Notes de cours de l'ISIMA, deuxième année, http://www.isima.fr/~leborgne

Mécanique : tenseurs 1ère partie – vecteurs et formes linéaires (contravariance et covariance), formules de changement de bases, tenseurs uniformes, contractions...

Gilles Leborgne

4 janvier 2021

Table des matières

1 Définition et premiers résultats (rappels)

But : comprendre les calculs tensoriels en mécanique (les calculs matriciels associés aux tenseurs).

1.1 Notation

Soit E une espace vectoriel réel de dimension finie $n \geq 1$. Par exemple $E = \mathbb{R}^n$, ou $E =$ un plan vectoriel dans $\mathbb{R}^3$...

On rappelle qu'une base d'un espace vectoriel est une famille $(\vec{e}_i)_{i=1,...,n}$ de vecteurs qui est libre (i.e., $\sum_{i=1}^n \alpha_i \vec{e}_i = 0 \Rightarrow \alpha_i = 0$ pour tout i, i.e., aucun des $\vec{e}_i$ n'est combinaison linéaire des autres) et génératrice (i.e., $\forall \vec{x} \in E$, $\exists x^1, ..., x^n \in \mathbb{R}$ t.q. $\vec{x} = \sum_{i=1}^n x^i \vec{e}_i$, i.e., tout vecteur de E est combinaison linéaires des $\vec{e}_i$). Et qu'alors $\dim E = n$.

Soit $(\vec{e}_i)_{i=1,...,n} =^{\text{noté}} (\vec{e}_i)$ une base dans E. Soit $\vec{x} \in E$ (un vecteur de E). Soit $x^1, ..., x^n$ ses composantes sur la base $(\vec{e}_i)$, i.e. les n réels tels que :

$$\vec{x} = \sum_{i=1}^n x^i \vec{e}_i, \quad \text{et} \quad [\vec{x}]_{|\vec{e}} := \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} \tag{1.1}$$

est la matrice colonne des composantes de $\vec{x}$ relativement à la base $(\vec{e}_i)$.

On note $\mathcal{M}_{np}$ l'ensemble des matrices $n * p$. Ainsi $[\vec{x}]_{|\vec{e}} \in \mathcal{M}_{n1}$ (matrice colonne).

Convention d'Einstein :

1- Les composantes x^i d'un vecteur $\vec{x} = \sum_{i=1}^n x^i \vec{e}_i$ dans une base $(\vec{e}_i)$ sont notés avec des index i en exposant (en haut).

2- Les vecteurs $(\vec{e}_i)$ d'une base sont avec des index en indices (en bas).

3- Une somme comme $\sum_{i=1}^n x^i \vec{e}_i$ où l'index répété i est à la fois en exposant (en haut) et en indice (en bas) représente une quantité indépendante d'un observateur ; ici c'est un vecteur $\vec{x} = \sum_{i=1}^n x^i \vec{e}_i$.

(Et la convention d'Einstein ajoute que dans ce cas on note simplement

$$\sum_{i=1}^n x^i \vec{e}_i \overset{\text{noté}}{=} x^i \vec{e}_i, \tag{1.2}$$

i.e., on omet le signe $\sum$. Avec l'apparition de l'ordinateur et des traitements de texte comme LaTeX, l'écriture du signe somme $\sum$ ne pose plus de problème, et dans ce poly le signe $\sum$ ne sera pas omis.)

1.2 Produit scalaire et métrique

1.2.1 Forme bilinéaire

Soit E et F deux espaces vectoriels de dimensions finies n et m. Une forme bilinéaire $g(\cdot,\cdot)$: $E \times F \to \mathbb{R}$ est une fonction linéaire par rapport à chaque variable, i.e. telle que, pour tout $\vec{u}_1, \vec{u}_2, \vec{u} \in E$, $\vec{v}_1, \vec{v}_2, \vec{v} \in F$ et $\lambda \in \mathbb{R}$:

$$\begin{cases} g(\vec{u}_1 + \lambda \vec{u}_2, \vec{v}) = g(\vec{u}_1, \vec{v}) + \lambda g(\vec{u}_2, \vec{v}), \\ g(\vec{u}, \vec{v}_1 + \lambda \vec{v}_2) = g(\vec{u}, \vec{v}_1) + \lambda g(\vec{u}, \vec{v}_2). \end{cases}$$

Autrement dit, à $\vec{v}$ fixé quelconque, la fonction $g_{\vec{v}} : E \to \mathbb{R}$ définie par $g_{\vec{v}}(\vec{u}) = g(\vec{u}, \vec{v})$ est linéaire, et, à $\vec{u}$ fixé quelconque, la fonction $g_{\vec{u}} : F \to \mathbb{R}$ définie par $g_{\vec{u}}(\vec{v}) = g(\vec{u}, \vec{v})$ est linéaire.

On note $\mathcal{L}(E, F; \mathbb{R})$ l'ensemble des formes bilinéaires de $E \times F$ dans $\mathbb{R}$. Quand $F = E$, on note $\mathcal{L}(E, E; \mathbb{R}) =^{\text{noté}} L^2(E; \mathbb{R})$.

Exemple 1.1 $g(\cdot,\cdot) : (x, y) \in \mathbb{R} \times \mathbb{R} \to g(x, y) = xy \in \mathbb{R}$ est la forme bilinéaire appelée produit dans $\mathbb{R}$ (c'est le produit scalaire dans $\mathbb{R}$). ▪

Représentation relative à une base $(\vec{a}_i)$ de E et à une base $(\vec{b}_i)$ de F : la matrice $[g(\cdot,\cdot)]_{|\vec{a},\vec{b}} = [g_{ij}]$ représentant un forme bilinéaire $g(\cdot,\cdot) \in \mathcal{L}(E, F; \mathbb{R})$ est donnée par

$$g_{ij} := g(\vec{a}_i, \vec{b}_j), \quad \text{donc} \quad [g(\cdot,\cdot)]_{|\vec{a},\vec{b}} := [g_{ij}]_{\substack{i=1,...,n \\ j=1,...,m}} = \begin{pmatrix} g_{11} & \cdots & g_{1m} \\ \vdots & & \\ g_{n1} & \cdots & g_{nm} \end{pmatrix} \overset{\text{noté}}{=} [g_{ij}], \tag{1.3}$$

Et, pour $\vec{u}=\sum_{i=1}^n u^i\vec{a}_i$ et $\vec{v}=\sum_{j=1}^m v^j\vec{b}_j$, la bilinéarité donne

$$g(\vec{u},\vec{v})=\sum_{i=1}^n\sum_{j=1}^m u^iv^jg(\vec{a}_i,\vec{b}_j),\quad \text{i.e.}\quad g(\vec{u},\vec{v})=[\vec{u}]_{|\vec{a}}^T.[g(\cdot,\cdot)]_{|\vec{a},\vec{b}}.[\vec{v}]_{|\vec{b}}. \tag{1.4}$$

Exemple 1.2 Dans $\mathbb{R}^2$ la matrice représentant le produit scalaire canonique dans la base canonique $(\vec{E}_i)$ est $[g]_{|\vec{E}}=\begin{pmatrix}1&0\\0&1\end{pmatrix}=I$ la matrice identité (donnée par $g(\vec{E}_i,\vec{E}_j)=\delta_{ij}$).

Ce même produit scalaire est représenté dans la base $(\vec{e}_1,\vec{e}_2)=(2\vec{E}_1,\vec{E}_2)$ par la matrice $[g]_{|\vec{e}}=\begin{pmatrix}4&0\\0&1\end{pmatrix}$ (donnée par $g_{ij}=g(\vec{e}_i,\vec{e}_j)$, e.g., $g_{11}=g(\vec{e}_1,\vec{e}_1)=4g(\vec{E}_1,\vec{E}_1)=4$).

On verra que les formules de changement de base pour les produits scalaires sont données par $[g]_{new}=P^T.[g]_{old}.[P]$ où P est la matrice de changement de base dont les colonnes sont données par les composantes de la nouvelle base dans l'ancienne. Soit ici $P=\begin{pmatrix}2&0\\0&1\end{pmatrix}$, et on retrouve bien $[g]_{|\vec{e}}=P^T.I.P=P^T.P$. ▪

1.2.2 Produit scalaire (et produits scalaires euclidiens)

Un produit scalaire sur E est une forme bilinéaire $g(\cdot,\cdot)=^{\text{noté}}(\cdot,\cdot)_g\in\mathcal{L}(E,E;\mathbb{R})$ qui est symétrique et définie positive, i.e., $g(\vec{u},\vec{v})=g(\vec{v},\vec{u})$ pour tout $\vec{u},\vec{v}\in E$, et $g(\vec{v},\vec{v})>0$ pour tout $\vec{v}\neq\vec{0}$.

La norme associée à un produit scalaire $g(\cdot,\cdot)$ est la fonction $||.||_g : E\to\mathbb{R}_+$ définie par

$$||\vec{v}||_g=\sqrt{g(\vec{v},\vec{v})}. \tag{1.5}$$

Représentation dans une base. Soit $(\vec{e}_i)$ une base de E. On note,

$$g_{ij}:=g(\vec{e}_i,\vec{e}_j)\quad\text{et}\quad [g]_{|\vec{e}}:=[g]_{|\vec{e},\vec{e}}=[g_{ij}]. \tag{1.6}$$

Donc si $\vec{u},\vec{v}\in E$, (1.4) donne

$$g(\vec{u},\vec{v})=\sum_{i,j=1}^n u^ig_{ij}v^j=[\vec{u}]_{|\vec{e}}^T.[g]_{|\vec{e}}.[\vec{v}]_{|\vec{e}}, \tag{1.7}$$

la dernière égalité avec les règles usuelles du produit matriciel.

Définition 1.3 Dans E muni d'un produit scalaire $g(\cdot,\cdot)$, une base $(\vec{a}_i)_{i=1,\dots,n}$ est dite orthonormée relativement à $g(\cdot,\cdot)$ ssi $g(\vec{a}_i,\vec{a}_j)=\delta_{ij}$ pour tout $i,j=1,\dots,n$, et alors $(\vec{a}_i)$ est dite être une $(\cdot,\cdot)_g$-b.o.n. (ou simplement une b.o.n. s'il n'y a pas d'ambiguïté).

Exercice 1.4 Montrer : si $g(\cdot,\cdot)$ est un produit scalaire, alors il existe une base $(\vec{a}_i)$ $(\cdot,\cdot)_g$-orthonormée.

Réponse. Soit $(\vec{e}_i)$ une base de E, et soit $M=[g]_{|\vec{e}}=[g(\vec{e}_i,\vec{e}_j)]=[g_{ij}]$, matrice symétrique (car $g(\cdot,\cdot)$ est symétrique) réelle. Donc M est diagonalisable dans une b.o.n. de $\mathbb{R}^n$ muni de sa base canonique $(\vec{E}_i)$ et de son produit scalaire canonique $(\cdot,\cdot)_{\mathbb{R}^n}$: il existe n réels $\lambda_1,\dots,\lambda_n$ (les valeurs propres) et n vecteurs $\vec{p}_1,\dots,\vec{p}_n$ formant une b.o.n. dans $\mathbb{R}^n$ (vecteurs propres associés) t.q. :

$$M.[\vec{p}_j]_{|\vec{E}}=\lambda_j[\vec{p}_j]_{|\vec{E}},\quad (\vec{p}_i,\vec{p}_j)_{\mathbb{R}^n}=\delta_{ij}=[\vec{p}_i]_{|\vec{E}}^T.[\vec{p}_j]_{|\vec{E}},$$

i.e.,

$$M.P=P.D\quad\text{et}\quad P^T.P=I,\quad\text{où}\quad P=[P_j^i]=\begin{pmatrix}[\vec{p}_1]_{|\vec{E}} & \cdots & [\vec{p}_n]_{|\vec{E}}\end{pmatrix}\quad\text{et}\quad D=\text{diag}(\lambda_1,\dots,\lambda_n).$$

Après ce passage dans $\mathbb{R}^n$: retour dans E. On définit alors la base $(\vec{b}_j)$ de E par

$$\vec{b}_j=\sum_i P_j^i\vec{e}_i\quad\in E.$$

Donc

$$g(\vec{b}_i,\vec{b}_j)=\sum_{k\ell}P_i^kP_j^\ell g(\vec{e}_k,\vec{e}_\ell)=\sum_{k\ell}(P^T)_k^iM_\ell^kP_j^\ell=(P^T.M.P)_j^i=D_j^i=\lambda_i\delta_j^i.$$

On pose $\vec{a}_i=\dfrac{\vec{b}_i}{\sqrt{\lambda_i}}$, on a donc $g(\vec{a}_i,\vec{a}_j)=\delta_{ij}$, vrai pour tout i,j. ▪

Définition 1.5 On appelle matrice de masse, ou matrice de Gram, la matrice d'un produit scalaire dans une base. I.e., si $g(\cdot,\cdot)$ est un produit scalaire et $(\vec{e}_i)$ est une base de E alors $[g]_{|\vec{e}}$ est une matrice de masse.

Exercice 1.6 1- Montrer qu'une matrice de masse est symétrique définie positive. (On rappelle qu'une matrice A est définie positive ssi $[\vec{x}]^T.A.[\vec{x}] > 0$ pour tout $[\vec{x}] \neq [\vec{0}]$.)

2- Et réciproquement, si A est une matrice $n*n$ symétrique définie positive, alors c'est la matrice d'un produit scalaire sur E.

Réponse. 1- Soit $M = [M_{ij}]$ une matrice de masse, i.e., il existe un produits scalaire $g(\cdot,\cdot)$ sur E et une base $(\vec{e}_i)$ dans E t.q. $M_{ij} = g(\vec{e}_i, \vec{e}_j)$. On a symétrie de g donc $g(\vec{e}_i, \vec{e}_j) = g(\vec{e}_j, \vec{e}_i)$, i.e. $M_{ij} = M_{ji}$. Puis, pour $\vec{v} \neq \vec{0}$, $0 < g(\vec{v}, \vec{v}) = [\vec{v}]^T.M.[\vec{v}]$, donc M est une matrice définie positive .

2- Réciproquement, on définit la forme bilinéaire $g(\cdot,\cdot)$ par $g(\vec{e}_i, \vec{e}_j) = A_{ij}$. Il est immédiat que g définit un produit scalaire quand A est symétrique définie positive.

Définition 1.7 En dimension finie, un espace vectoriel E muni d'un produit scalaire $g(\cdot,\cdot)$ est noté $(E, g(\cdot,\cdot))$ et est appelé un espace de Hilbert. Et s'il n'y a pas d'ambiguïté sur le produit scalaire on note simplement $(E, g(\cdot,\cdot)) = E$.

(En dimension infinie, $(E, g(\cdot,\cdot))$ est un Hilbert ssi E est de plus complet pour la norme $||.||_g$ associée au produit scalaire.)

Exercice 1.8 Soit $\mathbb{R}^n$ muni de son produit scalaire canonique $(\cdot,\cdot)_{\mathbb{R}^n}$, et soit $(\vec{a}_i)_{i=1,\dots,n}$ une base de $\mathbb{R}^n$. Soit A matrice $n*$n symétrique réelle. Montrer : on peut avoir $(A.\vec{a}_i, \vec{a}_i)_{\mathbb{R}^n} > 0$ pour tout i, et A non inversible, ou bien A a des valeurs propres < 0. Donc, il ne suffit pas de vérifier $(A.\vec{a}_i, \vec{a}_i)_{\mathbb{R}^n} > 0$ pour tout i pour avoir A définie positive : il faut vérifier $(A.\vec{x}, \vec{x})_{\mathbb{R}^n} > 0$ pour tout $\vec{x} \neq \vec{0}$.

Réponse. Dans $\mathbb{R}^2$, soit $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Alors $\lambda = 0$ est valeur propre (associé au vecteur propre $\vec{e}_2 = \vec{E}_2$) et A n'est pas inversible. Et on prend les deux vecteurs (indépendants) $\vec{a}_1$ et $\vec{a}_2$ donnés par $[\vec{a}_1]_{|\vec{E}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $[\vec{a}_2]_{|\vec{E}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. On a $[\vec{a}_k]^T.A.[\vec{a}_k] = (k \quad 1) . \begin{pmatrix} k \\ 0 \end{pmatrix} = k^2 > 0$ pour $k = 1, 2$.

Soit $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ (inversible), soit $\vec{a}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ et $\vec{a}_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$. On a $[\vec{a}_k]^T.A.[\vec{a}_k] = 3 > 0$ alors que $\lambda = -1$ est valeur propre et $(\vec{a}_1, \vec{a}_2)$ est une base de $\mathbb{R}^2$.

Exercice 1.9 Orthonormalisation de Gram–Schmidt : rappeler le procédé.

Réponse. Soit H espace vectoriel de dimension n muni d'un produit scalaire $(\cdot,\cdot)_H$. On cherche une b.o.n. $(\vec{b}_k)_{k=1,\dots,n}$ de $(H, (\cdot,\cdot)_H)$. On suppose qu'on connaît une famille $(\vec{a}_k)_{k=1,\dots,m}$ de vecteurs génératrice dans H (l'espace qu'elle engendre est H tout entier). Nécessairement $m \geq n$, et si $m > n$, alors la famille est liée, et si $m = n$ alors la famille $(\vec{a}_k)_{k=1,\dots,n}$ est une base de H (c'est souvent le cas : connaissant une base $(\vec{a}_i)$, on cherche une b.o.n.). Procédé d'orthonormalisation de Gram–Schmidt :

1) On choisit le premier vecteur non nul de la famille $\vec{a}_k$. Soit $\vec{a}_{k_1}$ ce vecteur, qu'on norme en posant :

$$\vec{b}_1 = \frac{\vec{a}_{k_1}}{||\vec{a}_{k_1}||_H}.$$

2) On choisit le premier vecteur suivant $\vec{a}_{k_1}$ qui est indépendant de $\vec{b}_1$. Soit $\vec{a}_{k_2}$ ce vecteur. On enlève sa composante suivant $\vec{b}_1$ et on norme :

$$\vec{p}_2 = \vec{a}_{k_2} - (\vec{a}_{k_2}, \vec{b}_1)_H \, \vec{b}_1, \qquad \vec{b}_2 = \frac{\vec{p}_2}{||\vec{p}_2||_H}.$$

Ainsi $(\vec{b}_2, \vec{b}_1)_H = 0 = (\vec{p}_2, \vec{b}_1)_H$ car $(\vec{p}_2, \vec{b}_1)_H = (\vec{a}_{k_2}, \vec{b}_1)_H - (\vec{a}_{k_2}, \vec{b}_1)_H (\vec{b}_1, \vec{b}_1)_H$, par bilinéarité du produit scalaire.

3) On choisit le premier vecteur suivant $\vec{a}_{k_2}$ qui est indépendant de $\vec{b}_1$ et $\vec{b}_2$. Soit $\vec{a}_{k_3}$ ce vecteur. On enlève sa composante sur Vect$\{\vec{b}_1, \vec{b}_2\}$ et on norme :

$$\vec{p}_3 = \vec{a}_{k_3} - (\vec{a}_{k_3}, \vec{b}_1)_H \, \vec{b}_1 - (\vec{a}_{k_3}, \vec{b}_2)_H \, \vec{b}_2, \qquad \vec{b}_3 = \frac{\vec{p}_3}{||\vec{p}_3||_H}.$$

On vérifie immédiatement que $(\vec{b}_3, \vec{p}_1)_H = 0 = (\vec{b}_3, \vec{p}_2)_H$.

$\vdots$

n) On poursuit jusqu'à $\vec{b}_n$: ainsi $(\vec{b}_k)_{k=1,\dots,n}$ est une b.o.n. de $(E, g(\cdot,\cdot))$.

1.2.3 Base euclidienne et produit scalaire euclidien

• Construction d'une **base euclidienne** :

1- on choisit une unité de mesure (par exemple le mètre, ou le pied anglais, ou une unité utilisée par Euclide...),

2- on prend des bâtons de longueur 3, 4 et 5 : il forment un triangle rectangle ($3^2+4^2=5^2=$ relation de Pythagore).

3- on dit que deux vecteurs $\vec{u}$ et $\vec{v}$ sont orthogonaux ssi on peut les mettre parallèles aux côtés de l'angle droit du triangle ci-dessus.

4- on choisit trois vecteurs $\vec{e}_1,\vec{e}_2,\vec{e}_3$ (dans $\mathbb{R}^3$) tous de longueur 1 (dans l'unité choisie) et 2 à 2 orthogonaux au sens de Pythagore : on a ainsi construit une base dite euclidienne dans $\mathbb{R}^3$,

• Un **produit scalaire euclidien** $g(\cdot,\cdot)=(\cdot,\cdot)_g$ est une forme bilinéaire définie dans $\mathbb{R}^3$ à l'aide d'une base euclidienne $(\vec{e}_i)$ (construite ci-dessus) et de $g(\vec{e}_i,\vec{e}_j)=\delta_{ij}$ pour tout $i,j=1,2,3$. Donc $g(\cdot,\cdot)$ (bilinéaire) est définie par $[g]_{|\vec{e}}=[\delta_{ij}]=I$ la matrice unité, et on a

$$\vec{u}=\sum_{i=1}^n u^i\vec{e}_i \quad\text{et}\quad \vec{v}=\sum_{i=1}^n v^i\vec{e}_i \quad\Longrightarrow\quad (\vec{u},\vec{v})_g:=g(\vec{u},\vec{v})=[\vec{u}]^T_{|\vec{e}}.[g]_{|\vec{e}}.[\vec{v}]_{|\vec{e}}=\sum_{i=1}^n u^iv^i, \tag{1.8}$$

et $||\vec{u}||_g=\sqrt{g(\vec{u},\vec{u})}=\sqrt{(u^1)^2+(u_2)^2}=$ la longueur (euclidienne) de $\vec{u}$.

Exemple 1.10 Expression d'un produit scalaire euclidien en coordonnées polaires (dans $\mathbb{R}^2$). Soit $(\vec{e}_1,\vec{e}_2)$ une base euclidienne (après choix d'une unité de mesure). On considère la fonction (le système de coordonnées polaires)

$$\vec{x}:\left\{\begin{array}{ll}\mathbb{R}_+\times[0,2\pi] & \to\mathbb{R}^2\\ (r,\theta) & \to\vec{x}=\vec{x}(r,\theta)=r\cos\theta\vec{e}_1+r\sin\theta\vec{e}_2\end{array}\right\},\quad\text{donc}\quad [\vec{x}(r,\theta)]_{|\vec{e}}=\begin{pmatrix}r\cos\theta\\ r\sin\theta\end{pmatrix}. \tag{1.9}$$

La base (holonome) du système polaire est donnée en $\vec{x}=\vec{x}(r,\theta)$ par les vecteurs (construction générique pour un système de coordonnées)

$$\vec{b}_1(\vec{x})=\frac{\partial\vec{x}}{\partial r}(r,\theta)=\cos\theta\vec{e}_1+\sin\theta\vec{e}_2,\quad\text{et}\quad \vec{b}_2(\vec{x})=\frac{\partial\vec{x}}{\partial\theta}(r,\theta)=-r\sin\theta\vec{e}_1+r\cos\theta\vec{e}_2, \tag{1.10}$$

i.e., $[\vec{b}_1(\vec{x})]_{|\vec{e}}=\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}$ et $[\vec{b}_2(\vec{x})]_{|\vec{e}}=\begin{pmatrix}-r\sin\theta\\ r\cos\theta\end{pmatrix}$. Donc

$$g(\vec{b}_1,\vec{b}_1)=1,\quad g(\vec{b}_1,\vec{b}_2)=0=g(\vec{b}_2,\vec{b}_1),\quad g(\vec{b}_2,\vec{b}_2)=r^2,\quad\text{et}\quad [g]_{|\vec{b}}=\begin{pmatrix}1&0\\0&r^2\end{pmatrix}. \tag{1.11}$$

C'est l'expression usuelle d'une métrique euclidienne en coordonnées polaires. Et

$$\vec{u}=u^1\vec{b}_1+u^2\vec{b}_2\quad\text{et}\quad\vec{v}=v^1\vec{b}_1+v^2\vec{b}_2\quad\text{donnent}\quad g(\vec{u},\vec{v})=[\vec{u}]^T_{|\vec{b}}.[g]_{|\vec{b}}.[\vec{v}]_{|\vec{b}}=u_1v_1+r^2u_2v_2, \tag{1.12}$$

et $||\vec{u}||_g=\sqrt{g(\vec{u},\vec{u})}=\sqrt{(u^1)^2+r^2(u_2)^2}=$ la longueur (euclidienne) de $\vec{u}$ exprimé dans la base polaire (relative l'unité euclidienne choisie). ▞

1.2.4 Métrique (dans un espace affine)

"Une métrique est un produit scalaire considéré aux points d'un espace affine" :

Soit $\mathcal{E}$ un espace affine (dans le cadre de la mécanique classique : l'espace des points de l'univers à trois dimensions dans lequel nous vivons). Soit E l'espace vectoriel associé (vu comme l'espace des vecteurs bi-points $\overrightarrow{AB}$ pour $A,B\in\mathcal{E}$).

Définition 1.11 Une métrique dans $\mathcal{E}$ est une application g qui à un point P de l'espace affine $\mathcal{E}$ associe un produit scalaire dans E. Donc $g:\left\{\begin{array}{ll}\mathcal{E}&\to\mathcal{L}(E,E;\mathbb{R})\\ P&\to g_P(\cdot,\cdot)\end{array}\right\}$ est telle que $g_P(\cdot,\cdot)$ est un produit scalaire dans E. Autrement dit, définir une métrique dans l'espace affine $\mathcal{E}$ c'est définir un produit scalaire en chaque point $P\in\mathcal{E}$.

Exemple 1.12 Dans notre espace affine à trois dimensions, la métrique euclidienne est définie par : en chaque point c'est le produit scalaire euclidien. Et ce produit scalaire euclidien peut être exprimé dans la base qui convient, par exemple une base euclidienne, ou par exemple une base polaire, voir exemple 1.10. ▞

1.2.5 Métrique sur une surface d'un espace affine

Définition rapide : Si S est une surface dans $\mathbb{R}^n$ affine, alors une métrique dans S est la restriction à S d'une métrique dans $\mathbb{R}^n$.

Mais supposons "qu'on ne peut pas sortir de la surface", e.g., sur terre on ne peut pas prendre de "hauteur" (on ne peut pas s'envoler). On ne dispose donc que des vecteurs tangents à la surface. Il faut donc définir la métrique "dans la surface", uniquement à l'aide des vecteurs tangents à la surface.

Démarche : on se donne un paramétrage de la surface et on utilise les vecteurs tangents $(\vec{e}_i(P))$ en chaque point P, pour définir un produit scalaire $g_P(\cdot,\cdot)$: donne la métrique dans la surface.

Exemple 1.13 Dans $\mathbb{R}^2$ affine et coordonnées polaires : on considère un repère $(O,(\vec{e}_1,\vec{e}_2)$, où $(\vec{e}_1,\vec{e}_2)$ est une base euclidienne, et on considère le cercle de centre (a,b) et de rayon R décrit par la fonction

$$\vec{x}:\left\{\begin{array}{ll}[0,2\pi] & \to \mathbb{R}^2\\ \qquad\theta & \to \vec{x}=\vec{x}(\theta)=(a+R\cos\theta)\vec{e}_1+(b+R\sin\theta)\vec{e}_2\end{array}\right\},\quad \text{i.e.}\quad [\vec{x}(\theta)]_{|\vec{e}}=\begin{pmatrix}a+R\cos\theta\\ b+R\sin\theta\end{pmatrix}. \tag{1.13}$$

(Exemple : on est sur une piste circulaire, et on n'a pas le droit de sortir de la piste.) Un vecteur tangent en $\vec{x}=\vec{x}(\theta)$ est (démarche générique)

$$\vec{b}(\vec{x})=\frac{\partial\vec{x}}{\partial\theta}(\theta)=-R\sin\theta\vec{e}_1+R\cos\theta\vec{e}_2,\quad \text{i.e.}\quad [\vec{b}]_{|\vec{e}}=\begin{pmatrix}-R\sin\theta\\ R\cos\theta\end{pmatrix}. \tag{1.14}$$

Ce vecteur $\vec{b}(\vec{x})$ défini une base en $\vec{x}$ sur la droite affine tangente au cercle (c'est la seule direction qu'on peut prendre quand on est en $\vec{x}$ et qu'on veut rester sur le cercle). Et la métrique (euclidienne) dans le cercle est donnée en $\vec{x}$ par

$$g_{\vec{x}}(\vec{b}(\vec{x}),\vec{b}(\vec{x}))=g(\begin{pmatrix}-R\sin\theta\\ R\cos\theta\end{pmatrix},\begin{pmatrix}-R\sin\theta\\ R\cos\theta\end{pmatrix})=R^2,\quad \text{et}\quad [g_{\vec{x}}]_{|\vec{b}(\vec{x})}=[R^2]\quad (\text{matrice } 1{*}1) \tag{1.15}$$

(le vecteur $\vec{b}(\vec{x})$ a pour longueur $\sqrt{g_{\vec{x}}(\vec{b}(\vec{x}),\vec{b}(\vec{x}))}=R$). Cette métrique sert par exemple à calculer une vitesse : exemple de la courbe paramétrée qui repère un coureur :

$$\vec{r}:\left\{\begin{array}{ll}[0,T] & \to \mathbb{R}^2\\ \qquad t & \to \vec{x}=\vec{r}(t)=(a{+}R\cos(\theta(t))\vec{e}_1+(b{+}R\sin(\theta(t)))\vec{e}_2\end{array}\right\},\quad \text{i.e.}\quad [\vec{r}(t)]_{|\vec{e}}=\begin{pmatrix}a{+}R\cos(\theta(t))\\ b{+}R\sin(\theta(t))\end{pmatrix}, \tag{1.16}$$

alors sa vitesse en $\vec{x}=\vec{r}(t)$ est

$$\vec{r}\,'(t)=R\theta'(t)(-\sin(\theta(t))\vec{e}_1+\cos(\theta(t)))\vec{e}_2=\theta'(t)\vec{b}(\vec{x}), \tag{1.17}$$

de longueur $||\vec{r}\,'(t)||=|\theta'(t)|\,||\vec{b}(\vec{x})||=R|\theta'(t)|=$ vitesse du coureur quand il se trouve en $\vec{x}=\vec{r}(t)$. ▪

1.2.6 * Relativité restreinte et pseudo produit scalaire de $\mathbb{R}^4$ (Minkowski)

L'espace-temps est ici approximé par l'espace vectoriel $\mathbb{R}^4=\mathbb{R}\times\mathbb{R}^3=$ temps $\times$ espace.

