Cours d'algèbre 3

Marc Pauly

Chapitre 1. Vecteurs propres et valeurs propres

Définition. *Soit f un endomorphisme d'un espace vectoriel V sur un corps commutatif K. On dit que $\lambda \in K$ est* **une valeur propre de** *f s'il existe un vecteur non nul $v \in V$ tel que*

$$f(v) = \lambda \cdot v$$

On appelle **vecteur propre de valeur propre** *λ tout vecteur v de V tel que*

$$f(v) = \lambda \cdot v$$

Exemples. 1. Soit V un espace vectoriel sur K non réduit à $\{0\}$. Si f désigne l'identité sur V, $\lambda = 1$ est une valeur propre, et tout vecteur v de V est vecteur propre de valeur propre 1.

2. Soit V un espace vectoriel sur K non réduit à $\{0\}$. Si f désigne l'endomorphisme nul sur V, $\lambda = 0$ est une valeur propre, et tout vecteur v de V est vecteur propre de valeur propre 0.

3. On considère $p : K^3 \to K^3 : (x, y, z) \mapsto (x, 0, 0)$. $\lambda = 1$ est une valeur propre, car par exemple

$$f(1, 0, 0) = 1 \cdot (1, 0, 0)$$

De même, $\lambda = 0$ est aussi une valeur propre, car

$$f(0, 1, 0) = 0 \cdot (0, 1, 0)$$

L'ensemble des vecteurs propres de valeur propre 1 est $\{(x, 0, 0)|x \in K\}$.
L'ensemble des vecteurs propres de valeur propre 0 est $\{(0, y, z)|y, z \in K\}$.

Théorème. *Soit f un endomorphisme de V.*
1. Si λ est une valeur propre de f, alors l'ensemble de tous les vecteurs propres de valeur propre λ est un sous-espace vectoriel de V.
2. Soit v un vecteur non nul de V. Le vecteur v est un vecteur propre de f si et seulement si $f(Vect(\{v\})) \subset Vect(\{v\})$.

Preuve. 1. $f(v) = \lambda v$ équivaut à $(f - \lambda \cdot id_V)(v) = 0$ ou encore

$$v \in \ker(f - \lambda \cdot id)$$

Or $f - \lambda \cdot id$ est un endomorphisme de V. Donc son noyau est un sous-espace vectoriel de V.

2. «seulement si» (condition nécessaire) :
On suppose que $f(v) = \lambda v$ pour un certain scalaire λ. Or $Vect(\{v\})$ est l'ensemble Kv de tous les multiples de v. Si αv est un multiple quelconque de v, on voit par linéarité de f que

$$f(\alpha v) = \alpha f(v) = \alpha(\lambda v) = (\alpha\lambda)v$$

Donc $f(\alpha v)$ reste bien un multiple de v.
«si» (condition suffisante) : Comme v est un multiple de v, on sait par hypothèse que $f(v)$ est un multiple de v. Donc il existe $\lambda \in K$ avec

$$f(v) = \lambda v$$

Comme v est non nul, on peut donc dire que λ est une valeur propre de f, et que v est un vecteur propre de valeur propre λ. □

Définition. *Soit f un endomorphisme de V, et λ une valeur propre de f. On appelle* **sous-espace propre de** *λ le sous-espace vectoriel de V de tous les vecteurs propres de valeur propre λ. On le note $E_\lambda(f)$.*

On remarquera que pour une valeur propre λ de f, l'espace propre $E_\lambda(f)$ n'est jamais réduit à l'espace $\{0\}$.

Observons aussi le lien étroit existant entre la valeur propre 0 et le noyau d'un endomorphisme : $\lambda = 0$ est valeur propre de f si et seulement si le noyau de f n'est pas réduit au vecteur nul. Dans ce cas, l'espace propre associé E_0 est le noyau de f.

Théorème. *Soit f un endomorphisme d'un espace vectoriel V de dimension finie. Le scalaire λ est une valeur propre de f si et seulement si* $\det(f - \lambda \cdot id) = 0$.

Preuve. λ est valeur propre de f si et seulement si le noyau de $f - \lambda \cdot id$ n'est pas réduit à $\{0\}$. Cela équivaut à ce que l'endomorphisme $f - \lambda \cdot id$ ne soit pas injectif. Or, en dimension finie, un endomorphisme est injectif si et seulement si son déterminant est non nul. Par contraposée, on trouve le résultat. □

Théorème. *Soient f un endomorphisme de V, et g un endomorphisme de W. On suppose qu'il existe un isomorphisme $\phi : V \to W$ tel que $f = \phi^{-1} \circ g \circ \phi$. Alors f et g ont les mêmes valeurs propres. En particulier, deux endomorphismes semblables d'un espace vectoriel V ont les mêmes valeur propres.*

Preuve. Il suffit de montrer que toute valeur propre de f est aussi une valeur propre de g. En effet, pour l'autre implication, il suffit d'échanger le rôle de f et g.

Soit λ une valeur propre de f. Alors il existe v non nul avec

$$f(v) = \lambda v$$

Il vient alors

$$\phi^{-1}(g(\phi(v))) = \lambda v$$

et donc

$$g(\phi(v)) = \phi(\lambda v) = \lambda \phi(v)$$

Or $\phi(v)$ est non nul, car v est non nul et ϕ est injective. Donc $\phi(v)$ est un vecteur propre de valeur propre λ de g. □

Définition. *Soit f un endomorphisme d'un espace vectoriel V. On appelle* **spectre de** *f l'ensemble des valeurs propres de f. Le spectre de f est donc un sous-ensemble de K. On le note $Spec(f)$.*

Définition. *Soit A une matrice carrée $n \times n$ à coefficients dans K. On dit que λ est une* **valeur propre de** *A s'il existe une matrice colonne X non nulle à n lignes avec*

$$AX = \lambda X$$

Si λ est une valeur propre, on appelle vecteur propre de λ toute matrice-colonne $X \in \mathcal{M}_{n,1}(K)$ telle que $AX = \lambda X$. **Le spectre de** *A, noté $Spec(A)$, est l'ensemble des valeurs propres de A.*

Comme pour les endomorphismes, on montre que l'ensemble de tous les vecteurs propres de la matrice A, associés à une valeur propre fixée λ, est un espace vectoriel. C'est en l'occurrence un sous-espace vectoriel (de dimension non nulle) de $\mathcal{M}_{n,1}(K)$. On l'appelle **espace propre de** λ (pour la matrice A). On a aussi un critère similaire pour décider si $\lambda \in K$ est une valeur propre d'une matrice A :

Théorème. *Soit A une matrice $n \times n$. Un scalaire λ est une valeur propre de A si et seulement si*

$$\det(A - \lambda I_n) = 0$$

Si λ est valeur propre de A, l'espace propre de λ est l'espace des matrices-colonnes $X \in \mathcal{M}_{n,1}(K)$ qui vérifient l'équation matricielle $(A - \lambda I_n)X = 0$.

Preuve. Évidente □

Toujours par analogie avec la théorie pour les endomorphismes, nous avons le résultat sur les valeurs propres de deux matrices semblables.

Théorème. *Soient A, B deux matrices semblables de taille $n \times n$, c'est-à-dire il existe une matrice $P \in GL(n, K)$ telle que $A = P^{-1}BP$. Alors A et B ont les mêmes valeurs propres. Autrement dit, leurs spectres sont égaux.*

Preuve. Soit λ un scalaire quelconque. On a un enchaînement d'équivalence : λ est valeur propre de A si et seulement si $\det(A - \lambda I_n) = 0$. Or

$$\det(A - \lambda I_n) = \det(P^{-1}(B - \lambda I_n)P) = (\det P)^{-1} \det(B - \lambda I_n) \det P = \det(B - \lambda I_n)$$

La condition $\det(A - \lambda I_n) = 0$ équivaut alors à $\det(B - \lambda I_n) = 0$, ce qui exprime de manière équivalente que λ est valeur propre de B. □

Exemple. Cherchons les valeurs propres et vecteurs propres de la matrice carrée réelle

$$A = \begin{pmatrix} 2 & -1 & 5 \\ 4 & -3 & 5 \\ 4 & 1 & 1 \end{pmatrix}$$

Le réel λ est valeur propre de A si et seulement si

$$\det(A - \lambda I_3) = 0$$

c'est-à-dire

$$\det \begin{pmatrix} 2-\lambda & -1 & 5 \\ 4 & -3-\lambda & 5 \\ 4 & 1 & 1-\lambda \end{pmatrix} = 0$$

Un calcul montre que le déterminant est égal à $-\lambda^3 + 28\lambda + 48$. Il faut par conséquent résoudre l'équation **polynomiale** d'inconnue λ :

$$\lambda^3 - 28\lambda - 48 = 0$$

Avec un brin de chance, nous observons que $\lambda = -2$ est une racine de ce polynôme. Or

$$\lambda^3 - 28\lambda - 48 = (\lambda + 2)(\lambda^2 - 2\lambda - 24) = (\lambda + 2)(\lambda + 4)(\lambda - 6)$$

Nous avons trouvé trois valeurs propres. Le spectre de la matrice A est l'ensemble $\{-2, -4, 6\}$.

Il nous reste à trouver les vecteurs propres.

Le cas de la valeur propre $\lambda = -2$. Résolvons l'équation matricielle

$$\begin{pmatrix} 2-(-2) & -1 & 5 \\ 4 & -3-(-2) & 5 \\ 4 & 1 & 1-(-2) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Cela débouche sur le système

$$\begin{cases} 4x - y + 5z = 0 \\ 4x - y + 5z = 0 \\ 4x + y + 3z = 0 \end{cases}$$

Les première et deuxième équations sont identiques. On supprime l'une d'entre elles, et on trouve

$$\begin{cases} x = -t \\ y = t \\ z = t \end{cases}$$

avec t un réel quelconque. L'espace propre E_{-2} est de dimension 1, et il est engendré par le vecteur $(-1, 1, 1)$.

Le cas de la valeur propre $\lambda = -4$. Résolvons l'équation matricielle

$$\begin{pmatrix} 2-(-4) & -1 & 5 \\ 4 & -3-(-4) & 5 \\ 4 & 1 & 1-(-4) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Cela débouche sur le système

$$\begin{cases} 6x - y + 5z = 0 \\ 4x + y + 5z = 0 \\ 4x + y + 5z = 0 \end{cases}$$

Les deuxième et troisième équations sont identiques. On supprime l'une d'entre elles, et on trouve

$$\begin{cases} x = t \\ y = t \\ z = -t \end{cases}$$

avec t un réel quelconque. L'espace propre E_{-4} est de dimension 1, et il est engendré par le vecteur $(1, 1, -1)$.

Le cas de la valeur propre $\lambda = 6$. Résolvons l'équation matricielle

$$\begin{pmatrix} 2-6 & -1 & 5 \\ 4 & -3-6 & 5 \\ 4 & 1 & 1-6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Cela débouche sur le système

$$\begin{cases} -4x - y + 5z = 0 \\ 4x - 9y + 5z = 0 \\ 4x + y - 5z = 0 \end{cases}$$

Les première et troisième équations sont identiques. On supprime l'une d'entre elles, et on trouve

$$\begin{cases} x = t \\ y = t \\ z = t \end{cases}$$

avec t un réel quelconque. L'espace propre E_6 est de dimension 1, et il est engendré par le vecteur $(1, 1, 1)$.

***

Nous allons maintenant nous intéresser aux liens qui existent entre valeurs propres/vecteurs propres d'endomorphismes en dimension finie et les vecteurs propres/valeurs propres de matrices carrées. Nous résumons toutes ces propriétés dans le théorème suivant.

Théorème. *Soit u un endomorphisme d'un espace vectoriel E de dimension finie sur K. On considère $e = (e_1, \ldots, e_n)$ une base quelconque de E.*
1. Le spectre de u est égal au spectre de la matrice $\mathcal{M}_e(u)$.
2. Soit λ une valeur propre de u (et donc de $\mathcal{M}_e(u)$). Alors il existe un isomorphisme canonique entre $E_\lambda(\mathcal{M}_e(u))$ et $E_\lambda(u)$:

$$E_\lambda(\mathcal{M}_e(u)) \to E_\lambda(u) : \qquad \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \mapsto \sum_{j=1}^{n} X_j e_j$$

Preuve.
1. Il suffit de comprendre que la condition $\det(u - \lambda id) = 0$ est équivalente à $\det(\mathcal{M}_e(u) - \lambda I_n) = 0$. Pour le voir, il suffit d'observer que

$$\mathcal{M}_e(u - \lambda id) = \mathcal{M}_e(u) - \lambda I_n$$

et de se souvenir que pour tout endomorphisme v, on a $\det v = \det \mathcal{M}_e(v)$.
2. L'application décrite est évidemment un isomorphisme linéaire de $\mathcal{M}_{n,1}(K)$ vers E, puisque e est une base de E. Pour montrer que l'application $\phi : E_\lambda(\mathcal{M}_e(u)) \to E_\lambda(u)$ est un isomorphisme, il suffit de montrer que les phrases «$X \in E_\lambda(\mathcal{M}_e(u))$» et «$\phi(X) \in E_\lambda(u)$» sont équivalentes :
Or $X \in E_\lambda(\mathcal{M}_e(u))$ si et seulement si $\mathcal{M}_e(u)X = \lambda X$, si et seulement si $\phi(\mathcal{M}_e(u)X) = \phi(\lambda X)$. En effet, ϕ est une bijection. Or c'est aussi une application linéaire, et on trouve

$$\sum_{j=1}^{n} (\mathcal{M}_e(u) \cdot X)_j e_j = \lambda \phi(X)$$

En développant et réarrangeant le produit matriciel dans le membre de gauche, cela peut aussi s'écrire sous la forme

$$\sum_{i=1}^{n} X_i \left(\sum_{j=1}^{n} \left(\mathcal{M}_e(u)\right)_{ji} e_j \right) = \lambda \phi(X)$$

La parenthèse intérieure est par définition $u(e_i)$. On trouve alors $u(\sum_{i=1}^{n} X_i e_i) = \lambda \phi(X)$ et enfin $u(\phi(X)) = \lambda \phi(X)$. Cela veut bien dire que $\Phi(X) \in E_\lambda(u)$. □

Corollaire. *Soit A une matrice $n \times n$ à coefficients dans un corps commutatif K. On appelle u_A l'endomorphisme de $\mathcal{M}_{n,1}(K)$ défini par la règle*

$$\forall X \in \mathcal{M}_{n,1}(K), \quad u_A(X) = AX$$

Alors A et u_A ont même spectre, et pour toute valeur propre λ commune, les espaces propres de λ pour A et u_A sont égaux.

Preuve. Il suffit d'observer que la matrice de u_A relativement à la base canonique de $\mathcal{M}_{n,1}(K)$ est la matrice A. □

Nous pouvons désormais introduire un outil important de l'algèbre linéaire, à savoir le polynôme caractéristique d'une matrice respectivement d'un endomorphisme en dimension finie. Nous commençons par définir le polynôme caractéristique d'une matrice.

Théorème. *Soit A une matrice $n \times n$ à coefficients dans K. On considère la matrice $XI_n - A$, à coefficients dans l'anneau des polynômes $K[X]$. Alors le déterminant de cette matrice est un polynôme unitaire de degré n.*

Preuve. Soit $B = XI_n - A$. Nous avons

$$\det B = \sum_{\sigma \in Sym(n)} \epsilon(\sigma) B_{\sigma(1)1} \cdots B_{\sigma(n)n}$$

Or chaque B_{ij} est un polynôme en X de degré 1 (si $i = j$) ou de degré 0 (si $i \neq j$).
Donc, pour toute permutation σ, la quantité $\epsilon(\sigma) B_{\sigma(1)1} \cdots B_{\sigma(n)n}$ est un polynôme en X de degré au plus n, et de même pour $\sum_{\sigma \in Sym(n)} \epsilon(\sigma) B_{\sigma(1)1} \cdots B_{\sigma(n)n}$.

Ceci prouve déjà que $\det B$ est un polynôme de degré au plus n.

Il reste à montrer que le coefficient de X^n vaut 1.
Il suffit de voir que, parmi les $n!$ termes de la somme $\sum_\sigma$, un seul est de degré égal à n, c'est celui qui correspond au cas où σ est l'identité. Pour $\sigma = id$, ce terme vaut

$$\epsilon(id)(X - A_{11})(X - A_{22}) \cdots (X - A_{nn})$$

La signature de l'identité vaut $+1$, il s'agit bien d'un polynôme unitaire. □

Définition. *Soit A une matrice $n \times n$ à coefficients dans K.* **Le polynôme caractéristique de** *A est le polynôme* $\det(XI_n - A)$.

Exemple. Le polynôme caractéristique de la matrice réelle $A = \begin{pmatrix} 2 & -1 & 5 \\ 4 & -3 & 5 \\ 4 & 1 & 1 \end{pmatrix}$ est

$$X^3 - 28X - 48 \in \mathbb{R}[X]$$

Remarque. Si deux matrices sont semblables, elles ont même polynôme caractéristique. C'est essentiellement l'argument qui montre que de telles matrices ont même spectre.

Définition. *Soit f un endomorphisme d'un espace vectoriel V de dimension finie. Le* **polynôme caractéristique de** *f est le polynôme caractéristique de la matrice $\mathcal{M}_e(f)$, où e est une base quelconque de V.*

Il faut se convaincre que c'est bien indépendant du choix de e. Or un changement de base transforme $\mathcal{M}_e(f)$ en une matrice semblable. Donc les polynômes caractéristiques de ces matrices sont égaux, ce qui assure que notre définition a bien du sens.

Le résultat suivant est un corollaire évident de ce que nous avons démontré plus haut.

Théorème. *Soit A une matrice carrée à coefficients dans un corps commutatif K. Le scalaire $\lambda \in K$ est valeur propre de la matrice A si et seulement si λ est racine du polynôme caractéristique de A.*
Soit f un endomorphisme d'un espace vectoriel de dimension finie sur K. Le scalaire $\lambda \in K$ est valeur propre de f si et seulement si λ est racine du polynôme caractéristique de f.
Autrement dit, le spectre de la matrice A est l'ensemble des racines du polynôme caractéristique de A, et le spectre de l'endomorphisme f est l'ensemble des racines du polynôme caractéristique de f.

Définition. *Soit f un endomorphisme de V (de dimension finie), et g un endomorphisme de W (de dimension finie) sur le même corps K. On dit que f et g sont* **semblables** *s'il existe un isomorphisme $\phi : V \to W$ tel que $f = \phi^{-1} \circ g \circ \phi$.*
On dit que deux matrices carrées A, B de taille $n \times n$ à coefficients dans K sont **semblables** *s'il existe une matrice inversible $P \in GL(n, K)$ telle que $A = P^{-1}BP$.*

Remarquons que deux matrices semblables sont toujours équivalentes, mais que la réciproque est fausse.

Rappelons le résultat principal : Le polynôme caractéristique est **invariant par similitude**.

Théorème.
1. Si deux endomorphismes sont semblables, alors ils ont le même polynôme caractéristique.
2. Si deux matrices sont semblables, alors elles ont le même polynôme caractéristique.

Le fait qu'un polynôme de degré n possède au plus n racines a un corollaire intéressant sur le spectre d'une matrice ou d'un endomorphisme en dimension finie :

Théorème.
1. Soit f un endomorphisme d'un espace vectoriel de dimension n. Alors f possède au plus n valeurs propres.
2. Soit A une matrice carrée $n \times n$. Alors A possède au plus n valeurs propres.

Preuve. 1. Les valeurs propres sont exactement les racines du polynôme caractéristique. Comme il est de degré n, il possède au plus n racines distinctes dans K. C'est fini.
2. De même. □

Le prochain résultat établit qu'une famille de vecteurs propres non nuls et de valeurs propres distinctes est toujours libre.

Théorème. *Soit f un endomorphisme d'un espace vectoriel V sur le corps K (pas nécessairement de dimension finie). On prend un sous-ensemble $I \subset Spec(f)$, et on se donne une famille de vecteurs*

$$(v_\lambda)_{\lambda \in I}$$

indexée par l'ensemble I. On suppose que

$$\forall \lambda \in I, v_\lambda \in E_\lambda - \{0\}$$

Autrement dit, dans cette famille de vecteurs, il y a seulement des vecteurs propres non nuls, et les valeurs propres associées sont deux à deux distinctes.

Alors la famille $(v_\lambda)_{\lambda \in I}$ est une famille libre.

Preuve. Pour montrer qu'une famille de vecteurs est libre, il suffit de montrer que toute sous-famille finie est libre. Nous allons démontrer cela en faisant un raisonnement par récurrence sur le nombre d'éléments de la famille.

Si la famille est vide, c'est évident. Supposons maintenant que le résultat est connu pour une famille ayant $n-1$ vecteurs, et démontrons-le pour une famille ayant n vecteurs. Appelons $v_1, \cdots, v_n$ ces vecteurs, de valeurs propres respectives $\lambda_1, \lambda_2, \ldots, \lambda_n$ (deux à deux distinctes). Il faut montrer l'implication

$$\sum_{k=1}^{n} \alpha_k v_k = 0 \Rightarrow v_1 = v_2 = \cdots = v_n = 0$$

Comme on suppose $\sum_{k=1}^{n} \alpha_k v_k = 0$, on peut appliquer l'endomorphisme f à cette équation. Nous trouvons

$$\sum_{k=1}^{n} \alpha_k f(v_k) = f(0) = 0$$

Comme v_k est un vecteur propre de f de valeur propre λ_k, nous obtenons

$$\sum_{k=1}^{n} \alpha_k \lambda_k v_k = 0$$

Mais il est vrai aussi que

$$\lambda_n \cdot (\sum_{k=1}^{n} \alpha_k v_k) = 0$$

c'est-à-dire

$$\sum_{k=1}^{n} \alpha_k \lambda_n v_k = 0$$

et en faisant la soustraction, nous voyons alors que

$$\sum_{k=1}^{n} \alpha_k (\lambda_k - \lambda_n) v_k = 0$$

ce qui équivaut à

$$\sum_{k=1}^{n-1} \alpha_k (\lambda_k - \lambda_n) v_k = 0$$

On peut alors appliquer l'hypothèse de récurrence, qui dit que la famille $(v_1, \cdots, v_{n-1})$ est libre. Par conséquent

$$\forall k \in [[1, n-1]], \alpha_k(\lambda_k - \lambda_n) = 0$$

Or nous avons supposé que $\lambda_k - \lambda_n \neq 0$. Donc $\alpha_k = 0$ pour toutes les valeurs de k entre 1 et $n-1$. Il ne reste plus qu'à montrer que $\alpha_n = 0$. Mais cela est très facile : en remplaçant dans l'égalité

$$\sum_{k=1}^{n} \alpha_k v_k = 0$$

on trouve simplement

$$\alpha_n v_n = 0$$

Enfin, par hypothèse, le vecteur v_n n'est pas nul. Il faut donc que $\alpha_n = 0$. Ceci achève la démonstration par récurrence, et par la même occasion, celle du théorème. □

Le théorème précédent possède aussi une autre version, qui décrit ce qui se passe lorsqu'une somme de vecteurs propres de valeurs propres deux à deux distinctes est égale au vecteur nul.

Théorème. *Soit f un endomorphisme d'un espace vectoriel V, et $I \subset Spec(f)$ un sous-ensemble (pas nécessairement fini) du spectre de f. On considère une famille $(v_\lambda)_{\lambda \in I}$ de vecteurs tels que*
(i) L'ensemble des $\lambda \in I$ tels que $v_\lambda \neq 0$ est fini,
(ii) Pour tout $\lambda \in I$, on a $v_\lambda \in E_\lambda(f)$,
(iii) La somme des v_λ est nulle : $\sum_{\lambda \in I} v_\lambda = 0$.
Alors tous les v_λ sont nuls.

Preuve. Observons que la condition (i) sert à garantir que la somme du (iii) ait bien un sens (les sommes infinies ne sont admises que si elles ne comportent qu'un nombre fini de termes non nuls). On note $J = \{\lambda \in I | v_\lambda \neq 0\}$. C'est par hypothèse un ensemble fini. Il faut montrer que J est l'ensemble vide. Par (iii) on peut aussi dire que

$$\sum_{\lambda \in J} v_\lambda = 0$$

Supposons par l'absurde que $J \neq \varnothing$. Alors la famille $(v_\lambda)_{\lambda \in J}$ est liée. Mais dans cette famille, tous les vecteurs sont non nuls, et les valeurs propres sont deux à deux distinctes. Donc c'est aussi une famille libre, grâce au théorème précédent. Or une famille de vecteurs ne peut être en même temps liée et libre. Contradition. On a bien $J = \varnothing$, ce qui prouve le résultat. □

Rappelons à ce stade la définition d'une somme directe de sous-espace vectoriels d'un espace vectoriel. Si V est un espace vectoriel sur un corps commutatif K, et $W_1, \ldots, W_n$ des sous-espaces de V (en nombre fini), on dit que V est la somme directe des W_i (où i parcourt $[[1,n]]$) et on note $V = W_1 \oplus \cdots \oplus W_n$, lorsque d'une part $V = W_1 + \cdots + W_n$ et que d'autre part la seule décomposition du vecteur nulle en somme de vecteurs appartenant aux W_i est la décomposition triviale. Il est équivalent de dire que tout vecteur de V admet une décomposition unique en somme de vecteurs appartenant aux W_i. On peut alors définir n projecteurs $p_1, \ldots, p_n$ associés à la somme directe $V = W_1 \oplus \cdots \oplus W_n$. Le projecteur p_j envoie le vecteur $w_1 + \ldots + w_n$ sur le vecteur w_j.

On peut étendre cette définition à la situation où le nombre de sous-espaces vectoriels de V est infini. Le seul changement par rapport à la situation précédente consiste à prendre seulement des sommes contenant un nombre fini de termes non nuls.

Avec ce langage, la proposition précédente peut se réécrire de la façon suivante : Si f est un endomorphisme, alors la somme de tous les sous-espaces propres de f est en fait la somme **directe** de ces sous-espaces propres. En langage symbolique, si $\mathrm{Spec}(f)$ désigne le spectre de f, on a

$$\sum_{\lambda \in \mathrm{Spec}(f)} E_\lambda(f) = \bigoplus_{\lambda \in \mathrm{Spec}(f)} E_\lambda(f).$$

Définition.
On dit qu'une matrice carrée A de taille $n \times n$ est **diagonalisable** *s'il existe une matrice inversible $P \in GL(n,K)$ telle que la matrice $P^{-1}AP$ soit une matrice diagonale.*

On dit qu'un endomorphisme f d'un espace vectoriel V est **diagonalisable** *s'il existe une base $(e_i)_{i\in I}$ de V dans laquelle chaque vecteur e_i est un vecteur propre de f.*

Faisons tout de suite le lien entre les deux sens donnés (a priori différents) au même mot «diagonalisable».

