i\alpha}$ et il suffit de prendre $\theta = \alpha$. L'inégalité devient $0 \leq 2 - \frac{2\rho}{\|x\|\|y\|}$, ce qui est l'inégalité de Schwarz voulue. En cas d'égalité, on a nécessairement

$$\frac{x}{\|x\|} = e^{-i\alpha}\frac{y}{\|y\|}$$

et x et y sont bien proportionnels. Réciproquement, il est évident que l'inégalité de Schwarz devient une égalité si $y = \lambda x$ pour un nombre complexe λ.

Pour l'inégalité du triangle on écrit

$$\begin{aligned} (\|x\| + \|y\|)^2 - \|x + y\|^2 &= \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| - \langle x + y, x + y\rangle \\ &= 2\|x\|\|y\| - 2\Re\langle x, y\rangle \geq 2\|x\|\|y\| - 2|\langle x, y\rangle| \geq 0, \end{aligned}$$

la dernière inégalité étant celle de Schwarz, et la précédente venant de $\Re z \leq |z|$ si $z = \langle x, y\rangle$. En cas d'égalité dans l'inégalité du triangle, la chaîne précédente d'inégalités devient une chaîne d'égalités. Egalité dans Schwarz implique ou bien $x = 0$ ou bien l'existence d'un nombre complexe λ tel que $y = \lambda x$. L'égalité $\Re z = |z|$ implique que z est un réel ≥ 0 (pour le voir on écrit $z = a + ib$ et $a = \sqrt{a^2 + b^2}$). Puisque $z = \langle x, y\rangle = \lambda\|x\|^2$ alors $\lambda \geq 0$ et la proposition est montrée.

Exercice 1.1. 1) Si a et b sont deux nombres complexes, montrer que $|a+b|^2 \leq 2|a|^2+2|b|^2$ et que $|\overline{a}b| \leq \frac{1}{2}(|a|^2 + |b|^2)$. 2) Soit H l'ensemble des suites complexes $z = (z_n)_{n\geq 0}$ telles que la série $\sum_{n\geq 0} |z_n|^2$ converge (on note par $S(z)$ sa somme). Montrer que $\lambda z + \lambda' z' = (\lambda z_n + \lambda' z'_n)_{n\geq 0}$ munit H d'une structure d'espace vectoriel complexe (Méthode : à l'aide du 1), montrer que $S(z + z') \leq 2S(z) + 2S(z')$). 3) Si $z = (z_n)_{n\geq 0}$ et $z' = (z'_n)_{n\geq 0}$ sont dans H montrer à l'aide du 1) que la série $\sum_{n\geq 0} \overline{z_n} z'_n$ converge (on note par $\langle z, z'\rangle$ sa somme). Montrer que $\langle z, z'\rangle$ munit H d'une structure d'espace préhilbertien complexe.

Exercice 1.2. Si x et y sont dans un espace préhilbertien complexe, calculer

$$\frac{1}{2\pi}\int_0^{2\pi} e^{-i\theta}\|x + e^{i\theta} y\|^2 d\theta.$$

En déduire une autre démonstration que celle donnée par la Proposition 1.1 du fait que la connaissance de la norme donne la connaissance du produit scalaire.

II Orthogonalité, dualité et adjoints

Cette section dit à peu près les mêmes choses que les analogues euclidiens, et nous n'allons même pas faire les démonstrations.

Définition : Dans un espace préhilbertien complexe H, on dit que deux vecteurs x et y de H sont orthogonaux si $\langle x, y\rangle = 0$. Une partie U de H est dite orthogonale si tout couple de U de vecteurs distincts est orthogonal, et elle est dite orthogonale si de plus tous les vecteurs de U sont de norme 1. Une base orthonormale est abréviée en bon.

Proposition 2.1. Soit H préhilbertien complexe et $V \subset H$ telle que V soit orthogonale et ne contienne pas 0. Alors V est libre. Si $V = \{v_1, \ldots, v_k\}$ et si $x = \lambda_1 v_1 + \ldots + \lambda_k v_k$ alors $\lambda_j\|v_j\|^2 = \langle v_j, x\rangle$ et on a le théorème de *Pythagore* :

$$\|v_1 + \cdots + v_k|\|^2 = \|v_1\|^2 + \cdots + \|v_k\|^2.$$

Enfin si H est de dimension finie q, si e est une bon et si $[x]^e = [x_1, \ldots, x_q]^T$, $[y]^e = [y_1, \ldots, _q]^T$ alors $x_j = \langle x, e_j\rangle$ et

$$\langle x, y\rangle = \overline{x}_1 y_1 + \overline{x}_2 y_2 + \cdots + \overline{x}_q y_q.$$

Proposition 2.2. (Procédé d'orthonormalisation de Schmidt) Soit H un espace préhilbertien complexe et $f = (f_k)_{k=1}^q$ une suite finie de q vecteurs de H *indépendants*. Soit F_k

le sous espace vectoriel de dimension k engendré par $\{f_1, f_2, \ldots, f_k\}$. Alors il existe une suite (appelée *base de Schmidt* associée à f) *orthonormale unique* $e_f = (e_1, \ldots, e_q)$ telle que

- Pour tout k, la suite $(e_1, \ldots, e_k)$ est une base de F_k.
- Le nombre $\langle e_k, f_k\rangle$ est strictement positif.

De plus, si H est hermitien de dimension q, de bases f et e où e est orthonormale, alors e est la base de Schmidt associée à f si et seulement si $[id_H]_f^e$ est triangulaire supérieure avec diagonale formée d'éléments > 0.

Proposition 2.3. (Projection orthogonale) Soit H un espace préhilbertien complexe, F un sous espace de H de *dimension finie*, et soit $F^\perp$ l'ensemble de tous les vecteurs de E orthogonaux à tous les éléments de F. Alors

1. $F^\perp$ est un sous espace, F et $F^\perp$ sont en somme directe et $H = F \oplus F^\perp$.
2. Soit p_F la projection de H sur F parallèlement à $F^\perp$. Si $x \in H$ et $y_0 \in F$ alors il y a équivalence entre les trois propriétés suivantes : (a) $y_0 = p_F(x)$; (b) $\|x-y_0\| \leq \|x-y\|$ pour tout $y \in F$; (c) $\langle x - y_0, y\rangle = 0$ pour tout $y \in F$.

Proposition 2.4. (Dualité) Soit H un espace hermitien et soit H^* son dual. Si $y \in H$, on note par F_y l'élément de H^* défini par $F_y(x) = \langle y, x\rangle$. Alors $\varphi :\ y \mapsto F_y = \varphi(y)$ est une bijection entre H et H^* telle que $\varphi(y + y') = \varphi(y) + \varphi(y')$ et $\varphi(\lambda y) = \overline{\lambda}\varphi(y)$.

Pour la dernière proposition on introduit la notion de *matrice adjointe*, qui est l'outil correspond à la transposée quand on passe de l'euclidien à l'hermitien. Si

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

est une matrice complexe à p lignes et q colonnes, la matrice adjointe de A est $A^* = (\overline{A})^T$, parfois dite transposée-conjuguée. Elle a donc q lignes et p colonnes soit, si $b_{ij} = \overline{a_{ji}}$:

$$A^* = (b_{ij})_{1\leq i\leq q, 1\leq j\leq p}.$$

Proposition 2.5. (Adjoint) Soit H et F deux espaces hermitiens. Alors pour tout $a \in L(H, F)$ il existe un unique $a^* \in L(F, H)$ appelé *adjoint* de a tel que pour tout $x \in E$ et tout $y \in F$ on ait

$$\langle a(x), y\rangle = \langle x, a^*(y)\rangle. \tag{2.3}$$

Dans ces conditions, $a \mapsto a^*$ est une bijection entre les espaces vectoriels $L(H, F)$ et $L(F, H)$ telle que $(a + b)^* = a^* + b^*$, $\lambda a = \overline{\lambda}a^*$ et $a^{**} = a$. De plus, si e et f sont des *bon* de H et F, alors $[a^*]_f^e = ([a]_e^f)^*$. Ensuite, si $\dim H = \dim F$ et si a^{-1} existe, alors $(a^{-1})^* = (a^*)^{-1}$. Si G est hermitien, si $a \in L(H, F)$, si $b \in L(F, G)$ alors $(b \circ a)^* = a^* \circ b^*$. Enfin si $a \in L(H)$ alors $\det a^* = \overline{\det a}$ et $(\exp a)^* = \exp(a^*)$.

III Endomorphismes normaux.

Théorème 3.1. Soit H un espace hermitien de dimension q et $a \in L(H)$. Les propriétés suivantes sont équivalentes

1. $aa^* = a^*a$.
2. a est diagonalisable en base orthonormale.
3. Pour tout $x \in H$ on a $\|a(x)\|^2 = \|a^*(x)\|^2$.
4. a^* est un polynôme en a.
5. Si $\mu_1, \dots, \mu_q$ sont les q racines du polynôme caractéristique P_a alors

$$\mathrm{trace}\,(aa^*) = \sum_{j=1}^{q} |\mu_j|^2.$$

Remarques. Les endomorphismes de H qui commutent avec leur adjoint sont dits *normaux*. Le théorème ci dessus en donne donc 4 propriétés caractéristiques. Les 4) et 5) sont des curiosités moins importantes que les 2) et 3), et dont les démonstrations sont plus instructives que les énoncés.

Démonstration : 2) $\Rightarrow$ 1) est évident. 1) $\Rightarrow$ 3) vient de

$$\|a(x)\|^2 = \langle a(x), a(x)\rangle = \langle x, a^*a(x)\rangle = \langle x, aa^*(x)\rangle = \langle a^*(x), a^*(x)\rangle = \|a^*(x)\|^2.$$

3) $\Rightarrow$ 1) vient de la polarisation (Proposition 1.1) :

$$\langle x, aa^*(y)\rangle = \langle a^*(x), a^*(y)\rangle = \frac{1}{4}\sum_{k=0}^{3} i^{-k}\|a^*(x+i^k y)\|^2 \tag{3.4}$$

$$\langle x, a^*a(y)\rangle = \langle a(x), a(y)\rangle = \frac{1}{4}\sum_{k=0}^{3} i^{-k}\|a(x+i^k y)\|^2. \tag{3.5}$$

Si 3) est vrai, les deux membres de droite de (3.4) et de (3.5) sont donc égaux. Donc $\langle x, aa^*(y) - a^*a(y)\rangle = 0$ pour tous x et y de H. Donc $aa^* - a^*a = 0$ (rappel : si $b \in L(H)$ et si $\langle x, b(y)\rangle = 0$ pour tous x et y, en prenant $x = b(y)$ on voit que $\|b(y)\|^2 = 0$ pour tout y et donc $b = 0$). 1) $\Rightarrow$ 2) est plus créatif et se montre par récurrence sur q. C'est trivial pour $q = 1$. Supposons que 1) $\Rightarrow$ 2) soit vrai pour $\dim H < q$. Soit maintenant $\dim H = q$. Alors a a au moins une valeur propre, soit λ. Soit alors une bon e_1 de l'espace propre E_λ, de dimension p avec $1 \leq p \leq q$. Soit e_2 une bon quelconque de son orthogonal et soit $e = e_1 \cup e_2$. Alors

$$[a]_e^e = A = \begin{bmatrix} \lambda I_p & B \\ 0 & C \end{bmatrix},\ A^* = \begin{bmatrix} \overline{\lambda} I_p & 0 \\ B^* & C^* \end{bmatrix}.$$

Comme e est une bon, on a $A^* = [a^*]_e^e$. Comme 1) est vrai on a

$$0 = AA^* - A^*A = \begin{bmatrix} BB^* & B(C-\lambda I_p)^* \\ (C-\lambda I_p)B^* & CC^* - C^*C - B^*B \end{bmatrix}.$$

On en déduit que $BB^* = 0$ et donc que $B = 0$ (pour vérifier ce point, observer que si $(b_{i,1}, b_{i,2}, \ldots, b_{i,q-p})$ est la i ème ligne de B, avec $i = 1, \ldots p$, alors le coefficient (i,i) de BB^* est $\sum_{j=1}^{q-p} |b_{i,j}|^2 = 0$). Donc finalement $[a]_e^e = \begin{bmatrix} \lambda I_p & 0 \\ 0 & C \end{bmatrix}$ avec $CC^* - C^*C = 0$. Donc $E_\lambda^\perp$ est stable par a et la restriction c de a à $E_\lambda^\perp$ est un endomorphisme normal puisque $[c]_{e_2}^{e_2} = C$, $[c^*]_{e_2}^{e_2} = C^*$ (car e_2 est une bon) et $0 = CC^* - C^*C = [cc^* - c^*c]_{e_2}^{e_2}$. Comme $\dim E_\lambda^\perp = q - p < q$ on peut appliquer l'hypothèse de récurrence et 1) $\Rightarrow$ 2) est montré.

4) $\Rightarrow$ 1) est évident. Montrons 2) $\Rightarrow$ 4). Soit $\{\lambda_1, \ldots, \lambda_p\}$ le spectre de a. A l'aide des polynômes de Lagrange (Chap.1 section 4) on voit qu'il existe un polynôme P de degré $p-1$ tel que $P(\lambda_j) = \overline{\lambda_j}$ pour tout $j = 1, \ldots, p$. Montrons que $P(a) = a^*$. Il suffit d'utiliser le fait que H est la somme directe des espaces propres E_{λ_j}. Si donc $x = x_1 + \ldots + x_p$ avec $x_j \in E_{\lambda_j}$ alors

$$P(a)(x) = P(\lambda_1)x_1 + \cdots + P(\lambda_p)x_p = \overline{\lambda_1}x_1 + \cdots + \overline{\lambda_p}x_p = a^*(x).$$

Donc 2) $\Rightarrow$ 4) est montré.

2) $\Rightarrow$ 5) est évident. Montrons 5) $\Rightarrow$ 2). D'après le Chap. 1 section 7 il existe une base f de triangulation pour a. Appliquant le procédé de Schmidt à f, on obtient une bon e telle que $A = [a]_e^e$ soit triangulaire supérieure. Soit $D = \mathrm{diag}(\mu_1, \ldots, \mu_q)$ la diagonale de A, qui est aussi formée des racines du polynôme caractéristique P_a. Soit $N = A - D$. Comme N est triangulaire supérieure et nilpotente alors les diagonales de ND^* et DN^* sont nulles. Donc, puisque 5) est vrai

$$\mathrm{trace}\, DD^* = \mathrm{trace}\, AA^* = \mathrm{trace}\,(D+N)(D+N)^* = \mathrm{trace}\, DD^* + \mathrm{trace}\, NN^*.$$

Par conséquent $\mathrm{trace}\, NN^* = 0$ et donc $N = 0$ par un raisonnement analogue à celui fait pour 1) $\Rightarrow$ 2). Donc $A = D$, ce qui montre que a est diagonalisable dans une bon et achève la démonstration du théorème.

Définitions. Soit H un espace hermitien et $a \in L(H)$. On dit que a est
- *unitaire* (ou isométrie vectorielle) si $aa^* = a^*a = \mathrm{id}_E$;
- *hermitien* (ou autoadjoint) si $a = a^*$;
- *hermitien positif* s'il est hermitien et tel que $\langle a(x), x\rangle \geq 0$ pour tout $x \in H$.
- *hermitien défini-positif* s'il est hermitien et tel que $\langle a(x), x\rangle > 0$ pour tout $x \in H \setminus \{0\}$.
- *antihermitien* si $a^* = -a$.

On remarque que toutes ces définitions entraînent *qu'il s'agit alors d'endomorphismes normaux* et le théorème de diagonalisation 3.1 simplifie beaucoup leur étude. Commencons par celle des endomorphismes unitaires.

IV Endomorphismes unitaires, $\mathbb{SU}(2)$ et $\mathbb{SO}(3)$

Théorème 4.1 Soit H un espace hermitien et soit a un endomorphisme de H. Alors les propriétés suivantes sont équivalentes

1. Pour tous x et y de H on a $\langle a(x), a(y)\rangle = \langle x, y\rangle$ (*Conservation du produit scalaire*).
2. Pour tout $x \in H$ on a $\|a(x)\|^2 = \|x\|^2$ (*Conservation de la norme*).
3. a^{-1} existe et $a^{-1} = a^*$ (*L'adjoint égale l'inverse*).
4. a est diagonalisable dans une bon de H et les valeurs propres de a sont toutes de module 1.
5. Il existe une bon $e = (e_1, \ldots, e_q)$ de H telle que $(a(e_1), \ldots, a(e_q))$ soit aussi une bon.
6. *Pour toute* bon $e = (e_1, \ldots, e_q)$ de H alors $(a(e_1), \ldots, a(e_q))$ est aussi une bon (*Conservation des bon).*

Démonstration : C'est la copie conforme du Théorème 6.1 du chapitre 2 à l'exception du 4). Nous nous contentons donc de montrer 4) $\Leftrightarrow$ 3). Si 3) est vrai alors a est normal donc diagonalisable dans une bon e. Si $D = [a]_e^e = \mathrm{diag}(z_1, \ldots, z_q)$ et puisque 3) est vrai alors $DD^* = I_q$ et donc $z_j\overline{z_j} = 1$. La partie 4 $\Rightarrow$ 3) est semblable.

