Chapitre 26 : Espaces préhilbertiens réels.

MPSI Lycée Camille Jullian

14 juin 2024

Monsieur et Madame Bertienne ont un fils, comment s'appelle-t-il ?

Basile.

Si quelqu'un, en l'éveil de son intelligence, n'a pas
été capable de s'enthousiasmer pour une telle architecture,
alors jamais il ne pourra réellement s'initier à la recherche théorique.

Albert Einstein, à propos de la géométrie euclidienne.

L'objectif de ce dernier chapitre d'algèbre de l'année (mais oui, déjà) est de munir nos espaces vectoriels préférés d'une structure supplémentaire permettant de vraiment pouvoir faire de la géométrie de façon satisfaisante : celle de produit scalaire. Jusqu'ici, le produit scalaire était pour vous un moyen d'étudier l'orthogonalité, et découlait donc des notions naturelles d'angle et de perpendicularité. Dans le cadre plus général des espaces vectoriels réels, on procèdera en fait exactement en sens inverse : il n'existe pas de notion « naturelle » d'orthogonalité ou même de distance dans un espace vectoriel, ces notions sont en fait une conséquence de l'introduction d'un outil théorique appelé produit scalaire. Il est surtout essentiel de bien comprendre qu'un tel produit scalaire n'a aucune raison d'être unique, et qu'il existe donc plusieurs façons différentes (tout aussi valables les unes que les autres) de définir des distances (ou une notion d'orthogonalité) dans un espace vectoriel comme $\mathcal{M}_3(\mathbb{R})$ ou $\mathbb{R}[X]$. Même dans $\mathbb{R}^n$, le produit scalaire usuel n'est pas le seul disponible, même s'il sera privilégié en pratique. Une fois ce choix de produit scalaire effectué, on pourra vraiment faire de la géométrie dans notre espace, et notamment parler de projections ou de symétries orthogonales sur un sous-espace vectoriel donné. On reviendra en particulier sur la notion d'isométrie, qu'on verra sous un jour nouveau (et beaucoup plus algébrique) par rapport à la première étude faite dans le chapitre sur les complexes.

Objectifs du chapitre :

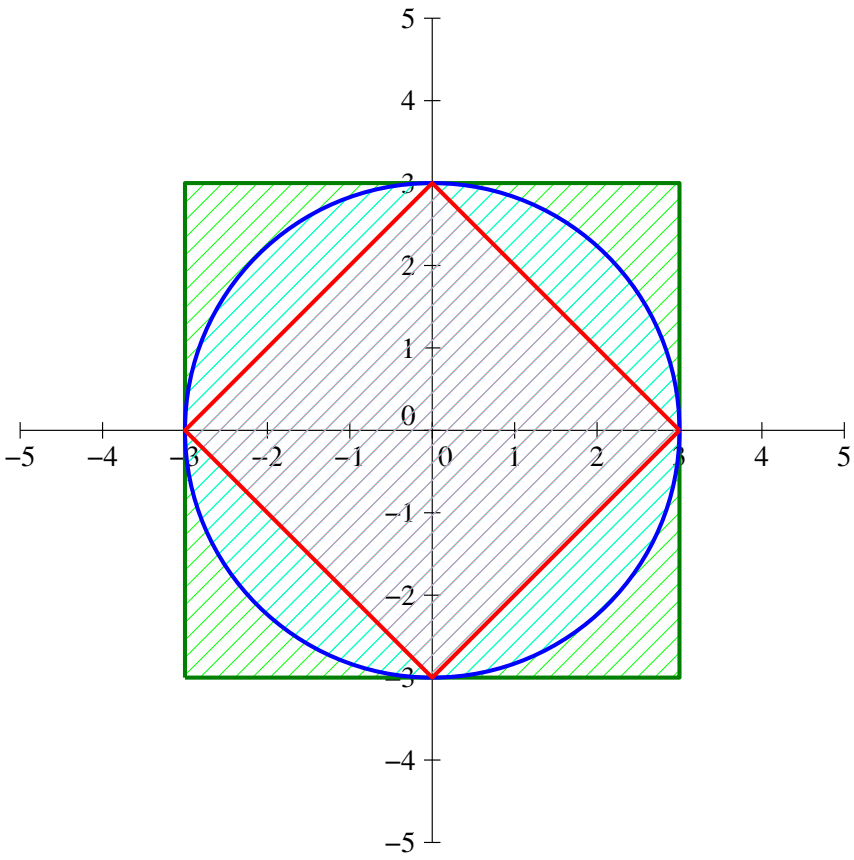
- comprendre la notion de produit scalaire et de distance dans un espace vectoriel autre que $\mathbb{R}^2$ et $\mathbb{R}^3$.
- savoir exploiter les bases orthonormales (et en calculer) pour effectuer des calculs de projections.
- connaître les différents types d'isométries dans le plan et dans l'espace.

Introduction.

Avant de rentrer dans le vif du sujet, quelques mots sur des concepts qui vont irriguer les deux derniers chapitres de l'année concernant la géométrie du plan (ce chapitre sera directement centré sur la géométrie dans les espaces vectoriels réels, le suivant sera consacré aux fonctions de deux variables, sujet a priori assez éloigné du précédent, mais qui nécessite une bonne compréhension de la topologie dans $\mathbb{R}^2$, ce qui n'est pas sans lien avec ce dont je vais parler ici). Vous avez l'habitude depuis que vous avez commencé à faire de la géométrie de considérer certains concepts géométriques basiques comme « naturels » et de vous appuyer dessus pour comprendre les notions plus compliquées progressivement mises en place dans votre cours de mathématiques. Ainsi, la notion « visuelle » d'angles et de perpendicularité dans le plan fait partie des premières rencontrées au début du collège, et mène par exemple, après quelques péripéties, à la construction d'outils comme le produit scalaire, qui permet de caractériser de façon très pratique l'orthogonalité. Sauf que, comme souvent, les mathématiciens ont une vision plus abstraite des choses, et aiment bien baser leurs théories sur des constructions logiques solides, et qu'en l'occurrence cela risque de perturber un peu vos habitudes. Commençons donc par deux remarques fondamentales avant de donner les premières définitions :

- la notion d'orthogonalité n'a en fait rien de naturel dans les espaces vectoriels (y compris les plus simples comme $\mathbb{R}^2$ ou $\mathbb{R}^3$), et surtout rien d'**intrinsèque** : autant la colinéarité découle logiquement de la structure même d'espace vectoriel, autant l'orthogonalité nécessite l'ajout d'une structure supplémentaire qui n'a aucune raison d'être initialement présente dans notre espace. C'est d'ailleurs pour cela que nous verrons dans ce chapitre le produit scalaire comme l'outil fondamental dont vont découler toutes les autres notions, et non le contraire.
- toutes les notions que nous allons étudier dans ce chapitre (produit scalaire, orthogonalité, distance, projections) n'ont pas la moindre raison d'être uniques ! Il existe ainsi des dizaines de possibilités de créer des produits scalaires, et donc des distances, différentes sur $\mathbb{R}^2$, qui vont « déformer » notre vision intuitive des choses (par exemple, les boules ne seront plus rondes, cf exemple ci-dessous). Mais ces différentes façons de procéder définiront malgré tout la même **topologie** sur $\mathbb{R}^2$ (ce ne serait pas forcément le cas dans des espaces de dimension infinie, mais vous étudierez ça plus en détail l'an prochain), nous essaierons d'en redire deux mots dans le prochain chapitre.

Un exemple concret : dans le plan, on appelle **boule ouverte** centrée en A et de rayon R l'ensemble des points du plan dont la distance à A est strictement inférieure à R. Pour la distance usuelle, les boules ouvertes sont effectivement des boules (des disques en l'occurrence). Mais considérons une autre définition possible de la distance : si $A(x, y)$ et $M(x', y')$, on va décréter que la distance AM sera désormais égale à $\max(|x - x'|, |y - y'|)$. Cette distance un peu inhabituelle a des propriétés mathématiques suffisamment sympathiques pour qu'elle mérite effectivement le nom de distance. Mais à quoi ressemble, pour cette distance, la boule ouverte centrée en A de rayon R ? Eh bien, à un carré (ouvert). Un autre exemple : on considère maintenant qu'on travaille avec la distance usuelle, mais en ayant uniquement le droit de se déplacer suivant des segments de droite parallèles aux axes (imaginez le plan des rues de Manhattan, on ne peut pas traverser les immeubles pour se rendre d'un point à un autre). Les boules ouvertes pour cette nouvelle distance ressembleraient à nouveau à des carrés, mais dont les côtés seraient « diagonaux ». Une illustration ci-dessous : en bleu, la boule ouverte centrée en l'origine de rayon 3, en rouge la même boule ouverte pour la distance « de Manhattan », et en vert toujours une boule de rayon 3 pour la deuxième distance définie.



On pourra remarquer que, quitte à faire une homothétie, chaque type de boule peut être entièrement incluse dans une boule d'un autre type. C'est ce qui fait que les topologies associées sont les mêmes.

1 Définition du produit scalaire.

Dans tout le chapitre, nous ne manipulerons que des espaces vectoriels réels, même si les notions étudiées peuvent se généraliser (avec quelques petites modifications) à des espaces complexes. Ces espaces vectoriels peuvent par contre très bien être de dimension infinie.

Définition 1. On appelle **produit scalaire** dans un espace vectoriel réel E toute forme bilinéaire symétrique définie positive, autrement dit toute application $\varphi : E \times E \to \mathbb{R}$ vérifiant les trois propriétés fondamentales suivantes :

- bilinéarité : $\forall (u,v,w) \in E^2$, $\forall (\lambda,\mu) \in \mathbb{R}^2$, $\varphi(\lambda u + \mu v, w) = \lambda\varphi(u,w) + \mu\varphi(v,w)$ et $\varphi(u, \lambda v + \mu w) = \lambda\varphi(u,v) + \mu\varphi(u,w)$.
- symétrie : $\forall (u,v) \in E^2$, $\varphi(u,v) = \varphi(v,u)$.
- définie positivité : $\forall u \in E$, $\varphi(u,u) \geqslant 0$, et $\varphi(u,u) = 0 \Leftrightarrow u = 0$.

Si φ est un produit scalaire, $\varphi(u,v)$ sera souvent noté $u \cdot v$, ou $\langle u,v\rangle$ ou même $\langle u \mid v\rangle$ (mais cette dernière notation étant privilégiée par les physiciens, on l'évitera dans ce cours de maths).