Quand deux observateurs, se déplaçant l'un par rapport à l'autre, "mesurent un même vecteur", ils comparent leurs mesures à l'aide des formules de changement de base données par les transformations de Lorentz. Ces formules ne sont pas des changements de bases orthonormaux usuel dans $\mathbb{R}^4$: ils ne sont ni normés ni orthogonaux au sens d'un produit scalaire $(\cdot,\cdot)_{\mathbb{R}^4}$ de $\mathbb{R}^4$

Exemple : Soit un référentiel $K=(O,\vec{e}_0,\vec{e}_1)$ et un référentiel $K'=(O',\vec{e}_0{}',\vec{e}_1{}')$ (l'indice 0 se réfère à une base temporelle et l'indice 1 se réfère à une base spatiale), le référentiel K' se déplaçant à la vitesse $\vec{v}=v\vec{e}_1$ dans K (déplacement en espace), avec $|v|<c$ où c est la vitesse de la lumière. On note :

$$\beta=\frac{v}{c}\ (<1)\qquad \text{et}\qquad \gamma=\frac{1}{\sqrt{1-\beta^2}}\ (>1), \tag{1.18}$$

soit $\gamma=\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$. La matrice P de passage de la base $(\vec{e}_i)$ vers la base $(\vec{e}_i{}')$ est

$$P=\gamma\begin{pmatrix}1&+\beta\\+\beta&1\end{pmatrix},\quad \text{donc}\quad [\vec{e}_0{}']_{|\vec{e}}=\gamma\begin{pmatrix}1\\ \beta\end{pmatrix},\quad [\vec{e}_1{}']_{|\vec{e}}=\gamma\begin{pmatrix}\beta\\ 1\end{pmatrix}, \tag{1.19}$$

soit $\vec{e}_0{}'=\gamma(\vec{e}_0+\beta\vec{e}_1)$ et $\vec{e}_1{}'=\gamma(\beta\vec{e}_0+\vec{e}_1)$. (Remarque : $\det(P)=1$: conservation des "volumes de l'espace temps".) Donc, avec $\vec{v} = t\vec{e}_0 + x\vec{e}_1 = t'\vec{e}_0{}' + x'\vec{e}_1{}'$,

$$P^{-1} = \gamma \begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} ct' \\ x' \end{pmatrix} = P^{-1}. \begin{pmatrix} ct \\ x \end{pmatrix}, \quad \text{i.e.} \quad \begin{cases} ct' = \gamma(ct - \frac{v}{c}\,x), \\ x' = \gamma(x - \frac{v}{c}\,ct). \end{cases} \tag{1.20}$$

P est symétrique mais $P^{-1} \neq P^T$ (attendu : on n'utilise pas le produit scalaire euclidien usuel, mais on utilise le pseudo-produit scalaire de Lorentz–Minkowski).

Dans $\mathbb{R}^4 = \mathbb{R} \times \mathbb{R}^3$ et sa base canonique $(\vec{E}_0, \vec{E}_1, \vec{E}_2, \vec{E}_3)$, toujours pour $\vec{v} = v\vec{e}_1$, comme y et z sont inchangés, la transformation se lit :

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = P^{-1}. \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}, \qquad P^{-1} = \begin{pmatrix} \gamma & -\gamma\frac{v}{c} & 0 & 0 \\ -\gamma\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.21}$$

Minkowski garde comme postulat, dans un référentiel donné (relativité restreinte) :

« à chaque instant t, l'espace géométrique est l'espace $\mathbb{R}^3$ usuel ».

Et il introduit un pseudo-produit scalaire $\eta(\cdot,\cdot)$ dans lequel la transformation de Lorentz est "pseudo-orthonormale", i.e. la matrice de passage P donnée en (1.19) est "pseudo-orthonormale" : notant $\vec{X} = (x^0, \vec{x}) = (ct, x^1, x^2, x^3) \in \mathbb{R} \times \mathbb{R}^3$, le pseudo-produit de Minkowski est la forme bilinéaire symétrique non dégénérée $\eta(\cdot,\cdot)$ (i.e., t.q. : si $\forall\, \vec{X}$ on a $\eta(\vec{X}, \vec{Y}) = 0$ alors $\vec{Y} = \vec{0}$) donnée par :

$$\eta(\vec{X}, \vec{Y}) = x^0y^0 - (\vec{x}, \vec{y})_{\mathbb{R}^3}, \quad [\eta]_{|\vec{E}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \overset{\text{noté}}{=} \begin{pmatrix} 1 & 0 \\ 0 & -I \end{pmatrix}, \tag{1.22}$$

où I est la matrice identité de l'espace géométrique $\mathbb{R}^3$. La pseudo-norme de Minkowski est :

$$||\vec{X}||_\eta = \sqrt{\eta(\vec{X}, \vec{X})} = \sqrt{(x^0)^2 - ||\vec{x}||^2_{\mathbb{R}^3}} \qquad (= \sqrt{c^2t^2 - ||\vec{x}||^2_{\mathbb{R}^3}}), \tag{1.23}$$

où en particulier $||\vec{X}||_\eta = 0$ quand $||\vec{x}||_{\mathbb{R}^3} = c|t|$, i.e. quand la particule atteint la vitesse de la lumière. Et $||\vec{X}||_\eta \in \mathbb{R}_+$ (non complexe) car aucune particule ne dépasse la vitesse de la lumière (dans la théorie de la relativité restreinte).

Proposition 1.14 *La transformation de Lorentz (1.20) conserve le pseudo-produit scalaire :*

$$\eta(\vec{X}, \vec{Y}) = \eta(\vec{X}', \vec{Y}'). \tag{1.24}$$

Donc le changement de base conserve la pseudo-orthonormalité : $\vec{X} \perp_\eta \vec{Y}$ ssi $\eta(\vec{X}, \vec{Y}) = 0$ ssi $\eta(\vec{X}', \vec{Y}') = 0$ ssi $\vec{X}' \perp_\eta \vec{Y}'$. Et conserve les pseudo-normes : $||\vec{X}||_\eta = ||\vec{X}'||_\eta$.

Preuve. La vérification est immédiate : $(ct')^2 - (x')^2 = \gamma^2((c^2t^2 + \beta^2x^2 - 2vxt - x^2 - \beta^2c^2t^2 + 2vxt) = \gamma^2((1-\beta^2)c^2t^2 - (1-\beta^2)x^2) = c^2t^2 - x^2$. ∎

Et la matrice de changement de base P donnée en (1.19) est bien pseudo-orthonormée, au sens où pour le pseudo-produit scalaire de Minkowski on a :

$\eta(\vec{e}_0{}', \vec{e}_1{}') = \gamma^2(\frac{v}{c} - \frac{v}{c}) = 0$ ("pseudo-orthogonalité temps-espace"),

$\eta(\vec{e}_0{}', \vec{e}_0{}') = \gamma^2(1 - \frac{v^2}{c^2}) = 1$ (pseudo-normée "en temps"), et

$\eta(\vec{e}_1{}', \vec{e}_1{}') = \gamma^2(\frac{v^2}{c^2} - 1) = -1$ (pseudo-normée "en espace").

Remarque 1.15 La représentation de Minkowski est une représentation sous forme hyperbolique $(x^0)^2 - x^2$, soit $c^2t^2 - x^2$ (physiquement "non observable"), en complément de la représentation sous forme elliptique précédente $(x^0)^2 + x^2$, soit $c^2t^2 + x^2$ adaptée à la mesure de vitesse et d'accélération (observable au sens de la mesure newtonienne).

Une représentation hyperbolique est souvent utilisée pour représenter des lois de conservation, ce qui est le cas de la représentation de Minkowski. ∎

Remarque 1.16 Suivant les auteurs, le signe peut changer : $\eta(\vec{X}, \vec{Y}) = -x_0y_0 + (\vec{x}, \vec{y})_{\mathbb{R}^3}$. ∎

1.2.7 * Continuité d'une forme bilinéaire

Définition 1.17 Une forme bilinéaire $g(\cdot,\cdot) : E \times F \to \mathbb{R}$ est dite continue ssi :

$$\exists c > 0, \quad \forall (\vec{x},\vec{y}) \in E \times F, \quad \text{on a} \quad |g(\vec{x},\vec{y})| \leq c\, ||\vec{x}||_E ||\vec{y}||_F. \tag{1.25}$$

Et on note $||g||$ la plus petite constante c qui convient, i.e., $||g|| = \sup_{||\vec{x}||=||\vec{y}||=1} |g(\vec{x},\vec{y})|$.

Par bilinéarité de $g(\cdot,\cdot)$ la propriété (1.25) est équivalente à :

$$\begin{aligned} &\exists c > 0, \quad \forall (\vec{x},\vec{y}) \in E \times F \text{ t.q. } ||\vec{x}||_E \leq 1 \text{ et } ||\vec{y}||_F \leq 1, \quad \text{on a} \quad |g(\vec{x},\vec{y})| \leq c, \\ &\exists c > 0, \quad \forall (\vec{x},\vec{y}) \in E \times F \text{ t.q. } ||\vec{x}||_E = 1 \text{ et } ||\vec{y}||_F = 1, \quad \text{on a} \quad |g(\vec{x},\vec{y})| \leq c, \end{aligned} \tag{1.26}$$

i.e. $g(\cdot,\cdot)$ bornée "sur le carré unité" ou bornée "sur le bord du carré unité".

Proposition 1.18 *En dimension finie, toute forme bilinéaire est continue. Faux en dimension infinie.*

Preuve. Dans E, soit une base $(\vec{a}_i)_{i=1,\dots,m}$ de E. Les normes étant équivalentes (on est en dimension finie), choisissons la norme $||.||_\infty$ donnée par, quand $\vec{x} = \sum_{i=1}^m x_i \vec{a}_i$:

$$||\vec{x}||_\infty = \sup_{i=1,\dots,m} |x_i|.$$

C'est bien une norme (1- définie positive : immédiat ; 2- homogène : immédiat, 3- inégalité triangulaire : immédiat puisque $\sup_{i=1,\dots,m} |x_i+y_i| \leq \sup_{i=1,\dots,m} |x_i| + \sup_{j=1,\dots,m} |y_j|$). Idem dans F : soit une base $(\vec{b}_i)_{i=1,\dots,n}$ et la norme encore notée $||.||_\infty$. On a $g(\vec{x},\vec{y}) = \sum_{ij} x_i g_{ij} y_j$ où $[g]_{|\vec{a},\vec{b}} = [g_{ij} = g(\vec{a}_i,\vec{b}_j)]$, et donc :

$$|g(\vec{x},\vec{y})| \leq (\sup_i |x_i|)(\sup_j |y_j|) \sum_{ij} |g_{ij}|,$$

et la constante $c = \sum_{ij} |g_{ij}|$ dans (1.25) convient. Pour la dimension infinie, voir exercice 1.26. ∎

Proposition 1.19 *Quelle que soit la dimension (finie ou infinie), si $g(\cdot,\cdot) =^{\text{noté}} (\cdot,\cdot)_g$ est un produit scalaire sur E et $||.||_g$ est sa norme, alors le produit scalaire $g(\cdot,\cdot)$ est continu relativement à sa norme associée $||.||_g$. On dit qu'un produit scalaire est "continu par rapport à lui-même".*

Preuve. Cauchy–Schwarz donne $|(\vec{x},\vec{y})_g| \leq ||\vec{x}||_g ||\vec{y}||_g$ pour tout $\vec{x},\vec{y} \in E$: donc $c = 1$ convient dans (1.25). ∎

Exercice 1.20 Soit $g(\cdot,\cdot)$ bilinéaire. À $\vec{x} \in E$ fixé, on note $g_{\vec{x}} : \vec{y} \in F \to g_{\vec{x}}(\vec{y}) =^{\text{déf}} g(\vec{x},\vec{y}) \in \mathbb{R}$ (forme linéaire : trivial). De même, à $\vec{y} \in F$ fixé, on note $g_{\vec{y}} : \vec{x} \in E \to g_{\vec{y}}(\vec{x}) =^{\text{déf}} g(\vec{x},\vec{y}) \in \mathbb{R}$ (forme linéaire : trivial). 1-Montrer que si $g(\cdot,\cdot)$ est continue (cf. (1.25)) alors :

$$\forall \vec{x} \in E \text{ et } \forall \vec{y} \in F, \text{ les formes linéaires } g_{\vec{x}} \text{ et } g_{\vec{y}} \text{ sont continues.} \tag{1.27}$$

(On rappelle qu'une forme linéaire $\ell : E \to \mathbb{R}$) est continue ssi $\exists \gamma > 0, \forall \vec{x} \in E$, on a $||\ell(\vec{x})||_E < \gamma\, ||\vec{x}||_E$; et on note $||\ell||$ la plus petite des constantes γ.)

2- Montrer que la réciproque est vraie si E (ou F) est de dimension finie.

Réponse. 1- Si $g(\cdot,\cdot)$ est continue, cf. (1.25), $|g(\vec{x},\vec{y})| \leq ||g||\, ||\vec{x}||_E\, ||\vec{y}||_F$, donc à $\vec{y} \in F$ fixé, $|g(\vec{x},\vec{y})| \leq \gamma ||\vec{x}||_E$ avec $\gamma = ||g||\, ||\vec{y}||_F$, donc $g_{\vec{x}}$ est continu, et à $\vec{x} \in E$ fixé, $|g(\vec{x},\vec{y})| \leq \gamma ||\vec{y}||_F$ avec $\gamma = ||g||\, ||\vec{x}||_E$, donc $g_{\vec{y}}$ est continu. D'où (1.27).

2- Soit $(\vec{a}_i)_{i=1,\dots,n}$ une base de E, soit $\vec{x} \in E$, $\vec{x} = \sum_{i=1}^n x^i \vec{a}_i$, et on prend $||\vec{x}|_E = \sum_{i=1}^n (x^i)^2$ (toutes les normes sont équivalentes en dimension finie). Alors $g(\vec{x},\vec{y}) = \sum_{i=1}^n x^i g(\vec{e}_i,\vec{y})$, d'où (avec Cauchy–Schwarz dans $\mathbb{R}^n$) $|g(\vec{x},\vec{y})| \leq (\sum_{i=1}^n (x^i)^2)(\sum_{i=1}^n g(\vec{a}_i,\vec{y})^2)^{\frac12} = ||\vec{x}||_E (\sum_{i=1}^n g_{\vec{a}_i}(\vec{y})^2)^{\frac12} \leq ||\vec{x}||_E (\sum_{i=1}^n ||g_{\vec{a}_i}||^2 ||\vec{y}||_F^2)^{\frac12} \leq c\, ||\vec{x}||_E ||\vec{y}||_F$ où on a posé $c = (\sum_{i=1}^n ||g_{\vec{a}_i}||^2)^{\frac12}$. D'où (1.25). ∎

Exercice 1.21 Montrer que : $f : (\vec{v},\vec{w}) \in E \times F \to f(\vec{v},\vec{w}) = ||\vec{v}||_E ||\vec{w}||_F$ n'est pas une norme sur $E \times F$. Pas plus que $g : (\vec{v},\vec{w}) \in E \times F \to g(\vec{v},\vec{w}) = \sqrt{||\vec{v}||_E ||\vec{w}||_F}$.

Réponse. On prend $\vec{v} \neq \vec{0}$ et $\vec{w} = 0$, et on obtient $f(\vec{v},\vec{w}) = 0$ bien que $(\vec{v},\vec{w})$ est non nul.

(On n'a pas non plus l'homogénéité, car $f(\lambda(\vec{v},\vec{w})) = |\lambda|^2 f(\vec{v},\vec{w})$, et on n'a pas non plus l'inégalité triangulaire, car $f((\vec{v}_1,\vec{w}_1)+(\vec{v}_2,\vec{w}_2)) = ||\vec{v}_1+\vec{v}_2||_E ||\vec{w}_1+\vec{w}_2||_F$, alors que $f(\vec{v}_1,\vec{w}_1)+f(\vec{v}_2,\vec{w}_2) = ||\vec{v}_1||_E ||\vec{w}_1||_F + ||\vec{v}_2||_E ||\vec{w}_2||_F$, et si on prend $\vec{v}_1 = \vec{0}$ et $\vec{w}_2 = \vec{0}$, on a $f((\vec{0},\vec{w}_1)+(\vec{v}_2,\vec{0})) \not\leq f(\vec{0},\vec{w}_1)+f(\vec{v}_2,\vec{0})$ car $||\vec{v}_2||_E ||\vec{w}_1||_F \not\leq 0$ en général.)

Pour g on n'a ni la nullité ni l'inégalité triangulaire (adaptée la réponse pour f). ∎

On munit le produit cartésien $Z = E \times F$ de la norme $||.||_{E\times F}$ (voir exercice 1.23) définie par :

$$||\vec{z}||_Z = ||\vec{z}||_{E\times F} := \sup(||\vec{x}||_E, ||\vec{y}||_F) \quad \text{quand} \quad \vec{z} = (\vec{x}, \vec{y}). \tag{1.28}$$

Proposition 1.22 *Soit $g(\cdot,\cdot)$ une forme bilinéaire sur $E \times F$.*

Quand $g(\cdot,\cdot)$ est continue (au sens (1.25)), alors l'application $f : Z \to \mathbb{R}$ définie par $f(\vec{z}) = g(\vec{x}, \vec{y})$ (qui n'est pas linéaire) est continue sur $(Z, ||.||_Z)$. (Donc f est toujours continue en dimension finie.)

N.B. : penser à la forme bilinéaire = le produit simple dans $\mathbb{R}$ donné par $(x, y) \in \mathbb{R}\times\mathbb{R} \to xy \in \mathbb{R}$, associée la fonction $f : (x, y) \in \mathbb{R}^2 \to xy \in \mathbb{R}$ dont le graphe $G(f) = \{(x, y, f(x, y)) \in \mathbb{R}^3\}$ est la "selle de cheval" : $f(x, y) = 0$ si $x = 0$ ou $y = 0$, $f(x, x) = x^2$ (parabole "vers le haut") et $f(x, -x) = -x^2$ (parabole "vers le bas").

Preuve. Montrons la continuité en un point $\vec{z}_0 = (\vec{x}_0, \vec{y}_0) \in Z$. Soit $\vec{z} = (\vec{x}, \vec{y}) \in Z$. On a :

$$\begin{aligned}
|f(\vec{z}) - f(\vec{z}_0)| &= |g(\vec{x}, \vec{y}) - g(\vec{x}_0, \vec{y}_0)| = |g(\vec{x} - \vec{x}_0, \vec{y}) + g(\vec{x}_0, \vec{y} - \vec{y}_0)| \\
&= |g(\vec{x} - \vec{x}_0, \vec{y} - \vec{y}_0) + g(\vec{x} - \vec{x}_0, \vec{y}_0) + g(\vec{x}_0, \vec{y} - \vec{y}_0)| \\
&\le ||g||\,||\vec{x} - \vec{x}_0||_E||\vec{y} - \vec{y}_0||_F + ||g||\,||\vec{x} - \vec{x}_0||_E||\vec{y}_0||_F + ||g||\,||\vec{x}_0||_E||\vec{y} - \vec{y}_0||_F \\
&\le ||g||\sup(1, ||\vec{y}_0||_F, ||\vec{x}_0||_E)\sup(||\vec{x} - \vec{x}_0||_E, ||\vec{y} - \vec{y}_0||_F).
\end{aligned}$$

Donc quand $\vec{z} \to \vec{z}_0$, i.e. quand $||\vec{z} - \vec{z}_0||_\infty \to 0$, i.e. quand $\sup(||\vec{x} - \vec{x}_0||_E, ||\vec{y} - \vec{y}_0||_F) \to 0$, on a $|f(\vec{z}) - f(\vec{z}_0)| \to 0$, et donc f est continue en $\vec{z}_0$. Ce pour tout $\vec{z}_0$, donc f est continue sur Z. ∎

Exercice 1.23 Vérifier que $||.||_Z = ||.||_{E\times F}$ défini en (1.28) est bien une norme.

Réponse. Positivité : immédiat. Stricte positivé :$||(\vec{x}, \vec{y})||_{E\times F} = 0$ implique $(\vec{x}, \vec{y})$ nul : immédiat. Homogénéité : $||\lambda(\vec{x}, \vec{y})||_{E\times F} = ||(\lambda\vec{x}, \lambda\vec{y})||_{E\times F} = |\lambda|\,||(\vec{x}, \vec{y})||_{E\times F}$: immédiat. Inégalité triangulaire : on a :

$$\begin{aligned}
||(\vec{x}_1, \vec{y}_1) + (\vec{x}_2, \vec{y}_2)||_{E\times F} &= ||(\vec{x}_1 + \vec{x}_2, \vec{y}_1 + \vec{y}_2)||_{E\times F} = \sup(||\vec{x}_1 + \vec{x}_2||_E, ||\vec{y}_1 + \vec{y}_2||_F) \\
&\le \sup(||\vec{x}_1||_E + ||\vec{x}_2||_E, ||\vec{y}_1||_F + ||\vec{y}_2||_F).
\end{aligned}$$

Et est-ce :

$$\overset{?}{\le} ||(\vec{x}_1, \vec{y}_1)||_{E\times F} + ||(\vec{x}_2, \vec{y}_2||_{E\times F} = \sup(||\vec{x}_1||_E, ||\vec{y}_1||_F) + \sup(||\vec{x}_2||_E, ||\vec{y}_2||_F)\ ?$$

Autrement dit, a-t-on dans $\mathbb{R}$ (on pose $X_1 = ||\vec{x}_1||_E$, $Y_1 = ||\vec{x}_2||_E$, $X_2 = ||\vec{y}_1||_E$, $Y_2 = ||\vec{y}_2||_E$) :

$$\sup(|X_1| + |Y_1|, |X_2| + |Y_2|) \overset{?}{\le} \sup(|X_1|, |X_2|) + \sup(|Y_1|, |Y_2|)\ ? \tag{1.29}$$

pour des réels X_i, Y_i, pour $i = 1, 2$. Soit, notant $\vec{X} = (X_1, X_2)$ et $||\vec{X}||_\infty = \sup(X_1, X_2)$ et $\vec{Y} = (Y_1, Y_2)$:

$$||\vec{X} + \vec{Y}||_\infty \le ||\vec{X}||_\infty + ||\vec{Y}||_\infty\ ? \tag{1.30}$$

Mais $||.||_\infty$ est une norme dans $\mathbb{R}^2$, donc vérifie l'inégalité triangulaire (voir exercice suivant). Cqfd. ∎

Exercice 1.24 Rappeler, dans $\mathbb{R}^n$, la démonstration de l'inégalité triangulaire (1.30).

Réponse. On a $|X_i| + |Y_i| \le \sup_{j=1,...,n}(|X_i|) + \sup_{j=1,...,n}(|Y_j|)$, pour tout $i = 1, ..., n$, donc le sup du membre de gauche est inférieur au membre de droite, soit (1.29). ∎

Exercice 1.25 Donner un exemple en dimension infinie où deux normes ne sont pas équivalentes.

Réponse. On rappelle que dans $\mathbb{R}^n$ on note $||\vec{x}||_1 = \sum_{i=1}^n |x_i|$ et $||\vec{x}||_\infty = \sup_{i=1,...,}|x_i|$, et que $||.||_1$ et $||.||_\infty$ sont deux normes, et que $||\vec{x}||_1 \le n||\vec{x}||_\infty$ pour tout $\vec{x}$ (immédiat), l'égalité étant obtenue lorsque $x_i = x_j$ pour tout i, j. On "voit" qu'il y a un problème quand $n \to \infty$:

Soit $S = \{\vec{x} = (x_n)_{\mathbb{N}^*} \in \mathbb{R}^{\mathbb{N}^*}\}$ l'ensemble des suites $\vec{x} = (x_n)_{n\in\mathbb{N}^*}$, soit $||\vec{x}||_1 = \sum_{i=1}^\infty |x_i|$, soit $||\vec{x}||_\infty = \sup_{i\in\mathbb{N}^*}|x_i|$, et soit T le sous-ensemble de S des suites t.q. $||\vec{x}||_1 < \infty$ et $||\vec{x}||_\infty < \infty$. Si les normes $||.||_1$ et $||.||_\infty$ étaient équivalentes dans T on aurait :

$$\exists c > 0, \quad \forall\vec{x} \in T, \quad ||\vec{x}||_1 \le c||\vec{x}||_\infty. \tag{1.31}$$

Montrons le contraire : on prend la suite $\vec{x}^{(N)} = (1, ..., 1, 0, ...)$ dont les N premiers termes valent 1 et tous les suivants valent 0. Il est immédiat que $\vec{x}^{(N)} \in T$, ce quel que soit N. Et on a $||\vec{x}^{(N)}||_1 = N$ et $||\vec{x}^{(N)}||_\infty = 1$. Donc, quitte à prendre N assez grand on a :

$$\forall c > 0, \quad \exists\vec{x} \in T, \quad \text{prendre } \vec{x}^{(N)} \text{ avec } N > c, \quad \text{t.q.} \quad ||\vec{x}||_1 > c||\vec{x}||_\infty. \tag{1.32}$$

Donc les normes ne sont pas équivalentes. ∎

Exercice 1.26 Donner en dimension infinie une forme bilinéaire qui n'est pas continue.

Réponse. Soit $\ell^2 = \{ \vec{x} \in \mathbb{R}^{\mathbb{N}^*} : \sum_{i=1}^\infty x_i^2 < \infty\}$ l'ensemble des suites de carrés sommables muni de sa norme usuelle $||\vec{x}||_{\ell^2} = (\sum_{i=1}^\infty x_i^2)^{\frac{1}{2}}$.

Soit $a(\cdot,\cdot)$ donnée par "sa matrice généralisée" $A = \mathrm{diag}(1,2,...,n,...)$, i.e. par $a(\vec{E}_n,\vec{E}_m) = n\delta_{nm}$ pour tout n,m, où $(\vec{E}_n)_{n\in\mathbb{N}^*}$ est la base canonique de ℓ^2.

$a(\vec{E}_n,\vec{E}_n) = n \underset{n\to\infty}{\longrightarrow} \infty$, avec $||\vec{E}_n||_{\ell^2} = 1$, et donc $a(\cdot,\cdot)$ n'est pas continue relativement à la norme $||.||_{\ell^2}$. ▪

1.3 Applications linéaires, formes linéaires, endomorphismes

1.3.1 Définitions

Soit E et F deux espaces vectoriels réels, de dimensions finies respectives n et m.

Définition 1.27 1- Une application linéaire L de E dans F est une application $L \in \mathcal{F}(E;F)$ telle que :

$$\forall \vec{u},\vec{v} \in E, \quad \forall \lambda \in \mathbb{R}, \quad L(\vec{u}+\lambda\vec{v}) = L(\vec{u}) + \lambda L(\vec{v}). \tag{1.33}$$

Et on note alors, pour tout $\vec{u} \in E$:

$$L(\vec{u}) = L.\vec{u}, \tag{1.34}$$

notation du "produit" puisqu'on a la "distributivité", cf. (1.33). On note

$$\mathcal{L}(E;F) := \{\text{l'ensemble des applications linéaires de } E \text{ dans } F\}. \tag{1.35}$$

2- Cas particulier $F = \mathbb{R}$ (ou $\mathbb{C}$) : une application linéaire $\ell \in \mathcal{L}(E;\mathbb{R})$ à valeurs réelles (ou $\ell \in \mathcal{L}(E;\mathbb{C})$ à valeurs complexes) est appelée une forme linéaire. Et on note

$$E^* := \mathcal{L}(E;\mathbb{R}), \quad \text{appelé le dual de } E. \tag{1.36}$$

3- Cas particulier $F = E$: une application linéaire $L \in \mathcal{L}(E;E)$ est appelé un endomorphisme de E. Et on note

$$\mathcal{L}(E) := \mathcal{L}(E;E). \tag{1.37}$$

Ayant $\mathcal{L}(E;F) \subset \mathcal{F}(E;F)$, on munit $\mathcal{L}(E;F)$ des opérations induites d'addition interne ($f+g$ est définie par $(f+g)(\vec{u}) := f(\vec{u}) + g(\vec{u})$) et de multiplication externe par un scalaire (λf est définie par $(\lambda f)(\vec{u}) := \lambda\, f(\vec{u})$).

