Préparation à l'agrégation –

Notes sur l'algèbre bilinéaire

Peter Haïssinsky, Université de Provence

2007–2008

Ces notes sont en cours d'élaboration. Tout commentaire sur le contenu sera apprécié. Pour plus de précisions, on peut se reporter aux ouvrages cités en référence.

Contents

1 Formes bilinéaires et variantes

Soient $\mathbb{K}$ un corps commutatif et E un espace vectoriel sur $\mathbb{K}$. On note E^* le dual de E et, si $\mathcal{B}$ est une base de E, on notera $\mathcal{B}^*$ la base duale associée. On rappelle que si $\mathcal{B}$ et $\mathcal{C}$ sont des bases de E (de dimension finie), P la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$, alors la matrice de passage P^* de $\mathcal{B}^*$ à $\mathcal{C}^*$ vérifie ${}^t\mathrm{P}^* \cdot \mathrm{P} = \mathrm{I}$.

Définition 1.1. — *Une forme bilinéaire est une application $b : \mathrm{E} \times \mathrm{E} \to \mathbb{K}$ telle que, pour tout $y \in \mathrm{X}$, $b_1 : x \mapsto b(x, y)$ est linéaire et pour tout $x \in \mathrm{X}$, $b_2 : y \mapsto b(x, y)$ est linéaire.*

On peut associer à b l'application linéaire

$$\begin{aligned} \bar{b} : \quad & \mathrm{E} \to \mathrm{E}^* \\ & y \mapsto (x \mapsto b(x, y)) \end{aligned}$$

Dimension finie. Si E est de dimension finie et si $\mathcal{B} = \{\mathrm{e}_i\}_{1 \leq i \leq n}$ est une base, on définit

$$\mathrm{Mat}(b, \mathcal{B}) = (b(\mathrm{e}_i, \mathrm{e}_j))_{i,j}.$$

Si X et Y sont les coordonnées de vecteurs x et y dans la base $\mathcal{B}$, alors $b(x, y) = {}^t\mathrm{X} \cdot \mathrm{Mat}(b, \mathcal{B}) \cdot \mathrm{Y}$.

On remarque que $\mathrm{Mat}(b, \mathcal{B}) = \mathrm{Mat}(\bar{b}, \mathcal{B}, \mathcal{B}^*)$.

Si $\mathcal{C}$ est une autre base et P est la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$, alors $\mathrm{Mat}(b, \mathcal{C}) = {}^t\mathrm{P}\,\mathrm{Mat}(b, \mathcal{B})\mathrm{P}$.

Définition 1.2. — *On dit que b est non dégénérée si $\bar{b}$ est injective. On définit le noyau de b par*

$$\mathrm{Ker}\, b = \mathrm{Ker}\, \bar{b} = \{y \in \mathrm{E},\ \forall\, x \in \mathrm{E},\ b(x, y) = 0\}\,.$$

Si b est dégénérée, alors il existe une base telle que la matrice associée ait $\dim \mathrm{Ker}\, b$ colonnes nulles.

Définition 1.3. — *Si b est une forme bilinéaire, on dit que b est symétrique si $b(x, y) = b(y, x)$, antisymétrique si $b(x, y) = -b(y, x)$, et alternée si $b(x, x) = 0$ pour tous x, y.*

Remarque. — Pour tout corps, $\mathbb{K}$, si b est non nulle, alors b alternée implique b antisymétrique. En revanche, b antisymétrique implique b alternée seulement si $car(\mathbb{K}) \neq 2$; si $car(\mathbb{K}) = 2$ alors les notions de symétrie et d'antisymétrie sont équivalentes.

On a la variante suivante : si σ est un automorphisme de $\mathbb{K}$, on dit que b est σ-sesquilinéaire si b est linéaire en la première variable, additive en la seconde et si $b(x, \lambda y) = \sigma(\lambda) b(x, y)$.

L'application $\bar{b}$ devient maintenant σ-semi-linéaire, mais on peut encore parler de noyau, de formes dégénérées ou non, ...

Définition 1.4. — *Une forme σ-sesquilinéaire est dite hermitienne si $\sigma \neq \mathrm{Id}$ et si $b(y, x) = \sigma(b(x, y))$. En particulier, σ est une involution.*

Si $(h_{i,j})$ sont les coefficients de la matrice d'une forme σ-hermitienne dans une base fixée, alors les coefficients vérifient $h_{ji} = \sigma(h_{ij})$ et $h_{ii} = \sigma(h_{ii})$. On rappelle que $\mathbb{K}(\sigma) = \{x \in \mathbb{K},\ \sigma(x) = x\}$ est un sous-corps de $\mathbb{K}$.

Remarque. — Une forme σ-sesquilinéaire alternée est bilinéaire ($\sigma = \mathrm{Id}$) et antisymétrique.

La théorie des formes bilinéaires sert de fondement à la géométrie différentielle : les formes bilinéaires symétriques conduisent à la géométrie riemannienne, les formes alternées à la géométrie symplectique et les formes hermitiennes à la géométrie complexe.

2 Formes quadratiques

Dans cette partie et dans la suite, on suppose $car(\mathbb{K}) \neq 2$.

Définition 2.1. — *Une application $q : \mathrm{E} \to \mathbb{K}$ est une forme quadratique s'il existe une forme bilinéaire symétrique b telle que $q(x) = b(x, x)$. On appelle b la forme polaire associée à q.*

Si q est une forme quadratique, alors sa forme polaire est définie par

$$b(x, y) = \frac{1}{2}(q(x + y) - q(x) - q(y)) = \frac{1}{4}(q(x + y) - q(x - y)) .$$

On transfère les notions associées aux formes bilinéaires (symétriques) aux formes quadratiques. Par exemple, on définit $\mathrm{Mat}(q, \mathcal{B}) = \mathrm{Mat}(b, \mathcal{B})$, $\mathrm{Ker}\, q = \mathrm{Ker}\, b$, ...

Exemples.

- Etant donnée une matrice symétrique $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$, on définit sur $\mathbb{K}^n$ la forme $q(\mathrm{X}) = {}^t\mathrm{XAX}$. La forme polaire est $b(\mathrm{X}, \mathrm{Y}) = {}^t\mathrm{XAY}$.
- Si $q : \mathbb{K}^n \to \mathbb{K}$ est une forme quadratique, alors il existe une matrice symétrique $\mathrm{A} = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ telle que, si
$$\mathrm{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$
alors
$$q(\mathrm{X}) = \sum_{1 \leq i,j \leq n} a_{ij} x_i x_j .$$
Démonstration. — Soit b la forme polaire de q et notons $\mathcal{B} = (\mathrm{e}_1, \ldots, \mathrm{e}_n)$ la base canonique de $\mathbb{K}^n$. On a, pour tous $\mathrm{X}, \mathrm{Y} \in \mathbb{K}^n$,
$$b(\mathrm{X}, \mathrm{Y}) = \sum_{i,j} b(\mathrm{e}_i, \mathrm{e}_j) x_i y_j .$$
On pose $a_{ij} = b(\mathrm{e}_i, \mathrm{e}_j)$. La matrice $\mathrm{A} = (a_{ij})$ vérifie les propriétés voulues. ∎
- Si (X, μ) est un espace mesuré, alors on définit $q : \mathrm{L}^2(\mathrm{X}, \mu; \mathbb{R}) \to \mathbb{R}$ par
$$q(f) = \int f^2 d\mu.$$
La forme polaire est
$$b(f, g) = \int fg d\mu.$$
- Si $\mathrm{L}_1, \ldots, \mathrm{L}_r \in \mathrm{E}^*$ sont r formes indépendantes et si $(\lambda_i) \in (\mathbb{K}^*)^r$, alors la formule
$$q(x) = \sum_{1 \leq j \leq r} \lambda_j (\mathrm{L}_j(x))^2$$
définit une forme quadratique sur E de forme polaire
$$b(x, y) = \sum_{1 \leq j \leq r} \lambda_j \mathrm{L}_j(x) \mathrm{L}_j(y) .$$
- Si $\mathrm{M} \in \mathcal{M}_n(\mathbb{K})$, alors $q(\mathrm{M}) = \mathrm{tr}({}^t\mathrm{MM})$ détermine une forme quadratique (définie positive) de forme polaire $b(\mathrm{M}, \mathrm{N}) = \mathrm{tr}({}^t\mathrm{MN})$

– On définit sur $\mathcal{M}_2(\mathbb{K})$ la forme $q(\mathrm{M}) = \det \mathrm{M}$. On remarque que q est un polynôme quadratique homogène en les coefficients de M. De plus, par le théorème de Cayley-Hamilton, on a

$$\mathrm{M}^2 - (\operatorname{tr} \mathrm{M}) \cdot \mathrm{M} + (\det \mathrm{M}) \cdot \mathrm{I}_2 = 0\,.$$

En prenant la trace, on trouve

$$q(\mathrm{M}) = \frac{(\operatorname{tr} \mathrm{M})^2 - \operatorname{tr} \mathrm{M}^2}{2}\,.$$

Du coup, la forme polaire associée est

$$b(\mathrm{M}, \mathrm{N}) = \frac{(\operatorname{tr} \mathrm{M})(\operatorname{tr} \mathrm{N}) - \operatorname{tr}(\mathrm{MN})}{2}\,.$$

– Soit $\mathbb{K}$ un corps de caractéristique 2. On pose $q(x) = x^2$. Alors q est une forme quadratique ainsi qu'une application linéaire !

Remarque. — On peut comme dans le cas symétrique associer à une forme hermitienne une forme quadratique hermitienne. L'exemple fondamental est, sur $\mathbb{C}^n$, la forme $b(\mathrm{X}, \mathrm{Y}) = \sum x_j \bar{y_j}$.

2.1 Orthogonalité – Isotropie

Définition 2.2. — *On dit que deux vecteurs x et y sont orthogonaux, et l'on note $x \perp y$ si $b(x, y) = 0$. Si $\mathrm{F} \subset \mathrm{E}$ est un sous-ensemble, alors on définit*

$$\mathrm{F}^\perp = \{x \in \mathrm{E},\ \forall y \in \mathrm{F},\ b(x, y) = 0\}\,.$$

C'est toujours un sous-espace vectoriel. Enfin, une base orthogonale $\mathcal{B}$ est une base telle que $b(\mathrm{e}_i, \mathrm{e}_j) = 0$ dès que $i \neq j$.

Théorème 2.1. — *Soient* E *un espace vectoriel de dimension finie et q un forme quadratique. Il existe une base orthogonale $\mathcal{B} = \{\mathrm{e}_1, \dots, \mathrm{e}_n\}$. Dans cette base,* $\mathrm{Mat}(q)$ *est diagonale, et*

$$q(x) = \sum q(\mathrm{e}_i)(\mathrm{e}_i^*(x))^2\,.$$

De plus, le nombre de termes non nuls est le rang de q. En particulier, q est non dégénérée si et seulement si $q(\mathrm{e}_i) \neq 0$ pour tout $i \in \{1, \dots, n\}$.

Complément. Une base orthogonale peut être obtenue par l'algorithme de Gauss.

Démonstration. — Par récurrence. Si $n = 1$, alors il n'y a rien à faire. Supposons que tout espace de dimension n admette une base orthogonale et considérons un espace E de dimension $n + 1$ muni d'une forme quadratique q. Si $q \equiv 0$, alors il n'y a rien à faire non plus. Sinon, il existe $\mathrm{e}_{n+1} \in \mathrm{E}$ tel que $q(\mathrm{e}_{n+1}) \neq 0$. Soit $\mathrm{H} = \mathrm{e}_{n+1}^\perp$. Il s'agit d'un espace vectoriel, et $\mathrm{H} = \operatorname{Ker}\{x \mapsto b(x, \mathrm{e}_{n+1})\}$. Comme $q(\mathrm{e}_{n+1}) \neq 0$, $\dim \mathrm{H} = n$, et l'hypothèse de récurrence nous construit une base orthogonale $(\mathrm{e}_1, \dots, \mathrm{e}_n)$ de H. Vérifions que

$$\mathcal{B} = (\mathrm{e}_1, \dots, \mathrm{e}_n, \mathrm{e}_{n+1})$$

est une base. Soient $\lambda_1, \dots, \lambda_{n+1}$ des scalaires tels que $\sum \lambda_j \mathrm{e}_j = 0$. Alors

$$b\left(\sum \lambda_j \mathrm{e}_j, \mathrm{e}_{n+1}\right) = \sum \lambda_j b(\mathrm{e}_j, \mathrm{e}_{n+1}) = \lambda_{n+1} q(\mathrm{e}_{n+1}) = 0$$

et $\lambda_{n+1} = 0$. Du coup, ces vecteurs sont bien indépendants. ∎

Algorithme de Gauss. On part de

$$q(\mathrm{X}) = \sum_{1 \leq i \leq j \leq n} a_{ij} x_i x_j\,.$$

A chaque étape, l'algorithme fait apparaître un carré et fait disparaître une coordonnée dans le reste de la forme.

– Si $a_{11} \neq 0$, alors on écrit

$$\begin{aligned} q(\mathrm{X}) &= a_{11}\left(x_1^2 + \sum_{j=2}^{n} \frac{a_{1j}}{a_{11}} x_1 x_j\right) + \sum_{2\leq i\leq j\leq n} a_{ij}x_i x_j \\ &= a_{11}\left(x_1 + \frac{1}{2}\sum_{j=2}^{n} \frac{a_{1j}}{a_{11}} x_j\right)^2 - \frac{a_{11}}{4}\left(\sum_{j=2}^{n} \frac{a_{1j}}{a_{11}} x_j\right)^2 + \sum_{2\leq i\leq j\leq n} a_{ij}x_i x_j \end{aligned}$$

– Si, pour tout indice j, $a_{jj} = 0$, alors on considère un terme du type xy. On écrit

$$xy = \frac{(x+y)^2 - (x-y)^2}{4}\,.$$

On pose alors

$$\begin{cases} u = \dfrac{x+y}{2} \\ \\ v = \dfrac{x-y}{2} \end{cases}$$

On remplace x et y dans q par ces formes. On se ramène ainsi au cas précédent.

Au bout du compte, on obtient $k\,(\leq n)$ formes linéaires non nulles (L_j) indépendantes et k scalaires (λ_j) tels que

$$q(\mathrm{X}) = \sum_j \lambda_j (\mathrm{L}_j(\mathrm{X}))^2\,.$$

Si on complète ces formes en une base de E^*, alors elles représentent la base duale d'une base orthogonale de E.

Exemple. On considère sur $\mathbb{R}^3$ la forme

$$q(x,y,z) = xy + yz + xz\,.$$

Dans la base canonique $\mathcal{B}$, on a donc

$$\mathrm{Mat}(q,\mathcal{B}) = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

On pose alors

$$\begin{cases} u = \dfrac{x+y}{2} \\ \\ v = \dfrac{x-y}{2} \end{cases}$$

Il vient

$$q(x,y,z) = u^2 - v^2 + (u-v)z + (u+v)z = u^2 + 2uz - v^2\,;$$

du coup,

$$q(x,y,z) = (u+z)^2 - v^2 - z^2\,.$$

Notons $\mathcal{B}$ la base canonique, et

$$\mathrm{P}^* = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & -1/2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

la matrice de passage de $\mathcal{B}^*$ à une base $\mathcal{C}^*$ base duale d'une base orthogonale $\mathcal{C}$ que l'on veut déterminer. On calcule donc $\mathrm{P} = ({}^t\mathrm{P}^*)^{-1}$ et on obtient

$$\mathrm{P} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

Dans cette base, $q(x,y,z) = x^2 - y^2 - z^2$.

Remarque. — Si q est dégénérée, alors il existe une base orthogonale $\mathcal{B}$ telle $\mathrm{Mat}(q,\mathcal{B})$ soit diagonale avec les $\dim \mathrm{Ker}\, q$ derniers termes nuls. On peut alors restreindre la forme à l'espace engendré par les $n - \dim \mathrm{Ker}\, q$ premiers vecteurs.

Remarque. — On a toujours $\mathrm{F} \subset (\mathrm{F}^\perp)^\perp$. En effet, si $x \in \mathrm{F}$ et si $y \in \mathrm{F}^\perp$, alors $b(x,y) = 0$ donc $x \in (\mathrm{F}^\perp)^\perp$.

Proposition 2.1. — *Supposons q non dégénérée et* E *de dimension finie. Si* F *est un sous-espace de* E *alors on a*

1. $\dim \mathrm{F} + \dim \mathrm{F}^\perp = \dim \mathrm{E}$,
2. $\mathrm{F} = (\mathrm{F}^\perp)^\perp$,
3. $(\mathrm{F}+\mathrm{G})^\perp = \mathrm{F}^\perp \cap \mathrm{G}^\perp$ *et* $(\mathrm{F}\cap\mathrm{G})^\perp = \mathrm{F}^\perp + \mathrm{G}^\perp$.

Avant de donner la démonstration, on rappelle la notion d'orthogonalité entre E et E^*.

Rappel. Si $\mathrm{F} \subset \mathrm{E}$ est un sous-espace vectoriel, on pose

$$\mathrm{F}^\circ = \{\mathrm{L} \in \mathrm{E}^*,\ \mathrm{L}(x) = 0\ \forall\, x \in \mathrm{F}\}\,.$$

On a $\dim \mathrm{F} + \dim \mathrm{F}^\circ = \dim \mathrm{E}$. En effet, si $(\mathrm{e}_1,\dots,\mathrm{e}_k)$ est une base de F, que l'on complète en une base de E, on trouve

$$\mathrm{F}^\circ = \mathrm{Vect}\{\mathrm{e}^*_{k+1},\dots,\mathrm{e}^*_n\}\,.$$

Démonstration. —

1. On part de l'observation que $x \in \mathrm{F}^\perp$ si et seulement si $\bar{b}(x) \in \mathrm{F}^\circ$, où $\bar{b} : \mathrm{E} \to \mathrm{E}^*$ est associée à la forme polaire de q. En effet, $x \in \mathrm{F}^\perp$ est équivalent à $b(x,y) = 0$ pour tout $y \in \mathrm{F}$, ce qui s'écrit aussi $\bar{b}(x)(y) = 0$ pour tout $y \in \mathrm{F}$, soit $\bar{b}(x) \in \mathrm{F}^\circ$.

   Or, q est non dégénérée, donc $\bar{b}$ est un isomorphisme et $\mathrm{F}^\circ = \bar{b}(\mathrm{F}^\perp)$. D'où

   $$\dim \mathrm{F}^\perp = \dim \mathrm{F}^\circ = \dim \mathrm{E} - \dim \mathrm{F}\,.$$

2. On a vu que $\mathrm{F} \subset (\mathrm{F}^\perp)^\perp$. De plus, $\dim \mathrm{F} = \dim(\mathrm{F}^\perp)^\perp$ par l'identité que l'on vient de montrer. Donc $\mathrm{F} = (\mathrm{F}^\perp)^\perp$.

3. est laissé à titre d'exercice.

■

Définition 2.3. — *On dit qu'un vecteur non nul x est* isotrope *si $q(x) = 0$. On note par $\mathcal{C}(q)$ le* cône *des vecteurs isotropes. On remarque en effet que si $q(x) = 0$ alors, pour tout $\lambda \in \mathbb{K}$, $q(\lambda x) = 0$. On a toujours* $\mathrm{Ker}\, q \subset \mathcal{C}(q)$.

Un espace vectoriel F *est* isotrope *si* $\mathrm{F} \cap \mathrm{F}^\perp \neq \{0\}$.

Enfin, on dit que F *est* totalement isotrope *si* $\mathrm{F} \subset \mathrm{F}^\perp$, *autrement dit, si* $q|_\mathrm{F} \equiv 0$.

Exemple. Si $\mathrm{E} = \mathbb{R}^n$ et $q(\mathrm{X}) = x_1x_2$, alors les vecteurs de la base canonique sont tous isotropes.

Remarque. — Si F est totalement isotrope, alors $\dim \mathrm{F} \leq \dim \mathrm{E}/2$. De plus, si F est isotrope, alors $\mathrm{F} \cap \mathrm{F}^\perp$ est totalement isotrope. Si q est non dégénérée et si F est non isotrope, alors on a $\mathrm{F} \oplus \mathrm{F}^\perp = \mathrm{E}$.

Définition 2.4. — *Etant donnée une forme quadratique, on définit son* indice ν *comme le maximum des dimensions des espaces totalement isotropes. Si $\nu = 0$* i.e.*, si $q(x) = 0 \Rightarrow x = 0$, on dit que q est anisotrope, ou définie.*

Remarque. — Si E est un espace réel, et si q est définie, alors, ou bien pour tout $x \neq 0$ on a $q(x) > 0$, ou bien, pour tout $x \neq 0$ on a $q(x) < 0$.

Le cas de la dimension 2. Soit E un espace vectoriel de dimension 2 muni d'une forme quadratique q. De quatre choses l'une :

1. la forme est anisotrope ;
2. il existe exactement une seule droite isotrope, et dans une base appropriée, $q(x, y) = ax^2$, pour $a \neq 0$;
3. il existe exactement deux droites isotropes, et dans une base appropriée,
$$q(x, y) = xy = (1/4)\{(x + y)^2 - (x - y)^2\} ;$$
dans ce cas, on dit que E est *hyperbolique*.
4. il existe au moins trois droites isotropes, et $q \equiv 0$.

Démonstration. — On suppose que la forme n'est pas définie. Il existe donc une base $(\mathrm{e}_1, \mathrm{e}_2)$ telle que $q(\mathrm{e}_1) = 0$. Notons $\alpha = q(\mathrm{e}_2)$ et $\beta = b(\mathrm{e}_1, \mathrm{e}_2)$. Il vient
$$q(x\mathrm{e}_1 + y\mathrm{e}_2) = 2\beta xy + \alpha y^2 .$$
Si $x\mathrm{e}_1 + y\mathrm{e}_2 \in \mathcal{C}(q)$, alors, ou bien $y = 0$, ou bien $2\beta x + \alpha y = 0$. On distingue plusieurs cas selon les valeurs de α et β.

1. Si $\alpha = \beta = 0$. Alors on a au moins trois droites isotropes et $q \equiv 0$.
2. Si $\alpha \neq 0$ et $\beta = 0$, on obtient $q(x\mathrm{e}_1 + y\mathrm{e}_2) = \alpha y^2$, donc on a exactement une seule droite isotrope, et la forme voulue.
3. Si $\beta \neq 0$, on obtient deux droites isotropes : $y = 0$ et $2\beta x + \alpha y = 0$. De plus, si $\alpha = 0$, alors
$$q(x\mathrm{e}_1 + y\mathrm{e}_2) = 2\beta xy .$$
et si $\alpha \neq 0$ alors
$$q(x\mathrm{e}_1 + y\mathrm{e}_2) = \frac{1}{4}\left\{(2\beta x + (\alpha + 1)y)^2 - (2\beta x - (\alpha - 1)y)^2\right\} .$$

■

On en déduit la caractérisation suivante des plans hyperboliques.

