# Aide à la compréhension sur le cours sur les espaces préhilbertiens réels

## Les normes, les produits scalaires et les normes euclidiennes

Dans ce chapitre, on oublie les espaces vectoriels sur les corps $\mathbb{Q}$, $\mathbb{C}$ ou $\mathbb{Z}/7\mathbb{Z}$ pour ne retenir que les $\mathbb{R}$-espaces vectoriels.
Dans la suite, on se donne $E$ un $\mathbb{R}$-espace vectoriel.
La définition d'un produit scalaire est très importante. À retenir : un produit scalaire est une forme bilinéaire symétrique définie positive. On traduit tout ceci.
La définition des formes bilinéaires $\Phi : E^2 \longrightarrow \mathbb{R}$ a été vue dans le chapitre sur les déterminants. La définition des formes symétriques également. Comme ici, on considère des couples de vecteurs, il n'y a que deux permutations dans $\mathfrak{S}_2$, à savoir id et la transposition $(1, 2)$. C'est uniquement la transposition qui est utile et la définition de symétrique donnée en définition 1 coïncide alors avec ce que l'on connaissait, en ne faisant intervenir que la transposition $(1, 2)$.
On dit que la forme $\Phi : E^2 \longrightarrow \mathbb{R}$ est positive lorsque pour tout vecteur $x \in E$, on a :

$$\Phi(x) \geqslant 0.$$

On dit que la forme $\Phi : E^2 \longrightarrow \mathbb{R}$ est définie si le seul vecteur $x \in E$ tel que $\Phi(x, x) = 0_{\mathbb{R}}$ est le vecteur nul.
On peut condenser la définition de « définie positive » en :

$$\forall x \in E,\ x \neq 0_E \Longrightarrow \Phi(x, x) > 0_{\mathbb{R}}.$$

Je conseille cependant de scinder la difficulté en deux en montrant que $\Phi$ est d'abord positive, puis définie.

On retiendra que pour montrer que $\Phi : E^2 \longrightarrow \mathbb{R}$ est un produit scalaire, il faut montrer **dans cet ordre** les choses suivantes :

1. montrer que $\Phi$ est symétrique
2. montrer que $\Phi$ est linéaire par rapport à sa première variable – l'application $\Phi$ sera alors linéaire par rapport à sa deuxième variable par symétrie
3. montrer que $\Phi$ est positive
4. montrer que $\Phi$ est définie – en général, c'est le plus difficile à montrer.

On note $\Phi(\vec{u},\vec{v}) = (\vec{u} \mid \vec{v})$ ou parfois $\vec{u}\cdot\vec{v}$ le produit scalaire. Nous privilégierons la notation $\Phi(\vec{u},\vec{v}) = (\vec{u} \mid \vec{v})$ car l'autre peut être confondue avec la multiplication par un scalaire.
On définit également les notions d'espace préhilbertien réel (ou espace préhilbertien) et d'espace euclidien pour signifier respectivement que l'espace $E$ est de dimension quelconque ou finie, mais surtout que l'espace $E$ est muni d'un produit scalaire $(\cdot \mid \cdot)$.

Passons aux premiers exemples, à travers l'exemple 1.

• Sur $\mathbb{R}^n$, on pose :

$$\forall (x,y) \in \mathbb{R}^n,\ \Phi(x,y) = \sum_{i=1}^{n} x_i \cdot y_i,$$

où $x = (x_1, \cdots, x_n)$ et $y = (y_1, \cdots, y_n)$.
D'une part, il est évident que $\Phi$ est symétrique car pour tout entier $i$ entre 1 et $n$, on a : $x_i \cdot y_i = y_i \cdot x_i$.
Ensuite, il est évident que $\Phi$ est linéaire par rapport à la première variable puisque si $y = (y_1, \cdots, y_n)$ est fixé dans $\mathbb{R}^n$, alors l'application :

$$\Phi_1 : x \longmapsto \sum_{i=1}^{n} x_i \cdot y_i,$$

est clairement linéaire. Il s'agit en fait de la forme linéaire :

$$\Phi_1 = \sum_{i=1}^{n} y_i \cdot e_i^*,$$

où $(e_1^*, \cdots, e_n^*)$ est la base duale de la base canonique de $\mathbb{R}^n$. Par symétrie, l'application $\Phi$ est bilinéaire.
Il est de plus facile de voir que $\Phi$ est positive car si $x \in E$,

$$\Phi(x,x) = \sum_{i=1}^{n} x_i^2 \geqslant 0.$$

On termine par le caractère défini de $\Phi$. Si $x$ est un vecteur de $\mathbb{R}^n$ tel que :

$$\Phi(x,x) = 0,$$

alors $\displaystyle\sum_{i=1}^{n} x_i^2 = 0$.
Cette dernière somme est nulle et constituée de termes positifs : chaque terme est nul amenant le fait que chaque composante $x_i$ du vecteur $x$ est nulle : le vecteur $x$ est nul dans $\mathbb{R}^n$.
Ce produit scalaire est très important dans $\mathbb{R}^n$. On l'appelle le produit scalaire habituel sur l'espace $\mathbb{R}^n$. On verra pourquoi il est si important plus tard, mais lorsqu'en Physique vous travaillez avec des vecteurs et lorsque vous les projetez, c'est ce produit scalaire habituel que vous utilisez dans le plan $\mathbb{R}^2$ ou dans l'espace $\mathbb{R}^3$.

• Pour le deuxième produit scalaire $\Psi : \left|\begin{array}{ccc} C^0([a,b],\mathbb{R}) & \longrightarrow & \mathbb{R} \\ (f,g) & \longmapsto & \displaystyle\int_a^b f \times g \end{array}\right.$, il faut déjà impérativement que $a$ soit strictement inférieur ou $b$. On verra que cela intervient dans le théorème

aux quatre hypothèses. On précise que la notation $C^0([a,b],\mathbb{R})$ désigne l'espace vectoriel de toutes les fonctions $f:[a,b]\longrightarrow\mathbb{R}$ qui sont continues.
Vérifions que $\Psi$ est un produit scalaire.
Comme pour toutes fonctions continues $f$ et $g$, on a $f\times g=g\times f$ – c'est bien une multiplication et non pas une composition qui est d'ailleurs interdite ici – et donc en intégrant $f\times g=g\times f$ entre $a$ et $b$, on obtient la symétrie.
Ensuite, par linéarité de l'intégrale, si $g$ est une fonction continue sur $[a,b]$, l'application $\Psi_1 : f\longmapsto\displaystyle\int_a^b f\times g$ est bien linéaire. Par symétrie, l'application $\Psi$ est bilinéaire.
De plus, si $f\in C^0([a,b],\mathbb{R})$, alors $\Psi(f,f)=\displaystyle\int_a^b f^2$ est positive car l'intégrande est positive et par positivité de l'intégrale et aussi parce que l'on intègre dans le bon sens : $a\leqslant b$.
Enfin, si $\Psi(f,f)=0$, alors on évoque le théorème aux quatre hypothèses. En effet, $a<b$, l'intégrande $f^2$ est continue, positive et d'intégrale nulle. L'intégrande $f^2$ est nulle, et donc pour tout $t\in[a,b]$,

$$f^2(t)=0, \text{ donc } (f(t))^2=0 \text{ et donc } f(t)=0.$$

La fonction $f$ est nulle sur $[a,b]$.

• En notant $\Theta : \left| \begin{array}{ccc} \mathbb{R}[X] & \longrightarrow & \mathbb{R} \\ (P,Q) & \longmapsto & \displaystyle\int_0^1 P(t)\cdot Q(t)\ dt \end{array}\right.$, on va encore détailler le fait que $\Theta$ est un produit scalaire sur $\mathbb{R}[X]$.
Ce genre de produit scalaire qui lie intégrale et polynômes est d'utilisation courante. Il faut absolument retenir les bons arguments qui vont suivre. Je vous donne ici une rédaction typique.

Comme pour tous polynômes $P(X)$ et $Q(X)$ dans $\mathbb{R}[X]$, $P\times Q=Q\times P$, alors pour tout $t\in[0,1]$, $P(t)\cdot Q(t)=Q(t)\cdot P(t)$ et en intégrant, l'application $\Theta$ est bien symétrique.
Si $Q(X)$ est fixé dans $\mathbb{R}[X]$, si $\lambda\in\mathbb{R}$ et $P_1$ puis $P_2$ sont deux polynômes dans $\mathbb{R}[X]$, alors :

$$\begin{aligned}
\Theta(\lambda\cdot P_1+P_2,Q) &= \int_0^1 \Big(\lambda\cdot P_1(t)+P_2(t)\Big)\cdot Q(t)\ dt\\
&= \int_0^1 \Big(\lambda\cdot P_1(t)\cdot Q(t)+P_2(t)\cdot Q(t)\Big)\ dt\\
&= \lambda\cdot\int_0^1 P_1(t)\cdot Q(t)\ dt+\int_0^1 P_2(t)\cdot Q(t)\ dt, \text{ par linéarité de l'intégrale}\\
&= \lambda\cdot\Theta(P_1,Q)+\Theta(P_2,Q).
\end{aligned}$$

L'application $\Theta$ est bien linéaire par rapport à la première variable, donc bilinéaire car elle est symétrique.
Si $P(X)\in\mathbb{R}[X]$, alors on a directement $\Theta(P,P)=\displaystyle\int_0^1 P^2(t)\ dt\geqslant 0$, par positivité de l'intégrale.
Si $P(X)\in\mathbb{R}[X]$ vérifie $\Theta(P,P)=0$, alors $\displaystyle\int_0^1 P^2(t)\ dt=0$. On utilise dans un premier temps le théorème aux quatre hypothèses. Comme $0<1$, comme $t\longmapsto P^2(t)$ est continue, positive et

d'intégrale nulle, alors l'intégrande $t \longmapsto P^2(t)$ est nulle. Par conséquent, pour tout $t \in [0,1]$, $P^2(t) = 0$, donc $(P(t))^2 = 0$, puis $P(t) = 0$. On utilise ensuite les racines du polynôme $P(X)$. Comme pour tout $t \in [0,1]$, $P(t) = 0$, alors le polynôme $P(X)$ admet une infinité de racines : au moins tous les réels entre 0 et 1. Le polynôme $P(X)$ est le polynôme nul.

- On pose :

$$\Lambda : \left| \begin{array}{ccc} \mathbb{R}_n[X] & \longrightarrow & \mathbb{R} \\ (P,Q) & \longmapsto & \displaystyle\sum_{i=0}^{n} P(i) \cdot Q(i) \end{array} \right. .$$

Il est évident que l'application $\Lambda$ est symétrique par commutativité du produit dans $\mathbb{R}$.
Si $Q(X)$ est dans $\mathbb{R}_n[X]$, il est facile de voir que l'application $P(X) \longmapsto \displaystyle\sum_{i=0}^{n} P(i) \cdot Q(i)$ est linéaire. En fait, il s'agit de la forme :

$$\sum_{i=0}^{n} Q(i) \cdot \mathrm{ev}_i.$$

Ensuite, il est évident que $\Lambda$ est positive car si $P$ est dans $\mathbb{R}_n[X]$, alors :

$$\Lambda(P,P) = \sum_{i=0}^{n} \Big(P(i)\Big)^2 \geqslant 0.$$

Enfin, si $P$ est un polynôme dans $\mathbb{R}_n[X]$ tel que $\Lambda(P,P) = 0$, alors la somme $\displaystyle\sum_{i=0}^{n} \Big(P(i)\Big)^2$ est nulle et composée de termes positifs. Chaque terme $P(i)$ est nul, pour $i$ variant dans l'ensemble à $(n+1)$ éléments $[\![0,n]\!]$. Le polynôme $P(X)$ admet au moins $(n+1)$ racines et est de degré au maximum $n$, donc admet strictement plus de racines que son degré : c'est le polynôme nul.

**Remarque** : l'application $\Sigma : \left| \begin{array}{ccc} \mathbb{R}_n[X] & \longrightarrow & \mathbb{R} \\ (P,Q) & \longmapsto & \displaystyle\sum_{i=1}^{n} P(i) \cdot Q(i) \end{array} \right.$ n'est pas un produit scalaire.

Elle vérifie pratiquement tout, sauf le caractère défini. En effet, en posant $P(X) = \displaystyle\prod_{i=1}^{n}(X-i)$, alors $P$ n'est pas le polynôme nul dans $\mathbb{R}_n[X]$ et pourtant :

$$\Sigma(P,P) = 0.$$

On termine par le produit scalaire sur $\mathscr{M}_n(\mathbb{R})$.
On va effectuer deux rédactions possibles pour ensuite les comparer en terme d'efficacité.
On pose dans la suite :

$$\Phi : \left| \begin{array}{ccc} \mathscr{M}_n(\mathbb{R})^2 & \longrightarrow & \mathbb{R} \\ (A,B) & \longmapsto & \mathrm{Tr}(A^T B) \end{array} \right. .$$

**méthode 1 : la méthode scolaire**
On vérifie les points les uns après les autres.
Si $A$ et $B$ sont deux matrices dans $\mathscr{M}_n(\mathbb{R})$, alors :

$$\Phi(A,B)=\mathrm{Tr}(A^TB)=\mathrm{Tr}\Big((A^TB)^T\Big)=\mathrm{Tr}(B^TA)=\Phi(B,A).$$

Ensuite, si $B$ est une matrice dans $\mathscr{M}_n(\mathbb{R})$, l'application $A\longmapsto\Phi(A,B)$ est linéaire car la transposée est linéaire, le produit est distributif et la trace est linéaire. Par symétrie, l'application $\Phi$ est bilinéaire.
Ensuite, si $A$ est une matrice dans $\mathscr{M}_n(\mathbb{R})$, alors :

$$\begin{aligned}\Phi(A,A) &= \mathrm{Tr}(A^TA)\\ &= \sum_{i=1}^{n}\Big(A^TA\Big)_{i,i}\\ &= \sum_{i=1}^{n}\left(\sum_{k=1}^{n}(A^T)_{i,k}\cdot A_{k,i}\right)\\ &= \sum_{i,k=1}^{n}A_{k,i}^2\geqslant 0.\end{aligned}$$

Si $A$ est une matrice dans $\mathscr{M}_n(\mathbb{R})$ telle que $\Phi(A,A)=0$, alors :

$$\sum_{i,k=1}^{n}A_{k,i}^2=0,$$

et chaque terme positif de la somme nulle précédente est nul : tous les coefficients de la matrice $A$ sont nuls et $A$ est la matrice nulle.

**méthode 2 : la méthode la plus efficace : celle à privilégier**
On commence par calculer $\mathrm{Tr}(A^TB)$ en fonction des coefficients des matrices $A$ et $B$. On obtient :

$$\begin{aligned}\Phi(A,B) &= \mathrm{Tr}(A^TB)\\ &= \sum_{i=1}^{n}\Big(A^TB\Big)_{i,i}\\ &= \sum_{i=1}^{n}\left(\sum_{k=1}^{n}(A^T)_{i,k}\cdot B_{k,i}\right)\\ &= \sum_{i,k=1}^{n}A_{k,i}\cdot B_{k,i}\\ &= \sum_{i,j=1}^{n}A_{i,j}\cdot B_{i,j}\end{aligned}$$

En identifiant les espaces $\mathscr{M}_n(\mathbb{R})$ et $\mathbb{R}^{n^2}$, on observe alors que l'application $\Phi$ est le produit habituel sur $\mathbb{R}^{n^2}$, en référence au premier exemple de l'exemple 1 en remplaçant $n$ par $n^2$.

La formule $\Phi(A,B) = \sum_{i,j=1}^{n} A_{i,j} \cdot B_{i,j}$ est clairement symétrique, linéaire par rapport à la première composante $A$, positive et définie positive, en retrouvant la formule déjà trouvée dans la méthode 1, à savoir :

$$\Phi(A,A) = \mathrm{Tr}(A^T A) = \sum_{i,j=1}^{n} A_{i,j}^2.$$

La définition d'une norme est également très importante. Elle servira énormément en seconde année.
Pour montrer qu'une application $N : E \longrightarrow \mathbb{R}$ définie sur un $\mathbb{R}$-espace vectoriel $E$ est une norme, on vérifiera en général les points dans cet ordre, pour aborder les choses en difficulté croissante :

1. On montre que la norme est positive, à savoir que pour tout vecteur $u \in E$, on a :
$$N(u) \geqslant 0_{\mathbb{R}}.$$
2. On montre ensuite que la norme est homogène, à savoir que pour tout scalaire réel $\lambda$ et tout vecteur $u \in E$,
$$N(\lambda \cdot u) = |\lambda| \cdot N(u).$$
3. On montre ensuite que la norme est séparante, à savoir que si $N(u) = 0_{\mathbb{R}}$, alors $u = 0_E$.
4. On termine par l'inégalité triangulaire, à savoir que pour tous vecteurs $u$ et $v$,
$$N(u+v) \leqslant N(u) + N(v).$$

On remarque des similarités entre les définitions de produit scalaire et de norme : ne vous mélangez pas les pinceaux !

La propriété 1 nous apprend que les produits scalaires permettent de définir des normes, normes que l'on qualifiera de normes euclidiennes.

Démontrons la propriété 1.
On se donne un espace préhilbertien réel $\Big(E, (\cdot \mid \cdot)\Big)$, c'est-à-dire un $\mathbb{R}$-espace vectoriel $E$ (de dimension quelconque) et d'un produit scalaire $(\cdot \mid \cdot)$.
On pose l'application :

$$N : \left| \begin{array}{ccl} E & \longrightarrow & \mathbb{R} \\ x & \longmapsto & N(x) = \sqrt{(x \mid x)} \end{array} \right. .$$

Première chose : cette application est bien définie car par positivité du produit scalaire, pour tout vecteur $x$ dans $E$, on sait que $(x \mid x) \geqslant 0$, donc on peut prendre la racine carrée.

Deuxième chose : il devient alors évident que pour tout vecteur $x \in E$, la quantité $N(x) = \sqrt{(x \mid x)}$ est positive.

Troisième chose : si $x$ est un vecteur de $E$ et si $\lambda \in \mathbb{R}$, alors on peut écrire :

$$\begin{aligned}
N(\lambda \cdot x) &= \sqrt{\left(\lambda \cdot x \mid \lambda \cdot x\right)} \\
&= \sqrt{\lambda \cdot \left(x \mid \lambda \cdot x\right)} \text{ , produit scalaire linéaire première variable} \\
&= \sqrt{\lambda^2 \cdot \left(x \mid x\right)} \text{ , produit scalaire linéaire deuxième variable} \\
&= \sqrt{\lambda^2} \cdot N(x) \\
&= |\lambda| \cdot N(x).
\end{aligned}$$

Quatrième chose : si $x$ est un vecteur de $E$ tel que $N(x) = 0_{\mathbb{R}}$, alors $(x \mid x) = 0$ et par le caractère défini du produit scalaire, le vecteur $x$ est nul dans $E$.

