Géométrie 2

R. Taillefer
Université Jean Monnet
Licence de Mathématiques 2007-2008
Semestre 4

Table des matières

Formes bilinéaires et sesquilinéaires

I Formes bilinéaires symétriques

Soit $\mathbb{K}$ un corps arbitraire dans lequel 2 est inversible. Soit E un espace vectoriel sur $\mathbb{K}$ de dimension finie $n \geqslant 1$.

A Généralités

A.1 Formes bilinéaires symétriques

Définition A.1. *Une forme bilinéaire symétrique sur E est une application $f : E \times E \to \mathbb{K}$ telle que :*

(i) $f(\lambda u + \mu v, w) = \lambda f(u, w) + \mu f(v, w)$ pour tous $u, v, w \in E$ et tous $\lambda, \mu \in \mathbb{K}$; cela signifie que f est linéaire en la première variable.

(ii) $f(v, u) = f(u, v)$.

Remarque A.2. Ces deux propriétés impliquent la propriété suivante :

$$\text{(iii) } f(u, \lambda v + \mu w) = \lambda f(u, w) + \mu f(u, v)$$

En effet, $f(u, \lambda v + \mu w) = f(\lambda v + \mu w, u) = \lambda f(v, u) + \mu f(w, u) = \lambda f(u, v) + \mu f(u, w)$.

Exemples A.3. • Soit $E = \mathbb{R}^2$ et $f(u, v) = u_1 v_1 + u_2 v_2$ où $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ et $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$. C'est le produit scalaire usuel sur $\mathbb{R}^2$. On le note parfois $u \cdot v$.

- Soit $E = \mathbb{R}^4$ et soit $f(u, v) = u_1 v_1 + u_2 v_2 + u_3 v_3 - c^2 u_4 v_4$ où c est une constante, $u = t(u_1, u_2, u_3, u_4)$ et $v = t(v_1, v_2, v_3, v_4)$. (Si c est la vitesse de la lumière, cette forme bilinéaire symétrique joue un rôle important en théorie de la relativité).
- Soit E l'espace vectoriel des fonctions polynomiales de degré inférieur ou égal à $n - 1$, et soit $f(p, q) = \int_{-1}^{1} p(t) q(t) dt$. C'est une forme bilinéaire symétrique.

Définition A.4. *Soit f une forme bilinéaire symétrique sur E. On lui associe l'application linéaire*

$$\begin{array}{rcl} \theta_f : E & \to & E^* \\ u & \mapsto & f(u, \cdot), \end{array}$$

c'est-à-dire que pour chaque $u \in E$, $\theta_f(u)$ est une forme linéaire sur E qui associe à $v \in E$ le scalaire $f(u, v)$.

On appelle rang de f le rang de l'application linéaire θ_f.

Si le rang de f est égal à n, on dit que f est non dégénérée.

Remarque A.5. Grâce au théorème du rang et au fait que $\dim E^* = \dim E$, on voit que f est non dégénérée si et seulement si θ_f est injective, ce qui est équivalent à dire que θ_f est un isomorphisme.

Fixons une base $\mathcal{B} = \{e_1, \ldots, e_n\}$ de E.

Définition A.6. *Considérons la matrice $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ de taille $n \times n$ définie de la façon suivante :*

$$a_{ij} = f(e_i, e_j),\ 1 \leqslant i, j \leqslant n.$$

On l'appelle la matrice de f dans la base $\mathcal{B}$.

Remarque A.7. Notons que la matrice de f dans la base $\mathcal{B}$ est égale à la matrice de θ_f dans les bases $\mathcal{B}$ et $\mathcal{B}^*$, où $\mathcal{B}^*$ est la base duale de $\mathcal{B}$. En effet, $\theta_f(e_j) = \sum_{i=1}^n \theta_f(e_j)(e_i)e_i^* = \sum_{i=1}^n f(e_i, e_j)e_i^*$.

En particulier, le rang de f est égal au rang de sa matrice.

Propriétés A.8. On constate les propriétés immédiates suivantes :

(i) La matrice A est *symétrique*, c'est-à-dire que la transposée tA est égale à A. Nous noterons $\mathcal{S}ym_n(\mathbb{K})$ l'espace vectoriel des matrices symétriques $n \times n$ à coefficients dans $\mathbb{K}$.

(ii) Fixons une base $\mathcal{B} = \{e_1, \ldots, e_n\}$ de E. Soient $u = \sum_{i=1}^n x_i e_i$ et $v = \sum_{i=1}^n y_i e_i$. Alors $f(u, v) = \sum_{1 \leqslant i,j \leqslant n} a_{ij} x_i y_j$. Autrement dit, se donner f équivaut à se donner une fonction polynôme symétrique homogène de degré 2 en les x_i et y_j.

On peut réécrire cela matriciellement sous la forme :

$$f(u, v) = tXAY \tag{A.1}$$

où $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$.

(iii) Soit $A \in \mathcal{S}ym_n(\mathbb{K})$. Alors A définit une forme bilinéaire sur $\mathbb{K}^n$ (ou sur E en fixant une base de E) par la formule (**??**).

Finalement, se donner une forme bilinéaire symétrique revient à se donner une matrice symétrique.

Que se passe-t-il si nous changeons de base ? Soit $\mathcal{B}' = \{e'_1, \ldots, e'_n\}$ une autre base de E, et posons : $a'_{ij} = f(e'_i, e'_j)$. Soit A' la matrice dont les coefficients sont les a'_{ij}, et notons $u = \sum_{i=1}^n x'_i e'_i$ et $v = \sum_{i=1}^n y'_i e'_i$. Ces vecteurs ont maintenant pour vecteurs colonne $X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ et $Y' = \begin{pmatrix} y'_1 \\ \vdots \\ y'_n \end{pmatrix}$. On a donc

$$f(u, v) = tX'A'Y'. \tag{A.2}$$

Soit P la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$, donc P est une matrice $n \times n$ inversible telle que $X = PX'$ et $Y = PY'$. En reportant dans (**??**), on obtient $f(u, v) = tX'tPAPY'$, et en comparant à (**??**) on obtient

$$A' = tPAP.$$

Définition A.9. *Soient A et A' appartenant à $\mathcal{S}ym_n(\mathbb{K})$. On dit que A' est congruente à A s'il existe une matrice inversible P telle que $A' = tPAP$.*

Remarque A.10. La relation de congruence est une relation d'équivalence dans $\mathcal{S}ym_n(\mathbb{K})$:

(i) A est congruente à elle-même : prendre $P = I_n$.

(ii) Si A' est congruente à A, alors A est congruente à A', car si $A' = tPAP$, on a bien $A = t(P^{-1})A'P^{-1}$.

(iii) Si A' est congruente à A et si A'' est congruente à A', alors A'' est congruente à A; en effet, si $A' = tPAP$ et si $A'' = tQA'Q$, alors $A'' = t(PQ)A(PQ)$.

Remarque A.11. Si A et A' sont congruentes, alors elles ont le même rang. Le rang est donc un invariant pour la relation de congruence.

Le rang suffit-il à décrire l'ensemble des classes de congruence ? On verra plus tard que la réponse est OUI si $\mathbb{K} = \mathbb{C}$ et NON si $\mathbb{K} = \mathbb{R}$.

A.2 Formes quadratiques, forme quadratique associée à une forme bilinéaire symétrique

Définition A.12. *Une forme quadratique sur E est une application $q : E \to \mathbb{K}$ vérifiant :*

(i) $\forall \lambda \in \mathbb{K}, \forall u \in E, q(\lambda u) = \lambda^2 q(u)$,

(ii) l'application

$$\begin{array}{rcl} f_q : E \times E & \to & \mathbb{K} \\ (u,v) & \mapsto & \frac{1}{2}(q(u+v) - q(u) - q(v)) \end{array}$$

est une forme bilinéaire symétrique.

f_q s'appelle la forme polaire de q.

Exemple A.13. On pose $E = \mathbb{R}^2$ et $q(u) = u_1^2 + u_2^2$ pour tout $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \in E$; alors q est une forme quadratique.

Définition-Proposition A.14. Soit f une forme bilinéaire symétrique sur E. L'application $q_f : E \to \mathbb{K}$ définie par

$$\forall u \in E,\ q_f(u) = f(u,u)$$

est une forme quadratique, qui s'appelle la *forme quadratique associée à f*.

Démonstration. Soient $\lambda \in \mathbb{K}$ et $u \in E$. Alors $q(\lambda u) = f(\lambda u, \lambda u) = \lambda^2 f(u,u) = \lambda^2 q(u)$.

D'autre part, calculons la forme polaire de q :

$$\begin{aligned} 2f_q(u,v) &= f(u+v,u+v) - f(u,u) - f(v,v) \\ &= f(u,u) + f(u,v) + f(v,u) + f(v,v) - f(u,u) - f(v,v) \\ &= 2f(u,v) \end{aligned}$$

donc $f_q = f$ est une forme bilinéaire symétrique. □

Proposition A.15. Se donner une forme quadratique équivaut à se donner une forme bilinéaire symétrique, c'est-à-dire que l'on a une bijection :

$$\begin{array}{ccc} \{\text{formes bilinéaires symétriques}\} & \longleftrightarrow & \{\text{formes quadratiques}\} \\ f & \mapsto & q_f \\ f_q & \leftarrow\!\shortmid & q \end{array}$$

Comment s'exprime une forme quadratique si l'on fixe une base ?

Soit $\mathcal{B} = \{e_1, \ldots, e_n\}$ une base de E, et soit A la matrice de f dans la base $\mathcal{B}$. Si $u = \sum_{i=1}^n x_i e_i$, alors

$$q(u) = f(u,u) = \sum_{i,j} a_{ij} x_i x_j = \sum_{i=1}^n a_{ii} x_i^2 + 2 \sum_{i<j} a_{ij} x_i x_j.$$

On remarque que si A est diagonale, alors $q(u) = \sum_{i=1}^n a_{ii} x_i^2$ est une combinaison linéaire de carrés, d'où le nom de forme quadratique. Nous allons voir (théorème **??** et les remarques qui suivent) que, moyennant un changement de base, on peut toujours se ramener à ce cas.

A.3 Formes quadratiques et orthogonalité

Définition A.16. *Soit q une forme quadratique associée à une forme bilinéaire symétrique f sur E. On appelle noyau de q le noyau de l'application θ_f. Il s'agit donc de l'ensemble*

$$E^{\perp_q} := \{x \in E \mid \forall y \in E,\ f(x,y) = 0\}.$$

On le note $E^\perp$ lorsqu'il n'y a pas d'ambiguïté.

On remarque que q est non dégénérée si et seulement si son noyau est $\{0\}$ (théorème du rang).

Définition A.17. *Soit f une forme bilinéaire symétrique sur E. On dit que deux éléments x et y de E sont orthogonaux (pour f) si $f(x,y) = 0$.*

Si H est une partie de E, l'orthogonal de H (pour f) est le sous-ensemble

$$H^{\perp_f} := \{x \in E \mid \forall y \in H,\ f(x,y) = 0\}.$$

On le note $H^{\perp}$ lorsqu'il n'y a pas d'ambiguïté.

Remarque A.18. Supposons que H soit un sous-espace vectoriel. Considérons l'application linéaire

$$\begin{array}{rcl} \theta_f^H : E & \to & H^* \\ u & \mapsto & \theta_f(u)_{|H}. \end{array}$$

L'orthogonal de H est le noyau de θ_f^H. En effet, identifions le noyau de θ_f^H : un vecteur $u \in E$ est dans $\operatorname{Ker}\theta_f^H$ si et seulement si $\theta_f(u)(v) = 0$ pour tout $v \in H$. Donc $u \in \operatorname{Ker}\theta_f$ si et seulement si $f(u,v) = 0$ pour tout $v \in H$. Par définition, c'est équivalent à dire que u appartient à $H^{\perp}$. Donc $\operatorname{Ker}\theta_f^H = H^{\perp}$.

Remarque A.19. Soit H un sous-espace vectoriel de E et soit $\{e_1, \dots, e_r\}$ une base de H. Alors

$$H^{\perp} = \{x \in E \mid \forall i = 1, \dots r,\ f(x, e_i) = 0\}.$$

Proposition A.20. Soit f une forme bilinéaire symétrique sur E et soit H une partie de E. Alors :

(i) $H^{\perp}$ est un sous-espace vectoriel de E.

(ii) Si $H \subseteq K$, alors $K^{\perp} \subseteq H^{\perp}$.

(iii) Si H est un *sous-espace vectoriel* de E, et si f est non dégénérée, alors $\dim H + \dim H^{\perp} = \dim E$ et $(H^{\perp})^{\perp} = H$.

Démonstration. (TD) □

Remarque A.21. $H^{\perp} = \operatorname{vect}(H)^{\perp}$. (TD).

Définition A.22. *Soit f une forme bilinéaire symétrique sur E. On dit qu'une famille $\{e_1, \dots, e_r\}$ de E est orthogonale pour f (ou pour la forme quadratique associée q) si $f(e_i, e_j) = 0$ dès que $i \neq j$.*

On dit que la famille $\{e_1, \dots, e_r\}$ est orthonormale pour f (ou pour q) si elle est orthogonale pour f et si $f(e_i, e_i) = 1$ pour tout $i = 1, \dots, r$.

Remarque A.23. Si $\mathcal{B}$ est une base orthogonale pour f, alors la matrice de f dans la base $\mathcal{B}$ est diagonale.

Si $\mathcal{B}$ est orthonormale pour f, alors la matrice de f dans la base $\mathcal{B}$ est I_n.

Théorème A.24. Soit E un espace vectoriel de dimension finie sur $\mathbb{K}$. Alors il existe des bases orthogonales pour toute forme bilinéaire symétrique sur E.
Matriciellement, cela signifie que pour $A \in \mathcal{S}ym_n(\mathbb{K})$, il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $tPAP = D$ est diagonale de la forme

$$D = \begin{pmatrix} \alpha_1 & & & & & \\ & \ddots & & & \mathbf{0} & \\ & & \alpha_r & & & \\ & & & 0 & & \\ & \mathbf{0} & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

avec $\alpha_1 \neq 0, \dots, \alpha_r \neq 0$. Nécessairement, r est égal au rang de A.

Remarque A.25. Il n'existe pas nécessairement de base orthonormale pour f.

Démonstration. Soit f une forme bilinéaire symétrique sur E. On veut trouver une base $\mathcal{B} = \{e_1, \ldots, e_n\}$ telle que $f(e_i, e_j) = 0$ si $i \neq j$. On raisonne par récurrence sur n.

Si $n = 1$, il n'y a rien a démontrer, puisque toute base est orthogonale et tout vecteur non nul est un vecteur propre.

Supposons donc que pour tout espace vectoriel de dimension inférieure ou égale à $n - 1$ et pour toute forme bilinéaire symétrique sur cet espace vectoriel il existe une base orthogonale.

Soit E un espace vectoriel de dimension n et soit f une forme bilinéaire symétrique sur E.

Si $f = 0$, alors toute base est orthogonale.

Si $f \neq 0$, alors il existe $e_1 \in E$ tel que $f(e_1, e_1) \neq 0$ (en effet, si $f(x, x) = 0$ pour tout $x \in E$, alors $f(x, y) = \frac{1}{2}(f(x + y, x + y) - f(x, x) - f(y, y)) = 0$ pour tout $x, y \in E$, et donc $f = 0$). Posons $H = \mathrm{vect}(e_1)$, et montrons que $E = H \oplus H^\perp$:

- $H \cap H^\perp = \{0\}$: en effet, soit $u \in H \cap H^\perp$; $u = \lambda e_1$ pour un $\lambda \in \mathbb{K}$, donc $f(u, u) = \lambda^2 f(e_1, e_1)$; mais aussi $f(u, u) = 0$ puisque $u \in H^\perp$, ce qui implique que $\lambda = 0$ et donc $u = 0$.
- $\dim H + \dim H^\perp = \dim E$: considérons $\theta_f^H : E \to H^*$. On sait que $H^\perp = \mathrm{Ker}\, \theta_f^H$. Donc $\dim H^\perp = \dim E - \dim \mathrm{Im}\, \theta_f^H$ par le théorème du rang. Déterminons donc $\dim \mathrm{Im}\, \theta_f^H$: on sait déjà que $\dim \mathrm{Im}\, \theta_f^H \leqslant \dim H^* = \dim H = 1$. De plus, $\theta_f^H(e_1)(e_1) = f(e_1, e_1) \neq 0$, donc θ_f^H n'est pas l'application linéaire nulle. Donc $\dim \mathrm{Im}\, \theta_f^H \geqslant 1$. On en déduit donc que $\dim H^\perp = \dim E - 1 = \dim E - \dim H$.

On a donc $\dim H^\perp = n - 1$.

On définit une forme bilinéaire symétrique g sur $H^\perp$ en posant $g(u, v) = f(u, v)$ pour tous $u, v \in H^\perp$.