Proposition 2.2. — *Un plan* (E, q) *muni d'une forme quadratique est hyperbolique si et seulement si* q *est non dégénérée et s'il existe un vecteur isotrope.*

On en déduit la décomposition générale d'un espace (E, q).

Théorème 2.2 (décomposition d'un espace). — *Soit* (E, q) *un espace de dimension finie muni d'une forme quadratique non dégénérée. Il se décompose sous la forme*
$$\mathrm{E} = (\oplus_{1 \leq j \leq r} \mathrm{P}_j) \oplus \mathrm{F}$$
où les (P_j, q) *sont des plans hyperboliques et* (F, q) *est anisotrope.*

On commence par un lemme.

Lemme 2.1. — *Soit* (E, q) *un espace de dimension finie muni d'une forme quadratique non dégénérée. Si* x *est isotrope, il existe un plan hyperbolique contenant* x*.*

Démonstration. — Il existe y tel que $b(x, y) \neq 0$. Du coup, le plan engendré par x et y est hyperbolique puisque la matrice de q est de rang 2 et x est isotrope. ■

Démonstration du Théorème 2.2. — Si q est anisotrope, alors c'est bon. Sinon, on procède par récurrence. Le cas $n = 1$ est sans intérêt, le cas $n = 2$ provient du Lemme 2.1.

Soit E un espace de dimension $n \geq 3$. Si $\mathcal{C}(q) \neq \varnothing$, alors le Lemme 2.1 produit un plan hyperbolique P qui contient une unique droite isotrope $\mathbb{K}x_0$.

Montrons que $\mathrm{E} = \mathrm{P} \oplus \mathrm{P}^\perp$. Si $v \in \mathrm{P} \cap \mathrm{P}^\perp$ alors $x \in \mathrm{F} \mapsto b(v, x)$ est identiquement nul. Ceci contredit que q est non dégénérée. Donc $\mathrm{E} = \mathrm{P} \oplus \mathrm{P}^\perp$. L'hypothèse de récurrence s'applique à $\mathrm{P}^\perp$. ■

2.2 Le groupe orthogonal

On suppose que $car(\mathbb{K}) \neq 2$.

Etant donné un espace vectoriel muni d'une forme quadratique non dégénérée, on dégage des endomorphismes particuliers qui respectent la structure ajoutée. Plus précisément, on dit que $u \in \mathrm{End(E)}$ est une *isométrie* si, pour tous $x, y \in \mathrm{E}$, on a $b(u(x), u(y)) = b(x, y)$.

Lemme 2.2. — *Une isométrie d'un espace de dimension finie est inversible. Un endomorphisme u est une isométrie si et seulement si $q \circ u = q$.*

L'ensemble des isométries est un groupe appelé *groupe orthogonal* et noté $\mathrm{O}(q)$.

Remarque. — Si b est hermitienne ou alternée, on peut encore parler d'isométries. Dans le cas hermitien, on parle du groupe unitaire $\mathrm{U}(b)$ et dans le cas alterné du groupe symplectique $\mathrm{S}p(b)$.

Démonstration. — La première assertion vient du fait que b est non dégénérée. Si $x \in \operatorname{Ker} u$, alors, pour tout $y \in \mathrm{E}$, on a $b(x, y) = b(u(x), u(y)) = b(0, u(y)) = 0$, donc $x \in \operatorname{Ker} b$ et $x = 0$.

Supposons que $q \circ u = q$. On utilise la relation reliant la forme quadratique à sa forme polaire. ■

Ce résultat reste vrai pour les formes hermitiennes.

Soit (E, q) muni d'une base $\mathcal{B}$. Si $u \in \mathrm{O}(q)$, alors

$$ {}^t\mathrm{Mat}(u, \mathcal{B}) \cdot \mathrm{Mat}(q, \mathcal{B}) \cdot \mathrm{Mat}(u, \mathcal{B}) = \mathrm{Mat}(q, \mathcal{B}). $$

Il s'ensuit que $(\det u)^2 = 1$. On définit alors le groupe spécial orthogonal $\mathrm{SO}(q)$ comme le sous-groupe normal des isométries de déterminant 1.

Symétries Orthogonales. Une symétrie est un endomorphisme u tel que $u \circ u = \mathrm{Id}$. En particulier, elles sont inversibles. De plus, les valeurs propres sont ± 1, et il existe une décomposition de E en somme directe $\mathrm{E} = \mathrm{E}_+ \oplus \mathrm{E}_-$, où E_+ est l'espace propre associé à la valeur propre 1, et E_- à la valeur propre -1.

Proposition 2.3. — *Une symétrie est orthogonale si et seulement si les E_+ et E_- sont orthogonaux. Dans ce cas, ces espaces sont non isotropes.*

Réciproquement, si F est un sous-espace non isotrope, alors il existe une unique symétrie orthogonale telle que F soit exactement l'espace propre associé à la valeur propre 1.

Démonstration. — Si u est orthogonale, alors, pour $x \in \mathrm{E}_+$ et $y \in \mathrm{E}_-$, on a

$$ b(x, y) = b(u(x), u(y)) = -b(x, y) $$

donc $b(x, y) = 0$ car $car\, \mathbb{K} \neq 2$.

Réciproquement, si ces espaces sont orthogonaux, alors, soient $x, y \in \mathrm{E}$. On écrit

$$ \begin{cases} x = x_+ + x_-, & (x_+, x_-) \in \mathrm{E}_+ \times \mathrm{E}_- \\ y = y_+ + y_-, & (y_+, y_-) \in \mathrm{E}_+ \times \mathrm{E}_- \end{cases} $$

Il vient

$$ b(u(x), u(y)) = b(x_+, y_+) + b(x_-, y_-) = b(x, y)\,. $$

Soit F non isotrope. On note $\mathrm{H} = \mathrm{F}^\perp$. On a donc $\mathrm{E} = \mathrm{F} \overset{\perp}{\oplus} \mathrm{H}$, et on peut définir $u \in \mathrm{O}(q)$ par $u|_{\mathrm{F}} = \mathrm{Id}$ et $u|_{\mathrm{H}} = -\mathrm{Id}$. ■

Remarque. — La conjugaison d'une symétrie orthogonale est encore une symétrie orthogonale.

Définition 2.5. — *Lorsque* $\dim \mathrm{E}_- = 1$*, on dit que u est une* réflexion*, et quand* $\dim \mathrm{E}_- = 2$*, on parle de* renversement*. En dimension trois, si* $\dim \mathrm{E}_+ = 1$*, on parle de demi-tour par rapport à x, si x est un vecteur directeur de* E_+*.*

Similitudes. Il s'agit d'endomorphismes u tels qu'il existe un scalaire $\lambda \in \mathbb{K}^*$ tel que $b(u(x), u(y)) = \lambda \cdot b(x, y)$. Ils forment un groupe $\mathrm{GO}(q)$ et on a la suite exacte

$$1 \to \mathrm{O}(q) \to \mathrm{GO}(q) \to \mathbb{K}^*$$

où la dernière flèche est donnée par le scalaire λ.

Matriciellement, on obtient l'identité suivante :

$$^t\mathrm{Mat}(u, \mathcal{B}) \cdot \mathrm{Mat}(q, \mathcal{B}) \cdot \mathrm{Mat}(u, \mathcal{B}) = \lambda \cdot \mathrm{Mat}(q, \mathcal{B})\,.$$

Du coup, $\det^2 u = \lambda^n$.

Lorsque $\mathbb{K}$ est algébriquement clos, ou lorsque $(\mathbb{K}^*)^2 = \mathbb{K}^*$, on a la suite exacte courte :

$$1 \to \mathrm{O}(q) \to \mathrm{GO}(q) \to \mathbb{K}^* \to 1.$$

En effet, si $\mu \in \mathbb{K}^*$, on considère $\lambda \in \mathbb{K}$ tel que $\lambda^2 = \mu$, et $u = \lambda \mathrm{I}$. Il vient $b(u(x), u(y)) = \lambda^2 \cdot b(x, y) = \mu \cdot b(x, y)$. En général, il n'est pas évident de trouver une section qui donne la racine carrée d'un scalaire.

On a la caractérisation suivante des similitudes.

Proposition 2.4. — *Soit* E *un* $\mathbb{K}$*-espace vectoriel de dimension finie muni d'une forme quadratique* q *non dégénérée. Soit* $u \in \mathrm{GL(E)}$*. Alors,* u *est une similitude si et seulement si* u *préserve l'orthogonalité, soit*

$$\forall\, x, y \in \mathrm{E},\ x \perp y \iff u(x) \perp u(y)\,.$$

Démonstration. — Il est aisé de vérifier qu'une similitude préserve l'orthogonalité. Inversement, on considère une base orthogonale $\mathcal{B} = (\mathrm{e}_1, \ldots, \mathrm{e}_n)$ de E. On considère $\varepsilon_i = u(\mathrm{e}_i)$, $i = 1, \ldots, n$, qui forment aussi une base orthogonale par hypothèse.

Comme q est non dégénérée, $q(\mathrm{e}_i)$, $q(\varepsilon_i)$ sont non nuls, donc il existe $\lambda_i \in \mathbb{K}^*$ tel que $q(\varepsilon_i) = \lambda_i q(\mathrm{e}_i)$. Il suffit de montrer que λ_i est indépendant de i pour conclure que u est une similitude. On se donne deux indices $i \neq j$ et on pose $\lambda = -q(\mathrm{e}_i)/q(\mathrm{e}_j)$. Il vient

$$b(\mathrm{e}_i + \mathrm{e}_j, \mathrm{e}_i + \lambda \mathrm{e}_j) = q(\mathrm{e}_i) + \lambda q(\mathrm{e}_j) = 0$$

donc ces vecteurs sont orthogonaux. Du coup, $u(\mathrm{e}_i + \mathrm{e}_j) = \varepsilon_i + \varepsilon_j$ et $u(\mathrm{e}_i + \lambda \mathrm{e}_j) = \varepsilon_i + \lambda \varepsilon_j$ sont aussi orthogonaux et on en déduit

$$\lambda = -\frac{q(\varepsilon_i)}{q(\varepsilon_j)} = -\frac{\lambda_i}{\lambda_j} \cdot \frac{q(\mathrm{e}_i)}{q(\mathrm{e}_j)} = \lambda \cdot \frac{\lambda_i}{\lambda_j}\,.$$

Ceci montre bien que λ_i est une fonction constante de i. ∎

2.3 Classification des formes quadratiques

Définition 2.6. — *Deux espaces* (E, q) *et* (E', q') *sont équivalents s'il existe un isomorphisme* $u : \mathrm{E} \to \mathrm{E}'$ *tel que, pour tout* $x \in \mathrm{E}$*, on a* $q'(u(x)) = q(x)$*.*

On remarque que la conjugaison par u permet d'identifier $\mathrm{O}(q)$ à $\mathrm{O}(q')$.

Si q est une forme quadratique sur E, alors son expression dans des bases différentes fournit des exemples de formes équivalentes. Réciproquement, la classe d'équivalence d'une forme q définie sur E est en correspondance avec ses différentes expressions dans des bases différentes.

Définition 2.7. — *Un* invariant *est une application définie sur* $\{(\mathrm{E}, q)\}/\sim$.

Exemples. Soient $(\mathrm{E}, \mathcal{B}, q)$ et $\mathrm{M} = \mathrm{Mat}(q, \mathcal{B})$.

- Le rang de la matrice M est un invariant. On note $\mathrm{rg}(q) = \mathrm{rg}(\mathrm{M})$.
- L'indice ν est un invariant aussi.

– **Discriminant.** Soit $\mathcal{C}$ une autre base de E, et soit $\mathrm{N} = \mathrm{Mat}(q, \mathcal{C})$. On a donc $\mathrm{N} = {}^t\mathrm{PMP}$, où P est la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$. Du coup, $\det \mathrm{N} = \det \mathrm{M} (\det \mathrm{P})^2$. On remarque que le déterminant n'est pas un invariant. En revanche, si $q \sim q'$, alors leurs déterminants dans la base $\mathcal{B}$ diffèrent d'un carré de $\mathbb{K}^*$ car P est inversible. On définit ainsi le *discriminant* de q comme $\det \mathrm{M} \in \mathbb{K}/(\mathbb{K}^*)^2$. Notons que dans un corps commutatif, $\mathbb{K}^*/(\mathbb{K}^*)^2$ a une structure de groupe, et si q est non dégénérée, alors son discriminant vit dans $\mathbb{K}^*/(\mathbb{K}^*)^2$.

  Cette quantité peut s'avérer utile quand $\mathbb{K}^*/(\mathbb{K}^*)^2$ est non triviale.

  Notons par exemple que $\mathbb{C}/(\mathbb{C}^*)^2 \approx \{0, 1\}$ et $\mathbb{R}/(\mathbb{R}^*)^2 \approx \{-1, 0, 1\}$.

Si E est un espace vectoriel de dimension (n) finie sur $\mathbb{Q}$, alors on a une infinité de classes d'équivalence, qui peuvent se différencier par leur discriminant : on identifie E à $\mathbb{Q}^n$, et on définit, pour tout nombre premier p,

$$q_p(x) = \sum_{1 \leq j < n} x_j^2 + p x_n^2 \,.$$

Le discriminant est p *mod* $(\mathbb{Q}^2)$. Montrons que si p_1, p_2 sont deux nombres premiers différents, alors leur discriminants sont différents : si ce n'était pas le cas, on trouverait $a, b \in \mathbb{Z}$ premiers entre eux tels que $p_1 a^2 = p_2 b^2$. Ceci implique d'une part que p_1 divise b^2, donc b ; par suite, p_1^2 divise b^2. Par conséquent, p_1 doit aussi diviser a^2, donc a, ce qui contredit que a et b sont premiers entre eux.

Théorème 2.3. — *Si* $\mathbb{K}$ *est algébriquement clos, alors, pour tout* q*, il existe une base telle que*

$$q(x) = \sum_{1 \leq j \leq \mathrm{rg}(q)} (\mathrm{e}_j^*(x))^2.$$

On a exactement $\dim \mathrm{E} + 1$ *classes d'équivalences, qui se différencient à l'aide du rang de* q*.*

Démonstration. — Dans une base orthogonale $\mathcal{B} = (\mathrm{e}_1, \ldots, \mathrm{e}_n)$, on a

$$q(x) = \sum_{1 \leq j \leq \mathrm{rg}(q)} \lambda_j (\mathrm{e}_j^*(x))^2,$$

avec $\lambda_j \in \mathbb{K}^*$. Comme $\mathbb{K}$ est algébriquement clos, il existe $\mu_j \in \mathbb{K}$ tel que $\mu_j^2 = \lambda_j$. En posant $f_i = \mathrm{e}_i/\mu_i$, on obtient

$$q(x) = \sum \lambda_j (\mathrm{e}_j^*(x))^2 = \sum (\mu_j \mathrm{e}_j^*(x))^2 = \sum (f_j^*(x))^2 \,.$$

■

Théorème 2.4. — *Si* $\mathbb{K} = \mathbb{R}$*, alors, pour tout* q*, il existe un unique couple d'entiers naturels* (s, t)*, appelé la signature de* q *tel que* $s + t = \mathrm{rg}(q)$*, et tel qu'il existe une base dans laquelle*

$$q(x) = \sum_{1 \leq j \leq s} (\mathrm{e}_j^*(x))^2 - \sum_{s+1 \leq j \leq \mathrm{rg}(q)} (\mathrm{e}_j^*(x))^2.$$

On a exactement $(\dim \mathrm{E} + 1)(\dim \mathrm{E} + 2)/2$ *classes d'équivalence, qui se distinguent à l'aide de la signature de* q*.*

Démonstration. — Dans une base orthogonale $\mathcal{B} = (\mathrm{e}_1, \ldots, \mathrm{e}_n)$, on a

$$q(x) = \sum_{1 \leq j \leq \mathrm{rg}(q)} \lambda_j (\mathrm{e}_j^*(x))^2,$$

avec $\lambda_j \in \mathbb{R}^*$. On suppose que les s premiers scalaires sont positifs et les t derniers négatifs. On note $\mu_j = \sqrt{|\lambda_j|}$. En posant $f_j = \mathrm{e}_j/\mu_j$, on obtient

$$q(x) = \sum \lambda_j (\mathrm{e}_j^*(x))^2 = \sum_{1 \leq j \leq s} (\mu_j \mathrm{e}_j^*(x))^2 - \sum_{s+1 \leq j \leq \mathrm{rg}(q)} (\mu_j \mathrm{e}_j^*(x))^2 = \sum_{1 \leq j \leq s} (f_j^*(x))^2 - \sum_{s+1 \leq j \leq \mathrm{rg}(q)} (f_j^*(x))^2 \,.$$

Il reste à montrer que la signature est bien définie, et qu'elle distingue les classes d'équivalences.

Supposons qu'on ait deux bases $\mathcal{B}$ et $\mathcal{B}'$ avec des nombres (s,t) et (s',t') qui vérifient $s > s'$. On note $\mathrm{F} = \mathrm{Vect}\{\mathrm{e}_1, \dots, \mathrm{e}_s\}$ et $\mathrm{G} = \mathrm{Vect}\{\mathrm{e}_{s'+1}, \dots, \mathrm{e}_n\}$. Par suite, pour tout $x \in \mathrm{F} \setminus \{0\}$, on a $q(x) > 0$ et $q(x) \leq 0$ pour $x \in \mathrm{G}$. Du coup, on a $\mathrm{F} \cap \mathrm{G} = \{0\}$. Mais alors,

$$\dim \mathrm{E} = n \geq \dim \mathrm{F} + \dim \mathrm{G} = s + (n - s') > s' + (n - s') = n\,.$$

Ceci est une contradiction. Donc $(s,t) = (s',t')$. Cet argument montre aussi que deux formes quadratiques avec des signatures différentes ne peuvent être équivalentes. ■

Extrema locaux et position d'une hypersurface par rapport à son plan tangent. Si $f : \mathbb{R}^n \to \mathbb{R}$ est une application de classe $\mathcal{C}^2$, alors le théorème de Schwarz implique que la matrice des dérivées partielles secondes est symétrique. Cela nous permet d'appliquer ce qui précède au calcul des variations.

Lemme de Morse. — *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une application de classe $\mathcal{C}^k$, $k \geq 2$ telle que $f(0) = \mathrm{D}_0 f = 0$ et telle que $\mathrm{D}_0^2 f$ soit inversible. Alors il existe un voisinage V de l'origine et des applications de classe $\mathcal{C}^{k-2}$ $y_1, \dots, y_r$ et $z_1, \dots, z_s$ définies sur V telles que $r + s = n$ et, pour $x \in \mathrm{V}$, on ait*

$$f(x) = \sum_{1 \leq j \leq r} y_j^2 - \sum_{1 \leq j \leq s} z_j^2\,.$$

Démonstration. — On considère le développement limité avec reste intégrale de f au voisinage de l'origine. On a

$$f(x) = \int_0^1 (1-t)\mathrm{D}^2_{tx} f(x,x)dt = {}^t\mathrm{X} \cdot \int_0^1 (1-t)\mathrm{D}^2_{tx} f dt \cdot \mathrm{X}\,.$$

On note

$$\mathrm{A}(x) = \left(\int_0^1 (1-t) \frac{\partial^2 f}{\partial x_i \partial x_j}(tx)dt \right)_{i,j} \qquad \text{et} \qquad \mathrm{A}_0 = \mathrm{A}(0)\,.$$

On utilise alors le lemme suivant.

Lemme 2.3. — *Si $\mathrm{A}_0 \in \mathcal{S}_n(\mathbb{R}) \cap \mathrm{GL}_n(\mathbb{R})$, il existe un voisinage U de A_0 dans $\mathcal{S}_n(\mathbb{R})$ et une application infiniement différentiable $\psi : \mathrm{U} \to \mathcal{M}_n(\mathbb{R})$ tels que $\psi(\mathrm{A}_0) = \mathrm{I}$ et, pour tout $\mathrm{A} \in \mathrm{U}$,*

$${}^t\psi(\mathrm{A}) \cdot \mathrm{A}_0 \cdot \psi(\mathrm{A}) = \mathrm{A}\,.$$

Du coup, si x est assez proche de l'origine, alors $f(x) = {}^t x {}^t\psi(\mathrm{A}(x))\mathrm{A}_0\psi(\mathrm{A}(x))x$. On pose $\psi_1(x) = \psi(\mathrm{A}(x))x$, et on obtient

$$f(x) = {}^t\psi_1(x)\mathrm{A}_0\psi_1(x)\,.$$

Or il existe une base P dans laquelle A_0 soit diagonale avec r valeurs sur la diagonale égales à 1 et s égales à -1. On note J cette matrice et on pose $\psi_2 = \mathrm{P} \cdot \psi_1$. Il vient $f(x) = {}^t\psi_2(x)\mathrm{J}\psi_2(x)$. Si on appelle $y_1, \dots, y_r$ et $z_1, \dots, z_s$ les coordonnées de ψ_2, on obtient la forme recherchée.

Ceci établit le lemme de Morse modulo le Lemme 2.3. ■

Démonstration du Lemme 2.3. — On considère l'application $h : \mathcal{M}_n(\mathbb{R}) \to \mathcal{S}_n(\mathbb{R})$ définie par $h(\mathrm{M}) = {}^t\mathrm{M}\mathrm{A}_0\mathrm{M}$. On calcule la différentielle à l'identité de h.

$$h(\mathrm{I}+\mathrm{M}) = {}^t(\mathrm{I}+\mathrm{M})\mathrm{A}_0(\mathrm{I}+\mathrm{M}) = \mathrm{A}_0 + ({}^t\mathrm{M}\mathrm{A}_0 + \mathrm{A}_0\mathrm{M}) + {}^t\mathrm{M}\mathrm{A}_0\mathrm{M} = \mathrm{A}_0 + ({}^t\mathrm{M}\mathrm{A}_0 + \mathrm{A}_0\mathrm{M}) + \mathrm{O}(\|\mathrm{M}\|^2)\,.$$

Donc $\mathrm{D}_\mathrm{I} h(\mathrm{M}) = {}^t\mathrm{M}\mathrm{A}_0 + \mathrm{A}_0\mathrm{M}$. Cette application n'est pas inversible puisque $\dim \mathcal{S}_n(\mathbb{R}) < \dim \mathcal{M}_n(\mathbb{R})$. En revanche, le noyau consiste en les matrices M telles que $\mathrm{A}_0\mathrm{M}$ soit antisymétrique. On considère l'espace E des matrices M telles que $\mathrm{A}_0\mathrm{M}$ soit symétrique. Cet espace est supplémentaire à $\operatorname{Ker} \mathrm{D}_\mathrm{I} h$, et la restriction de $\mathrm{D}_\mathrm{I} h$ à E devient maintenant inversible (injective car A_0 est inversible et espaces source et but de même dimension).