Remarque 1. En pratique, pour démontrer qu'une application φ est un produit scalaire, on se contente de prouver la linéarité à gauche (première des deux relations données pour la bilinéarité), celle à droite découlant ensuite de la symétrie.

Définition 2. Un **espace préhilbertien** est un espace vectoriel E muni d'un produit scalaire. Si E est de dimension finie, on parlera plutôt d'**espace euclidien**.

Proposition 1. L'application $\varphi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ définie par $\varphi(u,v) = \sum_{i=1}^{n} u_i v_i$ (en notant u_i et v_i les coordonnées respectives des vecteurs u et v dans la base canonique) est un produit scalaire appelé **produit scalaire canonique** sur $\mathbb{R}^n$.

L'application φ définie sur $(\mathcal{M}_{n,p}(\mathbb{R}))^2$ par $\varphi(A,B) = \mathrm{Tr}(A^\top B) = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{i,j}b_{i,j}$ est un produit scalaire appelé **produit scalaire canonique** sur $\mathcal{M}_{n,p}(\mathbb{R})$.

Démonstration. Tout est essentiellement trivial : la symétrie est évidente dans les deux cas (avec la formule explicite faisant intervenir les coefficients dans le cas des matrices), la bilinéarité aussi à partir des formules données, la positivité est triviale ($u \cdot u = \sum_{i=1}^{n} u_i^2$ et $A \cdot A = \sum_{i,j=1}^{n} a_{i,j}^2$). Enfin, le caractère défini découle du fait que toutes nos coordonnées (ou coefficients dans le cas matriciel) doivent être nulles pour que la somme de leur carré soit égale à 0, ce qui ne peut donc arriver que pour le vecteur nul. □

Remarque 2. On peut en fait calculer le produit scalaire de deux vecteurs de $\mathbb{R}^n$ de façon matricielle : en notant X et Y les matrices-colonnes des coordonnées des deux vecteurs u et v dans la base canonique, alors $u \cdot v = X^\top Y$.

Exemple : Il existe bien d'autres produits scalaires définis sur $\mathbb{R}^2$. En notant $u = (x,y)$ et $v = (x',y')$, le produit scalaire canonique correspond à la formule usuelle $u \cdot v = xx' + yy'$. Mais on peut par exemple prouver aisément que $u \cdot v = 2xx' + xy' + x'y + 2yy'$ est également un produit scalaire : la symétrie et la bilinéarité son évidents, et avec cette définition, $u \cdot u = 2x^2 + 2y^2 + 2xy = x^2 + y^2 + (x+y)^2$ est toujours positif et ne s'annule que si $x = y = 0$.

Exemple : Dans l'espace vectoriel $\mathbb{R}_n[X]$, l'application définie par $\varphi(P,Q) = \sum_{k=0}^{n} P(k)Q(k)$ est un produit scalaire : elle est bien symétrique et bilinéaire, et $\varphi(P,P) = \sum_{k=0}^{n} P^2(k) \geqslant 0$. De plus, si $\varphi(P,P) = 0$, alors on peut affirmer que $\forall k \in \{0,\dots,n\}$, $P(k) = 0$, et donc P est un polynôme de degré n admettant (au moins) $n+1$ racines distinctes, il est donc nécessairement nul. Remarquons tout de même qu'il est indispensable de faire intervenir $n+1$ valeurs distinctes dans le calcul de $\varphi(P,Q)$ pour qu'il s'agisse d'un produit scalaire (peu importe par contre qu'il s'agisse des entiers de 0 à n ou d'autres valeurs). En particulier, il s'agit d'un produit scalaire qu'on ne pourrait pas « étendre » facilement à $\mathbb{R}[X]$. On verra plus bas quelques autres exemples de produits scalaires sur les espaces de polynômes.

Exemple : Sur l'espace vectoriel E des fonctions continues sur $[0,1]$, on peut définir un produit scalaire par la formule suivante : $f \cdot g = \int_0^1 f(t)g(t)\,dt$. Là encore, les vérifications de symétrie, bilinéarité et positivité sont triviales. De plus, si $f \cdot f = 0$, alors $\int_0^1 f(t)^2\,dt = 0$. Comme il s'agit de l'intégrale d'une fonction continue et positive sur $[0,1]$, elle ne peut s'annuler que si la fonction f est nulle sur tout l'intervalle $[0,1]$. Attention tout de même, cette même définition ne fournirait pas un produit scalaire sur l'ensemble des fonctions continues sur $\mathbb{R}$ tout entier car l'application ne serait pas définie (il existe plein de fonctions continues sur $\mathbb{R}$ et nulles entre 0 et 1, mais qui ne sont pas nulles partout). Ce produit scalaire est en fait une sorte de « généralisation continue » du produit scalaire canonique de $\mathbb{R}^n$ défini plus haut. On peut d'ailleurs effectuer exactement la même

construction sur tout segment $[a, b]$ autre que $[0, 1]$. Notons en passant que, sur ce même segment $[0, 1]$, on peut construire facilement d'autres produits scalaires en ajoutant une fonction **positive** à l'intérieur de l'intégrale, par exemple $f \cdot g = \int_0^1 f(t)g(t)e^t \, dt$, ou $f \cdot g = \int_0^1 t^2 f(t)g(t) \, dt$ (la preuve du fait qu'il s'agit toujours de produits scalaires est laissée en exercice).

Définition 3. Soit $(E, \cdot)$ un espace vectoriel préhilbertien, et $(u, v) \in E^2$. La **norme** du vecteur u est le réel positif $\|u\| = \sqrt{u \cdot u}$. La **distance** entre les vecteurs u et v est le réel positif $d(u, v) = \|u - v\|$.

Remarque 3. Il est essentiel de bien comprendre que la distance entre deux vecteurs dépend du choix initial du produit scalaire. On peut très bien avoir des distances entre deux mêmes vecteurs qui vont être très différentes selon le produit scalaire utilisé. En fait, la notion de distance est donc initimement liée à celle d'orthogonalité, ce que certains des résultats énoncés plus loin dans ce cours vont illustrer.

Exemple : plaçons-nous dans l'espace vectoriel $E = \mathbb{R}_3[X]$ des polynômes de degré inférieur ou égal à 3. On souhaite définir un produit scalaire sur cet espace vectoriel, notamment pour pouvoir calculer des distances entre polynômes. Cette notion de distance entre deux polynômes peut tout à fait s'interpréter intuitivement : on peut par exemple avoir envie de dire que deux polynômes sont « proches » s'ils ont des coefficients proches, ou bien s'ils prennent des valeurs proches à certains endroits, ou encore si leurs courbes ne s'éloignent pas trop l'une de l'autre (tout ça restant bien sûr assez flou pour l'instant). Eh bien, ces différentes notions de proximité entre polynômes peuvent en fait être matérialisées par différents choix de produits scalaires sur l'espace E, qui amèneront à des définitions différentes des distances.

On peut ainsi prouver que l'application φ définie par $\varphi(P, Q) = \sum_{i=0}^{3} a_i b_i$ (où a_i et b_i désignent les coefficients des polynômes P et Q) est un produit scalaire. C'est même le produit scalaire naturel sur $\mathbb{R}_3[X]$ puisqu'il correspond au produit scalaire canonique sur $\mathbb{R}^4$, en identifiant les coefficients du polynôme à des coordonnées. À partir de ce produit scalaire, la distance entre deux polynômes P et Q sera définie par la formule $\|P - Q\| = \sqrt{\sum_{i=0}^{3} (a_i - b_i)^2}$. On retrouve ici une notion de distance où des polynômes sont à une faible distance l'un de l'autre si leurs coefficients sont proches. Par exemple, la distance entre $P = 1$ et $Q = 2X^2 - X$ sera égale à $\sqrt{1 + 1 + 4} = \sqrt{6}$, mais la distance entre $P = X$ et $Q = 2X + X^2$ sera égale à $\sqrt{1 + 1} = \sqrt{2}$ et donc plus petite que la précédente (ce qui devrait vous paraitre vaguement logique).

Définissons maintenant sur ce même espace vectoriel un deuxième produit scalaire : on pose $\psi(P, Q) = \int_0^1 P(t)Q(t) \, dt$. Je vous laisse vérifier que cette définition est bien celle d'un produit scalaire, à partir duquel on peut donc définir une notion de distance qui n'est plus du tout la même que la précédente. Cette fois-ci, la distance entre deux polynômes P et Q sera égale à $\|P - Q\| = \sqrt{\int_0^1 (P(t) - Q(t))^2 \, dt}$. Pour ce nouveau produit scalaire, la distance entre $P = 1$ et $Q = 2X^2 - X$ sera égale à $\sqrt{\int_0^1 (2t^2 - t - 1)^2 \, dt} = \frac{2}{\sqrt{5}}$, et la distance entre $P = X$ et $Q = 2X + X^2$ vaut $\sqrt{\int_0^1 (t^2 + t)^2 \, dt} = \sqrt{\frac{31}{30}}$. Cette fois-ci c'est la deuxième distance qui est plus grande, ce qui peut se traduire par le fait que les courbes des polynômes P et Q sont « plus proches l'une de l'autre sur l'intervalle $[0, 1]$ » dans le deuxième cas. Bien entendu, ce genre de calcul peut sembler sans intérêt, mais toute la puissance de la géométrie euclidienne appliquée aux espaces vectoriels réside dans le fait qu'on peut réellement exploiter toutes les notions et tous les calculs classiques de géométrie dans ce cadre. Un exemple : on cherche une courbe qui approche le plus possible celle d'une fonction

(allez, disons celle de l'exponentielle pour fixer les idées) sur l'intervalle $[0,1]$, en imposant que la fonction recherchée soit un polynome de degré 3. Les développements limités fournissent une réponse si on cherche la meilleure approximation possible en un point donné, mais pas sur un intervalle. En fait, le produit scalaire dont nous venons de parler fournit une façon de faire ce calcul : il « suffit » d'effectuer une projection orthogonale de la fonction exponentielle sur le sous-espace vectoriel $\mathbb{R}_3[X]$, et ce calcul est réellement un calcul de projection orthogonale tout à fait classique (au détail près que la notion d'orthogonalité utilisée est bien entendu celle qui découle de notre produit scalaire). Nous allons voir dans plus loin dans ce cours comment effectuer de tels calculs.

Proposition 2. Lien entre normes et produit scalaire.