Par hypothèse de récurrence, il existe une base orthogonale $\{e_2, \ldots, e_n\}$ de $H^\perp$ orthogonale pour g. Il est clair que $\{e_1, e_2, \ldots, e_n\}$ est une base de E. Il reste a vérifier qu'elle est bien orthogonale pour f :

- Si $2 \leqslant i \leqslant n$, on a $e_i \in H^\perp$, donc $f(e_1, e_i) = 0 = f(e_i, e_1)$.
- Si $2 \leqslant i \neq j \leqslant n$, on a $f(e_i, e_j) = g(e_i, e_j) = 0$ par construction. □

Remarque A.26. Pour la forme quadratique q associée à f, le théorème se traduit par le fait que $q(u)$ est une combinaison linéaire de carrés.

Comment fait-on dans la pratique pour exprimer $q(u)$ comme une combinaison linéaire de carrés ? Une méthode algorithmique est donnée par la *méthode de Gauss* suivante : soit à réduire la forme quadratique q définie sur E par :

$$q(u) = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j.$$

On suit alors pas à pas l'algorithme suivant :

$\boxed{1}$ On prend un terme en x_i^2 (avec un coefficient le plus simple possible afin d'alléger les calculs) et on considère tous les termes contenant x_i que l'on écrit comme le début du développement d'un carré. On poursuit cette étape jusqu'à épuisement des termes en x_i^2. Il y a alors deux possibilités : soit le problème est résolu et c'est terminé, soit ce n'est pas le cas et on passe à l'étape 2.

$\boxed{2}$ S'il n'y a pas (ou plus) de termes en x_i^2, on considère un terme "simple" en $x_i x_j$ et on écrit tous les termes sous la forme :

$$\lambda x_i x_j + x_i R + x_j S = \lambda \underbrace{(x_i + \frac{S}{\lambda})}_{a} \underbrace{(x_j + \frac{R}{\lambda})}_{b} - \frac{RS}{\lambda}$$

et on utilise le fait que $ab = \frac{1}{4}\left[(a + b)^2 - (a - b)^2\right]$.

Ensuite, on reprend la méthode à partir de l'étape 1.

Exemple A.27. Soit $E = \mathbb{R}^4$ et soit $q(u) = x_1^2 + x_2^2 + x_4^2 + 2x_1x_2 + 2x_2x_3 + 2x_4x_1 - x_3x_4$. On est dans le premier cas, avec le coefficient de x_1^2 non nul. Donc

$$\begin{aligned} q(u) &= x_1^2 + 2x_1(x_2+x_4) + x_2^2 + x_4^2 + 2x_2x_3 - x_3x_4 \\ &= (x_1+x_2+x_4)^2 - (x_2+x_4)^2 + x_2^2 + x_4^2 + 2x_2x_3 - x_3x_4 \\ &= (x_1+x_2+x_4)^2 + 2x_2x_3 - 2x_2x_4 - x_3x_4. \end{aligned}$$

La forme quadratique qui reste ne contient plus de carrés, on passe donc à la deuxième étape :

$$\begin{aligned} q(u) &= (x_1+x_2+x_4)^2 - x_3x_4 + x_3(2x_2) - x_4(2x_2) \\ &= (x_1+x_2+x_4)^2 + (x_3+2x_2)(-x_4+2x_2) - 4x_2^2 \\ &= (x_1+x_2+x_4)^2 + \tfrac{1}{4}(4x_2+x_3-x_4)^2 - \tfrac{1}{4}(x_3+x_4)^2 - 4x_2^2. \end{aligned}$$

B Réduction sur un corps arbitraire, sur $\mathbb{C}$, sur $\mathbb{R}$.

On cherche à classifier les formes bilinéaires symétriques à congruence près (la congruence étant celle des matrices correspondantes).

Nous avons déjà vu une première réduction dans le théorème **??**. Mais lorsque l'on choisit des corps plus particuliers, on peut en dire plus.

Théorème B.1. Supposons que tout élément de $\mathbb{K}$ ait une racine carrée dans $\mathbb{K}$ (c'est le cas par exemple si $\mathbb{K} = \mathbb{C}$). Soit f une forme bilinéaire symétrique sur E sur $\mathbb{K}$. Alors il existe une base (nécessairement orthogonale pour f) de E dans laquelle la matrice associée à f est égale à

$$\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

En particulier, deux matrices A et A' dans $\mathcal{S}ym_n(\mathbb{K})$ sont congruentes si et seulement si elles ont même rang. Il y a donc $n+1$ classes de congruence.

Démonstration. Soit $\{e_1, \ldots, e_n\}$ une base dans laquelle la matrice de la forme bilinéaire symétrique f est de la forme

$$D = \begin{pmatrix} \alpha_1 & & & & & \\ & \ddots & & & \mathbf{0} & \\ & & \alpha_r & & & \\ & & & 0 & & \\ & \mathbf{0} & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

avec $\alpha_1 \neq 0, \ldots, \alpha_r \neq 0$ (théorème **??**). Pour chaque α_i, il existe par hypothèse $\beta_i \in \mathbb{K}$ tel que $\beta_i^2 = \alpha_i$. Posons $\varepsilon_i = \begin{cases} \beta_i^{-1} e_i & \text{si } i \leqslant r \\ e_i & \text{si } i > r. \end{cases}$

Alors $\{\varepsilon_1, \ldots, \varepsilon_n\}$ est une base de E. De plus, $f(\varepsilon_i, \varepsilon_j)$ est un multiple de $f(e_i, e_j)$, donc la matrice associée à f dans cette nouvelle base est encore diagonale. De plus, si $i > r$, on a $f(\varepsilon_i, \varepsilon_i) = f(e_i, e_i) = 0$, et si $i \leqslant r$ on a $f(\varepsilon_i, \varepsilon_i) = \beta_i^{-2} f(e_i, e_i) = \alpha_i^{-1}\alpha_i = 1$. Donc la matrice est bien celle requise.

Nous savons que deux matrices congruentes ont même rang. Supposons donc que A et A' soient deux matrices symétriques ayant le même rang, r. Alors A est congruente à

$$\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

et A' est congruente à

$$\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},$$

donc par transitivité, A est congruente à A'.

Il y a donc autant de classes de congruence que de valeurs possibles pour le rang, $0 \leqslant r \leqslant n$. $\square$

Théorème B.2. (Loi d'inertie de Sylvester) Supposons que $\mathbb{K} = \mathbb{R}$. Soit q une forme quadratique de rang r. Alors il existe un entier s avec $0 \leqslant s \leqslant r$ et une base $\{e_1, \ldots, e_n\}$ tels que :

$$q(u) = x_1^2 + \ldots + x_s^2 - x_{s+1}^2 - \ldots - x_r^2, \quad \text{avec } u = \textstyle\sum_{i=1}^n x_i e_i, \tag{B.1}$$

où l'entier s ne dépend pas de la base $\{e_1, \ldots, e_n\}$ choisie.
Le couple $(s, r-s)$ s'appelle la *signature* de q (ou de f, ou de la matrice A).
Autrement dit, toute matrice $A \in \mathcal{S}ym_n(\mathbb{R})$ de signature (s,t) avec $s, t \geqslant 0$ est congruente à

$$\begin{pmatrix} I_s & 0 & 0 \\ 0 & -I_t & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

En particulier, deux matrices A et A' dans $\mathcal{S}ym_n(\mathbb{R})$ sont congruentes si et seulement si elles ont même signature. Il y a donc $\frac{(n+1)(n+2)}{2}$ classes de congruence.

Démonstration. On raisonne de manière similaire au théorème précédent.

Quitte à réordonner la base, nous pouvons supposer que $\alpha_i > 0$ pour $1 \leqslant i \leqslant s$, que $\alpha_i < 0$ pour $s < i \leqslant r$ et que $\alpha_i = 0$ pour $i > r$. Pour tout $i \leqslant r$, il existe $\beta_i \in \mathbb{R}$ tel que $\beta_i^2 = |\alpha_i|$. La démonstration de l'existence d'une matrice de la forme requise se termine comme dans le cas précédent.

Il nous faut donc montrer l'indépendance de s du choix de la base. Remarquons que dans la construction ci-dessus, $s = \#\{e_i \mid q(e_i) > 0\}$.

Supposons que l'on puisse écrire $q(u) = \sum_{i=1}^{s'} x_i^2 - \sum_{i=s'+1}^{r} x_i^2$, dans une base orthogonale $\mathcal{B}' = \{e_1', \ldots, e_n'\}$. On a alors $s' = \#\{e_i' \mid q(e_i') > 0\}$. Posons

$$\begin{aligned} F &= \operatorname{vect}\{e_1, \ldots, e_s\}, & G &= \operatorname{vect}\{e_{s+1}, \ldots, e_n\} \\ F' &= \operatorname{vect}\{e_1', \ldots, e_s'\}, & G' &= \operatorname{vect}\{e_{s+1}', \ldots, e_n'\}. \end{aligned}$$

Alors $E = F \oplus G$ et $E = F' \oplus G'$.

Supposons maintenant que $u \in F \cap G'$. Alors, si $u \neq 0$, comme $u \in F$ on a $q(u) > 0$. D'autre part, puisque $u \in G'$, on a $q(u) \leqslant 0$. Donc on obtient une contradiction, et $F \cap G' = \{0\}$. Donc $F + G' = F \oplus G'$, et

$$\dim(F \oplus G') = \dim F + \dim G' = s + (n - s')$$

et ceci doit être inférieur à n, donc $s \leqslant s'$.

De même, en considérant F' et G, on montre que $s' \leqslant s$, et donc $s' = s$.

Le nombre de classes de congruence est donc le nombre de signatures, ce qui est égal au cardinal de l'ensemble $\left\{(r,s) \in \mathbb{N}^2 \mid 0 \leqslant r \leqslant n,\ 0 \leqslant s \leqslant r\right\}$, qui est $\frac{(n+1)(n+2)}{2}$. $\square$

Remarque B.3. Après avoir effectué la réduction de Gauss sur une forme quadratique, la signature de la forme quadratique est (nombre de carrés à coefficients strictements positifs, nombre de carrés à coefficients strictement négatifs).

Définition B.4. *La forme bilinéaire symétrique $f : E \times E \to \mathbb{R}$ (ou la forme quadratique q, ou la matrice symétrique A) est dite définie positive si $r = s = n$.*

Remarque B.5. Une forme quadratique q est définie positive si et seulement si $q(u) > 0$ pour tout $u \neq 0$:

- Supposons que q soit définie positive. Alors il existe une base $\{e_1, \ldots, e_n\}$ dans laquelle la matrice associée à q est I_n. Donc si $u = \sum_i x_i e_i \neq 0$, on a $q(u) = \sum_i q(e_i) x_i^2 = \sum_{i=1}^n x_i^2 > 0$.
- Réciproquement, soit $\{e_1, \ldots, e_n\}$ une base dans laquelle la matrice associée à q est de la forme

$$D = \begin{pmatrix} \alpha_1 & & & & \\ & \ddots & & & \mathbf{0} \\ & & \alpha_r & & \\ & \mathbf{0} & & \ddots & \\ & & & & \alpha_n \end{pmatrix}$$

avec $\alpha_i \in \mathbb{K}$. Par hypothèse, $\alpha_i = q(e_i) > 0$ pour tout i, donc on peut multiplier chaque élément de la base par un scalaire pour obtenir une base dans laquelle la matrice associée à q est I_n.

Remarque B.6. Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est symétrique définie positive si et seulement s'il existe $P \in \mathrm{GL}_n(\mathbb{R})$ telle que $A = {}^tPP$.

Remarque B.7. Si $\mathbb{K} = \mathbb{Q}$, le problème est beaucoup plus compliqué. Il y a une infinité dénombrable de classes de congruence dans $\mathcal{S}ym_n(\mathbb{Q})$. (*c.f.* [J.P. Serre, Cours d'arithmétique]).

C Espaces vectoriels euclidiens, endomorphismes symétriques, orthogonaux

C.1 Généralités sur les espaces euclidiens

Définition C.1. *Soit E un espace vectoriel sur $\mathbb{R}$. Un produit scalaire sur E est une forme bilinéaire symétrique dont la forme quadratique associée est définie positive sur E.*

Autrement dit, $f : E \times E \to \mathbb{R}$ vérifie :

(i) $f(u + \lambda v, w) = f(u, w) + \lambda f(v, w)$

(ii) $f(v, u) = f(u, v)$

(iii) $f(u, u) > 0$ *si* $u \neq 0$

pour tous $u, v, w \in E$ et tout $\lambda \in \mathbb{R}$.

On notera le produit scalaire de la façon suivante : $f(u, v) = \langle u, v \rangle$.

Définition C.2. *Un espace euclidien est un couple (E, f) où E est un $\mathbb{R}$-espace vectoriel de dimension finie $n \geqslant 1$ et où f est un produit scalaire sur E.*

Remarque C.3. On a $\langle u, u \rangle = 0 \iff u = 0$.

Exemples C.4.
- Sur $\mathbb{R}^n$, on a le produit scalaire usuel ou *canonique* : $\langle u, v \rangle = u \cdot v = \sum_{i=1}^n u_i v_i$. C'est l'unique produit scalaire défini sur $\mathbb{R}^n$ pour lequel la base canonique est orthonormale. L'espace $\mathbb{R}^n$ muni de ce produit scalaire est appelé *espace euclidien canonique. (Laissé en exercice).*
- Sur l'espace vectoriel des matrices $n \times n$ symétriques, $\langle A, B \rangle = \mathrm{tr}(AB)$ est un produit scalaire. (TD)
- Sur l'espace vectoriel des polynômes à coefficients réels de degré inférieur à $n-1$, $\langle p, q \rangle = \int_0^1 p(t)q(t)dt$ est un produit scalaire. (*c.f.* TD)

Définition-Proposition C.5. On pose $\|u\| = \sqrt{\langle u, u\rangle}$ pour tout $u \in E$.
L'application $\|\cdot\| : E \to \mathbb{R}$ est alors une *norme* sur E, c'est-à-dire qu'elle vérifie, pour tous $u, v \in E$:

(i) $\|u\| \geqslant 0$

(ii) $\|u\| = 0 \iff u = 0$

(iii) $\|\lambda u\| = |\lambda| \cdot \|u\|$

(iv) $\|u+v\| \leqslant \|u\| + \|v\|$ *(inégalité triangulaire).*

$\|\cdot\|$ est appelée la *norme euclidienne* de E.

Démonstration. Les trois premières propriétés sont faciles à vérifier. Pour l'inégalité triangulaire, on utilise l'inégalité de Cauchy-Schwarz qui suit. □

Proposition C.6. (Inégalité de Cauchy-Schwarz) Soit $(E, \langle,\rangle)$ un espace euclidien. Alors, pour tous $u, v \in E$, on a
$$|\langle u, v\rangle| \leqslant \|u\| \ \|v\| .$$
De plus, cette inégalité devient une égalité si et seulement si u et v sont colinéaires.

Démonstration. Soit $\lambda \in \mathbb{R}$. Alors :
$$\begin{aligned} 0 &\leqslant \langle u+\lambda v, u+\lambda v\rangle \\ &= \|u\|^2 + \lambda\langle u, v\rangle + \lambda\langle v, u\rangle + \lambda^2 \|v\|^2 \\ &= \|u\|^2 + 2\langle u, v\rangle\lambda + \|v\|^2 \lambda^2. \end{aligned} \tag{C.1}$$
Il s'agit d'un polynôme en λ du second degré qui est toujours positif, donc son discriminant doit être négatif (pas de changement de signe, donc pas de racines distinctes). Ainsi $\langle u, v\rangle^2 - \|u\|^2 \|v\|^2 \leqslant 0$ et le résultat suit.

Maintenant supposons que l'on ait égalité dans l'inégalité de Cauchy-Schwarz. Alors le discriminant est nul, c'est-à-dire que le polynôme (**??**) possède une racine, disons λ_0. En remontant les égalités, on voit que $\|u + \lambda_0 v\| = 0$, ce qui implique que $u + \lambda_0 v = 0$ et donc u et v sont colinéaires.

Si u et v sont colinéaires, il est clair que l'on a égalité. □

Démonstration de l'inégalité triangulaire : On applique l'inégalité de Cauchy-Schwarz pour obtenir :
$$\|u+v\|^2 = \|u\|^2 + 2\langle u, v\rangle + \|v\|^2 \leqslant \|u\|^2 + 2\|u\| \|v\| + \|v\|^2 = (\|u\| + \|v\|)^2.$$

Exemples C.7. En appliquant l'inégalité de Cauchy-Schwarz à des produits scalaires classiques, nous pouvons obtenir des inégalités intéressantes :

(i) Si $E = \mathbb{R}^n$ muni du produit scalaire usuel, $u \in E$ et $v = t(1, \ldots, 1)$, alors
$$\left(\sum_{i=1}^n x_i\right)^2 \leqslant n \sum_{i=1}^n x_i^2.$$

(ii) Si E est l'espace des polynômes de degré au plus $n-1$ muni du produit scalaire défini par $\langle p, q\rangle = \int_{-1}^1 p(t)q(t)\,dt$, si $p \in E$ et si $q \equiv 1$, alors
$$\left(\int_{-1}^1 p(t)dt\right)^2 \leqslant 2\int_{-1}^1 p(t)^2 dt.$$

On vérifie facilement :

Propriétés C.8. (i) $\langle u, v\rangle = \frac{1}{2}[\|u+v\|^2 - \|u\|^2 - \|v\|^2]$ *(polarisation)*

(ii) $\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2)$ *(identité du parallélogramme)*

Démonstration. (i)

$$\begin{aligned} \|u+v\|^2 - \|u\|^2 - \|v\|^2 &= \langle u+v, u+v\rangle - \langle u,u\rangle - \langle v,v\rangle \\ &= \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle - \langle u,u\rangle - \langle v,v\rangle \\ &= 2\langle u,v\rangle \end{aligned}$$

□

Un produit scalaire permet aussi de définir des angles : si u et v sont des éléments non nuls d'un espace euclidien, alors on sait d'après l'inégalité de Cauchy-Schwarz que $-\|u\|\ \|v\| \leqslant \langle u,v\rangle \leqslant \|u\|\ \|v\|$. Comme u et v sont non nuls, on sait que leurs normes sont non nulles, et on a donc

$$-1 \leqslant \frac{\langle u,v\rangle}{\|u\|\ \|v\|} \leqslant 1.$$

Il existe donc un unique $\theta \in [0,\pi]$ tel que $\frac{\langle u,v\rangle}{\|u\|\|v\|} = \cos\theta$.