Le théorème d'inversion locale appliqué à $h|_\mathrm{E}$ montre qu'il existe des voisinages U de A_0 et V de I et un difféomorphisme infiniment différentiable $\psi : \mathrm{U} \to \mathrm{V}$ qui inverse $h|_\mathrm{E}$. ■

Corollaire 2.1. — *Sous ces hypothèses* 0 *est un maximum local strict de* f *si et seulement si* $\mathrm{D}^2 f$ *est définie négative, et est un minimum local strict de* f *si et seulement si* $\mathrm{D}^2 f$ *est définie positive.*

Corollaire 2.2. — *On étudie dans* $\mathbb{R}^{n+1}$ *l'hypersurface* $\mathcal{S}$ *définie par* $x_{n+1} = \mathrm{F}(x_1, \ldots, x_n)$ *où* F *est de classe* $\mathcal{C}^k(\mathbb{R}^n)$, $k \geq 2$, *et* $\mathrm{D}^2\mathrm{F}$ *est non dégénérée. Le plan tangent au point* $p = (x_0, \mathrm{F}(x_0))$ *sépare localement* $\mathcal{S}$ *d'un demi-espace si et seulement si* $\mathrm{D}^2_{x_0}\mathrm{F}$ *est définie.*

Démonstration. — On considère l'application $g(x) = \mathrm{F}(x) - (\mathrm{F}(x_0) + \mathrm{D}_{x_0}\mathrm{F}(x - x_0))$. Cette application vérifie les hypothèses du lemme de Morse. On en déduit que

$$\mathrm{F}(x) = \mathrm{F}(x_0) + \mathrm{D}_{x_0}\mathrm{F}(x - x_0) + \sum_{1 \leq j \leq r} y_j^2 - \sum_{1 \leq j \leq s} z_j^2 \,.$$

■

2.4 Classification sur les corps finis.

Si $\mathbb{K}$ est un corps à q éléments, alors il existe un nombre premier p tel que $q = p^n$ et $\mathbb{K}$ est le corps de décomposition de $\mathrm{X}^{p^n} - \mathrm{X}$ sur $\mathbb{Z}/p\mathbb{Z}$. On écrit $\mathbb{K} = \mathbb{F}_q$. On a de plus $\mathbb{F}_q^* \approx \mathbb{Z}/(q-1)\mathbb{Z}$ en tant que groupe.

Carrés dans $\mathbb{F}_q$. Soit

$$\begin{array}{rccc} f: & \mathbb{F}_q^* & \to & \mathbb{F}_q^* \\ & x & \mapsto & x^2 \end{array}$$

C'est un morphisme de groupes. Si $p = 2$, alors f est une bijection, sinon $f(x) = 1$ si et seulement si $x = \pm 1$. Du coup, $\operatorname{Ker} f = \{-1, 1\}$. On en déduit que

$$|(\mathbb{F}_q^*)^2| = \frac{q-1}{2} \,.$$

En effet, en caractéristique 2, on a $2x = 0$ donc $x = -x$.

Proposition 2.5. — *On suppose que* $p > 2$. *Alors* x *est un carré dans* $\mathbb{F}_q^*$ *si et seulement si* $x^{(q-1)/2} = 1$.

On rappelle que si $p \neq 2$, alors q est impaire.

Démonstration. — Puisque x est inversible, on a $x^{q-1} = 1$ par le petit théorème de Fermat. Du coup,

$$x^{q-1} - 1 = (x^{(q-1)/2} - 1)(x^{(q-1)/2} + 1) = 0 \,.$$

On note R_+ les racines de $\mathrm{X}^{(q-1)/2} = 1$ et R_- celles de $\mathrm{X}^{(q-1)/2} = -1$. Si $x = y^2$ alors $x^{(q-1)/2} = y^{q-1} = 1$ donc $x \in \mathrm{R}_+$. Donc $(\mathrm{F}_q^*)^2 \subset \mathrm{R}_+$. Or $(\mathrm{F}_q^*)^2$ a $(q-1)/2$ éléments et R_+ au plus $(q-1)/2$, donc $(\mathrm{F}_q^*)^2 = \mathrm{R}_+$. ■

On en déduit deux corollaires intéressants en soi.

Corollaire 2.3. — (-1) *est un carré dans* $\mathbb{F}_p$ *si et seulement si* $p \equiv 1 \mod 4$.

Démonstration. — (-1) est un carré si et seulement si $(-1)^{(p-1)/2} = 1$ dans $\mathbb{F}_p$, ce qui équivaut à la parité de $(p-1)/2$, soit $p \equiv 1 \mod 4$. ■

Corollaire 2.4. — *Si* $p > 2$ *alors il existe* $\alpha \in \mathbb{F}_q^*$ *tel que*

$$\mathbb{F}_q^*/(\mathbb{F}_q^*)^2 \approx \{1, \alpha\}.$$

Démonstration. — Le sous-groupe $(\mathbb{F}_q^*)^2$ agit librement sur $\mathbb{F}_q^*$ par multiplication (si $x^2 y = y$ alors $x^2 = 1$), donc toutes les orbites ont le même cardinal. Puisque la classe de 1 a $(q-1)/2$ éléments, $\mathbb{F}_q$ n'a que deux classes. Soit $\alpha \in \mathrm{R}_-$. On a $\mathbb{F}_q^*/(\mathbb{F}_q^*)^2 \approx \{1, \alpha\}$. ■

Classification des formes quadratiques. On peut maintenant énoncer le résultat principal de ce paragraphe.

Théorème 2.5. — *Soit* E *un* $\mathbb{F}_q$*-espace vectoriel de dimension finie. Il existe exactement deux classes d'équivalence de formes quadratiques non dégénérées. Elles sont représentées par*

$$q(x) = \sum_{1\leq i\leq n} (\mathrm{e}_i^*(x))^2,$$

et par

$$q(x) = \sum_{1\leq i\leq n-1} (\mathrm{e}_i^*(x))^2 + \alpha(\mathrm{e}_n^*(x))^2,$$

où $\alpha \in \mathbb{F}_q^*$ *n'est pas un carré de* $\mathbb{F}_q$*.*

Ces classes sont distinguées par leurs discriminants.

On commence par un petit lemme.

Lemme 2.4. — *Soient* $\alpha_1, \alpha_2 \in \mathbb{F}_q^*$*. L'équation*

$$a^2\alpha_1 = 1 - b^2\alpha_2$$

en $a, b \in \mathbb{F}_q$ *admet au moins une solution.*

Démonstration. — Puisque $\mathbb{F}_q$ a $(q+1)/2$ carrés, l'ensemble $\{a^2\alpha_1,\ a \in \mathbb{F}_q\}$ a autant d'éléments car $\alpha_1 \neq 0$. De même pour $\{1 - b^2\alpha_2,\ b \in \mathbb{F}_q\}$. Du coup, ces ensembles s'intersectent. ■

Démonstration du théorème. — On procède par récurrence.

Cas $n = 2$. Soit $(\mathrm{e}_1, \mathrm{e}_2)$ une base orthogonale telle que $q(\mathrm{e}_1)$ et $q(\mathrm{e}_2)$ soient non nuls. D'après le lemme, il existe $a, b \in \mathbb{F}_q$ tel que $a^2q(\mathrm{e}_1) = 1 - b^2q(\mathrm{e}_2)$. On note $f_1 = a\mathrm{e}_1 + b\mathrm{e}_2$. Du coup, $q(f_1) = 1$. Soit f_2 un vecteur non nul et orthogonal à f_1. Ou bien $q(f_2)$ est un carré, en ce cas, on peut supposer que $q(f_2) = 1$, ou bien non, et on peut se ramener au cas $q(f_2) = \alpha$.

Cas général. On construit comme ci-dessus un vecteur e_1 tel que $q(\mathrm{e}_1) = 1$. Soit $\mathrm{H} = \mathrm{e}_1^\perp$. Alors H est de dimension une de moins que E. Du coup, l'hypothèse de récurrence s'applique.

On a vu que le discriminant est un invariant. Dans la première classe, on trouve 1, et dans la seconde α. Donc ces classes sont distinctes. ■

2.5 Le théorème de Witt et ses conséquences

Théorème 2.6 (de Witt). — *Soit* q *une forme quadratique sur un espace* E *de dimension finie. Soient* F *un sous-espace de* E *et* $s : \mathrm{F} \to \mathrm{E}$ *une application linéaire injective telle que* $q(s(x)) = q(x)$ *pour tout* $x \in \mathrm{E}$*. Alors il existe* $u \in \mathrm{O}(q)$ *tel que* $u|_\mathrm{F} = s$*.*

Démonstration. — On étudie deux cas.

Premier cas. $q|_\mathrm{F}$ est non dégénérée.

On procède par récurrence sur la dimension de F.

Si F est une droite $\mathbb{K}x$, on note $y = s(x)$. Puisque

$$q(x+y) + q(x-y) = 2(q(x) + q(y)) = 4q(x) \neq 0,$$

il existe $\varepsilon = \pm 1$ tel que $q(x + \varepsilon y) \neq 0$. Notons $\mathrm{G} = \mathbb{K}(x + \varepsilon y)$. Cet espace est donc non isotrope, donc la Proposition 2.3 implique l'existence d'une symétrie orthogonale σ par rapport à G.

On remarque que $b(x + \varepsilon y, x - \varepsilon y) = q(x) - q(x) + \varepsilon(b(x, y) - b(x, y)) = 0$ donc on a

$$2\sigma(x) = \sigma(x + \varepsilon y) + \sigma(x - \varepsilon y) = x + \varepsilon y - (x - \varepsilon y) = 2\varepsilon y,$$

soit $\sigma(x) = \varepsilon s(x)$. L'application $u = \varepsilon\sigma$ répond à la question.

Supposons le théorème vrai pour tout sous-espace non dégénéré de dimension au plus n. Soit maintenant F de dimension $n + 1$ non dégénéré. On considère une base orthogonale de F, ce qui nous

permet d'écrire $\mathrm{F} = \mathrm{F}_1 \overset{\perp}{\oplus} \mathrm{F}_2$ avec $\dim \mathrm{F}_j \leq n$. Soit $u_1 \in \mathrm{O}(q)$ telle que $u_1|_{\mathrm{F}_1} = s|_{\mathrm{F}_1}$. On a $s \circ u_1^{-1}(\mathrm{F}_2) \subset \mathrm{F}_1^\perp$ car $\mathrm{F}_1 \perp \mathrm{F}_2$. L'hypothèse de récurrence nous permet de trouver $u_2 \in \mathrm{O}(q|_{\mathrm{F}_1^\perp})$ qui prolonge $s \circ u_1^{-1}|_{\mathrm{F}_2}$.

On note $u = u_2 \circ u_1$ sur $\mathrm{F}_1^\perp$ et $u = u_1$ sur F_1. Si $x \in \mathrm{F}_2$ alors $u(x) = u_2 \circ u_1(x) = s(x)$. De plus, u est orthogonal sur si F_1 et $\mathrm{F}_1^\perp$. Si maintenant $x \in \mathrm{F}_1$ et $y \in \mathrm{F}_1^\perp$ alors $b(u(x), u(y)) = b(u_1(x), u_2 \circ u_1(y)) = b(x, u_2(y))$. Or $u_2(\mathrm{F}_1^\perp) = \mathrm{F}_1^\perp$ donc $b(u(x), u(y)) = 0$ et $u \in \mathrm{O}(q)$.

Second cas. $q|_{\mathrm{F}}$ est dégénérée.

On procède aussi par récurrence, en montrant que l'on peut toujours prolonger s à un espace de dimension une de plus. on obtient alors un plongement isométrique défini sur un espace dégénéré de dimension plus grande. En réitérant le processus, on finit par couvrir tout E.

Soit $x \in \mathrm{F} \cap \mathrm{F}^\perp$. Il existe $y_1 \in \mathrm{E}$ tel que $b(x, y_1) = 1$. Du coup, $y \notin \mathrm{F}$. Si on note $y = y_1 - (1/2)q(y_1)x$, alors $q(y) = q(y_1) - q(y_1)b(y_1, x) = 0$ et $b(x, y) = b(y_1, x) - (1/2)q(y_1)q(x) = 1$. On note $\mathrm{G} = \mathrm{F} \oplus \mathbb{K}y$.

On prolonge s à E en un automorphisme σ. Puisque q est non dégéńérée, il existe $y_1' \in \mathrm{E}$ tel que $b(\sigma^{-1}(z), y) = b(z, y_1')$ pour tout $z \in s(\mathrm{F})$ (voir le calcul matriciel sous-jacent). Notons $y' = y_1' - (1/2)q(y_1')s(x)$. On a $q(s(x)) = 0$ et $b(s(x), y_1') = b(x, y) = 1$, donc $q(y') = q(y_1') - q(y_1') = 0$ et, pour $z \in s(\mathrm{F})$, $b(z, y') = b(z, y_1') - (1/2)q(y_1')b(z, s(x)) = b(z, y_1')$ car $s(x) \in s(\mathrm{F}) \cap s(\mathrm{F})^\perp$.

On note $v : \mathrm{G} \to \mathrm{E}$ défini par $v|_{\mathrm{F}} = s$ et $v(y) = y'$. Si $w \in \mathrm{F}$, alors $b(v(w), v(y)) = b(s(w), y') = b(w, y)$. Donc $v \in \mathrm{O}(q|_{\mathrm{G}})$. ■

3 Espaces euclidiens

Définition 3.1. — *Un espace euclidien est un espace vectoriel de dimension finie sur $\mathbb{R}$ muni d'une forme quadratique q définie positive.*

Exemples.

– Si (X, μ) est un espace mesuré, alors on définit $q : \mathrm{L}^2(\mathrm{X}, \mu; \mathbb{R}) \to \mathbb{R}$ par

$$q(f) = \int f^2 d\mu.$$

On a bien $q(f) > 0$ pour tout $f \neq 0$ dans L^2. Toute restriction à un sous-espace de dimension finie confère une structure d'espace euclidien.

– Si $\mathrm{M} \in \mathcal{M}_n(\mathbb{K})$, alors $q(\mathrm{M}) = \mathrm{tr}({}^t\mathrm{MM})$ détermine une forme quadratique définie positive. Pour cela, on note $\mathrm{M} = (a_{ij})$ et on calcule les termes diagonaux de ${}^t\mathrm{MM} = (b_{ij})$:

$$b_{jj} = \sum_{k=1}^{n} a_{kj} a_{kj} = \sum_{k=1}^{n} a_{kj}^2$$

et

$$\mathrm{tr}({}^t\mathrm{MM}) = \sum_{1 \leq i,j \leq n} a_{ij}^2$$

donc $q(\mathrm{M}) \geq 0$ et $q(\mathrm{M}) = 0$ implique $\mathrm{M} = 0$.

Théorème 3.1. — *Un espace euclidien est naturellement un espace vectoriel normé.*

Pour montrer ce résultat, on commence par montrer l'inégalité de Cauchy-Schwarz.

Inégalité de Cauchy-Schwarz. — *Soit (E, q) un $\mathbb{R}$-espace vectoriel de dimension finie muni d'une forme positive (pour tout $x \in \mathrm{E}$, on a $q(x) \geq 0$). On a, pour tout $x, y \in \mathrm{E}$,*

$$b(x, y)^2 \leq q(x) q(y)\,.$$

On note une conséquence directe de cette inégalité.

Corollaire 3.1. — *Le noyau de q coïncide avec son cône isotrope.*

Démonstration. — On a toujours $\operatorname{Ker} q \subset \mathcal{C}(q)$. Si $x \in \mathcal{C}(q)$, alors, pour tout $y \in \mathrm{E}$, on a $0 \leq b(x, y)^2 \leq q(x)q(y) \leq 0$ donc $b(x, y) = 0$ pour tout $y \in \mathrm{E}$. ∎

Démonstration de l'inégalité de Cauchy-Schwarz. — Pour $t \in \mathbb{R}$, on a $q(xt + y) \geq 0$. Or $q(tx+y) = q(x)t^2 + 2tb(x, y) + q(y)$. Comme ce polynôme quadratique ne prend que des valeurs positives, il ne peut s'annuler au plus qu'une fois, donc son discriminant doit être négatif :

$$b(x, y)^2 - q(x) \cdot q(y) \leq 0\,.$$

∎

Démonstration du Théorème 3.1. — On pose, pour $x \in \mathrm{E}$, $\|x\| = \sqrt{q(x)}$. Pour montrer que $\|\cdot\|$ est une norme, il suffit de vérifier l'inégalité triangulaire.

$$\begin{aligned}(\|x\| + \|y\|) - \|x + y\| &= \frac{q(x) + q(y) + 2\sqrt{q(x)q(y)} - q(x + y)}{\|x + y\| + \|x\| + \|y\|} \\ &\geq \frac{2\sqrt{q(x)q(y)} - 2b(x, y)}{\|x + y\| + \|x\| + \|y\|} \geq 0\,.\end{aligned}$$

∎

Définition 3.2. — *Une base orthonormée est une base* $\mathcal{B} = (\mathrm{e}_1, \dots, \mathrm{e}_n)$ *telle que* $b(\mathrm{e}_i, \mathrm{e}_j) = \delta_{i,j}$*, où* $\delta_{i,j}$ *est le symbole de Kronecker.*

Théorème 3.2. — *Dans un espace euclidien, il existe toujours une base orthonormée.*

Dans une telle base, la matrice de q est l'identité. Du coup, les éléments orthogonaux sont représentés par des matrices M telles que ${}^t\mathrm{M} \cdot \mathrm{M} = \mathrm{I}$. Dans cette situation, on appelle la forme polaire *un produit scalaire* et on le note communément $\langle \cdot, \cdot \rangle$.

Démonstration. — Ou bien on applique le théorème de classification des formes quadratiques sur $\mathbb{R}$, ou bien on peut utiliser l'algorithme de Gram-Schmidt. ∎

L'orthogonalisation de Gram-Schmidt implique aussi que, pour tout sous-espace F de E, il existe une base orthonormée dont les premiers $\dim \mathrm{F}$ vecteurs forment une base de F.

Proposition 3.1. — *Soit* $b : \mathrm{E}^2 \to \mathbb{R}$ *une forme bilinéaire sur un espace euclidien. Il existe des endomorphismes* u *et* v *tels que, pour tout* $x, y \in \mathrm{E}$*, on ait*

$$b(x, y) = \langle x, u(y) \rangle = \langle v(x), y \rangle .$$

Ces endomorphismes sont uniques.

Démonstration. — Soit $\mathcal{B} = (\mathrm{e}_1, \dots, \mathrm{e}_n)$ une base orthonormée ; on définit, pour tout $j \in \{1, \dots, n\}$,

$$\begin{cases} u(\mathrm{e}_j) = \displaystyle\sum_{i=1}^{n} b(\mathrm{e}_i, \mathrm{e}_j)\mathrm{e}_i \,, \\ \\ v(\mathrm{e}_i) = \displaystyle\sum_{j=1}^{n} b(\mathrm{e}_i, \mathrm{e}_j)\mathrm{e}_j \,. \end{cases}$$

On a bien

$$\begin{cases} \langle \mathrm{e}_i, u(\mathrm{e}_j) \rangle = \displaystyle\sum_{k=1}^{n} b(\mathrm{e}_k, \mathrm{e}_j)\langle \mathrm{e}_k, \mathrm{e}_i \rangle = b(\mathrm{e}_i, \mathrm{e}_j) \,, \\ \\ \langle v(\mathrm{e}_i), \mathrm{e}_j \rangle = \displaystyle\sum_{k=1}^{n} b(\mathrm{e}_i, \mathrm{e}_k)\langle \mathrm{e}_j, \mathrm{e}_k \rangle = b(\mathrm{e}_i, \mathrm{e}_j) \,. \end{cases}$$

L'unicité provient aussi de ces relations. ∎

Définition 3.3. — *Si* $u \in \mathrm{End}(\mathrm{E})$*, on appelle* adjoint de u*, que l'on note* u^**, l'endomorphisme tel que, pour tout* $x, y \in \mathrm{E}$*,*

$$\langle u(x), y \rangle = \langle x, u^*(y) \rangle .$$

L'adjoint existe toujours d'après la proposition précédente puisque $b : (x, y) \mapsto \langle x, u(y) \rangle$ est bilinéaire. De plus, $u^{**} = u$.

Si M est la matrice de u dans une base orthonormée, alors

$$\langle x, u^*(y) \rangle = \langle u(x), y \rangle = {}^t(\mathrm{M} \cdot \mathrm{X}) \cdot \mathrm{Y} = {}^t\mathrm{X} \cdot {}^t\mathrm{M} \cdot \mathrm{Y}$$

donc la matrice de u^* est ${}^t\mathrm{M}$.

Remarque. — Les polynômes caractéristiques de u et u^* sont les mêmes.

Définition 3.4. — *Un endormorphisme* $u \in \mathrm{End}(\mathrm{E})$ *est* auto-adjoint *si* $u = u^*$.

Dans une base orthonormée, la matrice de u est alors symétrique, et l'application

$$x \mapsto \langle u(x), x \rangle$$

détermine une forme quadratique.

3.1 Réduction des endomorphismes auto-adjoints

Le résultat principal de ce paragraphe est le suivant.

Théorème 3.3. — *Pour tout endomorphisme auto-adjoint u d'un espace euclidien, il existe une base orthonormée de* E *qui diagonalise u.*

Autrement dit, si q et q' sont deux formes quadratiques et si q est définie positive, il existe une base de E *qui soit orthonormée pour q et orthogonale pour q'.*

Lemme 3.1. — *Soit u auto-adjoint.*

1. *Si* $\mathrm{F} \subset \mathrm{E}$ *est un espace vectoriel tel que* $u(\mathrm{F}) \subset \mathrm{F}$*, alors* $u(\mathrm{F}^\perp) \subset \mathrm{F}^\perp$.
2. *Il existe une valeur propre réelle λ de u.*

Démonstration. —

1. Soit $\mathrm{F} \subset \mathrm{E}$ un espace vectoriel tel que $u(\mathrm{F}) \subset \mathrm{F}$. Pour tout $x \in \mathrm{F}^\perp$ et tout $y \in \mathrm{F}$, on a
$$\langle u(x), y\rangle = \langle x, u(y)\rangle = 0\,.$$
Donc $u(x) \in \mathrm{F}^\perp$.
2. On se fixe une base orthonormée de E et on note A la matrice de u dans cette base. Supposons que toutes les valeurs propres sont complexes. Si λ est l'une d'elles, alors $\bar\lambda$ aussi. Soit X un vecteur propre (complexe) de λ. On a
$$\mathrm{A}\cdot\bar{\mathrm{X}} = \overline{\mathrm{AX}} = \overline{\lambda \mathrm{X}} = \bar\lambda\cdot\bar{\mathrm{X}}$$
donc $\bar{\mathrm{X}}$ est un vecteur propre associé à $\bar\lambda$. Or, d'une part ${}^t\mathrm{X}\cdot\mathrm{A}\cdot\bar{\mathrm{X}} = \bar\lambda\cdot{}^t\mathrm{X}\cdot\bar{\mathrm{X}}$, et d'autre part, ${}^t\mathrm{X}\cdot\mathrm{A}\cdot\bar{\mathrm{X}} = {}^t(\mathrm{A}\cdot\mathrm{X})\cdot\bar{\mathrm{X}} = \lambda\cdot{}^t\mathrm{X}\cdot\bar{\mathrm{X}}$, donc $\lambda = \bar\lambda$.