Théorème. *Soit f un endomorphisme d'un espace vectoriel E de dimension finie sur un corps commutatif K. On considère une base quelconque e de E.*
L'endomorphisme f est diagonalisable si et seulement si la matrice $\mathcal{M}_e(f)$ est diagonalisable.

Preuve. «seulement si» : On suppose f diagonalisable. Alors il existe une base $d = (d_1, \ldots, d_n)$ de E telle que

$$\forall i \in [[1,n]], \exists \lambda_i \in K, \quad f(d_i) = \lambda_i d_i$$

Mais alors la matrice $\mathcal{M}_d(f)$ est une matrice diagonale : le coefficient en position (i,i) est la valeur propre λ_i. Or $\mathcal{M}_e(f)$ et la matrice diagonale $\mathcal{M}_d(f)$ sont semblables, car elles représentent le même endomorphisme. Il en découle que $\mathcal{M}_e(f)$ est diagonalisable.
«si» : On suppose la matrice $\mathcal{M}_e(f)$ diagonalisable. Alors il existe une matrice diagonale D et une matrice inversible P telle que

$$P^{-1}\mathcal{M}_e(f)P = D$$

Or il existe une unique base d de E telle que $P_{e\to d} = P$. En l'occurrence il s'agit de la base d définie par

$$\forall i \in [[1,n]], \quad d_i = \sum_{j=1}^{n} P_{ji} e_j$$

C'est bien une base, puisque P est inversible. Avec cette définition de d on trouve alors

$$D = P_{e\to d}^{-1}\mathcal{M}_e(f)P_{e\to d} = \mathcal{M}_d(f)$$

Donc la matrice de f dans la base d est diagonale, ce qui signifie : $\forall i \in [[1,n]], f(d_i) = D_{ii}d_i$. La base d est bien une base de E dont tous les vecteurs sont des vecteurs propres de f. □

Si f est un endomorphisme de V (de dimension éventuellement infinie), on peut alors dire que f est diagonalisable si et seulement si $V = \sum_{\lambda \in \mathrm{Spec}(f)} E_\lambda(f)$ si et seulement si $V = \bigoplus_{\lambda \in \mathrm{Spec}(f)} E_\lambda(f)$.

Théorème.
Soit u un endomorphisme d'un espace vectoriel de dimension finie n sur un corps commutatif K. On suppose que le polynôme caractéristique de u possède n racines **distinctes** *dans le corps K. Alors u est diagonalisable.*

Preuve. Notons $r_1, \dots, r_n$ les n racines du polynôme caractéristique. Ce sont les valeurs propres de l'endomorphisme u. Pour chaque valeur propre r_j, choisissons un vecteur propre non nul v_j de u associé a cette valeur propre.
Nous savons que ces n vecteurs forment une famille libre de l'espace vectoriel. Comme l'espace vectoriel est de dimension égale a n, cette famille de vecteurs propres $(v_1, \dots, v_n)$ en est une base. La matrice associée à u dans la cette base est

$$\begin{pmatrix} r_1 & 0 & \cdots & 0 & 0 \\ 0 & r_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & r_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & r_n \end{pmatrix}$$

Puisque cette matrice est diagonale, on peut dire que l'endomorphisme u est diagonalisable. □

Corollaire.
Soit A une matrice carrée $n \times n$ à coefficients dans un corps commutatif K. On suppose que le polynôme caractéristique de A possède n racines **distinctes** *dans le corps K.*
Alors A est diagonalisable.

Pour terminer, donner également un critère simple pour savoir si un endomorphisme est diagonalisable dans une situation plus générale.

Théorème. *Soit u un endomorphisme d'un espace vectoriel de dimension finie n sur un corps commutatif K. L'endomorphisme u est diagonalisable si et seulement si :*
i) son polynôme caractéristique est scindé sur K,
ii) Pour chaque racine λ du polynôme caractéristique, sa multiplicité est égale à la dimension de l'espace propre associé à λ.

Preuve. «seulement si». Comme u est diagonalisable, il existe une base e dans laquelle la matrice représentant u est une matrice diagonale de diagonale $(d_1, d_2, \dots, d_n)$. Le polynôme caractéristique de u est alors

$$(-1)^n (X - d_1)(X - d_2) \cdots (X - d_n)$$

ce qui prouve évidemment que c'est un polynôme scindé. En plus, si λ apparaît k fois sur la diagonale de la matrice, alors la multiplicité de la racine λ est égale à k, et l'espace propre associé à λ est l'espace engendré par les k vecteurs de la base correspondant aux apparitions de λ. Cet espace est aussi de dimension k. Les conditions i) et ii) sont vérifiées.

«si». On appelle $\lambda_1, \dots, \lambda_p$ les racines **distinctes** du polynôme caractéristique de u. On note $m_1, \dots, m_p$ leurs multiplicités. Comme le polynôme caractéristique est scindé, la somme des multiplicités vaut n. En plus, l'espace propré associé à λ_j est de dimension m_j. Comme la somme des m_j est égale à la dimension de l'espace vectoriel, on peut en déduire que l'espace vectoriel de dimension n se décompose en somme directe des p espaces propres. Cela implique en particulier qu'on peut trouver une base de vecteurs propres de u. Mais s'il existe une base de vecteurs propres, cela signifie que u est diagonalisable. □

Corollaire. *Soit A une matrice $n \times n$ à coefficients dans un corps commutatif K. La matrice A est diagonalisable si et seulement si :*
i) son polynôme caractéristique est scindé sur K,
ii) Pour chaque racine λ du polynôme caractéristique, sa multiplicité est égale à la dimension de l'espace propre associé à λ.

Il est bien connu que si le corps K est le corps $\mathbb{C}$ des nombres complexes, tout polynôme est scindé ; c'est le théorème de Gauss-d'Alembert. En particulier, un polynôme caractéristique est toujours scindé sur $\mathbb{C}$. Mais il ne faut pas en déduire que tout endomorphisme d'un espace vectoriel sur $\mathbb{C}$ est diagonalisable, car la condition (ii) n'est pas toujours vérifiée. Nous donnons ici un contre-exemple.

Soit la matrice

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathcal{M}_{2,2}(\mathbb{C})$$

Son polynôme caractéristique est

$$\det \begin{pmatrix} X & -1 \\ 0 & X \end{pmatrix} = X^2,$$

qui a une seule racine ($X = 0$), de multiplicité égale à 2. Mais l'espace propre associé à la valeur propre $\lambda = 0$ est l'espace des vecteurs de la forme

$$\begin{pmatrix} t \\ 0 \end{pmatrix}$$

avec t complexe quelconque. Donc l'espace propre est de dimension 1. Cette dimension n'est pas égale à la multiplicité de la racine zéro. Conclusion : La matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ n'est pas diagonalisable.

***

Rappelons enfin que pour V de dimension finie, une décomposition en somme directe $V = W_1 \oplus \cdots \oplus W_n$ implique l'égalité dimensionnelle

$$\dim V = \sum_{i=1}^{n} \dim W_i.$$

Il en découle alors que f endomorphisme de V est diagonalisable si et seulement si $\dim V$ est égale à la somme des dimensions de tous les espaces propres. Le cas "f diagonalisable" correspond exactement au cas d'égalité de l'inégalité suivante, qui est toujours valable (comme on peut le voir en utilisant $\sum E_\lambda(f) = \bigoplus E_\lambda(f)$) :

$$\sum_{\lambda \in \mathrm{Spec}(f)} \dim E_\lambda(f) \leqslant \dim V.$$

Chapitre 2. Espaces pré-hilbertiens. Espaces euclidiens.

Dans ce chapitre, tous les espaces vectoriels sont des espaces vectoriels sur le corps $\mathbb{R}$ des nombres réels (de dimension finie ou infinie).

Le concept essentiel que nous développons dans ce chapitre est celui de **produit scalaire** sur un espace vectoriel sur $\mathbb{R}$. Rappelons rapidement ce qu'est une forme bilinéaire symétrique.

Définition. *Soit E un espace vectoriel sur $\mathbb{R}$.* **Une forme bilinéaire symétrique sur** *E est une application*

$$E \times E \to \mathbb{R} : (x, y) \mapsto \langle x, y \rangle$$

qui obéit aux relations ci-dessous :
i) Pour tout $x \in E$, l'application $y \mapsto \langle x, y \rangle$ est une forme linéaire sur E
ii) Pour tout $y \in E$, l'application $x \mapsto \langle x, y \rangle$ est une forme linéaire sur E
(ce sont les propriétés de bilinéarité)
iii) Pour tous $x, y \in E$, on a $\langle x, y \rangle = \langle y, x \rangle$.
(c'est la propriété de symétrie)

Exemples.
1. Le produit scalaire géométrique sur le plan $\mathbb{R}^2$ et le produit scalaire géométrique sur l'espace $\mathbb{R}^3$ sont des formes bilinéaires symétriques.
2. L'application $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} : ((x, y), (x', y')) \mapsto xx' - yy'$ est une forme bilinéaire symétrique.
3. L'application $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} : ((x, y), (x', y')) \mapsto xx' - yx'$ est une forme bilinéaire, mais pas symétrique.

A chaque forme bilinéaire symétrique on peut associer sa forme quadratique associée :

Définition. *Soit $\langle\ ,\ \rangle$ une forme bilinéaire symétrique sur un espace vectoriel E. Sa* **forme quadratique associée** *est l'application*

$$q : E \to \mathbb{R} : x \mapsto \langle x, x \rangle$$

Ainsi, la forme quadratique associée au produit scalaire géométrique sur l'espace $\mathbb{R}^3$ est l'application

$$\mathbb{R}^3 \to \mathbb{R} : (x, y, z) \mapsto x^2 + y^2 + z^2$$

C'est le carré de la norme du vecteur (x, y, z). Cette forme quadratique prend seulement des valeurs positives.

La forme quadratique associée à la forme bilinéaire symétrique $((x, y), (x', y')) \mapsto xx' - yy'$ sur $\mathbb{R}^2$ est l'application

$$\mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto x^2 - y^2$$

Cette forme quadratique ne prend pas seulement des valeurs positives.

Lorsqu'on connaît la forme quadratique associée à une forme bilinéaire symétrique, on peut facilement retrouver la forme bilinéaire symétrique, grâce à l'identité de **polarisation**.

Théorème (Identité de polarisation).
Soit $\langle\ ,\ \rangle$ une forme bilinéaire symétrique sur E, et q sa forme bilinéaire associée. Alors

$$\forall x, y \in E, \quad 2\langle x, y \rangle = q(x + y) - q(x) - q(y)$$

Preuve. Il suffit de calculer le membre de droite.

$$\begin{aligned} q(x+y) - q(x) - q(y) &= \langle x+y, x+y \rangle - \langle x, x \rangle - \langle y, y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle - \langle x, x \rangle - \langle y, y \rangle \\ &= \langle x, y \rangle + \langle y, x \rangle \\ &= 2\langle x, y \rangle \end{aligned}$$

□

En d'autres termes, la forme quadratique associée à une forme bilinéaire symétrique détermine la forme bilinéaire de façon unique.

Définition.
On dit qu'une forme quadratique q sur E est **positive** *si pour tout $x \in E$, on a $q(x) \geqslant 0$.*
On dit qu'elle est **définie positive** *si elle est positive et si $\forall x \in E, q(x) = 0 \Rightarrow x = 0$.*
On dit qu'une forme bilinéaire symétrique est positive (respectivement définie positive) si sa forme quadratique associée est positive (respectivement définie positive).

Par exemple, le produit scalaire géométrique sur $\mathbb{R}^2$ est une forme bilinéaire symétrique définie positive.
De même, le produit scalaire géométrique sur $\mathbb{R}^3$ est une forme bilinéaire symétrique définie positive.
La forme quadratique sur $\mathbb{R}^3$ définie par $(x, y, z) \mapsto x^2 + y^2$ est une forme quadratique positive, mais pas définie positive. La forme quadratique sur $\mathbb{R}^2$ définie par $(x, y) \mapsto x^2 - y^2$ n'est pas une forme quadratique positive.

Les formes bilinéaires symétriques positives impliquent une inégalité importante, celle de Cauchy-Schwarz.

Théorème (Inégalité de Cauchy-Schwarz).
Soit $\langle , \rangle$ une forme bilinéaire symétrique positive sur E. Alors

$$\forall x, y \in E, \langle x, y\rangle^2 \leqslant q(x)q(y)$$

En plus, si la forme est définie positive, alors le cas d'égalité $\langle x, y\rangle^2 = q(x)q(y)$ implique nécessairement que les vecteurs x, y sont colinéaires.

Preuve. Evidemment, pour tout t réel, on peut dire que

$$q(x + ty) \geqslant 0$$

puisque la forme quadratique associée est positive. En développant, cela donne

$$\forall t, \quad q(y)t^2 + 2\langle x, y\rangle t + q(x) \geqslant 0$$

Si $q(y) = 0$, il faut obligatoirement que $\langle x, y\rangle = 0$, car sinon l'inégalité ne peut pas être vérifiée. Dans ce cas, l'inégalité est clairement vérifiée (et c'est même l'égalité $0 = 0$).

Si $q(y) \neq 0$, l'expression $q(y)t^2 + 2\langle x, y\rangle t + q(x)$ est un polynôme en t de degré 2 qui prend seulement des valeurs positives. Dès lors, il est impossible que son discriminant soit strictement positif. Donc son discriminant est $\leqslant 0$, ce qui donne le résultat annoncé :

$$4\langle x, y\rangle^2 - 4q(y)q(x) \leqslant 0$$

Il faut encore examiner le cas où la forme est définie positive. On suppose qu'on a l'égalité $\langle x, y\rangle^2 = q(x)q(y)$. Si $q(y) = 0$, alors $y = 0$, et il est clair que x, y sont colinéaires. Si on suppose maintenant que $q(y)$ est non nul, l'hypothèse signifie que le discriminant du polynôme $q(y)t^2 + 2\langle x, y\rangle t + q(x)$ est égal à zéro. Donc ce polynôme a une racine, ce qui veut dire qu'il existe t réel tel que

$$q(x + ty) = 0$$

Et comme q est définie positive, on peut en déduire que $x + ty = 0$. Bref, les vecteurs x, y sont colinéaires. □

A chaque forme bilinéaire symétrique (pas forcément positive) sur un espace vectoriel de dimension finie n, on peut associer une matrice carrée $n \times n$ de la façon suivante :

Définition.
Soit E un espace vectoriel de dimension n, $\langle , \rangle$ une forme bilinéaire symétrique sur E, et soit $e = (e_1, \ldots, e_n)$ une base de E. **La matrice associée à la forme bilinéaire symétrique par par rapport à la base** *e est la matrice $G(e)$ de taille $n \times n$ définie par*

$$\forall i, j \in [[1, n]], G(e)_{ij} := \langle e_i, e_j\rangle$$

On remarque évidemment que $G(e)$ est une matrice symétrique.

La matrice associée au produit scalaire géométrique de l'espace par rapport à une base **orthonormée** est la matrice identité

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

Il suffit de connaître la matrice $G(e)$ pour connaître complètement la forme bilinéaire symétrique :

Théorème. *Soient x, y deux vecteurs de E, et on note X, Y les matrices-colonne dont les coefficients sont les coordonnées de x, y par rapport à la base e. Alors*

$$\langle x, y\rangle = {}^tX \cdot G(e) \cdot Y$$

(il faut observer que le membre de droite est une matrice 1×1, que l'on voit comme un réel.)

Preuve.

$$\begin{aligned}
\langle x, y\rangle &= \langle \sum_{i=1}^{n} X_i e_i, \sum_{j=1}^{n} Y_j e_j \rangle \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} X_i Y_j \langle e_i, e_j\rangle \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} X_i G_{ij} Y_j \\
&= \sum_{i=1}^{n} X_i \sum_{j=1}^{n} G_{ij} Y_j \\
&= \sum_{i=1}^{n} X_i (GY)_{i1} \\
&= \sum_{i=1}^{n} ({}^tX)_{1i} (GY)_{i1} \\
&= ({}^tXGY)_{11} \\
&= {}^tXGY
\end{aligned}$$

□

Nous allons maintenant étudier la loi de transformation de la matrice $G(e)$ lorsqu'on fait varier la base e.

Théorème. *Soit une forme bilinéaire symétrique sur un espace vectoriel de dimension finie n, et deux bases e, e' de ces espace. On note $P_{e\to e'}$ la matrice de passage de la base e vers la base e'. Alors*

$$G(e') = {}^tP_{e\to e'} \cdot G(e) \cdot P_{e\to e'}$$

Preuve. Soient i, j des entiers entre 1 et n.

$$\begin{aligned}
G(e')_{ij} &= \langle e'_i, e'_j\rangle \\
&= \sum_{k=1}^{n}\sum_{l=1}^{n} P_{ki} G(e)_{kl} P_{lj} \\
&= ({}^tP \cdot G(e) \cdot P)_{ij}
\end{aligned}$$

On peut alors en déduire que les matrices $G(e')$ et ${}^tP_{e\to e'} \cdot G(e) \cdot P_{e\to e'}$ sont égales. □

Nous arrivons à la définition d'un produit scalaire.

Définition.
Soit E un espace vectoriel sur le corps des réels.
Un produit scalaire sur *E est une forme bilinéaire symétrique et définie positive.*
Un espace pré-hilbertien *est un espace vectoriel muni d'un produit scalaire.*
Un espace euclidien *est un espace préhilbertien de dimension finie.*

Donc, dans un espace pré-hilbertien, l'inégalité de Cauchy-Schwarz est toujours vraie.

Voyons maintenant les exemples les plus courants d'espaces pré-hilbertiens et euclidiens.

Exemples.
1. Si $E = \mathbb{R}^n$, on peut définir sur E le **produit scalaire canonique** défini par la loi naturelle ci-dessous :
Si $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$ sont deux vecteurs de $\mathbb{R}^n$, on pose

$$\langle x, y\rangle := x_1y_1 + x_2y_2 + \cdots + x_ny_n = \sum_{i=1}^{n} x_iy_i$$

Il est facile de vérifier qu'il s'agit d'une forme bilinéaire, qu'elle est symétrique, et qu'elle est définie positive. L'espace vectoriel $\mathbb{R}^n$, avec ce produit scalaire, est donc un espace euclidien (car il est de dimension finie).

2. On appelle E l'ensemble de toutes les fonctions continues du segment $[a, b]$ vers $\mathbb{R}$. C'est un espace vectoriel de dimension infinie (un bon exercice consiste à montrer pourquoi). On définit le produit scalaire L^2 sur $[a, b]$ par la formule

$$\forall f, g \in E, \langle f, g\rangle = \int_a^b f(x)g(x)dx$$

C'est bien une forme bilinéaire et symétrique, clairement positive, et on peut montrer (même si c'est un peu plus compliqué) que c'est une forme définie positive. C'est bien un produit scalaire. L'espace E, muni de ce produit scalaire L^2, est un espace pré-hilbertien.

Avec un produit scalaire, on peut définir les notions usuelles de norme, de distance et d'orthogonalité :

Définition.
Soit E un espace pré-hilbertien, et q la forme quadratique associée au produit scalaire. Si x est un vecteur de E, **la norme de** *x est le nombre réel positif $\sqrt{q(x)}$. On la note $||x||$.*
Si x, y sont deux vecteurs de E, **la distance de** *x* **à** *y est le nombre $||y - x||$. On la note $d(x, y)$.*

La norme possède les propriétés suivantes :
1. Quel que soit x, la norme de x est positive, et ne s'annule que si $x = 0$.
2. Quel que soit x (vecteur) et k (réel), on a $||kx|| = |k| \cdot ||x||$.
Ces deux propriétés sont triviales.
3. Quels que soient les vecteurs x, y, on a l'inégalité triangulaire

$$||x + y|| \leqslant ||x|| + ||y||$$

Pour la prouver, il suffit d'utiliser l'inégalité de Cauchy-Schwarz (valable dans les espaces pré-hilbertiens).

Le distance possède les propriétés suivantes :
1. Une distance est toujours positive, et si $d(x, y) = 0$, alors $x = y$.
2. Une distance est symétrique : $d(x, y) = d(y, x)$.
3. Une distance obéit à l'inégalité triangulaire : $d(x, z) \leqslant d(x, y) + d(y, z)$
Avec ce que nous savons sur la norme, ces trois propriétés sont faciles à prouver.

Définition. *Soit E un espace pré-hilbertien, et $\langle\ ,\ \rangle$ son produit scalaire.*
On dit qu'un vecteur x de E est **unitaire ou normé** *si sa norme vaut 1.*
On dit que deux vecteurs x, y de E sont **orthogonaux** *si $\langle x, y\rangle = 0$. On dit que deux sous-espaces vectoriels F_1, F_2 de E sont* **orthogonaux** *si tout vecteur de F_1 est orthogonal à tout vecteur de F_2.*
Si F est un sous-espace vectoriel de E, **l'orthogonal de** *F, noté $F^\perp$, est l'ensemble de tous les vecteurs qui sont orthogonaux à tous les vecteurs de F :*

$$F^\perp := \{x \in E | \forall y \in F, \langle x, y\rangle = 0\}$$

Définition.
On dit qu'une famille $(x_i)_{i\in I}$ de vecteurs de E est **une famille orthogonale** *si ces vecteurs sont deux à deux orthogonaux.*
On dit qu'une famille $(x_i)_{i\in I}$ de vecteurs de E est **une famille orthonormale** *si cette famille est orthogonale et si tous ces vecteurs sont normés.*

Par exemple, si $E = \mathbb{R}^5$ est muni du produit scalaire canonique, la famille

$$((1,0,0,0,0),(0,0,1,0,0),(0,0,0,1,0))$$

est une famille orthonormée, mais ce n'est pas une base orthonormée (il y a seulement 3 vecteurs, alors qu'une base en a forcément 5).

Théorème.
L'orthogonal $F^\perp$ d'un sous-espace vectoriel F de E est aussi un sous-espace vectoriel de E.
On a aussi $F \cap F^\perp = \{0\}$.

Preuve.
1. Le vecteur nul de E est dans $F^\perp$, puisque le vecteur nul est orthogonal à tout vecteur (cela découle de la bilinéarité du produit scalaire).
2. La somme de deux vecteurs x, x' de $F^\perp$ est encore dans $F^\perp$. En effet, pour tout $y \in F$, on a

$$\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle = 0 + 0 = 0$$

3. Le produit d'un vecteur x de $F^\perp$ avec un réel k est encore dans $F^\perp$. En effet, pour tout $y \in F$, on a

$$\langle kx, y\rangle = k\langle x, y\rangle = k \cdot 0 = 0$$

Donc $F^\perp$ est effectivement un sous-espace vectoriel de E.
Il reste à montrer que l'intersection de F et de son orthogonal est réduite au vecteur nul. D'un côté, il est évident que le vecteur nul appartient aux deux sous-espaces. De l'autre côté, si x appartient à la fois à F et à $F^\perp$, alors en particulier, les vecteurs x et x sont orthogonaux, donc

$$\langle x, x\rangle = 0$$

Mais cela veut dire $q(x) = 0$ et comme la forme q est définie positive (nous sommes dans un espace préhilbertien) on a forcément $x = 0$. □

Et maintenant, un grand classique dans une version moderne :

Théorème de Pythagore. *Soit E un espace pré-hilbertien, et $(x_1, x_2, \ldots, x_n)$ une* **famille orthogonale** *finie de vecteurs de E. Alors on a l'égalité*

$$||x_1 + x_2 + \cdots + x_n||^2 = ||x_1||^2 + ||x_2||^2 + \cdots + ||x_n||^2$$

On peut aussi l'écrire sous la forme

$$q(\sum_i x_i) = \sum_i q(x_i)$$

Preuve. Il suffit d'utiliser les définitions.

$$\begin{aligned} ||x_1 + x_2 + \cdots + x_n||^2 &= \langle \sum_{i=1}^n x_i, \sum_{j=1}^n x_j\rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n \langle x_i, x_j\rangle \\ &= \sum_{i=1}^n \langle x_i, x_i\rangle \qquad \text{(c'est une famille orthogonale)} \\ &= ||x_1||^2 + \cdots + ||x_n||^2 \end{aligned}$$

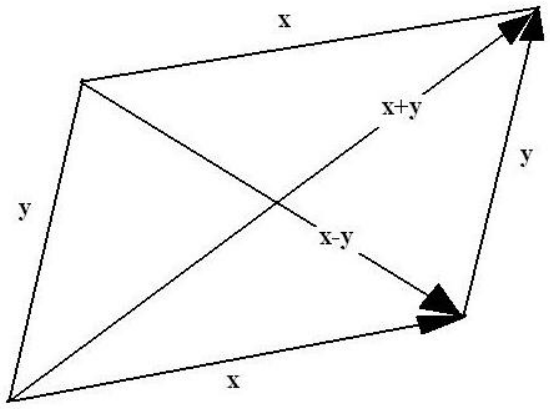
□

Pour terminer ce chapitre, citons une petite identité amusante, appelée **l'identité du parallélogramme**. Elle dit que dans tout espace pré-hilbertien, et pour tout couple de vecteurs x, y de E, on a

$$||x+y||^2 + ||x-y||^2 = 2(||x||^2 + ||y||^2)$$

La preuve en est très simple, il suffit par exemple d'utiliser deux fois l'identité de polarisation.

Pourquoi cette égalité s'appelle-t-elle l'identité du parallélogramme ? En fait, elle possède une interprétation géométrique simple : Dans le plan, dessinez un parallélogramme. L'identité dit alors que la somme des carrés des longueurs des 4 côtés du parallélogramme est égale à la somme des carrés des longueurs des 2 diagonales du parallélogramme (Si le parallélogramme est un rectangle, que retrouve-t-on ?)



Chapitre 3. Bases orthonormées. Orthogonal d'un sous-espace vectoriel.

Rappelons qu'une **base orthonormée** d'un espace euclidien est une base qui est aussi une famille orthonormée.

Au début de ce chapitre, nous allons apprendre comment construire une base orthonormée à partir d'une base quelconque d'un espace euclidien. Cette construction s'appelle le procédé de Gram-Schmidt.