Cela définit donc une notion d'angle, non orienté, en géométrie euclidienne (notons que l'angle entre u et v est égal à l'angle entre v et u).

C.2 Compléments sur les bases orthogonales

Proposition C.9. (Théorème de Pythagore) Soient u et v deux éléments d'un espace euclidien $(E,\langle,\rangle)$. Alors u et v sont orthogonaux (pour $\langle,\rangle$) si et seulement si $\|u+v\|^2 = \|u\|^2 + \|v\|^2$.

Démonstration. Claire □

Remarque C.10. Si $\{e_1,\ldots,e_r\}$ est une famille orthogonale de vecteurs *non nuls*, alors elle est libre. En effet, supposons qu'il existe une combinaison linéaire triviale $\sum_{i=1}^n \lambda_i e_i = 0$ avec $\lambda_i \in \mathbb{R}$. Fixons un indice j et prenons le produit scalaire avec e_j :

$$0 = \langle 0, e_j\rangle = \langle \sum_{i=1}^n \lambda_i e_i, e_j\rangle = \sum_{i=1}^n \lambda_i \langle e_i, e_j\rangle = \lambda_j \left\|e_j\right\|^2.$$

Comme e_j est non nul, sa norme est non nulle, donc $\lambda_j = 0$. Ceci est vrai pour tout j, donc les vecteurs sont linéairement indépendants.

Ainsi, pour affirmer qu'une famille orthogonale de vecteurs non nuls est une base orthogonale de E, il suffit qu'elle ait $n = \dim E$ éléments.

Remarque C.11. Si $\{e_1,\ldots,e_r\}$ est une famille orthogonale de vecteurs *non nuls*, alors $\left\{\frac{e_1}{\|e_1\|},\ldots,\frac{e_r}{\|e_r\|}\right\}$ est une famille orthonormale.

En particulier, il existe toujours des bases orthonormales dans un espace euclidien.

Remarque C.12. Si $\{e_1,\ldots,e_n\}$ est une base orthonormale de E, alors pour tout $u \in E$ on a :

$$u = \sum_{i=1}^n \langle u, e_i\rangle e_i.$$

En effet, posons $v = u - \sum_{i=1}^n \langle u,e_i\rangle e_i$. Alors pour tout j avec $1 \leqslant j \leqslant n$, on a $\langle v, e_j\rangle = \langle u, e_j\rangle - \sum_{i=1}^n \langle u,e_i\rangle\langle e_i,e_j\rangle = \langle u,e_j\rangle - \langle u,e_j\rangle = 0$. On en déduit donc que $\langle v,w\rangle = 0$ pour tout $w \in E$, et en particulier que $\langle v,v\rangle = 0$, donc que $v = 0$.

C.3 Adjoint d'un endomorphisme

Soit $(E, \langle , \rangle)$ un espace euclidien. Nous avons rencontré l'application linéaire

$$\begin{array}{rcl} \theta : E & \to & E^* \\ u & \mapsto & \theta(u) = \langle u, \cdot \rangle \end{array}$$

(dans le cas d'une forme bilinéaire symétrique générale). De plus, dans le cas présent où la forme bilinéaire symétrique est non dégénérée, θ est un isomorphisme.

En particulier, pour toute forme linéaire φ dans E^*, il existe un unique $u \in E$ tel que $\varphi = \theta(u)$, c'est-à-dire tel que $\varphi(v) = \langle u, v \rangle$ pour tout $v \in E$.

Soit maintenant ψ un endomorphisme de E. Alors l'application $\langle u, \psi(\cdot) \rangle$ est une forme linéaire. Il existe donc pour chaque $u \in E$ un unique $u_\psi \in E$ tel que $\langle u, \psi(v) \rangle = \langle u_\psi, v \rangle$ pour tout $v \in E$. Notons $u_\psi = \psi^*(u)$. Nous allons montrer que ψ^* est un endomorphisme linéaire de E, et que c'est le seul endomorphisme linéaire de E vérifiant

$$\forall u, v \in E,\ \langle u, \psi(v) \rangle = \langle \psi^*(u), v \rangle. \tag{C.2}$$

Montrons d'abord la linéarité. Soient $u, v \in E$ et soit $\lambda \in \mathbb{R}$. Alors, pour tout $w \in E$, nous avons

$$\begin{aligned} \langle \psi^*(u + \lambda v), w \rangle &= \langle u + \lambda v, \psi(w) \rangle \\ &= \langle u, \psi(w) \rangle + \lambda \langle v, \psi(w) \rangle \\ &= \langle \psi^*(u), w \rangle + \lambda \langle \psi^*(v), w \rangle \\ &= \langle \psi^*(u) + \lambda \psi^*(v), w \rangle \end{aligned}$$

On en déduit que le vecteur $\psi^*(u + \lambda v) - [\psi^*(u) + \lambda \psi^*(v)]$ est dans le noyau $E^\perp = \operatorname{Ker} \theta$, qui est nul. Donc $\psi^*(u + \lambda v) = \psi^*(u) + \lambda \psi^*(v)$.

Pour l'unicité, supposons qu'il existe une application linéaire $\widetilde{\psi} : E \to E$ telle que $\forall u, v \in E,\ \langle u, \psi(v) \rangle = \langle \widetilde{\psi}(u), v \rangle$. Nous savons déjà que ψ^* satisfait à (**??**).

Alors, pour tous $u, v \in E$, nous avons :

$$\begin{aligned} \langle \psi^*(u) - \widetilde{\psi}(u), v \rangle &= \langle \psi^*(u), v \rangle - \langle \widetilde{\psi}(u), v \rangle \\ &= \langle u, \psi(v) \rangle - \langle u, \psi(v) \rangle \\ &= 0 \end{aligned}$$

Ainsi $\psi^*(u) - \widetilde{\psi}(u)$ est dans le noyau $E^\perp$ pour tout u, donc $\psi^*(u) - \widetilde{\psi}(u) = 0$ pour tout $u \in E$, et finalement $\widetilde{\psi} = \psi^*$.

Définition-Proposition C.13. Soit $(E, \langle , \rangle)$ un espace euclidien et soit ψ un endomorphisme de E. Alors il existe un unique endomorphisme ψ^* de E tel que

$$\forall u, v \in E,\ \langle u, \psi(v) \rangle = \langle \psi^*(u), v \rangle.$$

L'endomorphisme ψ^* est appelé *endomorphisme adjoint* de ψ.

Propriétés C.14. Soit $(E, \langle , \rangle)$ un espace euclidien et soit ψ un endomorphisme de E. Alors :

(i) Si $\mathcal{B}$ est une base orthonormale de E, alors la matrice de ψ^* dans la base $\mathcal{B}$ est égale à la transposée de la matrice de ψ dans la base $\mathcal{B}$.

(ii) $(\lambda \psi + \mu \chi)^* = \lambda \psi^* + \mu \chi^*$ pour des endomorphismes ψ et χ et des scalaires λ, μ.

(iii) $(\psi^*)^* = \psi$.

(iv) $(\psi \circ \chi)^* = \chi^* \circ \psi^*$.

(v) $\forall u, v \in E,\ \langle \psi(u), v \rangle = \langle u, \psi^*(v) \rangle$.

Démonstration. Exercice (TD) □

C.4 Isométries d'un espace euclidien

Définition C.15. *Soient $(E, \langle,\rangle_E)$ et $(E', \langle,\rangle_{E'})$ deux espaces euclidiens. Une isométrie de $(E, \langle,\rangle_E)$ dans $(E', \langle,\rangle_{E'})$ est une application linéaire $\Phi : E \to E'$ vérifiant :*

$$\forall u \in E,\ \|\Phi(u)\|_{E'} = \|u\|_E$$

Si $E = E'$, on appelle aussi Φ un endomorphisme orthogonal de E.

Remarque C.16. Par définition, une isométrie est une application linéaire qui préserve les distances (normes). C'est également une application linéaire qui préserve les angles : une application linéaire $\Phi : E \to E'$ est une isométrie si et seulement si $\langle \Phi(u), \Phi(v)\rangle_{E'} = \langle u, v\rangle_E$ pour tous $u, v \in E$.

En effet, supposons que Φ préserve les angles. Alors

$$\|\Phi(u)\| = \sqrt{\langle \Phi(u), \Phi(u)\rangle} = \sqrt{\langle u, u\rangle} = \|u\|$$

pour tout $u \in E$.

Réciproquement, supposons que Φ préserve les normes. Alors

$$2\langle \Phi(u), \Phi(v)\rangle = \|\Phi(u+v)\|^2 - \|\Phi(u)\|^2 - \|\Phi(v)\|^2 = \|u+v\|^2 - \|u\|^2 - \|v\|^2 = 2\langle u, v\rangle.$$

Proposition C.17. Soit Φ une isométrie de $(E, \langle,\rangle_E)$ dans $(E', \langle,\rangle_{E'})$. Alors Φ est injective. En particulier, si Φ est une isométrie de $(E, \langle,\rangle_E)$ dans $(E, \langle,\rangle_E)$ (on dit que Φ est une isométrie de E), alors elle est bijective.

Démonstration. Si $u \in E$ est tel que $\Phi(u) = 0$, alors $\|u\| = \|\Phi(u)\| = \|0\| = 0$, donc $u = 0$. Ainsi le noyau de Φ est nul.

La fin découle du théorème du rang. □

Théorème C.18. Tout espace euclidien de dimension n est isométrique à l'espace euclidien canonique $\mathbb{R}^n$ (c'est-à-dire qu'il existe une isométrie (bijective) entre les deux).

Démonstration. Soit $(E, \langle,\rangle)$ un espace euclidien, et soit $\mathcal{B} = \{e_1, \dots, e_n\}$ une base orthonormale de E. Soit $\{\varepsilon_1, \dots, \varepsilon_n\}$ la base canonique de $\mathbb{R}^n$ (où ε_i est le vecteur colonne qui a un 1 à la $i^{\text{ème}}$ place et des 0 ailleurs). On définit un isomorphisme $\Phi : E \to \mathbb{R}^n$ en posant $\Phi(e_i) = \varepsilon_i$ et en prolongeant par linéarité. Il faut montrer que c'est une isométrie. Soit donc $u = \sum_{i=1}^n x_i e_i$; on a

$$\|\Phi(u)\|^2_{\mathbb{R}^n} = \langle \Phi(u), \Phi(u)\rangle_{\mathbb{R}^n} = \langle \sum_{i=1}^n x_i\varepsilon_i, \sum_{j=1}^n x_j\varepsilon_j\rangle_{\mathbb{R}^n} = \sum_{i=1}^n\sum_{j=1}^n x_i x_j \langle \varepsilon_i, \varepsilon_j\rangle_{\mathbb{R}^n} = \sum_{i=1}^n x_i^2$$

d'une part, et

$$\|u\|^2_E = \langle u, u\rangle_E = \langle \sum_{i=1}^n x_i e_i, \sum_{j=1}^n x_j e_j\rangle_E = \sum_{i=1}^n\sum_{j=1}^n x_i x_j \langle e_i, e_j\rangle_E = \sum_{i=1}^n x_i^2$$

d'autre part, d'où l'égalité $\|\Phi(u)\|_{\mathbb{R}^n} = \|u\|_E$. □

Propriétés C.19. Soit $(E, \langle,\rangle)$ un espace euclidien.

(i) Soit Φ une isométrie de E. Alors Φ^{-1} est une isométrie de E.

(ii) Si Φ et Ψ sont des isométries de E alors $\Psi \circ \Phi$ est une isométrie de E.

(iii) id_E est une isométrie de E.

Démonstration. Exercice (TD). □

Nous avons donc :

Définition-Proposition C.20. L'ensemble des isométries de l'espace euclidien $(E, \langle,\rangle_E)$ est un sous-groupe de $\mathrm{GL}(E)$. On l'appelle le *groupe orthogonal* et on le note $\mathrm{O}(E)$.
L'ensemble des isométries de E dont le déterminant est positif est également un sous-groupe de $\mathrm{GL}(E)$. On l'appelle le *groupe spécial orthogonal* et on le note $\mathrm{SO}(E)$.
Si $E = \mathbb{R}^n$ muni de sa structure canonique d'espace euclidien, on note $\mathrm{O}(E) = \mathrm{O}_n(\mathbb{R})$ et $\mathrm{SO}(E) = \mathrm{SO}_n(\mathbb{R})$.

Démonstration. Pour terminer la démonstration, il suffit de remarquer que le déterminant est multiplicatif (c'est-à-dire que $\det(\Psi \circ \Phi) = \det \Psi \det \Phi$) et on en déduit facilement que la propriété d'avoir un déterminant positif est conservée par toutes les opérations dans un sous-groupe de $\mathrm{GL}(E)$. □

Proposition C.21. Soit $(E, \langle,\rangle)$ un espace euclidien et soit $\Phi : E \to E$ une application linéaire. Alors les assertions suivantes sont équivalentes :

(i) Φ est une isométrie.

(ii) $\Phi^* = \Phi^{-1}$.

(iii) Si $\mathcal{B}$ est une base orthonormale de E et si M est la matrice de Φ dans la base $\mathcal{B}$, alors $tMM = I_n = MtM$.

(iv) Si $\mathcal{B}$ est une base orthonormale de E et si M est la matrice de Φ dans la base $\mathcal{B}$, alors M est inversible et $M^{-1} = tM$.

(v) Si $\mathcal{B}$ est une base orthonormale de E et si M est la matrice de Φ dans la base $\mathcal{B}$, alors les vecteurs colonne de M forment une base orthonormale de $\mathbb{R}^n$ avec sa structure euclidienne canonique.

(vi) Si $\mathcal{B}$ est une base orthonormale de E, alors $\Phi(\mathcal{B})$ est une base orthonormale de E.

Lorsque les conditions équivalentes (iii)-(v) ci-dessus sont vérifiées pour une matrice M, on dit que M est une *matrice orthogonale.*

Démonstration. (TD) □

Corollaire C.22. (a) Soit Φ une isométrie de E. Alors

(i) $|\det \Phi| = 1$

(ii) si λ est valeur propre de Φ alors $|\lambda| = 1$.

(b) Soit M une matrice. Alors M est orthogonale si et seulement si M est la matrice de passage d'une base orthonormale à une autre.

Démonstration. (a) (i) D'après le (iii) de la proposition précédente, si M est la matrice de Φ dans une base orthonormale,

$$1 = \det I_n = \det(tMM) = \det(tM)\det(M) = \det(M)^2$$

donc $\det(\Phi) = \det(M) = \pm 1$.

(ii) Soit λ une valeur propre de Φ, et soit u un vecteur propre associé. Alors

$$\|u\| = \|\Phi(u)\| = \|\lambda u\| = |\lambda|\ \|u\|$$

donc $|\lambda| = 1$ et $\lambda = \pm 1$.

(b) Soit M une matrice orthogonale. Fixons une base orthonormale $\mathcal{B}$ de E. Alors M est la matrice d'une transformation Φ dans la base $\mathcal{B}$, et d'après la proposition précédente (vi), $\Phi(\mathcal{B})$ est une base orthonormale, c'est-à-dire que M est la matrice de passage d'une base orthonormale à une autre.

Maintenant, si P est la matrice de passage d'une base orthonormale $\mathcal{B} = \{e_1, \ldots, e_n\}$ à une base orthonormale $\mathcal{B}' = \{e_1, \ldots, e_n\}$, définissons $\Phi : E \to E$ en posant $\Phi(e_i) = e'_i$ et en prolongeant par linéarité. Alors $\mathcal{B}' = \Phi(\mathcal{B})$ donc Φ est une isométrie de E d'après le (vi) de la proposition précédente. De plus, P est la matrice de Φ dans la base $\mathcal{B}$ (puisque chaque colonne est obtenue en écrivant $e'_i = \Phi(e_i)$ dans la base $\mathcal{B}$), donc d'après la proposition précédente, P est orthogonale. $\square$

Remarque C.23. Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases orthonormales, et soient M et M' les matrices d'un endomorphisme de E dans les bases $\mathcal{B}$ et $\mathcal{B}'$ respectivement. Soit P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Alors $M' = P^{-1}MP = tPMP$.