■

Démonstration du Théorème 3.3.— On procède par récurrence sur la dimension de E. Pour $n = 1$, il n'y a rien à dire. Supposons que ce soit vrai pour tout espace de dimension n, et que E est de dimension $n+1$. D'après le lemme précédent, il existe une valeur propre réelle λ, et un vecteur propre (normé) e_{n+1}. On note $\mathrm{F} = \mathbb{R}\mathrm{e}_{n+1}$. Cet espace est stable par u donc $\mathrm{F}^\perp$ aussi, qui est un supplémentaire de F. L'hypothèse de récurrence s'applique à $u|_{\mathrm{F}^\perp}$. Cette base se complète en une base orthonormée de E, et la matrice de u dans cette base est diagonale. ■

Corollaire 3.2. — *L'application* $\exp : \mathcal{S}_n \to \mathcal{S}_n^{++}$ *est un homéomorphisme.*

Démonstration. — L'application exp est continue, et si $\mathrm{M} \in \mathcal{S}_n(\mathbb{R})$, alors il existe $\mathrm{O} \in \mathrm{O}(n)$ telle que $\mathrm{O}^{-1}\mathrm{MO}$ soit diagonale. On en déduit que $\exp\mathrm{M} = \mathrm{O}(\exp\mathrm{O}^{-1}\mathrm{MO})\mathrm{O}^{-1} \in \mathcal{S}_n^{++}(\mathbb{R})$. Réciproquement, si $\mathrm{M} \in \mathcal{S}_n^{++}(\mathbb{R})$, il existe $\mathrm{O} \in \mathrm{O}(n)$ telle que $\mathrm{O}^{-1}\mathrm{MO}$ soit diagonale avec des valeurs propres strictement positives. On peut alors considérer la matrice N diagonale formée des logarithmes des valeurs propres de M. On a $\mathrm{ONO}^{-1} \in \mathcal{S}_n(\mathbb{R})$ et $\exp(\mathrm{ONO}^{-1}) = \mathrm{M}$.

Il reste à voir que l'application est injective (d'inverse continu). Tout d'abord, le théorème de réduction nous permet de diagonaliser une matrice $\mathrm{M} \in \mathcal{S}_n$. Sous cette forme, $\exp\mathrm{M}$ est aussi diagonale et M et $\exp\mathrm{M}$ ont la même décomposition en espaces propres, et les valeurs propres sont liées via l'exponentielle (numérique). Du coup, si $\exp\mathrm{M} = \exp\mathrm{N}$, alors la décomposition en sous-espaces propres nous permet de conclure que $\mathrm{M} = \mathrm{N}$.

Enfin, pour voir que l'application réciproque est continue, il suffit de montrer que exp est propre. Pour cela, on munit $\mathcal{S}_n$ de la norme associée à la forme $q(\mathrm{M}) = \mathrm{tr}\,{}^t\mathrm{MM}$. Restreinte à $\mathcal{S}_n$, elle prend la forme $q(\mathrm{M}) = \mathrm{tr}\,\mathrm{M}^2$, qui se traduit par la somme des carrés des valeurs propres de M. Par suite, si $\exp\mathrm{M}$ reste dans un compact de $\mathcal{S}_n^{++}$, les valeurs propres de $\exp\mathrm{M}$ restent dans un compact de $\mathbb{R}_+^*$, donc les valeurs propres restent dans un compact de $\mathbb{R}$, et il s'ensuit que M aussi reste dans un compact de $\mathcal{S}_n$. Du coup, on en déduit que exp est continue, propre et injective, donc un homéomorphisme sur son image.

■

Applications aux côniques et aux quadriques de $\mathbb{R}^2$ et $\mathbb{R}^3$. Une cônique est donnée par une équation de la forme $q(x,y)=1$, où q est un polynôme homogène de degré 2. Autrement dit, q est une forme quadratique. D'après le théorème 3.3, il existe une base orthonormée de $\mathbb{R}^2$ telle que P ait une forme canonique, qui nous donne la notion d'ellipse, d'hyperbole...

On définit une quadrique de $\mathbb{R}^3$ comme le lieu

$$\mathcal{Q}=\{(x,y,z)\in\mathbb{R}^3,\ q(x,y,z)=1\}$$

où q est une forme quadratique non dégénérée. On discute selon la signature de q de la forme de la quadrique.

sig(q)=(0,3) Dans une base adaptée de $\mathbb{R}^3$, on a $q(x,y,z)=-x^2-y^2-z^2$, donc $\mathcal{Q}=\varnothing$.

sig(q)=(1,2) Dans une base adaptée de $\mathbb{R}^3$, on a $q(x,y,z)=x^2-y^2-z^2$. Donc $\mathcal{Q}$ a deux composantes connexes selon que $x\geq 1$ ou $x\leq -1$. La quadrique coupe le plan $\{x=cste\}$, pour $|x|\geq 1$, en un cercle de rayon x^2-1. On dit que $\mathcal{Q}$ est un hyperboloïde à deux nappes.

sig(q)=(2,1) Dans une base adaptée de $\mathbb{R}^3$, on a $q(x,y,z)=x^2+y^2-z^2$. Donc $\mathcal{Q}$ est connexe. La quadrique coupe le plan $\{z=0\}$ en un cercle. On dit que $\mathcal{Q}$ est un hyperboloïde à une nappe.

Une propriété importante de cette quadrique est qu'elle est très exactement doublement réglées.

Un point appartient à $\mathcal{Q}$ si $(y-z)(y+z)=(1-x)(1+x)$. On devine l'équation de deux familles de droites incluses dans $\mathcal{Q}$.

$$\Delta_a\ \begin{cases} y-z=a(1-x)\\ (y+z)a=1+x\end{cases}\quad a\in\mathbb{R}\quad\text{et }\Delta_\infty\ \begin{cases} y=-z\\ x=1\end{cases}$$

ainsi que

$$\mathrm{D}_b\ \begin{cases} y+z=b(1-x)\\ (y-z)b=1+x\end{cases}\quad b\in\mathbb{R}\quad\text{et }\mathrm{D}_\infty\ \begin{cases} y=z\\ x=1\end{cases}$$

Un simple calcul montre que ces familles sont transverses, et que seule une droite par famille passe par un point donné de $\mathcal{Q}$.

De plus, si L est une droite incluse dans $\mathcal{Q}$ passant par un point p, et si v est un vecteur directeur, on a, pour tout $t\in\mathbb{R}$,

$$1=q(p+tv)=q(p)+2tb(p,v)+t^2q(v)\,,$$

donc $b(p,v)=0$ et $q(v)=0$. Ceci implique que $v\in(\mathbb{R}p)^\perp\cap\mathcal{C}(q)$ (il s'agit d'une équivalence). Or, puisque $q(p)\neq 0$, on a $\mathbb{R}^3=(\mathbb{R}p)\oplus(\mathbb{R}p)^\perp$; de plus $sig(q|_{\mathbb{R}p})=(1,0)$ donc $sig(q|_{(\mathbb{R}p)^\perp})=(1,1)$. Or on a vu qu'une forme définie sur un plan avec cette signature avait exactement deux droites isotropes.

sig(q)=(3,0) Dans une base adaptée de $\mathbb{R}^3$, on a $q(x,y,z)=x^2+y^2+z^2$. Donc $\mathcal{Q}$ est connexe. On dit que $\mathcal{Q}$ est un ellipsoïde. Dans une base orthonormée de $\mathbb{R}^3$, l'équation prend la forme plus générale

$$\left(\frac{x}{a}\right)^2+\left(\frac{y}{b}\right)^2+\left(\frac{z}{c}\right)^2=1\,.$$

On remarque qu'un ellipsoïde est la sphère unité pour une structure euclidienne appropriée.

On a

Théorème de John. — *Si* K *est un compact de* $\mathbb{R}^n$ *dont l'intérieur contient l'origine, alors il existe un unique ellipsoïde de volume minimal contenant* K.

En conséquence, si G est un sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$, il est conjugué à un sous-groupe de $\mathrm{O}(n,\mathbb{R})$ (voir le théorème 3.7, ainsi que [1]).

On note $\mathcal{S}_n(\mathbb{R})$ l'ensemble des matrices symétriques, $\mathcal{S}_n^+(\mathbb{R})$ celles qui sont positives et $\mathcal{S}_n^{++}(\mathbb{R})$ celles qui sont définies positives. Avant de démontrer le théorème de John, nous établissons quelques lemmes.

Lemme 3.2. — *Soit* $\mathcal{E}_\mathrm{S}=\{\mathrm{X}\in\mathbb{R}^n,\ {}^t\mathrm{XSX}\leq 1\}$ *où* $\mathrm{S}\in\mathcal{S}_n^{++}(\mathbb{R})$. *Alors* $\mathrm{vol}\,\mathcal{E}_\mathrm{S}=\mu(\mathrm{S})\,\mathrm{vol}\,\mathcal{E}_\mathrm{I}$ *où* $\mu:\mathrm{S}\in\mathcal{S}_n^{++}(\mathbb{R})\mapsto(\det\mathrm{S})^{-1/2}$.

DÉMONSTRATION. — D'après le théorème de réduction, il existe $\mathrm{O} \in \mathrm{O}_n(\mathbb{R})$ telle que ${}^t\mathrm{OSO} = \mathrm{D}$ soit une matrice diagonale dont les coefficients (diagonaux) $\lambda_1, \dots, \lambda_n$ sont tous strictement positifs. On considère D' la matrice diagonale dont les termes diagonaux sont $1/\sqrt{\lambda_j}$, et on note $\mathrm{R} = \mathrm{OD}'\mathrm{O}^{-1}$, qui est inversible et symétrique. On a donc $\mathrm{RSR} = \mathrm{I}_n$. De plus

$$\mathcal{E}_\mathrm{S} = \{\mathrm{X} \in \mathbb{R}^n,\ {}^t\mathrm{X}\mathrm{R}^{-2}\mathrm{X} \leq 1\} = \{\mathrm{X} \in \mathbb{R}^n,\ {}^t(\mathrm{R}^{-1}\mathrm{X})(\mathrm{R}^{-1}\mathrm{X}) \leq 1\} = \{\mathrm{X} \in \mathbb{R}^n,\ \mathrm{R}^{-1}(\mathrm{X}) \in \mathcal{E}_\mathrm{I}\} = \mathrm{R}(\mathcal{E}_\mathrm{I})\,.$$

Par la formule de changement de variables, on obtient le résultat recherché :

$$\mathrm{vol}\,\mathcal{E}_\mathrm{S} = \det \mathrm{R}\,\mathrm{vol}\,\mathcal{E}_\mathrm{I} = \mu(\mathrm{S})\,\mathrm{vol}\,\mathcal{E}_\mathrm{I}\,.$$

■

Lemme 3.3. — *Les espaces $\mathcal{S}_n^{++}(\mathbb{R})$ et $\mathcal{S}_n^{+}(\mathbb{R})$ sont convexes.*

DÉMONSTRATION. — On traite le cas de $\mathcal{S}_n^{++}(\mathbb{R})$. Soient $\mathrm{S}_0, \mathrm{S}_1 \in \mathcal{S}_n^{++}(\mathbb{R})$. On note, pour $s \in [0,1]$, $\mathrm{S}_s = (1-s)\mathrm{S}_0 + s\mathrm{S}_1$. On a $\mathrm{S}_s \in \mathcal{S}_n(\mathbb{R})$, et, pour tout $\mathrm{X} \in \mathbb{R}^n \setminus \{0\}$,

$${}^t\mathrm{X}\mathrm{S}_s\mathrm{X} = (1-s)({}^t\mathrm{X}\mathrm{S}_0\mathrm{X}) + s({}^t\mathrm{X}\mathrm{S}_1\mathrm{X}) > 0\,.$$

■

Lemme 3.4. — *L'application $\mu : \mathrm{S} \in \mathcal{S}_n^{++}(\mathbb{R}) \mapsto (\det \mathrm{S})^{-1/2}$ est strictement convexe.*

DÉMONSTRATION. — Soient $\mathrm{S}_0, \mathrm{S}_1 \in \mathcal{S}_n^{++}(\mathbb{R})$ distinctes. On note, pour $s \in [0,1]$, $\mathrm{S}_s = (1-s)\mathrm{S}_0 + s\mathrm{S}_1$. D'après le théorème de réduction, il existe $\mathrm{P} \in \mathrm{GL}_n(\mathbb{R})$ telle que ${}^t\mathrm{P}\mathrm{S}_0\mathrm{P} = \mathrm{I}_n$ et ${}^t\mathrm{P}\mathrm{S}_1\mathrm{P} = \mathrm{D}$ soit une matrice diagonale de coefficients (diagonaux) $\lambda_1, \dots, \lambda_n$. Si $\mathrm{D} = \mathrm{I}_n$ alors on aurait ${}^t\mathrm{P}\mathrm{S}_0\mathrm{P} = {}^t\mathrm{P}\mathrm{S}_1\mathrm{P}$ et $\mathrm{S}_0 = \mathrm{S}_1$, ce qui est contraire à l'hypothèse. Donc $\mathrm{D} \neq \mathrm{I}_n$, et on peut donc supposer que $\lambda_1 \neq 1$.

Du coup,

$$\det \mathrm{S}_s = \frac{1}{\det^2 \mathrm{P}} \det({}^t\mathrm{P}\mathrm{S}_s\mathrm{P}) = \frac{1}{\det^2 \mathrm{P}} \det((1-s)\mathrm{I}_n + s\mathrm{D}) = \frac{1}{\det^2 \mathrm{P}} \prod[(1-s) + s\lambda_j]\,.$$

Posons $\mathrm{A}_j(s) = (1-s) + s\lambda_j$ et $u(s) = \mu(\mathrm{S}_s)/|\det \mathrm{P}|$. Cette application est différentiable, et

$$u'(s) = \sum_{j=1}^n \frac{-1}{2} \frac{\lambda_j - 1}{\mathrm{A}_j(s)^{3/2}} \frac{1}{\prod_{i \neq j} \mathrm{A}_i(s)^{1/2}} = \frac{-1}{2} u(s) \sum_{j=1}^n \frac{\lambda_j - 1}{\mathrm{A}_j(s)}$$

et

$$u''(s) = \frac{1}{4} u(s) \left(\sum_{j=1}^n \frac{\lambda_j - 1}{\mathrm{A}_j(s)} \right)^2 + \frac{1}{2} u(s) \sum_{j=1}^n \left(\frac{\lambda_j - 1}{\mathrm{A}_j(s)} \right)^2 \geq \frac{1}{2} u(s) \left(\frac{\lambda_1 - 1}{\mathrm{A}_1(s)} \right)^2 > 0\,.$$

■

Nous pouvons maintenant nous atteler à la démonstration du théorème de John.

DÉMONSTRATION. — Par hypothèse, il existe $\rho_1, \rho_2 > 0$ tel que $\mathrm{B}(0,\rho_1) \subset \mathrm{K} \subset \mathrm{B}(0,\rho_2)$.

Assertion. L'ensemble

$$\mathcal{C} = \{\mathrm{S} \in \mathcal{S}_n^{++}(\mathbb{R}),\ \det \mathrm{S} \geq 1/\rho_2^{2n},\ \mathrm{K} \subset \mathcal{E}_\mathrm{S}\}$$

est convexe compact non vide.

Si cette assertion est vraie, alors le minimum de μ est atteint en un unique point, qui définit un ellipsoïde de volume minimal (si $\det \mathrm{S} \leq 1/\rho_2^{2n}$, alors $\mu(\mathrm{S}) \geq \rho_2^n$, donc ne peut être minimal). Montrons donc cette assertion.

On a

$$\mathcal{E}_{(1/\rho_2^2)\mathrm{I}_n} = \{(1/\rho_2^2)\,{}^t\mathrm{X}\mathrm{X} \leq 1\} = \mathrm{B}(0,\rho_2)$$

donc $(1/\rho_2^2)\mathrm{I}_n \in \mathcal{C}$.

On note $f : \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$ définie par $f(\mathrm{M}) = \max_{\mathrm{X}\in\mathrm{K}}({}^t\mathrm{XMX})$. Cette application est continue et on a

$$\mathcal{C} = \mathcal{S}_n^+(\mathbb{R}) \cap \det{}^{-1}[1/\rho_2^{2n}, \infty[\cap f^{-1}[0,1])$$

donc $\mathcal{C}$ est fermé.

Soit $\mathrm{S} \in \mathcal{C}$, il existe une base orthonormée (e_j) de vecteurs propres de S. Puisque $\mathrm{B}(0,\rho_1) \subset \mathrm{K}$, on a $\mathrm{B}(0,\rho_1) \subset \mathcal{E}_\mathrm{S}$ donc $\rho_1 \mathrm{e}_j \in \mathcal{E}_\mathrm{S}$, soit

$$\langle \mathrm{S}(\rho_1 \mathrm{e}_j), \rho_1 \mathrm{e}_j \rangle = \rho_1^2 \lambda_j \leq 1$$

donc $\lambda_j \leq 1/\rho_1^2$. On en déduit que $\mathcal{C}$ est borné.

Si $\mathrm{S}_0, \mathrm{S}_1 \in \mathcal{C}$ et $\mathrm{S}_s = (1-s)\mathrm{S}_0 + s\mathrm{S}_1$ alors, pour $s \in [0,1]$, on a

$$\begin{cases} f(\mathrm{S}_s) \leq s f(\mathrm{S}_1) + (1-s) f(\mathrm{S}_0) \leq 1 \\ \mu(\mathrm{S}_s) \leq s\mu(\mathrm{S}_1) + (1-s)\mu(\mathrm{S}_0) \leq \rho_2^n \end{cases}$$

donc $\mathrm{S}_s \in \mathcal{C}$ et $\mathcal{C}$ est bien convexe. ■

3.2 Endomorphismes normaux

Définition 3.5. — *Un endomorphisme est normal s'il commute avec son adjoint.*

Les exemples proviennent, dans une base orthonormée, des matrices symétriques, antisymétriques et orthogonales. En petite dimension, il est facile de les caractériser.

Proposition 3.2. — *Si u est un endomorphisme normal d'un plan euclidien sans valeurs propres réelles, alors il existe $\lambda \in \mathbb{R}_+$ et $\theta \in \mathbb{R} \setminus 2\pi\mathbb{Z}$ tels que $u = \lambda \cdot \mathrm{R}_\theta$ où R_θ désigne la rotation d'angle θ. Sinon, u est diagonalisable en base orthonormée.*

Démonstration. — On écrit u dans une base orthonormée :

$$\mathrm{Mat}(u) = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

On traduit sur ses coefficients le fait d'être normal : on obtient le système

$$\begin{cases} (c-b)(c+b) = 0 \\ (a-d)(b-c) = 0 \end{cases}$$

Sachant que u n'a pas de valeurs propres réelles, on sait aussi que $b \neq c$, car alors A est symétrique, donc diagonalisable dans une base orthonormée. En particulier, $b, c \neq 0$.

Du coup, $b = -c$, et donc $a = d$ puisque $b, c \neq 0$. Posons $\lambda^2 = a^2 + b^2$. Il existe alors θ tel que $a = \lambda\cos\theta$ et $b = \lambda\sin\theta$. Du coup,

$$\mathrm{Mat}(u) = \lambda \cdot \begin{pmatrix} a/\lambda & -b/\lambda \\ b/\lambda & a/\lambda \end{pmatrix} = \lambda \cdot \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Si u a une valeur propre réelle alors u est diagonalisable. En effet, on a vu que $b = -c$ entraînait que u était de la forme $\lambda\mathrm{R}_\theta$. Sinon, u est symétrique, donc diagonalisable en base orthonormée. ■

Proposition 3.3. — *Soit u un endomorphisme normal.*

1. *On a l'égalité* $\operatorname{Ker} u = \operatorname{Ker} u^*$.
2. *u et u^* ont mêmes valeurs propres et espaces propres. Ceux-ci sont orthogonaux.*
3. *u fixe au moins une droite ou un plan.*
4. *Si* F *est stable par u, alors* F *est aussi stable par u^*, et* $\mathrm{F}^\perp$ *est stable par u et u^*.*

Démonstration. —

1. Si u est normal, alors, pour tout $x \in \mathrm{E}$, on a $\|u(x)\| = \|u^*(x)\|$. En effet,

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

Donc $u(x)$ et $u^*(x)$ s'annulent simultanément.

2. provient du fait que $u+\lambda\mathrm{I}$ est normal si u est normal, donc $\mathrm{Ker}\,(u+\lambda\mathrm{Id}) = \mathrm{Ker}\,(u^*+\lambda\mathrm{Id})$ d'après ci-dessus. Soient λ, μ deux valeurs propres distinctes, et x, y deux vecteurs propres associés : $u(x) = \lambda x$ et $u(y) = \mu y$. Il vient

$$\lambda\langle x, y\rangle = \langle u(x), y\rangle = \langle x, u^*(y)\rangle = \mu\langle x, y\rangle .$$

Du coup $(\lambda - \mu)\langle x, y\rangle = 0$ et $\langle x, y\rangle = 0$.

3. Si u n'a pas de valeurs propres réelles, alors son polynôme caractéristique χ_u s'écrit

$$\chi_u(\lambda) = \prod_{j=1}^{k} (\lambda^2 + a_j\lambda + b_j) .$$

Or, le théorème de Cayley-Hamilton implique que $\chi_u(u) = 0$. Donc il existe un indice j tel que l'endomorphisme $u^2 + a_j u + b_j\mathrm{Id}$ n'est pas injectif. Si $x \in \mathrm{Ker}\,(u^2 + a_j u + b_j\mathrm{Id})$, alors $u^2(x) = (-b_j)x + (-a_j)u(x)$. Comme u n'a pas de valeur propre réelle, $(x, u(x))$ engendre un plan. Il est stable puisque $u^2(x)$ est une combinaison linéaire de x et de $u(x)$.

