Cours d'agrégation : algèbre et géométrie euclidienne et hermitienne

Frédéric Paulin

Professeur à la Faculté des sciences d'Orsay

Cours de seconde année de Master, parcours :
Formation à l'enseignement supérieur en mathématique à Orsay,

Université Paris-Saclay

Année 2022-2023

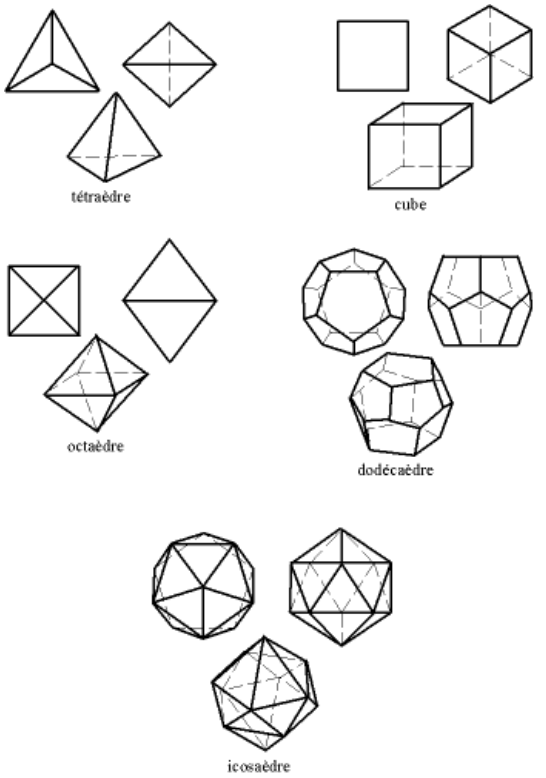


Table des matières

Préambule.

Le but de ce cours est de traiter les parties 4 (c), 4 (d) et 4 (e) du programme 2023, identique au programme 2022, de l'agrégation externe de mathématiques, correspondant à l'extrait officiel de programme suivant.

4 Formes bilinéaires et quadratiques sur un espace vectoriel

(c) Espaces vectoriels euclidiens, espaces vectoriels hermitiens. Isomorphisme d'un espace vectoriel euclidien avec son dual. Supplémentaire orthogonal. Inégalité de Cauchy-Schwarz. Norme. Bases orthonormales.

(d) Groupe orthogonal, groupe spécial orthogonal. Exemple de générateurs du groupe orthogonal : décomposition d'un automorphisme orthogonal en produit de réflexions. Endomorphismes symétriques, endomorphismes normaux. Diagonalisation d'un endomorphisme symétrique. Décomposition en valeurs singulières d'une matrice à coefficients réels. Réduction simultanée de deux formes quadratiques réelles, l'une étant définie positive. Décomposition polaire dans $\mathrm{GL}_n(\mathbb{R})$. Espaces vectoriels euclidiens de dimension deux, classification des éléments de $\mathrm{O}(2,\mathbb{R})$. Espaces vectoriels euclidiens de dimension trois, classification des éléments de $\mathrm{O}(3,\mathbb{R})$; produit mixte, produit vectoriel.

(e) Groupe unitaire, groupe spécial unitaire. Diagonalisation des endomorphismes normaux. Décomposition polaire dans GL(n, C).

Nous supposons acquis les points 1 à 3 dudit programme, mais nous retrouverons les points 4 (a), 4 (b) comme cas particuliers.

Les références recommandées sont les chapitres V à VIII de [Per1], les chapitres 8, 9 (hors affine), 11 et 12 de [Ber1, Ber2], et les chapitres V, VI, VII, X de [Deh]. Il est aussi utile de consulter [MT] et [AB], voire [Gou].

L'illustration en couverture est extraite de l'article de Wikipedia sur les polyèdres réguliers https ://fr.wikipedia.org/wiki/Polyèdres_régulier

1 Espaces quadratiques et quadratiques hermitiens

Nous renvoyons par exemple à [Per1, §V] pour le contenu des parties 1.1, 1.2, 1.3, 1.5, 1.6 et 1.7, et par exemple à [Ber1, §8.11] pour le contenu de la partie 1.11. Voir aussi [Gou, Chap. 5].

Soit k un corps commutatif[1] muni d'un automorphisme[2] involutif[3] σ de k. Nous utiliserons la notation exponentielle $\sigma : \lambda \mapsto \lambda^\sigma$ pour alléger les notations.[4] Nous serons uniquement intéressés par les cas où (k, σ) vaut $(\mathbb{R}, \mathrm{id})$ ou $(\mathbb{C}, z \mapsto \overline{z})$. Nous notons $k^\times = (k - \{0\}, \times)$ le groupe multiplicatif de k.

Lemme 1.1. *Si la caractéristique de k est différente de 2, et si l'involution σ n'est pas l'identité, alors*

- *si $k_0 = \{x \in k \ : \ x^\sigma = x\}$ est le sous-corps fixe de σ, alors l'extension de k_0 par k est de degré $[k : k_0] = 2$,*
- *il existe un élément $\iota \in k - k_0$ (en particulier ι est non nul) tel que $\iota^2 \in k_0$,*
- *nous avons $k = k_0[\iota] = \{a + \iota b \ : \ a, b \in k_0\}$*

et

- *nous avons $\iota^\sigma = -\iota$.*

Si $(k, \sigma) = (\mathbb{C}, z \mapsto \overline{z})$, nous avons $k_0 = \mathbb{R}$ et nous pouvons prendre $\iota = i$. Nous renvoyons à [Per1, §V] pour le cas de la caractéristique 2.

Démonstration. Puisque σ est un automorphisme de corps, l'ensemble k_0 de ses points fixes est un sous-corps de k. Il est distinct de k puisque $\sigma \neq \mathrm{id}$. Fixons $\iota_0 \in k - k_0$ et posons $\iota = \iota_0 - \iota_0^\sigma$. Alors $\iota \neq 0$ et $\iota^\sigma = -\iota$, donc $\iota \in k - k_0$. De plus $\iota^2 = -\,\iota\,\iota^\sigma$ est fixe par σ, donc appartient à k_0. Pour tout $x \in K$, nous pouvons écrire

$$x = \frac{x + x^\sigma}{2} + \iota\, \frac{x - x^\sigma}{2\,\iota} \ . \tag{1}$$

Puisque les éléments $\frac{x+x^\sigma}{2}$ et $\frac{x-x^\sigma}{2\,\iota}$ sont fixes par σ, le résultat en découle. □

1.1 Formes sesquilinéaires

Soit E un espace vectoriel sur k.

Définition 1.2. *Une* forme sesquilinéaire[5] *sur E est une application $f : E \times E \to k$ telle que*

1. En français, un corps n'est pas forcément commutatif. La structure des formes sesquilinéaires et des groupes unitaires sur les corps non commutatifs est tout à fait passionnante mais tout à fait hors programme de l'agrégation. Nous renvoyons par exemple à [Die, BT].
2. Un *automorphisme* d'un corps commutatif k est une bijection $f : k \to k$ telle que $f(x + y) = f(x) + f(y)$ et $f(xy) = f(x)f(y)$ pour tous les $x, y \in k$.
3. Une application g d'un ensemble dans lui-même est *involutive* si

$$g \circ g = \mathrm{id} \ ,$$

ou, de manière équivalente, si elle est bijective, et égale à son application réciproque. Le fait que l'automorphisme σ de k soit involutif est nécessaire pour que la relation d'orthogonalité soit toujours symétrique, et pour pouvoir toujours définir des formes sesquilinéaires hermitiennes, voir [Per1, §V].
4. Attention, si τ est un autre automorphisme involutif de corps de k, alors $\lambda^{\sigma\circ\tau} = (\lambda^\tau)^\sigma$.
5. On dit aussi *σ-sesquilinéaire* lorsqu'il convient de préciser σ.

(1) *pour tout $y \in E$, l'application $x \mapsto f(x,y)$ est linéaire,*

(2) *pour tout $x \in E$, l'application $y \mapsto f(x,y)$ est une forme* semi-linéaire,[6] *c'est-à-dire pour tous les $y, y' \in E$ et $\lambda \in k$, nous avons*

$$f(x, y + \lambda y') = f(x,y) + \lambda^\sigma f(x,y') \ .$$

Lorsque $\sigma = \mathrm{id}$, on dit simplement une forme bilinéaire (voir le cours de Joël Riou). Si $(k,\sigma) = (\mathbb{C}, z \mapsto \overline{z})$, on dit parfois *anti-linéaire* au lieu de semi-linéaire. Certaines références définissent les formes sesquilinéaires comme étant semi-linéaire à gauche et linéaire à droite, il vaut mieux surveiller (et expliciter) la convention utilisée.

Si f' est une forme sesquilinéaire sur un espace vectoriel E' sur k, nous dirons que f et f' sont *équivalentes* s'il existe un isomorphisme linéaire $v : E \to E'$ tel que

$$\forall\, x, y \in E, \quad f'(v(x), v(y)) = f(x,y) \ .$$

Un tel élément v s'appelle une *conjugaison* entre f et f'. La relation « être équivalent à » est une relation d'équivalence sur tout ensemble de formes sesquilinéaires sur k, en particulier sur l'ensemble des formes sesquilinéaires sur E. Le problème de la classification à équivalence près des formes sesquilinéaires sur les espaces vectoriels sur k est un problème fondamental. Il dépend fortement du corps involutif (k,σ). Nous y reviendrons dans la partie 1.7.

Interprétation matricielle.

Supposons que l'espace vectoriel E sur k soit de dimension $n \in \mathbb{N}$ et muni d'une base $\mathscr{B} = (e_1, \dots, e_n)$. Nous allons montrer qu'une forme sesquilinéaire f sur E est entièrement déterminée par la matrice[7]

$$A = \big(f(e_i, e_j)\big)_{1 \leq i,j \leq n} \in \mathscr{M}_n(k) \ ,$$

appelée la *matrice de* f dans la base $\mathscr{B}$, notée $\mathrm{Mat}_{\mathscr{B}}(f)$. Notons respectivement X et Y les matrices colonnes des coordonnées de deux vecteurs x et y de E dans cette base. Pour toute matrice $Z = \big(z_{i,j}\big)_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}}$ à coefficients dans k, notons $Z^\sigma = \big((z_{i,j})^\sigma\big)_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}}$ la matrice dont les coefficients sont les images par σ des coefficients de Z.[8] Avec notre convention sur les formes sesquilinéaires de la définition 1.2, nous avons alors

$$f(x,y) = {}^t X A\, Y^\sigma \ .$$

Cette formule montre que la forme sesquilinéaire f est déterminée par sa matrice A. Elle montre aussi que toute matrice $A \in \mathscr{M}_n(k)$ est la matrice dans la base $\mathscr{B}$ d'une forme sesquilinéaire sur E, en définissant f par la formule ci-dessus.

Soit f' une forme sesquilinéaire sur un espace vectoriel E' sur k de dimension finie $n' \in \mathbb{N}$, muni d'une base $\mathscr{B}' = (e'_1, \dots, e'_{n'})$. Le résultat suivant donne l'interprétation matricielle de l'équivalence des formes sesquilinéaires.

6. On dit aussi *σ-linéaire* lorsqu'il convient de préciser σ.

7. Si $m, n \in \mathbb{N} - \{0\}$, et si k' est un corps commutatif, nous notons $\mathscr{M}_n(k')$ l'algèbre des matrices $n \times n$ à coefficients dans k' et $\mathscr{M}_{n,m}(k')$ l'espace vectoriel des matrices $n \times m$ à coefficients dans k'.

8. Remarquons que $(AB)^\sigma = A^\sigma B^\sigma$ pour toutes les matrices multipliables A et B (c'est-à-dire telles que le nombre de colonnes de A soit égal au nombre de lignes de B), et que $\det(A^\sigma) = (\det A)^\sigma$ si A est une matrice carrée.

Proposition 1.3. *Les formes sesquilinéaires f et f' sont équivalentes si et seulement s'il existe une matrice $P \in \mathrm{GL}_n(k)$ telle que*

$$\mathrm{Mat}_{\mathscr{B}'}(f') = {}^tP \ \mathrm{Mat}_{\mathscr{B}}(f) \ P^\sigma \ .$$

La relation sur l'ensemble $\mathscr{M}_n(k)$ définie par $A \sim A'$ si et seulement s'il existe une matrice $P \in \mathrm{GL}_n(k)$ telle que

$$A' = {}^tP \ A \ P^\sigma$$

est une relation d'équivalence.[9] **Le lecteur ou la lectrice prendra bien garde** à ne pas confondre cette relation d'équivalence avec la relation « être semblable à » sur $\mathscr{M}_n(k)$, définie par $A \sim A'$ si et seulement s'il existe une matrice $Q \in \mathrm{GL}_n(k)$ telle que

$$A' = Q \ A \ Q^{-1} \ .$$

Démonstration. Notons $A = \mathrm{Mat}_{\mathscr{B}}(f)$ et $A' = \mathrm{Mat}_{\mathscr{B}'}(f')$.

Si $v : E \to E'$ est une conjugaison entre f et f', notons P la matrice de v dans les bases $\mathscr{B}$ de E et $\mathscr{B}'$ de E'. Comme v est un isomorphisme linéaire, les dimensions de E et de E' sont égales, et $P \in \mathrm{GL}_n(k)$. Pour tous les vecteurs x et y de E, en notant respectivement X et Y leur colonne de coordonnées dans la base $\mathscr{B}$, de sorte que PX et PY sont les colonnes de coordonnées de $v(x)$ et $v(y)$ dans la base $\mathscr{B}'$, nous avons

$$f'(v(x), v(y)) = {}^t(PX)A'(PY)^\sigma = {}^tX({}^tP \, A' \, P^\sigma)Y^\sigma \ .$$

Puisque

$$f(x, y) = {}^tX \, A \, Y^\sigma \ ,$$

la relation $f'(v(x), v(y)) = f(x, y)$ pour tous les $x, y \in E$ implique que $A = {}^tP \ A' \ P^\sigma$.

Réciproquement, supposons qu'il existe $P \in \mathrm{GL}_n(k)$ telle que $A = {}^tP \ A' \ P^\sigma$. En particulier, nous avons $n = n'$. Soit $v : E \to E'$ l'application linéaire dont la matrice dans les bases $\mathscr{B}$ de E et $\mathscr{B}'$ de E' est égale à P. Alors v est un isomorphisme linéaire car P est inversible, et les deux formules centrées ci-dessus montrent que $f'(v(x), v(y)) = f(x, y)$ pour tous les $x, y \in E$. □

Soient $\mathscr{B}' = (e'_1, \ldots, e'_n)$ une autre base de E et $P = P_{\mathscr{B},\mathscr{B}'}$ la *matrice de passage* de la base $\mathscr{B}$ à la base $\mathscr{B}'$, dont la j-ème colonne est la matrice colonne des coordonnées dans la base $\mathscr{B}$ du j-ème vecteur de $\mathscr{B}'$, pour $j = 1 \ldots, n$. Ceci implique que si X' est la matrice colonne de tout élément $x \in E$ dans la base $\mathscr{B}'$, alors $X = PX'$. Par conséquent[10],

$$\mathrm{Mat}_{\mathscr{B}'}(f) = {}^tP \ \mathrm{Mat}_{\mathscr{B}}(f) \ P^\sigma \ . \tag{2}$$

En particulier,

$$\det \mathrm{Mat}_{\mathscr{B}'}(f) = (\det P)(\det P)^\sigma \ \det \mathrm{Mat}_{\mathscr{B}}(f) \ . \tag{3}$$

9. Ceci découle du fait que la relation « être équivalente à » sur les formes sesquilinéaires est une relation d'équivalence, mais cela se montre facilement en utilisant le fait que pour tous les $P, Q \in \mathrm{GL}_n(k)$, nous avons ${}^t(P^{-1}) = ({}^tP)^{-1}$, $(P^{-1})^\sigma = (P^\sigma)^{-1}$, ${}^t(Q\,P) = {}^tP \, {}^tQ$ et $(Q\,P)^\sigma = Q^\sigma P^\sigma$.

10. Ceci découle de la proposition 1.3, mais un argument direct est le suivant : avec les notations ci-dessus, si $A' = \mathrm{Mat}_{\mathscr{B}'}(f)$, pour tous les $x, y \in E$, nous avons

$${}^tX' \, A' \, Y'^\sigma = f(x, y) = {}^tX A \, Y^\sigma = {}^t(PX') \, A \, (PY')^\sigma = {}^tX'({}^tP A \, P^\sigma) \, Y'^\sigma \ .$$

Discriminant.

Supposons encore dans cette sous-partie que l'espace vectoriel E est de dimension $n \in \mathbb{N}$ et muni d'une base $\mathscr{B} = (e_1, \dots, e_n)$.

Appelons *norme*[11] de k un élément de k de la forme $N(\lambda) = \lambda\lambda^\sigma$ où $\lambda \in k^\times$. Remarquons que $1 = 1\,1^\sigma$ est une norme et que

$$\forall\, x, y \in k, \quad N(xy) = (xy)(xy)^\sigma = xx^\sigma yy^\sigma = N(x)N(y)\ .$$

Notons $N(k^\times)$ le groupe multiplicatif[12] des normes non nulles de k, qui agit par multiplication sur k. La classe dans l'ensemble quotient $k/N(k^\times)$ du déterminant de la matrice de f dans la base $\mathscr{B}$ est indépendante du choix de la base $\mathscr{B}$ par la formule (3). Elle est appelée le *discriminant* de f, et notée $\mathrm{Disc}(f)$:

$$\mathrm{Disc}(f) = \left[\det \mathrm{Mat}_{\mathscr{B}}(f)\right] \ \in\ k/N(k^\times)\ .$$

Si $\sigma = \mathrm{id}$, il est donc bien défini à un carré près. Si (k, σ) vaut $(\mathbb{R}, \mathrm{id})$ ou $(\mathbb{C}, z \mapsto \overline{z})$, le discriminant est bien défini modulo multiplication par un réel strictement positif, et $k/N(k^\times)$ s'identifie respectivement à $\{-1, 0, 1\}$ et au cercle $\mathbb{S}_1 = \{z \in \mathbb{C} : |z| = 1\}$.

Isomorphisme de dualité.

Nous renvoyons par exemple à l'exercice E.20 pour des rappels sur la dualité linéaire. Un espace vectoriel F de dimension finie et son espace vectoriel dual $F^* = \mathscr{L}(F, k)$, ayant même dimension, sont isomorphes. Mais il existe en général de nombreux isomorphismes linéaires entre F et F^*. La donnée d'une forme sesquilinéaire non dégénérée (au sens ci-dessous, comme pour les produits scalaires que nous introduirons dans la partie suivante) permet de définir un isomorphisme particulier, du moins après une petite modification liée à la présence de l'involution σ.

Notons E^σ l'espace vectoriel sur k ayant le même groupe additif sous-jacent que E (et en particulier même ensemble sous-jacent),

$$(E^\sigma, +) = (E, +)\ ,$$

et dont la multiplication externe est l'application de $k \times E$ dans E définie par

$$(\lambda, x) \mapsto \lambda^\sigma\, x\ .$$

Notons que $(E^\sigma)^\sigma = E$, que E et E^σ ont les mêmes sous-espaces vectoriels, et que les formes sesquilinéaires sur E sont les formes bilinéaires sur l'espace vectoriel produit $E \times E^\sigma$.

Pour tout $y \in E$, l'application $f_y : x \mapsto f(x, y)$ est une forme linéaire sur E. Si E^* est l'espace vectoriel dual de E, nous avons une application linéaire[13]

$$\begin{array}{rcl} \widetilde{f} : E^\sigma & \to & E^* \\ y & \mapsto & f_y : x \mapsto f(x, y)\ , \end{array}$$

qui caractérise f, puisque pour tous les $x, y \in E$, nous avons $f(x, y) = \widetilde{f}(y)(x)$.

Remarque 1.4. Si l'espace vectoriel E est de dimension $n \in \mathbb{N}$ et muni d'une base $\mathscr{B} = (e_1, \dots, e_n)$, si $\mathscr{B}^* = (e_1^*, \dots, e_n^*)$ est la base de l'espace vectoriel dual E^* duale[14] à

11. Attention, une norme dans le corps involutif $(\mathbb{C}, z \mapsto \overline{z})$ est un carré de valeur absolue : $N(z) = |z|^2$ pour tout $z \in \mathbb{C}$. On ne la confondra pas avec la norme associé au produit scalaire euclidien usuel de $\mathbb{C}$, qui est $\|z\| = |z| = \sqrt{(\mathrm{Re}\ z)^2 + (\mathrm{Im}\ z)^2}$.

12. C'est un sous-groupe du groupe multiplicatif $k_0^\times$.

13. C'est l'intérêt d'avoir introduit E^σ, car en tant qu'application de E dans E^*, cette application $\widetilde{f}$ est seulement semi-linéaire : $\widetilde{f}(y + \lambda y') = \widetilde{f}(y) + \lambda^\sigma\, \widetilde{f}(y')$ pour tous les $y, y' \in E$.

14. Voir l'exercice E.20 pour des rappels sur la dualité linéaire.

$\mathscr{B}$, alors $\mathscr{B}$ est encore une base de l'espace vectoriel E^σ, et la matrice de $\widetilde{f}$ dans les bases $\mathscr{B}$ de E^σ et $\mathscr{B}^*$ de E^* est liée à la matrice de f dans la base $\mathscr{B}$ de la manière suivante :

$$\mathrm{Mat}_{\mathscr{B},\mathscr{B}^*}(\widetilde{f}) = {}^t\mathrm{Mat}_{\mathscr{B}}(f)\ .$$

Démonstration. Soient x et y des vecteurs de E, de matrices colonnes des coordonnées X et Y respectivement dans la base $\mathscr{B}$ de E. La matrice colonne des coordonnées dans la base $\mathscr{B}^*$ de la forme linéaire $\widetilde{f}(x)$ est $\mathrm{Mat}_{\mathscr{B},\mathscr{B}^*}(\widetilde{f})X$. La matrice colonne des coordonnées dans la base $\mathscr{B}$ de E^σ du vecteur $y \in E^\sigma$ est Y^σ. Donc le scalaire $\widetilde{f}(x)(y)$ est égal à ${}^t\big(\mathrm{Mat}_{\mathscr{B},\mathscr{B}^*}(\widetilde{f})X\big)Y^\sigma$. Comme $\widetilde{f}(x)(y) = f(x,y) = {}^tX\ \mathrm{Mat}_{\mathscr{B}}(f)\ Y^\sigma$, et ceci pour tous les $x, y \in E$, le résultat en découle. □

Une forme sesquilinéaire f est dite *non dégénérée* si l'application $\widetilde{f}$ est injective, ou, de manière équivalente, si le noyau de $\widetilde{f}$

$$\ker\ \widetilde{f} = \{y \in E\ :\ \forall\, x \in E,\quad f(x,y) = 0\}$$

est nul (réduit à $\{0\}$). Ce noyau est appelé par abus le *noyau* de f, et noté $\ker\ f$.[15] Le *rang* de f est défini comme la dimension de l'image de $\widetilde{f}$.

Lorsque la dimension de E est finie, le rang de f est la codimension du noyau de f, ainsi que le rang de la matrice de f (dans n'importe quelle base de E). De plus, demander que f soit non dégénérée équivaut alors à demander que $\widetilde{f}$ soit bijective, ou que la matrice de f (dans n'importe quelle base de E) soit inversible, ou que le discriminant de f (dans n'importe quelle base de E) soit non nul.

Ainsi, si la dimension de E est finie et si f est non dégénérée, alors l'application

$$\widetilde{f} : E^\sigma \xrightarrow{\sim} E^*$$

ci-dessus est un isomorphisme linéaire, appelé la *dualité* induite par f.

1.2 Formes symétriques, alternées ou hermitiennes

Soient E un espace vectoriel sur k et f une forme sesquilinéaire sur E.

• La forme f est dite *symétrique* si $\sigma = \mathrm{id}$ et si $f(y,x) = f(x,y)$ pour tous les $x, y \in E$. L'application $q : E \to k$ définie par $x \mapsto q(x) = f(x,x)$ est appelée la *forme quadratique associée* à f, et f est appelée une *forme polaire* de q.

Définition 1.5. *Une* forme quadratique *est une application $q : E \to k$ telle qu'il existe une forme bilinéaire symétrique f qui est une forme polaire de q.*

Si E' est un espace vectoriel sur k, deux formes quadratiques q sur E et q' sur E' sont dites *équivalentes* s'il existe un isomorphisme linéaire $v : E' \to E$ tel que $q' = q \circ v$. En particulier, si q et q' ont des formes polaires f et f' équivalentes, alors q et q' sont équivalentes. La relation « être équivalente à » est une relation d'équivalence sur tout ensemble de formes quadratiques sur k, en particulier sur l'ensemble $\mathscr{Q}(E)$ de toutes les formes quadratiques sur E.

15. Bien noter qu'il s'agit d'une convention de notation, car si E n'est pas nul, alors f n'est pas une application linéaire !

En caractéristique différente de 2, nous avons les *formules de polarisation*

$$f(x,y) = \frac{1}{2}(q(x+y) - q(x) - q(y)) \tag{4}$$
$$= \frac{1}{4}(q(x+y) - q(x-y)) \ , \tag{5}$$

donc q détermine f, et q admet une unique forme polaire. De plus, deux formes quadratiques sont équivalentes si et seulement si leurs formes polaires sont équivalentes. [16]

- La forme f est dite *antisymétrique* si $\sigma = \mathrm{id}$ et si $f(y,x) = -f(x,y)$ pour tous les $x, y \in E$.
- La forme f est dite *alternée* si $f(x,x) = 0$ pour tout $x \in E$.
- La forme f est dite *hermitienne* si $f(y,x) = (f(x,y))^\sigma$ pour tous les $x, y \in E$. L'application $q : E \to k$ définie par $x \mapsto q(x) = f(x,x)$ est appelée la *forme quadratique hermitienne associée* à f, et f est appelée une *forme polaire* de q.
- La forme f est dite *antihermitienne* si $f(y,x) = -(f(x,y))^\sigma$ pour tous les $x, y \in E$.
- La forme f est dite *positive* si $f(x,x) \in \{\lambda\lambda^\sigma \ : \ \lambda \in k\}$ pour tout $x \in E$. Lorsque $(k,\sigma) = (\mathbb{R}, \mathrm{id})$ ou $(k,\sigma) = (\mathbb{C}, z \mapsto \overline{z}\,)$, ceci revient à demander que $f(x,x) \geq 0$ pour tout $x \in E$.

Définition 1.6. *Une* forme quadratique hermitienne *est une application $q : E \to k$ telle qu'il existe une forme sesquilinéaire hermitienne f qui est une forme polaire de q.*

Certains ouvrages omettent le mot « quadratique » dans la terminologie. Si E' est un espace vectoriel sur k, deux formes quadratiques hermitiennes q sur E et q' sur E' sont dites *équivalentes* s'il existe un isomorphisme linéaire $u : E' \to E$ tel que $q' = q \circ u$. En particulier, si q et q' ont des formes polaires f et f' équivalentes, alors q et q' sont équivalentes. La relation « être équivalente à » est une relation d'équivalence sur tout ensemble de formes quadratiques hermitiennes sur k, en particulier sur l'ensemble $\mathscr{H}(E)$ de toutes les formes quadratiques hermitiennes sur E.

Soient q et f comme dans la définition 1.6. Notons que $q(x)$ appartient au sous-corps fixe de σ pour tout $x \in E$ (donc à $\mathbb{R}$ si $(k,\sigma) = (\mathbb{C}, z \mapsto \overline{z}\,)$). [17] En caractéristique différente de 2 avec $\sigma \neq \mathrm{id}$, si ι est comme dans le lemme 1.1, alors nous avons, par la formule (1), les *formules de polarisation* (6) et (7) ci-dessous de la forme sesquilinéaire hermitienne f :

$$f(x,y) = \frac{f(x,y) + f(x,y)^\sigma}{2} + \iota\,\frac{f(x,y) - f(x,y)^\sigma}{2\iota}$$
$$= \frac{f(x,y) + f(y,x)}{2} + \frac{f(\iota x, y) + f(y, \iota x)}{2\iota}$$
$$= \frac{1}{2}(q(x+y) - q(x) - q(y)) + \frac{1}{2\iota}(q(\iota x + y) - q(\iota x) - q(y)) \tag{6}$$
$$= \frac{1}{4}(q(x+y) - q(x-y)) + \frac{1}{4\iota}(q(\iota x + y) - q(\iota x - y)) \ . \tag{7}$$

Donc q détermine f, et q admet une unique forme polaire. En caractéristique 2, il est important de travailler avec le couple (q,f), car q ne détermine pas forcément f.

16. En caractéristique 2, il est important de travailler avec le couple (q,f), car f n'est pas forcément déterminée par q, voir par exemple [BT].

17. En effet, puisque f est hermitienne, pour tout $x \in E$, nous avons $q(x)^\sigma = f(x,x)^\sigma = f(x,x) = q(x)$.

Remarques. (1) Si $\sigma = \mathrm{id}$, les formes sesquilinéaires hermitiennes sont les formes bilinéaires symétriques.

(2) En caractéristique 2, antisymétrique équivaut à symétrique. En caractéristique différente de 2, antisymétrique implique alterné, car $f(x,x) = -f(x,x)$, donc $2f(x,x) = 0$.

(3) Si f est alternée non nulle, alors $\sigma = \mathrm{id}$ et f est antisymétrique. En effet, pour tous les $x, y \in E$, nous avons

$$0 = f(x+y, x+y) = f(x,x) + f(x,y) + f(y,x) + f(y,y) = f(x,y) + f(y,x) \; .$$

De plus, soient $x, y \in E$ tels que $f(x,y) \neq 0$. Alors pour tout $\lambda \in k$, nous avons

$$\lambda f(x,y) = f(\lambda x, y) = -f(y, \lambda x) = -\lambda^\sigma f(y,x) = \lambda^\sigma f(x,y) \; ,$$

d'où par simplification $\lambda = \lambda^\sigma$ et $\sigma = \mathrm{id}$.

(4) Si f est symétrique ou hermitienne, nous appellerons *noyau, matrice, rang, discriminant* de la forme quadratique ou quadratique hermitienne associée q ceux de f, et nous dirons que q est *non dégénérée* si f l'est.

Pour tous les $x \in E$ et $\lambda \in k$, nous avons

$$q(\lambda x) = \lambda \lambda^\sigma q(x) \; . \tag{8}$$

En particulier, $q(\lambda x) = \lambda^2 \, q(x)$ si f est symétrique et

$$q(\lambda x) = |\, \lambda \,|^2 \, q(x)$$

si $(k, \sigma) = (\mathbb{C}, z \mapsto \overline{z})$.

Espaces euclidiens et hermitiens.

Supposons que $(k, \sigma) = (\mathbb{R}, \mathrm{id})$ (respectivement $(k, \sigma) = (\mathbb{C}, z \mapsto \overline{z})$). Nous dirons qu'une forme quadratique (respectivement quadratique hermitienne) q sur E, et sa forme polaire, sont

- *positives*[18] si $q(x) \geq 0$ pour tout $x \in E$,
- *définies positives* si $q(x) > 0$ pour tout $x \in E - \{0\}$.

Dans ce dernier cas, la forme polaire de q est appelée un *produit scalaire (euclidien)* (respectivement *produit scalaire (hermitien)*) sur E, noté $(x,y) \mapsto \langle x, y \rangle$.

Il est remarquable[19] que si E' est un sous-espace vectoriel de E, alors la restriction d'une forme quadratique (respectivement quadratique hermitienne) définie positive q à E' est encore définie positive, et la restriction du produit scalaire à $E' \times E'$ est encore un produit scalaire de E'.

La *norme associée* au produit scalaire est l'application $\| \cdot \| : E \to \mathbb{R}$ définie par

$$\|x\| = \sqrt{q(x)} = \sqrt{\langle x, x \rangle} \; .$$

La *distance associée* au produit scalaire de E est la distance associée à la norme associée au produit scalaire, c'est-à-dire l'application $d : E \times E \to \mathbb{R}$ définie par

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

Les propriétés suivantes sont bien connues.

18. Certaines références disent *semi-définies positives*. Cette définition coïncide avec celle donnée juste avant la définition 1.6 lorsque $(k, \sigma) = (\mathbb{R}, \mathrm{id})$ ou $(k, \sigma) = (\mathbb{C}, z \mapsto \overline{z}\,)$

19. par exemple parce que le caractère non dégénéré n'est pas forcément préservé par restriction à un sous-espace vectoriel

(i) L'application $\|\cdot\| : E \to \mathbb{R}$ est une norme : elle est positive et ne s'annule que sur le vecteur nul, et pour tous les vecteurs x, y et pour tout scalaire λ,

$$\|x+y\| \leq \|x\| + \|y\|, \quad \|\lambda x\| = |\lambda| \, \|x\|$$

Comme toute distance associé à une norme, l'application $d : E \times E \to \mathbb{R}$ est une distance.

Munissons E d'un produit scalaire $\langle\ ,\ \rangle$ de norme associée $\|\ \|$ et de distance associéee d, et soit E' un espace vectoriel sur k muni d'un produit scalaire $\langle\ ,\ \rangle'$ de norme associée $\|\ \|'$ et de distance associée d'. Une *isométrie vectorielle* de E dans E' est un isomorphisme linéaire $\varphi : E \to E'$ qui est une *isométrie* entre les distances associées d et d', c'est-à-dire qui vérifie

$$\forall\, x, y \in E, \quad d'(\varphi(x), \varphi(y)) = d(x, y)\ .$$

De manière équivalente par linéarité, une isométrie vectorielle est un isomorphisme linéaire $\varphi : E \to E'$ qui préserve les normes de E et E' :

$$\forall\, x \in E, \quad \|\, \varphi(x) \,\|' = \|\, x \,\|\ .$$

Une *similitude vectorielle* de E dans E' est un isomorphisme linéaire $\varphi : E \to E'$ tel qu'il existe une constante $\rho > 0$ telle que

$$\forall\, x, y \in E, \quad d'(\varphi(x), \varphi(y)) = \rho\, d(x, y)\ .$$

Cette constante ρ est unique si E est non nul, appelée le *rapport* de la similitude φ.

(ii) **(Inégalité de Cauchy-Schwarz et cas d'égalité)** Pour tous les vecteurs x et y de E, nous avons l'inégalité [20]

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

avec égalité si et seulement si x et y sont colinéaires.

(iii) **(Formule du parallélogramme)** [21] Pour tous les vecteurs x et y de E, nous avons

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

20. Il est important de remarquer que cette inégalité sans son cas d'égalité n'utilise que la positivité : si q est une forme quadratique ou quadratique hermitienne positive de forme polaire f, alors

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

En particulier, pour une forme quadratique ou quadratique hermitienne q positive, l'ensemble de ses vecteurs isotropes (c'est-à-dire annulant q) coïncide avec son noyau : pour tout $x \in E$, nous avons $q(x) = 0$ si et seulement si $f(x, y) = 0$ pour tout $y \in E$.