Cinquième et dernière chose : on regarde l'inégalité triangulaire. Celle-ci va nous prendre un peu plus de temps.
On fixe deux vecteurs $x$ et $y$ dans $E$. On distingue deux cas.

- Si $y$ est le vecteur nul, alors les inégalités suivantes :

$$\boxed{|(x \mid y)| \leqslant N(x) \cdot N(y)}$$

et :

$$\boxed{N(x+y) \leqslant N(x) + N(y)}$$

sont en fait des égalités, car par linéarité du produit scalaire par rapport à la deuxième variable, alors :

$$(x \mid y) = (x \mid 0_E) = O_{\mathbb{R}}.$$

- Si le vecteur $y$ n'est pas nul, on pose la fonction :

$$f : t \longmapsto N(x + t \cdot y)^2 = (x + t \cdot y \mid x + t \cdot y).$$

Par bilinéarité du produit scalaire, on peut tout développer – un peu comme on pourrait le faire avec un produit de nombres. On obtient :

$$\forall t \in \mathbb{R},\ f(t) = (x \mid x) + t \cdot (x \mid y) + t \cdot (y \mid x) + t^2 \cdot (y \mid y) = N(x)^2 + 2t\,(x \mid y) + t^2 \cdot N(y)^2.$$

On vient d'utiliser également la symétrie du produit scalaire pour regrouper les deux termes $(x \mid y)$ et $(y \mid x)$. De plus, comme $y$ n'est pas le vecteur nul, alors $N(y)$ est un réel strictement positif. La fonction $f$ est donc une fonction polynomiale de degré 2. Par ailleurs, pour tout $t \in \mathbb{R}$, la quantité $f(t)$ est positive par positivité du produit scalaire. Il est impossible que la fonction polynomiale admette deux racines réelles différentes, car sinon, il y aurait changement de signe et la fonction $f$ prendrait des valeurs strictement négatives entre ces deux racines. Il ne reste que les deux possibilités : la fonction $f$ ne

s'annule qu'une seule fois (le discriminant $\Delta$ sera alors nul) ou la fonction $f$ ne s'annule pas (le discriminant $\Delta$ sera alors strictement négatif).
Quoiqu'il arrive, le discriminant $\Delta$ est négatif ou nul. Or, le discriminant $\Delta$ vaut :

$$\Delta = \Big(2(x \mid y)\Big)^2 - 4\ N(x)^2 \cdot N(y)^2 = 4\Big((x \mid y)^2 - N(x)^2 \cdot N(y)^2\Big) \leqslant 0.$$

On obtient ainsi, en prenant les racines carrées, l'inégalité :

$$\boxed{|(x \mid y)| \leqslant N(x) \cdot N(y)}$$

D'autre part, on va comparer $N(x+y)$ à $N(x)+N(y)$, ou encore, comme tout est positif, on va comparer $(N(x+y))^2$ à $(N(x)+N(y))^2$. On pose $A$ la différence. On obtient alors :

$$\begin{aligned} A &= (N(x) + N(y))^2 - (N(x+y))^2 \\ &= (N(x))^2 + (N(y))^2 + 2N(x)\ N(y) - (x+y \mid x+y) \\ &= (x \mid x) + (y \mid y) + 2N(x)\ N(y) - \Big((x \mid x) + 2(x \mid y) + (y \mid y)\Big) \\ &= 2\Big(N(x)\ N(y) - (x \mid y)\Big) \geqslant 0. \end{aligned}$$

On obtient ainsi l'inégalité voulue :

$$\boxed{N(x+y) \leqslant N(x) + N(y)}$$

Quoiqu'il arrive, on a bien l'inégalité triangulaire, avec en prime l'inégalité que l'on appelle inégalité de Cauchy-Schwarz :

$$\forall (x,y) \in E^2,\ \Big|(x \mid y)\Big| \leqslant N(x) \cdot N(y).$$

On vient de montrer la première moitié de la propriété 3 et de terminer la démonstration de la propriété 1.

Il est pratique de noter $\|\cdot\|$ la norme d'un vecteur, qui est une notation traditionnelle.
Il est à noter que lorsque l'on vous donne un $\mathbb{R}$-espace vectoriel $E$, avec un produit scalaire $(\cdot \mid \cdot)$, la notation $\|\cdot\|$ fait implicitement référence à une norme. Ensuite, en l'absence d'autres informations, cette norme est la norme euclidienne issue du produit scalaire, et donc définie par la formule :

$$\forall x \in E,\ \|x\| = \sqrt{(x \mid x)}.$$

Bref, connaissant le produit scalaire, on a directement la norme euclidienne associée.

La propriété 2 est relativement évidente. Il suffit de développer par bilinéarité du produit scalaire. On refait le calcul. Pour tous vecteurs $x$ et $y$ dans $E$,

$$\begin{aligned} \|x+y\|^2 &= (x+y \mid x+y) \\ &= (x \mid x) + (x \mid y) + (y \mid x) + (y \mid y) \\ &= \|x\|^2 + 2\ (x \mid y) + \|y\|^2. \end{aligned}$$

On découvre que connaissant uniquement la norme euclidienne $\|\cdot\|$, on peut retrouver le produit scalaire grâce par exemple à la formule :

$$\forall (x,y) \in E^2,\ (x \mid y) = \frac{1}{2}\Big(\|x+y\|^2 - \|x\|^2 - \|y\|^2\Big).$$

On dispose d'autres formules permettant d'obtenir le produit scalaire à partir de la norem euclidienne. Par exemple,

$$\forall (x,y) \in E^2,\ (x \mid y) = \frac{\|x+y\|^2 - \|x-y\|^2}{4}.$$

En effet, si $x$ et $y$ sont deux vecteurs de $E$, alors en utilisant la bilinéarité du produit scalaire :

$$\|x+y\|^2 = \|x\|^2 + 2\ (x \mid y) + \|y\|^2 \qquad [\text{ ligne } L_1\ ]$$

et :

$$\|x-y\|^2 = \|x\|^2 - 2\ (x \mid y) + \|y\|^2 \qquad [\text{ ligne } L_2\ ]$$

Il suffit alors d'effectuer l'opération $L_1 - L_2$, pour avoir $4\ (x \mid y)$ et obtenir ce qu'il faut.
Tout cela pour dire qu'il existe plein de formules exprimant le produit scalaire uniquement à l'aide de la norme euclidienne. Toutes ces formules plus ou moins aussi inutiles les unes que les autres portent cependant des noms ; c'est ce que l'on appelle des identités de polarisation. Cela ne sert pas souvent à grand chose...

L'inégalité de Cauchy-Schwarz de la propriété 3 a déjà été montrée. Cette inégalité provient directement d'un calcul de discriminant négatif ou nul dans un polynôme de degré 2.
Passons au cas d'égalité dans l'inégalité de Cauchy-Schwarz.
Supposons les vecteurs $x$ et $y$ colinéaires. Par exemple, on peut écrire $y = \lambda \cdot x$. On en déduit :

$$(x \mid y) = (x \mid \lambda \cdot x) = \lambda \cdot (x \mid x),$$

donc $|(x \mid y)| = |\lambda| \cdot \|x\|^2$.
Or,

$$\begin{array}{rcl} \|x\| \cdot \|y\| & = & \|x\| \cdot \|\lambda \cdot x\| \\ & = & \|x\| \cdot |\lambda| \cdot \|x\|, \end{array}$$

et l'égalité $|(x \mid y)| = \|x\| \cdot \|y\|$ est acquise.
Pour la réciproque, supposons donnés deux vecteurs $x$ et $y$ tels que :

$$|(x \mid y)| = \|x\| \cdot \|y\|.$$

Si le vecteur $y$ est nul, alors le vecteur $y$ est colinéaire à tous les vecteurs, en particulier, $y$ est colinéaire à $x$.
Si le vecteur $y$ est non nul, en reprenant la fonction polynomiale :

$$f : t \longmapsto \|x + t \cdot y\|^2 = \|x\|^2 + 2t\ (x \mid y) + t^2\ \|y\|^2,$$

alors le discriminant $\Delta$ vaut :

$$\Delta = 4\Big((x \mid y)^2 - \|x\|^2 \cdot \|y\|^2\Big) = 0.$$

Cela signifie que la fonction polynomiale $f$ admet une racine double $t_0$, et donc que :

$$f(t_0) = 0.$$

Ceci se réécrit :

$$\|x + t_0 \cdot y\|^2 = 0, \text{ puis } \|x + t_0 \cdot y\| = 0$$

et donc par définie positivité du produit scalaire :

$$x + t_0 \cdot y = 0_E.$$

La famille $(x, y)$ est donc liée et les vecteurs $x$ et $y$ sont bien colinéaires.

Cette inégalité de Cauchy-Schwarz est quelque chose d'important dans ce chapitre.

Examinons maintenant le corollaire du cas d'égalité dans l'inégalité triangulaire.
Supposons donnés deux vecteurs $x$ et $y$ dans $E$, colinéaires et de même sens. Par exemple, pour changer on peut écrire :

$$x = \alpha \cdot y \text{ , avec } \alpha \text{ un réel positif ou nul.}$$

On en déduit

$$\begin{array}{rcl}
\|x + y\| & = & \|(1 + \alpha) \cdot x\| \\
 & = & |1 + \alpha| \cdot \|x\| \\
 & = & (1 + \alpha) \cdot \|x\|, \text{ car } 1 + \alpha \geqslant 0 \\
 & = & \|x\| + \alpha \cdot \|x\| \\
 & = & \|x\| + |\alpha| \cdot \|x\|, \text{ car } \alpha \geqslant 0 \\
 & = & \|x\| + \|\alpha \cdot x\| \\
 & = & \|x\| + \|y\|.
\end{array}$$

On a bien égalité dans l'inégalité triangulaire.
Réciproquement, supposons donnés deux vecteurs $x$ et $y$ dans $E$ tels que :

$$\|x + y\| = \|x\| + \|y\|.$$

Si l'un des deux vecteurs $x$ ou $y$ est nul, comme le vecteur nul est colinéaire et de même sens que n'importe quel autre vecteur, on a ce qu'il faut.
Plaçons-nous maintenant dans le cas où les deux vecteurs $x$ et $y$ sont non nuls.
On en déduit en élevant l'égalité triangulaire sur les normes au carré, puis en utilisant le bilinéarité du produit scalaire :

$$\begin{array}{rcl}
\|x\|^2 + 2\ \|x\| \cdot \|y\| + \|y\|^2 & = & \Big(\|x\| + \|y\|\Big)^2 \\
 & = & \|x + y\|^2 \\
 & = & \|x\|^2 + 2\ (x \mid y) + \|y\|^2.
\end{array}$$

On en déduit :

$$(x \mid y) = \|x\| \cdot \|y\|.$$

La quantité $(x \mid y)$ est positive et on a égalité dans l'inégalité de Cauchy-Schwarz. Les vecteurs $x$ et $y$ sont donc colinéaires. On ne sait pas encore qu'ils sont de même sens. On écrit par exemple :

$$y = \beta \cdot x,$$

pour un certain réel $\beta$.
La quantité positive $(x \mid y)$ vaut :

$$\beta \; \|x\|^2,$$

par linéarité du produit scalaire par rapport à la seconde variable.
Finalement,

$$\beta = \frac{(x \mid y)}{\|x\|^2} \geqslant 0,$$

et les vecteurs $x$ et $y$ sont bien colinéaires de même sens.

Passons en revue l'exemple 2.
Voici un premier exemple d'inégalité de Cauchy-Schwarz dans $\mathbb{R}^n$ muni du produit scalaire habituel :

$$\forall \Big( x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_n) \Big) \in \mathbb{R}^n \times \mathbb{R}^n, \; \left| \sum_{i=1}^n x_i \cdot y_i \right| \leqslant \sqrt{\sum_{i=1}^n x_i^2} \times \sqrt{\sum_{j=1}^n y_j^2}.$$

• Le premier point de l'exemple 2 provient directement de l'inégalité de Cauchy-Schwarz pour les vecteurs $f$ et $g$ dans l'espace préhilbertien $E = C^0([a,b], \mathbb{R})$ muni du produit scalaire :

$$\Phi : (h_1, h_2) \longmapsto \int_a^b h_1 \times h_2.$$

• Le produit scalaire habituel $(A, B) \longmapsto \mathrm{Tr}(A^T B)$ sur $\mathscr{M}_n(\mathbb{R})$ définit la norme euclidienne associée :

$$\forall A \in \mathscr{M}_n(\mathbb{R}), \; \|A\|_2 = \sqrt{\mathrm{Tr}(A^T A)} = \sqrt{\sum_{i,j=1}^n (A_{i,j})^2}.$$

Montrons qu'il s'agit d'une norme d'algèbres, autrement dit que la norme d'un produit est inférieure au produit des normes.
Fixons $A$ et $B$ dans $\mathscr{M}_n(\mathbb{R})$. Alors,

$$\begin{aligned} \|AB\|_2^2 &= \sum_{i,j=1}^n (AB)_{i,j}^2 \\ &= \sum_{i,j=1}^n \left( \sum_{k=1}^n A_{i,k}, B_{k,j} \right)^2. \end{aligned}$$

On utilise maintenant l'inégalité de Cauchy-Schwarz sur l'espace euclidien hanituel $\mathbb{R}^n$ muni de fait du produit scalaire habituel.
Fixons deux indices $i$ et $j$ entre 1 et $n$. On pose le vecteur :

$$u = (A_{i,k})_{1 \leqslant k \leqslant n}$$

et le vecteur :

$$v = (B_{k,j})_{1 \leqslant k \leqslant n}.$$

On en déduit par l'inégalité de Cauchy-Schwarz :

$$\left( \sum_{k=1}^{n} A_{i,k}, B_{k,j} \right)^2 = (u \mid v)^2 \leqslant \left( \sum_{r=1}^{n} (A_{i,r})^2 \right) \times \left( \sum_{s=1}^{n} (B_{s,j})^2 \right).$$

Par conséquent,

$$\begin{aligned} \|AB\|_2^2 &\leqslant \sum_{i,j=1}^{n} \left( \sum_{r=1}^{n} (A_{i,r})^2 \right) \times \left( \sum_{s=1}^{n} (B_{s,j})^2 \right) \\ &= \left( \sum_{i,r=1}^{n} (A_{i,r})^2 \right) \times \left( \sum_{j,s=1}^{n} (B_{s,j})^2 \right) \\ &= \|A\|_2^2 \times \|B\|_2^2. \end{aligned}$$

On obtient ce qu'il faut en prenant la racine carrée croissante sur $[0, +\infty[$.

En ce qui concerne l'exemple 3, on propose trois normes et on souhaite vérifier s'il s'agit d'une norme euclidienne. Autrement dit, si $N : E \longrightarrow \mathbb{R}$ est une norme, la question qui se pose est : *« existe-t-il un produit scalaire* $(\cdot \mid \cdot)$ *sur* $E$ *tel que :*

$$\forall x \in E,\ \|x\| = \sqrt{(x \mid x)} \quad ? \text{»}$$

On peut éventuellement utiliser des identités de polarisation. On va plutôt tester l'égalité dans l'inégalité triangulaire...
Les trois normes $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ portent des notations officielles sur l'espace $\mathbb{R}^n$.
Vérifions par exemple que $\|\cdot\|_\infty$ est une norme.
Premièrement, pour tout $x = (x_1, \cdots, x_n)$, la quantité $\max\{|x_i| \ ; \ 1 \leqslant i \leqslant n\}$ est bien positive.
Ensuite, si $\lambda$ est un scalaire, en notant $|x_i| = \|x\|_\infty$, alors le nombre $|\lambda \ x_i|$ sera la composante maximale en module dans le vecteur $\lambda \cdot x$, d'où :

$$\|\lambda \cdot x\|_\infty = |\lambda| \cdot \|x\|_\infty.$$

Ensuite, si $\|x\|_\infty = 0$, alors pour tout entier $i$ entre 1 et $n$,

$$|x_i| \leqslant 0,$$

donc $x_i = 0$ et $x$ est le vecteur nul.

De plus, si $x$ et $y$ sont dans $\mathbb{R}^n$, pour tout entier $i$ entre 1 et $n$, en utilisant l'inégalité triangulaire sur $\mathbb{R}$ :
$$|x_i + y_i| \leqslant |x_i| + |y_i| \leqslant \|x\|_\infty + \|y\|_\infty.$$
Ceci étant valable pour tout indice $i$, alors :
$$\|x + y\|_\infty \leqslant \|x\|_\infty + \|y\|_\infty.$$
A priori, il est plus compliqué de montrer que la norme $\|\cdot\|_2$ est une norme sur $\mathbb{R}^n$. On remarque – on commence à avoir l'habitude – que la norme $\|\cdot\|_2$ est la norme euclidienne associée au produit scalaire habituel sur $\mathbb{R}^n$.
Lorsque $n = 1$, cette norme euclidienne coïncide avec la valeur absolue et lorsque $n = 1$, les trois normes proposées dans l'exemple 3 sont les mêmes, donc ce sont des normes euclidiennes dans ce cas.
On pourrait vérifier que $\|\cdot\|_1$ est bien une norme : c'est facile à faire.
On va montrer que lorsque l'entier $n$ est supérieur ou égal à 2, alors les normes $\|\cdot\|_1$ et $\|\cdot\|_\infty$ ne sont pas euclidiennes.
Pour la norme $\|\cdot\|_1$, prenons la base canonique $(e_1, \cdots, e_n)$ de $\mathbb{R}^n$. On constate que :
$$\|e_1 + e_2\|_1 = 1 + 1 = 2 = \|e_1\|_1 + \|e_2\|_1.$$
On a égalité dans l'inégalité triangulaire et pourtant les deux vecteurs $e_1$ et $e_2$ ne sont pas colinéaires de même sens. La norme $\|\cdot\|_1$ ne peut être une norme euclidienne.
Pour la norme $\|\cdot\|_\infty$, on constate par exemple que :
$$\|2\ e_1 + e_2\|_\infty = 2 = \|e_1 + e_2\|_\infty + \|e_1\|_\infty.$$
On a égalité dans l'inégalité triangulaire et pourtant les deux vecteurs $e_1 + e_2$ et $e_1$ ne sont toujours pas colinéaires de même sens. La norme $\|\cdot\|_\infty$ ne peut pas non plus être une norme euclidienne.