C.5 Isométries de $\mathbb{R}^2$

Nous considérons $\mathbb{R}^2$ muni de sa structure canonique d'espace euclidien. Soit $\Phi \in \mathrm{O}_2(\mathbb{R})$. Dans la base canonique, la matrice de Φ s'ecrit $M = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. Les vecteurs $\begin{pmatrix} a \\ b \end{pmatrix}$ et $\begin{pmatrix} c \\ d \end{pmatrix}$ forment une base orthonormale, et nous avons vu que $\det M = \pm 1$, donc nous obtenons les équations :

$$\begin{cases} a^2 + b^2 = 1 \\ c^2 + d^2 = 1 \\ ac + bd = 0 \\ ad - bc = \varepsilon \text{ avec } \varepsilon = \pm 1. \end{cases}$$

Supposons d'abord que $a \neq 0$. Alors la troisième équation donne $c = -\frac{bd}{a}$. De la quatrième équation on déduit que $d = a\varepsilon$, donc $c = -\frac{bd}{a} = -b\varepsilon$. La deuxième équation est automatiquement satisfaite. Ainsi,

$$M = \begin{pmatrix} a & -b\varepsilon \\ b & a\varepsilon \end{pmatrix}.$$

Supposons maintenant que $a = 0$. Le système devient donc :

$$\begin{cases} b^2 = 1 \\ c^2 + d^2 = 1 \\ bd = 0 \\ -bc = \varepsilon \text{ avec } \varepsilon = \pm 1. \end{cases}$$

Comme $b = \pm 1$, la dernière équation donne $c = -\varepsilon b$. La troisième donne $d = 0 = \varepsilon a$. Ainsi la matrice M est encore de la forme ci-dessus.

Comme $a^2 + b^2 = 1$, il existe $\alpha \in \mathbb{R}$ tel que $a = \cos\alpha$ et $b = \sin\alpha$. Donc $M = \begin{pmatrix} \cos\alpha & -\varepsilon\sin\alpha \\ \sin\alpha & \varepsilon\cos\alpha \end{pmatrix}$.

Si le déterminant de M est égal à 1, c'est-à-dire si $M \in \mathrm{SO}_2(\mathbb{R})$, alors $\varepsilon = 1$ et Φ est la rotation de centre O et d'angle α (l'expression en nombres complexes de la transformation est $z \mapsto e^{i\alpha}z$).

Si le déterminant de M est égal à -1, alors $\varepsilon = -1$ et le polynôme caractéristique de Φ est égal à $x^2 - 1 = (x-1)(x+1)$. Donc Φ admet deux valeurs propres réelles, les sous-espaces propres sont nécessairement de dimension 1, et $\left\{ \begin{pmatrix} \cos\frac{\alpha}{2} \\ \sin\frac{\alpha}{2} \end{pmatrix} ; \begin{pmatrix} -\sin\frac{\alpha}{2} \\ \cos\frac{\alpha}{2} \end{pmatrix} \right\}$ est une base orthonormale de $\mathbb{R}^2$ formée de vecteurs propres pour Φ. Dans cette base, la matrice de Φ s'écrit $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, donc Φ est la symétrie orthogonale par rapport à la droite engendrée par $\begin{pmatrix} \cos\frac{\alpha}{2} \\ \sin\frac{\alpha}{2} \end{pmatrix}$.

Résumons tout cela :

Proposition C.24. Soit Φ une isométrie de $\mathbb{R}^2$. Alors sa matrice dans la base canonique de $\mathbb{R}^2$ est de la forme

$$\begin{pmatrix} \cos\alpha & -\varepsilon\sin\alpha \\ \sin\alpha & \varepsilon\cos\alpha \end{pmatrix}$$

pour un $\alpha \in \mathbb{R}$, avec $\varepsilon = \pm 1$.
Φ est dans $\mathrm{SO}(\mathbb{R}^2)$ si et seulement si Φ est la rotation de centre O et d'angle α.
Φ est dans $\mathrm{O}(\mathbb{R}^2) \setminus \mathrm{SO}(\mathbb{R}^2)$ si et seulement si Φ est la symétrie orthogonale par rapport à la droite engendrée par $\begin{pmatrix} \cos\frac{\alpha}{2} \\ \sin\frac{\alpha}{2} \end{pmatrix}$.

C.6 Isométries de $\mathbb{R}^3$

Proposition C.25. Si $\Phi \in \mathrm{O}_3(\mathbb{R})$, il s'agit d'une rotation de $\mathbb{R}^3$ éventuellement suivie d'une symétrie orthogonale. Il existe une base orthonormale de $\mathbb{R}^3$ dans laquelle sa matrice est de la forme

$$\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & \pm 1 \end{pmatrix}.$$

Démonstration. en TD. □

Remarque C.26. Afin de déterminer α au signe près, remarquons que

$$\operatorname{tr}\Phi = \begin{cases} 1 + 2\cos\alpha & \text{si } \det\Phi = 1 \\ -1 + 2\cos\alpha & \text{si } \det\Phi = -1. \end{cases}$$

D Le théorème spectral des endomorphismes symétriques ; application aux coniques et quadriques

Définition D.1. *Un endomorphisme ψ de E est dit symétrique ou auto-adjoint si $\psi^* = \psi$.*

Remarque D.2. Un endomorphisme est symétrique si et seulement si sa matrice dans une (ou dans toute) base *orthonormale* est symétrique (c'est-à-dire égale à sa transposée).

Proposition D.3. Soit ψ un endomorphisme symétrique, et soient λ et μ des valeurs propres de ψ avec $\lambda \neq \mu$. Soit u un vecteur propre associé à la valeur propre λ et soit v un vecteur propre associé à la valeur propre μ. Alors u et v sont orthogonaux.

Démonstration. $(\lambda - \mu)\langle u, v\rangle = \langle \lambda u, v\rangle - \langle u, \mu v\rangle = \langle \psi(u), v\rangle - \langle u, \psi(v)\rangle = \langle \psi(u), v\rangle - \langle \psi(u), v\rangle = 0$ puisque ψ est symétrique. Comme $\lambda - \mu \neq 0$, on a bien $\langle u, v\rangle = 0$. □

Théorème D.4. (Théoreme spectral des endomorphismes symétriques) Soit $(E, \langle , \rangle)$ un espace euclidien et soit ψ un endomorphisme symétrique. Alors ψ est diagonalisable dans une base *orthonormale*. Autrement dit, il existe une base orthonormale de E formée de vecteurs propres de ψ.

Démonstration. Nous la verrons à la fin du Chapitre II. □

Du théorème précédent, on déduit :

Corollaire D.5. (Théoreme spectral pour les matrices symétriques) Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique. Alors il existe une matrice orthogonale P telle que ${}^tPAP$ soit diagonale.

Démonstration. La matrice A représente un endomorphisme symétrique dans la base canonique (orthonormale) de $\mathbb{R}^n$. Donc d'après le théorème spectral ci-dessus, il existe une matrice $P \in \mathcal{M}_n(\mathbb{R})$ inversible telle que $P^{-1}AP$ soit diagonale. De plus, toujours d'après le théorème spectral, cette matrice est la matrice de passage d'une base orthonormale à une autre base orthonormale. Elle est donc bien orthogonale, et $P^{-1} = {}^tP$. □

D.1 Application : étude des hypersurfaces du second degré

Définition D.6. *On se place dans un espace euclidien de dimension n, et on fixe une base orthonormale. On peut donc supposer qu'il s'agit de $\mathbb{R}^n$ avec sa base canonique.*

On appelle hypersurface du second degré l'ensemble $\mathcal{H}$ des points $u \in \mathbb{R}^n$ vérifiant :

$$F(u) := q(u) + \ell(u) + c = 0$$

où q est une forme quadratique non nulle, ℓ est une forme linéaire, et c est une constante (réelle).

Lorsque $n = 2$, une hypersurface de degré 2 s'appelle une conique ; lorsque $n = 3$ une hypersurface de degré 2 s'appelle une quadrique.

On cherche à déterminer la forme de l'hypersurface. Nous commencerons dans le cas général, puis déterminerons la classification complète dans le cas des coniques et des quadriques.

$\boxed{1}$ La première étape est de réduire la forme quadratique q dans une base orthonormale.
Soit A la matrice associée à q, et soit ψ l'endomorphisme de $\mathbb{R}^n$ dont la matrice dans la base canonique est A. La matrice A est symétrique.
Si $u \in \mathbb{R}^n$, soit U le vecteur colonne associé ; alors

$$q(u) = tUAU = t(AU)U = \langle \psi(u), u \rangle.$$

Comme ψ est symétrique, on peut lui appliquer le théorème spectral : il existe donc une base orthonormale $\{\varepsilon_1, \ldots, \varepsilon_n\}$ formée de vecteurs propres pour ψ. Soit λ_i la valeur propre associée au vecteur propre ε_i.
Notons $u = \sum_{i=1}^n x_i \varepsilon_i$. Alors $q(u) = \langle \psi(u), u \rangle = \sum_{i=1}^n \lambda_i x_i^2$, donc $F(u) = \sum_{i=1}^n \lambda_i x_i^2 - \sum_{i=1}^n \tau_i x_i + c$.

$\boxed{2}$ La deuxième étape est de réduire F. Pour cela on "complète les carrés", c'est-à-dire que l'on écrit, lorsque $\lambda_i \neq 0$,

$$\lambda_i x_i^2 - \tau_i x_i = \lambda_i (x_i^2 - \frac{\tau_i}{\lambda_i} x_i) = \lambda_i \left(x_i - \frac{\tau_i}{2\lambda_i} \right)^2 - \frac{\tau_i^2}{4\lambda_i}.$$

Ceci revient à faire un changement d'origine (si $\lambda_1, \ldots, \lambda_r$ sont non nuls et $\lambda_{r+1} = \cdots = \lambda_n = 0$, la nouvelle origine est $\Omega = \left(\frac{\tau_1}{2\lambda_1}, \ldots, \frac{\tau_r}{2\lambda_r}, 0, \ldots, 0 \right)$). L'équation devient donc

$$\sum_{i=1}^r \lambda_i x_i'^2 = \sum_{r+1}^n \tau_i x_i + c'. \tag{D.1}$$

Notons que si l'un des τ_i est non nul, par exemple τ_n, alors on peut faire un nouveau changement d'origine pour que l'équation ne contienne plus de constante ($\Omega' = \Omega + (0, \ldots, 0, -\frac{c'}{\tau_n})$).

$\boxed{3}$ Pour étudier l'hypersurface $\mathcal{H}$, on se placera dans le repère orthonormal privilégié $(\Omega \mid \varepsilon_1, \ldots, \varepsilon_n)$ dans lequel l'équation a l'expression réduite (**??**).

a) Classification des coniques

Soit $\mathcal{C}$ une conique (c'est-à-dire une hypersurface de degré 2 dans un espace euclidien de dimension 2). Notons $\begin{pmatrix} x \\ y \end{pmatrix}$ les coordonnées d'un point dans le repère privilégié $(\Omega \mid \varepsilon_1, \varepsilon_2)$, et λ, μ les valeurs propres de ψ.

- Si q est de rang 2, *i.e.* les valeurs propres sont non nulles, l'équation réduite de $\mathcal{C}$ est $\lambda x^2 + \mu y^2 = \delta$. Notons que quitte à diviser par $|\delta|$, on peut supposer que $\delta = 1, 0$ ou -1. On peut également, quitte à tout multiplier par -1 ou à échanger les rôles de x et y, supposer que $\lambda = \frac{1}{\alpha^2} > 0$. On obtient alors :

Equation réduite	*Nature de la conique*
$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1$	Ellipse
$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 0$	Point Ω
$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = -1$	Vide
$\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} = 0$	Réunion de deux droites sécantes d'équations $y = \pm\frac{\beta}{\alpha}x$
$\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} = 1$	Hyperbole

- Si q est de rang 1, *i.e.* si l'une des valeurs propres est nulle, par exemple μ, l'équation réduite de $\mathcal{C}$ est $\lambda x^2 + \tau y = \rho$. Comme auparavant, nous pouvons supposer que $\lambda = \frac{1}{\alpha^2} > 0$ et $\rho = 0, 1$ ou -1, et même que $\rho = 0$ lorsque $\tau \neq 0$. On obtient alors :

Equation réduite	*Nature de la conique*
$\frac{x^2}{\alpha^2} = 1$	Réunion de deux droites parallèles d'équations $x = \pm\alpha$
$\frac{x^2}{\alpha^2} = 0$	Droite d'équation $x = 0$
$\frac{x^2}{\alpha^2} = -1$	Vide
$\frac{x^2}{\alpha^2} - \frac{y}{\beta} = 0$	Parabole

Exemple D.7. Considérons la conique $\mathcal{C}$ d'équation $x^2 + 10\sqrt{3}xy + 11y^2 - 16 = 0$. Nous voulons déterminer sa nature.

La forme quadratique est $x^2 + 10\sqrt{3}xy + 11y^2$. La matrice de ψ dans la base canonique est donc

$$A = \begin{pmatrix} 1 & 5\sqrt{3} \\ 5\sqrt{3} & 11 \end{pmatrix}.$$

Nous voulons trouver ses valeurs propres et ses vecteurs propres.

Le polynôme caractéristique est

$$\begin{vmatrix} 1-\lambda & 5\sqrt{3} \\ 5\sqrt{3} & 11-\lambda \end{vmatrix} = \lambda^2 - 12\lambda - 64 = (\lambda - 16)(\lambda + 4)$$

donc les valeurs propres sont 16 et -4, non nulles.

Le noyau de $A + 4I = 5\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & 3 \end{pmatrix}$ est engendré par $\begin{pmatrix} 3 \\ -\sqrt{3} \end{pmatrix}$ et le noyau de $A - 16I = 5\begin{pmatrix} -3 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$ est engendré par $\begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix}$. Donc nous obtenons un repère orthonormal en normalisant ces vecteurs :

$$\left(O \mid \begin{pmatrix} \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{pmatrix}; \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}\right).$$

Dans ce repère, $\mathcal{C}$ a pour équation $16X^2 - 4Y^2 = 16$, ou

$$X^2 - \frac{Y^2}{4} = 1.$$

Il s'agit donc d'une hyperbole.

b) Classification des quadriques

Soit $\mathcal{Q}$ une quadrique (c'est-à-dire une hypersurface de degré 2 dans un espace euclidien de dimension 3). Notons x, y, z les coordonnées dans le repère privilégié $(\Omega \mid \varepsilon_1, \varepsilon_2, \varepsilon_3)$ et λ, μ, ν les valeurs propres de ψ.

- Si q est de rang 3, *i.e.* si les valeurs propres sont toutes non nulles, alors l'équation réduite de $\mathcal{Q}$ est $\lambda x^2 + \mu y^2 + \nu z^2 = \delta$. Comme dans le cas des coniques, on peut supposer que $\lambda = \frac{1}{\alpha^2} > 0$ et $\mu = \frac{1}{\beta^2} > 0$, et que $\delta = 0, 1$ ou -1. On obtient alors :

Equation réduite	*Nature de la quadrique*
$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} + \frac{z^2}{\gamma^2} = 1$	Ellipsoïde
$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - \frac{z^2}{\gamma^2} = 1$	Hyperboloïde à une nappe
$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - \frac{z^2}{\gamma^2} = -1$	Hyperboloïde à deux nappes
$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - \frac{z^2}{\gamma^2} = 0$	Cône du second degré à base elliptique
$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} + \frac{z^2}{\gamma^2} = 0$	Point Ω
$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} + \frac{z^2}{\gamma^2} = -1$	Vide

- Si q est de rang 2, *i.e.* si exactement une des valeurs propres, par exemple ν, est nulle, l'équation réduite de $\mathcal{Q}$ est $\lambda x^2 + \mu y^2 = \tau z + \delta$. Lorsque $\tau \neq 0$, on peut supposer que $\delta = 0$. Donc on suppose que $\tau = 1$ et $\delta = 0$ ou que $\tau = 0$ et que $\delta = 0, 1$ ou -1. On obtient alors :

Equation réduite	*Nature de la quadrique*
$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = z$	Paraboloïde elliptique
$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1$	Cylindre à base elliptique
$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 0$	Droite d'équations $x = 0;\ y = 0$
$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = -1$	Vide
$\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} = z$	Paraboloïde hyperbolique
$\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} = 1$	Cylindre à base hyperbolique
$\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} = 0$	Réunion de deux plans d'équations $y = \pm\frac{\beta}{\alpha}x$

- Si q est de rang 1, *i.e.* si exactement deux valeurs propres sont nulles, disons μ et ν, l'équation réduite est $\frac{x^2}{\alpha^2} = \tau y + \sigma z + \omega$. Comme auparavant, on peut supposer que lorsque τ ou σ est non nul alors $\omega = 0$, et que $\omega = 0, 1$ ou -1 sinon. De plus, lorsque $\tau \neq 0$ et $\sigma \neq 0$, on peut changer la variable y pour mettre l'équation sous la forme $\lambda x^2 + \tau y' = 0$ avec $y' = (y + \frac{\sigma}{\tau}z)$ (en supposant déjà que $\omega = 0$). On obtient :

Equation réduite	*Nature de la quadrique*
$x^2 = 2p$	Cylindre à base parabolique
$\frac{x^2}{\alpha^2} = 1$	Réunion de deux plans parallèles distincts d'équations $x = \pm\alpha$
$\frac{x^2}{\alpha^2} = 0$	Plan d'équation $x = 0$
$\frac{x^2}{\alpha^2} = -1$	Vide

Exemple D.8. Considérons une quadrique $\mathcal{Q}$ d'équation $-7x^2 + 25y^2 + 7z^2 + 48xz + 5x - k = 0$.