Corollaire 4.2 Soit $\mathbb{U}(H)$ l'ensemble des endomorphismes unitaires de l'espace hermitien H et soit $\mathbb{SU}(H)$ l'ensemble de ceux qui de plus sont de déterminant 1. Alors $|\det u| = 1$ si $u \in \mathbb{U}(H)$, et $\mathbb{U}(H)$ et $\mathbb{SU}(H)$ sont des sous groupes connexes de $\mathbf{GL}(H)$. Enfin, si $u \in L(H)$, alors u est dans $\mathbb{SU}(H)$ si et seulement si il existe un endomorphisme antihermitien b de trace nulle tel que $u = \exp b$.

Démonstration : $|\det u| = 1$ vient du 4) du Théorème 4.1. Que $\mathbb{U}(H)$ et $\mathbb{SU}(H)$ soient des sous groupes de $\mathbf{GL}(H)$ est standard. Pour voir qu'ils sont connexes, on montre qu'ils sont connexes par arc encore grâce au 4) du Théorème 4.1 : si $u \in \mathbb{U}(H)$ est diagonalisable dans la bon e on a

$$[u]_e^e = \mathrm{diag}(\exp(i\theta_1), \ldots, \exp(i\theta_q)), \tag{4.6}$$

Pour $0 \leq t \leq 1$, définissons $u_t \in \mathbb{U}(H)$ par

$$[u_t]_e^e = \mathrm{diag}(\exp(it\theta_1), \ldots, \exp(it\theta_q)).$$

Il est clair que $t \mapsto u_t$ est continu et définit un chemin de id_H à u. De plus, si $u \in \mathbb{SU}(H)$ alors on peut prendre $(\theta_1, \ldots, \theta_q)$ tel que $\theta_1 + \ldots + \theta_q = 0$, donc $t\theta_1 + \ldots + t\theta_q = 0$ et $u_t \in \mathbb{SU}(H)$. Donc $\mathbb{SU}(H)$ est aussi connexe. Enfin si $u \in SU(H)$ est de la forme (4.6) avec $\theta_1 + \ldots + \theta_q = 0$, définissons l'endomorphisme antihermitien b par

$$[b]_e^e = \mathrm{diag}(i\theta_1, \ldots, i\theta_q).$$

Il est clair $u = \exp b$. La réciproque est immédiate.

Définition. Une matrice carrée complexe U est dite *unitaire* si elle est inversible et si $U^{-1} = U^*$. L'ensemble des matrices unitaires d'ordre q est noté $\mathbb{U}(q)$, l'ensemble des matrices unitaires d'ordre q à déterminant 1 est noté $\mathbb{SU}(q)$.

Théorème 4.3. Soit H un espace hermitien de dimension q et soit U une matrice carrée complexe d'ordre q. Alors les propriétés suivantes sont équivalentes :

1. $U \in \mathbb{U}(q)$.
2. Il existe une bon e de H et $u \in \mathbb{U}(E)$ tels que $[u]_e^e = U$.
3. Pour toute bon e de H il existe $u \in \mathbb{U}(E)$ tel que $[u]_e^e = U$.
4. Il existe deux bon e et f de H telles que $[\mathrm{id}_H]_f^e = U$.
5. Pour toute bon e de E il existe une bon f de H telle que $[\mathrm{id}_H]_f^e = U$.
6. Pour toute bon f de H il existe une bon e de E telle que $[\mathrm{id}_H]_f^e = U$.
7. Il existe une matrice unitaire V et une matrice diagonale D dont les éléments diagonaux sont de module 1, telles que $U = VDV^*$.

Démonstration : C'est la copie du Théorème 6.3 du chapitre 2, à l'exception du 7). Pour voir 7) $\Rightarrow$ 1) on écrit $U^* = V^*D^*V^{**} = V^*D^*V$ et donc $UU^* = I_q$. Pour voir 2) $\Rightarrow$ 7) on prend $H = \mathbb{C}^q$ muni de sa base canonique e et on prend u défini par $U = [u]_e^e$.

Théorème 4.4. (Description de $\mathbb{SU}(2)$.) Les matrices de $\mathbb{SU}(2)$ sont les matrices de la forme

$$U_{a,b} = \begin{bmatrix} a & b \\ -\overline{b} & \overline{a} \end{bmatrix}$$

où a et b sont deux nombres complexes tels que $|a|^2 + |b|^2 = 1$. En particulier, si $\theta \in [0, \pi]$ est tel que $\cos\theta = \Re a$ alors il existe $V \in \mathbb{SU}(2)$ tel que

$$U_{a,b} = V \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix} V^* \tag{4.7}$$

De plus, si S est la sphère unité des quaternions, si $(\vec{i}, \vec{j}, \vec{k})$ est une bon directe de l'espace E des quaternions purs, soit $a = a_0 + ia_1$ et $b = b_0 + ib_1$ et $A : \ S \to SU(2)$ défini par

$$A(a_0\mathbf{1} + a_1\vec{i} + b_0\vec{j} + b_1\vec{k}) = U_{a,ib}.$$

Alors $x \mapsto A(x)$ est un homomorphisme bijectif entre les groupes S et $\mathbb{SU}(2)$.

Démonstration : Soit $U = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ une matrice carrée complexe d'ordre 2. Si elle est dans $\mathbb{SU}(2)$ elle est inversible, et comme son déterminant $ad - bc$ est 1 on a :

$$U^{-1} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}, \ \mathbb{U}^* = \begin{bmatrix} \overline{a} & \overline{c} \\ \overline{b} & \overline{d} \end{bmatrix},$$

et on voit que $d = \overline{a}$ et $c = -\overline{b}$. Enfin $ad - bc = 1$ entraîne $|a|^2 + |b|^2 = 1$. La réciproque est immédiate.

Pour la seconde partie, on voit que le polynôme caractéristique de $U_{a,b}$ est $X^2 - 2X\cos\theta + 1$ et donc que les valeurs propres sont $e^{\pm i\theta}$. Comme $U_{a,b}$ est unitaire, il est normal et diagonalisable en base orthonormale et V dans $\mathbb{U}(E)$ satisfaisant (4.7) existe. Mais on peut multiplier V par λ tel que $\lambda^2 = (\det V)^{-1}$ pour qu'il soit dans $\mathbb{SU}(2)$ et on a (4.7).

Pour montrer la dernière partie on utilise l'isomorphisme $x \mapsto A(x)$ défini au Chapitre 2 (8.17) entre l'algèbre des quaternions $\mathbf{H}$ et l'algèbre $\mathcal{A}$. Si $A(x) = \begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}$ il est clair que $\|x\|^2 = |a|^2 + |b|^2$ et sa restriction à S a bien les propriétés voulues.

On voit que $\mathbb{SU}(2)$ est isomorphe au groupe S des quaternions de norme 1. Comme on a vu qu'il existe un homomorphisme surjectif entre les groupes S et $\mathbb{SO}(3)$ de noyau $\pm\mathbf{1}$ on voit que $\mathbb{SO}(3)$ est isomorphe à $\mathbb{SU}(2)/\pm\mathbf{1}$. Pour étoffer ce chapitre, qui n'est jusqu'ici qu'une copie un peu pâle du précédent, donnons un résultat substantiel sur $\mathbb{SU}(2)$ et $\mathbb{SO}(3)$. Il est algébrique, mais sa démonstration utilise l'analyse :

Théorème 4.5. Soit N un sous groupe de $\mathbb{SU}(2)$ tel que pour tout $U \in N$ et tout $V \in \mathbb{SU}(2)$ alors VUV^{-1} est dans N. Alors N est égal à $\{I_2, -I_2\}$, à $\{I_2\}$ ou à $\mathbb{SU}(2)$.

Démonstration : On suppose que N contient $U \neq \pm I_2$ et on cherche à montrer qu'alors $N = \mathbb{SU}(2)$. Soit $t \mapsto V_t$ une application continue de $[0,1]$ dans $\mathbb{SU}(2)$ telle que $V_0 = I_2$ et V_1 satisfasse $V_1UV_1^{-1} \neq U$ (ce qui est possible parce que $U \neq \pm I_2$) et posons $U_t = V_tUV_t^{-1}$. Par définition de N, $U_t \in N$. Comme N est un groupe, alors $P_t = U^{-1}U_t$ est dans N pour tout $t \in [0,1]$ avec $P_0 = I_2$ et $P_1 \neq I_2$. Donc en appliquant (4.7), $\operatorname{trace} P_1 < 2$. Puisque la fonction $t \mapsto \operatorname{trace} P_t$ est continue sur $[0,1]$ son image est donc un intervalle $[c,2]$ avec $-2 \leq c < 2$. Par conséquent, pour tout θ dans $[0,\pi]$ tel que $c \leq 2\cos\theta \leq 2$ il existe un élément de N de valeurs propres $e^{\pm i\theta}$. Soit N_θ l'ensemble des éléments de $\mathbb{SU}(2)$ de la forme (4.7). La définition de N montre que si $c \leq 2\cos\theta \leq 2$ alors $N_\theta \subset N$. Notons par $\mathcal{V}$ la réunion de ces N_θ : elle contient I_2 dans son intérieur. On voit alors que N est un ouvert de $\mathbb{SU}(2)$. En effet pour tout $U \in N$ alors $U\mathcal{V} \subset N$. On peut en fait visualiser $\mathcal{V}$ ainsi : d'après le Théorème 4.4, les éléments de $\mathbb{SU}(2)$ sont de la forme

$$U_{a,b} = \begin{bmatrix} a_0 + ia_1 & b_0 + ib_1 \\ -b_0 + ib_1 & a_0 - ia_1 \end{bmatrix}$$

où a_0, a_1, b_0, b_1 sont des nombres réels tels que $a_0^2 + a_1^2 + b_0^2 + b_1^2 = 1$. La trace de $U_{a,b}$ est alors $2a_0$. L'ensemble $\mathcal{V}$ n'est autre que l'ensemble des $U_{a,b}$ de $\mathbb{SU}(2)$ tels que $2a_0 \geq c$, une sorte de calotte sphérique sur une sphère de $\mathbb{R}^4$ dont le pôle est I_2.

D'autre part,

$$\mathbb{SU}(2) \setminus N = \cup_{V \notin N} VN$$

est ouvert comme réunion d'ouverts. Or $SU(2)$ est connexe (Corollaire 4.2). Donc $N = \mathbb{SU}(2)$ et le théorème est montré.

Corollaire 4.6. Soit N_0 un sous groupe de $\mathbb{SO}(3)$ tel que pour tout $U \in N_0$ et tout $V \in \mathbb{SO}(3)$ alors VUV^{-1} est dans N_0. Alors N_0 est égal à $\{I_3\}$ ou à $\mathbb{SO}(3)$.

Démonstration : Elle utilise une technique classique en algèbre. On a vu en conséquence du Théorème 4.4 qu'il existe un homomorphisme surjectif φ du groupe $\mathbb{SU}(2)$ vers le groupe $\mathbb{SO}(3)$ tel son noyau soit $\pm I_2$, c'est à dire que $\varphi(U) = I_3$ entraîne $U = \pm I_2$. Soit N l'image inverse de N_0 par φ :

$$N = \{U \in \mathbb{SU}(2)\ ;\ \varphi(U) \in N_0\}.$$

Alors si $V \in \mathbb{SU}(2)$ et $U \in N$ on a $\varphi(VUV^{-1}) = \varphi(V)\varphi(U)(\varphi(V))^{-1}$ car φ est un homomorphisme. Donc N satisfait les hypothèses du Théorème 4.5. Que N soit égal à $\{I_2, -I_2\}$ ou à $\{I_2\}$, il est clair que son image N_0 par φ est $\{I_3\}$. Si en revanche $N = \mathbb{SU}(2)$ alors l'image est tout $\mathbb{SO}(3)$. Le corollaire est montré.

Remarques. En général, si N est un sous groupe d'un groupe G on dit que N est *normal* (ou *distingué*) si pour tout $x \in N$ et tout $y \in G$ alors yxy^{-1} est dans N. Les cas où N est G ou est réduit à l'identité sont des exemples triviaux de sous groupes distingués. On dit que G est *simple* s'il n'a pas de sous groupes normaux non triviaux.

Avec ces définitions, une manière plus compacte d'énoncer le Théorème 4.5 et son corollaire est alors de dire que le seul sous groupe normal de $\mathbb{SU}(2)$ non trivial est $\{\pm I_2\}$, et de dire que $\mathbb{SO}(3)$ est un groupe simple.

Exercice 4.1. Soit λ un nombre complexe de module 1. Trouver tous les U de $\mathbb{U}(2)$ de déterminant λ. En déduire que si $U \in \mathbb{U}(2)$ alors

1. $U = U^*$ si et seulement si $\det U = -1$.
2. Si $J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ alors $U^T JU = (\det U)J$.

V Théorème spectral hermitien et conséquences

La suite de ce chapitre est sans surprise et continue de copier le cas euclidien. Les démonstrations des 5.1 et 5.2 sont triviales parce qu'un endomorphisme hermitien est normal, donc diagonalisable dans une bon. Les théorèmes 5.3 et 5.4 sont un exercice d'imitation.

Théorème 5.1. Soit H hermitien et $a \in L(H)$ hermitien. Alors

1. Il existe une bon e telle que $[a]_e^e$ soit diagonale réelle (théorème spectral hermitien).
2. Pour tout $x \in E$ $\langle a(x), x\rangle$ est réel ≥ 0 si et seulement si les valeurs propres de a sont ≥ 0.
3. Pour tout $x \in E \setminus \{0\}$ on a $\langle a(x), x\rangle$ est réel strictement positif si et seulement si les valeurs propres de a sont > 0.

On dit qu'une matrice carrée complexe A est *hermitienne* si $A = A^*$.

Corollaire 5.2. Soit A une matrice hermitienne d'ordre q. Alors il existe des matrices diagonale D et unitaire U d'ordre q telles que $A = UDU^{-1} = PDU^*$.

Définition. Un endomorphisme *hermitien* a de l'espace hermitien H est dit *positif* s'il satisfait la propriété suivante : pour tout $x \in H$ on a $\langle a(x), x\rangle \geq 0$. Il est dit *défini-positif* s'il satisfait la propriété suivante : pour tout $x \in H \setminus \{0\}$ on a $\langle a(x), x\rangle > 0$. Une matrice hermitienne A d'ordre q est dite *positive* si elle satisfait la propriété suivante : pour tout $X \in \mathbb{C}^q$ on a $X^*AX \geq 0$. Elle est dite *définie-positive* si elle satisfait la propriété suivante : pour tout $X \in \mathbb{C}^q \setminus \{0\}$ on a $X^*AX > 0$.

Remarques. Il est clair que si a est hermitien et si e est une bon, alors $A = [a]_e^e$ est positive ou définie positive en même temps que a. Par le Théorème 5.1 parties 3) et 4), l'endomorphisme hermitien a (ou la matrice $A = [a]_e^e$) est positif si et seulement si ses valeurs propres sont toutes ≥ 0, et il est défini-positif si et seulement si ses valeurs propres sont toutes > 0. Le Théorème 9.4 du chapitre 2 reste valable. Par exemple si on considère la matrice hermitienne

$$\begin{bmatrix} x & c & b \\ \overline{c} & y & a \\ \overline{b} & \overline{a} & z \end{bmatrix}$$

avec x, y, z réels et a, b, c complexes alors elle est positive si et seulement si $x, y, z \geq 0$ et

$$xy \geq |c|^2,\ yz \geq |a|^2,\ zx \geq |b|^2,\ xyz + 2\Re(abc) - (x|a|^2 + y|b|^2 + z|c|^2) \geq 0.$$

Notons aussi que l'analogue hermitien de la Proposition 9.6 du chapitre 2 sur les matrices de Gram est vrai : $(\langle f_i, f_j\rangle)_{1\leq i,j\leq k}$ est positive, et est définie positive si et seulement si les $f_1, \ldots, f_k$ sont indépendants.

Théorème 5.3. (Racine carrée) Soit a un endomorphisme hermitien positif de l'espace hermitien H. Alors il existe un unique endomorphisme hermitien positif b, tel que $b^2 = a$.

Théorème 5.4. (Décomposition polaire) Soit a un endomorphisme inversible de l'espace hermitien H. Alors il existe un couple unique (p, u) d'endomorphismes tel que $a = pu$, p est hermitien défini-positif et u est unitaire.

Exercice 5.1. Soit $z = \rho e^{i\theta}$ avec $\rho > 0$ et θ réel. Trouver D diagonale et U unitaire carrées d'ordre 2 telles que $UDU^* = \begin{bmatrix} 0 & z \\ \overline{z} & 0 \end{bmatrix}$.

Exercice 5.2. Soit a un endomorphisme de l'espace hermitien H. Montrer que a est hermitien si et seulement si pour tout $x \in H$ le nombre $\langle a(x), x\rangle$ est réel. Méthode : si $B(x, y) = \langle a(x), y\rangle$ montrer que $B(x, y) + B(y, x)$ et $iB(x, y) - iB(y, x)$ sont égaux à leur conjugués pour voir que $B(x, y) = \overline{B(x, y)}$.

