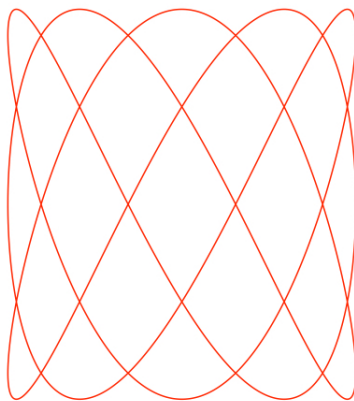


Polycopié de L2 MA223: Géométrie

Par Frédéric Naud



Septembre 2025

Avant propos...

Ce polycopié est destiné aux étudiants de L2 maths à Sorbonne Université. Sous l'intitulé "Géométrie" on a regroupé l'étude des espaces Euclidiens et Hermitiens et leurs isométries, puis un chapitre consacré à l'étude des courbes planes : courbes paramétrées, implicites et leurs propriétés métriques (courbure, longueur). Il contient aussi une introduction rudimentaire aux surfaces, en faisant appel au moins de calcul différentiel possible.

Ce polycopié a été rédigé intégralement par un être humain et non pas par Chat-GPT ou autre Deep-Seek. Il contient donc des typos que vous ne manquerez pas de signaler à l'auteur pour améliorer le texte. En revanche, contrairement aux LLM, les preuves de ce poly sont garanties sans hallucinations et sont en principe correctes. Certaines preuves faciles sont laissées en exercice et souvent reprises en TD, mais les plus difficiles sont détaillées au maximum.

Les prérequis pour ce cours sont l'algèbre linéaire de L1 et l'analyse réelle de L1. Tout est élémentaire, des compléments sur l'intégration des fonctions à valeur complexes sont donnés en annexe.

Table des matières

1	Espaces Euclidiens et Hermitiens	4
1.1	Produit scalaire, définitions	4
1.2	Bases orthonormées et algorithme de Gram-Schmidt	9
1.3	Distance et projection orthogonale	12
1.4	Endomorphismes et adjoints	15
1.5	Le théorème de représentation de Riesz	16
2	Groupes d'isométries	19
2.1	Isométries	19
2.2	Groupes unitaires et orthogonaux	21
2.3	Angles et orientation	22
2.4	Etude de $O_2(\mathbb{R})$ et $O_3(\mathbb{R})$: les isométries du plan et de l'espace	24
2.5	Le produit vectoriel dans \mathbb{R}^3	28
2.6	Produit vectoriel et rotations	30
2.7	Compléments : le groupe $SU_2(\mathbb{C})$	32
3	Réduction des endomorphismes auto-adjoints	35
3.1	Sur les déterminants	35
3.2	Rappels sur la diagonalisation des endomorphismes	40
3.3	Diagonalisation des endomorphismes auto-adjoints	46
3.4	Une application : la décomposition polaire dans $GL_n(\mathbb{R})$	49
4	Courbes planes	53
4.1	Courbes paramétrées, généralités	53
4.2	Etude locale dans le plan	56
4.3	Exemples de courbes implicites, coniques	63
4.3.1	Cas où on peut isoler une variable.	64
4.3.2	Exemples de courbes unicursales.	65
4.3.3	Cas quadratique : les coniques.	66
4.4	Courbure dans le plan	70
4.5	La formule de Gauss-Bonnet en dimension 1	73

5	Courbes et surfaces dans \mathbb{R}^3	77
5.1	Courbes gauches	77
5.1.1	Courbure dans \mathbb{R}^3	77
5.1.2	Torsion dans \mathbb{R}^3	78
5.2	Introduction aux surfaces	81
5.2.1	Rudiments de calcul différentiel en deux variables.	81
5.2.2	Surfaces paramétrées, exemples	84
5.2.3	La première forme fondamentale	87
5.2.4	Deuxième forme fondamentale, courbure	89
A	Intégration à valeur complexe et relèvement C^1.	96

Chapitre 1

Espaces Euclidiens et Hermitiens

1.1 Produit scalaire, définitions

Dans ce qui suit \mathbb{K} désigne un corps qui en pratique sera $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . Si $z = x + iy \in \mathbb{C}$, avec $x, y \in \mathbb{R}$, on note son conjugué par $\bar{z} = x - iy$. Si $M = (m_{ij}) \in \mathcal{M}_{p,q}(\mathbb{K})$ est une matrice à p lignes et q colonnes, on note tM sa transposée, qui est une matrice à q lignes et p colonnes définie par $({}^tM)_{ij} = m_{ji}$, pour tout $1 \leq i \leq q$ et $1 \leq j \leq p$.

Soit E un \mathbb{K} -espace vectoriel de dimension finie. Le vecteur nul de E sera noté 0_E si il y a une ambiguïté.

Définition 1.1.1 Une forme bilinéaire φ sur E est une application

$$\varphi : E \times E \rightarrow \mathbb{K},$$

qui vérifie

1. Pour tout $u_1, u_2, v \in E$, pour tout $\lambda \in \mathbb{K}$,

$$\varphi(u_1 + \lambda u_2, v) = \varphi(u_1, v) + \lambda \varphi(u_2, v).$$

2. Pour tout $u, v_1, v_2 \in E$, pour tout $\lambda \in \mathbb{K}$,

$$\varphi(u, v_1 + \lambda v_2) = \varphi(u, v_1) + \lambda \varphi(u, v_2).$$

Définition 1.1.2 Dans le cas où $K = \mathbb{C}$, une forme sesquilinéaire φ sur E est une application

$$\varphi : E \times E \rightarrow \mathbb{K},$$

qui vérifie

1. (linéarité à gauche) Pour tout $u_1, u_2, v \in E$, pour tout $\lambda \in \mathbb{K}$,

$$\varphi(u_1 + \lambda u_2, v) = \varphi(u_1, v) + \lambda \varphi(u_2, v).$$

2. (semi-linéarité à droite) Pour tout $u, v_1, v_2 \in E$, pour tout $\lambda \in \mathbb{K}$,

$$\varphi(u, v_1 + \lambda v_2) = \varphi(u, v_1) + \bar{\lambda} \varphi(u, v_2).$$

Remarque. La bilinéarité (ou la sesquilinearité) entraîne automatiquement que $\varphi(0_E, x) = 0$ pour tout $x \in E$. En effet, $\varphi(0_E, x) = \varphi(0_E + 0_E, x) = \varphi(0_E, x) + \varphi(0_E, x)$ et donc en simplifiant $\varphi(0_E, x) = 0$. Même chose pour $\varphi(x, 0_E)$.

Définition 1.1.3 1. Si φ est une forme bilinéaire et si on a pour tout $x, y \in E$, $\varphi(x, y) = \varphi(y, x)$, on dit que φ est symétrique.
 2. Si φ est une forme sesquilinéaire et si on a pour tout $x, y \in E$, $\varphi(x, y) = \overline{\varphi(y, x)}$, on dit que φ est hermitienne symétrique.

Exemples.

1. Si $E = \mathbb{R}^n$, et si $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$, alors $\varphi(x, y) = \sum_{i=1}^n x_i y_i$ est bien une forme bilinéaire symétrique.
2. Si $E = \mathbb{C}^n$, et si $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$, alors $\varphi(x, y) = \sum_{i=1}^n x_i \bar{y}_i$ définit une forme hermitienne symétrique.
3. Si $S = (S_{ij})_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbb{R})$ est une matrice réelle de taille $n \times n$ qui est symétrique i.e. $S_{ij} = S_{ji}$ pour tout i, j , alors pour

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

$$\varphi(X, Y) = {}^t X S Y,$$

définit une forme bilinéaire symétrique sur \mathbb{R}^n .

4. Si $E = \mathbb{R}_n[X]$ est le \mathbb{R} -espace vectoriel des polynômes réels de degré inférieur ou égal à n , la formule

$$\varphi(P, Q) := \int_{-1}^{+1} P(t) Q(t) dt$$

définit une forme bilinéaire symétrique sur E .

5. Si $E = \mathcal{T}_n$, le \mathbb{C} -espace des polynômes trigonométriques P sur \mathbb{R} de la forme

$$P(t) = \sum_{k=-n}^{+n} a_k e^{ikt},$$

où les coefficients $a_k \in \mathbb{C}$. Alors

$$\varphi(P, Q) := \frac{1}{2\pi} \int_0^{2\pi} P(t) \overline{Q(t)} dt$$

est une forme sesquilinéaire sur \mathcal{T}_n .

Définition 1.1.4 Si φ est une forme bilinéaire (si $K = \mathbb{R}$) ou sesquilinéaire (si $K = \mathbb{C}$) on dit que φ est positive ssi pour tout $x \in E$, $\varphi(x, x) \in \mathbb{R}^+$ et définie ssi pour tout $x \in E$,

$$\varphi(x, x) = 0 \Rightarrow x = 0.$$

Remarque. Si φ est une forme sesquilinéaire hermitienne symétrique, alors pour tout $x \in E$, $\varphi(x, x) = \overline{\varphi(x, x)}$, donc $\varphi(x, x) \in \mathbb{R}$.

Définition 1.1.5 Soit E un K -espace vectoriel de dimension finie. E est appelé espace Euclidien (resp. Hermitien) s'il est muni d'une forme bilinéaire symétrique définie positive (resp. sesquilinéaire symétrique définie positive). Une telle forme est appelée **produit scalaire**.

Exemples : les exemples 1), 2), 4) et 5) ci-dessus sont des produits scalaires.

Proposition 1.1.6 (Inégalité de Cauchy-Schwarz). Soit (E, φ) un espace Euclidien ou Hermitien. On a pour tout $x, y \in E$,

$$|\varphi(x, y)| \leq \sqrt{\varphi(x, x)} \sqrt{\varphi(y, y)},$$

avec égalité ssi x, y sont colinéaires.

Preuve. On commence par traiter le cas réel ou $K = \mathbb{R}$. Se fixant $x, y \in E$, On étudie la fonction

$$\lambda \mapsto f(\lambda) := \varphi(x + \lambda y, x + \lambda y).$$

Par bilinéarité et symétrie de φ , on a donc

$$f(\lambda) = \varphi(x, x) + 2\lambda\varphi(x, y) + \lambda^2\varphi(y, y).$$

Comme φ est positive, on a donc $f(\lambda) \geq 0$ pour tout $\lambda \in \mathbb{R}$, ainsi le discriminant de cette fonction polynomiale de degré au plus 2 doit être négatif. Ceci donne $4(\varphi(x, y))^2 - 4\varphi(y, y)\varphi(x, x) \leq 0$, ce qui entraîne

$$(\varphi(x, y))^2 \leq \varphi(y, y)\varphi(x, x),$$

puis en prenant la racine carrée

$$|\varphi(x, y)| \leq \sqrt{\varphi(x, x)}\sqrt{\varphi(y, y)}.$$

Si il y a égalité, c'est que le discriminant est nul et que $f(\lambda_0) = 0$ pour un certain $\lambda_0 \in \mathbb{R}$. Comme φ est définie, ceci entraîne $x + \lambda_0 y = 0$, (x, y) est bien liée.

Cas complexe $\mathbb{K} = \mathbb{C}$. On s'inspire des mêmes idées. On pose pour $\lambda \in \mathbb{R}$,

$$f(\lambda) := \varphi(x + \lambda y, x + \lambda y).$$

Quitte à multiplier x par $e^{i\theta}$ avec $\theta \in \mathbb{R}$ bien choisi, on peut supposer $\varphi(x, y) \in \mathbb{R}$. On a alors, comme dans le cas réel,

$$f(\lambda) = \varphi(x, x) + 2\lambda\varphi(x, y) + \lambda^2\varphi(y, y),$$

et on suit exactement le même raisonnement. \square

Proposition 1.1.7 *Soit (E, φ) un espace Euclidien ou Hermitien. On pose pour tout $x \in E$,*

$$\|x\| := \sqrt{\varphi(x, x)}.$$

Alors $\|\cdot\|$ est une norme sur E , i.e. elle vérifie :

1. *Pour tout $x \in E$, $\|x\| = 0 \Rightarrow x = 0$.*
2. *Pour tout $(x, \lambda) \in E \times \mathbb{K}$, $\|\lambda x\| = |\lambda|\|x\|$.*
3. *Pour tout $x, y \in E$, $\|x + y\| \leq \|x\| + \|y\|$.*

Preuve. Les points 1) et 2) sont évidents. Prouvons 3) dans le cas complexe. Le cas réel étant analogue. Par bilinéarité et symétrie, on a

$$\|x + y\|^2 = \|x\|^2 + 2\operatorname{Re}(\varphi(x, y)) + \|y\|^2.$$

Mais par Cauchy-Schwarz on a aussi

$$|\operatorname{Re}(\varphi(x, y))| \leq |\varphi(x, y)| \leq \|x\|\|y\|.$$

Ainsi on a

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

la preuve est finie. \square

Lorsque deux vecteurs x, y vérifient $\varphi(x, y) = 0$, on dit qu'ils sont *orthogonaux*. On a l'importante remarque suivante.

Proposition 1.1.8 (*Thm de Pythagore*) Si $v_1, \dots, v_n \in E$ sont deux à deux orthogonaux, alors

$$\left\| \sum_{i=1}^n v_i \right\|^2 = \sum_{i=1}^n \|v_i\|^2.$$

Preuve. On écrit par bilinéarité

$$\left\| \sum_{i=1}^n v_i \right\|^2 = \varphi \left(\sum_i v_i, \sum_j v_j \right) = \sum_{i=1}^n \sum_{j=1}^n \varphi(v_i, v_j),$$

comme pour $i \neq j$ on a $\varphi(v_j, v_i) = 0$, la somme double ci-dessus se réduit à

$$\left\| \sum_{i=1}^n v_i \right\|^2 = \sum_{i=1}^n \varphi(v_i, v_i) = \sum_{i=1}^n \|v_i\|^2,$$

la preuve est finie. \square

La norme $\| \cdot \|$ associée à un produit scalaire est appelée norme euclidienne (ou hermitienne). La connaissance de la norme détermine complètement le produit scalaire. En effet on a les identités suivantes (dites de **polarisation**).

Proposition 1.1.9 (*Formules de polarisation*)

— (*Cas euclidien*) On a pour tout $x, y \in E$,

$$\varphi(x, y) = \frac{1}{2}(\|x + y\|^2 - \|x\|^2 - \|y\|^2).$$

— (*Cas hermitien*) On a pour tout $x, y \in E$,

$$\varphi(x, y) = \frac{1}{2}(\|x + y\|^2 - \|x\|^2 - \|y\|^2 + i(\|x + iy\|^2 - \|x\|^2 - \|y\|^2)).$$

La preuve (facile) est laissée en exercice. Les normes euclidiennes vérifient aussi d'autres identités remarquables dont *l'identité du parallélogramme* : pour tout $x, y \in E$, on a

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

On peut montrer (voir TD) que cette identité caractérise les normes euclidiennes.

1.2 Bases orthonormées et algorithme de Gram-Schmidt

Définition 1.2.1 Soit (E, φ) un espace Euclidien ou Hermitien. Une base (e_1, \dots, e_n) de E est dite orthonormée ssi on a les propriétés suivantes.

1. (Normalisation) Pour tout $i = 1, \dots, n$ on a $\|e_i\| = 1$.
2. (Orthogonalité) Pour tout $i \neq j \in \{1, \dots, n\}$, on a $\varphi(e_i, e_j) = 0$.

Remarque. Une famille orthogonale de vecteurs est automatiquement libre (à vérifier en exercice).

Exemples. Dans \mathbb{R}^n , muni du produit scalaire

$$\langle x, y \rangle := \sum_{i=1}^n x_i y_i,$$

la base canonique $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, ..., $e_n = (0, \dots, 0, 1)$ est une base orthonormée (le vérifier).

Dans l'espace des polynômes trigonométriques \mathcal{T}_n , la famille

$$\{t \mapsto e^{ikt} : -n \leq k \leq n\}$$

est orthonormée (exercice). On en déduit donc par Pythagore que si

$$P(t) = \sum_k a_k e^{ikt},$$

on a

$$\|P\|^2 = \frac{1}{2\pi} \int_0^{2\pi} |P(t)|^2 dt = \sum_k |a_k|^2,$$

c'est l'identité de Parseval.

L'intérêt des bases orthonormées est de ramener tout les calculs de norme et de produit scalaire à des calculs simples sur les coordonnées.

Proposition 1.2.2 Soit (E, φ) un espace Euclidien ou Hermitien muni d'une base (e_1, \dots, e_n) orthonormée. Alors on a les faits suivants.

1. Pour tout $x, y \in E$ tels que $x = \sum_i x_i e_i$ et $y = \sum_i y_i e_i$, on a (cas réel)

$$\varphi(x, y) = \sum_{i=1}^n x_i y_i,$$

ou dans le cas complexe,

$$\varphi(x, y) = \sum_{i=1}^n x_i \overline{y_i}.$$

2. Pour tout $x = \sum_i x_i e_i$, on a

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

3. Pour tout $x \in E$, on a

$$x = \sum_{i=1}^n \varphi(x, e_i) e_i.$$

Preuve. Exercice. \square

Que se passe-t-il dans une base non-orthogonale? On peut toujours donner une expression pour calculer le produit scalaire via la matrice de Gram.

Proposition 1.2.3 Soit (E, φ) un espace Euclidien ou Hermitien muni d'une base (e_1, \dots, e_n) , pas forcément orthonormée. Considérons la matrice G définie par

$$G := \begin{pmatrix} \varphi(e_1, e_1) & \dots & \varphi(e_1, e_n) \\ \vdots & \ddots & \vdots \\ \varphi(e_n, e_1) & \dots & \varphi(e_n, e_n) \end{pmatrix}.$$

Alors pour tout $x, y \in E$, avec $x = \sum_i x_i e_i$ et $y = \sum_j y_j e_j$, on a (dans le cas euclidien)

$$\varphi(x, y) = (x_1 \dots x_n) G \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

et dans le cas hermitien

$$\varphi(x, y) = (x_1 \dots x_n) G \begin{pmatrix} \overline{y_1} \\ \vdots \\ \overline{y_n} \end{pmatrix}.$$

Preuve. Cas euclidien. On écrit simplement par bilinéarité :

$$\varphi(x, y) = \varphi\left(\sum_i x_i e_i, \sum_j y_j e_j\right) = \sum_i \sum_j x_i y_j \varphi(e_i, e_j),$$

et on reconnaît bien le produit matriciel

$$\varphi(x, y) = (x_1 \dots x_n) G \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Le cas hermitien est pareil (avec des conjugués). \square

On remarque au passage qu'une base est orthonormée ssi la matrice de Gram est l'identité. On sait que tout espace vectoriel E de dimension finie admet une base. Si E est de plus doté d'une structure euclidienne ou hermitienne, il n'est pas évident qu'il possède une base orthonormée. On va décrire ci-dessous un algorithme, dit de **Gram-Schmidt**, qui permet, partant d'une base quelconque, de fabriquer une base orthonormée.

Soit donc (v_1, \dots, v_n) une base de E , on va construire par récurrence une base orthonormée (e_1, \dots, e_k) de E telle que pour tout $k = 1, \dots, n$

$$\text{Vect}(e_1, \dots, e_k) = \text{Vect}(v_1, \dots, v_k).$$

- Etape $k = 1$. Il n'y a presque rien à faire, on pose $e_1 = \frac{v_1}{\|v_1\|}$ de telle sorte que $\|e_1\| = 1$.
- Etape $k = 2$. Posons $e'_2 = v_2 - \varphi(v_2, e_1)e_1$. On a bien $e'_2 \in \text{Vect}(v_1, v_2)$ et $\varphi(e'_2, e_1) = 0$. Il reste à normaliser en posant $e_2 = \frac{e'_2}{\|e'_2\|}$.
- Supposons avoir construit e_1, \dots, e_k orthonormée de telle sorte que

$$\text{Vect}(e_1, \dots, e_k) = \text{Vect}(v_1, \dots, v_k).$$

On pose alors

$$e'_{k+1} = v_{k+1} - \sum_{j=1}^k \varphi(v_{k+1}, e_j)e_j.$$

Il est clair que $e'_{k+1} \in \text{Vect}(v_1, \dots, v_{k+1})$ et on vérifie sans peine que

$$\varphi(e'_{k+1}, u) = 0$$

pour tout $u \in \text{Vect}(e_1, \dots, e_k)$. On termine en posant $e_{k+1} = \frac{e'_{k+1}}{\|e'_{k+1}\|}$.

- On poursuit ce procédé jusqu'à $k = n$.

Corollaire 1.2.4 *Tout espace euclidien ou hermitien possède une base orthonormée.*

Illustrons la méthode de Gram-Schmidt sur un exemple concret. On considère ici $E = \mathbb{R}^3$, muni du produit scalaire standard

$$\langle x, y \rangle = \sum_{i=1}^3 x_i y_i.$$

Considérons la famille $v_1 = (1, 0, 0)$, $v_2 = (1, 1, 0)$ et $v_3 = (1, 1, 1)$. On laisse le soin au lecteur de justifier qu'il s'agit bien d'une famille libre et donc d'une base de \mathbb{R}^3 . On a $\|v_1\| = 1$, donc on pose $e_1 = v_1$. On calcule ensuite $e'_2 = v_2 - \langle v_2, e_1 \rangle e_1 = (1, 1, 0) - (1, 0, 0) = (0, 1, 0)$. Comme on a $\|e'_2\| = 1$, on a $e_2 = (0, 1, 0)$. Enfin, on calcule $e'_3 = v_3 - \langle v_3, e_1 \rangle e_1 - \langle v_3, e_2 \rangle e_2 = (1, 1, 1) - (1, 0, 0) - (0, 1, 0) = (0, 0, 1)$. Finalement on a obtenu

$$e_1 = (1, 0, 0), \quad e_2 = (0, 1, 0), \quad e_3 = (0, 0, 1),$$

c'est la base canonique de \mathbb{R}^3 !

1.3 Distance et projection orthogonale

Définition 1.3.1 Soit (E, φ) un espace euclidien ou hermitien. Soit $A \subset E$ une partie de E . On appelle orthogonal de A , noté A^\perp , la partie définie par

$$A^\perp := \{x \in E : \forall a \in A, \varphi(x, a) = 0\}.$$

A^\perp est automatiquement un sous-espace vectoriel de E (exercice).

Proposition 1.3.2 Soit (E, φ) un espace euclidien ou hermitien, et soit F un sous-espace vectoriel de E ¹. On a toujours

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

De plus, il existe un unique endomorphisme $P_F : E \rightarrow F$ tel que $P_F \circ P_F = P_F$ avec $\text{IM}(P_F) = F$ et $\text{Ker}(P_F) = F^\perp$. On l'appelle projection orthogonale sur F . Elle est donnée par l'expression suivante, valable pour toute base orthonormée (e_1, \dots, e_p) de F ,

$$P_F(x) := \sum_{i=1}^p \varphi(x, e_i) e_i.$$

Preuve. Si $x \in F \cap F^\perp$, il vient $\varphi(x, x) = 0$, et comme φ est définie, on a donc $x = 0$. Ainsi $F \cap F^\perp = \{0\}$. On considère (e_1, \dots, e_p) une base orthonormée de F , et on pose pour tout $x \in E$, $P_F(x) := \sum_{i=1}^p \varphi(x, e_i) e_i$. Clairement $P_F(x) \in F$. On observe de plus que $x - P_F(x) \in F^\perp$. En effet, on a pour tout $y \in F$,

$$\varphi(x - P_F(x), y) = \varphi(x, y) - \varphi(P_F(x), y) = \varphi(x, y) - \sum_{i=1}^p \varphi(x, e_i) \varphi(e_i, y).$$

1. On exclut les cas pathologiques $F = \{0\}$ et $F = E$.

De plus on a (dans le cas hermitien)

$$\begin{aligned} \sum_{i=1}^p \varphi(x, e_i) \varphi(e_i, y) &= \varphi \left(x, \sum_i \overline{\varphi(e_i, y)} e_i \right) = \varphi \left(x, \sum_i \varphi(y, e_i) e_i \right) \\ &= \varphi(x, y), \end{aligned}$$

ainsi $\varphi(x - P_F(x), y) = 0$. L'écriture $x = P_F(x) + x - P_F(x)$ montre donc que $F + F^\perp = E$, et le premier point est prouvé. On vérifie sans peine que $x \mapsto P_F(x)$ est une application linéaire telle que $P_F \circ P_F = P_F$. On observe que $x \in \text{Ker}(P_F)$ ssi pour tout $i = 1, \dots, p$, on a $\varphi(x, e_i) = 0$, ce qui équivaut à avoir $\varphi(x, y) = 0$ pour tout $y \in F$ (car (e_1, \dots, e_p) est une base de F). On a donc bien $\text{Ker}(P_F) = F^\perp$. Comme $\text{IM}(P_F) \subset F$, et que $\dim(F) + \dim(F^\perp) = \dim(E)$, le théorème du rang appliqué à P_F montre que $\dim(\text{IM}(P_F)) = \dim(F)$ et donc $\text{IM}(P_F) = F$. Il reste à voir que P_F est unique. Observons d'abord que si $z \in F$, on a $z = P_F(x)$ pour un certain $x \in E$, et donc $P_F(z) = P_F \circ P_F(x) = P_F(x) = z$. Soit maintenant $x \in E$. Comme $E = F \oplus F^\perp$, on a une décomposition unique $x = x_F + x_{F^\perp}$, et ainsi

$$P_F(x) = P_F(x_F) + P_F(x_{F^\perp}) = P_F(x_F) = x_F.$$

L'application P_F est donc unique. \square

Théorème 1.3.3 *Soit (E, φ) un espace euclidien ou hermitien, et soit F un sous-espace vectoriel de E . Se fixant $x \in E$, on a*

$$d(x, F) := \inf_{y \in F} \|x - y\| = \|x - P_F(x)\|,$$

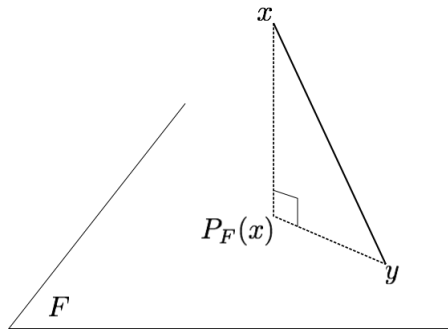
et cet inf est atteint uniquement en $y = P_F(x)$. Autrement dit, le projeté orthogonal de x sur F minimise la distance de x au sous-espace F .

Preuve. Il est clair que puisque $P_F(x) \in F$,

$$\inf_{y \in F} \|x - y\| \leq \|x - P_F(x)\|.$$

Soit $y \in F$, on écrit

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



Projection orthogonale sur F

En observant que $P_F(x) - y \in F$ et sachant que $x - P_F(x) \in F^\perp$, on a donc par Pythagore (voir figure ci-dessus)

$$\|x - y\|^2 = \|x - P_F(x)\|^2 + \|P_F(x) - y\|^2 \geq \|x - P_F(x)\|^2,$$

et donc en passant à l'inf on a

$$\inf_{y \in F} \|x - y\| \geq \|x - P_F(x)\|,$$

d'où l'égalité. On remarque de plus que si $y \in F$ est tel que $\|x - y\| = \|x - P_F(x)\|$, alors Pythagore montre que $P_F(x) = y$. \square .

Exemple. On se place dans $E = \mathbb{R}^3$ muni du produit scalaire usuel noté $\langle ., . \rangle$, et on considère

$$F = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}.$$

L'ensemble F est un sous-espace vectoriel et on va calculer la distance $d((1, 2, 1), F)$.

On observe d'abord que $(x, y, z) \in F$ ssi $(x, y, z) = (-y - z, y, z) = y(-1, 1, 0) + z(-1, 0, 1)$. Ainsi $F = \text{Vect}\{(-1, 1, 0), (-1, 0, 1)\}$ et $v_1 = (-1, 1, 0)$, $v_2 = (-1, 0, 1)$ constitue une base de F . Pour calculer la projection orthogonale P_F , il faut orthonormaliser cette base. On a $\|v_1\| = \sqrt{2}$ donc on pose $e_1 = \frac{1}{\sqrt{2}}(-1, 1, 0)$.

On calcule ensuite

$$e'_2 = v_2 - \langle v_2, e_1 \rangle e_1 = (-1, 0, 1) - \frac{1}{2}(-1, 1, 0) = (-1/2, -1/2, 1).$$

On a donc $e_2 = \frac{e'_2}{\|e'_2\|} = \sqrt{\frac{2}{3}}(-1/2, -1/2, 1)$. On calcule donc maintenant

$$\begin{aligned} P_F(x, y, z) &= \langle (x, y, z), e_1 \rangle e_1 + \langle (x, y, z), e_2 \rangle e_2 \\ &= \frac{1}{2}(y - x)(-1, 1, 0) + \frac{2}{3}(z - x/2 - y/2)(-1/2, -1/2, 1). \end{aligned}$$

Ainsi $P_F((1, 2, 1)) = (-1/3, 2/3, -1/3)$, et

$$d((1, 2, 1), F) = \|(1, 2, 1) - P_F((1, 2, 1))\| = \|(4/3, 4/3, 4/3)\| = 4 \frac{\sqrt{3}}{3}.$$

1.4 Endomorphismes et adjoints

Pour tout endomorphisme T d'un espace euclidien ou hermitien, on va définir l'adjoint T^* via le produit scalaire.