Si $(E,\cdot)$ est un espace vectoriel préhilbertien, et $(u,v)\in E^2$, on a les relations suivantes :

- $\|u+v\|^2 = \|u\|^2 + 2\langle u,v\rangle + \|v\|^2$
- $\|u-v\|^2 = \|u\|^2 - 2\langle u,v\rangle + \|v\|^2$
- $\langle u+v, u-v\rangle = \|u\|^2 - \|v\|^2$
- $\langle u,v\rangle = \dfrac{1}{2}(\|u+v\|^2-\|u\|^2-\|v\|^2) = \dfrac{1}{4}(\|u+v\|^2-\|u-v\|^2)$ (identités de polarisation)
- $\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2+\|v\|^2)$ (identité du parallélogramme)
- $|\langle u,v\rangle| \leqslant \|u\| \times \|v\|$ (inégalité de Cauchy-Schwartz), avec égalité si et seulement si u et v sont colinéaires
- $\|u+v\| \leqslant \|u\| + \|v\|$ (inégalité triangulaire), avec égalité si et seulement si u et v sont colinéaires de même sens

Démonstration. La plupart de ces formules sont des conséquences directes de la bilinéarité et de la symétrie du produit scalaire :

- $\|u+v\|^2 = (u+v)\cdot(u+v) = uu + u\cdot v + v\cdot u + v\cdot v = \|u\|^2 + 2u\cdot v + \|v\|^2$.
- Pareil que ci-dessus avec des changements de signes.
- Encore à peu près pareil : $\langle u+v, u-v\rangle = \langle u,u\rangle + \langle v,u\rangle - \langle u,v\rangle - \langle v,v\rangle = \|u\|^2 - \|v\|^2$.
- Les identités de polarisation découlent de façon triviale des deux premières formules prouvées.
- L'identité du parallalogramme est obtenue directement en additionnant ces deux mêmes formules.
- Il y a un peu plus de travail pour démontrer l'inégalité de Cauchy-Schwartz. Rappelons la démonstration classique déjà vue dans le cours d'intégration : on pose $P(t) = \langle u+tv, u+tv\rangle = \|u\|^2 + 2t\langle u,v\rangle + t^2\|v\|^2$. Comme le produit scalaire calculé est toujours positif, le discriminant du trinôme obtenu doit être négatif, ce qui revient à dire que $4\langle u,v\rangle^2 - 4\|u\|^2\|v\|^2 \leqslant 0$. On en déduit immédiatement l'inégalité. De plus, l'inégalité est une égalité si le discriminant calculé est nul, ce qui se produit si $P(t)$ s'annule pour une certaine valeur de t, autrement dit s'il existe un réel t tel que $u+tv=0$, ce qui implique bien la colinéarité des deux vecteurs.
- L'inégalité triangulaire découle de celle de Cauchy-Schwartz : $\|u+v\|^2 = \|u\|^2 + 2\langle u,v\rangle + \|v\|^2 \leqslant \|u\|^2 + 2\|u\|\times\|v\| + \|v\|^2 = (\|u\|+\|v\|)^2$, on passe tout à la racine carrée pour obtenir l'inégalité.

□

Exemple : L'inégalité de Cauchy-Schwartz permet de démontrer énormément d'inégalités, y compris certaines qui n'ont a priori pas grand rapport avec la géométrie euclidienne. Par exemple, si on l'applique dans $\mathbb{R}^2$ aux vecteurs $u=(x_1,\dots,x_n)$ et $v=(1,\dots,1)$, on obtient la majoration $\left(\sum_{k=1}^n x_k\right)^2 \leqslant n\sum_{k=1}^n x_k^2$, avec égalité si et seulement si $x_1=\dots=x_n$.

Définition 4. Un vecteur $u\in E$ préhilbertien est **unitaire** (ou **normé**) si $\|u\|=1$.

2 Orthogonalité.

2.1 Bases orthogonales.

Définition 5. Soit E un espace vectoriel préhilbertien, alors :

- deux vecteurs $(u, v) \in E^2$ sont **orthogonaux** si $u \cdot v = 0$. On le notera $u \perp v$.
- deux sous-ensembles F et G de E (pas nécessairement des sous-espaces vectoriels) sont **orthogonaux** si $\forall (u, v) \in F \times G$, $u \cdot v = 0$. On le notera $F \perp G$.
- une famille de vecteurs $\mathcal{F} = (u_1, \ldots, u_k)$ est **orthogonale** si elle est constituée de vecteurs deux à deux orthogonaux, **orthonormale** si de plus les vecteurs de la famille sont tous unitaires.
- une **base orthonormale** (ou simplement **orthogonale**) dans un espace euclidien est une base qui forme une famille orthonormale (respectivement orthogonale).

Remarque 4. On ne parlera jamais de perpendicularité pour des sous-ensembles d'espaces préhilbertiens, mais toujours d'orthogonalité. En effet, la notion de droites perpendiculaires dans un espace préhilbertien (droites orthogonales ayant un point en commun) n'a aucun intérêt dans la mesure où deux droites vectorielles ont toujours un élément en commun (le vecteur nul de l'espace). La notion de perpendicularité serait adaptée uniquement si on considérait des sous-espaces affines, ce que nous ne ferons pas.

Exemple : La base canonique de $\mathbb{R}^n$ est une base orthonormale pour le produit scalaire canonique. De même pour la base canonique de $\mathcal{M}_{n,p}(\mathbb{R})$. Si on reprend par contre les deux exemples de produits scalaires sur $\mathbb{R}_3[X]$ donnés dans le paragraphe précédent, la base canonique de $\mathbb{R}_3[X]$ est orthonormale pour φ, mais pas du tout pour ψ. En effet, elle n'est même pas normée pour ce produit scalaire, par exemple $\|1\| = \sqrt{\displaystyle\int_0^1 1 \; dt} = 1$, mais $\|X\| = \sqrt{\displaystyle\int_0^1 t^2 \; dt} = \dfrac{1}{\sqrt{3}}$. Elle n'est pas non plus orthogonale puisque $\psi(1, X) = \displaystyle\int_0^1 t \; dt = 1$. On va voir dans le paragraphe suivant comment construire une base orthonormale pour ce produit scalaire en partant de la base canonique.

Exemple : En notant $E = \{f : \mathbb{R} \to \mathbb{R} \text{ continues et } 2\pi\text{-périodiques}\}$, on peut définir un produit scalaire sur E par $f \cdot g = \dfrac{1}{\pi} \displaystyle\int_0^{2\pi} f(t)g(t) \; dt$. En notant $u_n : t \mapsto \cos(nt)$ et $v_n : t \mapsto \sin(nt)$ (pour $n \geqslant 1$), la famille $\mathcal{F}_n = (u_1, v_1, \ldots, u_n, v_n)$ est une famille orthonormale dans E quel que soit $n \geqslant 1$. En effet, via formules de duplication, $\dfrac{1}{\pi} \displaystyle\int_0^{2\pi} \cos^2(nt) \; dt = \dfrac{1}{\pi} \int_0^{2\pi} \dfrac{1 + \cos(2nt)}{2} \; dt = \dfrac{1}{2\pi} \left[t + \dfrac{\sin(2nt)}{2n} \right]_0^{2\pi} = 1$, donc $\|u_n\| = 1$. Le calcul est quasiment le même pour les sinus. De plus, avec un peu de transformations sommes-produits, si $n \neq p$, $u_n \cdot u_p = \dfrac{1}{\pi} \displaystyle\int_0^{2\pi} \cos(nt)\cos(pt) \; dt = \dfrac{1}{2\pi} \int_0^{2\pi} \cos((n+p)t) + \cos((n-p)t) \; dt = 0$ (les deux cosinus ont des primitives en sinus qui s'annulent en 0 et en 2π). Un calcul identique montre que $v_n \cdot v_p = 0$, et un dernier calcul pas plus intéressant montre enfin que $u_n \cdot v_p$ est toujours nul (même pour $n = p$). On aurait bien envie de dire que la famille infinie obtenue en prenant toutes les fonctions u_n et v_n (et une fonction constante de norme 1) est une base orthonormale de l'espace E, mais il faudrait pour cela que toute fonction de E puisse s'écrire comme combinaison linéaire de fonctions cosinus et sinus, ce qui n'est pas le cas (si c'était le cas, on dirait que la famille est une **base hilbertienne** de l'espace E, définition qui n'est pas à votre programme cette année mais qui vous permettra de comprendre l'hilarante blague donnée en préambule de ce chapitre). On peut par contre approcher n'importe quelle fonction de E par un élément de $\text{Vect}(\mathcal{F}_n)$, quel que soit l'entier n, en effectuant une projection orthogonale sur ce sous-espace vectoriel (cf paragraphes suivants). Plus n est grand, plus cette projection sera « proche »

de la fonction de départ. C'est exactement le principe du développement en série de Fourier d'une fonction périodique.

Théorème 1. Théorème de Pythagore :
Deux vecteurs $(u, v) \in E$ sont orthogonaux si et seulement si $\|u+v\|^2 = \|u\|^2 + \|v\|^2$.

Démonstration. C'est trivial à partir du développement de $\|u+v\|^2$ donné dans le paragraphe précédent. On peut généraliser le principe du théorème de Pythagore à une somme de k vecteurs : si la famille $(u_1, \ldots, u_k)$ est orthogonale, alors $\left\|\sum_{i=1}^{k} u_i\right\|^2 = \sum_{i=1}^{k} \|u_i\|^2$ (propriété qui sera exploitée dans la démonstration du corollaire ci-dessous). □

Corollaire 1. Une famille orthogonale de vecteurs non nuls dans un espace préhilbertien est toujours libre.

Démonstration. Si $(u_1, \ldots, u_k)$ est orthogonale, supposons $\sum_{i=1}^{k} \lambda_i u_i = 0$, alors $\left\|\sum_{i=1}^{k} \lambda_i u_i\right\|^2$ $= \sum_{i=1}^{k} \|\lambda_i u_i\|^2 = 0$, ce qui impose que $\forall i \in \{1, \ldots, k\}$, $\|\lambda_i u_i\|^2 = 0$, et donc $\lambda_i = 0$ (les vecteurs n'ayant pas le droit d'être nuls). Ce raisonnement prouve bien la liberté de la famille. □

Proposition 3. Soit $u \in E$ et $\mathcal{B} = (e_1, \ldots, e_n)$ une base orthonormale de E, alors $u = \sum_{k=1}^{n} \langle u, e_k \rangle e_k$. Autrement dit, les coordonnées de u dans la base $\mathcal{B}$ sont ses produits scalaires avec les vecteurs de $\mathcal{B}$.