Proposition 1.28 *$\mathcal{L}(E;F)$ est un espace vectoriel, sous-espace vectoriel de $\mathcal{F}(E;F)$.*

Preuve. Montrons que c'est un sous-espace vectoriel de $(\mathcal{F}(E;F),+,.)$. Soit $L,M \in \mathcal{L}(E;F)$ et $\mu \in \mathbb{R}$ (ou $\mathbb{C}$) : il s'agit de montrer que $L+\mu M \in \mathcal{L}(E;F)$, cf. (1.33). Soit $\lambda \in \mathbb{R}$ et $\vec{u},\vec{v} \in E$; on a $(L+\mu M)(\vec{u}+\lambda\vec{v}) = L(\vec{u}+\lambda\vec{v}) + \mu M(\vec{u}+\lambda\vec{v}) = L(\vec{u}) + \lambda L(\vec{v}) + \mu M(\vec{u}) + \mu\lambda M(\vec{v}) = (L+\mu M)(\vec{u}) + \lambda(L+\mu M)(\vec{v})$. Donc $L+\lambda M$ est bien linéaire. ▪

1.3.2 Bases et caractérisation d'une application linéaire, matrices

Soit $(\vec{a}_i)_{i=1,...,n} \overset{\text{noté}}{=} (\vec{a}_i)$ une base de E.

Proposition 1.29 (et définitions) *Une application $L \in \mathcal{L}(E;F)$ est entièrement déterminée par la donnée de tous les $L(\vec{a}_j)$ pour tout $j = 1,...,n$.*

Et donc, si $(\vec{b}_i)_{i=1,...,m}$ est une base de F alors L est entièrement déterminée par la donnée de tous les scalaires L^i_j définis par (les composantes des $L(\vec{a}_j)$ dans la base $(\vec{b}_i)$) :

$$L(\vec{a}_j) = \sum_{i=1}^m L^i_j \vec{b}_i, \quad \textit{et} \quad [L]_{|\vec{a},\vec{b}} := [L^i_j], \tag{1.38}$$

noté $L.\vec{a}_j = \sum_{i=1}^m L^i_j \vec{b}_i$. La matrice $[L]_{|\vec{a},\vec{b}} := [L^i_j]_{\substack{i=1,...,n\\ j=1,...,m}}$ est appelée la matrice de L relativement aux bases $(\vec{a}_i)$ et $(\vec{b}_i)$. Et donc, pour tout $\vec{x} \in E$,

$$[L(\vec{x})]_{|\vec{b}} = [L]_{\vec{a},\vec{b}}.[\vec{x}]_{\vec{a}}, \tag{1.39}$$

soit $[L.\vec{x}]_{|\vec{b}} = [L]_{|\vec{a},\vec{b}}.[\vec{x}]_{|\vec{a}}$.

Cas particulier $F=\mathbb{R}$, i.e. $\ell \in \mathcal{L}(E;\mathbb{R}) = E^*$ (forme linéaire) : Alors, sauf indication contraire, dans $F=\mathbb{R}$ on prend la base canonique $(\vec{b}_1)=(1)$ (l'unité pour la multiplication dans $\mathbb{R}^n$), et on note

$$[\ell]_{|\vec{a}} := [\ell]_{|\vec{a},1} = \text{ matrice (ligne) d'une forme linéaire,} \tag{1.40}$$

où donc, cf. (1.38), $[\ell]_{|\vec{a}} = [\ell^1_j] \stackrel{noté}{=} \begin{pmatrix} \ell_1 & \cdots & \ell_n \end{pmatrix}$ (matrice ligne) où $\ell(\vec{a}_j)=\ell_j$ pour tout j.

Cas particulier $F=E$, i.e. $L\in\mathcal{L}(E;E)=\mathcal{L}(E)$ (endomorphisme), alors, sauf indication contraire, on prend $(\vec{b}_i)=(\vec{a}_i)$, et on note

$$[L]_{|\vec{a}} := [L]_{|\vec{a},\vec{a}} = \text{ matrice (carrée) d'un endomorphisme,} \tag{1.41}$$

où donc $[L]_{|\vec{a}} = [L^i_j]$ où $L(\vec{a}_j)=\sum_{i=1}^n L^i_j\vec{a}_i$ pour tout j.

Preuve. Soit $\vec{v}\in E$, $\vec{v}=\sum_{i=1}^n v^i\vec{a}_i$. L est linéaire, donc $L(\sum_{i=1}^n v^i\vec{a}_i)=\sum_{i=1}^n v^iL(\vec{a}_i)$. Donc si on connaît tous les $L(\vec{a}_i)$, on connaît $L(\vec{v})$, et ce pour tout $\vec{v}\in E$.

Et $\vec{x}=\sum_{j=1}^n x^j\vec{a}_j$ donne $L(\vec{x})=\sum_{j=1}^n x^jL(\vec{a}_j)=\sum_{i=1}^m L^i_jx^j\vec{b}_i$, soit (1.39). ∎

Corollaire 1.30 *Si on se donne n vecteurs $\vec{b}_j\in F$ pour $j=1,...,n$, alors il existe une unique application linéaire L qui vérifie $L(\vec{a}_j)=\vec{b}_j$ pour tout $j=1,...,n$.*

Preuve. On définit L par : L est linéaire et $L(\vec{a}_j):=\vec{b}_j$ pour tout $j=1,...,n$, donc $L(\vec{x}):=\sum_{j=1}^m x^j\vec{b}_j$ quand $\vec{x}=\sum_{j=1}^n x^j\vec{a}_j$. ∎

1.3.3 Exemple de la différentielle, et la jacobienne

Soient E et F deux espaces vectoriels normés de dimension finie, et soit Ω un ouvert de E. Soit $\vec{f}:\Omega\to F$ une application à valeurs vectorielles.

Définition 1.31 On dit que $\vec{f}$ est différentiable en un point $\vec{x}\in\Omega$ ssi $\vec{f}$ admet un développement limité à l'ordre 1 au voisinage de $\vec{x}\in E$, i.e. ssi il existe une application linéaire $L_{\vec{x}}\in\mathcal{L}(E;F)$ telle que, pour tout $\vec{v}\in\mathbb{R}^n$:

$$\vec{f}(\vec{x}+h\vec{v})=\vec{f}(\vec{x})+h\,L_{\vec{x}}(\vec{v})+o(h) \quad \in F, \tag{1.42}$$

pour h dans un voisinage de 0.

Et si $\vec{f}$ est différentiable en $\vec{x}$, alors $L_{\vec{x}} \stackrel{\text{noté}}{=} d\vec{f}(\vec{x})\in\mathcal{L}(E;F)$ est appelée la différentielle de f en $\vec{x}$, ou encore l'application linéaire tangente à $\vec{f}$ en $\vec{x}$.

Définition 1.32 Si pour tout $\vec{x}\in\Omega$ l'application linéaire $d\vec{f}(\vec{x})$ existe, alors $d\vec{f}:\left\{\begin{array}{ll}\Omega & \to \mathcal{L}(E;F)\\ \vec{x} & \mapsto d\vec{f}(\vec{x})\end{array}\right\}$ est appelée la différentielle de $\vec{f}$. Si de plus $d\vec{f}$ est continue sur Ω, on dit que $\vec{f}$ est $C^1(\Omega;F)$.

Si $\vec{f}$ différentiable en $\vec{x}$, alors la linéarité de $d\vec{f}(\vec{x})$ permet de noter, pour $\vec{v}\in E$,

$$d\vec{f}(\vec{x})(\vec{v}) \stackrel{\text{noté}}{=} d\vec{f}(\vec{x}).\vec{v} \quad \Big(= \lim_{h\to 0}\frac{\vec{f}(\vec{x}+h\vec{v})-\vec{f}(\vec{x})}{h} \quad \in F\Big) \tag{1.43}$$

est appelé la dérivée directionnelle de $\vec{f}$ en $\vec{x}$ dans la direction $\vec{v}$.

Soit $(\vec{a}_i)$ une base de E. La linéarité de $d\vec{f}(\vec{x})$ indique que la connaissance des $d\vec{f}(\vec{x}).\vec{a}_i$ suffit à connaître $d\vec{f}(\vec{x})$.

Soit $(\vec{a}_i)$ et $(\vec{b}_i)$ des bases de E et F. Soit L^i_j les composantes de $L.\vec{a}_j$ dans la base $(\vec{b}_i)$, i.e.,

$$d\vec{f}(\vec{x}).\vec{a}_j=\sum_{i=1}^m L^i_j(\vec{x})\vec{b}_i, \quad [d\vec{f}(\vec{x})]_{|\vec{a},\vec{b}} := [L^i_j] = \begin{pmatrix} [d\vec{f}(\vec{x}).\vec{a}_1]_{|\vec{b}} & \cdots & [d\vec{f}(\vec{x}).\vec{a}_n]_{|\vec{b}}\end{pmatrix}. \tag{1.44}$$

Définition 1.33 La matrice $[d\vec{f}(\vec{x})]_{|\vec{a},\vec{b}}$ est appelée la matrice jacobienne de f en $\vec{x}$ relativement aux bases choisies (sa j-ème colonne est la matrice colonne $[d\vec{f}(\vec{x}).\vec{a}_j]_{|\vec{b}} = \begin{pmatrix} L^1_j \\ \vdots \\ L^n_j\end{pmatrix}$ appelée abusivement "vecteur colonne").

(Notation tensorielle explicitant les bases utilisées, avec (a^i) la base duale de la base $(\vec{a}_i)$, voir § 3.4 : $d\vec{f}(\vec{x}) = \sum_{i=1}^m L^i_j(\vec{x})\vec{b}_i \otimes a^j$, et on utilise les règles de contraction tensorielle $(\vec{b}_i \otimes \underbrace{a^j).\vec{u}} = (a^j.\vec{u})\vec{b}_j$.)

Donc si $\vec{f}(\vec{x}) = \sum_{i=1}^m f^i(\vec{x})\vec{b}_i$, i.e. $[\vec{f}(\vec{x})]_{|\vec{b}} = \begin{pmatrix} f^1(\vec{x}) \\ \vdots \\ f^m(\vec{x}) \end{pmatrix}_{|\vec{b}}$ (les $f^i(\vec{x})$ les composantes du vecteur $\vec{f}(\vec{x}) \in F$ dans la base $(\vec{b}_i)$), alors

$$d\vec{f}(\vec{x}) = \sum_{i=1}^m \vec{b}_i \otimes df^i(\vec{x}), \quad \text{au sens} \quad d\vec{f}(\vec{x}).\vec{a}_j = \sum_{i=1}^m (df^i(\vec{x}).\vec{a}_j)\,\vec{b}_i, \tag{1.45}$$

et on a, cf (1.40),

$$[df^i(\vec{x})]_{|\vec{a}} = (\; df^i(\vec{x}).\vec{a}_1 \quad \cdots \quad df^i(\vec{x}).\vec{a}_1 \;) \quad , \tag{1.46}$$

donc

$$[df(\vec{x})]_{|\vec{a},\vec{b}} = \begin{pmatrix} [df^1(\vec{x})] \\ \vdots \\ [df^m(\vec{x})] \end{pmatrix} = \Big(\; [d\vec{f}(\vec{x}).\vec{a}_1]_{|\vec{b}} \quad \cdots \quad [d\vec{f}(\vec{x}).\vec{a}_n]_{|\vec{b}} \Big). \tag{1.47}$$

Définition 1.34 Si $(\vec{a}_i)$ est base cartésienne (la même base en tout point de l'espace), alors

$$d\vec{f}(\vec{x}).\vec{a}_i \overset{\text{noté}}{=} \partial_i f(\vec{x}) \overset{\text{noté}}{=} \frac{\partial \vec{f}}{\partial x^i}(\vec{x}) \quad \in F \tag{1.48}$$

est appelé la i-ème dérivée partielle de $\vec{f}$.

NB : Ambiguïté : (1.48) est également noté $d\vec{f}(\vec{y}).\vec{a}_i =^{\text{noté}} \frac{\partial \vec{f}}{\partial y^i}(\vec{y})$ si $\vec{y}$ est le nom de la variable. Une telle notation qui dépend du nom d'une variable est ambigüe, et en cas d'ambiguïté on revient à la notation $d\vec{f}(\vec{x}).\vec{a}_i$, ou à la notation $\partial_i f(\vec{x})$; sinon $\frac{\partial f}{\partial x^i}(\vec{x}_0)$ ou bien $\frac{\partial f}{\partial x^i}(\vec{z})$ peut poser problème : sans ambiguïté, $\partial_i f(\vec{x}_0) := d\vec{f}(\vec{x}_0).\vec{a}_i = \lim_{h\to 0} \frac{\vec{f}(\vec{x}_0+h\vec{a}_i)-\vec{f}(\vec{x}_0)}{h}$.

Exercice 1.35 Dans $\mathbb{R}^2$, on choisit de travailler avec une base cartésienne. Soit $f(x,y) = x\sin(y)$. Calculer $\frac{\partial f}{\partial x^2}(3,\pi)$.

Réponse. On demande de calculer $\partial_2 f(\vec{x}) = df(\vec{x}).\vec{a}_2 = \frac{\partial f}{\partial y}(x,y) = x\cos(y)$ donc $\partial_2 f(\vec{x})(3,\pi) = 3\cos(\pi) = -3$ également noté $\frac{\partial f}{\partial x^2}(3,\pi) = \frac{\partial f}{\partial y}(3,\pi)$. ▪

Remarque 1.36 Après s'être donné une base $(\vec{b}_i)$ dans F, dire que $\vec{f}$ est différentiable en $\vec{x}$, c'est dire que toutes les $f^i : \Omega \to \mathbb{R}$ sont différentiables en $\vec{x}$, i.e. que les graphes de toutes les f^i ont un plan tangent en $\vec{x}$. ▪

1.3.4 Application linéaire inversible

Définition 1.37 L'application linéaire $A \in \mathcal{L}(E;F)$ est inversible (ou bijective) ssi il existe une application linéaire $B \in \mathcal{L}(F;E)$ tel que :

$$A \circ B = I_F \quad \text{et} \quad B \circ A = I_E \tag{1.49}$$

où I_E, resp. I_F, est l'endomorphisme identité de E, resp. de F. Et B est alors appelé inverse de A et noté $B = A^{-1}$

C'est un endomorphisme inversible quand $F = E$.

On note $L_i(E;F)$ l'espace des applications linéaires inversibles, sous ensemble de $\mathcal{L}(E;F)$ (ce n'est pas un sous-espace vectoriel : si A est inversible alors $A - A = 0$ n'est pas inversible).

(Rappel : l'inverse dans un anneau, ici $\mathcal{L}(E;F),+,\circ)$, est unique s'il existe : si C est un autre inverse, alors $A \circ B = I_F$ et $C \circ A = I_E$ donne d'une part $C \circ (A \circ B) = C \circ I_F = C$, et d'autre part, avec l'associativité de la composée, $C \circ (A \circ B) = (C \circ A) \circ B = I_E \circ B = B$).

Notation :

$$A \circ B \overset{\text{noté}}{=} A.B, \tag{1.50}$$

notation du produit (car A et B étant linéaire on a la "distributivité") ou encore notation de la contraction tensorielle (voir plus loin).

Proposition 1.38 *Si E est de dimension finie, si $A,B \in \mathcal{L}(E)$ vérifient $A.B = I$ alors $B.A = I$. C'est faux en dimension infinie.*

Preuve. Soit $(\vec{a}_i)_{i=1,...,n}$ une base de E et supposons $A.B = I$. Posons $\vec{b}_j = B.\vec{a}_j$ pour tout j. Donc $A.\vec{b}_j = A.B.\vec{a}_j = \vec{a}_j$ pour tout j. Donc $B.A.\vec{b}_j = B.\vec{a}_j = \vec{b}_j$ pour tout j. Donc $B.A = I$.

En dimension infinie, soit $\mathbb{R}^{\mathbb{N}^*}$ l'ensemble des suites réelles $(x_i)_{i\in\mathbb{N}^*}$. On prend B le "shift à droite" : $B(x_1,...,x_n,...) = (0,x_1,...,x_{n-1},...)$. On prend A le "shift à gauche" : $A(x_1,...,x_n,...) = (x_2,...,x_{n+1},...)$ (on a perdu l'information x_1).

On a $A.B = I$ mais $B.A \neq I$ puisque $B.A(x_1,x_2,...,x_n,...) = (0,x_2,...,x_n,...)$. ▪

Exercice 1.39 Montrer que si $A \in \mathcal{L}(E;F)$ et si son inverse B existe dans l'espace $\mathcal{F}(E;F)$, i.e. $A \circ B = I_F$ et $B \circ A = I_E$, alors nécessairement B est linéaire.

Réponse. Soient $\vec{w}_1, \vec{w}_2 \in F$. Comme A est bijective dans $\mathcal{F}(E;F)$, elle est surjective. Donc il existe deux vecteurs $\vec{v}_1, \vec{v}_2 \in E$ tels que $\vec{w}_1 = A(\vec{v}_1)$ et $\vec{w}_2 = A(\vec{v}_2)$. Et si $\lambda \in \mathbb{R}$ on a $B(\vec{w}_1 + \lambda\vec{w}_2) = B(A(\vec{v}_1) + \lambda A(\vec{v}_2)) = B(A(\vec{v}_1 + \lambda\vec{v}_2))$ par linéarité de A, donc $B(\vec{w}_1 + \lambda\vec{w}_2) = \vec{v}_1 + \lambda\vec{v}_2 = B(\vec{w}_1) + \lambda B(\vec{w}_2)$, donc B est linéaire. ▪

1.4 Matrice transposée

Si $A = [A_{ij}]_{\substack{i=1,...,m\\ j=1,...,n}}$, matrice $m*n$, alors sa transposée $A^T = [B_{ij}]_{\substack{i=1,...,n\\ j=1,...,m}}$, matrice $n*m$, est définie par $B_{ij} = A_{ji}$ pour tout i,j.

Avec d'autres notations : si $A = [A^i_j]_{\substack{i=1,...,m=ligne\\ j=1,...,n=colonne}}$, matrice $m*n$, alors, sa transposée est la matrice $n*m$ donnée par $A^T = [A^j_i]_{\substack{i=1,...,n=ligne\\ j=1,...,m=colonne}}$.

E.g., $A = \begin{pmatrix} 1\\ 3\end{pmatrix}$ donne $A^T = (\,1 \quad 3\,)$ (poser $A^1_1 = 1$ et $A^2_1 = 3$).

1.5 Application linéaire symétrique (si produits scalaires)

1.5.1 Application linéaire transposé (si produits scalaires)

On suppose E muni d'un produit scalaire $g(\cdot,\cdot) = (\cdot,\cdot)_g$ et F muni d'un produit scalaire $h(\cdot,\cdot) = (\cdot,\cdot)_h$.

Définition 1.40 Soit $L \in \mathcal{L}(E;F)$. L'application linéaire transposée $L^T_{gh} \in L(F;E)$ relativement aux produits scalaires $(\cdot,\cdot)_g$ et dd_h est définie par, pour tout $\vec{x},\vec{y} \in E$:

$$(L^T_{gh}.\vec{y},\vec{x})_g = (\vec{y}, L.\vec{x})_h. \tag{1.51}$$

(L^T_{gh} dépend du choix des produits scalaires, voir exercice 1.45.)

Dans la suite on se restreint (pour alléger la présentation) aux endomorphismes (donc $E = F$ et $(\cdot,\cdot)_g = (\cdot,\cdot)_h$).

Définition 1.41 Soit $L \in \mathcal{L}(E)$ un endomorphisme. L'endomorphisme transposée $L^T_g \in L(E)$ relativement au produit scalaire $g(\cdot,\cdot) = (\cdot,\cdot)_g$ est défini par, pour tout $\vec{x},\vec{y} \in E$:

$$(L^T_g.\vec{y},\vec{x})_g = (\vec{y}, L.\vec{x})_g. \tag{1.52}$$

(L^T_g dépend donc du choix du produit scalaire $g(\cdot,\cdot) = (\cdot,\cdot)_g$, voir exercice 1.45.)

Proposition 1.42 *On se donne un endomorphisme $L \in L(E)$ et une base $(\vec{a}_i)$ dans E.*

Soit $[L]_{|\vec{a}} = [L^i_j]$ la matrice de L dans cette base, i.e. $L.\vec{a}_j = \sum_i L^i_j \vec{a}_i$ pour tout j.

Soit $[L^T_g]_{|\vec{a}} = [(L^T_g)^i_j]$ la matrice de L^T_g dans cette base, i.e. $L^T_g.\vec{a}_j = \sum_i (L^T_g)^i_j \vec{a}_i$ pour tout j.

On a :

$$[g]_{|\vec{a}}.[L^T_g]_{|\vec{a}} = [L]^T_{|\vec{a}}.[g]_{|\vec{a}}, \quad \text{i.e.} \quad [L^T_g]_{|\vec{a}} = [g]^{-1}_{|\vec{a}}.[L]^T_{|\vec{a}}.[g]_{|\vec{a}}. \tag{1.53}$$

En particulier :

1- Si la base $(\vec{a}_i)$ est $(\cdot,\cdot)_g$ orthonormée, alors $[L^T_g]_{|\vec{a}} = [L]^T_{|\vec{a}}$.

2- Si $[g]_{|\vec{a}}.[L]_{|\vec{a}} = [L]_{|\vec{a}}.[g]_{|\vec{a}}$ (les matrices $[g]_{|\vec{a}}$ et $[L]_{|\vec{a}}$ commutent), alors $[L^T_g]_{|\vec{a}} = [L]^T_{|\vec{a}}$.

3- Si $[g]_{|\vec{a}}.[L]_{|\vec{a}} \neq [L]_{|\vec{a}}.[g]_{|\vec{a}}$ (les matrices $[g]_{|\vec{a}}$ et $[L]_{|\vec{a}}$ ne commutent pas), alors $[L^T_g]_{|\vec{a}} \neq [L]^T_{|\vec{a}}$.

Preuve. (1.52) donne $g(L_g^T.\vec{a}_j,\vec{a}_i)=g(L.\vec{a}_i,\vec{a}_j)$, d'où

$$\sum_k (L_g^T)^k_j g(\vec{a}_k,\vec{a}_i)=\sum_\ell L^\ell_i g(\vec{a}_\ell,\vec{a}_j), \tag{1.54}$$

soit, par symétrie de $g(\cdot,\cdot)$,

$$\sum_k g_{ik}(L_g^T)^k_j=\sum_\ell L^\ell_i g_{\ell j},$$

vrai pour tout i,j, donc

$$[g]_{|\vec{a}}.[L_g^T]_{|\vec{a}}=[L]^T_{|\vec{a}}.[g]_{|\vec{a}}, \tag{1.55}$$

i.e. (1.53).

1- Si la base $(\vec{a}_i)$ est dd_g-orthonormée, alors $[g]_{|\vec{a}}=I$ et donc $[L_g^T]_{|\vec{a}}=[L]^T_{|\vec{a}}$.

2- Si $[g]_{|\vec{a}}.[L]_{|\vec{a}}=[L]_{|\vec{a}}.[g]_{|\vec{a}}$ alors $[L]^T_{|\vec{a}}.[g]_{|\vec{a}}=[g]_{|\vec{a}}.[L]^T_{|\vec{a}}$ car $g(\cdot,\cdot)$ est symétrique, donc $[g]_{|\vec{a}}$ et $[L]^T_{|\vec{a}}$ commutent. D'où $[L_g^T]_{|\vec{a}}=[g]^{-1}_{|\vec{a}}.[L]^T_{|\vec{a}}.[g]_{|\vec{a}}=[g]^{-1}_{|\vec{a}}.[g]_{|\vec{a}}.[L]^T_{|\vec{a}}=[L]^T_{|\vec{a}}$.

3- Si $[g]_{|\vec{a}}.[L]_{|\vec{a}}\neq[L]_{|\vec{a}}.[g]_{|\vec{a}}$ alors $[L]^T_{|\vec{a}}.[g]_{|\vec{a}}\neq[g]_{|\vec{a}}.[L]^T_{|\vec{a}}$ et $[g]^{-1}_{|\vec{a}}.[L]^T_{|\vec{a}}.[g]_{|\vec{a}}\neq[g]^{-1}_{|\vec{a}}.[g]_{|\vec{a}}.[L]^T_{|\vec{a}}=[L]^T_{|\vec{a}}$, donc $[L_g^T]_{|\vec{a}}\neq[L]^T_{|\vec{a}}$. ∎

1.5.2 Endomorphisme symétrique (relativement à un produit scalaire)

Soit $g(\cdot,\cdot)$ un produit scalaire sur E.

Définition 1.43 Un endomorphisme $L\in\mathcal{L}(E)$ est dit symétrique relativement au produit scalaire $g(\cdot,\cdot)$, ou $(\cdot,\cdot)_g$-symétrique, ssi $L_g^T=L$, i.e. ssi, pour tout $\vec{x},\vec{y}\in E$,

$$(L.\vec{y},\vec{x})_g=(L.\vec{x},\vec{y})_g \tag{1.56}$$

Comme $L_g^T=L$, on a $[L_g^T]_{|\vec{a}}=[L]_{|\vec{a}}$ quelle que soit la base, et (1.53) donne :

Corollaire 1.44 *Si L est $(\cdot,\cdot)_g$-symétrique, alors quelque soit la base $(\vec{a}_i)$ on a*

$$[g]_{|\vec{a}}.[L]_{|\vec{a}}=[L]^T_{|\vec{a}}.[g]_{|\vec{a}},\quad \text{i.e.}\quad [L]_{|\vec{a}}=[g]^{-1}_{|\vec{a}}.[L]^T_{|\vec{a}}.[g]_{|\vec{a}}. \tag{1.57}$$

Et on note $[g].[L]=[L]^T.[g]$.

Et l'exercice 1.45 à montrer que la symétrie de L est dépendante du choix du produit scalaire : on peut pour un produit scalaire donné avoir L symétrique, et pour un autre produit scalaire L NON symétrique.

Exercice 1.45 1- Montrer que si deux matrices A et B sont symétriques alors le produit $A.B$ n'est pas nécessairement symétrique.

2- Donner un endomorphisme $L\in\mathcal{L}(E)$, une base $(\vec{a}_i)$, et deux produits scalaires $(\cdot,\cdot)_g$ et $(\cdot,\cdot)_h$ t.q. $[L_g^T]_{|\vec{a}}=[L]_{|\vec{a}}$ et $[L_h^T]_{|\vec{a}}\neq[L]_{|\vec{a}}$ (la symétrie est une notion dépend du produit scalaire).

3- Donner un endomorphisme $L\in\mathcal{L}(E)$ et deux bases $(\vec{a}_i)$ et $(\vec{b}_i)$ t.q. $[L]_{|\vec{a}}$ est symétrique et $[L]_{|\vec{b}}$ est non symétrique.

Réponse. 1- Soit $A=\begin{pmatrix}1&0\\0&2\end{pmatrix}$ et $B=\begin{pmatrix}0&1\\1&0\end{pmatrix}$. Alors $A.B=\begin{pmatrix}0&1\\2&0\end{pmatrix}$ n'est pas symétrique.

2- Dans $\mathbb{R}^2$ avec sa base canonique $(\vec{E}_i)$. On prend $(\vec{a}_i)=(\vec{E}_i)$. Soit L donnée par $[L]_{|\vec{E}}=\begin{pmatrix}0&1\\1&0\end{pmatrix}$.

Soit $g(\cdot,\cdot)$ le produit scalaire canonique, donc $[g]_{|\vec{E}}=I$. Donc $[L_g^T]_{|\vec{E}}=[L]_{|\vec{E}}$.