Proposition 1.4.1 *Soit (E, φ) un espace euclidien ou hermitien. Soit $f : E \rightarrow E$ une application linéaire. Il existe une unique application linéaire $f^* : E \rightarrow E$ telle que pour tout $x, y \in E$,*

$$\varphi(f(x), y) = \varphi(x, f^*(y)).$$

Preuve. Unicité. Si on a pour tout $x, y \in E$,

$$\varphi(x, f_1(y)) = \varphi(f(x), y) = \varphi(x, f_2(y)),$$

on a donc pour tout $x, y \in E$,

$$\varphi(x, f_1(y) - f_2(y)) = 0,$$

et en posant $x = f_1(y) - f_2(y)$ et en utilisant que φ est définie on a $f_1(y) = f_2(y)$ pour tout $y \in E$. Existence. On traite le cas hermitien, le cas euclidien est analogue, mais sans conjugués. On considère (e_1, e_2, \dots, e_n) une base orthonormée de (E, φ) . On définit alors l'endomorphisme f^* en posant pour tout $j = 1, \dots, n$,

$$f^*(e_j) := \sum_{i=1}^n \overline{\varphi(f(e_i), e_j)} e_i.$$

Avec cette définition, on a donc pour tout i, j ,

$$\varphi(e_i, f^*(e_j)) = \varphi(f(e_i), e_j).$$

On vérifie alors que si $x = \sum_i x_i e_i$, $y = \sum_j y_j e_j$, on a

$$\varphi(f(x), y) = \sum_{i,j} x_i \overline{y_j} \varphi(f(e_i), e_j) = \sum_{i,j} x_i \overline{y_j} \varphi(e_i, f^*(e_j)),$$

et par sesquilinearité et linéarité de f^* , on tombe sur

$$\varphi(f(x), y) = \varphi(x, f^*(y)).$$

La preuve est finie. \square

Proposition 1.4.2 *Soit (E, φ) un espace Euclidien ou Hermitien. Soit f un endomorphisme et \mathcal{B} une base orthonormée de E . Alors on a l'identité matricielle :*

$$\begin{aligned} \text{Mat}_{\mathcal{B}}(f^*) &= {}^t \overline{\text{Mat}_{\mathcal{B}}(f)} \quad (\text{cas Hermitien}), \\ \text{Mat}_{\mathcal{B}}(f^*) &= {}^t \text{Mat}_{\mathcal{B}}(f) \quad (\text{cas Euclidien}). \end{aligned}$$

Preuve. C'est fait dans le calcul précédent : dans toute base orthonormée (e_1, e_2, \dots, e_n) , on a

$$\varphi(f^*(e_j), e_i) = \overline{\varphi(f(e_i), e_j)},$$

d'où l'identité matricielle. \square

Terminons cette section par une identité bien utile.

Proposition 1.4.3 *Soit (E, φ) un espace euclidien ou hermitien. Soit $f : E \rightarrow E$ un endomorphisme. On a alors*

$$\text{Ker}(f) = \text{IM}(f^*)^\perp.$$

Preuve. On a $x \in \text{IM}(f^*)^\perp$ ssi pour tout $y \in E$, $\varphi(x, f^*(y)) = 0$. Mais on a donc pour tout $y \in E$,

$$\varphi(x, f^*(y)) = \varphi(f(x), y) = 0.$$

Ceci équivaut donc à $f(x) = 0$, car φ est définie, et donc $x \in \text{IM}(f^*)^\perp$ ssi $x \in \text{Ker}(f)$. \square

Remarque. Comme $(F^\perp)^\perp = F$, en remplaçant f par f^* , on a aussi

$$\text{IM}(f) = \text{Ker}(f^*)^\perp.$$

Terminologie. Si f est un endomorphisme tel que $f = f^*$, il est dit *auto-adjoint*. Si f est un isomorphisme et que $f^{-1} = f^*$, on dit qu'il est *unitaire*. Si on a $f \circ f^* = f^* \circ f$, on dit qu'il est *normal*. Toute application linéaire auto-adjointe ou unitaire est normale, mais il existe bien sur des endomorphismes normaux qui ne sont ni autoadjoint, ni unitaires, par exemple si U est unitaire et si $P \in \mathbb{C}[X]$ est un polynôme, $P(U)$ est normal mais en général pas unitaire ni auto-adjoint.

1.5 Le théorème de représentation de Riesz

Si E est un \mathbb{K} -espace vectoriel, on appelle *dual de E* , noté E^* , le \mathbb{K} -espace vectoriel des applications linéaires de $E \rightarrow \mathbb{K}$ (les formes linéaires).

Proposition 1.5.1 *Si E est de dimension finie, alors on a $\dim(E) = \dim(E^*)$.*

Preuve. Soit (e_1, \dots, e_n) une base de E . Pour tout $i = 1, \dots, n$, on note e_i^* la forme linéaire définie sans équivoque par

$$e_i^* \left(\sum_j x_j e_j \right) := x_i.$$

On va montrer que (e_1^*, \dots, e_n^*) est une base de E^* . C'est une famille libre. En effet si $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ sont tels que

$$\sum_i \lambda_i e_i^* = 0,$$

alors pour tout $x \in E$, on a

$$\sum_i \lambda_i e_i^*(x) = 0,$$

en particulier pour tout $j = 1, \dots, n$, on a

$$\sum_i \lambda_i e_i^*(e_j) = 0 = \lambda_j,$$

ainsi $\lambda_1 = \dots = \lambda_n = 0$. C'est une famille génératrice. Soit $\ell \in E^*$, on écrit pour tout $x \in E$,

$$\ell(x) = \ell\left(\sum_i x_i e_i\right) = \sum_i x_i \ell(e_i) = \sum_i e_i^*(x) \ell(e_i),$$

ainsi on a $\ell = \sum_i \ell(e_i) e_i^*$ et la preuve est finie. \square

Théorème 1.5.2 (*Riesz*) *On suppose ici que (E, φ) est euclidien ou hermitien, donc de dimension finie. Pour tout $\ell \in E^*$, il existe un unique vecteur $y \in E$ tel que pour tout $x \in E$ on a*

$$\ell(x) = \varphi(x, y).$$

Preuve. On se limite pour simplifier au cas $\mathbb{K} = \mathbb{R}$. L'idée est la même dans le cas de \mathbb{C} mais il faut faire attention à la semi-linéarité. On considère l'application $\Phi : E \rightarrow E^*$, donnée pour tout $y \in E$ par

$$\Phi(y)(x) = \varphi(x, y).$$

Cette application Φ est bien définie car $x \mapsto \varphi(x, y)$ est bien une forme linéaire. Elle est de plus linéaire car par bilinéarité de φ , on a bien

$$\Phi(y + \lambda z)(x) = \varphi(x, y + \lambda z) = \varphi(x, y) + \lambda \varphi(x, z) = \Phi(y)(x) + \lambda \Phi(z)(x).$$

Calculons le noyau de Φ . Si $\Phi(y) = 0$, cela veut dire que pour tout $x \in E$, $\varphi(x, y) = 0$, mais comme φ est définie, on a $y = 0$. Ainsi Φ est injective. Comme $\dim(E) = \dim(E^*)$, le théorème du rang montre que Φ est surjective. \square

Un exemple. On considère $E = \mathbb{R}_n[X]$, muni du produit scalaire

$$\langle P, Q \rangle = \sum_{i=0}^n a_i b_i,$$

où $P(X) = \sum_i a_i X^i$ et $Q(X) = \sum_i b_i X^i$. L'application $P \mapsto \int_0^1 P(t) dt$ est une forme linéaire. Le théorème de Riesz nous dit qu'il existe un unique polynôme $Q_0 = \sum_i b_i X^i$ tel que pour tout $P = \sum_i a_i X^i \in \mathbb{R}_n[X]$, on a

$$\int_0^1 P(t) dt = \langle P, Q_0 \rangle = \sum_i a_i b_i.$$

En effet on a

$$\int_0^1 P(t) dt = \sum_i a_i \int_0^1 t^i dt = \sum_i \frac{a_i}{i+1} = \langle P, Q_0 \rangle,$$

avec $Q_0(X) = \sum_i \frac{1}{i+1} X^i$.

Le théorème de Riesz, et son extension en dimension infinie dans les espaces de Hilbert, est un outil très puissant en algèbre et en Analyse, voir cours de L3 et M1. Comme son nom l'indique, il permet de représenter toute forme linéaire via le produit scalaire.

Chapitre 2

Groupes d'isométries

2.1 Isométries

Définition 2.1.1 Soit (E, φ) un espace euclidien ou hermitien. Une application linéaire $f : E \rightarrow E$ est appelée isométrie ssi on a pour tout $x \in E$, $\|f(x)\| = \|x\|$, où $\| \cdot \|$ est la norme euclidienne (ou hermitienne) associée au produit scalaire φ .

Dans le cas euclidien, on parle de transformation *orthogonale*, dans le cas hermitien, on parle de transformation *unitaire*.

Proposition 2.1.2 Soit (E, φ) un espace euclidien ou hermitien et f un endomorphisme. On a alors les faits équivalents suivants.

1. Pour tout $x \in E$, $\|f(x)\| = \|x\|$.
2. Pour tout $x, y \in E$, on a $\varphi(f(x), f(y)) = \varphi(x, y)$.
3. L'application f est un isomorphisme et $f^{-1} = f^*$.

Preuve. On fait une preuve circulaire. On commence par $1) \Rightarrow 2)$. C'est une conséquence immédiate des formules de polarisation, voir Proposition 1.1.9. Prouvons $2) \Rightarrow 3)$. Puisque pour tout $x \in E$, $\|f(x)\| = \|x\|$, si $x \in \text{Ker}(f)$ alors $\|f(x)\| = \|x\| = 0$ et donc $x = 0$. Ainsi $\text{Ker}(f) = \{0\}$, f est injective. Par le théorème du rang, f est donc surjective : c'est bien un isomorphisme, dont l'application réciproque f^{-1} est automatiquement linéaire. Par définition de l'adjoint, on a pour tout $x, y \in E$,

$$\varphi(f(x), f(y)) = \varphi(x, f^* \circ f(y)) = \varphi(x, y),$$

ainsi $f^* \circ f = \text{Id}$, et par unicité de l'inverse $f^{-1} = f^*$. Prouvons pour finir $3) \Rightarrow 1)$. Pour tout $x \in E$, écrivons

$$\|f(x)\|^2 = \varphi(f(x), f(x)) = \varphi(x, f^* \circ f(x)),$$

(par définition de l'adjoint) et comme $f^* = f^{-1}$, on tombe sur

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

et la preuve est bouclée. \square

L'ensemble des isométries (noté $U(E)$) d'un espace Euclidien ou Hermitien est naturellement muni d'une structure de *groupe* pour la loi \circ de composition des applications linéaires. Plus précisément on a les fait suivants.

- L'élément neutre de $U(E)$ pour la composition \circ est l'application identité Id .
- Tout $f \in U(E)$ admet un inverse à gauche et à droite $f^{-1} = f^*$ pour la composition \circ .
- La loi \circ est associative.

En général, ce groupe n'est pas commutatif !

Proposition 2.1.3 *Soit (E, φ) un espace euclidien ou hermitien et f un endomorphisme de E . Alors f est une isométrie ssi l'image par f d'une base orthonormée est orthonormée.*

Preuve. On a vu précédemment que f est une isométrie ssi pour tout $x, y \in E$, $\varphi(f(x), f(y)) = \varphi(x, y)$. Ainsi si (e_1, \dots, e_n) est une base orthonormée de E , $(f(e_1), \dots, f(e_n))$ est sans difficulté une famille orthonormée, donc automatiquement libre, et donc une base car elle a le bon cardinal. Réciproquement, supposons qu'il existe une base orthonormée (e_1, \dots, e_n) telle que $(f(e_1), \dots, f(e_n))$ soit une base orthonormée, alors on a pour tout $x = \sum_{i=1}^n x_i e_i$, par linéarité de f ,

$$\|f(x)\|^2 = \left\| \sum_{i=1}^n x_i f(e_i) \right\|^2,$$

on applique alors Pythagore (deux fois) qui nous donne

$$\|f(x)\|^2 = \left\| \sum_{i=1}^n x_i f(e_i) \right\|^2 = \sum_{i=1}^n |x_i|^2 = \|x\|^2,$$

et la preuve est finie. \square

Que dire de la **matrice** d'une isométrie dans une base orthonormée ? Soit donc f une isométrie et $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormée de E . Posons

$$M = \text{Mat}_{\mathcal{B}}(f).$$

Comme $f^* = f^{-1}$, on a donc $M^{-1} = {}^t M$, dans le cas réel, et $M^{-1} = {}^t \overline{M}$ (cas complexe). Si ${}^t M = M^{-1}$, on a

$$\det(M) \det({}^t M) = \det(M.M^{-1}) = \det(I_n) = 1,$$

mais comme $\det({}^t M) = \det(M)$, il vient $\det(M)^2 = 1$ d'où $\det(M) = \pm 1$. Par le même raisonnement, dans le cas complexe on a aussi $|\det(M)| = 1$. Ces propriétés conduisent naturellement au paragraphe suivant.

2.2 Groupes unitaires et orthogonaux

Définition 2.2.1 *On a les définition suivantes.*

- On note $O_n(\mathbb{R}) = \{M \in \mathcal{M}_n(\mathbb{R}) : M^t M = I_n\}$, c'est le groupe orthogonal.
- On note $U_n(\mathbb{C}) = \{M \in \mathcal{M}_n(\mathbb{C}) : M^t \overline{M} = I_n\}$, c'est le groupe unitaire.
- On note $SO_n(\mathbb{R}) = \{M \in O_n(\mathbb{R}) : \det(M) = 1\}$, c'est le groupe spécial orthogonal.
- On note $SU_n(\mathbb{C}) = \{M \in U_n(\mathbb{C}) : \det(M) = 1\}$, c'est le groupe spécial unitaire.

Remarques. On vérifie sans peine que si $M^t M = I_n$, alors M est forcément inversible car $\det(M) = \pm 1$ et $M^{-1} = {}^t M$. Ainsi la matrice de toute isométrie d'un espace euclidien de dimension n est orthogonale. De même la matrice de toute isométrie d'un espace hermitien de dimension n est unitaire. Les ensembles $O_n(\mathbb{R})$, $SO_n(\mathbb{R})$, $U_n(\mathbb{C})$ et $SU_n(\mathbb{C})$ sont des *groupes* pour la multiplication matricielle. Réciproquement, toute matrice orthogonale (resp. unitaire) peut être vue comme la matrice d'une isométrie de \mathbb{R}^n muni du produit scalaire canonique (resp. \mathbb{C}^n muni du produit scalaire hermitien canonique).

Si on munit \mathbb{R}^n de son produit scalaire canonique, une matrice est orthogonale ssi les colonnes forment une base orthonormée par la Proposition 2.1.3. Par exemple les matrices $(\theta \in \mathbb{R})$,

$$R(\theta) := \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

sont dans $SO_2(\mathbb{R})$. Ce sont les matrices de rotations d'angle θ dans le plan \mathbb{R}^2 .

Il grand temps d'aborder la notion délicate des angles.

2.3 Angles et orientation

Dans cette section, (E, φ) est un espace euclidien (donc $\mathbb{K} = \mathbb{R}$). On a vu par Cauchy-Schwarz que si $x, y \in E$, avec $x, y \neq 0$, alors

$$\frac{\varphi(x, y)}{\|x\| \|y\|} \in [-1, 1].$$

Définition 2.3.1 (*Angle non-orienté de deux vecteurs*) Si $x, y \neq 0$, l'angle non-orienté des deux vecteurs x, y est l'unique $\theta \in [0, \pi]$ tel que

$$\varphi(x, y) = \|x\| \|y\| \cos(\theta).$$

Autrement dit on a

$$\theta(x, y) = \arccos\left(\frac{\varphi(x, y)}{\|x\| \|y\|}\right),$$

avec $\theta = 0$ ou π ssi x, y sont colinéaires. On peut remarquer qu'avec cette définition on a $\theta(x, y) = \theta(y, x)$, l'angle est dit *non-orienté*.

Définition 2.3.2 Deux bases $\mathcal{B}, \mathcal{B}'$ d'un \mathbb{R} -espace vectoriel sont dites de même orientation ssi $\det(P_{\mathcal{B}, \mathcal{B}'}) > 0$, où $P_{\mathcal{B}, \mathcal{B}'}$ est la matrice de passage de la base \mathcal{B} à \mathcal{B}' .

L'espace euclidien \mathbb{R}^n muni de sa base canonique est muni d'une notion d'orientation naturelle : une famille libre de n vecteurs v_1, \dots, v_n sera dite directe (ou positivement orientée) si le déterminant de ces n vecteurs dans la base canonique est > 0 , ce qui revient à dire que si \mathcal{B} est la base canonique et $\mathcal{B}' = (v_1, \dots, v_n)$, ces deux bases ont même orientation. On voit alors que $SO_n(\mathbb{R})$ est le groupe des isométries de \mathbb{R}^n qui *préservent l'orientation*.

En dimension 2, si on se donne un vecteur non nul $e \in \mathbb{R}^2$, il existe un unique vecteur e^\wedge de même norme tel que (e, e^\wedge) soit une base orthogonale directe. En effet, si $e = (x, y)$ avec $x^2 + y^2 \neq 0$, comme e^\perp est une droite on doit avoir

$$e^\wedge = \lambda(-y, x),$$

avec $\lambda = \pm 1$ car on veut $\|e^\wedge\|^2 = x^2 + y^2$. La condition $\det(e, e^\wedge) > 0$ force $\lambda = 1$. Ainsi

$$e^\wedge = (-y, x).$$

On appellera dans la suite e^\wedge le **vecteur directement orthogonal** à e .

La notion d'angle orienté de deux vecteurs n'a de sens qu'en dimension $n = 2$.

Définition 2.3.3 Soit (E, φ) un espace euclidien orienté de dimension 2. Soient $x, y \in E$ avec $x, y \neq 0$. L'angle orienté $\widehat{\theta}(x, y)$ est l'unique $\theta \in \mathbb{R}/2\pi\mathbb{Z}$ tel que l'on a

$$\varphi(x, y) = \|x\| \|y\| \cos(\widehat{\theta}),$$

et

$$\det(x, y) = \|x\| \|y\| \sin(\widehat{\theta}).$$

Ici $\det(x, y)$ est calculé dans une base orthonormée orientée. L'angle $\widehat{\theta}(x, y)$ est défini modulo 2π .

On peut vérifier que maintenant,

$$\widehat{\theta}(x, y) = -\widehat{\theta}(y, x) [2\pi].$$

Cette définition nécessite une explication. Soit \mathcal{B} une base orthonormée directe de E et posons $\mathcal{B}' = (x, y)$. Alors un calcul montre que

$$B := \begin{pmatrix} \varphi(x, x) & \varphi(x, y) \\ \varphi(y, x) & \varphi(y, y) \end{pmatrix} = {}^t P_{\mathcal{B}, \mathcal{B}'} P_{\mathcal{B}, \mathcal{B}'}.$$

On a donc

$$\det(B) = \det(P_{\mathcal{B}, \mathcal{B}'})^2 = \|x\|^2 \|y\|^2 - \varphi(x, y)^2,$$

ce qui nous donne

$$\left(\frac{\det(x, y)}{\|x\| \|y\|} \right)^2 + \left(\frac{\varphi(x, y)}{\|x\| \|y\|} \right)^2 = 1.$$

On sait que si $a, b \in \mathbb{R}$ sont tels que $a^2 + b^2 = 1$ alors il existe une unique θ modulo 2π tel que $a = \cos(\theta)$ et $b = \sin(\theta)$.

On retiendra que $\det(x, y)$ est l'aire (algébrique) du parallélogramme donné par x, y et que

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

On a par ailleurs obtenu par ce calcul une autre preuve de *Cauchy-Schwarz*. En effet,

$$0 \leq \det(P_{\mathcal{B}, \mathcal{B}'})^2 = \|x\|^2 \|y\|^2 - \varphi(x, y)^2$$

ainsi

$$|\varphi(x, y)| \leq \|x\|^2 \|y\|^2,$$

avec égalité ssi $\det(P_{\mathcal{B}, \mathcal{B}'}) = 0$, i.e. (x, y) est liée.

Que se passe-t-il dans \mathbb{R}^3 ? Même si on a choisi une orientation de \mathbb{R}^3 , un plan vectoriel V dans \mathbb{R}^3 n'hérite pas canoniquement d'une orientation naturelle. Il faut se donner un vecteur unitaire e_3 normal à V . Une base (e_1, e_2) de V sera dite orientée positivement ssi la base (e_1, e_2, e_3) est positivement orientée.

2.4 Etude de $O_2(\mathbb{R})$ et $O_3(\mathbb{R})$: les isométries du plan et de l'espace

On va commencer par le théorème suivant qui classifie complètement les éléments de $O_2(\mathbb{R})$.

Théorème 2.4.1 *Soit $M \in O_2(\mathbb{R})$, vu comme isométrie de \mathbb{R}^2 . On a les possibilités suivantes.*

1. *Soit $\det(M) = -1$ et il existe une base orthonormée directe de \mathbb{R}^2 dans laquelle la matrice de M devient*

$$S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

i.e. $M = USU^{-1}$, où $U \in SO_2(\mathbb{R})$.

2. *Soit $\det(M) = 1$ et il existe $\theta \in \mathbb{R}$ tel que*

$$M = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Preuve. On commence par supposer que $\det(M) = 1$. Ecrivons

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Comme les colonnes de M sont orthogonales, on a $ab + cd = 0$ et donc les vecteurs $(a, -c)$ et (b, d) sont liés. Mais comme $1 = b^2 + d^2$, $(b, d) \neq (0, 0)$ et donc il existe $\lambda \in \mathbb{R}$ tel que $(a, -c) = \lambda(d, b)$. On a aussi $\det(M) = ad - bc = 1$ donc $\lambda(d^2 + b^2) = \lambda = 1$. Ainsi on a obtenu $a = d$ et $c = -b$ avec $a^2 + b^2 = 1$. Il existe donc $\theta \in \mathbb{R}$ tel que $\cos(\theta) = a$ et $\sin(\theta) = b$, et donc le point 2) est prouvé. Supposons que $\det(M) = -1$, et écrivons de nouveau

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Pour les mêmes raisons que précédemment, on a $(a, -c) = \lambda(d, b)$ et $ad - bc = -1$. Ceci entraîne que $\lambda = -1$ et donc on a

$$M = \begin{pmatrix} a & c \\ c & -a \end{pmatrix},$$

avec $a^2 + c^2 = 1$. Remarquons que si $a = 1$ alors $c = 0$ et donc

$$M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

et il n'y a plus rien à prouver. Supposons donc $a \neq 1$, alors le vecteur colonne

$$X = \begin{pmatrix} \frac{c}{1-a} \\ 1 \end{pmatrix},$$

vérifie $MX = X$. Le vecteur directement normal (de même norme) est

$$X^\wedge = \begin{pmatrix} -1 \\ \frac{c}{1-a} \end{pmatrix},$$

et un calcul direct montre que $MX^\wedge = -X^\wedge$. En posant $e_1 = \frac{X}{\|X\|}$ et $e_2 = \frac{X^\wedge}{\|X^\wedge\|}$, on obtient bien une base orthonormée (directe) telle que la matrice de M dans la base (e_1, e_2) est exactement

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

le théorème est prouvé. \square

On voit donc que les isométries d'un espace euclidien de dimension 2 sont de deux types : une *réflexion orthogonale*, c'est le cas 1) où l'orientation est renversée, ou une *rotation*, c'est le cas 2). En particulier les éléments de $SO_2(\mathbb{R})$ sont donc tous des rotations. Dans ce cas, si $X = \begin{pmatrix} x \\ y \end{pmatrix}$ avec $x^2 + y^2 = 1$, on a

$$\langle MX, X \rangle = \cos(\theta) \text{ et } \det(X, MX) = \sin(\theta),$$

donc l'angle orienté $\widehat{\theta}(X, MX)$ est exactement θ modulo 2π . Notons que $-I_2$ préserve l'orientation et correspond à une rotation d'angle orienté $+\pi$.

$SO_2(\mathbb{R})$ est un groupe *commutatif* : on peut vérifier que posant

$$R(\theta) := \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

on a pour tout $\theta_1, \theta_2 \in \mathbb{R}$, $R(\theta_1)R(\theta_2) = R(\theta_2)R(\theta_1) = R(\theta_1 + \theta_2)$, voir exercice en TD.

Abordons le cas de la dimension 3. On a le théorème suivant.

Théorème 2.4.2 *Soit $M \in O_3(\mathbb{R})$, vu comme isométrie de \mathbb{R}^3 . On a les possibilités suivantes.*

1. Soit $\det(M) = -1$ et il existe une base orthonormée directe de \mathbb{R}^3 dans laquelle la matrice de M devient de la forme ($\theta \in \mathbb{R}$)

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

2. Soit $\det(M) = 1$ et il existe une base orthonormée directe de \mathbb{R}^3 et $\theta \in \mathbb{R}$ tel que M se transforme sous la forme

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

Preuve. Considérons le polynôme caractéristique $\chi_M(\lambda) = \det(M - \lambda I_3)$. C'est un polynôme de degré 3, donc par le théorème des valeurs intermédiaires, il admet un zéro réel noté $\lambda_0 \in \mathbb{R}$. On rappelle, voir cours de L1, qu'il existe donc un vecteur propre $X_0 \in \mathbb{R}^3$, non nul, tel que $MX_0 = \lambda_0 X_0$. On peut bien sur supposer que $\|X_0\| = 1$. Comme M est une isométrie, on a $\|MX_0\| = \|X_0\| = |\lambda_0| \|X_0\|$, et ainsi $\lambda_0 = \pm 1$. Considérons maintenant l'orthogonal de X_0 , c'est à dire

$$X_0^\perp = \{Y \in \mathbb{R}^3 : \langle Y, X_0 \rangle = 0\}.$$

Pour tout $Y \in X_0^\perp$, on a

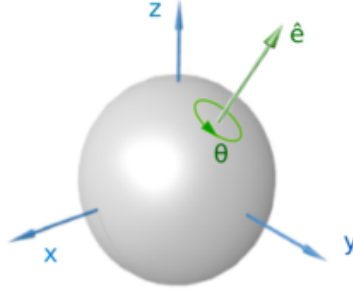
$$\langle MY, X_0 \rangle = \langle Y, M^* X_0 \rangle = \langle Y, M^{-1} X_0 \rangle = \pm \langle Y, X_0 \rangle = 0.$$

Ainsi $F = X_0^\perp$ est stable par M . C'est un espace euclidien de dimension 2, on est ramené au cas précédent. On voit ainsi que si $\det(M) = 1$, soit $\lambda_0 = 1$ et alors $\det(M|_F) = 1$ et on est donc dans le cas 2) : on choisit une base orthonormée (e_1, e_2) de F telle que (e_1, e_2, X_0) soit directe, et la matrice de M dans cette base est de la forme 2). Si $\lambda_0 = -1$, on doit avoir $\det(M|_F) = -1$, et en choisissant une base orthonormée (positivement orientée par X_0) de F la matrice de M devient

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

ce qui après une permutation circulaire des vecteurs de la base donne en fait le cas 2) avec $\theta = \pi$. Le cas où $\det(M) = -1$ est analogue. \square

Dans le cas 2), M est appelée *rotation d'angle orienté θ et d'axe dirigé par X_0* . Notons que l'orientation du plan X_0^\perp découle du choix de X_0 . Comme $SO_3(\mathbb{R})$ est un groupe, la composée de deux rotations est toujours une rotation. Il est en revanche non-commutatif (le vérifier en exercice).



Rotation d'angle θ , d'axe $\hat{e} = \text{Vect}(X_0)$.

Le Théorème 2.4.2 est une version moderne d'un théorème d'Euler (1776) que l'on cite ici en latin.

Theorema : Quomodocunque sphaera circa centrum suum conuertatur, semper assignari potest diameter, cuius directio in situ translato conueniat cum situ initiali.

La traduction française approximative serait : si on fait tourner une sphère autour de son centre, on peut toujours trouver un diamètre qui reste invariant lors de cette transformation. C'est la droite engendrée par le vecteur X_0 de la preuve précédente...

Un exemple de matrice de rotation. Calculer la matrice (dans la base canonique) de la rotation \mathcal{R} d'angle $+\frac{\pi}{4}$ et d'axe $\hat{e} = \text{Vect}((1, 1, 1))$. Remarquons que

$$\hat{e}^\perp = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}.$$

Une base évidente de \hat{e}^\perp est $v_1 = (1, -1, 0)$, $v_2 = (1, 0, -1)$. En appliquant le procédé d'orthonormalisation de Schmidt à (v_1, v_2) on trouve une base orthonormée donnée par $e_1 = \frac{1}{\sqrt{2}}(1, -1, 0)$ et $e_2 = \sqrt{\frac{2}{3}}(1/2, 1/2, -1)$, que l'on complète par $e_3 = \frac{1}{\sqrt{3}}(1, 1, 1)$. On vérifie que $\det(e_1, e_2, e_3) > 0$ et donc la matrice de \mathcal{R} dans cette base est (avec $\theta = \pi/4$),

$$M = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Il reste maintenant à calculer la matrice de \mathcal{R} dans la base canonique : la matrice de passage P de la base canonique à $\mathcal{B} = (e_1, e_2, e_3)$ est

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & -\sqrt{\frac{2}{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Comme P est orthogonale, on a $P^{-1} = {}^tP$, et il ne reste plus qu'à calculer

$$\text{Mat}_{b.c.}(\mathcal{R}) = PM^tP = \begin{pmatrix} \frac{1+\sqrt{2}}{3} & \frac{1-\sqrt{2}-\sqrt{6}}{6} & \frac{1-\sqrt{2}+\sqrt{6}}{6} \\ \frac{1-\sqrt{2}+\sqrt{6}}{6} & \frac{1+\sqrt{2}}{3} & \frac{1-\sqrt{2}-\sqrt{6}}{6} \\ \frac{1-\sqrt{2}-\sqrt{6}}{6} & \frac{1-\sqrt{2}+\sqrt{6}}{6} & \frac{1+\sqrt{2}}{3} \end{pmatrix}.$$

2.5 Le produit vectoriel dans \mathbb{R}^3

On se place ici dans l'espace euclidien \mathbb{R}^3 muni du produit scalaire canonique noté $\langle \cdot, \cdot \rangle$.

