Chapitre 11

Produits scalaires et espaces vectoriels euclidiens

11.1 Définitions fondamentales.

On considère un espace vectoriel réel V sur le corps des réels.

Définitions. Un *produit scalaire (généralisé)* sur V est une application

$$g : V \times V \to \mathbb{R}$$

qui est *bilinéaire, symétrique et définie-positive* :

(i.) g est bilinéaire.

(ii.) g est symétrique, c'est-à-dire $g(x, y) = g(y, x)$ pour tous $x, y \in V$.

(iii.) g est *positive*, c'est-à-dire $g(x, x) \geq 0$ pour tout $x \in V$.

(iv.) g est *définie*, c'est-à-dire $g(x, x) = 0 \Leftrightarrow x = 0$.

Un *espace euclidien* est un espace vectoriel réel de dimension finie muni d'un produit scalaire.

Remarque. Lorsqu'on s'est donné un produit scalaire g sur V, on note souvent $\langle x, y\rangle_g = g(x, y)$ (ou simplement $\langle x, y\rangle$ s'il n'y a pas de risque d'ambiguïté).

Exemples 1. Le *produit scalaire standard* sur $\mathbb{R}^n$ est défini par

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

2. Sur l'espace $M_n(\mathbb{R})$ on a un produit scalaire défini par

$$\langle A, B\rangle = \mathrm{Trace}(A^\top \cdot B).$$

3. On note ℓ_2 l'espace vectoriel des suites réelles de carré sommable

$$\ell_2 = \{\xi = (x_k)_{k=1}^{\infty} \mid x_k \in \mathbb{R},\ \sum_{k=1}^{\infty} |x_k|^2 < \infty\}.$$

Un produit scalaire naturel sur cet espace est défini par

$$(\xi \mid \eta)_{\ell_2} = \sum_{i=1}^{\infty} x_i y_i.$$

Observons qu'il s'agit d'une généralisation en dimension infinie du produit scalaire standard de $\mathbb{R}^n$.

4. Un produit scalaire naturel est défini sur l'espace vectoriel $C^0([a,b])$ des fonctions continues sur l'intervalle $[a,b]$ par l'intégration :

$$(f \mid h)_{L^2} = \int_a^b f(x)h(x)dx,$$

on l'appelle [1] le “produit scalaire L^2”.

Définition 11.1.1. Si g est un produit scalaire sur V, on définit la *norme* d'un vecteur $x \in V$ associée à g par :

$$\|x\|_g = \sqrt{g(x,x)} = \sqrt{\langle x, x\rangle}.$$

La norme est bien définie car $\langle x, x\rangle \geq 0$ pour tout x. Le résultat suivant est une propriété fondamentale des produits scalaires.

Proposition 11.1.2 (Inégalité de Cauchy-Schwarz.)**.** *Soit V est un espace vectoriel réel muni d'un produit scalaire $\langle\ ,\ \rangle$. Pour tous $x, y \in V$, on a*

$$|\langle x, y\rangle| \leq \|x\|\|y\|.$$

De plus il y a égalité si et seulement si x et y sont colinaires.

Preuve. On note $p(t) = \|xt + y\|^2$ et on observe qu'il s'agit un polynôme à coefficients réel de degré 2. Plus précisément on a

$$\begin{aligned} p(t) &= \|tx + y\|^2 \\ &= \langle tx + y, tx + y\rangle = t^2\langle x, x\rangle + t\langle x, y\rangle + t\langle y, x\rangle + \langle y, y\rangle \\ &= \|x\|^2 t^2 + 2\langle x, y\rangle t + \|y\|^2. \end{aligned}$$

Il est d'autre part que $p(t) \geq 0$ pour tout $t \in \mathbb{R}$, le polynôme $p(t)$ a donc au plus une racine réelle, ce qui implique que son discriminant Δ est ≤ 0. On a donc

$$\Delta = \left(\langle x, y\rangle^2 - \|x\|^2\|y\|^2\right) \leq 0,$$

c'est-à-dire $|\langle x, y\rangle| \leq \|x\|\|y\|$. De plus on a égalité si et seulement si $\Delta = 0$. Dans ce cas il existe $t \in \mathbb{R}$ tel que $p(t) = 0$, ce qui signifie que $y = -tx$.

$\square$

1. Cette terminologie est justifiée par le fait que ce produit scalaire peut être défini sur l'espace des fonctions de *carré intégrable au sens de Lebesgue*. La lettre L fait référence à Lebesgue et l'exposant 2 à la condition d'intégrabilité du carré de f. Il s'agit donc de l'espace vectoriel des fonctions $f : [a,b] \to \mathbb{R}$ qui vérifient $\int_a^b f(x)^2 dx < \infty$, l'intégrale étant prise au sens de Lebesgue (qui est plus générale que la notion d'intégrabilité au sens de Riemann).

Proposition 11.1.3. *La norme vérifie les propriétés suivantes pour tous $x, y \in V$ et $\lambda \in \mathbb{R}$:*
(a) $\|x\| \geq 0$ *et* $\|x\| = 0$ *si et seulement si* $x = 0$.
(b) $\|\lambda x\| = |\lambda|\|x\|$.
(c) $\|x + y\| \leq \|x\| + \|y\|$.

Preuve. Les deux premières propriétés suivent facilement des définitions. La troisième propriété est une conséquence de l'inégalité de Cauchy-Schwarz. On a en effet

$$\|x + y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 .$$

Comme les normes de x, y et $x + y$ sont positives ou nulles, on peut prendre la racine carrée dans l'inégalité ci-dessus, ce qui nous donne $\|x + y\| \leq \|x\| + \|y\|$.

□

Remarquons qu'on a également $\|x - y\| \leq \|x\| + \|y\|$, car $\| - y\| = \|y\|$ par la première propriété, et donc

$$\|x - y\| = \|x + (-y)\| \leq \|x\| + \| - y\| = \|x\| + \|y\|.$$

Définition. Si g est un produit scalaire sur l'espace vectoriel réel V, on définit :

(1.) La *distance* entre deux éléments x et y de V est

$$d(x, y) = \|y - x\|.$$

(2.) L'*angle* $\alpha \in [0, \pi]$ entre deux vecteurs non nuls $x, y \in V$ est défini par

$$\cos(\alpha) = \frac{\langle x, y\rangle}{\|x\|\|y\|}.$$

Cette notion est bien définie car d'une part $\|x\|\|y\| \neq 0$ lorsque x et y sont non nuls et d'autre part on a

$$-1 \leq \frac{\langle x, y\rangle}{\|x\|\|y\|} \leq +1$$

par l'inégalité de Cauchy-Schwarz. Notons que le produit scalaire est parfois défini géométriquement à partir de la notion d'angle via la formule

$$\langle x, y\rangle = \|x\|\|y\| \cos(\alpha),$$

mais du point de vue de l'algèbre linéaire, c'est le produit scalaire qui est la notion de base et l'angle est une notion dérivée, et non l'inverse.

(3.) L'*aire* du parallélogramme $\mathcal{P}(x, y)$ construit sur les vecteurs x et y est définie par

$$\text{Aire}(x, y) = \sqrt{\|x\|^2\|y\|^2 - \langle x, y\rangle^2}.$$

A nouveau, l'inégalité de Cauchy-Schwarz justifie aussi que $\text{Aire}(x, y)$ est bien définie. On vérifie facilement que

$$\text{Aire}(x, y) = \|x\|\|y\| \sin(\alpha).$$

Proposition 11.1.4. *Soit E un espace vectoriel euclidien. Alors pour tous $x, y, z \in E$ on a*

(i.) $d(x, z) \leq d(x, y) + d(y, z)$ (inégalité du triangle).

(ii.) Si x et y sont non nuls, alors l'angle θ entre x et y est égal à $\pi/2$ si et seulement si

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 \quad \text{(théorème de Pythagore).}$$

Preuve. (i) En utilisant la proposition 11.1.3 (c) on voit que

$$d(x, z) = \|z - x\| = \|(z - y) + (y - x)\| \leq \|(z - y)\| + \|(y - x)\| = d(x, y) + d(y, z).$$

(ii) Le théorème de Pythagore est une conséquence de la bilinéarité du produit scalaire et de la définition de l'angle. En effet on a d'une part

$$\theta = \pi/2 \quad \Leftrightarrow \quad \cos(\theta) = 0 \quad \Leftrightarrow \quad \langle x, y \rangle = 0.$$

D'autre part

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

□

Proposition 11.1.5. *Soient a, b deux vecteurs non nuls d'un espace vectoriel euclidien E. Alors il existe deux vecteurs c et d tels que a et d forment un angle de $\pi/2$, c est colinéaire à a, et $b = c + d$.*

Preuve. On cherche c et d sous la forme $c = \lambda a$ et $d = b - c$. On demande que a et d forment un angle droit, on a donc

$$0 = \langle d, a \rangle = \langle b - c, a \rangle = \langle b - \lambda a, a \rangle = \langle b, a \rangle - \lambda \langle a, a \rangle.$$

Par conséquent :

$$\lambda = \frac{\langle b, a \rangle}{\langle a, a \rangle} = \frac{\langle b, a \rangle}{\|a\|^2},$$

et donc

$$c = \frac{\langle b, a \rangle}{\|a\|^2} a \quad \text{et} \quad d = b - \frac{\langle b, a \rangle}{\|a\|^2} a.$$

□

Observons que dans cette décomposition $b = c + d$, le vecteur c représente la composante de b en direction de a et d représente la composante de b normale à a.