Remarquons aussi que si q est une forme quadratique sur $\mathbb{R}$ ou quadratique hermitienne sur $\mathbb{C}$, alors q ou $-q$ est positive si et seulement si le noyau de q coïncide avec l'ensemble de ses vecteurs isotropes (c'est-à-dire annulant q). En effet, supposons que ces deux ensembles coïncident, soit E' un supplémentaire du noyau de q. Si q n'est pas de signe constant, alors il existe $x, y \in E'$ vérifiant $q(x) < 0$ et $q(y) > 0$. En particulier x et y sont non colinéaires. Par le théorème des valeurs intermédiaires, q s'annule sur un point z du segment affine entre x et y (image de l'application de $[0, 1]$ dans E définie par $t \mapsto t\,x + (1-t)y$), qui n'est pas le vecteur nul car x et y sont non colinéaires. Donc z est un élément non nul qui appartient à la fois à E' et au noyau, ce qui contredit le fait que ces sous-espaces vectoriels sont supplémentaires.

21. La somme des carrés des longueurs des quatre côtés d'un parallélogramme est égale à la somme des carrés des longueurs de ses deux diagonales.

Un *espace euclidien*[22] (respectivement *hermitien*) est un espace vectoriel réel (respectivement complexe) de dimension finie munie d'un produit scalaire euclidien (respectivement hermitien) ou, de manière équivalente, d'une forme quadratique (respectivement quadratique hermitienne) définie positive.

Par exemple, l'*espace euclidien usuel* $\mathbb{R}^n$ est l'espace vectoriel réel $\mathbb{R}^n$ muni de son *produit scalaire usuel*

$$\langle (x_1,\dots,x_n),(y_1,\dots,y_n)\rangle = \sum_{i=1}^n x_i\, y_i \ ,$$

de norme associée $\|(x_1,\dots,x_n)\| = \sqrt{\sum_{i=1}^n x_i^2}$. L'*espace hermitien usuel* $\mathbb{C}^n$ est l'espace vectoriel complexe $\mathbb{C}^n$ muni de son *produit scalaire hermitien usuel*

$$\langle (w_1,\dots,w_n),(z_1,\dots,z_n)\rangle = \sum_{i=1}^n w_i\, \overline{z_i} \ ,$$

de norme associée $\|(z_1,\dots,z_n)\| = \sqrt{\sum_{i=1}^n |z_i|^2}$.

Interprétation matricielle.

Jusqu'à la fin de la partie 1.2, nous supposons que E est de dimension finie $n \in \mathbb{N}-\{0\}$, nous fixons une base $\mathscr{B} = (e_1,\dots,e_n)$ de E, et nous notons $A = (a_{i,j})_{1\leq i,j\leq n}$ la matrice de f dans la base $\mathscr{B}$. Alors

- f est symétrique si et seulement si $\sigma = \mathrm{id}$ et ${}^tA = A$.
- f est antisymétrique si et seulement si $\sigma = \mathrm{id}$ et ${}^tA = -A$.
- f est hermitienne si et seulement si ${}^tA = A^\sigma$ ou, de manière équivalente, si la matrice A est *auto-adjointe*, c'est-à-dire si ${}^tA^\sigma = A$. En particulier, si $(k,\sigma) = (\mathbb{C}, z\mapsto \overline{z})$, les coefficients diagonaux d'une matrice auto-adjointe sont réels.
- f est antihermitienne si et seulement si ${}^tA = -A^\sigma$ ou, de manière équivalente, si la matrice A est *anti-auto-adjointe*, c'est-à-dire si ${}^tA^\sigma = -A$. En particulier, lorsque $(k,\sigma) = (\mathbb{C}, z\mapsto \overline{z})$, les coefficients diagonaux d'une matrice anti-auto-adjointe A sont imaginaires purs.

Pour tout $n \in \mathbb{N}$, nous notons[23]

$$\mathrm{Sym}_n(k) = \{A = (a_{i,j})_{1\leq i,j\leq n} \in \mathscr{M}_n(k) \ : \ {}^tA = A \Leftrightarrow \forall\, i,j \in \{1,\dots,n\},\quad a_{ji} = a_{ij}\} \ ,$$

$$\begin{aligned}\mathrm{Asym}_n(k) = \{A = (a_{i,j})_{1\leq i,j\leq n} \in \mathscr{M}_n(k) \ : \ & {}^tA = -A \\ & \Leftrightarrow \forall\, i,j \in \{1,\dots,n\},\quad a_{ji} = -a_{ij}\} \ ,\end{aligned}$$

$$\begin{aligned}\mathrm{Herm}_n(k) = \{A = (a_{i,j})_{1\leq i,j\leq n} \in \mathscr{M}_n(k) \ : \ & {}^tA = A^\sigma \\ & \Leftrightarrow \forall\, i,j \in \{1,\dots,n\},\quad a_{ji} = (a_{ij})^\sigma\} \ ,\end{aligned}$$

$$\begin{aligned}\mathrm{Aherm}_n(k) = \{A = (a_{i,j})_{1\leq i,j\leq n} \in \mathscr{M}_n(k) \ : \ & {}^tA = -A^\sigma \\ & \Leftrightarrow \forall\, i,j \in \{1,\dots,n\},\quad a_{ji} = -(a_{ij})^\sigma\} \ .\end{aligned}$$

22. On dit *espace préhilbertien réel* (respectivement *complexe*) si on enlève l'hypothèse de dimension finie. Un *espace de Hilbert réel* (respectivement *complexe* est un espace préhilbertien réel (respectivement complexe) dont la norme (c'est-à-dire la distance associée) est complète.

23. Dans la littérature, en particulier agrégative, les espaces vectoriels $\mathrm{Sym}_n(k)$ et $\mathrm{Asym}_n(k)$ sont souvent notés simplement $\mathscr{S}_n(k)$ et $\mathscr{A}_n(k)$

Alors $\mathrm{Sym}_n(k)$ et $\mathrm{Asym}_n(k)$ sont des sous-espaces vectoriels de $\mathscr{M}_n(k)$, de dimension respectivement $\frac{n(n+1)}{2}$ et $\frac{n(n-1)}{2}$. Mais attention, $\mathrm{Herm}_n(k)$ et $\mathrm{Aherm}_n(k)$ sont des espaces vectoriels sur le corps fixe k_0 de σ, de dimension n^2 sur k_0, qui, si $\sigma \neq \mathrm{id}$ et $n \geq 1$, ne sont pas des sous-espaces vectoriels sur k de $\mathscr{M}_n(k)$. Nous avons les décompositions en somme directe

$$\mathscr{M}_n(k) = \mathrm{Sym}_n(k) \oplus \mathrm{Asym}_n(k)$$

en tant qu'espaces vectoriels sur k, donnée par $X = \frac{X+{}^tX}{2} + \frac{X-{}^tX}{2}$, et

$$\mathscr{M}_n(k) = \mathrm{Herm}_n(k) \oplus \mathrm{Aherm}_n(k)$$

en tant qu'espaces vectoriels sur k_0 (attention !), donnée par $X = \frac{X+X^\sigma}{2} + \frac{X-X^\sigma}{2}$.

- **Expression semi-homogène de degré** 2. Supposons que f soit symétrique ou hermitienne, et notons q sa forme quadratique ou quadratique hermitienne associée. Pour tout $x = \sum_{i=1}^n x_i e_i \in E$, nous avons

$$q(x) = \sum_{i,j=1}^n a_{i,j}\, x_i\, (x_j)^\sigma \ .$$

Par exemple, si $a_1, \ldots, a_n \in k$, alors

$$q(x) = \sum_{i=1}^n a_i\, x_i^2$$

est une forme quadratique, dite *diagonale* dans la base $\mathscr{B}$, associée à la forme bilinéaire symétrique $f(x,y) = \sum_{i=1}^n a_i\, x_i\, y_i$. Si $(k,\sigma) = (\mathbb{R}, \mathrm{id})$, cette forme quadratique est positive (respectivement définie positive) si et seulement si les coefficients a_i sont positifs (respectivement strictement positifs).

Par exemple, si $(k,\sigma) = (\mathbb{C}, z \mapsto \overline{z})$ et $a_1, \ldots, a_n \in \mathbb{R}$, [24] alors

$$q(z) = \sum_{i=1}^n a_i\, |z_i|^2$$

est une forme quadratique hermitienne, dite *diagonale* dans la base $\mathscr{B}$, associée à la forme sesquilinéaire hermitienne $f(w,z) = \sum_{i=1}^n a_i\, w_i\, \overline{z_i}$. Elle est positive (respectivement définie positive) si et seulement si les coefficients a_i sont positifs (respectivement strictement positifs).

Exercice E.1. [25] *Si f est une forme quadratique antisymétrique nondégénérée sur un espace vectoriel E sur k de dimension finie, montrer que la dimension de E est paire.*

1.3 Orthogonalité et isotropie

Soient E un espace vectoriel sur k et f une forme sesquilinéaire sur E, qui est symétrique, alternée ou hermitienne. Supposons la caractéristique de k différente de 2.

24. Attention à ne pas prendre les a_i complexes quelconques, une forme quadratique hermitienne sur un espace vectoriel complexe est à valeurs dans $\mathbb{R}$!

25. Voir la partie 1.13 pour des indications de solutions.

Orthogonalité.

Deux éléments $x, y \in E$ sont dits *orthogonaux* pour f si $f(x, y) = 0$. Lorsque f est sous-entendue, on écrit alors $x \perp y$. La relation « être orthogonal à » est symétrique.

Soient A et B deux parties de E. Nous dirons que A et B sont *orthogonales* pour f, et nous noterons $A \perp B$ lorsque f est sous-entendue, si $x \perp y$ pour tous les $x \in A$ et $y \in B$. L'*orthogonal* de A pour f est défini par

$$A^\perp = \{x \in E \ : \ \forall\, a \in A, \quad x \perp a\} \ .$$

Par exemple, le noyau de f est l'orthogonal de E :

$$\ker f = E^\perp \ .$$

Nous avons les propriétés élémentaires suivantes.

- La partie $A^\perp$ est un sous-espace vectoriel de E, par la linéarité à gauche de f.
- Nous avons $A \subset (A^\perp)^\perp$ par la symétrie de la relation « être orthogonal à », mais il n'y a pas toujours égalité, par exemple si A n'est pas un sous-espace vectoriel.
- Nous avons $A^\perp = (\operatorname{Vect} A)^\perp$, par la semi-linéarité à droite de f. En particulier, pour calculer l'orthogonal d'un sous-espace vectoriel, il suffit de calculer l'orthogonal d'une partie génératrice de ce sous-espace.
- Si $A \subset B$, alors $B^\perp \subset A^\perp$.

Si E est de dimension finie n et si f est non dégénérée, alors pour tous les sous-espaces vectoriels A et B de E,

- si la dimension de A est p, alors la dimension de $A^\perp$ est $n - p$, car si $(e_1, \ldots, e_p)$ est une base de A, alors $A^\perp$ est l'intersection des noyaux des p formes linéaires $x \mapsto f(x, e_i)$, qui sont linéairement indépendantes car f étant non dégénérée, l'application linéaire de dualité $\widetilde{f} : E^\sigma \to E^*$ est bijective,
- nous avons $A = (A^\perp)^\perp$, par inclusion et égalité des dimensions,
- nous avons $(A + B)^\perp = A^\perp \cap B^\perp$, par inclusion et égalité des dimensions,
- nous avons $A^\perp + B^\perp = (A \cap B)^\perp$, par inclusion et égalité des dimensions.

Isotropie.

Un élément $x \in E$ est dit *isotrope* pour f si $f(x, x) = 0$. Par exemple, f est alternée si et seulement si tout vecteur de E est isotrope pour f. Un sous-espace vectoriel A de E est dit *isotrope* si $A \cap A^\perp$ est non nul, ou, de manière équivalente, si la restriction de f à $A \times A$ est dégénérée. Notons que A est isotrope si et seulement si $A^\perp$ l'est. Un sous-espace vectoriel A de E est dit *totalement isotrope* si $A \subset A^\perp$ ou, de manière équivalente, si la restriction de f à $A \times A$ est nulle. Par exemple, le noyau de f est totalement isotrope.

Si E est de dimension finie, l'*indice* $\nu(f)$ de f est le maximum des dimensions des sous-espaces vectoriels A totalement isotropes de E. Si f est non dégénérée, comme nous avons $\dim A \leq \dim A^\perp$ si A est totalement isotrope et $\dim A^\perp = \dim E - \dim A$ comme vu ci-dessus, nous avons donc

$$\nu(f) \leq \ \frac{1}{2} \dim E \ . \tag{9}$$

La forme sesquilinéaire f est dite *anisotrope* ou *définie* si $\nu(f) = 0$, ou, de manière équivalente, si pour tout $x \in E$, nous avons $f(x, x) = 0$ si et seulement si $x = 0$. Elle est dite *isotrope* ou *indéfinie* si $\nu(f) > 0$. Par exemple, un produit scalaire euclidien ou hermitien est anisotrope.

Remarques. (1) Par définition, un sous-espace vectoriel A de E est non isotrope si et seulement si $A \cap A^\perp = \{0\}$. Donc si E est de dimension finie et f non dégénérée, nous avons

$$E = A \oplus A^\perp$$

si et seulement si A est non isotrope, et si $B = A^\perp$, nous noterons alors $E = A \overset{\perp}{\oplus} B$.

(2) On peut montrer que tous les sous-espaces vectoriels totalement isotropes maximaux (pour l'inclusion) ont la même dimension $\nu(f)$, voir par exemple [Per1, §VIII].

Cône isotrope.

Si f est symétrique ou hermitienne, de forme quadratique ou quadratique hermitienne associée q, nous dirons que q est *isotrope*, *anisotrope*, *définie* ou *indéfinie* si f l'est. Nous appellerons *cône isotrope* de q la partie

$$C(q) = \{x \in E \ : \ q(x) = 0\} \ .$$

Il est réduit à $\{0\}$ si et seulement si q est anisotrope. La formule (8) montre que $C(q)$ est un *cône* dans l'espace vectoriel E, c'est-à-dire une partie A de E telle que pour tous les $x \in A$ et $\lambda \in k$, nous avons $\lambda\, x \in A$.

Exercice E.2. *Soient $p, q \in \mathbb{N}$. Pour tout $n \in \mathbb{N}$, notons*

$$\mathbb{S}_n = \Big\{(x_0, x_1, \dots, x_n) \in \mathbb{R}^{n+1} \ : \ \sum_{i=0}^{n} x_i^2 = 1\Big\}$$

la sphère de dimension n, c'est-à-dire la sphère unité de l'espace vectoriel euclidien usuel $\mathbb{R}^{n+1}$.[26] *Par convention, $\mathbb{S}_{-1} = \emptyset$.*

(1) *Montrer que, dans l'espace vectoriel réel $\mathbb{R}^{p+q}$ muni de sa base canonique, le cône isotrope de la forme quadratique $-x_1^2 - \cdots - x_p^2 + x_{p+1}^2 + \cdots + x_{p+q}^2$, privé de l'origine 0, est homéomorphe à $\mathbb{S}_{p-1} \times \mathbb{S}_{q-1} \times \,]0, +\infty[$.*

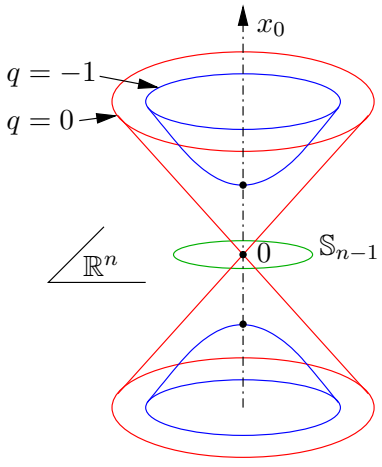
(2) *Montrer que le lieu des points où cette forme quadratique vaut -1 est homéomorphe à $\mathbb{S}_{p-1} \times \mathbb{R}^q$.*

(3) *Montrer que, dans l'espace vectoriel complexe $\mathbb{C}^{p+q}$ muni de sa base canonique, le cône isotrope de la forme quadratique hermitienne $-|z_1|^2 - \cdots - |z_p|^2 + |z_{p+1}|^2 + \cdots + |z_{p+q}|^2$, privé de l'origine 0, est homéomorphe à $\mathbb{S}_{2p-1} \times \mathbb{S}_{2q-1} \times \,]0, +\infty[$.*

Le dessin ci-contre (en rouge) représente le cône isotrope de la forme quadratique réelle

$$q(x_0, x_1, \dots x_n) = -\,x_0^2 + x_1^2 + \cdots + x_n^2 \ .$$

sur $\mathbb{R}^{n+1}$ pour $n \geq 1$. Plus précisément, c'est un dessin précis si $n = 2$. Si $n = 1$, comme la sphère unité de $\mathbb{R}$, qui est $\mathbb{S}_0 = \{-1, +1\}$, est réduite à deux points, ce cône isotrope est la réunion des deux droites vectorielles de pente -1 et $+1$ dans le plan réel $\mathbb{R}^2$. Enfin, le dessin donne une bonne idée pour $n \geq 2$, en remplaçant le cercle horizontal par la sphère de dimension $n-1$ (penser à la sphère terrestre sur laquelle le méridien de Greenwich tourne autour de l'axe passant par les pôles).



26. Attention, la sphère de dimension n vit dans l'espace euclidien de dimension $n+1$!

Pour comprendre le cône isotrope de la forme quadratique réelle

$$-x_0^2 - \cdots - x_p^2 + x_{p+1}^2 + \cdots + x_{p+q}^2$$

sur $\mathbb{R}^{p+q}$, il suffit de prendre la partie du cône isotrope de $-x_0^2 + x_{p+1}^2 + \cdots + x_{p+q}^2$ où $x_0 \geq 0$, et de faire tourner l'axe des x_0 positifs par le groupe orthogonal des p premières coordonnées (fixant les q dernières) : nous verrons dans la proposition 2.1 que le groupe orthogonal $\mathrm{O}(p)$ agit transitivement sur la sphère unité $\mathbb{S}_{p-1}$ de $\mathbb{R}^p$.

1.4 Adjoint

Soient E et F deux espaces vectoriels sur k de dimension finie, munis de formes sesquilinéaires non dégénérées f et g respectivement, qui sont simultanément symétriques ou hermitiennes.

Proposition 1.7. *Pour tout $u \in \mathscr{L}(E,F)$, il existe une unique application $u^* \in \mathscr{L}(F,E)$ telle que*

$$\forall\, x \in E,\ \ \forall\, y \in F, \quad f(x, u^*(y)) = g(u(x), y)\ . \tag{10}$$

L'application $u \mapsto u^*$ est involutive (c'est-à-dire qu'elle vérifie l'égalité $(u^*)^* = u$ pour tout $u \in \mathscr{L}(E,F)$) et semi-linéaire *(c'est-à-dire qu'elle vérifie $(u+\lambda v)^* = u^* + \lambda^\sigma v^*$ pour tous les $\lambda \in k$ et $u, v \in \mathscr{L}(E,F)$). Elle est aussi contravariante (c'est-à-dire qu'elle vérifie* $\mathrm{id}^* = \mathrm{id}$ *et $(u \circ v)^* = v^* \circ u^*$ pour tous les $u \in \mathscr{L}(F,G)$ et $v \in \mathscr{L}(E,F)$). En particulier $(u^n)^* = (u^*)^n$ pour tous les $n \in \mathbb{N}$ et $u \in \mathscr{L}(E)$. L'application linéaire u^* est inversible si et seulement si u l'est, et alors $(u^*)^{-1} = (u^{-1})^*$ et si $E = F$ alors $(u^*)^n = (u^n)^*$ pour tout $n \in \mathbb{Z}$.*

L'application u^* est appelée l'*adjoint* de u (relativement à f et g). Puisque f et g sont simultanément symétriques ou hermitiennes, la formule (10) est équivalente à

$$\forall\, x \in E,\ \ \forall\, y \in F, \quad f(u^*(y), x) = g(y, u(x))\ .$$

Démonstration. Pour l'existence, soient $\widetilde{f} : E^\sigma \to E^*$ et $\widetilde{g} : F^\sigma \to F^*$ les isomorphismes de dualité, de sorte que $\widetilde{f}(a)(b) = f(a,b)$ pour tous les $a, b \in E$. Soit $\breve{u} \in \mathscr{L}(F^*, E^*)$ l'application *duale* (on dit parfois *transposée*) de u, qui est l'application de précomposition par u des formes linéaires sur F, donc définie par $\ell \mapsto \ell \circ u$. Notons que si ℓ est une forme linéaire sur F, alors $\ell \circ u$ est bien une forme linéaire sur E, et que l'application $u \mapsto \breve{u}$ est linéaire.

Rappelons que les groupes additifs sous-jacents à E et à E^σ coïncident, ainsi que ceux de F et F^σ. Montrons que l'application

$$u^* = \widetilde{f}^{-1} \circ \breve{u} \circ \widetilde{g}\ , \tag{11}$$

considérée comme une application de F dans E convient. En effet, pour tous les $x \in E$ et $y \in F$, nous avons

$$\begin{aligned} f\big(x, \widetilde{f}^{-1} \circ \breve{u} \circ \widetilde{g}\,(y)\big)^\sigma &= f\big(\widetilde{f}^{-1} \circ \breve{u} \circ \widetilde{g}\,(y), x\big) = \widetilde{f}\big(\widetilde{f}^{-1} \circ \breve{u} \circ \widetilde{g}\,(y)\big)(x) \\ &= (\breve{u} \circ \widetilde{g}\,(y))(x) = \widetilde{g}\,(y)(u(x)) = g(y, u(x)) = g(u(x), y)^\sigma\ . \end{aligned}$$

Ceci montre la formule (10). De plus, u^* est un morphisme de groupes additifs, comme composé de trois tels morphismes. Comme l'application $\breve{u}$ est linéaire et les applications $\widetilde{f} : E \to E^*$ et $(\widetilde{f})^{-1} : E \to E^*$ sont semi-linéaires (et puisque σ est involutif), l'application u^* est linéaire.

L'unicité de u^* est claire, car un vecteur de E orthogonal à tout vecteur de E est nul puisque f est non dégénérée. Elle implique les propriétés d'involution, de semi-linéarité, et la relation de contravariance $(u \circ v)^* = v^* \circ u^*$. □

Un opérateur linéaire $u \in \mathscr{L}(E)$ est dit

- *auto-adjoint* si $u = u^*$ (on dit aussi *hermitien*, et *symétrique* dans le cas particulier où f est symétrique),
- *positif* si $f(u(x), x) \in \{\lambda\lambda^\sigma \ : \ \lambda \in k\}$ pour tout $x \in E$, [27]
- *normal* si $u \circ u^* = u^* \circ u$,
- *unitaire* si $u \circ u^* = u^* \circ u = \mathrm{id}$.

Remarques. (1) Si A (respectivement B) est la matrice de f (respectivement g) dans une base $\mathscr{B}$ (respectivement $\mathscr{C}$) de E (respectivement F), si M est la matrice de $u \in \mathscr{L}(E, F)$ dans ces bases, alors la matrice M^* de $u^* \in \mathscr{L}(F, E)$ dans ces bases est [28]

$$M^* = (A^\sigma)^{-1}\, {}^tM^\sigma B^\sigma \ . \tag{12}$$

En particulier, si E et F sont des espaces euclidiens (respectivement hermitiens), et si $\mathscr{B}$ et $\mathscr{C}$ sont orthonormées (voir la partie 1.7), [29] alors M^* est la matrice transposée de M (respectivement la matrice transposée de la matrice conjuguée de M) :

$$M^* = {}^tM \ \text{ (respectivement } \ M^* = {}^t\overline{M}\ ) \ .$$

Donc si $E = F$ et $\mathscr{B} = \mathscr{C}$, un opérateur linéaire $u \in \mathscr{L}(E)$ est auto-adjoint si et seulement si sa matrice M dans la base orthonormée $\mathscr{B}$ est symétrique ${}^tM = M$ (respectivement hermitienne ${}^tM = \overline{M}$).

(2) Un opérateur linéaire $u \in \mathscr{L}(E)$ auto-adjoint ou unitaire est normal. Si $u \in \mathscr{L}(E)$ est normal et $v \in \mathscr{L}(E)$ est unitaire, alors $v \circ u \circ v^{-1}$ est normal, car

$$(v \circ u \circ v^{-1})^* = (v \circ u \circ v^*)^* = v \circ u^* \circ v^* = v \circ u^* \circ v^{-1}$$

et les conjugués par un même automorphisme de deux endomorphismes qui commutent commutent encore.

27. lorsque $(k, \sigma) = (\mathbb{R}, \mathrm{id})$ ou $(k, \sigma) = (\mathbb{C}, z \mapsto \overline{z}\,)$, ceci s'écrit $f(u(x), x) \geq 0$ pour tout $x \in E$

28. En effet, si X et Y sont les vecteurs colonnes des coordonnées de n'importe quels $x \in E$ et $y \in F$ dans les bases $\mathscr{B}$ et $\mathscr{C}$ respectivement, alors

$${}^tX(A(M^*)^\sigma)Y^\sigma = {}^tXA(M^*Y)^\sigma = f(x, u^*(y)) = g(u(x), y) = {}^t(MX)BY^\sigma = {}^tX({}^tMB)Y^\sigma \ .$$

Donc $A(M^*)^\sigma = {}^tMB$ et $M^* = (A^{-1})^\sigma({}^tM)^\sigma B^\sigma$.

29. Attention à ne pas oublier cette hypothèse, elle est nécessaire. Par exemple, si $E = F$ est l'espace euclidien usuel $\mathbb{R}^2$, si $\mathscr{B} = \mathscr{C}$ est la base $\{(1, 0), (1, 1)\}$ (qui n'est pas orthonormée) et si p est la projection orthogonale sur la première droite de coordonnées $\mathbb{R}(1, 0)$, alors nous avons

$$\mathrm{Mat}_{\mathscr{B}}(p) = \mathrm{Mat}_{\mathscr{B}}(p^*) = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \ .$$

(3) Si $E = F$, pour tout $u \in \mathscr{L}(E)$, l'application $h : (x, y) \mapsto f(u(x), y)$ est une forme sesquilinéaire sur E, qui est hermitienne si et seulement si u est auto-adjoint, et qui est positive si et seulement si u est positif.

(4) Si $E = F$, si $\sigma \neq \mathrm{id}$ et si la caractéristique de k est différente de 2, alors un opérateur linéaire positif u est auto-adjoint. En effet, par ce qui précède, en posant $h : (x, y) \mapsto f(u(x), y)$, il suffit de montrer que la forme sesquilinéaire positive h est hermitienne. La forme h, étant positive, prend ses valeurs sur la diagonale de $E \times E$ dans l'ensemble $\{\lambda\lambda^\sigma : \lambda \in k\}$, qui est contenu dans le sous-corps fixe k_0 de σ. Donc pour tous les $x, y \in E$, nous avons

$$h(y, x) + h(x, y) = h(x + y, x + y) - h(x, x) - h(y, y) \in k_0 \ .$$

Notons $\iota \in k$ un élément donné par le lemme 1.1. En remplaçant y par $\iota^{-1}y$ dans la formule centrée ci-dessus, nous avons

$$h(y, x) - h(x, y) = \iota\big(h(\iota^{-1}y, x) + h(x, \iota^{-1}y)\big) \in \iota\, k_0 \ .$$

Puisque $z^\sigma = -z$ pour tout $z \in \iota k_0$ et $z^\sigma = z$ pour tout $z \in k_0$, nous avons donc

$$\begin{aligned}
&h(y, x) - h(x, y)^\sigma \\
&= \frac{h(y, x) + h(y, x)^\sigma}{2} + \iota\frac{h(y, x) - h(y, x)^\sigma}{2\iota} - \frac{h(x, y) + h(x, y)^\sigma}{2} + \iota\frac{h(x, y) - h(x, y)^\sigma}{2\iota} \\
&= \frac{(h(y, x) - h(x, y)) + (h(y, x) - h(x, y))^\sigma}{2} + \iota\frac{(h(y, x) + h(x, y)) - (h(y, x) + h(x, y))^\sigma}{2\iota} \\
&= 0 \ .
\end{aligned}$$

Ceci montre que h est hermitienne.

(5) Par exemple, $\mathrm{id} \in \mathscr{L}(E)$ est auto-adjoint et unitaire, et pour tout u dans $\mathscr{L}(E)$, les opérateurs linéaires $u + u^*$ et $\iota(u - u^*)$ (celui-ci n'étant défini que lorsque $\sigma \neq \mathrm{id}$ et la caractéristique de k est différente de 2) sont auto-adjoints, par les propriétés de semi-linéarité et d'involution de l'adjoint. Si de plus f est positive, alors id est positif. Si $u \in \mathscr{L}(E)$ est auto-adjoint et si $v \in \mathscr{L}(E)$ est unitaire, alors $v \circ u \circ v^{-1}$ est auto-adjoint.

De même, pour toute application linéaire $u \in \mathscr{L}(E, F)$, les opérateurs linéaires $u^* \circ u \in \mathscr{L}(E)$ et $u \circ u^* = (u^*)^* \circ u^* \in \mathscr{L}(F)$ sont auto-adjoints, car

$$\forall\, x, y \in E, \quad f(u^* \circ u(x), y) = g(u(x), u(y)) = f(x, u^* \circ u(y)) \ .$$

L'opérateur linéaire $u^* \circ u \in \mathscr{L}(E)$ est positif si la forme sesquilinéaire g est positive (et donc l'opérateur linéaire $u \circ u^* = (u^*)^* \circ u^* \in \mathscr{L}(F)$ est positif si la forme sesquilinéaire f est positive), car

$$\forall\, x \in E, \quad f(u^* \circ u(x), x) = g(u(x), u(x)) \in \{\lambda\, \lambda^\sigma : \lambda \in k\} \ .$$

Le résultat suivant regroupe les relations élémentaires entre les propriétés d'une application linéaire de E dans F et celles de son adjoint. [30]

30. Nous renvoyons à un cours d'algèbre linéaire pour la théorie générale de la réduction des endomorphismes (linéaires) u d'un espace vectoriel de dimension finie E' sur un corps commutatif k'. L'ensemble $\mathrm{Vp}(u)$ des *valeurs propres* de u est l'ensemble des $\lambda \in k'$ tels que l'endomorphisme $u - \lambda\,\mathrm{id}$ de E' soit non inversible dans l'algèbre $\mathscr{L}(E')$ des endomorphismes linéaires de E'. Pour tout $\lambda \in k'$, posons $E'_\lambda = \{x \in E' : u(x) = \lambda x\}$, que nous noterons $E'_\lambda(u)$ lorsqu'il convient de préciser u. Le scalaire λ appartient à $\mathrm{Vp}(u)$ si et seulement E'_λ n'est pas le sous-espace nul, et E'_λ est alors appelé l'*espace propre* de u associé à la valeur propre $\lambda \in \mathrm{Vp}(u)$.

Proposition 1.8. *Soit $u \in \mathscr{L}(E,F)$.*

(i) L'orthogonal de l'image de u est le noyau de son adjoint et le noyau de u est l'orthogonal de l'image de son adjoint :

$$\big(u(E)\big)^{\perp} = \ker(u^*) \quad \text{et} \quad \ker(u) = \big(u^*(F)\big)^{\perp} .$$

En particulier, u^* est injectif si et seulement si u est surjectif, et u^* est surjectif si et seulement si u est injectif.

(ii) Si f et g sont anisotropes, alors

$$\ker(u) = \ker(u^* \circ u) \quad \text{et} \quad \ker(u^*) = \ker(u \circ u^*) .$$

(iii) Si f et g sont anisotropes, alors les endomorphismes u, u^*, $u \circ u^*$ et $u^* \circ u$ ont le même rang.

Dans la suite de cet énoncé, nous supposons que $E = F$.

(iv) Soit $\lambda \in k$. Alors λ est valeur propre de u^* si et seulement si λ^σ est valeur propre de u :

$$\mathrm{Vp}(u^*) = \big(\mathrm{Vp}(u)\big)^{\sigma} .$$

(v) Si V est un sous-espace vectoriel de E invariant *(ou* stable*) par u (c'est-à-dire tel que $u(V) \subset V$), alors $V^\perp$ est invariant par u^*. En particulier, si u est auto-adjoint, et si V est un sous-espace vectoriel de E invariant par u, alors $V^\perp$ est aussi invariant par u.*

(vi) Si f est anisotrope et si u est auto-adjoint (respectivement anti-auto-adjoint (c'est-à-dire si $u^* = -u$) ; unitaire), alors l'ensemble $\mathrm{Vp}(u)$ des valeurs propres de u est contenu dans le sous-corps fixe $k_0 = \{\lambda \in k : \lambda^\sigma = \lambda\}$ de σ (respectivement dans $\{\lambda \in k : \lambda^\sigma = -\lambda\}$; dans $N^{-1}(1) = \{\lambda \in k : \lambda\lambda^\sigma = 1\}$). Si f est anisotrope et positive, et si u est positif, alors $\mathrm{Vp}(u)$ est contenu dans $\{\lambda\lambda^\sigma : \lambda \in k\}$.[31]

(vii) Si f est anisotrope, et si u est normal,[32] *alors $x \in E$ est un vecteur propre de u associé à la valeur propre λ si et seulement si x est un vecteur propre de u^* associé à la valeur propre λ^σ. En particulier $\ker(u) = \ker(u^*)$ et plus généralement, nous avons, pour tout $\lambda \in k$, l'égalité suivante entre les espaces propres $E_\lambda(u)$ de u pour la valeur propre λ et $E_{\lambda^\sigma}(u^*)$ de u^* pour la valeur propre λ^σ :*

$$E_\lambda(u) = E_{\lambda^\sigma}(u^*) .$$

Démonstration. (i) Pour tout vecteur $y \in F$, nous avons $y \in u(E)^\perp$ si et seulement si $g(y, u(x)) = 0$ pour tout x dans E, si et seulement si $f(u^*(y), x) = 0$ pour tout x dans E, donc (puisque f est non dégénérée) si et seulement si $u^*(y) = 0$, c'est-à-dire $y \in \ker(u^*)$. La seconde égalité de (i) découle de la première en remplaçant u par u^* et en utilisant la propriété d'involution de $u \mapsto u^*$.

(ii) L'inclusion $\ker(u) \subset \ker(u^* \circ u)$ est immédiate, car si $u(x) = 0$, alors $u^* \circ u(x) = 0$. Réciproquement, si $u^* \circ u(x) = 0$, alors $0 = f(u^* \circ u(x), x) = g(u(x), u(x))$, donc $u(x) = 0$ car g est anisotrope. La seconde égalité de (ii) découle de la première en remplaçant u par u^* et en utilisant la propriété d'involution de $u \mapsto u^*$.

31. En particulier, si E est un espace euclidien ou hermitien, et si u est auto-adjoint (respectivement anti-auto-adjoint, unitaire, positif), alors $\mathrm{Vp}(u)$ est contenu dans $\mathbb{R}$ (respectivement dans $i\mathbb{R}$ (et donc est nul si E est euclidien), dans le cercle $\mathbb{S}_1 = \{z \in \mathbb{C} : |z| = 1\}$ (et donc dans $\{\pm 1\}$ si E est euclidien), dans $[0, +\infty[$).

32. **Attention** à ne pas oublier l'hypothèse que u est normal dans cette assertion (vii) !

(iii) Comme f est non dégénérée, par l'assertion (i) et par le théorème du rang, nous avons

$$\dim(\text{im } u^*) = \dim E - \dim\big((\text{im } u^*)^\perp\big) = \dim E - \dim(\ker u) = \dim(\text{im } u) \ ,$$

ce qui est l'égalité des rangs de u et de u^*.