Exercice 5.3. Soit f une fonction continue positive sur $\mathbb{R}$ telle que l'intégrale $\int_{-\infty}^{\infty} f(x)dx$ converge et soit, pour $t \in \mathbb{R}$ le nombre $\varphi(t) = \int_{-\infty}^{\infty} e^{itx} f(x)dx$. Soit $(t_1, \ldots, t_q)$ une suite de $[0, \infty[$. Montrer que la matrice $A = (\varphi(t_j - t_k))_{1\leq j,k\leq q}$ est hermitienne positive. Méthode : si $X = (x_1, \ldots, x_q)^T \in \mathbb{C}^q$ exprimer X^*AX comme une intégrale utilisant la fonction $x \mapsto$

$|\sum_{j=1}^{q} x_j e^{it_j x}|^2$. Applications : en admettant que

$$e^{-|t|} = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{e^{itx}dx}{1+x^2},$$

montrer que pour $y \in]0,1[$ la matrice $(y^{|j-k|})_{1\leq j,k\leq q}$ est définie positive (Méthode : considérer $(t_1,\ldots,t_q)$ tels que $t_j = j\log y$). Montrer que pour $a>0$ la matrice

$$\left(\frac{1}{a+|j-k|}\right)_{1\leq j,k\leq q}$$

est définie positive (Méthode : considérer $\int_0^1 y^{|j-k|}y^{a-1}dy$).

Exercice 5.4. (Suites stationnaires dans un espace hermitien). Soit H un espace hermitien de dimension q et soit $(z_n)_{n\in\mathbb{Z}}$ une suite de H indexée par l'ensemble $\mathbb{Z}$ des entiers relatifs et qui engendre H, c'est à dire qu' il existe $n_1 < n_2 < \ldots < n_q$ tels que $f = (z_{n_1},\ldots,z_{n_q})$ est une base de H (non nécessairement orthonormale). Montrer qu'il existe $u \in \mathbb{U}(H)$ tel que pour tout $n\in\mathbb{Z}$ on a $z_n = u^n(z_0)$ si et seulement si pour tout $k\in\mathbb{Z}$ le nombre $\varphi(k) = \langle z_{n+k}, z_n\rangle$ ne dépend pas de $n\in\mathbb{Z}$. Montrer dans ces conditions qu'il existe des nombres $p_1,\ldots,p_r\geq 0$ et des nombres réels $\theta_1,\ldots,\theta_r$ tels que

$$\varphi(k) = \sum_{j=1}^{r} e^{ik\theta_j}p_j.$$

(Méthode pour $\Leftarrow$: montrer que $f' = (z_{n_1+1},\ldots,z_{n_q+1})$ est aussi une base en montrant qu'elle a la même matrice de Gram que f. Définir alors $u\in L(H)$ par $u(z_{n_j}) = z_{n_j+1}$ pour $j=1,\ldots,q$, montrer que le fait que f' soit une base entraîne que u^{-1} existe, et déduire de $\langle u(z_{n_j}), z_{n_k+1}\rangle = \langle z_{n_j}, u^{-1}(z_{n_k+1})\rangle$ que $u^{-1} = u^*$. Montrer que pour tout $n\in\mathbb{Z}$ on a $\langle u(z_n), z_{n_k+1}\rangle = \langle z_{n+1}, z_{n_k+1}\rangle$ et en déduire $u(z_n) = z_{n+1}$.)

Exercice 5.5. On considère une matrice circulante complexe $P(R)$ comme définie au chapitre 1 section 3. A quelle condition sur P la matrice $P(R)$ est elle hermitienne ? A l'aide de l'exemple 5.3 du chapitre 1, si $P(R)$ est hermitienne, montrer qu'elle est positive si et seulement si $P(e^{\frac{2ik\pi}{q}})\geq 0$ pour tout $k=0,\ldots,q-1$.

Exercice 5.6. En considérant le polynôme X^5-1 montrer que $\sum_{k=0}^{4} e^{\frac{2ik\pi}{5}} = 0$. En déduire $\cos\frac{2\pi}{5}$ et $\cos\frac{4\pi}{5}$ par considération de la partie réelle de cette égalité. Utiliser alors l'exercice 5.5 pour donner la condition nécessaire et suffisante sur le nombre complexe $z = u+iv$ pour que la matrice

$$\begin{bmatrix} 1 & z & 0 & \ldots & \overline{z} \\ \overline{z} & 1 & z & \ldots & 0 \\ 0 & \overline{z} & 1 & z & 0 \\ 0 & 0 & \overline{z} & 1 & z \\ z & 0 & 0 & \overline{z} & 1 \end{bmatrix}$$

soit positive. Dessiner la partie correspondante du plan complexe (c'est le pentagône régulier circonscrit au cercle $u^2+v^2 = 1/4$ dont un côté est $u=-1/2$, plus l'intérieur du pentagône).

VI Burnside et von Neumann*

Dans cette section [3]nous étudions des "algèbres d'opérateurs ". Plus précisément, nous nous donnons un espace hermitien H (insistons sur le fait qu'il est de dimension finie) et un sous espace vectoriel A de $L(H)$. On dit que A est une algèbre si de plus le produit ab est dans A pour tous a et b dans A. On dit que A est une *algèbre de von Neumann* si A est une algèbre telle que de plus l'adjoint a^* est dans A pour tout a de A et si A contient l'identité id_H.

Exemple 6.1 Si H est somme directe orthogonale de sous espaces :

$$H = H_1 \oplus \ldots \oplus H_p$$

de dimensions respectives $q_1, \ldots, q_p$ et si $a_j \in L(H_j)$ notons par $a = (a_1, \ldots, a_p)$ l'endomorphisme de H qui a $x = x_1 + \ldots + x_p$ (avec $x_j \in H_j$) fait correspondre $a(x) = a_1(x_1) + \ldots + a_p(x_p)$. Alors l'ensemble A de tous les a construits de cette manière pour la somme directe $H = H_1 \oplus \ldots \oplus H_p$ est une algèbre de von Neumann. On note un tel A par

$$A = L(H_1) \oplus \ldots \oplus L(H_p). \tag{6.8}$$

Si on prend des bon e_j pour chaque H_j alors $e = e_1 \cup \ldots \cup e_k$ est une bon de H et a est dans A si et seulement si $[a]_e^e$ est diagonale par blocs :

$$[a]_e^e = \mathrm{diag}(A_1, \ldots, A_p)$$

où la matrice carrée complexe A_j est d'ordre q_j.

Exemple 6.2 Soit a un endomorphisme normal de l'espace hermitien H (rappelons que cela signifie $aa^* = a^*a$) et soit A l'ensemble des endomorphismes de $L(H)$ de la forme $P(a, a^*)$ où P est un polynôme à coefficients complexes à deux variables. Il est clair que A est une algèbre de von Neumann commutative. Soit alors $H = E_1 \oplus \ldots \oplus E_k$ la décomposition en somme directe orthogonale par les espace propres E_j de a et soit $(\lambda_1, \ldots, \lambda_k)$ la suite des valeurs propres correspondantes de a, de multiplicités respectives $p_1 = \dim E_1, \ldots, p_r = \dim E_k$. Soit e_j une bon de E_j. Alors $e = e_1 \cup \ldots \cup e_k$ est une base de diagonalisation de a et plus généralement de l'élément typique $P(a, a^*)$ de A. Plus précisément on a

$$[P(a, a^*)]_e^e = \mathrm{diag}(P(\lambda_1, \overline{\lambda}_1)I_{p_1}, \ldots, P(\lambda_k, \overline{\lambda}_k)I_{p_k}).$$

Autrement dit les $b \in A$ sont les éléments de $L(H)$ tels qu'il existe des complexes $\mu_1, \ldots, \mu_k$ tels que $[b]_e^e = \mathrm{diag}(\mu_1 I_{p_1}, \ldots, \mu_r I_{p_k})$. Cet exemple est important : dans une algèbre de von Neumann quelconque A, il y a beaucoup d'éléments normaux puisque si $b \in A$ alors $a = bb^*$ est hermitien positif, donc diagonalisable en base orthonormale, c'est à dire normal. L'exemple montre que dès que $a \in A$ est normal alors les projections orthonormales sur les espaces propres de a font partie de A.

[3]Ma reconnaissance va à Hari Bercovici pour l'organisation de cette section.

Plus généralement on peut concevoir une algèbre de von Neumann A qui généralise les deux exemples précédents, en ce sens qu'il existe une bon e de H, des entiers k, $p_1, q_1, \ldots, p_k, q_k$ tels que $p_1q_1 + \cdots + p_kq_k = \dim H$ de sorte que les éléments b de A soient de la forme

$$[b]_e^e = \mathrm{diag}(A_1, \ldots, A_1, A_2, \ldots, A_2, \ldots, A_k, \ldots, A_k) \tag{6.9}$$

où les matrices carrées complexes A_j sont arbitraires d'ordre q_j et reproduites p_j fois dans 6.9. L'exemple 6.1 correspend à $k = 1$, l'exemple 6.2 correspond à $q_1 = \ldots = q_k = 1$. Le but de cette section est de montrer la réciproque, c'est à dire que *toute algèbre de von Neumann est de cette forme*. Ce résultat non trivial s'appuie sur le théorème de Burnside (1903), le théorème 6.6 ci dessous. Réglons auparavant le cas simple des algèbres de von Neumann commutatives : nous allons voir qu'elles coïncident avec l'exemple 6.2.

Proposition 6.1. Soit H un espace hermitien et soit A une algèbre de von Neumann c'est à dire un sous espace vectoriel de $L(H)$ contenant id_H tel que $ab \in A$ si a et b sont dans A et telle que $a^* \in A$ pour tout a de A. Alors A est commutative (c'est à dire que $ab = ba$ pour tous a et b de A) si et seulement si il existe une base orthonormale e de H telle que la matrice $[a]_e^e$ soit diagonale pour tout a de A.

Dans ces conditions si la dimension de l'espace vectoriel A est k il existe k entiers $p_j > 0$ tels que $q = p_1 + \ldots + p_k$ et il est possible de numéroter e de sorte que si $a \in L(H)$ alors $a \in A$ si et seulement si il existe une suite de complexes $(\mu_1, \ldots, \mu_k)$ telle que

$$[a]_e^e = \mathrm{diag}[\mu_1 I_{p_1}, \ldots, \mu_k I_{p_k}].$$

Démonstration. La partie $\Leftarrow$ est évidente. Pour voir $\Rightarrow$ on remarque que pour tout $a \in A$ on a $aa^* = a^*a$ et donc a est donc diagonalisable dans une base orthonormale d'après le théorème 3.1 sur les endomorphismes normaux. Le fait qu'il existe une même bon universelle e qui diagonalise chaque a de A résulte du théorème 6.1 du chapitre 1 de diagonalisation simultanée.

Montrons la seconde partie. Si $T \subset \{1, \ldots, q\}$ on note e_T l'élément de $L(H)$ telle que $[e_T]_e^e = \mathrm{diag}[\epsilon_1, \ldots, \epsilon_q]$ avec $\epsilon_j = 1$ si $j \in T$ et $\epsilon_j = 0$ sinon. Si T est non vide et si $e_T \in A$ on dit que T est un ensemble idempotent. Un ensemble idempotent T est dit minimal si $S \subset T$ et S idempotent entraine $S = T$.

On remarque que deux ensembles idempotents T et S satisfont $e_T e_S = e_{T \cap S}$ et donc $T \cap S$ est ou vide ou idempotent. Cela entraîne que deux ensembles idempotents minimaux distincts sont nécessairement disjoints. Soit $\{T_1, \ldots, T_p\}$ l'ensemble des idempotents minimaux. Alors $\cup_{j=1}^p T_j = \{1, \ldots, q\}$. Sinon $T_0 = \{1, \ldots, q\} \setminus \cup_{j=1}^p T_j$ ne serait pas vide et serait un ensemble idempotent. En effet id_H est dans A, ce qui entraine que

$$e_{T_0} = \mathrm{id}_H - \sum_{j=1}^p e_{T_j} \in A.$$

L'ensemble idempotent T_0 contiendrait un ensemble idempotent minimal, ce qui contredit la définition de $\{T_1, \ldots, T_p\}$.

On a donc déjà que A contient la sous algèbre A_1 des a de la forme $\sum_{j=1}^{p}\mu_j e_{T_j}$ où $(\mu_1,\ldots,\mu_p)\in\mathbb{C}^p$ est arbitraire. Pour voir qu'en fait $A=A_1$ et que $p=k$ montrons qu'un élément quelconque $a\in A$ défini par $[a]_e^e=\mathrm{diag}(a_1,\ldots,a_q)$ est dans A_1. Supposons qu'il n'en soit pas ainsi. Alors il existerait un j et deux éléments i_1 et i_2 de T_j tels que $a_{i_1}\neq a_{i_2}$. Soit $T=\{i;a_i=a_{i_1}\}$. Soit alors le polynôme de Lagrange L tel que $L(a_{i_1})=1$ et tel que $L(b)=0$ pour toutes les valeurs b prises par les a_i quand $i\notin T$. Alors $L(a)=e_T$ et donc T est un ensemble idempotent qui contient i_1 et pas i_2 pourtant situés dans le même ensemble idempotent minimal T_j : contradiction. Donc $A=A_1$. Le fait que $k=p$ découle de $\dim A_1=p$. Pour terminer, on note par p_j la taille de T_j et on numérote la bon e pour que $T_1\cup\ldots\cup T_j=\{1,\ldots,p_1+\cdots+p_j\}$ pour tout $j=1,\ldots,k$. La proposition est montrée.

Nous attaquons la description des algèbres de von Neumann générales par plusieurs définitions. Soit $A\subset L(H)$ une algèbre de von Neumann. L'algèbre de von Neumann suivante

$$A'=\{x\in L(H);\ ax=xa\ \forall a\in A\}$$

est appelée le *commutant* de A. L'algèbre de von Neumann commutative $C(A)=A\cap A'$ est appelée le *centre* de A. Si $C(A)$ est simplement formée des multiples de l'identité on dit que c'est un *facteur*. Voici des exemples de facteur :

Proposition 6.2. Si H est un espace hermitien alors $L(H)$ est un facteur. Si H est la somme directe orthogonale de k espaces identiques à H_1 et si A est l'algèbre de von Neumann des $a=\mathrm{diag}(a_1,\ldots,a_1)$ tels que $a_1\in L(H_1)$ alors A est un facteur.

Démonstration. Il suffit donc de montrer que si une matrice carrée $A=(a_{ij})$ d'ordre commute avec toute matrice B alors A est un multiple de I_q. Notons par E_{ij} la matrice carrée d'ordre q dont le coefficient (i,j) est 1 et dont tous les autres sont nuls. Comme la matrice $AE_{ii}-E_{ii}A=0$ on voit immédiatement que ligne et colonne i de A sont nulles en dehors de (i,i). Comme c'est vrai pour tout i, A est diagonale. Pour voir enfin que $a_{ii}=a_{11}$, écrire $AB-BA=0$ pour $B=E_{11}+E_{ii}$.

Pour la seconde partie, on cherche les $x\in L(H)$ tels que $x=\mathrm{diag}(x_1,\ldots,x_1)$ satisfasse $xa=ax$ pour tout $a\in A$ et donc $x_1a_1=a_1x_1$ pour tout $a_1\in L(H_1)$. La première partie montre que x_1 est multiple de l'identité de donc que A est un facteur.

En fait un long travail va consister à montrer au théorème 6.9 que tous les facteurs sont de la forme indiquée par la proposition 6.2. Cette proposition permet aussi de trouver que le commutant A' de l'exemple 6.1 est contenu dans A et est formé des x de la forme

$$[x]_e^e=\mathrm{diag}(\lambda_1 I_{q_1},\ldots,\lambda_p I_{q_p}).$$

Pour le voir on écrit x par blocs, c'est à dire $x=(x_{ij})$ avec $x_{ij}\in L(H_i,H_j)$. L'égalité $xa=ax$ se traduit par $(x_{ij})\mathrm{diag}(a_1,\ldots,a_p)=\mathrm{diag}(a_1,\ldots,a_p)(x_{ij})$ qui se traduit par $x_{ij}a_j=a_ix_{ij}$. Pour $i\neq j$ ceci entraine $x_{ij}=0$ (prendre $a_i=0$ et a_j arbitraire). Pour $i=j$ on applique la proposition 6.5 pour voir que x_{ii} est un multiple de l'identité. En revanche le commutant A' de l'exemple 6.2 est formé des $x=\mathrm{diag}(x_1,\ldots,x_q)$ avec x_j

élément arbitraire de $L(E_j)$. Cela se voit par une méthode de blocs analogue. Finalement, la même méthode de blocs montre que le commutant du facteur A de la proposition 6.2 est formé des $x = (x_{ij})_{1\leq,j\leq p}$ tels que $x_{ij} = \lambda_{ij}\mathrm{id}_H$ avec $(\lambda_{ij})_{1\leq,j\leq p}$ matrice de nombres complexes arbitraire. On utilise pour cela la première partie de la proposition 6.2. Voici enfin l'origine du mot facteur :

Proposition 6.3. Soit H un espace hermitien et $A \subset L(H)$ une algèbre de von Neumann. Alors il existe une décomposition en somme directe orthogonale $H = E_1 \oplus \cdots \oplus E_k$ et des facteurs $A_j \subset L(E_j)$ tels que A est formé des $a = \mathrm{diag}(a_1, \ldots, a_k)$ tels que $a_j \in A_j$.