Démonstration. Soient $(u_1, \ldots, u_n)$ les coordonnées de u dans la base $\mathcal{B}$, alors $\langle u, e_k \rangle = \sum_{i=1}^{n} \langle u_i e_i, e_k \rangle =$ u_i (tous les produits scalaire de la somme sont nuls, sauf celui avec e_i qui vaut u_i, par définition de ce qu'est une base orthonormale). □

Proposition 4. Si $\mathcal{B}$ est une base orthonormale de E, le produit scalaire des vecteurs $u = (u_1, \ldots, u_n)$ et $v = (v_1, \ldots, v_n)$ peut se calculer par la formule $u \cdot v = \sum_{i=1}^{n} u_i v_i$, où les réels u_i et v_i sont les cooordonnées des deux vecteurs dans la base $\mathcal{B}$. Par conséquent, $\|u\|^2 = \sum_{i=1}^{n} u_i^2$.

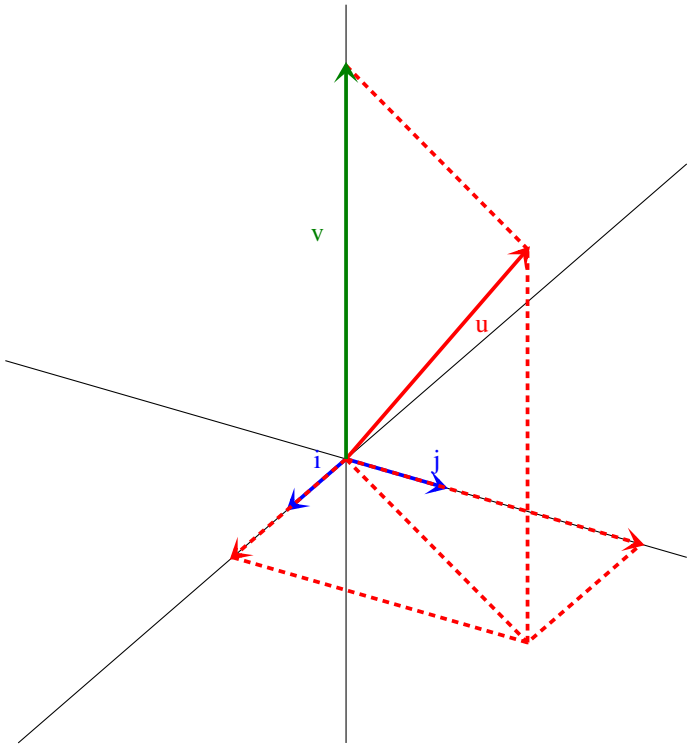
Démonstration. On utilise à nouveau très simplement la bilinéarité : $u \cdot v = \left(\sum_{i=1}^{n} u_i e_i\right) \cdot \left(\sum_{j=1}^{n} v_j e_j\right) = \sum_{i,j=1}^{n} u_i v_j (e_i \cdot e_j) = \sum_{i=1}^{n} u_i v_i$. □

2.2 Orthonormalisation de Gram-Schmidt.

Le but de ce paragraphe est très simple : proposer une construction explicite permettant de transformer n'importe quelle base d'une espace vectoriel euclidien en base orthonormale (on parlera plutôt de base orthonormée dans ce cas). Pour cela, on va s'appuyer sur le principe suivant : il suffit de « redresser » un vecteur pour le rendre orthogonal à tous les vecteurs d'une famille.

Proposition 5. Si la famille $(e_1, \ldots, e_k)$ est orthonormale, et $u \in E$, alors le vecteur $u - \sum_{i=1}^{k} \langle u, e_k \rangle e_k$ est orthogonal à tous les vecteurs e_1, e_2, …, e_k.

Sur la figure ci-dessous, la famille (i, j) est orthonormale, et le vecteur $v = u - \langle u, i \rangle i - \langle u, j \rangle j$ est orthogonal à la fois à i et à j. Les vecteurs $\langle u, i \rangle i$ et $\langle u, j \rangle j$ sont indiqués avec une flèche (ce sont les deux vecteurs colinéaires à i et j).



Démonstration. C'est encore une fois un calcul immédiat : $\left\langle u - \sum_{i=1}^{k} \langle u, e_i \rangle e_i, e_j \right\rangle = \langle u, e_j \rangle - \sum_{i=1}^{k} \langle u, e_i \rangle \langle e_i, e_j \rangle = \langle u, e_j \rangle - \langle u, e_j \rangle = 0$. □

Théorème 2. Procédé d'orthonormalisation de Gram-Schmidt.

Soit $(u_1, u_2, \ldots, u_n)$ une base d'un espace vectoriel euclidien E, on peut alors construire une base orthonormale $(e_1, \ldots, e_n)$ de E en procédant de la façon suivante :

- on pose $e_1 = \dfrac{u_1}{\|u_1\|}$
- on calcule $u_2' = u_2 - \langle u_2, e_1\rangle e_1$, puis on pose $e_2 = \dfrac{u_2'}{\|u_2'\|}$
- plus généralement, pour tout entier $k \leqslant n$, en supposant avoir déjà construit les vecteurs $(e_1, \ldots, e_{k-1})$, on calcule $u_k' = u_k - \displaystyle\sum_{i=1}^{k-1} \langle u_k, e_i\rangle e_i$, puis on normalise le vecteur en posant $e_k = \dfrac{u_k'}{\|u_k'\|}$.

De plus, on aura, $\forall k \in \{1, 2, \ldots, n\}$, $\mathrm{Vect}(e_1, \ldots, e_k) = \mathrm{Vect}(u_1, \ldots, u_k)$.

Démonstration. C'est en fait à peu près trivial en utilisant la propriété précédente. On démontre par récurrence sur k que la famille $(e_1, \ldots, e_k)$ est toujours libre, orthonormale, et vérifie $\mathrm{Vect}(e_1, \ldots, e_k) = \mathrm{Vect}(u_1, \ldots, u_k)$. C'est évident pour $k = 1$ (on a simplement transformé u_1 en vecteur unitaire en le divisant par sa norme), et si on le suppose au rang k, alors par construction e_{k+1} est normé, et orhogonal à tous les vecteurs construits précédemment, donc la famille est bien orthonormale. De plus, e_k ne peut pas appartenir à $\mathrm{Vect}(e_1, \ldots, e_{k-1})$ sinon ce serait aussi le cas de u_k (qui ne diffère de e_k que par un élément appartenant justement à $\mathrm{Vect}(e_1, \ldots, e_{k-1})$), ce qui contredirait la liberté de la famille initiale puisque par hypothèse de récurrence $\mathrm{Vect}(e_1, \ldots, e_{k-1}) = \mathrm{Vect}(u_1, \ldots, u_{k-1})$. La famille $(e_1, \ldots, e_k)$ est donc toujours une famille libre, et $u_k \in \mathrm{Vect}(e_1, \ldots, e_k)$ (par définition de u_k' puis de e_k), donc $\mathrm{Vect}(e_1, \ldots, e_k) = \mathrm{Vect}(u_1, \ldots, u_k)$ (il contient tous les vecteurs u_1, …, u_k et est de dimension k). À la fin de la construction, on a donc bien obtenu une famille orthonormale de n vecteurs qui est libre, c'est une base orthonormale de E. □

Corollaire 2. Il existe des bases orthonormales dans tout espace euclidien E. De plus, on peut compléter toute famille orthonormale de E en une base orthonormale de E (théorème de la base orthonormale incomplète).

Démonstration. Il suffit d'appliquer le procédé de Gram-Schmidt à une base quelconque de E (tout ev de dimension finie admet des bases). Pour compléter une famille orthonormale en base, on la complète en base, puis on applique encore une fois Gram-Schmidt : le procédé laissera intact les premiers vecteurs de la base qui constituaient une famille orthonormale. □

Exemple : Il est absolument essentiel de savoir appliquer l'agorithme de Gram-Schmidt en pratique. Si vous avez bien compris la figure faite un peu plus haut, les formules à appliquer doivent en fait vous paraître naturelles. Reprenons l'exemple de $E = \mathbb{R}_3[X]$, muni du produit scalaire $P \cdot Q = \displaystyle\int_0^1 P(t)Q(t)\, dt$, et appliquons le procédé de Gram-Schmidt à la base canonique pour obtenir une base orthonormale de E pour ce produit scalaire :

- On commence par normer le premier vecteur de la base canonique : $\|1\| = \sqrt{\displaystyle\int_0^1 1\, dt} = 1$, donc il n'y a rien à faire, on pose $e_1 = 1$.