Par deux utilisations du théorème du rang et par l'assertion (ii), nous avons

$$\dim(\text{im } (u^* \circ u)) = \dim E - \dim(\ker(u^* \circ u)) = \dim E - \dim(\ker u) = \dim(\text{im } u) \ .$$

Ceci montre l'égalité des rangs de u et de $u^* \circ u$. La dernière égalité entre les rangs de u^* et de $u \circ u^*$ en découle en remplaçant u par u^* et en utilisant la propriété d'involution de $u \mapsto u^*$.

Supposons maintenant $E = F$.

(iv) L'opérateur linéaire $u - \lambda \,\mathrm{id} \in \mathscr{L}(E)$ est inversible si et seulement si son adjoint, qui est $u^* - \lambda^\sigma \,\mathrm{id}$, est inversible. Donc $\mathrm{Vp}(u^*) = \big(\mathrm{Vp}(u)\big)^\sigma$.

(v) Soient u et V comme dans l'énoncé, et soit $x \in V^\perp$. Pour tout $y \in V$, puisque $u(y) \in V$, nous avons $f(u^*(x), y) = f(x, u(y)) = 0$. D'où $u^*(x) \in V^\perp$.

(vi) Soit $\lambda \in \mathrm{Vp}(u)$ une valeur propre de u associée à un vecteur propre non nul donc non isotrope $x \in E_\lambda - \{0\}$. Notons que si u est unitaire, alors u est inversible, donc de valeurs propres non nulles, et $u^* = u^{-1}$ donc $u^*(x) = \frac{1}{\lambda}\, x$ car

$$x = u^{-1}(u(x)) = u^{-1}(\lambda\, x) = \lambda\, u^*(x) \ .$$

Si u est autoadjoint (respectivement anti-auto-adjoint, unitaire), posons $\lambda' = \lambda$ (respectivement $\lambda' = -\lambda$, $\lambda' = \frac{1}{\lambda}$). Alors

$$\lambda f(x,x) = f(\lambda x, x) = f(u(x), x) = f(x, u^*(x)) = f(x, \lambda' x) = (\lambda')^\sigma f(x,x) \ .$$

Puisque $f(x,x) \neq 0$, nous avons donc $\lambda = (\lambda')^\sigma$. Si u est auto-adjoint, ceci signifie que λ appartient au sous-corps fixe k_0 de σ. Si u est unitaire, ceci signifie que $N(\lambda) = \lambda\lambda^\sigma = 1$.

Si u est positif, alors

$$\lambda\, f(x,x) = f(\lambda\, x, x) = f(u(x), x) \in \{\mu\mu^\sigma : \mu \in k\} \ .$$

Si f est positive, puisque x n'est pas isotrope, nous avons $f(x,x) \in \{\mu\mu^\sigma : \mu \in k\} - \{0\} = N(k^\times)$, donc $\lambda \in \{\mu\mu^\sigma : \mu \in k\}$.

(vii) L'égalité $\ker(u) = \ker(u^*)$ découle de l'assertion (ii) lorsque u est normal, et l'assertion (vii) en découle en appliquant cette égalité à l'opérateur linéaire normal $u - \lambda \,\mathrm{id}$, dont l'adjoint est $u^* - \lambda^\sigma \,\mathrm{id}$. □

Exercice E.3. *Pour tout polynôme $P = \sum_{i=0}^m a_i X^i \in k[X]$, notons $P^\sigma \in k[X]$ le polynôme $P^\sigma = \sum_{i=0}^m a_i^\sigma X^i$. Soit $u \in \mathscr{L}(E)$, de polynôme minimal unitaire π_u et de polynôme caractéristique χ_u. Montrer que le polynôme minimal unitaire de u^* est $\pi_{u^*} = (\pi_u)^\sigma$ et que le polynôme caractéristique de u^* est $\chi_{u^*} = (\chi_u)^\sigma$.*

1.5 Groupes orthogonaux, symplectiques et unitaires

Soient E un espace vectoriel sur k, muni d'une forme sesquilinéaire non dégénérée f, qui est symétrique, alternée ou hermitienne. Supposons la caractéristique de k différente de 2.

Une *isométrie* (vectorielle) de E (relativement à f) est un automorphisme linéaire u de E qui préserve la forme sesquilinéaire f, c'est-à-dire tel que

$$\forall\, x, y \in E, \quad f(u(x), u(y)) = f(x, y)\ .$$

Lorsque f est symétrique ou hermitienne, en remplaçant y par $u^{-1}(y)$ dans la formule ci-dessus, ceci équivaut à dire que l'endomorphisme u est unitaire (d'inverse égale son adjoint). Par les propriétés de linéarité de u et de sesquilinéarité de f, il suffit de vérifier cette condition pour les x et y dans une partie génératrice fixée de E. L'ensemble des isométries vectorielles de E est un sous-groupe du groupe linéaire $\mathrm{GL}(E)$, appelé et noté :

- le *groupe orthogonal* $\mathrm{O}(f)$ (ou parfois $\mathrm{O}(E)$) de f, si f est symétrique,
- le *groupe symplectique* $\mathrm{Sp}(f)$ (ou parfois $\mathrm{Sp}(E)$) de f, si f est alternée,
- le *groupe unitaire* $\mathrm{U}(f)$ (ou parfois $\mathrm{U}(E)$) de f, si f est hermitienne.

Remarques. (1) Si une isométrie (vectorielle) u de E préserve un sous-espace vectoriel F de E, alors u préserve aussi son orthogonal. En effet, si $y \in F^\perp$, alors pour tout $x \in F$, nous avons $f(u(y), x) = f(y, u^{-1}(x))$ qui est nul car $u^{-1}(x) \in F$, donc $u(y) \in F^\perp$.

(2) Si f est symétrique (respectivement hermitienne), il découle de la formule de polarisation (4) (respectivement de la formule de polarisation (6)) qu'un élément $u \in \mathrm{GL}(E)$ est une isométrie de E si et seulement si u préserve la forme quadratique (respectivement forme quadratique hermitienne) q associée à f, c'est-à-dire

$$\forall\, x \in E, \quad q(u(x)) = q(x)\ .$$

Nous noterons alors $\mathrm{O}(q) = \mathrm{O}(f)$ (respectivement $\mathrm{U}(q) = \mathrm{U}(f)$), que nous appellerons le *groupe orthogonal* (respectivement *unitaire*) de q.

(3) Si E est un espace euclidien, alors[33] une isométrie de la distance de E est la composée d'une isométrie vectorielle de E et d'une translation (ainsi que la composée d'une

33. Voir par exemple [Aud, Exercice II.5]. Un argument direct est le suivant. Soit $u : E \to E$ une isométrie de la distance euclidienne d de E. Quitte à composer (à droite ou à gauche) par une translation, nous pouvons supposer que u fixe 0. Pour tout $x \in E$ et $\lambda \in [0, 1]$, puisque $\lambda\, x$ appartient au segment $[0, x]$, le point $\lambda\, x$ est l'unique point z de E tel que

$$d(0, z) + d(z, x) = d(0, x) \quad \text{et} \quad d(0, z) = |\lambda|\, d(0, x)\ .$$

Alors $d(0, u(z)) + d(u(z), u(x)) = d(0, u(x))$ et $d(0, u(z)) = |\lambda|\, d(0, u(x))$. Donc

$$u(\lambda x) = \lambda\, u(x)\ . \tag{13}$$

On procède de manière similaire pour montrer cette égalité lorsque $\lambda \in [1, +\infty[$ et $]-\infty, 0]$. Puisque le milieu d'un segment $[y, z]$ est l'unique point m de E tel que

$$d(y, m) + d(m, z) = d(y, z) \quad \text{et} \quad d(y, m) = d(m, z)\ ,$$

nous avons $u(\frac{y+z}{2}) = \frac{u(y)+u(z)}{2}$. Donc $u(y + z) = u(y) + u(z)$ par la formule (13) avec $\lambda = 2$ et $x = \frac{y+z}{2}$. Par conséquent, u est une bijection linéaire qui préserve la distance à l'origine, donc qui préserve la norme. Ceci montre que u est une isométrie vectorielle de E.

translation et d'une isométrie vectorielle de E). De plus, une similitude de la distance de E est la composée d'une isométrie vectorielle, d'une translation et d'une homothétie.

(4) **Interprétation matricielle.** Supposons que E soit de dimension finie n, fixons une base $\mathscr{B}$ dans E, notons A la matrice de f dans la base $\mathscr{B}$, et pour tous les $x \in E$ et $u \in \mathrm{GL}(E)$, notons X la matrice colonne des coordonnées de x et U la matrice de u dans la base $\mathscr{B}$. Alors[34] u est une isométrie vectorielle de E si et seulement si

$$ {}^tU\, A\, U^\sigma = A\,. \tag{14}$$

Pour tout $n \in \mathbb{N}$, notons I_n la matrice identité de taille $n \times n$. Si E est l'espace euclidien $\mathbb{R}^n$ (respectivement hermitien $\mathbb{C}^n$) usuel, nous identifions $\mathscr{L}(\mathbb{R}^n)$ avec $\mathscr{M}_n(\mathbb{R})$ (respectivement $\mathscr{L}(\mathbb{C}^n)$ avec $\mathscr{M}_n(\mathbb{C})$) par l'application qui à un endomorphisme associe sa matrice dans la base canonique, ce qui induit une identification entre $\mathrm{O}(E)$ et

$$\mathrm{O}(n) = \{x \in \mathscr{M}_n(\mathbb{R}) :\ {}^t x\, x = I_n\} = \{x \in \mathrm{GL}_n(\mathbb{R}) : x^{-1} = {}^t x\}$$

(respectivement $\mathrm{U}(E)$ et

$$\mathrm{U}(n) = \{x \in \mathscr{M}_n(\mathbb{C}) :\ {}^t \overline{x}\, x = I_n\} = \{x \in \mathrm{GL}_n(\mathbb{C}) : x^{-1} = {}^t \overline{x}\}\ )\,.$$

(5) **Groupes spéciaux unitaire et orthogonal.** Si E est de dimension finie, la formule (14) (en rappelant que le discriminant de f est non nul car f est non dégénérée) montre que le déterminant d'une isométrie vectorielle u vérifie la propriété

$$(\det u)(\det u)^\sigma = 1\,. \tag{15}$$

En particulier $\det u = \pm 1$ si $\sigma = \mathrm{id}$, et $|\det u| = 1$ si $(k, \sigma) = (\mathbb{C}, z \mapsto \overline{z})$.

Si f est alternée, il est possible de montrer (voir [Art], et l'exercice E.5 lorsque $k = \mathbb{R}$) que le déterminant d'un élément du groupe symplectique $\mathrm{Sp}(f)$ est toujours 1.

Si f est respectivement symétrique ou hermitienne, le sous-groupe

$$\mathrm{SO}(f) = \mathrm{SO}(E) = \{u \in \mathrm{O}(f)\ :\ \det u = 1\}$$

$$\text{ou}\quad \mathrm{SU}(f) = \mathrm{SU}(E) = \{u \in \mathrm{U}(f)\ :\ \det u = 1\}$$

est un sous-groupe distingué de $\mathrm{O}(f)$ ou $\mathrm{U}(f)$, appelé le *groupe spécial orthogonal* (ou aussi le *groupe des rotations* si $k = \mathbb{R}$) ou le *groupe spécial unitaire* de f.

Si E est l'espace euclidien $\mathbb{R}^n$ (respectivement hermitien $\mathbb{C}^n$) usuel, l'identification de la fin de la remarque (4) ci-dessus induit une identification entre $\mathrm{SO}(E)$ et

$$\begin{aligned}\mathrm{SO}(n) &= \{x \in \mathscr{M}_n(\mathbb{R}) :\ {}^t x\, x = I_n \ \text{ et } \ \det x = 1\}\\ &= \{x \in \mathrm{GL}_n(\mathbb{R}) : x^{-1} = {}^t x \ \text{ et } \ \det x = 1\}\end{aligned}$$

(respectivement $\mathrm{SU}(E)$ et

$$\begin{aligned}\mathrm{SU}(n) &= \{x \in \mathscr{M}_n(\mathbb{C}) :\ {}^t \overline{x}\, x = I_n \ \text{ et } \ \det x = 1\}\\ &= \{x \in \mathrm{GL}_n(\mathbb{C}) : x^{-1} = {}^t \overline{x} \ \text{ et } \ \det x = 1\}\ )\,.\end{aligned}$$

34. En effet, $f(u(x), u(y)) = {}^t(UX)\, A\, (UY)^\sigma = {}^t X({}^t U\, A\, U^\sigma) Y^\sigma$.

Si f' est une forme sesquilinéaire sur un espace vectoriel E' sur k équivalente[35] à f, et si $v : E \to E'$ est une conjugaison entre f et f', alors

- u est une isométrie vectorielle de f si et seulement si $v \circ u \circ v^{-1}$ est une isométrie vectorielle de f', et f' est anisotrope si et seulement si f l'est ;
- f' est non dégénérée si et seulement si f l'est et alors f et f' ont même rang, même indice et même discriminant (modulo les normes, c'est-à-dire modulo un carré si $(k, \sigma) = (\mathbb{R}, \mathrm{id})$ ou le carré d'une valeur absolue si $(k, \sigma) = (\mathbb{C}, z \mapsto \overline{z})$) :

$$\mathrm{rang}(f') = \mathrm{rang}(f), \quad \nu(f') = \nu(f), \quad \mathrm{Disc}(f') = \mathrm{Disc}(f) \in k^\times / N(k^\times) ;$$

- f' est symétrique si et seulement si f l'est, et alors

$$\mathrm{O}(f') = v \circ \mathrm{O}(f) \circ v^{-1} ; \tag{16}$$

- f' est alternée si et seulement si f l'est, et alors

$$\mathrm{Sp}(f') = v \circ \mathrm{Sp}(f) \circ v^{-1} ;$$

- f' est hermitienne si et seulement si f l'est, et alors

$$\mathrm{U}(f') = v \circ \mathrm{U}(f) \circ v^{-1} .$$

Nous étudierons plus en détail les groupes orthogonaux des espaces euclidiens et les groupes unitaires des espaces hermitiens dans les parties 2 et 3.

Exercice E.4. *Quel est le groupe* $\mathrm{SO}(f)$ *pour* $f = x^2 + y^2$ *sur le corps fini* $\mathbb{Z}/p\mathbb{Z}$ *pour* $p = 3, 5, 7$ *?*

Exercice E.5. *Soit* k *un corps commutatif. Un* plan symplectique *sur* k *est un plan vectoriel sur* k*, muni d'une forme bilinéaire alternée non dégénérée. Soit* E *un espace vectoriel de dimension finie sur* k*, muni d'une forme bilinéaire alternée* f*.*

(1) *Montrer que* E *est somme orthogonale de droites isotropes et de plans symplectiques. En déduire que si* f *est non dégénérée, alors la dimension de* E *est paire. Montrer que deux formes bilinéaires alternées non dégénérées de même rang sont équivalentes.*

(2) *Montrer que tous les sous-espaces vectoriels totalement isotropes maximaux de* E *ont la même dimension, et la déterminer.*

(3) *Si* E *est un plan symplectique, montrer que* $\mathrm{Sp}(f)$ *est isomorphe à* $\mathrm{SL}_2(k)$*.*

(4) *Montrer que si* f *est non dégénérée, alors il existe* $n \in \mathbb{N}$ *et une base* $\mathscr{B}$ *de* E *dans laquelle la matrice de* f *est* $J_n = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$*. Montrer que l'application de* $\mathrm{Sp}(f)$ *dans*

$$\mathrm{Sp}_n(k) = \{X \in \mathrm{GL}_{2n}(k) : {}^tX\, J_n\, X = J_n\}$$

qui à un endomorphisme associe sa matrice dans la base $\mathscr{B}$ *est un isomorphisme de groupes.*

(5) *Montrer que* $\mathrm{Sp}_n(\mathbb{R}) \cap \mathrm{O}(2n) = \left\{X = \begin{pmatrix} A & -B \\ B & A \end{pmatrix} : A, B \in \mathscr{M}_n(k), X \in \mathrm{O}(2n)\right\}$*. Si* f *est non dégénérée et* $k = \mathbb{R}$*, pour tout* $u \in \mathrm{Sp}(f)$*, montrer, en utilisant le lemme 1.24 et la décomposition polaire (formule (28) de la partie 3.5), que* $\det(u) = +1$*.*

35. Voir la partie 1.1 pour la définition.

1.6 Symétries orthogonales

Soit E un espace vectoriel sur k de dimension finie, muni d'une forme bilinéaire non dégénérée symétrique f, de forme quadratique associée q. Supposons la caractéristique de k différente de 2. Nous introduisons maintenant des éléments particuliers de $\mathrm{O}(q)$, qui seront utiles pour étudier une partie génératrice de $\mathrm{O}(q)$ dans la partie 2.2.

Rappelons que pour toute *symétrie* de E, c'est-à-dire tout automorphisme linéaire involutif u de E, il existe deux uniques sous-espaces vectoriels E^- et E^+ de E tels que

- $E = E^- \oplus E^+$ est somme directe de E^- et E^+,
- $E^\pm$ est l'espace propre de u pour la valeur propre ± 1 : nous avons $u\mid_{E^-} = -\mathrm{id}_{E^-}$ et $u\mid_{E^+} = +\mathrm{id}_{E^+}$.

Nous dirons que u est la symétrie *par rapport à* E^+ *parallèlement à* E^-. Si $\dim E^- = 1$, la symétrie u est appelée une *réflexion*, et si $\dim E^- = 2$, la symétrie u est appelée un *renversement* (ou aussi un *retournement*). Une symétrie ou une réflexion ou un renversement qui appartient à $\mathrm{O}(q)$ est dit *orthogonal*.

Notons que les espaces propres E^- et E^+ d'une symétrie orthogonale u, étant en somme directe et orthogonaux (voir le premier point de la proposition suivante), sont non isotropes, et sont les orthogonaux l'un de l'autre :

$$(E^-)^\perp = E^+ \quad \text{et} \quad (E^+)^\perp = E^- \ .$$

En particulier l'un détermine l'autre, et nous dirons alors que u est la symétrie orthogonale *par rapport à* E^+.

Proposition 1.9. *Une symétrie $u \in \mathrm{GL}(E)$ est orthogonale si et seulement si ses espaces propres E^- et E^+ sont orthogonaux.*

Pour tout sous-espace vectoriel F non isotrope de E, il existe une et une seule symétrie orthogonale s_F dont le sous-espace vectoriel fixe est F. Pour tout $v \in \mathrm{O}(q)$, le conjugué par v de la symétrie orthogonale par rapport à F est la symétrie orthogonale par rapport à l'image de F par v :

$$v \circ s_F \circ v^{-1} = s_{v(F)} \ . \tag{17}$$

Démonstration. Soit $u \in \mathrm{GL}(E)$ une symétrie.

Supposons que $u \in \mathrm{O}(q)$. Pour tous les $x \in E^-$ et $y \in E^+$, nous avons

$$f(x,y) = f(u(x),u(y)) = f(-x,y) = -f(x,y) \ .$$

Donc $f(x,y) = 0$ car la caractéristique est différente de 2. Par conséquent, nous avons $x \perp y$, et E^- et E^+ sont orthogonaux.

Réciproquement, si E^- et E^+ sont orthogonaux, alors pour tous les $x, y \in E$, en écrivant $x = x_- + x_+$ et $y = y_- + y_+$ leur décomposition dans la somme directe orthogonale $E = E^- \overset{\perp}{\oplus} E^+$, nous avons

$$\begin{aligned} f(u(x),u(y)) = f(-x_- + x_+, -y_- + y_+) &= f(x_-,y_-) + f(x_+,y_+) \\ &= f(x_- + x_+, y_- + y_+) = f(x,y) \ . \end{aligned}$$

Donc u est orthogonale.

Si F est un sous-espace vectoriel non isotrope de E, alors $E = F \oplus F^\perp$ par la remarque (1) de la partie 1.3, et l'unique application linéaire $s_F \in \mathscr{L}(E)$ telle que $s_F\mid_F = \mathrm{id}_F$ et $s_F\mid_{F^\perp} = -\mathrm{id}_{F^\perp}$ convient.

Le conjugué $v \circ s_F \circ v^{-1}$ appartient à $\mathrm{O}(q)$, est involutif, et a pour espaces propres associés aux valeurs propres 1 et -1 respectivement $v(F)$ et $v(F^\perp) = v(F)^\perp$. Donc $v \circ s_F \circ v^{-1} = s_{v(F)}$ par unicité. □

Remarque. L'application $u \mapsto E^-$ (respectivement $u \mapsto E^+$) induit donc une bijection entre l'ensemble des réflexions orthogonales et l'ensemble des droites vectorielles non isotropes (respectivement entre l'ensemble des réflexions orthogonales et l'ensemble des hyperplans vectoriels non isotropes).

Exercice E.6. *Soient (E, f, q) comme dans le début de la partie 1.6. Pour tout vecteur non isotrope*[36] *x, notons s_x la réflexion orthogonale par rapport à l'hyperplan orthogonal à x. Montrer que*

$$s_x : y \mapsto y - 2\,\frac{f(x,y)}{f(x,x)}\,x\,. \tag{18}$$

Pour tout $u \in \mathrm{O}(q)$, montrer que $u \circ s_x \circ u^{-1} = s_{u(x)}$.

Dans la suite, nous supposons que E est un espace vectoriel (réel) euclidien. Un système de racines (réduit) *de E est une partie R de E telle que*

- *R est une partie finie de E, ne contenant pas 0 et engendrant le $\mathbb{R}$-espace vectoriel E,*
- *pour tout $\alpha \in R$, nous avons $s_\alpha(R) = R$,*
- *pour tous les $\alpha, \beta \in R$, les* constantes de structure *$n(\alpha,\beta) = 2\,\frac{\langle\alpha,\beta\rangle}{\langle\alpha,\alpha\rangle}$ appartiennent à $\mathbb{Z}$,*
- *pour tout $\alpha \in R$, les seuls éléments de R colinéaires à α sont α et $-\alpha$.*

Un système de racines R' d'un espace vectoriel (réel) euclidien E' de dimension finie, de constantes de structure $(n'(\alpha',\beta'))_{\alpha',\beta' \in R'}$, est dit isomorphe *au système de racine R s'il existe un isomorphisme linéaire $\phi : E \to E'$ tel que $\phi(R) = R'$ et $n'(\phi(\alpha), \phi(\beta)) = n(\alpha,\beta)$ pour tous les $\alpha, \beta \in R$.*

(1) *Pour tous les $\alpha, \beta \in R$, notons $\theta_{\alpha,\beta} \in [0,\pi]$ l'angle entre les vecteurs α et β.*
   - *Vérifier que si $\alpha \in R$, alors $-\alpha \in R$.*
   - *Montrer que*
   $$n(\alpha,\beta)\,n(\beta,\alpha) = 4\,\cos^2(\theta_{\alpha,\beta})\,.$$
   - *En déduire les valeurs possibles de $\theta_{\alpha,\beta}$.*
   - *Montrer que le couple $(n(\alpha,\beta), n(\beta,\alpha))$ ne peut pas prendre les valeurs $(1,4)$, $(4,1)$, $(-1,-4)$ et $(-4,-1)$.*
   - *Si $\theta_{\alpha,\beta} \neq \frac{\pi}{2}$, montrer que $\dfrac{\|\alpha\|^2}{\|\beta\|^2} = \dfrac{n(\beta,\alpha)}{n(\alpha,\beta)}$.*
   - *En supposant $\|\alpha\| \leq \|\beta\|$, donner un tableau des différentes valeurs possibles du quadruplet*
   $$\Big(n(\alpha,\beta), n(\beta,\alpha), \theta_{\alpha,\beta}, \frac{\|\alpha\|}{\|\beta\|}\Big)\,.$$

(2) *Supposons E de dimension deux, et munissons-le d'une base orthonormée (voir la partie 1.7). Soient α une racine de R de norme minimale, et β une racine de R non proportionnelle et non orthogonale à α. Montrer qu'à isomorphisme près, nous pouvons supposer que les coordonnées de α sont $(1,0)$ et que la deuxième coordonnée de β est strictement positive. Montrer que $n(\alpha,\beta) \neq 0$ et que $n(\alpha, s_\alpha(\beta)) = -n(\alpha,\beta)$.*

36. Rappelons que si f est anisotrope (on dit aussi définie), ceci équivaut à x non nul. Notons qu'un vecteur non nul est non isotrope si et seulement si la droite vectorielle qu'il engendre est non isotrope.

(3) *Montrer qu'il existe exactement quatre classes d'isomorphisme de systèmes de racines (réduits) dans le plan euclidien, et en donner un dessin.*

1.7 Bases orthogonales

Soit E un espace vectoriel sur k de dimension finie $n \in \mathbb{N} - \{0\}$, muni d'une forme sesquilinéaire non dégénérée f, qui est symétrique ou hermitienne. Supposons la caractéristique de k différente de 2.

Une suite $\mathscr{B} = (e_1, \ldots, e_n)$ de E est une *base orthogonale* de E pour f si

- $\mathscr{B}$ est une base de l'espace vectoriel E,
- pour tous les $i, j = 1, \ldots, n$ distincts, les vecteurs e_i et e_j sont orthogonaux, c'est-à-dire $f(e_i, e_j) = 0$, ou, de manière équivalente, la matrice de f dans la base $\mathscr{B}$ est diagonale.

Nous dirons de plus que $\mathscr{B}$ est *orthonormée* si $f(e_i, e_i) = 1$ pour tout $i = 1, \ldots, n$, ou, de manière équivalente, si la matrice de f dans la base $\mathscr{B}$ est la matrice identité. Par convention, la suite vide est l'unique base orthonormée de tout espace vectoriel nul.

Par exemple, le produit scalaire de l'espace euclidien usuel $\mathbb{R}^n$ est l'unique produit scalaire sur l'espace vectoriel reél $\mathbb{R}^n$ rendant sa base canonique orthonormée, et le produit scalaire hermitien de l'espace hermitien usuel $\mathbb{C}^n$ est l'unique produit scalaire hermitien sur l'espace vectoriel complexe $\mathbb{C}^n$ rendant sa base canonique orthonormée.

Proposition 1.10. *Il existe une base orthogonale de E pour f.*

Il n'existe par contre pas toujours de base orthonormée.

Notons que la condition d'être non dégénérée n'est pas nécessaire pour cette proposition, car il est possible de considérer un supplémentaire F du noyau de f, d'appliquer le résultat à la restriction de f à $F \times F$, et de compléter par une base du noyau de f.

Démonstration. Raisonnons par récurrence sur la dimension n de E. Le résultat est immédiat si $n = 1$, toute base étant orthogonale. Par la remarque (3) de la partie 1.2, la forme sesquilinéaire f n'est pas alternée (sinon, elle serait à la fois symétrique et antisymétrique en caractéristique différente de 2, donc nulle), donc il existe un vecteur e_1 non isotrope dans E. Alors son orthogonal $H = e_1^\perp$ est un hyperplan de E, de dimension $n-1$, et la restriction de f à H est encore non dégénérée, et symétrique ou hermitienne. Donc par l'hypothèse de récurrence, H admet une base orthogonale $(e_2, \ldots, e_n)$. Alors $(e_1, \ldots, e_n)$ convient. □

Lorsque f est un produit scalaire euclidien ou hermitien, voici une procédure effective pour construire des bases orthonormées.

Proposition 1.11. (Procédure d'orthonormalisation de Gram-Schmidt.) *Soit E un espace vectoriel (réel) euclidien ou un espace vectoriel (complexe) hermitien. Pour toute base $\mathscr{B} = (e_1, \ldots, e_n)$ de E, il existe une base orthonormée $\mathscr{B}' = (e'_1, \ldots, e'_n)$ de E telle que la matrice de passage P_n de $\mathscr{B}$ à $\mathscr{B}'$ soit triangulaire supérieure, à coefficients diagonaux strictement positifs.*

Remarques sur l'énoncé. (1) Nous avons $\mathrm{Vect}(e_1, \ldots, e_k) = \mathrm{Vect}(e'_1, \ldots, e'_k)$ pour tout $k = 1, \ldots, n$.

(2) Il est possible de passer de $\mathscr{B}$ à $\mathscr{B}'$ de manière continue dans l'espace des bases de E. Plus précisément, il existe une application continue f de $[0, 1]$ dans E^n telle que

$f(0) = (e_1, \dots, e_n)$, $f(1) = (e'_1, \dots, e'_n)$, et $f(t)$ est une base de E pour tout $t \in [0,1]$. En effet, si

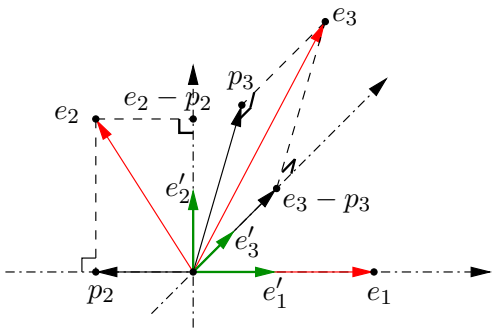
$$P_n = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ 0 & \cdots & 0 & a_{n,n} \end{pmatrix}$$

alors $f(t) = (P(t)e_1, \dots, P(t)e_n)$, où

$$P(t) = \begin{pmatrix} a_{1,1}^t & t\,a_{1,2} & \cdots & t\,a_{1,n} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & t\,a_{n-1,n} \\ 0 & \cdots & 0 & a_{n,n}^t \end{pmatrix},$$

convient (en rappelant que les $a_{i,i}$ sont strictement positifs).

Démonstration. Procédons par récurrence sur la dimension n de E. Si $n = 1$, posons $e'_1 = \frac{1}{\|e_1\|} e_1$, de sorte que $\mathscr{B}' = (e'_1)$ convient. Soit $n \geq 2$, et supposons le résultat vrai pour $n-1$. Soient E et $(e_1, \dots, e_n)$ comme dans l'énoncé. En appliquant l'hypothèse de récurrence à $E' = \mathrm{Vect}(e_1, \dots, e_{n-1})$, il existe une base orthonormée $(e'_1, \dots, e'_{n-1})$ de E' telle que la matrice de passage P_{n-1} de $(e_1, \dots, e_{n-1})$ à $(e'_1, \dots, e'_{n-1})$ soit triangulaire supérieure, à coefficients diagonaux strictement positifs.



Posons $p_n = \sum_{i=1}^{n-1} \langle e'_i, e_n \rangle\, e'_i$ la projection orthogonale de e_n sur l'hyperplan vectoriel $E' = \mathrm{Vect}(e'_1, \dots, e'_{n-1})$ de E, et

$$e'_n = \frac{1}{\|e_n - p_n\|} (e_n - p_n).$$

Il est immédiat de vérifier que $(e'_1, \dots, e'_n)$ est une base orthonormée qui convient, avec par récurrence

$$P_n = \begin{pmatrix} & & -\frac{\langle e'_1, e_n \rangle}{\|e_n - p_n\|} \\ & P_{n-1} & \vdots \\ & & -\frac{\langle e'_{n-1}, e_n \rangle}{\|e_n - p_n\|} \\ 0 & \cdots \quad 0 & \frac{1}{\|e_n - p_n\|} \end{pmatrix}.$$

□

Le théorème de classification suivant est l'un des théorèmes majeurs de ce chapitre.

Théorème 1.12.

(1) **Classification des formes quadratiques complexes.** *Si k est algébriquement clos, pour tout $n \in \mathbb{N}$, il existe une et une seule classe d'équivalence de formes bilinéaires f symétriques non dégénérées de rang n. Il existe alors une base*[37] *de l'espace vectoriel de définition de f dans laquelle la matrice de f est la matrice identité, et la forme quadratique associée est, en notant $(x_1, \dots, x_n)$ les coordonnées dans cette base d'un vecteur x,*

$$q(x) = x_1^2 + x_2^2 + \cdots + x_n^2 \ .$$

De plus, l'indice de f est $\nu(f) = \lfloor \frac{n}{2} \rfloor$.

(2) **Classification des formes quadratiques réelles.** *Si $(k, \sigma) = (\mathbb{R}, \mathrm{id})$, pour tout entier $n \in \mathbb{N}$, il existe exactement $n+1$ classes d'équivalence de formes bilinéaires f symétriques non dégénérées de rang n. Il existe un unique $p \in \{0, \dots, n\}$ tel qu'il existe une base*[38] *de l'espace vectoriel de définition de f dans laquelle la matrice de f soit la matrice par blocs*

$$I_{p,\,n-p} = \begin{pmatrix} -I_p & 0 \\ 0 & I_{n-p} \end{pmatrix}$$

où I_k est la matrice identité $k \times k$. La forme quadratique associée est, en notant $(x_1, \dots, x_n)$ les coordonnées dans cette base d'un vecteur x,

$$q(x) = -x_1^2 - x_2^2 - \cdots - x_p^2 + x_{p+1}^2 + \cdots + x_n^2 \ .$$

De plus, l'indice de f est $\nu(f) = \min\{p,\, n-p\}$.

(3) **Classification des formes hermitiennes complexes.** *Si $(k, \sigma) = (\mathbb{C}, z \mapsto \overline{z})$, pour tout $n \in \mathbb{N}$, il existe exactement $n+1$ classes d'équivalence de formes sesquilinéaires f hermitiennes non dégénérées de rang n. Il existe un unique $p \in \{0, \dots, n\}$ tel qu'il existe une base*[39] *de l'espace vectoriel de définition de f dans laquelle la matrice de f soit la matrice par blocs*

$$I_{p,\,n-p} = \begin{pmatrix} -I_p & 0 \\ 0 & I_{n-p} \end{pmatrix}$$

où I_k est la matrice identité $k \times k$. La forme quadratique hermitienne associée est, en notant $(z_1, \dots, z_n)$ les coordonnées dans cette base d'un vecteur z,

$$q(z) = -|z_1|^2 - |z_2|^2 - \cdots - |z_p|^2 + |z_{p+1}|^2 + \cdots + |z_n|^2 \ .$$

Dans les cas (2) et (3), le couple $(p, n-p)$ s'appelle la *signature* de la forme quadratique ou quadratique hermitienne non dégénérée q. Certains ouvrages appellent signature le couple $(n-p, p)$, et il convient de préciser ce choix pour lever toute ambiguité : notre convention est que le premier terme p représente le nombre de signes $-$, et le second terme $n-p$ le nombre de signes $+$. Ce théorème dit que la signature caractérise à équivalence près la forme quadratique réelle ou la forme quadratique hermitienne complexe non dégénérée. Le fait que la signature soit bien définie (c'est-à-dire indépendante de la base qui permet d'écrire q sous cette forme) s'appelle la *loi d'inertie de Sylvester*.

37. qui est donc une base orthonormée de f
38. Cette base est orthonormée si et seulement si $p = 0$.
39. Cette base est orthonormée si et seulement si $p = 0$.

Par exemple, une forme quadratique réelle ou une forme quadratique hermitienne complexe est définie positive si et seulement si sa signature est $(0, n)$, où n est la dimension de l'espace vectoriel considéré.

Nous renvoyons par exemple à [Per1, §5] pour la classification lorsque k est un corps fini, et à [Ser] si $(k, \sigma) = (\mathbb{Q}, \mathrm{id})$.

Interprétation matricielle. L'application qui à une isométrie vectorielle de f associe sa matrice dans une base (qui n'est pas forcément unique) donnée par le théorème précédent induit donc des isomorphismes de groupes suivants.