Théorème 2.5.1 1. Pour tout vecteurs $U, V, W \in \mathbb{R}^3$, la quantité $\det(U, V, W)$ est invariante par changement de base orthonormée directe.
2. Pour tout vecteurs $U, V \in \mathbb{R}^3$, il existe un unique vecteur dans \mathbb{R}^3 , noté $U \wedge V$, tel que pour tout $W \in \mathbb{R}^3$ on a

$$\det(U, V, W) = \langle U \wedge V, W \rangle.$$

Preuve. 1) Si U est le vecteur colonne des coordonnées dans la base canonique, alors dans toute autre base orthonormée directe on a $U = PU'$ où $P \in SO_3(\mathbb{R})$ est la matrice de passage à la nouvelle base et U' les coordonnées de U dans cette nouvelle base. Mais donc comme le déterminant d'un produit matriciel est le produit des déterminants, on a la formule

$$\det(U, V, W) = \det(PU', PV', PW') = \det(P) \det(U', V', W') = \det(U', V', W'),$$

car $\det(P) = 1$.

2) Se donnant $U, V \in \mathbb{R}^3$, l'application

$$W \mapsto \det(U, V, W)$$

est une forme linéaire sur \mathbb{R}^3 , donc par le théorème de Riesz, il existe un unique vecteur $U \wedge V$ tel que pour tout W ,

$$\det(U, V, W) = \langle U \wedge V, W \rangle.$$

La preuve est finie. \square

Le vecteur $U \wedge V$ est appelé produit vectoriel des vecteurs U, V . L'opération "produit vectoriel" vérifie les propriétés suivantes.

Proposition 2.5.2 — Pour tout $U, V \in \mathbb{R}^3$, on a $V \wedge U = -U \wedge V$.
— Pour tout $U, V, Z \in \mathbb{R}^3$ et $\lambda \in \mathbb{R}$, on a

$$(U + \lambda V) \wedge Z = U \wedge Z + \lambda(V \wedge Z).$$

— Pour tout $U, V \in \mathbb{R}^3$, on a $U \wedge V = 0$ ssi U, V sont colinéaires.
— Pour tout $U, V \in \mathbb{R}^3$, avec U, V linéairement indépendants, on a

$$(U \wedge V)^\perp = \text{Vect}(U, V).$$

Preuve. C'est un exercice facile basé sur l'unicité de $U \wedge V$ et les propriétés du déterminant. \square

Proposition 2.5.3 Pour tout $U, V \in \mathbb{R}^3$, on a

$$\|U \wedge V\| = \|U\| \cdot \|V\| \cdot |\sin(\widehat{\theta}(U, V))|,$$

où $\widehat{\theta}(U, V)$ est l'angle (orienté) des vecteurs U, V dans le plan $F = \text{Vect}(U, V)$, l'orientation de F étant donnée par le vecteur normal $U \wedge V$.

Preuve. On peut bien sur supposer que U, V sont non colinéaires. On a par définition

$$\|U \wedge V\|^2 = \det(U, V, U \wedge V).$$

On calcule alors ce déterminant dans une base orthonormée directe (e_1, e_2, e_3) où (e_1, e_2) est une base orthonormée de F et e_3 est colinéaire à $U \wedge V$. Dans cette base on a $U = x_1 e_1 + x_2 e_2$, $V = y_1 e_1 + y_2 e_2$ et $U \wedge V = \pm \|U \wedge V\| e_3$ de sorte que

$$\det(U, V, U \wedge V) = \begin{vmatrix} x_1 & y_1 & 0 \\ x_2 & y_2 & 0 \\ 0 & 0 & \pm \|U \wedge V\| \end{vmatrix} = \pm \det(U, V) \|U \wedge V\|.$$

En simplifiant par $\|U \wedge V\|$ et en prenant la valeur absolue on a bien

$$\|U \wedge V\| = |\det(U, V)| = \|U\| \cdot \|V\| \cdot |\sin(\widehat{\theta}(U, V))|.$$

La preuve est finie. \square

En pratique, on peut calculer le produit vectoriel en coordonnées par la recette suivante.

Proposition 2.5.4 Soit (e_1, e_2, e_3) une base orthonormée directe de \mathbb{R}^3 . Alors si $U = \sum_{i=1}^3 u_i e_i$ et $V = \sum_{i=1}^3 v_i e_i$, on a

$$U \wedge V = (u_2 v_3 - u_3 v_2) e_1 + (u_3 v_1 - u_1 v_3) e_2 + (u_1 v_2 - u_2 v_1) e_3.$$

Preuve. On commence par remarquer par les propositions précédentes que $e_1 \wedge e_2 = e_3$, $e_2 \wedge e_3 = e_1$ et $e_3 \wedge e_1 = e_2$. En utilisant l'antisymétrie et la bilinéarité du produit vectoriel, on a

$$U \wedge V = \left(\sum_i u_i e_i \right) \wedge \left(\sum_j v_j e_j \right) = \sum_{i < j} (u_i v_j - u_j v_i) e_i \wedge e_j,$$

et on obtient la formule en calculant composante par composante. \square

Attention, le produit vectoriel n'est pas associatif, mais il existe une formule dite du double produit vectoriel que l'on cite ici.

Proposition 2.5.5 Soit $U, V, W \in \mathbb{R}^3$, on les identités suivantes.

- (Double Produit Vectoriel) $U \wedge (V \wedge W) = \langle U, W \rangle V - \langle U, V \rangle W$.
- (Jacobi) $U \wedge (V \wedge W) + W \wedge (U \wedge V) + V \wedge (W \wedge U) = 0$

Preuve. Voir exercice en TD.

Il n'est pas possible de définir un produit vectoriel en dimension quelconque en conservant ces propriétés remarquables. On peut en revanche, en dimension n , définir le produit $x_1 \wedge \dots \wedge x_{n-1}$ par un déterminant analogue. La bonne généralisation de l'opération \wedge est celle du produit extérieur, voir cours de M1.

2.6 Produit vectoriel et rotations

On a vu précédemment que si U, V sont des vecteurs non-colinéaires, alors

$$\det(U, V, U \wedge V) = \|U \wedge V\|^2 > 0,$$

donc $(U, V, U \wedge V)$ est une base orientée positivement, avec $(U \wedge V)^\perp = \text{Vect}(U, V)$. On va utiliser ces propriétés pour établir une formule intrinsèque exprimant l'action d'une rotation sur un vecteur quelconque de \mathbb{R}^3 .

Proposition 2.6.1 (Formule de Rodrigues) Soit N un vecteur unitaire de \mathbb{R}^3 et $\theta \in \mathbb{R}$. Notons $\mathcal{R} \in SO_3(\mathbb{R})$ la rotation d'axe N et d'angle égal à θ modulo 2π . Pour tout $U \in \mathbb{R}^3$, on a

$$\mathcal{R}(U) = \cos(\theta)U + (1 - \cos(\theta))\langle U, N \rangle N + \sin(\theta)N \wedge U.$$

Preuve. Posons $V = U - \langle U, N \rangle N$. On a

$$\mathcal{R}(U) = \mathcal{R}(\langle U, N \rangle N + V) = \langle U, N \rangle N + \mathcal{R}(V).$$

Comme $V \in N^\perp$, l'action de \mathcal{R} sur V est celle d'une rotation de $SO_2(\mathbb{R})$ d'angle θ . On remarque de plus que $N \wedge V$ est un vecteur de N^\perp directement orthogonal à V (pour l'orientation de N^\perp induite par N). Comme N est unitaire et que N et V sont orthogonaux, on a de plus

$$\|N \wedge V\| = \|V\|.$$

Dans la base orthonormée $\mathcal{B} = (e_1, e_2)$ avec $e_1 = \frac{V}{\|V\|}$, $e_2 = \frac{N \wedge V}{\|V\|}$, la matrice de \mathcal{R} est donc

$$\text{Mat}_{\mathcal{B}}(\mathcal{R}) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

et ainsi

$$\mathcal{R}(V) = \cos(\theta)V + \sin(\theta)N \wedge V,$$

ce qui donne finalement

$$\begin{aligned} \mathcal{R}(U) &= \langle U, N \rangle N + \cos(\theta)V + \sin(\theta)N \wedge V \\ &= \cos(\theta)U + (1 - \cos(\theta))\langle U, N \rangle N + \sin(\theta)N \wedge U. \end{aligned}$$

La preuve est complète. \square

La formule de Rodrigues présente plusieurs intérêts : elle est intrinsèque, peut se composer... Elle donne une formule générale pour la matrice d'une rotation (exprimée dans la base canonique) d'angle θ et d'axe $N = (a, b, c)$ avec $a^2 + b^2 + c^2 = 1$.

$$\begin{aligned} \text{Mat}_{b.c.}(\mathcal{R}) &= \\ &\begin{pmatrix} \cos(\theta) + (1 - \cos(\theta))a^2 & ab(1 - \cos(\theta)) - c \sin(\theta) & ac(1 - \cos(\theta)) + b \sin(\theta) \\ ab(1 - \cos(\theta)) + c \sin(\theta) & \cos(\theta) + (1 - \cos(\theta))b^2 & bc(1 - \cos(\theta)) - a \sin(\theta) \\ ac(1 - \cos(\theta)) - b \sin(\theta) & bc(1 - \cos(\theta)) + a \sin(\theta) & \cos(\theta) + (1 - \cos(\theta))c^2 \end{pmatrix}. \end{aligned}$$

D'un point de vue numérique, le calcul des produits de matrices de rotation est couteux, on lui préfère pour les applications (animation, images générées par ordinateur) l'utilisation des *quaternions* dont on ne parlera pas trop ici, faute de temps...

2.7 Compléments : le groupe $SU_2(\mathbb{C})$

Le groupe $SU_2(\mathbb{C})$ est le groupe des matrices unitaires de déterminant 1 en dimension 2. On rappelle que \mathbb{C}^2 est muni du produit scalaire hermitien

$$\langle Z, W \rangle = z_1 \overline{w_1} + z_2 \overline{w_2},$$

avec $Z = (z_1, z_2)$ et $W = (w_1, w_2)$. Comme l'image de la base canonique par $U \in SU_2(\mathbb{C})$ est une base orthonormée de \mathbb{C}^2 , les colonnes de U doivent être orthogonales et de norme 1. On voit donc que toute matrice $U \in SU_2(\mathbb{C})$ est de la forme

$$U = \begin{pmatrix} \alpha & \overline{\beta} \\ -\beta & \overline{\alpha} \end{pmatrix},$$

avec $\alpha, \beta \in \mathbb{C}$ tels que $|\alpha|^2 + |\beta|^2 = 1$. En écrivant $\alpha = x_1 + iy_1$ et $\beta = x_2 + iy_2$ avec $x_1, x_2, y_1, y_2 \in \mathbb{R}$, on voit que

$$|\alpha|^2 + |\beta|^2 = 1 \Leftrightarrow x_1^2 + x_2^2 + y_1^2 + y_2^2 = 1,$$

ce qui montre que $SU_2(\mathbb{C})$ est en bijection avec la sphère unité S^3 de \mathbb{R}^4 (pour la norme euclidienne standard). En particulier la sphère S^3 peut être munie d'une structure de groupe (celle de $SU_2(\mathbb{C})$). Il n'est pas difficile de montrer, en calculant le polynôme caractéristique, que toute matrice $U \in SU_2(\mathbb{C})$ est diagonalisable, avec des valeurs propres conjuguées de module 1. La diagonalisation peut se faire en outre en base orthonormée (voir le chapitre suivant sur les endomorphismes normaux). En résumé, toute matrice $U \in SU_2(\mathbb{C})$ peut s'écrire sous la forme

$$U = V D_\theta V^*,$$

où $V \in SU_2(\mathbb{C})$ et

$$D_\theta = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix},$$

avec $\theta \in \mathbb{R}$.

On va voir que malgré les apparences, $SU_2(\mathbb{C})$ est en fait très voisin de $SO_3(\mathbb{R})$. Considérons l'ensemble \mathscr{V} suivant.

$$\mathscr{V} = \{M \in \mathcal{M}_2(\mathbb{C}) : \text{Tr}(M) = 0 \text{ et } M^* = M\}.$$

Ici $\text{Tr}(M)$ désigne la trace d'une matrice, c'est à dire la somme de ses éléments diagonaux. On va voir facilement que \mathscr{V} est un \mathbb{R} -espace vectoriel de dimension 3. En effet, si $M_1, M_2 \in \mathscr{V}$ et $\lambda \in \mathbb{R}$, alors il est clair que

$$\text{Tr}(M_1 + \lambda M_2) = \text{Tr}(M_1) + \lambda \text{Tr}(M_2) = 0,$$

ainsi que

$$(M_1 + \lambda M_2)^* = M_1^* + \lambda M_2^* = M_1 + \lambda M_2.$$

Par définition de \mathcal{V} , on voit que $M \in \mathcal{V}$ ssi il existe $a, x, y \in \mathbb{R}$ tel que

$$M = \begin{pmatrix} a & x + iy \\ x - iy & a \end{pmatrix}.$$

Donc on a

$$M = yT_1 + xT_2 + aT_3,$$

où T_1, T_2, T_3 est une base de \mathcal{V} donnée par

$$T_1 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad T_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad T_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

On va définir un produit scalaire sur \mathcal{V} qui en fait un espace euclidien de dimension 3. Posons pour tout $A, B \in \mathcal{V}$,

$$\varphi(A, B) := \frac{1}{2} \text{Tr}(AB).$$

Il s'agit bien d'une forme bilinéaire symétrique, car la trace est linéaire et on a toujours $\text{Tr}(AB) = \text{Tr}(BA)$. Pour voir qu'elle est définie on utilise le fait que si $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{V}$ alors

$$2\varphi(M, M) = \text{Tr}(MM^*) = |a|^2 + |b|^2 + |c|^2 + |d|^2.$$

On laisse le lecteur vérifier que (T_1, T_2, T_3) est en fait *une base orthonormée* de \mathcal{V} pour le produit scalaire φ . On va construire une application

$$\text{Ad} : SU_2(\mathbb{C}) \rightarrow \text{GL}(V)$$

en se basant sur l'espace (\mathcal{V}, φ) . Posons pour tout $U \in SU_2(\mathbb{C})$, pour tout $M \in \mathcal{V}$,

$$\text{Ad}_U(M) := UMU^*.$$

Cette application est bien définie : si $M \in \mathcal{V}$, on a

$$\text{Tr}(UMU^*) = \text{Tr}(U^*UM) = \text{Tr}(M) = 0,$$

ainsi que

$$(UMU^*)^* = UM^*U^* = UMU^*.$$

De plus, pour $U \in SU_2(\mathbb{C})$ fixée, l'application $M \mapsto \text{Ad}_U(M)$ est linéaire, et préserve le produit scalaire car

$$\varphi(\text{Ad}_U(M), \text{Ad}_U(M)) = \frac{1}{2} \text{Tr}(UMU^*UMU^*) = \frac{1}{2} \text{Tr}(M^2) = \varphi(M, M).$$

Ainsi la matrice de Ad_U dans la base (T_1, T_2, T_3) est dans $O_3(\mathbb{R})$. Examinons l'action de Ad_U sur cette base dans le cas simple où $U = D_\theta$. Un calcul direct montre que

$$\text{Ad}_U(T_3) = T_3, \quad \text{Ad}_U(T_1) = \begin{pmatrix} 0 & -ie^{2i\theta} \\ ie^{-2i\theta} & 0 \end{pmatrix} = \cos(2\theta)T_1 + \sin(2\theta)T_2,$$

$$\text{Ad}_U(T_2) = \begin{pmatrix} 0 & e^{2i\theta} \\ e^{-2i\theta} & 0 \end{pmatrix} = \cos(2\theta)T_2 - \sin(2\theta)T_1.$$

Ainsi la matrice de Ad_{D_θ} dans la base $\mathcal{B} = (T_1, T_2, T_3)$ est

$$\text{Mat}_{\mathcal{B}}(\text{Ad}_{D_\theta}) = \begin{pmatrix} \cos(2\theta) & -\sin(2\theta) & 0 \\ \sin(2\theta) & \cos(2\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Donc Ad_U est une rotation d'angle 2θ , une isométrie directe (pour l'orientation donnée par \mathcal{B}) ! Le cas général où $U = VD_\theta V^*$ n'est pas plus dur : il suffit de remplacer la base \mathcal{B} par la base conjuguée par V . En résumé, on a construit une application

$$\mathcal{J} : \begin{cases} SU_2(\mathbb{C}) \rightarrow SO_3(\mathbb{R}) \\ U \mapsto \text{Mat}_{\mathcal{B}}(\text{Ad}_U) \end{cases}.$$

Avec un peu plus de travail, on peut montrer que \mathcal{J} est surjective. C'est de plus un morphisme de groupes (le vérifier). Elle n'est pas injective, mais son noyau est $\text{Ker}(\mathcal{J}) = \{\pm I_2\}$. En effet, si $\text{Ad}_U = Id$, c'est que $\theta = 0$ modulo π , et donc $U = \pm I_2$.

En termes savants (voir M1), on dira que $SU_2(\mathbb{C})$ est un revêtement double de $SO_3(\mathbb{R})$...L'application Ad ne sort pas du chapeau. C'est la représentation adjointe de $SU_2(\mathbb{C})$, et \mathcal{V} n'est autre que l'algèbre de Lie réelle de $SU_2(\mathbb{C})$, voir cours de M1 sur les groupes et algèbres de Lie.

Chapitre 3

Réduction des endomorphismes auto-adjoints

3.1 Sur les déterminants

On rappelle dans cette section les propriétés de bases sur les déterminants (qu'on a déjà librement utilisés dans le chapitre précédent). L'ensemble des bijections $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ forme un groupe pour la composition o appelé groupe symétrique d'ordre n , et noté S_n . Le cardinal de S_n est donc $|S_n| = n!$. Une transposition $\tau = (i \ j) \in S_n$ est une bijection qui échange deux éléments de S_n et laisse les autres invariants : $\tau(i) = j$, $\tau(j) = i$ et pour tout $k \neq i, j$, $\tau(k) = k$.

Théorème 3.1.1 *Toute permutation $\sigma \in S_n$ différente de l'identité peut s'écrire comme un produit de au plus $n - 1$ transpositions : les transpositions engendrent S_n .*

Preuve. Par récurrence sur n . Si $n = 2$, on a $S_2 = \{Id, (1 \ 2)\}$, le théorème est bien vrai. Supposons avoir démontré le résultat au rang n . Considérons $\sigma \in S_{n+1}$, différente de l'identité. Il existe donc $i_0 \in \{1, \dots, n+1\}$ tel que $\sigma(i_0) = j_0 \neq i_0$. Soit $\tau = (i_0 \ j_0)$ la transposition qui échange i_0 et j_0 et considérons $\sigma' = \tau \circ \sigma$. On a $\sigma'(i_0) = i_0$, donc σ' induit une bijection $\sigma' : \{1, \dots, n+1\} \setminus i_0 \rightarrow \{1, \dots, n+1\} \setminus i_0$, on peut donc appliquer l'hypothèse de récurrence à σ' qui dit que $\sigma' = \tau_1 \circ \dots \circ \tau_k$, où τ_1, \dots, τ_k sont des transpositions et $k \leq n - 1$. Ainsi on a

$$\sigma = \tau \circ \sigma' = \tau \circ \tau_1 \circ \dots \circ \tau_k,$$

est le théorème est prouvé par récurrence. \square

Attention, cette décomposition n'est pas unique, par exemple dans S_3 , la permutation circulaire σ définie par $\sigma(1) = 2$, $\sigma(2) = 3$, $\sigma(3) = 1$ s'écrit $\sigma = (1 \ 2)(2 \ 3)$ mais aussi $\sigma = (3 \ 2)(1 \ 3)$.

Théorème 3.1.2 *Il existe un unique morphisme de groupes $\varepsilon : S_n \rightarrow \{-1, +1\}$ tel que $\varepsilon(\tau) = -1$ pour toute transposition τ . On l'appelle signature.*

Preuve. Unicité. Si $\sigma \in S_n$, alors par le théorème précédent, il existe une décomposition de σ en produit de transpositions $\sigma = \tau_1 \circ \dots \circ \tau_k$ et comme ε est un morphisme, on a

$$\varepsilon(\sigma) = \varepsilon(\tau_1) \dots \varepsilon(\tau_k) = (-1)^k,$$

donc $\varepsilon(\sigma)$ est complètement déterminée par k .

Existence. Posons pour tout $\sigma \in S_n$,

$$\varepsilon(\sigma) := \prod_{1 \leq i < j \leq n} \left(\frac{\sigma(j) - \sigma(i)}{j - i} \right).$$

Si on note \mathcal{P} l'ensemble des parties $\{i, j\}$ à deux éléments de $\{1, \dots, n\}$, l'application $T_\sigma(\{i, j\}) = \{\sigma(i), \sigma(j)\}$ est une bijection de $\mathcal{P} \rightarrow \mathcal{P}$. Ceci montre par changement d'indice que

$$|\varepsilon(\sigma)| = \prod_{\{i, j\} \in \mathcal{P}} \frac{|\sigma(j) - \sigma(i)|}{|j - i|} = \frac{\prod_{\{i, j\} \in \mathcal{P}} |\sigma(j) - \sigma(i)|}{\prod_{\{i, j\} \in \mathcal{P}} |j - i|} = 1,$$

donc $\varepsilon(\sigma) \in \{-1, 1\}$. Il faut ensuite vérifier que c'est un morphisme, ce qui découle d'un changement d'indice analogue. On montre ensuite en distinguant les cas que si σ est une transposition alors $\varepsilon(\sigma) = -1$. \square

Définition 3.1.3 *Soit E un \mathbb{K} -espace vectoriel. Une application $\psi : E^p \rightarrow \mathbb{K}$ est dite p -linéaire alternée ssi elle est linéaire par rapport à chaque variable et si ψ est nulle sur toute p -famille de vecteurs dont au moins deux sont égaux.*

Théorème 3.1.4 *Soit E un \mathbb{K} -espace vectoriel et $\psi : E^p \rightarrow \mathbb{K}$ une forme p -linéaire alternée. Alors pour tout $\sigma \in S_p$, pour tout $x_1, \dots, x_p \in E$, on a*

$$\psi(x_{\sigma(1)}, \dots, x_{\sigma(p)}) = \varepsilon(\sigma) \psi(x_1, \dots, x_p).$$

Preuve. On commence par faire l'observation suivante. Si $i, j \in \{1, \dots, p\}$ sont deux indices distincts, on a par multilinéarité

$$\begin{aligned} \psi(\dots, x_i + x_j, \dots, x_i + x_j, \dots) &= 0 = \psi(\dots, x_i, \dots, x_j, \dots) + \psi(\dots, x_j, \dots, x_i, \dots) \\ &\quad + \underbrace{\psi(\dots, x_i, \dots, x_i, \dots)}_{=0} + \underbrace{\psi(\dots, x_j, \dots, x_j, \dots)}_{=0}, \end{aligned}$$

ce qui montre que $\psi(\dots, x_i, \dots, x_j, \dots) = -\psi(\dots, x_j, \dots, x_i, \dots)$. En d'autres termes, si τ est une transposition,

$$\psi(x_{\tau(1)}, \dots, x_{\tau(p)}) = -\psi(x_1, \dots, x_p).$$

Le cas général s'obtient en décomposant σ en produit de transpositions. \square

Théorème 3.1.5 *On suppose ici $E = \mathbb{K}^n$. L'espace des formes n -linéaires alternées $f : E^n \rightarrow \mathbb{K}$ est de dimension 1. Si on impose $f(e_1, \dots, e_n) = 1$ où (e_1, \dots, e_n) est la base canonique de $E = \mathbb{K}^n$, on appelle $f(x_1, \dots, x_n)$ le **déterminant** des n -vecteurs x_1, \dots, x_n que l'on notera $\det(x_1, \dots, x_n)$.*

Preuve. On écrit pour chaque vecteur x_j , $x_j = \sum_{i=1}^n x_{ij} e_i$. On a donc

$$f(x_1, \dots, x_n) = \sum_{(i_1, \dots, i_n) \in \{1, \dots, n\}^n} x_{i_1 1} \dots x_{i_n n} f(e_{i_1}, \dots, e_{i_n}).$$

Comme la forme f est alternée, tout les termes qui ne correspondent pas à une indexation bijective $k \mapsto i_k$ s'annulent. On peut donc réécrire

$$\begin{aligned} f(x_1, \dots, x_n) &= \sum_{\sigma \in S_n} x_{\sigma(1)1} \dots x_{\sigma(n)n} f(e_{\sigma(1)}, \dots, e_{\sigma(n)}) \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma) x_{\sigma(1)1} \dots x_{\sigma(n)n} f(e_1, \dots, e_n) = f(e_1, \dots, e_n) \sum_{\sigma \in S_n} \varepsilon(\sigma) x_{\sigma(1)1} \dots x_{\sigma(n)n}. \end{aligned}$$

On voit donc que cet espace est de dimension 1 et cette n -forme f est unique si on impose $f(e_1, \dots, e_n) = 1$. Réciproquement, on vérifie (exercice) que si on pose

$$f(x_1, \dots, x_n) = \sum_{\sigma \in S_n} \varepsilon(\sigma) x_{\sigma(1)1} \dots x_{\sigma(n)n},$$

elle a bien les propriétés demandées. \square

Si $M = (m_{ij}) \in \mathcal{M}_n(\mathbb{K})$, on définit son déterminant par

$$\det(M) = \det(M_1, \dots, M_n) = \sum_{\sigma \in S_n} \varepsilon(\sigma) m_{\sigma(1)1} \dots m_{\sigma(n)n},$$

où M_1, \dots, M_n sont les n colonnes de la matrice M , vus comme vecteurs de \mathbb{K}^n . On le note aussi matriciellement

$$\det(M) = \begin{vmatrix} m_{11} & \dots & m_{1n} \\ \vdots & & \vdots \\ m_{n1} & \dots & m_{nn} \end{vmatrix}.$$

Proposition 3.1.6 *On a les propriétés suivantes.*

1. On a $\det(I_n) = 1$.
2. Pour tout $A \in \mathcal{M}_n(\mathbb{K})$, $\lambda \in \mathbb{K}$, on a $\det(\lambda A) = \lambda^n \det(A)$.
3. Pour tout $A, B \in \mathcal{M}_n(\mathbb{K})$, on a $\det(AB) = \det(A)\det(B)$.
4. Pour tout $A \in \mathcal{M}_n(\mathbb{K})$, on a $\det({}^t A) = \det(A)$.
5. Une matrice A est inversible ssi $\det(A) \neq 0$. Dans ce cas on a $\det(A^{-1}) = (\det(A))^{-1}$.

Preuve. Les points 1) et 2) sont faciles. Pour le point 3) on observe que l'application

$$(X_1, \dots, X_n) \mapsto \det(AX_1, \dots, AX_n)$$

est n -linéaire alternée. Comme c'est un espace de dimension 1, il existe $\lambda(A) \in \mathbb{K}$ tel que pour tout $(X_1, \dots, X_n) \in \mathbb{K}^n$, on a

$$\det(AX_1, \dots, AX_n) = \lambda(A) \det(X_1, \dots, X_n).$$

En prenant $X_1 = e_1, \dots, X_n = e_n$ on voit que $\lambda(A) = \det(A)$. Le résultat s'obtient alors en particulierisant avec $X_1 = B_1, \dots, X_n = B_n$ où B_1, \dots, B_n sont les vecteurs colonnes de B . Le point 4) est plus technique et se montre à partir de la définition en faisant le bon changement d'indice dans la somme portant sur toutes les permutations. Pour le point 5), si A n'est pas inversible alors son rang est $< n$, donc ses colonnes sont liées. Un des vecteurs colonnes s'exprime donc comme combinaison linéaire des autres ce qui par le fait que \det est multilinéaire alternée entraîne que $\det(A) = 0$. Par la contraposée, $\det(A) \neq 0$ implique donc que A est inversible. Clairement si A est inversible, $AA^{-1} = I_n$ et donc par 3) on a $\det(A)\det(A^{-1}) = \det(I_n) = 1$. \square

Exemples. En dimension $n = 2$, on a directement par la définition

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

En dimension $n = 3$, on a plus péniblement en listant les 6 permutations de S_3

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + dhc + gbf - (gec + dbi + ahf),$$

ce qu'on peut retenir par le moyen mnémotechnique des diagonales montantes et des diagonales descendantes (règle de Sarrus).

En dimension quelconque, il existe une technique de *développement* par rapport à une ligne ou une colonne qu'on va expliquer ici. Un peu de terminologie s'impose. Si $M = (m_{ij})$ est une matrice carrée d'ordre n , on appelle *cofacteur* d'indice i, j , noté Δ_{ij} , le déterminant de la matrice extraite de M en *rayant* la colonne d'indice j et la ligne d'indice i .