Démonstration. Le centre $C(A)$ étant une algèbre de von Neumann commutative, on lui applique la proposition 6.1 pour garantir qu'il existe une décomposition en somme directe orthogonale $H = E_1 \oplus \cdots \oplus E_k$ telle que les éléments de $C(A)$ sont de la forme $\mu_1 p_1 + \ldots + \mu_k p_k$ en notant $p_j \in L(H)$ la projection orthogonale de H sur E_j et où $(\mu_1, \ldots, \mu_k)$ est arbitraire dans $\mathbb{C}^k$. Définissons $A_j \subset L(E_j)$ comme l'ensemble des $p_j a p_j$ quand a parcourt A. Alors A_j est une algèbre de von Neumann puisque p_j commute avec les éléments de A. De plus si $i \neq j$ et a et b sont dans A alors $p_i a p_i p_j b p_j = 0$ car $p_i p_j = 0$. On le résume en disant $A_i A_j = 0$. Reste à vérifier que A_j est un facteur. Si $x \in A$ est tel que on a pour tout $a \in A$ l'égalité $p_j a p_j p_j x p_j = p_j x p_j p_j a p_j$ alors comme p_j commute avec tout A on a $a p_j x p_j = p_j x p_j a$ ce qui est dire que $p_j x p_j \in C(A)$ et donc que $p_j x p_j$ est proportionnel à l'identité id_{E_j}. L'algèbre A_j est donc un facteur et la proposition est montrée.

Proposition 6.4. Soit H un espace hermitien et soit $A \subset L(H)$ une algèbre de von Neumann. Si F est un sous espace vectoriel de H soit p_F la projection orthogonale de H sur F. Alors il y a équivalence entre les trois faits suivants :

1. F est invariant par A
2. p_F est dans A'
3. $I = p_F A p_F$ est tel que AI et IA sont contenus dans I.

En particulier si $p_F \in A$ alors F est invariant par A si et seulement si $p_F \in C(A)$ et si et seulement si $I = p_F A p_F$ est tel que AI et IA sont contenus dans $I \subset A$.

Finalement, A est un facteur si et seulement si il ne contient pas de sous espace vectoriel I non trivial tel que AI et IA sont contenus dans I et tel que I soit stable par $a \mapsto a^*$.

Démonstration. $1 \Rightarrow 2$: Pour tout $a \in A$ on a $p_F a = p_F a p_F = (p_F a^* p_F)^* = (p_F a^*)^* = a p_F$. $2 \Rightarrow 3$: puisque p_F commute avec tout élément de A on a pour a et b dans A que $a p_F b p_F = p_F a b p_F \in I$ et que $p_F a p_F b = p_F a b p_F \in I$. Enfin $3 \Rightarrow 1$ car si $v \in F$ et si $a \in A$ alors $a(v) = a p_F(v) \in I(v) \subset F$. L'application au cas particulier $p_F \in A$ est immédiate.

Montrons maintenant la dernière partie. $\Leftarrow$: Si A n'est pas un facteur, son centre $C(A)$ est non trivial. D'après la proposition 6.1 $C(A)$ contient donc une projection orthogonale p_F sur un sous espace vectoriel non trivial F de H. En considérant $I = p_F A p_F$ et la première partie de la proposition on a la contradiction voulue. La partie $\Rightarrow$ est moins simple. S'il existe un sous espace vectoriel non trivial I de A stable par $a \mapsto a^*$ et avec

$AI \subset I$ et $IA \subset I$ alors I est une algèbre qui ne contient pas id_H. Elle contient cependant des projections orthogonales p_F. Fixons F de dimension maximale telle que $p_F \in I$. Alors $\mathrm{id}_H - p_F \neq 0$ et il existe $a \in A$ tel que $(\mathrm{id}_H - p_F)a \neq 0$ (sinon $A = p_F A \subset I$ et I ne serait pas un sous espace vectoriel strict de A). Par conséquent l'endomorphisme symétrique positif $b = (\mathrm{id}_H - p_F)aa^*(\mathrm{id}_H - p_F)$ appartient à I et est non nul. Soit G un des espaces propres de b : il est dans l'orthogonal de F et pourtant $p_{F+G} = p_F + p_G$ est dans I. Cela contredit la maximalité de la dimension de F et achève la démonstration.

Voici maintenant le théorème du double commutant qui dit que $A = A''$ pour une algèbre de von Neumann. Quand on veut étendre la présente théorie aux espaces H de dimension infinie, ce résultat doit être pris comme axiome supplémentaire de la définition des algèbres de von Neumann [4] .

Théorème 6.5. Soit H un espace hermitien et soit $A \subset L(H)$ une algèbre de von Neumann. Alors $A = A''$.

Démonstration. Rappelons les définitions du commutant et du double commutant :

$$A' = \{x \in L(H);\ ax = xa\ \forall\ a \in A\} \quad A'' = \{y \in L(H);\ yx = xy\ \forall\ x \in A'\}.$$

Cela montre clairement que $A \subset A''$. Soit alors $y \in A''$. Montrons que $y \in A$. J'observe d'abord que pour tout $v \in H$ alors il existe au moins un $a \in A$ tel que $a(v) = y(v)$. En effet $F = Av$ est un sous espace invariant et donc la projection orthogonale p_F est dans A' (voir Proposition 6.4). Donc p_F commute avec y et donc $p_F y(v) = y p_F(v) = y(v)$. Donc $y(v)$ est dans $A(v)$.

Il ne reste plus qu'à montrer que le $a \in A$ que nous avons trouvé ne dépend pas de v, c'est à dire qu'en fait $y \in A$. Pour cela on utilise une astuce appelée *dilatation* de von Neumann. Si $n = \dim H$ je forme $H_1 = H^{\oplus n}$ de dimension n^2 et je forme l'algèbre de von Neumann $A_1 \subset L(H_1)$ des $a_1 = \mathrm{diag}(a, a, \ldots, a)$ avec $a \in A$. Son commutant A_1' est formé des $x = (x_{ij})_{1 \leq i,j \leq n}$ avec $x_{ij} \in A'$, et son double commutant est formé des $y_1 = \mathrm{diag}(y, y, \ldots, y)$ avec $y \in A''$. Appliquons alors la première partie de la démonstration à A_1 et à $v = e \in H_1$ quand $e = (e_1, \ldots, e_n)$ est une bon de H. Pour tout $y_1 \in A_1''$ il existe donc $a_1 \in A_1$ tel que $y_1(e) = a_1(e)$ et donc $y(e_i) = a(e_i)$ pour tout $i = 1, \ldots, n$. C'est dire $y = a$ et le résultat est montré.

Nous attaquons maintenant la seconde partie de cette section, en montrant d'abord le théorème de Burnside 6.6 et en développant les conséquences pour classer les facteurs (c'est à dire pour montrer la réciproque de la proposition 6.2) et donc avec la proposition 6.3 trouver toutes les algèbres de von Neumann.

Théorème 6.6. Soit H un espace hermitien et soit A une algèbre, c'est à dire un sous espace vectoriel de $L(H)$ tel que $ab \in A$ si a et b sont dans A. On suppose de plus que si F est un sous espace vectoriel de H tel que $A(F) \subset F$ (c'est à dire que $a(F) \subset F$ pour tout $a \in A$) alors $F = \{0\}$ ou H. Dans ces conditions on a toujours $A = L(H)$.

[4] Voir J. Dixmier, (1969) *Les algèbres d'opérateurs dans l'espace Hilbertien* Gauthier-Villars, Paris.

Remarques. Notez que les algèbres générales sont plus difficiles à classer que les algèbres de von Neumann . Dans l'énoncé du théorème de Burnside ci dessus, il n'y a rien d'hermitien mais seulement une structure sur les complexes, alors que la définition d'une algèbre de von Neumann utilise l'adjoint et donc une structure hermitienne. Toutefois l'apport un peu artificiel de cette structure hermitienne simplifiera la démonstration qui est plutôt rosse, en évitant le recours à l'espace dual. L'idée de cette démonstration (l'ingénieuse étape 3 ci dessous) est due à Halperin et Rosenthal (*Amer. Math. Monthly* (1980) page 810). Insistons enfin sur le fait qu'on ne suppose pas dans le théorème que $\mathrm{id}_H \in A$.

Démonstration du Théorème 6.6. Convenons de dire que le sous espace vectoriel F de H est *invariant* si $A(F) \subset F$. De plus, sans perte de généralité on suppose $\dim H \geq 2$ car le résultat est trivial si $\dim H = 0$ ou 1.
Etape 1. Soit $x_0 \in H \setminus \{0\}$ et $F = A(x_0) = \{a(x_0); a \in A\}$. Je dis que $F = H$. En effet F est invariant. L'espace F n'est pas $\{0\}$ car sinon $F_1 = \mathbb{C}x_0$ satisferait $A(F_1) = \{0\} \subset F_1$ et serait donc invariant et différent de $\{0\}$ et de H : c'est contraire à l'hypothèse. Par conséquent l'espace invariant F est égal à H.
Etape 2. Soit $y_0 \in H \setminus \{0\}$ et $G = A^*(x_0) = \{a^*(x_0); a \in A\}$. Je dis que $G = H$. En effet soit F l'orthogonal de G. Alors F est invariant, car dire $x_0 \in F$ est équivalent à dire que pour tout $a \in A$ on a

$$0 = \langle a^*(y_0), x_0\rangle = \langle y_0, a(x_0)\rangle. \tag{6.10}$$

Il est bien clair que $b(x_0)$ a la même propriété pour tout $b \in A$ et donc que $b(F) \subset F$: d'où l'invariance de F. Ensuite, si F comprend un vecteur x_0 non nul alors $F \supset A(x_0) = H$ d'après l'étape 1, et l'égalité 6.10 entraîne la contradiction $y_0 = 0$. Donc $F = \{0\}$ et son orthogonal G est H.
Etape 3. Rappelons que le rang d'un endomorphisme d'un espace vectoriel de dimension finie est la dimension de son image. Nous montrons que A contient nécessairement un élément de rang 1, c'est à dire un endomorphisme de la forme $x_0 \otimes y_0$. Par cette notation, où x_0 et y_0 sont dans $H \setminus \{0\}$, nous signifions l'application de H dans H définie par

$$x \mapsto x_0\langle y_0, x\rangle.$$

C'est bien de rang 1 comme on le voit en considérant sa matrice représentative dans une bon. Il est aussi facile de voir que tout endomorphisme de H de rang 1 est de cette forme.

Si $r(a)$ désigne le rang de $a \in A$, soit $r_0 = \min\{r(a);\ a \in A\setminus\{0\}\}$. Il existe certainement $a_0 \in A \setminus \{0\}$ tel que $r(a_0) = r_0$ et nous cherchons à montrer $r_0 = 1$. Si c'est faux, alors $r_0 \geq 2$ et il existe donc x_1 et x_2 dans H tels que $a_0(x_1)$ et $a_0(x_2)$ soient indépendants. D'après l'étape 1 appliquée à $x_0 = a_0(x_1)$, il existe $a_1 \in A$ tel que $a_1a_0(x_1) = x_2$. Par conséquent $a_0a_1a_0(x_1)$ et $a_0(x_1)$ sont indépendants, et donc pour tout $\lambda \in C$ le vecteur de H égal à $a_0a_1a_0(x_1) - \lambda a_0(x_1)$ est non nul. Cela entraîne enfin que pour tout $\lambda \in C$ l'élément de A égal à $a_0a_1a_0 - \lambda a_0$ est non nul.

Soit alors $F = a_0(H)$ l'image de a_0. On a $\dim F = r_0$ par définition. On observe que $a_0a_1(F) \subset F$ et on note par $b \in L(F)$ la restriction de a_0a_1 à F. L'endomorphisme b de l'espace *complexe* F a une valeur propre λ_0 associée au vecteur propre f_0 (Ce point est crucial : le théorème serait faux pour des espaces vectoriels réels). Le rang de $b - \lambda_0\mathrm{id}_F$

étant $\leq r_0 - 1$ on en déduit que le rang de $a_0a_1a_0 - \lambda_0 a_0$ est $\leq r_0 - 1$. Or A étant une algèbre, $a_0a_1a_0 - \lambda_0 a_0$ est un élément de A. On a vu qu'il est non nul. Comme il est de rang $\leq r_0 - 1$ cela contredit la définition de r_0. Par conséquent $r_0 \geq 2$ est impossible et l'étape 3 est montrée.

Etape 4. A contient tous les éléments de $L(H)$ de rang 1. En effet on a vu à l'étape 3 qu'il existait un $a_0 = x_0 \otimes y_0$ dans A, avec $x_0 \neq 0$ et $y_0 \neq 0$. Si $x_1 \in H$ on a vu à l'étape 1 qu'il existait $a_1 \in A$ tel que $a_1(x_0) = x_1$. Si $y_1 \in A$ on a vu à l'étape 2 qu'il existait $a \in A$ tel que $a^*(y_0) = y_1$. par conséquent $x \mapsto a_1a_0a(x)$ est l'endomorphisme

$$a_1a_0a(x) = a_1(x_0\langle y_0, a(x)\rangle) = a_1(x_0\langle a^*(y_0), x\rangle) = a_1(x_0\langle y_1, x\rangle) = x_1\langle y_1, x\rangle = (x_1 \otimes y_1)(x).$$

Donc $a_1a_0a = x_1 \otimes y_1 \in A$.

Etape 5. Nous montrons enfin que tout élément de $L(H)$ est somme d'éléments de rang 1, ce qui acchèvera la démonstration d'après l'étape 4 et le fait que A est un sous espace vectoriel de $L(H)$. En effet si c'était faux il existerait un $b \in L(H) \setminus \{0\}$ tel que $\mathrm{trace}\,(bx_1 \otimes y_1) = 0$ pour tous les x_1 et y_1 de H et donc

$$0 = \mathrm{trace}\,(bx_1 \otimes y_1) = \langle y_1, b(x_1)\rangle.$$

Appliquant cela à $y_1 = b(x_1)$ on obtient $\|b(x_1)\|^2 = 0$. Comme c'est vrai pour tout x_1 c'est dire que $b = 0$, ce qui est une contradiction. Le théorème de Burnside est démontré.

Corollaire 6.7. Toute sous algèbre A de $L(H)$ distincte de $L(H)$ admet au moins un sous espace invariant différent de $\{0\}$ et H. De plus, si F est un sous espace invariant de A minimal (en ce sens que si $F_1 \subset F$ est aussi invariant alors $F_1 = F$ ou $\{0\}$) alors la restriction de A à F est $L(F)$.

Remarques. La première partie du corollaire ci dessous est équivalente au théorème 6.6 et est aussi appelée théorème de Burnside par certains auteurs. Le corollaire se concentre sur la notion de sous espace F invariant de H par l'algèbre A ainsi que sur celle de sous espace invariant minimal. Il est utile de tester les exemples 6.1 et 6.2 : les sous espaces invariants minimaux de l'exemple 6.1 sont les H_j, alors que ceux de l'exemple 6.2 sont les sous espaces vectoriels de dimension 1 de chaque E_i.

Démonstration du corollaire 6.7. Si A n'avait pas de sous espace invariant non trivial, le théorème 6.3 entraînerait que l'algèbre est égale à $L(H)$. Ensuite la restriction de A au sous espace invariant F est une algèbre pour l'espace F. Comme elle n'a pas de sous espace invariant non trivial elle est égale à $L(F)$.

Corollaire 6.8. Si $A \subset L(H)$ est une algèbre de von Neumann alors son commutant A' est formé des multiples de l'identité si et seulement si $A = L(H)$.

Démonstration. $\Leftarrow$ découle de la première partie de la proposition 6.2. $\Rightarrow$ Si $A \neq L(H)$ alors A admet un espace invariant F non trivial (corollaire 6.7). Soit p_F la projection orthogonale de H sur F : ce n'est pas un multiple de l'identité puisque F est non trivial. Ensuite, observons que puisque A est non seulement une algèbre, mais une algèbre de von

Neumann, alors $F^\perp$ est aussi invariant par A : si $a \in A$, puisque F est aussi conservé par a^*, pour tous $w \in F^\perp$ et $v \in F$ on a $\langle a(w), v\rangle = \langle w, a^*(v)\rangle = 0$. Nous en déduisons que p_F est dans A' ce qui apportera la contradiction désirée. Pour $v \in F$ on a $a(v) \in F$ et donc $p_F a(v) = a(v) = ap_F(v)$. Même raisonnement si $v \in F^\perp$ d'où $p_F a = ap_F$. Le corollaire est montré.