4. On considère une base orthonormée de E telle que les premiers $p = \dim \mathrm{F}$ vecteurs forment une base de F et les derniers une base de $\mathrm{F}^\perp$. Dire que F est stable par u signifie qu'il existe des matrices $\mathrm{A} \in \mathcal{M}_p(\mathbb{R})$, $\mathrm{B} \in \mathcal{M}_{p,n-p}(\mathbb{R})$ et $\mathrm{C} \in \mathcal{M}_{n-p}(\mathbb{R})$ telles que la matrice de u s'écrive

$$\mathrm{Mat}(u) = \begin{pmatrix} \mathrm{A} & \mathrm{B} \\ 0 & \mathrm{C} \end{pmatrix}$$

Dire que u est normale signifie que $\mathrm{A}\cdot {}^t\mathrm{A} + \mathrm{B}\cdot {}^t\mathrm{B} = {}^t\mathrm{A}\cdot\mathrm{A}$. Or, un simple calcul montre que $\mathrm{tr}(\mathrm{A}\cdot {}^t\mathrm{A}) = \mathrm{tr}({}^t\mathrm{A}\cdot\mathrm{A})$, donc $\mathrm{tr}(\mathrm{B}\cdot {}^t\mathrm{B}) = 0$. Mais, si $\mathrm{B} = (b_{ij})_{1\le i\le p, 1\le j\le n-p}$, alors

$$\mathrm{tr}(\mathrm{B}\cdot {}^t\mathrm{B}) = \sum_{j=1}^{p}\left(\sum_{k=1}^{n-p} b_{jk}b_{jk}\right) = \sum_{j=1}^{p}\sum_{k=1}^{n-p} b_{jk}^2 .$$

Donc $b_{jk} = 0$ pour tout j, k et $\mathrm{B} = 0$. Du coup,

$$\mathrm{Mat}(u) = \begin{pmatrix} \mathrm{A} & 0 \\ 0 & \mathrm{C} \end{pmatrix}$$

avec A et C normaux. Ceci montre que F et $\mathrm{F}^\perp$ sont stables par u et u^*.

■

Théorème 3.4. — *Soit u un endomorphisme normal. Il existe $\alpha_1, \dots \alpha_p$, $\rho_1, \dots, \rho_q$ et $\theta_1, \dots, \theta_q$ et une base orthonormée dans laquelle*

$$\mathrm{Mat}(u) = \begin{pmatrix} \alpha_1 & & & & & & \\ & \ddots & & & 0 & & \\ & & \alpha_p & & & & \\ & & & \rho_1\cdot \mathrm{R}_{\theta_1} & & & \\ & 0 & & & \ddots & & \\ & & & & & \rho_q\cdot \mathrm{R}_{\theta_q} & \end{pmatrix}$$

Démonstration. — On raisonne par récurrence sur la dimension. Les cas de la dimension 1 et 2 sont faits. On suppose que le théorème est vrai jusqu'à la dimension n et on suppose que E est de dimension $n+1$. D'après le lemme, u fixe une droite ou un plan F. Dans cet espace, on sait trouver une base orthonormée qui donnera la forme voulue pour $u|_{\mathrm{F}}$. Ensuite, on sait, toujours par le lemme, que u et u^* préservent $\mathrm{F}^\perp$, donc $u|_{\mathrm{F}^\perp}$ est aussi normal. On peut appliquer l'hypothèse de récurrence. ■

Application aux formes antisymétriques. On se donne une forme bilinéaire antisymétrique non dégénérée b sur $\mathbb{R}^{2n}$. D'après la Proposition 3.1, il existe $u \in \mathrm{End(E)}$ telle que $b(x,y) = \langle x, u(y)\rangle$. On a donc

$$\langle x, u(y)\rangle = b(x,y) = -b(y,x) = -\langle y, u(x)\rangle = \langle -u(x), y\rangle\,.$$

Donc u est antisymétrique ($u = -u^*$). Il s'ensuit que u ne peut avoir de valeurs propres réelles, donc le théorème nous fournit une matrice $\mathrm{O} \in \mathrm{O}(2n)$, des nombres $\rho_1, \dots, \rho_n$ et $\theta_1, \dots, \theta_n$ tels que

$$\mathrm{Mat}(u) = \mathrm{O}^{-1} \cdot \begin{pmatrix} \rho_1 \cdot \mathrm{R}_{\theta_1} & & 0 \\ & \ddots & \\ 0 & & \rho_n \cdot \mathrm{R}_{\theta_n} \end{pmatrix} \cdot \mathrm{O}$$

Comme u est antisymétrique et $\mathrm{O} \in \mathrm{O}(2n)$, cette matrice doit aussi être antisymétrique. Autrement dit, $\theta_j = \pm\pi/2$ pour tout indice j, et, quitte à réordonner les paires convenablement, on obtient

$$\mathrm{Mat}(u) = \mathrm{O}^{-1} \cdot \begin{pmatrix} \rho_1 \mathrm{J} & & 0 \\ & \ddots & \\ 0 & & \rho_n \mathrm{J} \end{pmatrix} \cdot \mathrm{O}$$

où

$$\mathrm{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

ou encore, en renumérotant les vecteurs dans la base donnée par O,

$$\mathrm{Mat}(u) = \mathrm{O}^{-1} \cdot \begin{pmatrix} 0 & & -\mathrm{I} \\ & \ddots & \\ \mathrm{I} & & 0 \end{pmatrix} \cdot \mathrm{O}$$

Puisque O est orthogonale, cette matrice est aussi l'expression de b dans la base donnée par O.

3.3 Le groupe orthogonal euclidien

On suppose que E est un espace euclidien muni d'une base orthonormée $\mathcal{B}$.

Théorème 3.5. — *Soit u un endomorphisme de* E. *Les propriétés suivantes sont équivalentes.*

- $u \in \mathrm{O(E)}$ *;*
- *pour tout $x \in \mathrm{E}$, on a $\|u(x)\| = \|x\|$;*
- $u \circ u^* = \mathrm{Id}$ *;*
- $u^* \circ u = \mathrm{Id}$ *;*
- $\mathrm{Mat}(u, \mathcal{B}) \cdot {}^t\mathrm{Mat}(u, \mathcal{B}) = \mathrm{I}$ *;*
- ${}^t\mathrm{Mat}(u, \mathcal{B}) \cdot \mathrm{Mat}(u, \mathcal{B}) = \mathrm{I}$ *;*
- *les colonnes de* $\mathrm{Mat}(u, \mathcal{B})$ *forment une base orthonormée ;*
- *les lignes de* $\mathrm{Mat}(u, \mathcal{B})$ *forment une base orthonormée ;*
- *u transforme une base orthonormée en une base orthonormée.*

La démonstration est laissée en exercice.

Corollaire 3.3. — *Si $u \in \mathrm{O(E)}$, alors* $\det u = \pm 1$ *et donc* $\mathrm{SO(E)} = \det^{-1}\{1\} \cap \mathrm{O(E)}$ *est distingué d'indice* 2.

Remarque. — Si $u \in \mathrm{O(E)}$, alors u est normal, donc le théorème de réduction s'applique, et on obtient

Théorème 3.6. — *Soit u un endomorphisme orthogonal de* E. *Il existe une base orthonormée telle que*

$$\mathrm{Mat}(u) = \begin{pmatrix} \mathrm{I}_p & & & 0 & \\ & -\mathrm{I}_q & & & \\ & & \mathrm{R}_{\theta_1} & & \\ & 0 & & \ddots & \\ & & & & \mathrm{R}_{\theta_m} \end{pmatrix}$$

Si $u \in \mathrm{SO(E)}$, alors q est paire.

Corollaire 3.4. — *L'application* $\exp : \mathcal{A}_n(\mathbb{R}) \to \mathrm{SO(E)}$ *est surjective, où* $\mathcal{A}_n(\mathbb{R})$ *désigne l'ensemble des matrices antisymétriques.*

Démonstration. — Si A est antisymétrique, alors ${}^t\mathrm{A} = -\mathrm{A}$. Pour toute matrice A, $\exp(-\mathrm{A}) = (\exp \mathrm{A})^{-1}$ et $\exp({}^t\mathrm{A}) = {}^t(\exp \mathrm{A})$. Du coup, si A est antisymétrique, alors ${}^t\mathrm{A} = -\mathrm{A}$ et $(\exp \mathrm{A})^{-1} = {}^t(\exp \mathrm{A})$, donc $\exp(\mathcal{A}_n) \subset \mathrm{O(E)}$. Or $\det \exp \mathrm{A} = \exp \mathrm{tr}\, \mathrm{A}$ donc, pour $\mathrm{A} \in \mathcal{A}_n$, on a $\det \exp \mathrm{A} = 1$: $\exp \mathrm{A} \in \mathrm{SO(E)}$.

Soit

$$\mathrm{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

On a $\mathrm{J}^2 = -\mathrm{I}$, et, pour $\theta \in \mathbb{R}$, on obtient

$$\exp(\theta \mathrm{J}) = \cos\theta \cdot \mathrm{I} + \sin\theta \cdot \mathrm{J} = \mathrm{R}_\theta \,.$$

Par conséquent, si

$$\mathrm{A} = \begin{pmatrix} 0 & & & & & \\ & \ddots & & & 0 & \\ & & 0 & & & \\ & & & \theta_1 \cdot \mathrm{J} & & \\ & 0 & & & \ddots & \\ & & & & & \theta_q \cdot \mathrm{J} \end{pmatrix}$$

$$\exp \mathrm{A} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ & & & \mathrm{R}_{\theta_1} & & \\ & 0 & & & \ddots & \\ & & & & & \mathrm{R}_{\theta_q} \end{pmatrix}$$

Ceci permet de montrer la surjectivité. ∎

Cas de la dimension 2. Soit $\mathcal{B}$ une base orthonormée d'un espace euclidien de dimension 2. Si $u \in \mathrm{SO(E)}$, alors il existe $\theta \in \mathbb{R}$ tel que $\mathrm{Mat}(u, \mathcal{B}) = \mathrm{R}_\theta$. Si $u \in \mathrm{O(E)} \setminus \mathrm{SO(E)}$, alors il existe une base appropriée telle que

$$\mathrm{Mat}(u) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

En effet, si $u \in \mathrm{SO(E)}$, alors la Proposition 3.2 implique que, ou bien la matrice est symétrique et diagonalisable, donc $u = \mathrm{Id}$, ou bien $u = \lambda \mathrm{R}_\theta$, avec $\lambda^2 = 1$. On peut alors choisir $\lambda = 1$, quitte à changer θ en $\theta + \pi$. Sinon, on a $\det u = -1$, ce qui implique par la Proposition 3.2 que u est diagonalisable, avec valeurs propres ± 1.

Proposition 3.4. — $\mathrm{SO}(2)$ *est commutatif et* $\mathrm{SO}(2)$ *est isomorphe à* $\mathrm{U}(1)$.

Démonstration. — On définit $\varphi : \mathbb{R} \to \mathrm{SO}(2)$ par $\varphi(\theta) = \mathrm{R}_\theta$. On a $\mathrm{Ker}\, \varphi = 2\pi\mathbb{Z}$, donc $\mathrm{SO}(2) \approx \mathbb{R}/2\pi\mathbb{Z}$. On a

$$\mathrm{U}(1) = \{u \in \mathrm{End}(\mathbb{C}),\ \|u(z)\| = \|z\|\}\,.$$

Du coup, il existe $\lambda \in \mathbb{S}^1$ tel que $u(z) = \lambda z$. ∎

Du coup, pour tout $u \in \mathrm{SO}(2)$, il existe une matrice M telle que, quel que soit la base orthonormée orientée considérée, la matrice de u est M.

Cas de la dimension 3. Pour tout $u \in \mathrm{SO}(3)$, il existe une base orthonormée $\mathcal{B}$ telle que la matrice de u soit l'une des possibilités suivantes :

$$\mathrm{Mat}(u, \mathcal{B}) = \mathrm{I}_3 \,,$$

ou

$$\mathrm{Mat}(u, \mathcal{B}) = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{R}_\theta \end{pmatrix} ;$$

on a $\operatorname{tr} u = 1 + 2\cos\theta$, donc θ est déterminé par u ; on dit que u est une rotation d'axe e_1 ;

si $u \in \mathrm{O}(3) \setminus \mathrm{SO}(3)$, alors ou bien

$$\mathrm{Mat}(u, \mathcal{B}) = \begin{pmatrix} \mathrm{I} & 0 \\ 0 & -1 \end{pmatrix} ;$$

on parle alors de réflexion par rapport à $\langle \mathrm{e}_1, \mathrm{e}_2 \rangle$;

ou bien

$$\mathrm{Mat}(u, \mathcal{B}) = \begin{pmatrix} -1 & 0 \\ 0 & \mathrm{R}_\theta \end{pmatrix} .$$

Propriétés topologiques. On s'intéresse aux propriétés topologiques du groupe orthogonal.

Lemme 3.5. — $\mathrm{O}_n(\mathbb{R})$ *est compact.*

Démonstration. — On considère l'application $f : \mathrm{M} \in \mathcal{M}_n(\mathbb{R}) \mapsto {}^t\mathrm{MM}$. On a $\mathrm{O}_n(\mathbb{R}) = f^{-1}(\mathrm{I}_n)$, donc $\mathrm{O}_n(\mathbb{R})$ est fermé. Pour montrer que $\mathrm{O}(n)$ est borné, on considère la norme induite par $\operatorname{tr}({}^t\mathrm{MM})$. On a $\mathrm{O}_n(\mathbb{R}) \subset \mathrm{B}(0, \sqrt{n})$. ∎

Proposition 3.5. — *Le groupe* $\mathrm{O(E)}$ *a exactement deux composantes connexes qui sont homéormorphes :* $\mathrm{SO(E)}$ *et* $\mathrm{O(E)}^- = \mathrm{O(E)} \setminus \mathrm{SO(E)}$.

Démonstration. — On montre d'abord que SO(E) est connexe par arcs. Puisque q est pair, on peut remplacer $-\mathrm{I}_q$ par une matrice par blocs de rotations d'angle π. Du coup,

$$\mathrm{Mat}(u) = \begin{pmatrix} \mathrm{I}_p & & 0 & \\ & \mathrm{R}_{\theta_1} & & \\ & & \ddots & \\ & 0 & & \mathrm{R}_{\theta_q} \end{pmatrix}$$

Pour $t \in [0, 1]$, on pose

$$\mathrm{Mat}(u_t) = \begin{pmatrix} \mathrm{I}_p & & 0 & \\ & \mathrm{R}_{t\theta_1} & & \\ & & \ddots & \\ & 0 & & \mathrm{R}_{t\theta_q} \end{pmatrix}$$

On a $u_0 = \mathrm{Id}$ et $u_1 = u$. On vérifie que chaque $u_t \in \mathrm{SO(E)}$.

On considère l'application ι définie en base orthonormée par

$$\mathrm{Mat}(\iota) = \begin{pmatrix} -1 & & 0 & \\ 0 & 1 & & \\ & & \ddots & \\ & 0 & & 1 \end{pmatrix}$$

La multiplication à gauche par cette matrice définit un homéomorphisme entre SO(E) et $\mathrm{O}^-(\mathrm{E})$. Enfin, l'application $\det : \mathrm{O(E)} \to \{\pm 1\}$ montre que O(E) n'est pas connexe. ∎

Décomposition polaire. L'application $\Phi : \mathrm{O}(n) \times \mathcal{S}_n^{++}(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$ définie par $\Phi(\mathrm{O}, \mathrm{S}) = \mathrm{OS}$ est un homéomorphisme.

Lemme 3.6. — *Si* $\mathrm{M} \in \mathrm{GL}_n(\mathbb{R})$*, alors* ${}^t\mathrm{M} \cdot \mathrm{M}$ *est définie positive.*

Démonstration. — On considère une base orthonormée de $\mathbb{R}^n$ et on considère M comme la matrice d'un endomorphisme u (inversible) dans cette base. On a, pour $x \neq 0$,

$$\langle x, u^* \circ u(x) \rangle = \langle u(x), u(x) \rangle > 0\,.$$

Idée Si $\mathrm{M} = \mathrm{OS}$, alors ${}^t\mathrm{M} = \mathrm{SO}^{-1}$ donc ${}^t\mathrm{M} \cdot \mathrm{M} = \mathrm{S}^2$, avec S définie positive.

Démonstration de la décomposition : existence. — D'après le lemme, il existe $\widehat{\mathrm{O}} \in \mathrm{O}(n)$ telle que

$$\widehat{\mathrm{O}}^{-1} \cdot ({}^t\mathrm{M} \cdot \mathrm{M}) \cdot \widehat{\mathrm{O}} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

soit diagonale avec des valeurs propres strictement positives. On note

$$\mathrm{S} = \widehat{\mathrm{O}} \cdot \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} \cdot \widehat{\mathrm{O}}^{-1}$$

et $\mathrm{O} = \mathrm{MS}^{-1}$. Comme S est symétrique, S^{-1} aussi, donc

$${}^t\mathrm{O} \cdot \mathrm{O} = {}^t(\mathrm{S}^{-1}) \cdot ({}^t\mathrm{M} \cdot \mathrm{M}) \cdot \mathrm{S}^{-1} = \mathrm{S}^{-1}\mathrm{S}^2\mathrm{S}^{-1} = \mathrm{I}\,.$$

■

Démonstration de la décomposition : unicité. — Soit $\mathrm{P} \in \mathbb{R}[\mathrm{X}]$ tel que $\mathrm{P}(\lambda_j) = \sqrt{\lambda_j}$. On a

$$\mathrm{P}({}^t\mathrm{M} \cdot \mathrm{M}) = \sum a_k \widehat{\mathrm{O}} \cdot \begin{pmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix} \cdot \widehat{\mathrm{O}}^{-1} = \widehat{\mathrm{O}} \cdot \begin{pmatrix} \mathrm{P}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & \mathrm{P}(\lambda_n) \end{pmatrix} \cdot \widehat{\mathrm{O}}^{-1} = \mathrm{S}\,.$$

Donc, si M admet une seconde décomposition polaire $\mathrm{M} = \mathrm{O}'\mathrm{S}'$ alors ${}^t\mathrm{M} \cdot \mathrm{M} = (\mathrm{S}')^2$ donc S' commute avec ${}^t\mathrm{M} \cdot \mathrm{M}$, donc avec S puisque S est un polynôme en ${}^t\mathrm{M} \cdot \mathrm{M}$. Du coup, S et S' peuvent être diagonalisé simultanément (en effet, $\mathrm{E}_{\lambda_i}(\mathrm{S}) = \oplus(\mathrm{E}_{\lambda_i}(\mathrm{S}) \cap \mathrm{E}_{\mu_j}(\mathrm{S}'))$). Ceci oblige les valeurs propres à coïncider, donc ces matrices sont les mêmes. ■

Démonstration de la décomposition : homéomorphie. — L'application Φ est clairement continue. Réciproquement, si $\mathrm{M}_n = \Phi(\mathrm{O}_n, \mathrm{S}_n)$ converge vers une matrice $\mathrm{M} = \mathrm{OS}$, montrons que O_n tend vers O et S_n vers S. Comme O(E) est compact, quitte à extraire une sous-suite, on peut supposer que O_n tend vers une matrice orthogonale $\widehat{\mathrm{O}}$. Du coup, S_n tend vers une matrice (symétrique) $\widehat{\mathrm{O}}^{-1}\mathrm{M}$. Du coup, on a $\mathrm{OS} = \widehat{\mathrm{O}}(\widehat{\mathrm{O}}^{-1}\mathrm{M})$. Par l'unicité de la décomposition polaire, on obtient $\mathrm{O} = \widehat{\mathrm{O}}$ et S_n tend vers S. ■

On en déduit quelques résultats.

Proposition 3.6. — $\mathrm{GL}_n(\mathbb{R})$ *a exactement deux composantes connexes.*

Démonstration. — On a

$$\mathrm{GL}_n(\mathbb{R}) \approx \mathrm{O}_n(\mathbb{R}) \times \mathcal{S}_n^{++}(\mathbb{R}) = (\mathrm{SO}_n(\mathbb{R}) \times \mathcal{S}_n^{++}(\mathbb{R})) \cup (\mathrm{O}_n^-(\mathbb{R}) \times \mathcal{S}_n^{++}(\mathbb{R}))$$

par la décomposition polaire. ■

Proposition 3.7. — $\mathrm{SO}_n(\mathbb{R})$ *(resp.* $\mathrm{O}_n(\mathbb{R})$*) est un sous-groupe compact maximal de* $\mathrm{SL}_n(\mathbb{R})$ *(resp.* $\mathrm{GL}_n(\mathbb{R})$*).*

Démonstration. — Soit $\mathrm{G} \subset \mathrm{SL}_n(\mathbb{R})$ un sous-groupe compact qui contient $\mathrm{SO}_n(\mathbb{R})$, et soit $g \in \mathrm{G} \setminus \mathrm{SO}_n(\mathbb{R})$. On a $g = \mathrm{OS}$ avec $\mathrm{O} \in \mathrm{SO}_n(\mathbb{R})$ car $\det g > 0$ et $\mathrm{S} \neq \mathrm{I}$ car $g \notin \mathrm{SO}_n(\mathbb{R})$. Donc $\mathrm{S} \in \mathrm{G}$. Or si v est un vecteur propre associé à une valeur propre λ de S différente de 1, alors $\mathrm{Log}\,\|\mathrm{S}^n x\|$ tend vers l'infini, ce qui contredit la compacité de G. ∎

On montre que l'on peut améliorer ce résultat de la manière suivante.

Théorème 3.7. — *Un sous-groupe compact* G *de* $\mathrm{GL}_n(\mathbb{R})$ *est conjugué à un sous-groupe de* $\mathrm{O}_n(\mathbb{R})$*.*

Démonstration. — On remarque tout d'abord que si $\mathcal{E}$ est un ellipsoïde et si $\mathrm{M} \in \mathrm{GL}_n(\mathbb{R})$, alors $\mathrm{M}\mathcal{E}$ est aussi un ellipsoïde. En effet si $\mathcal{E} = \{{}^t\mathrm{XSX} = 1\}$ avec $\mathrm{S} \in \mathcal{S}_n^{++}(\mathbb{R})$, alors

$$\mathrm{M}\mathcal{E} = \{{}^t(\mathrm{M}^{-1}\mathrm{X})\mathrm{SM}^{-1}\mathrm{X} = 1\} = \{{}^t\mathrm{X}({}^t\mathrm{M}^{-1}\mathrm{SM}^{-1})\mathrm{X} = 1\}$$

Or ${}^t\mathrm{M}^{-1}\mathrm{SM}^{-1}$ est aussi définie positive car il ne s'agit que d'un changement de variables, donc $\mathrm{M}\mathcal{E}$ est aussi un ellipsoïde.

Soit B la boule unité fermée de $\mathbb{R}^n$. On note $\mathrm{K} = \cup_{g\in\mathrm{G}} g(\mathrm{B})$. Alors K est compact car G et B le sont, K est invariant par définition, et 0 est un point intérieur de K car K contient $\mathrm{I}(\mathrm{B}) = \mathrm{B}$. D'après le théorème de John, il existe un unique ellipsoïde $\mathcal{E}_\mathrm{S}$ qui contient K de volume minimal.