Si f est une forme bilinéaire symétrique non dégénérée réelle de signature $(p, n-p)$, alors

$$\mathrm{O}(f) \simeq \mathrm{O}(p, n-p) = \{X \in \mathrm{GL}_n(\mathbb{R}) \ : \ {}^t X \, I_{p,\, n-p} \, X = I_{p,\, n-p}\} \ ,$$

$$\mathrm{SO}(f) \simeq \mathrm{SO}(p, n-p) = \{X \in \mathrm{O}(p, n-p) \ : \ \det X = 1\} \ .$$

Comme déjà dit dans la remarque (4) de la partie 1.5, nous noterons

$$\mathrm{O}(n) = \mathrm{O}(0, n) = \{X \in \mathrm{GL}_n(\mathbb{R}) \ : \ {}^t X \, X = I_n\}$$

et $\mathrm{SO}(n) = \mathrm{SO}(0, n)$. Si f est une forme sesquilinéaire hermitienne non dégénérée complexe de signature $(p, n-p)$, alors

$$\mathrm{U}(f) \simeq \mathrm{U}(p, n-p) = \{X \in \mathrm{GL}_n(\mathbb{C}) \ : \ {}^t X \, I_{p,\, n-p} \, \overline{X} = I_{p,\, n-p}\} \ ,$$

$$\mathrm{SU}(f) \simeq \mathrm{SU}(p, n-p) = \{X \in \mathrm{U}(p, n-p) \ : \ \det X = 1\} \ .$$

Comme déjà dit dans la remarque (4) de la partie 1.5, nous noterons

$$\mathrm{U}(n) = \mathrm{U}(0, n) = \{X \in \mathrm{GL}_n(\mathbb{C}) \ : \ {}^t X \, \overline{X} = I_n\}$$

et $\mathrm{SU}(n) = \mathrm{SU}(0, n)$.

Démonstration. (Voir par exemple [Per1, §V].) Par la proposition 1.10, pour tout espace vectoriel E sur k, pour toute forme bilinéaire symétrique (respectivement forme sesquilinéaire hermitienne) non dégénérée f de rang n sur E (en particulier la dimension de E est égale à n), de forme quadratique (respectivement forme quadratique hermitienne) associée q, fixons une base orthogonale $\mathscr{B} = (e_1, \ldots, e_n)$ de E pour f et, pour $j = 1, \ldots, n$, notons $a_j = q(e_j)$, qui est non nul puisque q est non dégénérée.

(1) Si k est algébriquement clos, pour $j = 1, \ldots, n$, il existe $b_j \in k^\times$ tel que $b_j^2 = a_j$. Alors la matrice de f dans la base $(\frac{1}{b_1} e_1, \ldots, \frac{1}{b_n} e_n)$ est la matrice identité. Pour toute base de E, la forme quadratique dont la matrice dans cette base est égale à la matrice identité, est symétrique, non dégénérée, de rang n. Soit f' une autre forme quadratique non dégénérée de rang n sur un espace vectoriel E'. Alors E et E' ont même dimension n, et si $\mathscr{B}'$ est une base de E' dans laquelle la matrice de f' est la matrice identité, alors l'unique isomorphisme linéaire de E dans E' qui envoie $\mathscr{B}$ sur $\mathscr{B}'$ est une conjugaison entre f et f'.

Par la formule (9), nous savons que $\nu(f) \leq \nu = \lfloor \frac{n}{2} \rfloor$. Soit $i \in k^\times$ une racine carrée de -1 dans k. Alors les vecteurs $e_1 + i\, e_{\nu+1}, e_2 + i\, e_{\nu+2}, \ldots, e_\nu + i\, e_{2\nu}$ sont ν vecteurs isotropes linéairement indépendants deux à deux orthogonaux, donc qui engendrent un sous-espace totalement isotrope de dimension ν. Donc $\nu(f) \geq \nu$, d'où $\nu(f) = \nu$.

(2) et (3) Puisque les a_j sont des réels non nuls, il existe $p \in \{0, \dots, n\}$ tel que, quitte à permuter la base, nous ayons $a_1, \dots, a_p < 0$ et $a_{p+1}, \dots, a_n > 0$. Pour $j = 1, \dots, n$, il existe $b_j \in \mathbb{R}^\times$ tel que $-b_j^2 = a_j$ si $j \leq p$ et $b_j^2 = a_j$ si $j > p$. Alors la matrice de f dans la base $(\frac{1}{b_1} e_1, \dots, \frac{1}{b_n} e_n)$ est la matrice cherchée.

Si E admet une base $(e'_1, \dots, e'_n)$ dans laquelle la matrice de q est

$$\begin{pmatrix} -I_{p'} & 0 \\ 0 & I_{n-p'} \end{pmatrix}$$

avec $p' \in \{0, \dots, n\}$, alors, en notant $F = \mathrm{Vect}\{e_1, \dots, e_p\}$ et $F' = \mathrm{Vect}\{e'_{p'+1}, \dots, e'_n\}$, qui sont des sous-espaces vectoriels de dimension p et $n - p'$ respectivement, nous avons $q(x) < 0$ pour tout $x \in F - \{0\}$ et $q(x) > 0$ pour tout $x \in F' - \{0\}$, donc $F \cap F' = \{0\}$, et $p + n - p' \leq n$. Donc $p \leq p'$ et par symétrie $p = p'$. Par conséquent p est uniquement déterminé.

L'application $f \mapsto p$, bien définie par ce qui précède, induit une bijection de l'ensemble des classes d'isomorphisme de formes f vérifiant les hypothèses de l'assertion (2) (respectivement (3)) vers l'ensemble $\{0, \dots, n\}$ de cardinal $n+1$, par le même argument que pour l'assertion (1).

En supposant par exemple que $p \leq n - p$, les vecteurs $e_1 + e_{p+1}, e_2 + e_{p+2}, \dots, e_p + e_{2p}$ sont p vecteurs isotropes linéairement indépendants deux à deux orthogonaux, donc qui engendrent un sous-espace totalement isotrope de dimension p. Donc $\nu(f) \geq \min\{p, n-p\}$. Nous renvoyons à [Per1, §VIII] pour l'inégalité opposée. □

Exercice E.7. *(1) Soient $n \in \mathbb{N} - \{0\}$ et $\mathbb{C}^n$ l'espace vectoriel complexe hermitien usuel. Soient $c_0, \dots, c_{n-1}$ des nombres complexes, et*

$$C = C(c_0, c_1, \dots, c_{n-1}) = \begin{pmatrix} c_0 & c_1 & \cdots & c_{n-1} \\ c_{n-1} & c_0 & \cdots & c_{n-2} \\ \vdots & & \ddots & \vdots \\ c_1 & c_2 & \cdots & c_0 \end{pmatrix}$$

la matrice circulante *de $c_0, \dots, c_{n-1}$. Montrer que pour tout entier $k \in \mathbb{Z}$, si $\zeta = e^{2i\pi/n}$ est la racine primitive n-ème de l'unité standard, alors $(1, \zeta^k, \dots, \zeta^{(n-1)k})$ est un vecteur propre de C. Montrer que C est diagonalisable dans une base orthonormée de $\mathbb{C}^n$. Calculer le déterminant de C. La matrice C est-elle normale, c'est-à-dire commute-t-elle avec sa matrice adjointe $C^* = {}^t\overline{C}$?*

(2) Soient $n \in \mathbb{N} - \{0, 1\}$ et P un polygone quelconque du plan euclidien usuel à n sommets. Soit $(P_i)_{i \in \mathbb{N}}$ la suite de polygones définie par récurrence en posant $P_0 = P$ et P_{i+1} le polygone dont les sommets consécutifs sont les milieux des arêtes consécutives de P_i. Montrer que la suite $(P_i)_{i \in \mathbb{N}}$ converge pour la distance de Hausdorff vers le singleton constitué de l'isobarycentre de P. Que se passe-t-il si P_0 est convexe et si on s'autorise à dilater P_{i+1} pour qu'il soit de même aire que P_i ?

Exercice E.8. *Calculer le rang et la signature des formes quadratiques réelles ou formes quadratiques hermitiennes complexes suivantes :*

- $q(z_1, \dots, z_n) = \sum_{1 \leq i < j \leq n} z_i \overline{z_j}$ *sur* $\mathbb{C}^n$,
- $q(z_1, \dots, z_n) = \sum_{1 \leq i, j \leq n} (i + j)\, z_i \overline{z_j}$ *sur* $\mathbb{C}^n$,
- $q(z_1, \dots, z_n) = \sum_{1 \leq i, j \leq n} (i^2 + ij + j^2)\, z_i \overline{z_j}$ *sur* $\mathbb{C}^n$,

- $q(z_1, \dots, z_n) = \sum_{1 \leq i < j \leq n} |z_i - z_j|^2$ *sur* $\mathbb{C}^n$,
- $q(P) = \int_{-\infty}^{+\infty} e^{-t^2} P(t) \overline{P(-t)}\, dt$ *sur l'espace vectoriel complexe* $\mathbb{C}_n[X]$ *des polynômes complexes* P *de degré au plus* n,
- $q(A) = \mathtt{tr}(A^2)$ *sur* $\mathscr{M}_n(\mathbb{R})$, *ainsi qu'en restriction à son sous-espace vectoriel* $\mathrm{Sym}_n(\mathbb{R})$ *des matrices symétriques,*
- $q(A) = |\,\mathtt{tr}\, A|^2$ *sur* $\mathscr{M}_n(\mathbb{C})$.

1.8 Décomposition spectrale des opérateurs normaux en dimension finie

Soit E un espace vectoriel sur k de dimension finie n, muni d'une forme sesquilinéaire non dégénérée f, qui est symétrique ou hermitienne.[40] Supposons la caractéristique de k différente de 2.

Théorème 1.13. (Diagonalisation des opérateurs linéaires normaux des espaces hermitiens) *Supposons k algébriquement clos et f anisotrope.*

Un endomorphisme de E admet une base orthogonale formée de vecteurs propres si et seulement s'il est normal : pour tout $u \in \mathscr{L}(E)$, il existe une base orthogonale $\mathscr{B}$ de E dans laquelle la matrice de u est diagonale si et seulement si $u \circ u^* = u^* \circ u$.

En particulier, et c'est le seul cas à retenir pour l'agrégation de mathématique, un opérateur linéaire d'un espace hermitien (sur le corps algébriquement clos $\mathbb{C}$, et dont le produit scalaire est défini positif donc anisotrope) est diagonalisable en base orthonormée (en rendant orthonormée une base orthogonale, ce qui est possible sur $\mathbb{C}$) si et seulement s'il est normal.

Matriciellement, ceci se traduit de la manière suivante, en considérant, pour toute matrice $M \in \mathscr{M}_n(\mathbb{C})$, l'endomorphisme linéaire u de l'espace hermitien usuel $\mathbb{C}^n$ dont la matrice dans la base canonique est M, et en remarquant que la matrice de passage P de la base canonique à une base orthonormée $\mathscr{B}$ de $\mathbb{C}^n$ donnée par le théorème précédent (dans laquelle la matrice D de u est diagonale) est unitaire.[41]

Corollaire 1.14. *Pour tout $M \in \mathscr{M}_n(\mathbb{C})$, il existe une matrice diagonale $D \in \mathscr{M}_n(\mathbb{C})$ et une matrice unitaire $P \in \mathrm{U}(n)$ telles que*

$$M = P\, D\, P^* = P\, D\, P^{-1}$$

si et seulement si $M\, M^* = M^*\, M$. □

Démonstration. (Voir par exemple [Deh, page 256], [Gou, §5, page 270].) Montrons tout d'abord la nécessité de la condition de normalité[42]. Soient $u \in \mathscr{L}(E)$ et $\mathscr{B}$ une base orthogonale pour f dans laquelle la matrice M de u est diagonale. Alors la matrice A

40. Rappelons que « symétrique » est un cas particulier de « hermitienne », correspondant à l'égalité $\sigma = \mathrm{id}$. Nous ne mentionnons les deux cadres que pour des raisons mnémotechniques.

41. Une autre manière de montrer que la condition de normalité est nécessaire dans le corollaire 1.14 est la suivante. Soit $M \in \mathscr{M}_n(\mathbb{C})$. Supposons qu'il existe une matrice diagonale $D \in \mathscr{M}_n(\mathbb{C})$ et une matrice unitaire $P \in \mathrm{U}(n)$ telles que $M = P\, D\, P^*$. Alors, puisque deux matrices diagonales commutent et puisque $P^* = P^{-1}$, nous avons

$$M\, M^* = (P\, D\, P^*)(P\, D\, P^*)^* = P\, D\, D^*\, P^* = P\, D^*\, D\, P^* = (P\, D\, P^*)^*(P\, D\, P^*) = M^*\, M\,,$$

donc M et M^* commutent.

42. Cette implication n'utilise pas les hypothèses k algébriquement clos et f anisotrope.

de f dans la base $\mathscr{B}$ est diagonale, et la matrice M^* de u^* dans la base $\mathscr{B}$, qui est $M^* = (A^\sigma)^{-1}\,{}^tM^\sigma A^\sigma$ par la formule (12), est aussi diagonale. Comme deux matrices diagonales commutent, nous en déduisons donc que u et u^* commutent.

Réciproquement, montrons par récurrence sur la dimension n de E que tout endomorphisme normal de E est diagonalisable en base orthogonale de E. Si n vaut 0 ou 1, le résultat est immédiat. Supposons que $n \geq 2$, et soit $u \in \mathscr{L}(E)$ normal.

Puisque k est algébriquement clos, tout polynôme non constant à coefficients dans k admet au moins une racine dans k. Donc u admet au moins un vecteur propre non nul e_1 (associé à une valeur propre λ). Par la proposition 1.8 (vii), puisque f est anisotrope et u est normal, e_1 est aussi un vecteur propre de u^* (associé à la valeur propre λ^σ). Comme f est anisotrope, l'orthogonal de la droite vectorielle $k\,e_1$ (qui est stable par u et par u^*) est un hyperplan vectoriel F de E supplémentaire à la droite vectorielle $k\,e_1$. Par la proposition 1.8 (v), l'hyperplan vectoriel F est stable par u^* et $(u^*)^* = u$. La restriction de u à F est encore normale, et la restriction de f à F est encore anisotrope. Par récurrence, soit $(e_2, \ldots, e_n)$ une base orthogonale de F formée de vecteurs propres de $u_{|F}$. Alors $(e_1, e_2, \ldots, e_n)$ est une base orthogonale de E formée de vecteurs propres de u. □

Les corollaires suivants s'obtiennent en appliquant le théorème précédant respectivement dans les cas des endomorphismes hermitiens, antihermitiens, et unitaires d'un espace vectoriel (complexe) hermitien, par la proposition 1.8 (vi).

Corollaire 1.15.

(1) **(Théorème spectral hermitien)** *Soit E un espace hermitien. Soit $u \in \mathscr{L}(E)$ un endomorphisme hermitien (respectivement anti-hermitien ou unitaire). Alors u est diagonalisable dans une base orthonormée de E, de valeurs propres réelles (respectivement imaginaires pures ou de module 1).*

(2) Soit $M \in \mathscr{M}_n(\mathbb{C})$ une matrice hermitienne (respectivement hermitienne définie positive, anti-hermitienne, unitaire). Alors il existe une matrice unitaire $P \in \mathrm{U}(n)$ et une matrice diagonale $D \in \mathscr{M}_n(\mathbb{C})$ à coefficients diagonaux réels (respectivement strictement positifs, imaginaires purs, de module 1) telles que

$$M = P\,D\,P^* = P\,D\,P^{-1}\,. \quad \square$$

Nous reviendrons sur ce point dans la partie 3.3.

Le théorème 1.13 n'est plus valable sans l'hypothèse algébriquement clos. Dans le cas euclidien, nous avons tout de même un joli résultat de structure.

Théorème 1.16. (Décomposition spectrale des opérateurs linéaires normaux des espaces euclidiens) *Soit E un espace euclidien de dimension n.*

(1) Un endomorphisme u de E est normal si et seulement s'il est somme directe orthogonale d'homothéties et de similitudes planes, c'est-à-dire s'il existe des entiers $r, s \in \mathbb{N}$ tels que $r + 2s = n$ et une somme directe orthogonale

$$E = D_1 \overset{\perp}{\oplus} \ldots \overset{\perp}{\oplus} D_r \overset{\perp}{\oplus} P_1 \overset{\perp}{\oplus} \ldots \overset{\perp}{\oplus} P_s$$

où $D_1, \ldots, D_r$ sont des droites vectorielles invariantes par u, sur chacune desquelles u agit par une homothétie, et où $P_1, \ldots, P_s$ sont des plans vectoriels invariants par u, sur chacun desquels u agit par une similitude vectorielle différente d'une homothétie.

(2) Pour tout $M \in \mathscr{M}_n(\mathbb{R})$ telle que $M\,{}^tM = {}^tM\,M$, il existe des entiers $r, s \in \mathbb{N}$ tels que $r + 2s = n$, une matrice réelle

$$D = \begin{pmatrix} a_1 & & & & & & & \\ & \ddots & & & & & & \\ & & a_r & & & & & \\ & & & b_1 & -c_1 & & & \\ & & & c_1 & b_1 & & & \\ & & & & & \ddots & & \\ & & & & & & b_s & -c_s \\ & & & & & & c_s & b_s \end{pmatrix} \in \mathscr{M}_n(\mathbb{R})$$

diagonale par blocs de taille 1×1 ou 2×2, avec $a_1, \ldots, a_r, b_1, \ldots, b_s$ des réels et $c_1, \ldots, c_s$ des réels non nuls, de sorte que chaque bloc 2×2 appartienne à la partie $\mathbb{R}^*(\mathrm{O}(2) - \{\pm I_2\})$ de $\mathscr{M}_2(\mathbb{R})$, et une matrice orthogonale $P \in \mathrm{O}(n)$ tels que

$$M = P\,D\,{}^tP = P\,D\,P^{-1}\ .$$

Remarque. Une matrice P comme dans l'assertion (2) n'est pas forcément unique (par exemple si $M = \mathrm{id}$). Une matrice D comme dans l'assertion (2) est unique modulo permutation des blocs 1×1 et modulo permutation et transposition des blocs 2×2 (ou modulo permutation des blocs 1×1 entre eux et des blocs 2×2 entre eux, si on demande de plus que $c_1, \ldots, c_s > 0$). Ceci découle de la propriété d'unicité de la décomposition en espaces propres sur $\mathbb{C}$, et du fait que $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ est un élément d'ordre 2 de $\mathrm{O}(2)$, et que

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

Démonstration. (Voir par exemple [Deh, page 263], [Gou, §5, page 271].) (2) Rappelons [43] qu'une similitude vectorielle v est la composée d'une homothétie et d'une transformation orthogonale, donc est en dimension 2 de matrice, dans toute base orthonormée, de la forme $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ avec $a, b \in \mathbb{R}$, où $b \neq 0$ si et seulement si v n'est pas une homothétie. La seconde assertion découle de la première, en prenant pour u l'endomorphisme de $\mathbb{R}^n$ dont la matrice dans la base canonique de $\mathbb{R}^n$ est M, en remarquant que la matrice de passage P de la base canonique à la base orthonormée $\mathscr{B}$ de $\mathbb{R}^n$ obtenue par concaténation de bases orthornormées de $D_1, \ldots, D_r, P_1, \ldots, P_s$, est orthogonale.

(1) Si u est somme directe d'homothéties et de similitudes planes, alors u est normal, car les homothéties sont auto-adjointes et centrales, et les transformations orthogonales sont unitaires.

Réciproquement, supposons u normal. Montrons le résultat par récurrence sur n. Si n vaut 0 ou 1, le résultat est immédiat. Supposons que $n \geq 2$, et soit $u \in \mathscr{L}(E)$ normal.

Si u admet une valeur propre réelle λ, soit $e_1 \in E$ un vecteur propre non nul de u associé à λ. C'est aussi un vecteur propre de u^* (pour la valeur propre $\lambda^\sigma = \lambda$), par la

43. Voir la remarque (3) de la partie 1.5.

proposition 1.8 (vii). En particulier, la droite vectorielle $\mathbb{R}e_1$ est stable par u et u^*. Donc son orthogonal F est stable par u^* et $(u^*)^* = u$ par la proposition 1.8 (v), et la restriction de u à F, de dimension égale à $n-1$, est normale. Nous concluons par récurrence et par concaténation de blocs.

Supposons donc que u n'admette pas de valeur propre réelle. Quitte à conjuguer (deux espaces euclidiens de même dimension sont isomorphes (par exemple par le théorème de classification 1.12 (2)), et les conjugaisons entre produits scalaires préservent droites vectorielles, plans vectoriels, orthogonalité et similitudes vectorielles), nous pouvons supposer que E est l'espace euclidien usuel $\mathbb{R}^n$. Notons encore $u : z = x+i\,y \mapsto u(x)+i\,u(y)$ l'unique extension $\mathbb{C}$-linéaire de u à $\mathbb{C}^n = \mathbb{R}^n + i\,\mathbb{R}^n$. Pour tous les $z = x + i\,y$, $z' = x' + i\,y'$ dans $\mathbb{C}^n = \mathbb{R}^n + i\,\mathbb{R}^n$, notons de la même manière

$$\langle z, z'\rangle = (\langle x, x'\rangle + \langle y, y'\rangle) + i(\langle y, x'\rangle - \langle x, y'\rangle)$$

l'extension du produit scalaire euclidien usuel de $\mathbb{R}^n$ en un nouveau produit scalaire (car $\langle z, z\rangle = \|x\|^2 + \|y\|^2$) hermitien sur $\mathbb{C}^n = \mathbb{R}^n + i\,\mathbb{R}^n$.

Notons que les extensions linéaires à $\mathbb{C}^n$ de deux endomorphismes commutants de $\mathbb{R}^n$ commutent encore, et que l'extension linéaire de u^* est l'adjoint de l'extension de u pour l'extension du produit scalaire ci-dessus, [44] et donc que l'extension de u à $\mathbb{C}^n$ est encore normale.

Par la proposition 1.8 (vii), il existe un vecteur $z = x + iy$ de $\mathbb{C}^n = \mathbb{R}^n + i\,\mathbb{R}^n$ non nul (donc $x \neq 0$ ou $y \neq 0$) qui est propre pour u de valeur propre $\lambda = a + ib$ et qui est propre pour u^* de valeur propre $\overline{\lambda} = a - ib$. Notons que $b \neq 0$ puisque u n'admet pas de valeur propre λ réelle. Nous avons donc

$$u(z) = \lambda\, z \Leftrightarrow \left\{ \begin{array}{l} u(x) = a\,x - b\,y \\ u(y) = b\,x + a\,y \end{array} \right. \tag{19}$$

et

$$u^*(z) = \overline{\lambda}\, z \Leftrightarrow \left\{ \begin{array}{l} u^*(x) = a\,x + b\,y \\ u^*(y) = -b\,x + a\,y\ . \end{array} \right. \tag{20}$$

Notons V le sous-espace vectoriel (réel) de $\mathbb{R}^n$ engendré par x et y, qui est donc stable par u et par u^* (par les formules (19) et (20)), et non nul. Notons que $\overline{z} = x - iy$ est un vecteur propre de u pour la valeur propre $\overline{\lambda} \neq \lambda$, car

$$u(\overline{z}) = u(x) - i\,u(y) = (a\,x - b\,y) - i(b\,x + a\,y) = (a - i\,b)(x - i\,y) = \overline{\lambda}\,\overline{z}\ .$$

Comme l'extension de u à $\mathbb{C}^n$ est diagonalisable en base orthonormée par le théorème 1.13, les espaces propres associés à deux valeurs propres différentes sont orthogonaux, et nous avons

$$0 = \langle z, \overline{z}\rangle = (\|x\|^2 - \|y\|^2) + 2\,i\langle x, y\rangle\ .$$

Donc $\|x\|^2 = \|y\|^2$ et $\langle x, y\rangle = 0$. Quitte à diviser z par $\|x\|$, nous obtenons que (x, y) est une base orthonormée de V.

44. En effet, avec les notations ci-dessus, nous avons

$$\begin{aligned} \langle z, u^*(z')\rangle = \langle u(z), z'\rangle &= (\langle u(x), x'\rangle + \langle u(y), y'\rangle) + i(\langle u(y), x'\rangle - \langle u(x), y'\rangle) \\ &= (\langle x, u^*(x')\rangle + \langle y, u^*(y')\rangle) + i(\langle y, u^*(x')\rangle - \langle x, u^*(y')\rangle)\ . \end{aligned}$$

La matrice de la restriction de u à V est, dans la base orthonormée (x, y), égale à $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ par la formule (19). Elle appartient à $\mathbb{R}^*(\mathrm{SO}(2) - \{\pm I_2\})$, en mettant en facteur le scalaire $\sqrt{a^2+b^2} > 0$, puisque $b \neq 0$ et $\begin{pmatrix} a' & -b' \\ b' & a' \end{pmatrix}$ est une matrice de rotation euclidienne plane si ${a'}^2 + {b'}^2 = 1$. Donc la restriction de u à V est une similitude vectorielle de rapport $\sqrt{a^2+b^2}$. Si $n = 2$, ceci montre le résultat.

Si $n > 2$, l'orthogonal $V^\perp$ de V dans $\mathbb{R}^n$ est stable par u^* et $(u^*)^* = u$, et les restrictions de u aux espaces euclidiens V et $V^\perp$, de dimensions strictement inférieures à n, sont donc normales. Nous concluons aussi par récurrence et par concaténation de blocs. □

Le corollaire (1) ci-dessous est un cas particulier, à retenir prioritairement, du théorème 1.16 (1), puisqu'un endomorphisme autoadjoint est normal.

Les valeurs propres complexes des matrices symétriques, antisymétriques, et orthogonales sont respectivement réelles, imaginaires pures, et de module 1, par la proposition 1.8 (vi). Les corollaires (2), (3), (4) ci-dessous s'obtiennent donc en appliquant le théorème 1.16 (2) respectivement dans les cas des matrices symétriques, antisymétriques et orthogonales.

Corollaire 1.17. *(1)* **(Théorème spectral euclidien)** *Soit $u \in \mathscr{L}(E)$ un endomorphisme autoadjoint d'un espace euclidien E. Alors u est diagonalisable dans une base orthonormée de E.*

(2) Soit $M \in \mathscr{M}_n(\mathbb{R})$ une matrice symétrique (respectivement symétrique définie positive). Alors il existe une matrice orthogonale $P \in \mathrm{O}(n)$ et une matrice diagonale D à coefficients réels (respectivement strictement positifs) tels que

$$M = P\,D\,{}^tP = P\,D\,P^{-1}\,.$$

(3) Soit $M \in \mathscr{M}_n(\mathbb{R})$ une matrice antisymétrique. Alors il existe une matrice orthogonale $P \in \mathrm{O}(n)$, des entiers $r, s \in \mathbb{N}$ tels que $r + 2s = n$, et des réels $c_1, \ldots, c_s$ tels que si

$$D = \begin{pmatrix} 0 & & & & & & & \\ & \ddots & & & & & & \\ & & 0 & & & & & \\ & & & 0 & -c_1 & & & \\ & & & c_1 & 0 & & & \\ & & & & & \ddots & & \\ & & & & & & 0 & -c_s \\ & & & & & & c_s & 0 \end{pmatrix},$$

alors $M = P\,D\,{}^tP = P\,D\,P^{-1}$.

(4) Soit $M \in \mathscr{M}_n(\mathbb{R})$ une matrice orthogonale. Alors il existe une matrice orthogonale $P \in \mathrm{O}(n)$, des entiers $r, s \in \mathbb{N}$ tels que $r + 2s = n$, des éléments $a_1, \ldots, a_r$ de $\{\pm 1\}$ et des

matrices $A_1, \dots, A_s \in \mathrm{O}(2) - \{\pm I_2\}$ *tels que si*

$$D = \begin{pmatrix} a_1 & & & & & & \\ & \ddots & & & & & \\ & & a_r & & & & \\ & & & A_1 & & & \\ & & & & \ddots & & \\ & & & & & A_s \end{pmatrix},$$

alors $M = P\,D\,{}^tP = P\,D\,P^{-1}$. □

Nous reviendrons sur le point (4) dans la partie 2.3.

Exercice E.9.

(1) *Soit* $A = (a_{i,j})_{1\le i,j\le n} \in \mathscr{M}_n(\mathbb{C})$ *une matrice hermitienne positive. Montrer l'inégalité d'Hadamard*

$$\det A \le \prod_{i=1}^{n} a_{i,i}\ .$$

Quand y a-t-il égalité ?

(2) *Soient* E *un espace hermitien de dimension* n*, et* f *un endomorphisme auto-adjoint de* E *défini positif, de valeurs propres* $\lambda_1, \dots, \lambda_n$*. Notons* $\mathscr{A}$ *l'ensemble des bases orthonormées de* E*. Montrer que*

$$\prod_{i=1}^{n} \lambda_i = \min_{(e_1,\dots,e_n)\in\mathscr{A}} \prod_{i=1}^{n} \langle f(e_i), e_i\rangle\ .$$

1.9 Diagonalisation simultanée de formes quadratiques réelles

Soit E un espace vectoriel de dimension finie n sur k. Rappelons qu'une forme quadratique hermitienne q sur E est *diagonale* dans une base $\mathscr{B}$ de E s'il existe des scalaires $a_1, \dots, a_n$ (nécessairement dans le sous-corps fixe k_σ de k pour σ) tels que pour tout vecteur $x \in E$ de coordonnées $(x_1, \dots, x_n)$ dans la base $\mathscr{B}$, nous avons

$$q(x) = \sum_{i=1}^{n} a_i\ x_i x_i^\sigma\ ,$$

donc $q(x) = \sum_{i=1}^n a_i\ x_i^2$ si $\sigma = \mathrm{id}$. Nous dirons que q est *diagonalisable* s'il existe une base de E dans laquelle elle est diagonale. Nous dirons qu'un ensemble Q de formes quadratiques hermitiennes sur E est *simultanément diagonalisable* s'il existe une base de E dans laquelle tout élément de Q est diagonal.

Étant donné deux formes quadratiques hermitiennes q et q' sur E, il n'existe pas toujours de base $\mathscr{B}$ de E dans laquelle ces formes sont toutes les deux diagonales. Nous pouvons prendre par exemple $(k, \sigma) = (\mathbb{R}, \mathrm{id})$, $E = \mathbb{R}^2$ et les deux formes quadratiques non dégénérées

$$q(x, y) = x^2 - y^2 \quad \text{et} \quad q'(x, y) = 2\,xy\ .$$

La véritable explication découlera du lemme 1.18, mais voici un argument direct. Si deux vecteurs (x, y) et (x', y') sont orthogonaux pour les deux formes quadratiques q et q', alors

$xx' - yy' = 0$ et $xy' + x'y = 0$, ce qui implique (par les formules de Cramer) que l'un des deux vecteurs est nul. En particulier, il n'existe pas de base orthogonale commune à q et à q'. Remarquons que ces deux formes sont de signature $(1,1)$. En particulier, elles ne sont ni définies positives ni définies négatives. Le résultat 1.19 ci-dessous dit que c'est la seule obstruction sur les corps réel et complexe.

L'outil principal pour le problème de diagonalisation simultanée est le suivant. Soient q et q' deux formes quadratiques hermitiennes sur E, de forme polaires f et f', telles que q soit non dégénérée. Notons $\widetilde{f} : E \to (E^\sigma)^*$ l'isomorphisme de dualité (voir la partie 1.1) défini par f. Montrons qu'il existe un unique opérateur linéaire $u = u_{q,q'} \in \mathscr{L}(E)$, appelé *l'endomorphisme de passage de q à q'*, tel que pour tous les $x, y \in E$, nous ayons

$$f'(x,y) = f(u(x), y) \ .$$

Il suffit en effet de poser, comme il est nécessaire, lorsque x parcours E,

$$u : x \mapsto (\widetilde{f})^{-1}(y \mapsto f'(x,y)) \ ,$$

qui est bien définie par la semi-linéarité à droite de f' (qui implique que l'application $y \mapsto f'(x,y)$ est une forme linéaire sur E^σ) et qui est linéaire par la linéarité à gauche de f' et la linéarité de $\widetilde{f}$.

Interprétation matricielle. Si $\mathscr{B}$ est une base de E, si A (respectivement A') est la matrice de f (respectivement f') dans la base $\mathscr{B}$, alors la matrice M de l'endomorphisme de passage $u = u_{q,q'}$ dans la base $\mathscr{B}$ est donnée par [45]

$$M = ({}^tA)^{-1}\, {}^tA' \ . \tag{21}$$

Lemme 1.18. *(1) L'endomorphisme de passage $u_{q,q'}$ est inversible si et seulement si q' est non dégénérée.*

(2) L'endomorphisme de passage $u_{q,q'}$ est auto-adjoint pour la forme polaire f de q.

(3) Deux formes quadratiques hermitiennes q et q' sur E, telles que q soit non dégénérée, sont simultanément diagonalisables si et seulement si l'endomorphisme de passage $u_{q,q'}$ est diagonalisable en base orthogonale pour q.

Démonstration. (1) Ceci découle par exemple de la formule (21).

(2) Ceci découle du fait que f et f' sont hermitiennes : pour tous les $x, y \in E$, nous avons

$$f(u(x), y) = f'(x,y) = f'(y,x)^\sigma = f(u(y), x)^\sigma = f(x, u(y)) \ .$$

(3) Posons $u = u_{q,q'}$. S'il existe une base $\mathscr{B} = (e_1, \dots, e_n)$ de E orthogonale pour q et q', alors pour tout $1 \leq i \leq n$, le vecteur $u(e_i)$ est orthogonal pour f à tout vecteur e_j pour $j \neq i$, car $f'(e_i, e_j) = 0$. Donc $u(e_i)$ est colinéaire à e_i car q est non dégénérée et diagonale dans $\mathscr{B}$. Par conséquent, u est diagonalisable dans $\mathscr{B}$.

45. En effet, si X et Y sont les matrices colonnes des coordonnées de vecteurs quelconques x et y de E dans la base $\mathscr{B}$, alors nous avons

$${}^tX\, A'\, Y^\sigma = f'(x,y) = f(u(x), y) = {}^t(MX)\, A\, Y^\sigma = {}^tX\, ({}^tM\, A)\, Y^\sigma \ ,$$

donc $A' = {}^tM\, A$. Rappelons que A est inversible si et seulement si q est non dégénérée.

Réciproquement, si $\mathscr{B} = (e_1, \dots, e_n)$ est une base de E orthogonale pour q dans laquelle la matrice de u est diagonale, de coefficients diagonaux $\lambda_1, \dots, \lambda_n$, alors pour tous les $i \neq j$, nous avons

$$f'(e_i, e_j) = f(u(e_i), e_j) = \lambda_i \, f(e_i, e_j) = 0 \, .$$

Donc q' est diagonale dans $\mathscr{B}$. □

Proposition 1.19. (Théorème de diagonalisation simultanée des formes quadratiques réelles ou quadratiques hermitiennes complexes) *Supposons que nous ayons* $(k, \sigma) = (\mathbb{R}, \mathrm{id})$ *(respectivement* $(k, \sigma) = (\mathbb{C}, z \mapsto \overline{z})$*). Soit q une forme quadratique (respectivement quadratique hermitienne) non dégénérée sur E. Les assertions suivantes sont équivalentes.*

(1) Pour toute forme quadratique (respectivement quadratique hermitienne) q' sur E, il existe une base de E orthogonale pour q et q'.