11.2 Orthogonalité dans un espace vectoriel euclidien

Définitions. 1. Deux vecteurs $x, y \in E$ sont dit *orthogonaux* si $\langle x, y \rangle = 0$. On note cette relation $x \perp y$.

2. Deux sous-espaces vectoriels $W_1, W_2 \subset E$ sont dit *orthogonaux* si $x \perp y$ pour tout $x \in W_1$ et tout $y \in W_2$. On note cette relation $W_1 \perp W_2$.

3. Une base $\{v_1, \ldots v_n\}$ de E est dite *orthogonale* si $v_i \perp v_j$ pour tous $1 \leq i, j \leq n$ tels que $i \neq j$.

4. Une base $\{v_1, \ldots v_n\}$ de E est dite *orthonormée* si elle est orthogonale et si $\|v_i\| = 1$ pour tout i.

Lemme 11.2.1. *Soit (E, g) un espace vectoriel euclidien et $\{v_1, \dots v_n\}$ une base de E. Les conditions suivantes sont équivalentes :*

(a) $\{v_1, \dots, v_n\}$ est une base orthonormée de E.

(b) $g(v_i, v_j) = \langle v_i, v_j \rangle = \delta_{ij}$.

(c) La matrice de Gram G de g dans cette base est la matrice identité.

Ce lemme ne fait que traduire les définitions.

□

Théorème 11.2.2. *Tout espace vectoriel euclidien E admet des bases orthonormées.*

Preuve. Le preuve se fait par récurrence sur $n = \dim(E)$. Si $n = 1$ il suffit de choisir un vecteur $w \in E$ non nul. Alors $e = \frac{w}{\|w\|}$ est une base de E.
Admettons le théorème démontré pour $n - 1$ et supposons que $\dim(E) = n$. Choisissons de nouveau un vecteur $w \in E$ non nul et définissons un covecteur $\theta \in E^*$ par

$$\theta(x) = \langle w, x \rangle.$$

Alors $\theta : E \to \mathbb{R}$ est une application linéaire surjective (car $\theta(w) = \|w\|^2 \neq 0$) donc, par le théorème du rang, on a $\dim \mathrm{Ker}(\theta) = n - 1$. Notons ce sous-espace

$$E_1 = w^\perp = \{x \in E \mid \langle w, x \rangle = 0\} = \mathrm{Ker}(\theta) \subset E.$$

Par hypothèse de récurrence, il existe une base orthonormée de E_1. Notons $\{e_1, \dots, e_{n-1}\} \subset E_1$ cette base et $e_n = \frac{w}{\|w\|}$. Alors on a

$$\langle e_i, e_j \rangle = \delta_{ij}, \quad \text{pour } 1 \leq i, j \leq n.$$

par conséquent $\{e_1, \dots, e_n\}$ est une base orthonormée de E. Le théorème est démontré.

□

Remarque. Si $\{e_1, \dots, e_n\}$ est une base orthonormée de l'espace vectoriel euclidien E, et si $x, y \in E$ sont des vecteurs de composantes x_i, y_j dans cette base, alors on a

$$\langle x, y \rangle = \langle \sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j \rangle = \sum_{i,j=1}^n x_i y_j \langle e_i, e_j \rangle = \sum_{i,j=1}^n x_i y_j \delta_{ij} = \sum_{i=1}^n x_i y_i.$$

Ainsi dans une base orthonormée, le produit scalaire se calcule de la même manière que le produit scalaire standard de $\mathbb{R}^n$.

11.2.1 Projections orthogonales sur un sous-espace vectoriel

Soit V un espace vectoriel euclidien. dont on note $\langle \cdot, \cdot \rangle$ le produit scalaire. Tout sous-espace vectoriel $W \subset V$ est alors lui-même un espace euclidien pour le même produit scalaire restreint à W. En particulier W possède des bases orthonormées. Le théorème suivant nous permet de construire la *projection orthogonale* de V sur W.

Théorème 11.2.3. *Soit $\{w_1, \ldots, w_m\}$ une base orthonormée du sous-espace vectoriel W de l'espace vectoriel euclidien V. On note $P_W : V \to V$ l'application définie par*

$$\boxed{P_W(x) = \sum_{i=1}^{m} \langle w_i, x\rangle w_i.} \tag{11.1}$$

Cette application possède les propriétés suivantes :

(i) P_W *est linéaire.*

(ii) $P_W(x) = x$ *si et seulement si* $x \in W$.

(iii) $W = \mathrm{Im}(P_W)$.

(iv) $P_W \circ P_W = P_W$.

(v) Le noyau de P_W est l'ensemble des vecteurs de V qui sont orthogonaux à tous les vecteurs de W. On note

$$W^{\perp} = \mathrm{Ker}(P_W) = \{v \in V \mid \langle v, w\rangle = 0 \ \forall w \in W\}.$$

(vi) W et $W^{\perp}$ sont supplémentaires dans V, i.e. $V = W \oplus W^{\perp}$.

Définitions. On dit qu'un sous-espace vectoriel W d'un espace vectoriel euclidien V est la *somme directe orthogonale* de W_1 et W_2 si $W_1, W_2 \subset V$ sont des sous-espaces vectoriels tels que $W = W_1 \oplus W_2$ et $W_1 \perp W_2$. Dans ce cas on note

$$W = W_1 \boxplus W_2.$$

La proposition précédente nous dit en particulier que pour tout sous-espace vectoriel $W \subset V$ on a

$$V = W \boxplus W^{\perp}.$$

Preuve du théorème. (i) La linéarité de P_W découle de la linéarité de l'application $x \mapsto \langle w_i, x\rangle$.
(ii) Supposons $x \in W$, alors on peut s'écrire $x = \sum_{j=1}^{m} x_j w_j$ et donc

$$\langle w_i, x\rangle = \langle w_i, \sum_{j=1}^{m} x_j w_j\rangle = x_i.$$

Par conséquent

$$P_W(x) = \sum_{i=1}^{m} x_i w_i = x.$$

Inversément, si $y \notin W$, alors on a $y \neq P_W(y)$ (car clairement $P_W(y) \in W$).
(iii) Il est clair par construction que $\mathrm{Im}(P_W) \subset W$. D'autre part $W \subset \mathrm{Im}(P_W)$ par la condition précédente, car tout $x \in W$ vérifie $x = P_W(x) \in \mathrm{Im}(P_W)$.
(iv) Cette condition découle immédiatement de (ii) car pour tout x on a $P_W(x) \in W$, donc

$$P_W^2(x) = P_W(P_W(x)) = P_W(x).$$

(v) Montrons d'abord que $W^{\perp} \subset \mathrm{Ker}(P_W)$. Soit donc $x \in W^{\perp}$, alors $\langle w, x\rangle = 0$ pour tout $w \in W$, en particulier $\langle w_i, x\rangle = 0$ pour tout $i = 1, \ldots, m$, et donc $P_W(x) = \sum_{i=1}^{m} \langle w_i, x\rangle w_i = 0$.

Montrons maintenant l'inclusion réciproque. Supposons $x \in \mathrm{Ker}(P_W)$, alors

$$P_W(x) = \sum_{i=1}^{m} \langle w_i, x\rangle w_i = 0,$$

et donc $\langle w_i, x\rangle = 0$ pour tout $i = 1, \dots, m$ car les vecteurs w_i sont linéairement indépendants. Pour montrer que $x \in W^{\perp}$ on doit prouver que $\langle w_i, x\rangle = 0$ pour tout $w \in W$. Mais si $w \in W$, alors on peut écrire $w = \sum_{i=1}^{m} \lambda_i w_i$, et on a donc

$$\langle w, x\rangle = \langle \sum_{i=1}^{m} \lambda_i w_i, x\rangle = \sum_{i=1}^{m} \lambda_i \langle w_i, x\rangle = 0.$$

(vi) On remarque tout d'abord que $W \cap W^{\perp} = \{0\}$. En effet, si $x \in W \cap W^{\perp}$, alors $x \perp x$, c'est-à-dire $\|x\|^2 = \langle x, x\rangle = 0$ ce qui implique que $x = 0$ car tout produit scalaire est défini positif. Il nous reste à prouver que $V = W + W^{\perp}$. Or pour tout vecteur $x \in V$ on a $(x - P_W(x)) \in \mathrm{Ker}(P_W)$ car $P_W(x - P_W(x)) = P_W(x) - P_W^2(x) = P_W(x) - P_W(x) = 0$. On a donc

$$x = P_W(x) + \underbrace{(x - P_W(x))}_{\in \mathrm{Ker}(P_W) = W^{\perp}} \in W + W^{\perp}.$$

□

Remarques.