Puisque G est compact, on a, pour tout $g \in \mathrm{G}$, $|\det g| = 1$. Donc $\mathrm{vol}\, g(\mathcal{E}_\mathrm{S}) = \mathrm{vol}\,\mathcal{E}_\mathrm{S}$, et comme $\mathrm{K} = g(\mathrm{K}) \subset g(\mathcal{E}_\mathrm{S})$, on obtient $g(\mathcal{E}_\mathrm{S}) = \mathcal{E}_\mathrm{S}$. Par suite, $\mathcal{E}_\mathrm{S}$ est invariant par G. Soit T une racine carrée de S^{-1}. Alors $\mathcal{E}_\mathrm{S} = \mathrm{T}(\mathrm{B})$, et $\mathrm{TGT}^{-1} \subset \mathrm{O}_n(\mathbb{R})$. ∎

Voici un autre argument. Puisque G est un groupe compact métrique, il existe une mesure de Haar *i.e.*, une mesure de probabilité μ borélienne sur G invariante par multiplication à gauche et à droite dans G. Autrement dit, si $\varphi : \mathrm{G} \to \mathbb{R}$ est une fonction intégrable, alors

$$\int_\mathrm{G} \varphi(g)d\mu(g) = \int_\mathrm{G} \varphi(hg)d\mu(g) = \int_\mathrm{G} \varphi(gh)d\mu(g)$$

pour tout $h \in \mathrm{G}$. On définit sur $\mathbb{R}^n$ la forme

$$\langle x, y\rangle_\mathrm{G} = \int_\mathrm{G} \langle gx, gy\rangle d\mu(g)\,.$$

Il s'agit clairement d'une forme bilinéaire symétrique positive. Si $\langle x, x\rangle_\mathrm{G} = 0$ alors il existe $g \in \mathrm{G}$ tel que $\|g(x)\| = 0$, donc $x = 0$. Du coup, $\langle\cdot,\cdot\rangle_\mathrm{G}$ est un produit scalaire sur $\mathbb{R}^n$ représenté par une matrice symétrique M.

Or, si $h \in \mathrm{G}$, l'invariance de μ implique que

$$\langle h(x), h(y)\rangle_\mathrm{G} = \int_\mathrm{G} \langle ghx, ghy\rangle d\mu(g) = \int_\mathrm{G} \langle gx, gy\rangle d\mu(g) = \langle x, y\rangle_\mathrm{G}$$

donc h est une isométrie pour cette structure euclidienne.

Par le théorème de Sylvester, il existe une matrice $\mathrm{P} \in \mathrm{GL}_n(\mathbb{R})$ telle que ${}^t\mathrm{PMP} = \mathrm{I}_n$. Du coup, $\mathrm{PGP}^{-1} \subset \mathrm{O}_n(\mathbb{R})$.

Remarque. — Il est facile de voir que $\mathrm{O}_2(\mathbb{C})$ n'est pas compact, donc $\mathrm{O}_n(\mathbb{C})$ non plus.

Proposition 3.8. — *Les espaces* $\mathrm{GL}_n(\mathbb{R})$ *et* $\mathrm{SL}_n(\mathbb{R})$ *sont respectivement homéomorphes à* $\mathrm{O}_n(\mathbb{R}) \times \mathbb{R}^{\frac{n(n+1)}{2}}$ *et* $\mathrm{SO}_n(\mathbb{R}) \times \mathbb{R}^{\frac{n(n-1)}{2}}$*.*

Démonstration. — L'application $\exp : \mathcal{S}_n(\mathbb{R}) \to \mathcal{S}_n^{++}(\mathbb{R})$ est un homéomorphisme et $\mathcal{S}_n(\mathbb{R}) \approx \mathbb{R}^{\frac{n(n+1)}{2}}$, donc

$$\mathrm{GL}_n(\mathbb{R}) \approx \mathrm{O}_n(\mathbb{R}) \times \mathcal{S}_n^{++}(\mathbb{R}) \approx \mathrm{O}_n(\mathbb{R}) \times \mathbb{R}^{\frac{n(n+1)}{2}}\,.$$

De même, $\mathrm{SL}_n(\mathbb{R}) \approx \mathrm{SO}_n(\mathbb{R}) \times (\mathcal{S}_n^{++}(\mathbb{R}) \cap \mathrm{SL}_n(\mathbb{R}))$ et l'application $\exp : \mathcal{S}_n(\mathbb{R}) \cap \mathrm{tr}^{-1}\{0\} \to \mathcal{S}_n^{++}(\mathbb{R}) \cap \mathrm{SL}_n(\mathbb{R})$ est un homéomorphisme. ∎

Propriétés algébriques.

Théorème 3.8. — O(E) *est engendré par des réflexions. Plus précisément, si* $u \in \mathrm{O(E)}$*, alors* u *est produit d'au plus* $\dim \mathrm{E} - \dim \mathrm{Ker}\,(\mathrm{Id} - u)$ *réflexions.*

Démonstration. — On constate tout d'abord que le produit de 2 réflexions dans $\mathbb{R}^2$ par rapport à des droites e_1 et e_2 est une rotation d'angle 2 fois l'angle entre e_1 et e_2. Donc, si on écrit la forme réduite de u, chaque bloc R_θ compte pour deux réflexions, alors que chaque (-1) compte pour une seule. ■

Exercice. Montrer que SO(E) est engendré par des renversements.

Théorème 3.9. — $\mathrm{Z(O(E))} = \{\pm\mathrm{Id}\}$ *et,* $\mathrm{Z(SO(E))} = \{\mathrm{Id}\}$ *si* $\dim \mathrm{E}$ *est impaire,* $\mathrm{Z(SO(E))} = \{\pm\mathrm{Id}\}$ *si* $\dim \mathrm{E}$ *est paire et* $\dim \mathrm{E} \geq 4$*, et* $\mathrm{Z(SO(E))} = \mathrm{SO(E)}$ *si* $\dim \mathrm{E} = 2$*.*

Démonstration. — Soit x de norme 1. On complète en une base orthonormée. La symétrie par rapport à x s'écrit

$$\mathrm{Mat}(s_x, \mathcal{B}) = \begin{pmatrix} 1 & 0 \\ 0 & -\mathrm{I} \end{pmatrix}.$$

Si $u \in \mathrm{O(E)}$, alors $us_xu^{-1} = s_{u(x)}$, donc si $u \in \mathrm{Z(O(E))}$, alors $s_x = s_{u(x)}$, donc il existe $\lambda_x \in \mathbb{R}$ telle que $u(x) = \lambda_x x$. Comme $u \in \mathrm{O(E)}$, on a $\lambda_x = \pm 1$. Ceci implique que u est une homothétie de rapport $\lambda = \pm 1$. En effet, on a, pour x, y indépendants,

$$u(x+y) = \lambda_{x+y}x + \lambda_{x+y}y = \lambda_x x + \lambda_y y.$$

Quant au centre de SO(E), on raisonne de la même manière. ■

Théorème 3.10. — $\mathrm{D(O(E))} = \mathrm{SO(E)}$ *et,* $\mathrm{D(SO(E))} = \mathrm{SO(E)}$ *si* $\dim \mathrm{E} \geq 3$ *et* $\mathrm{D(SO(E))} = \{id\}$ *si* $\dim \mathrm{E} = 2$*.*

Démonstration. — Si $u, v \in \mathrm{O(E)}$ alors $\det uvu^{-1}v^{-1} = 1$, donc $\mathrm{D(O(E))} \subset \mathrm{SO(E)}$. Or les produits pairs de réflexions engendrent SO(E). Montrons que les produits de deux réflexions sont des commutateurs : soient x, y unitaires. Il existe $u \in \mathrm{O(E)}$ tel que $u(x) = y$. On a $s_y = s_{u(x)} = u \circ s_x \circ u^{-1}$, donc

$$s_x \circ s_y = s_x \circ u \circ s_x \circ u^{-1} = s_x \circ u \circ s_x^{-1} \circ u^{-1}.$$

■

Théorème 3.11. — SO(E) *est simple si* $\dim \mathrm{E} = 3$*.*

Démonstration. — Soit G un sous-groupe distingué de SO(E) non trivial. Pour montrer que $\mathrm{G} = \mathrm{SO(E)}$, il suffit de montrer que G contient un demi-tour. A ce moment-là, on saura qu'il les contient tous par conjugaison, et donc que $\mathrm{G} = \mathrm{SO(E)}$. Soit $g \in \mathrm{G}$ non trivial. Comme $g \in \mathrm{SO(E)}$, il s'agit d'une rotation d'axe x et d'angle θ. Si $\theta = \pi$, alors on a gagné.

Si $\theta \neq \pi$, on remarque que, pour $v \in \mathrm{SO(E)}$, on a $vgv^{-1}g^{-1} \in \mathrm{G}$. En particulier, si $v = s_y$, où $y \in \mathrm{E} \setminus \{0\}$, alors $s_y g s_y g^{-1} = s_y \circ s_{g(y)} \in \mathrm{G}$.

Si y et $g(y)$ sont orthogonaux, alors $s_y \circ s_{g(y)}$ serait un demi-tour. Pour voir cela, il suffit de considérer une base orthonormée contenant y et $g(y)$.

Pour conclure, on cherche donc $y \neq 0$ tel que $g(y) \perp y$. Soit $(x, \mathrm{e}_2, \mathrm{e}_3)$ une base orthonormée. On a

$$\mathrm{Mat}(g) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

donc si $y = y_1 x + y_2 \mathrm{e}_2 + y_3 \mathrm{e}_3$, on cherche à résoudre

$$y_1^2 + y_2(y_2\cos\theta - y_3\sin\theta) + y_3(y_2\sin\theta + y_3\cos\theta) = 0,$$

soit

$$y_1^2 + (y_2^2 + y_3^2)\cos\theta = 0.$$

Quitte à itérer g, on peut supposer que $\cos\theta \leq 0$. Du coup, une solution existe. ■

3.4 Sous-groupes finis d'isométries en petite dimension

On s'intéresse à la classification des sous-goupes finis de $\mathrm{O}_2(\mathbb{R})$ et de $\mathrm{SO}_3(\mathbb{R})$ et à leurs relations avec la géométrie.

Sous-groupes de $\mathrm{O}_2(\mathbb{R})$. On s'intéresse d'abord à $\mathrm{O}_2(\mathbb{R})$ en vue de $\mathrm{SO}_3(\mathbb{R})$.

Proposition 3.9. — *Soit* G *un sous-groupe fini de* $\mathrm{O}_2(\mathbb{R})$ *non trivial. Alors* G *préserve un* n*-gône régulier. Si* G *n'est constitué que de rotations, alors* G *est isomorphe à* $\mathbb{Z}/n\mathbb{Z}$*. Sinon,* G *est isomorphe à un groupe diédral* D_n*.*

Démonstration. — Si G n'est constitué que de rotations, alors G s'identifie à un sous-groupe fini de $\mathbb{R}/\mathbb{Z}$. On note $p : \mathbb{R} \to \mathbb{R}/\mathbb{Z}$ la projection canonique. Alors $p^{-1}(\mathrm{G})$ est un sous-groupe de $\mathbb{R}$ qui contient les entiers. Puisque G est fini, $p^{-1}(\mathrm{G})$ est discret, donc il s'agit de $(1/n)\mathbb{Z}$ pour un entier $n \geq 1$. Du coup G est isomorphe à $\mathbb{Z}/n\mathbb{Z}$. On en déduit que G préserve un n-gône régulier inscrit dans le disque unité.

Sinon, notons $\mathrm{SG} = \mathrm{G} \cap \mathrm{SO}_2(\mathbb{R})$. D'après ci-dessous, SG est isomorphe à $\mathbb{Z}/n\mathbb{Z}$. De plus, on a la sous-suite exacte courte

$$1 \to \mathrm{SG} \to \mathrm{G} \to \mathbb{Z}/2\mathbb{Z} \to 1\,.$$

Donc G est d'ordre $2n$.

Soit σ une symétrie de G. Elle fixe deux poins opposés x et $-x$ du cercle unité $\mathbb{S}^1$. On écrit alors $\sigma = \sigma_x = \sigma_{-x}$. On note X l'ensemble des points fixes de toutes les symétries de $\mathrm{G} \setminus \mathrm{SG}$. On a n symétries donc $2n$ points fixes.

D'autre part, on a, pour $g \in \mathrm{G}$,

$$g \circ \sigma_x \circ g^{-1}(g(x)) = g(x)$$

donc

$$g \circ \sigma_x \circ g^{-1} = \sigma_{g(x)}$$

et G opère sur X. Le stabilisateur de chaque point de X est d'ordre deux, donc chaque orbite est de cardinal n. Du coup, on a deux orbites sous l'action de SG.

On en déduit que G est isomorphe à D_n. ∎

Sous-groupes de $\mathrm{SO}_3(\mathbb{R})$. Soit G un sous-groupe de $\mathrm{SO}_3(\mathbb{R})$ d'ordre N. Chaque élément non trivial est une rotation, donc fixe deux points opposés sur la sphère $\mathbb{S}^2$. On note $\mathrm{X} = \mathrm{X}_\mathrm{G}$ l'ensemble de ces points. Comme ci-dessus, G opère sur X.

Le stabilisateur de chaque point x de X fixe le plan orthogonal $x^\perp$. Sa restriction est un sous-groupe fini de $\mathrm{SO}_2(\mathbb{R})$, donc isomorphe à $\mathbb{Z}/r_x\mathbb{Z}$. Son orbite est donc d'ordre $n_x = \mathrm{N}/r_x$. Chaque point x est le point fixe de $(r_x - 1)$ rotations, étant différentes pour tous les autres points excepté son opposé. Du coup, on a

$$2\mathrm{N} - 2 = \sum_{x \in \mathrm{X}} (r_x - 1) = \sum_{j \in \mathrm{X}/\mathrm{G}} n_j (r_j - 1)\,.$$

On en déduit que

$$2 - 2/\mathrm{N} = \sum_{j \in \mathrm{X}/\mathrm{G}} (1 - 1/r_j)\,.$$

Or $r_j \geq 2$ par définition, et donc

$$2 > 2 - 2/\mathrm{N} = \sum_{j \in \mathrm{X}/\mathrm{G}} (1 - 1/r_j) \geq |\mathrm{X}/\mathrm{G}|/2\,.$$

Par conséquent, on a au plus trois orbites. D'autre part, le groupe G n'opère pas transitivement sur X. En effet, on aurait $2 - 2/\mathrm{N} = 1 - 1/r$, soit $1 = 2/\mathrm{N} - 1/r \leq 1/\mathrm{N}$, puisque $r \leq \mathrm{N}$!!

Proposition 3.10. — *Si on a deux orbites, alors* G *est un groupe de rotations du plan, isomorphe à* $\mathbb{Z}/\mathrm{N}\mathbb{Z}$*.*

Démonstration. — On a $2-2/\mathrm{N}=2-(1/r_1+1/r_2)$ soit $2/\mathrm{N}=1/r_1+1/r_2$. Si $\mathrm{N}=r_1$ alors $r_2=\mathrm{N}$. Du coup, X a deux éléments, et G fixe leur orthogonal. Par conséquent, il opère comme un sous-groupe de $\mathrm{SO}_2(\mathbb{R})$ et il est isomorphe à $\mathbb{Z}/\mathrm{N}\mathbb{Z}$.

Sinon, on a $2r_1 \leq \mathrm{N}$, soit $(1/r_1) \geq 2/\mathrm{N}$ donc $r_2 \leq 0$!! ∎

Le cas de trois orbites comporte plusieurs cas. Notre équation s'écrit

$$\frac{1}{r_1}+\frac{1}{r_2}+\frac{1}{r_3}=1+\frac{2}{\mathrm{N}},$$

où on choisit $r_1 \leq r_2 \leq r_3$.

Si $r_1 \geq 3$, alors le terme de gauche est plus petit que 1 alors que le second est strictement plus grand. Donc $r_1=2$.

Si $r_2=2$, alors $1/r_3=2/\mathrm{N}$, soit $\mathrm{N}=2r_3$. On a donc $r_j=(2,2,\mathrm{N}/2)$ et $n_j=(\mathrm{N}/2,\mathrm{N}/2,2)$.

Si $r_2 \geq 3$ alors $1/r_3=1/2+2/\mathrm{N}-1/r_2>1/2-1/3$, donc $r_3<6$.

- Si $r_j=(2,3,3)$ alors $\mathrm{N}=12$ et $n_j=(6,4,4)$.
- Si $r_j=(2,3,4)$ alors $\mathrm{N}=24$ et $n_j=(12,8,6)$.
- Si $r_j=(2,3,5)$ alors $\mathrm{N}=60$ et $n_j=(30,20,12)$.

Si $r_2 \geq 4$, alors

$$\frac{1}{r_1}+\frac{1}{r_2}+\frac{1}{r_3} \leq 1 < 1+\frac{2}{\mathrm{N}},$$

donc on a la liste complète.

Proposition 3.11. — G *est un groupe diédral* D_r *dans le cas*

$$\mathrm{N}=2r \quad r_j=(2,2,r) \quad n_j=(r,r,2).$$

Démonstration. — Les points opposés ont les mêmes comportements donc la troisième orbite est constitué de deux points opposés. Par suite, G fixe leur orthogonale, et les isométries qui ne fixent pas ces points sont des symétries sur ce plan. Donc il s'agit du groupe diédral. ∎

Si on n'est pas dans un des cas précédents, alors aucune orbite n'est planaire. En effet, la restriction au plan nous ramenerait aux cas déjà traités.

Proposition 3.12. — *Si* $\mathrm{N}=12$, $r_j=(2,3,3)$, *et* $n_j=(6,4,4)$ *alors* G *est le groupe d'isométries d'un tétraèdre, et est isomorphe à* $\mathfrak{a}_4$.

Démonstration. — Soit $x \in o(2)$. Son stabilisateur est un groupe de rotations qui opère sur son orbite, donc sur trois points. Ces points forment un triangle équilatéral. Il s'agit donc d'un tétraèdre. De plus, G opère sur ces sommets : il s'identifie à un sous-groupe de permutation à 4 éléments $\mathfrak{S}_4$. Or le seul endomorphisme qui fixe ces quatre points non coplanaires est l'identité. Donc G est un sous-groupe d'indice 2 : il s'agit de $\mathfrak{a}_4$. ∎

Proposition 3.13. — *Si* $\mathrm{N}=24$, $r_j=(2,3,4)$, *et* $n_j=(12,8,6)$ *alors* G *est le groupe d'isométries d'un cube et d'un octaèdre, et est isomorphe à* $\mathfrak{S}_4$.

Démonstration. — Le stabilisateur d'un point $x \in o(3)$ opère sur $o(2)$, en deux orbites. Chacune forme un carré, et ces deux carrés ne peuvent pas être coplanaires. En changeant de point de $o(3)$, on voit que $o(2)$ sont les sommets d'un cube dont les faces sont dans la direction des points de $o(3)$ et les arêtes de $o(1)$.

Or G opère sur les paires de sommets opposés, donc on a un morphisme $\varphi : \mathrm{G} \to \mathfrak{S}_4$. Si $\varphi(g)$ est l'identité, et g échange deux sommets, alors, puisqu'il n'y a que deux points fixes, g échange au moins deux autres paires. Par suite, $g=-\mathrm{Id}$, mais g est une rotation, donc c'est impossible, et φ est injective.

Par dualité, l'envoloppe convexe de $o(3)$ est un octaèdre de faces centrées sur $o(2)$. ∎

On remarque que le cube contient deux tétraèdres "opposés", en considérant pour arêtes des diagonales des faces. Le groupe préserve ces tétraèdres ou les échange. On peut ainsi en déduire que $\mathfrak{a}_4$ est un sous-groupe d'indice 2 de G, donc distingué.

Proposition 3.14. — *Si* $\mathrm{N} = 60$*,* $r_j = (2, 3, 5)$*, et* $n_j = (30, 20, 12)$ *alors* G *est le groupe d'isométries d'un dodécaèdre et d'un icosaèdre, et est isomorphe à* $\mathfrak{a}_5$*.*

Démonstration. — Les trois orbites s'organisent par paires.

Le stabilisateur d'un point $x \in o(3)$ opère sur $o(2)$, en quatre orbites de cinq éléments. Les points de $o(2)$ les plus proches de x forment un pentagone (on ne peut avoir deux orbites sur un même plan en considérant un autre point de $o(3)$). En opérant sur toute l'orbite de x, on devine un dodécaèdre à 20 sommets et 30 arêtes. Par dualité, on obtient un icosaèdre.

On se fixe un sommet $x \in o(2)$. Il correspond à un sommet commun de trois pentagones. Le stabilisateur de x les permute, ainsi que les arêtes communes. Il opère aussi sur les autres sommets de ces pentagones, en deux orbites. Les trois segments qui joignent x à une de ces orbites se coupent à angle droit.

On devine ainsi un cube, si on considère $(-x)$. Les centres des faces correspondent à des points de $o(1)$, soit des centres des arêtes. On peut voir que chaque arête est de même longueur, et que l'on a bien 8 sommets.

A chaque paire de sommets correspond deux cubes, et chaque cube admet 8 sommets, soit 4 diagonales. Du coup, on obtient $10 \times 2/4 = 5$ cubes. De manière équivalente, chaque arête correspond à une face de cube, qui en comporte 6. Ce qui nous fait $30/6 = 5$ cubes.

Notre groupe G opère sur ces cubes par permutation. Supposons qu'un élément $g \in \mathrm{G}$ fixe tous ces cubes globalement. S'il est non trivial, il fixe donc son axe de rotation. Si son ordre est 2, il fixe une arête du dodécaèdre, donc une face de cube. Il ne peut fixer les autres cubes alors. Si son ordre est 3, alors son axe passe par les sommets de deux cubes. Les cubes induits par les sommets contigüs doivent aussi être préservés par g, ce qui est impossible. L'ordre ne peut être 5 puisqu'aucun élément qui préserve un cube n'est d'ordre multiple de 5.

Donc $g = \mathrm{Id}$. Notre groupe est donc un sous-groupe de $\mathfrak{S}_5$ d'indice 2 (pour une question d'ordre), il s'agit de $\mathfrak{a}_5$. ∎

Sous-groupes distingués. Si G est un sous-groupe fini de $\mathrm{SO}_3(\mathbb{R})$, et si H est un sous-groupe de G, alors $\mathrm{X_H} \subset \mathrm{X_G}$. Si $x \in \mathrm{X_H}$, $g \in \mathrm{G}$, alors il existe $h \in \mathrm{H}$ tel que $h(x) = x$, et $ghg^{-1}(gx) = gx$. Donc l'action de G induite par automorphismes intérieurs sur H est en correspondance avec l'action de G sur les images de $\mathrm{X_H}$.

Si de plus H est distingué, alors $\mathrm{X_H}$ est préservé. Donc G opère sur $\mathrm{X_H}$, et chaque point de $\mathrm{X_H}$ a une orbite par H qui est une sous-orbite pour G. Autrement dit, $o_{\mathrm{H}}(x)$ divise $o_{\mathrm{G}}(x)$, et $\mathrm{X_H}$ est une réunion d'orbites de G. Du coup, les orbites dans $\mathrm{X_H}$ de G se décomposent en orbites de H. On montre ainsi facilement que $\mathfrak{a}_5$ est simple.

Remarque. — Un polytope convexe est une intersection de demi-espaces affines d'intérieur non vide. Si P est un polytope convexe, on peut supposer que l'origine est dans son intérieur. Si on projette le bord de P sur $\mathbb{S}^2$, alors on obtient une triangulation de la sphère en sommets, arêtes et faces. Si on note s le nombre de sommets, a le nombre d'arêtes et f le nombre de faces, alors $s - a + f = 2$.