- On calcule ensuite $X \cdot 1 = \displaystyle\int_0^1 t\ dt = \frac{1}{2}$, puis on pose $u_2' = X - \frac{1}{2}$. Il reste à calculer $\left\|X - \frac{1}{2}\right\|^2 = \displaystyle\int_0^1 \left(t - \frac{1}{2}\right)^2 dt = \int_0^1 t^2 - t + \frac{1}{4}\ dt = \frac{1}{3} - \frac{1}{2} + \frac{1}{4} = \frac{1}{12}$ (on a calculé le carré de la norme pour éviter de trainer des racines carrées dans les calculs, ce qu'on fera aussi pour les calculs de normes suivants), puis à poser $e_2 = \dfrac{X - \frac{1}{2}}{\frac{1}{\sqrt{12}}} = 2\sqrt{3}X - \sqrt{3}$.
- On calcule ensuite $X^2 \cdot 1 = \displaystyle\int_0^1 t^2\ dt = \frac{1}{3}$, et (c'est moins agréable) $X^2 \cdot (2\sqrt{3}X - \sqrt{3}) = \displaystyle\int_0^1 2\sqrt{3}t^3 - \sqrt{3}t^2\ dt = \frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{3} = \frac{\sqrt{3}}{6}$, puis on pose $u_3' = X^2 - \frac{1}{3} - \frac{\sqrt{3}}{6}(2\sqrt{3}X - \sqrt{3}) = X^2 - X + \frac{1}{6}$. Reste à nouveau à calculer la norme : $\left\|X^2 - X + \frac{1}{6}\right\|^2 = \displaystyle\int_0^1 \left(t^2 - t + \frac{1}{6}\right)^2 dt = \int_0^1 t^4 - 2t^3 + \frac{4}{3}t^2 - \frac{1}{3}t + \frac{1}{36}\ dt = \frac{1}{5} - \frac{1}{2} + \frac{4}{9} - \frac{1}{6} + \frac{1}{36} = \frac{36 - 90 + 80 - 30 + 5}{180} = \frac{1}{180}$. On posera donc $e_3 = \sqrt{180}\left(X^2 - X + \frac{1}{6}\right) = 6\sqrt{5}X^2 - 6\sqrt{5}X + \sqrt{5}$.
- Normalement, vous devriez déjà vous être enfuis en anticipant que le dernier calcul va être immonde. Je vous rassure tout de suite : oui, ce sera le cas. Calculons donc $X^3 \cdot 1 = \displaystyle\int_0^1 t^3\ dt = \frac{1}{4}$, $X^3 \cdot (2\sqrt{3}X - \sqrt{3}) = \displaystyle\int_0^1 2\sqrt{3}t^4 - \sqrt{3}t^3\ dt = \frac{2\sqrt{3}}{5} - \frac{\sqrt{3}}{4} = \frac{3\sqrt{3}}{20}$, et enfin $X^3 \cdot (6\sqrt{5}X^2 - 6\sqrt{5}X + \sqrt{5}) = \displaystyle\int_0^1 6\sqrt{5}t^5 - 6\sqrt{5}t^4 + \sqrt{5}t^3\ dt = \sqrt{5} - \frac{6\sqrt{5}}{5} + \frac{\sqrt{5}}{4} = \frac{\sqrt{5}}{20}$. Il est temps de poser $u_4' = X^3 - \frac{1}{4} - \frac{3\sqrt{3}}{20}(2\sqrt{3}X - \sqrt{3}) - \frac{\sqrt{5}}{20}(6\sqrt{5}X^2 - 6\sqrt{5}X + \sqrt{5}) = X^3 - \frac{3}{2}X^2 + \frac{3}{5}X - \frac{1}{20}$. Allez, un dernier effort à fournir pour calculer la norme de cette horreur : $\|u_4'\|^2 = \displaystyle\int_0^1 \left(t^3 - \frac{3}{2}t^2 + \frac{3}{5}t - \frac{1}{20}\right)^2 dt = \int_0^1 t^6 - 3t^5 + \frac{69}{20}t^4 - \frac{19}{10}t^3 + \frac{51}{100}t^2 - \frac{3}{50}t + \frac{1}{400}\ dt = \frac{1}{2\ 800}$. Oui, je vous ai épargné une partie du détail car je n'ai tout simplement pas fini le calcul à la main, c'est trop horrible. Concluons quand même en posant $e_4 = \sqrt{2\ 800}\left(X^3 - \frac{3}{2}X^2 + \frac{3}{5}X - \frac{1}{20}\right) = 20\sqrt{7}X^3 - 30\sqrt{7}X^2 + 12\sqrt{7}X - \sqrt{7}$.

On conclut bien sûr brillament que la base (e_1, e_2, e_3, e_4) est bel et bien orthonormale pour notre produit scalaire. Ce calcul aura au moins servi à démontrer une chose : en pratique, un produit scalaire en apparence bien gentillet peut donner des bases orthonormales extrêmement laides. On appliquera de fait très rarement Gram-Schmidt en dimension supérieure à 3 ou 4, sauf cas particuliers de produits scalaires donnant des résultats très maniables.

2.3 Supplémentaires orthogonaux.

Définition 6. Soit E un espace préhilbertien et $F \subset E$, on appelle **orthogonal de** F, et on note $F^\perp$, l'ensemble $F^\perp = \{u \in E \mid \forall v \in F,\ \ u \cdot v = 0\}$.

Exemples : même si ce n'est pas explicitement imposé, on calculera quasiment exclusivement les orthogonaux de sous-espaces vectoriels de E. Comme le vecteur nul est le seul vecteur orthogonal à tous les vecteurs d'un espace préhilbertien (il est le seul vecteur à être orthogonal à lui-même), on a $E^\perp = \{0\}$, et « réciproquement » $\{0\}^\perp = E$.

Dans $\mathbb{R}^3$ muni du produit scalaire canonique, en notant $F = \{(1,1,1)\}$, on obtient $F^\perp = \{(x,y,z) \in \mathbb{R}^3 \mid x + y + z = 0\}$. On constate ici que l'orthogonal d'un vecteur (ou de la droite engendrée par ce vecteur, on obtiendra le même ensemble) est un plan décrit par une équation cartésienne dont les

coefficients sont les coordonnées du vecteur. On dit plutôt en géométrie que le vecteur est un vecteur **normal** au plan. La constatation se généralise en fait : si $u \in E$ a pour cooordonnées $(u_1, u_2, \dots, u_n)$ dans une base orthonormale de E (qu'on suppose donc ici euclidien), alors $\{u\}^\perp$ est un hyperplan de E décrit par l'équation cartésienne $\sum_{i=1}^{n} u_i x_i = 0$ (et on dira, comme dans l'espace, que u est un vecteur normal de l'hyperplan en question).

Dans $\mathbb{R}_3[X]$, on décide de se munir du produit scalaire suivant : $P \cdot Q = P(0)Q(0) + P(1)Q(1) + P(2)Q(2) + P(3)Q(3)$ (démonstration du fait qu'il s'agit d'un produit scalaire laissée en exercice), et on souhaite déterminer $(\mathbb{R}_1[X])^\perp$. Par bilinéarité du produit scalaire, un polynôme est orthogonal à $\text{Vect}(1, X)$ si et seulement si il est orthogonal à 1 et à X (plus généralement, un vecteur sera orthogonal à un sous-espace vectoriel de E si et seulement s'il est orthogonal à tous les vecteurs d'une famille génératrice de ce sous-espace). Soit donc $P = aX^3 + bX^2 + cX + d$, alors $P \cdot 1 = P(0)+P(1)+P(2)+P(3) = 36a+14b+6c+4d$, et $P \cdot X = P(1)+2P(2)+3P(3) = 98a+36b+14c+6d$. Notre orthogonal est donc constitué des polynômes vérifiant $36a+14b+6c+4d = 98a+36b+14c+6d = 0$. En multipliant la première équation par $\frac{3}{2}$ avant de la soustraire à la seconde, on trouve $44a + 15b + 5c = 0$, soit $c = -\frac{44}{5}a - 3b$. On peut alors reporter dans notre première équation : $36a + 14b - \frac{264}{5}a - 18b + 4d = 0$, donc $4d = \frac{84}{5}a + 4b$, ou encore $d = \frac{21}{5}a + b$. Finalement on obtient $(\mathbb{R}_1[X])^\perp = \text{Vect}\left(X^3 - \frac{44}{5}X + \frac{21}{5}, X^2 - 3X + 1\right)$.

Proposition 6. Propriétés de l'orthogonal.

- $F^\perp$ est toujours un sous-espace vectoriel de E.
- si F est un sous-espace vectoriel de E, alors F et $F^\perp$ sont en somme directe.
- si F est un sous-espace vectoriel de dimension finie de E, alors $F^\perp$ est un supplémentaire de F, appelé **supplémentaire orthogonal** de F.
- $F \subset F^{\perp\perp}$, avec égalité dans le cas où F est un sous-espace vectoriel de dimension finie de E.

Démonstration. La bilinéarité du produit scalaire fait que $F^\perp$ est naturellement stable par combinaisons linéaires, et donc qu'il s'agit d'un sous-espace vectoriel de E (il contient toujours le vecteur nul). Supposons maintenant que F est lui-même un sous-espace vectoriel, et que $u \in F \cap F^\perp$. Alors u doit en particulier vérifier $u \cdot u = 0$, ce qui implique $u = 0$. F et son orthogonal sont donc bien en somme directe. Passons au cas de la dimension finie pour F. On sait déjà que $F \cap F^\perp = \{0\}$, reste à prouver que $F + F^\perp = E$. Soit donc $u \in E$, on a vu plus haut que, si $(e_1, \dots, e_k)$ était une base orthonormale de F (et il en existe toujours), le vecteur $v = u - \sum_{i=1}^{k} (u \cdot e_k) e_k$ était orthogonal à tous les vecteurs de la famille $(e_1, \dots, e_k)$. Par linéarité, on en déduit que $v \in F^\perp$. Comme $\sum_{i=1}^{k} (u \cdot e_k) e_k \in F$ par construction, on a bien $u \in F + F^\perp$, ce qui prouve la supplémentarité des deux sous-espaces. L'inclusion $F \subset F^{\perp\perp}$ est complètement triviale. Dans le cas où F est de dimension finie, la supplémentarité prouvée plus haut montre que tout vecteur peut s'écrire sous la forme $u = f_1 + f_2$, avec $f_1 \in F$ et $f_2 \in F^\perp$. Supposons $u \in F^{\perp\perp}$, alors $u \perp f_2$, donc $(f_1 \cdot f_2) + (f_2 \cdot f_2) = 0$. Mais $f_1 \cdot f_2 = 0$ par définition de f_2, donc $(f_2 \cdot f_2) = 0$, ce qui impose $f_2 = 0$ et donc $u \in F$. □

Exemples : On se place dans $E = \{f : [0,1] \to \mathbb{R} \text{ continues}\}$, muni du produit scalaire intégral déjà croisé $f \cdot g = \int_0^1 f(t)g(t)\, dt$, et on pose $F = \{f \in E \mid f(0) = 0\}$. Cet ensemble est un sous-espace

vectoriel de E distinct de E tout entier (c'est même un hyperplan). Pourtant, on peut constater que $F^\perp = \{0\}$. En effet, soit $f \in F^\perp$, posons $g(t) = tf(t)$, alors $g(0) = 0$, donc $f \cdot g = 0$. Or, $f \cdot g = \displaystyle\int_0^1 tf^2(t)\ dt$, intégrale d'une fonction positive qui ne peut donc s'annuler que si $tf^2(t) = 0$ sur tout l'intervalle $[0, 1]$. Cela implique que f soit nulle sur $]0, 1]$, et donc aussi en 0 par continuité. Seule la fonction nulle appartient donc à $F^\perp$.

Un autre exemple extrêmement classique : pour le produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$, les sous-espaces $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont supplémentaires orthogonaux. Soient en effet A une matrice symétrique et B une matrice antisymétrique, alors $A \cdot B = \mathrm{Tr}(A^\top B) = \mathrm{Tr}(AB)$, mais $B \cdot A = \mathrm{Tr}(B^\top A) = \mathrm{Tr}(-BA) = -\mathrm{Tr}(BA)$. Or, on sait que $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, donc la symétrie du produit scalaire impose $\mathrm{Tr}(AB) = -\mathrm{Tr}(AB)$, donc $A \cdot B = 0$. Ce calcul prouve que $(\mathcal{S}_n(\mathbb{R})) \subset (\mathcal{A}_n(\mathbb{R}))^\perp$, et un raisonnement sur les dimensions permet alors de conclure.