Soit $h(\cdot,\cdot)$ le produit scalaire donné par $[h]_{|\vec{E}}=\begin{pmatrix}1&0\\0&2\end{pmatrix}$. Alors

$$[L_h^T]_{|\vec{E}}=[h]^{-1}_{|\vec{E}}.[L]^T_{|\vec{E}}.[h]_{|\vec{E}}=\begin{pmatrix}1&0\\0&\frac12\end{pmatrix}.\begin{pmatrix}0&1\\1&0\end{pmatrix}.\begin{pmatrix}1&0\\0&2\end{pmatrix}=\begin{pmatrix}1&0\\0&\frac12\end{pmatrix}.\begin{pmatrix}0&2\\1&0\end{pmatrix}=\begin{pmatrix}0&2\\\frac12&0\end{pmatrix}\neq[L]^T_{|\vec{E}}.$$

3- Dans $\mathbb{R}^2$ avec sa base canonique $(\vec{E}_i)$. On prend $(\vec{a}_i)=(\vec{E}_i)$. Soit L donnée par $[L]_{|\vec{E}}=\begin{pmatrix}0&1\\1&0\end{pmatrix}$. Donc $[L]_{|\vec{E}}=[L]^T_{|\vec{E}}$ (matrice symétrique). Soit $(\vec{b}_i)$ la base donnée par $\vec{b}_1=\vec{E}_1$ et $\vec{b}_2=\vec{E}_1+\vec{E}_2$, donc $P=\begin{pmatrix}1&1\\0&1\end{pmatrix}$ est la matrice de passage de $(\vec{E}_i)$ vers $(\vec{b}_i)$, donc $P^{-1}=\begin{pmatrix}1&-1\\0&1\end{pmatrix}$, donc

$$[L]_{\vec{b}}=P^{-1}.[L]_{\vec{a}}.P=\begin{pmatrix}1&-1\\0&2\end{pmatrix}$$

n'est pas symétrique. ∎

1.5.3 * Remarque : transposée vs adjointe

Soit $L \in \mathcal{L}(E;F)$ (une application linéaire). Pour définir l'application transposée, on a eu besoin d'introduire deux produits scalaires, cf. (1.52) : la définition L^T de la transposée de L n'est pas intrinsèque (à L), puisqu'elle dépend d'outils introduit par des observateurs (qui choisissent les produits scalaires).

Pour avoir une notion intrinsèque, donc sans faire appel à un produit scalaire, on définit l'application linéaire adjointe $L^* : \ell \in F^* \to L^*(\ell) =^{\text{noté}} L^*.\ell \in E^*$ par, pour tout $\vec{u} \in E$:

$$L^*(\ell)(\vec{u}) := \ell(L(\vec{u})), \quad \text{noté} \quad (L^*.\ell).\vec{u} = \ell.(L.\vec{u}), \quad \text{noté} \quad \langle L^*.\ell, \vec{u}\rangle_{E^*,E} = \langle \ell, L.\vec{u}\rangle_{F^*,F}, \tag{1.58}$$

où on a utilisé "le crochet de dualité" ⟨forme linéaire,vecteur⟩ relatif aux espaces considérés.

La confusion entre l'application adjointe et l'application transposée vient de l'usage courant du produit scalaire canonique dans le cas $E = F = \mathbb{R}^n$ qui fait qu'on utilise sans le voir le théorème de représentation de Riesz d'une forme linéaire par un vecteur (représentation non intrinsèque puisqu'elle n'a de sens qui si on dispose d'un produit scalaire).

Mais dans l'espace géométrique, ou dans un espace usuel de dimension n, il n'y a pas de produit scalaire canonique. Exemple : dans l'espace géométrique $\mathbb{R}^3$, l'espace $\mathbb{R}^3$ n'est pas un produit cartésien $\mathbb{R}\times\mathbb{R}\times\mathbb{R}$ puisqu'il n'y a pas de "direction privilégiée" : c'est un observateur qui décide de la direction (et de la longueur) des vecteurs de bases (et d'un repère). On confond parfois un "produit scalaire canonique" avec un produit scalaire euclidien, voir § 1.2.3... Mais un produit scalaire euclidien n'a rien de canonique, voir § 1.6.

Remarque : Il n'y a pas d'isomorphisme naturel entre E et E^* (voir poly suivant) et on ne peut pas identifier L^T et L^* : il y a bien un isomorphisme $L^* \leftrightarrow L^T$, mais un tel isomorphisme nécessite l'introduction d'un outil supplémentaire comme un produit scalaire ou une base, et donc un tel isomorphisme n'est pas une "identification".

1.6 Utilisation du mot canonique

1.6.1 Dans l'espace $\mathbb{R}^n$

Dans $\mathbb{R}^n = \mathbb{R}\times ... \times\mathbb{R}$ (produit cartésien n-fois), les vecteurs de base canonique sont

$$\vec{E}_1 = (1,0,...,0), \quad \vec{E}_2 = (0,1,0,...,0), \quad ..., \quad \vec{E}_n = (0,...,0,1). \tag{1.59}$$

Et $(\vec{E}_1,...,\vec{E}_n)$ est trivialement une base de $\mathbb{R}^n$ appelée la base canonique.

N.B : l'utilisation mot "canonique" signifie : on considère

1- un espace qui est le produit cartésien d'un corps avec lui-même, ici $E = \mathbb{R}\times ... \times\mathbb{R}$, et

2- on utilise uniquement le 0 élément neutre de l'addition, et le 1 élément neutre de la multiplication, ce dernier n'étant utilisé qu'une seule fois.

Et alors la forme bilinéaire $(\cdot,\cdot)_{\mathbb{R}^n}$ définie par $(\vec{E}_i,\vec{E}_j)_{\mathbb{R}^n} = \delta_{ij}$ pour tout i,j est appelé le produit scalaire canonique, bien qu'il ne soit pas canonique... puisqu'il utilise des vecteurs pour sa définition.

1.6.2 * Isomorphisme canonique

On verra l'expression "isomorphisme canonique" (entre E et E^{**}) : on a $\mathcal{J} : E \to E^{**}$, voir (9.2), qui est défini sans constante (au sens : on n'utilise que la constante = 1 élément neutre de la multiplication), et en particulier sans référence à une base particulière ou à un produit scalaire particulier. (De plus $\mathcal{J}$ est 'naturel" = indépendant d'un observateur, voir le poly suivant).

Et on verra qu'il n'a aura pas d'isomorphisme "canonique" entre E et E^* : il existe des isomorphismes entre E et E^*, mais leur considération exige l'introduction d'objets (outils pour quantifier), par exemple une base ou un produit scalaire (et donc un tel isomorphisme n'est pas naturel). (Voir aussi le paragraphe suivant (1.7) et le "caractère intrinsèque".)

Et on verra (poly suivant) que l'espace des formes bilinéaires $\mathcal{L}(E,E;\mathbb{R})$ est canoniquement (et naturellement) isomorphe à $\mathcal{L}(E;E^*)$ espace des applications linéaires de E dans E^*, voir (10.3). Et qu'il n'y a pas d'isomorphisme "canonique" ("naturel") entre $\mathcal{L}(E;E)$ espace des endomorphismes sur E et $\mathcal{L}(E,E;\mathbb{R})$ espace de formes bilinéaires sur E.

1.6.3 * Contre-exemples classiques

Lorsque E est un espace de dimension n qui n'est pas un produit cartésien, on ne peut plus parler de “base canonique”.

Exemple dans l'univers dans lequel nous vivons modélisé en mécanique classique par l'espace affine $\mathbb{R}^3$: ce serait quoi une base canonique ? Il n'y en a pas... chaque observateur définissant sa base... Et donc il n'y a pas non plus de produit scalaire “canonique”.

Exemple de l'approximation d'une fonction $f : [0,1] \to \mathbb{R}$ par une fonction continue $f_h : [0,1] \to \mathbb{R}$ qui est affine par morceaux (démarche de la méthode des éléments finis) : on commence par exemple par découper $[0,1]$ en n intervalles égaux en posant $x_i = i\,\frac{1}{n}$ pour $i = 0,...,n$ (on a défini ainsi $n{+}1$ points), et on note :

$$P_1 = \{f_h \in C^1([0,1];\mathbb{R}) : (f_h)_{|[x_{i-1},x_i]} \text{ affine},\ i = 1,...,n\}.$$

C'est espace n'est pas de type produit cartésien $K \times ... \times K$ où K est un corps. Et quand on prend une base, elle n'est donc pas canonique. P_1 est un espace de dimension $n{+}1$ dont une base est par exemple donnée par les fonctions (dites fonctions chapeau) définies par, pour tout $i, j = 0, ..., n$:

$$\varphi^i \in P_1, \qquad \varphi^i(x_j) = \delta^i_j.$$

(Voir cours d'éléments finis.) Bien que les φ^i soient définis à l'aide de 1 (une fois pour $\varphi^i(x_i) = 1$) et de 0 (n fois pour $\varphi^i(x_j) = 0$ pour $j \neq i$), on ne dit pas que $(\varphi^i)_{i=0,...,n}$ est une base canonique : l'espace P_1 n'a pas de base canonique car il ne peut pas être écrit comme produit cartésien $E \times ... \times E$. On ne peut donc pas parler de produit scalaire “canonique”, même si on utilise souvent le produit scalaire “usuel” de L^2 :

$$b_0(f,g) = \int_0^1 f(x)g(x)\,dx. \tag{1.60}$$

“Usuel” ne veut pas dire canonique. Par exemple, on utilise également souvent le produit scalaire “usuel” :

$$b_1(f,g) = b_0(f,g) + b_0(f',g') \tag{1.61}$$

qui donne “l'énergie de déformation” (voir cours d'éléments finis). Il n'est pas plus canonique que $b_0(\cdot,\cdot)$.

1.7 Utilisation du mot intrinsèque

On utilisera le mot intrinsèque dans le sens suivant :

« Une expression est intrinsèque (“à un objet”) ssi elle est indépendante d'un observateur. »

Dans ce cours, cela voudra dire que l'expression considérée est intrinsèque ssi elle est indépendante d'une base (qui dépend du choix d'un observateur) et d'un produit scalaire (qui dépend du choix d'un observateur).

Exemple 1.46 La trace d'une application linéaire $L : E \to E$ est une valeur intrinsèque.

Rappel : si $(\vec{e}_i)$ est une base de E, si $[L]_{|\vec{e}} = [L^i_j]$ est la matrice de L relativement à la base $(\vec{e}_i)$, i.e. si $L.\vec{e}_j = \sum_{i,j=1}^n L^i_j \vec{e}_i$ pour tout j, alors la trace de L est $\mathrm{Tr}(L) := \sum_{i=1}^n L^i_i$.

Rappel : si $(\vec{e}_{i,old})$ et $(\vec{e}_{i,new})$ sont deux bases, si $M = [M^i_j] = [L]_{|\vec{e}_{i,old}}$ et $N = [N^i_j] = [L]_{|\vec{e}_{i,new}}$ sont les matrices de L relativement aux base considérées, i.e. $L.\vec{e}_{j,old} = \sum_{i,j=1}^n M^i_j \vec{e}_{i,old}$ et $L.\vec{e}_{j,new} = \sum_{i,j=1}^n N^i_j \vec{e}_{i,new}$, alors $\sum_{i=1}^n M^i_i = \sum_{i=1}^n N^i_i =^{\text{noté}} \mathrm{Tr}(L)$. ▪

1.8 Sur les vecteurs en mécanique et en mathématique

L'introduction des vecteurs semble due à Bellavitis (1835), Hamilton, Grassmann. En mécanique, un vecteur $\vec{v}$ représente alors un “bâton” (objet matériel mono-dimensionnel) à l'aide d'un “vecteur bi-points” $\vec{v} = \overrightarrow{AB}$. C'est le même vecteur bi-point pour tous les observateurs ; mais sa longueur dépend d'un observateur (en pied ? en mètre ?) : la longueur n'est pas une valeur intrinsèque (au bâton).

Un vecteur pour un mathématicien est un élément d'un espace vectoriel, un espace vectoriel étant une ensemble muni de structures algébriques + et $\cdot$, espace qui peut être de dimension finie ou infinie : un vecteur mathématique ne représente pas un objet “matériel” en général.

Puis sont apparus les produits scalaires (Hamilton 1953), et au 20ième siècle est apparu la représentation d'une forme linéaire à l'aide d'un “vecteur de représentation de Riesz” (1907). Ce vecteur de représentation n'est pas un “vecteur matériel” : ce n'est qu'une quantification mathématique d'une forme linéaire à l'aide d'un produit scalaire (n'est pas intrinsèque à l'objet qu'il représente). E.g., pour une même fonction linéaire (un unique objet mathématique) dans $\mathbb{R}^3$, un anglais utilisant le pied ($= 0,3048$ mètre) obtient un vecteur de représentation qui est plus de 10 fois plus grand (exactement $(0,3048)^{-2}$ fois plus grand) que le vecteur de représentation d'un français qui utilise le mètre : le vecteur de représentation dépend de l'observateur.

2 Vecteur, vecteur covariant, vecteur contravariant

Les mots covariant et contravariant semblent avoir eue une définition inverse de la définition actuelle. Cf. Spivak [9], p.113 : "...Classical terminology used these same words, and it just happens to have reversed this... And no one has had the gall or authority to reverse terminology so sanctified by years of usage...". Ici on suit la définition de Spivak [9], qui semble être la définition actuelle, voir aussi Abraham et Marsden [1] et leur seconde édition (1978) où ils adoptent la définition de Spivak (ils ont inversé la terminologie par rapport à leur première édition de 1966).

Définition 2.1 Un élément $\vec{v} \in E$ d'un espace vectoriel E est appelé un vecteur.

Il est également appelé vecteur contravariant en référence à la formule de changement de base des vecteurs fait qui intervenir P^{-1} pour ses composantes. Voir (8.7).

Définition 2.2 Un élément $\ell \in E^*$ (une forme linéaire sur E) est appelé un vecteur covariant : c'est un vecteur au sens où l'espace de fonctions E^* est un espace vectoriel, et il est co-variant au sens où il agit sur $\vec{v}$ un vecteur de E, et en référence à la formule de changement de base des formes linéaires qui fait qui intervenir P pour ses composantes, cf. (8.12).

Un vecteur covariant est égal appelé un vecteur dual, ou encore un tenseur uniforme de type $\binom{0}{1}$.

Définition 2.3 Un élément du bidual $E^{**} = (E^*)^* = \mathcal{L}(E^*;\mathbb{R})$ est un vecteur contravariant, ou un tenseur de type $\binom{1}{0}$.

Donc un vecteur contravariant $v \in E^{**}$ varie avec (ou agit sur) une forme linéaire ℓ (pour donner le réel $v(\ell)$.

Le bidual E^{**} étant identifié à E, voir (9.2), cela justifie également l'appellation : Un vecteur est également appelé un vecteur contravariant.

3 Base duale

3.1 La base duale d'une base

La proposition 1.29 donne dans le cas particulier $F = \mathbb{R}$ (cas des formes linéaires) : Soit $\ell \in \mathcal{L}(E;R) = E^*$ une forme linéaire, soit $\vec{x} \in E$, soit $(\vec{a}_i)$ une base de E, et soit $x^1, ..., x^n$ les composantes de $\vec{x}$ dans la base, i.e., $\vec{x} = \sum_{i=1}^n x^i \vec{a}_i$. Alors

$$\ell_i := \ell(\vec{a}_i) \ (\in \mathbb{R}) \quad \text{donne} \quad \ell(\vec{x}) = \sum_{i=1}^n \ell_i x^i \ (\in \mathbb{R}), \tag{3.1}$$

soit, avec la notation pour les ℓ linéaire cf. (1.34), $\ell.\vec{x} = \sum_{i=1}^n \ell_i x^i$. Et avec la matrice (ligne) de ℓ relativement à la base $(\vec{a}_i)$, cf. (1.40), à savoir

$$[\ell]_{|\vec{a}} = (\ \ell_1 \quad ... \quad \ell_n\ )\,, \tag{3.2}$$

on a, pour tout $\vec{x} \in E$,

$$\ell.\vec{x} = [\ell]_{|\vec{a}}.[\vec{x}]_{|\vec{a}}, \tag{3.3}$$

le réel produit de $[\ell]_{|\vec{a}}$ matrice $1 * n$ (ligne), cf. (3.2), et de $[\vec{x}]_{[\vec{a}}$ matrice $n * 1$ (colonne), cf. (1.1).

Définition 3.1 La base duale $(a^i)_{i=1,...,n}$ d'une base $(\vec{a}_i)_{i=1,...,n}$ de E est l'unique famille des formes linéaires $a^i \in \mathcal{L}(E;\mathbb{R}) = E^*$, $i = 1, ..., n$, définies par, pour tout $j = 1, ..., n$:

$$a^i(\vec{a}_j) = \delta^i_j \quad (\overset{\text{noté}}{=} a^i.\vec{a}_j). \tag{3.4}$$

(La dernière notation car a^i est une application linéaire.)

Et (3.2) indique que la matrice représentant de la forme linéaire a^i relativement à la base $(\vec{a}_i)$ est la matrice ligne dont toutes les composantes sont nulles sauf la i-ème qui vaut 1 :

$$[a^1]_{|\vec{a}} = (\,1 \quad 0 \quad \dots \quad 0\,), \quad \dots \quad, [a^n]_{|\vec{a}} = (\,0 \quad \dots \quad 0 \quad 1\,), \tag{3.5}$$

soit $[a^i]_{|\vec{a}} = [\vec{a}_i]^T_{|\vec{a}}$ pour tout i.

Donc $\vec{x} = \sum_{i=1}^n x^i\vec{a}_i \in E$ donne $a^i(\vec{x}) = a^i(\sum_{j=1}^n x^j\vec{a}_j) = \sum_{i=1}^n x^j a^i(\vec{a}_j) = \sum_{i=1}^n x^i\delta^i_j = x^j$, donc :

$$\vec{x} = \sum_{i=1}^n x^i\vec{a}_i \in E \quad \Longrightarrow \quad a^i(\vec{x}) = x^i \quad (= [a^i]_{|\vec{a}}.[\vec{x}]_{|\vec{a}}). \tag{3.6}$$

Ou si on préfère :

$$\vec{x} \in E \quad \text{et} \quad a^i(\vec{x}) = x^i \quad \Longrightarrow \quad \vec{x} = \sum_{j=1}^n x^j\vec{a}_j \quad (= \sum_{j=1}^n a^i(\vec{x})\vec{a}_j). \tag{3.7}$$

Interprétation : a^i est l'opérateur de projection sur la droite $\mathrm{Vect}\{\vec{a}_i\}$ parallèlement aux directions $\vec{a}_j$ pour $j \neq i$, i.e., c'est la fonction $a^i : E \to \mathbb{R}$ (l'instrument de mesure) que donne la i-ème composante d'un vecteur relativement à la base $(\vec{a}_i)$, cf. (3.6). Dessin.

Remarque 3.2 Les a^i ne sont pas des projections "orthogonales". D'ailleurs, orthogonal n'a de sens que si on a préalablement introduit un produit scalaire sur E, ce qui n'est pas le cas ici. Et même si on avait introduit un produit scalaire, on aurait dû supposer la base $(\vec{a}_i)$ orthogonale (relativement au produit scalaire choisi) pour que les projections e^i soient orthogonales. ▟

Proposition 3.3 *La famille $(a^i)_{i=1,\dots,n}$ des fonctions linéaires définies en (3.4) est une base de E^*, comme annoncé dans la définition 3.1 (base dite base duale de la base $(\vec{a}_i)$). Et $\dim(E^*) = \dim(E)$. Ainsi, toute forme linéaire $\ell \in E^*$ s'exprime sur la base $(a^i)_{i=1,\dots,n}$:*

$$\exists \ell_1, \dots, \ell_n \in \mathbb{R} \quad t.q. \quad \ell = \sum_{i=1}^n \ell_i a^i \ \in E^*, \quad \text{et} \quad \ell_i = \ell(\vec{a}_i) \in \mathbb{R}. \tag{3.8}$$

(La i-ème composante ℓ_i de ℓ sur la base duale (a^i) est obtenue en calculant $\ell(\vec{a}_i)$.)

Preuve. La famille (a^i) est libre : si $\sum_{i=1}^n \alpha_i a^i = 0$, i.e., si $(\sum_{i=1}^n \alpha_i a^i)(\vec{x}) = 0$ pour tout $\vec{x}$ i.e., si $\sum_{i=1}^n \alpha_i a^i(\vec{x}) = 0$ pour tout $\vec{x}$, alors en particulier $\sum_{i=1}^n \alpha_i a^i(\vec{a}_j) = 0$ pour tout j, donc $\sum_{i=1}^n \alpha_i\delta^i_j = 0$ pour tout j, donc $\alpha_j = 0$ pour tout j.

Famille génératrice : on applique la prop. 1.29. ℓ est déterminée par ses valeurs $\ell_i := \ell(\vec{a}_i)$. Notons $\lambda \in E^*$ la forme linéaire donnée par $\lambda = \sum_{i=1}^n \ell_i a^i$. On a $\lambda(\vec{a}_i) = \ell_i = \ell(\vec{a}_i)$ pour tout i, donc $\lambda = \ell$ donc $\ell = \sum_{i=1}^n \ell_i a^i$: la famille (a^i) est génératrice, et on a (3.8).

Donc $(a^i)_{i=1,\dots,n}$ est donc une base de E^*, d'où $\dim(E^*) = n$. ▟

Exercice 3.4 Soit $F = E$ et $(\vec{a}_i)$ et $(\vec{b}_i)$ deux bases dans E. Soit P la matrice de passage de la base $(\vec{a}_i)$ vers la base $(\vec{b}_i)$, i.e., les P^i_j sont les composantes de $\vec{b}_j$ dans la base $(\vec{a}_i)$, i.e., pour $j \in [1,n]_{\mathbb{N}}$,

$$\vec{b}_j = \sum_{i=1}^n P^i_j\vec{a}_i, \quad P = [P^i_j]. \tag{3.9}$$

Soit $Q = [Q^i_j] := [P^i_j]^{-1} = P^{-1}$. Montrer que les composantes des b^i dans la base (a^i) sont les Q^i_j :

$$b^i = \sum_{j=1}^n Q^i_j a^j. \tag{3.10}$$

Réponse. $(\sum_{j=1}^n Q^i_j a^j)(\vec{b}_k) = \sum_{j=1}^n Q^i_j a^j(\vec{b}_k) = \sum_{j=1}^n Q^i_j a^j(\sum_{\ell=1}^n P^\ell_k\vec{a}_\ell) = \sum_{j,\ell=1}^n Q^i_j P^\ell_k a^j(\vec{a}_\ell) = \sum_{j,\ell=1}^n Q^i_j P^\ell_k\delta^j_\ell = \sum_{j=1}^n Q^i_j P^j_k = (QP)^i_k = \delta^i_k$, d'où $(\sum_{j=1}^n Q^i_j a^j) = b^i$. ▟

Exercice 3.5 Soit $L \in \mathcal{L}(E;E)$ (endomorphisme de E).

Montrer : si pour tout $\vec{v} \in E$ on a $L.\vec{v}$ est parallèle à $\vec{v}$, alors il existe $p \in \mathbb{R}$ t.q. $L = pI$ où I est l'identité.

Réponse. Soit $(\vec{e}_i)$ une base de E. Soit L donné par $L(\vec{e}_j) = \sum_i L^i_j \vec{e}_i$ pour tout j. Par hypothèse il existe $\lambda_j \in \mathbb{R}$ t.q. $L(\vec{e}_j) = \lambda_j \vec{e}_j$. Et donc, pour $i \neq j$:

$e^i(L(\vec{e}_i + \vec{e}_j)) = e^i(L(\vec{e}_i)) + e^i(L(\vec{e}_j)) = e^i(\lambda_i \vec{e}_i + \lambda_j \vec{e}_j) = \lambda_i + 0 = \lambda_i$.

Et par hypothèse il existe $\lambda_{i+j} \in \mathbb{R}$ t.q. $L.(\vec{e}_i + \vec{e}_j) = \lambda_{i+j}(\vec{e}_i + \vec{e}_j)$. Et donc, pour $i \neq j$:

$e^i(L(\vec{e}_i + \vec{e}_j)) = e^i(\lambda_{i+j}(\vec{e}_i + \vec{e}_j)) = \lambda_{i+j}(1+0) = \lambda_{i+j}$,

Donc $\lambda_{i+j} = \lambda_i$ pour tout i,j, donc $\lambda_{i+j} = \lambda_{j+i} = \lambda_j$ pour tout i,j, donc $\lambda_i = \lambda_j \overset{\text{noté}}{=} p$ pour tout i,j. ▪

Notation du crochet de dualité. On note également, si $\ell \in E^*$ et $\vec{x} \in E$:

$$\ell(\vec{x}) = \ell.\vec{x} = \langle \ell, \vec{x} \rangle_{E^*,E} = \langle \ell | \vec{x} \rangle, \tag{3.11}$$

et on omet l'indice ${}_{E^*,E}$ si le contexte est clair. La dernière notation est la notation de Dirac utilisée pour les espaces vectoriels complexes (mécanique quantique) ; et (3.6) s'écrit $\vec{x} = \sum_{i=1}^n \langle a^i | \vec{x} \rangle \vec{a}_i$.

3.2 Notation dx^i dans $\mathbb{R}^n$

Définition 3.6 (Notation). Quand $E = \mathbb{R}^n$, quand $(\vec{a}_i)_{i=1,\dots,n} = (\vec{E}_i)_{i=1,\dots,n}$ est une base cartésienne, alors on note également :

$$E^i \overset{\text{noté}}{=} dx^i, \tag{3.12}$$

i.e. on note $(dx^i)_{i=1,\dots,n}$ la base duale de $(\vec{E}_i)_{i=1,\dots,n}$. Donc, quand $\vec{x} = \sum_{i=1}^n x^i \vec{E}_i$, on a $dx^i(\vec{x}) = x^i$: dx^i est la projection sur $\mathrm{Vect}\{\vec{E}_i\}$ parallèlement aux autres directions ($dx^i.\vec{E}_j = 0$ pour tout $j \neq i$, et $dx^i.\vec{E}_i = 1$).

Dans $\mathbb{R}^2$ on note aussi $(dx^1, dx^2) = (dx, dy)$, et dans $\mathbb{R}^3$ on note aussi $(dx^1, dx^2, dx^3) = (dx, dy, dz)$.

Exercice 3.7 Dans $\mathbb{R}^2$. Soit $\vec{e}_1 = \vec{E}_1$ et $\vec{e}_2 = \vec{E}_1 + \vec{E}_2$, donc $[\vec{e}_1]_{|\vec{E}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $[\vec{e}_2]_{|\vec{E}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Faire un dessin. Montrer que $e^1 = dx^1 - dx^2$ et $e^2 = dx^2$, i.e., $[e^1]_{|\vec{E}} = (\,1 \quad -1\,)$ et $[e^2]_{|\vec{E}} = (\,0 \quad 1\,)$.

Réponse. 1- Calcul direct : Soit α, β les composantes de e^1 dans la base (dx^1, dx^2), i.e., $e^1 = \alpha dx^1 + \beta dx^2$. On a $1 = \delta^1_1 = e^1(\vec{e}_1) = \alpha dx^1(\vec{e}_1) + \beta dx^2(\vec{e}_1) = \alpha dx^1(\vec{E}_1) + \beta dx^2(\vec{E}_1) = \alpha + 0$ donc $\alpha = 1$, et on a $0 = \delta^1_2 = e^1(\vec{e}_2) = \alpha dx^1(\vec{e}_2) + \beta dx^2(\vec{e}_2) = \alpha dx^1(\vec{E}_1 + \vec{E}_2) + \beta dx^2(\vec{E}_1 + \vec{E}_2) = \alpha(dx^1(\vec{E}_1) + dx^1(\vec{E}_2)) + \beta(dx^2(\vec{E}_1) + dx^2(\vec{E}_2)) = \alpha(1+0) + \beta(0+1)$ donc $\alpha + \beta = 0$, donc $\beta = -\alpha = -1$, donc $e^1 = dx^1 - dx^2$

Soit α, β les composantes de e^2 dans la base (dx^1, dx^2), ... (même démarche) ..., d'où $e^2 = dx^2$.

2- On applique (3.9)-(3.10) : ici $P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ (les colonnes sont les $[\vec{e}_j]_{|\vec{E}}$), donc $P^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = Q$ dont les lignes donnent $[e^1]_{|\vec{E}} = (\,1 \quad -1\,)$ et $[e^2]_{|\vec{E}} = (\,0 \quad 1\,)$, soit $e^1 = dx - dy$ et $e^2 = dy$. ▪

3.3 Remarque : les bases duales vectorielles d'une base

Une "base duale vectorielle" d'une base $(\vec{a}_i)$ n'a de sens que si on dispose d'un produit scalaire $(\cdot,\cdot)_g$ sur E (exemple un produit scalaire euclidien en pied, ou en mètre). C'est la base des vecteurs $(\vec{a}_{ig})$ de E relative au produit scalaire $(\cdot,\cdot)_g$ caractérisée par, pour tout $i \in [1,n]_{\mathbb{N}}$,

$$\forall j \in [1,n]_{\mathbb{N}}, \quad (\vec{a}_{ig}, \vec{a}_j)_g = \delta_{ij}. \tag{3.13}$$

Si on change de produit scalaire alors on change de base duale vectorielle. E.g., un anglais et son produit scalaire euclidien $(\cdot,\cdot)_g$ en pied aura un vecteur $\vec{a}_{ig}$, un français et son produit scalaire euclidien $(\cdot,\cdot)_h$ en mètre aura un vecteur $\vec{a}_{ih}$, et $\vec{a}_{ig} = (0,3048)^{-2}\vec{a}_{ih}$ (donc $\vec{a}_{ig}$ est plus de 10 fois plus grand que $\vec{a}_{ih}$).

Donc il y a autant de bases duales vectorielles que de produits scalaires (donc il y en a une infinité) Alors qu'il n'y a qu'une seule base duale (fonctionnelle) (a^i), cf. (3.4). (On peut noter que les vecteurs $\vec{a}_{ig}$ de la $(\cdot,\cdot)_g$ base vectorielle sont les $(\cdot,\cdot)_g$ représentant de Riesz des a^i : on a $a^i.\vec{a}_j = (\vec{a}_{ig}, \vec{a}_j)_g$ pour tout i,j, cf. (3.4) et (3.13).)