Proposition 3.1.7 *Pour tout $j = 1, \dots, n$, on a*

$$\det(M) = \sum_{i=1}^n (-1)^{i+j} \Delta_{ij}.$$

Comme $\det({}^t M) = \det(M)$, on a aussi pour tout $i = 1, \dots, n$,

$$\det(M) = \sum_{j=1}^n (-1)^{i+j} \Delta_{ij}.$$

Preuve. On commence par examiner le cas où M est de la forme :

$$M = \begin{pmatrix} m_{11} & m_{12} & \dots & m_{1n} \\ 0 & m_{22} & \dots & m_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & m_{n2} & \dots & m_{nn} \end{pmatrix}.$$

Par définition on a

$$\det(M) = \sum_{\sigma \in S_n} \varepsilon(\sigma) m_{\sigma(1)1} \dots m_{\sigma(n)n},$$

mais $m_{\sigma(1)1} = 0$ pour toute permutation σ telle que $\sigma(1) \neq 1$. La somme se réécrit donc

$$\begin{aligned} \det(M) &= m_{11} \sum_{\sigma \in S_n : \sigma(1)=1} m_{\sigma(2)2} \dots m_{\sigma(n)n} \\ &= m_{11} \det(\tilde{M}), \end{aligned}$$

où on a posé

$$\tilde{M} = \begin{pmatrix} m_{22} & \dots & m_{2n} \\ \vdots & & \vdots \\ m_{n2} & \dots & m_{nn} \end{pmatrix}.$$

Fixons maintenant $j \in \{1, \dots, n\}$. En écrivant la colonne M_j de M sous la forme $M_j = \sum_{i=1}^n m_{ij} e_i$, on a par linéarité par rapport à la j -ème colonne,

$$\det(M) = \det(M_1, \dots, M_j, \dots, M_n) = \sum_{i=1}^n m_{ij} \det(M_1, \dots, e_i, \dots, M_n).$$

En permutant i fois les lignes et j fois les colonnes de $[M_1, \dots, e_i, \dots, M_n]$ on se ramène au cas précédent et on a

$$\det(M_1, \dots, e_i, \dots, M_n) = (-1)^{i+j} \Delta_{ij}.$$

La preuve est finie. \square

Corollaire 3.1.8 *Si une matrice est triangulaire supérieure ou inférieure, son déterminant est le produit des coefficients diagonaux.*

La preuve est immédiate par récurrence en développant par rapport à la première colonne (cas triangulaire supérieure), ou la première ligne (cas triangulaire inférieure).

Définition 3.1.9 *Soit f un endomorphisme d'un \mathbb{K} -espace vectoriel de dimension finie. Soit \mathcal{B} une base de E . Le déterminant de f , noté $\det(f)$, est par définition*

$$\det(f) := \det(\text{Mat}_{\mathcal{B}}(f)).$$

Cette définition ne dépend pas du choix de la base \mathcal{B} . En effet, si \mathcal{B}' est une autre base, on a par changement de base

$$\text{Mat}_{\mathcal{B}'}(f) = P_{\mathcal{B}, \mathcal{B}'}^{-1} \text{Mat}_{\mathcal{B}}(f) P_{\mathcal{B}, \mathcal{B}'},$$

où $P_{\mathcal{B}, \mathcal{B}'}$ est la matrice de passage de \mathcal{B} à \mathcal{B}' . Ainsi on a

$$\begin{aligned} \det(\text{Mat}_{\mathcal{B}'}(f)) &= \det(P_{\mathcal{B}, \mathcal{B}'}^{-1} \text{Mat}_{\mathcal{B}}(f) P_{\mathcal{B}, \mathcal{B}'}) \\ &= \det(P_{\mathcal{B}, \mathcal{B}'})^{-1} \det(\text{Mat}_{\mathcal{B}}(f)) \det(P_{\mathcal{B}, \mathcal{B}'}) = \det(\text{Mat}_{\mathcal{B}}(f)). \end{aligned}$$

De même que pour les matrices, un endomorphisme f est un isomorphisme ssi $\det(f) \neq 0$.

3.2 Rappels sur la diagonalisation des endomorphismes

Définition 3.2.1 *Soit E un \mathbb{K} -espace vectoriel de dimension finie. Soient $F_1, \dots, F_k \subset E$ des sous-espaces vectoriels de E . On dit que la somme $F_1 + \dots + F_k$ est directe, et on la note $F_1 \oplus \dots \oplus F_k$, ssi pour tout vecteurs $(x_1, \dots, x_k) \in F_1 \times \dots \times F_k$,*

$$x_1 + x_2 + \dots + x_k = 0 \Rightarrow x_1 = x_2 = \dots = x_k = 0.$$

Proposition 3.2.2 *La somme $F_1 + \dots + F_k$ est directe ssi pour tout $x \in F_1 + \dots + F_k$, il existe un unique k -uplet $(x_1, \dots, x_k) \in F_1 \times \dots \times F_k$ tel que $x = x_1 + \dots + x_k$. De plus, on a*

$$\dim(F_1 \oplus \dots \oplus F_k) = \dim(F_1) + \dots + \dim(F_k).$$

Preuve. Si la somme est directe, supposons que l'on ait deux telles décompositions

$$x = x_1 + \dots + x_k = x'_1 + \dots + x'_k,$$

on a donc

$$x_1 - x'_1 + \dots + x_k - x'_k = 0,$$

avec pour tout $i = 1, \dots, k$, $x_i - x'_i \in F_i$. Par définition de la somme directe on a donc $x_i = x'_i$ pour tout i . Réciproquement si $x_1 + \dots + x_k = 0$ avec $x_i \in F_i$ pour tout i , par unicité de l'écriture on doit avoir $x_i = 0$ pour tout i .

Pour chaque sous espace F_i , avec $\dim(F_i) = m_i$, considérons $\mathcal{B}_i = (e_1^i, \dots, e_{m_i}^i)$ une base de F_i . On va montrer que

$$\mathcal{B} = (e_1^1, \dots, e_{m_1}^1, e_1^2, \dots, e_{m_2}^2, \dots, e_1^k, \dots, e_{m_k}^k)$$

est une base de $F_1 \oplus \dots \oplus F_k$. C'est une famille libre car si on a des scalaires $\lambda_{ij} \in \mathbb{K}$ tels que

$$\sum_{i=1}^k \sum_{j=1}^{m_i} \lambda_{ij} e_j^i = 0,$$

alors par définition d'une somme directe on a pour tout i

$$\sum_{j=1}^{m_i} \lambda_{ij} e_j^i = 0$$

et ainsi $\lambda_{ij} = 0$ pour tout i, j car les vecteurs e_j^i sont libres. C'est évidemment une famille génératrice car tout élément $x \in F_1 \oplus \dots \oplus F_k$ s'écrit sous la forme $x = x_1 + \dots + x_k$ avec $x_i \in F_i$. Comme \mathcal{B}_i est une base de F_i , chacun des x_i s'écrit comme combinaison linéaire des vecteurs e_j^i et c'est fini. On peut conclure, la dimension de $F_1 \oplus \dots \oplus F_k$ est donc $m_1 + \dots + m_k = \dim(F_1) + \dots + \dim(F_k)$. \square

Définition 3.2.3 *Soit f un endomorphisme d'un \mathbb{K} -espace vectoriel E . On dit que $\lambda \in \mathbb{K}$ est une valeur propre de f s'il existe un vecteur non nul $x \in E$ tel que $f(x) = \lambda x$. Un tel vecteur x est appelé vecteur propre associé à λ . Le sous-espace vectoriel*

$$E_\lambda(f) := \text{Ker}(f - \lambda \text{Id}),$$

est appelé sous-espace propre de f associé à λ .

Remarques. $E_\lambda(f)$ est un sous-espace vectoriel puisque c'est un noyau d'application linéaire. Si λ est une valeur propre de f , alors par définition, $E_\lambda(f) \neq \{0\}$. Encore par définition, le sous-espace $E_\lambda(f)$ est stable par f et la restriction de f à ce sous-espace est égale à λId . L'existence d'une valeur propre pour un endomorphisme donné n'est pas une question évidente et dépend beaucoup du corps de base \mathbb{K} . L'outil pour calculer les valeurs propres est le polynôme caractéristique.

Proposition 3.2.4 *Soit f un endomorphisme d'un \mathbb{K} -espace vectoriel E de dimension finie n . La fonction de λ définie par*

$$\chi_f(\lambda) := \det(f - \lambda Id),$$

est un polynôme de degré n en λ . De plus $\lambda_0 \in \mathbb{K}$ est une valeur propre de f ssi on a $\chi_f(\lambda_0) = 0$. Le polynôme $\chi_f(\lambda)$ est appelé polynôme caractéristique de f . Ses zéros dans \mathbb{K} sont exactement les valeurs propres de f .

Preuve. La définition du déterminant, qui n'utilise que des sommes et des produits, montre facilement que $\chi_f(\lambda)$ est un polynôme. Le terme de plus haut degré provient de $\sigma = Id$ et vaut $(-1)^n \lambda^n$. Par définition, on a λ valeur propre de $f \Leftrightarrow \text{Ker}(f - \lambda Id) \neq \{0\} \Leftrightarrow f - \lambda Id$ non inversible $\Leftrightarrow \det(f - \lambda Id) = \chi_f(\lambda) = 0$. C'est fini. \square

Définition 3.2.5 *Un endomorphisme f est dit diagonalisable ssi il existe une base \mathcal{B} de E constituée de vecteurs propres. Dans cette base \mathcal{B} , on a alors*

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix},$$

où $\lambda_1, \dots, \lambda_n$ sont les valeurs propres (pas forcément distinctes).

Remarque importante. Il existe des endomorphismes non diagonalisables, et ceci même en dimension 2. Considérons l'endomorphisme $f : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ dont la matrice dans la base canonique est

$$M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Alors $\chi_f(\lambda) = \lambda^2$, donc la seule valeur propre est 0. Si f était diagonalisable, sa matrice serait donc nulle dans une base de \mathbb{C}^2 , ce qui entraînerait que f est nulle, absurde.

On a le fait suivant qui est l'observation la plus importante de la théorie.

Théorème 3.2.6 Notons $\text{Sp}(f) \subset \mathbb{K}$, l'ensemble des valeurs propres distinctes de f , appelé "spectre" de f . Si cet ensemble est non vide, alors la somme

$$\bigoplus_{\lambda \in \text{Sp}(f)} E_{\lambda}(f)$$

est toujours directe.

Preuve. Notons $\lambda_1, \dots, \lambda_k$ les valeurs propres distinctes de f . On va montrer par récurrence sur k que si x_1, \dots, x_k avec $x_j \in E_{\lambda_j}(f)$ sont tels que $x_1 + \dots + x_k = 0$ alors $x_1 = \dots = x_k = 0$. Si $k = 1$, il n'y a rien à faire. Supposons avoir montré cette propriété au rang k , et écrivons

$$x_1 + \dots + x_k + x_{k+1} = 0, \quad (3.1)$$

avec pour tout $j = 1, \dots, k+1$, $x_j \in E_{\lambda_j}(f)$. En appliquant f , on a aussi

$$\lambda_1 x_1 + \dots + \lambda_k x_k + \lambda_{k+1} x_{k+1} = f(x_1 + \dots + x_k + x_{k+1}) = f(0) = 0.$$

En multipliant (3.1) par λ_{k+1} et en la soustrayant à la formule ci-dessus, on obtient

$$(\lambda_1 - \lambda_{k+1})x_1 + \dots + (\lambda_k - \lambda_{k+1})x_k = 0.$$

Par hypothèse de récurrence on a donc pour tout $j = 1, \dots, k$, $(\lambda_j - \lambda_{k+1})x_j = 0$, mais comme $\lambda_j - \lambda_{k+1} \neq 0$ (les valeurs propres sont distinctes), on obtient donc que pour tout $j = 1, \dots, k$, $x_j = 0$. En revenant à (3.1), on déduit que $x_{k+1} = 0$ et la preuve est finie. \square

Un corollaire facile est le suivant.

Théorème 3.2.7 Soit f un endomorphisme d'un \mathbb{K} -espace vectoriel E , de dimension n , admettant n valeurs propres distinctes dans \mathbb{K} , alors f est diagonalisable.

Preuve. Si f admet n valeurs propres distinctes, comme les sous-espaces sont en somme directe, on a

$$n \geq \dim \left(\bigoplus_{\lambda \in \text{Sp}(f)} E_{\lambda}(f) \right) = \sum_{\lambda \in \text{Sp}(f)} \dim(E_{\lambda}(f)) \geq n,$$

donc

$$\bigoplus_{\lambda \in \text{Sp}(f)} E_{\lambda}(f) = E.$$

En prenant un vecteur propre x_λ (non nul) dans chaque espace $E_\lambda(f)$, on obtient donc une base $\{x_\lambda : \lambda \in \text{Sp}(f)\}$ de E , et c'est fini. \square

Exemples. On considère l'endomorphisme $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dont la matrice dans la base canonique est

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

On a $\chi_f(\lambda) = \lambda^2 - \lambda - 1$, les valeurs propres sont $\lambda = \frac{1 \pm \sqrt{5}}{2}$, elles sont distinctes, donc f est diagonalisable. Considérons maintenant l'endomorphisme $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dont la matrice dans la base canonique est

$$B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

On constate alors que $\chi_g(\lambda) = \lambda^2 + 1$, qui n'a pas de racine sur \mathbb{R} , donc g n'est pas diagonalisable sur \mathbb{R} . En revanche, si on considère g comme agissant sur \mathbb{C}^2 , alors les valeurs propres sont i et $-i$, qui sont distinctes, et g est diagonalisable sur \mathbb{C} .

Le cas complexe. Dans le cas où $\mathbb{K} = \mathbb{C}$, on sait (par le théorème fondamental de l'algèbre) que tout polynôme non nul se décompose comme produit de facteurs irréductibles i.e. pour tout $P \in \mathbb{C}[X]$ de degré n , on peut écrire

$$P(z) = a_n \prod_{j=1}^k (z - z_k)^{m_j},$$

où $a_n \neq 0$ est le coefficient dominant, $m_1 + \dots + m_k = n$ et z_1, \dots, z_k sont les racines distinctes. Si f est un endomorphisme d'un \mathbb{C} -espace vectoriel de dimension finie n , on écrira donc

$$\chi_f(z) = (-1)^n \prod_{\lambda \in \text{Sp}(f)} (z - \lambda)^{m_\lambda}.$$

L'entier m_λ est appelé *multiplicité algébrique* de λ . On posera $d_\lambda = \dim(E_\lambda(f))$, on l'appelle *multiplicité géométrique* de λ .

Théorème 3.2.8 Soit E un \mathbb{K} -espace vectoriel et f un endomorphisme de E . Il est diagonalisable ssi on a pour tout $\lambda \in \text{Sp}(f)$, $d_\lambda = m_\lambda$.

Preuve. Si f est diagonalisable, alors dans une base \mathcal{B} bien choisie sa matrice est

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix},$$

où les $\lambda = \lambda_i \in \text{Sp}(f)$ sont répétées d_λ fois. On calcule alors $\chi_f(z)$ dans cette base et on trouve

$$\chi_f(z) = (-1)^n \prod_{\lambda \in \text{Sp}(f)} (z - \lambda)^{d_\lambda},$$

donc par unicité de cette écriture irréductible, $d_\lambda = m_\lambda$ pour tout $\lambda \in \text{Sp}(f)$. Réciproquement, si on a pour tout $\lambda \in \text{Sp}(f)$, $d_\lambda = m_\lambda$, alors comme

$$\sum_{\lambda \in \text{Sp}(f)} d_\lambda = \sum_{\lambda \in \text{Sp}(f)} m_\lambda = n,$$

on a

$$\dim \left(\bigoplus_{\lambda \in \text{Sp}(f)} E_\lambda(f) \right) = \sum_{\lambda \in \text{Sp}(f)} d_\lambda = n,$$

et donc on obtient

$$\bigoplus_{\lambda \in \text{Sp}(f)} E_\lambda(f) = E,$$

f est donc diagonalisable. \square

Exemple. On considère l'endomorphisme $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dont la matrice dans la base canonique est

$$A = \begin{pmatrix} 1 & 4 & -2 \\ 0 & 6 & -3 \\ -1 & 4 & 0 \end{pmatrix}.$$

En observant que la somme des coefficients sur chaque ligne fait 3, on peut calculer le polynôme caractéristique sous forme factorisée et on trouve

$$\chi_f(\lambda) = -(\lambda - 3)(\lambda - 2)^2.$$

Les valeurs propres sont donc 3, valeur propre simple, et 2, de multiplicité algébrique 2. L'endomorphisme f est donc diagonalisable (sur \mathbb{C}) ssi $\dim(E_2(f)) = 2$. En résolvant le système linéaire associé à $E_2(f)$, on trouve que

$$E_2(f) = \text{Vect}\{(4, 3, 4)\},$$

et donc $d_2 = 1 \neq m_2 = 2$, f n'est pas diagonalisable sur \mathbb{C} .

3.3 Diagonalisation des endomorphismes auto-adjoints

Dans un espace euclidien ou hermitien (E, φ) , la structure additionnelle donnée par le produit scalaire permet d'en dire plus sur les questions de diagonalisation. On rappelle qu'un endomorphisme f est auto-adjoint ssi $f = f^*$, en d'autres termes ssi on a pour tout $x, y \in E$,

$$\varphi(f(x), y) = \varphi(x, f(y)).$$

On a déjà rencontré des applications auto-adjointes concrètes : si F est un sous-espace vectoriel de E , alors la projection orthogonale P_F est auto-adjointe. En effet, si (e_1, \dots, e_p) est une base orthonormée de F , on a pour tout $x \in E$,

$$P_F(x) = \sum_{i=1}^p \varphi(x, e_i) e_i,$$

et donc pour tout $x, y \in E$,

$$\varphi(P_F(x), y) = \sum_i \varphi(x, e_i) \varphi(e_i, y) = \varphi\left(x, \sum_i \varphi(y, e_i) e_i\right) = \varphi(x, P_F(y)).$$

On va voir de plus que

$$E = \text{IM}(P_F) \oplus \text{Ker}(P_F),$$

comme P_F restreint à $\text{IM}(P_F)$ est l'identité, en fait $\text{Ker}(P_F - I) = \text{IM}(P_F)$. Ainsi P_F est diagonalisable, et ses valeurs propres sont 1 et 0. La somme directe ci-dessus étant orthogonale, P_F est donc diagonalisable en base orthonormée. On va voir qu'en fait ce phénomène est général.

Proposition 3.3.1 *Soit (E, φ) un espace hermitien ou euclidien et $f : E \rightarrow E$ une application linéaire auto-adjointe.*

1. *Les valeurs propres sur \mathbb{C} de f sont toutes réelles.*
2. *Les sous espaces-propres associés à des valeurs propres distinctes sont orthogonaux.*

Preuve. 1). On fait d'abord la preuve dans le cas hermitien, qui est le plus facile. Soit $\lambda \in \mathbb{C}$ une valeur propre de f et $v \neq 0$ un vecteur propre associé. On a donc $f(v) = \lambda v$ et on peut écrire

$$\varphi(f(v), v) = \lambda \varphi(v, v),$$

mais comme f est auto-adjoint, on a aussi

$$\varphi(v, f(v)) = \lambda \varphi(v, v).$$

Comme φ est semi-linéaire à droite, on a donc

$$\overline{\lambda}\varphi(v, v) = \lambda\varphi(v, v),$$

et comme $\varphi(v, v) = \|v\|^2 \neq 0$, on obtient $\lambda = \overline{\lambda}$ donc $\lambda \in \mathbb{R}$. Dans le cas où (E, φ) est euclidien, on se ramène au cas précédent en complexifiant. On se donne une base orthonormée (e_1, \dots, e_n) de E et on considère A la matrice de f dans cette base. On la fait agir sur \mathbb{C}^n muni du produit scalaire hermitien canonique, A est autoadjointe et son spectre sur \mathbb{C} coïncide avec celui de f , il est donc réel.

2). Soient $\lambda \neq \mu \in \text{Sp}(f)$ (forcément réelles) et $u \in \text{Ker}(f - \lambda Id)$, $v \in \text{Ker}(f - \mu Id)$ des vecteurs propres. On a

$$\lambda\varphi(u, v) = \varphi(f(u), v) = \varphi(u, f(v)) = \mu\varphi(u, v),$$

d'où $(\lambda - \mu)\varphi(u, v) = 0$, et comme $\mu \neq \lambda$ on a donc $\varphi(u, v) = 0$. \square

Théorème 3.3.2 *Soit (E, φ) un espace hermitien ou euclidien et $f : E \rightarrow E$ une application linéaire auto-adjointe, alors il existe une base orthonormée de E dans laquelle la matrice de f est diagonale, à coefficient réels.*

Preuve. On fait la preuve par récurrence sur la dimension de l'espace, notée n . On va avoir besoin du lemme suivant.

Lemme 3.3.3 *Soit $f : E \rightarrow E$ une application linéaire auto-adjointe et $F \subset E$ un sous-espace vectoriel stable par f i.e. $f(F) \subset F$. Alors F^\perp est aussi stable par f .*

Preuve du lemme. On se fixe $y \in F^\perp$, c'est à dire que pour tout $x \in F$, $\varphi(x, y) = 0$. Comme $f(F) \subset F$ et utilisant le fait que $f = f^*$, on a donc pour tout $x \in F$,

$$\varphi(f(x), y) = 0 = \varphi(x, f(y)),$$

donc $f(y) \in F^\perp$. \square

Retour à la preuve du théorème. Si $n=1$, il n'y a rien à faire, toute application linéaire est diagonale. Supposons avoir démontré le thm pour $\dim(E) \leq n$, et plaçons nous dans le cas $\dim(E) = n + 1$. Par le thm fondamental de l'algèbre, $\chi_f(\lambda)$ a au moins une racine complexe λ_0 , qui est donc réelle par la Proposition 3.3.1. Le sous espace propre $E_{\lambda_0}(f) = \text{Ker}(f - \lambda_0 Id)$ est stable par f , de dimension ≥ 1 , et donc $(E_{\lambda_0}(f))^\perp$ est de dimension $\leq n$, stable par f . On applique alors l'hypothèse de récurrence à la restriction de f à $(E_{\lambda_0}(f))^\perp$ pour trouver une base orthonormée \mathcal{B} de $(E_{\lambda_0}(f))^\perp$ qui diagonalise f . On choisit ensuite une base orthonormée de $E_{\lambda_0}(f)$ que l'on concatène à \mathcal{B} pour obtenir cqfd. \square

Corollaire 3.3.4 Soit $S \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique réelle, alors il existe une matrice diagonale à coefficients réels D et $P \in O_n(\mathbb{R})$ tel que

$$S = PD^tP.$$

Preuve. On considère \mathbb{R}^n muni du produit scalaire standard. On considère l'application $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ dont la matrice dans la base canonique est S , comme f est auto-adjointe, le théorème précédent s'applique et nous fournit l'existence d'une base orthonormée qui diagonalise f . La matrice de passage P de la base canonique à cette nouvelle base est orthogonale. \square

Suivant le même principe, on a aussi le fait suivant dans le cas hermitien.

Corollaire 3.3.5 Soit $S \in \mathcal{M}_n(\mathbb{C})$ une matrice hermitienne i.e. ${}^t\overline{S} = S$, alors il existe une matrice diagonale à coefficients réels D et $P \in U_n(\mathbb{C})$ tel que

$$S = PD^tP.$$

Preuve. Même chose dans le cas hermitien. \square

Un exemple. Soit S la matrice symétrique donnée par

$$S = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 2 \\ 0 & 2 & 1 \end{pmatrix}.$$

On calcule son polynôme caractéristique $\chi_S(\lambda)$ qui est

$$\chi_S(\lambda) = \begin{vmatrix} 1-\lambda & 2 & 0 \\ 2 & 3-\lambda & 2 \\ 0 & 2 & 1-\lambda \end{vmatrix},$$

et en développant par rapport à la première colonne on a

$$\chi_S(\lambda) = -(\lambda - 1)(\lambda - 5)(\lambda + 1).$$

Les valeurs propres sont donc $1, 5, -1$ et on sait que S est diagonalisable. En calculant les sous espaces propres associés, dont on sait qu'ils sont orthogonaux, on trouve une base orthonormée de \mathbb{R}^3 constituée de vecteurs propres :

$$X_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix}, X_5 = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix}, X_{-1} = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Ainsi $S = PD^tP$ avec

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

La réduction des matrices symétriques réelles a de nombreuses applications géométriques et analytiques : redressement des coniques et quadriques, calculs de courbures, calcul de directions de rotations propres en mécanique du solide...

3.4 Une application : la décomposition polaire dans $GL_n(\mathbb{R})$

On va utiliser la théorie précédente pour démontrer un célèbre résultat, appelé *décomposition polaire*. On va parler au préalable d'endomorphismes auto-adjoint *positifs* et *défini positifs*.

Définition 3.4.1 Soit (E, φ) un espace euclidien ou hermitien. Un endomorphisme auto-adjoint $f : E \rightarrow E$ est dit :

- Positif ssi pour tout $x \in E$, $\varphi(f(x), x) \geq 0$.
- Défini positif ssi pour tout $x \in E$, $x \neq 0 \Rightarrow \varphi(f(x), x) > 0$.

Remarque. Il revient au même de dire que f est défini positif ou que la forme bilinéaire $(x, y) \mapsto \varphi(f(x), y)$ est un produit scalaire. Une matrice symétrique réelle est dite positive (resp. définie positive) ssi l'application linéaire associée agissant sur \mathbb{R}^n munit du produit scalaire standard est positive (resp. définie positive).

Proposition 3.4.2 Soit (E, φ) un espace euclidien ou hermitien. Un endomorphisme auto-adjoint $f : E \rightarrow E$ est positif ssi toutes ses valeurs propres sont positives. Il est défini positif ssi toutes ses valeurs propres sont strictement positives.

Preuve. Supposons f positif. Soit $\lambda \in \text{Sp}(f)$ une valeur propre et v un vecteur propre associé (non nul). On a donc

$$0 \leq \varphi(f(v), v) = \lambda \|v\|^2,$$

ainsi $\lambda \geq 0$. On voit facilement de plus que si f est défini positif, alors $\varphi(f(v), v) > 0$ et donc $\lambda > 0$. Réciproquement, on prend (e_1, \dots, e_n) une base

orthonormée de E qui diagonalise f . Dans cette base, on a pour tout $x \in E$, avec $x = \sum_i x_i e_i$,

$$\varphi(f(x), x) = \sum_i \sum_j \lambda_i x_i \overline{x_j} \varphi(e_i, e_j) = \sum_{i=1}^n \lambda_i |x_i|^2,$$

où $\lambda_1, \dots, \lambda_n$ sont les valeurs propres de f . Il est maintenant clair que si $\lambda_i \geq 0$ pour tout i , alors $\varphi(f(x), x) \geq 0$, et que si $\lambda_i > 0$ pour tout i , alors si $x \neq 0$, $\varphi(f(x), x) > 0$. \square

Théorème 3.4.3 *Soit $M \in \mathcal{M}_n(\mathbb{R})$, supposée **inversible**. Alors il existe une unique matrice symétrique définie positive S et une matrice orthogonale O tel que*

$$M = SO.$$

Preuve. On va faire agir les matrices sur \mathbb{R}^n , équipé du produit scalaire usuel noté $\langle \cdot, \cdot \rangle$. On commence par montrer l'unicité si existence. Si $M = SO$ avec S symétrique et O orthogonale, alors on a

$$MM^* = SOO^*S^* = S^2.$$

On va voir que si on impose à S d'être définie positive, il n'y a qu'un seul choix possible pour résoudre $S^2 = MM^*$. Par suite, $O = S^{-1}M$ et donc O est aussi unique. Existence. Considérons la matrice $R = MM^*$. C'est une matrice symétrique définie positive : en effet, pour tout $x \in \mathbb{R}^n$, on a

$$\langle MM^*x, x \rangle = \langle M^*x, M^*x \rangle = \|M^*x\|^2 \geq 0,$$

et de plus

$$\langle MM^*x, x \rangle = 0 \Rightarrow M^*x = 0,$$

mais on a vu que $\text{Ker}(M^*) = \text{IM}(M)^\perp = \{0\}$, car M étant inversible son image est \mathbb{R}^n (théorème du rang). Ainsi $x = 0$. D'après la Proposition 3.4.2, les valeurs propres de MM^* sont donc strictement positives. On sait que MM^* est diagonalisable en base orthonormée, donc il existe P une matrice orthogonale telle que

$$MM^* = PDP^*,$$

avec D diagonale à coefficients strictement positifs, notée

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

Posons donc

$$S = P\sqrt{D}P^*,$$

avec

$$\sqrt{D} := \begin{pmatrix} \sqrt{\lambda_1} & 0 & \dots & 0 \\ 0 & \sqrt{\lambda_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sqrt{\lambda_n} \end{pmatrix}.$$

On a bien sur $S^2 = PDP^* = MM^*$. On observe au passage que les solutions de $T^2 = MM^*$ sont toutes du type

$$T = P \begin{pmatrix} \pm\sqrt{\lambda_1} & 0 & \dots & 0 \\ 0 & \pm\sqrt{\lambda_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \pm\sqrt{\lambda_n} \end{pmatrix} P^*,$$

et que la seule solution définie positive est S . Calculons maintenant pour tout $x \in \mathbb{R}^n$,

$$\|S^{-1}Mx\|^2 = \langle S^{-1}Mx, S^{-1}Mx \rangle = \langle Mx, S^{-2}Mx \rangle,$$

mais $S^{-2} = (MM^*)^{-1} = (M^*)^{-1}M^{-1}$, donc

$$\|S^{-1}Mx\|^2 = \langle Mx, (M^*)^{-1}x \rangle = \|x\|^2.$$

Ainsi $S^{-1}M$ conserve la norme euclidienne, et c'est donc une matrice orthogonale O , donc $M = SO$. \square

On peut remarquer en plus que $\det(O) = \det(S^{-1})\det(M)$ et comme on a

$$\det(S^{-1}) > 0,$$

il vient que $O \in SO_n(\mathbb{R})$ ssi $\det(M) > 0$. Le théorème de décomposition polaire généralise l'écriture polaire des nombres complexes. En effet, notons

$$\mathcal{E} = \text{Vect}_{\mathbb{R}}\{I_2, J\} \subset \mathcal{M}_2(\mathbb{R}),$$

où

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad J^2 = -I_2.$$

Les éléments de \mathcal{E} sont les matrices de la forme

$$M = a + bJ = \begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

qui sont donc inversibles quand $\det(M) = a^2 + b^2 \neq 0$. Il est facile de voir que \mathcal{E} est une algèbre commutative, de dimension 2 sur \mathbb{R} , isomorphe à \mathbb{C} . Un calcul direct montre que $MM^* = (a^2 + b^2)I_2$ donc la décomposition polaire dit que

$$M = \sqrt{a^2 + b^2}R(\theta),$$

où bien sur

$$R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = \cos(\theta)I_2 + \sin(\theta)J = \sum_{k=0}^{\infty} \frac{(\theta J)^k}{k!} = \exp(\theta J).$$

est la matrice de rotation d'angle θ . On a bien retrouvé la décomposition usuelle $z = |z|e^{i\theta}$ dans \mathbb{C} .