Nous sommes maintenant en mesure de terminer la classification des algèbres de von Neumann, c'est à dire de montrer qu'il n'y a pas d'autres facteurs que ceux qui sont décrits dans la proposition 6.2. Couplé avec la proposition 6.3 cela donne l'affirmation faite en 6.9.

Théorème 6.9. Soit $A \subset L(H)$ un facteur. Alors H est la somme directe orthogonale de r espaces isomorphes telle que si $H = H_1^{\oplus r}$ alors A est formé des $a = \mathrm{diag}(a_1, \ldots, a_1)$ où a_1 est un élément arbitraire de $L(H_1)$.

Démonstration. Appliquons la proposition 6.3 au commutant A', qui est aussi une algèbre de von Neumann. Il existe donc une décomposition en somme directe orthogonale $H = H_1 \oplus \cdots \oplus H_k$ et des facteurs $B_j \subset L(H_j)$ tels que A' est formé des $x = \mathrm{diag}(x_1, \ldots, x_k)$ tels que $x_j \in B_j$. Notons pour simplifier par p_j la projection orthogonale de H sur H_j. Puisque p_j est dans A' on peut dire que $A_j = p_j A p_j \subset L(H_j)$ est une algèbre de von Neumann sur H_j et que l'application φ de A dans A_1 définie par $\varphi(a) = p_1 a p_1$ satisfait $\varphi(ab) = \varphi(a)\varphi(b)$ pour tous a et b de A. De plus le sous espace vectoriel de A suivant

$$I = \{a \in A;\ p_1 a p_1 = 0\},$$

c'est à dire le noyau de l'application linéaire φ, satisfait $AI \subset I$ et $IA \subset I$ et est stable par $a \mapsto a^*$. Comme A est un facteur, d'après la proposition 6.4, I est ou bien égal à A (ce qui donnerait l'absurdité $p_1 = 0$ puisque A contient id_H) ou bien $I = \{0\}$. Donc φ est injectif. Comme φ est surjectif par définition de A_1 on en déduit que A et A_1 sont isomorphes -aussi bien que A et A_j.

Dernière étape : arriver à montrer que $A_1 = L(H_1)$. Pour cela on écrit

$$p_1 A' p_1 = (p_1 A' p_1)'' = (p_1 A p_1)' = A_1'$$

La première égalité est le théorème 6.5 du double commutant appliqué à l'algèbre de von Neumann $p_1 A' p_1 \subset L(H_1)$. La deuxième résulte de $(p_1 A' p_1)' = p_1 A p_1 = A_1$ qui résulte à son tour de $A'' = A$. Maintenant supposons que $p_1 A' p_1$ contienne un $x_1 \in L(H_1)$ qui ne soit pas un multiple de id_{H_1}. Alors $x_1 x_1^*$ a un espace propre $F \subset H_1$ non trivial, et donc p_F est dans $p_1 A' p_1$. Cela contredit le fait que $H = H_1 \oplus \cdots \oplus H_k$ est la décomposition de la proposition 6.3 pour l'algèbre A'. Par conséquent $A_1' = p_1 A' p_1$ n'est formée que de multiples de l'unité. Appliquant alors le corollaire 6.8 du théorème de Burnside on en déduit que $A_1 = L(H_1)$ et la démonstration est achevée.

Exercice 6.1 Si $H = \mathbb{C}^2$ est l'espace hermitien canonique de dimension 2, on identifie $L(H)$ aux matrices complexes (2,2). Quelle est la plus petite algèbre de von Neumann qui contient

$a = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$? Considérons ensuite avec $H = \mathbb{C}^3$

$$a = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Quelle est la plus petite algèbre de von Neumann qui contient a? Méthode : puisque $a^2 = 0$ former toutes les combinaisons linéaires de $\mathrm{id}, a, a^*, aa^*, a^*a$. Est ce que b est normal ? Montrer que $L(H)$ est la plus petite algèbre de von Neumann qui le contienne. Méthode : calculer $b^2, b^2b^*, bb^*, bb^*b^*$.

Chapitre 4

Formes quadratiques

Ce chapitre est court, abstrait et difficile. Historiquement, les formes quadratiques sont apparues comme des polynômes homogènes à q variables à coefficients dans $\mathbb{R}$ ou $\mathbb{C}$, comme pour $q = 2$ ou 3

$$Q(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2, \tag{0.1}$$

$$Q(x_1, x_2, x_3) = a_1x_1^2 + a_2x_2^2 + a_3x_3^2 + 2b_3x_1x_2 + 2b_1x_3x_1 + 2b_2x_2x_3. \tag{0.2}$$

Mais dans 4 cas sur 5 on les voyait apparaitre dans un contexte euclidien traité à la facon du XIX ème siècle, c'est à dire avec coordonnées. Nous avons appris au chapitre 2 section 9 à les étudier avec le théorème spectral en leur associant un endomorphisme symétrique. Toutefois cette méthode ne permet pas l'étude dans le cas complexe et dans d'autres corps. Ensuite, même dans le cas réel imposer une structure euclidienne peut être artificiel : le médecin qui collecte des données de patients du genre (poids en kilos, taille en cm)$= (x, y)$ n'a que faire d'une structure euclidienne avec une norme de la forme $\sqrt{ax^2 + 2bxy + cy^2}$. Enfin les formes quadratiques apparaissent de facon naturelle en géométrie différentielle, en théorie des nombres, en mathématiques appliquées et en informatique (codes correcteurs d'erreur). En résumé, bien que les formes quadratiques ne forment pas une industrie aussi importante que celle des endomorphismes, leur étude est utile et intéressante.

I Matrices représentatives d'une forme bilinéaire ou quadratique

On sait (Chapitre 2, section 2) que si E et F sont des espaces vectoriels sur un même corps K, une forme bilinéaire est une application $(x, y) \mapsto B(x, y)$ de $E \times F$ dans K telle que $x \mapsto B(x, y)$ est linéaire sur E pour $y \in F$ fixé, et telle que $y \mapsto B(x, y)$ est linéaire sur F pour $x \in E$ fixé. On pourrait interpréter B comme la donnée d'une application φ_B linéaire de E à valeurs dans le dual F^* définie par $B(x, y) = \varphi_B(x)(y)$, un point de vue fécond qui est développé en licence pour introduire les produits tensoriels d'espaces.

Si $E = F$ et si de plus $B(x, y)$ est symétrique, c'est à dire que $B(x, y) = B(y, x)$ alors la fonction $x \mapsto Q_B(x)$ sur E à valeurs dans K est appelée forme quadratique associée à

B. En fait, la connaissance de Q_B donne la connaissance de B par polarisation si on a le droit de diviser par 2 dans le corps K :

Proposition 1.1. Soit K un corps tel que si $\lambda \in K$ est tel que $\lambda + \lambda = 0$ alors $\lambda = 0$. Soit E un espace de dimension finie sur K et soit $Q : E \to K$ telle que

- Pour tout $\lambda \in K$ et tout $x \in E$ on a $Q(\lambda x) = \lambda^2 Q(x)$.
- $B(x, y) = \frac{1}{2}(Q(x + y) - Q(x) - Q(y))$ est une forme bilinéaire symétrique sur E.

Alors Q est la forme quadratique associée à B (et B s'appelle la *polarisée* de Q). De plus si B_1 est une autre forme bilinéaire symétrique telle que $Q_B = Q_{B_1}$ alors $B = B_1$.

Démonstration. La première propriété joue un rôle crucial :

$$B(x, x) = \frac{1}{2}(Q(2x) - 2Q(x)) = \frac{1}{2}(4Q(x) - 2Q(x)) = Q(x).$$

Ensuite, on voit facilement que puisque B_1 est bilinéaire symétrique alors

$$B_1(x, y) = \frac{1}{2}(B_1(x + y, x + y) - B_1(x, x) - B_1(y, y))$$

et donc $B = B_1$.

Soit à nouveau E et F des espaces vectoriels sur K, et de dimensions finies p et q, équipés de bases e et f. Soit B une forme bilinéaire sur $E \times F$. Alors la matrice

$${}_e[B]_f = (B(e_i, f_j))_{1 \leq i \leq p, 1 \leq i \leq p}$$

est appelée *matrice représentative* de B dans les bases (e, f). Elle permet de calculer immédiatement $B(x, y)$ connaissant les matrices de composantes $X = [x]^e$ et $Y = [y]^f$:

$$B(x, y) = X^T \; {}_{e'}[B]_{f'} Y,$$

et le changement de base se fait immédiatement : si e' et f' sont de nouvelles bases de E et F avec $P = [\mathrm{id}_E]^e_{e'}$ et $Q = [\mathrm{id}_F]^f_{f'}$ pour matrices de changement de base, alors

$${}_e[B]_f = P^T \; {}_e[B]_f Q. \tag{1.3}$$

Remarques. Il est inutile d'apprendre cette égalité par coeur, mais il faut savoir la retrouver. Il est important de noter qu'elle est différente de la formule de changement de base pour une application linéaire vue au chapitre 1, section 1. En mathématiques pures ou appliquées on rencontre souvent des matrices carrées ou rectangulaires : il est important de savoir s'il faut les interpréter comme des matrices représentatives d'applications linéaires (cas le plus fréquent) ou comme des matrices représentatives de formes bilinéaires.

Finalement, bien que le présent chapitre se concentre sur les formes quadratiques et donc les formes bilinéaires symétriques, il faut savoir que la prochaine catégorie intéressante est celle des formes bilinéaires antisymétriques (ou alternées) sur $E \times E$ satisfaisant $B(x, y) = -B(y, x)$ à la base de la géométrie symplectique et de la mécanique moderne.

Supposons maintenant $E = F$ de dimension q muni de la base e et soit B une forme bilinéaire. Alors B est symétrique si et seulement si la matrice représentative $A = {}_e[B]_e = (B(e_i, e_j) = (a_{ij})$ l'est. Dans ce cas, ${}_e[B]_e$ est appelée aussi la matrice représentative de la forme quadratique $Q = Q_B$ et on a pour $X = [x]^e = (x_1, \ldots, x_q)$

$$Q(x) = X^T \ {}_e[B]_e X = \sum_{i,j=1}^{q} a_{ij} x_i x_j.$$

Par exemple les matrices représentatives des formes quadratiques sur K^2 et K^3 en base canonique des exemples (0.1) et (0.2) sont respectivement

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}, \quad \begin{bmatrix} a_1 & b_3 & b_2 \\ b_3 & a_2 & b_1 \\ b_2 & b_1 & a_3 \end{bmatrix}.$$

Quant à la formule de changement d'une base e à une base f de E, si $P = [\mathrm{id}_E]^e_f$ c'est donc

$${}_e[B]_e = P^T \ {}_f[B]_f P.$$

On utilise parfois le vocabulaire suivant : on dit que deux matrices carrées d'ordre q symétriques A et B sur K sont *congruentes* si il existe une matrice inversible P d'ordre q telle que $A = PBP^T$. C'est à distinguer des matrices *semblables* pour lesquelles $A = PBP^{-1}$. La congruence est liée à la représentation d'une même forme quadratique dans différentes bases, la similitude est liée à la représentation d'un même endomorphisme dans différentes bases.

Dans le corps K introduisons la relation d'équivalence suivante $\lambda \sim \lambda'$ si et seulement si il existe $\mu \neq 0$ dans K tel que $\lambda = \mu^2 \lambda'$. Par exemple si $K = \mathbb{R}$ il y a trois classes d'équivalence, celles de 1, de 0 et de -1. Si $K = \mathbb{C}$ il n'y a que celles de 1 et de 0 et si $K = \mathbf{Q}$ il y a celles de 0, de n et de $-n$ où n est un entier positif produit de nombres premiers *distincts.*

On remarque alors que le déterminant de ${}_e[B]_e$ dépend de la base puisque $\det({}_e[B]_e) = (\det P)^2 (\det \ {}_f[B]_f)$. Ce qui est intrinsèque à la forme quadratique Q indépendamment de la base dans ce déterminant est donc seulement sa *classe d'équivalence* pour la relation d'équivalence précédente. Cette classe d'équivalence est appelée *discriminant* de Q.

Exercice 1.1 Soit A une matrice symétrique d'ordre q inversible sur le corps K. Si $X \in K^q$ est écrit comme une matrice colonne on considère la forme quadratique

$$Q(X) = -\det \begin{bmatrix} A & X \\ X^T & 0 \end{bmatrix}.$$

Montrer que la matrice représentative de Q dans la base canonique de K^q est $A^{-1} \det A$, c'est à dire la matrice des cofacteurs de A. Méthode : pour calculer $Q(X)$ utiliser l'exercice 2.1 du chapitre 1 appliqué à $B = X$, $C = X^T$ et $D = 0$.

II Orthogonalité, noyau, rang, diagonalisation

Soit Q une forme quadratique dans E espace de dimension finie sur K. On suppose *désormais* que si $\lambda \in K$ est tel que $\lambda + \lambda = 0$ alors $\lambda = 0$. Soit B la forme polaire de Q définie à la Proposition 1.1. On dit que x et y dans E sont *orthogonaux* pour Q si $B(x, y) = 0$. On dit qu'une base $e = (e_1, \ldots, e_q)$ est orthogonale si $B(e_i, e_j) = 0$ (c'est à dire que ${}_e[B]_e$ est diagonale). On dit que $x \in E$ est *isotrope* si $Q(x) = 0$.

Notez que l'ensemble des vecteurs isotropes *n'est pas un sous espace vectoriel* en général. Exemples : $E = \mathbb{R}^2$ et $Q(x_1, x_2) = x_1 x_2$; l'ensemble des vecteurs isotropes est ici la réunion de deux droites. Si $E = \mathbb{R}^3$ et $Q(x_1, x_2, x_3) = x_1^2 + x_2^3 - x_3^3$ l'ensemble des vecteurs isotropes est ici la réunion de deux cônes de révolution. L'ensemble des x tels que $Q(x) = 0$ est un espace vectoriel dans le cas très exceptionnel où $K = R$ et Q est définie positive, comme vu au chapitre 2, section 3.

Ce mot "orthogonal" est un choix malheureux qui n'indique qu'une vague analogie avec le cas euclidien. Si U est une partie de E on appelle orthogonal de U pour la forme quadratique Q l'ensemble

$$U^0 = \{x \in E \ ; \ B(x, y) = 0 \ \forall y \in U\}.$$

Cette fois, U^0 est un sous espace vectoriel de E, c'est facile à vérifier. Le sous espace E^0 est alors appelé le *noyau*[1] de Q. Il ne faut *pas le confondre* avec l'ensemble des vecteurs isotropes. Le *rang* de Q est l'entier $\dim E - \dim E^0$. On dit que Q est non dégénérée si $E^0 = \{0\}$. Q est alors de rang $q = \dim E$.

Proposition 2.1. Le rang de Q est égal au rang de sa matrice représentative $A = {}_e[B]_e$ dans une base quelconque e de E. En particulier Q est non dégénérée si et seulement si $\det A \neq 0$.

Démonstration : Par définition $y \in E^0$ si et seulement si $B(x, y) = 0$ pour tout $x \in E$. Passant à la base e avec $Y = [y]^e$ on a $X^T A Y = 0$ pour tout X et donc $AY = 0$, ce qui montre que Y est dans le noyau de l'application linéaire de K^q de matrice A dans la base canonique. Le rang de A est donc bien $q - \dim E^0$.

Théorème 2.2. Soit Q une forme quadratique sur E de dimension finie q. Alors il existe une base e de E telle que la matrice représentative de Q dans E soit diagonale, c'est à dire telle que e soit orthogonale pour Q.

Démonstration : On procède par récurrence sur q. C'est trivial pour $q = 1$. Supposons que ce soit vrai pour $q - 1$ et prenons E de dimension q. Si $Q(x) = 0$ pour tout x alors toute base est orthogonale. Sinon il existe un vecteur e_1 non isotrope. Soit F l'orthogonal de Ke_1. Il ne contient pas e_1, qui est non isotrope. C'est le noyau de l'application linéaire $x \mapsto B(x, e_1)$ et il est donc de dimension $q-1$. Donc $E = Ke_1 \oplus F$. Appliquons l'hypothèse de récurrence à F : il possède donc une base orthogonale f et donc $e = \{e_1\} \cup f$ est une base orthogonale pour Q.

---

[1] Certains auteurs disent *radical*.

Corollaire 2.3. Si Q est une forme quadratique de rang p sur un espace E de dimension q alors il existe des nombres non nuls $a_1, \ldots, a_p$ et des formes linéaires sur E indépendantes $f_1, \ldots, f_p$ tels que

$$Q(x) = a_1(f_1(x))^2 + \cdots + a_p(f_p(x))^2. \tag{2.4}$$

Démonstration : En effet, si e est une base de diagonalisation de Q, si $A = {}_e[B]_e = \mathrm{diag}(a_1, \ldots, a_q)$ et si $X = [x]^e = (x_1, \ldots, x_q)^T$ alors $Q(x) = a_1x_1^2 + \cdots + a_qx_q^2$. Puisque le rang p de la matrice représentative de Q ne dépend pas de la base, il y a donc exactement p nombres parmi les (a_i) qui ne sont pas nuls. Sans perte de généralité on suppose que ce sont les p premiers. Donc $Q(x) = a_1x_1^2 + \cdots + a_px_p^2$. Enfin, soit f_j la forme linéaire définie par $f_j(x) = x_j$. Le fait que e soit une base entraîne que les $f_1, \ldots, f_q$ sont indépendants et le corollaire est montré.