Malheureusement, bien que les $\vec{a}_{ig}$ sont des vecteurs (contravariants), notés avec un index ${}_i$ "en bas" (en indice) avec la convention d'Einstein, les mécaniciens le note parfois $\vec{a}^i_g$: avec un index "en haut" (en exposant). Et ils oublient l'indice g. Il devient alors impossible de savoir si on parle de fonctions (la fonction a^i) ou bien de vecteur (le vecteur $(\cdot,\cdot)_g$-dual $\vec{a}_{ig}$). Et les notions de covariance (notion fonctionnelle) et les notions de contravariance (notion vectorielle) deviennent incompréhensibles. (Le respect de la convention d'Einstein est une aide à la compréhension.) Et donc confondre a^i et $\vec{a}_{ig}$ fait que les anglais et les français ne peuvent pas travailler ensemble, et donc que les français ne peuvent pas utiliser les travaux de Newton (qui ne connaissait pas le mètre)... ! ... Absurd isn't it ? Ou alors cela même à un accident de type Mars Climate Orbiter...

3.4 Dimensions en physique

Une question essentielle en physique est : quelle est la longueur de ... ? Quelle est la température de ... ?

Pour répondre, il faut préciser l'unité de mesure : en mètre, en pied ? En degré Celsius, en degrés Fahrenheit ?

Modélisation :

1- On se donne donc un "objet matériel réel" de référence, objet modéliser par un vecteur $\vec{e}$ auquel on donne la valeur 1 (par exemple le "mètre étalon" qui est un objet matériel réel en platine).

2- Soit un objet dont on veut connaître la longueur, objet représenté par un vecteur $\vec{v}$; déterminer sa longueur c'est donner une valeur à $\vec{v}$ (valeur de la mesure), à l'aide d'une fonction $\ell : \vec{v} \to \ell(\vec{v}) =$ longueur de $\vec{v}$ (fonction appelée appareil de mesure). Avec ici ℓ linéaire (donne une longueur "proportionnelle à la taille de l'objet") : si $\vec{v} = \lambda\vec{e}$, on aura $\ell(\vec{v}) = \lambda\,\ell(\vec{e}) = \lambda$ où λ est en unité $\vec{e}$.

Exemple 3.8 En 1-D. Mesurer la taille d'un être humain, représentée par le vecteur bipoint $\vec{v}$ =la "hauteur sous la toise".

On prend un objet de référence (par exemple le mètre étalon ou le pied anglais), et on l'appelle $\vec{e}$.

On définit la fonction linéaire $\ell : \mathbb{R} \to \mathbb{R}$ par $\ell(\vec{e}) = 1$ (une forme linéaire est déterminée dès qu'on connaît ses valeurs sur une base, et ici on est en dimension 1).

On pose $\lambda = \ell(\vec{v})$, donc on a $\vec{v} = \lambda\vec{e}$, et λ est la taille de $\vec{v}$ relativement à $\vec{e}$ (l'unité de mesure choisie).

Exemple : notons $\vec{e}_m$ et $\vec{e}_p$ les vecteurs représentant respectivement le mètre et le pied anglais. On leur associe ℓ_m et ℓ_p les formes linéaires définies par $\ell_m(\vec{e}_m) = 1$ et $\ell_p(\vec{e}_p) = 1$. Et donc $\ell_m(\vec{v}) = \lambda_m$ la taille en mètre, et $\ell_p(\vec{v}) = \lambda_p$ la taille en pied. Et comme $\ell_m(\vec{e}_p) = 0,3048$, on a $\lambda_m = 0.3048\lambda_p$. ▟

Exemple 3.9 En 3-D. En aviation, la convention internationale est :

• les "distances horizontales" sont données en mile nautique NM = Nautical Mile (où $1\,\text{NM} = 1852\,m$), et

• les "distances verticales" sont données en pied anglais ft = foot (où $1\,\text{ft} = 0.3048\,m$).

Une tour de contrôle qui demande au pilote sa position reçoit les informations "d NM" dans un direction (vue sur le radar) et "h ft". On suppose qu'on est dans un aéroport disposant de deux pistes d'atterrissage, disons pour fixer les idées, une vers le nord et l'autre vers le nord-est.

Modélisation : on se donne deux vecteurs horizontaux de longueur 1 NM $\vec{e}_1$ vers le nord et $\vec{e}_2$ vers le nord-est, tous deux de longueur 1 NM, et un vecteur $\vec{e}_3$ vertical de longueur 1 pied. On dispose de notre base de $\mathbb{R}^3$ adapté à notre aéroport. Et l'origine du repère est la tour de contrôle.

Pour obtenir les distances par rapport à la tour de contrôle et les pistes d'atterrissage, on définit les trois formes linéaires e^1, e^2, e^3 t.q. (e^1, e^2, e^3) est la base duale de $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$, i.e., les formes linéaires définies par, pour tout i, j,

$$e^i(\vec{e}_j) = \delta^i_j.$$

Ainsi, si $\vec{x} = \sum_i x^i\vec{e}_i$ est la position de l'avion (donc dans le repère $(O, (\vec{e}_1, \vec{e}_2, \vec{e}_3))$ de l'aéroport), l'altitude en pieds sera donnée :

$$e^3(\vec{x}) = x^3 \text{ ft, altitude en pieds.}$$

Et par exemple si l'avion en phase d'atterrissage vient du nord, sa distance à l'aéroport sera donnée par :

$$\begin{cases} e^1(\vec{x}) = x^1 \text{ NM, distance en miles entre l'avion et l'aéroport,} \\ e^2(\vec{x}) = x^2 = 0 \text{ car l'avion qui vient du nord.} \end{cases}$$

N.B. : on n'a pas utilisé de produit scalaire ici, ni de base orthonormée (normée par rapport à quoi, le pied ? le NM ? le mètre ?, pour quelle critère d'orthogonalité, les pistes ?, une orthogonalité euclidienne ?). ▟

4 Produit tensoriel de fonctions

Si $f : E \to \mathbb{R}$ et $g : F \to \mathbb{R}$ sont deux fonctions, alors leur produit tensoriel est la fonction dite "à variables séparées" $f \otimes g : E \times F \to \mathbb{R}$ définie par, pour $(\vec{x}, \vec{y}) \in E \times F$,

$$(f \otimes g)(\vec{x}, \vec{y}) = f(\vec{x})g(\vec{y}). \tag{4.1}$$

Attention : le produit tensoriel n'est pas commutatif : la commutativité n'a pas de sens si $F \neq E$, et par exemple $E = F = \mathbb{R}$ et $f(x) = x$ et $g(x) = x^2$ donnent $(f \otimes g)(x, y) = xy^2$ et $(g \otimes f)(x, y) = x^2 y$.

5 Expression tensorielle d'un produit scalaire

5.1 Base $a^i \otimes b^j$ dans $\mathcal{L}(E, F; \mathbb{R})$

Si $\varphi \in \mathcal{L}(E; \mathbb{R}) = E^*$ et $\psi \in \mathcal{L}(F; \mathbb{R}) = F^*$ (deux formes linéaires) alors (4.1) implique que la fonction $\varphi \otimes \psi : E \times F \to \mathbb{R}$ est une forme bilinéaire (immédiat) :

$$\forall (\vec{u}, \vec{v}) \in E \times F, \quad (\varphi \otimes \psi)(\vec{u}, \vec{v}) = \varphi(\vec{u})\psi(\vec{v}), \quad \text{et} \quad \varphi \otimes \psi \in \mathcal{L}(E, F; \mathbb{R}). \tag{5.1}$$

En particulier, si $(\vec{a}_i)$ est une base de E de base duale (a^i) (base dans E^*), et si $(\vec{b}_i)$ est une base de F de base duale (b^i) (base dans F^*), alors

$$\vec{u} = \sum_{i=1}^{n} u^i \vec{a}_i \quad \text{et} \quad \vec{v} = \sum_{j=1}^{m} v^j \vec{b}_j \quad \Longrightarrow \quad (a^i \otimes b^j)(\vec{u}, \vec{v}) = u^i v^j, \tag{5.2}$$

car $(a^i \otimes b^j)(\vec{u}, \vec{v}) = \sum_{k\ell} u^k v^\ell a^i(\vec{a}_k) b^j(b^\ell) = \sum_{k\ell} u^k v^\ell \delta^i_k \delta^j_\ell = u^i v^j$. En particulier $(a^i \otimes b^j)(\vec{a}_k, \vec{b}_\ell) = \delta^i_k \delta^j_\ell = 0$ pour tout $k \neq i$ et tout $\ell \neq j$, et $(a^i \otimes b^j)(\vec{a}_i, \vec{b}_j) = 1$, donc $[a^i \otimes b^j]_{|\vec{a},\vec{b}}$ est la matrice dont tous les termes sont nuls sauf celui à l'intersection de la ligne i et de la colonne j qui vaut 1.

Et quand $\varphi = \sum_{i=1}^n \varphi_i a^i$ et $\psi = \sum_{j=1}^m \psi_j b^j$, on a $\varphi \otimes \psi = \sum_{i=1}^n \sum_{j=1}^m \varphi_i \psi_j a^i \otimes b^j$, donc

$$[\varphi \otimes \psi]_{|\vec{a},\vec{b}} = [\varphi_i \psi_j]_{\substack{i=1,\dots,n \\ j=1,\dots,m}} \quad (= \sum_{i=1}^{n} \sum_{j=1}^{m} \varphi_i \psi_j [a^i \otimes b^j]_{|\vec{a},\vec{b}}). \tag{5.3}$$

5.2 Expression tensorielle d'une forme bilinéaire de $\mathcal{L}(E, F; \mathbb{R})$

Corollaire 5.1 *Si $(\vec{a}_i)$ est une base de E de base duale (a^i) (base dans E^*), et si $(\vec{b}_i)$ est une base de F de base duale (b^i) (base dans F^*), alors*

$$(a^i \otimes b^j)_{\substack{i=1,\dots,n \\ j=1,\dots,m}} \text{ est une base dans } \mathcal{L}(E, F; \mathbb{R}). \tag{5.4}$$

Donc, pour $g(\cdot, \cdot) \in \mathcal{L}(E, F; \mathbb{R})$ (une forme bilinéaire), on a "l'expression tensorielle de g" (où les bases utilisées sont explicites)

$$g(\cdot, \cdot) = \sum_{i=1}^{n} \sum_{j=1}^{m} g_{ij} a^i \otimes b^j \quad \text{où} \quad g_{ij} = g(\vec{a}_i, \vec{b}_j) \quad \text{et} \quad [g]_{|\vec{a},\vec{b}} = [g_{ij}], \tag{5.5}$$

les g_{ij} étant les composantes de $g(\cdot, \cdot)$ dans la base $(a^i \otimes b^j)$. Et si $\vec{u} = \sum_{i=1}^n u^i \vec{a}_i$ et $\vec{v} = \sum_{i=1}^m v^i \vec{b}_i$ alors :

$$g(\vec{u}, \vec{v}) = \sum_{i=1}^{n} \sum_{j=1}^{m} u^i g_{ij} v^j = [\vec{u}]^T_{|\vec{a}} . [g]_{|\vec{a},\vec{b}} . [\vec{v}]_{|\vec{b}} = [\vec{v}]^T_{|\vec{b}} . [g]_{|\vec{a},\vec{b}}{}^T . [\vec{u}]_{|\vec{a}} \quad (\in \mathbb{R}). \tag{5.6}$$

Preuve. Notons $g_{ij} = g(\vec{a}_i, \vec{b}_j)$; on a $\sum_{i=1}^n \sum_{j=1}^m g_{ij} a^i \otimes b^j \in \mathcal{L}(E, F, \mathbb{R})$ (est une forme bilinéaire car combinaison linéaire de formes bilinéaires), et

$$\begin{aligned}
(\sum_{k=1}^{n} \sum_{\ell=1}^{m} g_{k\ell} a^k \otimes b^\ell)(\vec{a}_i, \vec{b}_j) = & \sum_{k=1}^{n} \sum_{\ell=1}^{m} g_{k\ell} (a^k \otimes b^\ell)(\vec{a}_i, \vec{b}_j) = \sum_{k=1}^{n} \sum_{\ell=1}^{m} g_{k\ell} a^k(\vec{a}_i) b^\ell(\vec{b}_j) \\
= & \sum_{k=1}^{n} \sum_{\ell=1}^{m} g_{k\ell} \delta^k_i \delta^\ell_j = g_{ij} = g(\vec{a}_i, \vec{b}_j), \quad \forall i, j.
\end{aligned} \tag{5.7}$$

D'où $\sum_{k=1}^n \sum_{\ell=1}^m g_{k\ell} a^k \otimes b^\ell = g(\cdot, \cdot)$ (les deux formes bilinéaires ont mêmes valeurs sur les vecteurs de base). D'où la famille $(a^i \otimes b^j)_{\substack{i=1,\dots,n \\ j=1,\dots,m}}$ est génératrice.

Et si $\sum_{i=1}^n \sum_{j=1}^m g_{ij} a^i \otimes b^j = 0$, alors $(\sum_{i=1}^n \sum_{j=1}^m g_{ij} a^i \otimes b^j)(\vec{a}_k, \vec{b}_\ell) = 0$ pour tout k, ℓ, i.e., $g_{k\ell} = 0$ pour tout k, ℓ, d'où la famille $(a^i \otimes b^j)_{\substack{i=1,\dots,n \\ j=1,\dots,m}}$ est libre. D'où (5.4). ∎

6 Contractions tensorielles

Soit E, F et G trois espaces vectoriels.

6.1 D'une forme et d'un vecteur

Définition 6.1 Si $\ell \in E^* = \mathcal{L}(E;\mathbb{R})$ et $\vec{u} \in E$ alors le réel $\ell(\vec{u}) =^{\text{noté}} \ell.\vec{u} \in \mathbb{R}$ est également appelée l'opération de contraction de ℓ et $\vec{u}$.

6.2 D'un vecteur et d'une forme

Si $u \in E^{**} = \mathcal{L}(E^*;\mathbb{R})$ alors on lui associe naturellement et canoniquement le vecteur $\vec{u} \in E$ qui vérifie $u(\ell) = \ell(\vec{u})$ pour tout $\ell \in E^*$, cf. (9.2).

Définition 6.2 Si $u \in E^{**} = \mathcal{L}(E^*;\mathbb{R})$ et $\ell \in E^*$ alors le réel $u(\ell) = \ell(\vec{u}) =^{\text{noté}} u.\ell \in \mathbb{R}$ est appelée l'opération de contraction de u et ℓ ou de $\vec{u}$ et ℓ.

6.3 D'une forme bilinéaire (compatible) et d'un vecteur

Définition 6.3 Soit $\ell \in E^*$, $\vec{u} \in E$ et $\vec{v} \in F$. Et $\vec{v}$ est associé canoniquement à la forme linéaire $v \in F^{**}$, cf. (9.2).

La contraction de la forme bilinéaire $\vec{v} \otimes \ell \in \mathcal{L}(F^*, E;\mathbb{R})$ et du vecteur $\vec{u} \in E$ est le vecteur dans F défini par

$$(\vec{v} \otimes \underbrace{\ell)(\vec{u}}_{\text{contraction}}) = \ell(\vec{u})\,\vec{v} \ \in F. \tag{6.1}$$

6.4 De formes bilinéaires compatibles

Définition 6.4 Soit $\ell \in E^*$, $\vec{u} \in E$, $\vec{v} \in F$ et $\varphi \in G^*$. La contraction des formes bilinéaires $\vec{v} \otimes \ell \in \mathcal{L}(F^*, E;\mathbb{R})$ et $\vec{u} \otimes \varphi \in \mathcal{L}(E^*, G;\mathbb{R})$ est la forme bilinéaire dans $\mathcal{L}(F^*, G;\mathbb{R})$ définie par est le vecteur dans F défini par

$$(\vec{v} \otimes \underbrace{\ell)(\vec{u}}_{\text{contraction}} \otimes \varphi) = \ell(\vec{u})\,\vec{v} \otimes \varphi \overset{\text{noté}}{=} (\vec{v} \otimes \ell).(\vec{u} \otimes \varphi) \ \in \mathcal{L}(F^*, G;\mathbb{R}). \tag{6.2}$$

C'est contractions sont immédiatement généralisables à toutes formes multilinéaires compatibles.

7 Expressions tensorielles d'une application linéaire

Soit E, F deux espaces vectoriels, $\dim(E) = n$, $\dim(F) = m$.

7.1 Définition

Définition 7.1 Soit $\ell \in E^*$ et $\vec{v} \in F$. On définit l'application linéaire $(\vec{v} \otimes \ell)_L \in \mathcal{L}(E;F)$ à l'aide de (6.1) par

$$\forall \vec{u} \in E, \quad (\vec{v} \otimes \ell)_L(\vec{u}) := (\vec{v} \otimes \ell).\vec{u}, \tag{7.1}$$

i.e.,

$$\forall \vec{u} \in E, \quad (\vec{v} \otimes \ell)_L(\vec{u}) := \ell(\vec{u})\,\vec{v} \overset{\text{noté}}{=} (\vec{v} \otimes \ell).\vec{u} \ \in F. \tag{7.2}$$

Et l'opération $\vec{v} \otimes \ell.\vec{u}$ est appelée la contraction de l'application linéaire $\vec{v} \otimes \ell$ et du vecteur $\vec{u}$.

(On a l'isomorphisme canonique naturel $\mathcal{L}(E;F) \leftrightarrow \mathcal{L}(F^*, E;\mathbb{R})$ donné par (7.1), voir poly suivant.)

Expression dans des bases $(\vec{a}_i)$ et $(\vec{b}_i)$ de E et F :

$$[\vec{v}\otimes\ell]_{|\vec{a},\vec{b}} = [v^i\ell_j]_{\substack{i=1,\dots,n\\ j=1,\dots,m}} = [\vec{v}]_{|\vec{a}}.[\ell]_{|\vec{b}} \quad \text{quand} \quad \vec{v}=\sum_{i=1}^m v^i\vec{b}_i \quad \text{et} \quad \ell=\sum_{j=1}^n \ell_i a^i. \tag{7.3}$$

($[\vec{v}]_{|\vec{a}}.[\ell]_{|\vec{b}}$ est le produit d'une matrice $n*1$ par une matrice $1*m$: c'est une matrice $n*m$.) En effet, $(\vec{v}\otimes\ell).\vec{a}_j = (\ell.\vec{a}_j)\vec{v} = \ell_j\vec{v} = \ell_j\sum_{i=1}^m v^i\vec{b}_j = \sum_{i=1}^m (v^i\ell_j)\vec{b}_i$ donne la j-ème colonne de $[\vec{v}\otimes\ell]_{|\vec{a},\vec{b}}$. Et le produit de $[\vec{v}]_{|\vec{a}}$ matrice $m*1$ et de $[\ell]_{|\vec{b}}$ matrice $1*n$ est $[\vec{v}]_{|\vec{a}}.[\ell]_{|\vec{b}}$ matrice $m*n$.

Noter que $[\vec{v}\otimes\ell]_{|\vec{a},\vec{b}}$ est une matrice dégénérée puisque $\mathrm{rang}(\mathrm{Im}(\vec{v}\otimes\ell)) = \dim(\mathrm{Vect}\{\vec{v}\}) = 1$.

Exercice 7.2 Soit V un hyperplan vectoriel dans $\mathbb{R}^n$, muni d'un produit scalaire euclidien $(\cdot,\cdot)_{\mathbb{R}^n}$, dont un vecteur normal unitaire est noté $\vec{n}$: donc $V=\mathrm{Vect}\{\vec{n}\}^\perp$ et $V^\perp=\mathrm{Vect}\{\vec{n}\}$.

Soit $n^\flat:\mathbb{R}^n\to\mathbb{R}$ la forme linéaire définie par $n^\flat(\vec{v}) = (\vec{n},\vec{v})_{\mathbb{R}^n}$.

Montrer que $\vec{n}\otimes n^\flat$ est l'application linéaire de projection orthogonale sur $\mathrm{Vect}\{\vec{n}\} = V^\perp$ = la droite vectorielle engendrée par $\vec{n}$. Et que $L = I-\vec{n}\otimes n^\flat$ est l'application linéaire de projection orthogonale sur V.

Réponse. Soit $\vec{v}\in\mathbb{R}^n$. On a $\vec{n}\otimes n^\flat.\vec{v} = (n^\flat.\vec{v})\vec{n} \overset{\text{déf } n^\flat}{=} (\vec{n},\vec{v})_{\mathbb{R}^n}\vec{n} = \vec{v}_\perp \in V^\perp$: c'est le projeté de $\vec{v}$ le long de $\vec{n}$. Et $(\vec{v}-\vec{v}_\perp,\vec{n})_{\mathbb{R}^n} = (\vec{v},\vec{n})_{\mathbb{R}^n} - (\vec{v}_\perp,\vec{n})_{\mathbb{R}^n} = (\vec{v},\vec{n})_{\mathbb{R}^n} - ((\vec{n},\vec{v})_{\mathbb{R}^n}\vec{n},\vec{n})_{\mathbb{R}^n} = 0$, donc $\vec{v}-\vec{v}_\perp = (I-\vec{n}\otimes n^\flat).\vec{v}\in V$, donc $I-\vec{n}\otimes n^\flat$ est l'opérateur de projection orthogonale sur V. ∎

7.2 Base $(\vec{b}_i\otimes a^j)$ de $\mathcal{L}(E;F)$

La définition 7.1 donne en particulier : si $(\vec{a}_i)$ et $(\vec{b}_i)$ sont des bases de E et F, alors $\vec{b}_i\otimes a^i\in\mathcal{L}(E;F)$ est donnée par

$$(\vec{b}_i\otimes a^j)(\vec{u}) = (\vec{b}_i\otimes a^j)(\vec{u}) = u^j\vec{b}_i \quad \text{quand} \quad \vec{u}=\sum_{j=1}^n u^j\vec{a}_j, \tag{7.4}$$

et

$$[\vec{b}_i\otimes a^j]_{|\vec{a},\vec{b}} = \begin{array}{c} \text{colonne } j \\ \downarrow \\ \begin{pmatrix} & & \vdots & & \\ \cdots & 0 & 0 & 0 & \cdots \\ \cdots & 0 & 1 & 0 & \cdots \\ \cdots & 0 & 0 & 0 & \cdots \\ & & \vdots & & \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \\ \\ \leftarrow \text{ligne } i, \\ \\ \\ \end{array} \tag{7.5}$$

tous les termes étant nuls sauf le terme à l'intersection de la ligne i et de la colonne j qui vaut 1. Autrement dit, $[\vec{b}_i\otimes a^j]_{|\vec{a},\vec{b}} = [\delta_i^k\delta_\ell^j]_{\substack{k=1,\dots,n\\ \ell=1,\dots,m}}$ (écriture assez peu lisible...).

7.3 Expression tensorielle d'une application linéaire de $\mathcal{L}(E;F)$

Proposition 7.3 *Les applications linéaires $\vec{b}_i\otimes a^j\in\mathcal{L}(E;F)$ forment une base de $\mathcal{L}(E;F)$. Et si $L\in\mathcal{L}(E;F)$ alors, avec (7.4),*

$$L.\vec{a}_j = \sum_{i=1}^m L^i_j\vec{b}_i \iff L = \sum_{i=1}^m\sum_{j=1}^n L^i_j\vec{b}_i\otimes a^j \iff L^i_j = b^i.L.\vec{a}_j, \tag{7.6}$$

et $[L]_{|\vec{a},\vec{b}} = [L^i_j] = \sum_{i=1}^m L^i_j\,[\vec{b}_i\otimes a^j]_{|\vec{a},\vec{b}}$, cf. (1.38). Et on retrouve (1.39) : pour $\vec{u}=\sum_{i=1}^n u^i\vec{a}_i\in E$,

$$L.\vec{u} = \sum_{i=1}^m\sum_{j=1}^n L^i_j u^j\vec{b}_i, \quad \text{et} \quad [L.\vec{u}]_{|\vec{b}} = [L.\vec{u}]_{|\vec{a},\vec{b}}.[\vec{u}]_{|\vec{a}}. \tag{7.7}$$

Preuve. $\sum_{i=1}^m\sum_{j=1}^n L^i_j\vec{b}_i\otimes a^j$ est une somme d'applications linéaires, donc est linéaire.

$(\vec{b}_i\otimes a^j)$ est une famille libre : $(\sum_{ij}\alpha_{ij}\vec{b}_i\otimes a^j)(\vec{a}_k) = 0$ donne $\sum_i\alpha_{ik}\vec{b}_i = 0$ donne $\alpha_{ik}=0$ pour tout i, vrai pour tout k.

Et $(\sum_{ij}L^i_j\vec{b}_i\otimes a^j)(\vec{a}_k) = \sum_{ij}L^i_j(\vec{b}_i\otimes a^j)(\vec{a}_k) = \sum_{ij}L^i_j\vec{b}_i a^j(\vec{a}_k) = \sum_{ij}L^i_j\vec{b}_i\delta^j_k = \sum_i L^i_k\vec{b}_i = L.\vec{a}_k$, vrai pour tout k, donc $\sum_{ij}L^i_j\vec{b}_i\otimes a^j = L$, d'où $(\vec{b}_i\otimes a^j)$ est une génératrice, et (7.6) et (7.7). ∎

Exercice 7.4 Soit $L \in \mathcal{L}(E;F)$ et $M \in \mathcal{L}(F;G)$, et $[L]_{|\vec{a},\vec{b}} = [L^i_j]$ et $[M]_{|\vec{b},\vec{c}} = [M^i_j]$. Retrouver, à l'aide du calcul tensoriel,

$$M.L = \sum_{ijk} M^i_k L^k_j \vec{c}_i \otimes a^j, \quad \text{et} \quad [M.L]_{|\vec{a},\vec{c}} = [\sum_k M^i_k L^k_j]_{\substack{i \\ j}}. \tag{7.8}$$

Réponse. $M.L = \sum_{ijk\ell} M^i_j L^k_\ell (\vec{c}_i \otimes b^j).(\vec{b}_k \otimes a^\ell) = \sum_{ijk\ell} M^i_j L^k_\ell (b^j.\vec{b}_k)\vec{c}_i \otimes a^\ell = \sum_{ijk\ell} M^i_j L^k_\ell \delta^j_k \vec{c}_i \otimes a^\ell = \sum_{ij\ell} M^i_j L^j_\ell \vec{c}_i \otimes a^\ell$, i.e. (7.8). ▪

8 Formules de changement de base

8.1 Matrice de passage P

Soit $(\vec{e}_{i,old})$ et $(\vec{e}_{i,new})$ deux bases dans E, et soit $\mathcal{P} \in \mathcal{L}(E)$ l'endomorphisme défini par, pour tout j,

$$\mathcal{P}(\vec{e}_{j,old}) = \vec{e}_{j,new} \overset{\text{noté}}{=} \sum_{i=1}^n P^i_j \vec{e}_{i,old}, \quad \text{i.e.} \quad [\vec{e}_{j,new}]_{|old} = \begin{pmatrix} P^1_j \\ \vdots \\ P^n_j \end{pmatrix}, \quad \text{et} \quad [\mathcal{P}]_{|old} = [P^i_j] \overset{\text{noté}}{=} P, \tag{8.1}$$

où on a noté $[.]_{|old} := [.]_{|\vec{e}_{old}}$. Donc la matrice P stocke dans sa j-ème colonne les coordonnées du vecteur $\vec{e}_{j,new}$ dans la base $(\vec{e}_{i,old})$.

Définition 8.1 $[\mathcal{P}]_{|\vec{e}_{old}} = P = [P^i_j]$ est appelé matrice de passage de la base $(\vec{e}_{i,old})$ à la base $(\vec{e}_{i,new})$, ou matrice de changement de base.

Exemple 8.2 On fait la rotation d'angle θ et la dilatation 3, i.e. on se place dans $\mathbb{R}^2$, on prend $(\vec{e}_1, \vec{e}_2)$ b.o.n. pour le produit scalaire euclidien, et $P = 3\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. La base $(\vec{e}_1, \vec{e}_2)$ est transformée en la base $(\vec{b}_1, \vec{b}_2)$ où $\vec{b}_1 = L(\vec{e}_1) = P.\begin{pmatrix} 1 \\ 0 \end{pmatrix}_{|(\vec{e})} = 3\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}_{|(\vec{e})} = 3\cos\theta\vec{e}_1 + 3\sin\theta\vec{e}_2$ (première colonne de P) et $\vec{b}_2 = L(\vec{e}_2) = P.\begin{pmatrix} 0 \\ 1 \end{pmatrix}_{|(\vec{e})} = 3\begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}_{|(\vec{e})} = -3\sin\theta\vec{e}_1 + 3\cos\theta\vec{e}_2$ (deuxième colonne de P). Faire un dessin. ▪

8.2 Matrice de passage inverse : $Q = P^{-1}$

Un endomorphisme de changement de base est inversible (car $\dim(\mathrm{Im}(\mathcal{P})) = n = \dim(E)$). Avec (8.1) on note

$$\mathcal{Q} := \mathcal{P}^{-1} \in \mathcal{L}(E), \tag{8.2}$$

donc, pour tout j,

$$\mathcal{Q}.\vec{e}_{j,new} = \vec{e}_{j,old} \overset{\text{noté}}{=} \sum_{i=1}^n Q^i_j \vec{e}_{i,new}, \quad \text{i.e.} \quad [\vec{e}_{j,old}]_{|new} = \begin{pmatrix} Q^1_j \\ \vdots \\ Q^n_j \end{pmatrix}, \quad \text{et} \quad [\mathcal{Q}]_{|new} = [Q^i_j] \overset{\text{noté}}{=} Q. \tag{8.3}$$

Donc la matrice Q stocke dans sa j-ème colonne les coordonnées du vecteur $\vec{e}_{j,old}$ dans la base $(\vec{e}_{i,new})$.