Chapitre 4

Courbes planes

4.1 Courbes paramétrées, généralités

Dans ce chapitre, \mathbb{R}^n est équipé du produit scalaire euclidien usuel. La norme associée est notée $\| \cdot \|$. Dans tout ce qui suit I désigne un intervalle de \mathbb{R} . Une application $\gamma : I \rightarrow \mathbb{R}^n$ est dite de classe C^k ssi on a pour tout $t \in I$, $\gamma(t) = (\gamma_1(t), \dots, \gamma_n(t))$ où pour tout $i = 1, \dots, n$, $\gamma_i : I \rightarrow \mathbb{R}$ est une fonction de classe C^k sur I . On rappelle qu'une fonction f de classe C^k sur un intervalle de \mathbb{R} est une fonction k fois dérivable dont la dérivée k -ème $f^{(k)}$ est continue sur I .

Définition 4.1.1 On appelle courbe paramétrée (ou arc paramétré) C^k dans \mathbb{R}^n un couple (I, γ) où I est un intervalle de \mathbb{R} et $\gamma : I \rightarrow \mathbb{R}^n$ est une application de classe C^k . La courbe est dite régulière ssi pour tout $t \in I$, $\gamma'(t) \neq 0$. Un point $t_0 \in I$ tel que $\gamma'(t_0) = 0$ est dit singulier.

Remarques. L'application γ n'est pas forcément injective : l'image $\gamma(I)$ peut présenter des points de croisement (appelés points multiples). Si $k = 0$ i.e. l'application γ est juste supposée continue, l'image $\gamma(I)$ peut être très compliquée : il existe des courbes C^0 définies sur $I = [0, 1]$ dont l'image est le carré $[0, 1] \times [0, 1]$, appelées courbes de Peano. Dans ce cours on supposera toujours que $k \geq 1$, voire $k \geq 2$.

Deux courbes (I, f) et (J, g) peuvent avoir même image si on passe de l'une à l'autre par un "changement de paramétrisation". C'est ce qui motive la définition suivante.

Définition 4.1.2 Deux courbes C^k , (I, f) et (J, g) sont dites C^k -équivalentes ssi il existe un C^k -difféomorphisme $\theta : I \rightarrow J$ tel que $f = g \circ \theta$.

Un C^k -difféomorphisme est une bijection $\theta : I \rightarrow J$ dont l'application réciproque $\theta^{-1} : J \rightarrow I$ est aussi C^k . Donnons un exemple simple dans le plan. Soit

$f(t) = (\cos(2t), \sin(2t))$ avec $I = [0, \pi]$ et $g(t) = (\cos(t), \sin(t))$ avec $J = [0, 2\pi]$. Alors les courbes (I, f) et (J, g) sont C^∞ -équivalentes avec $\theta(x) = 2x$. Les images $f(I)$ et $g(J)$ sont les mêmes : c'est le cercle unité. On a simplement changé le paramétrage. Si θ est croissant, on dit que (I, f) et (J, g) ont la même *orientation*.

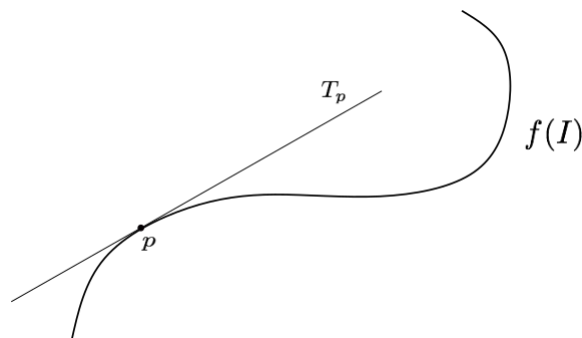
Si un arc est C^1 , on peut définir la droite tangente en un point *régulier* de la courbe.

Définition 4.1.3 Soit (I, f) un arc C^1 et $p = f(t_0)$ un point **régulier** de la courbe image. La droite tangente en p à (I, f) est par définition la droite affine $T_p := p + \mathbb{R}f'(t_0)$. Elle est indépendante du paramétrage.

Si (I, f) et (J, g) sont C^1 -équivalents avec $f = g \circ \theta$, posant $r_0 = \theta(t_0)$, on a par dérivation des fonctions composées $f'(t_0) = g'(\theta(t_0))\theta'(t_0) = g'(r_0)\theta'(t_0)$. Comme θ est un C^1 -difféo, on doit avoir $\theta'(t_0) \neq 0$ et donc

$$\mathbb{R}f'(t_0) = \mathbb{R}g'(r_0)\theta'(t_0) = \mathbb{R}g'(r_0),$$

les droites sont les mêmes.



Droite tangente T_p .

Définition 4.1.4 Soit $I = [a, b]$ et (I, f) un arc C^1 . L'abscisse curviligne, ou la longueur d'arc $s(t)$ est définie pour tout $t \in I$ par

$$s(t) = \int_a^t \|f'(x)\| dx.$$

C'est la longueur de la courbe paramétrée sur l'intervalle $[a, t]$. Une courbe est dite paramétrée par la longueur d'arc ssi pour tout $x \in I$, on a $\|f'(x)\| = 1$.

Exemple. Soit $I = [0, +\infty[$ et $f(t) = (e^{-t} \cos(t), e^{-t} \sin(t))$. L'arc (I, f) est appelé spirale logarithmique. Pour tout $t \geq 0$ on a

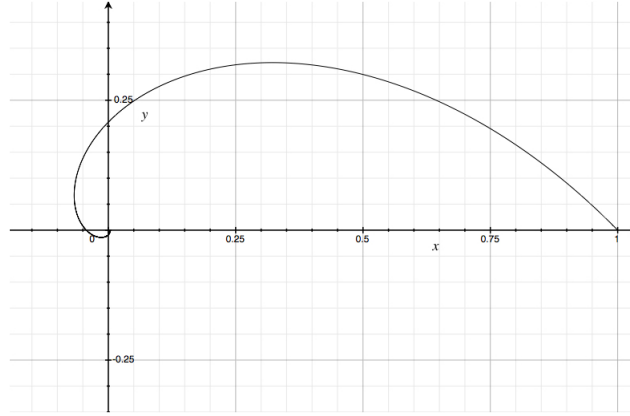
$$f'(t) = e^{-t}(-\sin(t) - \cos(t), \cos(t) - \sin(t)),$$

ainsi

$$\|f'(t)\| = e^{-t} \sqrt{(\sin(t) + \cos(t))^2 + (\cos(t) - \sin(t))^2} = \sqrt{2}e^{-t}.$$

On a donc

$$s(t) = \sqrt{2} \int_0^t e^{-x} dx = \sqrt{2} \int_0^t e^{-x} dx = \sqrt{2}(1 - e^{-t}).$$



Spirale Logarithmique.

Proposition 4.1.5 *Tout arc (I, f) de classe C^1 **régulier** est C^1 -équivalent à un arc paramétré par sa longueur d'arc. De plus si deux arcs paramétrés par la longueur d'arc sont équivalents, les paramétrages sont les mêmes à translation et changement d'orientation près.*

Preuve. Soit $a \in I$ et posons pour tout $t \in I$,

$$s(t) := \int_a^t \|f'(x)\| dx.$$

Alors $s : I \rightarrow \mathbb{R}$ est une fonction de classe C^1 , croissante, et comme (I, f) est régulier, on a

$$s'(t) = \|f'(t)\| > 0.$$

Donc $s : I \rightarrow s(I)$ est un C^1 -difféomorphisme et notons $\theta := s^{-1} : J = s(I) \rightarrow I$ son application réciproque. Posons $g = f \circ \theta$, définit sur J . L'arc (J, g) est C^1 -équivalent à (I, f) et on a de plus pour tout $x \in J$,

$$\|g'(x)\| = \|f'(\theta(x))\| \cdot |\theta'(x)| = \frac{\|f'(\theta(x))\|}{|s'(\theta(x))|} = 1.$$

Supposons maintenant que (J, g) et (U, h) soient deux arcs C^1 -équivalents tous deux paramétrés par la longueur d'arc, par exemple $g = h \circ \theta$ avec $\theta : J \rightarrow U$ un C^1 -difféo. On a alors pour tout $x \in J$,

$$\|g'(x)\| = 1 = \|h'(\theta(x))\| \cdot |\theta'(x)| = |\theta'(x)|,$$

donc $\theta'(x) = \pm 1$, pour tout $x \in J$. Comme θ est continu, le théorème des valeurs intermédiaires implique que θ' est constante sur J . Ainsi par intégration on a pour tout $x \in J$, $\theta(x) = x + c$ ou $\theta(x) = -x + c$, où c est une constante. \square

Pour comprendre l'aspect *local* des courbes paramétrées, l'outil essentiel est la formule de Taylor-Young vectorielle, que l'on rappelle ici. On rappelle que si $f : I \rightarrow \mathbb{R}^n$ et $g : I \rightarrow \mathbb{R}$ sont deux fonctions, on dit que au voisinage d'un point $t_0 \in I$ on a

$$f(t) = o(g(t)),$$

ssi pour tout $\epsilon > 0$, il existe $\eta > 0$ tel que pour tout $t \in [t_0 - \eta, t_0 + \eta]$,

$$\|f(t)\| \leq \epsilon |g(t)|.$$

On dira aussi qu'au voisinage de $t_0 \in I$ on a

$$f(t) = O(g(t)),$$

ssi il existe $M > 0$ et $\eta > 0$ tel que pour tout $t \in [t_0 - \eta, t_0 + \eta]$,

$$\|f(t)\| \leq M |g(t)|.$$

Proposition 4.1.6 (*Taylor-Young*) Soit $f : I \rightarrow \mathbb{R}^n$ une fonction de classe C^k . Soit $t_0 \in I$, on a alors pour tout h suffisamment petit,

$$f(t_0 + h) = f(t_0) + hf'(t_0) + \frac{h^2}{2}f''(t_0) + \dots + \frac{h^k}{k!}f^{(k)}(t_0) + o(h^k).$$

Preuve. C'est simplement la formule de Taylor-Young scalaire appliquée coordonnées par coordonnées, voir cours de L1. \square

4.2 Etude locale dans le plan

Dans la suite de ce chapitre, on supposera que l'on est dans \mathbb{R}^2 .

Définition 4.2.1 Soit (I, f) un arc C^2 dans le plan. Un point t_0 est dit **birégulier** ssi $f'(t_0)$ et $f''(t_0)$ forment une base de \mathbb{R}^2 .

Définition 4.2.2 Soit (I, f) un arc C^2 dans le plan. Soit $t_0 \in I$ un point birégulier. Le demi-plan de concavité fermé C_p^+ en $p = f(t_0) \in f(I)$ est par définition

$$C_p^+ := p + \mathbb{R}f'(t_0) + \mathbb{R}^+ f''(t_0).$$

Le demi-plan de concavité **ouvert** est défini par

$$C_p^{+,0} := p + \mathbb{R}f'(t_0) + \mathbb{R}_*^+ f''(t_0).$$

Il n'est pas difficile de voir que le demi-plan de concavité ne dépend pas du choix du paramétrage (exercice). L'intérêt de cette définition réside dans le résultat suivant.

Théorème 4.2.3 Soit (I, f) un arc C^2 dans le plan. Soit $t_0 \in I$ un point birégulier, et posons $p = f(t_0)$. Alors il existe $\eta > 0$ tel que pour tout $t \in [t_0 - \eta, t_0 + \eta]$ $f(t) \in C_p^+$. De plus, si $t \in [t_0 - \eta, t_0 + \eta] \setminus \{t_0\}$, alors $f(t) \in C_p^{+,0}$.

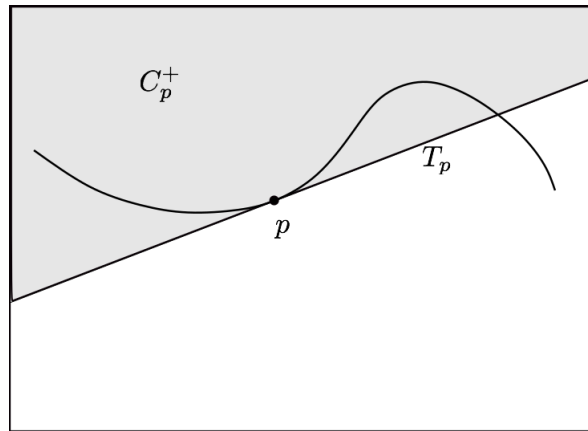
Preuve. On applique la formule de Taylor-Young à l'ordre 2 en t_0 . On a pour tout h voisin de 0,

$$f(t_0 + h) - f(t_0) = hf'(t_0) + \frac{h^2}{2}f''(t_0) + o(h^2).$$

En décomposant le reste $o(h^2)$ dans la base $(f'(t_0), f''(t_0))$, on a alors

$$f(t_0 + h) - f(t_0) = h(1 + o(h))f'(t_0) + \frac{h^2}{2}(1 + o(1))f''(t_0),$$

et il est alors clair que pour h suffisamment petit, la composante de $f''(t_0)$ est positive, i.e. $f(t_0 + h) \in C_p^+$. Si de plus $h \neq 0$, pour tout h assez petit on a $f(t_0 + h) \in C_p^{+,0}$. \square



Demi-plan de concavité C_p^+ et position locale de la courbe.

Courbes polaires. Un arc paramétré (I, f) est dit polaire ssi il est de la forme

$$f(t) = (\rho(t) \cos(t), \rho(t) \sin(t)),$$

pour tout $t \in I$. Alternativement, on peut aussi utiliser les nombres complexes et écrire

$$f(t) = \rho(t)e^{it}.$$

On peut observer que pour les courbes polaires, un point est singulier ssi $\rho(t) = 0$ et $\rho'(t) = 0$. En particulier il ne peut y avoir de point singulier qu'à l'origine. On peut appliquer le théorème précédent aux courbes polaires pour obtenir le critère de concavité suivant.

Proposition 4.2.4 *Soit une courbe polaire $(I, \rho(t)e^{it})$, de classe C^2 . Un point où $\rho(t_0) \neq 0$ est birégulier ssi $\rho^2 + 2\rho'^2 - \rho\rho'' \neq 0$. La courbe tourne sa concavité vers l'origine ssi*

$$\rho^2 + 2\rho'^2 - \rho\rho'' > 0.$$

Preuve. On cherche à déterminer $\lambda_1, \lambda_2 \in \mathbb{R}$ tels que

$$0 = f(t_0) + \lambda_1 f'(t_0) + \lambda_2 f''(t_0).$$

En dérivant deux fois on a

$$f'(t) = \rho'(t)e^{it} + i\rho(t)e^{it}, \quad f''(t) = \rho''(t)e^{it} + 2\rho'(t)ie^{it} - \rho(t)e^{it}.$$

Dire que t_0 est birégulier c'est dire que $\det(f'(t_0), f''(t_0)) \neq 0$, ce qui dans la base e^{it_0}, ie^{it_0} s'écrit

$$\begin{vmatrix} \rho' & \rho'' - \rho \\ \rho & \rho' \end{vmatrix} = 2\rho'^2 + \rho^2 - \rho\rho'' \neq 0.$$

En résolvant le système par substitution, on trouve dans le cas birégulier

$$\lambda_2 = \frac{\rho^2}{2\rho'^2 + \rho^2 - \rho\rho''},$$

et donc $0 \in C_p^{+,0}$ ssi $\lambda_2 > 0$ i.e. $2\rho'^2 + \rho^2 - \rho\rho'' > 0$. Cqfd. \square

Un premier exemple d'étude : le limaçon de Pascal. On se propose d'étudier la courbe polaire

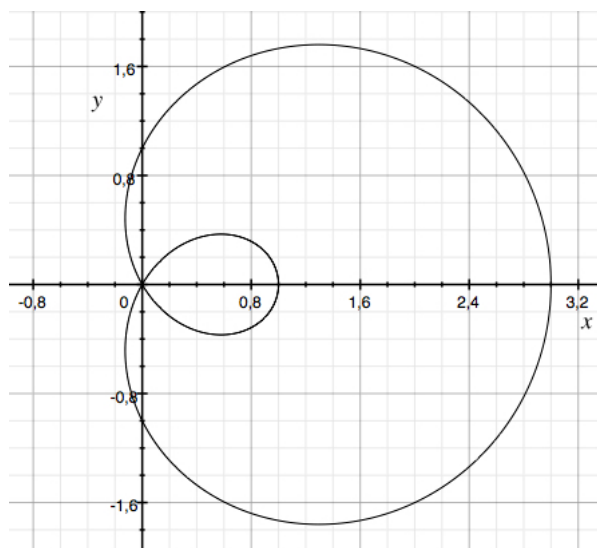
$$f(t) = \rho(t)e^{it} = (1 + 2\cos(t))e^{it},$$

avec $I = [0, 2\pi]$. On observe d'abord que $f(2\pi - t) = \overline{f(t)}$, donc la courbe est symétrique par rapport à l'axe des abscisses et on a $f(0) = f(2\pi) = 3$,

$f(2\pi/3) = f(4\pi/3) = 0$, $f(\pi) = 1$. La courbe passe par l'origine uniquement pour $t = 2\pi/3, 4\pi/3$, et en ces points $\rho' \neq 0$. Donc tout point de cette courbe est régulier. Pour tout $t \neq 2\pi/3, 4\pi/3$, on a par calcul direct que

$$2\rho'^2 + \rho^2 - \rho\rho' = 9 + 6\cos(t) > 0,$$

donc tout point est birégulier et la concavité de la courbe est toujours tournée vers l'origine. Comme $\rho'(t) = -2\sin(t)$, ρ décroît sur $[0, \pi]$ et croît sur $[\pi, 2\pi]$. Attention, le changement de signe de ρ en $t = 2\pi/3$ induit un déphasage de $+\pi$ pour l'argument de $f(t)$ sur l'intervalle $[2\pi/3, 4\pi/3]$. On peut donc esquisser un tracé de la courbe, qui est dessinée ci-dessous.



Le limaçon de Pascal.

Que se passe-t-il en présence de point non-birégulier, voire non-régulier ? En l'absence d'hypothèses supplémentaires, même dans le cas C^∞ , tout peut arriver. Donnons un exemple. Considérons la fonction $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ définie par

$$\varphi(x) = \begin{cases} e^{-\frac{1}{1-x^2}} & \text{si } |x| < 1 \\ 0 & \text{sinon.} \end{cases}$$

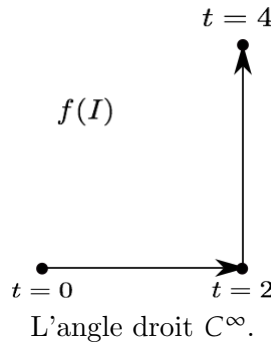
Cette fonction est positive, strictement positive sur $] -1, 1[$. On peut montrer, par récurrence, que φ est C^∞ sur \mathbb{R} . On définit alors la fonction $F_0 : \mathbb{R} \rightarrow \mathbb{R}$ par

$$F_0(x) = \frac{\int_{-\infty}^x \varphi(t) dt}{\int_{-1}^1 \varphi(t) dt}.$$

La fonction F_0 est aussi C^∞ , croissante, vaut $F_0(x) = 0$ pour $x \leq -1$ et vaut $F_0(x) = 1$ pour $x \geq 1$. A partir de F_0 , on construit alors l'arc paramétré (I, f) suivant. On pose $I = [0, 4]$ et

$$f(t) = (F_0(t-1), F_0(t-3)).$$

On observe que pour tout $t \in [0, 2]$, $f(t) = (F_0(t-1), 0)$ avec $t \mapsto F_0(t-1)$ qui croit de 0 à 1. Puis pour tout $t \in [2, 4]$, $f(t) = (1, F_0(t-3))$ avec $t \mapsto F_0(t-3)$ qui croit de 0 à 1. La courbe décrite fait un angle droit !



On peut observer qu'en $t = 2$, on a un point singulier et $f^{(k)}(2) = 0$ pour tout $k \geq 1$. Pour exclure ce type de pathologies, on va définir une notion de point singulier non-dégénéré.

Définition 4.2.5 Soit (I, f) un arc paramétré de classe C^n , avec $n \geq 3$. Soit $t_0 \in I$ un point singulier. On dit qu'il est non dégénéré s'il existe $2 \leq k, \ell \leq n$ tels que $(f^{(k)}(t_0), f^{(\ell)}(t_0))$ soit libre.

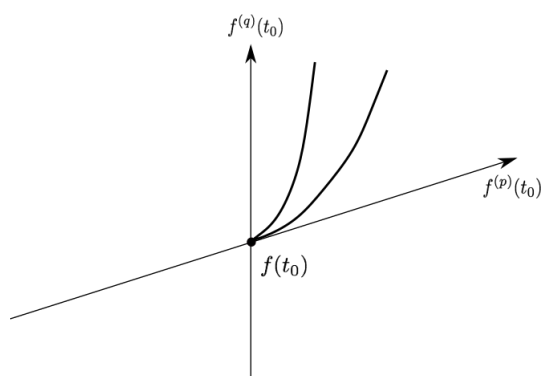
Théorème 4.2.6 Soit (I, f) un arc paramétré de classe C^n , avec $n \geq 3$, et soit $t_0 \in I$ un point singulier non dégénéré. Posons alors

$$p = \min\{2 \leq k \leq n : f^{(k)}(t_0) \neq 0\}, \quad q = \min\{p \leq \ell \leq n : (f^{(p)}(t_0), f^{(\ell)}(t_0)) \text{ libre}\}.$$

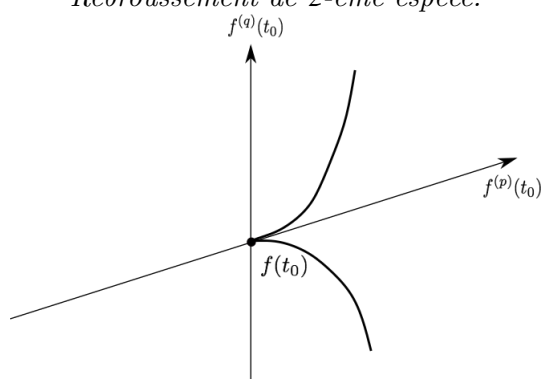
On a alors 4 cas possibles, en fonction des parités de p et q .

Parité de p	Parité de q	Type de singularité
Pair	Pair	Rebroussement de deuxième espèce
Pair	Impair	Rebroussement de première espèce
Impair	Pair	Point de concavité ordinaire
Impair	Impair	Point d'inflexion

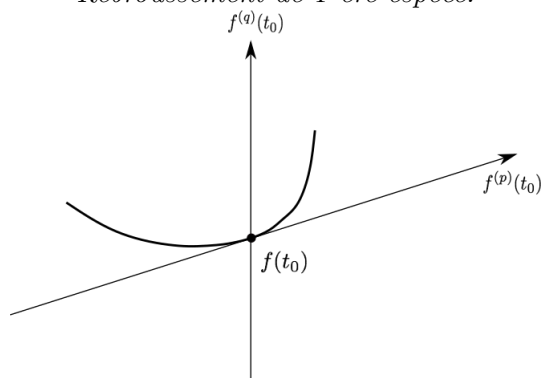
Chaque type de singularité est décrite dans les dessins suivants.



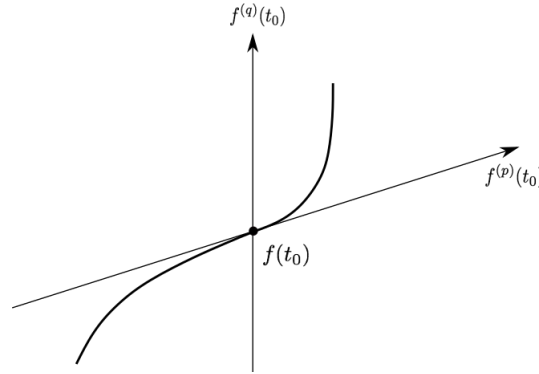
Rebroussement de 2-ème espèce.



Rebroussement de 1-ère espèce.



Concavité ordinaire.



Point d'inflexion.

Preuve. On applique la formule de Taylor-Young. On a pour tout h voisin de 0,

$$f(t_0 + h) - f(t_0) = \frac{h^p}{p!} f^{(p)}(t_0) + \frac{h^q}{q!} f^{(q)}(t_0) + o(h^q),$$

Comme les deux vecteurs $f^{(p)}(t_0), f^{(q)}(t_0)$ forment une base, on peut décomposer le reste $o(h^q)$ dans cette base pour écrire

$$f(t_0 + h) - f(t_0) = \frac{h^p(1 + o(h^{q-p}))}{p!} f^{(p)}(t_0) + \frac{h^q(1 + o(1))}{q!} f^{(q)}(t_0),$$

et les parités de p et q déterminent le signe des coordonnées de $f(t_0 + h) - f(t_0)$ exprimé dans la base $(f^{(p)}(t_0), f^{(q)}(t_0))$ pour h voisin de 0. Il y a donc 4 cas possibles qui sont résumés dans les dessins ci-dessus. Dans chaque cas, la direction de la tangente à la courbe est $f^{(p)}(t_0)$. \square

Un exemple : l'astroïde. En pratique, pour déterminer la nature d'un point singulier, on utilise des développements limités, plus facile à manipuler que la formule directe de Taylor-Young. Regardons l'exemple de (I, f) avec $I = [0, 2\pi]$ et

$$f(t) = (\cos^3(t), \sin^3(t)).$$

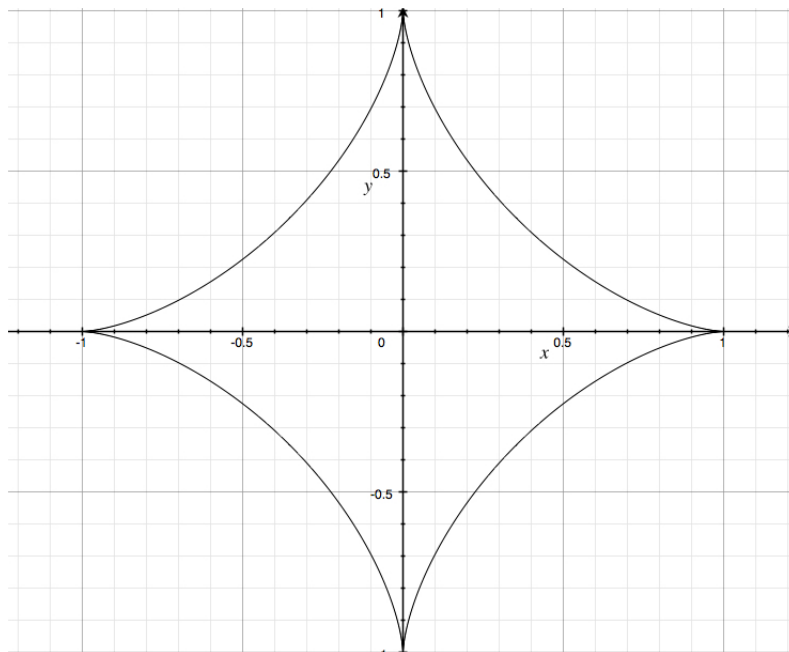
On observe qu'en $t = 0$, $f'(t) = 0$, on a donc un point singulier. En faisant un développement limité en $t = 0$ de $\cos^3(t)$ et $\sin^3(t)$ on trouve

$$\cos^3(t) = 1 - \frac{3}{2}t^2 + O(t^4), \quad \sin^3(t) = t^3 - \frac{t^6}{2} + O(t^7),$$

ce qui, mis sous forme vectorielle, donne

$$f(t) = (1, 0) + t^2 \left(-\frac{3}{2}, 0 \right) + t^3 (0, 1) + o(t^3).$$

On donc $p = 2$ et $q = 3$ et il s'agit d'un point de rebroussement de première espèce.