2.4 Projections et symétries orthogonales.

Définition 7. Soit E un espace préhilbertien et F un sous-espace vectoriel de dimension finie de E. La **projection orthogonale sur** F est la projection sur F parallèlement à $F^\perp$, la **symétrie orthogonale par rapport à** F est la symétrie par rapport à F parallèlement à $F^\perp$.

Proposition 7. Si $(e_1, \ldots, e_k)$ est une base orthonormale de F, alors le projeté orthogonal de u sur F est égal à $\displaystyle\sum_{i=1}^{k} \langle u, e_k\rangle e_k$.

Démonstration. C'est une conséquence immédiate du calcul effectué plus haut pour prouver que $F + F^\perp = E$ dans le cas d'un sev de dimension finie. □

Remarque 5. Attention, le fait que la base soit orthonormale est essentiel. Si on ne dispose que d'une base quelconque de F, on aura deux possibilités pour calculer les projections orthogonales sur F : orthonormaliser la base à l'aide du procédé de Gram-Schmidt pour pouvoir appliquer la proposition, ou revenir tout simplement à la définition du projeté orthogonal pour s'en sortir sans. Par définition, en notant v le projeté recherché du vecteur u, on doit avoir $v \in F$ et $u - v \in F^\perp$, ce qui fournit suffisamment d'informations pour avoir un système à résoudre.

Exemple : On se place dans l'espace E des fonctions continues sur $[0, 2\pi]$, muni classiquement du produit scalaire intégral $f \cdot g = \displaystyle\int_0^{2\pi} f(t)g(t)\ dt$. On pose dans cet espace $f(t) = t$ et on souhaite calculer le projeté orthogonal de f sur le sous-espace vectoriel F de E engendré par les fonctions cos et sin (qui est certainement de dimension finie). La base $(\cos, \sin)$ n'est pas une base orthonormale de F puisque $\|\cos\| = \sqrt{\displaystyle\int_0^{2\pi} \cos^2(t)\ dt} = \sqrt{\displaystyle\int_0^{2\pi} \frac{\cos(2t)+1}{2}\ dt} = \sqrt{\left[\frac{\sin(2t)}{2} + \frac{t}{2}\right]_0^{2\pi}} = \sqrt{\pi}$. Comme $\displaystyle\int_0^{2\pi} \cos^2(t) + \sin^2(t)\ dt = \int_0^{2\pi} 1\ dt = 2\pi$, on en déduit facilement que $\|\sin\| = \sqrt{\pi}$ également. Par contre, la famille $(\cos, \sin)$ est bien orthogonale puisque $\displaystyle\int_0^{2\pi} \cos(t)\sin(t)\ dt = \int_0^{2\pi} \frac{\sin(2t)}{2}\ dt = \left[-\frac{\cos(2t)}{4}\right]_0^{2\pi} = 0$. Le calcul d'une base orthonormale de F est alors assez rapide puisqu'il suffit de normer les fonctions sin et cos et qu'on a déjà tous les éléments pour le faire. Mais on va plutôt, volontairement, passer par la définition. Notons g le projeté orthogonal de f sur F. Par définition, $g \in F$ donc $g = a\cos + b\sin$. De plus, $f - g$ doit être orthogonale aux deux fonctions cos et sin, ce qui

impose $g \cdot \cos = (a\cos + b\sin)\cdot\cos = a\pi = f\cdot\cos = \displaystyle\int_0^{2\pi} t\cos(t)\ dt = [t\sin(t)]_0^{2\pi} - \int_0^{2\pi} \sin(t)\ dt = 0 - [-\cos(t)]_0^{2\pi} = 0$ (on a effectué une IPP pour le calcul de cette dernière intégrale). Autrement dit $a = 0$. On effectue un calcul similaire avec le sinus : $g\cdot\sin = (b\sin)\cdot\sin = b\pi = f\cdot\sin = \displaystyle\int_0^{2\pi} t\sin(t)\ dt = [-t\cos(t)]_0^{2\pi} + \int_0^{2\pi} \cos(t)\ dt = -2\pi$. En découle $b = -2$, et le projeté recherché est donc défini par $g(t) = -2\sin(t)$.

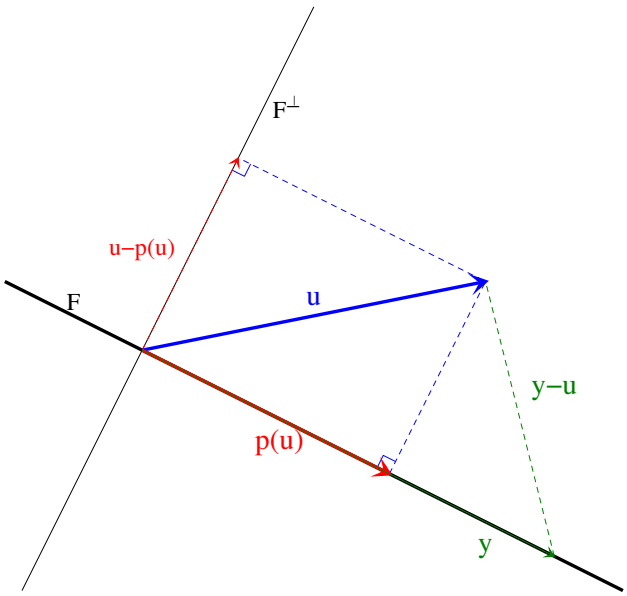
Remarque 6. Si p est la projection orthogonale sur un sous-espace F, alors $\mathrm{id}_E - p$ est la projection orthogonale sur $F^\perp$.

Définition 8. Une symétrie orthogonale par rapport à un hyperplan de E (ici supposé euclidien) est appelée **réflexion**. En notant a un vecteur normal unitaire de l'hyperplan, cette réflexion a pour expression $u \mapsto u - 2\langle u, a\rangle a$, et la projection orthogonale correspondante a pour expression $u \mapsto u - \langle u, a\rangle a$.

Démonstration. C'est une conséquence des propriétés et remarques précédentes : le projeté orthogonal de u sur $H^\perp$ est égal à $\langle u, a\rangle a$, puisque a constitue à lui tout seul une base orthonormale de $H^\perp$. La propriété en découle. □

Définition 9. Soit E un espace préhilbertien et A un sous-ensemble quelconque (non vide) de E. Pour tout vecteur $u \in E$, on appelle **distance de u à A** le réel positif $d(u, A) = \inf_{y\in A} d(x, y) = \inf_{y\in A} \|x - y\|$.

Proposition 8. Si F est un sous-espace vectoriel de dimension finie de E, et $p(u)$ le projeté orthogonal de u sur F, alors $d(u, F) = \|u - p(u)\|$, et ce minimum est atteint uniquement pour le projeté orthogonal. De plus, $d(u, F)^2 = \|u\|^2 - \|p(u)\|^2$.



Démonstration. C'est une conséquence directe du théorème de Pythagore : si $y \in F$ (cf illustration ci-dessus), alors $\|y - u\|^2 = \|u - p(u)\|^2 + \|p(u) - y\|^2$, ce qui prouve que $\|y - u\| \geqslant \|u - p(u)\|$, avec égalité si et seulement si $y = p(u)$. □

Exemple : le résultat précédent est fondamental car il permet de calculer facilement des distances à des sous-espaces vectoriels (il suffit de savoir projeter). Or, il y a énormément de calculs en mathématiques qu'on peut interpréter comme de tels calculs de distances, y compris dans des domaines a priori assez éloignés de la géométrie. On cherche par exemple à déterminer les réels a et b pour lesquels l'intégrale $I = \int_0^{2\pi} (t - a\cos(t) - b\sin(t))^2 \, dt$ est minimale, ainsi que la valeur de cette intégrale. Il s'agit exactement de calculer la distance de la fonction identité à l'espace vectoriel de dimension finie $F = \text{Vect}(\cos, \sin)$ pour le produit scalaire $f \cdot g = \int_0^{2\pi} f(t)g(t) \, dt$ déjà utilisé dans notre exemple précédent. On a déjà calculé le projeté orthogonal correspondant, qui nous assure que le minimum recherché est atteint pour $a = 0$ et $b = -2$. Il ne reste plus qu'à calculer $\| \text{id} \|^2 - \| -2\sin \|^2 = \int_0^{2\pi} t^2 \, dt - \int_0^{2\pi} 4\sin^2(t) \, dt = \frac{8\pi^3}{3} - \int_0^{2\pi} 2 - 2\cos(2t) \, dt = \frac{8\pi^3}{3} - 4\pi$.

3 Isométries d'un espace euclidien.

Cette dernière partie de cours est techniquement hors programme depuis que nos chers décideurs ont décidé de vraiment supprimer tout ce qui ressemblait trop à de la géométrie des programmes de première année. Cette étude permet toutefois de bien mieux comprendre l'intérêt de toutes les constructions faites dans ce chapitre en se penchant sur les applications linéaires « naturellement intéressantes » dans un espace euclidien : celles qui conservent produit scalaire et norme. Dans toute cette partie, E désignera donc un espace vectoriel euclidien, hypothèse qui ne sera pas rappelée par la suite.

3.1 Endomorphismes orthogonaux.

Définition 10. Un endomorphisme $f \in \mathcal{L}(E)$ est appelé **endomorphisme orthogonal** ou **isométrie** si $\forall (u, v) \in E^2$, $f(u) \cdot f(v) = u \cdot v$. On note $\mathcal{O}(E)$ l'ensemble des isométries de E.

Remarque 7. Un endomorphisme est donc orthogonal s'il conserve le produit scalaire, ce qui est une condition naturelle. Attention tout de même au vocabulaire, une symétrie orthogonale est un endomorphisme orthogonal, mais pas une projection orthogonale !

Proposition 9. Un endomorphisme est orthogonal si et seulement s'il conserve la norme : $\forall u \in E$, $\|f(u)\| = \|u\|$.

Démonstration. En effet, si f est orthogonale, on aura toujours $f(u) \cdot f(u) = u \cdot u$, donc $\|f(u)\|^2 = \|u\|^2$. La réciproque découle de l'identité de polarisation. Cette propriété explique le terme d'isométrie, ce sont des applications qui conservent les longueurs. On retrouve bien entendu la définition classique des isométries dans le plan par exemple, que nous avions eu le bonheur d'étudier dans le chapitre consacré aux nombres complexes. Il n'existe pas d'isométries qui ne soient pas des applications linéaires dans E (exercice laissé au lecteur très motivé, ce n'est pas trivial), la restriction aux applications linéaires n'en est donc pas vraiment une. □

Proposition 10. Un endomorphisme orthogonal est nécessairement bijectif. Plus précisément, $\mathcal{O}(E)$ est un sous-groupe de $GL(E)$.