## Les familles orthonormales et les bases orthonormales

Dans la suite, si $E$ est un espace muni d'un produit scalaire, on dit qu'un vecteur $u \in E$ est unitaire lorsque $\|u\| = 1$, où $\|\cdot\|$ est la norme euclidienne associée au produit scalaire sur $E$.
On dit que deux vecteurs $u$ et $v$ sont orthogonaux si leur produit scalaire est nul. On peut noter alors en abgrégé : « $u \perp v$ ».
Une remarque sur le mot « unitaire » que l'on connaît également pour les polynômes : problème, ce mot ne désigne pas toujours la même chose. Par exemple, dans l'espace préhilbertien $E = \mathbb{R}[X]$ muni du produit scalaire :
$$(P, Q) \longmapsto \int_0^3 P(t)\ Q(t)\ dt,$$
le polynôme $P(X) = X$ est un polynôme unitaire dans le sens où son coefficient dominant vaut 1, mais la norme de ce vecteur $P(X)$ vaut :
$$\|P(X)\| = \sqrt{\int_0^3 t^2\ dt} = \sqrt{9} = 3,$$

qui ne vaut pas 1. Dans cette situation, on fera bien attention à se rendre intelligible car le mot « unitaire » peut signifier deux choses différentes. En général, les choses se passent assez bien et cette situation est assez peu courante.

La définition d'une famille orthogonale est assez intuitive.
Une famille orthonormale (en abgrégé une f.o.n.) est une famille orthogonale à vecteurs unitaires.

Soit $\mathscr{F} = (x_i)_{i \in I}$ une famille orthonormale (ici quelconque, pas forcément finie).
Soit $\displaystyle\sum_{k=1}^{n} \lambda_k \cdot x_{i_k} = 0_E$ une combinaison linéaire nulle entre les vecteurs de la famille $\mathscr{F}$.
Fixons un entier $j$ entre 1 et $n$. On effectue le produit scalaire par le vecteur $x_{i_j}$ dans la combinaison linéaire nulle. Voici ce que cela donne :

$$\begin{aligned}
0_{\mathbb{R}} &= (0_E \mid x_{i_j}) \\
&= \left( \sum_{k=1}^{n} \lambda_k \cdot x_{i_k} \mid x_{i_j} \right) \\
&= \sum_{k=1}^{n} \lambda_k \cdot (x_{i_k} \mid x_{i_j}) \\
&= \sum_{k=1}^{n} \lambda_k \cdot \delta_{i_k, i_j} \\
&= \lambda_j.
\end{aligned}$$

On en déduit que chaque scalaire $\lambda_j$ est nul et donc que la famille $\mathscr{F}$ est libre. On a ici utilisé la linéarité du produit scalaire par rapport à la première variable. On aurait pu aussi effectuer dès le départ le produit scalaire avec le vecteur $x_{i_j}$ à gauche et utiliser la linéarité du produit scalaire par rapport à la seconde variable.
On remarque que la démonstration précédente montre la chose suivante : toute famille orthogonale composée de vecteurs non nuls est toujours libre.

Traitons maintenant l'exemple 4 qui vous rappelle peut-être les séries de Fourier si vous en avez fait en Physique...
On montre déjà que la formule proposée donne bien un produit scalaire sur l'espace $E$ des fonctions $f : \mathbb{R} \longrightarrow \mathbb{R}$ continues et $2\pi$-périodiques (c'est bien un $\mathbb{R}$-espace vectoriel).
La symétrie ne fait aucun doute et la linéarité de l'intégrale nous indique immédiatement que la formule proposée est linéaire par rapport à la première variable. Par symétrie, la formule proposée est bilinéaire.
Ensuite la quantité $\displaystyle\frac{1}{2\pi} \int_0^1 (f(t))^2 \; dt$ est toujours positive et si cette quantité est nulle, le théorème aux quatre hypothèses s'applique et la fonction $f$ est nulle sur l'intervalle $[0, 2\pi]$. Comme la fonction $f$ est $2\pi$-périodique, alors la fonction $f$ déjà nulle sur la période $[0, 2\pi]$ est en fait nulle partout.

On note la famille $\mathscr{F} = \Big( c_n ; n \in \mathbb{N} ; s_n ; n \in \mathbb{N}^* \Big)$.

On va montrer que la famille $\mathscr{F}$ est orthogonale à vecteurs non nuls : elle sera donc libre.
Détaillons déjà le calcul de $(c_k \mid c_\ell)$, lorsque $k \neq \ell$ dans $\mathbb{N}$, puis $(s_k \mid s_\ell)$, lorsque $k \neq \ell$ dans $\mathbb{N}^*$ et enfin $(c_k \mid s_\ell)$, lorsque $k \in \mathbb{N}$ et $\ell \in \mathbb{N}^*$.

Fixons $k \neq \ell$ dans $\mathbb{N}$. On utilise la formule :

$$\cos a \cdot \cos b = \frac{\cos(a-b) + \cos(a+b)}{2}.$$

Ainsi,

$$\begin{aligned}
(c_k \mid c_\ell) &= \frac{1}{2\pi} \int_0^{2\pi} \cos(kt)\ \cos(\ell t)\ dt \\
&= \frac{1}{2\pi} \int_0^{2\pi} \frac{\cos((k-\ell)t) + \cos((k+\ell)t)}{2}\ dt \\
&= \frac{1}{4\pi} \left[ \frac{\sin((k-\ell)t)}{k-\ell} + \frac{\sin((k+\ell)t)}{k+\ell} \right]_0^{2\pi}, \text{ car } k-\ell \neq 0 \text{ et } k+\ell \neq 0 \\
&= 0.
\end{aligned}$$

De même, fixons $k \neq \ell$ dans $\mathbb{N}^*$. On utilise la formule :

$$\sin a \cdot \sin b = \frac{\cos(a-b) - \cos(a+b)}{2}.$$

Ainsi,

$$\begin{aligned}
(s_k \mid s_\ell) &= \frac{1}{2\pi} \int_0^{2\pi} \sin(kt)\ \sin(\ell t)\ dt \\
&= \frac{1}{2\pi} \int_0^{2\pi} \frac{\cos((k-\ell)t) - \cos((k+\ell)t)}{2}\ dt \\
&= \frac{1}{4\pi} \left[ \frac{\sin((k-\ell)t)}{k-\ell} - \frac{\sin((k+\ell)t)}{k+\ell} \right]_0^{2\pi}, \text{ car } k-\ell \neq 0 \text{ et } k+\ell \neq 0 \\
&= 0.
\end{aligned}$$

Enfin, fixons deux entiers $k \in \mathbb{N}$ et $\ell$ dans $\mathbb{N}^*$ pas forcément différents.
On utilise la formule :

$$\cos a \cdot \sin b = \frac{\sin(a+b) - \sin(a-b)}{2}.$$

Ainsi,

$$\begin{aligned}
(c_k \mid s_\ell) &= \frac{1}{2\pi} \int_0^{2\pi} \cos(kt)\ \sin(\ell t)\ dt \\
&= \frac{1}{2\pi} \int_0^{2\pi} \frac{\sin((k+\ell)t) - \sin((k-\ell)t)}{2}\ dt \\
&= \frac{1}{4\pi} \left[ \frac{\cos((k-\ell)t)}{k-\ell} - \frac{\cos((k+\ell)t)}{k+\ell} \right]_0^{2\pi}, \text{ si } k-\ell \neq 0 \text{ et car } k+\ell \neq 0 \\
&= 0.
\end{aligned}$$

Lorsque $k = \ell$, le calcul donne la même chose car $\sin((k-\ell)t) = 0$.

On découvre que la famille $\mathscr{F}$ est une famille orthogonale à vecteurs non nuls : cette famille est libre.
Pour information, en utilisant $\cos^2\theta = \dfrac{1+\cos(2\theta)}{2}$ et $\sin^2\theta = \dfrac{1-\cos(2\theta)}{2}$, on obtient tous calculs faits :

$$\|c_0\| = 1 \text{ et pour tout } k \in \mathbb{N}^*,\ \|c_k\| = \|s_k\| = \frac{1}{\sqrt{2}}.$$

Le procédé d'orthonormalisation de Gram-Schmidt qui est détaillé ensuite est très important dans ce chapitre. Examinons cela de plus près.
Donnons-nous une famille libre finie $(u_1, \cdots, u_p)$ dans un espace préhilbertien $E$. Le procédé d'orthonormalisation va permettre de construire explicitement une famille orthonormale $(e_1, \cdots, e_p)$ avec certaines conditions, à savoir :

- pour tout $k$ entre 1 et $p$, les espaces vectoriels engendrés par $u_1, \cdots, u_k$ d'une part et $e_1, \cdots, e_k$ d'autre part sont les mêmes
- pour tout entier $k$ entre 1 et $p$, le produit scalaire $(e_k \mid u_k)$ est strictement positif.

Cette dernière condition est là pour garantir l'unicité.
La construction explicite figure dans la méthode. C'est un procédé de construction par récurrence.
Détaillons cette construction.
Le vecteur $u_1$ est non nul, car il fait partie d'une famille libre. En divisant le vecteur $u_1$ par sa norme strictement positive, alors le vecteur $e_1$ ainsi obtenu est assurément unitaire et colinéaire de même sens que $u_1$ : $\mathrm{Vect}(e_1) = \mathrm{Vect}(u_1)$ et $(e_1 \mid u_1) = \|u_1\| > 0$.
Supposons les vecteurs $e_1, \cdots, e_k$ construits et répondant aux conditions requises.
On pose le vecteur :

$$v_{k+1} = u_{k+1} - \sum_{j=1}^{k} (u_{k+1} \mid e_j) \cdot e_j.$$

D'une part, le vecteur $\displaystyle\sum_{j=1}^{k} (u_{k+1} \mid e_j) \cdot e_j$ appartient à $\mathrm{Vect}(e_1, \cdots, e_k) = \mathrm{Vect}(u_1, \cdots, u_k)$ et comme la famille $(u_1, \cdots, u_p)$ est libre, il est impossible que le vecteur $u_{k+1}$ appartienne à l'espace $\mathrm{Vect}(u_1, \cdots, u_k)$ : le vecteur $v_{k+1}$ n'est pas le vecteur nul.
Ensuite, le vecteur $v_{k+1}$ appartient clairement à $\mathrm{Vect}(u_1, \cdots, u_k, u_{k+1})$.
Si $i$ est un entier entre 1 et $k$, alors :

$$\begin{aligned}
(v_{k+1} \mid e_i) &= \left( u_{k+1} - \sum_{j=1}^{k} (u_{k+1} \mid e_j) \cdot e_j \mid e_i \right) \\
&= (u_{k+1} \mid e_i) - \sum_{j=1}^{k} (u_{k+1} \mid e_j) \cdot (e_j \mid e_i) \\
&= (u_{k+1} \mid e_i) - \sum_{j=1}^{k} (u_{k+1} \mid e_j) \cdot \delta_{i,j} \\
&= (u_{k+1} \mid e_i) - (u_{k+1} \mid e_i) = 0.
\end{aligned}$$

Le vecteur $v_{k+1}$ est orthogonal aux vecteurs $e_1, \cdots, e_k$.
Le vecteur non nul $v_{k+1}$ n'est pas forcément unitaire. En le divisant par sa norme strictement positive, on obtient le vecteur $e_{k+1}$ qui est unitaire. Le vecteur $e_{k+1}$ reste orthogonal aux vecteurs $e_1, \cdots, e_k$, donc la famille $(e_1, \cdots, e_k, e_{k+1})$ est bien une famille orthonormale.
Par ailleurs, on peut écrire :

$$\mathrm{Vect}(e_1, \cdots, e_k, e_{k+1}) = \mathrm{Vect}(e_1, \cdots, e_k, v_{k+1}) \subset \mathrm{Vect}(u_1, \cdots, u_{k+1}),$$

et par l'égalité des dimensions finies (égales à $k+1$), on a égalité des espaces.
Enfin, on teste la stricte positivité du produit scalaire $(e_{k+1} \mid u_{k+1})$ :

$$\begin{aligned}
(e_{k+1} \mid u_{k+1}) &= \left(e_{k+1} \mid v_{k+1} + \sum_{j=1}^{k} (u_{k+1} \mid e_j) \cdot e_j\right) \\
&= (e_{k+1} \mid v_{k+1}) + \sum_{j=1}^{k} (u_{k+1} \mid e_j) \cdot (e_{k+1} \mid e_j) \\
&= (e_{k+1} \mid v_{k+1}) + \sum_{j=1}^{k} (u_{k+1} \mid e_j) \cdot 0_{\mathbb{R}} \\
&= (e_{k+1} \mid v_{k+1}) \\
&= (e_{k+1} \mid \|v_{k+1}\| \cdot e_{k+1}) \\
&= \|v_{k+1}\| > 0.
\end{aligned}$$

Le procédé de construction utilise récursivement les vecteurs $e_i$ déjà calculés précédemment.
Il reste à terminer la démonstration du théorème 1 en prouvant l'unicité d'une telle famille, la construction garantissant son existence.

Soit $(\varepsilon_1, \cdots, \varepsilon_p)$ une famille orthonormée telle que pour tout entier $k$ entre 1 et $p$, alors :

$$\mathrm{Vect}(\varepsilon_1, \cdots, \varepsilon_k) = \mathrm{Vect}(u_1, \cdots, u_k) \text{ et } (\varepsilon_k \mid u_k) > 0.$$

On va montrer par récurrence forte sur l'entier $k$ que :

$$\varepsilon_k = e_k,$$

où les vecteurs $e_k$ sont deux qui ont été construits par la méthode.
• Lorsque $k = 1$, comme $\mathrm{Vect}(\varepsilon_1) = \mathrm{Vect}(u_1)$ et $(\varepsilon_1 \mid u_1) > 0$, alors le vecteur $\varepsilon_1$ est colinéaire de même sens que le vecteur $u_1$. De même, le vecteur $e_1$ est colinéaire de même sens que le vecteur $u_1$.
Les deux vecteurs $\varepsilon_1$ et $e_1$ sont colinéaires de même sens et comme ils sont unitaires, alors ils sont égaux.
• Supposons que pour tout $i$ entre 1 et $k$, on ait :

$$\varepsilon_i = e_i.$$

Alors, comme $\varepsilon_{k+1}$ appartient à $\mathrm{Vect}(\varepsilon_1, \cdots, \varepsilon_{k+1}) = \mathrm{Vect}(u_1, \cdots, u_{k+1}) = \mathrm{Vect}(u_1, \cdots, u_k) + \mathrm{Vect}(u_{k+1}) = \mathrm{Vect}(\varepsilon_1, \cdots, \varepsilon_k) + \mathrm{Vect}(u_{k+1})$, on peut écrire :

$$\varepsilon_{k+1} = \alpha \cdot u_{k+1} + \sum_{j=1}^{k} \beta_j \cdot \varepsilon_j.$$

On effectue le produit scalaire avec $\varepsilon_i$, où $i$ est un entier entre 1 et $k$, ce qui donne :

$$\begin{aligned} 0 &= (\varepsilon_{k+1} \mid \varepsilon_i) \\ &= \left( \alpha \cdot u_{k+1} + \sum_{j=1}^{k} \beta_j \cdot \varepsilon_j \mid \varepsilon_i \right) \\ &= \alpha \cdot (u_{k+1} \mid \varepsilon_i) + \beta_i, \end{aligned}$$

donc :

$$\beta_i = -\alpha \cdot (u_{k+1} \mid \varepsilon_i) = -\alpha \cdot (u_{k+1} \mid e_i),$$

et finalement :

$$\varepsilon_{k+1} = \alpha \cdot \left( u_{k+1} - \sum_{j=1}^{k} (u_{k+1} \mid e_j) \cdot e_j \right) = \alpha \cdot v_{k+1} = (\alpha \cdot \|v_{k+1}\|) \cdot e_{k+1}.$$

Les vecteurs $\varepsilon_{k+1}$ et $e_{k+1}$ sont colinéaires, de même norme, donc :

$$\varepsilon_{k+1} = \pm\, e_{k+1}.$$

Il reste à montrer que le scalaire $\alpha$ est positif, ce qui montrera que $\varepsilon_{k+1} = e_{k+1}$ et achèvera l'unicité par récurrence forte.
Or, on sait que $(\varepsilon_{k+1} \mid u_{k+1}) > 0$. Ce produit scalaire vaut :

$$\begin{aligned} (\varepsilon_{k+1} \mid u_{k+1}) &= \left( \varepsilon_{k+1} \mid \frac{1}{\alpha} \cdot \varepsilon_{k+1} + \sum_{j=1}^{k} (u_{k+1} \mid \varepsilon_i) \cdot \varepsilon_i \right) \\ &= \frac{1}{\alpha}. \end{aligned}$$

Le coefficient de colinéarité $\alpha \cdot \|v_{k+1}\|$ entre les vecteurs unitaires $\varepsilon_{k+1}$ et $e_{k+1}$ est positif : ce coefficient de colinéarité vaut 1 et $\varepsilon_{k+1} = e_{k+1}$.

Voici ce qu'il faut retenir sur ce procédé d'orthonormalisation :
- pour les exercices plutôt théoriques : l'énoncé du théorème 1
- pour les exercices plutôt pratiques : le procédé de construction de la méthode. Ce qu'il faut savoir sur cette méthode est que les calculs pratiques sont en général assez pénibles, avec inexorablement des racines carrées désagréables qui bourgeonnent dans les calculs. De plus, comme le procédé de construction est récursif, dès qu'une erreur de calcul apparaît, elle se propage irrémédiablement sur les vecteurs suivants.

Bien qu'il faille avoir en mémoire ces deux points, le premier point est souvent plus utile. On aura l'occasion de l'expérimenter à travers les exercices...

On va en déduire des résultats très importants, via le corollaire 2.
Démontrons l'existence de bases orthonormales dans n'importe quel espace euclidien.
Soit $E$ un espace euclidien, c'est-à-dire un $\mathbb{R}$-espace vectoriel de dimension finie $p$, cet espace étant muni d'un produit scalaire.

On sait que l'espace $E$ admet au moins une base $\mathscr{B} = (u_1, \cdots, u_p)$. La famille $\mathscr{B}$ est une famille libre. On peut orthonormaliser cette famille libre en une famille orthonormale $(e_1, \cdots, e_p) = \mathscr{C}$ de l'espace $E$.
Or, cette famille orthonormale $\mathscr{C}$ comporte $p = \dim(E)$ vecteurs : il s'agit d'une base orthonormale de $E$ (en abrégé : b.o.n.).

Pour le second point du corollaire 2, considérons une famille orthonormale $\mathscr{F} = (e_1, \cdots, e_r)$ dans l'espace $E$. La famille $\mathscr{F}$ est une famille libre dans $E$. Par le théorème de la base incomplète, on peut compléter cette famille $\mathscr{F}$ en une base :

$$\mathscr{B} = (\mathscr{F}, \chi_1, \cdots, \chi_s)$$

avec $r + s = p$.
La famille $\mathscr{B}$ est une base. On peut l'orthonormaliser par le procédé de Gram-Schmidt. Cela donne une famille orthonormale :

$$\mathscr{D} = (\varepsilon_1, \cdots, \varepsilon_r, \theta_1, \cdots, \theta_s).$$

On va montrer que pour tout $k \in [\![1, r]\!]$,

$$\varepsilon_k = e_k.$$

Considérons la famille :

$$\mathscr{E} = (e_1, \cdots, e_r, \theta_1, \cdots, \theta_s).$$

Soyez attentifs à ce qui va suivre. On va montrer que les familles $\mathscr{D}$ et $\mathscr{E}$ sont égales en exploitant l'unicité de la famille orthonormalisée. Autrement dit, on va montrer que la famille $\mathscr{E}$ est la famille obtenue à partir de la famille $\mathscr{B}$ après l'avoir orthonormalisée par le procédé de Gram-Schmidt.