Remarque. Il est important de noter que si la base de diagonalisation e de Q est telle que

$${}_e[B]_e = \mathrm{diag}(a_1, \ldots, a_p, 0, \ldots, 0)$$

avec $a_1a_2 \ldots a_p \neq 0$ alors $(e_{p+1}, \ldots, e_q)$ est une base de E^0. En effet si $p < j$ alors e_j est orthogonal à tous les e_i *y compris lui même*. Comme e est une base de E cela entraîne que e_j est dans E^0. Comme les $(e_{p+1}, \ldots, e_q)$ sont indépendants et que $\dim E^0 = q - p$ on a le résultat.

L'algorithme de Gauss de décomposition d'une forme quadratique en carrés. Cet algorithme, sur lequel on peut baser une seconde démonstration au Théorème 2.2, considère une forme quadratique Q sur l'espace K^q (où le corps K est tel que $2\lambda = 0$ entraîne $\lambda = 0$) définie par une matrice symétrique $A = (a_{ij})$ d'ordre q à coefficients dans K ainsi :

$$Q(x_1, \ldots, x_q) = \sum_{i,j=1}^{q} a_{ij}x_ix_j = C + R$$

où les sommes C et R, dites des carrés et des rectangles, sont respectivement

$$C = \sum_{i=1}^{q} a_{ii}x_i^2x_j \quad R = 2 \sum_{1 \leq i < j \leq q} a_{ij}x_ix_j.$$

Cet algorithme calcule le rang p et construit des formes linéaires f_j indépendantes telles que (2.4) soit vrai. Nous n'allons pas expliquer le programme informatique, mais nous contenter d'explications informelles.

1. Ou bien $C \neq 0$. Dans ce cas il existe un i tel que $a_{ii} \neq 0$. Prenons $i = 1$ sans perte de généralité. Alors $Q(x) = a_{11}x_1^2 + 2x_1f(x_2, \ldots, x_q) + S(x_2, \ldots, x_q)$ où f est une forme linéaire et S est une forme quadratique sur K^{q-1}. On pose alors $f_1(x_1, \ldots, x_q) = x_1 + \frac{1}{a_{11}}f_2$ et $Q_1 = S - \frac{1}{a_{11}}f^2$. Alors $Q = a_{11}f_1^2 + Q_1$ et la forme quadratique Q_1 est par rapport aux variables $x_2, \ldots, x_q$. Tout cela n'est que la bonne vieille technique consistant à compléter le carré dans un trinôme du second degré.

2. Ou bien $C = 0$. Le cas $R = 0$ étant trivial, supposons $R \neq 0$. Sans perte de généralité on suppose $a_{12} \neq 0$. Alors
$$Q(x) = 2a_{12}x_1x_2 + 2x_1g(x_3, \dots, x_q) + 2x_2f(x_3, \dots, x_q) + S(x_3, \dots, x_q)$$
où f et g sont des formes linéaires et S est une forme quadratique sur K^{q-2}. L'*astuce* est alors d'introduire les deux formes linéaires sur K^q suivantes
$$\begin{aligned} f_1(x_1, \dots, x_q) &= \frac{1}{2}(x_1 + x_2) + \frac{1}{2a_{12}}(f + g), & (2.5) \\ f_2(x_1, \dots, x_q) &= \frac{1}{2}(x_1 - x_2) + \frac{1}{2a_{12}}(f - g) & (2.6) \end{aligned}$$
et la forme quadratique $Q_2 = S - \frac{2}{a_{12}}fg$. On obtient alors $Q = 2a_{12}f_1^2 - 2a_{12}f_2^2 + Q_2$.

Pratiquons quelques exemples : si $Q(x_1, x_2, x_3) = x_1^2 + x_2^2 + 2x_3^2 - 4x_1x_2 + 6x_2x_3$ on est dans le cas 1) et on obtient
$$Q = (x_1^2 - 4x_1x_2 + 4x_2^2) - 3x_2^2 + 2x_3^2 + 6x_2x_3 = (x_1 - 2x_2)^2 + Q_1$$
avec $Q_1 = -3x_2^2 + 2x_3^2 + 6x_2x_3$. On réapplique la procédure 1) : $Q_1 = -3(x_2 - 2x_3)^2 + 14x_3^2$. et on a la décomposition de Q en une combinaison linéaire de carrés de formes linéaires indépendantes
$$Q = (x_1 - 2x_2)^2 - 3(x_2 - 2x_3)^2 + 14x_3^2.$$
Autre exemple, où il faut utiliser la procédure 2) : $Q(x_1, x_2, x_3) = 2x_1x_2 + 2x_2x_3 + 2x_1x_3$. Ici $f = g = x_3$ et l'application des formules (2.5) et (2.6) donne, cette fois ci en une seule étape
$$Q = \frac{1}{2}(x_1 + x_2 + 2x_3)^2 - \frac{1}{2}(x_1 - x_2)^2 - 2x_3^2.$$
Dans ces deux exemples le rang est 3.

Exercice 2.1 Diagonaliser les formes quadratiques suivantes dans $\mathbb{R}^4$ et $\mathbb{R}^3$ par l'algorithme de Gauss : $Q = 2(x_1x_2 + x_2x_3 + x_3x_4)$; $Q = x_1^2 + x_2^2 + x_3^2 - 2(x_1x_2 + x_2x_3 + x_3x_1)$.

Exercice 2.2 Diagonaliser par l'algorithme de Gauss la forme quadratique dans $\mathbb{R}^n$ avec $n \geq 2$, dont la matrice représentative dans la base canonique est :
$$\begin{bmatrix} 0 & 0 & \dots & 0 & 1 & 1 \\ 0 & 0 & \dots & 1 & 1 & 0 \\ 0 & 0 & \dots & 1 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 1 & \dots & 0 & 0 & 0 \\ 1 & 1 & \dots & 0 & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 & 0 \end{bmatrix}$$

III La signature d'une forme quadratique réelle

Le corollaire 2.3 montre qu' une forme quadratique est combinaison linéaire d'un nombre fixe p de carrés de formes linéaires indépendantes, bien qu'il y ait de nombreuses manières différentes de le faire. Dans cette section, nous allons supposer que $K = \mathbb{R}$ et montrer une invariance plus forte appelée *loi d'inertie de Sylvester.*

Théorème 3.1. Soit Q une forme quadratique de rang p sur l'espace réel E de dimension finie q. Si e est une base de diagonalisation de Q soit $\ _e[B]_e = \mathrm{diag}(a_1, \ldots, a_q)$. Soit r le nombre de j tels que $a_j > 0$, soit $s = p - r$ le nombre de j tels que $a_j < 0$ et $t = q - p$ le nombre de j tels que $a_j = 0$. Alors le triplet d'entiers (r, s, t), appelé *signature* de Q, est indépendant de la base de diagonalisation e choisie.

Démonstration : On a déjà vu que $r+s = p$ ne dépend pas de e puisque c'est le rang de Q. Il suffit donc de montrer que r est constant. Soit e' une autre base de diagonalisation avec $\ _{e'}[B]_{e'} = \mathrm{diag}(a'_1, \ldots, a'_q)$ correspondant au triplet (r', s', t) avec, sans perte de généralité $r' \geq r$, avec $a_j > 0$ pour $j = 1, \ldots, r$, avec $a'_j > 0$ pour $j = 1, \ldots, r'$ et avec $a_{p+1} = \ldots = a_q = a'_{p+1} = \ldots = a'_q = 0$.

Montrons qu'alors $e'_1, \ldots, e'_{r'}, e_{r+1}, \ldots, e_q$ sont linéairement indépendants. En effet, s'ils ne le sont pas on peut trouver des nombres $\lambda_1, \ldots, \lambda_{r'}, \mu_{r+1}, \ldots, \mu_q$ tels que

$$\lambda_1 e'_1 + \cdots + \lambda_{r'} e'_{r'} = -\mu_{r+1} e_{r+1} - \cdots - \mu_q e_q. \tag{3.7}$$

Notons par x la valeur commune des deux membres de l'égalité précédente et calculons de deux manières le nombre $Q(x) = B(x, x)$: à cause de l'orthogonalité des e_j entre eux et des e'_j entre eux on obtient

$$Q(x) = a'_1(\lambda_1)^2 + + \cdots + a'_{r'}(\lambda_{r'})^2 = a_{r+1}(\mu_{r+1})^2 + \cdots + a_p(\mu_p)^2.$$

Le membre de gauche est ≥ 0, le membre de droite est ≤ 0, ils sont donc tous deux nuls. Donc $\lambda_1 = \ldots = \lambda_{r'} = \mu_{r+1} = \ldots = \mu_p = 0$. Donc (3.7) se transforme en $0 = -\mu_{p+1} e_{p+1} - \cdots - \mu_q e_q$. Comme les $(e_{p+1}, \ldots, e_q)$ sont indépendants, les derniers μ sont nuls et donc les $e'_1, \ldots, e'_{r'}, e_{r+1}, \ldots, e_q$ sont indépendants. En conséquence $r' + (q - r) \leq q$ et donc $r' \leq r$. Comme nous avions choisi $r' \geq r$ on a bien $r = r'$ et le théorème de Sylvester est montré.

Proposition 3.2. Soit Q quadratique sur un espace réel E, de signature (r, s, t). Alors il existe $p = r + s$ formes linéaires indépendantes sur E telles que

$$Q(x) = (f_1(x))^2 + \cdots + (f_r(x))^2 - (f_{r+1}(x))^2 \cdots - (f_p(x))^2.$$

Démonstration : On sait déjà qu'il existe une base e telle que

$$_e[B]_e = \mathrm{diag}(a_1, \ldots, a_p, 0, \ldots, 0)$$

avec $a_1, \ldots, a_r > 0$ et $a_{r+1}, \ldots, a_p < 0$. Si $[x]^e = (x_1, \ldots, x_q)^T$, considérons les formes linéaires g_j définies par $g_j(x) = x_j$: elles sont indépendantes puisque e est une base. Il suffit maintenant de définir $f_j(x) = |a_j|^{1/2} g_j(x)$ pour $j = 1, \dots, p$. Les f_j ont la propriété annoncée.

Remarques. Pour comprendre ce chapitre, assez abstrait, il est tout à fait essentiel de comparer les résultats de cette section avec ce qu'on sait déjà des formes quadratiques sur un espace euclidien étudiées au chapitre 2 section 9. On y a vu que si Q est quadratique sur l'espace euclidien E alors il est associé à Q un endomorphisme symétrique a tel que $Q(x) = \langle a(x), x \rangle$. Si on applique à a le théorème spectral, on révèle une bon (au sens euclidien) de diagonalisation de a. Soit $[a]_e^e = \mathrm{diag}(\mu_1, \dots, \mu_q))$. Si $[x]^e = (x_1, \dots, x_q)^T$ on a

$$Q(x) = \mu_1 x_1^2 + \cdots + \mu_q x_q^2,$$

c'est à dire que e est aussi une base de diagonalisation au sens des formes quadratiques. Mais il y a beaucoup plus de bases de diagonalisation de la forme quadratique Q que de bases de diagonalisation pour l'endomorphisme a associé par la structure ambiante d'espace euclidien. Par exemple si $E = \mathbb{R}^3$ euclidien canonique et $Q(x) = 2x_1x_2 + 2x_2x_3 + 2x_1x_3$, alors l'endomorphisme associé a a pour matrice A dans la base canonique e

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

de polynôme caractéristique $(X+1)^2(X-2)$. La signature est bien $(r, s, t) = (1, 2, 0)$. Ce n'est pas une surprise, car on avait vu en exemple à la section 2 de l'algorithme de Gauss que

$$2x_1x_2 + 2x_2x_3 + 2x_1x_3 = \frac{1}{2}(x_1 + x_2 + 2x_3)^2 - \frac{1}{2}(x_1 - x_2)^2 - 2x_3^2.$$

Parfois même l'étude de la forme quadratique donne des renseignements sur les valeurs propres : l'algorithme de Gauss est de nature élémentaire alors que la recherche de valeurs propres ne l'est pas, puisqu'il faut chercher les racines d'un polynôme. Reprenons l'exemple $Q(x) = x_1^2 + x_2^2 + 2x_3^2 - 4x_1x_2 + 6x_2x_3$ toujours dans $\mathbb{R}^3$ euclidien canonique. L'endomorphisme associé a a pour matrice A dans la base canonique e

$$A = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 1 & 3 \\ 0 & 3 & 2 \end{bmatrix}.$$

Ici les racines du polynôme caractéristique sont compliquées. Pourtant, comme on a vu que $Q = (x_1 - 2x_2)^2 - 3(x_2 - 2x_3)^2 + 14x_3^2$. on voit que la signature est $(2, 1, 0)$ et donc que il y a 2 valeurs propres positives et une négative. En particulier, l'algorithme de Gauss permet de dire si un endomorphisme symétrique est défini positif, ou positif.

Exercice 3.1 Quelle est la signature de la forme quadratique de l'exercice 2.2 ?

Exercice 3.2 Soit a un nombre réel. Discuter suivant a la signature de la forme quadratique

$$(x_1 + \cdots + x_n)(y_1 + \cdots + y_n) - a(x_1y_1 + \cdots + x_ny_n)$$

en lui associant un endomorphisme symétrique de l'espace euclidien canonique $\mathbb{R}^n$. Si $e = (e_1, \ldots, e_n)$ est la bon canonique de $\mathbb{R}^n$, il est intéressant de diagonaliser cet endomorphisme dans une base orthonormale dont un des vecteurs est proportionnel à $f = e_1 + \cdots + e_n$.

Exercice 3.3 Soit F le sous espace vectoriel de $\mathbb{R}^n$ formé par les $x = (x_1, \ldots, x_n)$ tels que $x_1 + \cdots + x_n = 0$.

1. Soit $a_1, \ldots, a_n \in \mathbb{R}$. Soit la forme quadratique sur $\mathbb{R}^n$

$$Q(x) = -\frac{1}{2}\sum_{i,j=1}^{n} |a_i - a_j| x_i x_j.$$

Montrer que la restriction de Q à F est positive, c'est à dire que $Q(x) \geq 0$ pour tout x de E. Méthode : supposer sans perte de généralité que $a_1 = 0 \leq a_2 \leq \ldots \leq a_n$, introduire $p_k^2 = a_k - a_{k-1}$ et montrer

$$Q(x) = \sum_{k=2}^{n} p_k^2 (\sum_{j=k}^{n} x_j)^2.$$

A quelle condition sur les a_j la restriction de Q à F est elle définie positive ?

2. Soit E un espace euclidien quelconque et soit $v_1, \ldots, v_n \in E$. Montrer que la restriction à F de la forme quadratique sur $\mathbb{R}^n$

$$Q(x) = -\frac{1}{2}\sum_{i,j=1}^{n} \|v_i - v_j\|^2 x_i x_j.$$

est positive. Méthode : montrer $Q(x) = \| \sum_{j=1}^{n} x_j v_j\|^2$. Montrer que la restriction de Q à F est définie positive si et seulement si les vecteurs $v_1, \ldots, v_n$ engendrent un espace affine de dimension $n-1$ (voir chapitre 5).

3. Soit $A = (a_{ij})_{1 \leq i,j \leq n}$ une matrice symétrique réelle telle que $a_{ii} = 0$ pour tout i. On suppose que la restriction de $Q(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$ à F est positive. On veut montrer la réciproque du 2, c'est à dire qu'il existe un espace euclidien E et des vecteurs $v_1, \ldots, v_n \in E$ tels que $a_{ij} = -\frac{1}{2}\|v_i - v_j\|^2$. En fait on va le montrer avec $E = \mathbb{R}^n$ euclidien canonique. Pour cela, on introduit une base $e = (e_1, \ldots, e_{n-1})$ de F qui diagonalise Q c'est à dire que si $x = f_1(x)e_1 + \cdots + f_{n-1}(x)e_{n-1}$ où les f_j sont des formes linéaires sur F, alors

$$Q(x) = \lambda_1 (f_1(x))^2 + \cdots + \lambda_{n-1}(f_{n-1}(x))^2,$$

avec par hypothèse $\lambda_i \geq 0$. On considère l'endomorphisme φ de F défini par

$$\varphi(x) = \lambda_1^{1/2} f_1(x) e_1 + \cdots + \lambda_{n-1}^{1/2} f_{n-1}(x) e_{n-1},$$

qui satisfait $Q(x) = \|\varphi(x)\|^2$. Prendre alors $v_1, \ldots, v_{n-1}$ dans F tels que $\varphi(x) = x_1 v_1 + \ldots + x_{n-1} v_{n-1}$ et montrer que $v_1, \ldots, v_{n-1}$ complété par $v_n = 0$ répond à la question.