Proposition 8.3 *On a :*

$$Q = P^{-1}, \quad \text{i.e.} \quad [\mathcal{P}^{-1}]_{|\vec{e}_{new}} = [\mathcal{P}]_{|\vec{e}_{old}}{}^{-1}. \tag{8.4}$$

Et, pour tout j :

$$\mathcal{P}.\vec{e}_{j,new} = \sum_{i=1}^n P^i_j \vec{e}_{i,new} \quad \text{et} \quad \mathcal{Q}.\vec{e}_{j,old} = \sum_{i=1}^n Q^i_j \vec{e}_{i,old}, \tag{8.5}$$

i.e., P est la matrice de $\mathcal{P}$ à la fois relativement à la base $(\vec{e}_{i,old})$ et à la base $(\vec{e}_{i,new})$, et Q est la matrice de $\mathcal{Q}$ à la fois relativement à la base $(\vec{e}_{i,new})$ et à la base $(\vec{e}_{i,old})$. (C'est trivial avec (8.13) : $[\mathcal{P}]_{|\vec{e}_{new}} = P^{-1}.[\mathcal{P}]_{|\vec{e}_{old}}.P$.)

Preuve. Vérifions $Q = P^{-1}$. Avec $\vec{e}_{j,new} = \sum_i P^i_j \vec{e}_{i,old}$ et $\vec{e}_{i,old} = \sum_k Q^k_i \vec{e}_{k,new}$ on a :

$$\vec{e}_{j,new} = \sum_i P^i_j (\sum_k Q^k_i \vec{e}_{k,new}) = \sum_k (\sum_i Q^k_i P^i_j) \vec{e}_{k,new} = \sum_k (Q.P)^k_j \vec{e}_{k,new},$$

d'où $(Q.P)^k_j = \delta^k_j$, car $(\vec{e}_{i,new})$ est une base, soit $Q.P = I$.

Montrons (8.5). Notons $Z = [\mathcal{P}^{-1}]_{\vec{e}_{old}}$, i.e. $\mathcal{P}^{-1}.\vec{e}_{j,old} = \sum_i Z^i_j \vec{e}_{i,old}$. Alors :

$$\vec{e}_{j,old} = \mathcal{P}^{-1}.\vec{e}_{j,new} = L^{-1}.(\sum_k P^k_j \vec{e}_{k,old}) = \sum_k P^k_j L^{-1}.\vec{e}_{k,old} = \sum_k P^k_j \sum_i Z^i_k \vec{e}_{i,old} = \sum_i (ZP)^i_j \vec{e}_{i,old},$$

d'où $Z.P = I$, i.e. $Z = Q$, d'où $(8.5)_2$. Notons $Y = [\mathcal{P}]_{|\vec{e}_{new}}$, i.e. $\mathcal{P}.\vec{b}_j = \sum_i Y^i_j \vec{e}_{i,new}$. Alors :

$$\vec{e}_{j,new} = \mathcal{P}^{-1}.(\mathcal{P}.\vec{e}_{j,new}) = \mathcal{P}^{-1}.(\sum_i Y^i_j \vec{e}_{i,new}) = \sum_i Y^i_j \mathcal{P}^{-1}.\vec{e}_{i,new} = \sum_{ik} Y^i_j Q^k_i \vec{e}_{k,new} = \sum_k (QY)^k_j \vec{e}_{k,new},$$

d'où $QY = I$, i.e. $Y = Q^{-1} = P$, d'où $(8.5)_1$. ∎

8.3 Pour les composantes des vecteurs : $[\vec{v}]_{|new} = P^{-1}.[\vec{v}]_{|old}$

Soit $\vec{v} \in E$, et on note

$$\vec{v} = \sum_{i=1}^n v^i_{old} \vec{e}_{i,old} = \sum_i v^i_{new} \vec{e}_{i,new}, \quad \text{i.e.} \quad [\vec{v}]_{|old} = \begin{pmatrix} v^1_{old} \\ \vdots \\ v^n_{old} \end{pmatrix} \quad \text{et} \quad [\vec{v}]_{|new} = \begin{pmatrix} v^1_{new} \\ \vdots \\ v^n_{new} \end{pmatrix}. \tag{8.6}$$

Proposition 8.4 *On a, pour tout $i = 1, ..., n$:*

$$\forall i = 1, ..., n, \quad v^i_{new} = \sum_{j=1}^n Q^i_j \, v^j_{old}, \quad \text{i.e.} \quad [\vec{v}]_{|new} = P^{-1}.[\vec{v}]_{|old} \tag{8.7}$$

((8.7) est une formule de “contravariance” : quand les vecteurs de base sont transformés par P, cf. (8.1), les composantes des vecteurs sont transformées par $P^{-1} = Q$, cf. (8.7).)

Preuve. On a :

$$\sum_{i=1}^n v^i_{new} \vec{e}_{i,new} = \vec{v} = \sum_{j=1}^n v^j_{old} \vec{e}_{j,old} = \sum_{j=1}^n v^j_{old} (\sum_{i=1}^n Q^i_j \vec{e}_{i,new}) = \sum_{i=1}^n (\sum_{j=1}^n Q^i_j v^j_{old}) \vec{e}_{i,new},$$

d'où (8.7). ∎

8.4 Formules de changement de bases pour les bases duales

Soit (e^i_{old}), resp. (e^i_{new}), la base duale de $(\vec{e}_{i,old})$, resp. $(\vec{e}_{i,new})$.

Proposition 8.5 *On a :*

$$e^i_{new} = \sum_{j=1}^n Q^i_j \, e^j_{old} \in E^*, \quad \text{i.e.} \quad [e^i_{new}]_{|old} = \begin{pmatrix} Q^i_1 & \cdots & Q^i_n \end{pmatrix}. \tag{8.8}$$

i.e. la ligne i de $Q = P^{-1}$ stocke les composantes de e^i_{new} dans la base (e^i_{old}). Ou encore :

$$e^i_{old} = \sum_{j=1}^n P^i_j \, e^j_{new} \in E^* \quad (\text{quand } \vec{e}_{j,new} = \sum_{i=1}^n P^i_j \vec{e}_{i,old} \in E), \tag{8.9}$$

i.e. la ligne i de P stocke les composantes de e^i dans la base (b^j).

Et on retrouve les formules de changement de coordonnées (8.7) à l'aide du produit matriciel usuel :

$$v^i_{new} = e^i_{new}(\vec{v}) = e^i_{new}.\vec{v} = \sum_j Q^i_j v^j_e \quad (= [e^i_{new}]_{|old}.[\vec{v}]_{|old} = \begin{pmatrix} Q^i_1 & \dots & Q^i_n \end{pmatrix} . \begin{pmatrix} v^1_{old} \\ \vdots \\ v^n_{old} \end{pmatrix}). \tag{8.10}$$

Preuve. On a :

$$e^i_{new}(\vec{e}_{j,old}) = e^i_{new}(\sum_{k=1}^n Q^k_j \vec{e}_{k,new}) = \sum_{k=1}^n Q^k_j e^i_{new}(\vec{e}_{k,new}) = \sum_{k=1}^n Q^k_j \delta^i_k = Q^i_j,$$

d'où $e^i_{new} = \sum_{j=1}^n Q^i_j \, e^j_{old}$, ce pour tout i, i.e. (8.8). ▪

Exemple 8.6 Dans $\mathbb{R}^2$ soient $(\vec{a}_1, \vec{a}_2)$ (old) et $(\vec{b}_1, \vec{b}_2)$ (new) deux bases avec $[\vec{b}_1]_{|\vec{a}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $[\vec{b}_2]_{|\vec{a}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Donc la matrice de passage de $(\vec{a}_i)$ vers $(\vec{b}_i)$ est $P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, et $Q = P^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$.

(8.8) donne (lecture de lignes de Q) $[b^1]_{|old} = (\,1 \quad -1\,)$, i.e. $b^1 = e^1 - e^2$, et $[b^2]_{|old} = (\,0 \quad 1\,)$, i.e. $b^2 = e^2$.

Vérification : $(e^1{-}e^2)(\vec{b}_1) = (e^1{-}e^2)(\vec{e}_1) = 1$ et $(e^1{-}e^2)(\vec{b}_2) = (e^1{-}e^2)(\vec{e}_1{+}\vec{e}_2) = 0$ donc $b^1 = e^1 - e^2$; et $e^2(\vec{b}_1) = e^2(\vec{e}_1) = 0$ et $e^2(\vec{b}_2) = e^2(\vec{e}_1{+}\vec{e}_2) = 1$ donc $b^2 = e^2$. ▪

Exemple 8.7 Soit $\theta \in \mathbb{R}$ fixé, et soient $[\vec{b}_1]_{|\vec{a}} = \alpha \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$ et $[\vec{b}_2]_{|\vec{a}} = \alpha \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}_{|(\vec{e})}$. La matrice de passage de $(\vec{a}_1, \vec{a}_2)$ à $(\vec{b}_1, \vec{b}_2)$ est la matrice de rotation-homothétie $P = \alpha \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ (exemple : passage entre anglais et français). Donc $Q = P^{-1} = \frac{1}{\alpha} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$. D'où $b^1 = \frac{1}{\alpha} \cos\theta \, a^1 + \frac{1}{\alpha} \sin\theta \, a^2$ et $b^2 = -\frac{1}{\alpha} \sin\theta \, a^1 + \frac{1}{\alpha} \cos\theta \, a^2$ (lecture des lignes de Q). ▪

8.5 Pour les composantes des formes linéaires : $[\ell]_{new} = [\ell]_{old}.P$

Soit $\ell \in E^*$ et soit $(\ell_{i,old})_{i=1,\dots,n}$ ses composantes sur la base $(\vec{e}_{i,old})$ et soit $(\ell_{i,new})_{i=1,\dots,n}$ ses composantes sur la base $(\vec{e}_{i,new})$, i.e. :

$$\ell = \sum_{i=1}^n \ell_{i,old} e^i_{old} = \sum_{i=1}^n \ell_{i,new} e^i_{new}. \tag{8.11}$$

Proposition 8.8 *On a, pour tout j (formule de covariance) :*

$$[\ell]_{|new} = [\ell]_{|old}.P, \quad \text{i.e.} \quad \ell_{j,new} = \sum_{i=1}^n \ell_{i,old} P^i_j, \ \forall j = 1, ..., n. \tag{8.12}$$

(On rappelle qu'une forme linéaire est représentée par une matrice ligne.)

Preuve. $\sum_j \ell_{j,old} e^j_{old} = \ell = \sum_i \ell_{i,new} e^i_{new} = \sum_{ij} \ell_{i,new} Q^i_j e^j_{old} = \sum_j (\sum_i \ell_{i,new} Q^i_j) e^j_{old}$ donne $\ell_{j,old} = \sum_i \ell_{i,new} Q^i_j$, soit $[\ell]_{|old} = [\ell]_{|new}.Q$, soit $[\ell]_{|old}.P = [\ell]_{|new}$, i.e. (8.12). ▪

Remarque 8.9 On a aussi $\begin{pmatrix} \ell_{b1} \\ \vdots \\ \ell_{bn} \end{pmatrix} = P^T . \begin{pmatrix} \ell_{e1} \\ \vdots \\ \ell_{en} \end{pmatrix}$, mais il est préférable de ne pas utiliser les matrices colonnes pour représenter les formes linéaires : sous peine de confondre covariance et contravariance. ▪

8.6 Pour les composantes d'un endomorphisme : $[L]_{new} = P^{-1}.[L]_{old}.P$

Soit $L \in \mathcal{L}(E)$ (un endomorphisme de E).

Proposition 8.10 *On a :*

$$[L]_{new} = P^{-1}.[L]_{old}.P. \tag{8.13}$$

Preuve. Soit $[L]_{new} = [A^i_j]$ et $[L]_{old} = [B^i_j]$ les matrices de ses composantes relativement aux bases $(\vec{e}_{i,old})$ et $(\vec{e}_{i,new})$:

$$L.\vec{e}_{j,old} = \sum_{i=1}^n A^i_j \vec{e}_{i,old}, \quad \text{et} \quad L.\vec{e}_{j,new} = \sum_{i=1}^n B^i_j \vec{e}_{i,new}. \tag{8.14}$$

Donc

$$\begin{aligned}\sum_\ell B^\ell_j \vec{e}_{\ell,new} = L.\vec{e}_{j,new} &= L.\sum_k P^k_j \vec{e}_{k,old} = \sum_k P^k_j L.\vec{e}_{k,old} = \sum_{ki} P^k_j A^i_k \vec{e}_{i,old} = \sum_{ki\ell} P^k_j A^i_k Q^\ell_i \vec{e}_{\ell,new} \\ &= \sum_\ell (Q.A.P)^\ell_j \vec{e}_{\ell,new},\end{aligned}$$

d'où $B^\ell_j = (Q.A.P)^\ell_j$ pour tout ℓ, j, soit $B = Q.A.P$, i.e. (8.13).

$$M = \sum_{ij} A^i_j \vec{e}_i \otimes e^j = \sum_{ij} \sum_{k\ell} A^i_j Q^k_i P^j_\ell \vec{b}_k \otimes b^\ell = \sum_{k\ell} (\sum_{ij} Q^k_i B^i_j P^j_\ell) \vec{b}_k \otimes b^\ell,$$

avec également $M = \sum_{k\ell} B^k_\ell \vec{b}_k \otimes b^\ell$, d'où $B^k_\ell = \sum_{ij} Q^k_i A^i_j P^j_\ell$, comme annoncé. ▪

Exercice 8.11 Démontrer la proposition précédente avec le calcul tensoriel.

Réponse.

$$\begin{aligned}\sum_{k\ell} B^k_\ell \vec{e}_{k,new} \otimes e^\ell_{new} = L = \sum_{ij} A^i_j \vec{e}_{i,old} \otimes e^j_{old} &= \sum_{ij} A^i_j \sum_{k\ell} Q^k_i P^j_\ell \vec{e}_{k,new} \otimes e^\ell_{new} \\ &= \sum_{k\ell} (\sum_{ij} Q^k_i A^i_j P^j_\ell) \vec{b}_k \otimes b^\ell = \sum_{k\ell} (Q.A.P)^k_\ell \vec{b}_k \otimes b^\ell,\end{aligned}$$

d'où $B = Q.A.P$. ▪

8.7 Pour les composantes d'un produit scalaire : $[g]_{new} = P^T.[g]_{old}.P$

Soit $g \in \mathcal{L}(E, E; \mathbb{R})$ une forme bilinéaire sur E.

Proposition 8.12 *On a :*

$$[g]_{|new} = P^T.[g]_{|old}.P. \tag{8.15}$$

Preuve. Soit $A = [A_{ij}]$ la matrice de $g \in T^0_2$ relativement à la base $(\vec{e}_{i,old})$ et soit $B = [B_{ij}]$ la matrice de g relativement à la base $(\vec{e}_{i,new})$, i.e., $A_{ij} = g(\vec{e}_{i,old}, \vec{e}_{j,old})$ et $B_{ij} = g(\vec{e}_{i,new}, \vec{e}_{j,new})$. On a

$$\begin{aligned}\sum_{k\ell} B_{k\ell} e^k_{new} \otimes e^\ell_{new} = g = \sum_{ij} A_{ij} e^i_{old} \otimes e^j_{old} &= \sum_{ij} A_{ij} \sum_{k\ell} P^i_k P^j_\ell e^k_{new} \otimes e^\ell_{new} \\ &= \sum_{k\ell} (\sum_{ij} P^i_k A_{ij} P^j_\ell) e^k_{new} \otimes e^\ell_{new},\end{aligned}$$

d'où $B_{k\ell} = \sum_{ij} P^i_k A_{ij} P^j_\ell = \sum_{ij} (P^T)^k_i A_{ij} P^j_\ell = (P^T.A.P)^k_\ell$ pour tout k, ℓ, d'où (8.15). ▪

Exercice 8.13 Soit $\underline{\underline{t}} \in \mathcal{L}(E^*)$ (tenseur uniforme $\binom{2}{0}$) : sur les bases

$$\underline{\underline{t}} = \sum_{ij} A^{ij} \vec{e}_{i,old} \otimes \vec{e}_{j,old} = \sum_{ij} B^{ij} \vec{e}_{i,new} \otimes \vec{e}_{j,new}. \tag{8.16}$$

Montrer

$$[\underline{\underline{t}}]_{|new} = P^{-1}.[\underline{\underline{t}}]_{|old}.P^{-T}. \tag{8.17}$$

▪

8.8 Loi d'inertie de Sylvester et signature

La loi d'inertie de Sylvester concerne les formes quadratiques $Q : E \to \mathbb{R}$, ou encore les formes bilinéaires symétriques $b(\cdot,\cdot) \in \mathcal{L}(E,E;\mathbb{R})$ associées données par $Q(\vec{x}) = b(\vec{x},\vec{x})$.

Définition 8.14 Deux matrices A et B sont congruentes ssi il existe une matrice P inversible telle que :

$$B = P^T.A.P. \tag{8.18}$$

Deux matrices congruentes correspondent donc aux matrices d'une même forme bilinéaire $g(\cdot,\cdot) : E \times E \to \mathbb{R}$ exprimée dans deux bases $(\vec{a}_i)$ et $(\vec{b}_i)$, P étant la matrice de passage de $(\vec{a}_i)$ vers $(\vec{b}_i)$:

$$g(\cdot,\cdot) = \sum_{ij} A_{ij}\, a^i \otimes a^j = \sum_{ij} B_{ij}\, b^i \otimes b^j.$$

Définition 8.15 Deux matrices A et B sont semblables ssi il existe une matrice P inversible telle que :

$$B = P^{-1}.A.P. \tag{8.19}$$

Deux matrices semblables correspondent donc aux matrices d'une même endomorphisme $L : E \to E$ exprimée dans deux bases $(\vec{a}_i)$ et $(\vec{b}_i)$, P étant la matrice de passage de $(\vec{a}_i)$ vers $(\vec{b}_i)$:

$$L = \sum_{ij} A^i_j\, \vec{a}_i \otimes a^j = \sum_{ij} B^i_j\, \vec{b}_i \otimes b^j.$$

Remarque 8.16 Différence essentielle :

1- Cas des formes bilinéaires. Si $A = [A_{ij}] = [g(\vec{a}_j,\vec{a}_j)]$ est la matrice d'une forme bilinéaire $g(\cdot,\cdot) \in L(E,E;\mathbb{R})$ relativement à une base $(\vec{a}_i)$, les valeurs propres de cette matrice A sont les racines λ_i du polynôme $p(\lambda) = \det(A - \lambda I)$. Et relativement à une autre base $(\vec{b}_i)$, si $B = [B_{ij}] = [g(\vec{b}_j,\vec{b}_j)]$ est la matrice de la même forme linéaire, les valeurs propres de cette matrice sont les racines μ_i de $\det(B - \mu I)$. Mais les μ_i sont différents des λ_i en général, car :

$$\det(B - \mu I) = \det(P^T.A.P - \mu I) = \det(P)^2 \det(A - \mu M) \tag{8.20}$$

où $M = P^{-T}.P^{-1}$, autrement dit les μ_i sont les valeurs propres généralisées de A relativement à la matrice M (voir polycopié "diagonalisation...").

2- Cas des endomorphismes. Si $A = [A^i_j] = [a^i.(\mathcal{L}.\vec{a}_j)]$ est la matrice d'un endomorphisme $\mathcal{L} \in L(E;E)$ relativement à une base $(\vec{a}_i)$, les valeurs propres de cette matrice A sont les racines λ_i du polynôme $p(\lambda) = \det(A - \lambda I)$. Et relativement à une autre base $(\vec{b}_i)$, si $B = [B^i_j] = [b^i.\mathcal{L}.\vec{b}_j]$ est la matrice de la même forme linéaire, les valeurs propres de cette matrice sont les racines μ_i de $\det(B - \mu I)$. Avec :

$$\det(B - \mu I) = \det(P^{-1}.A.P - \mu I) = \det(P)^2 \det(A - \mu I), \tag{8.21}$$

et donc A a les mêmes valeurs propres que B : l'ensemble $\{\lambda_i : i = 1,...,n\}$ des valeurs propres d'un endomorphisme est conservé par changement de base. ▪

Malgré cette remarque, on a quand même :

Proposition 8.17 Loi d'inertie de Sylvester et définition de la signature.

"Le nombre s_+ de valeurs propres positives de la matrice d'une forme bilinéaire symétrique $g(\cdot,\cdot) \in L(E,E;\mathbb{R})$ est conservé par changement de base. De même le nombre s_- de valeurs propres négatives de la matrice d'une forme bilinéaire est conservé par changement de base".

Le couple (s_+,s_-) est appelé la signature de la forme bilinéaire symétrique. Et le réel $r = s_+ + s_-$ est appelé le rang de la forme bilinéaire symétrique.

Autrement dit, A est B t.q. (8.18) ont même nombre de valeurs propres > 0 et < 0.

De plus, il existe une base $(\vec{\varphi}_i)$ de E telle que, (φ^i) étant la base duale :

$$g(\cdot,\cdot) = \sum_{i=1}^{s_+} \varphi^i \otimes \varphi^i - \sum_{i=s_+ +1}^{s_+ + s_-} \varphi^i \otimes \varphi^i, \tag{8.22}$$

la matrice de $g(\cdot,\cdot)$ dans la base $(\vec{\varphi}_i)$ s'écrivant $[g]_{|\vec{\varphi}} = \begin{pmatrix} I_{s_+} & 0 & 0 \\ 0 & -I_{s_-} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$

Preuve. Soit $(\vec{a}_i)_{i=1,...,n}$ une base de E. La matrice $A = [g(\vec{a}_i, \vec{a}_j)]$ est une matrice $n*n$ symétrique réelle car $g(\cdot,\cdot)$ est symétrique. Ce tableau de réel sera également noté $A = [M^i_j]$ où donc $M^i_j =^{\text{déf}} A_{ij}$, la matrice $M = [M^i_j]$ étant considérée être la matrice d'un endomorphisme $L \in \mathcal{L}(\mathbb{R}^n)$ relativement à la base canonique $(\vec{E}_i)$ de $\mathbb{R}^n$: $L.\vec{E}_j = \sum_{i=1}^n M^i_j \vec{E}_i$ où $M^i_j = M^j_i$ pour tout i,j. Donc L est diagonalisable dans une b.o.n. de $\mathbb{R}^n$ muni son produit scalaire canonique $(\cdot,\cdot)_{\mathbb{R}^n}$: il existe n réels $\lambda_1,...,\lambda_n$ (les valeurs propres) et n vecteurs $\vec{p}_1,...,\vec{p}_n$ formant une b.o.n. dans $\mathbb{R}^n$ (vecteurs propres associés) t.q. :

$$L.\vec{p}_j = \lambda_j \vec{p}_j, \quad \vec{p}_j = \sum_i P^i_j \vec{E}_i, \quad P = [P^i_j], \quad \text{diag}(\lambda_1,...,\lambda_n) = D = P^{-1}.M.P, \quad P^{-1} = P^T.$$

On définit alors la base $(\vec{b}_j)$ de E par (formule de changement de base) :

$$\vec{b}_j = \sum_i P^i_j \vec{a}_i \quad \in E.$$

On obtient, pour tout i,j :

$$g(\vec{b}_i, \vec{b}_j) = \sum_{k\ell} P^k_i P^\ell_j g(\vec{a}_k, \vec{a}_\ell) = \sum_{k\ell} (P^T)^i_k M^k_\ell P^\ell_j = D^i_j = \lambda_i \delta^i_j,$$

et donc :

$$g(\cdot,\cdot) = \sum_{i,j=1}^n \lambda_i b^i \otimes b^i = \sum_{i=1}^{s_+} \varphi^i \otimes \varphi^i - \sum_{i=s_++1}^{s_++s_-} \varphi^i \otimes \varphi^i, \tag{8.23}$$

où on a posé $\varphi^i = \sqrt{\lambda_i}\, a^i$ si $\lambda_i > 0$, $\varphi^i = -\sqrt{|\lambda_i|}\, a^i$ si $\lambda_i < 0$ et $\varphi^i = a^i$ quand $\lambda_i = 0$.

Soit $(\vec{c}_i)_{i=1,...,n}$ une autre base de E. La matrice $[g(\vec{c}_i, \vec{c}_j)]$ est une matrice $n*n$ symétrique réelle. De même on obtient

$$g(\cdot,\cdot) = \sum_{i,j=1}^n \mu_i c^i \otimes c^i = \sum_{i=1}^{t_+} \psi^i \otimes \psi^i - \sum_{i=t_++1}^{t_++t_-} \psi^i \otimes \psi^i, \tag{8.24}$$

Il s'agit de montrer que $s_+ = t_+$ et $s_- = t_-$.

Montrons que la famille $(\varphi_1, ..., \varphi_{s_+}, \psi_{t_++1}, ..., \psi_n)$ est une famille libre (donc $s_+ + n - t_+ \le n$, et donc $s_+ \le t_+$). Soit une combinaison linéaire nulle $\alpha_1\varphi_1 + ... + \alpha_{s_+}\varphi_{s_+} + \beta_{t_++1}\psi_{t_++1} + ... + \beta_n\psi_n = 0$, i.e. les α_i et β_i vérifient :

$$\sum_{i=1}^{s_+} \alpha_i \varphi_i = -\sum_{i=t_++1}^{n} \beta_i \psi_i. \tag{8.25}$$

Donc :

$$g(\sum_{i=1}^{s_+} \alpha_i \varphi_i, \sum_{i=1}^{s_+} \alpha_i \varphi_i) = g(-\sum_{i=t_++1}^{n} \beta_i \psi_i, -\sum_{i=t_++1}^{n} \beta_i \psi_i),$$

soit :

$$\sum_{i,j=1}^{s_+} \alpha_i \alpha_j g(\varphi_i, \varphi_j) = \sum_{i,j=t_++1}^{n} \beta_i \beta_j g(\psi_i, \psi_j).$$

Et pour $i,j \in [1, s_+]_{\mathbb{N}}$ on a :

$$g(\varphi_i, \varphi_j) = \sum_{k=1}^{s_+} \varphi^k(\varphi_i)\varphi^k(\varphi_j) + 0 = \sum_{k=1}^{s_+} \delta^k_i \delta^k_j = \delta^i_j.$$

D'où $\sum_{i,j=1}^{s_+} \alpha_i \alpha_j b(\varphi_i, \varphi_j) = \sum_{i=1}^{s_+} \alpha_i^2$. De même $\sum_{i,j=t_++1}^{n} \beta_i \beta_j b(\psi_i, \psi_j) = -\sum_{i=t_++1}^{t_++t_-} \beta_i^2 + 0$.

D'où $\sum_{i=1}^{s_+} \alpha_i^2 = -\sum_{i=t_-+1}^{t_++t_-} \beta_i^2$. Donc tous les carrés sont nuls, donc tous les réels sont nuls : la famille $(\varphi_1, ..., \varphi_{s_+}, \psi_{t_++1}, ..., \psi_n)$ est libre. Donc $s_+ \le t_+$.

De même $t_+ \le s_+$, donc $s_+ = t_+$.

De même $s_- = t_-$. ∎

8.9 * Signature d'une métrique de Lorentz

La métrique lorentzienne a pour signature $(+,-,-,-)$ (signature $(1,3)$), cf. (1.22).

9 Bidual E^{**}

9.1 Espace bidual E^{**} et l'isomorphisme canonique $\mathcal{J}:E\to E^{**}$

Pour E e.v., le dual $E^*=\mathcal{L}(E;\mathbb{R})$ est un e.v. ; et donc $(E^*)^*=\mathcal{L}(E^*;\mathbb{R})$ est un e.v. (dual de E^*). On pose :

$$E^{**}\overset{\text{déf}}{=}(E^*)^*=\mathcal{L}(E^*;\mathbb{R}). \tag{9.1}$$

(E^{**} est l'ensemble des dérivées directionnelles, cf. (1.43).)