Théorème (Orthonormalisation de Gram-Schmidt).
Soit E un espace euclidien de dimension n et $e = (e_1, \dots, e_n)$ une base quelconque de E. Alors il existe une unique base orthonormée f de E telle que la matrice de passage $P_{e\to f}$ soit triangulaire supérieure et à des coefficients diagonaux > 0. En particulier on a pour tout $i \in [[1, n]]$:

$$Vect(e_1, \dots, e_i) = Vect(f_1, \dots, f_i)$$

Preuve.
Nous commençons par montrer l'unicité. Si f et f' sont deux bases orthonormées qui conviennent, alors $P_{f\to f'} = P_{e\to f}^{-1} P_{e\to f'}$ est encore une matrice triangulaire supérieure à coefficients diagonaux strictement positifs. En effet, l'ensemble des matrices triangulaires supérieures à coefficients diagonaux strictment positifs, muni de la multiplication matricielle, est un groupe; il s'agit d'un sous-groupe de $GL(n, \mathbb{R})$. D'un autre côté, $P_{f\to f'}$ est une matrice de passage entre deux bases orthonormées, donc

$$I_n = {}^t P_{f\to f'} \cdot I_n \cdot P_{f\to f'}$$

Si on note $A = P_{f\to f'}$, on sait que A est triangulaire supérieure à diagonale > 0 et que ${}^t AA = I$. En particulier $A_{11}^2 = 1$, d'où $A_{11} = 1$ par positivité. La première colonne de A est le premier vecteur de la base canonique de $\mathcal{M}_{n,1}(\mathbb{R})$. Ensuite $A_{11}A_{12} = 0$, ce qui donne $A_{12} = 0$. Puis on a aussi $A_{12}^2 + A_{22}^2 = 1$, et comme $A_{22} > 0$, on trouve que la deuxième colonne de A est le deuxième vecteur de la base canonique de $\mathcal{M}_{n,1}(\mathbb{R})$. De proche en proche, on montre alors que les colonnes de A sont exactement les vecteurs de la base canonique. Donc $A = I$, ce qui prouve bien que $f = f'$. L'unicité est établie.

Pour l'existence, nous allons construire une base f par récurrence. D'abord, on construit un vecteur f_1 en posant

$$f_1 := \frac{e_1}{||e_1||}$$

Bien sûr, on peut diviser par la norme de e_1, puisque e_1 n'est pas le vecteur nul. Le vecteur f_1 est de norme 1, et les vecteurs e_1, f_1 sont colinéaires. D'où

$$Vect(e_1) = Vect(f_1)$$

En outre, le coefficient $(1,1)$ de $P_{e\to f}$ est $1/||e_1|| > 0$.
Supposons maintenant que pour un certain $k \in [[1, n-1]]$, nous ayons déjà construit une famille orthonormée $(f_1, \dots, f_k)$ telle que : (1) $\forall j \leqslant k, Vect(f_1, \dots, f_j) = Vect(e_1, \dots, e_j)$, et
(2) $\forall j \leqslant k, f_j = \sum_{i=1}^{j} t_{ij} e_i$ avec $t_{jj} > 0$. Posons alors $g_{k+1} = e_{k+1} - \sum_{j=1}^{k} \langle e_{k+1}, f_j \rangle f_j$.
Le vecteur g_{k+1} est orthogonal à $f_1, f_2, \dots, f_k$, puisque pour tout $i \in [[1, k]]$,

$$\langle g_{k+1}, f_i \rangle = \langle e_{k+1}, f_i \rangle - \sum_{j=1}^{k} \langle e_{k+1}, f_j \rangle \langle f_j, f_i \rangle = \langle e_{k+1}, f_i \rangle - \langle e_{k+1}, f_i \rangle \cdot 1 = 0$$

En outre, le vecteur g_{k+1} n'est pas nul, car sinon on aurait

$$e_{k+1} = \sum_{j=1}^{k} \langle e_{k+1}, f_j \rangle f_j \in Vect(f_1, \dots, f_k)$$

et comme $Vect(f_1, \ldots, f_k) = Vect(e_1, \ldots, e_k)$, on aurait que la famille $(e_1, \ldots, e_k, e_{k+1})$ est liée, ce qui est évidemment impossible. Comme g_{k+1} est non nul, sa norme est aussi non nulle, et on peut définir

$$f_{k+1} := \frac{g_{k+1}}{||g_{k+1}||}$$

La famille $(f_1, \ldots, f_k, f_{k+1})$ est alors clairement une famille orthonormée. Il est également clair que

$$g_{k+1} \in Vect(e_1, \ldots, e_{k+1})$$

et donc que

$$f_{k+1} \in Vect(e_1, \ldots, e_{k+1})$$

On peut en déduire que

$$Vect(f_1, \ldots, f_{k+1}) \subset Vect(e_1, \ldots, e_{k+1})$$

D'un autre côté, il est possible d'exprimer e_{k+1} comme combinaison linéaire des vecteurs $f_1, \ldots, f_{k+1}$, ce qui montre l'inclusion

$$Vect(e_1, \ldots, e_{k+1}) \subset Vect(f_1, \ldots, f_{k+1})$$

D'où l'égalité entre ces deux ensembles. En outre, si on décompose f_{k+1} dans la base $(e_1, \ldots, e_{k+1})$ de $Vect(f_1, \ldots, f_{k+1})$, on trouve que le coefficient numéro $k+1$ est égal à $1/||g_{k+1}||$, un nombre strictement positif. Le pas de récurrence s'achève.

Après n étapes du processus décrit ci-dessus, on a construit une famille orthonormée de n vecteurs $(f_1, \ldots, f_n)$ telle que

$$Vect(f_1, \ldots, f_n) = Vect(e_1, \ldots, e_n)$$

Or e est une base, donc elle engendre tout l'espace vectoriel E. On en déduit que (f_i) est une famille génératrice ayant n vecteurs. C'est donc aussi une base, et comme c'est en même temps une famille orthonormée, nous pouvons dire que c'est une base orthonormée. □

On en tire deux corollaires immédiats :

Corollaire. *Dans tout espace euclidien, il existe une base orthonormée.*

Preuve. Prendre une base quelconque (on sait qu'elle existe) et lui appliquer le procédé de Gram-Schmidt. □

Corollaire. *Si $(f_1, \ldots, f_k)$ est une famille orthonormée d'un espace euclidien, alors on peut compléter cette famille en une base orthonormée.*

Preuve. Toute famille orthonormée est libre. En effet, si on a $\sum_{j=1}^{k} \lambda_j f_j = 0$, alors en particulier, on peut prendre la norme au carré des deux membres. Comme il s'agit d'une famille orthonormée, on trouve $\sum_{j=1}^{k} \lambda_j^2 = 0$. Tous les λ_j sont réels, ce qui implique qu'ils sont tous nuls.

Par le théorème de la base incomplète, on peut donc la compléter en une base de E (qui n'est pour l'instant pas orthonormée). Or si on applique à cette base le procédé de Gram-Schmidt, on obtient une base orthonormée, et en plus, les k premiers vecteurs de cette base sont toujours les vecteurs $f_1, \ldots, f_k$, car le procédé de Gram-Schmidt ne les change pas (ils forment déjà une famille orthonormée). □

Grâce à un produit scalaire, on peut associer à tout vecteur d'un espace pré-hilbertien E une forme linéaire sur E. En d'autres termes, on obtient une application canonique de E vers l'ensemble de toutes les formes linéaires sur E. On sait déjà que cet ensemble est lui aussi un espace vectoriel sur $\mathbb{R}$. Cet espace vectoriel est un objet important, car il intervient dans la dualité vecteurs/formes linéaires.

Définition. *Soit E un espace vectoriel sur $\mathbb{R}$. On appelle* **espace dual de** *E, noté E^*, l'espace vectoriel de toutes les formes linéaires sur E, autrement dit*

$$E^* := \mathcal{L}(E, \mathbb{R})$$

La définition s'étend sans problème au cas d'un corps commutatif K quelconque.

En dimension finie, on peut remarquer que la dimension de E et la dimension de E^* sont égales. En effet, nous savons déjà que si V, W sont deux espaces vectoriels de dimension finie, alors

$$\dim \mathcal{L}(V, W) = \dim V \cdot \dim W$$

En particulier,

$$\dim \mathcal{L}(E, \mathbb{R}) = \dim E \cdot \dim \mathbb{R} = \dim E$$

D'où le résultat : en dimension finie

$$\dim E^* = \dim E$$

Théorème. *Soit E un espace euclidien muni du produit scalaire $\langle\ ,\ \rangle$. On définit l'application*

$$\psi : E \mapsto E^* : a \mapsto (b \mapsto \langle a, b\rangle)$$

L'application ψ est un isomorphisme de l'espace vectoriel E vers l'espace dual E^*.

Preuve. Tout d'abord, il faut se convaincre que $\psi(a)$ est bien un élément de E^*, c'est-à-dire une forme linéaire sur E. Cela revient à dire que pour $a \in E$ fixé, l'application

$$E \to \mathbb{R} : b \mapsto \langle a, b\rangle$$

est une application linéaire. Evidemment cela est vrai par la bilinéarité du produit scalaire.

Ensuite, l'application ψ est linéaire, car

$$\begin{aligned} \psi(ka + k'a')(b) &= \langle ka + k'a', b\rangle \\ &= k\langle a, b\rangle + k'\langle a', b\rangle \\ &= k\psi(a)(b) + k'\psi(a')(b) \\ &= (k\psi(a) + k'\psi(a'))(b') \end{aligned}$$

où k, k' sont deux réels, et a, a', b trois vecteurs de E.

Montrons maintenant que ψ est injective. Pour cela il suffit de montrer que son noyau est réduit au vecteur nul. Si on a $\psi(a) = 0$, cela signifie que pour tout $b \in E$, on a $\langle a, b\rangle = 0$. En particulier, c'est vrai pour $b = a$, et donc $q(a) = 0$, ce qui implique que $a = 0$ (un produit scalaire est défini positif !).

Donc ψ est une application linéaire injective de E vers E^*. Mais E est euclidien, donc de dimension finie, et dans ce cas, E et E^* ont la même dimension. On peut en déduire que ψ est une application linéaire bijective. C'est bien un isomorphisme. □

Ce théorème est souvent utilisé, et de différentes manières. Par exemple, comme ψ est surjective, on peut dire : Quelle que soit la forme linéaire ω sur E, il existe toujours un vecteur $a \in E$ tel que

$$\forall b \in E, \omega(b) = \langle a, b\rangle$$

Ce vecteur a est (l'unique) antécédent de ω par l'isomorphisme ψ.

Notons aussi qu'une autre manière de définir l'isomorphisme ψ est d'écrire :

$$\forall a, b \in E, (\psi(a))(b) = \langle a, b\rangle$$

Théorème.
Soit E un espace pré-hilbertien, et F un sous-espace vectoriel **de dimension finie** *de E. Alors*

$$E = F \oplus F^\perp$$

En d'autres termes, chaque vecteur de E s'écrit d'une seule manière comme somme d'un vecteur de F et d'un vecteur de $F^\perp$.

Preuve. Introduisons l'application

$$\phi : E \to F^* : a \mapsto (b \mapsto \langle a, b\rangle)$$

où b est un vecteur quelconque de F. Comme dans la preuve précédente, il est facile de voir que $\phi(a)$ est bien dans F^*, et que ϕ est une application linéaire.

Cherchons le noyau de ϕ. On a

$$\begin{aligned} \phi(a) = 0 &\iff \forall b \in F, \langle a, b\rangle = 0 \\ &\iff a \in F^\perp \end{aligned}$$

Par conséquent, $\ker\phi = F^\perp$.
On sait aussi que l'application $\psi : F \to F^*$ du théorème précédent est un isomorphisme (car F est de dimension finie, et en tant que sous-espace d'un espace pré-hilbertien, c'est un espace euclidien). On peut alors former une application linéaire

$$\alpha = \psi^{-1} \circ \phi : E \to F$$

Le noyau de α est le noyau de ϕ car ψ est une bijection. Bref $\ker\alpha = F^\perp$. Ensuite montrons que $\alpha \circ \alpha = \alpha$, c'est-à-dire que α fixe tous les vecteurs de l'image de α. Comme l'image de α est contenue dans F, il suffit de prouver que

$$\forall y \in F, \alpha(y) = y$$

ce qui signifie encore

$$\forall y \in F, \phi(y) = \psi(y)$$

Or d'après la définition de ψ, cela est évident. Nous pouvons donc dire que α est un projecteur de E et que son noyau est $F^\perp$. En outre $F = \mathrm{Im}\ \alpha$. Puisqu'il s'agit d'un projecteur on a

$$E = \mathrm{Im}\ \alpha \oplus \ker\alpha = F \oplus F^\perp$$

□

Comme cas particulier fréquemment utilisé, nous avons ceci :

Corollaire. *Soit E un espace euclidien, et F un sous-espace vectoriel de E. Alors on a toujours*

$$E = F \oplus F^\perp$$

En effet, puisque E est supposé de dimension finie, F l'est forcément aussi.

Corollaire. *Soit F un sous-espace vectoriel d'un espace euclidien. Alors*

$$F^{\perp\perp} = F$$

Preuve. Comme F et $F^\perp$ sont supplémentaires, la somme de leurs dimenions vaut $\dim E$. De même, la somme des dimensions de $F^\perp$ et $F^{\perp\perp}$ vaut $\dim E$. On peut en déduire que les espaces F et $F^{\perp\perp}$ ont la même dimension. Nous allons maintenant montrer que F est inclus dans le double orthogonal $F^{\perp\perp}$. En effet, soit $x \in F$. Alors

$$\forall y \in F^\perp, \langle x, y\rangle = 0$$

Donc x est orthogonal à tous les vecteurs du sous-espace $F^\perp$, ce qui implique que

$$x \in (F^\perp)^\perp$$

De l'inclusion

$$F \subset F^{\perp\perp}$$

et de l'égalité des dimensions on peut déduire que ces deux sous-espaces sont égaux. □

Chapitre 4. Projecteurs orthogonaux.

Dans le chapitre précédent, nous avons vu que pour un espace E pré-hilbertien, et un sous-espace F de **dimension finie** de E, l'orthogonal de F est toujours un sous-espace vectoriel supplémentaire de F. On peut alors définir des projecteurs.

Définition.
Soit E un espace vectoriel pré-hilbertien, et F un sous-espace vectoriel de dimension finie de E. On appelle **projecteur orthogonal de E sur F** *le projecteur sur le sous-espace F parallèlement à son supplémentaire $F^\perp$.*

Théorème. *Soit $p : E \to E$ le projecteur orthogonal de E préhilbertien sur un sous-espace de dimension finie F (on pourrait aussi écrire $p : E \to F$). On considère une base orthonormée $(e_1, \ldots, e_n)$ de F. Alors*

$$\forall x \in E, p(x) = \sum_{i=1}^{n} \langle x, e_i \rangle e_i$$

Preuve. Il est évident que $\sum_{i=1}^{n} \langle x, e_i \rangle e_i$ appartient à F. Pour être certain que c'est $p(x)$, il suffit de montrer que $x - \sum_{i=1}^{n} \langle x, e_i \rangle e_i$ est dans $F^\perp$, ou encore que

$$\forall j \in [[1, n]], \langle x - \sum_{i=1}^{n} \langle x, e_i \rangle e_i, e_j \rangle = 0$$

On calcule alors

$$\begin{aligned} \langle x - \sum_{i=1}^{n} \langle x, e_i \rangle e_i, e_j \rangle &= \langle x, e_j \rangle - \sum_{i=1}^{n} \langle x, e_i \rangle \langle e_i, e_j \rangle \\ &= \langle x, e_j \rangle - \sum_{i=1}^{n} \langle x, e_i \rangle \delta_{ij} \\ &= \langle x, e_j \rangle - \langle x, e_j \rangle \\ &= 0 \end{aligned}$$

puisque la base (e_i) est supposée orthonormée. □

Théorème. *Soit E un espace pré-hilbertien, et F un sous-espace de dimension finie. On fixe un vecteur $x \in E$. Alors il existe un seul vecteur y de F tel que $d(x, y)$ soit minimale, et y est l'image de x par le projecteur orthogonal sur F.*

Preuve. Appelons p le projecteur sur F. Quel que soit $y \in F$, on a

$$y - x = p(y - x) + (y - x - p(y - x))$$

et comme les deux vecteurs dans le membre de droite sont orthogonaux, on peut utiliser le théorème de Pythagore pour trouver

$$||y - x||^2 = ||p(y - x)||^2 + ||y - x - p(y - x)||^2$$

Ensuite on observe que $||y - x||^2 \geqslant ||y - x - p(y - x)||^2$ et que

$$y - x - p(y - x) = y - x - p(y) + p(x) = y - x - y + p(x) = p(x) - x$$

Ce vecteur ne dépend donc pas de y, bref

$$||y - x||^2 \geqslant ||p(x) - x||^2$$

ou encore

$$d(x, y) \geqslant ||p(x) - x||$$

L'égalité est atteinte si et seulement si $||p(y - x)|| = 0$ si et seulement si $p(y - x) = 0$ si et seulement si $y = p(x)$. □

Dès lors, on peut définir **la distance** $d(x,F)$ **entre le vecteur** x **et le sous-espace** F comme étant le minimum de la distance $d(x,y)$, où y parcourt F. Cette distance est donc la norme de $p(x)-x$, où p est le projecteur orthogonal sur F.

Théorème (Inégalité de Bessel). *Soit E un espace pré-hilbertien, et $(e_1,\dots,e_n)$ une famille orthonormée finie de vecteurs de E. Alors pour tout $x\in E$, on a l'inégalité de Bessel :*

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

Preuve. Appelons F le sous-espace vectoriel engendré par la famille orthonormée. Alors cette famille est une base de F, et F est de dimension finie. On peut définir le projecteur orthogonal sur F. On a évidemment

$$||x||^2 = ||p(x)||^2 + ||x-p(x)||^2$$

par le théorème de Pythagore, et donc

$$||p(x)||^2 \leqslant ||x||^2$$

Or par un des théorèmes précédents, nous savons que

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

En appliquant à nouveau le théorème de Pythagore, on trouve

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

D'où l'inégalité voulue

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

□

Définition. *Soit E un espace pré-hilbertien, et F un sous-espace vectoriel de dimension finie de E.* **La symétrie orthogonale par rapport à** F *est la symétrie par rapport à F parallèlement au supplémentaire orthogonal $F^\perp$. On rappelle que la symétrie par rapport à la décomposition en somme directe $E=F\oplus F^\perp$ est l'application linéaire*

$$F\oplus F^\perp \to F\oplus F^\perp : x+y \mapsto x-y$$

Le lien entre le projecteur orthogonal p sur F et la symétrie orthogonale s par rapport à F est simple : on a les relations

$$s = 2p - id, \qquad p = \frac{1}{2}(s+id)$$

Définition. Une réflexion d'un espace euclidien E *est une symétrie orthogonale par rapport à un hyperplan de E (un hyperplan de E est un sous-espace vectoriel de dimension* $\dim E-1$*).*

Exemples. Prenons $E=\mathbb{R}^3$, muni du produit scalaire canonique. Il s'agit d'un espace euclidien. Soit F le sous-espace vectoriel de E engendré par les deux vecteurs $(1,1,0)$ et $(0,1,1)$. On peut alors définir $p:\mathbb{R}^3\to\mathbb{R}^3$ le projecteur orthogonal sur F, et s la symétrie orthogonale par rapport à F. Ce sont deux endomorphimes de $\mathbb{R}^3$. Le projecteur p n'est pas une injection, car son noyau est $F^\perp$, qui est de dimension 1. Ce n'est pas une surjection non plus, car son image est $F\neq E$. La symétrie s est un automorphisme, car comme pour toutes les symétries (même non orthogonales) on a $s\circ s = id$. Dans cet exemple, s est une réflexion, car F est un hyperplan de E.

On peut chercher les matrices de p et s dans la base canonique de $\mathbb{R}^3$. D'abord, cherchons la matrice de p. Pour cela, nous allons d'abord remplacer la base $((1,1,0),(0,1,1))$ de F par une base orthonormée de F. Le procédé de Gram-Schmidt nous permet de faire trouver une telle base orthonormée. On prend

$$f_1 = \frac{1}{\sqrt{2}}(1,1,0)$$

puis

$$g_2 = (0,1,1) - \langle (0,1,1), f_1\rangle f_1 = (-1/2, 1/2, 1)$$

Il suffit maintenant de poser $f_2 = g_2/\sqrt{3/2}$ pour obtenir une base orthonormée, à savoir

$$(f_1, f_2) = ((\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (-\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \sqrt{\frac{2}{3}}))$$

Nous savons que pour tout $x \in \mathbb{R}^3$ on a

$$p(x) = \langle x, f_1\rangle f_1 + \langle x, f_2\rangle f_2$$

En remplaçant successivement x par les trois vecteurs de la base canonique de $\mathbb{R}^3$, nous trouvons les trois colonnes de la matrice de p, qui est donc égale à

$$\frac{1}{3}\begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}$$

Une autre méthode consiste à trouver un vecteur qui engendre la droite orthogonale $F^\perp$. On peut pour cela calculer le produit vectoriel

$$(1,1,0) \wedge (0,1,1) = (1,-1,1)$$

Dans la base $((1,1,0),(0,1,1),(1,-1,1))$, la matrice de p est alors $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Pour se ramener à la base canonique de $\mathbb{R}^3$, il suffit alors de calculer

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix}^{-1}$$

Enfin, pour obtenir la matrice de s, il suffit de savoir que $s = 2p - id$, donc la matrice de s est

$$\frac{2}{3}\begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix} - I_3 = \frac{1}{3}\begin{pmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{pmatrix}$$

* *

*

Chapitre 5. Automorphismes orthogonaux. Groupe orthogonal.

Dans tout ce chapitre, il sera supposé que E est un espace euclidien, c'est-à-dire un espace vectoriel de dimension finie sur le corps $\mathbb{R}$ muni d'un produit scalaire.

Définition.
Un automorphisme orthogonal de *E est une application linéaire u de E dans E qui vérifie*

$$\forall x, y \in E, \langle u(x), u(y)\rangle = \langle x, y\rangle$$

A priori, il est abusif d'appeler automorphisme quelque chose qui est seulement défini comme un endomorphisme. Mais il est en fait facile de voir que u est nécessairement un automorphisme. Comme nous sommes sur un espace E de dimension finie, il suffit de comprendre pourquoi u est un endomorphisme injectif : Si $u(x) = 0$, alors en particulier $\langle x, x\rangle = \langle u(x), u(x)\rangle = 0$, ce qui donne tout de suite $x = 0$ puisqu'un produit scalaire est défini positif. Bref, nous avons le droit d'appeler u un automorphisme orthogonal.

Théorème. *Soit E un espace euclidien. Les quatre affirmations suivantes sont équivalentes :*
i) u est un automorphisme orthogonal
ii) u préserve la norme : $\forall x \in E, ||u(x)|| = ||x||$
iii) l'image de toute base orthonormée par u est une base orthonormée
iv) l'image d'une base orthonormée par u est une base orthonormée

Preuve. i) implique évidemment ii). La réciproque est également vraie, comme on peut le voir grâce à l'identité de polarisation

$$2\langle x, y\rangle = ||x + y||^2 - ||x||^2 - ||y||^2$$

Ensuite i) implique iii). En effet, soit $(e_1, \ldots, e_n)$ une base orthonormée quelconque de E. Comme u est un automorphisme orthogonal, on a

$$\langle u(e_i), u(e_j)\rangle = \langle e_i, e_j\rangle = \delta_{ij}$$

Cela signifie que la famille $(u(e_i))$ est une famille orthonormée, et comme c'est l'image d'une base par un automorphisme, c'est aussi une base. Il s'agit bien d'une base orthonormée.

Il est évident que iii) implique iv).

Il ne reste plus qu'à prouver que iv) implique i) : S'il existe une base orthonormée $(e_1, \ldots, e_n)$ qui est envoyée par u sur une base orthonormée, alors u est un automorphisme orthogonal. Prenons deux vecteurs quelconques x, y de E. On introduit les coordonnées de x, y par rapport à la base (e_i) :

$$x = \sum_{i=1}^{n} x_i e_i \qquad y = \sum_{j=1}^{n} y_j e_j$$

Alors

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

Mais on a supposé que la famille $(u(e_1), \ldots, u(e_n))$ est une base orthonormée de E. Donc

$$\langle u(x), u(y)\rangle = \sum_{i=1}^{n} x_i y_i$$

D'un autre côté, il est très facile de voir que

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$$

puisque $(e_1, \ldots, e_n)$ est aussi une base orthonormée. Donc

$$\langle u(x), u(y)\rangle = \langle x, y\rangle$$

□

Une observation essentielle est que la composée de deux automorphismes orthogonaux est encore un automorphismle orthogonal (pourquoi ?) et que la réciproque d'un automorphisme orthogonal est aussi un automorphisme orthogonal. L'ensemble des automorphismes orthogonaux d'un espace euclidien, muni de la loi interne de composition $\circ$, est un groupe.

Définition. *Soit E un espace euclidien. On appelle* **groupe orthogonal de** *E le groupe de tous les automorphismes orthogonaux de E, muni de la composition. On le note $O(E)$.*

Comme il existe un lien étroit entre les endomorphismes et les matrices carrées, nous pouvons aussi définir des matrices orthogonales.

Définition. *Soit A une matrice réelle $n \times n$. On dit que A est* **une matrice orthogonale** *s'il existe un automorphisme orthogonal u d'un espace euclidien E de dimension n et une base orthonormée e de E telle que $A = \mathcal{M}_e(u)$.*

Théorème. *Soit A une matrice réelle $n \times n$. Les 4 affirmations ci-dessous sont équivalentes :*
i) A est une matrice orthogonale.
ii) ${}^tA \cdot A = I$
iii) Les colonnes de A forment une base orthonormée de $\mathcal{M}_{n,1}(\mathbb{R})$ pour le produit scalaire canonique.
iv) Les lignes de A forment une base orthonormée de $\mathcal{M}_{1,n}(\mathbb{R})$ pour le produit scalaire canonique.