(i) L'application P_W ne dépend que du sous-espace $W \subset V$ et pas du choix de la base orthonormée $\{w_1, \dots, w_m\} \subset W$. Cette application s'appelle la *projection orthogonale de V sur W*. On dit aussi que P_W est un *projecteur orthogonal.*

(ii) Si on note $P_{W^{\perp}}$ la projection sur $W^{\perp}$, alors on a

$$P_W + P_{W^{\perp}} = \mathrm{Id}_V, \quad P_W \circ P_{W^{\perp}} = P_{W^{\perp}} \circ P_W = 0.$$

(iii) Si $\{w_1, \dots, w_m\} \subset W$ est une base de W et $\{w_{m+1}, \dots, w_n\} \subset W^{\perp}$ est une base de $W^{\perp}$, alors $\{w_1, \dots, w_n\}$ est une base de V et la matrice de P_W dans cette base est

$$M(P_W) = I_m \oplus 0_{n-m} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

Un autre raisonnement pour obtenir cette matrice est le suivant : La relation $P_W^2 = P_W$ nous dit que le polynôme minimal de P_W est $\mu(t) = t^2 - t = t(t-1)$, il est scindé à racine simple donc P_W est diagonalisable et les multiplicités géométriques des valeurs propres sont m pour $\lambda = 1$ et $(n-m)$ pour $\lambda = 0$.

Proposition 11.2.4. *Soit V un espace vectoriel euclidien et W un sous-espace vectoriel. Alors pour tout $x \in V$, le point $P_W(x)$ est le point de W le plus proche de x. Plus précisément, si on note $x' = P_W(x)$, alors $x' \in W$ et*

$$\|x - x'\| \leq \|x - y\| \quad \text{pour tout } y \in W,$$

avec égalité si et seulement si $y = x'$.

Preuve. Observons que $P_W(x-x') = P_W(x) - P_W(P_W(x)) = 0$, donc $(x-x') \in \mathrm{Ker}(P_W) = W^{\perp}$. On a donc pour tout $y \in W$, en utilisant le théorème de Pythagore

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

On a donc $\|x-y\|^2 \geq \|x-x'\|^2$, avec égalité si et seulement si $\|x'-y\| = 0$, c'est-à-dire si $y = x'$. $\square$

Corollaire 11.2.5. *La distance d'un point x d'un espace vectoriel euclidien E à un sous-espace vectoriel $W \subset E$ est donnée par*

$$\mathrm{dist}(x, W) = \|x - P_W(x)\| = \left\| x - \sum_{i=1}^{m} \langle w_i, x \rangle w_i \right\|,$$

où $\{w_1, \dots, w_m\}$ une base orthonormée de W.

11.2.2 Symétries orthogonales

Soit V une espace vectoriel euclidien et $W \subset V$ est un sous-espace vectoriel,

Définition. On appelle *symétrie orthogonale* à travers W l'endomorphisme $S_W : V \to V$ défini par

$$S_W = 2P_W - \mathrm{Id}_V. \tag{11.2}$$

Si $\{w_1, \dots, w_m\}$ est une base orthonormée de W, alors on peut écrire explicitement

$$S_W(x) = -x + 2\sum_{i=1}^{m} \langle w_i, x \rangle w_i. \tag{11.3}$$

Le théorème 11.2.3 implique le corollaire suivant dont la preuve est très simple :

Corollaire 11.2.6. *La symétrie orthogonale S_W possède les propriétés suivantes :*

(i) *S_W est linéaire.*

(ii) *$S_W(x) = x$ pour tout $x \in W$ et $S_W(y) = -y$ pour tout $y \in W^{\perp}$.*

(iii) *$S_W^2 = \mathrm{Id}_V$, en particulier S_W est inversible et égale à son propre inverse.*

$\square$

Remarquons que la décomposition de V en somme orthogonale $V = W \boxplus W^{\perp}$ signifie que tout vecteur $v \in V$ s'écrit d'une manière unique $v = x + y$ avec $x \in W$ et $y \in W^{\perp}$. On a alors

$$S_W(v) = S_W(x + y) = x - y,$$

En effet, si $x \in W$ et $y \in W^\perp$, alors

$$(2P_W - \mathrm{Id}_V)(x+y) = 2P_W(x+y) - \mathrm{Id}_V(x+y) = 2x - (x+y) = x - y.$$

Si $\{w_1, \ldots, w_n\}$ est une base de V telle que $\{w_1, \ldots, w_m\} \subset W$ et $\{w_{m+1}, \ldots, w_n\} \subset W^\perp$ alors la matrice de S_W dans cette base est

$$M(S_W) = I_m \oplus (-I_{n-m}) = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{pmatrix}.$$

11.3 Le procédé d'orthonormalisation de Gram-Schmidt

La proposition suivante nous donne une méthode explicite pour construire une base orthonormée d'un espace euclidien.

Proposition 11.3.1. *Soit $\{v_1, \ldots, v_m\}$ des vecteurs linéairement indépendants d'un espace vectoriel euclidien V. Alors il existe des vecteurs $\{u_1, \ldots, u_m\}$ tels que*

(i) $\{u_1, \ldots, u_m\}$ est un système de vecteurs orthonormé, i.e. $\langle u_i, u_j \rangle = \delta_{ij}$ pour tous $i, j \in \{1, \ldots, m\}$.

(ii) Pour tout $k = 1, ..., m$ on a

$$u_k \in \mathrm{Vec}(\{v_1, \ldots, v_k\}), \quad \textit{c'est à dire } u_k \textit{ est combinaison linéaire de } v_1, \ldots, v_k.$$

(iii) $\langle u_i, v_i \rangle > 0$ pour tout $i = 1, \ldots, m$.

De plus cette famille $\{u_1, \ldots, u_m\}$ est unique et la construction est algorithmique.

Preuve. Le premier vecteur u_1 doit être un multiple positif de v_1 et on doit avoir $\|u_1\| = 1$. On a donc

$$u_1 = \frac{v_1}{\|v_1\|}.$$

Supposons qu'on a construit les vecteurs $u_1, \ldots, u_{k-1}$, et notons

$$W_{k-1} = \mathrm{Vec}(\{v_1, \ldots, v_{k-1}\}) = \mathrm{Vec}(\{u_1, \ldots, u_{k-1}\}).$$

On note alors

$$\widehat{v}_k = v_k - P_{W_{k-1}}(v_k) = v_k - \sum_{i=1}^{k-1} \langle u_i, v_k \rangle u_i,$$

On vérifie facilement les propriétés suivantes :

(i) $\widehat{v}_k \perp W_{k-1}$.

(ii) $\{u_1, \ldots u_{k-1}, \widehat{v}_k\}$ est une famille libre de V, en particulier $\|\widehat{v}_k\|$ est non nul.

Le vecteur u_k cherché est alors défini par

$$u_k = \frac{\widehat{v_k}}{\|\widehat{v_k}\|},$$

le procédé s'arête après m étapes. $\square$

Notons que $\{u_1, \dots, u_m\}$ est une base orthonormée du sous-espace $W = \mathrm{Vec}(\{v_1, \dots, v_m\})$ engendré par les vecteurs donnés.

Définition 11.3.2. On dit que cette base orthonormée a été obtenue à partir de $\{v_1, \dots, v_m\}$ par le *procédé d'orthonormalisation de Gram-Schmidt.*

Le procédé d'orthonormalisation peut se résumer dans les formules suivantes :

$$\widehat{v_1} = v_1, \quad \widehat{v_2} = v_2 - \frac{\langle \widehat{v_1}, v_2 \rangle}{\|\widehat{v_1}\|^2}\widehat{v_1}, \ \cdots \ , \widehat{v_k} = v_k - \sum_{i=1}^{k-1} \frac{\langle \widehat{v_i}, v_k \rangle}{\|\widehat{v_i}\|^2}\widehat{v_i},$$

puis pour tout k on pose $u_k = \dfrac{\widehat{v_k}}{\|\widehat{v_k}\|}$.

11.4 Isométries d'un espace vectoriel euclidien.

Définition. Soit (E, g) un espace vectoriel euclidien. Une *isométrie* de E est une application bijective $f : E \to E$ qui respecte les distances, c'est-à-dire

$$\|f(y) - f(x)\| = \|y - x\|, \qquad \forall x, y \in E,$$

où $\|\cdot\| = \|\cdot\|_g$ est la norme associée au produit scalaire g.
Nous laissons au lecteur le soin de vérifier à partir de cette définition que les isométries de $\mathbb{E}^n$ forment un groupe.

Théorème 11.4.1. *L'application $f : E \to E$ est une isométrie si et seulement s'il existe un vecteur $b \in \mathbb{E}^n$ et une application linéaire $f_0 : E \to E$ tels que $f(x) = f_0(x) + b$ pour tout $x \in E$ et*

$$\|f_0(x)\| = \|x\| \qquad \forall x \in E.$$

On dit que f_0 est la *partie linéaire* de l'isométrie f et b est le *vecteur de translation* de f. Remarquons que ce vecteur est donné par $b = f(0)$.