(2) La forme q est définie positive ou définie négative.

Démonstration. Supposons que q ne soit ni définie positive ni définie négative. Par le théorème de classification 1.12 (2) (respectivement (3)), la dimension de E est au moins 2 et il existe une base $(e_1, \dots, e_n)$ de E orthogonale pour q telle que $q(e_1) = +1$ et $q(e_2) = -1$. Considérons la forme quadratique hermitienne

$$q' : x \mapsto (x_1 + x_2)x_1^\sigma - x_2 x_2^\sigma + \sum_{i=3}^{n} q(e_i) \, x_i x_i^\sigma \, ,$$

où $(x_1, \dots, x_n)$ sont les coordonnées de x dans cette base. Alors l'endomorphisme $u = u_{q,q'}$ de passage de q à q' a pour matrice dans cette base

$$\begin{pmatrix} 1 & 1 & \\ 0 & 1 & \\ & & 0 \\ 0 & & I_{n-2} \end{pmatrix} \, .$$

Comme cette matrice (qui est sous forme de Jordan) n'est pas diagonalisable sur k (par exemple car la seule matrice $n \times n$ à coefficients dans k, diagonalisable sur k, dont toutes les valeurs propres sont égales à 1, est la matrice identité I_n), il découle du lemme 1.18 (3) ci-dessus que q et q' ne sont pas simultanément diagonalisables.

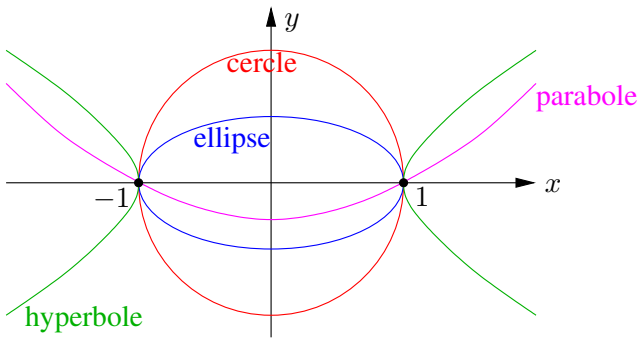
Supposons, quitte à remplacer q par $-q$, que q soit définie positive. Soit q' une forme quadratique hermitienne sur E. Alors l'endomorphisme de passage $u = u_{q,q'}$, qui est autoadjoint dans l'espace euclidien (respectivement hermitien) (E, f) par le lemme 1.18 (2), est diagonalisable en base orthonormée pour q par le corollaire 1.17 (1) (respectivement 1.15 (1)). Donc par le lemme 1.18 (3), les formes q et q' sont simultanément diagonalisables. □

Interprétation géométrique. Ce résultat a une interprétation géométrique intéressante, y compris dans le cas où les deux formes sont définies positives. Nous renvoyons par exemple à [Ber4, Aud] pour l'étude des classes de similitude de coniques[46] du plan euclidien. À isométries de la distance euclidienne et à homothéties près, une conique non dégénérée est

- une ellipse d'équation $x^2 + a\,y^2 = 1$ avec $a > 1$,

46. Une *conique* est une ligne de niveau d'une fonction polynomiale sur $\mathbb{R}^2$ de degré 2, comme par exemple les formes quadratiques sur $\mathbb{R}^2$.

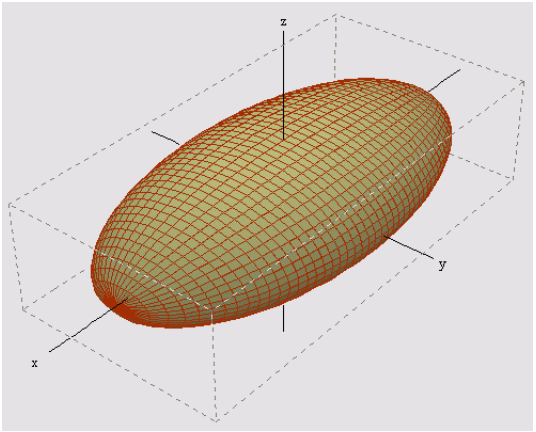
- un cercle d'équation $x^2 + y^2 = 1$,
- une parabole d'équation $x^2 - a\,y = 1$ avec $a > 0$,
- une hyperbole d'équation $x^2 - a\,y^2 = 1$ avec $a > 0$.

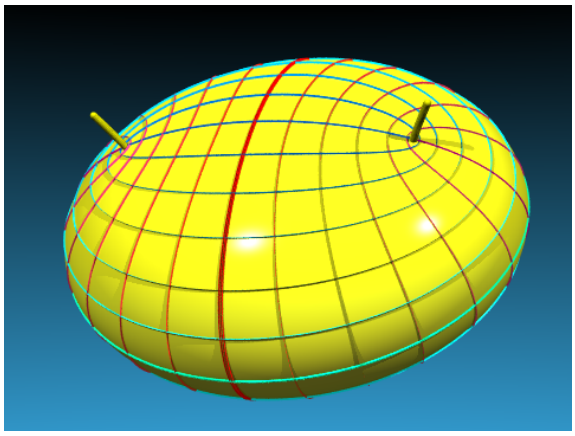


Fixons $n \in \mathbb{N} - \{0\}$. Pour toute forme quadratique définie positive q sur l'espace euclidien usuel $\mathbb{R}^n$, la boule unité $B_q = \{x \in \mathbb{R}^n : q(x) \leq 1\}$ de q est donc un *ellipsoïde* de centre 0, dont les *axes* sont les droites $\mathbb{R}e_i$ pour une base orthonormée $(e_1, \dots, e_n)$ de $\mathbb{R}^n$ dans laquelle q est diagonale et les *rayons* sont (avec multiplicité) les réels strictement positifs $\frac{1}{\sqrt{q(e_i)}}$ pour $1 \leq i \leq n$:

$$B_q = \Big\{ \sum_{i=1}^{n} x_i\, e_i : (x_1, \dots, x_n) \in \mathbb{R}^n \quad \text{et} \quad \sum_{i=1}^{n} q(e_i)\, x_i^2 \leq 1 \Big\}\,.$$

Notons que les axes ne sont définis de manière unique que si les rayons sont deux à deux distincts. Dans le cas contraire, l'ellipsoïde est invariant par des rotations. Voici des images, extraites respectivement de Wikipedia et du joli article en ligne [GL] dont je recommande les animations.





Remarque. Notons $\mathscr{Q}_n$ l'ensemble des formes quadratiques définies positives sur $\mathbb{R}^n$, muni de l'action naturelle (à gauche) de $\mathrm{GL}_n(\mathbb{R})$, de précomposition des formes quadratiques par l'inverse des automorphismes linéaires.

$$(g, q) \mapsto q \circ g^{-1}\,.$$

Pour tous les q, q' dans $\mathscr{Q}_n$, soit $\mathscr{B}$ une base (qui existe par la proposition 1.19 ci-dessus) de $\mathbb{R}^n$ dans laquelle q et q' sont diagonales, de coefficients diagonaux $(a_1, \dots, a_n)$ et $(a'_1, \dots, a'_n)$ respectivement (qui sont strictement positifs). Posons

$$d(q, q') = \sqrt{\sum_{i=1}^{n} \ln \frac{a_i}{a'_i}} \, .$$

Il est possible de montrer (voir par exemple [Pau3]) que $d(q, q')$ ne dépend pas du choix de $\mathscr{B}$, et définit une distance sur $\mathscr{Q}_n$, invariante par l'action du groupe $\mathrm{GL}_n(\mathbb{R})$.

1.10 Décomposition en valeurs singulières et applications

Soient $m, n \in \mathbb{N} - \{0\}$, et soient E et F des espaces tous deux euclidiens ou tous deux hermitiens,[47] de dimension n et m respectivement.

L'un des but de cette partie est de donner des résultats de décomposition de nature spectrale des matrices rectangulaires, pas forcément carrées. C'est en particulier important lorsque nous étudions un grand nombre de points $(X_j)_{1 \leq j \leq n}$ dans $\mathbb{R}^m$ ou $\mathbb{C}^m$, auxquels nous pouvons associer la matrice dans $\mathscr{M}_{m,n}(\mathbb{R})$ ou $\mathscr{M}_{m,n}(\mathbb{C})$ dont les n colonnes sont les coordonnées de $X_1, \dots, X_n$ dans la base canonique de $\mathbb{R}^m$ ou $\mathbb{C}^m$. Ces points peuvent représenter par exemple m valeurs de gris ou de couleurs associés à n pixels dans une image numérique (et le nombre de pixels est bien plus grand que le nombre de couleurs). Ils peuvent représenter des valeurs d'un grand nombre de variables aléatoires (indépendantes, identiquement distribuées) à valeurs dans un espace euclidien donné, tirées au hasard suivant une loi donnée. La décomposition en valeurs singulières (dont l'acronyme anglais est **SVD**) est ainsi utile dans des domaines aussi variés que le traitement du signal, la reconnaissance d'image, l'analyse sémantique, la météorologie, la robotique, l'analyse des grosses données (et particulièrement en analyse en composantes indépendantes).[48]

Nous ne traiterons pas des aspects numériques de la décomposition en valeurs singulières de $A \in \mathscr{M}_{m,n}(\mathbb{R})$ dans ce cours, pour lesquel nous renvoyons à [GL, Chap. 8], mais ils sont intéressants pour l'agrégation. La démonstration que nous donnons n'est pas algorithmique, ni d'utilisation pratique, car elle commence par une diagonalisation complète de la matrice symétrique ${}^t\!A\,A \in \mathscr{M}_{n,n}(\mathbb{R})$, problème notoirement difficile du point de vue calculatoire, surtout lorsque n est grand. L'algorithme le plus couramment utilisé, disons lorsque $m \geq n$ quitte à passer à la matrice transposée, est celui de Golub-Kahan (voir loc. cit.), qui commence par utiliser la méthode de bidiagonalisation de Householder (voir loc. cit., algorithme 5.4.2) pour transformer la matrice A en une matrice par blocs ${}^t\!U\,A\,V = \begin{pmatrix} B \\ 0 \end{pmatrix}$ où $B \in \mathscr{M}_{nn}(\mathbb{R})$ est bidiagonale supérieure, puis calcule la décomposition en valeurs singulières de B, par une succession de rotations de Givens (voir loc. cit., algorithme 8.6.1). Des méthodes de calcul encore plus efficaces (algorithme de Lanczos, voir par exemple https ://en.wikipedia.org/wiki/Lanczos_algorithm) existent lorsque l'on ne souhaite calculer qu'une partie de la décomposition en valeurs singulières. Les décompositions en valeurs singulières sont aussi utiles pour d'autres calculs (voir par exemple

47. Seul le cas euclidien est au programme de l'agrégation, mais le traitement du cas hermitien est presque le même.

48. Et un raton laveur ! Inventaire. Jacques Prévert.

https ://en.wikipedia.org/wiki/Low-rank_approximation), comme des problèmes de minimisation de normes sur les matrices (norme spectrale ou norme de Fröbenius). [49]

Nous commençons par les définitions et rappels utiles dans cette partie 1.10. Soit $\mathbb{K}$ un corps commutatif.

Une matrice rectangulaire [50] $A = (a_{i,j})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} \in \mathscr{M}_{m,n}(\mathbb{K})$ est dite *diagonale* si pour tous les $i \in \{1,\dots,m\}$ et $j \in \{1,\dots,n\}$ tels que $i \neq j$, nous avons $a_{i,j} = 0$. Ce sont donc les matrices de la forme

$$\begin{pmatrix} a_1 & & & 0 \\ & a_2 & & \\ & & \ddots & \\ 0 & & & a_n \\ 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix} \quad \text{si } n \leq m \text{ (donc s'il y a plus de lignes que de colonnes)}$$

$$\text{ou} \quad \begin{pmatrix} a_1 & & & 0 & 0 & \cdots & 0 \\ & a_2 & & & \vdots & & \vdots \\ & & \ddots & & \vdots & & \vdots \\ 0 & & & a_m & 0 & \cdots & 0 \end{pmatrix} \quad \text{si } n \geq m\,, \quad \begin{array}{l}\text{(donc s'il y a plus de colonnes}\\ \text{que de lignes)}\end{array}$$

avec $a_i \in \mathbb{K}$ pour $1 \leq i \leq \min\{m,n\}$. Notons que ${}^tA \neq A$ si $m \neq n$, même pour une matrice diagonale !

Exercice E.10. *Montrer que si $A \in \mathscr{M}_{m,n}(\mathbb{K})$ et $B \in \mathscr{M}_{n,p}(\mathbb{K})$ sont deux matrices diagonales multipliables, alors leur produit $AB \in \mathscr{M}_{m,p}(\mathbb{K})$ est aussi une matrice diagonale.*

Soit $u \in \mathscr{L}(E,F)$ une application linéaire, et notons r son rang (qui vérifie en particulier $r \leq \min\{m,n\}$). Rappelons que les opérateurs linéaires $u^* \circ u \in \mathscr{L}(E)$ et $u \circ u^* \in \mathscr{L}(F)$ sont auto-adjoints positifs (voir la remarque (5) de la partie 1.4), et que leurs valeurs propres sont positives ou nulles (voir la proposition 1.8 (vi)). De plus, par la proposition 1.8 (iii), les applications linéaires u, u^*, $u^* \circ u$ et $u \circ u^*$ ont le même rang r. Comme $u^* \circ u$ et $u \circ u^*$ sont diagonalisables en base orthonormée (voir les corollaires 1.17 (1) dans le cas euclidien et 1.15 (1) dans le cas hermitien), le nombre de leurs valeurs propres non nulles (donc strictement positives) est exactement égal à leur rang r.

Définition 1.20. *Soit $u \in \mathscr{L}(E,F)$ de rang n. Les* valeurs singulières $\mu_1, \mu_2, \dots, \mu_r$ *de u sont les racines carrées des valeurs propres non nulles de $u^* \circ u$ (avec multiplicité).*

Les valeurs singulières sont uniquement définies à l'ordre près. Nous verrons dans le théorème 1.22 que ce sont aussi les racines carrées des valeurs propres non nulles de $u \circ u^*$. Elles sont souvent ordonnées de manière décroissante, ce qui assure en particulier leur unicité.

Si $A \in \mathscr{M}_{m,n}(\mathbb{R})$ (respectivement $A \in \mathscr{M}_{m,n}(\mathbb{C})$) est une matrice rectangulaire réelle (respectivement complexe), les *valeurs singulières* de A sont les racines carrées des valeurs

49. Je remercie Christina et Mattis Paulin pour leur tutorat !
50. Attention à l'ordre des indices m (nombre de lignes) et n (nombre de colonnes).

propres non nulles de ${}^tA\,A \in \mathscr{M}_{n,n}(\mathbb{R})$ (respectivement $A^*\,A \in \mathscr{M}_{n,n}(\mathbb{C})$). Nous verrons que ce sont aussi celles de $A\,{}^tA \in \mathscr{M}_{m,m}(\mathbb{R})$ (respectivement $A\,A^* \in \mathscr{M}_{m,m}(\mathbb{C})$). Ce sont donc les valeurs singulières de l'endomorphisme de l'espace euclidien usuel $\mathbb{R}^n$ (respectivement hermitien usuel $\mathbb{C}^n$) dont la matrice dans la base canonique est A.

Le résultat principal de cette partie 1.10 utilisera le lemme suivant, sur lequel nous reviendrons au moment de l'utilisation de l'application exponentielle (voir la proposition 3.7 (3)).

Lemme 1.21. *Si $u' \in \mathscr{L}(E)$ est auto-adjoint positif, alors il existe un unique élément $v \in \mathscr{L}(E)$ auto-adjoint positif tel que $u' = v^2$.*

Démonstration. Par le théorème de diagonalisation en base orthonormée de l'endomorphisme auto-adjoint positif u' (voir les corollaires 1.17 dans le cas euclidien et 1.15 dans le cas hermitien), il existe une base orthonormée $\mathscr{B}$ de E dans laquelle la matrice M de u' est diagonale à coefficients diagonaux positifs ou nuls.

Pour montrer l'existence, il suffit de poser $v \in \mathscr{L}(E)$ l'application linéaire dont la matrice dans $\mathscr{B}$ est diagonale à coefficients diagonaux les racines carrées de ceux de M.

Pour montrer l'unicité, si v est comme dans l'énoncé, alors une diagonalisation en base orthonormée de l'endomorphisme auto-adjoint positif v montre que λ est valeur propre de v si et seulement si λ^2 est valeur propre de u', et qu'alors $\lambda \geq 0$ et $E_\lambda(v) = E_{\lambda^2}(u')$. Donc v est uniquement déterminée par le fait qu'elle vaut l'homothétie de rapport $\sqrt{\mu}$ sur l'espace propre $E_\mu(u')$ de u', pour toute valeur propre μ de u'. □

Théorème 1.22. (Décomposition en valeurs singulières) *Soient $m, n \in \mathbb{N} - \{0\}$, et E, F des espaces tous deux euclidiens ou tous deux hermitiens, de dimension n et m respectivement, et soit $u \in \mathscr{L}(E, F)$. Les valeurs singulières de u et de u^* sont égales. Il existe des bases orthonormées $\mathscr{B}$ de E et $\mathscr{C}$ de F, formées de vecteurs propres de $u^* \circ u$ et $u \circ u^*$ respectivement, telles que la matrice M_u de u dans les bases $\mathscr{B}$ et $\mathscr{C}$, et la matrice M_{u^*} de u^* dans les bases $\mathscr{C}$ et $\mathscr{B}$, soient des matrices rectangulaires diagonales (transposées l'une de l'autre : $M_{u^*} = {}^tM_u$), de coefficients diagonaux non nuls les valeurs singulières de u.*

Remarques. (1) Il n'y a en général pas unicité de telles bases. Si l'on ordonne les valeurs singulières de u de manière décroissante, et si on demande que les r premiers coefficients diagonaux de la matrice diagonale M_u sont ceux non nuls et sont ainsi ordonnés, alors la matrice diagonale M_u est unique.

(2) **Interprétation matricielle.** En prenant l'endomorphisme u entre les espaces euclidiens usuels $\mathbb{R}^n$ et $\mathbb{R}^m$ (respectivement entre les espaces hermitiens usuels $\mathbb{C}^n$ et $\mathbb{C}^m$), dont la matrice dans les bases canoniques est la matrice A donnée, et en utilisant le fait que la matrice de passage entre deux bases orthonormées est orthogonale (respectivement unitaire), nous avons une formulation matricielle équivalente du théorème précédent.

Corollaire 1.23. *Pour tout $A \in \mathscr{M}_{m,n}(\mathbb{R})$ (respectivement $A \in \mathscr{M}_{m,n}(\mathbb{C})$), il existe une matrice $D \in \mathscr{M}_{m,n}(\mathbb{R})$ diagonale, dont les coefficients diagonaux non nuls sont les valeurs singulières de A, et des matrices $P \in \mathrm{O}(n)$ et $P' \in \mathrm{O}(m)$ (respectivement $P \in \mathrm{U}(n)$ et $P' \in \mathrm{U}(m)$) telles que*

$$A = P'\,D\,{}^tP = P'\,D\,P^{-1} \quad \text{(respectivement } A = P'\,D\,P^* = P'\,D\,P^{-1}\text{)}\,.$$

Si les coefficients diagonaux de D sont ordonnés de manière décroissante, alors la matrice $D = D(A)$ est unique, et $D({}^tA) = {}^tD(A)$ (respectivement $D(A^*) = D(A)^*$). □

Par contre, les matrices P et P' ne sont pas forcément uniques. Tout tel triplet (P', D, P) est appelé une *décomposition en valeurs singulières* de A (et souvent noté (U, D, V) dans la littérature).

(3) Si $m = n$ et en restriction aux matrices inversibles, cette décomposition matricielle s'appelle la *décomposition de Cartan* de $G = \mathrm{GL}_n(\mathbb{R})$ (respectivement $G = \mathrm{GL}_n(\mathbb{C})$). Si $K = \mathrm{O}(n)$ (respectivement $K = \mathrm{U}(n)$) et si A^+ est l'ensemble des matrices dans $\mathrm{GL}_n(\mathbb{R})$ diagonales à coefficients diagonaux strictement positifs et décroissants, la décomposition de Cartan

$$G = K\, A\, K$$

dit que pour tout élément $g \in G$, il existe des éléments $a \in A^+$ et $k_1, k_2 \in K$ tels que $g = k_1\, a\, k_2$ (avec a unique, mais k_1, k_2 pas forcément uniques). Nous renvoyons à [Bor] pour des généralisations. Par exemple, si $n = 2$, une conséquence de la décomposition de Cartan dit que pour tout $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R})$, il existe un unique $\lambda \geq 1$ et des $\theta_1, \theta_2 \in \mathbb{R}/2\pi\mathbb{Z}$ tels que

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & \frac{1}{\lambda} \end{pmatrix} \begin{pmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{pmatrix} .$$

La décomposition de Cartan de $\mathrm{GL}_n(\mathbb{R})$ et $\mathrm{GL}_n(\mathbb{C})$ est un raffinement de la décomposition polaire décrite dans la partie 3.5, et peut aussi s'en déduire.

Démonstration. Nous renvoyons par exemple à [FGN, page 123, Exer. 2.24].

Par le lemme 1.21, soit $v \in \mathscr{L}(E)$ l'unique endomorphisme linéaire auto-adjoint positif de E tel que $u^* \circ u = v^2$. Soit $\mathscr{B} = (e_1, \ldots, e_n)$ une base orthonormée de E dans laquelle v (et donc $u^* \circ u$) est diagonale. À permutation près, nous pouvons supposer que $v(e_i) = \mu_i\, e_i$ pour $1 \leq i \leq r$, et que $\ker v = \mathrm{Vect}(e_{r+1}, \ldots, e_n)$, en notant r le rang de u et $\mu_1, \mu_2, \ldots, \mu_r$ les valeurs singulières de u. Rappelons que celles-ci sont strictement positives, donc non nulles.

Posons $f_i = \frac{1}{\mu_i}\, u(e_i)$ pour $1 \leq i \leq r$. La suite $(f_1, \ldots, f_r)$ dans F est orthonormée, car si $1 \leq i, j \leq r$, alors

$$\langle u(e_i), u(e_j) \rangle = \langle u^* \circ u(e_i), e_j \rangle = \langle v^2(e_i), e_j \rangle = \mu_i^2\, \langle e_i, e_j \rangle .$$

Complétons la suite orthonormée $(f_1, \ldots, f_r)$ en une base orthonormée $\mathscr{C} = (f_1, \ldots, f_m)$ de F. Montrons que les bases $\mathscr{B}$ et $\mathscr{C}$ conviennent.

Nous avons $u(e_i) = \mu_i\, f_i$ pour $1 \leq i \leq r$ par construction. Par la proposition 1.8 (ii) et les propriétés des noyaux des endomorphismes diagonalisables, nous avons

$$\ker u = \ker(u^* \circ u) = \ker(v^2) = \ker(v) = \mathrm{Vect}(e_{r+1}, \ldots, e_n) .$$

Donc $u(e_i) = 0$ si $r + 1 \leq i \leq n$.

Nous avons

$$u^*(f_i) = u^*\big(\frac{1}{\mu_i}\, u(e_i)\big) = \frac{1}{\mu_i} v^2(e_i) = \mu_i\, e_i$$

pour $1 \leq i \leq r$. Par la proposition 1.8 (i), nous avons

$$\ker(u^*) = (\mathrm{im}\, u)^\perp = \big(\mathrm{Vect}(f_1, \ldots, f_r)\big)^\perp = \mathrm{Vect}(f_{r+1}, \ldots, f_m) .$$

Donc $u^*(f_i) = 0$ si $r+1 \leq i \leq m$. En particulier, $u \circ u^*(f_i) = \mu_i\, u(e_i) = \mu_i^2\, f_i$ si $1 \leq i \leq r$ et $u \circ u^*(f_i) = 0$ si $r+1 \leq i \leq m$.

Ceci montre d'une part que $\mathscr{B}$ est formée de vecteurs propres de $u^* \circ u$, que $\mathscr{C}$ est formée de vecteurs propres de $u \circ u^*$, que les matrices de u et de u^* sont diagonales dans ces bases (et transposées l'une de l'autre), et d'autre part que les valeurs singulières de u et de u^* coïncident. □

Exercice E.11. *Soient $m, n \in \mathbb{N} - \{0\}$ et $A \in \mathscr{M}_{m,n}(\mathbb{R})$ une matrice de valeurs singulières décroissantes $\sigma_1, \ldots, \sigma_r$. Notons (U, Σ, V) une décomposition en valeurs singulières de A, de sorte que $A = U\,\Sigma\,{}^tV$ avec $U \in \mathrm{O}(m)$, $V \in \mathrm{O}(n)$, Σ diagonale, à coefficients diagonaux décroissants et dont les r premiers sont les valeurs singulières de A ordonnées. Pour tout $k \in \{1, \ldots, r-1\}$, notons $\Sigma_k = \begin{pmatrix} \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_k \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix} \in \mathscr{M}_{m,n}(\mathbb{R})$ la matrice diagonale dont les k premiers coefficients diagonaux sont les k premières valeurs singulières $\sigma_1, \ldots, \sigma_k$ de A dans cet ordre, et les autres coefficients diagonaux sont nuls.*

Montrer que $A_k = U\,\Sigma_k\,{}^tV$ est la matrice de rang k qui minimise la distance (pour la norme d'opérateurs)[51] *à A des matrices de rang k, et que $d(A_k, A) = \sigma_{k+1}$.*

Exercice E.12. *Soient $m, n \in \mathbb{N} - \{0\}$ et $k = \min\{m, n\}$.*

(1) Montrer qu'il existe un algorithme qui prend une matrice $A \in \mathscr{M}_{m,n}(\mathbb{R})$ et, en la multipliant à droite et à gauche par des matrices de réflexions de Householder (voir l'exercice E.40), la transforme en une matrice de la forme $\begin{pmatrix} B & 0 \end{pmatrix}$ ou $\begin{pmatrix} B \\ 0 \end{pmatrix}$ avec $B = (b_{i,j})_{1 \leq i,j \leq k} \in \mathscr{M}_k(\mathbb{R})$ une matrice bidiagonale supérieure (c'est-à-dire triangulaire supérieure et vérifiant $b_{i,j} = 0$ si $j > i+1$). Montrer que les valeurs singulières de A et de B sont égales.

(2) Si $C = \begin{pmatrix} 0 & B \\ {}^tB & 0 \end{pmatrix} \in \mathscr{M}_{2k}(\mathbb{R})$, montrer que les valeurs propres strictement positives de C sont les valeurs singulières de B, et qu'il existe un algorithme transformant toute telle matrice C en une matrice $T = (t_{i,j})_{1 \leq i,j \leq k} \in \mathscr{M}_{2k}(\mathbb{R})$ tridiagonale de diagonale nulle (c'est-à-dire vérifiant $t_{i,i} = 0$ et $t_{i,j} = 0$ si $|i-j| > 1$), ayant les mêmes valeurs propres.

(3) Donner un algorithme donnant les valeurs propres d'une matrice tridiagonale de diagonale nulle.

1.11 Produit mixte et produit vectoriel

Soient $n \in \mathbb{N}$ et E un espace vectoriel réel de dimension finie n.

Rappelons que l'espace vectoriel réel $\Lambda^n E^*$ des formes n-linéaires alternées[52] sur E est de dimension 1, engendré, pour toute base $\mathscr{B}$ de E, par l'application déterminant $\det_{\mathscr{B}}$

51. La *norme d'opérateur* sur $\mathscr{M}_{m,n}(\mathbb{R})$ (identifié de manière usuelle à $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$) – aussi appelée la *norme matricielle subordonnée* aux normes euclidiennes usuelles standards de $\mathbb{R}^m$ et $\mathbb{R}^n$ – d'un élément $A' \in \mathscr{M}_{m,n}(\mathbb{R})$ est définie par $\|A'\| = \max_{x \in \mathbb{R}^n\,:\,\|x\|=1} \|A'x\|$.

52. Si K est un corps commutatif et si $E_1, \ldots, E_{n+1}$ sont $n+1$ espaces vectoriels sur K, une application $f : E_1 \times \cdots \times E_n \to E_{n+1}$ est dite *multilinéaire* si elle est linéaire en chacune des n variables, et *alternée* si elle s'annule sur les n-uplets d'éléments dont deux au moins sont égaux. C'est une *forme* multilinéaire alternée si de plus $E_{n+1} = K$.

dans la base $\mathscr{B}$ des n-uplets de vecteurs de E : pour tous les $x_1, \dots, x_j, \dots, x_n$ dans E, de coordonnées $(x_{1,1}, \dots, x_{n,1}), \dots, (x_{i,j})_{1\le i\le n}, \dots, (x_{1,n}, \dots, x_{n,n})$ respectivement dans la base $\mathscr{B}$, nous avons

$$\det{}_{\mathscr{B}}(x_1, \dots, x_n) = \det\left((x_{i,j})_{1\le i,j\le n}\right) .$$

De plus, soit $P = P_{\mathscr{B},\mathscr{B}'}$ la matrice de passage de la base $\mathscr{B}$ à une autre base $\mathscr{B}'$ (dont les colonnes sont les vecteurs colonnes des coordonnées des éléments de $\mathscr{B}'$ dans la base $\mathscr{B}$). Alors

$$\det P = \det P_{\mathscr{B},\mathscr{B}'} = \det{}_{\mathscr{B}}\mathscr{B}'$$

et $X = P\,X'$ où X et X' sont les vecteurs colonnes des coordonnées d'un vecteur $x \in E$ dans les bases $\mathscr{B}$ et $\mathscr{B}'$ respectivement. Donc [53]

$$\det{}_{\mathscr{B}} = (\det{}_{\mathscr{B}}\mathscr{B}')\ \det{}_{\mathscr{B}'} . \tag{22}$$

Si $\mathscr{B}$ est la base canonique de $\mathbb{R}^n$, nous notons $\det{}_{\mathscr{B}} = \det : (\mathbb{R}^n)^n \to \mathbb{R}$. Remarquons que si $x_1, \dots, x_n$ sont n vecteurs linéairement indépendants de $\mathbb{R}^n$, alors le volume pour la mesure de Lebesgue de $\mathbb{R}^n$ du parallélépipède

$$P(x_1, \dots, x_n) = \Big\{ \sum_{i=1}^{n} \lambda_i x_i :\ \forall\, i = 1, \dots, n, \quad \lambda_i \in [0,1] \Big\}$$

défini par $x_1, \dots, x_n$ vérifie (voir par exemple la formule (25))

$$\mathrm{vol}(P(x_1, \dots, x_n)) = |\det{}_{\mathscr{B}}(x_1, \dots, x_n)| .$$

Rappelons qu'une *orientation* de E est une classe d'équivalence de bases de E pour la relation $\mathscr{B} \sim \mathscr{B}'$ si et seulement si le déterminant de la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$ est (strictement) positif, qu'il y a deux orientations possibles, et que E est *orienté* s'il est muni d'une orientation. Une base de E est alors dite *directe* (ou parfois *positive*) si elle appartient à l'orientation choisie.

Exemples. (1) L'*espace euclidien orienté usuel* $\mathbb{R}^n$ est l'espace euclidien usuel $\mathbb{R}^n$ muni de l'orientation rendant directe la base canonique de $\mathbb{R}^n$.

(2) Si E est un espace euclidien ou hermitien, si $\mathscr{B}'$ est la base obtenue par le procédé d'orthonormalisation de Gram-Schmidt à partir d'une base $\mathscr{B}$, alors $\mathscr{B}$ et $\mathscr{B}'$ définissent la même orientation. En effet, si $f : [0,1] \to E^n$ est un chemin continu de bases entre $\mathscr{B}$ et $\mathscr{B}'$ (voir le commentaire qui suit l'énoncé de la proposition 1.11), alors l'application $t \mapsto \det_{\mathscr{B}}(f(t))$, qui vaut 1 en $t = 0$ et qui ne s'annule pas, est, par le théorème des valeurs intermédiaires, positive en $f(1) = \det_{\mathscr{B}}(\mathscr{B}')$.

(3) Tout espace vectoriel complexe de dimension finie E, considéré comme un espace vectoriel réel, est muni d'une orientation canonique, de la manière suivante.

53. En effet, pour tous les $x_1, \dots, x_n$ dans E, dont les colonnes des coordonnées sont $X_1, \dots, X_n$ dans $\mathscr{B}$ et $X'_1, \dots, X'_n$ dans $\mathscr{B}'$, nous avons, en écrivant les matrices comme juxtapositions de leurs colonnes,

$$\begin{aligned}\det{}_{\mathscr{B}}(x_1, \dots, x_n) &= \det(X_1 \dots X_n) = \det(PX'_1 \dots PX'_n) = (\det P)\det(X'_1 \dots X'_n)\\ &= (\det{}_{\mathscr{B}}\,\mathscr{B}')\det{}_{\mathscr{B}'}(x_1, \dots, x_n) .\end{aligned}$$

Notons $E_{\mathbb{R}}$ l'espace vectoriel réel, dit obtenu de E par *restriction des scalaires* à $\mathbb{R}$, qui est le groupe additif sous-jacent à E muni de la restriction à $\mathbb{R} \times E$ de la multiplication externe de E. Si $\mathscr{B} = (e_1, \dots, e_n)$ est une base de E (en tant qu'espace vectoriel complexe), alors $\mathscr{B}_{\mathbb{R}} = (e_1, \dots, e_n, i\,e_1, \dots, i\,e_n,)$ est une base de $E_{\mathbb{R}}$ (en tant qu'espace vectoriel réel). L'orientation de $E_{\mathbb{R}}$ définie par $\mathscr{B}_{\mathbb{R}}$ est indépendante du choix de la base $\mathscr{B}$. En effet, soit $\mathscr{B}'$ une autre base, soit $P \in \mathscr{M}_n(\mathbb{C})$ la matrice de passage de la base $\mathscr{B}$ à la base $\mathscr{B}'$, et soient $A \in \mathscr{M}_n(\mathbb{R})$ et $B \in \mathscr{M}_n(\mathbb{R})$ les parties réelles et imaginaires de P (de sorte que $P = A + iB$). Alors $P_{\mathbb{R}} = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}$ est la matrice de passage de $\mathscr{B}_{\mathbb{R}}$ à $\mathscr{B}'_{\mathbb{R}}$, et son déterminant, égal à

$$\det(P_{\mathbb{R}}) = |\det P|^2$$

par le lemme suivant, est strictement positif, ce qui montre le résultat.