Preuve.
Prouvons d'abord l'équivalence entre (i) et (iii).
Soit A une matrice orthogonale. Alors il existe un automorphisme orthogonal u d'un espace euclidien E et une base orthonormée e de E avec $A = \mathcal{M}_e(u)$. Comme l'image de e par u est orthonormée, on peut dire que

$$\forall i, j \in [[1, n]], \langle u(e_i), u(e_j)\rangle = \delta_{ij}$$

où $\delta_{ij} = 1$ si $i = j$, et $\delta_{ij} = 0$ si $i \neq j$. Donc

$$\forall i, j \in [[1, n]], \langle \sum_{k=1}^{n} A_{ki} e_k, \sum_{k=1}^{n} A_{kj} e_k\rangle = \delta_{ij}$$

Nous savons que e est une base orthonormée, ce qui simplifie le membre de gauche en

$$\forall i, j \in [[1, n]], \sum_{k=1}^{n} A_{ki} A_{kj} = \delta_{ij}$$

Ceci signifie que les colonnes de la matrice A forment une base orthonormée pour le produit scalaire canonique sur $\mathcal{M}_{n,1}(\mathbb{R})$. Nous avons $(i) \Rightarrow (iii)$.
Réciproquement, on suppose (iii) vraie. Considérons l'endomorphisme $u : \mathcal{M}_{n,1}(\mathbb{R}) \to \mathcal{M}_{n,1}(\mathbb{R})$ défini par $u(X) = AX$. La matrice de u dans la base canonique e de $\mathcal{M}_{n,1}(\mathbb{R})$ est A, et la base e est orthonormée pour le produit scalaire canonique sur $\mathcal{M}_{n,1}(\mathbb{R})$. Il suffit maintenant de comprendre que u est un automorphisme orthogonal, c'est-à-dire

$$\forall X, Y \in \mathcal{M}_{n,1}(\mathbb{R}), \quad \langle AX, AY\rangle = \langle X, Y\rangle$$

Or

$$\langle AX, AY\rangle = \sum_{k=1}^{n} (AX)_k (AY)_k = \sum_{k,i,j \in [[1,n]]} A_{ki} X_i A_{kj} Y_j = \sum_{i,j \in [[1,n]]} X_i Y_j \left(\sum_{k=1}^{n} A_{ki} A_{kj} \right) =$$

$$= \sum_{i,j \in [[1,n]]} X_i Y_j \delta_{ij} = \sum_{i=1}^{n} X_i Y_i = \langle X, Y\rangle$$

Nous avons maintenant l'équivalence entre (i) et (iii).

Ensuite, quelque soient $i, j \in [[1, n]]$, il est facile de voir que $({}^tA \cdot A)_{ij} = \sum_{k=1}^{n} A_{ki} A_{kj}$ et qu'il s'agit donc du produit scalaire de la colonne i avec la colonne j. Or ce produit scalaire vaut I_{ij} si et seulement si ces colonnes forment une base orthonormée. D'où l'équivalence entre (ii) et (iii).

Il suffit maintenant de montrer que iii) et iv) sont équivalentes. Pour cela il suffit de montrer que si A est orthogonale, alors sa transposée est aussi orthogonale (la réciproque étant évidente puisque la transposition est une involution).

Or, pour A orthogonale on a

$$^tAA = I$$

Cela veut alors dire que A et sa transposée ^tA sont inverses l'une de l'autre, et on a aussi l'égalité

$$A\,^tA = I$$

qui peut encore s'écrire

$$^t(^tA)\,^tA = I$$

Or cela signifie simplement que la transposée de A est aussi orthogonale. □

Au cours de la preuve précédente, nous avons montré le corollaire suivant, permettant de construire explicitement un automorphisme orthogonal à partir d'une matrice orthogonale A.

Corollaire.
Soit A une matrice orthogonale de taille $n \times n$. Alors l'application $u : \mathcal{M}_{n,1}(\mathbb{R}) \to \mathcal{M}_{n,1}(\mathbb{R}) : X \mapsto AX$ est un automorphisme orthogonal de l'espace euclidien $\mathcal{M}_{n,1}(\mathbb{R})$, muni du produit scalaire canonique.

On peut donc définir aussi un groupe de matrices $O(n)$, qui est **le groupe des matrices orthogonales de taille** $n \times n$, muni de la multiplication des matrices. C'est un groupe isomorphe à $O(E)$, où E est un espace euclidien quelconque de dimension n. Pour avoir un isomorphisme de groupes, il suffit de choisir une base orthonormée e de E. L'application

$$\varphi : O(E) \to O(n) : u \mapsto \mathcal{M}_e(u)$$

est alors un isomorphisme de groupes.

Théorème.
Le déterminant de toute matrice orthogonale vaut +1 ou -1.
Le déterminant de tout automomorphisme orthogonal vaut +1 ou -1.

Preuve. Si A est une matrice orthogonale, on peut appliquer le déterminant à l'égalité matricielle

$$^tAA = I$$

pour trouver

$$\det {}^tA \cdot \det A = \det I = 1$$

et comme $\det {}^tA = \det A$, on montre que

$$(\det A)^2 = 1$$

Il en découle $\det A = \pm 1$.

La seconde partie est un corollaire de la première, puisque tout automorphisme orthogonal peut être représenté par une matrice orthogonale. □

Comme le déterminant du produit de deux matrices de déterminant 1 est encore une matrice de déterminant 1, on peut définir un nouveau groupe, qui sera le sous-groupe des éléments de déterminant 1 du groupe orthogonal. Il s'agit d'un groupe, car l'application $O(n) \mapsto \{\pm 1\}$ définie par $A \mapsto \det A$ est un morphisme de groupes (avec la multiplication pour loi sur les deux groupes), et que l'ensemble des matrices de déterminant 1 est le noyau de ce morphisme de groupes.

Définition.
Soit E un espace euclidien. L'ensemble des automorphismes orthogonaux de E de déterminant +1 forme un sous-groupe de $O(E)$ que l'on appelle **groupe orthogonal spécial de** *E et qu'on note $SO(E)$. C'est un sous-groupe distingué de $O(E)$.*
Soit n un entier naturel. L'ensemble des matrices orthogonales $n \times n$ de déterminant +1 forme un sous-groupe de $O(n)$ que l'on appelle **groupe orthogonal spécial de taille** *$n \times n$ et qu'on note $SO(n)$. C'est un sous-groupe distingué de $O(n)$.*

Pour terminer, nous allons rappeler quelques résutats de géométrie pour comprendre les groupes $SO(1), O(1), SO(2), O(2), SO(3), O(3)$.

Le groupe $SO(1)$ n'est pas très intéressant, il a un seul élément, la matrice 1×1 de coefficient 1.

Le groupe $O(1)$ a deux éléments : (1) et (-1). Ces deux matrices correspondent aux deux isométries de la droite qui fixent l'origine (l'identité et la symétrie par rapport à l'origine).

Le groupe $O(2)$ est le groupe de matrices qui correspondent aux isométries du plan fixant l'origine (après les identifications usuelles entre points et vecteurs, et entre applications linéaires et matrices).

Nous savons qu'une isométrie de déterminant 1 qui fixe O est forcément une rotation de centre O. Les matrices associées, de déterminant 1, sont alors de la forme

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

où θ est un réel quelconque (qu'on peut restreindre par exemple à $[0, 2\pi[$ pour des raisons de périodicité). Le groupe $SO(2)$ est donc l'ensemble de toutes ces matrices, muni de la multiplication matricielle. C'est un groupe commutatif :

$$SO(2) = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \middle| \theta \in \mathbb{R} \right\}$$

Le groupe $O(2)$ contient $SO(2)$, mais aussi les matrices orthogonales de déterminant -1. Géométriquement, elles correspondent aux anti-déplacements du plan qui fixent O. Or ce sont exactement les symétries par rapport aux droites passant par O. Nous savons que les matrices correspondantes sont de la forme

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

Le groupe non commutatif $O(2)$ s'écrit alors comme la réunion de deux parties disjointes, dont l'une est $SO(2)$:

$$O(2) = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \middle| \theta \in \mathbb{R} \right\} \cup \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \middle| \theta \in \mathbb{R} \right\}$$

Le groupe $SO(3)$ est le groupe des matrices orthogonales 3×3 de déterminant $+1$. Géométriquement, ce sont les déplacements de l'espace qui fixent l'origine O. Or nous savons que tout déplacement de l'espace est un vissage, et comme O est fixé, un déplacement de l'espace fixant O doit être une rotation de centre O. Si on choisit une base orthonormée directe de sorte que le premier vecteur soit sur l'axe de la rotation, alors la matrice qui représente la rotation dans cette base est de la forme

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

Cela veut dire : pour toute matrice $A \in SO(3)$, il existe une matrice $P \in SO(3)$ et un réel θ telle que

$$A = P^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} P$$

Le cosinus de θ est déterminé seulement par A, puisque deux matrices semblables ont même trace.

Enfin, le groupe $O(3)$ est le groupe de toutes les matrices orthogonales 3×3. Il contient le groupe $SO(3)$. Essayons de décrire les matrices de déterminant -1. Comme la matrice $-I$ est une matrice orthogonale de déterminant -1, on peut écrire toute matrice $A \in O(3)$ avec $\det A = -1$ sous la forme $A = -IB = -B$ avec $B \in SO(3)$.

On peut donc dire

$$O(3) = \left\{ \pm P^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} P \ \middle| \theta \in \mathbb{R}, P \in SO(3) \right\}$$

Chapitre 6. Trace d'une matrice. Trace d'un endomorphisme.

Définition. *Soit A une matrice carrée $n \times n$. On appelle* **trace de la matrice** *A la somme de ses éléments diagonaux :*

$$\operatorname{tr}(A) = \sum_{i=1}^{n} A_{ii}$$

Il convient de remarquer que la trace de la matrice ${}^tA$ est égale à la trace de la matrice A.

L'application qui associé à une matrice A sa trace $\operatorname{tr}(A)$ est une application linéaire de l'espace vectoriel des matrices $n \times n$ vers $\mathbb{R}$. C'est donc une forme linéaire sur $\mathcal{M}_{n,n}(\mathbb{R})$.

La trace de la matrice identité I_n est égale à n.

On rappelle la propriété fondamentale de la trace : si A, B sont deux matrices carrées $n \times n$, alors

$$\operatorname{tr}(AB) = \operatorname{tr}(BA)$$

Cette propriété nous permet alors de définir la trace d'un endomorphisme quelconque d'un espace vectoriel de dimension finie sur un corps commutatif K.

Définition. *Soit u un endomorphisme d'un espace vectoriel E de dimension finie n. On appelle* **trace de** *u la trace de la matrice $A = \mathcal{M}_e(u)$, où e est une base quelconque de E. On la note* $\operatorname{tr} u$.

Afin de s'assurer que la trace de u est bien définie, il faut montrer que

$$\operatorname{tr} \mathcal{M}_e(u) = \operatorname{tr} \mathcal{M}_{e'}(u)$$

Si on appelle P la matrice de passage de la base e à la base e', on a l'égalité matricielle

$$\mathcal{M}_{e'}(u) = P^{-1}\mathcal{M}_e(u)P$$

Donc les traces de ces deux matrices sont égales, et en appliquant la propriété fondamentale de la trace :

$$\operatorname{tr} \mathcal{M}_{e'}(u) = \operatorname{tr}(P^{-1}\mathcal{M}_e(u)P) = \operatorname{tr}(PP^{-1}\mathcal{M}_e(u)) = \operatorname{tr} \mathcal{M}_e(u)$$

On voit bien que la trace de la matrice $\mathcal{M}_e(u)$ ne dépend pas du choix de la base e.

Théorème. *Si u est un endomorphisme diagonalisable, alors la trace de u est la somme des valeurs propres de u, comptées avec leurs multiplicités.*

Preuve. On choisit une base de vecteurs propres e. Alors la matrice $\mathcal{M}_e(u)$ est une matrice diagonale, et sur la diagonale apparaissent les valeurs propres de u. Si λ est une racine d'ordre m du polynôme caractéristique de u, alors λ apparaît m fois sur la diagonale. Comme la trace de u est la somme des éléments diagonaux de $\mathcal{M}_e(u)$, on trouve le résultat. □

Nous allons voir maintenant ce qu'on peut dire en perticulier lorsqu'on travaille **avec une base orthonormée.**

Théorème. *Soit u un endomorphisme d'un espace euclidien E de dimension n, et $e = (e_1, \ldots, e_n)$ une base orthonormée de E. Alors*

$$\forall i, j \in [[1, n]], \quad (\mathcal{M}_e(u))_{ij} = \langle e_i, u(e_j) \rangle$$

Preuve. D'après la définition de la matrice $A = \mathcal{M}_e(u)$ on a

$$\forall j, u(e_j) = \sum_{k=1}^{n} A_{kj} e_k$$

Prenons le produit scalaire de cette égalité avec le vecteur e_i. On obtient

$$\begin{aligned}\langle e_i, u(e_j)\rangle &= \langle e_i, \sum_{k=1}^{n} A_{kj} e_k\rangle \\ &= \sum_{k=1}^{n} A_{kj}\langle e_i, e_k\rangle \\ &= \sum_{k=1}^{n} A_{kj} I_{ik} \\ &= A_{ij}\end{aligned}$$

□

On voit qu'il est très facile de décomposer un vecteur x d'un espace euclidien dans une base orthonormée : la i-ème coordonnée est simplement le produit scalaire $\langle x, e_i\rangle$. On se souvient que si on travaille dans une base quelconque, il est plus difficile de trouver les coordonnées de x, car il faut résoudre un système linéaire à n équations et n inconnues, où n désigne la dimension de l'espace.

Corollaire. *Soit E un espace euclidien, u un endomorphisme de E, et $(e_1, \ldots, e_n)$ une base orthonormée de E. Alors la trace de u est donnée par la formule*

$$\operatorname{tr}(u) = \sum_{i=1}^{n} \langle e_i, u(e_i)\rangle$$

Preuve. Il suffit d'appliquer le théorème précédent à la somme des A_{ii}. □

Observons que la formule pour la trace d'un produit de matrices se généralise à des produits de plusieurs matrices. Ainsi, si A, B, C sont trois matrices $n \times n$, on peut dire

$$\operatorname{tr} ABC = \operatorname{tr} BCA = \operatorname{tr} CAB \qquad \operatorname{tr} BAC = \operatorname{tr} CBA = \operatorname{tr} ACB$$

En revanche, il n'est pas toujours vrai que $\operatorname{tr} ABC = \operatorname{tr} ACB$.

Il convient également de remarquer que la propriété $\operatorname{tr} AB = \operatorname{tr} BA$ se traduit au niveau des endomorphismes : Quels que soient $u, v \in \mathcal{L}(E)$, avec E de dimension finie sur un corps quelconque K, on a

$$\operatorname{tr}(u \circ v) = \operatorname{tr}(v \circ u)$$

Attention, on a $\operatorname{tr}(u \circ v) \neq \operatorname{tr} u \cdot \operatorname{tr} v$ en général. Il ne faut pas confondre les propriétés de la trace avec celles du déterminant.

Mais il est toujours vrai que la trace d'une somme d'endomorphismes est la somme des traces. Cela découle de la linéarité de l'application «trace» sur l'espace des matrices :

$$\operatorname{tr}(u + v) = \operatorname{tr} u + \operatorname{tr} v$$

Chapitre 7. Adjoint d'un endomorphisme.

Dans ce chapitre, tous les espaces vectoriels considérés sont euclidiens (de dimension finie sur le corps $\mathbb{R}$ des réels, et munis d'un produit scalaire).

Théorème. *Soit u un endomorphisme d'un espace euclidien E. Alors il existe une seule application $u^* : E \to E$ telle que*

$$\forall x, y \in E, \langle u(x), y\rangle = \langle x, u^*(y)\rangle$$

En outre, cette application est linéaire, donc u^* est un endomorphisme.

Preuve. Fixons $y \in E$. L'application

$$E \to \mathbb{R} : x \mapsto \langle u(x), y\rangle$$

est une forme linéaire sur E. C'est donc un élément de E^*. Or nous savons que grâce au produit scalaire, nous pouvons construire un isomorphisme entre E et E^*. Cela veut dire qu'il existe un unique vecteur $a_y \in E$ tel que

$$\forall x \in E, \langle u(x), y\rangle = \langle x, a_y\rangle$$

Nous définissons alors l'application u^* par

$$u^*(y) := a_y$$

Cette application est bien définie, et vérifie

$$\forall x, y \in E, \langle u(x), y\rangle = \langle x, u^*(y)\rangle$$

Il reste à montrer que u^* est linéaire. Soient y, y' deux vecteurs quelconques de E. Alors

$$\begin{aligned}\langle x, u^*(y+y')\rangle &= \langle u(x), y+y'\rangle \\ &= \langle u(x), y\rangle + \langle u(x), y'\rangle \\ &= \langle x, u^*(y)\rangle + \langle x, u^*(y')\rangle \\ &= \langle x, u^*(y) + u^*(y')\rangle\end{aligned}$$

Donc le vecteur $u^*(y+y') - u^*(y) - u^*(y')$ est orthogonal à tous les vecteurs x de E. Il est nécessairement nul. Nous avons prouvé que pour y, y' vecteurs quelconques,

$$u^*(y+y') = u^*(y) + u^*(y')$$

On montre de manière analogue que pour tout réel k, on a

$$u^*(ky) = ku^*(y)$$

L'application u^* est bien un endomorphisme de E. □

Définition. *Soit E un espace euclidien, et u un endomorphisme de E. On appelle* **l'adjoint** *de u l'endomorphisme u^* du théorème précédent.*

Théorème. *Soit E un espace euclidien.*
Pour tout endomorphisme u de E, on a $u^{**} = u$
Pour tous endomorphismes u, v de E, on a $(u \circ v)^* = v^* \circ u^*$

Preuve. Quels que soient les vecteurs x, y de E,

$$\begin{aligned}\langle u(x), y\rangle &= \langle x, u^*(y)\rangle \\ &= \langle u^*(y), x\rangle \\ &= \langle y, u^{**}(x)\rangle \\ &= \langle u^{**}(x), y\rangle\end{aligned}$$

Le vecteur $u(x) - u^{**}(x)$ est donc orthogonal à tous les vecteurs y de E, ce qui implique que ce vecteur est nul. Il en découle que pour tout vecteur x on a

$$u(x) = u^{**}(x)$$

ce qui établit la première partie du théorème.

On a d'autre part

$$\begin{aligned}\langle (u\circ v)^*(x), y\rangle &= \langle x, u(v(y))\rangle \\ &= \langle u^*(x), v(y)\rangle \\ &= \langle v^*(u^*(x)), y\rangle \\ &= \langle (v^*\circ u^*)(x), y\rangle\end{aligned}$$

Cette égalité étant vraie pour tous les vecteurs x, y, on voit que le vecteur

$$(u\circ v)^*(x) - (v^*\circ u^*)(x)$$

est nul. La deuxième partie est prouvée. □

Nous allons nous intéresser maintenant au noyau, à l'image et au rang de l'adjoint d'un endomorphisme

Théorème.
Le noyau de u^* est le sous-espace orthogonal à l'image de u.
L'image de u^* est le sous-espace orthogonal au noyau de u.
Le rang de u^* est égal au rang de u.

Preuve. Le vecteur y est dans le noyau de u^* si et seulement si $u^*(y) = 0$. C'est équivalent à :

$$\forall x \in E, \langle u^*(y), x\rangle = 0$$

Par la définition de l'adjoint, c'est la même chose que

$$\forall x \in E, \langle y, u(x)\rangle = 0$$

Cette phrase signifie que y est orthogonal à tous les vecteurs de l'image de u. Donc

$$\ker u^* = (\mathrm{Im}\ u)^\perp$$

Montrons maintenant que l'image de u^* est l'orthogonal du noyau de u. Pour commencer, choisissons un élément y de l'image de u^*. Alors il existe z avec $y = u^*(z)$. Or pour tout $x \in \ker u$, on a

$$\langle y, x\rangle = \langle u^*(z), x\rangle = \langle z, u(x)\rangle = \langle z, 0\rangle = 0$$

Donc y est bien un élément de l'orthogonal du noyau de u. Nous avons prouvé que l'image de u^* est un sous-espace vectoriel de $(\ker u)^\perp$.

Appelons n la dimension de E. Nous savons que la dimension de l'image de u^* est égale à $n - \dim\ker u^*$ par le théorème du rang. Or $\ker u^*$ est l'orthogonal de l'image de u, donc $\dim\ker u^* = n - \dim \mathrm{Im}\ u$. Il en découle que l'image de u^* et l'image de u ont la même dimension. Cela prouve en passant la troisième partie du théorème sur l'égalité des rangs.

Par le théorème du rang, la dimension de l'image de u vaut $n - \dim\ker u$, ce qui est égal à la dimension de $(\ker u)^\perp$. Nous avons donc prouvé que les deux sous-espaces

$$\mathrm{Im}\ u^* \qquad \text{et} \qquad (\ker u)^\perp$$

ont la même dimension. Comme le premier est un sous-espace vectoriel du second, on peut conclure qu'ils sont égaux. □

Définition.
Soir E un espace vectoriel quelconque sur un corps commutatif quelconque. On dit qu'un sous-espace vectoriel F de E est **stable par l'endomorphisme** *u si $u(F) \subset F$.*

L'adjoint de u a une jolie propriété par rapport aux sous-espaces vectoriel stables par u.

Théorème. *Soit E un espace euclidien, et u un endomorphisme.*
Un sous-espace vectoriel F est stable par u si et seulement si $F^\perp$ est stable par u^*.

Preuve. «seulement si» : Soit x un élément de $F^\perp$ et montrons que $u^*(x) \in F^\perp$. Cela revient à montrer

$$\forall y \in F, \langle u^*(x), y\rangle = 0$$

Or $\langle u^*(x), y\rangle = \langle x, u(y)\rangle$. Nous avons pris y dans F, alors par hypothèse, on a $u(y) \in F$. comme x est dans l'orthogonal de F, on a forcément

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

La propriété est démontrée.

Pour montrer la réciproque, nous supposons que $F^\perp$ est stable par u^*. Alors par la première implication, nous pouvons dire que $F^{\perp\perp}$ est stable par u^{**}. Mais nous savons que $F^{\perp\perp} = F$ et $u^{**} = u$. □

Théorème. *Soit u un endomorphisme d'un espace euclidien E. Soit $e = (e_1, \ldots, e_n)$ une base orthonormée de E. Alors*

$$\mathcal{M}_e(u^*) = {}^t\mathcal{M}_e(u)$$

Preuve. Soient i, j deux entiers dans $[[1, n]]$. Alors

$$\begin{aligned}(\mathcal{M}_e(u^*))_{ij} &= \langle e_i, u^*(e_j)\rangle \\ &= \langle u(e_i), e_j\rangle \\ &= \mathcal{M}_e(u)_{ji} \\ &= ({}^t\mathcal{M}_e(u))_{ij}\end{aligned}$$

Il en découle que

$$\mathcal{M}_e(u^*) = {}^t\mathcal{M}_e(u)$$

□

Au niveau des matrices, le passage de u à u^* se traduit par le passage de A à ${}^tA$ (dans une base orthonormée). On retrouve alors facilement les propriétés

$$u^{**} = u \qquad \text{et} \qquad (u \circ v)^* = v^* \circ u^*$$

Corollaire. *Soit u un endomorphisme d'un espace euclidien E. Alors*

$$\operatorname{tr}(u^*) = \operatorname{tr}(u)$$

et

$$\det(u^*) = \det u$$

Preuve. Ces égalités sont vraies, parce que nous savons que pour toute matrice carrée A, on a

$$\operatorname{tr}(A) = \operatorname{tr}({}^tA)$$

$$\det(A) = \det({}^tA)$$

□

Définition.
Un endomorphisme u d'un espace euclidien E est appelé **autoadjoint ou symétrique** *si*

$$u^* = u$$

Cele revient à dire que

$$\forall x, y \in E, \langle u(x), y\rangle = \langle x, u(y)\rangle$$

Théorème (caractérisation des projecteurs orthogonaux).
Un endomorphisme p de E est un projecteur orthogonal si et seulement si $p^2 = p$ et $p^* = p$.

Preuve. «seulement si». Comme p est un projecteur, on a clairement $p^2 = p$. Comme p est un projecteur orthogonal, les deux sous-espaces définissant p sont des sous-espaces vectoriels orthogonaux. De $p^2 = p$ on déduit que $(p \circ p)^* = p^*$, et donc $(p^*)^2 = p^*$. Il en découle que p^* est un projecteur. Or le noyau de p^* est l'orthogonal de l'image de p, qui est à son tour l'orthogonal du noyau de p. Donc

$$\ker p^* = \ker p$$

On prouve de la même façon que

$$\text{Im } p^* = \text{Im } p$$

Or p, p^* sont deux projecteurs, et comme ils ont tous deux le même noyau et la même image, ces deux projecteurs sont égaux. Bref $p = p^*$.

«si» : Nous savons déjà qu'un endomorphisme p vérifiant $p^2 = p$ est un projecteur. Il reste à montrer que le noyau de p et l'image de p sont des sous-espaces orthogonaux. Pour cela nous utiliserons $p^* = p$. Le noyau de p est l'orthogonal de l'image de p^*, donc l'orthogonal de l'image de p. Les deux sous-espaces vectoriels définissant p sont bien orthogonaux. Bref, p est un projecteur orthogonal. □

On peut donc dire que les seuls projecteurs autoadjoints sont les projecteurs orthogonaux.

Définition. *On dit qu'un endomorphisme u d'un espace euclidien E est* **autoadjoint positif** *si*
(i) u est autoadjoint
(ii) $\forall x \in E, \langle u(x), x\rangle \geqslant 0$
On dit qu'un endomorphisme u d'un espace euclidient E est **autoadjoint défini positif** *si*
(i) u est autoadjoint positif
(ii) $\forall x \in E, (\langle u(x), x\rangle = 0 \Rightarrow x = 0)$

L'identité sur E est un endomorphisme auto-adjoint défini positif.

Tout projecteur orthogonal est un endomorphisme auto-adjoint positif. En effet, prenons $x \in E = F \oplus F^\perp$, et décomposons $x = a + b$ avec $a \in F, b \in F^\perp$. Alors

$$\langle p(x), x\rangle = \langle p(a+b), a+b\rangle = \langle a, a+b\rangle = ||a||^2 \geqslant 0$$

Nous finissons avec une manière d'exprimer «être un automorphisme orthogonal» à l'aide de l'adjoint.

Théorème.
Un endomorphisme u d'un espace euclidien E est un automorphisme orthogonal si et seulement si

$$u^* \circ u = u \circ u^* = id_E$$

Preuve. u est un automorphisme orthogonal équivaut à :

$$\forall x, y \in E, \langle u(x), u(y)\rangle = \langle x, y\rangle$$

ce qui peut encore s'écrire

$$\forall x, y \in E, \langle x, u^*(u(y))\rangle = \langle x, y\rangle$$

Il faut et il suffit alors que

$$u^*(u(y)) = y$$

ou encore

$$u^* \circ u = id_E$$

Or, en dimension finie, ceci équivaut aux deux relations

$$u^* \circ u = u \circ u^* = id_E$$

□

Evidemment, on reconnaît ici la traduction (au niveau des endomorphismes) de la relation matricielle ${}^tAA = A{}^tA = I$ pour les matrices orthogonales.