La matrice associée à la forme quadratique est $A = \begin{pmatrix} -7 & 0 & 24 \\ 0 & 25 & 0 \\ 24 & 0 & 7 \end{pmatrix}$. Ses valeurs propres sont -25 (simple) et 25 (double). On choisit une base orthonormale du plan $\mathrm{Ker}(A - 25I)$ et on complète avec un vecteur normal ε_3 de la droite $\mathrm{Ker}(A + 25I)$. On obtient donc un nouveau repère orthonormal $(0, \varepsilon_1 = \begin{pmatrix} \frac{3}{5} \\ 0 \\ \frac{4}{5} \end{pmatrix}; \varepsilon_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}; \varepsilon_3 = \begin{pmatrix} \frac{4}{5} \\ 0 \\ -\frac{3}{5} \end{pmatrix})$ par exemple. L'équation de $\mathcal{Q}$ devient $25X^2 + 25Y^2 - 25Z^2 + 3X + 4Z = k$, soit

$$25X'^2 + 25Y'^2 - 25Z'^2 = k - \frac{7}{100}$$

en posant $X' = X + \frac{3}{50}$, $Y' = Y$ et $Z' = Z - \frac{4}{50}$, c'est-à-dire que l'on a pris $\Omega = \left(-\frac{3}{50}; 0; \frac{2}{25}\right)$ dans le repère $(O \mid \varepsilon_1, \varepsilon_2, \varepsilon_3)$.

Donc,

- si $k > \frac{7}{100}$, alors $\mathcal{Q}$ est un hyperboloïde à une nappe ;
- si $k < \frac{7}{100}$, alors $\mathcal{Q}$ est un hyperboloïde à deux nappes ;
- si $k = \frac{7}{100}$, alors $\mathcal{Q}$ est un cône à base circulaire (les coefficients de X' et Y' sont égaux).

II Formes sesquilinéaires hermitiennes

A Introduction, généralités

Le but de ce chapitre est d'adapter la théorie des formes bilinéaires symétriques au cas d'espaces vectoriels sur $\mathbb{C}$.

On a sur $\mathbb{R}^n$ un produit scalaire canonique : si $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, on a $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$.
Ce produit scalaire définit une norme sur $\mathbb{R}^n$ par $\|x\| = \sqrt{\sum_{i=1}^n x_i^2}$.

Sur $\mathbb{C}^n$, l'application bilinéaire symétrique donnée par $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$ ne définit pas une norme, car x_i^2 n'est en général pas un réel positif.

Mais $\mathbb{C}^n$ s'identifie canoniquement à $\mathbb{R}^{2n}$, et sur $\mathbb{R}^{2n}$ on a une norme canonique :

$$\begin{array}{ccc} \mathbb{C}^n & \xrightarrow{\sim} & \mathbb{R}^{2n} \\ z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} & \mapsto & x = \begin{pmatrix} x_1 \\ y_1 \\ \vdots \\ y_n \\ x_n \end{pmatrix} \end{array} \text{ si } z_k = x_k + i y_k.$$

On a

$$\|x\| = \sqrt{\sum_{k=1}^n (x_k^2 + y_k^2)} = \sqrt{\sum_{k=1}^n |z_k|^2} = \sqrt{\sum_{k=1}^n z_k \cdot \bar{z}_k}.$$

Donc $\|x\| = \sqrt{f(z,z)}$ avec $f(z,z') = \sum_{k=1}^n z_k \cdot \bar{z}'_k$. L'application $f : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ n'est plus bilinéaire symétrique. Elle vérifie les propriétés suivantes :

(i) f est linéaire en la première variable : $f(z + \lambda z', z'') = f(z, z'') + \lambda f(z', z'')$ pour tous $z, z', z'' \in \mathbb{C}^n$ et tout $\lambda \in \mathbb{C}$;

(ii) $f(z', z) = \overline{f(z, z')}$ pour tous $z, z' \in \mathbb{C}^n$;

(iii) f est additive en la deuxième variable : $f(z, z' + z'') = f(z, z') + f(z, z'')$ pour tous $z, z', z'' \in \mathbb{C}^n$;

(iv) f est $\mathbb{C}$-antilinéaire en la deuxième variable : $f(z, \lambda z') = \bar{\lambda} f(z, z')$ pour tous $z, z' \in \mathbb{C}^n$ et tout $\lambda \in \mathbb{C}$.

Remarque A.1. Les propriétés (iii) et (iv) découlent des propriétés (i) et (ii).

Définition A.2. *Soit H un espace vectoriel sur $\mathbb{C}$ de dimension finie n. Une forme sesquilinéaire hermitienne sur H est une application $f : H \times H \to \mathbb{C}$ linéaire en la première variable et vérifiant $f(z, z') = \overline{f(z', z)}$ pour tous $z, z' \in H$.*

Remarque A.3. Soit $z \in H$. Alors $f(z, z) = \overline{f(z, z)}$, donc $f(z, z) \in \mathbb{R}$.

Remarque A.4. Dans le cas réel, il y a une bijection entre l'ensemble des formes bilinéaires symétriques et l'ensemble des formes quadratiques : si f est une forme bilinéaire symétrique, alors $x \mapsto f(x, x)$ est une forme quadratique ; si q est une forme quadratique, alors sa forme polaire $(x, y) \mapsto \frac{1}{2}[q(x + y) - q(x) - q(y)]$ est une forme bilinéaire symétrique.

Dans le cas complexe, on peut également retrouver une forme sesquilinéaire hermitienne à partir de sa restriction à la diagonale de $H \times H$:

$$f(z,z') = \frac{1}{2}[f(z+z',z+z') + if(z+iz',z+iz') - (1+i)(f(z,z)+f(z',z'))].$$

Expression dans une base : Soit H un espace vectoriel sur $\mathbb{C}$ et soit $\mathcal{B} = \{e_1,\ldots,e_n\}$ une base de H. Soient $x = \sum_{i=1}^n x_i e_i$ et $y = \sum_{i=1}^n y_i e_i$ deux éléments de H. Alors

$$f(x,y) = f(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j) = \sum_{i,j=1}^n x_i \overline{y}_j f(e_i,e_j).$$

On considère la *matrice associée à f*, $M = (f(e_i,e_j))_{i,j}$ et alors

$$f(x,y) = tXM\overline{Y} \tag{A.1}$$

si $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$.

On constate que $tM = \overline{M}$, et que la donnée d'une matrice A telle que $tA = \overline{A}$ définit une forme hermitienne sur H par la formule (**??**).

Définition A.5. *Une matrice M est dite hermitienne si elle vérifie $tM = \overline{M}$.*

En particulier, une matrice réelle symétrique est hermitienne.

B Espaces vectoriels hermitiens

Proposition B.1. Soit f une forme sesquilinéaire hermitienne. L'application

$$\begin{array}{rcl} \theta_f : H & \to & H^* \\ z & \mapsto & \theta_f(z) = f(\cdot,z) \end{array}$$

est une application $\mathbb{C}$-antilinéaire (*i.e* elle verifie $\theta_f(a+b) = \theta_f(a) + \theta_f(b)$ et $\theta_f(\lambda a) = \overline{\lambda}\theta_f(a)$ pour tous $a,b \in H$ et tout $\lambda \in \mathbb{C}$).

Remarque B.2. Le théorème du rang est valable pour les endomorphismes $\mathbb{C}$-antilinéaires.

Démonstration. Elle est analogue au cas réel. □

Définition B.3. *De la même façon que dans le cas réel, le rang d'une forme sesquilinéaire est le rang de l'application θ_f associée. Il est égal au rang de sa matrice dans une base de H.*

On dit qu'une forme sesquilinéaire hermitienne f est :

- *non dégénérée si sa matrice est de rang n. Dans ce cas, l'application θ_f est bijective (théorème du rang pour les applications $\mathbb{C}$-antilinéaires).*
- *définie positive si $f(z,z) > 0$ pour tout $z \neq 0$ dans H.*

Un produit (scalaire) hermitien est une forme sesquilinéaire hermitienne définie positive. Il est noté $\langle,\rangle$.

Un espace hermitien est un espace vectoriel sur $\mathbb{C}$ muni d'un produit scalaire hermitien.

Remarque B.4. Si f est définie positive, alors elle est non dégénérée. En effet, si f est définie positive, on voit facilement que θ_f est injective, donc surjective (théorème du rang) et donc f est bien de rang maximal.

Définition-Proposition B.5. Dans un espace hermitien H, on peut définir une *norme* par $\|z\| = \sqrt{\langle z, z\rangle}$ pour tout $z \in H$. Alors $\|\cdot\| : H \to \mathbb{R}$ vérifie :

$$\begin{cases} \|z\| \geqslant 0 \\ \|z\| = 0 \iff z = 0 \\ \|\lambda z\| = |\lambda|\ \|z\| \\ \|z + z'\| \leqslant \|z\| + \|z'\| \end{cases}$$

Démonstration. Les trois premières propriétés découlent de la définition d'un produit scalaire hermitien. La dernière découle comme dans le cas réel de l'inégalité de Cauchy-Schwarz qui va suivre. □

Proposition B.6. (Inégalité de Cauchy-Schwarz) Soit $(H, \langle,\rangle)$ un espace hermitien. Pour tous z, z' dans H, on a

$$|\langle z, z'\rangle| \leqslant \|z\|\ \|z'\|.$$

Démonstration. Pour chaque nombre complexe λ, nous avons

$$\begin{aligned} 0 \leqslant \|z + \lambda z'\|^2 &= \langle z + \lambda z', z + \lambda z'\rangle \\ &= \langle z, z\rangle + \lambda\langle z', z\rangle + \overline{\lambda}\langle z, z'\rangle + |\lambda|^2 \langle z', z'\rangle \\ &= \|z\|^2 + 2\Re(\lambda\langle z', z\rangle) + |\lambda|^2 \|z'\|^2. \end{aligned}$$

Ceci est vrai pour tout $\lambda \in \mathbb{C}$, donc si $z' \neq 0$ c'est vrai en particulier pour $\lambda = -\frac{\langle z, z'\rangle}{\|z'\|^2}$. Comme $\langle z, z'\rangle\langle z', z\rangle = \langle z, z'\rangle\overline{\langle z, z'\rangle} = |\langle z, z'\rangle|^2$, nous avons

$$0 \leqslant \|z\|^2 + 2\Re\left(-\frac{|\langle z, z'\rangle|^2}{\|z'\|^2}\right) + \frac{|\langle z, z'\rangle|^2}{\|z'\|^4}\|z'\|^2 = \|z\|^2 - \frac{|\langle z, z'\rangle|^2}{\|z'\|^2}.$$

Par conséquent,

$$|\langle z, z'\rangle| \leqslant \|z\|\ \|z'\|.$$

De plus, si $z' = 0$ le résultat est évident. □

Propriétés B.7. Citons d'autres propriétés de la norme :

(i) $\|z + z'\|^2 = \|z\|^2 + \|z\|^2 + 2\Re(\langle z, z'\rangle)$.

(ii) $\langle z, z'\rangle = \frac{1}{2}[\|z + z'\|^2 + i\,\|z + iz'\|^2 - (1 + i)(\|z\|^2 + \|z'\|^2)]$ *(polarisation).*

Démonstration. Exercice (TD). □

Exemples B.8. • Si $H = \mathbb{C}^n$, $\langle z, z'\rangle = \sum_{i=1}^n z_i\overline{z}_i'$ est un produit hermitien, dit canonique. La norme associée est $\|z\| = \sqrt{\sum_{i=1}^n |z_i|^2}$. L'inégalité de Cauchy-Schwarz donne donc $|\sum_{i=1}^n z_i\overline{z}_i'|^2 \leqslant \left(\sum_{i=1}^n |z_i|^2\right)\left(\sum_{i=1}^n |z_i'|^2\right)$

• Si H est l'espace vectoriel des polynômes complexes de degré inférieur ou égal à $n-1$, alors $\langle p, q\rangle = \int_0^1 p(t)\overline{q(t)}dt$ définit un produit hermitien. On voit par exemple, en prenant $q \equiv 1$ dans Cauchy-Schwarz, que $\left|\int_0^1 p(t)dt\right|^2 \leqslant \int_0^1 |p(t)|^2\, dt$.

B.1 Orthogonalité et bases orthogonales

Définition B.9. *Soit $(H, \langle,\rangle)$ un espace hermitien, et soient z et z' dans H. On dit que z et z' sont orthogonaux si $\langle z, z'\rangle = 0$.*

Remarque B.10. Si z et z' sont orthogonaux, on a l'égalité de Pythagore

$$\left\|z + z'\right\|^2 = \|z\|^2 + \left\|z'\right\|^2.$$

En revanche, contrairement au cas réel, la réciproque est fausse. Par exemple, $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} i \\ 0 \end{pmatrix}$ ne sont pas orthogonaux mais satisfont à l'identité ci-dessus.

On définit comme dans le cas réel des familles orthogonales, orthonormales, et des bases orthogonales et orthonormales.

Propriété B.11. Soit $z \in H$. Si $\{e_1, \ldots, e_n\}$ est une base orthonormale de H, alors

$$z = \sum_{i=1}^{n} \langle z, e_i\rangle e_i.$$

Attention, le z est placé dans la première composante du produit hermitien.

Définition-Proposition B.12. Soit $(H, \langle,\rangle)$ un espace hermitien. Soit A une partie de H. On définit l'*orthogonal* $A^\perp$ de A dans H comme dans le cas réel, par

$$A^\perp = \{z \in H \mid \forall a \in A,\ \langle z, a\rangle = 0\}.$$

C'est un sous-espace vectoriel de H. On a encore :

(i) $A \subseteq B \Rightarrow A^\perp \supseteq B^\perp$

(ii) $A^\perp = \text{vect}\{A\}^\perp$

(iii) $\{0\}^\perp = H$ et $H^\perp = \{0\}$

(iv) $\dim F + \dim F^\perp = \dim H$ pour tout sous-espace vectoriel F de H

(v) $H = F \oplus F^\perp$ pour tout sous-espace vectoriel F de H

(vi) $(F^\perp)^\perp$ pour tout sous-espace vectoriel F de H

(vii) $(F + G)^\perp = F^\perp \cap G^\perp$ pour tous sous-espaces vectoriels F et G de H

(viii) $(F \cap G)^\perp = F^\perp + G^\perp$ pour tous sous-espaces vectoriels F et G de H.

Démonstration. Comme dans le cas réel pour (i) à (iv) et pour (vi). (*cf.* TD).

(v) Soit $x \in F \cap F^\perp$: alors $\langle x, x\rangle = 0$ donc $x = 0$. Avec (iv), on a bien $F \oplus F^\perp = E$.

(vii)

- Soit $z \in F^\perp \cap G^\perp$. Soit $x = u + v \in F + G$ avec $u \in F$ et $v \in G$. Alors $\langle z, x\rangle = \langle z, u\rangle + \langle z, v\rangle = 0 + 0 = 0$, donc $z \in (F + G)^\perp$, d'où l'inclusion $F^\perp \cap G^\perp \subseteq (F + G)^\perp$.
- On a $F \subseteq F + G$ donc $(F + G)^\perp \subseteq F^\perp$. De même, $(F + G)^\perp \subseteq G^perp$. Finalement on a bien $(F + G)^\perp \subseteq F^\perp \cap G^\perp$.

On a donc l'égalité $F^\perp \cap G^\perp = (F + G)^\perp$.

(viii) On applique (vii) à $F^\perp$ et $G^\perp$:

$$F^\perp + G^\perp = \left[\left(F^\perp + G^\perp\right)^\perp\right]^\perp = \left[(F^\perp)^\perp \cap (G^\perp)^\perp\right]^\perp = (F \cap G)^\perp. \quad \square$$

Proposition B.13. Soit $(H, \langle,\rangle)$ un espace hermitien. Alors il existe une base orthonormale pour $\langle,\rangle$.

Démonstration. Même démonstration que dans le cas réel. $\square$

B.2 Adjoint d'un endomorphisme

Dans toute la suite, H est un espace vectoriel sur $\mathbb{C}$ de dimension n.

Définition-Proposition B.14. Soit ψ un endomorphisme de H. Il existe un endomorphisme ψ^* de H et un seul tel que

$$\langle \psi(z), w\rangle = \langle z, \psi^*(w)\rangle \text{ pour tous } z, w \in H.$$

ψ^* est appelé l'*adjoint* de ψ.