Lemme 1.24. *Pour tous les $A, B \in \mathscr{M}_n(\mathbb{R})$, nous avons*

$$\det \begin{pmatrix} A & -B \\ B & A \end{pmatrix} = \left|\det(A + iB)\right|^2 .$$

Démonstration. En multipliant par i les n dernières lignes puis les n dernières colonnes, en ajoutant la première ligne par blocs à la dernière ligne par blocs, puis la dernière colonne par blocs à la première colonne par blocs, et en utilisant la formule des déterminants des matrices triangulaires supérieures par blocs, [54] nous avons

$$\begin{aligned}
\begin{vmatrix} A & -B \\ B & A \end{vmatrix} &= i^{-n} \begin{vmatrix} A & -B \\ iB & iA \end{vmatrix} = (-1)^n \begin{vmatrix} A & -iB \\ iB & -A \end{vmatrix} = (-1)^n \begin{vmatrix} A & -iB \\ A+iB & -A-iB \end{vmatrix} \\
&= (-1)^n \begin{vmatrix} A & -iB \\ A+iB & -A-iB \end{vmatrix} = (-1)^n \begin{vmatrix} A-iB & -iB \\ 0 & -A-iB \end{vmatrix} \\
&= (-1)^n \det(A-iB)\det(-A-iB) = \det(\,\overline{A+iB}\,)\ \det(A+iB) .
\end{aligned}$$

Le résultat en découle. □

Jusqu'à la fin de la partie 1.11, nous supposons que E est un espace euclidien orienté de dimension n. Il découle des rappels ci-dessus et du fait que la matrice de passage P d'une base orthonormée à une autre base orthonormée est une matrice orthogonale, donc de déterminant ± 1, que si $\mathscr{B}$ et $\mathscr{B}'$ sont deux bases orthonormées directes de E, si P est la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$, alors $P \in \mathrm{SO}(n)$ et $\det_{\mathscr{B}} = \det_{\mathscr{B}'}$: l'application

54. Celle-ci dit que pour tous les $n, m \in \mathbb{N} - \{0\}$, pour tout corps commutatif k' et pour tous les $A \in \mathscr{M}_{m,m}(k')$, $B \in \mathscr{M}_{m,n}(k')$ et $D \in \mathscr{M}_{n,n}(k')$, nous avons

$$\det \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \det(A)\det(D) .$$

Pour démontrer cette formule, nous pouvons supposer que A est inversible (sinon les deux termes sont nuls, celui de gauche car son rang, vu comme le rang de l'espace vectoriel engendré par les colonnes, est strictement inférieur à $n+m$) ; nous écrivons

$$\begin{vmatrix} A & B \\ 0 & D \end{vmatrix} = \begin{vmatrix} A & 0 \\ 0 & I_n \end{vmatrix} \begin{vmatrix} I_m & A^{-1}B \\ 0 & D \end{vmatrix} ;$$

et nous calculons les deux derniers déterminants en développant respectivement par rapport aux n dernières lignes dans l'ordre décroissant et par rapport aux m premières colonnes dans l'ordre croissant.

$\det_{\mathscr{B}} : E^n \to \mathbb{R}$ dans une base orthonormée directe $\mathscr{B}$ ne dépend pas de son choix. Elle est appelée la *forme volume* de E, et notée ω_E. Si E^{op} est l'espace vectoriel euclidien E muni de l'orientation opposée, alors

$$\omega_{E^{\mathrm{op}}} = -\,\omega_E \ .$$

Le choix indifférent d'une base orthonormée directe $\mathscr{B}$ de espace euclidien orienté E de dimension n permet ainsi d'associer à tout n-uplet $(x_1, \dots, x_n)$ de vecteurs de E un scalaire réel

$$x_1 \wedge \cdots \wedge x_n = \omega_E(x_1, \dots, x_n) = \det{}_{\mathscr{B}}(x_1, \dots, x_n) \in \mathbb{R}$$

appelé le *produit mixte* de $(x_1, \dots, x_n)$. La notation $x_1 \wedge \cdots \wedge x_n$ pour ce nombre réel n'est pas standard (voir l'exercice E.20, avec certes un danger de confusion de notation, qui donne une explication profonde à la notation). Il ne faut pas confondre le scalaire réel associé à un n-uplet avec le vecteur associé à un $(n-1)$-uplet, voir ci-dessous. Ce produit mixte dépend de manière n-linéaire alternée de $(x_1, \dots, x_n)$ et vérifie les propriétés élémentaires suivantes (qui sont immédiates).

Proposition 1.25. *Soit $(x_1, \dots, x_n) \in E^n$.*

(1) Le produit mixte $\omega_E(x_1, \dots, x_n)$ est non nul si et seulement si $(x_1, \dots, x_n)$ est une base de E.

(2) Le produit mixte $\omega_E(x_1, \dots, x_n)$ est strictement positif si et seulement si $(x_1, \dots, x_n)$ est une base directe de E.

(3) Si $(x_1, \dots, x_n)$ est une base orthonormée de E, alors son produit mixte $\omega_E(x_1, \dots, x_n)$ vaut $+1$ si cette base est directe, et -1 sinon. □

Remarque. Par l'exercice E.39, pour tout $(x_1, \dots, x_n) \in E^n$, nous avons

$$|\ \omega_E(x_1, \dots, x_n)\ | \leq \prod_{i=1}^{n} \|x_i\| \ ,$$

avec égalité si et seulement si les vecteurs $x_1, \dots, x_n$ sont deux à deux orthogonaux. En particulier, dans un espace euclidien, le volume d'un parallélépipède de longueurs de côtés données est maximal lorsque ce parallélépipède est à angles droits.

Dans la proposition suivante, nous associons, à tout $(n-1)$-uplet $(x_1, \dots, x_{n-1})$ de vecteurs d'un espace euclidien orienté E de dimension n, un vecteur $x_1 \wedge \cdots \wedge x_{n-1}$ de E.

Proposition 1.26. *Pour tout $(n-1)$-uplet $(x_1, \dots, x_{n-1})$ de vecteurs de E, il existe un et un seul vecteur*

$$x_1 \wedge \cdots \wedge x_{n-1} \in E$$

tel que

$$\forall\, y \in E, \quad \langle x_1 \wedge \cdots \wedge x_{n-1}, y \rangle = \omega_E(x_1, \dots, x_{n-1}, y) \ . \tag{23}$$

De plus, les propriétés suivantes sont vérifiées :

(i) l'application $(x_1, \dots, x_{n-1}) \mapsto x_1 \wedge \cdots \wedge x_{n-1}$ de E^{n-1} dans E est $(n-1)$-linéaire alternée,

(ii) $x_1 \wedge \cdots \wedge x_{n-1} = 0$ si et seulement si $x_1, \dots, x_{n-1}$ sont linéairement dépendants,

(iii) $x_1 \wedge \cdots \wedge x_{n-1} \in \mathrm{Vect}(x_1, \dots, x_{n-1})^{\perp}$,

(iv) *si* $(e_1, \dots, e_n)$ *est une base orthonormée directe, alors*

$$e_1 \wedge \cdots \wedge e_{n-1} = e_n \ . \tag{24}$$

(v) *si* $x_1, \dots, x_{n-1}$ *sont linéairement indépendants, alors* $(x_1, \dots, x_{n-1}, x_1 \wedge \cdots \wedge x_{n-1})$ *est une base directe de* E.

(vi) *si* $\mathscr{B}$ *est une base orthonormée directe de* E, *si* $A = (x_{i,j})_{1 \le i \le n,\, 1 \le j \le n-1}$ *est la matrice* $n \times (n-1)$ *dont les colonnes sont les vecteurs colonnes des coordonnées de* $x_1, \dots, x_{n-1}$ *dans la base* $\mathscr{B}$, *alors la* i*-ème coordonnée* $(x_1 \wedge \cdots \wedge x_{n-1})_i$ *de* $x_1 \wedge \cdots \wedge x_{n-1}$ *dans la base* $\mathscr{B}$ *est* $(-1)^{n-i}$ *fois le déterminant de la matrice* A_i *obtenue en enlevant la* i*-ème ligne de* A *:*

$$(x_1 \wedge \cdots \wedge x_{n-1})_i = (-1)^{n-i} \begin{vmatrix} x_{1,1} & \cdots\cdots & x_{1,n-1} \\ \vdots & \vdots & \vdots \\ x_{i-1,1} & \cdots\cdots & x_{i-1,n-1} \\ x_{i+1,1} & \cdots\cdots & x_{i+1,n-1} \\ \vdots & \vdots & \vdots \\ x_{n,1} & \cdots\cdots & x_{n,n-1} \end{vmatrix} .$$

Le vecteur $x_1 \wedge \cdots \wedge x_{n-1}$ est appelée le *produit vectoriel* de $x_1, \dots, x_{n-1}$. Il dépend non seulement de la structure euclidienne de E, mais aussi de l'orientation de E : changer l'orientation change son signe. La propriété (24) du produit vectoriel s'appelle en dimension 3 la *règle des trois doigts de la main droite*. Les propriétés (i), (iii) et (iv) caractérisent le produit vectoriel.

Démonstration. Puisque l'application $y \mapsto \omega_E(x_1, \dots, x_{n-1}, y)$ est une forme linéaire sur E par multilinéarité du déterminant, l'existence et l'unicité de $x_1 \wedge \cdots \wedge x_{n-1}$ vient de l'isomorphisme de dualité $E \to E^*$ défini par $x \mapsto \{y \mapsto \langle x, y \rangle\}$.

Les propriétés (i), (ii), (iii), (iv) et (v) découlent de la définition et du fait que le déterminant est multilinéaire alterné.

Montrons la propriété (vi). Par développement d'un déterminant par rapport à sa dernière colonne, pour tous les $y \in E$ de coordonnées $(y_1, \dots, y_n)$ dans la base $\mathscr{B}$, en notant $(z_1, \cdots, z_n)$ les coordonnées de $z = x_1 \wedge \cdots \wedge x_{n-1}$, nous avons

$$\sum_{i=1}^{n} y_i \, z_i = \langle z, y \rangle = \omega_E(x_1, \dots, x_{n-1}, y) = \begin{vmatrix} x_{1,1} & \cdots\cdots & x_{1,n-1} & y_1 \\ \vdots & & \vdots & \vdots \\ x_{n,1} & \cdots\cdots & x_{n,n-1} & y_n \end{vmatrix}$$
$$= \sum_{i=1}^{n} (-1)^{n-i} \, y_i \ \det A_i \ .$$

Ceci étant vrai pour tous les $y_1, \dots, y_n$ dans $\mathbb{R}$, le résultat en découle. □

Exercice E.13. *Dans l'espace euclidien orienté usuel* $\mathbb{R}^3$, *considérons des vecteurs* a, b *et* c.

(1) *Si* $a = (x, y, z)$ *et* $b = (x', y', z')$, *montrer que*

$$a \wedge b = \left(\begin{vmatrix} y & y' \\ z & z' \end{vmatrix}, - \begin{vmatrix} x & x' \\ z & z' \end{vmatrix}, \begin{vmatrix} x & x' \\ y & y' \end{vmatrix} \right) .$$

(2) *Montrer que*

$$a \wedge (b \wedge c) = \langle a, c\rangle\, b - \langle a, b\rangle\, c\ ,$$

$$(a \wedge b) \wedge (a \wedge c) = \det(a, b, c)\ a\ ,$$

$$\det(a \wedge b, a \wedge c, b \wedge c) = (\det(a, b, c))^2\ .$$

(3) *Montrer que* $\|a \wedge b\| = \|a\|\ \|b\|\ |\sin \angle(a, b)|$ *est l'aire du parallélogramme de côtés* a *et* b*, et que* $\|a \wedge b\|^2 + \langle a, b\rangle^2 = \|a\|^2\ \|b\|^2$.

Exercice E.14. *Soient* E *un espace euclidien ou hermitien,* $n \in \mathbb{N} - \{0\}$ *et* $x_1, \dots, x_n$ *des éléments de* E*. On appelle* matrice de Gram *de* $(x_1, \dots, x_n)$ *la matrice de produits scalaires* $(\langle x_i, x_j\rangle)_{1\le i,j\le n}$*, et on note*

$$\mathrm{Gram}(x_1, \dots, x_n) = \det\big((\langle x_i, x_j\rangle)_{1\le i,j\le n}\big)$$

son déterminant, appelé le déterminant de Gram *de* $(x_1, \dots, x_n)$.

(1) *Montrer que si* E *est de dimension* n *et si* $\mathscr{B}$ *est une base orthonormée de* E*, alors*

$$\mathrm{Gram}(x_1, \dots, x_n) = (\det{}_{\mathscr{B}}(x_1, \dots, x_n))^2\ .$$

Montrer que $\mathrm{Gram}(x_1, \dots, x_n) \ge 0$ *avec égalité si et seulement si les vecteurs* $x_1, \dots, x_n$ *sont linéairement dépendants. En calculant le déterminant de Gram de deux vecteurs, en déduire l'inégalité de Cauchy-Schwarz, avec cas d'égalité, dans tout espace préhilbertien.*

(2) *Soient* $x_1, \dots, x_n \in E$ *et* $r \in \mathbb{N}$*. Montrer que les conditions suivantes sont équivalentes :*

(i) le rang du système $\{x_1, \dots, x_n\}$ *est inférieur ou égal à* r*,*

(ii) pour toute partie $\{x_{i_1}, \dots, x_{i_{r+1}}\}$ *de* $\{x_1, \dots, x_n\}$ *de cardinal* $r + 1$*, nous avons* $\mathrm{Gram}(x_{i_1}, \dots, x_{i_{r+1}}) = 0$*,*

(iii) le rang de la matrice de Gram de $(x_1, \dots, x_n)$ *est inférieur ou égal à* r*.*

(3) *Si* E *est de dimension finie, et si* $\mathscr{B}$ *et* $\mathscr{B}'$ *sont deux bases de* E*, quelle est la relation entre les matrices de Gram de* $\mathscr{B}$ *et de* $\mathscr{B}'$ *?*

(4) *Soit* $(x_1, \dots, x_n)$ *une base d'un sous-espace vectoriel* F *de* E*. Montrer que pour tout* $x \in E$*,*

$$d(x, F)^2 = \frac{\mathrm{Gram}(x, x_1, \dots, x_n)}{\mathrm{Gram}(x_1, \dots, x_n)}$$

En déduire par exemple la valeur de

$$A = \min_{a_1, \dots, a_n \in \mathbb{R}} \int_0^1 (1 + a_1 x + a_2 x^2 + \cdots + a_n x^n)^2\, dx\ .$$

(5) *Montrer que si* E *est l'espace vectoriel euclidien usuel* $\mathbb{R}^n$ *et* $x_1, \dots, x_n \in \mathbb{R}^n$*, alors*

$$\mathrm{Gram}(x_1, \dots, x_n) = \mathrm{vol}(P(x_1, \dots, x_n))^2$$

où $P(x_1, \dots, x_n)$ *est le parallélépipède*

$$P(x_1, \dots, x_n) = \Big\{ \sum_{i=1}^n t_i\, x_i\ :\ \forall\ i = 1, \dots, n,\ \ t_i \in [0, 1] \Big\}\ .$$

(6) *Montrer que si E est un espace euclidien orienté de dimension n, alors pour tous les $x_1, \dots, x_{n-1} \in E$, nous avons*

$$\|x_1 \wedge \cdots \wedge x_{n-1}\|^2 = \mathrm{Gram}(x_1, \dots, x_{n-1}) \, .$$

(7) *Pour tout $n \in \mathbb{N} - \{0\}$, le* déterminant de Cayley-Menger $\Gamma_n \in \mathbb{Z}[X_{i,j} \ : \ 0 \leq i < j \leq n]$ *est le polynôme à coefficients entier, en $\frac{n(n+1)}{2}$ indéterminées $X_{i,j}$, défini, en posant $X_{i,i} = 0$ et $X_{-1,i} = 1$ pour $0 \leq i \leq n$ et $X_{i,j} = X_{j,i}$ pour $-1 \leq i, j \leq n$, par*

$$\Gamma_n = \det\left((X_{i,j}^2)_{-1 \leq i,j \leq n}\right) \, .$$

(i) Montrer que Γ_n est un polynôme symétrique en les indéterminées $X_{i,j}$ pour $0 \leq i < j \leq n$, et que

$$\Gamma_n = \begin{vmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & X_{0,1}^2 & \dots & X_{0,n}^2 \\ 1 & X_{0,1}^2 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & X_{n-1,n}^2 \\ 1 & X_{0,n}^2 & \dots & X_{n-1,n}^2 & 0 \end{vmatrix}$$

$$= (-1)^{n+1} \begin{vmatrix} 2X_{0,1}^2 & X_{0,1}^2 + X_{0,2}^2 - X_{1,2}^2 & \dots & X_{0,1}^2 + X_{0,n}^2 - X_{1,n}^2 \\ X_{0,1}^2 + X_{0,2}^2 - X_{1,2}^2 & 2X_{0,2}^2 & \dots & X_{0,2}^2 + X_{0,n}^2 - X_{2,n}^2 \\ \vdots & \vdots & \ddots & \vdots \\ X_{0,1}^2 + X_{0,n}^2 - X_{1,n}^2 & X_{0,2}^2 + X_{0,n}^2 - X_{2,n}^2 & \dots & 2X_{0,n}^2 \end{vmatrix} .$$

(ii) Dans la suite de cette question (7), notons E' un espace affine euclidien orienté, $x_0, \dots, x_n$ des points de E', et $d_{i,j} = d(x_i, x_j)$ la distance entre x_i et x_j. Montrer que

$$\mathrm{Gram}(\overrightarrow{x_0x_1}, \dots, \overrightarrow{x_0x_n}) = \frac{(-1)^{n+1}}{2^n} \, \Gamma_n(d_{i,j} : 0 \leq i < j \leq n) \, .$$

(iii) Montrer que $x_0, \dots, x_n$ appartiennent à un même sous-espace affine de E' de dimension $n-1$ si et seulement si $\Gamma_n(d_{i,j} \ : \ 0 \leq i < j \leq n) = 0$. En déduire que trois points x, y, z d'un espace affine euclidien sont alignés si et seulement si $d(y,z) = |d(x,y) - d(x,z)|$ ou $d(y,z) = d(x,y) + d(x,z)$.

(iv) Si E' est de dimension n, montrer que le volume V du simplexe[55] de sommets $x_0, \dots, x_n$ vérifie

$$V^2 = \frac{(-1)^{n+1}}{2^n (n!)^2} \, \Gamma_n(d_{i,j} \ : \ 0 \leq i < j \leq n) \, .$$

(v) Soit ABC un triangle d'un plan affine euclidien, de longueurs des côtés a, b, c, et d'aire $\mathscr{A}$. Montrer la formule de Héron

$$\mathscr{A} = \frac{1}{4} \sqrt{(a+b+c)(b+c-a)(a+c-b)(a+b-c)} \, .$$

(8) *Considérons un tétraèdre T de $\mathbb{R}^3$ ayant un sommet A telle que la longueur des trois arêtes ayant A pour sommet soit 1, et les angles en A des trois faces ayant A pour sommet soient $\frac{\pi}{p}, \frac{\pi}{q}, \frac{\pi}{r}$, où $p, q, r \in \mathbb{N} - \{0, 1\}$ vérifient $\frac{1}{p} + \frac{\pi}{q} + \frac{\pi}{r} > 1$. Calculer le volume de T.*

55. c'est-à-dire de leur enveloppe convexe $\{\sum_{i=0}^n t_i x_i \ : \ t_i \geq 0, \ \sum_{i=0}^n t_i = 1\}$

1.12 Exercices complémentaires

Exercice E.15. Soient E un espace hermitien de dimension finie, et $u \in \mathscr{L}(E)$. Montrer que u est trigonalisable en base orthonormée.

Exercice E.16. Soient $n \in \mathbb{N}$ et $\mathbb{C}_n[X]$ l'espace vectoriel des polynômes complexes de degré au plus n.

(1) Montrer que

$$\langle P, Q\rangle = \frac{1}{2\pi}\int_0^{2\pi} P(e^{it})\,\overline{Q(e^{it})}\,dt$$

est un produit scalaire hermitien sur $\mathbb{C}_n[X]$, dont $(1, X, X^2, \dots, X^n)$ est une base orthonormée.

(2) Soient $k \in \{1, \dots, n\}$ et $a_1, \dots, a_k \in \mathbb{C}$ deux à deux distincts. Montrer qu'il existe une base orthonormée $(P_0, P_1, \dots, P_n)$ de $\mathbb{C}_n[X]$ telle que

$$P_i(a_1) = P_i(a_2) = \cdots = P_i(a_k) = 0$$

pour tout $i = k, \dots, n$.

Exercice E.17. Soient E un espace euclidien ou hermitien et $u \in \mathscr{L}(E)$. Montrer que u est normal si et seulement si $\|u(x)\| = \|u^*(x)\|$ pour tous les $x \in E$.

Exercice E.18.

(1) Montrer que l'application $(A, B) \mapsto \mathtt{tr}(B^*\,A)$ est un produit scalaire sur l'espace vectoriel complexe $\mathscr{M}_n(\mathbb{C})$, appelé produit scalaire de Hilbert-Schmidt sur $\mathscr{M}_n(\mathbb{C})$, invariant par l'action du groupe unitaire $\mathrm{U}(n)$ par multiplication à gauche. Notons $\|\cdot\|_{HS}$ la norme associée à ce produit scalaire.

(2) Calculer la norme $\|A\|_{HS}$ d'une matrice $A = (a_{i,j})_{1\leq i,j\leq n}$ en fonction de ses coefficients.

(3) Montrer que pour toute base orthonormée $(e_1, \dots, e_n)$ de l'espace hermitien standard $\mathbb{C}^n$, en identifiant de manière usuelle $\mathscr{M}_n(\mathbb{C})$ et $\mathscr{L}(\mathbb{C}^n)$, nous avons

$$\|A\|_{HS}^2 = \sum_{i=1}^{n} \|A\,e_i\|^2\ .$$

(4) Montrer que la base canonique $(E_{i,j})_{1\leq i,j\leq n}$ de $\mathscr{M}_n(\mathbb{C})$, formée des *matrices indicatrices* (où la matrice $E_{i,j}$ est à coefficients tous nuls sauf celui d'indice (i, j) égal à 1), est une base orthonormée de $\mathscr{M}_n(\mathbb{C})$ pour le produit scalaire de Hilbert-Schmidt.

(5) Calculer la norme de Hilbert-Schmidt d'une matrice unitaire.

(6) Montrer que, pour toute matrice symétrique $A = (a_{i,j})_{1\leq i,j\leq n}$ dans $\mathscr{M}_n(\mathbb{R})$, de valeurs propres $\lambda_1, \dots, \lambda_n$, nous avons

$$\sum_{i=1}^{n} \lambda_i^2 = \sum_{i,j=1}^{n} a_{i,j}^2\ .$$

Exercice E.19. Le but de cet exercice est la détermination de la droite de régression linéaire, et la détermination de la meilleure approximation polynomiale de degré donné, pour un nuage de points du plan réel euclidien.

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace de Hilbert (réel ou complexe), de norme associée $\|\cdot\|$.

(1) Soit $p \in \mathscr{L}(E)$. Montrer que les propriétés suivantes sont équivalentes

(i) p est un opérateur linéaire auto-adjoint *idempotent* (c'est-à-dire tel que $p \circ p = p$),

(ii) p est idempotent, et $\ker p = (\operatorname{im} p)^{\perp}$.

L'opérateur linéaire p est appelé un *projecteur orthogonal* (sur le sous-espace vectoriel $\operatorname{im} p = \ker(\mathrm{id} - p)$, parallèlement à $\ker p = (\operatorname{im} p)^{\perp}$). Montrer que $\operatorname{im} p = \ker(\mathrm{id} - p)$, que $\operatorname{im} p$ est fermé et que p est *positif* (c'est-à-dire $\langle p(x), x\rangle \geq 0$ pour tout $x \in E$).

(2) Soit p le projecteur orthogonal sur un sous-espace vectoriel fermé F de E. Montrer que, pour tout x dans E,

$$\|p(x) - x\| = \inf_{y \in F} \|y - x\| = \min_{y \in F} \|y - x\| = \inf_{x' \in E} \|p(x') - x\| = \min_{x' \in E} \|p(x') - x\| \, .$$

(3) Soit $n \in \mathbb{N} - \{0\}$. Considérons les points $A_i = (x_i, y_i)$ pour $i = 1, ..., n$ dans $\mathbb{R}^2$, et cherchons une droite de $\mathbb{R}^2$ d'équation $y = ax + b$ qui soit « la plus proche possible » des points A_i. Il s'agit de trouver des réels a et b minimisant la quantité

$$\sum_{i=1}^{n} (y_i - (ax_i + b))^2 \, .$$

Considérons les vecteurs $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ et $\underline{1} = (1, 1, \ldots, 1)$ de $\mathbb{R}^n$ et f l'application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^n$ définie par $(a, b) \mapsto a\, x + b\, \underline{1}$.

(a) Montrer que le problème consiste à déterminer (a, b) tel que la norme $\|f(a, b) - y\|$ soit minimale.

(b) Montrer que le couple (a, b) recherché vérifie

$$\begin{cases} a \langle x, x\rangle + b \langle \underline{1}, x\rangle = \langle x, y\rangle \\ a \langle \underline{1}, x\rangle + b \langle \underline{1}, \underline{1}\rangle = \langle \underline{1}, y\rangle \, . \end{cases}$$

(c) Montrer que si $x_1, \ldots, x_n$ ne sont pas tous égaux, alors

$$a = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

et

$$b = \frac{(\sum_{i=1}^n x_i^2)(\sum_{i=1}^n y_i) - (\sum_{i=1}^n x_i)(\sum_{i=1}^n x_i y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \, .$$

(4) Soient $p, n \in \mathbb{N} - \{0\}$, $A \in \mathscr{M}_{n,p}(\mathbb{R})$ une matrice réelle $n \times p$ de rang p, et $B \in \mathbb{R}^n$. Nous identifions de manière usuelle un vecteur de $\mathbb{R}^k$ avec le vecteur colonne de ses coordonnées dans la base canonique, et nous munissons $\mathbb{R}^k$ de sa norme euclidienne usuelle, pour tout $k \in \mathbb{N} - \{0\}$.

(a) Montrer que $p \leq n$, qu'il existe un et un seul élément $X \in \mathbb{R}^p$ tel que $\|AX - B\|$ soit minimum, et que c'est l'unique solution du système linéaire ${}^tAAX = {}^tAB$.

(b) Étant donné n points $(x_i, y_i) \in \mathbb{R}^2$ pour $1 \leq i \leq n$, déterminer un polynôme réel P de degré strictement inférieur à p qui minimise la quantité $\sum_{i=1}^n (P(x_i) - y_i)^2$. Si $p = 2$, montrer que le graphe de P dans $\mathbb{R}^2$ passe par le barycentre de points (x_i, y_i) pour $1 \leq i \leq n$.

Exercice E.20. A **Rappels sur la dualité linéaire.** Soient k un corps commutatif et E un espace vectoriel de dimension finie sur k. Soient $\mathscr{B} = (e_1, \dots, e_n)$ une base de E et $\mathscr{B}^* = (e_1^*, \dots, e_n^*)$ sa *base duale*[56] dans le *dual* $E^* = \mathscr{L}(E, k)$ de E.

(1) Montrer que les coordonnées d'une forme linéaire $\ell \in E^*$ dans la base $\mathscr{B}^*$ sont $(\ell(e_1), \dots, \ell(e_n))$.

(2) Montrer que si M est la matrice dans la base $\mathscr{B}$ d'un endomorphisme $u \in \mathscr{L}(E)$, alors la matrice $\widetilde{M}$ dans la base duale $\mathscr{B}^*$ de l'*endomorphisme dual* ${}^t u \in \mathscr{L}(E^*)$ de précomposition par u, c'est-à-dire défini par $\ell \mapsto {}^t u(\ell) = \ell \circ u$, est égale à ${}^t M$.

(3) Montrer que si P est la matrice de passage de $\mathscr{B}$ à une autre base $\mathscr{C}$ de E, alors la matrice de passage de la base $\mathscr{B}^*$ à la base duale $\mathscr{C}^*$ de $\mathscr{C}$ est $\breve{P} = ({}^t P)^{-1}$.

(4) Nous identifions E et son bidual $E^{**} = (E^*)^*$ par l'isomorphisme linéaire canonique en dimension finie $x \mapsto \{\mathrm{ev}_x : \ell \mapsto \ell(x)\}$, qui à un élément x de E associe la forme linéaire ev_x sur E^* d'évaluation en x des formes linéaires sur E. Pour toute base $\mathscr{B}$ de E, montrer que la base duale de la base duale de $\mathscr{B}$ est $\mathscr{B}$:

$$(\mathscr{B}^*)^* = \mathscr{B} \, .$$

(5) Dans cette question, $n \in \mathbb{N} - \{0\}$ et $k_n[X]$ est l'espace vectoriel des polynômes à coefficients dans k en une indéterminée X de degré au plus n. Considérons $n+1$ éléments deux à deux distincts $a_0, a_1, \dots, a_n$ dans k. Montrer que les formes linéaires $\ell_0 : P \mapsto P(a_0), \dots, \ell_n : P \mapsto P(a_n)$ forment une base $\mathscr{B}^*$ de l'espace vectoriel dual $k_n[X]^*$. Expliciter une base $\mathscr{B} = (P_0, \dots, P_n)$ de E dont la base duale est $\mathscr{B}^*$, et calculer les coordonnées dans cette base d'un polynôme $P \in k_n[X]$.

Pour toute fonction $f : k \to k$, le polynôme $P_f = \sum_{i=1}^n f(a_i) P_i$ est appelé le *polynôme d'interpolation de Lagrange* de f sur $a_0, a_1, \dots, a_n$.

B **Dualité de Hodge euclidienne.** Soit E un espace vectoriel euclidien orienté[57] de dimension $n \in \mathbb{N} - \{0\}$. Pour tout $k \in \{1, \dots, n\}$, notons $\Lambda^k E$ l'espace vectoriel réel des formes k-linéaires alternées sur E^*, de sorte que $\Lambda^1 E = (E^*)^* = E$. Pour tout k-uplet $(x_1, \dots, x_k)$ d'éléments de E, notons

$$x_1 \wedge \cdots \wedge x_k \in \Lambda^k E \, ,$$

appelé un *k-vecteur* de E, l'élément de $\Lambda^k E$ défini par

$$\forall (\ell_1, \dots, \ell_k) \in (E^*)^k, \quad x_1 \wedge \cdots \wedge x_k(\ell_1, \dots, \ell_k) = \det\big((\ell_i(x_j))_{1 \le i,j \le k} \big) \, .$$

(1) Soient $\mathscr{B} = (e_1, \dots, e_n)$ une base de E et $\mathscr{B}^* = (e_1^*, \dots, e_n^*)$ sa base duale dans E^*.

i) Montrer que

$$e_1 \wedge \cdots \wedge e_n = \det{}_{\mathscr{B}^*}$$

et que l'application de $\mathbb{R}$ dans $\Lambda^n E$ définie par $\lambda \mapsto \lambda \, e_1 \wedge \cdots \wedge e_n$ est un isomorphisme linéaire.

56. c'est-à-dire l'unique base de E^* telle que, pour tous les $i, j = 1, \dots, n$, nous ayons $e_i^*(e_j) = \delta_{i,j}$ avec $\delta_{i,i} = 1$ et $\delta_{i,j} = 0$ si $i \neq j$.

57. Le choix de l'orientation est crucial dans cet exercice.

ii) Montrer que $(e_{i_1} \wedge \cdots \wedge e_{i_k})_{1 \leq i_1 < \cdots < i_k \leq n}$ est une base de l'espace vectoriel $\Lambda^k E$, qui est donc de dimension égale au coefficient binomial $\binom{n}{k}$.

(2) Pour tous les entiers $k \in \{0, \ldots, n\}$ et $1 \leq i_1 < \cdots < i_k \leq n$, calculer la signature $\varepsilon_{\{i_1,\ldots,i_k\}} \in \{-1, +1\}$ de la permutation $(i_1, \ldots, i_k, 1, 2, \ldots, \widehat{i_1}, \ldots, \widehat{i_k}, \ldots, n)$ de l'ensemble $\{1, \ldots, n\}$ (avec la convention usuelle que les termes avec un chapeau n'apparaissent pas).

(3) Posons $\Lambda^0 E = \mathbb{R}$. Notons $\mathbf{1} \in \Lambda^0 E$ l'élément 1 de $\mathbb{R}$, et

$$\Lambda E = \bigoplus_{k=0}^{n} \Lambda^k E \, .$$

Montrer qu'il existe un unique opérateur linéaire $* \in \mathscr{L}(\Lambda E)$, appelé la dualité de Hodge, tel que pour toute base orthonormée directe $\mathscr{B} = (e_1, \ldots, e_n)$ de E nous avons

$$*\mathbf{1} = e_1 \wedge \cdots \wedge e_n, \quad *(e_1 \wedge \cdots \wedge e_n) = \mathbf{1} \, ,$$

et si $k \in \{1, \ldots n-1\}$, alors

$$*(e_1 \wedge \cdots \wedge e_k) = e_{k+1} \wedge \cdots \wedge e_n \, .$$

(4) Montrer que $** = (-1)^{k(n-k)}\,\mathrm{id}$ sur $\Lambda^k E$, et en particulier que la dualité de Hodge est un automorphisme linéaire de ΛE, envoyant $\Lambda^k E$ sur $\Lambda^{n-k} E$ pour tout $k \in \{0, \ldots n\}$.

(5) Remarquer que si $(x_1, \ldots, x_{n-1})$ est un $(n-1)$-uplet d'éléments de E, alors

$$*^{-1}(x_1 \wedge \cdots \wedge x_{n-1}) \, ,$$

qui est un élément de $\Lambda^1 E = E$, est exactement le produit vectoriel du $(n-1)$-uplet $(x_1, \ldots, x_{n-1})$. Remarquer que si $(x_1, \ldots, x_n)$ est un n-uplet d'éléments de E, alors

$$*^{-1}(x_1 \wedge \cdots \wedge x_n) \, ,$$

qui est un élément de $\Lambda^0 E = \mathbb{R}$, est exactement le produit mixte du n-uplet $(x_1, \ldots, x_n)$.

1.13 Indications pour la résolution des exercices

Correction de l'exercice E.1. Notons n la dimension de E et A la matrice de f dans une base de E, qui est antisymétrique par l'interprétation matricielle. Alors

$$\det(A) = \det({}^tA) = \det(-A) = (-1)^n \det(A) \; .$$

Comme A est inversible (donc de déterminant non nul) car f est non dégénérée, l'entier n est pair.