En effet, chaque fois que l'on supprime un sommet avec les arêtes qui le contient, on enlève autant de faces que d'arêtes, sauf que le sommet se transforme en face : $s - a + f$ reste constant lorsque l'on diminue le nombre de sommets. Lorsqu'il ne reste plus que 4 sommets, alors on peut vérifier la formule.

4 Espaces hermitiens complexes

Définition 4.1. — *Un espace hermitien complexe est un espace vectoriel de dimension finie sur* $\mathbb{C}$ *muni d'une forme quadratique hermitienne* q *définie positive.*

Théorème 4.1. — *Un espace hermitien est naturellement un espace vectoriel normé.*

Pour montrer ce résultat, on commence par montrer l'inégalité de Cauchy-Schwarz.

Inégalité de Cauchy-Schwarz. — *Soit* (E, q) *un* $\mathbb{C}$*-espace vectoriel de dimension finie muni d'une forme hermitienne positive (pour tout* $x \in \mathrm{E}$*, on a* $q(x) \geq 0$*). On a, pour tout* $x, y \in \mathrm{E}$*,*

$$(\mathrm{Re}\, b(x,y))^2 \leq q(x)q(y)\,.$$

Démonstration de l'inégalité de Cauchy-Schwarz. — Pour $t \in \mathbb{R}$, on a $q(xt+y) \geq 0$. Or $q(tx+y) = q(x)t^2 + 2t\mathrm{Re}\, b(x,y) + q(y)$. Comme ce polynôme quadratique ne prend que des valeurs positives, il ne peut s'annuler au plus qu'une fois, donc son discriminant doit être négatif :

$$(\mathrm{Re}\, b(x,y))^2 - q(x)\cdot q(y) \leq 0\,.$$

■

Démonstration du Théorème 4.1. — On pose, pour $x \in \mathrm{E}$, $\|x\| = \sqrt{q(x)}$. Pour montrer que $\|\cdot\|$ est une norme, il suffit de vérifier l'inégalité triangulaire.

$$\begin{aligned}(\|x\| + \|y\|) - \|x+y\| &= \frac{q(x) + q(y) + 2\sqrt{q(x)q(y)} - q(x+y)}{\|x+y\| + \|x\| + \|y\|} \\ &\geq \frac{2\sqrt{q(x)q(y)} - 2\mathrm{Re}\, b(x,y)}{\|x+y\| + \|x\| + \|y\|} \geq 0\,.\end{aligned}$$

■

Définition 4.2. — *Une base orthonormée est une base* $\mathcal{B} = (\mathrm{e}_1, \ldots, \mathrm{e}_n)$ *telle que* $b(\mathrm{e}_i, \mathrm{e}_j) = \delta_{i,j}$*, où* $\delta_{i,j}$ *est le symbole de Kronecker.*

Théorème 4.2. — *Dans un espace hermitien, il existe toujours une base orthonormée.*

Dans une telle base, la matrice de q est l'identité. Du coup, les éléments unitaires sont représentés par des matrices M telles que ${}^t\overline{\mathrm{M}} \cdot \mathrm{M} = \mathrm{I}$. Dans cette situation, on appelle la forme polaire *un produit scalaire hermitien* et on le note communément $\langle\cdot,\cdot\rangle$. Dans une base orthonormée, on a $\langle \mathrm{X}, y\rangle = {}^t\mathrm{X}\cdot\overline{\mathrm{Y}}$.

Démonstration. — On applique l'algorithme de Gram-Schmidt.

L'orthogonalisation de Gram-Schmidt implique aussi que, pour tout sous-espace F de E, il existe une base orthonormée dont les premiers $\dim \mathrm{F}$ vecteurs forment une base de F.

Proposition 4.1. — *Soit* $b : \mathrm{E}^2 \to \mathbb{C}$ *une forme linéaire en* x *et anti-linéaire en* y *sur un espace hermitien. Il existe des endomorphismes* u *et* v *tels que, pour tout* $x, y \in \mathrm{E}$*, on ait*

$$b(x,y) = \langle x, u(y)\rangle = \langle v(x), y\rangle\,.$$

Ces endomorphismes sont uniques.

Démonstration. — Soit $\mathcal{B} = (\mathrm{e}_1, \ldots, \mathrm{e}_n)$ une base orthonormée ; on note $\mathrm{M} = (b(\mathrm{e}_i, \mathrm{e}_j))_{i,j}$ la matrice de b. On note X, Y les vecteurs coordonnées de $x, y \in \mathrm{E}$. On a

$$b(x,y) = \sum_{i,j} x_i\overline{y_j} b(\mathrm{e}_i, \mathrm{e}_j) = {}^t\mathrm{X}\cdot\mathrm{M}\overline{\mathrm{Y}} = \langle {}^t\mathrm{M}\mathrm{X}, \mathrm{Y}\rangle = \langle \mathrm{X}, \overline{\mathrm{M}}\mathrm{Y}\rangle\,.$$

On définit donc, pour tout $j \in \{1, \dots, n\}$,

$$\begin{cases} u(\mathrm{e}_j) = \displaystyle\sum_{i=1}^{n} b(\mathrm{e}_i, \mathrm{e}_j)\mathrm{e}_i \,, \\ \\ v(\mathrm{e}_i) = \displaystyle\sum_{j=1}^{n} b(\mathrm{e}_i, \mathrm{e}_j)\mathrm{e}_j \,. \end{cases}$$

On a bien

$$\begin{cases} \langle \mathrm{e}_i, u(\mathrm{e}_j)\rangle = \displaystyle\sum_{k=1}^{n} b(\mathrm{e}_k, \mathrm{e}_j)\langle \mathrm{e}_k, \mathrm{e}_i\rangle = b(\mathrm{e}_i, \mathrm{e}_j) \,, \\ \\ \langle v(\mathrm{e}_i), \mathrm{e}_j\rangle = \displaystyle\sum_{k=1}^{n} b(\mathrm{e}_i, \mathrm{e}_k)\langle \mathrm{e}_j, \mathrm{e}_k\rangle = b(\mathrm{e}_i, \mathrm{e}_j) \,. \end{cases}$$

L'unicité provient aussi de ces relations. ■

Définition 4.3. — *Si $u \in \mathrm{End}(\mathrm{E})$, on appelle* adjoint de u, *que l'on note u^*, l'endomorphisme tel que, pour tout $x, y \in \mathrm{E}$,*

$$\langle u(x), y\rangle = \langle x, u^*(y)\rangle \,.$$

L'adjoint existe toujours d'après la proposition précédente puisque $b : (x, y) \mapsto \langle x, u(y)\rangle$ est sesquilinéaire. De plus, $u^{**} = u$.

Si M est la matrice de u dans une base orthonormée, alors

$$\langle x, u^*(y)\rangle = \langle u(x), y\rangle = {}^t(\mathrm{M}\cdot\mathrm{X})\cdot\overline{\mathrm{Y}} = {}^t\mathrm{X}\cdot{}^t\mathrm{M}\cdot\overline{\mathrm{Y}}$$

donc la matrice de u^* est $\overline{{}^t\mathrm{M}}$.

On dit que u est hermitien si $u = u^*$, antihermitien si $u = -u^*$ et unitaire si $uu^* = \mathrm{Id}$.

Remarque. — Les polynômes caractéristiques de u et u^* sont conjugués.

4.1 Endomorphismes normaux

Définition 4.4. — *Un endomorphisme est normal s'il commute avec son adjoint.*

Le résultat principal de ce paragraphe est le suivant.

Théorème 4.3. — *Pour tout endomorphisme normal u d'un espace hermitien, il existe une base orthonormée de* E *qui diagonalise u. Si u est hermitien, alors les valeurs propres sont réelles, si u est antihermitien, elles sont imaginaires pures et si u est unitaire, alors elles sont de module* 1.

Proposition 4.2. — *Soit u un endomorphisme normal.*

1. *On a l'égalité* $\mathrm{Ker}\, u = \mathrm{Ker}\, u^*$.
2. *u et u^* ont leurs valeurs propres conjuguées et partagent les mêmes espaces propres. Ceux-ci sont orthogonaux deux à deux.*
3. *u fixe au moins une droite.*
4. *Si* F *est stable par u, alors* F *est aussi stable par u^*, et* $\mathrm{F}^\perp$ *est stable par u et u^*.*

Démonstration. —

1. Si u est normal, alors, pour tout $x \in \mathrm{E}$, on a $\|u(x)\| = \|u^*(x)\|$. En effet,

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

Donc $u(x)$ et $u^*(x)$ s'annulent simultanément.

2. provient du fait que $u+\lambda\mathrm{I}$ est normal si u est normal, donc $\mathrm{Ker}\,(u+\lambda\mathrm{Id}) = \mathrm{Ker}\,(u^*+\bar{\lambda}\mathrm{Id})$ d'après ci-dessus. Soient λ, μ deux valeurs propres distinctes, et x, y deux vecteurs propres associés : $u(x) = \lambda x$ et $u(y) = \mu y$. Il vient

$$\lambda\langle x, y\rangle = \langle u(x), y\rangle = \langle x, u^*(y)\rangle = \mu\langle x, y\rangle\,.$$

Du coup $(\lambda-\mu)\langle x, y\rangle = 0$ et $\langle x, y\rangle = 0$.

3. Puisque $\mathbb{C}$ est algébriquement clos, il existe toujours une valeur propre et un vecteur propre associé.

4. On considère une base orthonormée de E telle que les premiers $p = \dim \mathrm{F}$ vecteurs forment une base de F et les derniers une base de $\mathrm{F}^\perp$. Dire que F est stable par u signifie qu'il existe des matrices $\mathrm{A} \in \mathcal{M}_p(\mathbb{C})$, $\mathrm{B} \in \mathcal{M}_{p,n-p}(\mathbb{C})$ et $\mathrm{C} \in \mathcal{M}_{n-p}(\mathbb{C})$ telles que la matrice de u s'écrive

$$\mathrm{Mat}(u) = \begin{pmatrix} \mathrm{A} & \mathrm{B} \\ 0 & \mathrm{C} \end{pmatrix}$$

Dire que u est normal signifie que $\mathrm{A}\cdot\,{}^t\overline{\mathrm{A}} + \mathrm{B}\cdot{}^t\overline{\mathrm{B}} = {}^t\overline{\mathrm{A}}\cdot\mathrm{A}$. Or, un simple calcul montre que $\mathrm{tr}(\mathrm{A}\cdot{}^t\overline{\mathrm{A}}) = \mathrm{tr}({}^t\overline{\mathrm{A}}\cdot\mathrm{A})$, donc $\mathrm{tr}(\mathrm{B}\cdot{}^t\overline{\mathrm{B}}) = 0$. Mais, si $\mathrm{B} = (b_{ij})_{1\leq i\leq p, 1\leq j\leq n-p}$, alors

$$\mathrm{tr}(\mathrm{B}\cdot{}^t\overline{\mathrm{B}}) = \sum_{j=1}^{p}\left(\sum_{k=1}^{n-p} b_{jk}\overline{b_{jk}}\right) = \sum_{j=1}^{p}\sum_{k=1}^{n-p}|b_{jk}|^2\,.$$

Donc $b_{jk} = 0$ pour tout j, k et $\mathrm{B} = 0$. Du coup,

$$\mathrm{Mat}(u) = \begin{pmatrix} \mathrm{A} & 0 \\ 0 & \mathrm{C} \end{pmatrix}$$

avec A et C normaux. Ceci montre que F et $\mathrm{F}^\perp$ sont stables par u et u^*.

■

Démonstration du Théorème de réduction. — On procède par récurrence sur la dimension de E. Pour $n = 1$, il n'y a rien à dire. Supposons que ce soit vrai pour tout espace de dimension n, et que E est de dimension $n+1$. D'après la proposition précédente, il existe une valeur propre λ, et un vecteur propre (normé) e_{n+1}. On note $\mathrm{F} = \mathbb{C}\mathrm{e}_{n+1}$. Cet espace est stable par u donc $\mathrm{F}^\perp$ aussi, qui est un supplémentaire de F. L'hypothèse de récurrence s'applique à $u|_{\mathrm{F}^\perp}$. Cette base se complète en une base orthonormée de E, et la matrice de u dans cette base est diagonale. ■

Corollaire 4.1. — *L'application* $\exp : \mathcal{H}_n(\mathbb{C}) \to \mathcal{H}_n^{++}(\mathbb{C})$ *est un homéomorphisme.*

Démonstration. — L'application est exp est continue, et si $\mathrm{M} \in \mathcal{H}_n(\mathbb{C})$, alors il existe $\mathrm{U} \in \mathrm{U}(n)$ telle que $\mathrm{U}^{-1}\mathrm{MU}$ soit diagonale. On en déduit que $\exp \mathrm{M} = \mathrm{U}(\exp \mathrm{U}^{-1}\mathrm{MU})\mathrm{U}^{-1} \in \mathcal{H}_n^{++}(\mathbb{C})$. Réciproquement, si $\mathrm{M} \in\in \mathcal{H}_n^{++}(\mathbb{C})$, il existe $\mathrm{U} \in \mathrm{U}(n)$ telle que $\mathrm{U}^{-1}\mathrm{MU}$ soit diagonale avec des valeurs propres strictement positives. On peut alors considérer la matrice N diagonale formée des logarithmes des valeurs propres de M. On a $\mathrm{UNU}^{-1} \in \mathcal{H}_n(\mathbb{C})$ et $\exp(\mathrm{UNU}^{-1}) = \mathrm{M}$.

Il reste à voir que l'application est injective. On suppose donc que $\exp \mathrm{M} = \exp \mathrm{N}$. Quitte à changer de base, on peut supposer que M est diagonale. Par suite, $\exp \mathrm{M}$ est diagonale aussi. Ceci montre que les valeurs propres de M et $\exp \mathrm{M}$ sont liées, ainsi que les espaces propres associés. Il en est donc de même pour N. On en déduit que $\mathrm{M} = \mathrm{N}$.

Enfin, pour voir que l'application réciproque est continue, on observe que $\mathcal{H}_n(\mathbb{C})$ admet une exhaustion par des compacts en considérant celles dont le spectre est contenu dans un intervalle compact. Du coup, chaque restriction est un homéomorphisme sur son image. ■

4.2 Le groupe unitaire

On suppose que E est un espace hermitien muni d'une base orthonormée $\mathcal{B}$.

Théorème 4.4. — *Soit u un endomorphisme de* E. *Les propriétés suivantes sont équivalentes.*

- $u \in \mathrm{U(E)}$;
- *pour tout* $x \in \mathrm{E}$, *on a* $\|u(x)\| = \|x\|$;
- $u \circ u^* = \mathrm{Id}$;
- $u^* \circ u = \mathrm{Id}$;
- $\mathrm{Mat}(u, \mathcal{B}) \cdot {}^t\overline{\mathrm{Mat}(u, \mathcal{B})} = \mathrm{I}$;
- ${}^t\overline{\mathrm{Mat}(u, \mathcal{B})} \cdot \mathrm{Mat}(u, \mathcal{B}) = \mathrm{I}$;
- *les colonnes de* $\mathrm{Mat}(u, \mathcal{B})$ *forment une base orthonormée ;*
- *les lignes de* $\mathrm{Mat}(u, \mathcal{B})$ *forment une base orthonormée ;*
- u *transforme une base orthonormée en une base orthonormée.*

La démonstration est laissée en exercice.

Corollaire 4.2. — *Si* $u \in \mathrm{U(E)}$, *alors* $|\det u| = 1$ *et donc* $\mathrm{SU(E)} = \det^{-1}\{1\} \cap \mathrm{U(E)}$ *est distingué.*

Propriétés topologiques. On s'intéresse aux propriétés topologiques du groupe unitaire.

Lemme 4.1. — $\mathrm{U}_n(\mathbb{C})$ *est compact.*

Démonstration. — On considère l'application $f : \mathrm{M} \in \mathcal{M}_n(\mathbb{C}) \mapsto {}^t\overline{\mathrm{M}}\mathrm{M}$. On a $\mathrm{U}_n(\mathbb{C}) = f^{-1}(\mathrm{I}_n)$, donc $\mathrm{U}_n(\mathbb{C})$ est fermé. Pour montrer que $\mathrm{U}(n)$ est borné, on considère la norme induite par $\mathrm{tr}\,{}^t\overline{\mathrm{M}}\mathrm{M}$. On a $\mathrm{U}_n(\mathbb{C}) \subset \mathrm{B}(0, \sqrt{n})$. ■

Proposition 4.3. — *Les groupes* U(E) *et* SU(E) *sont connexes.*

Démonstration. — On montre d'abord que U(E) est connexe par arcs. Le þde réduction nous permet de diagonaliser en base orthonormée tout éléments de U(E) : étant donné $\mathrm{U} \in \mathrm{U(E)}$, il existe $\widehat{\mathrm{U}} \in \mathrm{U(E)}$ et $\theta_1, \dots, \theta_n \in \mathbb{R}$ tels que

$$\mathrm{U} = \widehat{\mathrm{U}} \begin{pmatrix} \mathrm{e}^{i\theta_1} & & 0 \\ & \ddots & \\ 0 & & \mathrm{e}^{i\theta_n} \end{pmatrix} \widehat{\mathrm{U}}^{-1}.$$

On note, pour $s \in [0, 1]$,

$$\mathrm{U}_s = \widehat{\mathrm{U}} \begin{pmatrix} \mathrm{e}^{is\theta_1} & & 0 \\ & \ddots & \\ 0 & & \mathrm{e}^{is\theta_n} \end{pmatrix} \widehat{\mathrm{U}}^{-1}.$$

Pour SU(E), il faut s'y prendre un peu différemment. On sait que $\sum \theta_j \in 2\pi\mathbb{Z}$. On fait une récurrence sur la dimension pour ramener chaque valeur propre à 1 sans changer le déterminant. Pour cela, on définit $\theta_1(s) = \theta_1 - s$, $\theta_2(s) = \theta_1 + s$ et $\theta_j(s) = \theta_j$, pour $s \in [0, \theta_1]$. ■

Décomposition polaire. L'application $\Phi : \mathrm{U}(n) \times \mathcal{H}_n^{++}(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{C})$ définie par $\Phi(\mathrm{U}, \mathrm{H}) = \mathrm{UH}$ est un homéomorphisme.

Lemme 4.2. — *Si* $\mathrm{M} \in \mathrm{GL}_n(\mathbb{C})$, *alors* ${}^t\overline{\mathrm{M}} \cdot \mathrm{M}$ *est définie positive.*

Démonstration. — On considère une base orthonormée de $\mathbb{C}^n$ et on considère $\overline{\mathrm{M}}$ comme la matrice d'un endomorphisme u (inversible) dans cette base. On a, pour $x \neq 0$,

$$\langle x, u^* \circ u(x) \rangle = \langle u(x), u(x) \rangle > 0\,..$$

■

Démonstration de la décomposition : existence. — D'après le lemme, il existe $\widehat{\mathrm{U}} \in \mathrm{U}(n)$ telle que

$$\widehat{\mathrm{U}}^{-1} \cdot ({}^t\overline{\mathrm{M}} \cdot \mathrm{M}) \cdot \widehat{\mathrm{U}} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

soit diagonale avec des valeurs propres strictement positives. On note

$$\mathrm{H} = \widehat{\mathrm{U}} \cdot \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} \cdot \widehat{\mathrm{U}}^{-1}$$

et $\mathrm{U} = \mathrm{M}\mathrm{H}^{-1}$. Comme H est hermitienne, H^{-1} aussi, donc

$${}^t\overline{\mathrm{U}} \cdot \mathrm{U} = {}^t(\overline{\mathrm{H}^{-1}}) \cdot ({}^t\overline{\mathrm{M}} \cdot \mathrm{M}) \cdot \mathrm{H}^{-1} = \mathrm{H}^{-1}\mathrm{H}^2\mathrm{H}^{-1} = \mathrm{I}\,.$$

■

Démonstration de la décomposition : unicité. — Soit $\mathrm{P} \in \mathbb{C}[\mathrm{X}]$ tel que $\mathrm{P}(\lambda_j) = \sqrt{\lambda_j}$. On a

$$\mathrm{P}({}^t\overline{\mathrm{M}} \cdot \mathrm{M}) = \sum a_k \widehat{\mathrm{U}} \cdot \begin{pmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix} \cdot \widehat{\mathrm{U}}^{-1} = \widehat{\mathrm{U}} \cdot \begin{pmatrix} \mathrm{P}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & \mathrm{P}(\lambda_n) \end{pmatrix} \cdot \widehat{\mathrm{U}}^{-1} = \mathrm{H}\,.$$

Donc, si M admet une seconde décomposition polaire $\mathrm{M} = \mathrm{U}'\mathrm{H}'$ alors ${}^t\overline{\mathrm{M}} \cdot \mathrm{M} = (\mathrm{H}')^2$ donc H′ commute avec ${}^t\overline{\mathrm{M}} \cdot \mathrm{M}$, donc avec H puisque H est un polynôme en ${}^t\overline{\mathrm{M}} \cdot \mathrm{M}$. Du coup, H et H′ peuvent être diagonalisées simultanément (en effet, $\mathrm{E}_{\lambda_i}(\mathrm{H}) = \oplus(\mathrm{E}_{\lambda_i}(\mathrm{H}) \cap \mathrm{E}_{\mu_j}(\mathrm{H}'))$). Ceci oblige les valeurs propres à coïncider, donc ces matrices sont les mêmes. ■

Démonstration de la décomposition : homéomorphie. — L'application Φ est clairement continue. Réciproquement, si $\mathrm{M}_n = \Phi(\mathrm{U}_n, \mathrm{H}_n)$ converge vers une matrice $\mathrm{M} = \mathrm{U}\mathrm{H}$, montrons que U_n tend vers U et H_n vers H. Comme U(E) est compact, quitte à extraire une sous-suite, on peut supposer que U_n tend vers une matrice unitaire $\widehat{\mathrm{U}}$. Du coup, H_n tend vers une matrice (hermitienne) $\widehat{\mathrm{U}}^{-1}\mathrm{M}$. Du coup, on a $\mathrm{U}\mathrm{H} = \widehat{\mathrm{U}}(\widehat{\mathrm{U}}^{-1}\mathrm{M})$. Par l'unicité de la décomposition polaire, on obtient $\mathrm{U} = \widehat{\mathrm{U}}$ et H_n tend vers H. ■

On en déduit quelques résultats.

Proposition 4.4. — *Deux matrices réelles* A *et* B *unitairement semblables sont orthogonalement semblables.*

Démonstration. — On suppose qu'il existe $\mathrm{U} \in \mathrm{U}_n(\mathbb{C})$ telle que $\mathrm{A}\mathrm{U} = \mathrm{U}\mathrm{B}$. On écrit $\mathrm{U} = \mathrm{U}_1 + i\mathrm{U}_2$, où U_1 et U_2 sont des matrices réelles. Puisque ${}^t\overline{\mathrm{U}} = \mathrm{U}^{-1}$, on a $\mathrm{U}^{-1} = {}^t\mathrm{U}_1 - i\,{}^t\mathrm{U}_2$.