Preuve. Nous démontrons d'abord le théorème dans le cas particulier où f est une isométrie fixant l'origine, i.e. $f(0) = 0$. Nous devons prouver que dans ce cas, f est linéaire.
On remarque d'abord que pour tout $x \in E$, on a

$$\|f(x)\| = \|f(x) - f(0)\| = d(f(x), f(0)) = d(x, 0) = \|x\|.$$

On montre maintenant que f respecte les produit scalaires, i.e. $\langle f(x), f(y)\rangle = \langle x, y\rangle$ pour tous $x, y \in E$. Cela découle du calcul suivant :

$$\begin{aligned}
2\langle f(x), f(y)\rangle &= \|f(x)\|^2 + \|f(y)\|^2 - \|f(y) - f(x)\|^2 \\
&= d(f(x), 0)^2 + d(f(y), 0)^2 - d(f(x), f(y))^2 \\
&= d(x, 0)^2 + d(y, 0)^2 - d(x, y)^2 \\
&= \|x\|^2 + \|y\|^2 - \|x - y\|^2 \\
&= 2\langle x, y\rangle.
\end{aligned}$$

Nous pouvons maintenant montrer la linéarité de f. Soient $x \in E$ un vecteur quelconque et $\alpha \in \mathbb{R}$, alors

$$\begin{aligned}\|f(\alpha x) - \alpha f(x)\|^2 &= \|f(\alpha x)\|^2 - 2\langle f(\alpha x), \alpha f(x)\rangle + \alpha^2\|f(x)\|^2 \\ &= \|f(\alpha x)\|^2 - 2\alpha\langle f(\alpha x), f(x)\rangle + \alpha^2\|f(x)\|^2 \\ &= \|\alpha x\|^2 - 2\alpha\langle \alpha x, x\rangle + \alpha^2\|x\|^2 \\ &= 0,\end{aligned}$$

ce qui prouve que $f(\alpha x) = \alpha f(x)$.
D'autre part, si $x, y \in E$ sont deux vecteurs, alors

$$\begin{aligned}&\|f(x) + f(y) - f(x+y)\|^2 = \langle f(x) + f(y) - f(x+y), f(x) + f(y) - f(x+y)\rangle \\ &\quad = \|f(x)\|^2 + \|f(y)\|^2 + \|f(x+y)\|^2 + 2\langle f(x), f(y)\rangle - 2\langle f(x), f(x+y)\rangle - 2\langle f(x+y), f(y)\rangle \\ &\quad = \|x\|^2 + \|y\|^2 + \|x+y\|^2 + 2\langle x, y\rangle - 2\langle x, x+y\rangle - 2\langle x+y, y\rangle \\ &\quad = \langle x + y - (x+y), x + y - (x+y)\rangle = 0,\end{aligned}$$

ce qui prouve que $f(x+y) = f(x) + f(y)$. On a donc démontré qu'*une isométrie de E qui fixe l'origine est une application linéaire.*
Pour le cas d'une isométrie générale, on définit une application $f_0 : E \to E$ par $f_0(x) = f(x) - f(0)$. Alors il est clair que $f_0(0) = 0$ et f_0 est une isométrie car

$$\begin{aligned}d(f_0(x), f_0(y)) &= \|f_0(x) - f_0(y)\| \\ &= \|\,(f(x) - f(0)) - (f(y) - f(0))\,\| \\ &= \|f(x) - f(y)\| \\ &= d(x, y).\end{aligned}$$

On a donc montré que l'application f s'écrit $f(x) = f_0(x) + b$. où $b = f(0) \in E$ est constant et f_0 est une isométrie linéaire.

□

Corollaire 11.4.2. *Si g est un produit scalaire sur $\mathbb{R}^n$ et $f : \mathbb{R}^n \to \mathbb{R}^n$ est une isométrie pour la distance associée à ce produit scalaire, alors on a*

$$f(x) = Ax + b,$$

où $b = f(0)$ et $A \in GL_n(\mathbb{R})$ est une matrice vérifiant $A^\top G A = G$ (où G est la matrice de Gram de g dans la base canonique de $\mathbb{R}^n$).

Preuve. On a vu dans la preuve du théorème précédent que l'application $x \mapsto Ax$ préserve le produit scalaire, i.e. on a

$$g(Ax, Ay) = g(x, y)$$

pour tous $x, y \in \mathbb{R}^n$. D'autre part $Ae_r = \sum_{i=1}^n a_{ir}e_i$ et $Ae_s = \sum_{j=1}^n a_{js}e_j$. Rappelons que par définition de la matrice de Gram, on a $g_{ij} = g(e_i, e_j)$, par conséquent

$$g_{rs} = g(e_r, e_s) = g(Ae_r, Ae_s) = g\left(\sum_{i=1}^n a_{ir}e_i, \sum_{j=1}^n a_{js}e_j\right) = \sum_{i,j=1}^n a_{ir}g_{ij}a_{js} = \left(A^\top GA\right)_{rs},$$

ce qui prouve que $G = A^\top G A$.

Voici une autre preuve : on peut représenter les vecteurs de $\mathbb{R}^n$ par des matrice-colonnes et écrire $g(x,y) = X^\top GY$. L'égalité $g(x,y) = g(AX, AY)$ s'écrit alors $X^\top GY = (AX)^\top G(AY)$ et donc

$$X^\top GY = (AX)^\top G(AY) = (X^\top A^\top)\, G(AY) = X^\top (A^\top GA)Y,$$

pour tous X, Y. Ceci implique que $G = A^\top GA$.

□

Ce résultat justifie la définition importante suivante :

Définition 11.4.3. Une matrice $A \in M_n(\mathbb{R})$ est *G-orthogonale* si $A^\top GA = G$. On note

$$O(G) = \{A \in M_n(\mathbb{R}) \mid A^\top GA = G\}.$$

Remarques.
1. Il est facile de vérifier que $\det(A) = \pm 1$ pour tout $A \in O(G)$. De plus $O(G)$ est un sous-groupe de $GL_n(\mathbb{R})$.
2. Lorsque g est le produit scalaire standard de $\mathbb{R}^n$, alors on a $G = I_n$ et on note le groupe orthogonal correspondant simplement

$$O(n) = O(I_n) = \{A \in M_n(\mathbb{R}) \mid A^\top A = I_n\}.$$

Observer que $A \in O(n)$ si et seulement si A est inversible et $A^\top = A^{-1}$.

11.5 Le groupe orthogonal

Dans cette section, nous étudions les matrices orthogonales en détail.

Proposition 11.5.1. *Pour toute matrice $A \in M_n(\mathbb{R})$ les propriétés suivantes sont équivalentes :*

(i) $A \in O(n)$, *c'est-à-dire* $A^\top A = I_n$.
(ii) A *est inversible et* $A^{-1} = A^\top$.
(iii) $\|Ax\| = \|x\|$ *pour tout* $x \in \mathbb{R}^n$.
(iv) $\langle Ax, Ay\rangle = \langle x, y\rangle$ *pour tous* $x, y \in \mathbb{R}^n$.
(v) Les colonnes de A forment une base orthonormée de $\mathbb{R}^n$.
(vi) Les lignes de A forment une base orthonormée de $\mathbb{R}^n$.
(vii) Pour tout vecteur $b \in \mathbb{R}^n$, l'application affine $f : \mathbb{R}^n \to \mathbb{R}^n$ définie par $f(x) = Ax + b$ est une isométrie.

De plus $O(n)$ est un sous-groupe de $GL_n(\mathbb{R})$ et pour tout $A \in O(n)$ on a $\det(A) = \pm 1$.

Dans cette proposition, le produit scalaire est le produits scalaire standard de $\mathbb{R}^n$ et la norme et la distance sont associées à ce produit scalaire. Nous laissons la preuve de cette proposition en exercice.

Remarquons que l'application déterminant définit un homomorphisme de groupes

$$\det : O(n) \to \{\pm 1\}.$$

Le noyau de cet homomorphisme est le *groupe spécial orthogonal* :

$$SO(n) = O(n) \cap SL_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid A^\top A = I_n \text{ et } \det(A) = +1\}.$$

La proposition suivante décrit les 2×2 matrices orthogonales.

Proposition 11.5.2. *Pour toute matrice $A \in O(2)$, il existe un angle θ tel que*

$$A = R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}, \text{ si } \det(A) = +1,$$

et

$$A = S_{\theta/2} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}, \text{ si } \det(A) = -1.$$

La matrice R_θ représente une rotation d'angle θ et $S_{\theta/2}$ représente la réflexion à travers la droite vectorielle formant un angle $\theta/2$ avec le premier vecteur e_1 de la base canonique.

Preuve. Les colonnes d'une matrices orthogonale $A \in O(2)$ doivent former une base orthonormée de $\mathbb{R}^2$. Il existe donc $\theta \in (-\pi, \pi]$ tel que la première colonne s'écrive $\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$. La deuxième colonnes de A doit-être un vecteur de norme 1 orthogonal à la première colonne, c'est-à dire $\pm \begin{pmatrix} \sin(\theta) \\ \cos(\theta) \end{pmatrix}$. Ceci démontre que ou bien $A = R_\theta$ ou bien $A = S_{\theta/2}$.

Finalement R_θ est une matrice de rotation car l'angle entre tout vecteur non nul $\mathbf{x}$ et $R_\theta(\mathbf{x})$ est égal à θ et $S_{\theta/2}$ est une symétrie car cette matrice possède deux vecteurs propres orthogonaux de valeurs propre $+1$ et -1 respectivement. Ces vecteurs propres sont

$$\begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} \sin(\theta/2) \\ -\cos(\theta/2) \end{pmatrix}.$$

Nous laissons la vérification de ces deux dernières affirmations en exercice.

□

Décrivons maintenant les 3×3 matrices orthogonales.

Proposition 11.5.3. *Toute matrice $A \in O(3)$ est semblable à une matrice du type*

$$\begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Nous allons démontrer le résultat plus général suivant qui donne la structure des isométries linéaires d'un espace euclidien.