Démonstration. Elle est analogue à la démonstration dans le cas réel. $\square$

Propriétés B.15. Soit $(H, \langle,\rangle)$ un espace hermitien et soit ψ un endomorphisme de H. Alors :

(i) Si $\mathcal{B}$ est une base orthonormale de H, alors la matrice de ψ^* dans la base $\mathcal{B}$ est égale à la transposée de la conjuguée de la matrice de ψ dans la base $\mathcal{B}$.

(ii) $(\lambda\psi + \mu\chi)^* = \overline{\lambda}\psi^* + \overline{\mu}\chi^*$ pour des endomorphismes ψ et χ et des scalaires λ et μ.

(iii) $(\psi^*)^* = \psi$.

(iv) $(\psi \circ \chi)^* = \chi^* \circ \psi^*$.

(v) $\forall z, w \in H$, $\langle z, \psi(w)\rangle = \langle \psi^*(z), w\rangle$.

Démonstration. Analogue au cas réel. $\square$

B.3 Endomorphismes hermitiens ou auto-adjoints

Définition-Proposition B.16. Soit M une matrice dans $\mathcal{M}_n(\mathbb{C})$. On appelle *adjointe* de M la matrice $M^* = t\overline{M}$. On dit que M est *hermitienne* ou *auto-adjointe* si $M^* = M$.
Un endomorphisme ψ de H est dit *hermitien* ou *auto-adjoint* si $\psi^* = \psi$. Cela équivaut à dire que la matrice de ψ dans une (ou toute) base *orthonormale* est auto-adjointe.

Démonstration. Analogue au cas réel. $\square$

Proposition B.17. Soit ψ un endomorphisme hermitien. Alors ses valeurs propres sont réelles.

Démonstration. Soit λ une valeur propre pour ψ, et soit z un vecteur propre associé. Alors d'une part $\langle \psi(z), z\rangle = \langle \lambda z, z\rangle = \lambda \|z\|^2$ et d'autre part, puisque ψ est hermitien, $\langle \psi(z), z\rangle = \langle z, \psi^*(z)\rangle = \langle z, \psi(z)\rangle = \langle z, \lambda z\rangle = \overline{\lambda}\|z\|^2$. Comme $z \neq 0$ (c'est un vecteur propre), on a $\|z\| \neq 0$ et donc $\lambda = \overline{\lambda}$. $\square$

B.4 Isométries ou endomorphismes unitaires

Soient $(H, \langle,\rangle)$ et $(H', \langle,\rangle')$ deux espaces hermitiens. Notons $\|\cdot\|$ et $\|\cdot\|'$ les normes associées.

Définition B.18. *Une isométrie de $(H, \langle,\rangle)$ vers $(H', \langle,\rangle')$ est une application linéaire $\Phi : H \to H'$ qui préserve la norme, c'est-à-dire que $\|\Phi(z)\|' = \|z\|$ pour tout $z \in H$. Si $H = H'$, on dit aussi que Φ est un endomorphisme unitaire.*

Les propriétés suivantes se démontrent comme dans le cas euclidien.

Propriété B.19. Une application linéaire $\Phi : H \to H'$ est une isométrie si et seulement si elle préserve le produit hermitien.

Proposition B.20. Si Φ est une isométrie de H dans H', alors Φ est injective. De plus, si $\Phi : H \to H$ est un endomorphisme unitaire, alors Φ est un isomorphisme.

Démonstration. Elle est analogue au cas euclidien. □

Théorème B.21. Tout espace hermitien de dimension n est isométrique à l'espace hermitien canonique $\mathbb{C}^n$ (c'est-à-dire qu'il existe une isométrie bijective entre les deux).

Démonstration. Analogue au cas réel. □

Propriétés B.22. Soit $(H, \langle,\rangle)$ un espace hermitien.

(i) Soit Φ une isométrie de H. Alors Φ^{-1} est une isométrie de H.

(ii) Si Φ et Ψ sont des isométries de H alors $\Psi \circ \Phi$ est une isométrie de H.

(iii) id_H est une isométrie de H.

Définition-Proposition B.23. L'ensemble des isométries de l'espace hermitien $(H, \langle,\rangle_H)$ est un sous-groupe de $\mathrm{GL}(H)$. On l'appelle le *groupe unitaire* et on le note $\mathrm{U}(H)$.
L'ensemble des isométries de H dont le déterminant est égal à 1 est également un sous-groupe de $\mathrm{GL}(H)$. On l'appelle le *groupe spécial unitaire* et on le note $\mathrm{SU}(H)$. C'est le noyau de l'application déterminant sur $\mathrm{U}(H)$.
Si $H = \mathbb{C}^n$ est muni de sa structure canonique d'espace hermitien, on note $\mathrm{U}(H) = \mathrm{U}_n(\mathbb{C})$ et $\mathrm{SU}(H) = \mathrm{SU}_n(\mathbb{C})$.

Proposition B.24. Soit $(H, \langle,\rangle)$ un espace hermitien et soit $\Phi : H \to H$ une application linéaire. Alors les assertions suivantes sont équivalentes :

(i) Φ est une isométrie.

(ii) $\Phi^* = \Phi^{-1}$.

(iii) Si $\mathcal{B}$ est une base orthonormale de H et si M est la matrice de Φ dans la base $\mathcal{B}$, alors $t\overline{M}M = I_n = Mt\overline{M}$.

(iv) Si $\mathcal{B}$ est une base orthonormale de H et si M est la matrice de Φ dans la base $\mathcal{B}$, alors M est inversible et $M^{-1} = t\overline{M}$.

(v) Si $\mathcal{B}$ est une base orthonormale de H et si M est la matrice de Φ dans la base $\mathcal{B}$, alors les vecteurs colonne de M forment une base orthonormale de $\mathbb{C}^n$ avec sa structure hermitienne canonique.

(vi) Si $\mathcal{B}$ est une base orthonormale de H, alors $\Phi(\mathcal{B})$ est une base orthonormale de H.

Lorsque les conditions équivalentes (iii)-(v) ci-dessus sont vérifiées pour une matrice M, on dit que M est une *matrice unitaire.*

Démonstration. Analogue au cas réel. □

Proposition B.25. Soit Φ une isométrie. Alors

(i) les valeurs propres de Φ sont de module 1 ;

(ii) $|\det \Phi| = 1$.

Démonstration. (i) Soit λ une valeur propre pour Φ et soit z un vecteur propre ($z \neq 0$) associé. Alors, puisque Φ est une isométrie, $\|z\| = \|\Phi(z)\| = \|\lambda z\| = |\lambda| \, \|z\|$, donc $|\lambda| = 1$.

(ii) Soit M la matrice de Φ dans une base orthonormale. Alors $Mt\overline{M} = I_n$, donc

$$1 = \det(I_n) = \det(Mt\overline{M}) = \det(M)\det(t\overline{M}) = \det(M)\overline{\det(M)} = |\det(M)|^2 = |\det(\Phi)|^2,$$

donc $|\det(\Phi)| = 1$. □

B.5 Isométries de $\mathbb{C}^2$

Proposition B.26.

$$\mathrm{SU}_2(\mathbb{C}) = \left\{ \begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix} \mid |a|^2 + |b|^2 = 1 \right\}$$

$$\mathrm{U}_2(\mathbb{C}) = \left\{ \xi \begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix} \mid \xi \in \mathbb{C}, |\xi| = 1, \ |a|^2 + |b|^2 = 1 \right\}$$

Remarque B.27. Dans le second ensemble, la même matrice peut apparaître plusieurs fois avec des expressions différentes.

(TD)

B.6 Endomorphismes normaux

Soit $(H, \langle,\rangle)$ un espace hermitien de dimension n.

Définition B.28. *Un endomorphisme χ de H est dit normal s'il commute avec son adjoint, i.e. $\chi \circ \chi^* = \chi^* \circ \chi$.*

Une matrice A est dite normale si $A^*A = AA^*$.

Remarque B.29. Un endomorphisme χ est normal si et seulement si sa matrice dans une (toute) base *orthonormale* est normale.

Exemples B.30. (i) Tout endomorphisme hermitien est normal.

(ii) Tout endomorphisme unitaire (isométrie) est normal.

(iii) $A = \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix}$: alors $A^* = \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix}$. Donc A n'est ni hermitienne, ni unitaire. Cependant, A est normale car

$$AA^* = \begin{pmatrix} 2 & 2i \\ -2i & 2 \end{pmatrix} = A^*A.$$

Autre exemple $\begin{pmatrix} 2i & 0 \\ 0 & 0 \end{pmatrix}$.

C Théorèmes spectraux. Démonstration du théorème spectral pour les endomorphismes symétriques d'un espace euclidien.

Théorème C.1. (Théorème spectral des endomorphismes normaux) Soit H un espace hermitien. Un endomorphisme χ de H est normal si et seulement s'il est diagonalisable dans une base orthonormale de H.

Démonstration. Supposons d'abord que χ soit diagonalisable dans une base orthonormale, c'est-à-dire qu'il existe une base orthonormale $\mathcal{B}$ telle que la matrice D de χ dans $\mathcal{B}$ soit diagonale. Alors la matrice de χ^* est égale à $\overline{D}$. Or $D\overline{D} = \overline{D}D$, donc $\chi^* \circ \chi = \chi \circ \chi^*$, c'est-à-dire que χ est normal. On notera que les valeurs propres de χ sont des nombres complexes arbitraires.

On montre la réciproque par récurrence sur $n = \dim H$.

Le cas $n = 1$ est évident.

Supposons donc que $n \geqslant 2$ et que le résultat soit vrai pour un espace hermitien de dimension $n-1$. Soit χ un endomorphisme normal de H, où H est un espace hermitien de dimension n. Comme H est un espace vectoriel complexe, il existe une valeur propre λ pour χ. Soit z un vecteur propre associé. Posons $e_1 = \frac{z}{\|z\|}$: alors e_1 est un vecteur propre de norme 1 associé à la valeur propre λ.

Nous voulons montrer que $\operatorname{vect}\{e_1\}^\perp$ est stable par χ. Montrons d'abord que $\chi^*(e_1) = \overline{\lambda}e_1$. Pour cela, posons $w = \chi^*(e_1) - \overline{\lambda}e_1$, et calculons

$$\begin{aligned}
\langle w, w\rangle &= \langle \chi^*(e_1), \chi^*(e_1)\rangle - \langle \chi^*(e_1), \overline{\lambda}e_1\rangle - \langle \overline{\lambda}e_1, \chi^*(e_1)\rangle + \langle \overline{\lambda}e_1, \overline{\lambda}e_1\rangle \\
&= \langle e_1, \chi \circ \chi^*(e_1)\rangle - \lambda\langle e_1, \chi(e_1)\rangle - \overline{\lambda}\langle \chi(e_1), e_1\rangle + |\lambda|^2 \|e_1\|^2 \\
&= \langle e_1, \chi^* \circ \chi(e_1)\rangle - \lambda\langle e_1, \lambda e_1\rangle - \overline{\lambda}\langle \lambda e_1, e_1\rangle + |\lambda|^2 \\
&= \langle \chi(e_1), \chi(e_1)\rangle - |\lambda|^2 - |\lambda|^2 + |\lambda|^2 \\
&= \langle \lambda e_1, \lambda e_1\rangle - |\lambda|^2 \\
&= |\lambda|^2 \|e_1\|^2 - |\lambda|^2 \\
&= 0,
\end{aligned}$$

D'où $w = 0$ et donc $\chi^*(e_1) = \overline{\lambda}e_1$.

Soit $F = \operatorname{vect}\{e_1\}^\perp$ et soit $z' \in F$. Alors $\langle \chi(z'), e_1\rangle = \langle z', \chi^*(e_1)\rangle = \langle z', \overline{\lambda}e_1\rangle = \lambda\langle z', e_1\rangle = 0$ donc $\chi(z') \in F$. Ainsi la restriction de χ à F définit un endomorphisme $\chi_{|F}$ de F. C'est encore un endomorphisme normal (car on voit facilement que $(\chi_{|F})^* = (\chi^*)_{|F}$). L'hypothèse de récurrence montre qu'il existe une base orthonormale $\{e_2, \ldots, e_n\}$ de F formée de vecteurs propres pour $\chi_{|F}$ et donc pour χ. Alors $\{e_1, e_2, \ldots, e_n\}$ est une base orthonormale de H formée de vecteurs propres pour χ. □

De la même façon que dans le cas réel, on déduit la version matricielle :

Corollaire C.2. (Théorème spectral pour les matrices normales) Soit $A \in \mathcal{M}_n(\mathbb{C})$. Alors A est une matrice normale si et seulement s'il existe une matrice unitaire P telle que P^*AP soit diagonale.

Corollaire C.3. (Théorème spectral des endomorphismes hermitiens) Soit H un espace hermitien. Soit ψ un endomorphisme hermitien. Alors les valeurs propres de ψ sont réelles. De plus, ψ est diagonalisable dans une base orthonormale, c'est-à-dire qu'il existe une base orthonormale de H formée de vecteurs propres pour ψ.

Démonstration. Si ψ est hermitien, alors $\psi^* = \psi$ et donc ψ est normal. On peut donc appliquer le théorème spectral pour les endomorphismes normaux et la proposition **??**. □

Corollaire C.4. (Théorème spectral des endormorphismes unitaires) Soit $(H, \langle,\rangle)$ un espace hermitien et soit Φ une isométrie de H. Alors les valeurs propres de Φ sont de module 1. De plus, Φ est diagonalisable dans une base orthonormale de H.

Démonstration. Si Φ est unitaire, alors $\Phi^* = \Phi^{-1}$ donc Φ est normal. On peut donc appliquer le théorème spectral pour les endomorphismes normaux et la proposition **??**. □

Corollaire C.5. (Caractérisation des endomorphismes hermitiens et unitaires) Soit χ un endomorphisme normal de H. Pour que χ soit hermitien (*resp.* unitaire), il faut et il suffit que ses valeurs propres soient réelles (*resp.* de module 1).

Démonstration. Les conditions nécessaires ont déjà été prouvées (propositions **??** et **??**).

Réciproquement, d'après le théorème spectral pour les endomorphismes normaux, il existe une base orthonormale $\mathcal{B}$ dans laquelle la matrice de χ est diagonale, notons-la D.

Si les coefficients diagonaux sont réels, il est clair que $t\overline{D} = D$, donc χ est hermitien.

Si les coefficients diagonaux sont des complexes de module 1, il est clair que $t\overline{D} = \overline{D} = D^{-1}$, donc χ est unitaire. □

Nous pouvons maintenant également démontrer le théorème spectral pour les endomorphismes symétriques d'un espace euclidien (Chapitre I, théorème **??**).

Démonstration. Soit ψ un endomorphisme symétrique d'un espace euclidien E.

Montrons d'abord que ψ a au moins une valeur propre réelle. Soit A la matrice de ψ dans une base orthonormale. Alors $tA = A$. Soit $c(x)$ le polynôme caractéristique de A. On peut voir A comme une matrice de $\mathcal{M}_n(\mathbb{C})$, et nous savons donc que $c(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n)$ avec $\lambda_i \in \mathbb{C}$. De plus, A est hermitienne, donc toutes ses valeurs propres sont réelles. Donc ψ a au moins une valeur propre réelle.

Maintenant nous allons démontrer le théorème par récurrence sur $n = \dim E$.

Si $n = 1$, alors il n'y a rien à montrer. Supposons donc que $n \geqslant 2$. Nous savons donc qu'il existe une valeur propre réelle, λ. Soit v_1 un vecteur propre associé à la valeur propre λ ci-dessus. Quitte à le diviser par sa norme, nous pouvons supposer que $\|v_1\| = 1$. Nous avons une décomposition orthogonale $E = \operatorname{vect}\{v_1\} \oplus \operatorname{vect}\{v_1\}^{\perp}$ qui est stable par ψ. Nous pouvons donc restreindre ψ à $\operatorname{vect}\{v_1\}^{\perp}$ pour obtenir $\tilde{\psi} : \operatorname{vect}\{v_1\}^{\perp} \to \operatorname{vect}\{v_1\}^{\perp}$. Ce sous-espace est de dimension $n - 1$, donc par hypothèse de récurrence, il existe une base orthonormale $\{v_2, \ldots, v_n\}$ de $\operatorname{vect}\{v_1\}^{\perp}$ formée de vecteurs propres pour $\tilde{\psi}$ et donc pour ψ. Donc $\{v_1, \ldots, v_n\}$ est une base orthonormale de E formée de vecteurs propres pour ψ. □

Courbes

III Courbes paramétrées de l'espace

A Généralités, exemples

Idée Une courbe paramétrée est une déformation d'une droite ou d'un morceau de droite.

Définition A.1. *On appelle courbe paramétrée de classe $\mathcal{C}^k$, $k \geqslant 1$, toute application $\alpha : I \to \mathbb{R}^3$ de classe $\mathcal{C}^k$, où I est un intervalle ouvert non vide de $\mathbb{R}$.*

On identifiera $\mathbb{R}^2$ à $\mathbb{R}^2 \times \{0\}$, si bien que $\mathbb{R}^2 \subseteq \mathbb{R}^3$. La courbe α est dite plane si $\alpha(I) \subseteq \mathbb{R}^2$.