Proposition 9.1 *La fonction $\mathcal{J}\in\mathcal{F}(E;E^{**})$ définie par :*

$$\mathcal{J}:\begin{cases}E&\to E^{**},\\ \vec{v}&\mapsto \mathcal{J}(\vec{v}),\end{cases}\quad\text{telle que}\quad \mathcal{J}(\vec{v})(\ell)=\ell(\vec{v}),\quad \forall\ell\in E^*,\quad\forall\vec{v}\in E, \tag{9.2}$$

est un isomorphisme qui est canonique et "naturel", i.e. sa définition ne dépend pas du choix d'un observateur (choix d'une base ou du choix d'un produit scalaire par exemple). Il permet l'identification de $\vec{v}\in E$ et de $\mathcal{J}(\vec{v})\in E^{**}$, et on peut ainsi noter :

$$\mathcal{J}(\vec{v}).\ell\overset{\text{noté}}{=}\vec{v}.\ell\overset{\text{déf}}{=}\ell.\vec{v}. \tag{9.3}$$

($\mathcal{J}(\vec{v})=^{noté}\vec{v}$ est la dérivation de ℓ dans la direction $\vec{v}$, voir la suite.)

Preuve. La linéarité est immédiate, et comme $\dim(E^{**})=\dim(E^*)$ $(=\dim(E))$, il suffit de démontrer l'injectivité pour avoir la bijectivité. Soit donc $\vec{v}\in E$ t.q. $\mathcal{J}(\vec{v})=0$ i.e. t.q. $\ell(\vec{v})=0$ pour tout $\ell\in E^*$. Alors $\vec{v}=0$ (prendre $\ell=e^i$ une vecteur d'une base duale qui donne $v^i=0$, ce pour tout i).

Naturel : voir poly suivant. ▪

9.2 Base (∂_i) du bidual : opérateurs de dérivation dans les directions $\vec{e}_i$

Soit $(\vec{e}_j)_{j=1,\dots,n}$ une base de E et soit $(e^i)_{i=1,\dots,n}$ sa base duale (base de E^*).

Définition 9.2 La base biduale $(\partial_j)_{j=1,\dots,n}\in E^{**}$ de la base $(\vec{e}_i)$ est la base duale de la base duale $(e^i)_{i=1,\dots,n}$, donc définie par, pour tout $i,j=1,...,n$:

$$\partial_j.e^i=\delta^i_j,\quad\text{donc}\quad =e^i.\vec{e}_j\quad\text{et}\quad\partial_j=\mathcal{J}(\vec{e}_j). \tag{9.4}$$

Proposition 9.3 *Les ∂_j forment une base de E^{**} et vérifient, pour tout $j=1,...,n$, et on peut noter $\partial_j=\mathcal{J}(\vec{e}_j)=\vec{e}_j$.*

Preuve. On a $\dim(E^*)^*=\dim E^*=\dim E=n$, donc il suffit de montrer que (∂_j) est une famille libre. Supposons $\sum_j\alpha_j\delta_j=0$. Alors $\sum_j\alpha_j\delta_j.e^i=0=\sum_j\alpha_j\delta^i_j=\alpha_i$ pour tout i, donc la famille est libre. ▪

Donc si $\ell\in E^*$, $\ell=\sum_{i=1}^n\ell_ie^i$, alors $\partial_j.\ell=\ell_i$, cf. (9.4), avec $\ell(\vec{x})=\sum_{i=1}^n\ell_ix^i$ quand $\vec{x}=\sum_{i=1}^n x^i\vec{E}_i$, donc

$$\partial_j.\ell=d\ell(\vec{x}).\vec{e}_j\overset{\text{noté}}{=}\vec{e}_j(\ell), \tag{9.5}$$

cette dernière notation étant très utilisée en géométrie différentielle : $\partial_j\in E^{**}=\mathcal{L}(E^*;\mathbb{R})$ est l'opérateur de dérivation en $\vec{x}$ dans la direction $\vec{e}_j$.

Exemple 9.4 Pour $\vec{x}\in\Omega\subset\mathbb{R}^n$ et $f\in C^1(\Omega;\mathbb{R})$, on a $df(\vec{x})\in(\mathbb{R}^n)^*$, et :

$$df(\vec{x})=\sum_i(df(\vec{x}).\vec{E}_i)\,dx^i=\sum_i\frac{\partial f}{\partial x^i}(\vec{x})\,dx^i,\quad\text{donc}\quad\partial_j(df(\vec{x}))=\frac{\partial f}{\partial x^j}(\vec{x})\overset{\text{noté}}{=}(\vec{E}_j.f)(\vec{x}) \tag{9.6}$$

donne la i-ème dérivée partielle de f en $\vec{x}$. ▪

10 Tenseur uniforme (première approche des tenseurs)

(Pour les tenseurs non uniformes, voir poly suivant.) Soit E un espace vectoriel et $r, s \in \mathbb{N}$, avec l'un des deux entiers non nuls.

Définition 10.1 Un tenseur uniforme de type (r, s) est un élément de :

$$L^r_s(E) = \mathcal{L}((E^*)^r, E^s; \mathbb{R}) = L(\underbrace{E^*, ..., E^*}_{\text{r times}}, \underbrace{E, ..., E}_{\text{s times}}; \mathbb{R}) \overset{\text{noté}}{=} T^r_s(E) \tag{10.1}$$

l'ensemble des formes multilinéaires sur les espaces E^* et E. Donc $T \in L^r_s(E)$ agit donc sur les $(r+s)$-uplets $(\ell_1, ...\ell_r, \vec{v}_1, ..., \vec{v}_s) \in (E^*)^r \times E^s$ (dans cet ordre) pour former le réel $T(\ell_1, ...\ell_r, \vec{v}_1, ..., \vec{v}_s) \in \mathbb{R}$, et T est linéaire par rapport à chaque ℓ_i et à chaque $\vec{v}_i$.

Cas $r = s = 0$: on pose $L^0_0(E) = \mathbb{R}$ (ou plus généralement $= K$ le corps pour lequel E est un espace vectoriel.)

Si $r = 0$ alors $T \in L^0_s$ est covariant (n'agit que sur des vecteurs), si $s = 0$ alors $T \in L^r_0$ est contravariant (n'agit que sur des formes), et $T \in L^r_s$ est dit r fois contravariant et s fois covariant (tenseur mixte).

Exemple 10.2 $L^0_1(E) = \mathcal{L}(E; \mathbb{R}) = E^*$ est l'ensemble des formes linéaires sur E (les formes covariantes, celles qui varient avec les vecteurs).

Dans une base $(\vec{e}_i)$ de E de base duale (e^i), pour $T =^{\text{noté}} \ell \in L^0_1(E)$, $\ell = \sum_i \ell_i e^i$, on a $[\ell]_{|\vec{e}} = (\,\ell_1 \;\; \cdots \;\; \ell_n\,)$ où les index i des ℓ_i sont "en bas", d'où la notation L^0_1 avec 1 en bas. Et ℓ est covariant. ▪

Exemple 10.3 $L^1_0(E) = \mathcal{L}(E^*; \mathbb{R}) = E^{**}$ est l'ensemble des formes linéaires sur E^*. Comme $E^{**} \simeq E$ grâce à l'isomorphisme $\mathcal{J}$, cf. (9.2),

$$L^1_0(E) \simeq E \quad \text{(identification).} \tag{10.2}$$

Et pour $T \in L^1_0(E)$, et avec $\vec{v} = \mathcal{J}^{-1}(T) \in E$, cf. (9.2), on note $T(\ell) = T.\ell = \vec{v}.\ell$ $(= \ell.\vec{v})$, et $\vec{v}$ est contravariant.

Dans une base $(\vec{e}_i)$ de E de base duale (e^i) et de base biduale (∂_i), $T = \sum_i v^i \partial_i \in L^1_0(E) \simeq E$ avec $T =^{\text{noté}} \vec{v}$, et $[\vec{v}]_{|\vec{e}} = [v^i]_{i=1,\dots,n}$ où les index i des v^i sont en haut, d'où la notation L^1_0 avec 1 en haut. ▪

Exemple 10.4 $L^0_2(E) = \mathcal{L}(E, E; \mathbb{R}) = L^2(E; \mathbb{R})$ est l'ensemble des formes bilinéaires sur E. En particulier un produit scalaire est un élément de $T^0_2(E)$.

L'indice $_2$ en bas dans $L^0_2(E)$ correspond à la position des indices "en bas" des composantes g_{ij} d'un $T =^{\text{noté}} g(\cdot, \cdot) \in T^0_2(E)$: après introduction d'une base (e^i) de E^*, on a $g = \sum_{ij} g_{ij} e^i \otimes e^j$ où donc $g(\vec{e}_i, \vec{e}_j) = g_{ij}$. Ces formes bilinéaires (par exemple un produit scalaire) sont covariantes : elles agissent sur des vecteurs $\vec{u}, \vec{v} \in E$ pour former le réel $g(\vec{u}, \vec{v})$. ▪

Exemple 10.5 $L^1_1(E) = \mathcal{L}(E^*, E; \mathbb{R})$ est identifiable (canoniquement et naturellement, voir plus poly suivant) à l'ensemble $\mathcal{L}(E; E) = \mathcal{L}(E)$ des endomorphismes de E. Pour $T \in L^1_1(E)$, l'endomorphisme associé canoniquement est $L_T \in \mathcal{L}(E)$ donné par, pour tout $(\ell, \vec{v}) \in E^* \times E$,

$$\ell_T(L.\vec{v}) \overset{\text{déf}}{=} T(\ell, \vec{v}) \tag{10.3}$$

Après introduction d'une base $(\vec{e}_i)$ de E et de sa base duale (e^i) de E^*, on a $T \in L^1_1(E)$ de la forme $T = \sum_{ij} T^i_j \vec{e}_i \otimes e^j$ de même que l'endomorphisme L_T a été caractérisé par $L_T = \sum_{ij} T^i_j \vec{e}_i \otimes e^j$ relativement à la base $(\vec{e}_i)$: défini par $L_T.\vec{e}_j = \sum_{ij} T^i_j \vec{e}_i$, cf. (7.6).

Ces formes bilinéaires sont de type mixte : elles agissent sur les couples $(\ell, \vec{v})$ qui sont des couples "mixtes" (pas dans le même espace, l'un dans E^* et l'autre dans E). ▪

Remarque 10.6 Il n'y a pas de notation particulière pour applications bilinéaires $b(\cdot, \cdot) \in \mathcal{L}(E, E^*; \mathbb{R})$ canoniquement isomorphe à $\mathcal{L}(E^*, E; \mathbb{R}) = L^1_1(E)$ donc à $\mathcal{L}(E; E)$ l'ensemble des endomorphismes de E. Si on veut néanmoins se ramener à $L^1_1(E)$, on se sert de $b^T(\cdot, \cdot)$ le transposé de la forme bilinéaire $b(\cdot, \cdot)$ qui est défini par $b^T(\ell, \vec{u}) =^{\text{déf}} b(\vec{u}, \ell)$, avec donc $b^T(\cdot, \cdot) \in L^1_1(E) = \mathcal{L}(E^*, E; \mathbb{R})$. ▪

11 Vecteur de représentation d'une forme linéaire

La représentation d'une forme linéaire ℓ (appelé vecteur covariant) par un "vecteur de représentation" à l'aide du théorème de Riesz est non intrinsèque (à ℓ) : Deux observateurs différents obtiennent deux vecteurs différents.

11.1 Théorème de représentation de Riesz et vecteur de représentation

11.1.1 Le théorème de représentation de Riesz

Théorème 11.1 *(Théorème de représentation de Riesz) Soit $g(\cdot,\cdot) = (\cdot,\cdot)_g$ un produit scalaire sur un espace vectoriel E, et soit $||.||_g$ la norme associée. Si $\ell \in E^*$ (forme linéaire) est continue alors ℓ peut être représentée par un unique vecteur $\vec{\ell}^\sharp_g \in E$ relativement au produit scalaire $g(\cdot,\cdot)$:*

$$\exists!\, \vec{\ell}^\sharp_g \in E, \quad \ell(\vec{v}) = (\vec{\ell}^\sharp_g, \vec{v})_g, \qquad \forall \vec{v} \in \mathbb{R}^n. \tag{11.1}$$

De plus $||\vec{\ell}^\sharp_g||_g = ||\ell||$. (N.B. : Le vecteur $\vec{\ell}^\sharp_g$ dépend de la personne qui a imposé $g(\cdot,\cdot)$.)

Et donc dans une base $(\vec{e}_i)$ de base duale (e^i), si $\ell = \sum_i \ell_i e^i$ et si $\vec{\ell}^\sharp_g = \sum_i \ell^i_g \vec{e}_i$, et si $[g]_{|\vec{e}} = [g(\vec{e}_i, \vec{e}_j)]$, alors,

$$[\vec{\ell}^\sharp_g]_{|\vec{e}} = [g]^{-1}_{|\vec{e}} . [\ell]^T_{|\vec{e}} \tag{11.2}$$

(on rappelle que $[\ell]_{|\vec{e}}$ est une matrice ligne : matrice d'une forme linéaire), soit, pour tout i :

$$\ell^i_g = \sum_j g^{ij} \ell_j, \tag{11.3}$$

où on a noté :

$$[g^{ij}] \overset{\text{déf}}{=} [g]^{-1}. \tag{11.4}$$

En particulier pour la base duale (e^i), les vecteurs $(\cdot,\cdot)_g$-associés par Riesz sont les $\vec{e^i}^\sharp_g$ donnés par :

$$\vec{e^i}^\sharp_g = \sum_{i=1}^n g^{ij} \vec{e}_j, \tag{11.5}$$

et $(\vec{e^i}^\sharp_g)_{i=1,\dots,n}$ est une base de E appelée la base vectorielle $g(\cdot,\cdot)$-duale (elle vérifie $g(\vec{e}_i, \vec{e^j}^\sharp) = \delta^j_i$).

Preuve. Exercice classique, ou voir poly suivant. ∎

Remarque 11.2 La notation "dièse" dans $\vec{\ell}^\sharp_g$ fait référence à la "montée des indices" : si $\ell \in E^*$ est donnée par $\ell = \sum_i \ell_i e^i = [\ell_i]_{|(e)}$, alors $\vec{\ell}^\sharp_g = \sum_i \ell^i \vec{e}_i = [\ell^i]_{|(\vec{e})}$ donné par (11.3). ∎

Définition 11.3 Le vecteur $\vec{\ell}^\sharp_g$ est appelé vecteur de représentation de Riesz de ℓ pour le produit scalaire $g(\cdot,\cdot)$, ou $(\cdot,\cdot)_g$-vecteur de représentation.

On a ainsi défini l'application linéaire $\sharp_g$ qui dépend de $g(\cdot,\cdot)$:

$$\sharp_g : \left\{ \begin{array}{ll} E^* & \to E \\ \ell & \mapsto \sharp_g(\ell) = \vec{\ell}^\sharp_g \end{array} \right\}, \quad \text{avec} \quad ||\vec{\ell}^\sharp_g||_g = ||\ell||, \tag{11.6}$$

où $||\ell|| = \sup_{\vec{v}\in E} \frac{|\ell(\vec{v})|}{||\vec{v}||_g}$ est la norme de la forme linéaire ℓ (bien définie car ℓ est supposée continue : toujours vrai en dimension finie).

Proposition 11.4 *L'application $\sharp_g$ est linéaire : $\sharp_g \in \mathcal{L}(E^*, E)$.*

Preuve. $g((\ell+m)^\sharp_g, \vec{v}) = (\ell+m)(\vec{v}) = \ell(\vec{v}) + m(\vec{v}) = g(\vec{\ell}^\sharp_g, \vec{v}) + g(\vec{m}^\sharp_g, \vec{v}) = g(\vec{\ell}^\sharp_g + \vec{m}^\sharp_g, \vec{v})$ pour tout $\vec{v}$ donne bien $\sharp_g(\ell+m) = \sharp_g(\ell) + \sharp_g(m)$. Et de même $g((\lambda\ell)^\sharp_g, \vec{v}) = \lambda g((\ell)^\sharp_g, \vec{v})$ pour tout $\vec{v}$. ∎

Remarque 11.5 L'utilisation d'un produit scalaire est incontournable en mécanique (notion d'angle). Mais son usage systématique pour représenter une forme linéaire à l'aide d'un vecteur peut créer de nombreuses confusions, aussi bien dans les calculs que dans les interprétations. Par exemple :

- A force de tout représenter sous forme de vecteurs, on a tendance à oublier ce qu'est une fonction, et ce qu'est un vecteur. De manière générale une forme linéaire est "un instrument de mesure" servant à évaluer des "vecteurs".

- Suite : mélange du caractère covariant (fonctionnel) et du caractère contravariant (vectoriel). Visualisation : les formes et les vecteurs ne suivent pas les mêmes règles de changement de base.

- Le choix d'un produit scalaire dépend de l'utilisateur : si on se sert du produit scalaire euclidien pour estimer des longueurs, c'est évident (mètres, pieds ?). Son choix n'est pas intrinsèque.

- Voir § 1.8. ▪

11.1.2 Réciproque

La réciproque du théorème de Riesz est immédiate. Elle permet entre autre d'introduire la notation "bémol".

Théorème 11.6 *Soit $g(\cdot,\cdot) = (\cdot,\cdot)_g$ un produit scalaire sur un espace vectoriel E. Si $\vec{v} \in E$ alors l'application $v_g^\flat$ définie par :*

$$\forall \vec{w} \in E, \quad v_g^\flat(\vec{w}) = g(\vec{v}, \vec{w}) \tag{11.7}$$

est linéaire et continue sur $(E, ||.||_g)$.

Et donc dans une base $(\vec{e}_i)$ de base duale (e^i), si $\vec{v} = \sum_i v^i \vec{e}_i$, si $v_g^\flat = \sum_i v_i e^i$, alors :

$$[v_g^\flat]_{|\vec{e}} = [\vec{v}]_{|\vec{e}}^T.[g]_{|\vec{e}}, \quad i.e. \quad v_j = \sum_i v^i g_{ij} \quad \forall j. \tag{11.8}$$

Preuve. Comme $g(\cdot,\cdot)$ est bilinéaire, la linéarité de $v^\flat$ est immédiate. Puis $|v^\flat(\vec{w})| \leq ||\vec{v}||_g ||\vec{w}||_g$, donc $v^\flat$ est continue dans $(E, ||.||_g)$. Puis (11.7) donne (11.8). ▪

Remarque 11.7 La notation "bémol" dans $v_g^\flat$ fait référence à : "on descend les index" : si $[\vec{v}]_{|\vec{e}} = [v^i]$ (up), alors on a $[\vec{v}_g^\flat]_{|\vec{e}} = [v_i]$ (down), cf. (11.8). ▪

11.1.3 * Cas du pseudo-produit scalaire de Lorentz

Le théorème de représentation de Riesz subsiste :

Théorème 11.8 *On munit $\mathbb{R}^n$ d'un pseudo-produit scalaire $\eta(\cdot,\cdot)$ de Lorentz défini sur la base canonique $(\vec{e}_i)$:*

$$\eta(\vec{e}_i, \vec{e}_j) = \eta_{ij}, \quad \text{où} \quad [\eta]_{|\vec{e}} = [\eta_{ij}] \text{ est une matrice symétrique inversible.}$$

(La signature de η est de type $(s, n-s)$.) Alors si $\vec{v} \in \mathbb{R}^n$, l'application $v_\eta^\flat$ définie par :

$$\forall \vec{w} \in \mathbb{R}^n, \quad v_\eta^\flat(\vec{w}) = \eta(\vec{v}, \vec{w}) \tag{11.9}$$

est linéaire ($v_\eta^\flat \in \mathbb{R}^{n*}$). Et $[v_\eta^\flat]_{|\vec{e}} = [\vec{v}]_{|\vec{e}}^T.[\eta]_{|\vec{e}}$.

Et réciproquement, si $v^\flat \in (\mathbb{R}^n)^*$, alors il existe un unique vecteur $\vec{v}_\eta^\sharp \in \mathbb{R}^n$ vérifiant

$$\forall \vec{w} \in \mathbb{R}^n, \quad v^\flat(\vec{w}) = \eta(\vec{v}_\eta^\sharp, \vec{w}). \tag{11.10}$$

Et $[v^\flat]_{|\vec{e}} = [\vec{v}_\eta^\sharp]_{|\vec{e}}^T.[\eta]_{|\vec{e}}$.

Preuve. Si $\vec{v} \in \mathbb{R}^n$, alors $v^\flat(\vec{w}) = \eta(\vec{v}, \vec{w})$ donne $[\vec{v}]_{|\vec{e}}^T.[\eta]_{|\vec{e}}$. ▪

Exemple 11.9 Dans $\mathbb{R}^4$ et la métrique usuelle de Minkowski donnée dans la base canonique $(\vec{E}_i)$ par $[\eta] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$, si $v^\flat = \sum_{j=0}^{3} v_j e^j \overset{\text{noté}}{=} (\, v_0 \quad v_1 \quad v_2 \quad v_3 \,)$, si $\vec{v} = \sum_{j=0}^{3} v^j \vec{e}_j \overset{\text{noté}}{=} \begin{pmatrix} v^0 \\ v^1 \\ v^2 \\ v^3 \end{pmatrix}$ alors $\vec{v} = \vec{v}_\eta^\sharp$ ssi $v^0 = v_0$, $v^1 = -v_1$, $v^2 = -v_2$ et $v^3 = -v_3$. ▪

11.2 Le vecteur de représentation est un vecteur (est contravariant)

On reprend les bases précédentes $(\vec{e}_{i,old})$ et $(\vec{e}_{i,new})$, où donc $\vec{e}_{j,new} = \sum_i P^i_j \vec{e}_{i,old}$, avec un produit scalaire $g(\cdot,\cdot)$ fixé.

Proposition 11.10 *Le vecteur covariant $\ell \in E^*$ (la forme linéaire) a son vecteur de représentation associé $\vec{\ell}^\sharp_g$ qui est contravariant au sens des règles de changement de base :*

$$\text{si} \quad [\ell]_{new} = [\ell]_{old}.P \quad \text{alors} \quad [\vec{\ell}^\sharp_g]_{new} = P^{-1}.[\vec{\ell}^\sharp_g]_{old}. \tag{11.11}$$

Preuve. Avec (11.2) on a :

$$[\vec{\ell}^\sharp_g]_{new} = [g]^{-1}_{new}.[\ell]^T_{new}, \qquad [\vec{\ell}^\sharp_g]_{old} = [g]^{-1}_{old}.[\ell]^T_{old}.$$

Formules de changement de base :

$$[\ell]_{new} = [\ell]_{old}.P \qquad \text{et} \qquad [g]_{new} = P^T.[g]_{old}.P.$$

Donc :

$$[\vec{\ell}^\sharp_g]_{new} = (P^T.[g]_{old}.P)^{-1}.([\ell]_{old}.P)^T = P^{-1}.[g]^{-1}_{old}.P^{-T}.P^T.[\ell]^T_{old} = P^{-1}.[g]^{-1}_{old}.[\ell]^T_{old} = P^{-1}.[\vec{\ell}^\sharp_g]_{old}.$$

▪

Exemple 11.11 Dans $\mathbb{R}^2$ muni de son produit scalaire euclidien. Soit $(\vec{E}_i) = (\vec{e}_{i,old})$ la base canonique et (dx^i) sa base duale.

Soit $\ell : \mathbb{R}^2 \to \mathbb{R}$ donnée par $\ell = \ell_{1,old}dx^1 + \ell_{2,old}dx^2 \overset{\text{noté}}{=} (\,\ell_{1,old} \quad \ell_{2,old}\,) = [\ell]_{old}$.

Soit $(\vec{e}_i) = (\vec{e}_{i,new})$ la base donnée par la matrice de passage $P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, i.e. $\vec{e}_1 = \vec{E}_1$ et $\vec{e}_2 = \vec{E}_1 + \vec{E}_2$ (les composantes sont données par les colonnes de P).

Sa base duale (e^1, e^2) est donnée par les lignes de $P^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ (car $P^{-1}.P = I = [e^i.\vec{e}_j] = [\delta^i_j]$), soit donc $e^1 = dx - dy$ et $e^2 = dy$.

Soit ℓ donné par $\ell = \ell_{1,new}e^1 + \ell_{2,new}e^2 \overset{\text{noté}}{=} (\,\ell_{1,new} \quad \ell_{2,new}\,) = [\ell]_{new}$ dans la nouvelle base.

Formule de changement de base : $(\,\ell_{1,new} \quad \ell_{2,new}\,) = (\,\ell_{1,old} \quad \ell_{2,old}\,).P$.

Donc $\ell_{1,new} = \ell_{1,old}$ et $\ell_{2,new} = \ell_{1,old} + \ell_{2,old}$, ou encore $\ell_{1,old} = \ell_{1,new}$ et $\ell_{2,old} = \ell_{2,new} - \ell_{1,new}$.

Soit $\vec{\ell}^\sharp_g$ le vecteur représentant ℓ relativement au produit scalaire canonique. Dans l'ancienne base $\vec{\ell}^\sharp_g = \ell^1_{old}\vec{E}_1 + \ell^2_{old}\vec{E}_2 \overset{\text{noté}}{=} \begin{pmatrix} \ell^1_{old} \\ \ell^2_{old} \end{pmatrix}$ et $\ell(\vec{v}) = (\vec{\ell}^\sharp_g, \vec{v})_{\mathbb{R}^2}$ donne immédiatement : $\ell^1_{old} = \ell_{1,old}$ et $\ell^2_{old} = \ell_{2,old}$.

Dans la nouvelle base $\vec{\ell}^\sharp_g = \ell^1_{new}\vec{e}_1 + \ell^2_{new}\vec{e}_2 \overset{\text{noté}}{=} \begin{pmatrix} \ell^1_{new} \\ \ell^2_{new} \end{pmatrix}$, et $\ell(\vec{v}) = (\vec{\ell}^\sharp_g, \vec{v})_{\mathbb{R}^2}$ soit $[\ell]_{new}.[\vec{v}]_{new} = [\vec{\ell}^\sharp_g]^T_{new}.\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.[\vec{v}]_{new}$, car la matrice du produit scalaire dans la base nouvelle est $[g]_{new} = [(\vec{e}_i, \vec{e}_j)_{\mathbb{R}^2}] = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. Donc en particulier avec $\vec{v} = \vec{e}_1$ puis avec $\vec{v} = \vec{e}_2$ on obtient $\ell_{1,new} = (\,\ell^1_{new} \quad \ell^2_{new}\,).\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \ell^1_{new} + \ell^2_{new}$ et $\ell_{2,new} = (\,\ell^1_{new} \quad \ell^2_{new}\,).\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \ell^1_{new} + 2\ell^2_{new}$.

Et la formule de changement de base pour les vecteurs donne $\begin{pmatrix} \ell^1_{new} \\ \ell^2_{new} \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.\begin{pmatrix} \ell^1_{old} \\ \ell^2_{old} \end{pmatrix}$.

Donc $\ell^1_{new} = \ell^1_{old} - \ell^2_{old}$ et $\ell^2_{new} = \ell^2_{old}$, soit $\ell^1_{old} = \ell^1_{new} + \ell^2_{new}$ et $\ell^2_{old} = \ell^2_{new}$.

Vérifions : on a $\ell.\vec{v} = (\vec{\ell}^\sharp_g, \vec{v})_{\mathbb{R}^2}$ qui doit donner d'une part $= [\ell]_{old}.[\vec{v}]_{old} = [\vec{\ell}^\sharp_g]^T_{old}.[\vec{v}]_{old}$ et d'autre part $= [\ell]_{new}.[\vec{v}]_{new} = [\vec{\ell}^\sharp_g]^T_{new}.\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.[\vec{v}]_{new}$.

Donc avec $\vec{v} = \vec{e}_1$ on a d'une part $\ell.\vec{e}_1 = \ell_{1,old} = (\vec{\ell}^\sharp_g, \vec{e}_1)_{\mathbb{R}^2} = \ell^1_{old}$, et d'autre part $\ell.\vec{e}_1 = \ell_{1,new} = (\vec{\ell}^\sharp_g, \vec{e}_1)_{\mathbb{R}^2} = (\,\ell^1_{new} \quad \ell^2_{new}\,).\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \ell^1_{new} + \ell^2_{new}$. OK.

Et avec $\vec{v} = \vec{e}_2$ on a d'une part $\ell.\vec{e}_2 = \ell_{1,old} + \ell_{2,old} = (\vec{\ell}^\sharp_g, \vec{e}_2)_{\mathbb{R}^2} = \ell^1_{old} + \ell^2_{old}$, et d'autre part $\ell.\vec{e}_2 = \ell_{2,new} = (\vec{\ell}^\sharp_g, \vec{e}_2)_{\mathbb{R}^2} = (\,\ell^1_{new} \quad \ell^2_{new}\,).\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \ell^1_{new} + 2\ell^2_{new}$. OK. ▪

11.3 Exemple : un "vecteur gradient" est un vecteur de représentation

Pour simplifier la présentation on prend $E = \mathbb{R}^n$.