Démonstration. Pour un endomorphisme d'un espace vectoriel de dimension finie, être bijectif ou injectif est équivalent. Or, si $f(u) = 0$, d'après la propriété précédente, on aura $\|u\| = 0$, donc $u = 0$, ce qui prouve l'injectivité et donc la bijectivité de f. De plus, la composée de deux endomorphismes conservant la norme conserve trivialement la norme, et de même pour la réciproque, ce qui prouve que $\mathcal{O}(E)$ est bien un sous-groupe de $GL(E)$. □

Proposition 11. Un endomorphisme f est orthogonal si et seulement s'il vérifie, au choix, l'une des deux conditions suivantes :

- L'image par f de toute base orthonormale de E est une base orthonormale.
- L'image par f d'une base orthonormale fixée de E est une base orthonormale.

Démonstration. Le fait que ces conditions soient nécessaire est évident : puisqu'une isométrie conserve à la fois les normes et les produits scalaires, l'image d'une base orthonormale par une isométrie sera toujours une base orthonormale. La réciproque découle de l'expression du produit scalaire dans une base orthonormale vue plus haut dans ce cours (et de la bilinéarité du produit scalaire), il suffit de décomposer les vecteurs u et v dans la base orthonormale qui a une image orthonormale par f, et de décomposer $f(u)$ et $f(v)$ dans cette même image, pour obtenir le résultat (je vous épargne le calcul, les démonstrations seront réduites au strict minimum dans cette partie de cours facultative). □

Définition 11. Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est **orthogonale** si elle est la matrice représentative d'un endomorphisme orthogonal f dans la base canonique (ou dans toute base orthonormale) de $\mathbb{R}^n$ (ou d'un espace euclidien E de dimension n). On note $\mathcal{O}_n(\mathbb{R})$ l'ensemble de toutes les matrices orthogonales.

Proposition 12. $A \in \mathcal{O}_n(\mathbb{R}) \Leftrightarrow\ A^\top A = I_n$.

Démonstration. Supposons que $A^\top A = I_n$, et notons f l'endomorphisme représenté par A dans une base $(e_1, e_2, \ldots, e_n)$ orthonormale (la base canonique convient très bien). Dire que $A^\top A = I_n$ signifie deux choses : pour chaque colonne de A, la somme des carrés des éléments de la colonne est égale à 1 (c'est le calcul qu'on fait pour obtenir $(A^\top A)_{ii}$), ce qui revient à dire que $f(e_i)$ est un vecteur de norme 1 , et si on effectue le calcul $\sum_{k=1}^{n} a_{ik}a_{jk}$ pour $i \neq j$, on obtient 0, ce qui revient cette fois-ci à dire que $f(e_i)$ et $f(e_j)$ sont orthogonaux. Autrement dit, l'image de notre base orthonormale est une base orthonormale, donc f est orthogonal. La réciproque se fait exactement de la même façon. □

Remarque 8. Attention dans la définition des matrices orthogonales à ne pas oublier qu'on doit se placer dans une base orthonormale. Dans une base quelconque, la matrice d'une application orthogonale peut ressembler à n'importe quoi (d'inversible tout de même puisque f est bijective). Au passage, notre dernière propriété confirme que A est une matrice inversible, et même que son inverse est égale à sa transposée.

Exemple : La matrice $A = \frac{1}{9}\begin{pmatrix} 8 & -1 & -4 \\ -1 & 8 & -4 \\ 4 & 4 & 7 \end{pmatrix}$ est une matrice orthogonale. Pour le vérifier, on peut évidement calculer $A^\top A$, mais on peut aussi utiliser le petit truc suivant : on vérifie que les colonnes « sont de norme 1 » et « orthogonales entre elles » comme expliqué dans la démonstration

ci-dessus. Ici, $\frac{1}{9}\|(8,-1,4)\| = \frac{1}{9}\sqrt{64+1+16} = 1$ et de même pour les deux autres colonnes, et $(8,-1,4)\cdot(-1,8,4) = -8-8+16 = 0$, et de même pour les deux autres produits scalaires de colonnes.

Proposition 13. Une matrice A est orthogonale si et seulement si ses vecteurs-colonnes forment une base orthonormale de E. De même pour les vecteurs-lignes.

Démonstration. On vient de voir que c'est une autre façon de dire exactement la même chose que dans la proposition précédente. Si ça marche pour les colonnes, ça marche pour les lignes, car A est orthogonale si et seulement si $A^\top$ l'est (cela découle de l'égalité $A^\top A = I_n$). □

Proposition 14. La matrice de passage entre deux bases orthonormales est une matrice orthogonale.

Démonstration. En effet, si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases orthonormales de E, il existe une application $f \in \mathcal{L}(E)$ envoyant $\mathcal{B}$ sur $\mathcal{B}'$, et sa matrice dans la base $\mathcal{B}$ est exactement la matrice de passage recherchée. Comme f est orthogonale (elle transforme par définition une base orthonormale en une autre base orthonormale), notre matrice de passage l'est aussi. □

Proposition 15. Si $A \in \mathcal{O}_n(\mathbb{R})$, alors $\det(A) = \pm 1$.

Démonstration. On sait que $A^\top A = I_n$, et que $\det(A^\top) = \det(A)$, donc $\det(A)^2 = \det(I_n) = 1$, la propriété en découle. □

Définition 12. L'ensemble des matrices orthogonales de déterminant 1 est noté $\mathcal{SO}_n(\mathbb{R})$ (groupe spécial orthogonal) ou $\mathcal{O}_n^+(\mathbb{R})$, l'ensemble de celles de déterminant -1 est noté $\mathcal{O}_n^-(\mathbb{R})$ (non, n'insistez pas, il n'y a pas d'équivalent à la notation $\mathcal{SO}_n(\mathbb{R})$ dans ce cas-là).

Définition 13. Une isométrie est **directe** si sa matrice dans une base orthonormale appartient à $\mathcal{SO}_n(\mathbb{R})$, elle est **indirecte** sinon.

Remarque 9. Ces définitions et notations ont bien sûr un lien avec la notion d'orientation de l'espace évoquées dans le chapitre sur les déterminants. Une isométrie directe est une isométrie qui transforme toute base orthonormale en une base de même orientation, et c'est le contraire pour une isométrie indirecte. En particulier, avec l'orientation canonique, une isométrie f est directe (ou indirecte) si et seulement l'image de la base canonique par f est une base orthonormale directe (ou indirecte).

3.2 Isométries du plan.

Théorème 3. Toute matrice $A \in \mathcal{SO}_2(\mathbb{R})$ peut s'écrire sous la forme $A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$, où $\theta \in \mathbb{R}$. L'application f ayant pour matrice A dans n'importe quelle base orthonormale est appelée **rotation d'angle** θ.

Démonstration. Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{SO}_2(\mathbb{R})$. On peut traduire l'égalité $A^\top A = I_2$ par le système d'équations suivant : $\begin{cases} a^2 + b^2 = 1 \\ ac + bd = 0 \\ c^2 + d^2 = 1 \end{cases}$ (deux des équations obtenues sont identiques). On a de plus la condition $\det(A) = ad - bc = 1$. La dernière équation signifie que le point de coordonnées (c, d) dans le plan appartient au cercle trigonométrique (il est à une distance 1 de l'origine), ce qui permet de poser $c = \cos(\theta)$ et $d = \sin(\theta)$, pour un certain réel θ. De même, la première équation permet de poser $a = \cos(\alpha)$ et $b = \sin(\alpha)$. La deuxième condition devient alors $\cos(\alpha)\cos(\theta) + \sin(\alpha)\sin(\theta) = 0$, soit $\cos(\alpha - \theta) = 0$, et celle sur le déterminant donne de même $\sin(\theta - \alpha) = 1$. Autrement dit, on aura $\alpha = \theta + \frac{\pi}{2}[2\pi]$, et donc $\cos(\alpha) = \cos(\theta)$ et $\sin(\alpha) = -\sin(\theta)$, ce qui donne bien la matrice annoncée. □

Remarque 10. L'application n'est évidemment pas appelée rotation par hasard : il s'agit bel et bien d'une rotation d'angle θ autour de l'origine. On vient en fait de prouver que les seules isométries directes dans $\mathbb{R}^2$ sont les rotations (on parle ici d'isométries vectorielles, qui doivent laisser fixe l'origine O, ce qui explique l'absence des translations). La matrice d'une rotation dans le plan est indépendante de la base orthonormale choisie. Bien évidemment, dans une base qui n'est pas orthonormale, la matrice peut changer.

Définition 14. Un **retournement** est une rotation plane d'angle $\theta = \pi$.

Remarque 11. C'est ce que vous avez appelé pendant des années une symétrie centrale, terme que nous n'utiliserons plus jamais même si l'application f est de fait dans ce cas une symétrie (on obtient l'identité si on la compose avec elle-même).

Théorème 4. Toute matrice $A \in \mathcal{O}_2^-(\mathbb{R})$ peut s'écrire sous la forme $A = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$. L'application u ayant pour matrice A dans une base orthonormale est une réflexion.

Démonstration. La preuve que la matrice peut se mettre sous cette forme est identique à celle vue pour les isométries directes, à un changement de signe près à un endroit, nous nous épargnerons les calculs. Il est facile de constater que u est une symétrie (nécessairement orthogonale) en calculant $A^2 = \begin{pmatrix} \cos^2(\theta) + \sin^2(\theta) & 0 \\ 0 & \sin^2(\theta) + \cos^2(\theta) \end{pmatrix} = I_2$. C'est nécessairement une symétrie par rapport à une droite vectorielle, donc une réflexion. En effet, si le sous-espace par rapport auquel on symétrise n'est pas une droite, il s'agit soit de $\{0\}$ et alors $u = -\mathrm{id}$, soit de $\mathbb{R}^2$ et $u = \mathrm{id}$. Dans les deux cas, u ne serait pas une isométrie indirecte. □

Remarque 12. Dans le cas des réflexions, la matrice de u n'est pas du tout la même dans toutes les bases orthonormales. Par ailleurs, l'angle θ est beaucoup moins facile à interpréter géométriquement que dans le cas d'une rotation.

Remarque 13. Toute rotation dans le plan peut s'écrire comme composée de deux réflexions. En effet, $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \times \begin{pmatrix} \cos(-\theta) & \sin(-\theta) \\ \sin(-\theta) & -\cos(-\theta) \end{pmatrix}$.