$\triangleright$ Premièrement, la famille $\mathscr{E}$ est orthonormale.
En effet si $i$ et $j$ sont deux indices entre 1 et $r$, alors

$$(e_i \mid e_j) = \delta_{i,j}.$$

Si $i$ et $j$ sont deux indices entre 1 et $s$, alors :

$$(\theta_i \mid \theta_j) = \delta_{i,j}.$$

Enfin, si $i \in [\![1, r]\!]$ et $j \in [\![1, s]\!]$, alors par construction de la famille $\mathscr{D}$, on sait que :

$$\mathrm{Vect}(\varepsilon_1, \cdots, \varepsilon_r) = \mathrm{Vect}(\mathscr{F}) = \mathrm{Vect}(e_1, \cdots, e_r).$$

On peut écrire :

$$e_i = \sum_{k=1}^{r} \lambda_k \cdot \varepsilon_k.$$

Par conséquent,

$$\begin{aligned}
(e_i \mid \theta_j) &= \left(\sum_{k=1}^{r} \lambda_k \cdot \varepsilon_k \mid \theta_j\right) \\
&= \sum_{k=1}^{r} \lambda_k \cdot (\varepsilon_k \mid \theta_j) \\
&= \sum_{k=1}^{r} \lambda_k \cdot 0_{\mathbb{R}} \\
&= 0_{\mathbb{R}}.
\end{aligned}$$

$\triangleright$ Deuxièmement on montre que pour tout $k \in [\![1, r]\!]$,

$$\mathrm{Vect}(e_1, \cdots, e_k) = \mathrm{Vect}(\varepsilon_1, \cdots, \varepsilon_k).$$

En effet, cela provient directement du fait que la famille $\mathscr{D}$ est la famille orthonormalisée issue de $\mathscr{B}$.
Ensuite, on montre que si $\ell \in [\![1, s]\!]$,

$$\mathrm{Vect}(e_1, \cdots, e_k, \chi_1, \cdots, \chi_\ell) = \mathrm{Vect}(\varepsilon_1, \cdots, \varepsilon_k, \theta_1, \cdots, \theta_\ell).$$

En effet,

$$\begin{aligned}
\mathrm{Vect}(e_1, \cdots, e_r, \theta_1, \cdots, \theta_\ell) &= \mathrm{Vect}(e_1, \cdots, e_r) + \mathrm{Vect}(\theta_1, \cdots, \theta_\ell) \\
&= \mathrm{Vect}(\varepsilon_1, \cdots, \varepsilon_r) + \mathrm{Vect}(\theta_1, \cdots, \theta_\ell), \text{ par construction de } \mathscr{D} \\
&= \mathrm{Vect}(\varepsilon_1, \cdots, \varepsilon_r, \theta_1, \cdots, \theta_\ell) \\
&= \mathrm{Vect}(e_1, \cdots, e_r, \chi_1, \cdots, \chi_\ell), \text{ par construction de } \mathscr{D}
\end{aligned}$$

$\triangleright$ Il reste à montrer que pour tout $i \in [\![1, r]\!]$, $(e_i \mid e_i) > 0$, ce qui est vrai car $(e_i \mid e_i) = 1 > 0$ et que pour tout $j \in [\![1, s]\!]$, $(\theta_j \mid \chi_j) > 0$, ce qui est le cas par construction de la famille orthonormalisée $\mathscr{D}$.

Résultat des courses, la famille $\mathscr{E}$ est exactement la famille orthonormalisée par Gram-Schmidt construite à partir de la famille libre $\mathscr{B}$. Cette famille est égale à $\mathscr{D}$.
Conclusion, la famille $\mathscr{E}$ est une base orthonormale qui complète la famille orthonormale $\mathscr{F}$ de départ.

On verra un peu plus tard l'importance des bases orthonormales...

Traitons l'exemple 5, qui donne deux applications de ce procédé d'orthonormalisation.

- Le premier point est du calcul. On applique la méthode :
  - $\longrightarrow$ on pose $\mathscr{F} = (u_1, u_2, u_3)$ la famille proposée ; on notera $\mathscr{B} = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ la famille orthonormale obtenue après orthonormalisation ;

$\longrightarrow$ le premier vecteur vaut :

$$\varepsilon_1 = \frac{u_1}{\|u_1\|} = \frac{1}{\sqrt{2}} \cdot (1, 1, 0).$$

$\longrightarrow$ pour le deuxième vecteur, on pose :

$$\begin{aligned} v_2 = u_2 - (u_2 \mid \varepsilon_1) \cdot \varepsilon_1 &= u_2 - \frac{1}{2} \cdot (u_2 \mid u_1) \cdot u_1 \\ &= u_2 - \frac{1}{2} \cdot u_1 \\ &= \frac{1}{2} \cdot (1, -1, 2). \end{aligned}$$

Le deuxième vecteur $\varepsilon_2$ vaut donc :

$$\varepsilon_2 = \frac{v_2}{\|v_2\|} = \frac{(1, -1, 2)}{\|(1, -1, 2)\|} = \frac{1}{\sqrt{6}} \cdot (1, -1, 2).$$

$\longrightarrow$ pour le troisième vecteur, on pose :

$$\begin{aligned} v_3 &= u_3 - (u_3 \mid \varepsilon_1) \cdot \varepsilon_1 - (u_3 \mid \varepsilon_2) \cdot \varepsilon_2 \\ &= u_3 - \frac{1}{2} (u_3 \mid u_1) \cdot u_1 - \frac{1}{6} (u_3 \mid (1, -1, 2)) \cdot (1, -1, 2) \\ &= (0, 1, 1) - \frac{1}{2} \cdot (1, 1, 0) - \frac{1}{6} \cdot (1, -1, 2) \\ &= \frac{1}{6} \cdot (-4, 4, 4) = \frac{2}{3} \cdot (-1, 1, 1). \end{aligned}$$

On obtient ainsi le vecteur $\varepsilon_3$ :

$$\varepsilon_3 = \frac{v_3}{\|v_3\|} = \frac{(-1, 1, 1)}{\|(-1, 1, 1)\|} = \frac{1}{\sqrt{3}} \cdot (-1, 1, 1).$$

Ouvrons une parenthèse hors programme en Mathématiques et parlons du produit vectoriel. Si $u = (a, b, c)$ et $v = (\alpha, \beta, \gamma)$ sont deux vecteurs dans $\mathbb{R}^3$, on définit le produit vectoriel :

$$w = u \wedge v$$

comme le vecteur :

$$w = \Big(b\ \gamma - c\ \beta, c\ \alpha - a\ \gamma, a\ \beta - b\ \alpha\Big).$$

Ce vecteur est géométriquement défini de la façon suivante : si $u$ et $v$ sont colinéaires, alors le produit vectoriel $w$ est nul. Si la famille $(u, v)$ est libre, alors le produit vectoriel est non nul. Le produit vectoriel $w$ est le seul vecteur à la fois orthogonal aux vecteurs $u$ et $v$ de telle sorte que la base $(u, v, w)$ soit directe, autrement dit de telle sorte que le déterminant $\det_{\mathscr{B}_c}(u, v, w)$ selon la base canonique soit strictement positif et pour finir tel que la norme du vecteur $w$ soit égale à :

$$\|w\| = \|u\| \cdot \|v\| \cdot |\sin\theta|,$$

où $\theta$ est l'angle géométrique entre les vecteurs $u$ et $v$ dans le plan vectoriel $\mathrm{Vect}(u, v)$.
Comme vous manipulez beaucoup de produit vectoriel en SI et bientôt en Physique électromagnétique, il n'est pas très coûteux d'avoir ces notions en tête.
Voyons comment ce produit vectoriel peut alléger les calculs du vecteur $\varepsilon_3$ plus haut.
On repart du calcul de $\varepsilon_3$.
La famille $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ doit être une base orthonormale. Par conséquent,

$$\varepsilon_3 = \pm\, \varepsilon_1 \wedge \varepsilon_2.$$

Or, on a calculé les vecteurs $\varepsilon_1$ et $\varepsilon_2$, ce qui donne :

$$\varepsilon_1 \wedge \varepsilon_2 = \frac{1}{\sqrt{12}} \cdot (2, -2, -2) = \frac{1}{\sqrt{3}}\,(1, -1, -1).$$

Pour savoir si $\varepsilon_3$ vaut ce vecteur ou son opposé, il suffit de calculer le produit scalaire entre le vecteur $u_3$ et le vecteur $\varepsilon_1 \wedge \varepsilon_2$. Comme on doit avoir $(u_3 \mid \varepsilon_3)$ strictement positif, cela nous indiquera le signe.
Or,

$$\left(u_3 \mid \frac{1}{\sqrt{3}}(1, -1, -1)\right) = \frac{1}{\sqrt{3}}\,\Big((0, 1, 1) \mid (1, -1, -1)\Big) = -\frac{2}{\sqrt{3}} < 0.$$

Conclusion, il faut retenir le signe moins et :

$$\varepsilon_3 = -\, \varepsilon_1 \wedge \varepsilon_2 = \frac{1}{\sqrt{3}}\,(-1, 1, 1).$$

On retrouve le même résultat qu'avant avec un peu plus d'efficacité.

• Traitons maintenant le deuxième point de l'exemple 5. Ce point concerne les polynômes orthogonaux et il est bien de retenir les idées qui vont suivre.

1. La fonction $\Phi$ définie dans cette question est clairement symétrique et linéaire par rapport à la première variable, donc bilinéaire.
   Si $P(X)$ est un polynôme dans $\mathbb{R}[X]$, alors :

   $$\Phi(P, P) = \int_a^b P^2 \times \omega.$$

   Par hypothèse sur l'application $\omega$, l'intégrande $P^2 \times \omega$ est positive sur $[a, b]$, donc l'intégrale est positive.
   En outre, si $P(X)$ est un polynôme dans $\mathbb{R}[X]$ tel que $\Phi(P, P) = 0$, alors on peut appliquer le théorème aux quatre hypothèses : $a < b$, l'intégrale $t \longmapsto (P(t))^2 \cdot \omega(t)$ est continue, positive et d'intégrale nulle sur $[a, b]$. L'intégrande est nulle, donc pour tout $t \in [a, b]$,

   $$(P(t))^2 \cdot \omega(t) = 0.$$

   La fonction $\omega$ n'est pas la fonction nulle. Il existe $t_0 \in [a, b]$ tel que $\omega(t_0) \neq 0$. Par continuité de la fonction $\omega$, en prenant $\varepsilon = \dfrac{|\omega(t_0)|}{2} > 0$, il existe $\eta > 0$ tel que :

   $$\forall t \in [a, b],\ |t - t_0| \leqslant \eta \Longrightarrow \Big|\omega(t) - \omega(t_0)\Big| \leqslant \varepsilon, \text{ donc } |\omega(t)| \geqslant |\omega(t_0)| - \varepsilon = \frac{|\omega(t_0)|}{2} > 0.$$

L'intervalle $I = [a,b] \cap [t_0 - \eta, t_0 + \eta]$ est de longueur strictement positive, donc est infini. Sur l'intervalle $I$, on peut écrire :

$$\forall t \in I,\ \omega(t) \neq 0, \text{ donc } (P(t))^2 = 0, \text{ puis } P(t) = 0.$$

On découvre que le polynôme $P(X)$ admet au moins tous les éléments de l'ensemble $I$ comme racines, ce qui en fait une infinité : le polynôme $P(X)$ est le polynôme.

On retrouve les deux arguments classiques du théorème aux quatre hypothèses et le fait que seul le polynôme nul admet une infinité de racines.

2. Fixons un entier $N \in \mathbb{N}$. Dans l'espace euclidien $\mathbb{R}_N[X]$ muni du produit scalaire de l'énoncé, on considère la base canonique $\mathscr{F} = (1, X, \cdots, X^N)$.
On orthonormalise cette famille libre $\mathscr{F}$ par le procédé d'orthonormalisation de Gram-Schmidt, ce qui donne une famille orthonormée $\mathscr{B} = (P_0, \cdots, P_N)$ de $\mathbb{R}_N[X]$.
On sait que pour tout $k \in [\![0, N]\!]$ :

$$\mathrm{Vect}(1, X, \cdots, X^k) = \mathrm{Vect}(P_0, P_1, \cdots, P_k).$$

On va montrer que pour tout entier $k$ entre 0 et $N$,

$$\deg(P_k) = k.$$

▷ Si $k = 0$, on sait que $\mathrm{Vect}(1) = \mathbb{R}_0[X] = \mathrm{Vect}(P_0)$, ce qui impose au polynôme $P_0(X)$ d'être constant non nul : $\deg(P_0) = 0$.
▷ Supposons que pour tout entier $j$ entre 0 et $k$, alors :

$$\deg(P_j) = j.$$

▷ On sait alors que :

$$\begin{array}{rcl} \mathbb{R}_{k+1}[X] & = & \mathrm{Vect}(1, \cdots, X^k, X^{k+1}) \\ & = & \mathrm{Vect}(P_0, \cdots, P_k, P_{k+1}) \\ & = & \mathrm{Vect}(P_0, \cdots, P_k) + \mathrm{Vect}(P_{k+1}) \\ & = & \mathrm{Vect}(1, \cdots, X^k) + \mathrm{Vect}(P_{k+1}) \\ & = & \mathbb{R}_k[X] + \mathrm{Vect}(P_{k+1}). \end{array}$$

Par conséquent, le vecteur $P_{k+1}(X)$ doit appartenir à $\mathbb{R}_{k+1}[X] \setminus \mathbb{R}_k[X]$, ce qui signifie que le polynôme $P_{k+1}(X)$ est exactement de degré $k+1$.

Il reste à montrer que tout ce que l'on a fait ne dépend pas de l'entier $N$.
Si l'on considère la famille libre $(1, X, \cdots, X^{N+1})$ dans $\mathbb{R}_{N+1}[X]$ et qu'on l'orthonormalise selon l'orthonormalisation de Gram-Schmidt, on obtient une famille orthonormale $(Q_0, \cdots, Q_N, Q_{N+1})$ de $\mathbb{R}_{N+1}[X]$.
La construction des vecteurs $Q_0, \cdots, Q_N$ est identique à celle des vecteurs $P_0, \cdots, P_N$, dont la construction ne dépend que des vecteurs initiaux $1, \cdots, X^N$ : pour tout entier $k$ entre 0 et $N$,

$$P_k(X) = Q_k(X).$$

On en déduit que l'on peut construire par récurrence une suite $(P_n)_{n\in\mathbb{N}}$ de polynômes en rajoutant à la construction de la famille finie déjà construite $(P_0, \cdots, P_N)$ orthonormale, un vecteur $P_{N+1}$ de telle sorte que la famille complétée $(P_0, \cdots, P_N, P_{N+1})$ reste une famille orthonormale.
On en déduit que la famille $(P_n)_{n\in\mathbb{N}}$ est une famille qui répond aux conditions requises : si $k$ est un entier naturel, alors le polynôme $P_k(X)$ qui fait partie de la famille $(P_0, \cdots, P_k)$ vérifie :

$$\deg(P_k) = k.$$

De plus, si $i \neq j$ sont deux entiers, on trouve un entier $N$ supérieur à la fois aux entiers $i$ et $j$. Les deux polynômes $P_i(X)$ et $P_j(X)$ font partie de la même famille orthonormale $(P_0, \cdots, P_N)$, ce qui autorise à écrire :

$$(P_i \mid P_j) = \delta_{i,j}.$$

En conclusion, la famille $(P_n)_{n\in\mathbb{N}}$ est à degrés échelonnés, donc constitue une famille libre.
De plus, si $Q \in \mathbb{R}[X]$, on trouve un entier $N$ tel que $N \geqslant \deg(Q)$. Par conséquent,

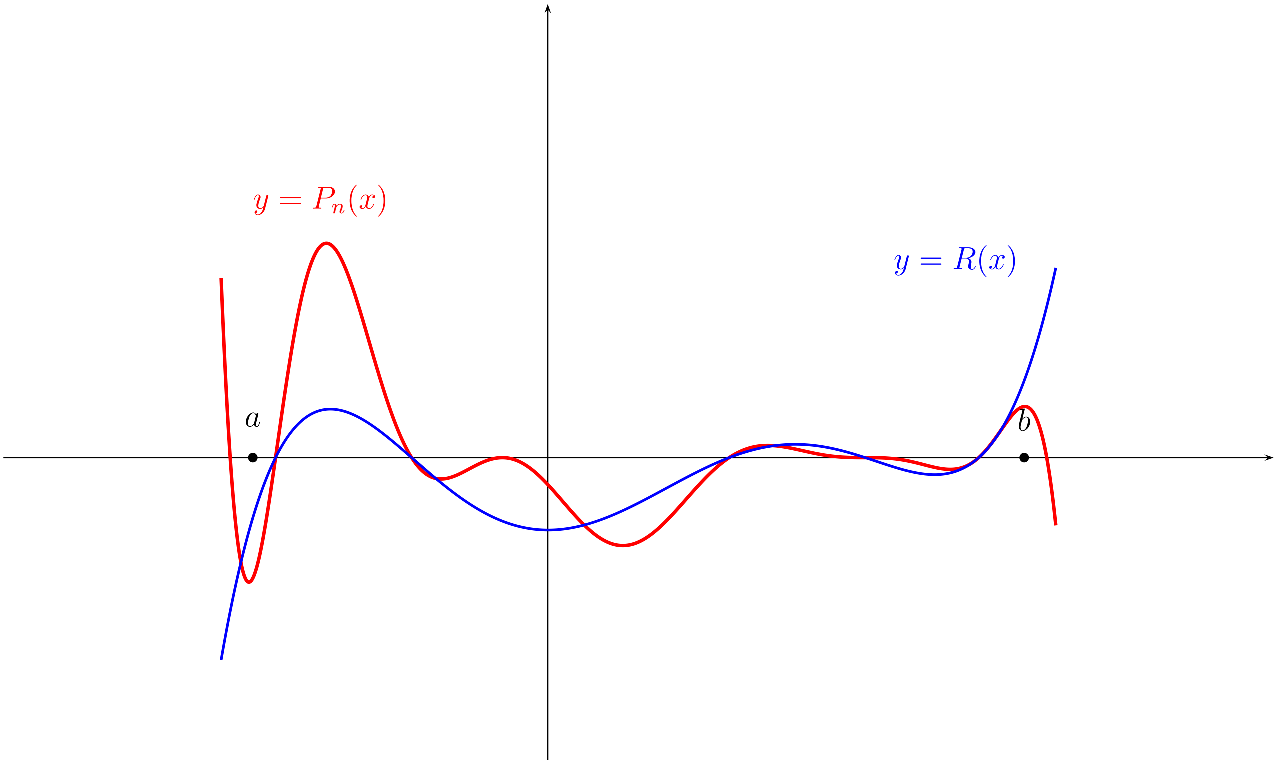
$$Q(X) \in \mathbb{R}_N[X] = \mathrm{Vect}(P_0, \cdots, P_N).$$

En définitive, le polynôme $Q(X)$ est une combinaison linéaire des vecteurs de la famille $(P_k)_{0\leqslant k\leqslant N}$, donc des vecteurs de la famille $(P_n)_{n\in\mathbb{N}}$.
**Remarque** : dans la rédaction de cette question, il y a sûrement un peu trop de détails... Cela ne fait pas de mal...