Enfin, remarquons qu'un endomorphisme u est auto-adjoint si et seulement si, pour toute base **orthonormée** e de E, la matrice associée à u dans la base e est une matrice symétrique, autrement dit :

$$\mathcal{M}_e(u) = {}^t\mathcal{M}_e(u)$$

Attention : Cette équivalence est valable si on travaille dans une base **orthonormée**, mais en général elle n'est plus vraie. Prenons en effet l'endomorphisme

$$u : \mathbb{R}^2 \to \mathbb{R}^2 : (x_1, x_2) \mapsto (x_1, 0)$$

Si on munit $\mathbb{R}^2$ du produit scalaire euclidien, on voit de suite que u est auto-adjoint, puisque pour deux vecteurs $a = (a_1, a_2)$ et $b = (b_1, b_2)$, on a

$$\langle u(a), b\rangle = a_1 b_1 = \langle a, u(b)\rangle$$

Travaillons maintenant avec la base $e = ((1,0),(1,1))$ de $\mathbb{R}^2$. Nous observons que e n'est pas une base orthogonale. La matrice $\mathcal{M}_e(u)$ est alors

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

et ce n'est pas une matrice symétrique.

Chapitre 8. Le théorème spectral

Dans ce chapitre, E désigne un espace euclidien. On rappelle qu'il s'agit d'un espace vectoriel sur $\mathbb{R}$ de dimension finie, et muni d'un produit scalaire.

Nous avons vu au chapitre 7 les endomorphismes auto-adjoints (aussi appelés symétriques), que l'on peut définir par la propriété suivante :

$$\forall x, y \in E, \langle u(x), y\rangle = \langle x, u(y)\rangle$$

ou encore, avec des notations plus légères,

$$u^* = u$$

Le théorème spectral, qui est l'objet de ce chapitre, est l'affirmation suivante :
Tout endomorphisme auto-adjoint d'un espace euclidien est diagonalisable dans une base orthonormée. Cela signifie qu'on peut toujours trouver une base orthonormée de l'espace euclidien par rapport à laquelle la matrice de l'endomorphisme est diagonale.

Nous avons déjà rencontré l'idée essentielle de la preuve lors de la classification des quadriques dans le cours de géométrie. Étant donnée l'importance de ces idées, nous allons les revisiter.

Tout d'abord, démontrons que tous les endomorphismes d'un espace euclidien sont des applications continues.

Lemme. *Soit E un espace euclidien, et u un endomorphisme quelconque de E.*
Alors il existe un réel C tel que pour tout vecteur x de E, on ait l'inégalité

$$||u(x)|| \leqslant C||x||$$

Il en découle que u est une application continue de E dans E.

Preuve. Choisissons une base orthonormée e. On appelle A la matrice carrée $n \times n$ (où n est la dimension de E) qui représente u dans la base e. On appelle $x_1, \ldots, x_n$ les coordonnées du vecteur x dans la base e.

$$\begin{aligned} ||u(x)||^2 &= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} A_{ij} x_j \right)^2 \\ &\leqslant \sum_{i=1}^{n} \left(\sum_{j=1}^{n} |A_{ij}||x_j| \right)^2 \end{aligned}$$

Appelons M le maximum de l'ensemble $\{|A_{ij}|\}$. C'est le plus grand coefficient de la matrice (en valeur absolue). Alors

$$||u(x)||^2 \leqslant \sum_{i=1}^{n} \left(\sum_{j=1}^{n} M|x_j| \right)^2 = nM^2 \left(\sum_{j=1}^{n} 1 \cdot |x_j| \right)^2 \leqslant nM^2 \left(\sum_{j=1}^{n} 1^2 \right) \cdot \left(\sum_{j=1}^{n} x_j^2 \right) = n^2 M^2 ||x||^2$$

On peut alors prendre $C = Mn$ et l'existence de C est garantie.

On peut en déduire immédiatement la continuité de u, puisque pour x, y des vecteurs quelconques de E, on aura

$$||u(x) - u(y)|| = ||u(x - y)|| \leqslant C||x - y||$$

Donc u est C-lipschitzienne. Cette inégalité entraîne clairement la continuité de l'application u. □

Le résultat suivant est essentiel pour la suite. Il utilise un peu d'analyse.

Lemme. *Soit u un endomorphisme auto-adjoint sur un espace euclidien E, avec $E \neq \{0\}$.*
Alors il existe au moins une valeur propre de u.

Preuve. On appelle S la sphère de centre 0 et de rayon 1, c'est-à-dire l'ensemble des vecteurs de norme 1. L'ensemble S est un ensemble borné et fermé (la démonstration est un bon exercice), donc un ensemble compact. On définit maintenant l'application

$$P : S \to \mathbb{R} : x \mapsto \langle x, u(x)\rangle$$

Cette application est continue. Pour le voir, il suffit de démontrer que l'application prolongée $E \to \mathbb{R}$ envoyant x sur $\langle x, u(x)\rangle$ est une application continue de E vers $\mathbb{R}$.

Or on voit que pour x, h des vecteurs de E :

$$\begin{aligned} |\langle x+h, u(x+h)\rangle - \langle x, u(x)\rangle| &= |\langle h, u(x+h)\rangle + \langle x, u(h)\rangle| \\ &\leqslant |\langle h, u(x+h)\rangle| + |\langle x, u(h)\rangle| \\ &\leqslant ||h|| \cdot ||u(x+h)|| + ||x|| \cdot ||u(h)|| \end{aligned}$$

La dernière majoration est l'inégalité de Cauchy-Schwarz. On utilise maintenant le lemme précédent :

$$\begin{aligned} |\langle x+h, u(x+h)\rangle - \langle x, u(x)\rangle| &\leqslant C||h|| \cdot ||x+h|| + C||x|| \cdot ||h|| \\ &= C||h||(||x+h|| + ||x||) \\ &\leqslant C||h||(2||x|| + ||h||) \end{aligned}$$

on voit clairement que lorsque $||h||$ tend vers zéro, alors $C||h||(2||x|| + ||h||)$ tend vers zéro. Il en découle que, pour x fixé,

$$\lim_{||h|| \to 0} |\langle x+h, u(x+h)\rangle - \langle x, u(x)\rangle| = 0$$

Ceci prouve que l'application P est continue.

Or toute application continue définie sur un compact non vide atteint sa borne supérieure. Or dans un espace euclidien non réduit à $\{0\}$, la sphère S est non vide. Donc il existe un vecteur x_0 de norme 1 tel que

$$\forall x \in S, P(x) \leqslant P(x_0)$$

Prenons un vecteur $y \in S$ qui soit en plus orthogonal à x_0. On définit l'application suivante :

$$f : \mathbb{R} \to \mathbb{R} : t \mapsto P(\cos t \cdot x_0 + \sin t \cdot y)$$

Observons que le vecteur $\cos t \cdot x_0 + \sin t \cdot y$ appartient à S par le théorème de Pythagore et l'identité classique $\cos^2 t + \sin^2 t = 1$. En outre, d'après la définition de x_0, la fonction f a un maximum en $t = 0$.
Calculons $f(t)$:

$$\begin{aligned} f(t) &= \langle \cos t \cdot x_0 + \sin t \cdot y, u(\cos t \cdot x_0 + \sin t \cdot y)\rangle \\ &= \langle \cos t \cdot x_0 + \sin t \cdot y, \cos t \cdot u(x_0) + \sin t \cdot u(y)\rangle \\ &= \cos^2 t \cdot f(0) + \cos t \sin t \langle x_0, u(y)\rangle + \cos t \sin t \langle y, u(x_0)\rangle + \sin^2 t \langle y, u(y)\rangle \\ &= \cos^2 t \cdot f(0) + 2\cos t \sin t \langle u(x_0), y\rangle + \sin^2 t \langle y, u(y)\rangle \end{aligned}$$

La dernière égalité s'obtient en utilisant le caractère auto-adjoint de u.
En raison du maximum en $t = 0$, on peut dire que $f'(0) = 0$. Le calcul montre que

$$f'(t) = 2\cos t \sin t \cdot f(0) + 2(\cos^2 t - \sin^2 t)\langle u(x_0), y\rangle + 2\sin t \cos t \langle y, u(y)\rangle$$

En posant $t = 0$ on trouve

$$0 = 2\langle u(x_0), y\rangle$$

Donc y et $u(x_0)$ sont orthogonaux.
Comme $y \in (\mathbb{R}x_0)^\perp$ est quelconque, on a donc prouvé que

$$u(x_0) \in (\mathbb{R}x_0)^{\perp\perp}$$

Nous savons que $(\mathbb{R}x_0)^{\perp\perp} = \mathbb{R}x_0$. Donc $u(x_0) \in \mathbb{R}x_0$. Cela veut dire qu'il existe un réel λ tel que

$$u(x_0) = \lambda x_0$$

Nous avons montré l'existence d'une valeur propre λ (puisque le vecteur x_0 est évidemment non nul). □

Lorsque nous aurons vu les espaces hermitiens, nous donnerons une autre preuve de ce résultat, en nous servant du théorème de Gauss-d'Alembert.

Maintenant nous sommes prêts à énoncer le théorème spectral pour les endomorphismes auto-adjoints d'un espace euclidien.

Théorème spectral. *Soit u un endomorphisme auto-adjoint d'un espace euclidien E. Alors il existe une base orthonormée de vecteurs propres de u. L'endomorphisme u est donc diagonalisable (par rapport à une base orthonormée).*

Preuve. On prouve le résultat par récurrence sur la dimension de E. Si E est de dimension nulle, il n'y a rien à prouver. Supposons que le résultat a été établi pour les espaces euclidiens de dimension $n-1$.

Soit E de dimension n. On sait par le lemme qui précède qu'il existe un vecteur u_0 unitaire (donc non nul) qui est un vecteur propre de valeur propre λ. On appelle F le sous-espace vectoriel de dimension 1 engendré par ce vecteur u_0. Comme F est stable par u, son sous-espace orthogonal $F^\perp$ (de dimension $n-1$) est stable par l'adjoint u^*. Or $u^* = u$, donc $F^\perp$ est stable par u. On peut alors définir la restriction v de u à ce sous-espace de dimension $n-1$. v est alors un endomorphisme auto-adjoint de $F^\perp$.

Par l'hypothèse de récurrence, il existe une base orthonormée $(g_1, \ldots, g_{n-1})$ de $F^\perp$ qui sont tous des vecteurs propres de v, et donc de u. Bien entendu, la famille

$$(u_0, g_1, \ldots, g_{n-1})$$

est une base orthonormée de E. C'est aussi une famille de vecteurs propres. Le théorème spectral est démontré. □

Du côté des matrices, le théorème spectral est traduit comme ceci :

Théorème spectral matriciel. *Soit A une matrice symétrique $n \times n$. Alors la matrice A est diagonalisable. En outre, il existe une matrice orthogonale $G \in O(n)$ telle que la matrice $G^{-1}AG$ soit diagonale.*

Preuve. Il suffit de relire l'énoncé précédent. □

Corollaire. *Toute matrice symétrique A peut s'écrire sous la forme*

$$A = G^{-1}DG = {}^tGDG$$

avec G une matrice orthogonale, et D une matrice diagonale.

Preuve. Triviale. □

Nous allons illustrer le théorème spectral en cherchant à diagonaliser la matrice symétrique

$$A = \begin{pmatrix} \frac{4}{3} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{18}} \\ -\frac{1}{\sqrt{6}} & \frac{3}{2} & -\frac{1}{\sqrt{12}} \\ \frac{1}{\sqrt{18}} & -\frac{1}{\sqrt{12}} & \frac{7}{6} \end{pmatrix}$$

Commençons par calculer le polynôme caractéristique de la matrice A :

$$\det(XI_3 - A) = (X - \frac{4}{3})(X - \frac{3}{2})(X - \frac{7}{6}) - \frac{2}{\sqrt{6 \cdot 12 \cdot 18}} - \frac{1}{18}(X - \frac{3}{2}) - \frac{1}{12}(X - \frac{4}{3}) - \frac{1}{6}(X - \frac{7}{6})$$

Après développement on trouve

$$\det(XI_3 - A) = X^3 - 4X^2 + 5X - 2$$

Ce polynôme se factorise

$$X^3 - 4X^2 + 5X - 2 = (X-1)^2(X-2)$$

Le spectre de la matrice A est donc $\{1, 2\}$. Cherchons maintenant pour chacune des deux valeurs propres le sous-espace propre correspondant.

Le cas $\lambda = 1$. Il faut résoudre le système

$$\left\{\begin{array}{rcl} \frac{x}{3} - \frac{y}{\sqrt{6}} + \frac{z}{\sqrt{18}} & = & 0 \\ -\frac{x}{\sqrt{6}} + \frac{y}{2} - \frac{z}{\sqrt{12}} & = & 0 \\ \frac{x}{\sqrt{18}} - \frac{y}{\sqrt{12}} + \frac{z}{6} & = & 0 \end{array}\right.$$

Il n'est pas difficile de voir que ces trois équations sont équivalentes. On peut par exemple supprimer les équations (2) et (3) pour ne garder que

$$\frac{x}{3} - \frac{y}{\sqrt{6}} + \frac{z}{\sqrt{18}} = 0$$

Le sous-espace propre E_1 est donc de dimension 2. C'est le plan vectoriel normal au vecteur

$$(\sqrt{2}, -\sqrt{3}, 1)$$

Traitons maintenant le cas $\lambda = 2$. Il faut résoudre le système

$$\left\{\begin{array}{rcl} -\frac{2x}{3} - \frac{y}{\sqrt{6}} + \frac{z}{\sqrt{18}} & = & 0 \\ -\frac{x}{\sqrt{6}} - \frac{y}{2} - \frac{z}{\sqrt{12}} & = & 0 \\ \frac{x}{\sqrt{18}} - \frac{y}{\sqrt{12}} - \frac{5z}{6} & = & 0 \end{array}\right.$$

La résolution de ce système non inversible - laissée en exercice - donne alors

$$\left\{\begin{array}{rcl} x & = & \sqrt{2}z \\ y & = & -\sqrt{3}z \end{array}\right.$$

En d'autres mots

$$E_2 = Vect(\sqrt{2}, -\sqrt{3}, 1)$$

Il se trouve donc que E_1 et E_2 sont deux sous-espaces supplémentaires de $\mathbb{R}^3$, et qu'en outre, ces sous-espaces sont orthogonaux. Ce n'est pas un accident. C'est une conséquence du théorème spectral, puisque la matrice A est symétrique.

On peut donc diagonaliser la matrice A dans une base orthonormée de vecteurs propres. Choisissons d'abord un vecteur propre unitaire dans E_2. On peut prendre par exemple

$$\frac{1}{\sqrt{6}}(\sqrt{2}, -\sqrt{3}, 1) = (\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{6}})$$

Ensuite, choisissons une base orthonormée de E_1. On peut par exemple prendre pour premier vecteur

$$(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{6}})$$

car il est orthogonal à E_2 et de norme 1.

On peut ensuite choisir comme deuxième vecteur de base de E_1 le produit vectoriel :

$$(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{6}}) \wedge (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{6}}) = (-\frac{1}{\sqrt{3}}, 0, \frac{2}{\sqrt{6}})$$

On a alors construit un exemple d'une base orthonormée de $\mathbb{R}^3$ formée d'un vecteur propre de valeur propre 2 et de deux vecteurs propres de valeur propre 1. Ce sont

$$\begin{array}{rcl} e_1 & = & (\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{6}}) \\ e_2 & = & (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{6}}) \\ e_3 & = & (-\frac{1}{\sqrt{3}}, 0, \frac{2}{\sqrt{6}}) \end{array}$$

La matrice de passage de la base canonique de $\mathbb{R}^3$ à cette base e est donc la matrice orthogonale suivante :

$$P = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{pmatrix}$$

Nous avons prouvé que

$$A = \begin{pmatrix} \frac{4}{3} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{18}} \\ -\frac{1}{\sqrt{6}} & \frac{3}{2} & -\frac{1}{\sqrt{12}} \\ \frac{1}{\sqrt{18}} & -\frac{1}{\sqrt{12}} & \frac{7}{6} \end{pmatrix} = P \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} P^{-1} = P \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} {}^tP$$

ou, en termes beaucoup plus explicites

$$\begin{pmatrix} \frac{4}{3} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{18}} \\ -\frac{1}{\sqrt{6}} & \frac{3}{2} & -\frac{1}{\sqrt{12}} \\ \frac{1}{\sqrt{18}} & -\frac{1}{\sqrt{12}} & \frac{7}{6} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}$$

Le lecteur avide de calculs est invité à vérifier cette égalité.

***

On peut décrire le spectre des endomorphismes auto-adjoints positifs (respectivement définis positifs).

Théorème.
Le spectre de tout endomorphisme auto-adjoint positif contient seulement des réels positifs.
Le spectre de tout endomorphisme auto-adjoint défini positif contient seulement des réels strictement positifs.

Preuve. Soit u auto-adjoint positif, et λ une valeur propre de u. Choisissons un vecteur propre non nul x associé à λ. On a alors

$$0 \leqslant \langle u(x), x\rangle = \langle \lambda x, x\rangle = \lambda ||x||^2$$

Comme $||x||^2 > 0$ par hypothèse, on en déduit que λ doit être positif.

Si on suppose en plus que u est défini positif, alors on a l'inégalité stricte

$$0 < \langle u(x), x\rangle$$

et on en déduit comme ci-dessus que $\lambda > 0$.

□

Lorsque nous avons étudié les quadriques, nous avons associé à l'équation d'une quadrique un endomorphisme auto-adjoint de $\mathbb{R}^3$. Cette construction va maintenant être généralisée :

Définition. *Soit ω une forme bilinéaire symétrique (pas forcément positive) sur un espace euclidien E. On appelle* **endomorphisme auto-adjoint associé à** *ω l'application $u : E \to E$ définie par*

$$\forall x, y \in E, \langle u(x), y\rangle = \omega(x, y)$$

Quelques remarques à propos de l'existence de u : Pour $x \in E$ fixé, l'application $y \mapsto \omega(x, y)$ est une forme linéaire sur E, donc de la forme $y \mapsto \langle a, y\rangle$ pour un unique vecteur $a \in E$. Ce vecteur a est $u(x)$. L'application u existe donc. Il est facile de démontrer que u est linéaire. C'est donc un endomorphisme.

Enfin, u est auto-adjoint, car

$$\langle u(x), y\rangle = \omega(x,y) = \omega(y,x) = \langle u(y), x\rangle = \langle x, u(y)\rangle$$

Exemple. Sur $\mathbb{R}^4$ muni du produit scalaire canonique, considérons la forme bilinéaire symétrique

$$\omega : ((x_1,x_2,x_3,x_4),(x_1',x_2',x_3',x_4')) \mapsto x_1x_1' + 2x_2x_3' + 2x_3x_2' - x_2x_4' - x_4x_2' + 3x_4x_4'$$

La matrice associée à cette forme bilinéaire symétrique par rapport à la base canonique est

$$G(e) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & -1 \\ 0 & 2 & 0 & 0 \\ 0 & -1 & 0 & 3 \end{pmatrix}$$

L'endomorphisme auto-adjoint associé à ω a pour matrice par rapport à la base canonique est alors la même matrice :

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & -1 \\ 0 & 2 & 0 & 0 \\ 0 & -1 & 0 & 3 \end{pmatrix}$$

Grâce au théorème spectral, nous pouvons toujours trouver une base orthonormée dans laquelle une forme bilinéaire symétrique prend une forme particulièrement simple.

Théorème. *Soit E un espace euclidien, et ω une forme bilinéaire symétrique sur E. Alors il existe une base orthonormée $(e_1,\dots,e_n)$ de E telle que*

$$\forall i,j \in [[1,n]], (i \neq j \Rightarrow \omega(e_i,e_j) = 0)$$

Preuve. On note u l'endomorphisme auto-adjoint associé à la forme bilinéaire symétrique ω. D'après le théorème spectral, il existe une base orthonormée $(e_1,\dots,e_n)$ dans laquelle la matrice de u est diagonale. Soit A cette matrice. Or nous savons, grâce au fait que e est une base orthonormée, que

$$A_{ij} = \langle u(e_j), e_i\rangle$$

Donc

$$A_{ij} = \omega(e_j,e_i) = \omega(e_i,e_j)$$

La matrice A est diagonale, donc $i \neq j \Rightarrow A_{ij} = 0$. L'énoncé est démontré. □

Théorème. *Soit ω une forme bilinéaire symétrique définie sur un espace vectoriel réel E de dimension finie n, et $e = (e_1,\dots,e_n)$ une base quelconque de E. La matrice $\Omega(e)$ est la matrice $n \times n$ définie par*

$$\Omega(e)_{ij} = \omega(e_i,e_j)$$

Alors le rang de $\Omega(e)$ ne dépend pas de la base e.

Preuve. Nous savons déjà que

$$\Omega(e') = {}^tP\Omega(e)P$$

où P est la matrice de passage de la base e à la base e'. Or P (et donc sa transposée) est une matrice inversible. Par conséquent, les deux matrices $\Omega(e)$ et $\Omega(e')$ sont équivalentes, ce qui suffit pour dire qu'elles ont même rang. □

Définition. Le rang d'une forme bilinéaire symétrique *sur un espace vectoriel réel de dimension finie est le rang de la matrice $\Omega(e)$, où e est une base quelconque de l'espace vectoriel (la proposition précédente montre que cette définition n'est pas ambiguë).*
Le rang d'une forme quadratique *sur un espace vectoriel réel de dimension finie est le rang de sa forme bilinéaire associée.*

Exemples.
1. Le rang du produit scalaire canonique sur $\mathbb{R}^n$ est égal à n.

2. Le rang de la forme quadratique de Minkowski sur $\mathbb{R}^4$ définie par

$$q(x,y,z,t) = x^2 + y^2 + z^2 - t^2$$

est égal à 4.

3. Le rang de la forme quadratique sur $\mathbb{R}^5$ défine par

$$q(a,b,c,d,e) = a^2 + 2c^2 - e^2$$

est égal à 3.

> **Définition.** *On dit qu'une forme bilinéaire symétrique (ou une forme quadratique) définie sur un espace vectoriel réel de dimension n est* **non dégénérée** *si son rang est égal à n. Sinon, on dit qu'elle est* **dégénérée**.

Exemple. En géométrie de l'espace, les quadriques correspondant aux formes quadratiques non dégénérées (c'est-à-dire de rang 3) sont d'un des types suivants :
1. ellipsoïde
2. hyperboloïde à une nappe
3. hyperboloïde à deux nappes
4. cône
5. un point
6. le vide

D'autres quadriques, comme par exemple le paraboloïde hyperbolique d'équation $x^2 - y^2 - z = 0$, sont associées à des formes quadratiques dégénérées. Dans le cas du paraboloïde hyperbolique, la forme quadratique associée est $q(x,y,z) = x^2 - y^2$, de matrice $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Cette matrice est seulement de rang 2.

***

> **Théorème.** *Une forme bilinéaire symétrique ω sur un espace vectoriel reél E de dimension finie est non dégénérée si et seulement si*
>
> $$\{x \in E | \forall y \in E, \omega(x,y) = 0\} = \{0\}$$

> **Preuve.** Choisissons une base e de E. Soit $\Omega(e)$ la matrice associée à ω dans la base e.
> Montrons d'abord «seulement si». Cela signifie que $\Omega(e)$ est de rang n, donc inversible. Soit x un vecteur tel que $\forall y \in E, \omega(x,y) = 0$. Alors, en notant X, Y les matrices-colonne des coordonnées de x, y dans la base e, on trouve
>
> $$\forall Y, \qquad {}^tX\Omega(e)Y = 0$$
>
> En particulier, on peut choisir Y tel que $Y = (\Omega(e))^{-1}X$. On obtient alors ${}^tX \cdot X = 0$, puis $X = 0$ et enfin $x = 0$.
>
> Réciproquement, supposons que seul le vecteur $x = 0$ possède la propriété $\forall y \in E, \omega(x,y) = 0$. Munissons E d'un produit scalaire, ce qui en fait un espace euclidien. Il existe alors une base orthonormée e de E telle que la matrice $\Omega(e)$ soit diagonale. Si X, Y sont les matrices-colonne des coordonnées des vecteurs x et y dans cette base e, on a
>
> $$\langle x, y \rangle = {}^tX\Omega(e)Y$$
>
> Pour montrer que $\Omega(e)$ est de rang n, il suffit de montrer que ses coefficients diagonaux $(d_i)_{i \in [\![1,n]\!]}$ sont tous non nuls. Supposons par l'absurde qu'il existe $k \in [\![1,n]\!]$ avec $d_k = 0$. Prenons X égal au k-ème vecteur de la base canonique de $\mathcal{M}_{n,1}(\mathbb{R})$. Alors quel que soit $Y \in \mathcal{M}_{n,1}(\mathbb{R})$, on peut dire
>
> $$ {}^tX\Omega(e)Y = 0 \cdot Y = 0$$
>
> Or d'après l'hypothèse il faudrait avoir $X = 0$, ce qui est contradictoire. □

Chapitre 9. Norme d'un endomorphisme

Dans ce chapitre, nous allons définir la norme d'un endomorphisme d'un espace euclidien E. Nous allons également définir la norme d'une matrice carrée, et explorer les liens existant entre ces deux normes.

Commençons par définir la norme d'un endomorphisme.

Définition. *Soit E un espace euclidien non réduit à $\{0\}$, et u un endomorphisme de E. On définit la norme de u, notée $||u||$, comme suit :*

$$||u|| = \sup_{x \in E - \{0\}} \frac{||u(x)||}{||x||}$$

Il faut d'abord justifier l'existence de $||u||$. Comme $E - \{0\}$ est non vide, l'ensemble

$$\{\frac{||u(x)||}{||x||} | x \in E - \{0\}\}$$

est également non vide. C'est donc un sous-ensemble non vide de $\mathbb{R}$ (et même de $\mathbb{R}_+$). Pour montrer l'existence de sa borne supérieure, il suffit de se convaincre que c'est un ensemble **majoré**, car tout sous-ensemble non vide et majoré de $\mathbb{R}$ possède une borne supérieure.

Or nous savons par le chapitre précédent que les endomorphismes des espaces euclidiens sont toujours continus, ce qui donne l'existence d'une constante C avec

$$\forall x, ||u(x)|| \leqslant C||x||$$

Par conséquent, l'ensemble

$$\{\frac{||u(x)||}{||x||} | x \in E - \{0\}\}$$

est bien majoré (par C). Bref, la norme de u existe.

Nous allons maintenant donner les principales propriétés de la norme d'un endomorphisme :

Théorème. *Soit E un espace euclidien.*
1. $\forall u \in \mathcal{L}(E), ||u|| = 0 \iff u = 0$
2. $\forall u \in \mathcal{L}(E), \forall r \in \mathbb{R}, ||ru|| = |r| \cdot ||u||$
3. $\forall u, v \in \mathcal{L}(E), ||u + v|| \leqslant ||u|| + ||v||$
4. $\forall u \in \mathcal{L}(E), \forall x \in E, ||u(x)|| \leqslant ||u|| \cdot ||x||$.