L'image $\alpha(I)$ est appelée la trace de α. L'application α est appelée paramétrisation de $\alpha(I)$.

Remarque A.2. Dans certaines situations que l'on précisera, I pourra être un intervalle non vide *quelconque*. Dans certains exemples, I pourra être une réunion d'intervalles ouverts non vides.

Définition A.3. *Soit $t \in I$ tel que $\alpha'(t) = 0$ (le vecteur dérivé $\alpha'(t)$ est obtenu en dérivant α composante à composante). On dit que t est une singularité de α et que $\alpha(t)$ est un point singulier de α.*

Une courbe paramétrée α sans singularité est dite régulière.

Définition A.4. *Supposons que $\alpha(t)$ ne soit pas singulier. Alors la tangente à $\alpha(I)$ en $\alpha(t)$ est la droite affine de vecteur directeur le vecteur tangent $\alpha'(t)$ et passant par $\alpha(t)$, c'est-à-dire $\{\alpha(t) + \lambda\alpha'(t) \mid \lambda \in \mathbb{R}\}$.*

Exemples A.5. **1)** Considérons

$$\begin{array}{ccc} \mathbb{R} & \to & \mathbb{R}^3 \\ t & \mapsto & A + tB \end{array}$$

où A et $B \neq 0$ sont des vecteurs de $\mathbb{R}^3$. Sa trace est une droite affine de $\mathbb{R}^3$.

2) Modifions un peu l'exemple précédent :

$$\begin{array}{ccc}]-\frac{\pi}{2}, \frac{\pi}{2}[& \to & \mathbb{R}^3 \\ u & \mapsto & A + (\tan u)B \end{array}$$

Sa trace est la même droite, mais avec une paramétrisation différente.

3) Considérons

$$\begin{array}{rcl} \alpha : \mathbb{R} & \to & \mathbb{R}^3 \\ t & \mapsto & (a\cos t, a\sin t, bt) \end{array}$$

avec $a > 0$ et $b \neq 0$. Alors α est une courbe régulière ($\alpha'(t) = (-a\sin t, a\cos t, b) \neq (0,0,0)$). La trace de α est une *hélice* de *pas* $2\pi b$.

4) Considérons

$$\begin{array}{rcl} \alpha : \mathbb{R} & \to & \mathbb{R}^2 \\ t & \mapsto & (t^3, t^2). \end{array}$$

C'est une courbe plane ayant une singularité en $t = 0$ ($\alpha'(t) = (3t^2, 2t)$). Nous étudierons ce type de courbe en TD.

5) Considérons

$$\begin{array}{rcl} \alpha : \mathbb{R} & \to & \mathbb{R}^2 \\ t & \mapsto & (t^3 - 4t, t^2 - 4). \end{array}$$

C'est une courbe régulière ($\alpha'(t) = (3t^2 - 4, 2t) \neq (0,0)$).
Nous étudierons ce type de courbe en TD. Il s'agit d'un folium parabolique.

Nous ne considérerons plus que des courbes régulières dans ce cours.

B Rappels et compléments de calcul différentiel et de géométrie

Dans ce chapitre, nous allons devoir utiliser du calcul différentiel pour des applications à valeurs dans $\mathbb{R}^n$. Nous allons donc faire quelques rappels et compléments sur le calcul différentiel.

L'espace $\mathbb{R}^n$ est considéré à la fois comme l'espace vectoriel euclidien canonique et comme un espace affine.

Définition B.1. *Soit I un intervalle de $\mathbb{R}$. Soit $f : I \to \mathbb{R}^n$ une application. On dit que*

- *f a une limite $A \in \mathbb{R}^n$ lorsque t tend vers t_0 si l'application à valeurs réelles $\|f(t) - A\|$ tend vers 0 lorsque t tend vers t_0.*
- *f est dérivable en t_0 si le vecteur*
$$\frac{f(t) - f(t_0)}{t - t_0}$$
a une limite lorsque t tend vers t_0. Cette limite s'appelle la dérivée en t_0, et est notée $f'(t_0)$ ou $\left.\frac{df}{dt}\right|_{t=t_0}$.

Remarque B.2. Si f est à valeurs dans un sous-espace affine F de $\mathbb{R}^n$, alors f' est à valeurs dans sa direction $\vec{F}$:

- Si F est un hyperplan affine, alors il est déterminé par une équation $a_1x_1 + \cdots + a_nx_n = c$. Puisque f est à valeurs dans F, pour tout $t \in I$, $f(t)$ vérifie cette équation : $a_1f_1(t) + \cdots + a_nf_n(t) = c$. En dérivant cette identité, on voit que f' est à valeurs dans $\{x \in \mathbb{R}^n \mid a_1x_1 + \cdots + a_nx_n = 0\}$, c'est-à-dire la direction de F.
- Si F est un sous-espace affine quelconque, on peut l'obtenir comme intersection d'hyperplans (les cas $n = 2$ et $n = 3$ qui nous concernent ici ont été démontrés au premier semestre), c'est-à-dire comme l'ensemble des points vérifiant une famille d'équations comme ci-dessus. On termine de même.

Remarque B.3. Notons que la dérivée d'une application à valeurs dans $\mathbb{R}^n$ est une application à valeurs dans la direction de $\mathbb{R}^n$, qui est encore $\mathbb{R}^n$. On peut donc itérer la dérivation aussi souvent que l'on veut.

Propriétés B.4. (i) Si $f = (f_1, \ldots, f_n)$ où les f_i sont les composantes de f (ce sont donc des applications de I dans $\mathbb{R}$), alors f est dérivable si et seulement si toutes ses composantes le sont, et on a $f' = (f_1', \ldots, f_n')$.

(ii) On déduit de la première propriété tous les résultats classiques sur la dérivation : la dérivée d'une fonction constante est nulle, la dérivation est linéaire, et on a les mêmes formules que pour les fonctions réelles pour la dérivée d'une composée (une fonction réelle suivie d'une fonction vectorielle) ou d'un produit par une fonction réelle.

(iii) Dérivation d'un produit scalaire : notons $\langle , \rangle$ le produit scalaire usuel de $\mathbb{R}^n$. Si f_1 et f_2 sont des applications dérivables à valeurs dans $\mathbb{R}^n$, alors $\langle f_1, f_2 \rangle$ est dérivable, et
$$\frac{d(\langle f_1, f_2 \rangle)}{dt} = \langle \frac{df_1}{dt}, f_2 \rangle + \langle f_1, \frac{df_2}{dt} \rangle.$$

Le lemme suivant sera utilisé à plusieurs reprises :

Lemme B.5. Soit $f : I \to \mathbb{R}^n$. Supposons que $\|f\| : I \to \mathbb{R}$ soit constante. Alors $f(t)$ est orthogonal à $f'(t)$ pour tout $t \in I$.

Démonstration. Soit $t \in I$. Nous avons $\|f(t)\| = k$ où k est une constante. Donc $\langle f(t), f(t)\rangle = \|f(t)\|^2 = k^2$. Si nous dérivons cela en utilisant les propriétés ci-dessus, nous obtenons : $0 = \langle f'(t), f(t)\rangle + \langle f(t), f'(t)\rangle = 2\langle f(t), f'(t)\rangle$, d'où le résultat. □

Proposition B.6. Le produit vectoriel $u \wedge v$ de deux vecteurs u et v de $\mathbb{R}^3$ est le vecteur déterminé par $\langle u \wedge v, x\rangle = \det(u, v, x)$ pour tout $x \in \mathbb{R}^3$, où le déterminant est pris dans la base canonique de $\mathbb{R}^3$. Si u et v ne sont pas colinéaires, $u \wedge v$ est un vecteur orthogonal au plan engendré par u et v, de norme égale à l'aire du parallélogramme déterminé par u et v, et dont le sens est déterminé par le fait que $\{u, v, u \wedge v\}$ doit être une base directe (c'est-à-dire telle que son déterminant dans la base canonique de $\mathbb{R}^3$ est strictement positif). Le produit vectoriel n'est pas associatif, et :

$$(u \wedge v) \wedge w = \langle u, w\rangle v - \langle v, w\rangle u.$$

Démonstration. Des propriétés du produit vectoriel sont données en document (et vous avez fait un contrôle dessus au semestre précédent). La dernière remarque est laissée en exercice. Nous allons montrer ici que la norme de $u \wedge v$ est l'aire du parallélogramme déterminé par u et v.

Montrons que

$$\langle u \wedge v, x \wedge y\rangle = \begin{vmatrix} \langle u, x\rangle & \langle v, x\rangle \\ \langle u, y\rangle & \langle v, y\rangle \end{vmatrix},$$

où u, v, x, y sont des vecteurs arbitraires. Remarquons que comme les deux côtés de l'égalité sont linéaires en u, v, x, y, il suffit de montrer que

$$\langle e_i \wedge e_j, e_k \wedge e_\ell\rangle = \begin{vmatrix} \langle e_i, e_k\rangle & \langle e_j, e_k\rangle \\ \langle e_i, e_\ell\rangle & \langle e_j, e_\ell\rangle \end{vmatrix}$$

pour tous $i, j, k, \ell = 1, 2, 3$, ce qui se fait facilement.

Ainsi,

$$\|u \wedge v\|^2 = \begin{vmatrix} \langle u, u\rangle & \langle u, v\rangle \\ \langle u, v\rangle & \langle v, v\rangle \end{vmatrix} = \|u\|^2 \|v\|^2 (1 - \cos^2 \theta) = \mathcal{A}^2$$

où θ is l'angle orienté entre u et v, et $\mathcal{A}$ est l'aire du parallélogramme déterminé par u et v. □

Proposition B.7. Soient $u, v \; :]a, b[\to \mathbb{R}^3$ deux applications dérivables sur $]a, b[$. Alors $u \wedge v$ est dérivable sur $]a, b[$, et

$$\frac{d}{dt}(u \wedge v) = \frac{du}{dt} \wedge v + u \wedge \frac{dv}{dt}.$$

Démonstration. Chaque composante de $u \wedge v$ est un polynôme en les composantes de u et v, donc $u \wedge v$ est dérivable. L'expression de la dérivée se vérifie aisément à partir de l'expression analytique du produit vectoriel (regarder les fonctions composante par composante). □

Définition B.8. *Une application f d'un intervalle ouvert non vide I dans un intervalle ouvert non vide J de $\mathbb{R}$ est un $\mathcal{C}^k$-difféomorphisme ($k \geqslant 1$) si c'est une application de classe $\mathcal{C}^k$ bijective* ***dont la réciproque est aussi de classe $\mathcal{C}^k$****.*

Remarque B.9. Si f est un $\mathcal{C}^k$-difféomorphisme, alors $f'(t) \neq 0$ pour tout $t \in I$ (il suffit de dériver l'identité $f^{-1} \circ f = \mathrm{id}_I$).

Théorème B.10. (Théorème d'inversion) Soit I un intervalle ouvert non vide de $\mathbb{R}$, et soit $f : I \to \mathbb{R}$ une application de classe $\mathcal{C}^k$ injective et telle que pour tout $t \in I$ on ait $f'(t) \neq 0$. Alors f est un C^k-difféomorphisme de I sur $f(I)$.

Changement de paramétrisation Soit $\alpha : I \to \mathbb{R}^3$ une courbe paramétrée de classe $\mathcal{C}^k$. Soit $f : J \to I$ un $\mathcal{C}^k$-difféomorphisme. Alors la trace de $\beta := \alpha \circ f : J \to \mathbb{R}^3$ est la même que la trace de $\alpha : I \to \mathbb{R}^3$. De plus, β est singulière au point t si et seulement si α est singulière au point $f(t)$: en effet, $\beta'(t) = \alpha'(f(t))f'(t)$ avec $f'(t) \neq 0$. Donc on peut changer la paramétrisation d'une courbe à l'aide d'un difféomorphisme (*cf* les deux premiers exemples du chapitre).

C Rectification et abscisse curviligne

Nous voulons définir la longueur (de la trace) d'une courbe paramétrée.

Définition C.1. *Soit $\alpha : I \to \mathbb{R}^3$ une courbe de classe $\mathcal{C}^k$, et soit $[a,b]$ un intervalle fermé borné contenu dans I. Pour toute subdivision σ de $[a,b]$ (c'est-à-dire que $\sigma = (t_i)_{0\leqslant i\leqslant n}$ avec $t_0 = a$, $t_n = b$ et $t_i < t_{i+1}$ pour $0 \leqslant i \leqslant n-1$), on considère la somme*

$$\ell(\alpha,\sigma) = \sum_{i=1}^{n} \|\alpha(t_i) - \alpha(t_{i-1})\| .$$

On dit que α est rectifiable sur $[a,b]$ si la borne supérieure des $\ell(\alpha,\sigma)$ est finie lorsque σ parcourt l'ensemble des subdivisions de $[a,b]$. Dans ce cas, cette borne supérieure s'appelle la longueur de α sur $[a,b]$ et se note $L(\alpha_{|[a,b]})$.

Remarque C.2. Géométriquement, on approche la trace de la courbe à l'aide de "droites brisées" dont on peut mesurer la longueur. Plus la droite brisée se rapproche de la courbe, plus cette longueur se rapproche de $L(\alpha_{|[a,b]})$.

Proposition C.3. La longueur est additive : soit $\alpha : I \to \mathbb{R}^3$ une courbe, et soient $a,b,c \in I$ avec $a < b < c$. Alors α est rectifiable sur $[a,c]$ si et seulement si α est rectifiable sur $[a,b]$ et sur $[b,c]$ et on a

$$L(\alpha_{|[a,c]}) = L(\alpha_{|[a,b]}) + L(\alpha_{|[b,c]}).$$

Démonstration. • Supposons que α soit rectifiable sur $[a,c]$. Soient σ' une subdivision de $[a,b]$ et σ'' une subdivision de $[b,c]$. Alors $\sigma = \sigma' \cup \sigma''$ est une subdivision de $[a,c]$ et

$$\ell(\alpha,\sigma') + \ell(\alpha,\sigma'') = \ell(\alpha,\sigma) \leqslant L(\alpha_{|[a,c]}).$$

On en déduit que α est rectifiable sur $[a,b]$ et sur $[b,c]$ et que

$$L(\alpha_{|[a,b]}) + L(\alpha_{|[b,c]}) \leqslant L(\alpha_{|[a,c]}).$$

• Réciproquement, supposons que α soit rectifiable sur $[a,b]$ et sur $[b,c]$. Soit θ une subdivision de $[a,c]$, et soit $\sigma = \theta \cup \{b\}$. Alors σ est encore une subdivision de $[a,c]$, et de plus on peut écrire $\sigma = \sigma' \cup \sigma''$ où σ' est une subdivision de $[a,b]$ et σ'' est une subdivision de $[b,c]$. De plus

$$\ell(\alpha,\theta) \leqslant \ell(\alpha,\sigma) = \ell(\alpha,\sigma') + \ell(\alpha,\sigma'') \leqslant L(\alpha_{|[a,b]}) + L(\alpha_{|[b,c]}).$$

Par conséquent, α est rectifiable sur $[a,c]$ et

$$L(\alpha_{|[a,c]}) \leqslant L(\alpha_{|[a,b]}) + L(\alpha_{|[b,c]}).$$

On en déduit le résultat. $\square$

Convention : Si $a = b$, nous posons $L(\alpha_{|[a,a]}) = 0$.