3. On a déjà vu le raisonnement qui va suivre.
Fixons un entier $n$ dans $\mathbb{N}$.
Si $n = 0$, alors $P_0$ est bien scindé à racines simples, avec aucune racine.
Dans la suite, on se place dans le cas où $n \geqslant 1$.
On pose la factorisation de $P_n(X)$ :

$$P_n(X) = \xi \prod_{k=1}^{r} (X - \lambda_k)^{\alpha_k} \times Q(X),$$

où $\lambda_1, \cdots, \lambda_r$ sont les racines réelles de $P_n(X)$, $\alpha_1, \cdots, \alpha_r$ sont leur multiplicité dans $\mathbb{N}^*$ et le polynôme $Q(X)$ produit de facteurs irréductibles de degré 2 n'admettant donc aucune racine réelle.
Faisons un dessin, pour y voir plus clair :




On a tracé en rouge la courbe $y = P_n(x)$. On ne va retenir que les racines $\lambda$ du polynôme $P_n(X)$ telles que $\lambda \in ]a, b[$ et la multiplicité de la racine $\lambda$ dans le polynôme $P_n(X)$ est impaire. Il s'agit en fait des abscisses où la fonction $t \longmapsto P_n(t)$ change de signe sur l'intervalle $]a, b[$. On note $\mathscr{A}$ l'ensemble de ces racines. Sur le schéma ci-dessus, le polynôme $P_n(X)$ admet des racines hors de $]a, b[$, que l'on ne prend pas en compte. Le polynôme $P_n(X)$ admet une racine de multiplicité paire : on a en cette racine une annulation mais la fonction garde le même signe localement en cette racine. On ne la prend pas en compte non plus.
On pose alors le polynôme :

$$R(X) = \prod_{\lambda \in \mathscr{A}} (X - \lambda).$$

Vous avez en bleu la courbe $y = R(x)$.
Lorsque l'on regarde les deux courbes uniquement sur l'intervalle $]a, b[$, alors les deux courbes changent de signe aux mêmes abscisses. Le produit $t \longmapsto P_n(t) \cdot R(t)$ garde donc un signe constant sur $]a, b[$, donc sur $[a, b]$, par continuité.
Par le théorème aux quatre hypothèses, comme la fonction $g : t \longmapsto P(t) \cdot R(t) \cdot \omega(t)$ n'est pas la fonction nulle – on rappelle que la fonction $\omega(\cdot)$ ne s'annule pas sur une partie infinie de $[a, b]$, donc il est impossible que sur cet ensemble infini, la fonction polynomiale non nulle $t \longmapsto P_n(t) \cdot R(t)$ soit tout le temps nulle, comme $a < b$, comme la fonction $g$ est continue et garde un signe constant, alors l'intégrale $\displaystyle\int_a^b g(t)\, dt$ n'est pas nulle.
On vient de montrer que le produit scalaire :

$$(P_n(X) \mid R(X))$$

est non nul.
Il est manifeste que le polynôme $R(X)$ divise le polynôme $P_n(X)$, donc le polynôme $R(X)$ est de degré inférieur à $\deg(P_n) = n$.

Imaginons un instant que les deux polynômes $R(X)$ et $P_n(X)$ soient différents. Dans ce cas, $\deg(R) < \deg(P_n)$, ou encore :

$$R(X) \in \mathbb{R}_{n-1}[X] = \mathrm{Vect}(P_0, \cdots, P_{n-1}).$$

On pose alors :

$$R(X) = \sum_{j=0}^{n-1} \alpha_j \cdot P_j(X).$$

On effectue le produit scalaire entre $R(X)$ et $P_n(X)$, ce qui donne par linéarité du produit scalaire par rapport à la première variable :

$$\begin{aligned}
(R(X) \mid P_n(X)) &= \left( \sum_{j=0}^{n-1} \alpha_j \cdot P_j(X) \mid P_n(X) \right) \\
&= \sum_{j=0}^{n-1} \alpha_j \cdot (P_j(X) \mid P_n(X)) \\
&= \sum_{j=0}^{n-1} \alpha_j \cdot 0_{\mathbb{R}}, \text{ car si } j < n,\ P_j(X) \perp P_n(X) \\
&= 0_{\mathbb{R}}.
\end{aligned}$$

On obtient alors une contradiction.
En definitive, les deux polynômes $P_n(X)$ et $R(X)$ sont égaux et on obtient alors tout immédiatement puisque le polynôme $R(X)$ est scindé à racines simples à racines dans $]a, b[$.

Nous allons voir dans cette section l'importance des bases orthonormales dans un espace euclidien.

La propriété 5 nous dit qu'il est très facile de calculer le produit scalaire entre deux vecteurs en fonction des coordonnées dans une base orthonormée, ainsi que la norme de tout vecteur. En effet, avec les notations qui sont proposées, on obtient :

$$\begin{aligned}
(x \mid y) &= \left( \sum_{i=1}^{n} x_i \cdot e_i \mid \sum_{j=1}^{n} y_j \cdot e_j \right) \\
&= \sum_{i,j=1}^{n} x_i \cdot y_j \cdot (e_i \mid e_j) \text{ [produit scalaire bilinéaire]} \\
&= \sum_{i,j=1}^{n} x_i \cdot y_j \cdot \delta_{i,j} \text{ [famille orthonormale]} \\
&= \sum_{i=1}^{n} x_i \cdot y_i.
\end{aligned}$$

En prenant $y = x$, alors :

$$\|x\|^2 = (x \mid x) = \sum_{i=1}^{n} x_i \cdot x_i = \sum_{i=1}^{n} x_i^2.$$

On peut voir à travers cette égalité une généralisation du théorème de Pythagore, ici en dimension $n$, dans un hypertriangle rectangle...

On définit maintenant l'orthogonal d'une partie $A$ non vide, où la partie $A$ n'est pas forcément un sous-espace vectoriel de $E$. L'ensemble $A^\perp$ est l'ensemble des vecteurs $x \in E$ qui sont orthogonaux à tout vecteur $a$ de $A$.

On démontre tranquillement la propriété 6.

• On a clairement $A^\perp \subset E$. Ensuite, le vecteur nul étant orthogonal à tout vecteur de $E$ est en particulier orthogonal à tout vecteur de $A$ :

$$0_E \in A^\perp.$$

Enfin, si $x$ et $y$ sont dans $A^\perp$ et si $\lambda$ est un scalaire réel, fixons un vecteur $a$ dans la partie $A$. On peut alors écrire :

$$\begin{array}{rcl}(\lambda \cdot x + y \mid a) & = & \lambda \cdot (x \mid a) + (y \mid a) \qquad \text{[linéarité première variable]} \\ & = & \lambda \cdot 0_{\mathbb{R}} + 0_{\mathbb{R}} \qquad \text{, car } x \perp a \text{ et } y \perp a \\ & = & 0_{\mathbb{R}}.\end{array}$$

Résultat des courses, le vecteur $\lambda \cdot x + y$ est orthogonal à tout vecteur $a$ de $A$ :

$$\lambda \cdot x + y \in A^\perp.$$

• Soit $F$ un sous-espace de $E$. On considère une base orthonormale de $F$. On peut rappeler qu'une telle base orthonormale existe. Pour ce faire, il suffit de considère une base $(u_1, \cdots, u_r)$ de $F$, puis de l'orthonormaliser par Gram-Schmidt, pour donner une famille orthonormale $(e_1, \cdots, e_r)$. On sait que :

$$F = \text{Vect}(u_1, \cdots, u_r) = \text{Vect}(e_1, \cdots, e_r).$$

La famille $(e_1, \cdots, e_r)$ est par conséquent une base orthonormale de $F$.
On peut compléter cette base orthonormale $(e_1, \cdots, e_r)$ en une base orthonormale

$$(e_1, \cdots, e_r, \varepsilon_1, \cdots, \varepsilon_s)$$

de l'espace euclidien $E$.
On pose :

$$G = \text{Vect}(\varepsilon_1, \cdots, \varepsilon_s).$$

On va montrer que $G = F^\perp$.
En effet, si $x \in G$, on pose :

$$x = \sum_{j=1}^{s} \alpha_j \cdot \varepsilon_j.$$

Soit $y$ un vecteur de $F$. On pose :

$$y = \sum_{i=1}^{r} \beta_i \cdot e_i.$$

On en déduit :

$$\begin{aligned}
(y \mid x) &= \left( \sum_{i=1}^{r} \beta_i \cdot e_i \mid \sum_{j=1}^{s} \alpha_j \cdot \varepsilon_j \right) \\
&= \sum_{i=1}^{r} \sum_{j=1}^{s} \beta_i \cdot \alpha_j \cdot (e_i \mid \varepsilon_j) \\
&= \sum_{i=1}^{r} \sum_{j=1}^{s} \beta_i \cdot \alpha_j \cdot 0_{\mathbb{R}} \\
&= 0_{\mathbb{R}}.
\end{aligned}$$

On vient de montrer que tout vecteur de $G$ est orthogonal à tout vecteur de $F$ :

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

Ensuite, comme $F + G = E$ et que $G \subset F^{\perp}$, alors :

$$F + F^{\perp} = E.$$

Soit $x$ dans l'intersection $F \cap F^{\perp}$. Alors, comme $x \in F^{\perp}$, alors $x$ est orthogonal à tout vecteur de $F$, en particulier le vecteur $x$ est orthogonal à lui-même car $x$ appartient aussi à $F$. Ainsi, $(x \mid x) = 0_{\mathbb{R}}$, donc $x = 0_E$.
Conclusion, on a bien $F \oplus F^{\perp} = E$. En prime, les espaces $F^{\perp}$ et $G$ sont deux supplémentaires de l'espace $F$ :

$$\dim(F^{\perp}) = \dim(E) - \dim(F) = \dim(G).$$

L'inclusion $G \subset F^{\perp}$ couplée à l'égalité des dimensions finies donne l'égalité :

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

Ensuite, si $x$ est un vecteur de $F$, alors si $y \in F^{\perp}$, on en déduit :

$$(y \mid x) = 0,$$

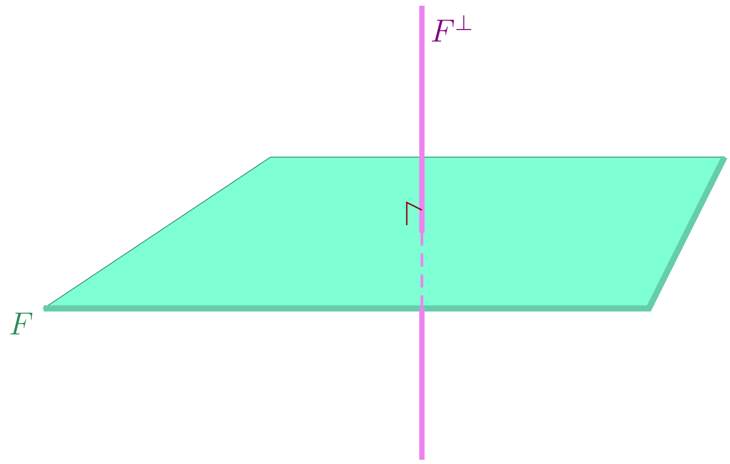
donc $(x \mid y) = 0$ et donc on apprend que le vecteur $x$ est orthogonal à tout vecteur de $F^{\perp}$ : le vecteur $x$ appartient bien à l'espace $(F^{\perp})^{\perp}$.
Les espaces $F$ et $(F^{\perp})^{\perp}$ sont deux supplémentaires de l'espace $F^{\perp}$, donc :

$$\dim(F) = \dim((F^{\perp})^{\perp}).$$

L'inclusion couplée à l'égalité des dimensions finies nous donne l'égalité :

$$F = (F^{\perp})^{\perp}.$$

Voici une représentation possible en dimension 3, lorsque $F$ est un plan vectoriel :




• Soit $\mathscr{B}_1 = (e_1, \cdots, e_r)$ une base orthonormale de $F$, soit $\mathscr{B}_2 = (\varepsilon_1, \cdots, \varepsilon_s)$ une base orthonormale de $F^{\perp}$.
Déjà, chaque vecteur $e_1, \cdots, e_r, \varepsilon_1, \cdots, \varepsilon_s$ est unitaire.
D'autre part, si $i \neq j$ sont deux entiers différents entre 1 et $r$, alors les vecteurs $e_i$ et $e_j$ sont orthogonaux.
Si $i$ et $j$ sont deux entiers différents entre 1 et $s$, les vecteurs $\varepsilon_i$ et $\varepsilon_j$ sont encore orthogonaux.
Il reste à tester si chaque vecteur $e_k$ et orthogonal à chaque vecteur $\varepsilon_\ell$. Si $k$ est un entier entre 1 et $r$ et si $\ell$ est un entier entre 1 et $s$, comme $\varepsilon_\ell$ appartient à $F^{\perp}$ et que $e_k$ appartient à $F$, alors ces deux vecteurs sont bien orthogonaux. La famille :

$$(\mathscr{B}_1, \mathscr{B}_2) = (e_1, \cdots, e_r, \varepsilon_1, \cdots, \varepsilon_s)$$

est bien une famille orthogonale à vecteurs unitaires, donc une famille orthonormale.

À retenir : lorsque l'on recolle deux bases orthonormales de deux sous-espaces orthogonaux $F$ et $F^{\perp}$, on obtient une base orthonormale de l'espace $E$ tout entier.

• Considérons maintenant une base orthonormale $(e_1, \cdots, e_n)$ de $E$ que l'on découpe en deux familles :

$$\mathscr{C}_1 = (e_1, \cdots, e_r) \text{ et } \mathscr{C}_2 = (e_{r+1}, \cdots, e_n).$$

Il est clair que ces deux sous-familles constituent des familles orthonormales.
On pose $F = \mathrm{Vect}(\mathscr{C}_1)$ et $G = \mathrm{Vect}(\mathscr{C}_2)$.
Si $x$ est dans $F$ et si $y$ est dans $G$, on peut poser :

$$x = \sum_{i=1}^{r} a_i \cdot e_i \text{ et } y = \sum_{j=r+1}^{n} b_j \cdot e_j,$$

de sorte que :

$$(x \mid y) = \sum_{i=1}^{r} \sum_{j=r+1}^{n} a_i \cdot b_j \cdot (e_i \mid e_j) = 0_{\mathbb{R}}.$$

Ainsi, tout vecteur de $F$ est orthogonal à tout vecteur de $G$, ce qui se réécrit au choix :

$$F \subset G^{\perp} \text{ ou encore } G \subset F^{\perp}.$$

L'espace $G$ admet deux supplémentaires qui sont $F$ et $G^{\perp}$. On sait que $F$ et $G^{\perp}$ sont de même dimension finie et on a une inclusion, donc on a égalité.
On peut reprendre le raisonnement précédent en intervertissant les rôles de $F$ et de $G$, ou alors on peut prendre l'orthogonal dans l'égalité :

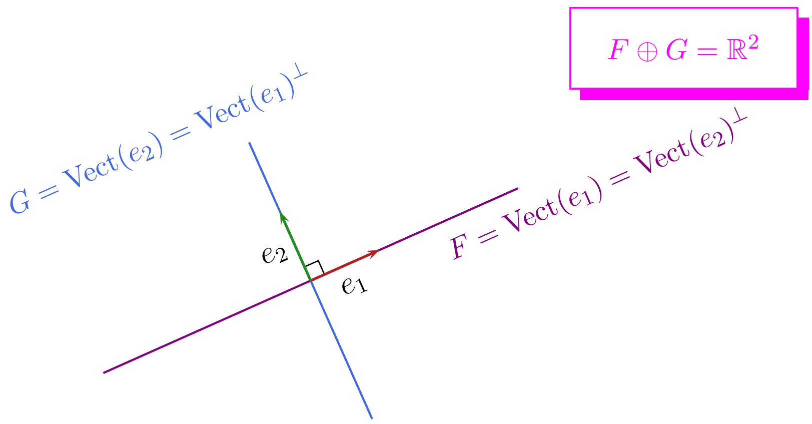
$$F = G^{\perp},$$

ce qui donne :

$$F^{\perp} = (G^{\perp})^{\perp} = G.$$

À retenir : lorsque l'on découpe une base orthonormale d'un espace euclidien $E$ en deux morceaux, les deux espaces engendrés par chaque morceau consituent deux espaces orthogonaux.

Voici par exemple une représentation en dimension deux où l'on prend une base $(e_1, e_2)$ orthonormale du plan $\mathbb{R}^2$ que l'on découpe en deux :




Passons à l'exemple 6.
• Si $E$ est un espace muni d'un produit scalaire, sans hypothèse sur la dimension finie, il y a des choses vues précédemment qui ne sont plus vraies.
Si $F$ est un sous-espace de $E$, pour tout $x \in F$ et pour tout $y \in F^{\perp}$, alors les vecteurs $x$ et $y$ sont orthogonaux, ce qui signifie que le vecteur $x$ est orthogonal à tout vecteur $y$ dans $F^{\perp}$ :

$$F \subset (F^{\perp})^{\perp}.$$

On sait que lorsque $E$ est de dimension finie, comme les espaces $F$ et $(F^{\perp})^{\perp}$ sont deux supplémentaires de l'espace $F^{\perp}$, alors :

$$\dim(F) = \dim(E) - \dim(F^{\perp}) = \dim((F^{\perp})^{\perp}).$$

L'inclusion vue précédemment se transforme en égalité.

Que se passe-t-il en dimension infinie ? Une chose est sûre : on ne peut pas utiliser l'argument précédent. On va voir que l'on peut avoir $F \neq (F^{\perp})^{\perp}$.