Preuve. De la partie 1., seule l'implication $\Rightarrow$ présente une difficulté. On suppose que $||u|| = 0$. Alors $||u(x)|| \leqslant 0$ pour tout x non nul. Donc $u(x) = 0$ pour tout x non nul, et on peut même dire que $u(x) = 0$ pour tout $x \in E$. Donc u est l'endomorphisme nul.

Partie 2 :

$$\begin{aligned} ||ru|| &= \sup_{x \neq 0} \frac{||(ru)(x)||}{||x||} \\ &= \sup_{x \neq 0} \frac{||r \cdot u(x)||}{||x||} \\ &= \sup_{x \neq 0} \frac{|r| \cdot ||u(x)||}{||x||} \\ &= |r| \sup_{x \neq 0} \frac{||u(x)||}{||x||} \\ &= |r| \cdot ||u|| \end{aligned}$$

Partie 3 : Pour tout vecteur x non nul,

$$\frac{||(u+v)(x)||}{||x||} = \frac{||u(x) + v(x)||}{||x||} \leqslant \frac{||u(x)||}{||x||} + \frac{||v(x)||}{||x||} \leqslant ||u|| + ||v||$$

Par conséquent

$$\sup_{x \neq 0} \frac{||(u+v)(x)||}{||x||} \leqslant ||u|| + ||v||$$

ce qui s'écrit encore

$$||u + v|| \leqslant ||u|| + ||v||$$

Partie 4 : C'est une conséquence triviale de la définition de la norme de u. □

De la même façon, on définit aussi la norme d'une matrice carrée A de taille $n \times n$ par la règle

$$||A|| = \sup_{X \in R^n - \{0\}} \frac{||AX||}{||X||}$$

où X parcourt les matrices-colonne non nulles, et où la norme est la norme euclidienne canonique :

$$\left|\left|\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}\right|\right| = \sqrt{X_1^2 + X_2^2 + \cdots + X_n^2}$$

Les propriétés de la norme d'une matrice sont évidemment semblables à celles que nous venons de démontrer pour la norme d'un endomorphisme :

1. $\forall A \in \mathcal{M}_n(\mathbb{R}), ||A|| = 0 \iff u = 0$
2. $\forall A \in \mathcal{M}_n(\mathbb{R}), \forall r \in \mathbb{R}, ||rA|| = |r| \cdot ||A||$
3. $\forall A, B \in \mathcal{M}_n(\mathbb{R}), ||A + B|| \leqslant ||A|| + ||B||$
4. $\forall A \in \mathcal{M}_n(\mathbb{R}), \forall X \in R^n, ||AX|| \leqslant ||A|| \cdot ||X||$.

Quel est le lien entre la norme d'un endomorphisme, et la norme d'une matrice ? Ces normes sont égales, si l'endomorphisme et la matrice sont associés par rapport à une base orthonormée.

Théorème.
Soit E un espace euclidien, u un endomorphisme de E, et e une base **orthonormée** *de E. Alors*

$$||u|| = ||\mathcal{M}_e(u)||$$

Preuve. Notons $A = \mathcal{M}_e(u)$. Soit x un vecteur quelconque non nul de E. On appelle $(x_1, \ldots, x_n)$ ses coordonnées par rapport à la base orthonormée e. Comme c'est une base orthonormée, on a

$$||x|| = \sqrt{x_1^2 + \cdots + x_n^2}$$

d'autre part, les coordonnées de $u(x)$ dans la base e sont données par la matrice $A \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$.

Par conséquent

$$||u|| = \sup_{X \neq 0} \frac{||AX||}{||X||}$$

où X parcourt les matrices-colonne non nulles. Il vient alors

$$||u|| = ||A||$$

□

La norme d'un endomorphisme se comporte bien par rapport à la composition. En l'occurrence, la norme est **sous-multiplicative :**

Théorème. *Soient u, v deux endomorphismes d'un espace euclidien E. Alors*

$$||u \circ v|| \leqslant ||u|| \cdot ||v||$$

Preuve. Soit x un vecteur quelconque de E. Alors

$$||(u \circ v)(x)|| = ||u(v(x))|| \leqslant ||u|| \cdot ||v(x)|| \leqslant ||u|| \cdot ||v|| \cdot ||x||$$

Donc, pour tout x non nul, on a la majoration

$$\frac{||(u \circ v)(x)||}{||x||} \leqslant ||u|| \cdot ||v||$$

D'où on tire immédiatement

$$||u \circ v|| \leqslant ||u|| \cdot ||v||$$

□

Bien entendu, cela fonctionne aussi pour les matrices ; si A, B sont deux matrices carrées, alors

$$||A \cdot B|| \leqslant ||A|| \cdot ||B||$$

Théorème. *Soit u un endomorphisme d'un espace euclidien E, et u^* son endomorphisme adjoint. Alors*

$$||u|| = ||u^*||$$

et aussi

$$||u||^2 = ||u^* \circ u||$$

Preuve. Montrons d'abord que $||u|| \leqslant ||u^*||$.

Prenons un vecteur quelconque x de E. Alors

$$\begin{aligned} ||u(x)||^2 &= \langle u(x), u(x)\rangle \\ &= \langle x, u^*(u(x))\rangle \\ &\leqslant ||x|| \cdot ||(u^* \circ u)(x)|| \end{aligned}$$

par l'inégalité de Cauchy-Schwarz. On peut maintenant dire que

$$||u(x)||^2 \leqslant ||x|| \cdot ||u^* \circ u|| \cdot ||x||$$

Donc, pour tout x non nul,

$$\left(\frac{||u(x)||}{||x||}\right)^2 \leqslant ||u^* \circ u||$$

ce qui implique bien évidemment

$$||u||^2 \leqslant ||u^* \circ u||$$

Or nous avons aussi

$$||u^* \circ u|| \leqslant ||u^*|| \cdot ||u||$$

D'où on tire

$$||u||^2 \leqslant ||u|| \cdot ||u^*||$$

Si $u = 0$, on a clairement $u^* = 0$ et l'inégalité est vraie. Sinon, u est de norme non nulle, et on peut simplifier l'inégalité ci-dessus en :

$$||u|| \leqslant ||u^*||$$

C'est la fin de la première partie de la preuve. On appliquant la même inégalité à u^*, on trouve

$$||u^*|| \leqslant ||u^{**}||$$

Mais nous savons aussi que $u^{**} = u$. Donc $||u|| = ||u^*||$.

Il ne nous reste plus qu'à prouver l'égalité $||u||^2 = ||u^* \circ u||$. Or nous avons établi ci-dessus que

$$||u||^2 \leqslant ||u^* \circ u||$$

et on sait également que

$$||u^* \circ u|| \leqslant ||u|| \cdot ||u^*|| = ||u||^2$$

Les deux inégalités, prises ensemble, débouchent évidemment sur

$$||u||^2 = ||u^* \circ u||$$

□

Il est découle en particulier que pour toute matrice carrée A à coefficients réels, on a les égalités suivantes :

$$||{}^tA|| = ||A|| \qquad ||A||^2 = ||{}^tAA||$$

Lorsque u est un endomorphisme auto-adjoint (ou symétrique), la norme de u s'exprime très facilement en termes du spectre de u :

Théorème. *Soit u un endomorphisme auto-adjoint d'un espace euclidien E non réduit à $\{0\}$. Alors la norme de u est égale à la plus grande valeur absolue des valeurs propres de u :*

$$||u|| = \max_{\lambda \in Spec(u)} |\lambda|$$

Preuve. Comme $\dim E \geqslant 1$, et que u est auto-adjoint, le spectre de u n'est pas vide, donc la plus grande valeur propre (en valeur absolue) existe. On appelle λ la valeur propre pour laquelle $|\lambda|$ est maximal.

Montrons d'abord que $||u|| \geqslant |\lambda|$. Pour cela, on choisit un vecteur propre non nul x associé à la valeur propre λ. Alors, évidemment

$$||u|| \geqslant \frac{||u(x)||}{||x||} = \frac{||\lambda x||}{||x||} = |\lambda|$$

Prouvons maintenant qu'on a aussi l'inégalité dans l'autre sens : $||u|| \leqslant |\lambda|$. Comme u est auto-adjoint, on sait par le théorème spectral qu'il existe une base orthonormée $(e_1, \ldots, e_n)$ de E, contenant seulement des vecteurs propres de u. On appelle $(\lambda_1, \ldots, \lambda_n)$ les valeurs propres associées. Soit x un vecteur quelconque de E. Appelons $(x_1, \ldots, x_n)$ ses coordonnées par rapport à la base orthonormée e. On trouve alors

$$\begin{aligned} ||u(x)|| &= \sqrt{(\lambda_1 x_1)^2 + \cdots + (\lambda_n x_n)^2} \\ &= \sqrt{\lambda_1^2 x_1^2 + \cdots + \lambda_n^2 x_n^2} \\ &\leqslant \sqrt{|\lambda|^2 x_1^2 + \cdots + |\lambda|^2 x_n^2} \\ &= |\lambda| \sqrt{x_1^2 + \cdots + x_n^2} \\ &= |\lambda| \cdot ||x|| \end{aligned}$$

On en déduit bien $||u|| \leqslant |\lambda|$.

Nous avons établi que $||u|| = |\lambda|$. □

Ce théorème se traduit aussi en langage matriciel : Si A est une matrice symétrique, alors sa norme $||A||$ est égale au maximum de $|\lambda|$, où λ parcourt le spectre de A.

Les deux derniers théorèmes, pris ensemble, permettent de calculer la norme d'un endomorphisme quelconque sur un espace euclidien. Soit u cet endomorphisme. On forme $v = u^* \circ u$, qui est auto-adjoint. On peut alors calculer facilement $||v||$ en étudiant la plus grande valeur propre de v en valeur absolue. Enfin, on peut en déduire la norme de u grâce à la formule

$$||u|| = ||v||^{1/2}$$

Chapitre 10. Espaces pré-hilbertiens complexes

Les espaces pré-hilbertiens complexes ressemblent aux espaces pré-hilbertiens. La différence vient de ce que le corps de base n'est plus $\mathbb{R}$, mais le corps des nombres complexes $\mathbb{C}$. Nous allons expliquer comment construire sur ces espaces quelque chose qui ressemble beaucoup à un produit scalaire, et nous allons introduire à peu près le même vocabulaire que dans les chapitres précédents.

D'abord, une particularité pour les espaces vectoriels sur $\mathbb{C}$, la définition d'une forme semi-linéaire sur un tel espace.

Définition. *Soit E un espace vectoriel sur le corps $\mathbb{C}$.* **Une forme semi-linéaire sur** *E est une application*

$$f : E \to \mathbb{C} : x \mapsto f(x)$$

qui vérifie

$$\forall x, y \in E, f(x+y) = f(x) + f(y)$$
$$\forall x \in E, \forall \lambda \in \mathbb{C}, f(\lambda x) = \overline{\lambda} f(x)$$

Les formes semi-linéaires sont presque des formes linéaires. Il y a juste une petite différence, la présence du conjugué $\overline{\lambda}$ dans la seconde condition.

Par exemple, pour des constantes complexes $(a_1, \ldots, a_n)$ l'application

$$f : \mathbb{C}^n \to \mathbb{C} : (z_1, \ldots, z_n) \mapsto \overline{a_1 z_1 + \cdots + a_n z_n}$$

est une forme semi-linéaire sur $\mathbb{C}^n$. Réciproquement, toute forme semi-linéaire sur $\mathbb{C}$ est de cette forme.

Définition. *Soit E un espace vectoriel sur $\mathbb{C}$.* **Un produit scalaire sur** *E est une application de $E \times E$ vers $\mathbb{C}$, envoyant le couple de vecteurs (x, y) sur un nombre complexe noté $\langle x, y\rangle$, et obéissant aux conditions ci-dessous :*
1. Pour tout $x \in E$, l'application $y \mapsto \langle x, y\rangle$ est une forme linéaire sur $\mathbb{C}$.
2. Pour tout $y \in E$, l'application $x \mapsto \langle x, y\rangle$ est une forme semi-linéaire sur $\mathbb{C}$.
3. Pour tout $x, y \in E$, $\overline{\langle x, y\rangle} = \langle y, x\rangle$.
4. Pour tout $x \in E$, $\langle x, x\rangle$ est un réel positif.
5. Si $\langle x, x\rangle = 0$, alors $x = 0$.

Les conditions 1. et 2. se traduisent comme suit :

$$\langle x, y + y'\rangle = \langle x, y\rangle + \langle x, y'\rangle$$
$$\langle x, \lambda y\rangle = \lambda \langle x, y\rangle$$
$$\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle$$
$$\langle \lambda x, y\rangle = \overline{\lambda} \langle x, y\rangle$$

Définition.
Un espace pré-hilbertien complexe *est un espace vectoriel sur $\mathbb{C}$ muni d'un produit scalaire.*

Faisons tout de suite connaissance avec un exemple très important, l'espace pré-hilbertien complexe $\mathbb{C}^n$. On définit une application de $\mathbb{C}^n \times \mathbb{C}^n$ vers $\mathbb{C}$ comme suit :

Pour $(w_1, \ldots, w_n), (z_1, \ldots, z_n)$ dans $\mathbb{C}^n$, on définit

$$\langle (w_1, \ldots, w_n), (z_1, \ldots, z_n)\rangle := \sum_{i=1}^{n} \overline{w_i} z_i$$

Théorème. *L'application ci-dessus est un produit scalaire sur $\mathbb{C}^n$.*

Preuve. 1. Si on fixe $(w_1, \ldots, w_n)$, l'application

$$(z_1, \ldots, z_n) \mapsto \sum_{i=1}^{n} \overline{w_i} z_i$$

est bien une forme linéaire sur $\mathbb{C}^n$.

2. Si on fixe $(z_1, \dots, z_n)$, l'application

$$(w_1, \dots, w_n) \mapsto \sum_{i=1}^{n} \overline{w_i} z_i$$

est bien une forme **semi**-linéaire sur $\mathbb{C}^n$.

3.

$$\begin{aligned} \overline{\langle (w_1, \dots, w_n), (z_1, \dots, z_n) \rangle} &= \overline{\sum_{i=1}^{n} \overline{w_i} z_i} \\ &= \sum_{i=1}^{n} w_i \overline{z_i} \\ &= \sum_{i=1}^{n} \overline{z_i} w_i \\ &= \langle (z_1, \dots, z_n), (w_1, \dots, w_n) \rangle \end{aligned}$$

4. On a

$$\langle (w_1, \dots, w_n), (w_1, \dots, w_n) \rangle = \sum_{i=1}^{n} \overline{w_i} w_i = \sum_{i=1}^{n} |w_i|^2$$

Clairement $\sum_{i=1}^{n} |w_i|^2$ est un réel positif.

5. Si $\langle (w_1, \dots, w_n), (w_1, \dots, w_n) \rangle = 0$, cela signifie que

$$\sum_{i=1}^{n} |w_i|^2 = 0$$

et donc $\forall i, |w_i| = 0$. Cela entraîne, $\forall i, w_i = 0$. Donc $(w_1, \dots, w_n) = (0, \dots, 0)$. □

Définition. *Le produit scalaire ainsi défini sur $\mathbb{C}^n$ est appelé* **produit scalaire canonique**.

***

En analyse, on utilise souvent un autre produit scalaire, défini sur l'espace vectoriel des fonctions continues d'un intervalle $[a, b]$ (avec $-\infty < a < b < +\infty$) vers $\mathbb{C}$. Si f, g sont deux fonctions continues $[a, b] \to \mathbb{C}$, on pose

$$\langle f, g \rangle := \int_a^b \overline{f(t)} g(t) dt$$

C'est un bon exercice de prouver qu'il s'agit bien d'un produit scalaire (et en particulier, que la condition 5 est vérifiée).

***

Théorème (Inégalité de Cauchy-Schwarz). *Soit E un espace pré-hilbertien complexe de produit scalaire $\langle , \rangle$. Alors, pour x, y deux vecteurs quelconques, on a*

$$|\langle x, y \rangle|^2 \leqslant \langle x, x \rangle \langle y, y \rangle$$

En outre, s'il y a égalité ci-dessus, alors les vecteurs x, y sont colinéaires.

Preuve. Comme beaucoup de preuves qui suivront, celle-ci utilise les mêmes idées que son équivalent pour le corps des réels : Pour tout t réel, on a

$$\langle x + ty, x + ty \rangle \geqslant 0$$

(La notation $\geqslant 0$ signifie que le nombre complexe $\langle x + ty, x + ty \rangle$ est un nombre réel positif. Cela ne veut pas dire qu'il existe une relation d'ordre naturelle sur $\mathbb{C}$).

D'après les relations vérifiées par un produit scalaire, cela devient

$$t^2\langle y,y\rangle + t\langle x,y\rangle + \bar{t}\langle y,x\rangle + \langle x,x\rangle \geqslant 0$$

ou encore

$$t^2\langle y,y\rangle + t(\langle x,y\rangle + \overline{\langle x,y\rangle}) + \langle x,x\rangle \geqslant 0$$

donc

$$t^2\langle y,y\rangle + 2t\mathrm{Re}\langle x,y\rangle + \langle x,x\rangle \geqslant 0$$

On peut supposer que y est non nul, car dans le cas $y=0$, l'inégalité est clairement vérifiée. Dès lors, le membre de gauche est une fonction polynomiale de degré 2 en t. Il faut donc que le discriminant soit négatif, ce qui mène à

$$(\mathrm{Re}\langle x,y\rangle)^2 \leqslant \langle x,x\rangle\langle y,y\rangle$$

Le cas d'égalité se présente seulement quand $x = ry$ avec r réel.
Pour tout nombre complexe z, il existe un réel θ tel que $z = |z|e^{i\theta}$. Donc il existe un reél θ tel que

$$\langle x,y\rangle = |\langle x,y\rangle|e^{i\theta}$$

Par conséquent

$$\langle e^{i\theta}x,y\rangle = \overline{e^{i\theta}}\langle x,y\rangle = e^{-i\theta}|\langle x,y\rangle|e^{i\theta} = |\langle x,y\rangle|$$

et en prenant la partie réelle puis en mettant au carré, on trouve donc

$$(\mathrm{Re}\langle e^{i\theta}x,y\rangle)^2 = |\langle x,y\rangle|^2$$

Or d'après ce qui précède,

$$(\mathrm{Re}\langle e^{i\theta}x,y\rangle)^2 \leqslant \langle e^{i\theta}x,e^{i\theta}x\rangle\langle y,y\rangle = e^{-i\theta}e^{i\theta}\langle x,x\rangle\langle y,y\rangle = \langle x,x\rangle\langle y,y\rangle$$

Cela prouve l'inégalité de Cauchy-Schwarz.

Il reste à étudier le cas d'égalité. Dans ce cas on a

$$(\mathrm{Re}\langle e^{i\theta}x,y\rangle)^2 \leqslant \langle e^{i\theta}x,e^{i\theta}x\rangle\langle y,y\rangle.$$

D'après la première partie, nous savons qu'il existe alors un réel r avec

$$e^{i\theta}x = ry$$

ou encore

$$x = re^{-i\theta}y$$

ce qui montre en effet que x, y sont des vecteurs colinéaires de E. □

Définition. *Soit E un espace pré-hilbertien complexe, de produit scalaire $\langle\ ,\ \rangle$.*
Pour tout vecteur x de E, on appelle **norme de** *x le nombre réel positif $||x|| := \sqrt{\langle x,x\rangle}$.*
Pour deux vecteurs x, y de E, on appelle **distance entre** *x, y le nombre $d(x,y) := ||y-x||$.*

Comme dans le cas des espaces pré-hilbertiens (réels), la norme et la distance obéissent aux classiques relations ci-dessous :
N1. $||x|| = 0 \iff x = 0$
N2. $\forall\lambda\in\mathbb{C}, ||\lambda x|| = |\lambda|\cdot||x||$
N3. $||x+x'|| \leqslant ||x|| + ||x'||$ (inégalité triangulaire)

D1. $d(x,y) = 0 \iff x = y$
D2. $d(x,y) = d(y,x)$
D3. $d(x,y) \leqslant d(x,z) + d(z,y)$ (inégalité triangulaire)

On a aussi les identités du parallélogramme et de polarisation (un peu différente !) pour les produits scalaires dans les espaces pré-hilbertiens complexes.

Théorème. *Pour deux vecteurs quelconques x, y de E, on a*

$$||x+y||^2 + ||x-y||^2 = 2(||x||^2 + ||y||^2)$$

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

Preuve. La preuve de l'identité du parallélogramme ne change pas.

Prouvons maintenant l'identité de polarisation. On voit d'abord que

$$||x+y||^2 - ||x-y||^2 = 4\mathrm{Re}\langle x, y\rangle$$

Remplaçons y par $-iy$. On trouve

$$||x-iy||^2 - ||x+iy||^2 = 4\mathrm{Re}\langle x, -iy\rangle = 4\mathrm{Re}(-i\langle x, y\rangle) = 4\mathrm{Im}\langle x, y\rangle$$

En utilisant les deux équations, on trouve

$$4\langle x, y\rangle = 4\mathrm{Re}\langle x, y\rangle + i \cdot 4\mathrm{Im}\langle x, y\rangle = ||x+y||^2 - ||x-y||^2 + i(||x-iy||^2 - ||x+iy||^2)$$

C'est le résultat annoncé dans l'énoncé.

□

Les prochaines définitions sont toutes familières :

Définition. *Soit E un espace pré-hilbertien complexe de produit scalaire $\langle\ ,\ \rangle$.*
Un vecteur x est appelé **unitaire** *si la norme de x vaut 1.*
Deux vecteurs x, y sont appelés **orthogonaux** *si $\langle x, y\rangle = 0$.*
Deux sous-espaces vectoriels F_1, F_2 de E sont dits orthogonaux si pour tout couple $(x_1, x_2) \in F_1 \times F_2$, les vecteurs x_1, x_2 sont orthogonaux.
L'orthogonal d'un sous-espace vectoriel *F de E est l'ensemble*

$$F^\perp := \{x \in E | \forall y \in F, \langle x, y\rangle = 0\}.$$

Une famille de vecteurs $(x_i)_{i\in I}$ est appelée **orthogonale** *si pour i, j distincts, les vecteurs x_i, x_j sont orthogonaux.*
Une famille de vecteurs est appelée **orthonormée** *si elle est orthogonale et que tous les vecteurs sont unitaires.*

Commentaires :
1. Deux vecteurs x, y sont orthogonaux si et seulement si y, x sont orthogonaux.

2. L'orthogonal d'un sous-espace vectoriel est toujours un sous-espace vectoriel, et leur intersection est toujours réduite à $\{0\}$. On peut aussi écrire

$$F^\perp := \{y \in E | \forall x \in F, \langle x, y\rangle = 0\}.$$

3. Une famille orthonormée peut être infinie (nous en verrons un exemple plus loin).

Théorème de Pythagore. *Soit $(x_i)_{i\in[[1,n]]}$ une famille orthogonale finie de vecteurs d'un espace pré-hilbertien complexe. Alors*

$$||x_1 + x_2 + \cdots + x_n||^2 = ||x_1||^2 + ||x_2||^2 + \cdots + ||x_n||^2$$

Preuve. Comme dans le cas des espaces pré-hilbertiens réels. □

Pour terminer ce chapitre, nous allons donner un exemple d'une famille orthonormée très importante, celle qui fonde la théorie des séries de Fourier.

On considère l'espace vectoriel $\mathcal{P}_{2\pi}(\mathbb{R},\mathbb{C})$ de toutes les fonction continues et 2π-périodiques de $\mathbb{R}$ vers $\mathbb{C}$. Cet espace vectoriel (sur le corps $\mathbb{C}$) devient un espace pré-hilbertien complexe lorsqu'on le munit du produit scalaire

$$\langle f,g\rangle := \frac{1}{2\pi}\int_0^{2\pi} \overline{f(t)}g(t)dt$$

Pour chaque entier $n\in\mathbb{Z}$, appelons e_n la fonction suivante :

$$e_n : \mathbb{R}\to\mathbb{C} : t\mapsto e^{nit}$$

La fonction e_n est évidemment continue, et elle est évidemment 2π-périodique.

Chaque fonction e_n est un élément de $\mathcal{P}_{2\pi}(\mathbb{R},\mathbb{C})$. La fonction e_n est donc un vecteur de cet espace vectoriel.

Considérons maintenant la famille de vecteurs $(e_n)_{n\in\mathbb{Z}}$. Nous allons d'abord montrer que c'est une famille orthogonale. En effet, pour m,n deux entiers distincts,

$$\begin{aligned}
\langle e_m,e_n\rangle &= \frac{1}{2\pi}\int_0^{2\pi}\overline{e^{mit}}e^{nit}dt\\
&= \frac{1}{2\pi}\int_0^{2\pi}e^{-mit}e^{nit}dt\\
&= \frac{1}{2\pi}\int_0^{2\pi}e^{(n-m)it}dt\\
&= \frac{1}{2\pi}\left[\frac{e^{(n-m)it}}{(n-m)i}\right]_0^{2\pi}\\
&= \frac{1}{2\pi}\left[\frac{1-1}{(n-m)i}\right]\\
&= 0
\end{aligned}$$

D'autre part, chaque vecteur e_n est de norme 1, puisque

$$\begin{aligned}
\langle e_n,e_n\rangle &= \frac{1}{2\pi}\int_0^{2\pi}\overline{e^{nit}}e^{nit}dt\\
&= \frac{1}{2\pi}\int_0^{2\pi}e^{-nit}e^{nit}dt\\
&= \frac{1}{2\pi}\int_0^{2\pi}dt\\
&= 1
\end{aligned}$$

Cela montre que la famille $(e_n)_{n\in\mathbb{Z}}$ est une famille orthonormée de $\mathcal{P}_{2\pi}(\mathbb{R},\mathbb{C})$. C'est «presqu'» une base de cet espace de dimension infinie. Mais elle n'engendre pas tous l'espace des fonctions continues. C'est pour cette raison qu'on admet les séries de Fourier, qui sont des sommes infinies, pour représenter davantage de fonctions de cet espace vectoriel.