Correction de l'exercice E.2. L'assertion (3) est en fait un cas particulier de l'assertion (1). En effet, en identifiant de manière usuelle $\mathbb{C}$ et $\mathbb{R}^2$, pour tout $n \in \mathbb{N}$, la sphère unité

$$\mathbb{S}_{2n-1} = \big\{(z_1, \ldots, z_n) \in \mathbb{C}^n \; : \; \sum_{i=1}^{n} |z_i|^2 = 1\big\} \; .$$

de l'espace hermitien standard $\mathbb{C}^n$ s'identifie avec la sphère unité de l'espace euclidien standard $\mathbb{R}^{2n}$, en utilisant l'application

$$(x_1, \ldots, x_{2n}) \mapsto (z_1 = x_1 + i\, x_2, \ldots, z_n = x_{n-1} + i\, x_{2n}) \; .$$

De plus, le cône isotrope de la forme quadratique hermitienne

$$-|z_1|^2 - \cdots - |z_p|^2 + |z_{p+1}|^2 + \cdots + |z_{p+q}|^2$$

s'identifie avec le cône isotrope de la forme quadratique

$$-x_1^2 - \cdots - x_{2p}^2 + x_{2p+1}^2 + \cdots + x_{2p+2q}^2 \; .$$

(1) Munissons aussi $\mathbb{R}^{p+q}$ de sa structure euclidienne usuelle, dont la sphère unité est

$$\mathbb{S}_{p+q-1} = \big\{(x_1, \ldots, x_{p+q}) \in \mathbb{R}^{p+q} \; : \; \sum_{i=1}^{p} x_i^2 + \sum_{i=p+1}^{p+q} x_i^2 = 1\big\} \; .$$

Notons que l'application $(t, x) \mapsto tx$ de $]0, +\infty[\times \mathbb{S}_{p+q-1}$ dans $\mathbb{R}^{p+q} - \{0\}$ est un homéomorphisme. Puisqu'un cône isotrope est un cône, il suffit de montrer que l'intersection du cône isotrope et de la sphère unité est homéomorphe à $\mathbb{S}_{p-1} \times \mathbb{S}_{q-1}$. Le résultat découle du fait que si a, b sont deux réels tels que $-a + b = 0$ et $a + b = 1$, alors $a = b = \frac{1}{2}$. L'application de $\mathbb{S}_{p-1} \times \mathbb{S}_{q-1} \times \,]0, +\infty[$ dans le cône isotrope étudié, définie par

$$((x_1, \cdots, x_p), (y_1, \cdots, y_q), t) \mapsto (\, t\, x_1, \cdots, t\, x_p, t\, y_1, \cdots, t\, y_q \,) \; ,$$

bijective d'inverse

$$\begin{aligned}(x'_1, \cdots, x'_p, x'_{p+1}, \cdots, x'_{p+q}) \mapsto \Big(& \big(\tfrac{\sqrt{2}\, x'_1}{\sqrt{(x'_1)^2 + \cdots + (x'_{p+q})^2}}, \cdots, \tfrac{\sqrt{2}\, x'_p}{\sqrt{(x'_1)^2 + \cdots + (x'_{p+q})^2}}\big), \\ & \big(\tfrac{\sqrt{2}\, x'_{p+1}}{\sqrt{(x'_1)^2 + \cdots + (x'_{p+q})^2}}, \cdots, \tfrac{\sqrt{2}\, x'_{p+q}}{\sqrt{(x'_1)^2 + \cdots + (x'_{p+q})^2}}\big), \\ & \tfrac{\sqrt{(x'_1)^2 + \cdots + (x'_{p+q})^2}}{\sqrt{2}} \Big) \; ,\end{aligned}$$

est un homéomorphisme cherché.

(2) L'application

$$(x_1,\dots,x_{p+q}) \mapsto \Big(\big(\frac{1}{\sqrt{1+\sum_{i=p+1}^{p+q} x_i^2}}\, x_1,\dots, \frac{1}{\sqrt{1+\sum_{i=p+1}^{p+q} x_i^2}}\, x_p \big), (x_{p+1},\dots,x_{p+q}) \Big)$$

est un homéomorphisme cherché.

Correction de l'exercice E.3. Pour tout polynome $P \in k[X]$, nous avons $(P^\sigma)^\sigma = P$ et $(P(u))^* = P^\sigma(u^*)$ par la semi-linéarité de l'application $u \mapsto u^*$ et le fait que $(u^n)^* = (u^*)^n$ pour tout $n \in \mathbb{N}$. Donc $P(u)$ est nul si et seulement si $P^\sigma(u^*)$ est nul, ce qui montre (avec le fait que $\sigma(1) = 1$) que le polynôme minimal unitaire de u^* est $\pi_{u^*} = (\pi_u)^\sigma$.

Soient M et M^* les matrices de u et u^* dans une même base, et A la matrice de f dans cette base. Par la formule (12), pour tout $\lambda \in k$, nous avons

$$\begin{aligned} \det(M^* - \lambda \,\mathrm{id}) &= \det((M - \lambda^\sigma \,\mathrm{id})^*) = \det((A^\sigma)^{-1}\, {}^t(M - \lambda^\sigma \,\mathrm{id})^\sigma A^\sigma) \\ &= \det({}^t(M - \lambda^\sigma \,\mathrm{id})^\sigma) = (\det(M - \lambda^\sigma \,\mathrm{id}))^\sigma = (\chi_u(\lambda^\sigma))^\sigma = (\chi_u)^\sigma(\lambda)\ . \end{aligned}$$

Donc le polynôme caractéristique de u^* est $\chi_{u^*} = (\chi_u)^\sigma$.

Correction de l'exercice E.4. En caractéristique différente de 2, le même calcul qu'avec des nombres réels pour SO(2) montre que

$$\mathrm{SO}(f) = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} : a, b \in \mathbb{Z}/p\mathbb{Z},\ \ a^2 + b^2 = 1 \right\}.$$

Si $p = 3, 5$, nous avons donc

$$\mathrm{SO}(f) = \left\{ \pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \pm \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\}.$$

Si $p = 7$, nous avons donc

$$\mathrm{SO}(f) = \left\{ \pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \pm \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \pm \begin{pmatrix} 2 & -2 \\ 2 & 2 \end{pmatrix}, \pm \begin{pmatrix} 2 & 2 \\ -2 & 2 \end{pmatrix} \right\}.$$

Correction de l'exercice E.5. (1) Si f est nulle, alors E est somme directe de droites isotropes. Sinon, il existe des vecteurs e_1 et e_2 de E tels que $f(e_1, e_2) \neq 0$. Quitte à remplacer e_2 par un multiple scalaire, nous pouvons supposer que $f(e_1, e_2) = 1$. Alors $P = \mathrm{Vect}(e_1, e_2)$ est un plan symplectique, et $\mathscr{B} = (e_1, e_2)$ est appelée une *base symplectique* de P. La matrice dans cette base de la restriction de f à P est $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Puisque P est non isotrope, nous avons $E = P \overset{\perp}{\oplus} P^\perp$. Par récurence sur la dimension, $P^\perp$ est somme orthogonale de droites isotropes et de plans symplectiques. Donc E l'est.

Si f est non dégénérée, alors E est somme directe de plans hyperboliques $P_1, \dots, P_n$. Donc la dimension de E est $2n$. De plus, la matrice de f dans une base concaténation de bases symplectiques des plans $P_1, \dots, P_n$ est diagonale par blocs 2-2, d'éléments diagonaux $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. L'unicité à équivalence près d'une forme bilinéaire alternée non dégénérée de rang $2n$ en découle.

(2) Par le résultat de structure précédent, si f est non dégénérée, alorstous les sous-espaces vectoriels totalement isotropes maximaux ont la même codimension, égale à la moitié du rang de f. En général, leur dimension est donc égale à la somme de la dimension du noyau de f et de la moitié du rang de f.

(3) Soit $\mathscr{B} = (e_1, e_2)$ une base symplectique de E, de sorte que la matrice de f dans cette base est $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Puisque

$$ {}^t\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & ad-bc \\ bc-ad & 0 \end{pmatrix} , $$

l'application qui à un automorphisme de E associe sa matrice dans la base $\mathscr{B}$ est un isomorphisme de groupes de $\mathrm{Sp}(f)$ dans $\mathrm{SL}_2(k)$.

(4) Comme vu dans la question (1), puisque f est non dégénérée, la dimension de E est paire, notée $2n$, et pour $i = 1, \dots, n$, il existe des bases symplectiques (e_i, e_i') de plans symplectiques P_i dont E est une somme directe orthogonale. Alors la matrice de f dans la base $\mathscr{B} = (e_1, \dots, e_n, e_1', \dots, e_n')$ est égale à J_n. Le fait que l'application donnée de $\mathrm{Sp}(f)$ dans $\mathrm{Sp}_n(k) = \{X \in \mathrm{GL}_{2n}(k) : {}^t X\, J_n\, X = J_n\}$ soit un isomorphisme de groupes découle de la formule (14).

(5) Une matrice $X \in \mathrm{O}(2n)$, donc telle que ${}^t X = X^{-1}$, appartient à $\mathrm{Sp}_n(\mathbb{R})$ si et seulement X commute avec J_n, donc, par un petit calcul par blocs, si et seulement s'il existe $A, B \in \mathscr{M}_n(\mathbb{R})$ tels que $X = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}$.

Par la décomposition polaire (28) pour les matrices réelles, pour tout $X \in \mathrm{Sp}_n(\mathbb{R})$, il existe une matrice $P \in \mathrm{O}(2n) \cap \mathrm{Sp}_n(\mathbb{R})$ et une matrice symétrique définie positive Q telle que $X = PQ$. Alors $\det P = 1$ par l'assertion précédente sur la forme des éléments de $\mathrm{O}(2n) \cap \mathrm{Sp}_n(\mathbb{R})$ et le lemme 1.24. De plus, $\det Q > 0$. Puisque $\det X = \pm 1$, nous avons donc $\det X = 1$.

Correction de l'exercice E.6. Voir la correction du problème de CAPES externe 2007 de mathématiques.

La formule (18) est vraie sur l'hyperplan vectoriel $x^\perp$ et sur la droite vectorielle engendrée par x, donc est vraie partout, par linéarité et somme directe $E = k\,x \oplus x^\perp$ (car x est non isotrope).

Pour tout $u \in \mathrm{O}(q)$, l'automorphisme linéaire $u \circ s_x \circ u^{-1}$ appartient à $\mathrm{O}(q)$, et il est involutif. Il fixe $u(x^\perp) = (u(x))^\perp$ et il vaut $-\mathrm{id}$ sur la droite engendrée par $u(x)$, donc $u \circ s_x \circ u^{-1} = s_{u(x)}$.

(1) Pour tout $\alpha \in R$, nous avons $-\alpha = s_\alpha(\alpha) \in R$. Nous avons

$$ n(\alpha,\beta)n(\beta,\alpha) = 4\,\frac{\langle \alpha, \beta\rangle^2}{\|\alpha\|^2\,\|\beta\|^2} . $$

et le résultat découle de la formule donnant $\cos(\theta_{\alpha,\beta})$. Puisque $4\cos^2(\theta_{\alpha,\beta})$ est un entier entre 0 et 4, il ne peut prendre que les valeurs $0, 1, 2, 3$ ou 4. Nous avons donc $\cos(\theta_{\alpha,\beta}) = 0, \pm\frac{1}{2}, \pm\frac{1}{\sqrt{2}}, \pm\frac{\sqrt{3}}{2}, \pm 1$, correspondant seulement à des angles $0, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \frac{3\pi}{4}, \frac{5\pi}{6}, \pi$.

Si $n(\alpha,\beta) = \pm 4$, alors $n(\beta,\alpha) = \pm 1$ et $\theta_{\alpha,\beta} = 0$ ou $\theta_{\alpha,\beta} = \pi$, donc $\beta = \pm\alpha$ puisque le système de racines est réduit, ce qui contredit le fait que $n(\alpha, \pm\alpha) = \pm 2$.

Si $\theta_{\alpha,\beta} \neq \frac{\pi}{2}$, c'est-à-dire si α et β ne sont pas orthogonaux, alors la définition donne immédiatement que $\frac{n(\beta,\alpha)}{n(\alpha,\beta)}$ existe et vaut $\frac{\|\alpha\|^2}{\|\beta\|^2}$.

Remarquons que $n(\alpha,\beta)$ et $n(\beta,\alpha)$ sont de même signe et simultanément nuls ou non nuls, et que $n(\alpha,\beta) \geq n(\beta,\alpha)$ si $\|\alpha\| \leq \|\beta\|$. Le tableau des valeurs possibles est donc le suivant.

$n(\alpha,\beta)$	$n(\beta,\alpha)$	$\theta_{\alpha,\beta}$	$\frac{\|\alpha\|}{\|\beta\|}$
-3	-1	$\frac{5\pi}{6}$	$\frac{1}{\sqrt{3}}$
-2	-2	π	1
-2	-1	$\frac{3\pi}{4}$	$\frac{1}{\sqrt{2}}$
-1	-1	$\frac{2\pi}{3}$	1
0	0	$\frac{\pi}{2}$	$t \in]0,1]$
1	1	$\frac{\pi}{3}$	1
2	1	$\frac{\pi}{4}$	$\frac{1}{\sqrt{2}}$
2	2	0	1
3	1	$\frac{\pi}{6}$	$\frac{1}{\sqrt{3}}$

(2) Nous pouvons transformer R par une rotation pour que la racine α soit de coordonnées $(t,0)$ avec $t>0$, puis par une homothétie pour que α soit de coordonnées $(1,0)$, puis, si la deuxième coordonnée de β est strictement négative, par la symétrie orthogonale par rapport à l'axe des premières coordonnées. Ces transformations préservent les constantes de structure. Nous avons $n(\alpha,\beta) \neq 0$ car la racine β n'est pas orthogonale à α. De plus, comme $s_\alpha(\alpha) = -\alpha$, nous avons

$$n(\alpha, s_\alpha(\beta)) = \frac{\langle \alpha, s_\alpha(\beta)\rangle}{\langle \alpha,\alpha\rangle} = -\frac{\langle s_\alpha(\alpha), s_\alpha(\beta)\rangle}{\langle \alpha,\alpha\rangle} = n(\alpha,\beta)\ .$$

(3) Supposons tout d'abord que R ne contienne pas de couple d'éléments non colinéaires non orthogonaux. Soient α,β des éléments non colinéaires, qui existent par la première hypothèse. Quitte à transformer R par une rotation, par la symétrie orthogonale par rapport à l'axe des premières coordonnées et par un automorphisme linéaire d'expression en coordonnées $(x,y) \mapsto (t\,x, t'\,y)$ où $t,t'>0$, qui préserve les constantes $n(\alpha,\beta)$ puisqu'elles sont nulles si α et β ne sont pas colinéaires, nous pouvons supposer que α et β sont de coordonnées $(1,0)$ et $(0,1)$. Donc R est de type $A_1 \times A_1$ comme sur le dessin ci-dessous.

Sinon, nous prenons α,β comme dans la question (2). Quitte à remplacer β par $s_\alpha(\beta)$, nous pouvons supposer que $n(\alpha,\beta) < 0$.

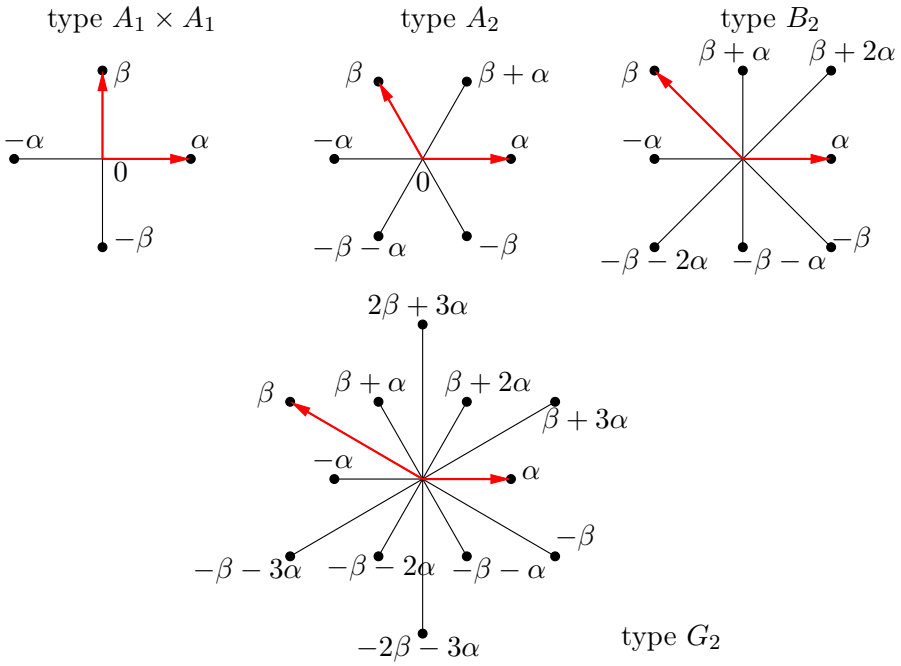
D'après le tableau de la question (1), trois cas peuvent se présenter.

Cas 1. Supposons que $\|\beta\| = \sqrt{2}\ \|\alpha\|$ et $\theta_{\alpha,\beta} = \frac{3\pi}{4}$. En calculant $s_\alpha(\beta) = \beta + 2\,\alpha$ et $s_\beta(\alpha) = \beta + \alpha$, et en utilisant l'invariance par l'antipodie $v \mapsto -v$, nous obtenons que R contient le système de racines de type B_2 du dessin ci-dessous. S'il existait une autre racine γ dans R qui n'est pas dans B_2, nous obtiendrions un angle entre γ et l'une des racines de B_2 qui n'est pas dans la liste possible. Donc $R = B_2$.

Cas 2. Supposons que $\|\beta\| = \sqrt{3}\ \|\alpha\|$ et $\theta_{\alpha,\beta} = \frac{5\pi}{6}$. En calculant $s_\alpha(\beta) = \beta + 3\,\alpha$, $s_\beta(\alpha) = \beta + \alpha$, $s_\beta \circ s_\alpha(\beta) = 2\,\beta + 3\,\alpha$ et $s_\alpha \circ s_\beta(\alpha) = \beta + 2\,\alpha$, et en utilisant l'invariance par l'antipodie $v \mapsto -v$, nous obtenons que R contient le système de racines de type G_2 du dessin ci-dessous. S'il existait une autre racine γ dans R qui n'est pas dans G_2, nous

obtiendrions un angle entre γ et l'une des racines de G_2 qui n'est pas dans la liste possible. Donc $R = G_2$.

Cas 3. Supposons que $\|\beta\| = \|\alpha\|$ et $\theta_{\alpha,\beta} = \frac{2\pi}{3}$. En calculant $s_\alpha(\beta) = \beta+\alpha$, nous obtenons que R contient le système de racines de type A_2 du dessin ci-dessous, qui convient. S'il existe une racine γ dans R qui n'est pas dans A_2, l'angle entre γ et les deux vecteurs de A_2 de part et d'autre de γ est $\frac{\pi}{6}$. Il en découle que $R = G_2$.

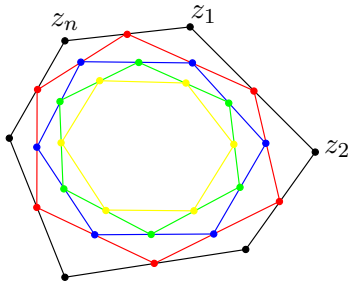


Correction de l'exercice E.7. (1) Pour $k = 0, \dots, n-1$, un calcul immédiat montre que le vecteur $(1, \zeta^k, \dots, \zeta^{(n-1)k})$ est un vecteur propre de C pour la valeur propre $\sum_{i=0}^{n-1} c_i \zeta^{ik}$. Puisque $\sum_{i=0}^{n-1} \zeta^{i(k-k')} = 0$ si $k \neq k'$, ces vecteurs sont deux à deux orthogonaux, et donc quitte à les renormaliser, ils forment une base orthonormée de $\mathbb{C}^n$, diagonalisant C. Le déterminant est alors

$$\det(C) = \prod_{k=0}^{n-1} \sum_{i=0}^{n-1} c_i \zeta^{ik} .$$

Toute matrice diagonalisable en base orthonormée est normale, comme vu dans le théorème 1.13.

(2) Nous identifions $\mathbb{R}^2$ et $\mathbb{C}$ de manière usuelle. Le résultat étant immédiat si $n = 2$, car alors P_i est réduit au milieu de P_0 pour $i > 0$, nous pouvons supposer que $n \geq 3$.



Notons $Z = (z_1, z_2, \dots, z_n) \in \mathbb{C}^n$ le n-uplet des sommets consécutifs de P (après le choix du premier d'entre eux). Soient C la matrice circulante $C = C(\frac{1}{2}, \frac{1}{2}, 0, \dots, 0)$, et $(Z_i)_{i\in\mathbb{N}}$ la suite dans $\mathbb{C}^n$ définie par récurrence par $Z_0 = Z$ et $Z_{i+1} = C\,Z_i$. Alors Z_i est le n-uplet des sommets consécutifs de P_i (pour un choix approprié du premier d'entre eux). Par la question (1), les valeurs propres de C sont $\frac{1}{2}(1+\zeta^k)$ pour $k = 0, \dots, n-1$ par la question (1). Elles sont donc ou bien égale à 1 avec multiplicité 1, ou bien de module strictement inférieur à 1. Par conséquent, dans une base orthonormée diagonalisant C de premier vecteur $\frac{1}{\sqrt{n}}\,(1, \dots, 1)$, la suite $(C^i)_{i\in\mathbb{N}}$ converge vers la matrice élémentaire $E_{1,1}$. La projection orthogonale de $\mathbb{C}^n$ sur la droite de vecteur directeur $\underline{1} = (1, \dots, 1)$ (dont la norme est $\sqrt{n}$) est $x \mapsto \langle x, \frac{1}{\sqrt{n}}\,\underline{1}\rangle\,\frac{1}{\sqrt{n}}\,\underline{1}$. Donc la suite $(Z_i = C^i Z)_{i\in\mathbb{N}}$ converge vers le vecteur $Z_\infty = (z, \dots, z)$ vérifiant $z = \frac{1}{n}(z_1 + z_2 + \cdots + z_n)$, c'est-à-dire l'isobarycentre de P. Puisqu'un polygone (pas forcément convexe) est contenu dans l'enveloppe convexe de ses sommets, le résultat en découle.

Correction de l'exercice E.8.

- La matrice de (la forme polaire de) $q = \sum_{1\le i\neq j\le n} z_i\overline{z_j}$ dans la base canonique de $\mathbb{C}^n$ est

$$\begin{pmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix}.$$

Cette matrice, qui est une matrice circulante, est, par l'exercice E.7, diagonalisable en base orthonormée, de valeur propre $n-1$ de multiplicité 1 pour le vecteur propre $(1, 1, \dots, 1)$ et de valeur propre -1 de multiplicité $n-1$ pour les vecteurs propres $(1, \zeta^k, \zeta^{2k}, \dots, \zeta^{(n-1)k})$ où $k = 1, \cdots, n-1$ et $\zeta = e^{2i\pi/n}$. Par conséquent, q est non dégénérée (donc de rang n), de signature $(n-1, 1)$.

- Si $n = 1$, la forme quadratique hermitienne $q = \sum_{1\le i,j\le n} (i+j)\, z_i\,\overline{z_j}$ est définie positive, sa signature est $(0, 1)$ et son rang est 1. Supposons donc dans ce qui suit que $n \ge 2$. Posons $Z = \sum_{1\le i\le n} i\, z_i$ et $W = \sum_{1\le j\le n} z_j$, qui sont des formes linéaires en $z = (z_1, \dots, z_n)$ telles que $Z + W$ et $Z - W$ sont linéairement indépendantes (car $n \ge 2$). Nous avons $\sum_{1\le i,j\le n} i\, z_i\,\overline{z_j} = Z\,\overline{W}$, donc

$$q(z) = Z\,\overline{W} + W\,\overline{Z} = \frac{1}{2}\,|Z+W|^2 - \frac{1}{2}\,|Z-W|^2\,.$$

Par conséquent, la signature est $(1, 1)$ et le rang est 2.

- Si $n = 1$, la forme quadratique hermitienne $q = \sum_{1\le i,j\le n} (i^2 + ij + j^2)\, z_i\,\overline{z_j}$ est définie positive, sa signature est $(0, 1)$ et son rang est 1. Si $n = 2$, alors

$$q = 3\,|z_1|^2 + 7\,(z_2\,\overline{z_1} + z_1\,\overline{z_2}) + 12\,|z_2|^2 = 3\,|z_1 + \frac{7}{3}\,z_2|^2 - \frac{13}{3}\,|z_2|^2\,,$$

donc la signature de q est $(1, 1)$ et le rang de q est 2. Supposons donc $n \ge 3$. Posons

$$X = \sum_{1\le i\le n} z_i, \quad Y = \sum_{1\le j\le n} j\, z_j \quad \text{et} \quad Z = \sum_{1\le k\le n} k^2\, z_k\,,$$

qui sont des formes linéaires en $z = (z_1, \dots, z_n)$ telles que $Z + X$, $Z - X$ et Y sont linéairement indépendantes (car $n \geq 3$ et par un petit calcul). Nous avons

$$q(z) = Z\,\overline{X} + |Y|^2 + X\,\overline{Z} = \frac{1}{2}\,|X+Z|^2 - \frac{1}{2}\,|X-Z|^2 + |Y|^2\,.$$

Donc la signature est $(1,2)$ et le rang est 3.

- La forme quadratique hermitienne $q = \sum_{1\leq i<j\leq n} |z_i - z_j|^2$ est positive, donc par l'inégalité de Cauchy-Schwarz (sans cas d'égalité, voir la note de bas de page 20), le noyau de q est égal à l'ensemble de ses vecteurs isotropes. Or nous avons $q(z) = 0$ si et seulement si $z_1 = z_2 = \cdots = z_n$, donc le rang de q est $n-1$ et la signature est $(0, n-1)$.

- Notons E^+ le sous-espace vectoriel complexe des polynômes pairs[58] sur $\mathbb{R}$, et E^- le sous-espace vectoriel complexe des polynômes impairs sur $\mathbb{R}$. Notons qu'un polynôme P appartient à E^+ si et seulement si ses coefficients d'indices impairs sont nuls, et à E^- si et seulement si ses coefficients d'indices pairs sont nuls. Donc E^+ est de dimension $\lfloor \frac{n+2}{2} \rfloor$ et E^- est de dimension $\lfloor \frac{n+1}{2} \rfloor$.

  Notons f la forme polaire de la forme quadratique hermitienne q. Si $P \in E^+$ et $Q \in E^-$, alors $f(P,Q) = \int_{-\infty}^{+\infty} e^{-t^2}\,P(t)\,\overline{Q(-t)}\;dt$ vaut $-\int_{-\infty}^{+\infty} e^{-t^2}\,P(t)\,\overline{Q(t)}\;dt$ en utilisant le fait que Q est impair, et $+\int_{-\infty}^{+\infty} e^{-t^2}\,P(t)\,\overline{Q(t)}\;dt$ par le changement de variable $t \mapsto -t$ et le fait que P est pair. Donc $f(P,Q) = 0$, et E^+ et E^- sont orthogonaux.

  Si $P \in E^+$, alors $q(P) = \int_{-\infty}^{+\infty} e^{-t^2}\,|P(t)|^2\;dt$, qui est strictement positif si P n'est pas le polynôme nul. Si $P \in E^-$, alors $q(P) = -\int_{-\infty}^{+\infty} e^{-t^2}\,|P(t)|^2\;dt$, qui est strictement négatif si P n'est pas le polynôme nul.

  Donc q a pour signature $\left(\lfloor \frac{n+1}{2} \rfloor, \lfloor \frac{n+2}{2} \rfloor\right)$, et rang $n+1$.

- Si $A = (a_{i,j})_{1\leq i,j\leq n}$, alors

$$\mathtt{tr}(A^2) = \sum_{i,j=1}^{n} a_{i,j}a_{j,i} = \sum_{i=1}^{n} a_{i,i}^2 + 2\sum_{1\leq i<j\leq n} \frac{1}{4}\big((a_{i,j}+a_{j,i})^2 - (a_{i,j}-a_{j,i})^2\big)\,.$$

Donc $A \mapsto \mathtt{tr}(A^2)$ est non dégénérée de signature $(\frac{n(n-1)}{2}, \frac{n(n+1)}{2})$ sur $\mathscr{M}_n(\mathbb{R})$. Elle est définie positive sur $\mathrm{Sym}_n(\mathbb{R})$.

- La forme quadratique hermitienne $A \mapsto |\,\mathtt{tr}\,A|^2$ a pour forme polaire la forme sesquilinéaire $(A,B) \mapsto (\mathtt{tr}\,A)(\,\overline{\mathtt{tr}\,B}\,)$, qui est non nulle, dégénérée, de noyau l'hyperplan des matrices de trace nulle, donc de rang 1, et de signature $(0,1)$ en restriction à toute droite vectorielle supplémentaire à son noyau.

Correction de l'exercice E.9. Cette correction est extraite de [FGN, page 133], voir aussi [Gou, §5.3.3].

(1) Nous pouvons supposer par continuité et densité que A est définie positive. Par le théorème de diagonalisation en base orthogonale, il existe $U = (u_{i,j})_{1\leq i,j\leq n} \in \mathrm{U}(n)$ et D une matrice diagonale de coefficients diagonaux strictement positifs $\lambda_1, \lambda_2, \dots, \lambda_n$ dans cet ordre telles que $A = U\,D\,U^*$. Par calcul de coefficients, pour tout $i = 1, \dots, n$, nous avons

$$a_{i,i} = \sum_{j=1}^{n} u_{i,j}\;\lambda_j\;\overline{u_{i,j}}\;\; = \sum_{j=1}^{n} |u_{i,j}|^2\;\lambda_j\,.$$

58. Une fonction $f : \mathbb{R} \to \mathbb{R}$ est *paire* si $f(-t) = f(t)$ pour tout $t \in \mathbb{R}$, et *impaire* si $f(-t) = -f(t)$ pour tout $t \in \mathbb{R}$

Puisque U est unitaire, ses vecteurs lignes sont de norme 1, donc $\sum_{j=1}^n |u_{i,j}|^2 = 1$ pour tout $i = 1, \dots, n$. Par concavité de la fonction ln, nous avons par conséquent $\ln a_{i,i} \geq \sum_{j=1}^n |u_{i,j}|^2 \ln \lambda_j$. D'où

$$\ln\Big(\prod_{i=1}^n a_{i,i}\Big) \geq \sum_{i,j=1}^n |u_{i,j}|^2 \ln \lambda_j = \sum_{j=1}^n \ln \lambda_j \sum_{i=1}^n |u_{i,j}|^2 = \sum_{j=1}^n \ln \lambda_j = \ln\Big(\prod_{i=1}^n \lambda_j\Big) ,$$

puisque U est unitaire, donc de vecteurs colonnes orthonormés. L'inégalité d'Hadamard en découle, puisque $\det A = \det D = \prod_{j=1}^n \lambda_i$.

Il est possible d'en déduire qu'il y a égalité dans l'inégalité d'Hadamard si et seulement si A est diagonale.

(2) Pour tout $(e_1, \dots, e_n) \in \mathscr{A}$, L'inégalité

$$\det f \leq \prod_{i=1}^n \langle f(e_i), e_i \rangle .$$

découle de l'assertion précédente en prenant la matrice de f dans la base $(e_1, \dots, e_n)$, et il y a égalité en prenant une base orthonormée dans laquelle f se diagonalise.

Correction de l'exercice E.10. Écrivons $A = (a_{i,k})_{\substack{1 \leq i \leq m \\ 1 \leq k \leq n}}$ et $B = (b_{k,j})_{\substack{1 \leq k \leq n \\ 1 \leq j \leq p}}$. Par la définition du produit des matrices, le résultat découle du fait que pour tous les éléments $i \in \{1, \dots, m\}$ et $j \in \{1, \dots, p\}$, la somme $\sum_{k=1}^n a_{i,k} b_{k,j}$ est nulle sauf peut-être s'il existe un $k \in \{1, \dots, n\}$ (forcément unique) tel que $i = k$ et $k = j$, ce qui force $i = j$.

Correction de l'exercice E.11. Il est immédiat que la matrice A_k, comme Σ_k, est de rang k. Puisque le groupe orthogonal préserve la norme euclidienne, et par définition de la norme subordonnée, nous avons $\|U' A' V'\| = \|A'\|$ pour tous les $A' \in \mathscr{M}_{m,n}(\mathbb{R})$, $U' \in \mathrm{O}(m)$ et $V' \in \mathrm{O}(n)$. Donc

$$\|A - A_k\| = \|U\,\Sigma\,{}^tV - U\,\Sigma_k\,{}^tV\| = \|\Sigma - \Sigma_k\| = \max_{k+1 \leq i \leq r} |\sigma_i| = \sigma_{k+1} .$$

Notons $(e_1, \dots, e_r)$ une suite libre de vecteurs propres de ${}^tA\,A$ associés aux valeurs propres $\sigma_1^2, \dots, \sigma_r^2$, de sorte que ${}^tV\,e_1, \dots, {}^tV\,e_r$ soient les r premiers vecteurs de la base canonique de $\mathbb{R}^n$. Le sous-espace vectoriel E_{k+1} de $\mathbb{R}^n$ engendré par $e_1, \dots, e_{k+1}$ est de dimension $k+1$. Si $B \in \mathscr{M}_{m,n}(\mathbb{R})$ est une matrice de rang k, alors son noyau N est de dimension $n-k$ par le théorème du rang. Donc l'intersection $E_{k+1} \cap N$, de dimension au moins 1, contient un vecteur unitaire $x \in \mathbb{R}^n$.

Par la définition de la norme d'opérateur, puisque x appartient au noyau de B, puisque U préserve la norme de $\mathbb{R}^m$, puisque ${}^tV\,x$ appartient au sous-espace vectoriel engendré par les $k+1$ premiers vecteurs de la base canonique de $\mathbb{R}^n$ (en restriction auquel la norme d'opérateur de la matrice diagonale Σ est au moins σ_{k+1}), et puisque V préserve la norme de $\mathbb{R}^n$, nous avons

$$\begin{aligned} \|A - B\| \geq \|(A-B)x\| = \|A\,x\| = \|U\,\Sigma\,{}^tV\,x\| = \|\Sigma\,{}^tV\,x\| \geq \sigma_{k+1}\|{}^tV\,x\| \\ = \sigma_{k+1}\|x\| = \sigma_{k+1} = \|A - A_k\| . \end{aligned}$$

Le résultat en découle.

Correction de l'exercice E.13. (1) Ceci découle de la proposition 1.26 (vi).

(2) Soit $\mathscr{B} = (e_1, e_2, e_3)$ la base canonique de $\mathbb{R}^3$.

Les deux membres de la formule $a \wedge (b \wedge c) = \langle a, c\rangle\, b - \langle a, b\rangle\, c$ sont multilinéaires en (a, b, c) et alternés en (b, c). Il suffit donc de la vérifier lorsque a, b et c appartiennent à $\mathscr{B}$, avec $b \neq c$. Elle est vraie si a est orthogonal à $\mathrm{Vect}\{b, c\}$, donc quitte à échanger b et c, il suffit de la vérifier pour $a = b$, et par permutation circulaire, si $a = b = e_1$. Le membre de droite vaut alors $-c$. Comme $e_1 \wedge (e_1 \wedge e_2) = e_1 \wedge e_3 = -e_2$ et $e_1 \wedge (e_1 \wedge e_3) = e_1 \wedge (-e_2) = -e_3$, le résultat en découle.

Les deux membres de la formule $(a \wedge b) \wedge (a \wedge c) = \det(a, b, c)\, a$ sont bilinéaires alternés en (b, c). Il suffit donc de la vérifier lorsque b et c appartiennent à $\mathscr{B}$, et, quitte à échanger b et c et à effectuer une permutation circulaire de $\mathscr{B}$, si $b = e_2$ et $c = e_3$. Si $a = (x, y, z)$, nous avons $\det(a, b, c) = x$ et $a \wedge e_2 = xe_3 - ze_1$ et $a \wedge e_3 = -xe_2 + ye_1$. L'assertion (1) de cet exercice montre que le membre de gauche est (x^2, xy, xz), comme voulu.