L'astroïde.

4.3 Exemples de courbes implicites, coniques

Définition 4.3.1 Soit $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction de deux variables. On dit que F définit une courbe implicite \mathcal{C} comme l'ensemble des solutions dans \mathbb{R}^2 de $F(x, y) = 0$. En d'autres termes,

$$\mathcal{C}_F = \{(x, y) \in \mathbb{R}^2 : F(x, y) = 0\}.$$

On peut avoir $\mathcal{C}_F = \emptyset$, par exemple si $F(x, y) = x^2 + y^2 + 1$. L'ensemble \mathcal{C}_F peut être bien différent d'une courbe : par exemple si

$$F(x, y) = F_0(x + y) - 1,$$

où F_0 est la fonction C^∞ "step" définie précédemment, alors on a

$$\mathcal{C}_F = \{(x, y) \in \mathbb{R}^2 : x + y \geq 1\},$$

c'est un demi-plan ! On va donc se restreindre à une classe raisonnable de fonction F , dites polynomiales et les courbes associées \mathcal{C}_F sont dites algébriques. On dit que F est un polynôme réel à deux variables ssi

$$F(x, y) = \sum_{i=0}^N \sum_{j=0}^N a_{i,j} x^i y^j,$$

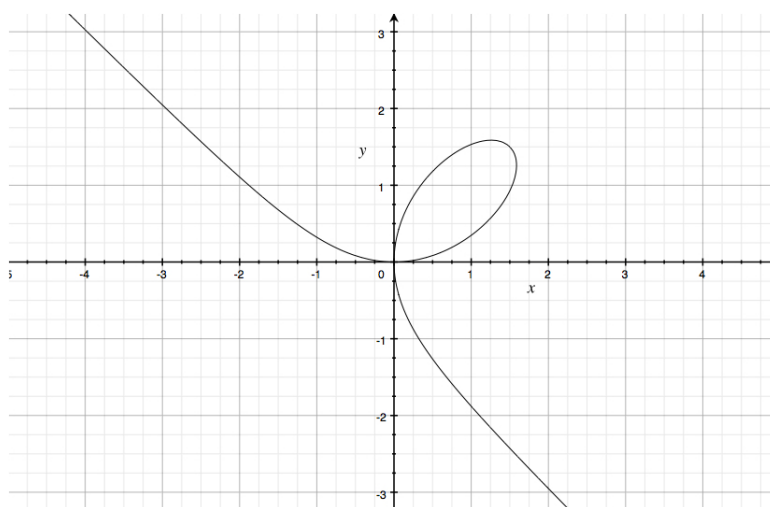
avec $N \in \mathbb{N}$ et $a_{i,j} \in \mathbb{R}$. Le degré de F est par définition

$$\deg(F) = \max\{i + j : a_{i,j} \neq 0\}.$$

Par exemple

$$F(x, y) = x^3 + y^3 - 3xy$$

est de degré 3 et la courbe \mathcal{C}_F associée est appelé "folium de Descartes".



Le folium de Descartes.

Il n'existe pas en général de méthode pour déterminer un paramétrage global ou une résolution des courbes implicites, même dans le cas algébrique. On va étudier quelques cas simples où il est possible de les paramétrer.

4.3.1 Cas où on peut isoler une variable.

C'est la situation où on peut exprimer y en fonction de x , on se ramène à l'étude d'une fonction d'une variable. Par exemple si

$$F(x, y) = y^2 - x^3 + 2x - 1,$$

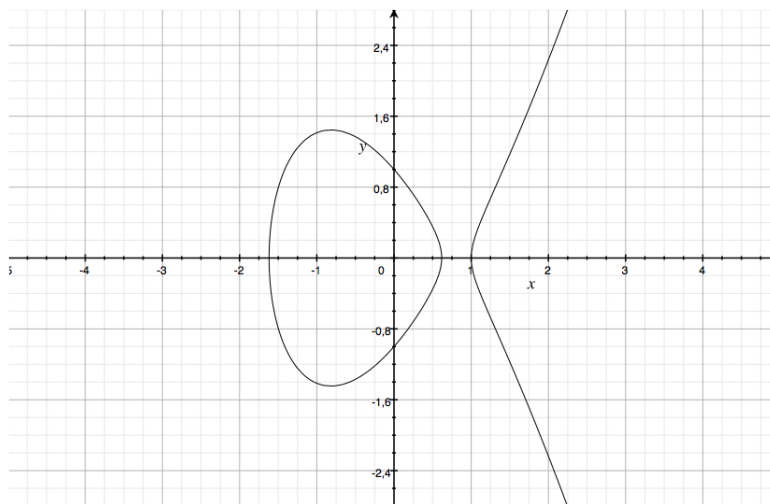
alors $F(x, y) = 0$ ssi $y = \pm\sqrt{x^3 - 2x + 1}$. On observe que

$$f(x) = x^3 - 2x + 1 = (x - 1)(x^2 + x - 1),$$

et les racines de l'équation du second degré $x^2 + x - 1 = 0$ sont $\frac{-1 \pm \sqrt{5}}{2}$. On a donc

$$x^3 - 2x + 1 \geq 0 \Leftrightarrow x \in \left[\frac{-1 - \sqrt{5}}{2}, \frac{-1 + \sqrt{5}}{2} \right] \cup [1, +\infty[.$$

L'étude de la dérivée montre que f est croissante sur $\left[\frac{-1-\sqrt{5}}{2}, -\sqrt{\frac{2}{3}}\right]$, décroissante sur $\left[-\sqrt{\frac{2}{3}}, \frac{-1+\sqrt{5}}{2}\right]$, puis croissante sur $[1, +\infty[$. En calculant quelques points on peut alors tracer la courbe, connue sous le nom de courbe elliptique.



Une courbe elliptique.

4.3.2 Exemples de courbes unicursales.

Les courbes unicursales sont des courbes algébriques qui admettent une paramétrisation rationnelle. On verra que les courbes de degré ≤ 2 sont toujours unicursales. Les courbes de degré 3 qui possèdent un point double sont aussi unicursales. Revenons sur l'exemple du folium de Descartes. Si on intersecte la courbe avec une droite passant par l'origine d'équation $y = tx$, où $t \in \mathbb{R}$ est la pente de la courbe, on voit que sauf pour certaines valeurs particulières de t , il n'y a qu'un seul point d'intersection. En remplaçant y par tx dans $x^3 + y^3 - 3xy = 0$, et en excluant le cas $x = 0$, on obtient

$$x(t) = \frac{3t}{1+t^3}, \quad y(t) = \frac{3t^2}{1+t^3}.$$

La valeur $t = -1$ correspond à l'asymptote oblique $y = -x$. Voir TD pour l'étude paramétrique.

Dans le cas des courbes de degré 4, la situation se complique, mais il existe des cas où des stratégies similaires d'abaissement du degré fonctionnent. On va examiner le cas de la Lemniscate de Bernoulli, de degré 4, dont l'équation est donnée par

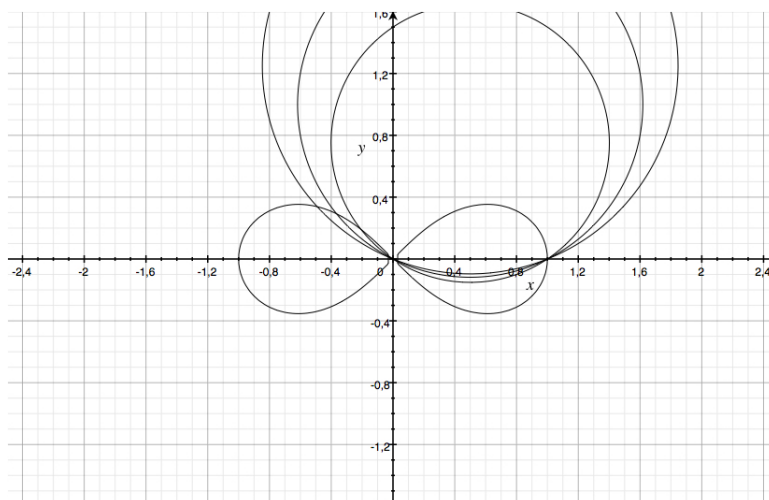
$$P(x, y) = (x^2 + y^2)^2 - x^2 + y^2.$$

Cette équation possède beaucoup de symétries car $P(\pm x, \pm y) = P(x, y)$, en particulier la courbe \mathcal{C}_P est invariante par réflexions par rapport aux axes de coordonnées usuelles et aussi par rotation d'angle π . Les points $(0,0)$ et $(1,0)$ sont dans \mathcal{C}_P . Intersectons donc cette courbe par une famille de cercles passant par ces deux points, donnés par l'équation $x^2 + y^2 = x + ty$, où $t \in \mathbb{R}$ est le paramètre. On résoud donc par substitution le système

$$\begin{cases} (x^2 + y^2)^2 - x^2 + y^2 = 0 \\ x^2 + y^2 = x + ty \end{cases}$$

ce qui en excluant le cas $x = 0$ et $x = 1$ donne

$$x(t) = \frac{1 - t^4}{t^4 + 6t^2 + 1}, \quad y(t) = \frac{2t(t^2 - 1)}{t^4 + 6t^2 + 1}.$$



Lemniscate intersectée avec une famille de cercles passant par $(0,0)$ et $(1,0)$. Le troisième point d'intersection a des coordonnées rationnelles en t .

4.3.3 Cas quadratique : les coniques.

Une conique est une courbe algébrique du type

$$ax^2 + by^2 + 2cxy + dx + ey + f = 0,$$

où $a, b, c, d, e, f \in \mathbb{R}$ sont des coefficients.

Théorème 4.3.2 *Pour toute conique définie comme ci-dessus, il existe un changement de variable affine-orthogonal*

$$\begin{cases} X = a_{11}x + a_{12}y + b_1 \\ Y = a_{21}x + a_{22}y + b_2 \end{cases},$$

avec $A = (a_{ij}) \in O_2(\mathbb{R})$ une matrice orthogonale, tel que l'on se ramène à l'un des 4 cas suivants.

Type	Equation réduite
Une conique dégénérée	\emptyset , un point ou une réunion de droites
Une ellipse	$\frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1$
Une hyperbole	$\frac{X^2}{a^2} - \frac{Y^2}{b^2} = 1$
Une parabole	$Y^2 = 2pX$

Preuve. On commence par examiner la forme quadratique

$$Q(x, y) = ax^2 + by^2 + 2cxy = (x \ y) \begin{pmatrix} a & c \\ c & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Comme la matrice de Q , notée M , est symétrique réelle, elle est diagonalisable en base orthonormée. Il existe donc une matrice orthogonale P telle que $M = {}^tPDP$ où

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

En notant

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = P \begin{pmatrix} x \\ y \end{pmatrix},$$

on a

$$Q(x, y) = \lambda_1(x')^2 + \lambda_2(y')^2.$$

Dans le nouveau système de coordonnées x', y' , on doit donc trouver les solutions de

$$\lambda_1(x')^2 + \lambda_2(y')^2 + d'x' + e'y' + f = 0,$$

où d', e' s'expriment en fonction de d, e via les coefficients de la matrice tP . Il faut alors discuter en fonction des valeurs propres λ_1, λ_2 .

Cas où $\lambda_1 \lambda_2 = 0$. Quitte à permuter les coordonnées, on peut supposer que $\lambda_1 = 0$. Si λ_2 est nul, on est dans un cas dégénéré (droite ou ensemble vide). Sinon, on a en divisant par λ_2 et appliquant l'identité remarquable

$$\left(y' + \frac{e'}{2\lambda_2}\right)^2 - \frac{e'^2}{4\lambda_2^2} + \frac{f}{\lambda_2} + \frac{d'}{\lambda_2}x' = \left(y' + \frac{e'}{2\lambda_2}\right)^2 + \alpha x' + \beta = 0.$$

Si $\alpha = 0$, on est encore dans un cas dégénéré, sinon on pose $Y = y' + \frac{e'}{2\lambda_2}$, $X = x' + \beta/\alpha$. On est alors dans le cas

$$Y^2 + \alpha X = 0,$$

c'est une parabole.

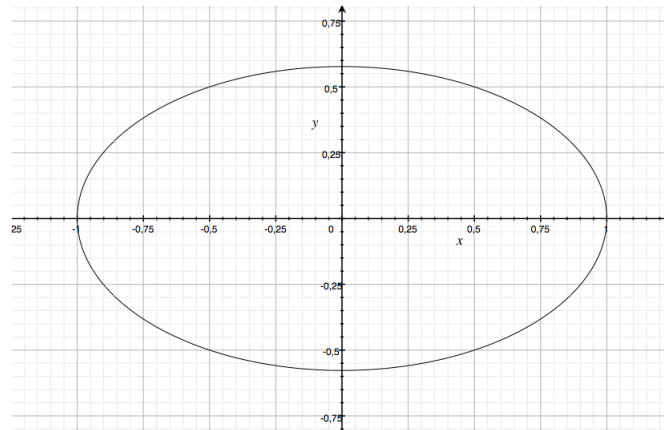
Cas où $\lambda_1 \lambda_2 \neq 0$. On écrit dans ce cas

$$\lambda_1 \left(x' + \frac{d'}{2\lambda_1} \right)^2 - \frac{d'^2}{4\lambda_1} + \lambda_2 \left(y' + \frac{e'}{2\lambda_2} \right)^2 - \frac{e'^2}{4\lambda_2} + f = 0$$

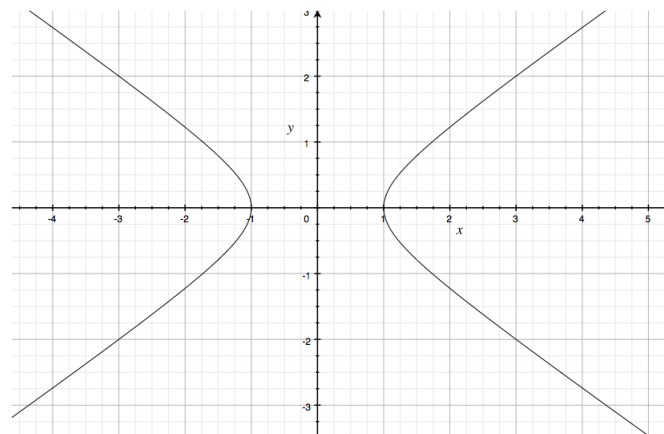
En posant $X = x' + \frac{d'}{2\lambda_1}$, $Y = y' + \frac{e'}{2\lambda_2}$, on tombe sur une équation du type

$$\lambda_1 X^2 + \lambda_2 Y^2 + \gamma = 0.$$

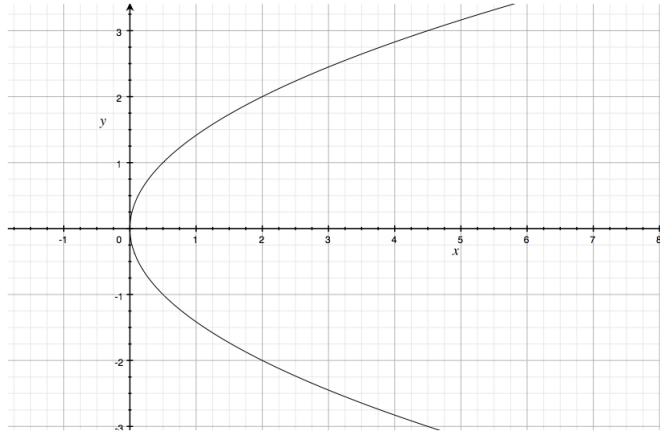
Si λ_1, λ_2 sont de même signe, c'est une ellipse ou un cas dégénéré. Sinon, c'est une hyperbole (ou un cas dégénéré). \square



Ellipse.



Hyperbole.



Parabole.

Remarques. Le produit des valeurs propres $\lambda_1 \lambda_2$ est égal au déterminant de M soit $ab - c^2$. On peut donc voir que si $ab - c^2 = 0$, on est dans le cas d'une parabole (ou un cas dégénéré). Si $ab - c^2 > 0$, c'est une ellipse (ou un cas dégénéré), et si $ab - c^2 < 0$, c'est une hyperbole (ou un cas dégénéré).

L'excentricité \mathcal{E} d'une conique est définie (pour $ab - c^2 \neq 0$) par

$$\mathcal{E} = \sqrt{1 - \frac{\lambda_1}{\lambda_2}},$$

où on a supposé $\lambda_1 \leq \lambda_2$ et $\lambda_2 \geq 0$. Ainsi dans le cas d'une ellipse, $0 \leq \mathcal{E} < 1$, et dans le cas d'une hyperbole on a $\mathcal{E} > 1$.

Paramétrages des coniques. L'ellipse sous sa forme réduite se paramétrise par

$$\gamma(\theta) = (a \cos(\theta), b \sin(\theta)) \quad \theta \in [-\pi, +\pi].$$

En posant $t = \tan(\theta/2)$, où $\theta \in]-\pi, +\pi[$, comme on a

$$\cos(\theta) = \frac{1 - t^2}{1 + t^2} \quad \sin(\theta) = \frac{2t}{1 + t^2},$$

on obtient un paramétrage rationnel de l'ellipse (moins un point) par

$$\gamma(t) = \left(a \frac{1 - t^2}{1 + t^2}, b \frac{2t}{1 + t^2} \right), \quad t \in]-\infty, +\infty[.$$

La branche positive de l'hyperbole se paramétrise par

$$\gamma(r) = (a \cosh(r), b \sinh(r)), \quad r \in]-\infty, +\infty[.$$

Comme on a $\cosh(r) = \frac{e^r + e^{-r}}{2}$, $\sinh(r) = \frac{e^r - e^{-r}}{2}$, en posant $t = e^r$, on obtient un paramétrage rationnel

$$\gamma(t) = \frac{1}{2} \left(at + \frac{a}{t}, bt - \frac{b}{t} \right), \quad t \in]0, +\infty[.$$

En coordonnées polaires, on peut montrer (voir TD), que les coniques non dégénérées s'écrivent sous la forme

$$\gamma(t) = \frac{p}{1 + \mathcal{E} \cos(t)} e^{it},$$

où $p > 0$ et \mathcal{E} est l'excentricité de la conique.

4.4 Courbure dans le plan

Considérons un arc C^2 , régulier, noté (I, f) . On posera pour tout $t \in I$,

$$\tau(t) = \frac{f'(t)}{\|f'(t)\|},$$

qui est le vecteur unitaire tangent à l'arc en $f(t)$. On pose aussi $N(t) = \tau(t)^\perp$, le vecteur directement orthogonal à $\tau(t)$.

Définition 4.4.1 Soit (I, f) un arc régulier de classe C^2 , paramétré par sa longueur d'arc. Pour tout $t \in I$, il existe un unique scalaire $\kappa(t)$ tel que

$$f''(t) = \kappa(t)N(t),$$

c'est la courbure algébrique de l'arc au point $f(t)$, qui est une fonction continue de $t \in I$.

Cette définition demande quelques commentaires. Comme on a pour tout $t \in I$, $\|f'(t)\|^2 = 1$, on obtient en dérivant,

$$\frac{d}{dt} (\|f'(t)\|^2) = 2\langle f''(t), f'(t) \rangle = 0,$$

ainsi $f''(t) \in \tau(t)^\perp$, i.e. $f''(t), N(t)$ sont colinéaires. Il existe donc bien un scalaire $\kappa(t)$ tel que $f''(t) = \kappa(t)N(t)$. La continuité découle du fait que $\kappa(t) = \langle f''(t), N(t) \rangle$ qui est bien continue car f est C^2 .

Cette définition ne dépend pas du choix du paramétrage par longueur d'arc : par la proposition 4.1.5, si on a deux paramétrages (I, f) et (J, g) par l.a. C^2 -équivalents, on a

$$f(t) = g(\pm t + c),$$

et donc

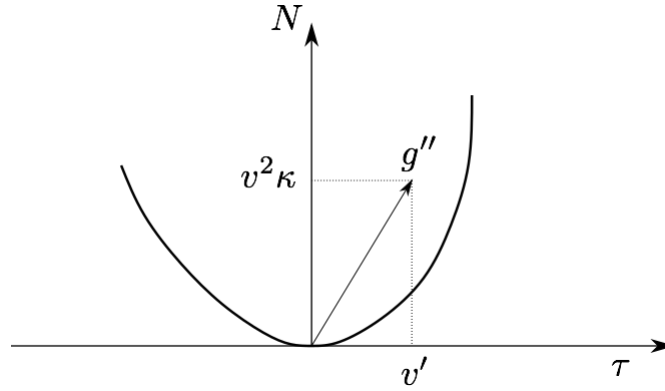
$$f''(t) = g''(\pm t + c).$$

En général, il est difficile de calculer la longueur des arcs et d'explicitier un paramétrage par l.a. On peut néanmoins quand même calculer la courbure de la façon suivante. Supposons que (I, f) soit un arc paramétré par l.a. et (J, g) un arc C^2 -équivalent tel que $g = f \circ \theta$. On a pour tout $t \in I$,

$$\begin{aligned} g''(t) &= (f'' \circ \theta(t))(\theta'(t))^2 + (f' \circ \theta(t))\theta''(t) \\ &= (\theta'(t))^2 \kappa(\theta(t))N(\theta(t)) + \theta''(t)\tau(\theta(t)). \end{aligned}$$

En posant $\theta' = v$, cette formule (dite de Frenet) s'écrit donc

$$g'' = v^2 \kappa N + v' \tau.$$



Repère mobile de Frenet et courbure.

Ainsi on a

$$\langle g''(t), N(\theta(t)) \rangle = (\theta'(t))^2 \kappa(\theta(t)),$$

et comme $\|g'(t)\| = |\theta'(t)|$, on déduit l'identité

$$\kappa(\theta(t)) = \frac{\langle g''(t), N(\theta(t)) \rangle}{\|g'(t)\|^2}.$$

De ce calcul on déduit directement les formules suivantes.

Proposition 4.4.2 Soit $\gamma(t) = (x(t), y(t))$, défini sur I , un arc régulier de classe C^2 . Alors la courbure algébrique est donnée par

$$\kappa(t) = \frac{\det(\gamma'(t), \gamma''(t))}{\|\gamma'(t)\|^3} = \frac{y''(t)x'(t) - x''(t)y'(t)}{((x'(t))^2 + (y'(t))^2)^{3/2}}.$$

Si on a $\gamma(t) = \rho(t)e^{it}$, alors la courbure s'exprime en polaire par

$$\kappa(t) = \frac{\rho^2(t) + 2(\rho'(t))^2 - \rho(t)\rho''(t)}{(\rho^2(t) + (\rho'(t))^2)^{3/2}}.$$

On observe donc que la courbure est non nulle ssi $\gamma(t)$ est birégulier. Le signe de la courbure est donné par l'orientation de la base $(\gamma'(t), \gamma''(t))$. En particulier, si on renverse l'orientation, la courbure algébrique est changée en son opposée. Dans le cas du cercle où $\rho(t) = R$ est constante, on a donc

$$\kappa(t) = \frac{1}{R}.$$

Dans le cas d'une ellipse, que l'on paramètre par

$$\gamma(t) = (a \cos(t), b \sin(t)), \quad t \in [0, 2\pi],$$

avec $0 < b < a$, on trouve

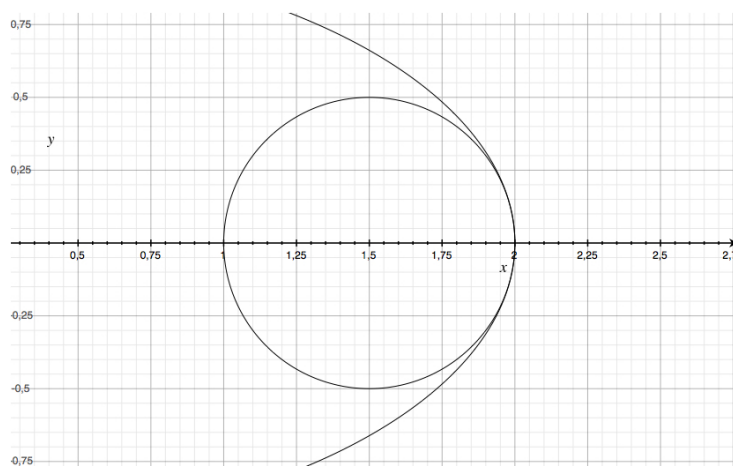
$$\kappa(t) = \frac{ab}{(a^2 \sin^2(t) + b^2 \cos^2(t))^{3/2}},$$

et donc

$$\frac{b}{a^2} \leq \kappa(t) \leq \frac{a}{b^2}.$$

Définition 4.4.3 (*Cercle osculateur*). Soit (I, f) un arc régulier de classe C^2 . Soit $t \in I$ tel que $\kappa(t) \neq 0$. Le centre de courbure est le point $C(t) = f(t) + R(t)N(t)$, où $R(t) = (\kappa(t))^{-1}$ est appelé rayon de courbure algébrique. Le cercle de centre $C(t)$ et de rayon $|R(t)|$ est appelé cercle osculateur. Le lieu des centres de courbure s'appelle la développée.

Par exemple, au point $(a, 0)$ de l'ellipse, le rayon de courbure est $\frac{b^2}{a}$ et le centre de courbure est en $(\frac{a^2-b^2}{a}, 0)$.



Cercle osculateur de l'ellipse au point $(a, 0)$.

Calculons maintenant la développée de l'ellipse. On a par calcul direct

$$N(t) = \frac{(-b \cos(t), -a \sin(t))}{\sqrt{a^2 \sin^2(t) + b^2 \cos^2(t)}},$$

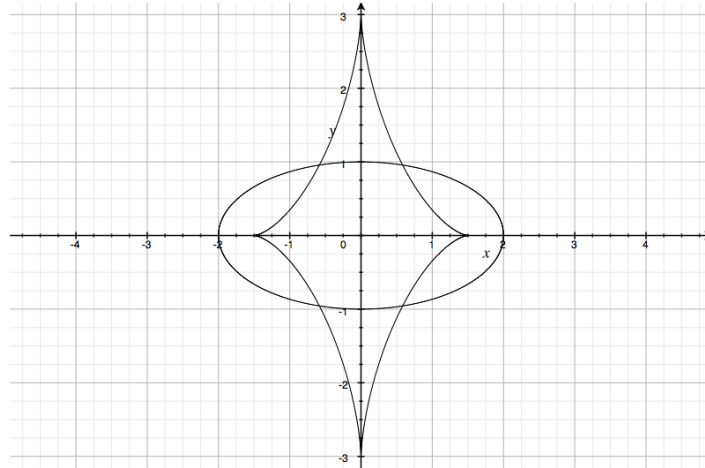
et $C(t) = \gamma(t) + R(t)N(t)$, avec

$$R(t) = \frac{(a^2 \sin^2(t) + b^2 \cos^2(t))^{3/2}}{ab}.$$

On obtient après calculs

$$C(t) = \left(\frac{a^2 - b^2}{a} \cos^3(t), -\frac{a^2 - b^2}{b} \sin^3(t) \right),$$

on reconnaît les équations paramétriques d'une astroïde.



Developpée de l'ellipse avec $a = 2$, $b = 1$.

4.5 La formule de Gauss-Bonnet en dimension 1

Dans ce paragraphe, on va voir le lien entre courbure et variation de l'angle fait par la tangente avec un vecteur fixé. On va avoir besoin du fait fondamental suivant, appelé "théorème du relèvement". Dans ce qui suit S^1 désigne le cercle unité du plan complexe \mathbb{C} , ou de \mathbb{R}^2 .

$$S^1 := \{z \in \mathbb{C} : |z| = 1\}.$$

Théorème 4.5.1 (*Relèvement C^1*). Soit I un intervalle de \mathbb{R} et $\gamma : I \rightarrow S^1$ une application C^1 . Alors il existe $\varphi : I \rightarrow \mathbb{R}$, de classe C^1 tel que pour tout $t \in I$,

$$\gamma(t) = e^{i\varphi(t)}.$$

On dit que φ est une relevée de γ . Si de plus $\tilde{\varphi}$ a les mêmes propriétés, alors il existe une constante $\alpha \in 2\pi\mathbb{Z}$ tel que pour tout $t \in I$, $\varphi(t) = \tilde{\varphi}(t) + \alpha$.

Preuve. Voir dans l'annexe pour les détails, c'est essentiellement de l'intégration de fonctions à valeurs complexes bien choisies. \square

Proposition 4.5.2 Soit (I, f) un arc régulier de classe C^2 , à valeurs dans \mathbb{C} . Il existe $\varphi : I \rightarrow \mathbb{R}$, de classe C^1 tel que pour tout $t \in I$, on a $\tau(t) = \frac{f'(t)}{|f'(t)|} = e^{i\varphi(t)}$. De plus pour tout $t \in I$, on a

$$\kappa(t)|f'(t)| = \frac{d\varphi}{dt}(t).$$

Preuve. Comme on a $|\tau(t)| = 1$ pour tout $t \in I$. Comme f est C^2 , f' est C^1 et on peut lui appliquer le théorème du relèvement ci-dessus. Il existe donc $\varphi : I \rightarrow \mathbb{R}$, de classe C^1 tel que $\tau(t) = e^{i\varphi(t)}$. Par ailleurs, on a la formule

$$\kappa(t) = \frac{\det(f'(t), f''(t))}{|f'(t)|^3},$$

et on a $f'(t) = |f'(t)|e^{i\varphi(t)}$, donc en dérivant on trouve

$$f''(t) = \varphi'(t)ie^{i\varphi(t)}|f'(t)| + \frac{\langle f'(t), f''(t) \rangle}{|f'(t)|}e^{i\varphi(t)}.$$

Ainsi on a

$$\det(f'(t), f''(t)) = |f'(t)|^2 \varphi'(t) \underbrace{\det(e^{i\varphi(t)}, ie^{i\varphi(t)})}_{=1} + \langle f'(t), f''(t) \rangle \underbrace{\det(e^{i\varphi(t)}, e^{i\varphi(t)})}_{=0},$$

donc $\kappa(t) = \frac{\varphi'(t)}{|f'(t)|}$, le calcul est terminé. \square

De ce lien entre courbure et variation d'angle, on peut tirer la première application suivante.