On peut utiliser cela pour calculer certaines intégrales, comme par exemple

$$I=\int_0^{2\pi}(\cos 2t+\sin 3t)^2dt$$

En effet, soit f la fonction $f : t\mapsto \cos 2t+\sin 3t$. L'intégrale cherchée est alors

$$I=\langle f,f\rangle = ||f||^2$$

Décomposons maintenant f dans la famille $(e_n)_{n\in\mathbb{Z}}$. Il est facile de voir que

$$f(t)=\frac{1}{2}e^{2it}+\frac{1}{2}e^{-2it}+\frac{1}{2i}e^{3it}-\frac{1}{2i}e^{-3it}$$

ou encore

$$f=\frac{1}{2}e_2+\frac{1}{2}e_{-2}+\frac{1}{2i}e_3-\frac{1}{2i}e_{-3}$$

La famille $(\frac{1}{2}e_2,\frac{1}{2}e_{-2},\frac{1}{2i}e_3,\frac{1}{2i}e_{-3})$ est une famille orthogonale, on peut alors appliquer le théorème de Pythagore pour trouver

$$I=||f||^2=||\frac{1}{2}e_2||^2+||\frac{1}{2}e_{-2}||^2+||\frac{1}{2i}e_3||^2+||\frac{1}{2i}e_{-3}||^2=\frac{1}{4}+\frac{1}{4}+\frac{1}{4}+\frac{1}{4}=1$$

Chapitre 11. Espaces hermitiens

Définition. Un espace hermitien *est un espace pré-hilbertien complexe de dimension finie.*

Dans ce chapitre, nous allons rapidement donner les théorèmes analogues à la situation des espaces euclidiens. Les preuves sont essentiellement les mêmes. C'est la raison pour laquelle nous n'allons pas souvent écrire les preuves.

Théorème. *Dans tout espace hermitien, il existe une base orthonormée.*
Plus généralement, si $(e_1, \dots, e_k)$ est une famille orthonormée d'un espace hermitien, alors on peut toujours la compléter en une base orthonormée.

Théorème. *Soit E un espace hermitien, et E^* son espace dual. L'application*

$$\Psi : \begin{array}{ccc} E & \to & E^* \\ a & \mapsto & (x \mapsto \langle a, x\rangle) \end{array}$$

est une bijection semi-linéaire.
Par semi-linéaire, nous entendons

$$\Psi(a+a') = \Psi(a) + \Psi(a') \qquad \Psi(\lambda a) = \overline{\lambda}\Psi(a)$$

Preuve. La semi-linéarité vient de ce que le produit scalaire sur un espace pré-hilbertien complexe est semi-linéaire à gauche, et linéaire à droite. Grâce à la linéarité à droite, on voit notamment que $\Psi(a)$ est bien une forme linéaire sur E.

Ψ est une injection, car si $\Psi(a) = \Psi(b)$, alors pour tout $x \in E$, on a

$$\langle a, x\rangle = \langle b, x\rangle$$

ou encore

$$\langle a-b, x\rangle = 0$$

Donc $a - b = 0$ ce qui veut dire $a = b$.

Soit $(e_1, \dots, e_n)$ une base de E. Montrons que $(\Psi(e_1), \dots, \Psi(e_n))$ est une famille libre de E^*. En effet, si on suppose

$$\sum_{k=1}^{n} \alpha_k \Psi(e_k) = 0,$$

alors par semi-linéarité de Ψ on obtient

$$\Psi(\sum_{k=1}^{n} \overline{\alpha_k} e_k) = 0$$

et par injectivité

$$\sum_{k=1}^{n} \overline{\alpha_k} e_k = 0$$

Mais $(e_1, \dots, e_n)$ est une base. Donc, pour tout k, on a $\overline{\alpha_k} = 0$, ce qui donne bien $\alpha_k = 0$. La famille $(\Psi(e_1), \dots, \Psi(e_n))$ est une famille libre.

Cette famille ayant n vecteurs, et puisque la dimension de E^* est égale à n, on peut affirmer que cette famille est une base de E^*. Maintenant nous sommes en mesure de démontrer que Ψ est surjective :

Soit $y \in E^*$ un vecteur quelconque. Comme $(\Psi(e_1), \dots, \Psi(e_n))$ est une base de E^*, on peut trouver des nombres complexes $\beta_1, \dots, \beta_n$ avec

$$y = \sum_{k=1}^{n} \beta_k \Psi(e_k)$$

ce qui peut ensuite s'écrire

$$y = \Psi(\sum_{k=1}^{n} \overline{\beta_k} e_k)$$

Cette égalité montre que y possède un antécédent par Ψ. Nous avons démontré la surjectivité de Ψ. Bref, Ψ est une application bijective semi-linéaire. □

Corollaire. *Pour toute forme linéaire ω sur un espace hermitien E, il existe un unique $a \in E$ vérifiant*

$$\forall x \in E, \qquad \omega(x) = \langle a, x\rangle$$

Théorème. *Soit E un espace pré-hilbertien complexe, et F un sous-espace vectoriel de dimension finie de E. Alors $F^\perp$ est un sous-espace vectoriel supplémentaire à F. On l'appelle parfois* **le supplémentaire orthogonal de** *F Autrement dit,*

$$F \oplus F^\perp = E.$$

On remarque que le supplémentaire orthogonal est **unique**, alors qu'un supplémentaire (sans la condition «orthogonal») n'est en général pas unique.

Définition. *Soit E un espace pré-hilbertien complexe, et F un sous-espace vectoriel de dimension finie de E. On appelle* **projection orthogonale de E sur F** *la projection sur le sous-espace vectoriel F parallèlement à son supplémentaire orthogonal $F^\perp$.*

Théorème. *Soit E un espace pré-hilbertien complexe, et F un sous-espace vectoriel de dimension finie de E. La distance d'un vecteur $x \in E$ au sous-espace F est égale à la distance de x à son image par la projection orthogonale p_F sur F. Autrement dit,*

$$\min_{y\in F} d(x,y) = d(x, p_F(x)).$$

Théorème. *Soit E un espace pré-hilbertien complexe, et F un sous-espace vectoriel de dimension finie de E. On considère la projection orthogonale p_F sur F. Enfin, soit $(e_1, \dots, e_n)$ une base orthonormée de F. Alors pour tout $x \in E$, on a*

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

Théorème (Inégalité de Bessel). *Soit E un espace pré-hilbertien complexe, et $(e_1, \dots, e_n)$ une famille orthonormée finie de vecteurs de E. Alors pour tout $x \in E$, on a l'inégalité de Bessel :*

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

Il faut observer qu'il y a une légère différence avec l'inégalité de Bessel dans les espaces pré-hilbertiens réels : ici on trouve le module dans le membre de gauche.

***

Pour terminer, nous allons donner une nouvelle preuve du théorème spectral, en tout cas de sa première partie. Nous allons pour cela nous appuyer sur la théorie des espaces hermitiens, et le théorème de Gauss-d'Alembert.

Théorème. *Soit E un espace euclidien non réduit à $\{0\}$, et u un endomorphisme auto-adjoint de E. Alors il existe une valeur propre de u.*

Preuve. E est un espace vectoriel sur $\mathbb{R}$ de dimension finie n. Choisissons une base orthonormée quelconque de E. Dans cette base, u est représenté par une matrice symétrique A carrée $n \times n$ à coefficients réels. Il faut prouver que A admet une valeur propre.

On considère l'espace vectoriel $\mathbb{C}^n$, muni du produit scalaire canonique

$$\langle (w_1, \dots, w_n), (z_1, \dots, z_n)\rangle = \sum_{k=1}^{n} \overline{w_k} z_k$$

L'application

$$u^{\mathbb{C}} : \mathbb{C}^n \to \mathbb{C}^n : W \mapsto A \cdot W$$

est un endomorphisme de $\mathbb{C}^n$ (les éléments de $\mathbb{C}^n$ sont ici vus comme des matrices-colonne). Donc le polynôme caractéristique de $u^{\mathbb{C}}$ est un polynôme de degré $n \geqslant 1$ sur $\mathbb{C}$. Par le théorème de Gauss-d'Alembert, ce polynôme admet au moins une racine complexe λ. En d'autres termes, il existe $W \in \mathbb{C}^n - \{0\}$ avec

$$A \cdot W = \lambda W$$

En travaillant avec le produit scalaire $\langle\ ,\ \rangle$ de $\mathbb{C}^n$, on a alors

$$\begin{aligned}
\lambda\langle W,W\rangle &= \langle W,\lambda W\rangle \\
&= \langle W,A\cdot W\rangle \\
&= {}^t\overline{W}(AW) \\
&= {}^t\overline{W}AW \\
&= {}^t\overline{W}\,{}^t\overline{A}W \qquad (A \text{ est symétrique et réelle}) \\
&= {}^t\overline{AW}\cdot W \\
&= \langle A\cdot W,W\rangle \\
&= \langle \lambda W,W\rangle \\
&= \overline{\lambda}\langle W,W\rangle
\end{aligned}$$

Comme W est non nul, $\langle W,W\rangle$ est également non nul, ce qui permet de dire

$$\lambda = \overline{\lambda}$$

Donc λ est réel. Il existe par conséquent un nombre réel λ et un vecteur non nul $W\in\mathbb{C}^n$ avec

$$AW=\lambda W$$

Pour montrer que λ est un valeur propre réelle de la matrice réelle A, il reste à montrer que l'on peut trouver W dans $\mathbb{R}^n$.

Comme il existe W non nul dans $\mathbb{C}^n$ avec $AW=\lambda W$, la matrice $A-\lambda I_n$ n'est pas inversible. Mais la matrice $A-\lambda I_n$ est une matrice **réelle**, puisque A et λ sont réels. Elle ne peut donc pas être inversible si on la voit comme matrice réelle. Par conséquent, il existe bien $W\in\mathbb{R}^n-\{0\}$ avec

$$AW=\lambda W$$

Nous avons prouvé que λ est une valeur propre réelle de la matrice réelle A. □

Sommaire du cours d'Algèbre 3

Annexe : Polynômes

1 Introduction

Si X est un nombre réel (ou complexe), tout le monde sait que

$$(X^2 + X - 2) + (X + 1) = X^2 + 2X - 1$$

et

$$(X^2 + X - 2) \cdot (X + 1) = X^3 + 2X^2 - X - 2$$

Nous appellerons polynôme toute expression de la forme

$$a_0 + a_1 X + a_2 X^2 + \cdots$$

à condition que la somme ne comporte qu'un nombre **fini** de termes. Nous savons comment additionner de telles expressions. La somme de deux polynômes est évidemment encore un polynôme.

On peut aussi multiplier deux polynômes. Si on écrit

$$P = \sum_{m=0}^{+\infty} a_m X^m \qquad\qquad Q = \sum_{n=0}^{+\infty} b_n X^n$$

(où il est sous-entendu que pour m et n assez grands, on a $a_m = 0$ et $b_n = 0$), alors

$$\begin{aligned} P \cdot Q &= (\sum_{m=0}^{+\infty} a_m X^m) \cdot (\sum_{n=0}^{+\infty} b_n X^n) \\ &= \sum_{m,n \geqslant 0} a_m X^m b_n X^n \\ &= \sum_{m,n \geqslant 0} a_m b_n X^{m+n} \\ &= \sum_{d=0}^{+\infty} (\sum_{m=0}^{d} a_m b_{d-m} X^d) \\ &= \sum_{d=0}^{+\infty} (\sum_{m=0}^{d} a_m b_{d-m}) X^d \end{aligned}$$

Bien sûr, si d est trop grand, la somme $\displaystyle\sum_{m=0}^{d} a_m b_{d-m}$ aura uniquement des termes nuls, car au moins un des nombres a_m ou b_{d-m} est égal à zéro. Donc $P \cdot Q$ est encore un polynôme.

On remarque aussi qu'il suffit de connaître la suite des nombres $a_0, a_1, a_2, \ldots$ pour décrire complètement le polynôme associé.

Ces observations de base nous permettent de donner notre définition des polynômes.

2 Définition des polynômes. Degré d'un polynôme.

Définition 1. *Soit K un corps commutatif.* **Un polynôme sur** K *est une suite $(a_n)_{n\geqslant 0}$ d'éléments de K telle que $a_n = 0$ pour tout n assez grand, autrement dit :*

$$\exists N \in \mathbb{N}, \forall n \geqslant N, a_n = 0$$

Les scalaires a_n sont appelés **les coefficients du polynôme.**
L'ensemble de tous les polynômes sur K est noté $K[X]$. On définit deux opérations $+$ et $\cdot$ sur $K[X]$ par les lois suivantes :

$$(a_n)_{n\geqslant 0} + (b_n)_{n\geqslant 0} = (a_n + b_n)_{n\geqslant 0}$$

$$(a_n)_{n\geqslant 0} \cdot (b_n)_{n\geqslant 0} = (\sum_{m=0}^{n} a_m b_{n-m})_{n\geqslant 0}$$

Pour noter un polynôme, on écrira presque toujours $\sum_{n=0}^{+\infty} a_n X^n$ au lieu de $(a_n)_{n\geqslant 0}$. Le symbole X peut être interprété comme une variable formelle. Il ne faut pas nécessairement penser que X est un élément de K.

Notre premier théorème décrit une partie de la structure algébrique de $K[X]$.

Théorème 1. *Avec les opérations $+$ et $\cdot$ définies ci-dessus, $(K[X], +, \cdot)$ est un anneau commutatif.*

Preuve. Il suffit de vérifier les différentes propriétés qui définissent les anneaux commutatifs.
Remarquons toutefois que le neutre additif est le polynôme $0 + 0X + 0X^2 + \cdots$, qu'on appelle aussi **le polynôme nul**. Le neutre multiplicatif est le polynôme $1+0X+0X^2+\cdots$.
□

Si $a \in K$ est un scalaire, on note souvent (par abus de notation) a au lieu de $a + 0X + 0X^2 + \cdots$. On appelle ces polynômes **les polynômes constants**. Le neutre de l'addition des polynômes est alors le polynôme constant 0, et le neutre de la multiplication des polynômes est le polynôme constant 1.

On peut aussi définir une multiplication externe des polynômes par un scalaire quelconque.
Si $P = a_0 + a_1X + a_2X^2 + \cdots$, et que $r \in K$ est un scalaire quelconque, on pose

$$r \cdot P = ra_0 + ra_1X + ra_2X^2 + \cdots$$

Bien sûr, le polynôme $r \cdot P$ est égal au produit du polynôme P par le polynôme constant r.

Théorème 2. *Avec l'opération $+$ et la multiplication externe $\cdot$ définies ci-dessus, $(K[X], +, \cdot)$ est un espace vectoriel sur K.*

Preuve. Simple routine. □

L'ensemble $K[X]$ possède à la fois une structure d'anneau et une structure d'espace vectoriel sur K. Un anneau qui est aussi un espace vectoriel est appelé *une algèbre.* (Ne pas confondre "une algèbre" avec "l'algèbre", qui est une branche des mathématiques).

La prochaine définition introduit le degré d'un polynôme. C'est une notion fondamentale.

Définition 2. *Soit $P \in K[X]$ un polynôme non nul sur K. On appelle a_n ses coefficients.*
Le degré de *P est le plus grand entier n tel que $a_n \neq 0$.*
Par convention, le degré du polynôme nul est égal à $-\infty$.
Si d est le degré d'un polynôme non nul P , le coefficient a_d est appelé **le coefficient dominant.**
Le terme $a_d X^d$ est alors appelé **le terme dominant de P.**
Un polynôme est dit **unitaire** *si son coefficient dominant vaut 1.*

Exemples.
Le polynôme $X^2 + X + 1$ est de degré 2 et il est unitaire.
Le polynôme $3X^4 - X^2 + X$ est de degré 4 et il n'est pas unitaire. Son coefficient dominant vaut 3.
Tout polynôme de degré 1 et unitaire est de la forme $X - a$ avec $a \in K$.

3 L'anneau des polynômes sur K est intègre

Parlons tout d'abord du degré d'un somme ou d'un produit de deux polynômes.

Théorème 3. *Soient $P, Q \in K[X]$ deux polynômes sur K.*
1. Le degré de $P + Q$ est inférieur ou égal à $\max(\deg P, \deg Q)$ *:*

$$\deg(P + Q) \leqslant \max(\deg P, \deg Q)$$

2. Si $\deg P \neq \deg Q$*, alors le degré de $P + Q$ est égal à* $\max(\deg P, \deg Q)$ *:*

$$\deg(P + Q) = \max(\deg P, \deg Q)$$

3. Le degré de $P \cdot Q$ est égal à la somme des degrés de P et Q :

$$\deg(PQ) = \deg P + \deg Q$$

Preuve. 1. et 2. sont évidents.
3. On note $u = \deg P, v = \deg Q$. Alors

$$P = a_0 + a_1 X + \cdots + a_u X^u \qquad Q = b_0 + b_1 X + \cdots + b_v X^v$$

avec $a_u \neq 0$ et $b_v \neq 0$. Si on multiplie P par Q, on obtient seulement des termes de la forme $* \cdot X^k$ avec $k \leqslant u + v$. On peut déjà dire que le degré de PQ est inférieur ou égal à $u + v$. Mais le seul terme de la forme $* \cdot X^{u+v}$ est le terme $a_u b_v X^{u+v}$.
Par hypothèse, a_u et b_v sont non nuls, et comme ils appartiennent à un corps, leur produit est également non nul. Il en découle que le degré de PQ vaut bien $u + v$. □

Remarque. Si P ou Q est le polynôme nul, le théorème ci-dessus reste encore vrai, il suffit simplement d'interpréter comme suit : $\max(n, -\infty) = n$, et $n + (-\infty) = -\infty$.

Théorème 4. *L'anneau $(K[X], +, \cdot)$ est intègre.*

Preuve. En effet, l'anneau est commutatif. Il faut encore vérifier que si $PQ = 0$, alors $P = 0$ ou $Q = 0$. Supposons par l'absurde que P et Q sont tous deux non nuls. Alors leurs degrés respectifs sont des entiers naturels u et v. On sait que le degré de PQ vaut $u + v \in \mathbb{N}$. Mais par hypothèse, le degré de PQ vaut $-\infty$. □

4 Division euclidienne des polynômes

De la même manière que nous pouvons faire une division euclidienne d'un entier par un entier (non nul), nous pouvons aussi faire la division euclidienne d'un polynôme par un polynôme (non nul).

Théorème 5. *Soient P_1, P_2 deux polynômes, et P_2 non nul. Alors il existe un unique couple de polynômes (Q, R) tel que*

$$P_1 = P_2Q + R \qquad et \qquad \deg R < \deg P_2$$

Preuve. Commençons par l'unicité. On suppose que $P_1 = P_2Q+R = P_2Q'+R'$. Montrons que $Q = Q'$ et $R = R'$.
Comme $P_2Q + R = P_2Q' + R'$, on peut dire

$$P_2(Q - Q') = R' - R$$

Par hypothèse, les degrés de R' et $-R$ sont strictement inférieurs à $\deg P_2$. Donc le degré de $R' - R$ est aussi strictement inférieur à $\deg P_2$. Mais le degré de $R' - R$ est aussi égal à $\deg P_2 + \deg(Q - Q')$. Il faut donc que $\deg(Q - Q') < 0$, c'est-à-dire $Q - Q' = 0$. On a déjà montré $Q = Q'$. Il en découle directement que $R' = R$.
Prouvons maintenant l'existence. Si $P_1 = 0$ c'est évident, il suffit de prendre $Q = R = 0$. Supposons maintenant P_1 non nul, donc $\deg P_1 \in \mathbb{N}$. Nous allons fixer P_2 et raisonner par récurrence sur le degré de P_1.

Si le degré de P_1 est nul, alors P_1 est une constante non nulle $a \in K$. Si P_2 est de degré nul, alors P_2 est aussi une constante non nulle $b \in K$. Dans ce cas, nous écrivons

$$P_1 = a = b \cdot \frac{a}{b} + 0$$

et on peut prendre $Q = \dfrac{a}{b}, R = 0$.
Si P_2 est de degré au moins 1, on peut écrire

$$P_1 = a = P_2 \cdot 0 + a$$

et on peut prendre $Q = 0, R = a$.

Supposons maintenant que la propriété est vraie pour tous les polynômes P_1 de degré inférieur ou égal à $n-1$, et montrons qu'alors elle est encore vraie pour tous les polynômes de degré n.
On prend P_1 de degré égal à n. Si P_2 est de degré $> n$, alors c'est fini, car il suffit de prendre $Q = 0$ et $R = P_1$.
Supposons dorénavant P_2 de degré $\leqslant n$. Alors on peut trouver $m \in \mathbb{N}$ et $c \in K$ tels que

$$P_1 - cX^m P_2$$

soit de degré $\leqslant n - 1$. En effet, si a_nX^n est le terme dominant de P_1 et b_pX^p le terme dominant de P_2, on peut prendre $m = n - p \geqslant 0$ et $c = \dfrac{a_n}{b_p} \in K$, ce qui supprime le terme de degré n dans $P_1 - cX^m P_2$. Par hypothèse de récurrence on peut alors affirmer l'existence d'un couple (Q, R) avec

$$P_1 - cX^m P_2 = P_2Q + R$$

et $\deg R < \deg P_2$. Dès lors

$$P_1 = P_2(Q + cX^m) + R$$

et on a trouvé un couple (Q, R) comme voulu. □

Exemple.

La division euclidienne de deux polynômes se fait comme la division de deux nombres, apprise à l'école primaire.

$$\begin{array}{rrrrcl}
X^4 & +2X^3 & +0X^2 & -X & \div & X^2+X-2 \;=\; X^2+X+1 \\
X^4 & X^3 & -2X^2 & & & \\
\cline{2-3}
 & X^3 & 2X^2 & -X & & \\
 & X^3 & X^2 & -2X & & \\
\cline{2-3}
 & & X^2 & X & +3 & \\
 & & X^2 & X & -2 & \\
\cline{2-3}
 & & & & 5 &
\end{array}$$

Donc $X^4+2X^3-X+3=(X^2+X-2)(X^2+X+1)+5$. Le quotient est X^2+X+1, le reste est le polynôme constant 5.

5 Fonction polynomiale associée à un polynôme. Racines

Définition 3. *Soit* $P=\sum_{n=0}^{\infty} a_n X^n \in K[X]$ *un polynôme.*
La fonction polynomiale associée à P *est la fonction*

$$f_P : K \to K : x \mapsto \sum_{n=0}^{\infty} a_n x^n$$

Remarque. Deux polynômes $P=\sum_{n=0}^{\infty} a_n X^n$ et $Q=\sum_{n=0}^{\infty} b_n X^n$ sont égaux si et seulement si $\forall n \geqslant 0, a_n = b_n$ (identification des coefficients de P et Q). En revanche, si les fonctions polynomiales associées f_P et f_Q sont égales, on ne peut pas toujours dire que $a_n = b_n$ pour tout $n \in \mathbb{N}$. Ce problème se présente lorsque K est un corps fini.

Prenons $K=\mathbb{Z}/2\mathbb{Z}$, qui est un corps à deux éléments. Les polynômes $P=X^2+X$ et $Q=0$ ne sont pas égaux. Pourtant, f_P et f_Q sont égales, car

$$0^2+0=0 \qquad \text{et} \qquad 1^2+1=1+1=2=0$$

Donc $f_P : \mathbb{Z}/2\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} : x \mapsto x^2+x$ est la fonction $f_Q : \mathbb{Z}/2\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} : x \mapsto 0$. Pourtant P n'est pas le polynôme nul.

Heureusement, nous verrons plus loin que si K est un corps infini (et c'est le cas de $\mathbb{R}, \mathbb{C}, \mathbb{Q}$), alors $f_P = f_Q$ si et seulement si $a_n = b_n$ pour tout n.

Définition 4. *Soit* $P=\sum_{n=0}^{\infty} a_n X^n \in K[X]$ *un polynôme.*
Une racine de P *est un élément* $x \in K$ *tel que*

$$f_P(x)=\sum_{n=0}^{\infty} a_n x^n = 0$$

Bien sûr, si P est le polynôme nul, tout élément de K est une racine de P.

Exemples.
Le polynôme $P = X^2 + 1 \in \mathbb{R}[X]$ n'a aucune racine.
En revanche, le polynôme $P = X^2 + 1 \in \mathbb{C}[X]$ a exactement deux racines, à savoir i et $-i$.
Les polynômes constants non nuls n'ont jamais de racine (quel que soit le corps K).
Le polynôme $P = X^2 + bX + c \in \mathbb{R}[X]$ a deux racines si $b^2 - 4c > 0$, une racine si $b^2 - 4c = 0$, aucune racine si $b^2 - 4c < 0$.
Pour tout $a \in K$, le polynôme $X - a$, de degré 1, possède une seule racine. C'est le scalaire a.

Définition 5. *On dit qu'un polynôme A est divisible par un polynôme B s'il existe un polynôme Q avec $A = BQ$. On dit aussi "B est un diviseur de A" ou "B divise A" ou "A est un multiple de B".*
Si B est non nul, il est équivalent de dire que le reste de la division euclidienne de A par B est nul.

Exemples.
Le seul polynôme divisible par 0 est le polynôme 0.
Le polynôme 0 est divisible par tout polynôme.
Tout polynôme est un multiple du polynôme constant 1.

Théorème 6. *Soit P un polynôme. Le scalaire $r \in K$ est une racine de P si et seulement si P est divisible par $X - r$.*

Preuve. Ecrivons la division euclidienne de P par $X - r$:

$$P = (X - r)Q + R$$

Comme $\deg R < \deg(X - r) = 1$, il faut que R soit un polynôme constant. Or on a évidemment

$$f_P(x) = (x - r)f_Q(x) + R$$

pour tout $x \in K$. En particulier $f_P(r) = 0$ si et seulement si $R = 0$, ce qui est équivalent au théorème énoncé ci-dessus. □

Soit P un polynôme non nul de degré n. Si r est une racine de P, alors $\dfrac{P}{X - r}$ est encore un polynôme, de degré $n - 1$. Parfois il arrive que r soit encore une racine de $\dfrac{P}{X - r}$. Dans ce cas, $\dfrac{P}{(X - r)^2}$ est encore un polynôme. On peut continuer à diviser par $X - r$ tant que r reste une racine. Mais après un nombre de fini de divisions, r ne peut plus être une racine (pourquoi ?).

Définition 6. *Soit P un polynôme non nul, et r une racine de P.*
La multiplicité de *r est le plus grand entier $m \geqslant 1$ tel que P soit divisible par $(X - r)^m$.*

Bien entendu, la multiplicité d'une racine de d'un polynôme non nul P ne peut pas dépasser le degré de P.

Exemples.
La racine 0 de $P = X^4 + X^3 - 2X^2$ est une racine de multiplicité 2.
Les racines de multiplicité 1 sont appelées racines **simples.**
Les racines de multiplicité 2 sont appelées racines **doubles.**
Lorsque r n'est pas une racine de P, on dit parfois que c'est une **racine de multiplicité nulle**.

6 Arithmétique des polynômes

Nous verrons dans cette section de nombreuses propriétés des polynômes qui ressemblent aux propriétés des nombres entiers rencontrées en arithmétique.

Théorème 7. *Soient A, B deux polynômes, avec B non nul. On appelle R le reste de la division euclidienne de A par B. Soit U un polynôme.*
Alors $(U|A$ et $U|B)$ si et seulement si $(U|B$ et $U|R)$.