Théorème 11.5.4. *Soit $f : V \to V$ une isométrie linéaire d'un espace vectoriel euclidien de dimension n. Alors il existe une base orthonormée de V dans laquelle la matrice de f prend la* *forme*

$$
\begin{aligned}
M(f) &= I_r \oplus (-I_s) \oplus R_{\theta_1} \oplus \cdots \oplus R_{\theta_m} \\
&= \begin{pmatrix}
I_r & & & & \\
& -I_s & & & \\
& & \begin{pmatrix} \cos(\theta_1) & -\sin(\theta_1) \\ \sin(\theta_1) & \cos(\theta_1) \end{pmatrix} & & \\
& & & \ddots & \\
& & & & \begin{pmatrix} \cos(\theta_m) & -\sin(\theta_m) \\ \sin(\theta_m) & \cos(\theta_m) \end{pmatrix}
\end{pmatrix}
\end{aligned}
$$

Pour la preuve, nous aurons besoin du lemme suivant :

Lemme 11.5.5. *Soit $f : V \to V$ une isométrie linéaire. Supposons que $W \subset V$ est invariant par f. Alors $W^\perp$ est aussi invariant par f.*

Preuve du lemme. Observons d'abord que si $W \subset V$ est invariant par f, i.e. $f(W) \subset W$, alors $f(W) = W$, i.e. la restriction de f à W est un isomorphisme de W vers W. Cela découle du théorème du rang et du fait que $\mathrm{Ker}(f) = 0$ pour toute isométrie linéaire d'un espace euclidien. Pour montrer que $W^\perp$ est invariant par f, on se donne $y \in W^\perp$ et $x \in W$ quelconque. On a alors $f^{-1}(x) \in W$, d'où l'on déduit que

$$\langle f(y), x\rangle = \langle f(y), f(f^{-1}(x))\rangle = \langle y, f^{-1}(x)\rangle = 0.$$

Cela prouve que $f(y) \in W^\perp$.

$\square$

Démonstration du théorème. Nous démontrons le théorème par récurrence sur la dimension de V. Si V est de dimension 1, alors le théorème affirme simplement que les seules isométries linéaires de V sont $f(v) = v$ et $f(v) = -v$, ce qui est évident. Le cas de la dimension 2 a été étudié dans la proposition 11.5.2. Supposons maintenant le théorème démontré pour toute isométrie d'un espace euclidien de dimension inférieure à $n = \dim(V)$. On sait par un résultat précédent que tout endomorphisme $f : V \to V$ d'un espace vectoriel de dimension finie admet un sous-espace invariant $W \subset V$ de dimension 1 ou 2. Par le lemme précédent, on sait que $W^\perp$ est alors aussi invariant par f. On distingue alors trois cas :

Cas 1. $\dim(W) = 1$. Par hypothèse de récurrence on peut trouver une base orthonormée $\{e_2, \ldots, e_n\}$ de $W^\perp$ telle que la matrice de la restriction de f à $W^\perp$ dans cette base prenne la forme

$$I_{r'} \oplus (-I_{s'}) \oplus R_{\theta_1} \oplus \cdots \oplus R_{\theta_m},$$

avec $r' + s' + 2m = n - 1$. Soit $e_1 \in W$ un vecteur de norme 1. Alors on a $f(e_1) = \pm e_1$ et la matrice de f dans la base $\{e_1, e_2, \ldots, e_n\}$ prend la forme

$$I_r \oplus (-I_s) \oplus R_{\theta_1} \oplus \cdots \oplus R_{\theta_m},$$

où $(r, s) = (r' + 1, s')$ si $f(e_1) = e_1$ et $(r, s) = (r', s' + 1)$ si $f(e_1) = -e_1$.

Cas 2. $\dim(W) = 2$ et la restriction de f à W est une symétrie. Alors il existe une droite $W_1 \subset W$ invariante par f et nous sommes ramenés au cas 1.

Cas 3. $\dim(W) = 2$ et la restriction de f à W est une rotation d'angle θ Par hypothèse de récurrence on peut trouver une base orthonormée $\{e_3, \ldots, e_n\}$ de $W^\perp$ telle que la matrice de la restriction de f $W^\perp$ dans cette base prenne la forme

$$I_{r'} \oplus (-I_{s'}) \oplus R_{\theta_1} \oplus \cdots \oplus R_{\theta_m},$$

avec $r' + s' + 2m = n - 2$. Soit $\{e_1, e_2\} \in W$ une base orthonormée de W, alors la matrice de la restriction de f au plan invariant W est la matrice de rotation R_θ et la matrice de f dans la base $\{e_1, e_2, \ldots, e_n\}$ prend la forme

$$R_\theta \oplus I_r \oplus (-I_s) \oplus R_{\theta_1} \oplus \cdots \oplus R_{\theta_m}.$$

Le théorème est démontré.

$\square$

On peut reformuler le théorème de la façon suivante : *Pour tout $A \in O(n)$, il existe une matrice $Q \in O(n)$ telle que*

$$A' := Q^\top AQ = Q^{-1}AQ = I_r \oplus (-I_s) \oplus R_{\theta_1} \oplus \cdots \oplus R_{\theta_m}.$$

11.6 Espace-temps Galiléen et référentiels inertiels

La mécanique classique, telle que développée depuis Galilée et Newton, et jusqu'à la fin du 19ème siècle, étudie des phénomènes telles que le mouvement dans un espace et pendant un intervalle de temps qui sont considérés comme des absolus. Le philosophe Emmanuel Kant considère que le temps et l'espaces sont des formes pures de l'intuition, des *catégories synthétiques a priori* de la connaissance. Pour Kant, dont l'un des projets est d'établir un cadre philosophique permettant d'intégrer la mécanique Newtonienne, l'espace et le temps sont donnés à notre intuition de façon indépendante de toute expérimentation (c'est ici le sens du mot *a priori*).

En mécanique classique, les événements sont donc étudiés dans un espace-temps de dimension 4 correspondant à une dimension temporelle et 3 dimensions spatiales. La mesure du temps est considérée comme absolue, pouvant se dérouler selon un axe réel $-\infty < t < \infty$, et l'espace est un espace euclidien de dimension 3, que nous identifions à $\mathbb{R}^3$. Il est commode d'appeler *événement* un élément [2] (x, y, z, t) de $\mathbb{R}^4$. On dit alors que $\mathbb{R}^4 = \mathbb{R}^3 \times \mathbb{R}$ est un *espace-temps galiléen* lorsqu'on le muni des deux structures suivantes :

(i) La *mesure du temps*, qui est la fonction $\mathbb{R}^4 \to \mathbb{R}$ donnée par $(x, y, z, t) \to t$. Elle est concrètement réalisée par l'horloge de référence de l'expérimentateur.

(ii) La *norme euclidienne*, qui est la fonction $\mathbb{R}^4 \to \mathbb{R}$ donnée par $(x, y, z, t) \to \sqrt{x^2 + y^2 + z^2}$. Cette norme est associée au produit scalaire euclidien de $\mathbb{R}^3$ et permet de reconstruire mathématiquement toute la géométrie euclidienne de l'espace (les distances, les angles, les aires etc.). Cette norme représente donc toutes les informations géométriques que fournissent les instruments de mesure disponibles à l'expérimentateur.

2. On dit alors que (x, y, z, t) est le *quadrivecteur* représentant les coordonnées spatiotemporelles de l'événement considéré. Notons qu'on peut parfois ignorer une ou deux coordonnées spatiales, par exemple si on étudie un mouvement dans un plan ou un mouvement rectiligne. L'événement est alors représenté par un élément $(x, y, t) \in R^3$ ou $(x, t) \in R^2$).

Nous avons alors les définitions naturelles suivantes :

Définitions 1. On appelle *durée* ou *intervalle temporel* entre deux événements $\xi_1 = (x_1, y_1, z_1, t_1)$ et $\xi_2 = (x_2, y_2, z_2, t_2)$ la quantité

$$\tau(\xi_1, \xi_2) = |t_1 - t_2|.$$

La durée est indépendante de la position spatiale des événements ξ_1 et ξ_2.

2. La *distance* entre ces événements est la quantité

$$d(\xi_1, \xi_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

La distance est indépendante du temps.

Définition. Une transformation $f : \mathbb{R}^4 \to \mathbb{R}^4$ est une *transformation galiléenne* si elle est vérifie les conditions suivantes :

(i) La transformation f est est bijective.

(ii) La transformation préserve les durées : pour tous $\xi_1, \xi_2 \in \mathbb{R}^4$ on a

$$\tau(f(\xi_1), f(\xi_2)) = \tau(\xi_1, \xi_2).$$

Cette condition correspond à l'hypothèse d'un temps absolu.

(iii) La transformation est *isométrique*, c'est-à-dire qu'elle préserve les distances : pour tous $\xi_1, \xi_2 \in \mathbb{R}^4$ on a

$$d(f(\xi_1), f(\xi_2)) = d(\xi_1, \xi_2).$$

Cette condition reflète l'hypothèse d'un espace absolu et uniforme, le même pour tous les observateurs.

(iv) La transformation est *inertielle* : elle transforme un mouvement rectiligne uniforme en un mouvement rectiligne uniforme.