3.3 Isométries de l'espace.

Définition 15. Tout élément de $\mathcal{SO}_3(\mathbb{R})$ est appelé **matrice de rotation** dans l'espace.

Théorème 5. Soit f une isométrie directe de l'espace, alors $F = \ker(f - \mathrm{id})$ est de dimension 1. Si (v, w) est une base orthonormale du plan orthogonal à F et a un vecteur directeur unitaire de F tel que (v, w, a) soit une base directe (par exemple $a = v \wedge w$), alors la matrice de f dans la base orthonormale $(v, w, v \wedge w)$ est $A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Définition 16. Une isométrie directe ayant pour matrice A dans une base orthonormale est appelée **rotation d'axe F et d'angle θ.**

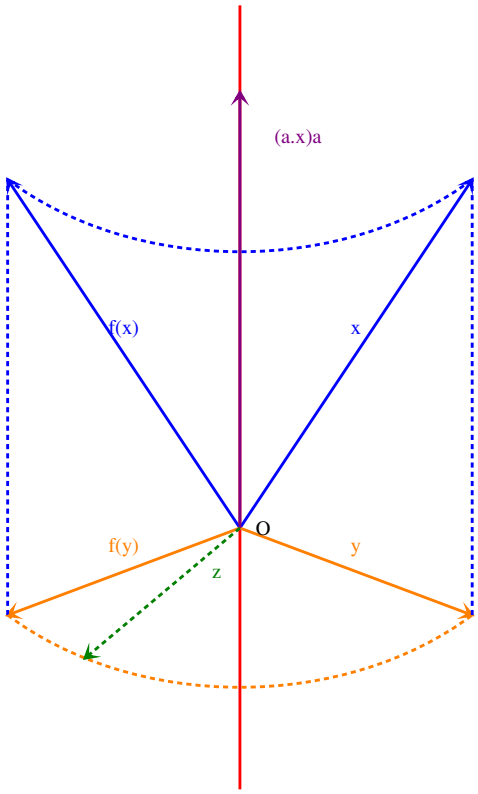
Remarque 14. Une rotation spatiale reste assez facile à visualiser : l'axe ne bouge pas, et le plan orthogonal à l'axe subit une rotation d'angle θ (autrement dit, on tourne d'un angle θ autour de l'axe). Voir le schéma plus bas. Notons quand même que la définition est un peu douteuse, car l'angle de la rotation dépend du vecteur choisi pour diriger l'axe. En effet, si on prend le vecteur opposé, l'angle va être également changé en son opposé.

Proposition 16. La composée de deux rotations dans l'espace est une rotation.

Démonstration. En effet, comme dans le plan, la composée reste une isométrie directe, donc une rotation. Mais pour le coup, ça n'a rien de géométriquement évident (et en particulier, l'angle de la rotation composée n'est pas évident à déterminer à partir de ceux des deux rotations). □

Proposition 17. Soit f la rotation d'axe $F = \mathrm{Vect}(a)$, avec a unitaire, et d'angle θ, alors $\forall u \in \mathbb{R}^3$, $f(u) = \cos(\theta)u + \sin(\theta)a \wedge u + (1 - \cos(\theta))(a \cdot u)a$.

Remarque 15. Dans le cas particulier où u appartient au plan orthogonal à F, on trouve plus simplement $f(u) = \cos(\theta)u + \sin(\theta)(a \wedge u)$, on peut donc retrouver facilement l'angle de la rotation, quasiment comme dans le cas d'une rotation plane, à l'aide des relations $u \cdot f(u) = \cos(\theta)$ et $\det(u, f(u), a) = \sin(\theta)$ (en choisissant un vecteur u unitaire). Sur la figure ci-dessous, l'axe de la rotation est en rouge, le plan orthogonal en marron, le vecteur noté z est $y \wedge a$.



Démonstration. La projection de u sur l'axe de la rotation est simplement donnée par $(a\cdot u)a$. Posons $y = u-(a\cdot u)a$, alors $(y, y\wedge a)$ est une base orthogonale directe du plan orthogonal à F constituée de deux vecteurs de même norme. De plus, $y\wedge a = (u-(a\cdot u)a)\wedge a = u\wedge a$ puisque $a\wedge a = 0$. On en déduit que $f(y) = \cos(\theta)y-\sin(\theta)(u\wedge a)$, donc $f(u) = f(y+(a\cdot u)a) = \cos(\theta)y+\sin(\theta)(a\wedge u)+(a\cdot u)a$ (puisque a est laissé fixe par la rotation). Autrement dit, $f(u) = \cos(\theta)u-\cos(\theta)(a\cdot u)a+\sin(\theta)(a\wedge u)+(a\cdot u)a$, ce qui est bien la formule donnée. □

Exemple : Soit f l'endomorphisme de $\mathbb{R}^3$ représenté dans la base canonique par la matrice $M = \frac{1}{3}\begin{pmatrix} 2 & 2 & 1 \\ -2 & 1 & 2 \\ 1 & -2 & 2 \end{pmatrix}$. La matrice est orthogonale puisque $\frac{1}{3}\|(2,2,1)\| = \frac{1}{3}\sqrt{4+4+1} = 1$, $\frac{1}{3}\|(-2,1,2)\| = 1$, $\frac{1}{3}\|(1,-2,2)\| = 1$, $(2,2,1)\cdot(-2,1,2) = -4+2+2 = 0$, $(2,2,1)\cdot(1,-2,2) = 0$ et $(-2,1,2)\cdot(1,-2,2) = 0$. Il s'agit donc de la matrice d'une isométrie.

On calcule ensuite le déterminant pour déterminer si l'isométrie est directe (dans le calcul qui suit, on effectue l'opération $L_3 \leftarrow L_3 - L_2$ puis on développe par rapport à la dernière ligne) : $\begin{vmatrix} 2 & 2 & 1 \\ -2 & 1 & 2 \\ 1 & -2 & 2 \end{vmatrix} = \begin{vmatrix} 2 & 2 & 1 \\ -2 & 1 & 2 \\ 3 & -3 & 0 \end{vmatrix} = 3\times\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} + 3\times\begin{vmatrix} 2 & 1 \\ -2 & 2 \end{vmatrix} = 3\times 3 + 3\times 6 = 27$, donc $\det(A) = \frac{1}{3^3}\times 27 = 1$. Il s'agit d'une isométrie directe, donc d'une rotation.

On cherche ensuite l'axe de la rotation, en déterminant simplement $\ker(f-\mathrm{id})$. Quitte à tout multiplier par 3, on doit résoudre le système $\begin{cases} 2x & + & 2y & + & z & = & 3x \\ -2x & + & y & + & 2z & = & 3y \\ x & - & 2y & + & 2z & = & 3z \end{cases}$. La deuxième équation donne $z = x+y$, et la dernière donne $2y = x-z$, soit $2y = -y$. Manifestement, cela implique $y = 0$, puis $z = x$. La première équation donne alors $3x = 3x$, donc $\ker(f-\mathrm{id}) = \mathrm{Vect}((1,0,1))$, qui est bien une droite F.

On choisit maintenant un vecteur unitaire v orthogonal à F. Ce n'est pas bien compliqué ici, il suffit de prendre $v = (0,1,0)$. On peut alors calculer, en notant θ l'angle de la rotation, $\cos(\theta) = v \cdot f(v)$. Puisque $f(0,1,0) = \frac{1}{3}(2,1,-2)$, on trouve $\cos(\theta) = \frac{1}{3}$. En posant $a = \frac{1}{\sqrt{2}}(1,0,1)$ (vecteur directeur unitaire de l'axe), on peut ensuite calculer $\det(v, f(v), a) = \frac{1}{3\sqrt{2}} \begin{vmatrix} 0 & 1 & 0 \\ 2 & 1 & -2 \\ 1 & 0 & 1 \end{vmatrix} = -\frac{1}{3\sqrt{2}} \begin{vmatrix} 2 & -2 \\ 1 & 1 \end{vmatrix} = -\frac{2\sqrt{2}}{3}$. On en déduit que $\sin(\theta) = -\frac{2\sqrt{2}}{3}$, donc $\theta = -\arccos\left(\frac{1}{3}\right)$. Seul le signe de $\sin(\theta)$ était nécessaire pour conclure, mais l'avantage de l'avoir calculé explicitement est de pouvoir vérifier que $\cos^2(\theta) + \sin^2(\theta) = \frac{1}{9} + \frac{8}{9} = 1$. On peut en tout cas désormais affirmer que f est la rotation d'axe $\mathrm{Vect}((1,0,1))$ et d'angle $-\arccos\left(\frac{1}{3}\right)$.

Remarque 16. Un autre moyen de vérifier la cohérence des calculs est d'utiliser la trace de la matrice. En effet, celle-ci est invariante par changement de base, donc est égale à $2\cos(\theta)+1$ dans n'importe quelle base. Ici, on pouvait donc calculer dès le départ $\mathrm{Tr}(A) = \frac{5}{3} = 2\cos(\theta)+1$, et en déduire que $\cos(\theta) = \frac{1}{3}$.

Théorème 6. Si f est une réflexion dans l'espace, elle a pour matrice dans une base orthonormale bien choisie $A = \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ \sin(\theta) & -\cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$. De plus, toute isométrie indirecte de l'espace est la composée d'une rotation (éventuellement égale à id) et d'une réflexion.

Remarque 17. On peut toujours écrire dans l'espace une rotation comme composée de deux réflexions (c'est exactement le même calcul que dans le plan, en ajoutant un 1 en bas à droite de la matrice et des 0 ailleurs sur la dernière ligne et la dernière colonne). Il existe donc quatre types d'isométries vectorielles dans l'espace :

- l'identité, qui laisse tout l'espace fixe.
- les réflexions, qui laissent tout un plan fixe et sont indirectes.
- les composées de deux réflexions, qui sont des rotations, laissent une droite fixe et sont directes.
- les composées de trois réflexions ne laissent que le vecteur nul invariant et sont indirectes.

Ainsi, dans l'espace, $-\mathrm{id}$ est une isométrie indirecte qui est une composée de trois réflexions (par exemple par rapport aux trois plans de coordonnées du repère canonique). Ce n'est absolument pas la rotation d'angle π autour de l'origine, cette application n'étant pas une rotation avec la définition que nous avons prise. Pour les plus curieux, le théorème précédent se généralise en dimension n, où les isométries peuvent toujours être écrites comme composées d'au plus n réflexions par rapport à des hyperplans (et laissant stable un sous-espace vectoriel de dimension $n-k$, où k est le nombre de réflexions composées pour obtenir l'isométrie).