Preuve. Il suffit d'utiliser $R = A - BQ$ et $A = BQ + R$. □

Nous pouvons donc dire qu'un polynôme U est un diviseur commun de A et B si et seulement si U est un diviseur commun de B et R.

Théorème 8. *Soient A, B deux polynômes, avec au moins un des deux polynômes non nul.*
Alors il existe un unique polynôme unitaire D tel que
i) D est un diviseur commun de A et B,
ii) Si U est un diviseur commun de A et B, alors $U|D$.

Preuve. Montrons d'abord l'unicité de D. Si D_1, D_2 ont cette propriété, alors il faut que $D_1|D_2$ et $D_2|D_1$. Mais alors D_1 et D_2 doivent avoir le même degré, et donc le polynôme D_2/D_1 est un polynôme constant unitaire. Mais le seul polynôme constant unitaire est le polynôme 1. Donc $D_1 = D_2$.

Montrons l'existence de D. Si $B = 0$ (et donc $A \neq 0$) il suffit de prendre D égal à A, divisé par son coefficient dominant. On va maintenant prouver la propriété par récurrence sur le degré de B.
Si le degré de B est nul, il est clair que $D = 1$ convient.
On suppose que si B est de degré $\leqslant n-1$, la propriété est vraie. Choisissons B de degré n et montrons qu'elle reste vraie.
On fait la division euclidienne de A par B :

$$A = BQ + R$$

Le polynôme R est de degré $\leqslant n-1$, donc par hypothèse de récurrence, il existe un polynôme unitaire D ayant les propriétés ci-dessus pour le couple (B, R).

Mais alors D divise B et R, et donc aussi A.
En plus, si U divise A et B, alors U divise B et R, et par la seconde condition, on peut alors dire que U divise D. Ceci montre que le polynôme D convient. L'existence est démontrée.
□

Définition 7. *Soient A, B deux polynômes. Si au moins un des polynômes est non nul, on appelle* **plus grand diviseur commun de A et B,** *ou $PGCD(A, B)$ ou $A \wedge B$, l'unique polynôme unitaire du théorème précédent. Par définition, $PGCD(0, 0) = 0$.*

Bien sûr, le PGCD de A et B est le plus "grand" diviseur commun de A et B, au sens suivant :

Théorème 9. *Soient A, B deux polynômes, avec A ou B non nul. L'ensemble des diviseurs communs unitaires de A et B contient un seul polynôme de degré maximal, et c'est le PGCD de A et B.*

Preuve. Soit n le degré du PGCD. Il ne peut pas y avoir de diviseur commun U de A et B avec un degré $> n$, car on a $U|PGCD(A,B)$.
D'autre part, si U est un diviseur de A et B de degré n, alors U divise le PGCD, et le quotient doit donc être un polynôme unitaire de degré 0. Le quotient est donc 1. Ceci prouve que U est égal au PGCD. □

Bien entendu, comme pour l'arithmétique des entiers, il y a l'algorithme d'Euclide, qui sert à calculer le PGCD de deux polynômes.

Théorème 10. *Soient A, B deux polynômes. Tant que B est non nul, on fait la division euclidienne de A par B, et on remplace le couple (A,B) par le couple (B,R), où R est le reste.*
1. L'algorithme s'arrête après un nombre fini d'étapes.
2. Le dernier reste non nul obtenu est égal au PGCD de A et B, à une constante multiplicative près.

Preuve. 1. Il suffit de remarquer que la suite des degrés des restes est strictement décroissante.
2. La preuve est identique à celle donnée pour $\mathbb{Z}$. □

Exemple. Calculons le PGCD des polynômes $A = X^4 + 2X^2 + 5X - 8$ et $B = X^3 + X - 2$.
Les divisions euclidiennes successives de l'algorithme d'Euclide donnent

$$\begin{aligned} X^4 + 2X^2 + 5X - 8 &= (X^3 + X - 2)X + (X^2 + 7X - 8) \\ X^3 + X - 2 &= (X^2 + 7X - 8)(X - 7) + (58X - 58) \\ X^2 + 7X - 8 &= (58X - 58)(\frac{1}{58}X + \frac{8}{58}) + 0 \end{aligned}$$

Le dernier reste non nul est $58(X-1)$, donc le PGCD, qui est unitaire et un multiple de $58(X-1)$, est égal à $X - 1$. On peut d'ailleurs calculer que

$$A = (X-1)(X^3 + X^2 + 3X + 8) \qquad B = (X-1)(X^2 + X + 2)$$

Cela confirme que $X - 1$ est bien un diviseur commun de A et B.

Théorème 11 (Bezout). *Soient A, B deux polynômes. Alors il existe deux polynômes U, V tels que*

$$AU + BV = A \wedge B$$

Preuve. Comme pour les entiers, il suffit de remonter l'algorithme d'Euclide. □

Théorème 12 (Gauss). *Soient A, B, C deux polynômes. Si $A \wedge B = 1$ et $A|BC$, alors $A|C$.*

Preuve. Comme pour les entiers. □

Théorème 13. *Soient A, B deux polynômes non nuls.*
Alors il existe un unique polynôme unitaire M tel que
i) M est un multiple commun de A et B,
ii) Si U est un multiple commun de A et B, alors $M|U$.

Preuve. L'unicité est facile à montrer.
Pour l'existence, on prend M égal à un multiple commun de degré minimal. Il est facile de voir que ce polynôme M convient. □

Définition 8. *Soient A, B deux polynômes non nuls, on appelle* **plus petit multiple commun de A et B,** *ou $PPCM(A,B)$ ou $A \vee B$, l'unique polynôme unitaire du théorème précédent.*
Par définition, si $A = 0$ ou $B = 0$, on pose $PPCM(A,B) = 0$.

Bien sûr, le PPCM de A et B est le plus "petit" multiple commun de A et B, au sens suivant :

Théorème 14. *Soient A, B deux polynômes, avec A et B non nuls. L'ensemble des multiples communs unitaires de A et B contient un seul polynôme de degré minimal, et c'est le PPCM de A et B.*

Preuve. Comme pour le PGCD. □

Pour terminer la section de l'arithmétique des polynômes, revenons aux racines. Le théorème suivant est très important.

Théorème 15.
1. Soit $P \in K[X]$ un polynôme non nul.
Si $r_1, r_2, \ldots, r_k$ sont des racines distinctes de P, de multiplicités respectives $m_1, m_2, \ldots, m_k$, alors

$$(X - r_1)^{m_1} \cdots (X - r_k)^{m_k} | P$$

2. Si $P \in K[X]$ est un polynôme de degré $n \geqslant 0$, le nombre de racines distinctes de P est fini, et il vaut au plus n. La somme des multiplicités des racines vérifie

$$\sum_{i=1}^{k} m_i \leqslant n$$

3. (Identification des coefficients)
Si K est un corps **infini**, *et si $\forall x \in K, \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} b_n x^n$, alors $\forall n \in \mathbb{N}, a_n = b_n$.*
4. Soient P, Q deux polynômes non nuls tels que $P|Q$. Soit $a \in K$ un scalaire. On note $m(a,P)$, respectivement $m(a,Q)$, la multiplicité de a dans le polynôme P, respectivement dans le polynôme Q. Alors $m(a,P) \leqslant m(a,Q)$.

Preuve. 1. On sait que $(X - r_1)^{m_1}$ divise P. Donc il existe un polynôme A tel que

$$P = (X - r_1)^{m_1} \cdot A$$

Comme r_2 est une racine de P, $X - r_2$ divise $(X - r_1)^{m_1} \cdot A$. Or les polynômes $X - r_1$ et $X - r_2$ sont premiers entre eux (car $r_1 \neq r_2$). Si on applique m_1 fois le théorème de Gauss, on peut alors dire que

$$X - r_2 | A$$

et écrire $A = (X - r_2)B$. Si r_2 est une racine double (ou plus), on peut alors dire

$$X - r_2 | (X - r_1)^{m_1} \cdot B$$

Par le théorème de Gauss, on arrive alors à $X - r_2 | B$, et ainsi de suite.
A la fin de ce processus nous aurons montré que

$$P = (X - r_1)^{m_1} (X - r_2)^{m_2} \cdot C$$

pour un certain polynôme C. On continue ensuite avec $r_3, r_4, \ldots, r_k$, et on prouve le résultat.

2. Grâce à 1. on voit que si on choisit un nombre fini de racines distinctes $r_1, \ldots, r_k$ de multiplicités $m_1, \ldots, m_k$, alors le degré de

$$(X - r_1)^{m_1} \cdots (X - r_k)^{m_k}$$

est inférieur ou égal au degré de P. Puisque chaque polynôme $X - r_i$ est de degré 1, on en déduit

$$\sum_{i=1}^{k} 1 \cdot m_i \leqslant n$$

On prouve ainsi directement l'inégalité la plus forte.

3. On regarde le polynôme $P = \sum_{n=0}^{\infty}(a_n - b_n)X^n$. D'après l'hypothèse, tout élément r de K est une racine de P. Or K est infini par hypothèse, donc P a une infinité de racines. Or nous savons qu'un polynôme non nul ne peut avoir qu'un nombre fini de racines. Il faut donc que P soit le polynôme nul. On en déduit que

$$\forall n \in \mathbb{N}, a_n - b_n = 0$$

4. est évident. □

7 Polynômes irréductibles. Polynômes scindés.

En arithmétique, nous avons vu que parmi les entiers, il y a des nombres particuliers : les nombres premiers. Ces nombres ne peuvent pas être écrits comme un produit non trivial de deux entiers.

En théorie des polynômes, il y a quelque chose qui ressemble aux nombres premiers. Ce sont les polynômes irréductibles.

Définition 9. *Un polynôme $P \in K[X]$ est appelé* **irréductible** *si*
i) P n'est pas constant,
ii) Si $P = AB$ avec A, B des polynômes, alors A ou B est constant.

Exemples.
1. Tout polynôme de degré 1 est irréductible.
2. Dans $\mathbb{R}[X]$, le polynôme $X^2 + 1$ est irréductible. S'il ne l'était pas, on pourrait écrire $X^2 + 1 = AB$ avec A et B de degré 1, et à coefficients réels. Mais alors A aurait une racine réelle, et $X^2 + 1$ aussi. C'est absurde.
3. Dans $\mathbb{C}[X]$, le polynôme $X^2 + 1$ n'est pas irréductible, car $X^2 + 1 = (X + i)(X - i)$.
4. Si P est un polynôme irréductible de degré $\geqslant 2$ dans $K[X]$, alors P n'a aucune racine.
5. Le polynôme $(X^2 + 1)^2 \in \mathbb{R}[X]$ n'a aucune racine ; pourtant il n'est pas irréductible.

Théorème 16 (Euclide). *Si P est irréductible, et $P|AB$, alors $P|A$ ou $P|B$.*

Preuve. Si $P|A$, c'est fini. On suppose que P ne divise pas A. Il suffit maintenant de montrer que P et A sont premiers entre eux, et on pourra conclure avec le théorème de Gauss pour les polynômes.
Soit D le PGCD de P et A. Comme $D|P$, et P irréductible, il faut que D soit constant, ou proportionnel à P. Mais si D est proportionnel à P, alors de $D|A$ on déduirait $P|A$, ce qui est absurde.
Donc D est constant, et puisque D est unitaire, $D = 1$. □

Nous savons que tout nombre entier positif peut s'écrire d'une manière unique comme produit de nombres premiers. Un résultat de ce type est encore vrai pour les polynômes.

Théorème 17. *Tout polynôme non nul $P \in K[X]$ peut être écrit sous la forme*

$$P = \lambda \prod_{i=1}^{k} P_i,$$

où $\lambda \in K$ et chaque P_i est un polynôme irréductible unitaire.
Cette décomposition de P en produit est unique à l'ordre près.

Preuve. Comme pour les entiers. □

Exemple. Soit $P = 2X^4 - 2 \in \mathbb{R}[X]$. Sa décomposition en produit de polynômes irréductibles unitaires est

$$P = 2(X^2 + 1)(X + 1)(X - 1)$$

Prenons $P = 2X^4 - 2 \in \mathbb{C}[X]$. Sa décomposition en produit de polynômes irréductibles unitaires est

$$P = 2(X + i)(X - i)(X + 1)(X - 1)$$

Définition 10. *Un polynôme $P \in K[X]$ est* **scindé** *si on peut écrire*

$$P = \lambda \prod_{i=1}^{k} (X - r_i)$$

avec $\lambda \in K - \{0\}$, $k \geqslant 0$ et $\forall i \in [[1, k]], r_i \in K$ (les r_i ne sont pas nécessairement distincts)

Exemples.
1. Le polynôme nul n'est pas scindé, alors que les polynômes constants non nuls sont scindés.
2. $X^2 + 1 \in \mathbb{R}[X]$ n'est pas scindé, alors que $X^2 + 1 \in \mathbb{C}[X]$ est scindé.
3. Un polynôme est scindé si et seulement si la somme des multiplicités de ses racines est égale à son degré.
4. Le produit de deux polynômes scindés est aussi un polynôme scindé.
5. Tout diviseur d'un polynôme scindé est encore scindé.

Définition 11. *Soient $r_1, r_2, \ldots, r_n$ des éléments de K. Soit $k \in [[1, n]]$.*
La k-ème fonction symétrique élémentaire des (r_i) *est*

$$\sigma_k = \sum_{1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n} r_{i_1} r_{i_2} \cdots r_{i_k}$$

Exemples. σ_1 est la somme des (r_j) : on a $\sigma_1 = r_1 + \cdots + r_n$.
La deuxième fonction symétrique élémentaire des (r_j) est la somme de tous les produits doubles : on a
$\sigma_2 = r_1 r_2 + r_1 r_3 + \cdots + r_{n-1} r_n$.
La n-ème fonction symétrique élémentaire σ_n est le produit : $\sigma_n = r_1 r_2 \cdots r_n$.

Si P est un polynôme scindé, les fonction symétriques élémentaires symétriques des racines sont étroitement liées aux coefficients du polynôme. C'est l'objet du prochain théorème.

Théorème 18. *Soit $P = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$ un polynôme scindé unitaire de degré n. On appelle $r_1, \ldots, r_n$ ses racines. Alors*

$$\forall k \in [[0, n-1]], \sigma_k = (-1)^k a_{n-k}$$

où σ_k est la k-ème fonction symétrique élémentaire des racines $r_1, \ldots, r_n$.

Preuve. Puisque P est scindé et unitaire, on peut écrire

$$P = (X - r_1)(X - r_2)\cdots(X - r_n)$$

En développant le produit dans le membre de droite on trouve

$$P = X^n - \sigma_1 X^{n-1} + \sigma_2 X^{n-2} + \cdots + (-1)^k \sigma_k X^{n-k} + \cdots + (-1)^n \sigma_n$$

En identifiant les coefficients de P avec les coefficients du polynôme dans le membre de droite, on a le résultat. □

En particulier, si $P = X^2 + AX + B \in \mathbb{C}[X]$, et r_1, r_2 ses deux racines complexes (éventuellement égales), alors on retrouve les résultats bien connus

$$r_1 + r_2 = -A \qquad r_1 r_2 = B$$

On peut donc généraliser à des polynômes scindés de degré arbitraire. On adaptera aussi le théorème ci-dessus à la situation où le polynôme P est scindé,mais pas unitaire.

Citons, sans démonstration (habituellement donnée en analyse), un théorème extrêmement important. C'est **le théorème de Gauss-d'Alembert**, aussi appelé **théorème fondamental de l'algèbre**.

Théorème 19 (Gauss-d'Alembert). *Tout polynôme non nul de $\mathbb{C}[X]$ est scindé.*

Grâce à ce théorème, on peut décrire avec précision les polynômes irréductibles de $\mathbb{C}[X]$ et de $\mathbb{R}[X]$.

Théorème 20. *1. Les polynômes irréductibles de $\mathbb{C}[X]$ sont exactement les polynômes de degré 1.*
2. Les polynômes irréductibles de $\mathbb{R}[X]$ sont exactement les polynômes de degré 1, et les polynômes de degré 2 à discriminant strictement négatif.

Preuve. 1. Nous savons déjà que les polynômes de degré 1 sont irréductibles. Il suffit maintenant de prouver que si $P \in C[X]$ est de degré $n > 1$, alors P n'est pas irréductible. En effet, d'après le théorème de Gauss-d'Alembert, P est scindé, et en particulier on peut écrire

$$P = (X - r)Q$$

Or $X - r$ est de degré 1, et Q de degré $n - 1 \neq 0$. Donc P n'est pas irréductible.

2. Il est facile de voir que les polynômes de degré 2 à discriminant strictement négatif sont irréductibles dans $\mathbb{R}[X]$. Montrons maintenant que si $P \in \mathbb{R}[X]$ est irréductible, alors P est de degré 1, ou P est de degré 2 à $\Delta < 0$.
Premier cas : P a une racine réelle r. Alors il existe $Q \in \mathbb{R}[X]$ avec

$$P = (X - r)Q$$

et comme P est irréductible, Q doit être constant, et alors P est de degré 1.

Second cas : P n'a aucune racine réelle. Si on voit P comme un élément de $\mathbb{C}[X]$, alors P est scindé dans $\mathbb{C}[X]$. Donc P a au moins une racine $z \in \mathbb{C}$. Par hypothèse z n'est pas réelle. Donc $\bar{z}$ est aussi une racine de P, et $z \neq \bar{z}$. Ceci permet d'affirmer que

$$(X - z)(X - \bar{z})|P$$

ou encore, il existe $Q \in \mathbb{C}[X]$ tel que

$$P = Q(X^2 - (2\mathrm{Re}\, z)X + |z|^2)$$

Mais le polynôme $X^2 - (2\mathrm{Re}\, z)X + |z|^2$ a des coefficients **réels**, d'où $Q \in \mathbb{R}[X]$. Comme P est irréductible, il faut que Q soit une constante réelle non nulle a. Mais alors P est de degré 2, et le discriminant de P vaut

$$4a^2(\mathrm{Re}\, z)^2 - 4a^2|z|^2 = -4a^2(\mathrm{Im}\, z)^2 < 0$$

puisque z n'est pas un réel. □

8 Dérivée d'un polynôme. Formule de Taylor.

Définition 12. *Soit $P = \sum_{n=0}^{\infty} a_n X^n$ un polynôme à coefficients dans un corps commutatif K.*
La dérivée de *P est le polynôme $\sum_{n=1}^{\infty} a_n n X^{n-1}$. On le note P' ou parfois $P^{(1)}$. On définit par récurrence la dérivée k-ème de P, notée $P^{(k)}$, comme étant la dérivée du polynôme $P^{(k-1)}$.*

Remarques.
1. Si P est un polynôme non constant et K *un corps de caractéristique nulle*, on a la formule $\deg P' = (\deg P) - 1$. La condition sur la caractéristique est importante. En effet, si K est de caractéristique non nulle, la différence entre le degré de P et le degré de P' peut être strictement supérieure à un. Prenons un exemple où $K = \mathbb{Z}/3Z$, qui est un corps de caractéristique 3. Si $P = X^3 + 2X$, alors $P' = 3X^2 + 2$. Mais $3 = 0$ dans un corps de caractéristique 3. Donc $P' = 2$, et la différence des degrés de P et P' vaut alors 3.
2. Si P est un polynôme de degré n, les dérivées d'ordre $\geqslant n + 1$ de P sont toutes nulles, quelle que soit la caractéristique.

Théorème 21 (Propriétés de la dérivée).
1. L'application $D : K[X] \to K[X] : P \mapsto P'$ est un endomorphisme d'espace vectoriel.
2. Pour tous polynômes P, Q, on a $(PQ)' = P'Q + PQ'$.

Preuve. 1. Montrons que D est linéaire. Soient $P = \sum_{n=0}^{\infty} a_n X^n$ et $Q = \sum_{n=0}^{\infty} b_n X^n$ deux polynômes, et r, s deux scalaires. Alors

$$D(rP + sQ) = (r\sum_{n=0}^{\infty} a_n X^n + s\sum_{n=0}^{\infty} b_n X^n)' = (\sum_{n=0}^{\infty}(ra_n + sb_n)X^n)' =$$

$$= \sum_{n=1}^{\infty}(ra_n + sb_n)nX^{n-1} = r\sum_{n=1}^{\infty} a_n n X^{n-1} + s\sum_{n=1}^{\infty} b_n n X^{n-1} = rP' + sQ' = rD(P) + sD(Q)$$

2. Comme D est linéaire, il est facile de voir qu'il suffit de démontrer la relation $(PQ)' = P'Q + PQ'$ seulement dans le cas où P et Q sont de la forme X^k et X^l. On calcule alors comme ceci :

$$\begin{aligned}(PQ)' &= (X^k X^l)' \\ &= (X^{k+l})' \\ &= (k+l)X^{k+l-1} \\ &= (kX^{k-1})X^l + X^k(lX^{l-1}) \\ &= P'Q + PQ'\end{aligned}$$

□

La relation $(PQ)' = P'Q + PQ'$ permet de prouver par récurrence **la formule de Leibniz**, utile pour calculer les dériveés d'ordre supérieur d'un produit de deux polynômes :

$$(PQ)^{(n)} = \sum_{k=0}^{n} C_n^k P^{(k)} Q^{(n-k)}$$

Théorème 22 (Formule de Taylor). *Soientt K un corps commutatif* **de caractéristique nulle**, *$P \in K[X]$ un polynôme, et $a \in K$ un scalaire. Alors on a la formule*

$$P = \sum_{n=0}^{\infty} \frac{P^{(n)}(a)}{n!}(X-a)^n$$

Remarques. 1. La somme infinie n'a en fait qu'un nombre fini de termes non nuls. En effet, si d est le degré de P, les valeurs $P^{(n)}(a)$ sont nulles pour tout $n > d$.
2. On doit supposer K de caractéristique nulle, car sinon la division par $n!$ est une division par zéro pour certains n. Supposons K de caractéristique p (c'est forcément un nombre premier). Alors pour tout $n \geqslant p$, l'élément $n!$ du corps K est égal à 0, puisque p est alors un facteur de $n!$. Or dans un corps, la division par zéro n'est pas définie. La formule de Taylor perdrait donc son sens si K est de caractéristique non nulle.

Preuve. Si $P = 0$, la formule est évidente. Raisonnons par récurrence sur le degré de P. Si P est de degré nul, la formule de Taylor est encore évidente. On suppose maintenant qu'elle est vraie pour tout les polynômes de degré $d-1$. Prenons P de degré d. Alors a est évidemment une racine du polynôme $P - P(a)$, donc

$$X - a | P - P(a)$$

Si on écrit $P - P(a) = (X-a)Q$, le polynôme Q est de degré $d-1$. On peut alors dire que

$$Q = \sum_{n=0}^{\infty} \frac{Q^{(n)}(a)}{n!}(X-a)^n$$

Si on dérive $n \geqslant 1$ fois la relation $P - P(a) = (X-a)Q$, on obtient grâce à la formule de Leibniz

$$P^{(n)} = (X-a)Q^{(n)} + nQ^{(n-1)}$$

En évaluant en $X = a$ on trouve alors

$$P^{(n)}(a) = nQ^{(n-1)}(a)$$

Dès lors on peut remplacer dans la formule de Taylor pour Q et obtenir

$$Q = \sum_{n=0}^{\infty} \frac{P^{(n+1)}(a)}{(n+1)n!}(X-a)^n = \sum_{n=0}^{\infty} \frac{P^{(n+1)}(a)}{(n+1)!}(X-a)^n$$

Il suffit maintenant de multiplier l'égalité par $X-a$ et d'ajouter $P(a)$. □

Théorème 23. *Soit P un polynôme non nul, et $r \in K$. On suppose K de caractéristique nulle.*
r est racine de multiplicité $m \in \mathbb{N}$ de P $\iff$ $\forall k \in [[0, m-1]], P^{(k)}(r) = 0$ et $P^{(m)}(r) \neq 0$.

Preuve. Montrons la partie "seulement si" par récurrence sur m. Si $m = 0$, il faut montrer que
(r n'est pas une racine de P) $\Rightarrow$ ($P(r) \neq 0$). C'est évident.
Supposons par récurrence que le théorème soit vrai pour $m-1$.
On prend une racine r de multiplicité m de P. Alors r est racine de multiplicité $m-1$ de $Q = \dfrac{P}{X-r}$. Donc $\forall k \in [[0, m-2]], Q^{(k)}(r) = 0$ et $Q^{(m-1)}(r) \neq 0$. Mais

$$P = Q(X-r)$$

et en appliquant la formule de Leibniz, on trouve

$$P^{(k)} = Q^{(k)}(X-r) + kQ^{(k-1)}$$

En posant $X = r$ on obtient la formule

$$\forall k \geqslant 1, P^{(k)}(r) = kQ^{(k-1)}(r) \qquad \text{et} \qquad P(r) = 0$$

On voit tout de suite que $\forall k \in [[0, m-1]], P^{(k)}(r) = 0$. Ensuite $P^{(m)}(r) = mQ^{(m-1)}(r)$. Par hypothèse $Q^{(m-1)}(r) \neq 0$, et $m \neq 0$ parce que nous avons supposé K de caractéristique nulle. Donc, on conclut que $P^{(m)}(r) \neq 0$.

Montrons maintenant la réciproque.
On suppose que $\forall k \in [[0, m-1]], P^{(k)}(r) = 0$ et $P^{(m)}(r) \neq 0$. En appliquant la formule de Taylor (valable car K est de caractéristique nulle), on trouve

$$P = \sum_{k=0}^{\infty} \frac{P^{(k)}(r)}{k!}(X-r)^k = \sum_{k=m}^{\infty} \frac{P^{(k)}(r)}{k!}(X-r)^k$$

Donc P est divisible par $(X-r)^m$, mais pas par $(X-r)^{m+1}$ (puisque $P^{(m)}(r) \neq 0$).

□

Application. Cherchons la multiplicité de la racine $X = 1$ du polynôme

$$P = X^4 - 2X^3 + 2X^2 - 2X + 1 \in \mathbb{R}[X]$$

On a bien $P(1) = 0$. Pour savoir si 1 est une racine au moins double, on peut dériver P.

$$P' = 4X^3 - 6X^2 + 4X - 2$$

Comme $P'(1) = 0$, on sait maintenant que 1 est une racine au moins double. Est-ce qu'elle est triple ?

$$P'' = 12X^2 - 12X + 4$$

Cette fois $P''(1) = 4 \neq 0$, ce qui veut dire que la racine 1 est de multiplicité 2.
D'ailleurs il est facile de vérifier que

$$P = (X-1)^2(X^2+1)$$

ce qui démontre d'une autre manière que 1 est une racine de P de multiplicité égale à 2.