Exemple. L'exemple le plus simple de transformation galiléenne est donné par la formule :

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \mapsto \begin{pmatrix} x + tv_1 \\ y + tv_2 \\ z + tv_3 \\ t + t_0 \end{pmatrix}$$

Théorème 11.6.1. *L'application $f : \mathbb{R}^4 \to \mathbb{R}^4$ est une transformation galiléenne si et seulement s'il existe $t_0 \in \mathbb{R}$, deux vecteurs $b, v \in \mathbb{R}^3$ et une matrice orthogonale $A \in O(3)$ telles que f se décompose sous la forme suivante :*

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto A \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} + t \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad \text{et} \quad t \mapsto t_0 \pm t.$$

Démonstration. La condition (ii) nous permet de décrire une transformation galiléenne en séparant la coordonnée temporelle et les coordonnées spatiales. La coordonnées spatiale se transforme selon la règle $t \mapsto t_0 \pm t$, et pour les coordonnées spatiales, nous avons à chaque instant t une bijection de $f_t : \mathbb{R}^3 \to \mathbb{R}^3$ qui doit être une isométrie pour la norme euclidienne standard.

Si on écrit les coordonnée spatiotemporelle d'un événement sous la forme $(x,t) = (x_1, x_2, x_3, t)$, alors f_t n'agit que sur les coordonnées spatiales $x = (x_1, x_2, x_3)$ par la formule

$$x \to f_t x = A(t)x + a(t),$$

où $A(t) \in O(3)$ est une matrice orthogonale qui dépend du temps et $a(t)$ est le vecteur de translation, qui dépend lui aussi du temps.

Admettons pour simplifier que la transformation temporelle est l'identité $t \mapsto t$ et supposons que $t \mapsto x(t)$ représente la trajectoire d'une particule. Notons $y(t) = f_t(x(t))$, nous avons alors

$$y(t) = A(t)x(t) + a(t), \quad \dot{y}(t) = A(t)\dot{x}(t) + \dot{A}(t)x(t) + \dot{a}(t),$$

ou le point représente la dérivée par rapport au temps ; et donc aussi

$$\ddot{y}(t) = A(t)\ddot{x}(t) + 2\dot{A}(t)\dot{x}(t) + \ddot{A}(t)x(t) + \ddot{a}(t). \tag{11.4}$$

L'hypothèse (iv) que la transformation est inertielle dit que pour toute trajectoire telle que $\ddot{x} = 0$ on doit aussi avoir $\ddot{y} = 0$. Cette condition nous dit que si $x(t) = c+tw$, où c et w sont des vecteurs constants quelconques de $\mathbb{R}^3$, alors on a $\dot{x}(t) = w$ et $\ddot{x}(t) = 0$. La condition (11.4) entraîne alors que

$$0 = \ddot{y}(t) = 2\dot{A}(t)w + \ddot{A}(t)(c + tw) + \ddot{a}(t).$$

Cette condition est valable pour tous $c, w \in \mathbb{R}^3$ constants. En posant $c = w = 0$ on obtient que $\ddot{a}(t) = 0$; il existe donc deux vecteurs constants $b, v \in \mathbb{R}^3$ tels que $a(t) = b+tv$. En posant $w = 0$, on voit que $\ddot{A}(t)c = 0$ pour tout vecteur c et donc $\ddot{A}(t) = 0$. Finalement en posant $c = 0$ on voit que $\dot{A}(t)w = 0$ pour tout vecteur w et donc $\dot{A}(t) = 0$.
On a ainsi prouvé que $A(t)$ est un élément constant de $O(3)$ et $a(t) = b + tv$, c'est-à-dire que

$$f_t(x) = Ax + b + tv,$$

avec $A \in O(3)$ et $b, v \in \mathbb{R}^3$.

□

Terminons ce paragraphe par quelques remarques :

Remarques. (i) Puisque la matrice A est constante et $\ddot{a} = 0$, l'équation (11.4) nous dit que l'accélération de $y(t) = f_t(x(t))$ vérifie

$$\ddot{y}(t) = A\ddot{x}(t).$$

Et comme $A \in O(3)$, nous avons en particulier $\|\ddot{y}(t)\| = \|\ddot{x}(t)\|$. Ceci implique que les transformations galiléennes respectent l'équation Newtonienne du mouvement (force = masse × accélération), ce que l'on formule sous la forme du principe de relativité galiléenne : *Le temps et l'espace sont absolus et les lois de la mécanique sont les mêmes dans tous les référentiels inertiels.*
(ii) Le lecteur attentif aura remarqué que les transformation galiléennes autorisent l'inversion du temps, i.e. la transformation $t \mapsto -t$ de la coordonnée temporelle. Cette convention est justifiée par le fait que l'inversion du temps est compatible avec la loi d'évolution Newtonienne d'une particule se mouvant dans un champ de force. Mais nous savons bien empiriquement que le temps s'écoule du passé vers l'avenir, et que la plupart des évolutions ne sont pas réversibles.

Ajoutons que l'inversion du temps contredit le second principe de la thermodynamique. Il est donc raisonnable de ne considérer que les transformations galiléennes qui respectent l'orientation temporelle (on dit d'une telle transformation qu'elle respecte la chronologie).

(iii) Un remarque similaire s'applique à l'orientation de l'espace. Si on considère que l'orientation de l'espace est une donnée essentielle de la physique, alors nous devons restreindre les transformations galiléenne au cas où $A \in SO(3)$. Notons que l'orientation de l'espace joue un rôle significatif en électromagnétisme (loi de Biot-Savard, force de Lorentz...)

(iv) Mentionnons pour finir que l'inverse d'une transformation galiléenne et la composition de deux transformations galiléennes sont encore des transformations galiléennes. Ces transformations forment donc un groupe, qu'on appelle le *groupe de Galilée.* C'est un sous-groupe du groupe des transformations affines de à $\mathbb{R}^4$.

"Le mouvement est comme rien" (un texte de Galillée)

La physique est une science expérimentale dont les lois sont formulées mathématiquement. Toutefois les expérimentations de laboratoire et les développements mathématiques sont complétés par un autre type de raisonnement, qu'on appelle des « expériences de pensées » (Gedankenexperimente). Ces raisonnements mettent en scène des schémas pseudo-expérimentaux et permettent de dégager les grands principes de la physique. Le grand génie des expériences de pensées est Albert Einstein, mais le procédé remonte à Galilée. Dans son Dialogue sur les deux grands systèmes du monde (1632)*, Galilée propose une expérience de pensée célèbre. Dans ce texte majeur de l'histoire des sciences, il nous donne la première formulation historique du principe de relativité sous la forme suivante.*

> « Enfermez-vous avec un ami dans la plus vaste cabine d'un grand navire, Et apportez des mouches, des papillons et d'autres petits animaux semblables. Amenez aussi un grand bocal d'eau contenant des poissons, suspendez au plafond un petit seau dont l'eau tombe goutte à goutte par un orifice étroit et tombe dans un vase posé sur le sol.
>
> Puis, alors que le navire est à l'arrêt, observez attentivement, comment ces petits animaux volent avec des vitesses égales quel que soit l'endroit de la cabine vers lequel ils se dirigent. Les poissons nagent indifféremment dans toutes les directions. Les gouttelettes d'eau tombent régulièrement dans le vase situé sur le sol. Si vous lancez un objet à un ami, vous n'avez pas besoin de le lancer plus fort dans une direction que dans une autre, si les distances sont égales, et si vous sautez à pieds joints, vous franchissez des distances égales dans toutes les directions. Il ne fait aucun doute que si le navire est à l'arrêt les choses doivent se passer ainsi.
>
> Une fois que vous aurez observé attentivement tout cela, faites avancer le bateau à l'allure qui vous plaira, pour autant que la vitesse soit uniforme et ne fluctue pas de-ci de-là. Vous ne discernerez alors aucun changement dans tous les effets précédents, et aucune observation ne vous renseignera si le navire est en marche ou s'il est arrêté.

> Si vous sautez, vous franchirez sur le plancher les mêmes distances qu'auparavant et, si le navire se déplace, vous n'en ferez pas pour autant des sauts plus grands vers la poupe que vers la proue, bien que, pendant que vous êtes en l'air, le plancher qui est en dessous ait glissé dans la direction opposée à celle de votre saut. Les gouttes d'eau tomberont comme précédemment dans le vase inférieur. Les poissons dans leur eau, et sans plus de fatigue, nageront d'un côté comme de l'autre.
>
> Enfin les papillons et les mouches continueront leur vol indifférent dans n'importe quel sens, sans être influencé par la marche et la direction du navire, on ne les verra pas s'accumuler du côté de la cloison qui fait face à la poupe ; ce qui ne manquerait pas d'arriver s'ils devaient s'épuiser à suivre le navire dans sa course rapide.
> La cause de la permanence de tous ces effets, c'est que le mouvement uniforme est commun au navire et à ce qu'il contient, y compris l'air. Le mouvement est mouvement, et agit comme mouvement, en tant et seulement qu'il est en rapport avec les choses qui en sont privées ; mais en ce qui concerne celles qui y participent toutes également, il est sans effet ; il est comme s'il n'était pas. Le mouvement est comme rien ! »

Ajoutons quelques commentaires à ce beau texte. Le Dialogue sur les deux grands systèmes du monde est rédigé sous forme de dialogue entre trois protagonistes. Dans ce texte, écrit en italien et non en latin, Galilée cherche à convaincre le lecteur de la supériorité du système héliocentrique (plaçant le soleil au centre de l'Univers) sur le géocentrisme qui place la terre au centre. L'héliocentrisme a été avancé par Copernic comme un modèle permettant de rendre compte du mouvement des planètes en décrivant leur orbites autour du soleil et non de la terre. Le modèle copernicien a été amélioré par Kepler qui, suite aux observations de Tycho-Brahé, énonça les trois lois qui portent son nom (les orbites des planètes sont des ellipses etc.). L'héliocentrisme fut condamné par l'Eglise en 1616 comme contraire aux enseignements de l'Ecriture. L'un des arguments contre l'héliocentrisme était que si la terre tourne autour du soleil, nous devrions nous en rendre compte (par exemple la chute des corps ne serait pas verticale etc.) Le texte ci-dessus réfute cet argument, mais une polémique s'en est suivie et Galilée fut condamné par l'inquisition à renier ses thèses et son ouvrage fut interdit de publication. C'était en 1633, un an après la publication du Dialogue (l'interdiction fut levée en en 1741 par le pape Benoît XIV).