Nous voulons un moyen plus commode de calculer la longueur d'une courbe rectifiable :

Proposition C.4. Soit $\alpha : I \to \mathbb{R}^3$ une courbe de classe $\mathcal{C}^k$ avec $k \geqslant 1$. Alors pour tout intervalle $[a,b]$ contenu dans I, α est rectifiable sur $[a,b]$ et

$$L(\alpha_{|[a,b]}) = \int_a^b \left\|\alpha'(t)\right\| dt. \tag{C.1}$$

Démonstration. • Soit $\sigma = (t_i)_{0\leqslant i\leqslant n}$ une subdivision d'un intevalle $[c,d]$ contenu dans $[a,b]$. La dérivée α' de α est continue sur l'intervalle fermé borné $[c,d]$, donc bornée : $\|\alpha'(t)\| \leqslant K$ pour tout $t \in [c,d]$. Nous pouvons appliquer le théorème des accroissements finis : $\|\alpha(t_i) - \alpha(t_{i-1})\| \leqslant K(t_i - t_{i-1})$ pour tout $1 \leqslant i \leqslant n$. Donc $\ell(\alpha,\sigma) \leqslant K\sum_{i=1}^n (t_i - t_{i-1}) = K(d-c)$. Le terme de droite est indépendant du choix de σ, donc le sup des $\ell(\alpha,\sigma)$ pris sur l'ensemble des subdivisions σ est fini et α est rectifiable sur $[c,d]$ avec

$$L(\alpha_{|[c,d]}) \leqslant \sup_{x\in[c,d]} \left\|\alpha'(x)\right\| |d-c| \,. \tag{C.2}$$

• Pour tout $t \in [a,b]$, on pose $\phi(t) = L(\alpha_{|[a,t]})$. Montrons que ϕ est dérivable sur $[a,b]$ et que $\phi'(t) = \|\alpha'(t)\|$ pour $t \in [a,b]$.
Soit $t \in [a,b]$ et soit $\varepsilon > 0$. Puisque α' est continue, il existe $\eta > 0$ tel que pour tout x avec $|x-t| < \eta$ et $x \in [a,b]$, on ait $\|\alpha'(x) - \alpha'(t)\| < \varepsilon$. Or $\|\alpha'(x) - \alpha'(t)\| \geqslant \big|\, \|\alpha'(x)\| - \|\alpha'(t)\| \,\big| \geqslant \|\alpha'(x)\| - \|\alpha'(t)\|$. Donc

$$\forall x, |x-t| < \eta \text{ et } x \in [a,b] \Rightarrow \left\|\alpha'(x)\right\| < \left\|\alpha'(t)\right\| + \varepsilon.$$

Maintenant, soit $h \geqslant 0$ tel que $h \leqslant \eta$ et $t+h \in [a,b]$. Alors, grâce à l'additivité de la longueur et à la relation (**??**), on a

$$|\phi(t+h) - \phi(t)| = L(\alpha_{|[t,t+h]}) \leqslant \sup_{x\in[t,t+h]} \left\|\alpha'(x)\right\| h < (\left\|\alpha'(t)\right\| + \varepsilon)h.$$

De même, lorsque $h \leqslant 0$, $|h| \leqslant \eta$ et $t+h \in [a,b]$, on a $|\phi(t+h) - \phi(t)| < (\|\alpha'(t)\| + \varepsilon)\,|h|$.
D'autre part, en prenant la subdivision triviale, on a $|\phi(t+h) - \phi(t)| = L(\alpha_{|[t,t+h]}) \geqslant \|\alpha(t+h) - \alpha(t)\|$ si $h \geqslant 0$, et la même minoration est vraie pour $h \leqslant 0$. On obtient donc

$$\frac{1}{|h|} \left\|\alpha(t+h) - \alpha(t)\right\| \leqslant \left|\frac{\phi(t+h) - \phi(t)}{h}\right| \leqslant \left\|\alpha'(t)\right\| + \varepsilon.$$

Puisque α est dérivable, on en déduit que ϕ est dérivable et que $\phi'(t) = \|\alpha'(t)\|$.

• On a donc

$$\int_a^b \left\|\alpha'(t)\right\| dt = \int_a^b \phi'(t)dt = \phi(b) - \phi(a) = L(\alpha_{|[a,b]}) - L(\alpha_{|[a,a]}) = L(\alpha_{|[a,b]}). \quad \square$$

Remarque C.5. On suppose que $f : J \to I$ est de classe $\mathcal{C}^k$ et que $f'(u) > 0$ pour tout $u \in J$. Si on pose $\beta = \alpha \circ f$, le changement de variable $t = f(u)$ dans l'intégrale (**??**) donne

$$\int_a^b \left\|\alpha'(t)\right\| dt = \int_{f^{-1}(a)}^{f^{-1}(b)} \left\|\alpha'(f(u))\right\| f'(u)du = \int_{f^{-1}(a)}^{f^{-1}(b)} \left\|\beta'(u)\right\| du$$

ce qui veut dire que l'intégrale (**??**) est indépendante du choix de la paramétrisation.

Définition C.6. *Soit $\alpha : I \to \mathbb{R}^3$ une courbe $\mathcal{C}^k$ avec $k \geqslant 1$. Fixons $t_0 \in I$. L'abscisse curviligne de α d'origine t_0 est l'application*

$$\begin{array}{rcl} s : I & \to & \mathbb{R} \\ t & \mapsto & s(t) = \int_{t_0}^t \|\alpha'(u)\|\, du. \end{array}$$

Remarque C.7. Supposons α régulière. On va pouvoir paramétrer la trace de α par s. En effet, $\frac{ds}{dt} = \|\alpha'(t)\| \neq 0$ partout. Clairement, s est de classe $\mathcal{C}^k$ (puisque $\frac{ds}{dt}$ est composée de α' qui ne s'annule pas et de la norme, de classe $\mathcal{C}^{k-1}$) et injective sur I (puisque $\|\alpha'\| > 0$). D'après le théorème d'inversion, $s : I \to s(I)$ est un $\mathcal{C}^k$-difféomorphisme. Notons $J = s(I)$. On pose

$$\begin{array}{rcl} \beta = \alpha \circ s^{-1} : J & \to & \mathbb{R}^3 \\ s & \mapsto & \beta(s). \end{array}$$

Alors β est une courbe paramétrée de même trace que α. En fait, une fois choisies l'origine $M_0 = \alpha(t_0)$ et l'orientation sur la trace, l'application $\beta : J \to \mathbb{R}^3$ est bien déterminée.

Propriété C.8. Soit $\alpha : I \to \mathbb{R}^3$ une courbe paramétrée, et soit β comme ci-dessus. Alors $\|\beta'(s)\| = 1$ pour tout $s \in J$. On dit que β est à *vitesse normalisée.* De plus, si α est une courbe à vitesse normalisée telle que $0 \in I$, alors $\beta = \alpha$ (si on choisit 0 comme origine).

En conclusion, *α est paramétrée par l'abscisse curviligne d'origine* 0 *si et seulement si elle est à vitesse normalisée.*

Démonstration. On a
$$\left\| \frac{d\alpha}{dt} \right\| = \left\| \frac{d\beta}{ds} \frac{ds}{dt} \right\| = \left\| \frac{d\beta}{ds} \right\| \left| \frac{ds}{dt} \right| = \left\| \frac{d\beta}{ds} \right\| \|\alpha'(t)\|$$
d'où $\|\beta'(s)\| = 1$.

Réciproquement, on a $s(t) = \int_0^t \|\alpha'(u)\|\, du = \int_0^t du = t$, donc $\beta = \alpha$. □

D Courbure, torsion, trièdre de Serret-Frénet

Fixons une courbe $\alpha : I \to \mathbb{R}^3$ régulière à vitesse normalisée (autrement dit, on paramètre α par l'abscisse curviligne s), de classe $\mathcal{C}^\infty$ pour simplifier (on pourrait prendre α de classe $\mathcal{C}^k$ avec k convenable).

Définition D.1. *Si $s \in I$, le nombre $\kappa(s) = \|\alpha''(s)\|$ s'appelle la courbure de α en s.*

Exemple D.2. Considérons le cercle de rayon r de paramétrisation

$$\begin{array}{rcl} \alpha : \mathbb{R} & \to & \mathbb{R}^2 \\ t & \mapsto & (r\cos t, r\sin t). \end{array}$$

Alors $\alpha'(t) = (-r\sin t, r\cos t)$ et $\|\alpha'(t)\|^2 = r^2$. Ceci signifie que $\frac{ds}{dt}(t) = \|\alpha'(t)\| = r$ et donc $s = rt$. Donc le cercle peut être paramétré par

$$\begin{array}{rcl} \beta : \mathbb{R} & \to & \mathbb{R}^2 \\ s & \mapsto & \left(r\cos\frac{s}{r}, r\sin\frac{s}{r}\right) \end{array}$$

et β est à vitesse normalisée. On obtient $\beta''(s) = \left(-\frac{1}{r}\cos\frac{s}{r}, -\frac{1}{r}\sin\frac{s}{r}\right)$ et donc $\kappa(s) = \frac{1}{r}$. Ainsi, pour un cercle, la courbure est l'inverse du rayon.

Proposition D.3. Soit α une courbe régulière de $\mathbb{R}^3$. Alors $\kappa \equiv 0$ si et seulement si la courbe est contenue dans une droite.

Démonstration. La courbure est identiquement nulle si et seulement si $\alpha'' \equiv 0$, ce qui est équivalent à $\alpha(s) = sA + B$ pour des vecteurs $A \neq 0$ (α est régulière) et B, ce qui est l'équation d'une droite. □

Définition D.4. *Supposons que $\kappa(s) \neq 0$ pour un $s \in I$. Alors*

- $\vec{t}(s) = \alpha'(s)$ *s'appelle le vecteur tangent de α en s.*
- $\vec{n}(s) = \frac{\alpha''(s)}{\kappa(s)}$ *s'appelle le vecteur normal de α en s.*

Remarque D.5. $\vec{t}(s)$ et $\vec{n}(s)$ sont unitaires et orthogonaux.

En effet, puisque α est à vitesse normalisée, $\vec{t}(s)$ est unitaire, et $\vec{n}(s)$ est unitaire par définition de la courbure. D'après le lemme **??**, $\vec{t}(s)$ et $\vec{n}(s)$ sont donc orthogonaux (puisque $\|\vec{t}\|$ est constante et $\vec{n}(s)$ est colinéaire à $\vec{t}'(s)$).

Dans toute la suite, on va supposer que la courbure κ est partout non nulle. L'application $\kappa : I \to]0, +\infty[$ est alors de classe $\mathcal{C}^\infty$.

Définition D.6. *Le plan déterminé par $\vec{t}(s)$ et $\vec{n}(s)$ est appelé le plan osculateur en s.*

Définition D.7. *Le vecteur $\vec{b}(s) = \vec{t}(s) \wedge \vec{n}(s)$ s'appelle le vecteur binormal en s.*

Le repère orthonormal direct $(\vec{t}(s), \vec{n}(s), \vec{b}(s))$ s'appelle le repère (ou trièdre) de Serret-Frénet (l'origine $\alpha(s)$ est sous-entendue).

Remarque D.8. Les applications $\vec{t}, \vec{n}, \vec{b} : I \to \mathbb{R}^3$ sont de classe $\mathcal{C}^\infty$.

On a $\vec{b}'(s) = \vec{t}'(s) \wedge \vec{n}(s) + \vec{t}(s) \wedge \vec{n}'(s) = \vec{t}(s) \wedge \vec{n}'(s)$ puisque $\vec{t}'(s)$ est colinéaire à $\vec{n}(s)$, donc $\vec{b}'(s)$ est orthogonal à $\vec{t}(s)$ pour tout s.

De plus, $\left\|\vec{b}(s)\right\| = \|\vec{t}(s)\| \, \|\vec{n}(s)\| = 1$ implique que $\vec{b}'(s)$ est orthogonal à $\vec{b}(s)$ pour tout s.

Donc $\vec{b}'(s)$ est colinéaire à $\vec{n}(s)$. D'où la définition :

Définition D.9. *On définit la torsion $\tau(s) \in \mathbb{R}$ en s par la formule*

$$\vec{b}'(s) = -\tau(s)\vec{n}(s).$$

Elle mesure les variations du plan osculateur.

Remarque D.10. On a $\tau = -\langle \vec{b}', \vec{n} \rangle$. En particulier, l'application $\tau : I \to \mathbb{R}$ est de classe $\mathcal{C}^\infty$.

Proposition D.11. Soit α une courbe régulière de $\mathbb{R}^3$. Alors $\tau \equiv 0$ si et seulement si la trace de α est contenue dans un plan.

Démonstration. Si la torsion est identiquement nulle, alors $\vec{b}'$ est indentiquement nul, donc $\frac{d}{ds}(\langle \alpha, \vec{b} \rangle) = \langle \vec{t}, \vec{b} \rangle = 0$ par orthogonalité. Ceci veut dire que $\langle \alpha, \vec{b} \rangle \equiv k$ pour une constante k, et donc la courbe est contenue dans le plan d'équation $\langle x, \vec{b} \rangle = k$.

Réciproquement, si la trace de α est contenue dans un plan, alors $\vec{t}$ et $\vec{t}' = \kappa\vec{n}$ sont parallèles à ce plan (en effet, si α est à valeurs dans un plan, ses dérivées successives sont à valeurs dans la direction de ce plan). Donc $\vec{b}$ est orthogonal à ce plan, et donc constant, donc $\vec{b}' = 0$ et $\tau = 0$. □

Lorsque s varie sur la courbe, le repère $(\vec{t}, \vec{n}, \vec{b})$ varie. Les formules de Serret-Frénet nous indiquent comment :

Théorème D.12. (Formules de Sérret-Frénet)

$$\frac{d}{ds}\begin{pmatrix} \vec{t} \\ \vec{n} \\ \vec{b} \end{pmatrix} = \begin{pmatrix} & \kappa\vec{n} & \\ -\kappa\vec{t} & & +\tau\vec{b} \\ & -\tau\vec{n} & \end{pmatrix}$$

Remarque D.13. Pour retenir ces formules, notons que $\frac{d}{ds}\begin{pmatrix} \vec{t} \\ \vec{n} \\ \vec{b} \end{pmatrix} = A\begin{pmatrix} \vec{t} \\ \vec{n} \\ \vec{b} \end{pmatrix}$ où A est la matrice antisymétrique

$$A = \begin{pmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{pmatrix}.$$

Démonstration. Puisque $(\vec{t}, \vec{n}, \vec{b})$ est une base orthonormale, la matrice P dont les lignes sont $\vec{t}, \vec{n}, \vec{b}$ est orthogonale, c'est-à-dire que

$$tPP = I = PtP, \tag{D.1}$$

et nous voulons trouver P'.

P' vérifie $P' = P'tPP$, ce qui, en posant $A = P'tP$, peut s'écrire $P' = AP$. Dériver (**??**) donne $P'tP + PtP' = 0$, soit $A + tA = 0$, donc A est antisymétrique. Les définitions de la courbure et de la torsion donnent la première et la dernière lignes de A, et l'antisymétrie donne la ligne manquante de A. □

Les formules donnant la courbure et la torsion d'une courbe si elle n'est pas à vitesse normalisée sont données à la fin du chapitre.

E Courbure des courbes planes

Théorème E.1. (Théorème de reconstruction pour les courbes planes) Soit $\kappa : I \to \mathbb{R}$ une application de classe $\mathcal{C}^2$ (au moins) ne s'annulant pas. Alors les courbes $\alpha : I \to \mathbb{R}^2$ à vitesse normalisée dont la courbure est κ sont données par

$$\alpha(s) = \left(\int_{s_0}^{s} \cos(\theta(u))\, du + a, \int_{s_0}^{s} \sin(\theta(u))\, du + b \right)$$

où $\theta(s) = \int_{s_0}^{s} \kappa(u) du + \varphi$ avec a, b, φ des constantes réelles et $s_0 \in I$. La courbe est donc déterminée à rotation et à translation près.

Démonstration. Notons $\alpha = (x, y)$. Alors $x'^2 + y'^2 = \|\alpha'\|^2 = 1$ (puisque α est à vitesse normalisée), donc il existe une fonction dérivable θ telle que $x' = \cos\theta$ et $y' = \sin\theta$. Alors $\kappa(s)^2 = \|\alpha''(s)\|^2 = \theta'(s)^2$, donc pour tout s on a $\theta'(s) = \pm\kappa(s)$ et par continuité $\theta' = \varepsilon\kappa$ où $\varepsilon = \pm 1$. Comme

$\{\vec{t} = (\cos\theta, \sin\theta); \vec{n} = (-\varepsilon \sin\theta, \varepsilon\cos\theta)\}$ est une base directe de $\mathbb{R}^2$, on a $\varepsilon = \begin{vmatrix} \cos\theta & -\varepsilon\sin\theta \\ \sin\theta & \varepsilon\cos\theta \end{vmatrix} = 1$, donc $\theta' = \kappa$. On en déduit l'expression de θ. Intégrer x' et y' donne le résultat. $\square$

F Théorème de reconstruction

Théorème F.1. (Théorème fondamental de la théorie locale des courbes) Soient $\kappa, \tau : I \to \mathbb{R}$ des applications de classe $\mathcal{C}^\infty$ avec $\kappa(s) > 0$ et $\tau(s) \neq 0$ pour tout $s \in I$. Alors il existe une courbe paramétrée régulière $\alpha : I \to \mathbb{R}^3$ telle que s est l'abscisse curviligne, $\kappa(s)$ est la courbure, et $\tau(s)$ est la torsion de α. De plus, toute autre courbe $\overline{\alpha}$ satisfaisant aux mêmes conditions est obtenue à partir de α en lui appliquant une isométrie positive, c'est-à-dire qu'il existe un endomorphisme orthogonal ρ de $\mathbb{R}^3$, de déterminant positif (c'est-à-dire que ρ est une rotation), et un vecteur A tel que $\overline{\alpha} = \rho \circ \alpha + A$.

(admis)

G Courbure et torsion de courbes qui ne sont pas à vitesse normalisée

Soit $\alpha : I \to \mathbb{R}^3$ une courbe paramétrée régulière de classe $\mathcal{C}^\infty$ avec une paramétrisation arbitraire. Notons s l'abscisse curviligne. Alors

$$\kappa(s) = \frac{\|\alpha'(t) \wedge \alpha''(t)\|}{\|\alpha'(t)\|^3} \text{ et } \tau(s) = \frac{\langle \alpha'(t) \wedge \alpha''(t), \alpha'''(t)\rangle}{\|\alpha'(t) \wedge \alpha''(t)\|^2}.$$

(Ces formules ne sont pas à connaître.)