4. Avec les notations de la question 1 et $v_i = (0, p_2, \dots, p_i, 0, \dots, 0) \in \mathbb{R}^n$ montrer que $|a_i - a_j| = \|v_i - v_j\|^2$.

Exercice 3.3 Soit F le sous espace vectoriel de $\mathbb{R}^n$ formé par les $x = (x_1, \dots, x_n)$ tels que $x_1 + \cdots + x_n = 0$. L'espace F est muni du produit scalaire $\langle x, y\rangle = x_1y_1 + \cdots + x_ny_n$. Soit $(a_1, \dots, a_n) \in \mathbb{R}^n$ et soit la forme quadratique sur F

$$Q(x) = a_1x^2 + \cdots + a_nx_n^2.$$

Montrer que l'endomorphisme de F associé à Q est

$$(x_1, \dots, x_n) \mapsto \frac{1}{n}((n-1)a_1x_1 - a_2x_2 - \cdots - a_nx_n, \dots, (n-1)a_nx_n - a_1x_1 - a_2x_2 - \cdots - a_{n-1}x_{n-1}).$$

IV Formes quadratiques unitaires et racines *

Comme le lecteur risque de s'ennuyer dans un chapitre un peu technique, complétons celui ci par une étude des formes quadratiques sur les réels à *coefficients entiers* pour la partie des termes rectangles et de la forme $\sum_{i=1}^n x_i^2$ pour la partie des termes carrés. La section suivante sur les graphes de Dynkin et leurs formes quadratiques associées fera rencontrer pour la première fois une des grandes divisions des objets mathématiques, la classification A, D, E qui se rencontre dans des parties des mathématiques aussi différentes que la théorie des catastrophes, la classification des 5 polyèdres réguliers de l'espace euclidien de dimension 3, les algèbres de Lie ou les fronts d'onde. [2]

Adoptons le vocabulaire suivant : une forme quadratique unitaire est une forme quadratique $q : \mathbb{Z}^n \to \mathbb{Z}$

$$x \mapsto \sum_{i=1}^n x_i^2 + \sum_{i<j} q_{ij}x_ix_j$$

avec $q_{ij} \in \mathbb{Z}$. Son extension canonique à $\mathbb{R}^n$ est notée $\overline{q}$. Nous noterons

$$(x, y) \mapsto q(x|y) = q(x+y) - q(x) - q(y) = 2\sum_{i=1}^n x_iy_i + \sum_{i\neq j} q_{ij}x_iy_j \tag{4.8}$$

ainsi que $\overline{q}(x|y)$ les formes bilinéaires correspondantes. Attention donc : $q(x|x) = 2q(x)$.

A une forme quadratique unitaire on associe un *bigraphe* dont les sommets sont les entiers $\{1, 2, \dots, n\}$ et tel que entre le sommet i et le sommet $j \neq i$ il y ait $|q_{ij}|$ arêtes (non orientées). Ces arêtes sont pleines si $q_{ij} < 0$ et en pointillé si $q_{ij} > 0$. On laisse le lecteur par exemple dessiner les bigraphes des formes unitaires suivantes

$$q_1 = \sum_{i=1}^5 x_i^2 - x_1(\sum_{i=2}^5 x_i)$$

[2] Voir V.I. Arnold *Catastrophe theory* page 103, Springer 1992.

$$\begin{aligned}
q_2 &= \sum_{i=1}^{7} x_i^2 - x_1(\sum_{i=2}^{7} x_i) + x_2x_3 + x_4x_5 + x_6x_7 \\
q_3 &= \sum_{i=1}^{8} x_i^2 - x_1(\sum_{i=2}^{8} x_i) + (x_2x_3 + x_3x_4 + x_4x_2) + (x_5x_6 + x_6x_7 + x_7x_5) \\
q_4 &= \sum_{i=1}^{9} x_i^2 - x_1(\sum_{i=2}^{9} x_i) + x_7x_8 + \frac{1}{2}\sum_{i\neq j, i,j=2,\ldots,6} x_ix_j \\
q_5 &= \sum_{i=1}^{9} x_i^2 - x_1(\sum_{i=2}^{9} x_i) + \frac{1}{2}\sum_{i\neq j, i,j=2,\ldots,5} x_ix_j + \sum_{6\leq i<j\leq 9} x_ix_j. \qquad (4.9)
\end{aligned}$$

On remarque qu'il n'y a pas d'arêtes multiples sur ces exemples. En revanche, si s est un entier > 0 alors le bigraphe correspondant à

$$q^{(s)}(x_1, x_2) = x_1^2 + x_2^2 - sx_1x_2 \qquad (4.10)$$

a s arêtes pleines entre les sommets 1 et 2.

Expliquons ensuite ce que sont les *composantes connexes* d'un bigraphe. Ignorons le fait que certaines arêtes soient pleines et d'autres pointillées. Un chemin de longueur m du sommet i au sommet j d'un bigraphe est la donnée d'une suite de sommets $i_0 = i, i_1, \ldots, i_m = j$ telle que i_{k-1}, i_k soit une arête pour tout $k = 1, \ldots, m$. Ceci a du sens même pour $m = 0$. Ceci introduit une relation d'équivalence $\sim$ sur les sommets, en convenant $i \sim j$ si il existe un chemin allant de i à j. Les composantes connexes du bigraphe sont précisément les classes d'équivalence de cette relation. Trivialement, chaque composante connexe est naturellement muni d'une structure de bigraphe. Par exemple le bigraphe associé à la forme unitaire $q(x) = \sum_{i=1}^5 x_i^2 - 2x_1x_2 + x_3x_4 - 3x_4x_5 - x_3x_5$ a deux composantes connexes. On dit qu'un bigraphe est *connexe* s'il n'a qu'une composante connexe. Un *sous bigraphe* B' d'un bigraphe B est le bigraphe obtenu en prenant un sous ensemble S des sommets et en gardant toutes les arêtes qui reliaient entre eux deux sommets de S. On dit aussi que le bigraphe B *contient* le bigraphe B'. Si le bigraphe n'a que des arêtes pleines, on parle plutôt de *graphe* et de *sous graphe* pour alléger.

Ensuite, une *racine* de la forme quadratique unitaire q est un élément x de $\mathbb{Z}^n$ tel que $q(x) = 1$. Il est clair que si $e = (e_1, \ldots, e_n)$ est la base canonique de $\mathbb{R}^n$ alors chaque e_i est une racine. Par exemple $q^{(1)}$ défini par (4.10) a six racines qui sont $(\pm 1, 0)$, $(0, \pm 1)$ et $\pm(1, 1)$, alors que $q^{(2)}$ en a une infinité. L'étude des racines de $x_1^2 + x_2^2 + x_3^3 - x_1x_2 - x_1x_3$ se fait par l'algorithme de Gauss et conduit à la discussion de l'équation $3a^2 + b^2 + 8c^2 = 12$ où les entiers a, b, c satisfont

$$a = 2x_1 - x_2 - x_3,\ b = 3x_2 - x_3,\ c = x_3.$$

On voit alors qu'ici il y a 8 racines : les $\pm e_i$ avec $i = 1, 2, 3$ et $\pm(1, 1, 1)$.

Remarquons maintenant que si $r = (r^1, \ldots, r^m)$ est une suite de m racines de q, on crée une nouvelle forme quadratique unitaire q_r, sur $\mathbb{Z}^m$ cette fois, définie par $q_r(y_1, \ldots, y_m) =$

$q(y_1r^1 + \cdots y_mr^m)$. En effet

$$\begin{aligned} q(y_1r^1 + \cdots y_mr^m) &= \sum_{i=1}^m q(y_ir^i) + \sum_{i<j} q(y_ir^i|y_jr^j) \\ &= \sum_{i=1}^m y_i^2 + \sum_{i<j} y_iy_jq(r^i|r^j) \end{aligned}$$

et on utilise le fait que $q(r^i|r^j)$ est un entier d'après (4.8).

V Graphes et formes quadratiques de Dynkin *

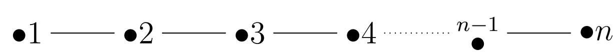
Une forme quadratique unitaire q est dite *positive* si $q(x) > 0$ pour tout $x \in \mathbb{Z}^n \setminus \{0\}$. C'est équivalent à dire que $\overline{q}$ est définie positive sur $\mathbb{R}^n$. C'est trivial dans un sens. Dans l'autre, si q est positive, on en déduit aisément que $\overline{q}(x) > 0$ pour tout $x \in \mathbf{Q}^n \setminus \{0\}$ (ici $\mathbf{Q}$ est le corps des rationnels). Par densité on en déduit que $\overline{q}$ est positive. Si $\overline{q}$ n'était pas définie positive, alors le sous espace de $\mathbb{R}^n$ égal à $\{x \in \mathbb{R}^n; \overline{q}(x) = 0\}$ est défini par des équations à coefficients *entiers* (en tant que noyau de la forme bilinéaire $(x, y) \mapsto q(x|y)$) Ce noyau contient donc des points à coordonnées rationnelles différents de 0 : contradiction. On déduit de cela que si q est positive alors elle n'a qu'un nombre fini de racines.

Nous définissons le *graphe de Dynkin* $G_{k,n}$ pour $k = 1, 2, 3$ et $n \geq 2k$ comme le graphe de sommets $\{1, \ldots, n\}$ et dont les arêtes sont :

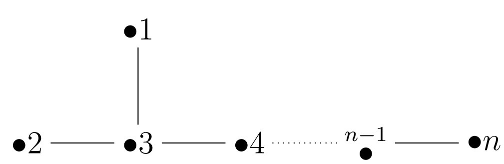
$$\{1, k+1\}, \{2, 3\}, \{3, 4\}, \{4, 5\}, \ldots \{n-1, n\},$$

avec la restriction $n = 6, 7, 8$ pour $k = 3$. Il est traditionnel d'appeler les graphes $G_{1,n}$ plutôt A_n, les $G_{2,n}$ plutôt D_n et les $G_{3,n}$ plutôt E_n. E_6, E_7 et E_8 sont dits exceptionnels. Par conséquent

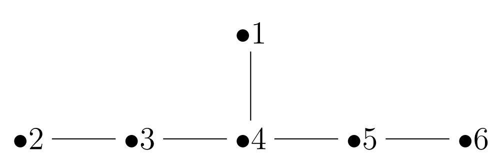
$G_{1,n} = A_n$



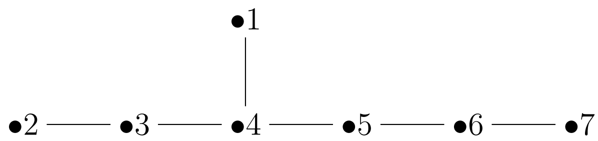
$G_{2,n} = D_n$



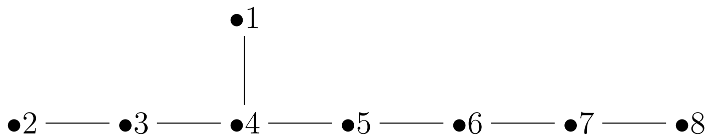
$G_{3,6} = E_6$



$G_{3,7} = E_7$



$G_{3,8} = E_8$



Proposition 5.1. Les graphes de Dynkin sont des bigraphes de formes quadratiques unitaires positives.

Démonstration. Pour A_n il faut montrer que la forme quadratique

$$q_{A_n}(x) = \sum_{i=1}^{n} x_i^2 - \sum_{i=2}^{n} x_{i-1}x_i$$

est définie positive, ou encore que la matrice symétrique M_{A_n} d'ordre n formée de 2 sur la diagonale, de -1 sur les deux diagonales voisines et de 0 ailleurs est définie positive. Le polynôme caractéristique P_n de M_{A_n} satisfait à la relation de récurrence

$$P_n(x) - (2-x)P_{n-1}(x) + P_{n-2}(x) = 0 \tag{5.11}$$

comme on le voit en développant le déterminant par rapport à la première ligne. De plus si on convient $P_0(x) = 1$ et $P_1(x) = 2 - x$, en posant $x = 2 - 2\cos\theta$ on découvre que

$$P_n(x) = \frac{\sin(n+1)\theta}{\sin\theta}.$$

Cela montre que les n racines de P_n sont $2 - 2\cos\frac{j\pi}{n+1}$ avec $j = 1, \ldots, n$. Donc elles sont dans $]0, 4[$ et donc la définie positivité de la matrice M_{A_n} est montrée.

Pour les autres graphes de Dynkin D_n, E_6, E_7, E_8 nous allons exploiter le calcul précédent en numérotant les sommets de la manière qui nous a servi à définir les graphes de Dynkin. Notons $M_{k,n}$ la matrice de la forme quadratique correspondante. Par exemple

$$M_{3,8} = M_{E_8} = \begin{bmatrix} 2 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 \end{bmatrix}$$

En développant par rapport à la première colonne on voit que

$$\begin{aligned}\det(M_{k,n}-xI_n) &= (2-x)\det(M_{A_{n-1}}-xI_{n-1})-\det(M_{A_{n-1-k}}-xI_{n-1-k})\det(M_{A_{k-1}}-xI_{k-1})\\ &= P_1(x)P_{n-1}(x)-P_{n-1-k}(x)P_{k-1}(x)\\ &= \frac{1}{\sin^2\theta}(\sin 2\theta\sin n\theta-\sin(n-k)\theta\sin k\theta),\end{aligned}$$

avec $2-x=2\cos\theta$ comme d'habitude. Pour $M_{D_n}=M_{2,n}$ on obtient

$$\det(M_{2,n}-xI_n)=\frac{2\cos\theta}{\sin\theta}(\sin n\theta-\sin(n-2)\theta)=\frac{1}{\sin\theta}(\sin(n+1)\theta-\sin(n-3)\theta)$$

ce qui permet facilement de calculer les racines de $\det(M_{2,n}-xI_n)$, de voir qu'elles sont dans $]0,4[$ et de montrer la définie positivité pour M_{D_n}.

Pour $k=3$ et $n=6,7,8$ il n'est pas facile de vérifier que l'équation

$$\sin 2\theta\sin n\theta-\sin(n-3)\theta\sin 3\theta=0$$

a n racines dans $]0,\pi[$[3]. Aussi prenons nous une méthode voisine pour montrer que M_{E_n} est définie positive. Du calcul précédent, en y faisant $\theta=0$ on tire que

$$\det M_{k,n}=2n-k(n-k)=(2-k)n+k^2.$$

On remarque d'ailleurs que ce nombre est positif si et seulement si (n,k) correspond aux graphes A_n,D_n,E_6,E_7,E_8 (en tenant compte du fait que (k,n) et $(n-k,n)$ donnent le même graphe).

De plus la matrice extraite de $M_{k,n}$ en supprimant les ligne et colonne 1 est $M_{A_{n-1}}$, qui est définie positive. Donc $M_{k,n}$ l'est aussi si et seulement si $\det M_{k,n}>0$. Cela est une conséquence immediate de la caractérisation de la définie positivité par les déterminants principaux (Chap 2, Th. 9.4 (3)).

On dit que $b=(b^1,\dots,b^n)$ est une base de $\mathbb{Z}^n$ si c'est une base de $\mathbb{R}^n$ formée d'éléments de $\mathbb{Z}^n$ telle que de plus si $P=[b^1,\dots,b^n]$ alors $\det P=\pm 1$. Ceci entraîne que P^{-1} est aussi à coefficients entiers. Nous sommes maintenant en position de montrer le remarquable résultat suivant, qui montre comment sont faites toutes les formes unitaires positives :

Théorème 5.2. Si la forme quadratique unitaire q sur $\mathbb{Z}^n$ est positive, alors il existe une base $b=(b^1,\dots,b^n)$ de $\mathbb{Z}^n$ formée de racines de q telle que les composantes connexes du bigraphe de q_b soient des graphes de Dynkin.

Démonstration. La base canonique e est une base de racines. On remarque d'abord que si b est une base de racines

$$0<q(b^i\pm b^j)=q(b^i)+q(b^j)\pm q(b^i|b^j)=2\pm q_{ij}.$$

[3] Par exemple, on peut montrer que les huit racines pour $\det(M_{3,8}-xI_8)$ sont données par $2-x=2\cos\theta$ où $30\theta/\pi$ parcourt les 8 entiers k premiers avec 30 avec $0\le k\le 30$ à savoir $\{1,7,11,13,17,19,23,29\}$.

Donc $q_{ij} = -1, 0, 1$ et le bigraphe de q_b n'a pas d'arêtes multiples quand b est une base de racines. A une telle base de racines, associons alors l'ensemble $E(b)$ formé des racines de q de la forme $\sum_{i=1} x_i b^i$ où les x_i sont des entiers entiers ≥ 0. Puisque q est positive, il n'y a qu'un nombre fini de racines et le le nombre $\pi(b)$ d'éléments de $E(b)$ est fini. De plus il existe donc une base de racines b qui maximise $\pi(b)$. Je dis que ce b maximal n'a pas de racines en pointillé dans son bigraphe. En effet, si $(q_b)_{ij} > 0$ c'est dire que $1 = q(b^i|b^j)$; donc en posant $b'^j = b^i - b^j$ on voit que b'^j est une racine (car $q(b'^j) = 2 - q(b^i|b^j) = 1$). Considérons alors la base de racines b' obtenue en remplacant dans b la racine b^j par la racine b'^j. Alors $E(b') \supset E(b)$. De plus $b'^j \in E(b') \setminus E(b)$. Donc $\pi(b') > \pi(b)$, ce qui contredit le fait que b est maximal.