Le principe d'inertie (ou de relativité) de Galilée, est à rapprocher de la première Loi de Newton (énoncée dans les *Principia* en 1687). *Tout corps persévère dans l'état de repos ou de mouvement uniforme en ligne droite dans lequel il se trouve, à moins que quelque force n'agisse sur lui, et ne le contraigne à changer d'état.*

En langage contemporain, le principe de relativité Galiléenne s'énonce ainsi : *Il existe une famille de référentiels qui sont mutuellement au repos ou en mouvement rectiligne uniforme. Les lois de la mécanique (classique) ont le même forme dans tout référentiel inertiel.* Le navire de Galilée est un exemple de référentiel inertiel.

11.7 Le théorème spectral (première version)

Le but de ce paragraphe est de démontrer le théorème suivant :

Théorème 11.7.1 (Théorème Spectral). *Soit $H \in M_n(\mathbb{R})$ une $n \times n$ matrice symétrique (i.e. $H^\top = H$) à coefficients réel. Alors on a*

(1.) Les valeurs propres de H sont réelles.

(2.) Les espaces propres de H associés à des valeurs propres distinctes sont deux-à-deux orthogonaux.

(3.) Il existe une base orthonormée de $\mathbb{R}^n$ formée de vecteurs propres.

Remarques. (i.) Dans les points (2) et (3) l'orthogonalité fait référence au produit scalaire standard de $\mathbb{R}^n$, que nous noterons ici $\langle \cdot, \cdot \rangle$.

(ii.) Le théorème nous dit en particulier que pour une matrice réelle symétrique, la multiplicité géométrique de chaque valeur propre est égale à sa multiplicité algébrique.

(iii.) Ce théorème est l'une des formes du théorème spectral, nous en verrons d'autres.

Le lemme suivant sera utilisé dans la preuve du théorème.

Lemme 11.7.2. *Pour toute matrice $A \in M_n(\mathbb{R})$ et tous $X, Y \in \mathbb{R}^n$ on a*

$$\langle X, AY \rangle = \langle A^\top X, Y \rangle = X^\top AY = \sum_{i,j=1}^{n} x_i a_{ij} y_j, .$$

en particulier $A^\top = A$ si et seulement si $\langle AX, Y \rangle = \langle X, AY \rangle$ pour tous $X, Y \in \mathbb{R}^n$. (on regarde les éléments $X, Y \in \mathbb{R}^n$ comme des vecteurs-colonnes).

Preuve : Exercice.

Démonstration du théorème spectral.

(1.) On démontre d'abord que les valeurs propres de H sont réelles. Considérons donc une valeur propre $\lambda \in \mathbb{C}$ et montrons que $\lambda \in \mathbb{R}$.

En effet, si $\lambda \in \sigma_{\mathbb{C}}(H)$, alors il existe $Z \in \mathbb{C}^n$ tel que $HZ = \lambda Z$. On a alors aussi $H\bar{Z} = \bar{\lambda}\bar{Z}$, car $\bar{H} = H$ puisque H est une matrice à coefficients réels. On a donc

$$H\bar{Z} = \bar{H}\bar{Z} = \overline{HZ} = \overline{\lambda Z} = \bar{\lambda}\bar{Z}.$$

Par conséquent on a

$$Z^\top(H\bar{Z}) = Z^\top(\bar{\lambda}\bar{Z}) = \bar{\lambda} \sum_{i=1}^{n} z_i \bar{z}_i. = \bar{\lambda} \sum_{i=1}^{n} |z_i|^2,$$

Mais on suppose que $H^\top = H$, donc

$$(Z^\top H)\bar{Z} = (H^\top Z)^\top \bar{Z} = (HZ)^\top \bar{Z} = (\lambda Z)^\top \bar{Z} = \lambda\, Z^\top \bar{Z} = \lambda \sum_{i=1}^{n} z_i \bar{z}_i = \lambda \sum_{i=1}^{n} |z_i|^2.$$

Ainsi
$$\lambda \sum_{i=1}^{n} |z_i|^2 = Z^\top H \bar{Z} = \bar{\lambda} \sum_{i=1}^{n} |z_i|^2.$$
Mais comme on a supposé que $Z \neq 0$, cette égalité implique $\lambda = \bar{\lambda}$. Cela prouve que toute valeur propre de H est réelle.

(2.) Montrons maintenant que les espaces propres associés à des valeurs propres distinctes sont orthogonaux, i.e.
$$\lambda, \mu \in \sigma(H),\ \lambda \neq \mu \ \Rightarrow\ E_\lambda \perp E_\mu.$$
En effet, si $x \in E_\lambda$ et $y \in E_\mu$, alors
$$\langle x, Hy\rangle = \langle x, \mu y\rangle = \mu\langle x, y\rangle.$$
Mais on a aussi
$$\langle Hx, y\rangle = \langle \lambda x, y\rangle = \lambda\langle x, y\rangle.$$
Et donc, en utilisant le lemme précédent :
$$\lambda\langle x, y\rangle = \langle Hx, y\rangle = \langle x, Hy\rangle = \lambda\langle x, y\rangle = \mu\langle x, y\rangle,$$
ce qui implique $\langle x, y\rangle = 0$ si $\lambda \neq \mu$. On a prouvé que des espaces propres associés à des valeurs propres distinctes sont orthogonaux.

(3.) Nous pouvons maintenant prouver qu'il existe une base orthonormée formée de vecteurs propres. Notons $\sigma(H) = \{\lambda_1, \dots, \lambda_r\}$ l'ensemble des valeurs propres de H. On peut choisir une base orthonormée $\{u_1, \dots, u_{m_1}\}$ de l'espace propre $E_{\lambda_1} \subset \mathbb{R}^n$ (où m_1 est la multiplicité géométrique de λ_1).

Par le point (2), on sait que pour tout $j \geq 2$, l'espace propre E_{λ_j} est orthogonal E_{λ_1}. C'est-à-dire
$$j \geq 2 \ \Rightarrow\ E_{\lambda_j} \subset E_{\lambda_1}^\perp = \{y \in \mathbb{R}^n \mid \langle x, y\rangle = 0,\, \forall x \in E_{\lambda_1}\}.$$

Affirmation. $E_{\lambda_1}^\perp$ est invariant par H.

Cette affirmation signifie que $y \in E_{\lambda_1}^\perp \Rightarrow Hy \in E_{\lambda_1}^\perp$, et se vérifie facilement grâce au lemme précédent. En effet, supposons que $y \in E_{\lambda_1}^\perp$, alors on a pour tout $x \in E_{\lambda_1}$
$$\langle x, Hy\rangle = \langle Hx, y\rangle = \lambda_1\langle x, y\rangle = 0,$$
ce qui entraîne que $Hy \in E_{\lambda_1}^\perp$. L'affirmation est démontrée.

Nous pouvons maintenant conclure la démonstration. Par hypothèse de récurrence, il existe alors une base orthonormée $\{u_{m_1+1}, \dots, u_n\}$ de $E_{\lambda_1}^\perp$ formée de vecteurs propres pour H. La réunion $\{u_1, \dots, u_n\}$ des deux bases est la base propre orthonormée de $\mathbb{R}^n$ cherchée.

$\square$

Diagonalisation Orthogonale

Définition 11.7.3. Deux matrices $A, A' \in M_n(\mathbb{R})$ sont dites *orthogonalement congruentes* s'il existe $P \in O(n)$ telle que

$$A' = P^\top AP.$$

Remarquer que dans ce cas les matrices sont aussi semblables car $P^\top = P^{-1}$. Le théorème spectral peut se reformuler ainsi :

Théorème 11.7.4. *Pour une matrice réelle $A \in M_n(\mathbb{R})$, les deux conditions suivantes sont équivalentes :*

(a) Toute matrice symétrique est orthogonalement diagonalisable, c'est-à-dire orthogonalement congruentes à une matrice diagonale.

(b) A est symétrique, i.e. $A^\top = A$.

Concrètement, cela signifie que si $A \in M_n(\mathbb{R})$ est symétrique, alors il existe une matrice orthogonale $P \in O(n)$ telle que

$$D = P^{-1}AP = P^\top AP$$

est diagonale.

Pour diagonaliser orthogonalement une matrice symétrique $A \in M_n(\mathbb{R})$, on procède selon les étapes suivantes :

1. On calcule le polynôme caractéristique de A et on cherche les valeurs propres $\{\lambda_1, \dots, \lambda_r\}$. Celles-si sont réelles car la matrice est symétrique.
2. $\mathbb{R}^n$ est somme directe des espaces propres E_{λ_i} car A est diagonalisable par le théorème spectral. De plus les espaces propres sont deux-à-deux orthogonaux car A est symétrique.
3. Pour chaque valeur propre λ_i on cherche une base orthonormée de l'espace propre E_{λ_i}.
4. La réunion des bases obtenues en (3) est une base propre orthonormée de l'espace vectoriel.