Prenons l'espace $E = C^0([0,1], \mathbb{R})$ des fonctions continue $f : [0,1] \longrightarrow \mathbb{R}$, muni du produit scalaire :

$$(f,g) \longmapsto \int_0^1 f \times g.$$

Considérons l'espace :

$$F = \Big\{ f \in E \mid f(0) = 0 \Big\}.$$

On remarque que $F$ est un hyperplan car c'est le noyau de la forme linéaire non nulle :

$$\varphi : \left| \begin{array}{ccc} E & \longrightarrow & \mathbb{R} \\ f & \longmapsto & f(0) \end{array} \right. .$$

On va montrer que $F^\perp = \{0\}$.
En effet, soit $f$ dans $F^\perp$. La fonction $g : t \longmapsto t \cdot f(t)$ appartient à l'espace $F$, donc les vecteurs $f$ et $g$ sont orthogonaux. Par conséquent,

$$\int_0^1 f(t)\ g(t)\ dt = 0, \text{ donc } \int_0^1 t \cdot f^2(t)\ dt = 0.$$

On peut utiliser le théorème aux quatre hypothèses, ce qui implique que l'intégrande $t \longmapsto t \cdot f^2(t)$ est nulle. Pour tout $t \in ]0,1]$, on sait alors que $f(t) = 0$ et la fonction $f$ est pour l'instant nulle sur $]0,1]$.
Or, $f(0) = \lim\limits_{t \longrightarrow 0^+} f(t) = \lim\limits_{t \longrightarrow 0^+} 0 = 0$ : la fonction $f$ est nulle sur $[0,1]$.
Résultat des courses, $F^\perp = \{0\}$ et $(F^\perp)^\perp = \{0\}^\perp = E$, car tout vecteur de $E$ est orthogonal au vecteur nul.
On a donc un contre-exemple où :

$$F \neq E = (F^\perp)^\perp.$$

On termine par l'exemple 7 connu sous le théorème de Riesz.
Soit $E$ un espace euclidien de dimension $n$.
On considère l'application :

$$\Phi : \left| \begin{array}{ccc} E & \longrightarrow & \mathscr{L}(E, \mathbb{R}) \\ \vec{u} & \longmapsto & \Big( x \longmapsto (x \mid \vec{u}) \Big) \end{array} \right. .$$

Cette application est bien définie. En effet, si $\vec{u}$ est dans $E$, comme le produit scalaire est linéaire par rapport à la première variable, alors l'application $x \longmapsto (x \mid \vec{u})$ est bien une forme linéaire de $E$ vers $\mathbb{R}$.
Ensuite, l'application $\Phi$ est linéaire par linéarité du produit scalaire par rapport à la seconde variable. On détaille un peu plus. Si $\vec{u}$ et $\vec{v}$ sont deux vecteurs de $E$ et si $\lambda$ est un scalaire réel, alors :

$$\begin{aligned} \forall x \in E,\ \Phi(\lambda \cdot \vec{u} + \vec{v})(x) &= (x \mid \lambda \cdot \vec{u} + \vec{v}) \\ &= \lambda \cdot (x \mid \vec{u}) + (x \mid \vec{v}) \\ &= \Big( \lambda \cdot \Phi(\vec{u}) + \Phi(\vec{v}) \Big)(x). \end{aligned}$$

Par conséquent, les formes linéaires $\Phi(\lambda \cdot \vec{u} + \vec{v})$ et $\Big(\lambda \cdot \Phi(\vec{u}) + \Phi(\vec{v})\Big)$ sont égales.
Enfin, on montre que l'application linéaire $\Phi$ est injective. En effet, si $\vec{u}$ est dans le noyau $\mathrm{Ker}(\Phi)$, alors pour tout vecteur $x$ de $E$, on sait alors que :

$$(\vec{u} \mid x) = 0.$$

En particulier, $\vec{u}$ est orthogonal à lui-même donc $\vec{u}$ est le vecteur nul.
En conclusion, comme les espaces $E$ et $\mathscr{L}(E, \mathbb{R})$ sont de même dimension finie, alors l'application $\Phi$ est un isomorphisme.
On peut alors reconsidérer ce que l'on doit montrer. Soit $\varphi : E \longrightarrow \mathbb{R}$ une forme linéaire. Il existe alors un seul vecteur $\vec{u} \in E$ tel que :

$$\forall x \in E, \; \varphi(x) = (x \mid \vec{u}).$$

Il s'agit du vecteur :

$$\vec{u} = \Phi^{-1}(\varphi).$$

## Les endomorphismes orthogonaux

Soit $E$ un espace euclidien. Soit $f \in \mathscr{L}(E)$ un endomorphisme.
On démontre l'équivalence entre les trois assertions suivantes :

**(1)** pour tous vecteurs $x$ et $y$ de $E$,

$$(f(x) \mid f(y)) = (x \mid y).$$

**(2)** pour tout vecteur $x$ de $E$, on a l'égalité entre normes euclidiennes :

$$\|f(x)\| = \|x\|.$$

**(3)** l'endomorphisme $f$ transforme une base orthonormale en une base orthonormale.

$\boxed{\mathbf{(1)} \Longrightarrow \mathbf{(2)}}$

Si le **(1)** est vérifié, alors pour tout vecteur $x$ de $E$,

$$\|f(x)\|^2 = (f(x) \mid f(x)) = (x \mid x) = \|x\|^2,$$

en utilisant l'hypothèse avec $y = x$. En prenant la racine carrée, comme tout est positif, on obtient le **(2)**.

$\boxed{\mathbf{(2)} \Longrightarrow \mathbf{(1)}}$

Si le **(2)** est vérifié, on utilise une identité de polarisation. Par exemple, si $x$ et $y$ sont deux vecteurs de $E$, alors :

$$\begin{aligned}
(f(x) \mid f(y)) &= \frac{1}{2}\Big(\|f(x) + f(y)\|^2 - \|f(x)\|^2 - \|f(y)\|^2\Big) \\
&= \frac{1}{2}\Big(\|f(x+y)\|^2 - \|f(x)\|^2 - \|f(y)\|^2\Big), \text{ car } f \text{ est linéaire} \\
&= \frac{1}{2}\Big(\|x+y\|^2 - \|x\|^2 - \|y\|^2\Big), \text{ par le } \mathbf{(2)} \\
&= (x \mid y), \text{ par l'identité de polarisation.}
\end{aligned}$$

$\boxed{\textbf{(1)} \Longrightarrow \textbf{(3)}}$

Si le **(1)** est vérifié, soit $\mathscr{B} = (e_1, \cdots, e_n)$ une b.o.n. de $E$. Cette base $\mathscr{B}$ est transformée par l'endomorphisme $f$ en la famille :

$$\mathscr{F} = (f(e_1), \cdots, f(e_n)).$$

Il est assez clair que pour tous indices $i$ et $j$ entre 1 et $n$,

$$(f(e_i) \mid f(e_j)) = (e_i \mid e_j) = \delta_{i,j}.$$

La famille $\mathscr{F}$ est une famille orthonormale à $n = \dim(E)$ vecteurs, donc est une b.o.n. de $E$.

$\boxed{\textbf{(3)} \Longrightarrow \textbf{(2)}}$

Si le **(3)** est vérifié, il existe une b.o.n. $\mathscr{B} = (\varepsilon_1, \cdots, \varepsilon_n)$ de $E$ telle que la famille $\mathscr{G} = (f(\varepsilon_1), \cdots, f(\varepsilon_n))$ soit encore une b.o.n. de $E$. On vérifie le **(2)** par le théorème de Pythagore. En effet, pour tout vecteur $x$ de $E$, on pose :

$$x = \sum_{i=1}^{n} x_i \cdot \varepsilon_i,$$

de sorte que :

$$\|x\|^2 = \sum_{i=1}^{n} x_i^2.$$

D'autre part, comme l'application $f$ est linéaire, alors :

$$f(x) = \sum_{i=1}^{n} x_i \cdot f(\varepsilon_i).$$

Toujours par le théorème de Pythagore utilisé cette fois-ci dans la b.o.n. $\mathscr{F}$, on obtient :

$$\|f(x)\|^2 = \sum_{i=1}^{n} x_i^2 = \|x\|^2.$$

En prenant la racine carrée, comme tout est positif, on obtient le **(2)**.

On a ainsi prouvé toutes les implications requises pour avoir l'équivalence entre les trois points. De plus, on remarque que si l'endomorphisme $f$ transforme une b.o.n. en une autre b.o.n., alors l'endomorphisme $f$ transformera toute b.o.n. en une autre b.o.n. de $E$.

La propriété 7 est alors assez immédiate.
Premièrement, si $f$ est endomorphisme orthogonal, alors $f$ transforme une b.o.n. en une b.o.n. de $E$, donc l'application linéaire $f$ est un isomorphisme : $f \in GL(E)$ et donc l'ensemble $O(E)$ de tous les endomorphismes orthogonaux sur $E$ est inclus dans l'ensemble $GL(E)$ de tous les isomorphismes de $E$. Une autre manière de faire est de dire que si $f$ est un endomorphisme orthogonal et si $x \in \mathrm{Ker}(f)$, alors $f(x) = 0_E$, donc :

$$\|x\| = \|f(x)\| = \|0_E\| = 0_{\mathbb{R}}.$$

Par conséquent, le vecteur $x$ est le vecteur nul et $f$ est injective puis surjective par le théorème du rang et car $E$ est un espace de dimension finie.

On sait par ailleurs que l'ensemble $GL(E)$ est un groupe pour la composition, en fait un sous-groupe de l'ensemble $\mathfrak{S}(E)$ de toutes les bijections sur $E$, muni de la composition.
On va montrer que l'ensemble $O(E)$ est un sous-groupe du groupe $(GL(E), \circ)$.
On vient d'expliquer l'inclusion :

$$O(E) \subset GL(E).$$

Ensuite, l'application $\mathrm{id}_E$ vérifie immédiatement le point **(2)** par exemple.
Enfin, si $f$ et $g$ sont dans $O(E)$, le plus simple – et ce que je vous conseille la plupart du temps est de vérifier le point **(2)** pour vérifier qu'un endomorphisme est orthogonal – est de vérifier que l'endomorphisme $f \circ g^{-1}$ vérifie le point **(2)**.
Soit donc $x$ un vecteur de $E$. On en déduit :

$$\begin{array}{rcl} \|f \circ g^{-1}(x)\| & = & \|f(g^{-1}(x))\| \\ & = & \|g^{-1}(x)\|, \text{ car } f \in O(E) \text{ et } \mathbf{(2)} \\ & = & \|g(g^{-1}(x))\|, \text{ car } g \in O(E) \text{ et } \mathbf{(2)} \\ & = & \|x\|. \end{array}$$

Résultat des courses :

$$f \circ g^{-1} \in O(E).$$

On a bien ce qu'il faut.
L'application :

$$\det : \left| \begin{array}{ccc} O(E) & \longrightarrow & \mathbb{R}^* \\ f & \longmapsto & \det(f) \end{array} \right.$$

est un morphisme de groupes entre $(O(E), \circ)$ et $(\mathbb{R}^*, \times)$ car pour tous endomorphismes $f$ et $g$ (pas forcément orthogonaux d'ailleurs) :

$$\det(f \circ g) = \det(f) \times \det(g).$$

L'ensemble $SO(E)$ est donc le noyau de ce morphisme de groupe et en tant que tel est un sous-groupe de $O(E)$.

On remarque avant de passer à autre chose qu'il y a assez peu de notions à retenir sur les endomorphismes orthogonaux. On verra des exemples plus tard, une fois que l'on aura vu les matrices orthogonales.

## Les matrices orthogonales

Lorsque l'on regarde la définition 6, on peut être légitimement surpris. On pourrait se dire que si $A$ est une matrice dans $\mathscr{M}_n(\mathbb{R})$, comme il y a un très fort lien entre matrice $A$ et endomorphisme $u_A$ canoniquement associé, alors on n'a qu'à définir le concept de « *matrice orthogonale* $A$ » comme le même que le concept de « *endomorphisme orthogonal* $u_A$ ». Une remarque à ce propos. On ne peut définir un endomorphisme orthogonal sur un espace $E$ qu'à condition que l'espace $E$ soit muni d'un produit scalaire. Ici, l'endomorphisme $u_A$ canoniquement associé à la matrice $A$ est un endomorphisme sur l'espace $\mathscr{M}_{n,1}(\mathbb{R})$ des matrices colonnes. On peut munir cet espace d'une infinité de produit scalaire. Lequel choisir ? Intuitivement, on choisit le plus facile qui est le produit scalaire habituel sur $\mathscr{M}_{n,1}(\mathbb{R})$, à savoir :

$$\forall \left( X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right) \in \mathscr{M}_{n,1}(\mathbb{R}) \times \mathscr{M}_{n,1}(\mathbb{R}),\ (X \mid Y) = \sum_{i=1}^{n} x_i\ y_i.$$

Lorsque l'on fait le produit matriciel $X^T \cdot Y$, on obtient une matrice à un seul coefficient qui vaut $(X \mid Y)$. On identifie comme d'habitude cette matrice dans $\mathscr{M}_1(\mathbb{R})$ à ce scalaire. En utilisant cette identification, on obtient comme produit scalaire habituel sur $\mathscr{M}_{n,1}(\mathbb{R})$ l'application :

$$(X, Y) \longmapsto X^T \cdot Y.$$

On verra que cette démarche intuitive correspond exactement à la définition 6.

On traite les points de la remarque 2.
L'égalité $A^T \cdot A = I_n$ est équivalente à dire que la matrice $A$ est inversible d'inverse $A^T$. On rappelle que si $B$ et $C$ sont deux matrices dans $\mathscr{M}_n(\mathbb{R})$ telles que $BC = I_n$, alors $u_B \circ u_C = \mathrm{id}$, donc l'endomorphisme $u_C$ est injectif puis bijectif en dimension finie, ici les espaces de départ et d'arrivée de $u_C$ étant $\mathscr{M}_{n,1}(\mathbb{R})$. En composant alors à droite par $u_C^{-1}$, on obtient :

$$u_B = u_C^{-1},$$

donc $u_C \circ u_B = \mathrm{id}$ et $CB = I_n$.
Il est donc équivalent de dire « *la matrice $A$ est orthogonale* » que d'écrire « $A \cdot A^T = I_n$ », car ces deux assertions sont équivalentes au fait que la matrice $A$ est inversible et d'inverse $A^T$.
Si $A$ est une matrice orthogonale, alors en prenant le déterminant dans l'égalité $A^T \cdot A = I_n$ par exemple, on obtient :

$$\det(A^T) \cdot \det(A) = \det(I_n) = 1,$$

et donc : $\det(A)^2 = 1$, ou encore $\det(A) = \pm 1$.
La réciproque est fausse comme le montre par exemple la matrice :

$$A = I_n + E_{1,2}.$$

La matrice $A$ est alors une matrice triangulaire supérieure à coefficients diagonaux valant tous 1 : $\det(A) = 1$.
Ensuite,

$$\begin{aligned} A^T \cdot A &= (I_n + E_{2,1}) \cdot (I_n + E_{1,2}) \\ &= I_n + E_{2,1} + E_{1,2} + E_{2,1} \cdot E_{1,2} \\ &= I_n + E_{2,1} + E_{1,2} + E_{2,2} \\ &\neq I_n, \end{aligned}$$

en se rappelant au passage la formule sur les matrices élémentaires :

$$\forall (i, j, k, \ell) \in [\![1, n]\!]^4,\ E_{i,j} \cdot E_{k,\ell} = \delta_{j,k} \cdot E_{i,\ell}.$$

La matrice $A = I_n + E_{1,2}$ de déterminant 1 n'est pas orthogonale.
Lorsque $n = 1$, il n'y a que deux matrices orthogonales : $(\pm 1) \in \mathscr{M}_1(\mathbb{R})$.
On verra grâce à la propriété 8 que l'ensemble $O_n(\mathbb{R})$ est un groupe pour la multiplication en faisant le lien avec les endomorphismes orthogonaux. On peut essayer de le montrer directement uniquement avec la définition donnée pour les matrices orthogonales.
Premièrement, on a l'inclusion $O_n(\mathbb{R}) \subset GL_n(\mathbb{R})$ car toute matrice orthogonale est inversible.
Deuxièmement, on a évidemment $I_n^T \cdot I_n = I_n \cdot I_n = I_n$, donc la matrice $I_n$ est orthogonale.
Troisièmement, si $A$ et $B$ sont deux matrices orthogonales, on peut écrire :

$$\begin{aligned}
(AB^{-1})^T \cdot (AB^{-1}) &= (B^{-1})^T \cdot A^T \cdot A \cdot B^{-1} \\
&= (B^{-1})^T \cdot I_n \cdot B^{-1}, \text{ car } A \text{ est orthogonale} \\
&= (B^{-1})^T \cdot B^{-1} \\
&= (B^T)^{-1} \cdot B^{-1} \\
&= (B \cdot B^T)^{-1} \\
&= I_n^{-1}, \text{ car } B \text{ est orthogonale} \\
&= I_n.
\end{aligned}$$

Ainsi, la matrice $AB^{-1}$ reste dans $O_n(\mathbb{R})$. De même l'ensemble $SO_n(\mathbb{R})$ est le noyau du morphisme de groupes $\det : O_n(\mathbb{R}) \longrightarrow \mathbb{R}^*$ ou de $\det : O_n(\mathbb{R}) \longrightarrow \{-1, 1\}$, le groupe d'arrivée étant muni de la multiplication. L'ensemble $SO_n(\mathbb{R})$ est par conséquent un sous-groupe de $O_n(\mathbb{R})$.

On reboucle la boucle en montrant la propriété 8, qui fait le lien entre « *matrices $A$ orthogonales* » et « *endomorphismes $u_A$ orthogonaux* ».

Il est à noter dans la suite que la base $\mathscr{B} = (e_1, \cdots, e_n)$ est une base orthonormale de $E$. Si l'on prend une base qui n'est pas orthonormale, l'équivalence décrite n'est plus du tout vraie.
On montre l'équivalence en traitant les deux implications ensemble.
Soient deux indices $i$ et $j$ entre 1 et $n$.
Or, les coordonnées de n'importe quel vecteur $f(e_k)$ selon la base $\mathscr{B}$ sont données par les coefficients présents dans la $k^{ème}$ colonne de la matrice $A$. Par conséquent,

$$f(e_i) = \sum_{\ell=1}^{n} A_{\ell,i} \cdot e_\ell \text{ et } f(e_j) = \sum_{m=1}^{n} A_{m,j} \cdot e_m.$$

Il nous reste à utiliser par exemple la formule du produit scalaire dans une b.o.n., en l'occurrence ici la b.o.n. $\mathscr{B}$ (propriété 5 du cours) ou de la redémontrer assez rapidement pour obtenir :

$$\begin{aligned}
(f(e_i) \mid f(e_j)) &= \sum_{\ell=1}^{n} A_{\ell,i} \cdot A_{\ell,j} \\
&= \sum_{\ell=1}^{n} (A^T)_{i,\ell} \cdot A_{\ell,j} \\
&= \left(A^T \cdot A\right)_{i,j}.
\end{aligned}$$

On a alors la suite d'équivalence suivante :

$$\begin{aligned}
f \in O(E) &\iff f \text{ transforme une b.o.n. en une b.o.n.} \\
&\iff \text{la famille } (f(e_1), \cdots, f(e_n)) \text{ est une b.o.n.} \\
&\iff \forall (i,j) \in [\![1,n]\!]^2,\ (f(e_i) \mid f(e_j)) = \delta_{i,j} \\
&\iff \forall (i,j) \in [\![1,n]\!]^2,\ \left(A^T \cdot A\right)_{i,j} = \delta_{i,j}, \text{ par le calcul fait plus haut} \\
&\iff \forall (i,j) \in [\![1,n]\!]^2,\ \left(A^T \cdot A\right)_{i,j} = (I_n)_{i,j} \\
&\iff A^T \cdot A = I_n.
\end{aligned}$$

On y voit maintenant plus clair entre les matrices orthogonales et les endomorphismes orthogonaux.