Soit $f \in C^1(\mathbb{R}^n;\mathbb{R})$. Sa différentielle en un point $\vec{x}$ est la forme linéaire définie par : pour tout $\vec{v} \in \mathbb{R}^n$, au voisinage de $h = 0$:

$$f(\vec{x}+h\vec{v}) = f(\vec{x}) + h\, df(\vec{x}).\vec{v} + o(h). \tag{11.12}$$

Soit un produit scalaire $g(\cdot,\cdot) = (\cdot,\cdot)_g$: on peut alors représenter $df(\vec{x})$ à l'aide de $\vec{\text{grad}}_g f(\vec{x})$ = le vecteur de $(\cdot,\cdot)_g$-représentation de Riesz de $df(\vec{x}) \in \mathbb{R}^{n*}$, appelé le gradient de f en $\vec{x}$: donc

$$df(\vec{x}).\vec{v} = (\vec{\text{grad}}_g f(\vec{x}), \vec{v})_g. \tag{11.13}$$

Cas $E = (\mathbb{R}^n, (\cdot,\cdot)_{\mathbb{R}^n})$ ($\mathbb{R}^n$ muni de son produit scalaire euclidien $g(\cdot,\cdot) = (\cdot,\cdot)_{\mathbb{R}^n}$). On note alors :

$$\vec{\text{grad}}_g = \vec{\text{grad}}. \tag{11.14}$$

Représentation dans une base euclidienne $(\vec{E}_i)$ donnée de $\mathbb{R}^n$ de base duale (dx^i) à l'aide d'un produit scalaire $(\cdot,\cdot)_g$. Si $\vec{x} = \sum_i x^i \vec{E}_i$, on a $df(\vec{x}) = \sum_j \frac{\partial f}{\partial x^j}(\vec{x})\, dx^j$, forme linéaire représentée par sa matrice ligne $[df(\vec{x})]_{|\vec{E}} = [\frac{\partial f}{\partial x^i}(\vec{x})]_{i=1,\dots n} = (\ \frac{\partial f}{\partial x^1}\ \ \dots\ \ \frac{\partial f}{\partial x^n}\ )$. Et (11.13) donne

$$[\vec{\text{grad}}_g f(\vec{x})] = [g]^{-1}.[df(\vec{x})]^T, \quad \text{i.e.} \quad (\vec{\text{grad}}_g f)^i = \sum_j g^{ij} \frac{\partial f}{\partial x^j} \quad \forall i. \tag{11.15}$$

En particulier, si $(\cdot,\cdot)_g$ est le produit scalaire associé à la base $(\vec{E}_i)$, alors $[\vec{\text{grad}} f(\vec{x})] = [df(\vec{x})]^T$ (car $[g^{ij}] = I$).

Exercice 11.12 Exprimer le gradient en coordonnées polaires (donc ici dans $(\mathbb{R}^n, (\cdot,\cdot)_{\mathbb{R}^n})$).

Réponse. On donne ici une démarche générique permettant de trouver l'expression du gradient dans tout système de coordonnées.

0- Le cadre initial est $\mathbb{R}^2$ muni d'un produit scalaire euclidien. On note $(\vec{E}_1, \vec{E}_2)$ une base euclidienne.

1- Soit

$$\vec{\varphi} : \begin{cases} \mathbb{R}_+ \times \mathbb{R} & \to \mathbb{R}^2 \\ \vec{q} = (r,\theta) \overset{\text{noté}}{=} \begin{pmatrix} r \\ \theta \end{pmatrix} & \to \vec{\varphi}(\vec{q}) = \vec{x} = r\cos\theta \vec{E}_1 + r\sin\theta \vec{E}_2, \quad \text{i.e.} \quad [\vec{x}]_{|\vec{E}} = \begin{pmatrix} x = r\cos\theta \\ y = r\sin\theta \end{pmatrix}, \end{cases}$$

le système de coordonnées polaires où la matrice colonne exprime les coordonnées dans la base canonique dans l'espace d'arrivée. Le vecteur $\vec{q}$ est le vecteur des coordonnées paramétriques et le vecteur $\vec{x}$ est le vecteur des coordonnées géométriques. (Pour avoir un difféomorphisme, on restreint par exemple l'espace de départ à $\mathbb{R}_+^* \times]-\pi,\pi[$ et l'espace d'arrivée à $\mathbb{R}^2 - \{(x \le 0, y = 0)\}$.)

Conventions :

Espace de départ des $\vec{q} = (r,\theta)$, base canonique $(\vec{E}_1, \vec{E}_2)$.

Espace d'arrivée des $\vec{x} = (x,y)$, une base euclidienne $(\vec{E}_1, \vec{E}_2)$, base duale (dx, dy) (d'après le nom des variables).

Si $f : \vec{q} \to \vec{x}$ est différentiable, on note $\frac{\partial f}{\partial q^i}(\vec{q}) \overset{\text{déf}}{=} df(\vec{q}).\vec{E}_i$ ($= \lim_{h\to 0} \frac{f(\vec{q}+h\vec{E}_i)-f(\vec{q})}{h}$), où donc ici $(\vec{E}_i)$ est la base canonique dans l'ensemble de départ (des paramètres).

2- La base des coordonnées polaires en un point $\vec{x} = \vec{\varphi}(\vec{q})$ et formée des vecteurs

$$\begin{cases} \vec{e}_1(\vec{x}) = d\vec{\varphi}(\vec{q}).\vec{E}_1 \overset{\text{noté}}{=} \dfrac{\partial \vec{\varphi}}{\partial r}(\vec{q}) \overset{\text{noté}}{=} \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \\ \vec{e}_2(\vec{x}) = d\vec{\varphi}(\vec{q}).\vec{E}_2 \overset{\text{noté}}{=} \dfrac{\partial \vec{\varphi}}{\partial \theta}(\vec{q}) \overset{\text{noté}}{=} \begin{pmatrix} -r\sin\theta \\ r\cos\theta \end{pmatrix}, \end{cases}$$

où les matrices colonnes exprime les coordonnées dans la base euclidienne $(\vec{E}_i)$ de l'espace d'arrivée.

Et $P = [d\vec{\varphi}(\vec{q})] = (\,(\frac{\partial \vec{\varphi}}{\partial r})\ \ (\frac{\partial \vec{\varphi}}{\partial \theta})\,) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$ = matrice de passage : les composantes des vecteurs $\vec{e}_1(\vec{x})$ et $\vec{e}_2(\vec{x})$ (de base du système en $\vec{x}$) sont données par les colonnes de P (dans la base canonique d'arrivée).

3- La base duale $(e^1(\vec{x}), e^2(\vec{x})) =^{\text{noté}} (dr(\vec{x}), d\theta(\vec{x}))$ de la base $(\vec{e}_1(\vec{x}), \vec{e}_2(\vec{x}))$ du système en $\vec{x} = \vec{\varphi}(\vec{q})$ est donnée par les lignes de $P^{-1} = \frac{1}{r}\begin{pmatrix} r\cos\theta & r\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ (puisque $[e^i.\vec{e}_j] = I = [\delta^i_j]$). Donc on a :

$$\begin{cases} dr(\vec{x}) = e^1(\vec{x}) = \cos\theta\,dx + \sin\theta\,dy = (\,\cos\theta \quad \sin\theta\,)\,, \\ d\theta(\vec{x}) = e^2(\vec{x}) = -\dfrac{1}{r}\sin\theta\,dx + \dfrac{1}{r}\cos\theta\,dy = (\,-\frac{1}{r}\sin\theta \quad \frac{1}{r}\cos\theta\,)\,, \end{cases}$$

expressions dans la base duale (dx, dy) de la base canonique (ici on est dans l'espace d'arrivée).

Noter que $dr(\vec{x})$ et $d\theta(\vec{x})$ sont bien les expressions des composantes de $d\vec{q}(\vec{x}) = d(\vec{\varphi}^{-1})(\vec{x}) = \begin{pmatrix} dr(\vec{x}) \\ dq(\vec{x}) \end{pmatrix}$. En effet $r = (x^2+y^2)^{\frac{1}{2}}$ et $\theta = \tan^{-1}(\frac{y}{x})$ pour les $\theta \in]-\frac{\pi}{2}, \frac{\pi}{2}[$, d'où $\frac{\partial r}{\partial x}(\vec{x}) = \frac{x}{(x^2+y^2)^{\frac{1}{2}}} = \frac{x}{r} = \cos\theta$ et $\frac{\partial r}{\partial y}(\vec{x}) = \frac{y}{r} = \sin\theta$ qui redonnent $dr(\vec{x}) = \frac{\partial r}{\partial x}(\vec{x})dx + \frac{\partial r}{\partial y}(\vec{x})dy$, et $\frac{\partial\theta}{\partial x}(\vec{x}) = -\frac{y}{x^2(1+(\frac{y}{x})^2)} = -\frac{y}{r^2} = -\frac{\sin\theta}{r}$, et $\frac{\partial\theta}{\partial y}(\vec{x}) = \frac{1}{x(1+(\frac{y}{x})^2)} = \frac{x}{r^2} = \frac{\cos\theta}{r}$ qui redonnent $d\theta(\vec{x}) = \frac{\partial\theta}{\partial x}(\vec{x})dx + \frac{\partial\theta}{\partial y}(\vec{x})dy$.

4- Soit une fonction $\beta : \vec{x} \to \beta(\vec{x})$ définie dans l'espace géométrique. De manière générique, étant donné une base $(\vec{e}_i(\vec{x}))$ de base duale associée $(e^i(\vec{x}))$, sa différentielle s'exprime comme :

$$d\beta(\vec{x}) = (d\beta(\vec{x}).\vec{e}_1(\vec{x}))e^1(\vec{x}) + (d\beta(\vec{x}).\vec{e}_2(\vec{x}))e^2(\vec{x}),$$

vérification triviale en appliquant $\vec{e}_i(\vec{x})$ aux deux termes et par définition de la base duale. Et ici :

$$d\beta(\vec{x}).\vec{e}_1(\vec{x}) = \frac{\partial(\beta\circ\vec{\varphi})}{\partial r}(\vec{x}) \overset{\text{noté}}{=} \frac{\partial\beta}{\partial r}(\vec{x}), \qquad d\beta(\vec{x}).\vec{e}_2(\vec{x}) = \frac{\partial(\beta\circ\vec{\varphi})}{\partial\theta}(\vec{x}) \overset{\text{noté}}{=} \frac{\partial\beta}{\partial\theta}(\vec{x})$$

sont les dérivées dans les directions des vecteurs de base $\vec{e}_1(\vec{x})$ et $\vec{e}_2(\vec{x})$ du système de coordonnées en $\vec{x}$ (pas dans les directions de vecteurs de base euclidienne de l'espace géométrique). Donc :

$$d\beta(\vec{x}) = \frac{\partial\beta}{\partial r}(\vec{x})\,dr(\vec{x}) + \frac{\partial\beta}{\partial\theta}(\vec{x})\,d\theta(\vec{x}).$$

N.B. : c'est bien une définition de notations car β est une fonction qui dépend de $\vec{x}$ et non de $\vec{q}$. Cette notation provient de la fonction f définie par $f = \beta\circ\vec{\varphi}$, i.e. par :

$$f(\vec{q}) \overset{\text{déf}}{=} \beta(\vec{x}), \quad \text{quand} \quad \vec{x} = \vec{\varphi}(\vec{q}).$$

Et on a :

$$df(\vec{q}) = \frac{\partial f}{\partial r}(\vec{q})dr + \frac{\partial f}{\partial\theta}(\vec{q})d\theta.$$

Et par dérivation de fonctions composées on a $df(\vec{q}) = d\beta(\vec{x}).d\vec{\varphi}(\vec{q})$ quand $\vec{x} = \vec{\varphi}(\vec{q})$. Ce qui appliqué aux vecteurs $\vec{E}_i$ donne :

$$\frac{\partial f}{\partial r}(\vec{q}) = d\beta(\vec{x}).\vec{e}_1(\vec{x}), \qquad \frac{\partial f}{\partial\theta}(\vec{q}) = d\beta(\vec{x}).\vec{e}_2(\vec{x}),$$

d'où les notations $\frac{\partial f}{\partial r}(\vec{q}) =^{\text{noté}} \frac{\partial\beta}{\partial r}(\vec{x})$ et $\frac{\partial f}{\partial\theta}(\vec{q}) =^{\text{noté}} \frac{\partial\beta}{\partial\theta}(\vec{x})$.

5- Par définition du gradient, avec le produit scalaire euclidien, on a (théorème de représentation de Riesz) :

$$(\vec{\text{grad}}\beta(\vec{x}), \vec{v})_{\mathbb{R}^2} = d\beta(\vec{x}).\vec{v},$$

pour tout vecteur $\vec{v} \in \mathbb{R}^2$. Notant $\begin{pmatrix} w^1 \\ w^2 \end{pmatrix}$ les composantes de $\vec{\text{grad}}\beta(\vec{x})$ dans la base du système polaire, c'est à dire $\vec{\text{grad}}\beta(\vec{x}) = w^1\vec{e}_1(\vec{x}) + w^2\vec{e}_2(\vec{x})$, puis prenant les $\vec{v} = \vec{e}_i(\vec{x})$, on obtient :

$$w^1||\vec{e}_1(\vec{x})||^2 = \frac{\partial\beta}{\partial r}(\vec{x}), \qquad w^2||\vec{e}_2(\vec{x})||^2 = \frac{\partial\beta}{\partial\theta}(\vec{x}).$$

Comme $||\vec{e}_1(\vec{x})||^2 = 1$ et $||\vec{e}_2(\vec{x})||^2 = r^2$ on obtient :

$$\vec{\text{grad}}\beta(\vec{x}) = \frac{\partial\beta}{\partial r}(\vec{x})\vec{e}_1(\vec{x}) + \frac{1}{r^2}\frac{\partial\beta}{\partial\theta}(\vec{x})\vec{e}_2(\vec{x}) = \begin{pmatrix} \frac{\partial\beta}{\partial r}(\vec{x}) \\ \frac{1}{r^2}\frac{\partial\beta}{\partial\theta}(\vec{x}) \end{pmatrix}_{|(\vec{e}_1(\vec{x}),\vec{e}_2(\vec{x}))}.$$

6- La base du système de coordonnées polaires n'est pas la base polaire choisie par les mécaniciens : ils lui préfèrent la base orthonormée :

$$\vec{e}_r(\vec{x}) = \vec{e}_1(\vec{x}) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \qquad \vec{e}_\theta(\vec{x}) = \frac{\vec{e}_2(\vec{x})}{r} = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}.$$

D'où l'expression usuelle du gradient en "coordonnées polaires" :

$$\vec{\text{grad}}\beta(\vec{x}) = \frac{\partial\beta}{\partial r}(\vec{x})\vec{e}_r(\vec{x}) + \frac{1}{r}\frac{\partial\beta}{\partial\theta}(\vec{x})\vec{e}_\theta(\vec{x}) = \begin{pmatrix} \frac{\partial\beta}{\partial r}(\vec{x}) \\ \frac{1}{r}\frac{\partial\beta}{\partial\theta}(\vec{x}) \end{pmatrix}_{(\vec{e}_r(\vec{x}),\vec{e}_\theta(\vec{x}))}. \tag{11.16}$$

7- Exemple : $\beta(\vec{x}) = r\tan\theta$, donc au sens $\beta(\vec{x}) = r(x,y)\tan(\theta(x,y)) = \sqrt{x^2+y^2}\,\dfrac{y}{x}$. On pose $\beta(\vec{x}) =$

$f(\vec{q}) = r\tan\theta$, et on obtient :

$$\vec{\text{grad}}\beta(\vec{x}) = \begin{pmatrix} \frac{\partial f}{\partial r}(\vec{q}) \\ \frac{1}{r}\frac{\partial f}{\partial \theta}(\vec{q}) \end{pmatrix} = \begin{pmatrix} \tan\theta \\ 1+\tan^2\theta \end{pmatrix}_{(\vec{e}_r,\vec{e}_\theta)},$$

quand $\vec{x} = \vec{\varphi}(\vec{q})$. ▪

11.3.1 Application à la méthode du gradient conjugué

La méthode du gradient conjuguée n'est autre que le méthode du gradient appliquée avec le produit scalaire adéquat. Traitons le cas $E = \mathbb{R}^2$ pour alléger les notations, $(\vec{E}_1, \vec{E}_2)$ une base cartésienne, et notons $\vec{x} = x\vec{E}_1 + y\vec{E}_2 = \begin{pmatrix} x \\ y \end{pmatrix}$ un point de $\mathbb{R}^2$ (considéré comme espace affine) et de manière générique $\vec{v} = v^1\vec{E}_1 + v^2\vec{E}_2 = \begin{pmatrix} v^1 \\ v^2 \end{pmatrix}$ un vecteur de $\mathbb{R}^2$ (considéré comme espace vectoriel associé).

On rappelle que si $\Phi : \mathbb{R}^n \to \mathbb{R}$ est C^1, on a le développement de Taylor :

$$\Phi(\vec{x}+h\vec{v}) = \Phi(\vec{x}) + h\, d\Phi(\vec{x}).\vec{v} + o(h). \tag{11.17}$$

Si on choisit le produit scalaire cartésien $g(\vec{v},\vec{w}) = (\vec{v},\vec{w}) = (\begin{pmatrix} v^1 \\ v^2 \end{pmatrix}, \begin{pmatrix} w^1 \\ w^2 \end{pmatrix}) = v^1w^1 + v^2w^2$, on définit le gradient usuel de Φ en $\vec{x}$ comme étant l'unique vecteur $\vec{\text{grad}}\Phi(\vec{x})$ vérifiant :

$$d\Phi(\vec{x}).\vec{v} = (\vec{\text{grad}}\Phi(\vec{x}),\vec{v}), \quad \forall \vec{v} \in \mathbb{R}^2.$$

L'existence et l'unicité sont donnés par le théorème de représentation de Riesz.

Si on prend un autre produit scalaire $g(\cdot,\cdot) = (\cdot,\cdot)_g$, notons $\vec{\text{grad}}_g\Phi$, toujours à l'aide du théorème de représentation de Riesz, l'unique vecteur tel que :

$$d\Phi(\vec{x}).\vec{v} = g(\vec{\text{grad}}_g\Phi(\vec{x}),\vec{v}), \quad \forall \vec{v} \in \mathbb{R}^2. \tag{11.18}$$

Les vecteurs $\vec{\text{grad}}_g\Phi$ et $\vec{\text{grad}}\Phi$ sont en général différents.

Application à la méthode du gradient conjugué.

On se place dans $\mathbb{R}^2$ pour simplifier la présentation, et on note $\vec{x} = (x,y)$. Il s'agit de minimiser la fonctionnelle :

$$\Phi(x,y) = ax^2 + by^2,$$

où $a, b \geq 0$ et $a \neq b$. Ici la réponse est immédiate : Φ est positive, et le minimum de Φ vaut 0, atteint pour $\vec{x} = \vec{0}$.

Si on applique la méthode du gradient, c'est qu'on considère le produit scalaire euclidien et la formule de Taylor (11.17) approximée :

$$\Phi(\vec{x}+h\vec{v}) \simeq \Phi(\vec{x}) + h\,(\vec{\text{grad}}\Phi(\vec{x}),\vec{v}),$$

et qu'on dit que $\vec{x}$ et h étant donnés, la direction $\vec{v}$ telle que $\Phi(\vec{x}+h\vec{v})$ soit le plus faible possible est donnée lorsque $(\vec{\text{grad}}\Phi(\vec{x}),\vec{v})$ est le plus négatif possible. C'est à dire lorsque $\vec{v} = -\alpha\vec{\text{grad}}\Phi(\vec{x})$ pour un $\alpha > 0$: on descend dans la direction de plus grande pente.

D'où on prend comme première direction de descente la direction donnée par $\vec{\text{grad}}\Phi(\vec{x})$: on calcule le point $\vec{x}_1 = \vec{x} - h\,\vec{\text{grad}}\Phi(\vec{x})$ où h réalise le minimum de :

$$\psi(h) = \Phi(\vec{x} - h\,\vec{\text{grad}}\Phi(\vec{x})), \tag{11.19}$$

puis on applique à nouveau la formule de Taylor au voisinage du point $\vec{x}_1$: on obtient un point $\vec{x}_2$, on recommence...

Le problème est que le gradient usuel de Φ au point $\vec{x}$ est $\vec{\text{grad}}\Phi(\vec{x}) = 2\begin{pmatrix} ax \\ by \end{pmatrix}$ et qu'il n'est pas dirigé selon le vecteur $\vec{x} - \vec{0} = \begin{pmatrix} x \\ y \end{pmatrix}$ qui, lui, pointe vers $\vec{0}$ ou se trouve le minimum. On note alors que :

$$\begin{pmatrix} ax \\ by \end{pmatrix} = A.\begin{pmatrix} x \\ y \end{pmatrix} \quad \text{où} \quad A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.$$

Et la matrice A est la matrice hessienne de Φ, i.e. $A = [\frac{\partial^2\Phi}{\partial x^i \partial x^j}] = [d^2\Phi(\vec{x})]$ (ici matrice constante car Φ est un polynôme de degré 2).

L'idée étant de trouver une meilleure direction de descente (une direction de descente adaptée à la fonction Φ), on choisit de prendre le produit scalaire donné par :

$$g(\vec{v},\vec{w}) = \vec{v}^{\,t}.A.\vec{w} = (\vec{v},\vec{w})_A \quad \text{où} \quad A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.$$

On a, avec (11.18), notant $\vec{\text{grad}}\Phi_g(\vec{x}) = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$:

$$\begin{aligned} g(\vec{\text{grad}}\Phi_g(\vec{x}),\vec{v}) &= (\alpha \quad \beta).\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.\begin{pmatrix} v^1 \\ v^2 \end{pmatrix} = a\alpha v^1 + b\beta v^2 \quad \text{avec} \\ &= d\Phi(\vec{x}).\vec{v} = (\vec{\text{grad}}\Phi,\vec{v}) = \frac{\partial\Phi}{\partial x}(\vec{x})v^1 + \frac{\partial\Phi}{\partial x}(\vec{x})v^2, \end{aligned}$$

pour tout v^1,v^2, d'où $\alpha = \frac{1}{a}\frac{\partial\Phi}{\partial x}(\vec{x})$ et $\beta = \frac{1}{b}\frac{\partial\Phi}{\partial y}(\vec{x})$, d'où $\vec{\text{grad}}_g\Phi(\vec{x}) = \begin{pmatrix} \frac{1}{a}\frac{\partial\Phi}{\partial x}(\vec{x}) \\ \frac{1}{b}\frac{\partial\Phi}{\partial y}(\vec{x}) \end{pmatrix} = 2\begin{pmatrix} x \\ y \end{pmatrix}$. Et cette direction pointe vers $\vec{0}$ comme désiré : on pose donc $\vec{x}_1 = \vec{x} + h\,\vec{\text{grad}}_g\Phi$. Et on a trouvé la "bonne" direction de descente : c'est celle associée au gradient conjugué $\vec{\text{grad}}_g\Phi$, i.e. celle associée au produit scalaire $g(\cdot,\cdot)$.

Remarque 11.13 On aurait pu trouver le résultat autrement, i.e. adopter la démarche : en $\vec{x}$ donnée, à h fixé (petit), on veut trouver la direction $\vec{v}$ telle que la différence $\Phi(\vec{x}+h\vec{v}) - \Phi(\vec{x})$ soit la plus grande possible, pour $\vec{v}$ vecteur normé.

Ici, Φ étant de degré 2, sa différentielle $d\Phi$ est linéaire en $\vec{x}$, et donc $d^2\Phi$ est constant en $\vec{x}$, et le développement limité de Φ est donné exactement par :

$$\Phi(\vec{x}+h\vec{v}) - \Phi(\vec{x}) = h\,d\Phi(\vec{x}).\vec{v} + \frac{h^2}{2}\,d^2\Phi.(\vec{v},\vec{v}),$$

et on va tenir compte de la correction du second ordre $\frac{h^2}{2}\,d^2\Phi.(\vec{v},\vec{v})$: le développement limité ci-dessus étant exact, le terme du second ordre est plus qu'une correction : la formule est exacte même pour h "grand", et on a implicitement besoin de h grand pour rapidement minimiser ψ en h, cf. (11.19).

Notons le membre de droite par, après division par h pour avoir la pente moyenne :

$$\Psi_{\vec{x}}(\vec{v}) = d\Phi(\vec{x}).\vec{v} + \frac{h}{2}\,d^2\Phi.(\vec{v},\vec{v}),$$

et il s'agit de trouver un optimum de $\Psi_{\vec{x}}$ (minimum ou maximum pour avoir la pente moyenne la plus favorable). La dérivée de $\Psi_{\vec{x}}$ en $\vec{v}$ est donnée dans toute direction $\vec{w}$ par :

$$d\Psi_{\vec{x}}(\vec{v}).\vec{w} = d\Phi(\vec{x}).\vec{w} + h\,d^2\Phi.(\vec{v},\vec{w}).$$

(Calcul immédiat à l'aide de la définition $= \lim_{k\to 0}\frac{\Psi_{\vec{x}}(\vec{v}+k\vec{w})-\Psi_{\vec{x}}(\vec{v})}{k}$, et on se sert de $d^2\Phi$ symétrique.)

Et $d\Psi_{\vec{x}}$ est nul pour tout $\vec{w}$ lorsque

$$d\Phi(\vec{x}).\vec{w} + h\,d^2\Phi.(\vec{v},\vec{w}) = 0.$$

Prenons alors comme produit scalaire, le produit scalaire donné par $g(\vec{v},\vec{w}) = d^2\Phi.(\vec{v},\vec{w})$. L'égalité ci-dessus se réécrit :

$$g(\vec{\text{grad}}_g\Phi(\vec{x}),\vec{w}) + h\,g(\vec{v},\vec{w}) = 0 = g(\vec{\text{grad}}_g\Phi(\vec{x}){+}h\,\vec{v}\ ,\ \vec{w}), \quad \forall\vec{w}\in\mathbb{R}^2,$$

d'où :

$$\vec{\text{grad}}_g\Phi(\vec{x}) = -h\vec{v}.$$

On a la bonne direction $\vec{v}$ (de meilleure descente moyenne), direction parallèle à $\vec{\text{grad}}_g\Phi(\vec{x})$. ▪

11.4 Double contraction de deux tenseurs de T^1_1

On effectue une première contraction pour obtenir, avec les notations de l'exemple précédent, $M.L = \sum_{i\ell}(ML)^i_\ell\vec{e}_i\otimes e^\ell$, puis on effectue une seconde contraction, i.e. on prend la trace du tenseur

$M.L \in T_1^1$, et on note (double contraction objective) :

$$M \odot L = \mathrm{Tr}(M.L) = \sum_{i=1}^{n} (ML)_i^i = \sum_{i,j=1}^{n} M_j^i L_i^j \qquad (= \mathrm{Tr}(M \circ L) \in \mathbb{R}). \tag{11.20}$$

Ici on note $\mathrm{Tr}(M \circ L)$ si on pense en terme d'endomorphismes (la trace est un invariant : même valeur qu'elle que soit la base utilisée), et on note $\mathrm{Tr}(M.L)$ si on pense en termes de tenseurs de T_1^1 (contraction).

Remarque 11.14 Attention $M \odot L$ n'est pas le produit termes à termes

$$M : L := \sum_{i,j=1}^{n} M_j^i L_j^i$$

souvent utilisé en mécanique. En particulier $M : L$ n'est pas un invariant. Il est également défini par $M : L = M^T \odot L$, où l'apparition du transposé suppose l'introduction préalable d'un produit scalaire (d'ailleurs $M : L = \sum_{i,j=1}^{n} M_j^i L_j^i = \sum_{i=1}^{n^2} A^i B^j$ est le produit scalaire de Frobenius dans $\mathbb{R}^{n^2}$, où on a noté $[A^i]$ est la matrice colonne obtenue en mettant les colonnes de M “les unes à la suite des autres”, idem pour $[B^i]$. Et on trouve la notation $\sum_{ij} M_j^i L_j^i = \mathrm{Tr}(M^T \circ L)$ (produit terme à terme). (Il n'y a plus d'ambiguïté si M ou L est symétrique.) Et on aura par exemple $(L.\vec{u}, M.\vec{v})_{\mathbb{R}^n} = (M^T.L.\vec{u}, \vec{v})_{\mathbb{R}^n}$.

Références

[1] Abraham R., Marsden J.E. : *Foundation of mechanics, 2nd edition*. Addison-Wesley, 1985.

[2] Arnold V. : *Mathematical Methods of Classical Mechanics*. Second Edition, Springer 1989.

[3] Cartan H. : *Cours de calcul différentiel*. Hermann 1990.

[4] Chenciner A. : 5 cours sur la géométrie différentielle, école d'été "Maths et Cerveau" de l'Institut de Mathématiques de Jussieu, IHP juin 2005, `http://www.imcce.fr/Equipes/ASD/person/chenciner/chen_preprint.php`

[5] Germain P. : *Cours de mécanique*. Ecole Polytechnique, édition 1983.

[6] Gurtin, M. E. : *Topics in Finite Elasticity*. Regional conference series in applied mathematics. Society for Industrial and Applied Mathematics. 1981. Philadelphia, Pennsylvania 19103.

[7] Marsden J.E., Hughes T.J.R. : *Mathematical Foundations of Elasticity*. Dover publications, 1993.

[8] Misner C.W., Thorne K.S., Wheeler J.A. : *Gravitation*. W.H. Freeman and Cie, 1973

[9] Spivak M. : *A comprehensive introduction to differential geometry, Volume 1*. Publish or Perish Inc., Third Edition, 1999.

[10] Truesdell C., Noll W. : *The non-linear field theory of mechanics*. Handbuch der physik Band III/3, Springer-Verlag 1965