Théorème 4.5.3 Soit I un intervalle de \mathbb{R} et $c : I \rightarrow \mathbb{R}$ une fonction continue. On se donne $t_0 \in I$ et $a, b \in \mathbb{C}$. Alors il existe un unique arc C^2 , noté (I, f) , paramétré par la longueur d'arc, tel que $f(t_0) = a$, $f'(t_0) = b$ et pour tout $t \in I$, $\kappa(t) = c(t)$. Autrement dit, la courbure algébrique détermine complètement l'arc à une isométrie affine près.

Preuve. Posons pour tout $t \in I$,

$$\varphi(t) := \alpha + \int_{t_0}^t c(t) dt,$$

avec $b = e^{i\alpha}$. C'est une fonction de classe C^1 puisque c est continue, et $\varphi'(t) = c(t)$. On pose ensuite

$$f(t) = a + \int_{t_0}^t e^{i\varphi(s)} ds,$$

f est bien C^2 et $f'(t) = \tau(t) = e^{i\varphi(t)}$, $f''(t) = c(t)i\tau(t) = c(t)N(t)$, donc $\kappa(t) = c(t)$. L'unicité se voit en faisant le raisonnement inverse : si un tel arc existe, on relève $\tau(t) = e^{i\varphi(t)}$ et on doit avoir $\varphi'(t) = c(t)$ et donc f doit vérifier la formule ci-dessus. \square

Le nombre d'enroulement d'un lacet. Soit (I, f) un arc régulier C^2 . Si $I = [a, b]$, on dit que (I, f) est un *lacet* ssi on a

$$f(a) = f(b) \text{ et } f'(a) = f'(b).$$

Définition 4.5.4 (*Enroulement d'un lacet*). Soit (I, f) un lacet comme ci-dessus, et $\varphi : I \rightarrow \mathbb{R}$ une relevée de $\tau : I \rightarrow S^1$. Alors la quantité

$$\text{Enroul}(I, f) := \frac{\varphi(b) - \varphi(a)}{2\pi} \in \mathbb{Z}$$

ne dépend pas du choix du relèvement.

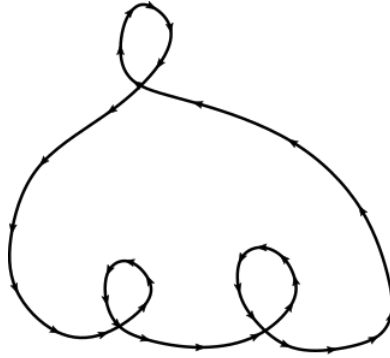
En effet, comme $f'(a) = f'(b)$, on a $\tau(a) = \tau(b)$ et $e^{i(\varphi(b) - \varphi(a))} = 1$ donc

$$\varphi(b) - \varphi(a) \in 2\pi\mathbb{Z}.$$

Si $\widetilde{\varphi}$ est une autre relevée de τ , alors $\widetilde{\varphi} = \varphi + \alpha$ et ainsi

$$\widetilde{\varphi}(b) - \widetilde{\varphi}(a) = \varphi(b) - \varphi(a).$$

L'enroulement d'un lacet compte le nombre de tours (avec orientation) que fait la tangente à la courbe quand on parcourt tout le lacet.



$$\text{Enroul}(I, f) = 2$$

Nombre d'enroulement d'un lacet.

Théorème 4.5.5 (*Gauss-Bonnet pour les courbes*) Soit (I, f) un lacet C^2 , avec $I = [\alpha, \beta]$. Alors on a l'identité

$$\int_{\alpha}^{\beta} \kappa(t) |f'(t)| dt = 2\pi \text{Enroul}(I, f).$$

Preuve. On a tout simplement

$$\int_{\alpha}^{\beta} \kappa(t) |f'(t)| dt = \int_{\alpha}^{\beta} \varphi'(t) dt = \varphi(\beta) - \varphi(\alpha) = 2\pi \text{Enroul}(I, f).$$

et c'est fini. \square

Cette formule remarquable montre que pour les lacets, l'intégrale de la courbure donne 2π fois le nombre d'enroulement. Ce qui est non-trivial, c'est que le résultat ne dépend au final que du nombre d'enroulement et pas des propriétés métriques de la courbe (longueur et courbure). Quand on l'applique par exemple à l'ellipse, on obtient

$$\int_0^{2\pi} \frac{ab}{a^2 \sin^2(t) + b^2 \cos^2(t)} dt = 2\pi,$$

et ceci quel que soit $a, b > 0$. Exercice : montrer cette identité directement.

Chapitre 5

Courbes et surfaces dans \mathbb{R}^3

5.1 Courbes gauches

Dans ce paragraphe, on va étendre l'étude métrique des arcs paramétrés au cas de la dimension 3. Les arcs paramétrés en dimension 3 sont appelés historiquement des "courbes gauches". On suppose *dans tout ce qui suit* que (I, f) est un arc paramétré C^3 , avec $f : I \rightarrow \mathbb{R}^3$, **birégulier** i.e. pour tout $t \in I$, $(f'(t), f''(t))$ sont linéairement indépendants.

Si $m = f(t)$, le plan $O_m := f(t) + \mathbb{R}f'(t) + \mathbb{R}f''(t)$ ne dépend pas du paramétrage (le vérifier en exercice) et est appelé plan *osculateur* à la courbe (I, f) au point m .

On rappelle que l'on a posé $\tau(t) = \frac{f'(t)}{\|f'(t)\|}$, qui est le vecteur tangent unitaire à la courbe.

5.1.1 Courbure dans \mathbb{R}^3

Définition 5.1.1 Soit (I, f) un arc paramétré comme ci dessus. On suppose qu'il est paramétré par sa longueur d'arc. Alors pour tout $t \in I$, $f''(t)$ et $f'(t)$ sont orthogonaux et on pose

$$K(t) = \|f''(t)\|, \quad \nu(t) = \frac{f''(t)}{\|f''(t)\|},$$

de sorte que $f''(t) = K(t)\nu(t)$. Le scalaire $K(t) > 0$ est appelé courbure au point $f(t)$.

Comme on a pour tout $t \in I$, $\|f'(t)\|^2 = 1$ (c'est l'hypothèse de paramétrisation l.a.), en dérivant on a directement

$$\frac{d}{dt} (\|f'(t)\|^2) = 0 = 2\langle f''(t), f'(t) \rangle,$$

et donc $\tau(t), \nu(t)$ sont orthogonaux. Attention, *contrairement au cas des courbes planes*, le plan osculateur O_m n'admet pas d'orientation canonique, c'est pour quoi on ne peut définir une notion de courbure algébrique (la notion de vecteur directement orthogonal n'a pas de sens en dimension 3).

La définition de courbure ne dépend pas du choix de la l.a. paramétrisation, comme pour les courbes planes. Comment calcule-t-on la courbure sans expliciter de l.a. paramétrisation ? C'est le même calcul qu'en dimension 2.

Si $g = f \circ \theta$ avec (I, f) une l.a. paramétrisation et $\theta : J \rightarrow I$ un difféomorphisme (au moins C^2), on a alors en dérivant deux fois :

$$g''(t) = (\theta'(t))^2 K(\theta(t)) \nu(\theta(t)) + \theta''(t) \tau(\theta(t)), \quad g'(t) = \theta'(t) \tau(t).$$

En faisant le produit vectoriel $g'(t) \wedge g''(t)$, on a donc (car $\tau(t) \wedge \tau'(t) = 0$),

$$g' \wedge g'' = K \cdot (\theta')^3 \tau \wedge \nu,$$

et en combinant ça avec $\|g'(t)\| = |\theta'|$, on obtient la formule suivante.

Proposition 5.1.2 *Soit (I, γ) un arc paramétré comme ci dessus, pas nécessairement paramétré par sa longueur d'arc. On a pour tout $t \in I$,*

$$K(t) = \frac{\|\gamma'(t) \wedge \gamma''(t)\|}{\|\gamma'(t)\|^3}.$$

5.1.2 Torsion dans \mathbb{R}^3

Soit (I, f) un arc paramétré, birégulier, de classe C^3 . On définit le vecteur

$$\beta(t) = \tau(t) \wedge \nu(t),$$

de sorte que $(\tau(t), \nu(t), \beta(t))$ soit un repère direct de \mathbb{R}^3 . Si (I, f) est l.a. paramétrisé, ce repère est **orthonormé direct** (voir définitions et propriétés du produit vectoriel). On l'appelle trièdre de Frenet au point $f(t)$.

En dérivant l'expression $\|\beta(t)\|^2 = 1$ on obtient pour tout $t \in I$, $\langle \beta'(t), \beta(t) \rangle = 0$. En dérivant l'expression $\langle \beta(t), \tau(t) \rangle = 0$ on obtient pour tout $t \in I$,

$$\langle \beta'(t), \tau(t) \rangle = -\langle \beta(t), \tau'(t) \rangle = -K(t) \langle \beta(t), \nu(t) \rangle = 0.$$

Par conséquent, $\beta'(t)$ et $\nu(t)$ sont colinéaires.

Définition 5.1.3 *Soit (I, f) un arc paramétré l.a., birégulier, de classe C^3 . La torsion $T(t)$ est l'unique réel tel que*

$$\beta'(t) = T(t) \nu(t).$$

En dérivant les identités $\|\nu(t)\|^2 = 1$ et $\langle \nu(t), \tau(t) \rangle = \langle \nu(t), \beta(t) \rangle = 0$ on obtient aussi

$$\nu'(t) = -K(t)\tau(t) - T(t)\beta(t).$$

En récapitulant, on a les **formules de Frenet** dans le trièdre direct (τ, ν, β) :

$$\begin{cases} \tau' &= & K \cdot \nu \\ \nu' &= & -K \cdot \tau & -T \beta \\ \beta' &= & T \cdot \nu \end{cases}.$$

Arrêtons nous un instant sur la signification géométrique de cette quantité "torsion". Si on a pour tout $t \in I$, $T(t) = 0$ alors $\beta'(t) = 0$ i.e. le vecteur $\beta(t) = B$ est constant. De plus on a donc pour tout $t \in I$,

$$\frac{d}{dt}(\langle f(t), B \rangle) = \langle f'(t), B \rangle = \langle \tau(t), B \rangle = 0,$$

ce qui montre que pour tout $t \in I$, $\langle f(t), B \rangle = Cte$. L'image de l'arc est donc incluse dans un plan affine : la courbe est plane. La torsion mesure donc le *défaut de planéité* de la courbe gauche.

Il nous faut expliciter des formules pour calculer la torsion quand l'arc n'est pas l.a. paramétrisé.

Proposition 5.1.4 *Soit (I, g) un arc birégulier de classe C^3 . On a pour tout $t \in I$,*

$$T(t) = -\frac{\det(g'(t), g''(t), g'''(t))}{\|g'(t) \wedge g''(t)\|^2}.$$

Preuve. Comme d'habitude, on écrit $g = f \circ \theta$ avec $\theta : J \rightarrow I$ un C^3 -difféomorphisme, et f un arc paramétrisé par la longueur d'arc. On note toujours (τ, ν, β) le trièdre de Frenet associé à (I, f) . Par les formules de Frenet, on a

$$\begin{cases} f' &= & \tau \\ f'' &= & K \nu \\ f''' &= & K' \nu + K \nu' = K' \nu - K^2 \tau - K T \beta \end{cases}.$$

Par conséquent,

$$\det(f', f'', f''') = \det(\tau, K \nu, K' \nu - K^2 \tau - K T \beta) = -K^2 T \underbrace{\det(\tau, \nu, \beta)}_{=1},$$

d'où

$$T = -\frac{\det(f', f'', f''')}{K^2} = -\frac{\det(f', f'', f''')}{\|f' \wedge f''\|^2}.$$

Dans le cas général, on écrit

$$\begin{aligned} g' &= \theta' f' \\ g'' &= \theta'' f' + (\theta')^2 f'' \\ g''' &= \theta''' f' + 2\theta'' \theta' f'' + (\theta')^3 f''' \end{aligned},$$

et en utilisant les propriétés habituelles du déterminant on a

$$\det(g', g'', g''') = (\theta')^6 \det(f', f'', f''') = -(\theta')^6 K^2 T,$$

et comme on a $\|g'\|^6 = (\theta')^6$ ainsi que

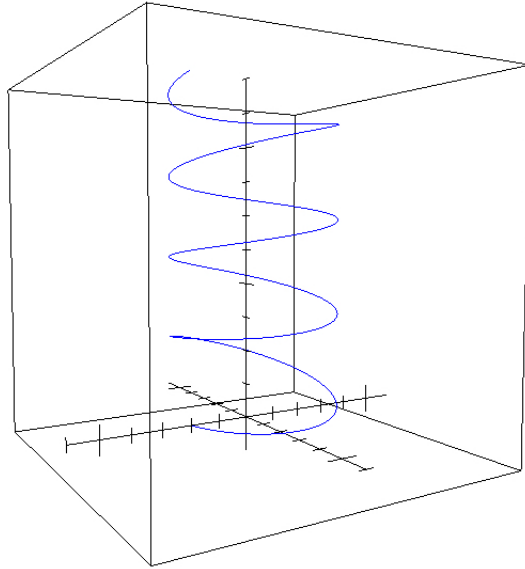
$$K^2 = \frac{\|g' \wedge g''\|^2}{\|g'\|^6},$$

on obtient la formule générale. \square

Un exemple : l'hélice elliptique. On considère l'arc paramétré $\gamma : \mathbb{R} \rightarrow \mathbb{R}^3$ avec

$$\gamma(t) = (a \cos(t), b \sin(t), kt),$$

où $a, b, k > 0$.



Hélice elliptique dans \mathbb{R}^3 .

En dérivant directement, on a

$$\|\gamma'(t) \wedge \gamma''(t)\| = \sqrt{(ab)^2 + k^2(a^2 \cos^2(t) + b^2 \sin^2(t))},$$

ce qui donne

$$K(t) = \frac{\sqrt{(ab)^2 + k^2(a^2 \cos^2(t) + b^2 \sin^2(t))}}{(a^2 \sin^2(t) + b^2 \cos^2(t) + k^2)^{3/2}}.$$

On a aussi

$$T(t) = -\left((ab)^2 + k^2(a^2 \cos^2(t) + b^2 \sin^2(t))\right)^{-1} \begin{vmatrix} -a \sin(t) & -a \cos(t) & a \sin(t) \\ b \cos(t) & -b \sin(t) & -b \cos(t) \\ k & 0 & 0 \end{vmatrix},$$

en développant le déterminant par rapport à la dernière ligne on tombe sur

$$T(t) = -\frac{kab}{(ab)^2 + k^2(a^2 \cos^2(t) + b^2 \sin^2(t))}.$$

Dans le cas particulier des hélices circulaires où $a = b$, on a

$$K(t) = \frac{a}{a^2 + k^2}, \quad T(t) = -\frac{k}{a^2 + k^2},$$

c'est une courbe gauche à courbure et torsion constante.

On peut montrer, comme pour les courbes planes, en utilisant la théorie des équations différentielles ordinaires (linéaires), que courbure et torsion déterminent *complètement les courbes gauches* à isométrie affine de \mathbb{R}^3 près. Voir programme de L3.

5.2 Introduction aux surfaces

5.2.1 Rudiments de calcul différentiel en deux variables.

Si $p \in \mathbb{R}^2$, et $r > 0$, on notera

$$B(p, r) = \{m \in \mathbb{R}^2 : \|m - p\| < r\},$$

appelée boule ouverte centrée en p et de rayon r . Un sous ensemble non-vidé $U \subset \mathbb{R}^2$ est dit **ouvert** ssi il est un voisinage de tout ses points : pour tout $p \in U$, il existe $\epsilon > 0$ tel que $B(p, \epsilon) \subset U$.

Définition 5.2.1 Soit U un ouvert de \mathbb{R}^2 et $f : U \rightarrow \mathbb{R}$ une fonction. Elle est dite continue ssi pour tout $p_0 \in U$, pour tout $\epsilon > 0$, il existe $\eta > 0$ tel que pour tout $x \in U$ tel que $\|x - p_0\| \leq \eta$, $|f(x) - f(p_0)| \leq \epsilon$. Autrement dit, se fixant $\epsilon > 0$, il existe $\eta > 0$ tel que $f(B(p, \eta)) \subset B(f(p), \epsilon)$.

La notion de continuité pour les fonctions de plusieurs variables est plus subtile qu'en dimension 1. Il n'est pas vrai que si $x \mapsto f(x, y)$ et $y \mapsto f(x, y)$ sont continues alors f l'est !

Définition 5.2.2 (Classe C^1). Soit U un ouvert de \mathbb{R}^2 , et $f : U \rightarrow \mathbb{R}$ une fonction. On dit que f admet en $p = (x_0, y_0) \in U$ des dérivées partielles ssi

$$\frac{\partial f}{\partial x}(p) := \lim_{h \rightarrow 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \text{ et } \frac{\partial f}{\partial y}(p) := \lim_{h \rightarrow 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h}$$

existent. Si les dérivées partielles de f existent en tout point de U et sont continues, on dit que f est C^1 . On peut montrer qu'elle est alors automatiquement continue.

Si f est C^1 , on dit que f admet des dérivées partielles secondes ssi les limites suivantes existent :

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2}(p) &:= \lim_{h \rightarrow 0} \frac{\frac{\partial f}{\partial x}(x_0 + h, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)}{h}, \quad \frac{\partial^2 f}{\partial y^2}(p) := \lim_{h \rightarrow 0} \frac{\frac{\partial f}{\partial y}(x_0, y_0 + h) - \frac{\partial f}{\partial y}(x_0, y_0)}{h}, \\ \frac{\partial^2 f}{\partial x \partial y}(p) &:= \lim_{h \rightarrow 0} \frac{\frac{\partial f}{\partial y}(x_0 + h, y_0) - \frac{\partial f}{\partial y}(x_0, y_0)}{h}, \quad \frac{\partial^2 f}{\partial y \partial x}(p) := \lim_{h \rightarrow 0} \frac{\frac{\partial f}{\partial x}(x_0, y_0 + h) - \frac{\partial f}{\partial x}(x_0, y_0)}{h}. \end{aligned}$$

Définition 5.2.3 (Classe C^2). Soit U un ouvert de \mathbb{R}^2 , et $f : U \rightarrow \mathbb{R}$ une fonction de classe C^1 . On dit que f est C^2 si les 4 dérivées partielles

$$\frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial y \partial x}, \frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial y^2}$$

existent en tout point de U et sont continues.

Proposition 5.2.4 (Lemme de Schwarz) Si $f : U \rightarrow \mathbb{R}$ est de classe C^2 alors on en tout point $p \in U$,

$$\frac{\partial^2 f}{\partial x \partial y}(p) = \frac{\partial^2 f}{\partial y \partial x}(p).$$

Preuve. Admis, preuve assez technique, voir cours de calcul différentiel. \square

Par récurrence, on peut ainsi définir les dérivées partielles d'ordre supérieur

$$\frac{\partial^k f}{\partial x_1 \partial x_2 \dots \partial x_k} := \frac{\partial}{\partial x_1} \left(\frac{\partial^{k-1} f}{\partial x_2 \dots \partial x_k} \right),$$

où $x_j \in \{x, y\}$.

Définition 5.2.5 (Classe C^k). Une fonction $f : U \rightarrow \mathbb{R}$ est dite de classe C^k si toutes les dérivées partielles existent et sont continues. On peut montrer que l'ordre de dérivation ne compte pas pour les dérivées croisées.

Si $f = (f_1, \dots, f_n) : U \rightarrow \mathbb{R}^n$ est une fonction à valeurs vectorielles, on dit qu'elle est de classe C^k ssi toutes les fonctions coordonnées f_1, \dots, f_n sont de classe C^k .

Définition 5.2.6 Soit $U \subset \mathbb{R}^2$ un ouvert, et $f = (f_1, f_2) : U \rightarrow \mathbb{R}^2$ une application de classe C^1 . On appelle **matrice jacobienne** au point $p \in U$ la matrice

$$D_p f := \begin{pmatrix} \frac{\partial f_1}{\partial x}(p) & \frac{\partial f_1}{\partial y}(p) \\ \frac{\partial f_2}{\partial x}(p) & \frac{\partial f_2}{\partial y}(p) \end{pmatrix}.$$

L'application linéaire associée est aussi notée (par abus d'écriture)

$$D_p f : \mathbb{R}^2 \rightarrow \mathbb{R}^2,$$

elle est appelée "différentielle" de f au point p .

Si une fonction est de classe C^1 , la différentielle donne une approximation linéaire d'une fonction au premier ordre, autrement dit un développement limité à l'ordre 1 d'une fonction de deux variables.

Proposition 5.2.7 Soit $f : U \rightarrow \mathbb{R}^2$ de classe C^1 , et $p \in U$. Alors $D_p f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ est l'unique application linéaire telle que l'on ait pour tout $h \in \mathbb{R}^2$ avec $\|h\|$ assez petit,

$$f(p+h) = f(p) + D_p f(h) + o(\|h\|).$$

Preuve. Ca n'est pas évident, en particulier il faut utiliser la continuité des dérivées partielles, voir cours de calcul diff. \square

Proposition 5.2.8 Soient $f : U \rightarrow \mathbb{R}^2$ et $g : V \rightarrow \mathbb{R}^2$ deux applications de classe C^1 telles que $f(U) \subset V$. Alors $g \circ f : U \rightarrow \mathbb{R}^2$ est aussi C^1 , et on a la relation matricielle

$$D_p(g \circ f) = (D_{f(p)}g)(D_p f),$$

ce qui composante par composante s'écrit pour $i, j \in \{1, 2\}$, (règle de la chaîne) :

$$\frac{\partial (g \circ f)_i}{\partial x_j} = \frac{\partial g_i}{\partial x}(f(p)) \frac{\partial f_1}{\partial x_j}(p) + \frac{\partial g_i}{\partial y}(f(p)) \frac{\partial f_2}{\partial x_j}(p).$$

Preuve. Ça découle assez directement de la proposition précédente en composant les développements limités. \square

Définition 5.2.9 Soit U, V deux ouverts de \mathbb{R}^2 . Une bijection $f : U \rightarrow V$ de classe C^k est dit être un C^k -difféomorphisme ssi $f^{-1} : V \rightarrow U$ est aussi de classe C^k .

Remarque. Si $f : U \rightarrow V$ est un C^1 -difféomorphisme, alors comme

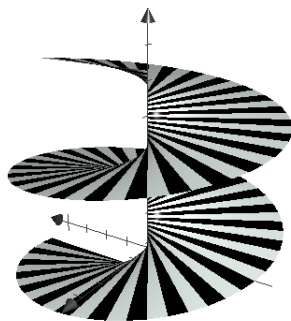
$$f^{-1} \circ f = Id_U,$$

la proposition précédente montre que $D_p f$ est un isomorphisme linéaire et que $D_{f(p)} f^{-1} = (D_p f)^{-1}$.

5.2.2 Surfaces paramétrées, exemples

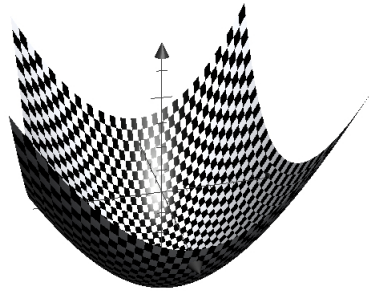
Définition 5.2.10 Une surface paramétrée $S = (U, f)$ de classe C^k dans \mathbb{R}^3 est la donnée d'un ouvert $U \subset \mathbb{R}^2$ et d'une fonction de classe C^k , $f : U \rightarrow \mathbb{R}^3$.

Exemples. On considère $U =]0, 1[\times]0, 4\pi[$ et $f(r, t) = (r \cos(t), r \sin(t), t)$, c'est une surface appelée hélicoïde.



Hélicoïde dans \mathbb{R}^3 .

On prend $U =]-1, +1[\times]-1, +1[$ et $g(x, y) = (x, y, x^2 + y^2)$.



Paraboloïde dans \mathbb{R}^3 .

C'est un cas particulier de surface où z est fonction de (x, y) : c'est le graphe (dans \mathbb{R}^3) de la fonction $z = f(x, y) = x^2 + y^2$.

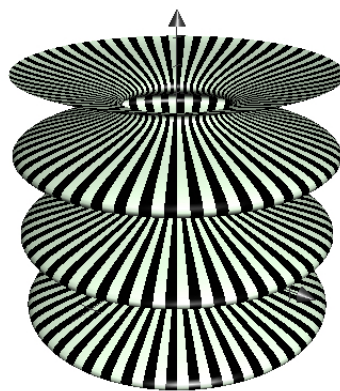
Considérons maintenant $U =]0, 2\pi[\times]0, 10[$ et

$$R(u, v) = ((2 + \sin(2v)) \cos(u), (2 + \sin(2v)) \sin(u), v),$$

où $u \in]0, 2\pi[$ et $v \in]0, 10[$. C'est une surface de *révolution* obtenue en faisant tourner le graphe de

$$f(x) = 2 + \sin(x)$$

autour de l'axe Oz .

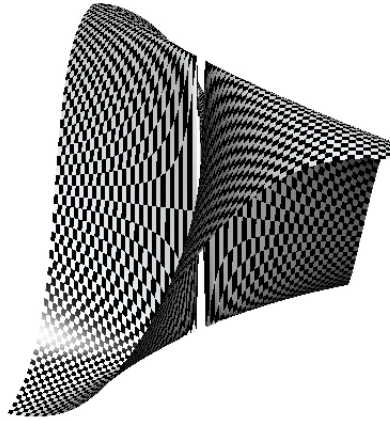


Surface de révolution dans \mathbb{R}^3 .

On termine cette série d'exemples par le conoïde de Plucker : c'est la surface d'équation

$$z = \frac{xy}{x^2 + y^2},$$

avec $(x, y) \in U = \mathbb{R}^2 \setminus \{(0, 0)\}$.



Conoïde de Plucker.

Définition 5.2.11 Deux surfaces (U, f) et (V, g) de classe C^k sont dites C^k -équivalentes ssi il existe un C^k -difféomorphisme $\theta : U \rightarrow V$ tel que $f = g \circ \theta$. On appelle θ un changement de carte.

Définition 5.2.12 Une surface $S = (U, f)$, de classe C^1 , est dite **régulière** en $m = f(p)$ ssi les deux dérivées partielles $\frac{\partial f}{\partial x}(p)$, $\frac{\partial f}{\partial y}(p)$ sont linéairement indépendantes. Si $m = f(p)$ est un point régulier, l'espace tangent à $S = (U, f)$ en m est le sous-espace vectoriel de \mathbb{R}^3 (un plan)

$$T_m S = \mathbb{R} \frac{\partial f}{\partial x}(p) + \mathbb{R} \frac{\partial f}{\partial y}(p) = \text{Vect} \left\{ \frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p) \right\} = \text{IM}(D_p f).$$

Remarques. L'espace *affine* tangent à la surface est $\mathcal{T}_m S := f(p) + T_m S$. Tout élément de $T_m S$ est appelé vecteur tangent. Les notions de point régulier et de plan tangent sont invariants par C^1 -équivalence des surfaces (ou par changement de carte). En effet, si $f = g \circ \theta$ avec θ un C^1 -difféomorphisme, alors on a

$$D_p f = (D_{\theta(p)} g) \circ (D_p \theta),$$

et comme $D_p \theta$ est un isomorphisme de $\mathbb{R}^2 \rightarrow \mathbb{R}^2$, on a bien

$$T_{f(p)} S = \text{IM}(D_p f) = \text{IM}(D_{\theta(p)} g).$$

5.2.3 La première forme fondamentale

Soit $S = (U, f)$ une surface et $m = f(p)$ un point régulier. La première forme fondamentale I_m est un produit scalaire canonique sur l'espace tangent $T_m S$ qui est défini comme la restriction du produit scalaire usuel $\langle \cdot, \cdot \rangle$ de \mathbb{R}^3 à $T_m S$.

Définition 5.2.13 *Pour tout vecteurs tangents $X = x_1 \frac{\partial f}{\partial x}(p) + x_2 \frac{\partial f}{\partial y}(p) \in T_m S$ et $Y = y_1 \frac{\partial f}{\partial x}(p) + y_2 \frac{\partial f}{\partial y}(p) \in T_m S$, on pose*

$$I_m(X, Y) = (x_1 \ x_2) \begin{pmatrix} \left\| \frac{\partial f}{\partial x}(p) \right\|^2 & \left\langle \frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p) \right\rangle \\ \left\langle \frac{\partial f}{\partial y}(p), \frac{\partial f}{\partial x}(p) \right\rangle & \left\| \frac{\partial f}{\partial y}(p) \right\|^2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Il s'agit bien d'un produit scalaire sur $T_m S$, puisque c'est la restriction du produit scalaire de \mathbb{R}^3 à $T_m S$. La première forme fondamentale encode les propriétés métriques de la surface. Elle ne dépend pas du changement de carte, mais son expression dans la base $\mathcal{B} = (\frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p))$ en dépend !