On termine par démontrer le dernier point de la propriété 8 qui est une série de deux équivalences.
Je vous propose de traiter ces équivalences plus en douceur pour manipuler les notions déjà développées précédemment.
On suppose que la matrice $A \in \mathscr{M}_n(\mathbb{R})$ est orthogonale.
On va montrer que l'endomorphisme $u_A \in \mathscr{L}(\mathscr{M}_{n,1}(\mathbb{R}))$ est un endomorphisme orthogonal lorsque l'on munit l'espace $\mathscr{M}_{n,1}(\mathbb{R})$ du produit scalaire habituel $(X, Y) \longmapsto X^T \cdot Y$.
Il faut retenir la chose très importante suivante concernant le produit scalaire habituel sur $\mathbb{R}^n$ ou sur $\mathscr{M}_{n,1}(\mathbb{R})$ :

> La base canonique de $\mathbb{R}^n$ ou de $\mathscr{M}_{n,1}(\mathbb{R})$ est une base orthonormale pour le produit scalaire habituel.

La vérification est presque immédiate mais c'est un point important qui confère à la base canonique le statut de « meilleure base » en général.
Prenons alors l'espace euclidien $\mathscr{M}_{n,1}(\mathbb{R})$ muni du produit scalaire habituel. Prenons la base canonique $\mathscr{B} = (e_1, \cdots, e_n)$ sur cet espace euclidien. C'est donc une b.o.n. Prenons l'endomorphisme $f = u_A$. On en déduit que la matrice $\mathscr{M}at_{\mathscr{B}}(f)$ représentant $f = u_A$ selon la base canonique est exactement la matrice $A$ et par l'équivalence déjà démontrée de la proposition 8, on sait que comme $A$ est dans $O_n(\mathbb{R})$, alors $f$ est dans $O(\mathscr{M}_{n,1}(\mathbb{R}))$ : l'endomorphisme $u_A$ est bien un endomorphisme orthogonal – je répète : à condition de munir l'espace des matrices colonnes du produit scalaire habituel, sinon, tout cela ne marche plus du tout car la base canonique n'est plus du tout une b.o.n. si l'on modifie le produit scalaire habituel contre un autre produit scalaire.
Pour la réciproque, on suppose que l'endomorphisme $u_A$ est un endomorphisme orthogonal sur l'espace euclidien habituel $\mathscr{M}_{n,1}(\mathbb{R})$. On sait que la base canonique $\mathscr{B} = (e_1, \cdots, e_n)$ est une b.o.n., donc la famille $(u_A(e_1), \cdots, u_A(e_n))$ est une b.o.n. de cet espace euclidien. Entre parenthèses, il s'agit de la famille des vecteurs colonnes de la matrice $A$. On reprendra ce point pour l'autre équivalence tout à l'heure.

Par conséquent, pour tous indices $i$ et $j$ entre 1 et $n$,

$$(u_A(e_i) \mid u_A(e_j)) = \delta_{i,j},$$

donc :

$$(A\ e_i \mid A\ e_j) = \delta_{i,j},$$

et en utilisant la formule du produit scalaire habituel :

$$(A\ e_i)^T \cdot A\ e_j = \delta_{i,j}.$$

Ainsi,

$$e_i^T\ (A^T A)\ e_i = \delta_{i,j}.$$

Or, pour toute matrice carrée $M \in \mathscr{M}_n(\mathbb{R})$, le produit matriciel montre que :

$$e_i^T\ M\ e_j = M_{i,j}.$$

Résultat des courses, la matrice $A^T \cdot A$ vaut $I_n$ et la matrice $A$ est orthogonale.
L'équivalence entre le fait que la matrice $A$ est orthogonale et le fait que la famille des colonnes de $A$ forme une b.o.n. provient directement du fait que la famille des vecteurs colonnes de $A$ est exactement la famille :

$$(u_A(e_1), \cdots, u_A(e_n)),$$

où l'on rappelle que la famille $(e_1, \cdots, e_n)$ est la base canonique de $\mathscr{M}_{n,1}(\mathbb{R})$.
Il suffit alors d'utiliser l'équivalence déjà montrée à savoir que l'endomorphisme $u_A$ est orthogonal si et seulement si la famille $(u_A(e_1), \cdots, u_A(e_n))$ est une b.o.n. de $\mathscr{M}_{n,1}(\mathbb{R})$.

Le paragraphe suivant est l'occasion de traiter des exemples de référence en matière d'endomorphismes orthogonaux ou de matrices orthogonales.

## Les exemples géométriques de référence

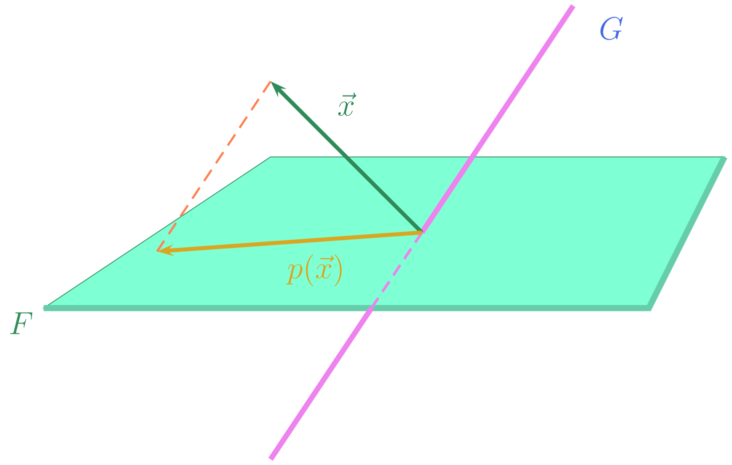
Dans un espace euclidien $E$, si $F$ est un sous-espace de $E$, on sait que $F^\perp$ est un supplémentaire de $F$ dans $E$ :

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

Ceci autorise à parler de la projection sur $F$ parallèlement à $F^\perp$ et à parler de la symétrie par rapport à $F$ parallèlement à $F^\perp$.
On appelle donc projection orthogonale sur $F$, la projection sur $F$ parallèlement à $F^\perp$.
On appelle symétrie orthogonale par rapport à $F$, la symétrie par rapport à $F$ parallèlement à $F^\perp$.

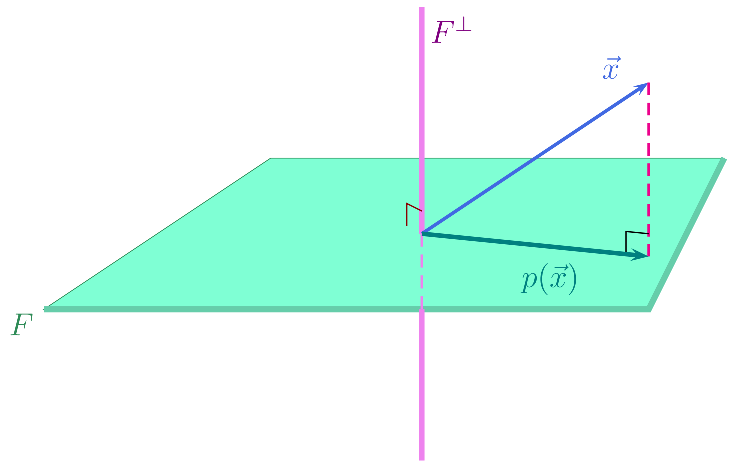
Voici la représentation que vous devez avoir concernant la projection sur $F$ parallèlement à $G$, lorsque :

$$F \oplus G = E.$$




Il est à noter que vous avez besoin de la donnée des deux sous-espaces supplémentaires $F$ et $G$ pour définir votre projection.

Voici maintenant la représentation que vous devez avoir en tête concernant une projection orthogonale sur $F$ :

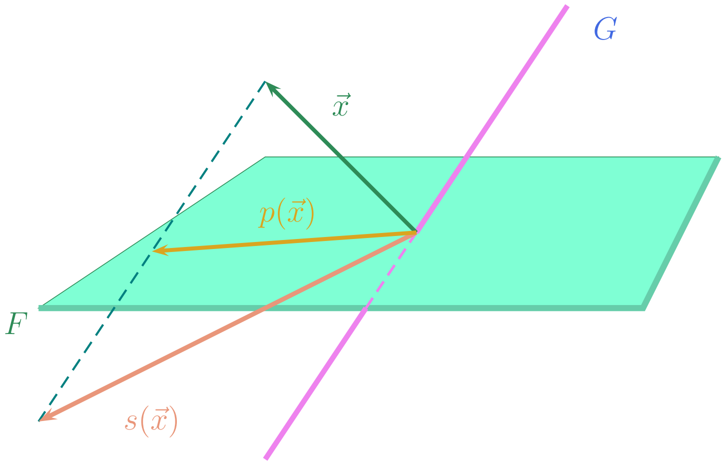



Remarquez que pour définir une projection orthogonale, il suffit de donner l'espace $F$ sur lequel on projète. La projection sera alors la projection sur l'espace $F$ parallèlement à $F^\perp$. Si $p : E \longrightarrow E$ est la projection orthogonale sur $F$, pour tout vecteur $x \in E$, alors $p(x) \in F$ et $x - p(x) \in \mathrm{Ker}(p) = F^\perp$, donc les vecteurs $p(x)$ et $x - p(x)$ sont orthogonaux. Sous ces notations, on peut écrire :

$$F = \mathrm{Im}(p) = \mathrm{Ker}(p - \mathrm{id}_E) \text{ et } F^\perp = \mathrm{Ker}(p).$$

Voici maintenant la représentation que vous devriez avoir en tête lorsque l'on vous parle de symétrie par rapport à $F$, parallèlement à $G$, avec :

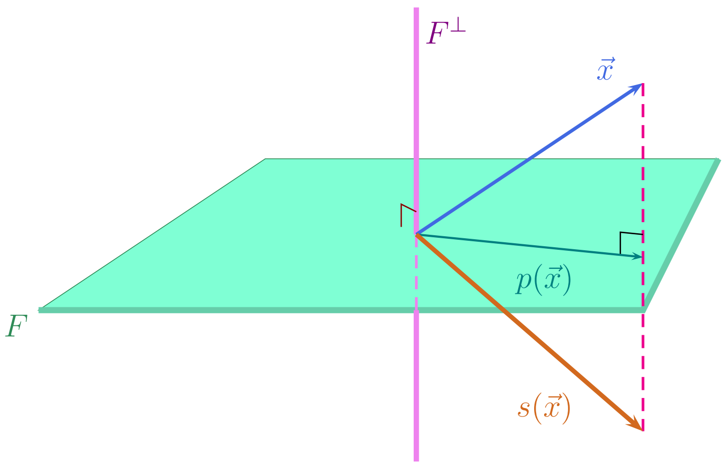
$$F \oplus G = E.$$




Là encore, pour définir notre symétrie, il faut la donnée des deux espaces supplémentaires $F$ et $F$.
Concernant les symétries orthogonales, il suffit de donner l'espace $F$ par rapport auquel s'effectue la symétrie $s$, car dans ce cas, on aura :

$$F = \mathrm{Ker}(s - \mathrm{id}_E) \text{ et } F^{\perp} = \mathrm{Ker}(s + \mathrm{id}_E)$$

selon la disposition géométrique suivante :




On observe le même lien entre la projection $p$ et la symétrie $s$ associée :

$$s = 2p - \mathrm{id}_E \text{ et } p = \frac{s + \mathrm{id}_E}{2}.$$

Remarquons qu'il n'existe qu'une seule projection $p \in \mathscr{L}(E)$ qui soit inversible :

$$p = \mathrm{id}_E.$$

Par conséquent, lorsque vous parlez de « projection orthogonale », je vous conseille de retenir votre respiration et de considérer cette expression comme un seul mot : une projection orthogonale n'est pas une projection qui est orthogonale. Autrement dit, une projection orthogonale est une projection, certes, mais qui n'est pas un endomorphisme orthogonal pour la simple et bonne raison qu'une projection orthogonale n'est pas inversible (sauf s'il s'agit de l'identité $\mathrm{id}_E$) ou une projection orthogonale n'est pas une isométrie : elle ne conserve pas les distances ou les normes comme devrait le faire n'importe quel élément de $O(E)$.

La propriété 9 nous apprend que l'on peut tout à fait prononcer l'expression « symétrie orthogonale » d'un seul coup ou en s'arrêtant entre les deux mots, car une symétrie est une symétrie orthogonale (en un seul mot) si et seulement si c'est une symétrie qui est un endomorphisme orthogonal.
Démontrons ceci.
Soit $s \in \mathscr{L}(E)$ une symétrie.
On suppose que $s$ est une symétrie orthogonale, par rapport à un sous-espace $F$ – sous-entendu parallèlement à $F^\perp$.
Soit $x \in E$. On pose avec des notations intuitives :

$$x = x_F + x_{F^\perp},$$

de sorte que :

$$s(x) = x_F - x_{F^\perp}.$$

Or, comme les vecteurs $x_F$ et $x_{F^\perp}$ sont orthogonaux, on peut écrire par Pythagore :

$$\begin{array}{rcl} \|s(x)\|^2 & = & (x_F - x_{F^\perp} \mid x_F - x_{F^\perp}) \\ & = & \|x_F\|^2 + \|x_{F^\perp}\|^2 \\ & = & \|x\|^2. \end{array}$$

On en déduit que $s$ conserve les normes : c'est un endomorphisme orthogonal.

Supposons maintenant que la symétrie $s \in \mathscr{L}(E)$ soit un endomorphisme orthogonal. La symétrie $s$ est la symétrie par rapport à $F$ parallèlement à $G$, avec

$$F \oplus G = E.$$

Il s'agit de montrer que :

$$G = F^\perp$$

et dans ce cas, la symétrie $s$ sera la symétrie orthogonale par rapport à l'espace $F$.
Soit $a \in G$. Soit $b \in F$. Comme $s$ est un endomorphisme orthogonal, alors :

$$(s(a) \mid s(b)) = (a \mid b).$$

Or, $s(a) = -a$ et $s(b) = b$, ce qui impose :

$$-(a \mid b) = (a \mid b),$$

et donc les vecteurs $a$ et $b$ sont orthogonaux.
On vient de montrer que le vecteur $a$ est orthogonal à n'importe quel vecteur $b$ de $F$ :

$$a \in F^\perp, \text{ puis } G \subset F^\perp.$$

Les sous-espaces $G$ et $F^\perp$ sont deux supplémentaires de l'espace $F$, d'où :

$$\dim(G) = \dim(E) - \dim(F) = \dim(F^\perp).$$

L'inclusion précédente couplée avec l'égalité des dimensions finies nous donne l'égalité voulue :

$$G = F^\perp.$$

La symétrie $s$ est bien la symétrie orthogonale par rapport à l'espace $F$.

Pour la définition 8, on justifie la bonne définition de la borne inférieure.
Si $x$ est un vecteur donné dans $E$ et si $F$ est un sous-espace de $E$, alors l'ensemble :

$$\mathscr{D} = \Big\{ \|x - y\| \mid y \in F \Big\}$$

est non vide car $F \neq \emptyset$, inclus dans $\mathbb{R}$ et minoré par 0. On note $d(x, F)$ cette borne inférieure appelée distance du vecteur $x$ à l'espace $F$.
Géométriquement, pour calculer cette distance, il faudra faire appel à la projection orthogonale $p$ sur $F$ – sous-entendu parallèlement à $F^\perp$ – et cette borne inférieure sera en fait un minimum de distance atteint en le seul vecteur

$$y = p(x).$$

On démontre ceci en démontrant en fait la méthode décrite dans le paragraphe 5.1.
On se donne un vecteur $x$ dans $E$. On se donne un sous-espace $F$ de $E$. On pose $p \in \mathscr{L}(E)$ la projection orthogonale sur $F$.
Soit $y$ un vecteur de $F$ différent du projeté orthogonal $p(x)$ du vecteur $x$ sur l'espace $F$.
On pose avec des notations intuitives :

$$x = x_F + x_{F^\perp},$$

de sorte que :

$$x_F = p(x) \text{ et } x - y = (x_F - y) + x_{F^\perp} = (p(x) - y) + x_{F^\perp}.$$

Comme le vecteur $p(x) - y$ appartient à $F$, alors ce vecteur est orthogonal à $x_{F^\perp}$ et donc :

$$\begin{array}{rcl} \|x - y\|^2 & = & \|p(x) - y\|^2 + \|x_{F^\perp}\|^2 \\ & > & \|x_{F^\perp}\|^2, \text{ car } p(x) - y \neq 0_E \\ & = & \|x - p(x)\|^2. \end{array}$$

En prenant la racine carrée qui est strictement croissante sur l'intervalle $[0, +\infty[$, on en déduit :

$$\forall y \in F \setminus \{p(x)\}, \ \|x - y\| > \|x - p(x)\|.$$

On en déduit que la borne inférieure est atteinte en le seul vecteur $p(x)$ qui réalise donc le minimum de distance parmi tous les vecteurs de l'espace $F$, par rapport au vecteur $x$.
Pour calculer ce minimum de distance, il faut donc passer par le calcul du projeté orthogonal $p(x)$.

On détaille la méthode décrite. Comme l'espace $F$ est donné, on trouve une base $(e_1, \cdots, e_r)$ de $F$, base pas forcément orthonormale. On sait que le vecteur $p(x)$ est combinaison linéaire de ces $r$ vecteurs :

$$p(x) = \sum_{k=1}^{r} \lambda_k \cdot e_k.$$

On sait que le vecteur $x - p(x)$ appartient à $\mathrm{Ker}(p) = F^{\perp}$, donc est orthogonal à tout vecteur de $F$, donc est orthogonal à chaque $e_k$. En disant cela, on n'a en fait perdu aucune information. Il s'agit de calculer les $r$ coordonnées $\lambda_1, \cdots, \lambda_r$ du vecteur $p(x)$ et pour cela on dispose de $r$ équations :

$$\forall k \in [\![1, r]\!],\ (x - p(x) \mid e_k) = 0.$$

On arrive tant bien que mal à résoudre ce système et on obtient les coordonnées $\lambda_k$, puis le vecteur $p(x)$, puis le vecteur $x - p(x)$ et enfin sa norme :

$$\|x - p(x)\| = d(x, F).$$

Je soulève une question : n'est-il pas plus simple d'orthonormaliser la base $(e_1, \cdots, e_r)$ pour avoir une b.o.n. $(\varepsilon_1, \cdots, \varepsilon_r)$ de l'espace $F$ ?
Il semble que oui, mais la réponse est non. En effet, il y a de très forts risques que lors de l'orthonormalisation se glisse au moins une erreur de calcul, avec de toutes façons des calculs pénibles. Je vous déconseille de faire comme cela.