De plus, la similitude de A et B implique $\mathrm{A}\mathrm{U}_1 = \mathrm{U}_1\mathrm{B}$ et $\mathrm{A}\mathrm{U}_2 = \mathrm{U}_2\mathrm{B}$. Par suite, si $\lambda \in \mathbb{R}$, alors $\mathrm{A}(\mathrm{U}_1 + \lambda\mathrm{U}_2) = (\mathrm{U}_1 + \lambda\mathrm{U}_2)\mathrm{B}$. On écrit $\mathrm{Q}(\lambda) = \det(\mathrm{U}_1 + \lambda\mathrm{U}_2) \in \mathbb{R}[\lambda]$. Comme $\mathrm{Q}(i) = \det \mathrm{U} \neq 0$, on en déduit que Q est un polynôme non nul, donc il existe $\lambda \in \mathbb{R}$ tel que $\mathrm{Q}(\lambda) \neq 0$ et $\mathrm{P} := \mathrm{U}_1 + \lambda\mathrm{U}_2 \in \mathrm{GL}_n(\mathbb{R})$.

D'autre part, ${}^t\mathrm{A} = {}^t\mathrm{U}^{-1}\,{}^t\mathrm{B}\,{}^t\mathrm{U} = \overline{\mathrm{U}}\,{}^t\mathrm{B}\,\overline{\mathrm{U}^{-1}}$ donc ${}^t\mathrm{A} = \mathrm{U}\,{}^t\mathrm{B}\mathrm{U}^{-1}$ et ${}^t\mathrm{A}\mathrm{P} = \mathrm{P}\,{}^t\mathrm{B}$. On a

$$\mathrm{P}\,{}^t\mathrm{B}\mathrm{P}^{-1} = \mathrm{A} = {}^t({}^t\mathrm{A}) = {}^t\mathrm{P}^{-1}\mathrm{B}\,{}^t\mathrm{P}$$

donc ${}^t\mathrm{P}\mathrm{P}\mathrm{B} = \mathrm{B}\,{}^t\mathrm{P}\mathrm{P}$. Or, la décomposition polaire nous donne $\mathrm{P} = \mathrm{O}\mathrm{S}$, où S est un polynôme en ${}^t\mathrm{P}\mathrm{P}$. Donc $\mathrm{B}\mathrm{S} = \mathrm{S}\mathrm{B}$ et on obtient

$$\mathrm{A} = \mathrm{P}\mathrm{B}\mathrm{P}^{-1} = \mathrm{O}\mathrm{S}\mathrm{B}\mathrm{S}^{-1}\mathrm{O}^{-1} = \mathrm{O}\mathrm{B}\mathrm{O}^{-1}\,.$$

■

Proposition 4.5. — $\mathrm{GL}_n(\mathbb{C})$ *est connexe.*

DÉMONSTRATION. — On a

$$\mathrm{GL}_n(\mathbb{C}) \approx \mathrm{U}_n(\mathbb{C}) \times \mathcal{H}_n^{++}(\mathbb{C})$$

par la décomposition polaire. ■

Remarque.— On a une démonstration plus simple : soient $\mathrm{P}, \mathrm{Q} \in \mathrm{GL}_n(\mathbb{C})$, et notons $\mathrm{R}(\lambda) = \det(\mathrm{P} + \lambda\mathrm{Q})$ qui est un polynôme complexe non nul puisque $\mathrm{R}(0) \neq 0$. Donc il existe un chemin dans $\mathbb{C} \setminus \mathrm{R}^{-1}\{0\}$ qui relie $\lambda = 0$ à $\lambda = 1$.

Proposition 4.6. — $\mathrm{SU}_n(\mathbb{C})$ *(resp.* $\mathrm{U}_n(\mathbb{C})$*) est un sous-groupe compact maximal de* $\mathrm{SL}_n(\mathbb{C})$ *(resp.* $\mathrm{GL}_n(\mathbb{C})$*).*

DÉMONSTRATION. — Soit $\mathrm{G} \subset \mathrm{SL}_n(\mathbb{C})$ un sous-groupe compact qui contient $\mathrm{SU}_n(\mathbb{C})$, et soit $g \in \mathrm{G} \setminus \mathrm{SU}_n(\mathbb{C})$. On a $g = \mathrm{UH}$ avec $\mathrm{U} \in \mathrm{SU}_n(\mathbb{C})$ car $\det g > 0$ et $\mathrm{H} \neq \mathrm{I}$ car $g \notin \mathrm{SU}_n(\mathbb{C})$. Donc $\mathrm{H} \in \mathrm{G}$. Or si v est un vecteur propre associé à une valeur propre λ de H de moule différent de 1, alors $\mathrm{Log}\,\|\mathrm{H}^n x\|$ tend vers l'infini, ce qui contredit la compacité de G. Donc $\mathrm{H} \in \mathrm{SU}_n(\mathbb{C})$. ■

Remarque. — Il est facile de voir que $\mathrm{O}_2(\mathbb{C})$ n'est pas compact, donc $\mathrm{O}_n(\mathbb{C})$ non plus.

Proposition 4.7. — *Les espaces* $\mathrm{GL}_n(\mathbb{C})$ *et* $\mathrm{SL}_n(\mathbb{C})$ *sont respectivement homéomorphes à* $\mathrm{U}_n(\mathbb{C}) \times \mathbb{C}^{n^2}$ *et* $\mathrm{SU}_n(\mathbb{C}) \times \mathbb{C}^{n^2-1}$.

DÉMONSTRATION. — L'application $\exp : \mathcal{H}_n(\mathbb{C}) \to \mathcal{H}_n^{++}(\mathbb{C})$ est un homéomorphisme et $\mathcal{H}_n(\mathbb{C}) \approx \mathbb{C}^{n^2}$, donc

$$\mathrm{GL}_n(\mathbb{C}) \approx \mathrm{U}_n(\mathbb{C}) \times \mathcal{H}_n^{++}(\mathbb{C}) \approx \mathrm{U}_n(\mathbb{C}) \times \mathbb{C}^{n^2}\,.$$

De même, $\mathrm{SL}_n(\mathbb{C}) \approx \mathrm{SU}_n(\mathbb{C}) \times (\mathcal{H}_n^{++}(\mathbb{C}) \cap \mathrm{SL}_n(\mathbb{C}))$ et l'application $\exp : \mathcal{H}_n(\mathbb{C}) \cap \mathrm{tr}^{-1}\{0\} \to \mathcal{H}_n^{++}(\mathbb{C}) \cap \mathrm{SL}_n(\mathbb{C})$ est un homéomorphisme. ■

Proposition 4.8. — *Une matrice* $\mathrm{A} \in \mathcal{M}_n(\mathbb{C})$ *appartient à* $\mathrm{O}_n(\mathbb{C})$ *si et seulement si il existe* $\mathrm{O} \in \mathrm{O}_n(\mathbb{R})$ *et* $\Theta \in \mathrm{A}_n(\mathbb{R})$ *telles que* $\mathrm{A} = \mathrm{O}\exp i\Theta$.

DÉMONSTRATION. — Si $\mathrm{A} = \mathrm{O}\exp i\Theta$ alors

$${}^t\mathrm{AA} = (\exp i\,{}^t\Theta)\,{}^t\mathrm{OO}\exp i\Theta = \exp i({}^t\Theta + \Theta) = \mathrm{I}_n\,.$$

Réciproquement, on utilise la décomposition polaire complexe : il existe $\mathrm{U} \in \mathrm{U}_n(\mathbb{C})$ et P hermitienne définie positive telle que $\mathrm{A} = \mathrm{UP}$. Or ${}^t\mathrm{AA} = \mathrm{I}$ implique que ${}^t\mathrm{P}\,{}^t\mathrm{UUP} = \mathrm{I}$, soit ${}^t\mathrm{UUP} = {}^t\,\mathrm{P}^{-1}$. Donc l'unicité de la décomposition nous montre que ${}^t\mathrm{UU} = \mathrm{I}$ et $\mathrm{P} = {}^t\,\mathrm{P}^{-1}$. La première condition nous conduit à $\mathrm{U} = \overline{\mathrm{U}}$ car U est unitaire, donc $\mathrm{U} \in \mathrm{O}_n(\mathbb{R})$. Quant à la seconde, on écrit $\mathrm{P} = \exp \mathrm{L}$, où L est hermitienne. On obtient $\mathrm{L} = -{}^t\mathrm{L}$, donc L est antisymétrique. De plus, comme L est hermitienne, on a ${}^t\mathrm{L} = \overline{\mathrm{L}} = -\mathrm{L}$ donc $\mathrm{L} = i\Theta$ avec $\Theta \in \mathrm{A}_n(\mathbb{R})$. ■

Comme corollaire, on en déduit que $\mathrm{O}_n(\mathbb{C})$ est homéomorphe à $\mathrm{O}_n(\mathbb{R}) \times \mathbb{R}^{\frac{n(n-1)}{2}}$ et que $\mathrm{O}_n(\mathbb{C})$ a exactement deux composantes connexes.

Théorème 4.5. — *On a* $\mathrm{SU}_2(\mathbb{C})/\{\pm\mathrm{I}_2\}$ *est homéomorphe à* $\mathrm{SO}_3(\mathbb{R})$.

Lemme 4.3. — *On considère le corps des quaternions sous la forme*

$$\mathbb{H} = \left\{ \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix},\ a, b \in \mathbb{C} \right\}$$

que l'on munit de $q(\mathrm{M}) = \det(\mathrm{M})$. *Alors* $\mathbb{H}$ *est un espace euclidien de dimension* 4 *et* $\mathrm{SU}_2(\mathbb{C})$ *s'identifie à la sphère unité de* $\mathbb{H}$.

DÉMONSTRATION. — Il est aisé de voir que $\mathbb{H}$ est un $\mathbb{R}$-espace vectoriel de dimension 4. Le déterminant est bien une forme quadratique par le théorème de Cayley-Hamilton : on constate que

$$\det \mathrm{M} = \frac{1}{2}((\operatorname{tr} \mathrm{M})^2 - \operatorname{tr} \mathrm{M}^2)$$

et $\det \mathrm{M} \geq 0$ pour toute $\mathrm{M} \in \mathbb{H}$. De plus $\det \mathrm{M} = 0$ si et seulement si $\mathrm{M} = 0$. Du coup, q est positive sans vecteur isotrope. On a bien une structure euclidienne.

Si $\mathrm{M} \in \mathrm{SU}_2$, alors les vecteurs colonnes sont orthogonaux, donc $a\bar{b} + c\bar{d} = 0$. Si $d = 0$ alors $bc = 1$ et $|b| = 1$ donc $c = -\bar{b}$ et $a = 0$. Sinon, $ad\bar{b} + c|d|^2 = 0$ et $(|b|^2 + |d|^2)c + \bar{b} = 0$ car $\det \mathrm{M} = 1$, soit $c + \bar{b} = 0$ car les vecteurs colonnes sont unitaires. En reprenant la première équation, on obtient $a - \bar{d} = 0$ si $b \neq 0$. Sinon, un argument similaire montre aussi que $d = \bar{a}$. Donc $\mathrm{M} \in \mathbb{H}$. ∎

On passe maintenant à la démonstration du théorème.

DÉMONSTRATION. — On note $\mathbb{H}_0 = \mathbb{H} \cap \operatorname{tr}^{-1}\{0\}$ l'ensemble des quaternions purs (on a $\mathrm{I}^\perp = \mathbb{H}_0$). Cet espace $\mathbb{H}_0$ est un $\mathbb{R}$-espace vectoriel de dimension 3. L'application

$$\mathcal{P} : \mathrm{X} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} ix_1 & -x_2 + ix_3 \\ x_2 + ix_3 & -ix_1 \end{pmatrix}$$

définit d'ailleurs un isomorphisme entre $\mathbb{R}^3$ et $\mathbb{H}_0$, et on observe que

$$\det \mathcal{P}(\mathrm{X}) = \|\mathrm{X}\|$$

donc $\mathcal{P}$ réalise en fait une isométrie (euclidienne) si on munit $\mathbb{H}_0$ de la forme "det".

On remarque que les éléments H de $\mathbb{H}_0$ vérifient $\mathrm{H} = -{}^t\overline{\mathrm{H}}$ et $\operatorname{tr} \mathrm{H} = 0$.

Si $\mathrm{H} \in \mathbb{H}_0$ et si $\mathrm{U} \in \mathrm{SU}_2(\mathbb{C})$, alors

$$-{}^t(\overline{\mathrm{UHU}^{-1}}) = {}^t\,\overline{\mathrm{U}^{-1}}(-{}^t\overline{\mathrm{H}})\overline{{}^t\mathrm{U}} = \mathrm{UHU}^{-1}$$

donc l'application φ_U définie par conjugaison par U est un endomorphisme de $\mathbb{H}_0$.

De plus, on a $\det \varphi_\mathrm{U}(\mathrm{H}) = \det \mathrm{H}$ donc $\varphi_\mathrm{U} \in \mathrm{O}(\mathbb{H}_0)$. Or $\varphi : \mathrm{SU}_2(\mathbb{C}) \to \mathrm{O}(\mathbb{H}_0)$ ainsi induite est un morphisme de groupes continu, et comme $\mathrm{SU}_2(\mathbb{C})$ est connexe, on a $\varphi\mathrm{SU}_2(\mathbb{C}) \subset \mathrm{SO}(\mathbb{H}_0)$.

Une matrice U est dans le noyau de φ si, pour tout $\mathrm{H} \in \mathbb{H}_0$, on a $\mathrm{UH} = \mathrm{HU}$. En considérant les matrices

$$\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

on trouve $\operatorname{Ker} \varphi = \{\pm \mathrm{Id}\}$.

Il reste à voir que φ est surjective. Pour cela, nous allons montré que $\operatorname{Im} \varphi$ contient tous les demi-tours. Soit $\mathrm{A} \in \mathbb{H}_0$ telle que $\det \mathrm{A} = 1$. Par le théorème de Cayley-Hamilton, on a $\mathrm{A}^2 = -\mathrm{I}$. Or, puisque $\mathrm{A} \in \mathbb{H}_0$, on a aussi $\mathrm{A} = -{}^t\overline{\mathrm{A}}$. Donc $\mathrm{A}\,{}^t\overline{\mathrm{A}} = -\mathrm{A}^2 = \mathrm{I}$ et $\mathrm{A} \in \mathrm{SU}_2(\mathbb{C})$. Par suite, φ_A n'est pas l'identité car $\mathrm{A} \neq \pm\mathrm{I}$, mais $\varphi_\mathrm{A} \circ \varphi_\mathrm{A}(\mathrm{H}) = \varphi_{\mathrm{A}^2}(\mathrm{H}) = \mathrm{H}$ car $\mathrm{A}^2 = -\mathrm{I}$ et $\varphi_\mathrm{A}(\mathrm{A}) = \mathrm{A}$. Donc φ_A est le demi-tour d'axe $\mathbb{R}\mathrm{A}$. ∎

Théorème 4.6. — *On a*

$$\mathrm{SU}_2 \times \mathrm{SU}_2/\{\pm(\mathrm{I}, \mathrm{I})\} \approx \mathrm{SO}_4\,.$$

DÉMONSTRATION. — On fait agir $\mathrm{SU}_2 \times \mathrm{SU}_2$ sur $\mathbb{H}$ par $(\mathrm{U}, \mathrm{V}) \cdot \mathrm{H} \mapsto \mathrm{UHV}^{-1}$. Cette fonction envoie $\mathrm{SU}_2 \times \mathrm{SU}_2$ dans O_4, mais comme $\mathrm{SU}_2 \times \mathrm{SU}_2$ est connexe, on définit ainsi une application linéaire $\Phi : \mathrm{SU}_2 \times \mathrm{SU}_2 \to \mathrm{SO}_4$.

Si $(\mathrm{U}, \mathrm{V}) \in \operatorname{Ker} \Phi$, alors, pour tout $\mathrm{A} \in \mathbb{H}$, on a $\mathrm{UAV}^{-1} = \mathrm{A}$. En particulier, pour $\mathrm{A} = \mathrm{I}$, on obtient $\mathrm{U} = \mathrm{V}$. Ceci nous ramène à un cas connu... On en déduit que $\operatorname{Ker} \Phi = \{\pm \mathrm{Id}\}$.

Montrons la surjectivité. Il suffit de montrer que tout renversement est dans l'image de Φ. On se fixe donc un plan de $\mathbb{H}$ engendré par deux éléments orthogonaux et unitaires M_1 et M_2. Il existe $\mathrm{U} \in \mathrm{SU}_2(\mathbb{C})$ telle que $\mathrm{UM}_1 \in \mathbb{H}_0^\perp = \mathbb{R}\mathrm{I}$ (prendre $\mathrm{U} = \mathrm{M}_1^{-1}$, en remarquant que $\mathrm{M}_1 \in \mathrm{SU}_2$). Du coup, $\mathrm{UM}_2 \in \mathbb{H}_0$: on s'est ramené à la situation précédente : on cherche $(\mathrm{U}, \mathrm{V}) \in \mathrm{SU}_2$ tel que $\Phi(\mathrm{U}, \mathrm{V})$ fixe $\mathbb{H}_0^\perp$ et un élément $\mathrm{H} \in \mathbb{H}_0$. Il suffit de prendre $\mathrm{U} = \mathrm{V} = \mathrm{H}$. Pour conclure, on conjugue cette application par $\mathrm{M} \mapsto \mathrm{U}^{-1}\mathrm{M}$ qui est une isométrie, donc conjugue notre renversement à celui que nous cherchions. ∎

A Caractéristique d'un corps

Pour tout corps $\mathbb{K}$, on peut définir un morphisme canonique de $\mathbb{Z}$ dans $\mathbb{K}$, par la formule

$$\varphi(n) = n.1.$$

L'application φ est un morphisme d'anneaux, et son noyau est donc un idéal de $\mathbb{Z}$, en particulier un sous-groupe de $(\mathbb{Z}, +)$. Mais un tel sous-groupe est de la forme $n\mathbb{Z}$, $n \in \mathbb{N}$, d'après le cours de théorie des groupes.

Définition A.1. — *Soit $\mathbb{K}$ un corps. On appelle* caractéristique *de $\mathbb{K}$ l'unique entier $n \geq 0$ tel que $n\mathbb{Z}$ soit le noyau du morphisme d'anneaux φ.*

B Changement de bases

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E.

Définition B.1. — *On appelle* matrice de passage *de la base $\mathcal{B}$ à la base $\mathcal{B}'$ la matrice dont les coefficients sont les coordonnées des vecteurs de $\mathcal{B}'$ dans la base $\mathcal{B}$.*

Ainsi, en notant P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, on lit dans les colonnes de P les vecteurs de $\mathcal{B}'$ en fonction de $\mathcal{B}$: si pour tout j on a $\mathrm{e}'_j = a_{1j}\mathrm{e}_1 + \cdots + a_{nj}\mathrm{e}_n$,

$$\mathrm{P} = \begin{array}{c} \begin{array}{ccc} \mathrm{e}'_1 & \dots & \mathrm{e}'_n \end{array} \\ \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \end{array} .$$

On constate alors

$$\boxed{\text{matrice de passage de } \mathcal{B} \text{ à } \mathcal{B}' \;=\; \mathrm{Mat}(\mathrm{Id}_\mathrm{E}, \mathcal{B}', \mathcal{B})} .$$

Théorème B.1. — *Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de l'espace vectoriel* E*, et* X, X' *les vecteurs colonnes associés à x dans les bases $\mathcal{B}, \mathcal{B}'$ respectivement. Soit* P *la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Alors les vecteurs colonnes vérifient*

$$\mathrm{X} = \mathrm{PX}'$$

et donc la matrice exprimant la nouvelle base $\mathcal{B}'$ en fonction de l'ancienne $\mathcal{B}$ permet d'exprimer les coordonnées dans l'ancienne base en fonction des coordonnées dans la nouvelle. On dit qu'il y a contravariance.

Théorème B.2. — *Soit* E *un espace vectoriel de dimension finie n muni d'une base $\mathcal{B}$. Alors toute matrice inversible* P *d'ordre n est la matrice de passage de la base $\mathcal{B}$ à une base $\mathcal{B}_\mathrm{P}$ de* E*, et toute base $\mathcal{B}'$ de* E *s'exprime à l'aide d'une matrice inversible d'ordre n.*

On a donc une correspondance bi-univoque entre l'ensemble des bases de E *et les matrices inversibles d'ordre n.*

Applications linéaires et changement de base. Soient E et F deux espaces vectoriels de dimension finie, et f une application linéaire de E dans F. Soient $\mathcal{B} = \{\mathrm{e}_1, \dots, \mathrm{e}_n\}, \mathcal{B}' = \{\mathrm{e}'_1, \dots, \mathrm{e}'_n\}$ deux bases de E, et $\mathcal{C} = \{\varepsilon_1, \dots, \varepsilon_p\}, \mathcal{C}' = \{\varepsilon'_1, \dots, \varepsilon'_p\}$ deux bases de F. Notons P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, Q la matrice de passage de $\mathcal{C}$ à $\mathcal{C}'$, A la matrice de f dans les bases $\mathcal{B}$ et $\mathcal{C}$ et enfin B la matrice de f dans les bases $\mathcal{B}'$ et $\mathcal{C}'$.

On a $\mathrm{Y}' = \mathrm{Q}^{-1}\mathrm{APX}'$, d'où la relation **fondamentale** :

$$\boxed{\mathrm{B} = \mathrm{Q}^{-1}\mathrm{AP}}$$

En explicitant un peu plus cette formule, elle devient encore plus claire :

$$\mathrm{Mat}(f, \mathcal{B}', \mathcal{C}') = \mathrm{passage}(\mathcal{C}' \to \mathcal{C})\,\mathrm{Mat}(f, \mathcal{B}, \mathcal{C})\mathrm{passage}(\mathcal{B} \to \mathcal{B}') .$$

En particulier, si $\mathrm{E} = \mathrm{F}$, $\mathcal{B} = \mathcal{C}$ et $\mathcal{B}' = \mathcal{C}'$, alors

$$\mathrm{B} = \mathrm{P}^{-1}\mathrm{AP} .$$

References

[1] Alessandri, Thèmes de géométrie pour l'agrégation.

[2] R. Mneimné & F. Testard, Introduction à la théorie des groupes de Lie classiques, Collection Méthodes, Hermann, Paris, 1986.

[3] D. Perrin, Cours d'algèbre, édité avec la collaboration de Marc Cabanes et de Martine Duchene. Collection de l'Ecole Normale Supérieure de Jeunes Filles, **18**. Ecole Normale Supérieure de Jeunes Filles, Paris, 1982.

[4] P. Tauvel, Mathématiques générales pour l'agrégation, seconde édition. Masson, Paris, 1997.