Longueur des courbes tracées sur S . Soit $S = (U, f)$ une surface régulière. Soit $I \subset \mathbb{R}$ un intervalle fermé borné et $\gamma = (\gamma_1, \gamma_2) : I \rightarrow U$ un arc C^1 , alors

$$f \circ \gamma : I \rightarrow \mathbb{R}^3$$

est une courbe gauche tracée sur S . On a de plus

$$f \circ \gamma'(t) = \gamma_1'(t) \frac{\partial f}{\partial x}(p) + \gamma_2'(t) \frac{\partial f}{\partial y}(p).$$

Sa longueur est donc

$$\begin{aligned} \ell(f \circ \gamma, I) &= \int_I \|(f \circ \gamma)'(t)\| dt = \int_I \|D_{\gamma(t)} f(\gamma'(t))\| dt \\ &= \int_I \sqrt{I_{f(\gamma(t))}((f \circ \gamma)'(t), (f \circ \gamma)'(t))} dt. \end{aligned}$$

On note traditionnellement

$$E = \left\| \frac{\partial f}{\partial x}(p) \right\|^2, \quad G = \left\| \frac{\partial f}{\partial y}(p) \right\|^2, \quad F = \left\langle \frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p) \right\rangle,$$

de sorte que

$$\text{Mat}_{\mathcal{B}}(I_m) = \begin{pmatrix} E & F \\ F & G \end{pmatrix},$$

$$I_m(X, X) = Ex_1^2 + 2Fx_1x_2 + Gx_2^2,$$

où $\mathcal{B} = (\frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p))$, et $X = x_1 \frac{\partial f}{\partial x}(p) + x_2 \frac{\partial f}{\partial y}(p)$. L'expression de la longueur est maintenant

$$\ell(f \circ \gamma, I) = \int_I \sqrt{E(\gamma'_1)^2 + 2F\gamma'_1\gamma'_2 + G(\gamma'_2)^2} dt.$$

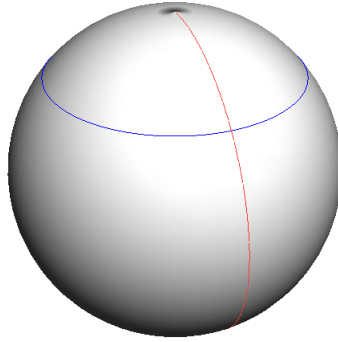
Exemple : courbes tracées sur la sphère. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ avec

$$f(\theta, \varphi) = (\cos(\theta) \sin(\varphi), \sin(\theta) \sin(\varphi), \cos(\varphi)).$$

C'est une paramétrisation de la sphère. Ce paramétrage est régulier pour

$$\varphi \neq 0 \bmod \pi.$$

On va calculer la longueur d'un *parallèle* correspondant à la courbe $\gamma(t) = (t, \varphi)$, où $\varphi \in]0, +\pi[$ est fixé ($\varphi = 0$ et $\varphi = \pi$ correspondent aux poles), et $t \in [0, 2\pi]$.



Parallèle (bleu) et méridien (rouge) sur la sphère.

On a par calcul direct

$$\frac{\partial f}{\partial \theta} = \begin{pmatrix} -\sin(\theta) \sin(\varphi) \\ \cos(\theta) \sin(\varphi) \\ 0 \end{pmatrix}, \quad \frac{\partial f}{\partial \varphi} = \begin{pmatrix} \cos(\theta) \cos(\varphi) \\ \sin(\theta) \cos(\varphi) \\ -\sin(\varphi) \end{pmatrix}.$$

Donc $E = \sin^2(\varphi)$, $G = 1$ et $F = 0$. La première forme fondamentale s'exprime donc dans ce système de coordonnées comme

$$I_m(X, X) = \sin^2(\varphi)x_1^2 + x_2^2.$$

La longueur cherchée est donc

$$\int_0^{2\pi} \sqrt{E(\gamma'_1)^2 + 2F\gamma'_1\gamma'_2 + G(\gamma'_2)^2} dt = \int_0^{2\pi} \sqrt{\sin^2(\varphi)} dt = 2\pi \sin(\varphi).$$

Que dire dans le cas d'un *méridien*? Dans ce cas la courbe est $\gamma(t) = (\theta, t)$ avec θ fixé et $t \in [0, \pi]$. La longueur est alors

$$\int_0^\pi \sqrt{E(\gamma'_1)^2 + 2F\gamma'_1\gamma'_2 + G(\gamma'_2)^2} dt = \int_0^\pi dt = \pi.$$

5.2.4 Deuxième forme fondamentale, courbure

Soit $S = (U, f)$ une surface régulière C^2 . Au point $m = f(p)$, on peut définir un vecteur unitaire normal directement orthogonal au plan tangent $T_m S$ en posant

$$N(p) := \frac{\frac{\partial f}{\partial x}(p) \wedge \frac{\partial f}{\partial y}(p)}{\|\frac{\partial f}{\partial x}(p) \wedge \frac{\partial f}{\partial y}(p)\|}.$$

On définit une application C^1 , appelée *application de Gauss*,

$$N : U \rightarrow S^2,$$

où $S^2 := \{x \in \mathbb{R}^3 : \|x\| = 1\}$ est la sphère unité de \mathbb{R}^3 . On remarquera que l'on pourrait aussi choisir $-N$ à la place de N : on fait ainsi un choix d'orientation de la surface. Par analogie avec le cas des courbes où on a vu que la courbure pouvait s'interpréter comme variation de l'angle de la normale (ou de la tangente), on aimerait définir une notion de courbure des surfaces qui mesure la variation de la normale unitaire au plan tangent quand le point p bouge. C'est le sens de ce paragraphe.

Proposition 5.2.14 *On a pour tout $p \in U$, $\frac{\partial N}{\partial x}(p), \frac{\partial N}{\partial y}(p) \in T_m S$. Autrement dit,*

$$\text{IM}(D_p N) \subset T_m S.$$

Preuve. Comme on a pour tout $p \in U$, $\|N(p)\|^2 = 1$, en calculant les deux dérivées partielles on a

$$\frac{\partial}{\partial x} (\|N(p)\|^2) = 2 \left\langle \frac{\partial N}{\partial x}(p), N(p) \right\rangle = 0$$

et

$$\frac{\partial}{\partial y} (\|N(p)\|^2) = 2 \left\langle \frac{\partial N}{\partial y}(p), N(p) \right\rangle = 0,$$

Ainsi $\text{IM}(D_p N) = \text{Vect} \left\{ \frac{\partial N}{\partial x}(p), \frac{\partial N}{\partial y}(p) \right\} \subset N(p)^\perp = T_m S$, c'est fini. \square

Définition 5.2.15 Soit $S = (U, f)$ une surface C^2 , régulière. Soit $m = f(p)$ avec $p \in U$. On définit l'endomorphisme de Weingarten $W_m : T_m S \rightarrow T_m S$ par

$$W_m(X) = -D_p N((D_p f)^{-1} X).$$

Cette définition a bien un sens : en un point régulier, $D_p f : \mathbb{R}^2 \rightarrow T_m S$ est un isomorphisme, on peut donc prendre son inverse $(D_p f)^{-1} : T_m S \rightarrow \mathbb{R}^2$, que l'on compose avec $-D_p N : \mathbb{R}^2 \rightarrow T_m S$. On peut montrer (exercice) que l'endomorphisme de Weingarten ne dépend pas du choix du paramétrage. Par contre il dépend du choix d'orientation que l'on a fait avec l'application de Gauss : il est changé alors en son opposé. On a par définition

$$W_m\left(\frac{\partial f}{\partial x}(p)\right) = -\frac{\partial N}{\partial x}(p), \quad W_m\left(\frac{\partial f}{\partial y}(p)\right) = -\frac{\partial N}{\partial y}(p).$$

Définition 5.2.16 La deuxième forme fondamentale, notée $\Pi_m(X, Y)$ est une forme bilinéaire sur $T_m S$ définie par

$$\Pi_m(X, Y) := I_m(W_m(X), Y) = \langle W_m(X), Y \rangle.$$

Théorème 5.2.17 L'endomorphisme de Weingarten W_m est auto-adjoint pour le produit scalaire I_m i.e. Π_m est bilinéaire symétrique.

Preuve. On va tout simplement calculer la matrice $\text{Mat}_{\mathcal{B}}(\Pi_m)$ et montrer qu'elle est symétrique, ce sera fini. On a par définition

$$\text{Mat}_{\mathcal{B}}(\Pi_m) = \begin{pmatrix} \Pi_m\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial x}\right) & \Pi_m\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \\ \Pi_m\left(\frac{\partial f}{\partial y}, \frac{\partial f}{\partial x}\right) & \Pi_m\left(\frac{\partial f}{\partial y}, \frac{\partial f}{\partial y}\right) \end{pmatrix}.$$

On a en outre

$$\Pi_m\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = -\left\langle \frac{\partial N}{\partial x}, \frac{\partial f}{\partial y} \right\rangle, \quad \Pi_m\left(\frac{\partial f}{\partial y}, \frac{\partial f}{\partial x}\right) = -\left\langle \frac{\partial N}{\partial y}, \frac{\partial f}{\partial x} \right\rangle.$$

Comme on a pour tout $p \in U$, $\langle N(p), \frac{\partial f}{\partial y}(p) \rangle = 0$, en dérivant par rapport la première variable on a donc

$$\frac{\partial}{\partial x} \left(\left\langle N, \frac{\partial f}{\partial y} \right\rangle \right) = 0 = \left\langle \frac{\partial N}{\partial x}, \frac{\partial f}{\partial y} \right\rangle + \left\langle N, \frac{\partial^2 f}{\partial x \partial y} \right\rangle,$$

donc

$$-\left\langle \frac{\partial N}{\partial x}, \frac{\partial f}{\partial y} \right\rangle = \left\langle N, \frac{\partial^2 f}{\partial x \partial y} \right\rangle.$$

En dérivant par rapport à la deuxième variable l'identité $\langle N(p), \frac{\partial f}{\partial y}(p) \rangle = 0$, on obtient de même :

$$-\langle \frac{\partial N}{\partial y}, \frac{\partial f}{\partial x} \rangle = \langle N, \frac{\partial^2 f}{\partial y \partial x} \rangle.$$

Le lemme de Schwarz sur l'égalité des dérivées croisées montre donc que

$$\Pi_m(\frac{\partial f}{\partial y}, \frac{\partial f}{\partial x}) = \Pi_m(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}) = \langle N, \frac{\partial^2 f}{\partial y \partial x} \rangle,$$

et la matrice de Π_m est bien symétrique. Un calcul analogue montre que l'on a aussi

$$\Pi_m(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial x}) = \langle N, \frac{\partial^2 f}{\partial x^2} \rangle, \quad \Pi_m(\frac{\partial f}{\partial y}, \frac{\partial f}{\partial y}) = \langle N, \frac{\partial^2 f}{\partial y^2} \rangle,$$

et la preuve est finie. \square

La matrice de la deuxième forme fondamentale est souvent notée

$$\text{Mat}_{\mathcal{B}}(\Pi_m) = \begin{pmatrix} L & M \\ M & P \end{pmatrix},$$

avec

$$L = \langle N, \frac{\partial^2 f}{\partial x^2} \rangle, \quad M = \langle N, \frac{\partial^2 f}{\partial x \partial y} \rangle, \quad P = \langle N, \frac{\partial^2 f}{\partial y^2} \rangle.$$

Définition 5.2.18 *L'endomorphisme de Weingarten est donc diagonalisable dans une base I_m -orthonormée.*

1. Les valeurs propres λ_1 et λ_2 sont appelées **courbures principales** de S au point m .
2. Le produit $K_m := \lambda_1 \lambda_2 = \det(W_m)$ est appelé **courbure de Gauss** au point m .
3. La demi-somme $\frac{1}{2}(\lambda_1 + \lambda_2) = \frac{1}{2}\text{Tr}(W_m)$ est appelée **courbure moyenne** en m .

Le signe des valeurs propres dépend du choix d'orientation de S . Par contre la courbure de Gauss n'en dépend pas. Il est important d'avoir en tête que la matrice de W_m dans la base $\mathcal{B} = (\frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p))$ n'est pas en général symétrique car la base \mathcal{B} n'est pas toujours I_m -orthogonale ! Pour calculer la courbure de Gauss, on utilise le fait suivant, bien pratique pour faire les calculs.

Proposition 5.2.19 On a $\text{Mat}_{\mathcal{B}}(\Pi_m) = \text{Mat}_{\mathcal{B}}(I_m)\text{Mat}_{\mathcal{B}}(W_m)$, ainsi

$$K_m = \frac{\det(\text{Mat}_{\mathcal{B}}(\Pi_m))}{\det(\text{Mat}_{\mathcal{B}}(I_m))} = \frac{LP - M^2}{EG - F^2} = \frac{\langle N, \frac{\partial^2 f}{\partial x^2} \rangle \langle N, \frac{\partial^2 f}{\partial y^2} \rangle - (\langle N, \frac{\partial^2 f}{\partial x \partial y} \rangle)^2}{\|\frac{\partial f}{\partial x}\|^2 \|\frac{\partial f}{\partial y}\|^2 - (\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \rangle)^2}.$$

Preuve. Si $X, Y \in T_m S$, notés comme vecteurs colonnes dans la base \mathcal{B} , on a

$$\Pi_m(X, Y) = {}^t X \text{Mat}_{\mathcal{B}}(\Pi_m) Y = I_m(X, W_m(Y)) = {}^t X \text{Mat}_{\mathcal{B}}(I_m) \text{Mat}_{\mathcal{B}}(W_m) Y.$$

La preuve est finie par identification des matrices. \square

En remplaçant N par son expression en termes de produit vectoriel, on peut donner une formule pour K_m ne faisant intervenir que f et ses dérivées, qui fait apparaître des produits mixtes (déterminants). Voir TD.

Exemple 1 : la selle de cheval. Soit $U = \mathbb{R}^2$ et $f(x, y) = (x, y, x^2 - y^2)$. C'est une surface en forme de selle de cheval, on va calculer la courbure en $(0, 0)$. On a

$$\frac{\partial f}{\partial x}(x, y) = (1, 0, 2x), \quad \frac{\partial f}{\partial y}(x, y) = (0, 1, -2y),$$

puis

$$\frac{\partial f}{\partial x}(x, y) \wedge \frac{\partial f}{\partial y}(x, y) = (-2x, 2y, 1),$$

de sorte que

$$N(x, y) = \frac{(-2x, 2y, 1)}{\sqrt{1 + 4x^2 + 4y^2}}.$$

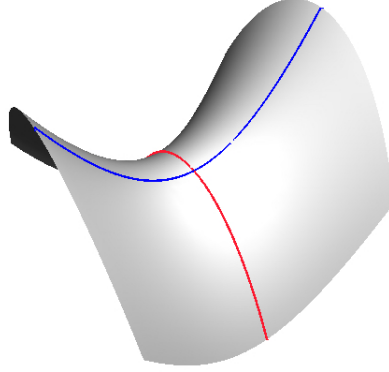
Ainsi

$$\frac{\partial N}{\partial x}(0, 0) = (-2, 0, 0) \quad \frac{\partial N}{\partial y}(0, 0) = (0, 2, 0).$$

Donc on a

$$W_{(0,0)}\left(\frac{\partial f}{\partial x}(0, 0)\right) = (2, 0, 0) = 2 \frac{\partial f}{\partial x}(0, 0), \quad W_{(0,0)}\left(\frac{\partial f}{\partial y}(0, 0)\right) = (0, -2, 0) = -2 \frac{\partial f}{\partial y}(0, 0).$$

L'endomorphisme de Weingarten $W_{(0,0)}$ est diagonal dans la base \mathcal{B} , les valeurs propres sont 2 et -2 , la courbure de Gauss vaut $K = -4$.



Selle de cheval et courbures principales en $(0, 0, 0)$.

Exemple 2 : la sphère. On revient sur le paramétrage de la sphère de rayon R , donné par $f(\theta, \varphi) = R(\cos(\theta)\sin(\varphi), \sin(\theta)\sin(\varphi), \cos(\varphi))$ avec $(\theta, \varphi) \in \mathbb{R} \times]0, \pi[$. La géométrie de la sphère fait que l'on peut prendre

$$N(\theta, \varphi) = (\cos(\theta)\sin(\varphi), \sin(\theta)\sin(\varphi), \cos(\varphi)).$$

On a fait un choix d'orientation qui simplifie les calculs. On a aussi

$$\frac{\partial f}{\partial \theta} = R \begin{pmatrix} -\sin(\theta)\sin(\varphi) \\ \cos(\theta)\sin(\varphi) \\ 0 \end{pmatrix}, \quad \frac{\partial f}{\partial \varphi} = R \begin{pmatrix} \cos(\theta)\cos(\varphi) \\ \sin(\theta)\cos(\varphi) \\ -\sin(\varphi) \end{pmatrix}.$$

Puis

$$\begin{aligned} \frac{\partial^2 f}{\partial \theta^2} &= R \begin{pmatrix} -\cos(\theta)\sin(\varphi) \\ -\sin(\theta)\sin(\varphi) \\ 0 \end{pmatrix}, \quad \frac{\partial^2 f}{\partial \varphi^2} = R \begin{pmatrix} -\cos(\theta)\cos(\varphi) \\ -\sin(\theta)\cos(\varphi) \\ -\cos(\varphi) \end{pmatrix}, \\ \frac{\partial^2 f}{\partial \varphi \partial \theta} &= R \begin{pmatrix} -\sin(\theta)\cos(\varphi) \\ \cos(\theta)\cos(\varphi) \\ 0 \end{pmatrix}. \end{aligned}$$

Ainsi on a

$$\langle \frac{\partial^2 f}{\partial \theta^2}, N \rangle = -R\sin^2(\varphi), \quad \langle \frac{\partial^2 f}{\partial \varphi^2}, N \rangle = -R, \quad \langle \frac{\partial^2 f}{\partial \theta \partial \varphi}, N \rangle = 0.$$

Finalement,

$$K(\theta, \varphi) = \frac{\langle N, \frac{\partial^2 f}{\partial \theta^2} \rangle \langle N, \frac{\partial^2 f}{\partial \varphi^2} \rangle - (\langle N, \frac{\partial^2 f}{\partial \theta \partial \varphi} \rangle)^2}{\|\frac{\partial f}{\partial \theta}\|^2 \|\frac{\partial f}{\partial \varphi}\|^2 - (\langle \frac{\partial f}{\partial \theta}, \frac{\partial f}{\partial \varphi} \rangle)^2} = \frac{1}{R^2}.$$

On peut voir qu'avec ce choix d'orientation, les deux courbures principales sont égales à $\frac{-1}{K}$, l'endomorphisme de Weingarten est diagonal dans la base naturelle.

Exemple 3 : la pseudo-sphère. On considère $U = \mathbb{R} \times \mathbb{R}^*$ et

$$f(\theta, t) = \left(\frac{\cos(\theta)}{\cosh(t)}, \frac{\sin(\theta)}{\cosh(t)}, t - \tanh(t) \right).$$

On peut vérifier par calcul que $S = (U, f)$ est une surface de révolution, partout régulière (en $t = 0$, il y a un cercle singulier), et que la courbure de Gauss de S est égale en tout point à -1 , voir TD.



La pseudo-sphère.

Remarques. Il est facile de voir que si l'application de Weingarten est partout nulle, alors l'application de Gauss est constante. En intégrant on obtient que la surface est incluse dans un plan affine de \mathbb{R}^3 , elle est plate. En revanche si la courbure de Gauss K_m est partout nulle, rien ne dit que la surface doit être plate. Considérons par exemple le cylindre (\mathcal{U}, f) avec $\mathcal{U} = \mathbb{R}^2$ et

$$f(u, v) = (\cos(u), \sin(u), v).$$

Un calcul simple montre que

$$\frac{\partial f}{\partial u} = \begin{pmatrix} -\sin(u) \\ \cos(u) \\ 0 \end{pmatrix}, \quad \frac{\partial f}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad N(u, v) = \begin{pmatrix} \cos(u) \\ \sin(u) \\ 0 \end{pmatrix}.$$

En dérivant une nouvelle fois, on a

$$W_m\left(\frac{\partial f}{\partial u}\right) = -\frac{\partial N}{\partial u} = -\frac{\partial f}{\partial u}, \quad W_m\left(\frac{\partial f}{\partial v}\right) = -\frac{\partial N}{\partial v} = 0,$$

ainsi on a

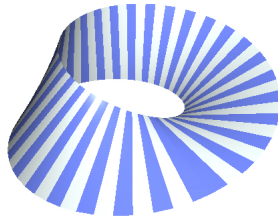
$$\text{Mat}_{\mathcal{B}}(W_m) = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \quad K_m = 0.$$

Le cylindre est à courbure de Gauss nulle !



Le cylindre.

Les surfaces à courbure de Gauss nulle sont loin d'être plates, on dit qu'elles sont développables : on peut les construire en courbant une bande de papier dans l'espace. C'est le cas par exemple du ruban de Moebius.



Ruban de Moebius.

Pour conclure. Notre définition de la courbure des surfaces passe par l'application de Gauss et la structure euclidienne naturelle de \mathbb{R}^3 , c'est une approche extrinsèque. Gauss a montré qu'en fait la courbure est complètement déterminée par la première forme fondamentale, elle est intrinsèque. C'est le "Theorema Egregium" de Gauss (1827).



Annexe A

Intégration à valeur complexe et relèvement C^1 .

Soit I un intervalle de \mathbb{R} et $f : I \rightarrow \mathbb{C}$ une fonction. On a vu que f est de classe C^k ssi ses fonctions coordonnées $\operatorname{Re}(f)$ et $\operatorname{Im}(f)$ sont elle-mêmes de classe C^k . On a de plus

$$f^{(k)}(t) = (\operatorname{Re}(f))^{(k)}(t) + i(\operatorname{Im}(f))^{(k)}(t).$$

Si $f, g : I \rightarrow \mathbb{C}$ sont dérivables, il est facile de voir (le vérifier) que la formule de dérivation des produits s'étend au cas *complexe* i.e. on a pour tout $t \in I$,

$$(fg)'(t) = f'(t)g(t) + f(t)g'(t).$$

Si $\varphi : I \rightarrow \mathbb{R}$ est une fonction dérivable alors $e^{i\varphi}$ l'est aussi et on a

$$(e^{i\varphi})'(t) = i\varphi'(t)e^{i\varphi(t)}.$$

En effet, on a

$$e^{i\varphi(t)} = \cos(\varphi(t)) + i\sin(\varphi(t)),$$

et donc en dérivant partie réelle et partie imaginaire on a

$$(e^{i\varphi})'(t) = -\sin(\varphi(t))\varphi'(t) + i\cos(\varphi(t))\varphi'(t) = i\varphi'(t)e^{i\varphi(t)}.$$

Si $I = [a, b]$ avec $a < b$ et $f : I \rightarrow \mathbb{C}$ est continue, alors on pose

$$\int_a^b f(t)dt := \int_a^b \operatorname{Re}(f(t))dt + i \int_a^b \operatorname{Im}(f(t))dt.$$

Il est direct de voir que l'intégrale ainsi définie est \mathbb{C} -linéaire : pour tout $f, g : I \rightarrow \mathbb{C}$ continues et $\lambda \in \mathbb{C}$, on a

$$\int_a^b (f(t) + \lambda g(t)) dt = \int_a^b f(t) dt + \lambda \int_a^b g(t) dt.$$

Tout comme pour les fonctions à valeur réelles, l'intégrale $F(x)$ d'une fonction continue f , vue comme fonction de la borne supérieure est dérivable i.e. pour tout $x \in I$,

$$F(x) = \int_a^x f(t) dt,$$

définit une fonction C^1 sur I et on a pour tout $x \in I$, $F'(x) = f(x)$. La preuve découle directement du cas réel appliqué à $\operatorname{Re}(f)$ et $\operatorname{Im}(f)$. On a également la version complexe de l'inégalité triangulaire.

Proposition A.0.1 *Soit $f : I \rightarrow \mathbb{C}$ continue. On a alors*

$$\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt.$$

Preuve. C'est moins évident. On va montrer d'abord le lemme suivant.

Lemme A.0.2 *Pour tout $z_0 \in \mathbb{C}$, il existe une forme \mathbb{R} -linéaire $L_0 : \mathbb{C} \rightarrow \mathbb{R}$ telle que $L_0(z_0) = |z_0|$ et pour tout $z \in \mathbb{C}$,*

$$L_0(z) \leq |z|.$$

Preuve du lemme. On peut bien sur supposer $z_0 \neq 0$, sinon c'est trivial, on prend $L_0 = 0$ qui convient. Posons alors

$$L_0(z) = \operatorname{Re} \left(\frac{\overline{z_0}}{|z_0|} z \right).$$

On a bien $L_0(z_0) = \operatorname{Re}(|z_0|) = |z_0|$, et pour tout $z \in \mathbb{C}$,

$$L_0(z) = \operatorname{Re} \left(\frac{\overline{z_0}}{|z_0|} z \right) \leq \left| \frac{\overline{z_0}}{|z_0|} z \right| = |z|.$$

Le lemme est prouvé. \square

On rappelle (par exemple par le thm de Riesz) que toute forme \mathbb{R} -linéaire $L : \mathbb{C} \rightarrow \mathbb{R}$ s'écrit sous la forme

$$L(z) = \operatorname{Re}(wz),$$

où $w \in \mathbb{C}$. Il est alors facile de voir que pour toute forme \mathbb{R} -linéaire L on a

$$L\left(\int_a^b f(t)dt\right) = \int_a^b L(f(t))dt.$$

En appliquant le lemme avec $z_0 = \int_a^b f(t)dt$ on a donc

$$\left|\int_a^b f(t)dt\right| = L_0\left(\int_a^b f(t)dt\right) = \int_a^b L_0(f(t))dt \leq \int_a^b |f(t)|dt,$$

la proposition est prouvée. \square

On va pouvoir utiliser l'intégration complexe pour démontrer le théorème suivant.

Théorème A.0.3 (*Relèvement C^1*). Soit I un intervalle de \mathbb{R} et $\gamma : I \rightarrow S^1$ une application C^1 . Alors il existe $\varphi : I \rightarrow \mathbb{R}$, de classe C^1 tel que pour tout $t \in I$,

$$\gamma(t) = e^{i\varphi(t)}.$$

On dit que φ est une relevée de γ . Si de plus $\tilde{\varphi}$ a les mêmes propriétés, alors il existe une constante $\alpha \in 2\pi\mathbb{Z}$ tel que pour tout $t \in I$, $\varphi(t) = \tilde{\varphi}(t) + \alpha$.

Preuve. Soit $a \in I$. Quitte à multiplier γ par un nombre complexe de module 1, on peut supposer que $\gamma(a) = 1$. Posons alors pour tout $t \in I$,

$$\varphi(t) := \int_a^t \frac{\gamma'(x)}{i\gamma(x)} dx.$$

Comme γ est C^1 , l'intégrale est bien définie et φ est C^1 de dérivée

$$\varphi'(t) = \frac{\gamma'(t)}{i\gamma(t)}.$$

Comme on a pour tout $t \in I$, $|\gamma(t)|^2 = 1$, en dérivant on trouve

$$\gamma'(t)\overline{\gamma(t)} + \gamma(t)\overline{\gamma'(t)} = 0,$$

et comme $\overline{\gamma(t)} = (\gamma(t))^{-1}$ on en déduit que

$$\frac{\gamma'(t)}{\gamma(t)} \in i\mathbb{R}.$$

Ainsi $\varphi'(t) \in \mathbb{R}$ et la fonction φ définie est bien à valeurs réelles. Considérons maintenant la fonction F définie sur I par

$$F(t) := \gamma(t)e^{-i\varphi(t)}.$$

Alors F est C^1 et on a pour tout $t \in I$,

$$F'(t) = \gamma'(t)e^{-i\varphi(t)} - i\gamma(t)\varphi'(t)e^{-i\varphi(t)} = 0.$$

Ce qui signifie que F est constante, mais comme $F(a) = \gamma(a)e^{-i\varphi(a)} = 1$, on a bien obtenu que pour tout $t \in I$, $F(t) = 1$ i.e.

$$\gamma(t) = e^{i\varphi(t)}.$$

Supposons maintenant qu'on dispose d'une autre relevée de γ notée $\widetilde{\varphi}$. Comme on a pour tout $t \in I$,

$$e^{i\widetilde{\varphi}(t)} = \gamma(t) = e^{i\varphi(t)},$$

on en déduit que $e^{i(\widetilde{\varphi}(t)-\varphi(t))} = 1$, donc pour tout $t \in I$ on a

$$\widetilde{\varphi}(t) - \varphi(t) \in 2\pi\mathbb{Z}.$$

Comme $\widetilde{\varphi} - \varphi$ est continue et I est un intervalle, le théorème des valeurs intermédiaires dit que $\widetilde{\varphi} - \varphi$ doit être constante. \square

Remarques. Le théorème du relèvement est valable si on suppose seulement $\gamma : I \rightarrow S^1$ continue, mais c'est plus dur...Le Lemme A.0.2 est une forme élémentaire du théorème de Hahn-Banach, voir cours de L3 d'analyse fonctionnelle.