Imaginons malgré tout que vous disposiez d'une b.o.n. $(\varepsilon_1, \cdots, \varepsilon_r)$ de l'espace $F$. Alors, dans ce cas les choses deviennent faciles. En effet, en posant :

$$p(x) = \sum_{i=1}^{r} \alpha_i \cdot \varepsilon_i,$$

on sait que pour tout $i \in [\![1, r]\!]$,

$$\alpha_i = (p(x) \mid \varepsilon_i).$$

Or, on se rappelle que $x = p(x) + (x - p(x))$ avec $x - p(x)$ orthogonal à tout vecteur de $F$, donc orthogonal à tout vecteur $\varepsilon_i$. Résultat des courses, pour tout entier $i$ entre 1 et $r$ :

$$\alpha_i = (x \mid \varepsilon_i),$$

et donc le projeté orthogonal de $x$ sur $F$, lorsque l'on a une b.o.n. $(\varepsilon_1, \cdots, \varepsilon_r)$ vaut :

$$p(x) = \sum_{i=1}^{r} (x \mid \varepsilon_i) \cdot \varepsilon_i.$$

Traitons maintenant l'exemple 8. On ne va pas détailler tous les calculs...
• Pour le premier point, on demande de calculer une borne inférieure. Vous apprendrez à avoir le coup d'œil et à interpréter cette borne inférieure comme un minimum de distance. Plusieurs questions jaillissent : quel vecteur $x$ prend-on [la notation $x$ n'a rien à voir avec la variable d'intégration dans l'intégrale proposée] ? quel espace $F$ prend-on ? et dans quel espace euclidien $E$ ? avec quel produit scalaire ?
Bref, cela fait beaucoup de questions et pas beaucoup de réponses pur l'instant.

Lorsqu'on lit ce qu'il y a dans la borne inférieure, on remarque que les réels $a$ et $b$ varient. Lorsque ceux-ci varient, à l'intérieur de l'intégrale, on a une expression du type $a\ \sin + b\ \cos$. On pressent qu'il sera une bonne idée de considérer l'espace :

$$F = \mathrm{Vect}(\sin, \cos).$$

Ensuite, on remarque que dans l'intégrale intervient la fonction exp qui fera office de vecteur dans l'espace euclidien :

$$E = \mathrm{Vect}(\exp, \cos, \sin),$$

espace de dimension finie inférieure ou égale à 3 – en fait de dimension égale à 3 exactement. Toutes les fonctions considérées sont définies sur le segment d'intégration $[0, \pi]$.
Reste à prendre un bon produit scalaire. Celui-ci découle de l'intégrale proposée. On choisit le produit scalaire :

$$\forall (f,g) \in E^2,\ (f \mid g) = \int_0^\pi f \times g,$$

dont on sait qu'il s'agit d'un produit scalaire. Il y a le théorème aux quatre hypothèses derrière tout cela.
Résultat des courses, avec les notations introduites, on observe que :

$$\forall (a,b) \in \mathbb{R}^2,\ \int_0^\pi \left(e^x - a\ \sin x - b\ \cos x\right)^2 dx = \| \exp - a\ \sin - b\ \cos \|^2,$$

et donc que la borne inféireure à calculer vaut :

$$\inf\left\{ \| \exp - f \|^2\ ;\ f \in F \right\} = \inf\left\{ \| \exp - f \|\ ;\ f \in F \right\}^2 = d(\exp, F)^2,$$

où la distance est calculée grâce au produit scalaire spécifié plus haut.
On introduit alors $p$ la projection orthogonale sur $F$.
On applique la méthode.
On a une base $(\sin, \cos)$ de l'espace $F$.
On pose :

$$p(\exp) = a\ \sin + b\ \cos .$$

On calcule $a$ et $b$ en utilisant les deux équations :

$$\begin{cases} (\exp - a\ \sin - b\ \cos \mid \cos) = 0 \\ (\exp - a\ \sin - b\ \cos \mid \sin) = 0 \end{cases}.$$

Pour résoudre ce système, il faut calculer les intégrales :

$$\int_0^\pi e^x \cos x\ dx,\ \int_0^\pi e^x\ \sin x\ dx, \int_0^\pi \cos^2 x\ dx,$$

$$\int_0^\pi \sin^2 x\ dx \text{ et } \int_0^\pi \cos x\ \sin x\ dx.$$

Pour les deux premières intégrales, le plus efficace est d'écrire :

$$e^x \cos x = \Re e(e^{(1+i)x}) \text{ et } e^x \sin x = \Im m(e^{(1+i)x}),$$

puis d'intégrer en complexe avant de prendre la partie réelle ou imaginaire.
Pour les autres intégrales, on calcule le tout à grand renfort de formules trigonométriques et de linéarisation. Là encore, on peut passer par les exponentielles complexes.
Une fois $a$ et $b$ calculés, ce que l'on ne fait pas – calculs malgré tout pénibles – on obtient $p(x)$, puis la borne inférieure à calculer :

$$\begin{aligned} \|\exp - a\ \sin - b\ \cos\|^2 &= (\exp - a\ \sin - b\ \cos \mid \exp - a\ \sin - b\ \cos) \\ &= (\exp - a\ \sin - b\ \cos \mid \exp), \text{ car } \exp - a\ \sin - b\ \cos \in F^{\perp}. \end{aligned}$$

On détaille les calculs pour le deuxième point.
Il faut très fortement sous-entendre dans ce second point que la distance va être calculée avec le produit scalaire habituel sur $\mathbb{R}^4$. On choisit d'emblée ce produit scalaire habituel.
Ensuite, on a le vecteur $x = (1, 1, 1, 1)$.
Enfin, on a une base $(e_1 = (1, 0, 2, 1), e_2 = (0, 0, 3, 2))$ de l'espace $F$.
Appliquons tranquillement la méthode.
On introduit la projection orthogonale $p \in \mathscr{L}(\mathbb{R}^4)$ sur l'espace $F$.
On pose :

$$p(x) = a\ e_1 + b\ e_2.$$

On trouve $a$ et $b$ via le système :

$$\begin{aligned} \left\{ \begin{array}{l} (x - a\ e_1 - b\ e_2) \mid e_1) = 0 \\ (x - a\ e_1 - b\ e_2) \mid e_2) = 0 \end{array} \right. &\iff \left\{ \begin{array}{l} 4 - 6a - 8b = 0 \\ 5 - 8a - 13b = 0 \end{array} \right. \\ &\iff \left( \begin{array}{cc|c} 3 & 4 & 2 \\ 8 & 13 & 5 \end{array} \right) \\ &\iff \left( \begin{array}{cc|c} 3 & 4 & 2 \\ 0 & \frac{7}{3} & -\frac{1}{3} \end{array} \right) \quad L_2 \longleftarrow L_2 - \frac{8}{3} L_1 \\ &\iff \left\{ \begin{array}{l} a = \dfrac{6}{7} \\ b = -\dfrac{1}{7} \end{array} \right. . \end{aligned}$$

On en déduit :

$$x - p(x) = \frac{1}{7}\ (1, 7, -2, 3).$$

La distance voulue vaut donc :

$$\|x - p(x)\| = \frac{1}{7}\ \|(1, 7, -2, 3)\| = \frac{\sqrt{63}}{7}.$$

Occupons-nous de l'équivalence de l'exemple suivant, qui est une caractérisation des projecteurs orthogonaux qu'il est bien de retenir. Cette notion sera de toutes façons largement reprise en deuxième année...

On suppose que la matrice $A \in \mathscr{M}_n(\mathbb{R})$ est une matrice de projection orthogonale, dans l'espace euclidien habituel $\mathscr{M}_{n,1}(\mathbb{R})$. La projection orthogonale $u_A$ est donc la projection orthogonale sur :

$$F = \mathrm{Im}(u_A),$$

parallèlement à :

$$F^{\perp} = \mathrm{Ker}(u_A).$$

En tant que projection, on a immédiatement $u_A^2 = u_A$, ou encore $A^2 = A$.
On va montrer que la matrice $A$ est symétrique.
On considère une b.o.n. $\mathscr{B}_1 = (e_1, \cdots, e_r)$ de l'espace $F$, puis une b.o.n. $\mathscr{B}_2 = (\varepsilon_1, \cdots, \varepsilon_s)$ de $F^{\perp}$. On sait que la famille :

$$\mathscr{C} = (\mathscr{B}_1, \mathscr{B}_2)$$

est une b.o.n. de l'espace $\mathscr{M}_{n,1}(\mathbb{R})$.
On note $P$ la matrice de passage de la base canonique de $\mathscr{M}_{n,1}(\mathbb{R})$ vers la base $\mathscr{C}$. En tant que matrice de passage entre deux b.o.n. – on rappelle que la base canonique est une b.o.n. pour l'espace euclidien envisagé ici – la matrice $P$ est automatiquement une matrice orthogonale. Pour les plus sceptiques, la matrice $P$ aura ses colonnes qui formeront une b.o.n. de l'espace euclidien.
D'autre part, il est clair que :

$$P^{-1}AP = \mathscr{M}at_{\mathscr{C}}(u_A) = J_r.$$

On en déduit :

$$A = PJ_rP^{-1}$$

et comme la matrice $P$ est orthogonale, alors $P^{-1} = P^T$, puis :

$$A = PJ_rP^T.$$

Par conséquent,

$$\begin{array}{rcl} A^T & = & (PJ_rP^T)^T \\ & = & (P^T)^T J_r^T P^T \\ & = & PJ_rP^T, \text{ car la matrice } J_r \text{ est diagonale donc symétrique} \\ & = & A. \end{array}$$

On a bien la première implication.
On suppose réciproquement que la matrice $A \in \mathscr{M}_n(\mathbb{R})$ vérifie :

$$A^2 = A \text{ et } A^T = A.$$

Comme $A^2 = A$, alors l'endomorphisme $u_A$ est une projection. On note $F = \mathrm{Im}(u_A)$ et $G = \mathrm{Ker}(u_A)$, de sorte que l'endomorphisme $u_A$ est la projection sur $F$ parallèlement à $G$.
Il s'agit de montrer que $G = F^{\perp}$.
Soit $X \in G$. Soit $Y \in F$.
On en déduit en utilisant très fortement la formule du produit scalaire habituel :

$$(Z_1 \mid Z_2) = Z_1^T \cdot Z_2$$

dans $\mathscr{M}_{n,1}(\mathbb{R})$ :

$$\begin{aligned}(X \mid Y) &= (X \mid AY), \text{ car } AY = Y, \text{ puisque } Y \in F \\ &= X^T \cdot AY \\ &= X^T\ A^T\ Y, \text{ car } A^T = A \\ &= (AX)^T \cdot Y \\ &= 0^T \cdot Y, \text{ car } X \in G = \operatorname{Ker}(u_A) \\ &= 0_{\mathbb{R}}.\end{aligned}$$

On vient de montrer que $G$ est inclus dans $F^\perp$ et par l'égalité des dimensions finies toutes les deux égales à $n - \dim(F)$, on obtient l'égalité. Résultat des courses, l'endomorphisme $u_A$ est la projection orthogonale sur l'espace $F$.

Il reste à examiner plus précisément le groupe $O_2(\mathbb{R})$ et à interpréter géométrique des matrices orthogonales de format $2\times2$. L'étude de $O_n(\mathbb{R})$ et en particulier de $O_3(\mathbb{R})$ sera faite en deuxième année.

On démontre la méthode du paragraphe 5.2.
Soit $A$ une matrice dans $O_2(\mathbb{R})$. Comme toute matrice orthogonale, le déterminant de la matrice $A$ vaut 1 ou $-1$.
On pose la matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in O_2(\mathbb{R})$.
La famille $(C_1, C_2)$ des deux colonnes forme une base orthonormée de $\mathscr{M}_{2,1}(\mathbb{R})$. En particulier, on dispose des égalités suivantes :

$$\begin{cases} a^2 + c^2 = 1 \\ b^2 + d^2 = 1 \\ a\ b + c\ d = 0 \end{cases}.$$

Le complexe $z = a + i\ c$ est de module 1 donc appartient à l'ensemble $\mathbb{U}$ et donc est de la forme :

$$z = e^{i\theta}.$$

Ainsi,

$$a = \cos\theta \text{ et } c = \sin\theta.$$

De même, en considérant le complexe $Z = b + i\ d$ appartenant au cercle unité, on trouve un angle $\theta'$ tel que :

$$b = \cos\theta' \text{ et } d = \sin\theta'.$$

L'égalité $a\ b + c\ d = 0$ donne :

$$\cos\theta \cdot \cos\theta' + \sin\theta \cdot \sin\theta' = 0,$$

ou encore :

$$\cos(\theta' - \theta) = 0$$

ou encore :

$$\theta' - \theta \equiv \frac{\pi}{2}\ [\pi].$$

On pose alors :

$$\theta' = \theta + \frac{\pi}{2} + k\ \pi,$$

pour un certain entier $k$.
Le déterminant de la matrice $A$ vaut :

$$\det(A) = a\ d - b\ c = \cos\theta \cdot \sin\theta' - \cos\theta' \cdot \sin\theta = \sin(\theta' - \theta) = \sin\left(\frac{\pi}{2} + k\ \pi\right) = (-1)^k.$$

On en déduit l'équivalence :

$$A \in SO_2(\mathbb{R}) \iff \text{l'entier } k \text{ est pair.}$$

On distingue les deux cas.

- on suppose que $\det(A) = 1$, ou encore que :

$$A \in SO_2(\mathbb{R}).$$

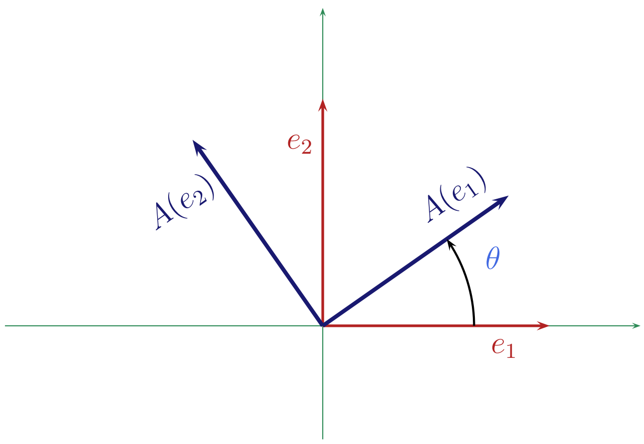
  Dans ce cas, $\theta' \equiv \theta + \dfrac{\pi}{2}\ [2\pi]$ et donc :

$$\cos\theta' = -\sin\theta \text{ et } \sin\theta' = \cos\theta.$$

  La matrice $A$ est donc de la forme :

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

  La matrice $A$ est donc la rotation vectorielle d'angle $\theta$, comme on peut s'en rendre compte sur la manière dont est transformée la base canonique de $\mathbb{R}^2$ identifié à $\mathscr{M}_{2,1}(\mathbb{R})$ :




- on suppose que $\det(A) = -1$, ou encore que :

$$A \notin SO_2(\mathbb{R}).$$

  Dans ce cas, $\theta' \equiv \theta + \dfrac{\pi}{2} + \pi\ [2\pi]$ et donc :

$$\cos\theta' = \sin\theta \text{ et } \sin\theta' = -\cos\theta.$$

La matrice $A$ est donc de la forme :

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

On remarque alors que la matrice $A$ est symétrique et comme la matrice $A$ est orthogonale, alors :

$$A^2 = A \cdot A = A^T \cdot A = I_2.$$

La matrice $A$ est une matrice de symétrie orthogonale. La matrice $A$ est donc la symétrie orthogonale par rapport à un sous-espace $F$ de $\mathscr{M}_{2,1}(\mathbb{R})$ et selon une base adaptée à :

$$F \oplus F^\perp = \mathscr{M}_{2,1}(\mathbb{R})$$

la matrice $A$ est semblable à une matrice diagonale $D$ avec des $\pm 1$ sur la diagonale. Il y a en fait $\dim(F)$ coefficients qui valent 1 et $2 - \dim(F) = \dim(F^\perp)$ coefficients valant $(-1)$ sur cette diagonale. D'autre part,

$$\mathrm{Tr}(A) = 0,$$

donc la trace de cette matrice diagonale $D$ est encore nulle. Cela impose :

$$\dim(F) = 1 = \dim(F^\perp).$$

Le sous-espace $F$ est une droite vectorielle.

En posant le vecteur $X = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix}$, on remarque que :

$$AX = \begin{pmatrix} \cos\theta \ \cos\frac{\theta}{2} + \sin\theta \ \sin\frac{\theta}{2} \\ \sin\theta \ \cos\frac{\theta}{2} - \cos\theta \ \sin\frac{\theta}{2} \end{pmatrix} = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix} = X.$$

Le vecteur invariant $X$ génère donc la droite vectorielle :

$$\mathrm{Ker}(A - I_2) = F.$$

La matrice $A$ est donc la symétrie orthogonale par rapport à $\mathrm{Vect}(X)$.

En résumé, on peut partager les matrices orthogonales de format $2 \times 2$ en deux catégories. L'ensemble $SO_2(\mathbb{R})$ correspond aux matrices de rotations, alors que l'ensemble $O_2(\mathbb{R}) \setminus SO_2(\mathbb{R})$ correspond à des matrices de symétries orthogonales par rapport à une droite vectorielle.

On termine le chapitre par une application numérique.
On remarque déjà que la première colonne

$$C_1 = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix},$$

avec l'angle $\theta = \frac{\pi}{6}$ est bien un vecteur unitaire.

La deuxième colonne de la matrice $A$ orthogonale doit être un vecteur unitaire orthogonal au vecteur $C_1$. Il y a deux possibilités :

$$C_2 = \pm \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}.$$

▷ Si on choisit le signe « $+$ », alors :

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

Dans ce cas, la matrice $A$ est la rotation vectorielle d'angle $\theta = \dfrac{\pi}{6}$.
▷ Si on choisit le signe « $-$ », alors :

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

Dans ce cas, la matrice $A$ est la symétrie orthogonale par rapport à :

$$\mathrm{Vect}\left(\begin{pmatrix} \cos(\pi/12) \\ \sin(\pi/12) \end{pmatrix}\right).$$

Pour les curieux qui voudraient avec une valeur autre, on obtient :

$$\cos\frac{\pi}{12} = \cos\left(\frac{\pi}{3} - \frac{\pi}{4}\right) = \frac{\sqrt{6}+\sqrt{2}}{4}$$

et :

$$\sin\frac{\pi}{12} = \sin\left(\frac{\pi}{3} - \frac{\pi}{4}\right) = \frac{\sqrt{6}-\sqrt{2}}{4}.$$