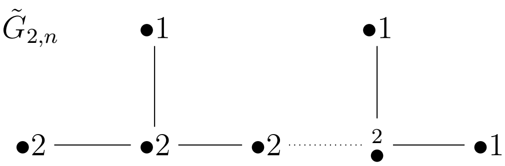
On est donc maintenant assuré qu'il existe une base de racines telle que son bigraphe n'a que des arêtes à la fois simples et pleines.

On montre alors le lemme suivant :

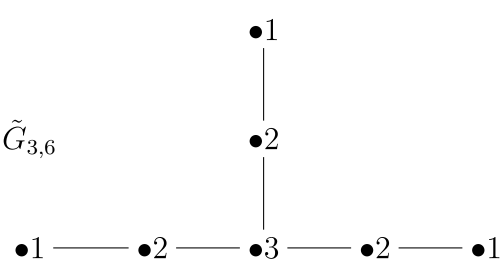
Lemme 5.3. Si un bigraphe connexe a arêtes simples a un graphe de Dynkin comme sous graphe et n'est pas lui même un graphe de Dynkin, alors la forme quadratique unitaire qui lui est associée n'est pas positive.

Démonstration du Lemme 5.3. Considérons les bigraphes suivants $\tilde{G}_{k,n}$ à arêtes pleines et simples, parfois appelés *graphes-étendus de Dynkin.* Les entiers à coté de chaque sommet ne sont pas le numéro du sommet mais les valeurs d'une certaine fonction δ définie sur les sommets de $\tilde{G}_{k,n}$. On remarque que $\tilde{G}_{k,n}$ a $n+1$ sommets. On n'a pas représenté le graphe $\tilde{G}_{1,n}$: il a $n+1$ sommets, disons $0, 1, \ldots, n$ et pour arêtes les $\{i-1, i\}$ avec $i = 1, \ldots, n$ plus l'arête $\{0, n\}$: c'est avec un polygône à $n+1$ cotés qu'on pourrait commodément dessiner ce graphe. La fonction δ est prise égale à 1 pour chaque sommet de $\tilde{G}_{1,n}$.

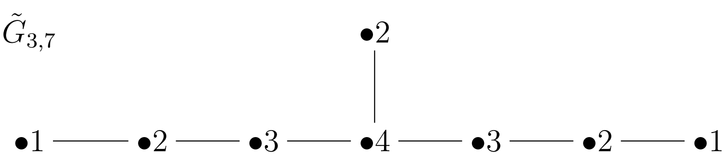
$\tilde{G}_{2,n}$



$\tilde{G}_{3,6}$



$\tilde{G}_{3,7}$



$\tilde{G}_{3,8}$

$$\begin{tikzcd}
 & & \bullet 3 \arrow[d, no head] & & & & & \\
\bullet 2 \arrow[r, no head] & \bullet 4 \arrow[r, no head] & \bullet 6 \arrow[r, no head] & \bullet 5 \arrow[r, no head] & \bullet 4 \arrow[r, no head] & \bullet 3 \arrow[r, no head] & \bullet 2 \arrow[r, no head] & \bullet 1
\end{tikzcd}$$

On constate avec patience que si un bigraphe satisfait à l'hypothèse du lemme, alors il contient comme sous graphe un graphe-étendu de Dynkin. Considérons alors la forme quadratique unitaire q associée à ce graphe-étendu de Dynkin et calculons $q(\delta)$. On trouve dans les 5 cas que $q(\delta) = 0$, ce qui montre que q n'est pas positive et achève la démonstration du lemme.

Fin de la démonstration du Théorème 5.2. Comme la forme quadratique unitaire associée est positive, on peut alors appliquer le Lemme 5.3 et la démonstration est achevée.

Remarque. On peut voir ce Théorème 5.2 sous un autre angle. Convenons de dire que les formes quadratiques unitaires q et q_1 sont *équivalentes* s'il existe une base b de $\mathbb{Z}^n$ formée de racines de q telle que $q_1 = q_b$. Si $P = [b^1, \ldots, b^n]$ c'est dire que les matrices représentatives M_q et M_{q_1} sont liées par $M_{q_1} = PM_qP^T$. Le fait que P^{-1} soit à coefficients entiers permet de voir que c'est bien une relation d'équivalence. Le Théorème 5.2 montre donc que *toute forme quadratique unitaire positive est équivalente à une somme directe de formes de Dynkin.*

Chapitre 5

Géométrie euclidienne affine

Ce chapitre est le plus concret de tous puisqu'il traite de géométrie élémentaire, celle du collège et du lycée, vue avec les outils déjà rassemblés. Nous commençons par définir l'espace affine, un outil qui connait des fortunes diverses suivant les modes. C'est en gros un espace vectoriel dans lequel l'origine perd toute importance. Si de plus l'espace vectoriel est de dimension 3 et est euclidien, c'est notre espace physique newtonien, dans lequel il n'y a *a priori* ni coordonnées ni origine, mais dans lequel parallélisme, orthogonalité, distance entre deux points et angles ont du sens.

I Espaces et variétés affines, barycentre et parallélisme.

Soit K un corps quelconque. Un espace affine est la donnée d'un ensemble $\mathcal{A}$ dont les éléments sont appelés points et les éléments de $\mathcal{A} \times \mathcal{A}$ sont appelés bipoints, d'un espace vectoriel E sur K appelé espace vectoriel associé et d'une application de $\mathcal{A} \times \mathcal{A}$ dans E notée $(A, B) \mapsto \overrightarrow{AB}$ qui satisfait aux axiomes suivants

1. Pour tout $A \in \mathcal{A}$ l'application de $\mathcal{A}$ dans E définie par $B \mapsto \overrightarrow{AB}$ est bijective.
2. Pour tous $A, B, C \in \mathcal{A}$ on a la relation de Chasles

$$\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}.$$

Remarques.

1. Compte tenu de ce qui a été appris, où l'éducation universitaire du lecteur a été faite d'abord en définissant un corps puis un espace vectoriel, si l'espace vectoriel E est donné, une manière de fabriquer un espace affine $\mathcal{A}$ d'espace vectoriel associé E est de prendre $\mathcal{A} = E$ en définissant pour deux vecteurs x et y de E l'élément $\overrightarrow{xy} = y - x$. Cette application de $E \times E$ dans E satisfait clairement aux deux axiomes : si x est fixé alors quel que soit $z \in E$ il existe un et un seul $y \in E$ tel que $z = y - x$, qui est $y = z + x$. Quant à $\overrightarrow{xz} = \overrightarrow{xy} + \overrightarrow{yz}$ c'est dire la chose évidente $(z - x) = (y - x) + (z - y)$.

2. Inversement choisissons dans l'espace affine $\mathcal{A}$ un point arbitraire A : la correspondance $B \mapsto \overrightarrow{AB}$ entre $\mathcal{A}$ et E permet de voir alors $\mathcal{A}$ comme un espace vectoriel d'origine A.
3. Bien entendu $\overrightarrow{AA} = 0$ et $\overrightarrow{AB} = -\overrightarrow{BA}$ comme on le déduit de la relation de Chasles.
4. Soit $x \in E$. Alors d'après l'axiome 1 pour tout point A il existe un unique B tel que $\overrightarrow{AB} = x$. Nous notons $B = \tau_x(A)$ et nous l'appelons le *translaté* du point A par le vecteur x.
5. La dimension de E est appelée la dimension de l'espace affine $\mathcal{A}$. Elle n'est pas nécessairement finie.

Le prochain concept important est celui de barycentre.

Proposition 1.1. Soit $\mathcal{A}$ un espace affine, soit $A_1, \dots, A_n$ des points de $\mathcal{A}$ pas nécessairement distincts et soit la suite de scalaires $(\lambda_1, \dots, \lambda_n)$ dans K^n telle que $\lambda_1 + \cdots + \lambda_n = 1$. Alors

1. (Existence) Il existe un unique point $B \in \mathcal{A}$ appelé barycentre des $A_1, \dots, A_n$ pour les poids $\lambda_1, \dots, \lambda_n$ tel que pour tout point $O \in \mathcal{A}$ on ait
$$\overrightarrow{OB} = \lambda_1 \overrightarrow{OA_1} + \cdots + \lambda_n \overrightarrow{OA_n}. \tag{1.1}$$
2. (Propriété d'associativité) Si $1 \leq k < n$ est tel que $p = \lambda_1 + \cdots + \lambda_k$ et $p' = 1 - p$ sont non nuls, soit B_1 et B_2 les barycentres de $\lambda_1, \dots, \lambda_k$ et $\lambda_{k+1}, \dots, \lambda_n$ pour les poids respectifs $\frac{1}{p}\lambda_1, \dots, \frac{1}{p}\lambda_k$ et $\frac{1}{p'}\lambda_{k+1}, \dots, \frac{1}{p'}\lambda_n$. Alors B est barycentre de (B_1, B_2) pour les poids (p, p').
3. (Repère affine) Si E est de dimension finie q et si $A_0, \dots, A_q$ sont tels que $\overrightarrow{A_0A_i}$ $i = 1, \dots, q$ est une base de E alors tout point B de $\mathcal{A}$ est barycentre de $A_0, \dots, A_q$ pour un système unique de poids.

Démonstration. Soit un point O fixé. D'après l'axiome 1 il existe certainement un point $B_O \in \mathcal{A}$ satisfaisant 1.1. Il s'agit de montrer que B_0 ne dépend pas de O. Si P est un autre point alors par la relation de Chasles et le fait que $\lambda_1 + \cdots + \lambda_n = 1$ on a

$$\begin{aligned} \overrightarrow{PB_O} &= \overrightarrow{PO} + \overrightarrow{OB_0} \\ &= \lambda_1(\overrightarrow{PO} + \overrightarrow{0A_1}) + \cdots + \lambda_n(\overrightarrow{PO} + \overrightarrow{OA_n}) \\ &= \overrightarrow{PB_P} \end{aligned}$$

ce qui par l'axiome 1 entraîne $B_0 = B_P$. La seconde partie est à peu près évidente. Pour la troisième on définit l'unique système de scalaires $(\lambda_1, \dots, \lambda_q)$ par

$$\overrightarrow{A_0B} = \lambda_1 \overrightarrow{A_0A_1} + \ldots + \lambda_q \overrightarrow{A_0A_q}$$

et on définit enfin $\lambda_0 = 1 - \lambda_1 + \cdots + \lambda_q$.

Remarques :

1. On note par convention le barycentre B de $A_1, \ldots, A_n$ pour les poids $\lambda_1, \ldots, \lambda_n$ par

$$B = \lambda_1 A_1 + \cdots + \lambda_n A_n.$$

   C'est une sténographie commode pour dire 1.1, puisque B ne dépend pas de O. Toutefois il faut se rappeler que ni la multiplication du point A_1 par le scalaire λ_1 ni la somme de points n'ont de sens dans l'espace affine.

2. Quand les poids sont tous égaux on parle d'*isobarycentre.* L'isobarycentre de (A_1, A_2) est donc calculé avec $(\lambda_1, \lambda_2) = (1/2, 1/2)$. Il est appelé le *milieu* du bipoint (A_1, A_2).

3. Si E est de dimension finie q et si $A_0, \ldots, A_q$ sont tels que $\overrightarrow{A_0A_i}$ $i = 1, \ldots, q$ est une base de E, alors $(A_0, \ldots, A_q)$ est appelé un repère affine. L'unique système $(\lambda_0, \ldots, \lambda_q)$ de scalaires tel que $\lambda_0 + \cdots + \lambda_q = 1$ et $B = \lambda_0 A_0 + \cdots + \lambda_q A_q$ s'appelle les *coordonnées barycentriques* de B. Il faut mentionner qu'alors que pour un $0 \leq j \leq q$ fixé, alors $\overrightarrow{A_jA_i}$ $i \neq j$ est aussi une base de E : Par exemple pour $j = q$ si c'était faux il existerait une suite non nulle de scalaires $(\lambda_0, \ldots, \lambda_{q-1})$ telle que

$$0 = \sum_{i=0}^{q-1} \lambda_i \overrightarrow{A_qA_i} = \sum_{i=0}^{q-1} \lambda_i (\overrightarrow{A_qA_0} + \overrightarrow{A_0A_i}) = -(\sum_{i=0}^{q-1} \lambda_i)\overrightarrow{A_0A_q} + \sum_{i=0}^{q-1} \lambda_i \overrightarrow{A_0A_i}.$$

   Comme $\overrightarrow{A_0A_i}$ $i = 1, \ldots, q$ est une base cela entraîne que $0 = \lambda_1 = \ldots = \lambda_{q-1} = -(\sum_{i=0}^{q-1} \lambda_i)$, et on tire $\lambda_0 = 0$ de la dernière égalité. D'où la contradiction.

Définition. Soit $\mathcal{A}$ un espace affine. Une variété affine de $\mathcal{A}$ est une partie $\mathcal{V}$de $\mathcal{A}$ telle que pour tout n et pour toute famille $A_1, \ldots, A_n$ de $\mathcal{V}$ alors tous les barycentres possibles de $A_1, \ldots, A_n$ sont dans $\mathcal{V}$. En particulier l'ensemble vide est une variété affine.

Proposition 1.2. Soit $\mathcal{A}$ un espace affine associé à un espace vectoriel E de dimension finie q.

1. Si $\emptyset \neq \mathcal{V} \subset \mathcal{A}$ alors V est une variété affine si et seulement si il existe un sous espace vectoriel $E_\mathcal{V}$ de E tel que pour tout $A, B \in V$ on a $\overrightarrow{AB} \in E_\mathcal{V}$. Un tel $E_\mathcal{V}$ est unique et est appelé la *direction* de $\mathcal{V}$ et $\mathcal{V}$ est un espace affine associé à $E_\mathcal{V}$.
2. L'intersection de deux variétés affines est une variété affine.

Démonstration. 1) $\Rightarrow$. Puisque $\mathcal{V}$ est non vide fixons $A \in \mathcal{V}$. Je dis que

$$E_\mathcal{V} = \{\overrightarrow{AB};\ B \in \mathcal{V}\}$$

est un sous espace vectoriel de E. En effet si B_1 et B_2 sont dans $\mathcal{V}$ alors le barycentre $B = \lambda_1 B_1 + \lambda_2 B_2 + (1 - \lambda_1 - \lambda_2)A$ est dans $\mathcal{V}$ (pour les poids $\lambda_1, \lambda_2, 1 - \lambda_1 - \lambda_2$ donc). En prenant $O = A$ dans 1.1 c'est dire que $\overrightarrow{AB} = \lambda_1 \overrightarrow{AB_1} + \lambda_2 \overrightarrow{AB_2}$ et donc $E_\mathcal{V}$ est un sous espace vectoriel de E. Reste à vérifier que $E_\mathcal{V}$ ne dépend pas du point A particulier choisi. Si $E'_\mathcal{V} = \{\overrightarrow{A'B};\ B \in \mathcal{V}\}$ alors $\overrightarrow{A'B} = \overrightarrow{AB} - \overrightarrow{AA'} \in E_\mathcal{V}$ et donc $E'_\mathcal{V} \subset E_\mathcal{V}$. Par symétrie $E_\mathcal{V} \subset E'_\mathcal{V}$. et on a le résultat.

La réciproque de 1) et l'unicité et le 2) sont faciles.

Définitions. La dimension d'une variété affine non vide est la dimension de sa direction. On convient de dire que -1 est la dimension de la variété vide.

Deux variétés affines non vides $\mathcal{V}_1$ et $\mathcal{V}_2$ de l'espace affine $\mathcal{A}$ sont dites parallèles elles sont disjointes et si $E_{\mathcal{V}_1} \subset E_{\mathcal{V}_2}$ ou si $E_{\mathcal{V}_2} \subset E_{\mathcal{V}_1}$

Proposition 1.3. Soit deux variétés affines non vides $\mathcal{V}_1$ et $\mathcal{V}_2$ de l'espace affine $\mathcal{A}$ telles que

$$\dim \mathcal{V}_1 + \dim \mathcal{V}_2 \geq \dim \mathcal{A}.$$

Alors elles sont parallèles si elles sont disjointes.

Démonstration. Rappelons la formule de première année

$$\dim E_{\mathcal{V}_1} + \dim E_{\mathcal{V}_2} = \dim E_{\mathcal{V}_1} \cap E_{\mathcal{V}_2} + \dim(E_{\mathcal{V}_1} + E_{\mathcal{V}_2}).$$

II Espace affine euclidien. Distance entre deux sous espaces

III Angles

IV Polyèdres réguliers et sous groupes finis de $\mathbf{SO}(3)$

V Coniques et quadriques de l'espace euclidien

VI Homographie et inversion