Des exemples seront vus aux exercices.

Application aux séries de Taylor des fonctions de n variables.

Soit $\Omega \subset \mathbb{R}^n$ un domaine ouvert de $\mathbb{R}^n$ et $f : \Omega \to \mathbb{R}$ une fonction continue ayant des dérivées continues jusqu'à l'ordre 3. Le *développement de Taylor l'ordre* 1 au voisinage d'un point $p \in \Omega$ s'écrit

$$f(p + v) = f(p) + df_p(v) + O(\|v\|^2).$$

où $df_p \in (\mathbb{R}^n)^*$ est la *différentielle de f en p*. C'est le covecteur défini par

$$df_p(v) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(p) v_i.$$

Ce covecteur représente l'approximation linéaire de f pour un petit "accroissement" $p + v$ de p. Le *développement de Taylor d"ordre* 2 en p s'écrit

$$f(p + v) = f(p) + df_p(v) + \frac{1}{2} h_p(v, v) + o(\|v\|^2),$$

où h_p est le *hessien* de f en p. C'est la forme bilinéaire symétrique définie par

$$h_p(v,w) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(p) v_i v_j.$$

La matrice de Gram de la forme bilinéaire h_p est la *matrice hessienne* de f en p, c'est-à-dire la matrice des dérivées secondes :

$$H_p = \left(\frac{\partial^2 f(p)}{\partial x_i \partial j} \right).$$

On peut réécrire le développement de Taylor sous la forme suivante :

$$f(x) = f(p) + \sum_{i=1}^{n} \frac{\partial f(p)}{\partial x_i}(x_i - p_i) + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f(p)}{\partial x_i \partial x_j}(x_i - p_i)(x_j - p_j) + o(\|x-p\|^3).$$

C'est une formule importante à comprendre et connaître.

Le théorème spectral, nous dit qu'on peut réduire le hessien à une forme diagonale en faisant un changement de coordonnées orthonormé. De façon plus précise, il existe un système de coordonnées $y_i = \sum_j p_{ij} x_j$, où (p_{ij}) est une matrice orthogonale, telles que dans ces coordonnées le développement de Taylor à l'ordre 2 s'écrit

$$\tilde{f}(y) = \tilde{f}(q) + \sum_{i=1}^{n} \frac{\partial \tilde{f}}{\partial y_i}(q)(y_i - q_i) + \frac{1}{2} \sum_{i=1}^{n} \frac{\partial^2 \tilde{f}}{\partial y_i^2}(q)(y_i - q_i)^2 + o(\|y-q\|^3).$$

Cette écriture permet de déterminer les points où la fonction f atteint un maximum local ou un minimum local (il faut que la différentielle en ce point possède soit nulle, puis on examine les signes des valeurs propres de H).

11.7.1 Application : le tenseur d'inertie et les moments d'inerties principaux

En mécanique, l'étude de la rotation d'un solide indéformable autour d'un axe met en évidence l'importance de la notion de *moment cinétique*. Considérons un solide indéformable qui est représenté par un domaine borné $D \subset \mathbb{R}^3$. La masse totale de ce solide est donnée par l'intégrale [3]

$$M = \int_D \rho(x) dx,$$

où la fonction $\rho : D \to \mathbb{R}_+$ représente la distribution (ou densité) de masse (si la masse se mesure en kg, alors l'unité de la fonction ρ est kg/m^3). Le *centre de gravité*, ou *barycentre* du solide D est le point de $\mathbb{R}^3$ défini par

$$C = \frac{1}{M} \int_D x \rho(x) dx.$$

3. Il s'agit ici d'une notation allégée pour l'intégrale triple

$$\iiint_D \rho(x_1 x_2 x_3) dx_1 dx_2 dx_3.$$

Ses coordonnées sont $C = (c_1, c_2, c_3)$, avec $c_i = \frac{1}{M}\int_D x_i \rho(x)dx$. Dans la suite on supposera que $C = 0$ (l'origine des coordonnées). Cette hypothèse revient à translater si nécessaire le solide D pour ramener son centre de gravité à l'origine des coordonnées.

Définition. On appelle *moment d'inertie*[4] du solide indéformable D en direction du vecteur unité u la quantité

$$\mathcal{I}_D(u) = \int_D \delta_u(x)^2 \rho(x)dx,$$

où $\delta_u(x)$ est la distance entre le point x et l'axe $\mathbb{R}\cdot u$.

Lemme 11.7.5. *Le moment d'inertie de D en direction de u peut aussi s'écrire*

$$\mathcal{I}_D(u) = \mathcal{J}_D(u,u),$$

où $\mathcal{J}_D$ est la forme bilinéaire définie sur $\mathbb{R}^3$ par

$$\mathcal{J}_D(u,v) = \int_D \left(\|x\|^2\langle u,v\rangle - \langle x,u\rangle\langle x,v\rangle\right)\rho(x)dx. \tag{11.5}$$

Preuve. On sait que la composante normale de x selon le vecteur u est $x - \langle x,u\rangle u$ (car on suppose $\|u\| = 1$), on a donc par le théorème de Pythagore

$$\|x\|^2 = \delta_u(x)^2 + \langle x,u\rangle^2.$$

Le moment d'inertie peut donc s'écrire

$$\mathcal{I}_D(u) = \int_D \delta_u(x)^2\rho(x)dx = \int_D \left(\|x\|^2 - \langle x,u\rangle^2\right)\rho(x)dx = \mathcal{J}_D(u,u).$$

□

Définition. La forme bilinéaire $\mathcal{J}_D : \mathbb{R}^3\times\mathbb{R}^3 \to \mathbb{R}$ définie par (11.5) s'appelle le *tenseur d'inertie* du solide D. Il ne dépend que de la forme de D et de la distribution de masse ρ.

Le tenseur d'inertie de D est clairement une forme bilinéaire symétrique sur $\mathbb{R}^3$. Le théorème spectral nous dit par conséquent qu'il existe une base orthonormée $\{u_1, u_2, u_3\}$ de $\mathbb{R}^3$ qui orthogonalise $\mathcal{J}_D$, c'est-à-dire :

$$\mathcal{J}_D(u_1,u_1) = \mathcal{J}_1, \quad \mathcal{J}_D(u_2,u_2) = \mathcal{J}_2, \quad \mathcal{J}_D(u_3,u_3) = \mathcal{J}_3, \quad \mathcal{J}_D(u_i,u_j) = 0 \text{ si } i \neq j.$$

On appelle $\mathcal{J}_1$, $\mathcal{J}_2$, $\mathcal{J}_3$ les *moments principaux d'inertie* et les directions de u_1, u_2, u_3 les *axes principaux d'inertie* du solide indéformable D. Dans la base $\{u_1, u_2, u_3\}$, la matrice de Gram de $\mathcal{J}_D$ est la matrice diagonale

$$\begin{pmatrix} \mathcal{J}_1 & 0 & 0 \\ 0 & \mathcal{J}_2 & 0 \\ 0 & 0 & \mathcal{J}_3 \end{pmatrix}.$$

Supposons maintenant que le solide D est en rotation autour d'un axe passant par le centre de gravité, et notons ω le vecteur de rotation instantanée. Alors le *moment cinétique* est le vecteur défini par

$$L_D(\omega) = \int_D \left(\|x\|^2\omega - \langle x,\omega\rangle x\right)\rho(x)dx.$$

4. Comparer avec la formule (1.57), page 54 du livre de mécanique de J.P. Ansermet. Noter que dans ce livre la formule est écrite pour un système fini de masses ponctuelles et non une densité continue de masse

Observons que pour tout vecteur $v \in \mathbb{R}^3$ on a

$$\mathcal{J}_D(\omega, v) = \langle L_D(\omega), v \rangle$$

Notons $\omega = \omega_1 u_1 + \omega_2 u_2 + \omega_3 u_3$, alors on a

$$L_D(\omega) = \sum_{i=1}^{3} \langle L_D(\omega), u_i \rangle u_i = \sum_{i=1}^{3} \mathcal{J}_D(\omega, u_i) u_i = \sum_{j=1}^{3} \sum_{j=1}^{3} \omega_j \mathcal{J}_D(u_j, u_i) u_i$$

Le moment cinétique se calcule donc dans la base $\{u_1, u_2, u_3\}$ à partir des moments principaux d'inertie par la formule

$$L_D(\omega) = \sum_{j=1}^{3} \mathcal{J}_i \cdot \omega_j.$$

En l'absence de force extérieure, le moment cinétique est conservé. Si des forces extérieures sont appliquées alors la variation du moment cinétique est égale à la somme des moments de forces. Cette relation s'écrit

$$\frac{dL_D}{dt} = \int_D x \times F(x) dx.$$

Cette équation détermine la dynamique du solide en rotation autour de son centre de gravité, qui est supposé fixe. Si le centre de gravité est lui aussi en mouvement, alors on doit ajouter l'équation de Newton :

$$M \frac{d^2 x_0}{dt^2} = \int_D F(x) dx = \text{force extérieure totale agissant sur le solide},$$

où $x_0(t)$ représente la position du centre de gravité au temps t. Ces deux dernières équations donnent une description complète de l'évolution d'un solide indéformable en mouvement.