Nous avons, par la définition du produit extérieur et l'assertion que nous venons juste de démontrer,

$$\det(a \wedge b, a \wedge c, b \wedge c) = \langle (a \wedge b) \wedge (a \wedge c), b \wedge c\rangle = \det(a, b, c)\langle a, b \wedge c\rangle = (\det(a, b, c))^2 \ .$$

(3) Par invariance par rotations et par homothéties, nous pouvons supposer que a et b sont dans le plan horizontal, et que $a = (1, 0, 0)$. Si $b = (x, y, 0)$, nous avons donc

$$\|a \wedge b\| = \left\| \left(0, 0, \begin{vmatrix} 1 & x \\ 0 & y \end{vmatrix}\right) \right\| = |y| = \sqrt{x^2 + y^2}\ \frac{|y|}{\sqrt{x^2 + y^2}} \ ,$$

et le résultat en découle. La dernière affirmation découle du fait que

$$\sin^2 \angle(a, b) + \cos^2 \angle(a, b) = 1 \ .$$

Correction de l'exercice E.14. Nous travaillons en hermitien, le cas euclidien est analogue. Sauf pour la question (7), voir par exemple [Ber1, §8.11], ou [Gou, §5.3.4].

(1) Soient E' le sous-espace vectoriel engendré par $x_1, \ldots, x_n$, et m sa dimension. Soit $E'' = E' \overset{\perp}{\oplus} \mathbb{C}^{n-m}$ la somme directe orthogonale de E' avec l'espace hermitien standard $\mathbb{C}^{n-m}$. Soit $\mathscr{B}'$ une base orthonormée de E' étendue par la base canonique de $\mathbb{C}^{n-m}$ en une base orthonormée $\mathscr{B}''$ de E''. Notons A la matrice $n \times n$ dont les lignes sont les vecteurs lignes des coordonnées de $x_1, \ldots, x_n$ dans la base $\mathscr{B}''$. Alors la matrice de Gram est exactement AA^*, par définition du produit matriciel. Donc

$$\mathrm{Gram}(x_1, \ldots, x_n) = \det AA^* = |\det A|^2 \geq 0$$

avec égalité si et seulement si A n'est pas inversible, c'est-à-dire si et seulement si $x_1, \ldots, x_n$ sont linéairement dépendants. Notons que nous avons alors

$$\det A = \det({}^t A) = \det{}_{\mathscr{B}''}(x_1, \ldots, x_n)$$

par définition de $\det_{\mathscr{B}''}$. La première affirmation de l'assertion (1) en découle en distinguant les cas $m \neq n$ et $m = n$.

Montrons maintenant la seconde affirmation. Nous avons

$$\mathrm{Gram}(x, y) = \begin{vmatrix} \|x\|^2 & \langle x, y\rangle \\ \langle y, x\rangle & \|y\|^2 \end{vmatrix} = \|x\|^2\|y\|^2 - |\langle x, y\rangle|^2 \ .$$

d'où l'inégalité de Cauchy-Schwarz avec égalité si et seulement si x et y sont linéairement dépendants.

(2) Notons E' l'espace vectoriel engendré par $x_1, \dots, x_n$. Il est immédiat que (iii) implique (ii), car les mineurs d'ordre au moins $r+1$ d'une matrice de rang r sont nuls.

Si l'assertion (ii) est vérifiée, alors par la question (1), toute partie de $\{x_1, \dots, x_n\}$ ayant $r+1$ éléments est liée, donc la dimension de E' est au plus r, ce qui montre l'assertion (i).

Si l'assertion (i) est vérifiée, alors les colonnes de toute matrice carrée extraite de la matrice de Gram de $x_1, \dots, x_n$ de taille au moins $r+1$ sont linéairement dépendantes, donc le mineur correspondant est nul.

(3) Si $P \in \mathrm{GL}(E)$ envoie une base $\mathscr{B} = (x_1, \dots, x_n)$ sur une base $\mathscr{B}' = (x'_1, \dots, x'_n)$, alors, avec $\mathscr{A}$ une base orthonormée de E, nous avons

$$\begin{aligned}\mathrm{Gram}(x'_1, \dots, x'_n) &= (\det{}_{\mathscr{A}}(x'_1, \dots, x'_n))^2 = (\det{}_{\mathscr{A}}(Px_1, \dots, Px_n))^2 \\ &= (\det P)^2 (\det{}_{\mathscr{A}}(x_1, \dots, x_n))^2 = (\det P)^2 \,\mathrm{Gram}(x_1, \dots, x_n)\ .\end{aligned}$$

(4) Soit $x \in E$. Nous pouvons supposer que x n'appartient pas à F, sinon les deux membres de l'égalité à démontrer sont nuls. Fixons une base orthonormée $\mathscr{A}$ de $F \oplus \mathbb{C}x$. Soit $y \in F$ la projection orthogonale de x sur F, de sorte que $d(x, F) = \|x - y\|$ et que $x - y$ est orthogonal à F. Nous avons par ce qui précède

$$\begin{aligned}\mathrm{Gram}(x, x_1, \dots, x_n) &= (\det{}_{\mathscr{A}}(x, x_1, \dots, x_n))^2 = (\det{}_{\mathscr{A}}(x - y, x_1, \dots, x_n))^2 \\ &= \mathrm{Gram}(x - y, x_1, \dots, x_n)\ .\end{aligned}$$

Comme $x - y$ est orthogonal à $x_1, \dots, x_n$, la matrice de Gram de $(x - y, x_1, \dots, x_n)$ a ses coefficients en première ligne et première colonne nuls sauf le coefficient $(1, 1)$, et

$$\mathrm{Gram}(x - y, x_1, \dots, x_n) = \|x - y\|^2 \,\mathrm{Gram}(x_1, \dots, x_n)\ ,$$

ce qu'il fallait démontrer.

Dans l'espace de Hilbert réel $\mathscr{H} = \mathbb{L}^2([0, 1])$, soit F le sous-espace vectoriel réel engendré par les fonctions $x, x^2, \dots, x^n$. Il s'agit de calculer le carré A de la distance de la fonction constante 1 à F dans $\mathscr{H}$. Par la formule précédente, nous avons donc

$$A = \frac{\det\left(\left(\int_0^1 x^{i+j}\,dx\right)_{0\le i,j\le n}\right)}{\det\left(\left(\int_0^1 x^{i+j}\,dx\right)_{1\le i,j\le n}\right)} = \frac{\det\left(\left(\frac{1}{i+j+1}\right)_{0\le i,j\le n}\right)}{\det\left(\left(\frac{1}{i+j+1}\right)_{1\le i,j\le n}\right)}\ .$$

(5) Le résultat est immédiat si $x_1, \dots, x_n$ sont linéairement dépendants, donc nous supposons que $(x_1, \dots, x_n)$ est une base de $\mathbb{R}^n$. L'assertion (5) découle alors de l'assertion (1) et du fait que si $\mathscr{B}$ est une base orthonormée de $\mathbb{R}^n$, alors

$$\mathrm{vol}(P(x_1, \dots, x_n)) = |\det{}_{\mathscr{B}}(x_1, \dots, x_n)|\ . \tag{25}$$

Ceci se démontre en utilisant le fait que le volume du cube unité est 1, et le fait que $\det_{\mathscr{B}}(x_1, \dots, x_n)$ est le déterminant de l'application de changement de base de la base canonique à la base $(x_1, \dots, x_n)$, par un changement de variable linéaire dans le calcul de l'intégrale de la fonction constante 1 sur le polyèdre $P(x_1, \dots, x_n)$ pour la mesure de Lebesgue.

(6) Notons $\mathscr{A}$ une base orthonormée de E et $z = x_1 \wedge \cdots \wedge x_{n-1}$. Le résultat est immédiat si $x_1, \dots, x_{n-1}$ sont linéairement dépendants, donc nous supposons que $z \neq 0$. Puisque z est orthogonal à $x_1, \dots, x_{n-1}$ par la proposition 1.26 (iii), par l'assertion (1), et par la définition du produit extérieur, nous avons

$$\begin{aligned}\|z\|^2 \operatorname{Gram}(x_1, \dots, x_{n-1}) &= \operatorname{Gram}(z, x_1, \dots, x_{n-1}) = (\det{}_{\mathscr{A}}(z, x_1, \dots, x_{n-1}))^2 \\ &= (\det{}_{\mathscr{A}}(x_1, \dots, x_{n-1}, z))^2 = (\langle x_1 \wedge \cdots \wedge x_{n-1}, z\rangle)^2 = \|z\|^4 \,,\end{aligned}$$

Le résultat s'en déduit car $z \neq 0$.

(7) Voir par exemple le problème 13 page 73 de [Zav], ou [Ber1, §9.7.3].

(i) La première affirmation s'obtient par permutation de deux lignes et de deux colonnes (dans un ordre indifférent, vu que les éléments diagonaux sont nuls.)

En enlevant la seconde ligne aux lignes suivantes, puis la seconde colonne aux colonnes suivantes, puis en développant par rapport à la première ligne et par rapport à la première colonne, et enfin en changeant tous les coefficients de signe, nous obtenons

$$\begin{aligned}\Gamma_n &= \begin{vmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & X_{0,1}^2 & \dots & X_{0,n}^2 \\ 1 & X_{0,1}^2 & & & \\ \vdots & \vdots & & \big(X_{i,j}^2\big)_{1\leq i,j\leq n} & \\ 1 & X_{0,n}^2 & & & \end{vmatrix} \\ &= \begin{vmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & X_{0,0}^2 & X_{0,1}^2 & \dots & X_{0,n}^2 \\ 0 & X_{0,1}^2 & & & \\ \vdots & \vdots & & \big(X_{i,j}^2 - X_{0,j}^2\big)_{1\leq i,j\leq n} & \\ 0 & X_{0,n}^2 & & & \end{vmatrix} \\ &= \begin{vmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & X_{0,0}^2 & X_{0,1}^2 & \dots & X_{0,n}^2 \\ 0 & X_{0,1}^2 & & & \\ \vdots & \vdots & & \big(X_{i,j}^2 - X_{0,j}^2 - X_{0,i}^2\big)_{1\leq i,j\leq n} & \\ 0 & X_{0,n}^2 & & & \end{vmatrix} \\ &= -\det\big((X_{i,j}^2 - X_{0,j}^2 - X_{0,i}^2)_{1\leq i,j\leq n}\big) = (-1)^{n+1}\det\big((X_{0,j}^2 + X_{0,i}^2 - X_{i,j}^2)_{1\leq i,j\leq n}\big) \,,\end{aligned}$$

ce qui donne le résultat sachant que $X_{i,i} = 0$.

(ii) Par le développement du carré de la norme d'une différence, et par la relation de Chasles, pour tous les $1 \leq i, j \leq n$, nous avons

$$\begin{aligned}\langle \overrightarrow{x_0x_i}, \overrightarrow{x_0x_j}\rangle &= \frac{1}{2}\big(\|\overrightarrow{x_0x_i}\|^2 + \|\overrightarrow{x_0x_j}\|^2 - \|\overrightarrow{x_0x_j} - \overrightarrow{x_0x_i}\|^2\big) \\ &= \frac{1}{2}\big(\|\overrightarrow{x_0x_i}\|^2 + \|\overrightarrow{x_0x_j}\|^2 - \|\overrightarrow{x_ix_j}\|^2\big) = \frac{1}{2}\big(d_{0,i}^2 + d_{0,j}^2 - d_{i,j}^2\big) \,.\end{aligned}$$

La formule découle donc de la définition d'un déterminant de Gram et de la question (i), après avoir mis en facteur la fraction $\frac{1}{2}$ sur chacune des n lignes.

(iii) Puisque les $n+1$ points $x_0, \dots, x_n$ appartiennent à un même sous-espace affine de E' de dimension $n-1$ si et seulement si les n vecteurs $\overrightarrow{x_0x_1}, \dots, \overrightarrow{x_0x_n}$ sont linéairement dépendants, la première affirmation découle de la seconde affirmation de la question (1).

Par conséquent, en posant $a = d(y,z)$, $b = d(x,z)$ et $c = d(x,y)$, les points x, y, z sont alignés si et seulement si

$$\begin{aligned} 0 = \Gamma_2(a,b,c) &= (-1)^3 \begin{vmatrix} 2\,c^2 & b^2+c^2-a^2 \\ b^2+c^2-a^2 & 2\,b^2 \end{vmatrix} \\ &= (b^2+c^2-a^2)^2 - (2bc)^2 = -(a^2-(b-c)^2)((b+c)^2-a^2)\,, \end{aligned}$$

donc si et seulement si $a = |b-c|$ ou $a = b+c$.

(iv) Si P est le volume du parallélépipède de sommets $x_0, \dots, x_n$, alors $V = \frac{1}{n!}\,P$, et le résultat découle donc des questions (5) et (ii) :

$$V^2 = \frac{1}{(n!)^2}\,P^2 = \frac{1}{(n!)^2}\,\operatorname{Gram}(\overrightarrow{x_0x_1}, \dots, \overrightarrow{x_0x_n}) = \frac{(-1)^{n+1}}{2^n(n!)^2}\,\Gamma_n(d_{i,j} : 0 \le i < j \le n)\,.$$

(v) Par les questions (iii) et (iv), nous avons

$$\begin{aligned} \mathscr{A}^2 &= \frac{(-1)^3}{2^2(2!)^2}\,\Gamma_2(a,b,c) = \frac{1}{4^2}\,(a^2-(b-c)^2)((b+c)^2-a^2) \\ &= \frac{1}{4^2}\,(a+b-c)(a-b+c)(a+b+c)(b+c-a)\,. \end{aligned}$$

(8) Nous pouvons supposer que A est l'origine de $\mathbb{R}^3$. Nous notons x, y, z les trois vecteurs définissant les arêtes en A. Rappelons que le produit scalaire de deux vecteurs unitaires est le cosinus de l'angle qu'ils forment, et qu'un parallélépipède P se subdivise en six tétraèdres de même volume dont l'un est le tétraèdre formé par un sommet de P et les trois arêtes de P qui en partent : voir une visualisation sur YouTube

https ://www.youtube.com/watch ?v=zf6zGe4kI0s !!

Donc

$$\begin{aligned} \operatorname{vol}(T) = \frac{1}{6}\,\operatorname{vol} P(x,y,z) = \frac{1}{6}\,\operatorname{Gram}(x,y,z)^{\frac12} &= \frac{1}{6} \begin{vmatrix} 1 & \cos\frac{\pi}{p} & \cos\frac{\pi}{r} \\ \cos\frac{\pi}{p} & 1 & \cos\frac{\pi}{q} \\ \cos\frac{\pi}{r} & \cos\frac{\pi}{q} & 1 \end{vmatrix}^{\frac12} \\ = \frac{1}{6}\,\big(1-\cos^2\frac{\pi}{p}-\cos^2\frac{\pi}{q}-\cos^2\frac{\pi}{r}+2\cos\frac{\pi}{p}\cos\frac{\pi}{q}\cos\frac{\pi}{r}\big)^{\frac12}\,. \end{aligned}$$

Correction de l'exercice E.15. Puisque le polynôme caractéristique de u est scindé sur $\mathbb{C}$, l'endomorphisme u est trigonalisable. Soit $(f_1, \dots, f_n)$ une base de E dans laquelle la matrice de u est triangulaire supérieure. Soit $(e_1, \dots, e_n)$ la base orthonormée de E obtenue par le procédé d'orthonormalisation de Gram-Schmidt. Puisque

$$\operatorname{Vect}\{e_1, \dots, e_k\} = \operatorname{Vect}\{f_1, \dots, f_k\}$$

pour $k = 1, \dots, n$, la matrice de u dans la base $(e_1, \dots, e_n)$ est aussi triangulaire supérieure.

Correction de l'exercice E.16. (1) L'application $(P,Q) \mapsto \langle P,Q\rangle$ est clairement sesquilinéaire, hermitienne et positive. Puisque le seul polynôme complexe qui est nul sur le cercle est le polynome nul, elle est définie. Puisque pour tout entier k, l'intégrale $\int_0^{2\pi} e^{kit}\,dt$ vaut 2π si $k=0$ et 0 sinon, la base canonique de $\mathbb{C}_n[X]$ est orthonormée.

(2) Les formes linéaires d'évaluations en $a_1,\dots,a_k$ sont linéairement indépendantes, car les a_i sont deux à deux distincts et le déterminant de la *matrice de Vandermonde*

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_k \\ a_1^2 & a_2^2 & \cdots & a_k^2 \\ \vdots & \vdots & \vdots & \vdots \\ a_1^{k-1} & a_2^{k-1} & \cdots & a_k^{k-1} \end{pmatrix}$$

est $\prod_{1\le i<j\le k}(a_j-a_i)$, donc non nul. L'intersection des noyaux de ces formes linéaires est donc un sous-espace vectoriel de dimension $n+1-k$. Elle admet donc une base orthonormée $(P_k,P_{k+1},\dots,P_n)$, que nous pouvons compléter en une base orthonormée de $\mathbb{C}_n[X]$ pour obtenir une base cherchée.

Correction de l'exercice E.17. Par les formules de polarisation (4) en euclidien et (6) en hermitien, et puisqu'un produit scalaire est non dégénéré, nous avons

$$\begin{aligned} & \forall\, x\in E,\quad \|u(x)\| = \|u^*(x)\| \\ \Leftrightarrow\quad & \forall\, x,y\in E,\quad \langle u(x),u(y)\rangle = \langle u^*(x),u^*(y)\rangle \\ \Leftrightarrow\quad & \forall\, x,y\in E,\quad \langle u^*\circ u(x),y\rangle = \langle u\circ u^*(x),y\rangle \\ \Leftrightarrow\quad & \forall\, x\in E,\quad u^*\circ u(x) = u\circ u^*(x) \quad\Leftrightarrow\quad u^*\circ u = u\circ u^*\,. \end{aligned}$$

Correction de l'exercice E.18. (1) et (2). L'application $\langle\cdot,\cdot\rangle_{HS} : (A,B)\mapsto \mathtt{tr}(B^*A)$ est, par la linéarité de la trace et son invariance par la transposition, clairement sesquilinéaire et hermitienne. Pour tout $A=(a_{ij})_{1\le i,j\le n}\in\mathscr{M}_n(\mathbb{C})$, nous avons $A^*A = (\sum_{1\le k\le n}\overline{a_{ik}}\,a_{kj})_{1\le i,j\le n}$, donc

$$\langle A,A\rangle_{HS} = \mathtt{tr}(A^*A) = \sum_{1\le i,j\le n} |\,a_{ij}\,|^2\,.$$

Ceci montre que la forme sesquilinéaire hermitienne $\langle\cdot,\cdot\rangle_{HS}$ est définie positive, et répond aussi à la question (2) : nous avons

$$\|A\|_{HS} = \Big(\sum_{1\le i,j\le n} |a_{ij}|^2\Big)^{\frac12}\,. \tag{26}$$

(3) Soit $\mathscr{B}=(e_1,\dots,e_n)$. Si $\mathscr{B}'=(e'_1,\dots,e'_n)$ est la base canonique de $\mathbb{C}^n$, alors $\|A\,e'_i\|^2 = \sum_{j=1}^n |a_{ij}|^2$, et le résultat découle lorsque $\mathscr{B}=\mathscr{B}'$ de la question précédente. Soit U la matrice de passage de la base canonique $\mathscr{B}'$ à la base $\mathscr{B}$. Par les propriétés de la trace et puisque U est unitaire, nous avons

$$\mathtt{tr}((A\,U)^*\,(A\,U)) = \mathtt{tr}(U^*\,A^*\,A\,U) = \mathtt{tr}(U\,U^*\,A^*\,A\,) = \mathtt{tr}(A^*\,A)\,.$$

Donc $\sum_{i=1}^n \|A\,e_i\|^2 = \sum_{i=1}^n \|A\,U\,e'_i\|^2 = \|A\,U\|_{HS} = \|A\|_{HS}$.

(4) Si $A = (a_{ij})_{1\leq i,j\leq n} \in \mathscr{M}_n(\mathbb{C})$, alors $A = \sum_{1\leq i,j\leq n} a_{ij}E_{i,j}$, et le résultat découle alors de la formule (26).

(5) Si $U \in \mathscr{M}_n(\mathbb{C})$ est unitaire, alors $\|U\|_{HS} = (\mathtt{tr}(U^*\,U))^{\frac{1}{2}} = (\mathtt{tr}\,I_n)^{\frac{1}{2}} = \sqrt{n}$.

(5) Si $A \in \mathscr{M}_n(\mathbb{R})$ (donc $A \in \mathscr{M}_n(\mathbb{C})$) est symétrique, alors A est diagonalisable en base orthonormée de $\mathbb{R}^n$ (donc de $\mathbb{C}^n$). Il existe donc une matrice réelle orthogonale (donc complexe unitaire) U et une matrice réelle diagonale D, dont les coefficients diagonaux sont les valeurs propres $\lambda_1, \ldots, \lambda_n$ de A, telles que $A = U\,D\,U^*$. Puisque les coefficients de A et D sont réels, nous avons donc

$$\sum_{i,j=1}^{n} a_{i,j}^2 = \sum_{i,j=1}^{n} |a_{i,j}|^2 = \|A\|_{HS} = \|U\,D\,U^*\|_{HS} = \|D\|_{HS} = \sum_{i=1}^{n} |\lambda_i|^2 = \sum_{i=1}^{n} \lambda_i^2 \; .$$

Correction de l'exercice E.19. (1) Soit $p \in \mathscr{L}(E)$ un opérateur idempotent tel que $\ker p = (\mathrm{im}\; p)^\perp$. Alors, en notant $F = \mathrm{im}\; p$, pour tous les x et y dans $\mathscr{H}$, les deux vecteurs $p(x) \in F$ et $y - p(y) \in \ker p = F^\perp$, ainsi que les deux vecteurs $p(x) - x \in \ker p = F^\perp$ et $p(y) \in F$ sont orthogonaux, et donc

$$\langle p(x), y\rangle = \langle p(x), p(y)\rangle = \langle x, p(y)\rangle \; .$$

Ceci montre que P est auto-adjoint, et positif, car $\langle p(x), x\rangle = \langle p(x), p(x)\rangle \geq 0$.

Réciproquement, soit $p \in \mathscr{L}(E)$ un opérateur auto-adjoint idempotent. Si $y = p(x)$, alors $p(y) = p^2(x) = p(x) = y$. Donc l'image $p(E)$ de p est contenue dans le noyau de $\mathrm{id} - p$, donc égal à ce noyau (car réciproquement, si $x \in \ker(\mathrm{id} - p)$, alors $x = p(x)$, donc x appartient à l'image de p). En particulier, $\mathrm{im}\; p$ est fermé par la continuité de $p \in \mathscr{L}(E)$. Puisque p est auto-adjoint, pour tous les $x \in \ker p$ et $y \in \mathrm{im}\; p$, si $x' \in E$ vérifie $p(x') = y$, nous avons

$$\langle x, y\rangle = \langle x, p(x')\rangle = \langle p(x), x'\rangle = 0 \; .$$

Donc $\ker p \subset (\mathrm{im}\; p)^\perp$. Réciproquement, si $x \in (\mathrm{im}\; p)^\perp$, alors pour tout $y \in E$, nous avons $\langle p(x), y\rangle = \langle x, p(y)\rangle = 0$. Donc $p(x) = 0$ et $x \in \ker p$.

(2) Il suffit de montrer que pour tout $y \in E$, nous avons $\|y - x\| \geq \|p(x) - x\|$. Puisque $y - p(x) \in F$ et $p(x) - x \in F^\perp$ sont orthogonaux, la formule de Pythagore nous donne

$$\|y - x\| = \|(y - p(x)) + (p(x) - x)\| = \|y - p(x)\| + \|p(x) - x\| \geq \|p(x) - x\| \; .$$

(3) a) Nous avons en effet

$$\|\; f(a,b) - y\;\| = \|\; a\,x + b\,\underline{1} - y\;\| = \sqrt{\sum_{i=1}^{n} (a\,x_i + b - y_i)^2} \; .$$

b) Si F_x est le sous-espace de $\mathbb{R}^n$ engendré par x et $\underline{1}$, et si p_x est la projection orthogonale sur F_x, alors le couple cherché (a, b) est tel que $f(a, b) = p_x(y)$ d'après la question (2). Cette égalité implique que $f(a, b) - y$ est orthogonal à F_x, c'est-à-dire à la fois à x et à $\underline{1}$. Ceci donne le système

$$\langle (a\,x + b\,\underline{1} - y, x\rangle = 0 \quad \text{et} \quad \langle (a\,x + b\,\underline{1} - y, \underline{1}\rangle = 0 \; ,$$

qui se réécrit comme cherché

$$a\,\langle x,x\rangle + b\,\langle \underline{1},x\rangle = \langle x,y\rangle \quad \text{et} \quad a\,\langle \underline{1},x\rangle + b\,\langle \underline{1},\underline{1}\rangle = \langle \underline{1},y\rangle\ .$$

c) Les formules de Cramer montrent le résultat.

(4) a) Puisque A est de rang p et puisque le rang d'une matrice est inférieure au nombre de ses lignes, nous avons $p \leq n$. L'application linéaire $A : \mathbb{R}^p \to \mathbb{R}^n$ est injective, sinon par le théorème du rang, son rang serait strictement inférieur à p. Notons F son image, π la projection orthogonale sur F, et X l'unique (par injectivité) élément de $\mathbb{R}^p$ tel que $A\,X = \pi(B)$. Alors X est l'unique minimum cherché. Pour tout $Y \in \mathbb{R}^p$, puisque $\pi(B) - B = A\,X - B$ est orthogonal à F, nous avons

$$\langle\, {}^tA\,A\,X,\ Y\rangle = \langle A\,X,\ A\,Y\rangle = \langle A\,X - B,\ A\,Y\rangle + \langle B,\ A\,Y\rangle = \langle\, {}^tA\,B,\ Y\rangle\ .$$

Ceci montre que X est une solution du système linéaire ${}^tA\,A\,X = {}^tA\,B$, car le produit scalaire de $\mathbb{R}^p$ est non dégénéré. L'unicité vient du fait que A étant injective, la matrice ${}^tA\,A$, symétrique de taille $p \times p$, est définie positive.

b) Soient $(x_i, y_i) \in \mathbb{R}^2$ pour $1 \leq i \leq n$, posons $B = (y_1, \dots, y_n)$ et

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{p-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{p-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{p-1} \end{pmatrix}\ .$$

Chercher un polynôme réel $P = a_0 + a_1 Z + \cdots + a_{p-1} Z^{p-1}$ de degré strictement inférieur à p qui minimise la quantité $\sum_{i=1}^n (P(x_i) - y_i)^2$ est équivalent à chercher un p-uplet de réels $X = (a_0, \dots, a_{p-1})$ tel que $\|AX - B\|$ soit minimum. Supposons que A soit de rang p, ce qui est possible par la formule du déterminant de Vandermonde, s'il y a au moins p éléments x_i deux à deux distincts, et sinon, on cherche un polynôme de degré plus petit. Par la question précédente, le polynôme P est donc l'unique polynôme de degré au plus $p-1$ dont le p-uplet de coefficients est $({}^tA\,A)^{-1}\,{}^tA\,B$.

Lorsque $p = 2$, on est ramené à la question (3).

Correction de l'exercice E.20. A (1) Pour tout $x \in E$ de coordonnées $(x_1, \dots, x_n)$ dans $\mathscr{B}$, par la linéarité de ℓ, nous avons

$$\ell(x) = \ell\Big(\sum_{i=1}^n x_i\,e_i\Big) = \sum_{i=1}^n \ell(e_i)\,x_i = \sum_{i=1}^n \ell(e_i)\,e_i^*(x)\ ,$$

donc $\ell = \sum_{i=1}^n \ell(e_i)\,e_i^*$.

(2) Pour tous les $x \in E$ et $\ell \in E^*$, notons X et L les colonnes des coordonnées de x et ℓ dans les bases $\mathscr{B}$ et $\mathscr{B}^*$ respectivement, de sorte que $\ell(x) = {}^tL\,X$. Alors

$${}^tu(\ell)(x) = \ell(u(x)) = {}^tL(M\,X) = {}^t({}^tM\,L)X\ ,$$

ce qui montre le résultat.

(3) Pour tout $x \in E$, notons X et X' les colonnes des coordonnées de x dans les bases $\mathscr{B}$ et $\mathscr{C}$ respectivement, de sorte que $X = P\,X'$. Pour tout $\ell \in E^*$, notons L et L' les colonnes des coordonnées de ℓ dans les bases $\mathscr{B}^*$ et $\mathscr{C}^*$ respectivement, de sorte que $L = \breve{P}\,L'$. Alors

$$\ell(x) = {}^tL\,X = {}^t(\breve{P}\,L')\,(P\,X') = {}^tL'({}^t\breve{P}\,P)X' \,.$$

Puisque nous avons aussi $\ell(x) = {}^tL'\,X'$, et ceci pour tous les $x \in E$ et $\ell \in E^*$, nous avons donc ${}^t\breve{P}\,P = \mathrm{id}$, ce qui donne le résultat.

(4) Si $\mathscr{B} = (e_1, \dots, e_n)$ est une base de E de base duale $\mathscr{B}^* = (e_1^*, \dots, e_n^*)$, alors pour tous les $i, j = 1, \dots, n$, nous avons $\mathrm{ev}_{e_i}(e_j^*) = e_j^*(e_i) = \delta_{j,i} = \delta_{i,j}$, donc la base duale de $\mathscr{B}^*$ est $(\mathrm{ev}_{e_1}, \dots, \mathrm{ev}_{e_n})$.

(5) Soient $\lambda_0, \dots, \lambda_n \in k$. Si $\sum_{i=0}^n \lambda_i \ell_i = 0$, alors en appliquant le membre de gauche de cette égalité aux polynômes $1, X \dots, X^n$, nous obtenons le système

$$\forall j = 0, \dots, n, \quad \sum_{i=0}^{n} \lambda_i\, a_i^j = 0$$

dont la matrice est la matrice de Vandermonde $(a_i^j)_{0 \le i,j \le n}$. Puisque $a_0, \dots, a_n$ sont deux à deux distincts, le *déterminant de Vandermonde* $\det\big((a_i^j)_{0 \le i,j \le n}\big) = \prod_{i<j}(a_j - a_i)$ est non nul. Donc $\lambda_0 = \cdots = \lambda_n = 0$, et les $n+1$ formes linéaires $\ell_0, \dots, \ell_n$ sont linéairement indépendantes dans un espace vectoriel de dimension $k+1$. Elles forment par conséquent une base de $k_n[X]^*$.

Posons $A_i = \prod_{0 \le j \le n,\, j \ne i}(X - a_j)$ et $P_i = \frac{1}{A_i(a_i)}\, A_i$. Notons que $A_i(a_i) \ne 0$ car $a_0, \dots, a_n$ sont deux à deux distincts. Alors $\ell_j(P_i) = P_i(a_j) = \delta_{j,i}$, donc $(P_0, \dots, P_n)$ est une base de $k_n[X]$ de base duale $(\ell_0, \dots, \ell_n)$.

Les coordonnées de P dans la base $(P_0, \dots, P_n)$ sont, par définition de la base duale, égales à $(\ell_0(P), \dots, \ell_n(P)) = (P(a_0), \dots, P(a_n))$.

B (1) i) Soient $\ell_1, \dots, \ell_n \in E^*$. Par la question A (1), les coordonnées de ℓ_j dans $\mathscr{B}^*$ sont $(\ell_j(e_1), \dots, \ell_j(e_n))$. Donc par la définition du produit extérieur, puisque le déterminant d'une matrice et de sa matrice transposée coïncident, et par la définition de l'application $\det_{\mathscr{B}^*}$ dans le début de la partie 1.11, nous avons

$$\begin{aligned} e_1 \wedge \cdots \wedge e_n(\ell_1, \dots, \ell_n) &= \det\big(\,(\ell_i(e_j))_{1 \le i,j \le n}\,\big) = \det\big(\,(\ell_j(e_i))_{1 \le i,j \le n}\,\big) \\ &= \det{}_{\mathscr{B}^*}(\ell_1, \dots, \ell_n)\,. \end{aligned}$$

Puisque l'espace vectoriel E^* est de dimension n et puisque $\mathscr{B}^*$ est une base de E^*, nous savons déjà (voir le cours d'algèbre linéaire sur le déterminant) que l'espace vectoriel des formes n-linéaires alternées sur E^*, qui est $\Lambda^n(E^*)^* = \Lambda^n E$, est une droite vectorielle, engendrée par $\det_{\mathscr{B}^*}$.

ii) Soient $w \in \Lambda^k E$ et $\ell_1, \dots, \ell_k \in E^*$. Puisque $(\ell_j(e_1), \dots, \ell_j(e_n))$ est le n-uplet des coordonnées de ℓ_j d'après la question A (1), par multilinéarité et par antisymétrie, nous

avons

$$\begin{aligned} w(\ell_1,\dots,\ell_k) &= w\Big(\sum_{i_1=1}^{n} \ell_1(e_{i_1})\, e^*_{i_1},\dots,\sum_{i_k=1}^{n} \ell_k(e_{i_k})\, e^*_{i_k}\Big) \\ &= \sum_{1\le i_1,\dots,i_k\le n} \ell_1(e_{i_1})\dots\ell_k(e_{i_k})\, w(e^*_{i_1},\dots,e^*_{i_k}) \\ &= \sum_{1\le i_1<\dots<i_k\le n} w(e^*_{i_1},\dots,e^*_{i_k})\ \det\big(\ (\ell_p(e_{i_q}))_{1\le p,q\le k}\ \big) \\ &= \sum_{1\le i_1<\dots<i_k\le n} w(e^*_{i_1},\dots,e^*_{i_k})\ e_{i_1}\wedge\dots\wedge e_{i_k}(\ell_1,\dots,\ell_k)\ . \end{aligned}$$

Il en découle que la famille $(e_{i_1}\wedge\dots\wedge e_{i_k})_{1\le i_1<\dots<i_k\le n}$ est une partie génératrice avec écriture unique de $\Lambda^k E$, donc une base de $\Lambda^k E$.

(2) Notons $\varepsilon : \mathscr{S}_n \to \{\pm 1\}$ le morphisme de signature sur le groupe symétrique $\mathscr{S}_n$ des permutations de l'ensemble $\{1,\dots,n\}$. Pour tous les éléments entiers $k\in\{0,\dots,n\}$ et $1\le i_1<\dots<i_k\le n$, en utilisant des transpositions pour remettre i_k à la bonne place, nous avons

$$\begin{aligned} &\varepsilon(i_1,\dots,i_k,1,2,\dots,\widehat{i_1},\dots,\widehat{i_k},\dots,n) \\ &= (-1)^{(i_k-1)-(k-1)}\varepsilon(i_1,\dots,i_{k-1},1,2,\dots,\widehat{i_1},\dots,\widehat{i_{k-1}},\dots,n)\ . \end{aligned}$$

Donc par récurrence, nous avons $\varepsilon_{\{i_1,\dots,i_k\}} = (-1)^{\sum_{\ell=1}^k (i_j-\ell)} = (-1)^{\sum_{\ell=1}^k i_j - \frac{k(k+1)}{2}}$.

(3) et (4) Fixons une base orthonormée directe $\mathscr{B}=(e_1,\dots,e_n)$ de E.

Montrons l'unicité d'un endomorphisme $*$ vérifiant les conditions demandées dans l'assertion (3). Notons que $\Lambda^0 E$ et $\Lambda^n E$ de E sont de dimension 1, engendrés par $\mathbf{1}$ et $e_1\wedge\dots\wedge e_n$ respectivement, d'après la question (1) i). De plus, si $k\in\{1,\dots,n-1\}$, alors l'espace vectoriel $\Lambda^k E$ est engendré par $(e_{i_1}\wedge\dots\wedge e_{i_k})_{1\le i_1<\dots<i_k\le n}$, par la question (1) ii). Comme $k<n$, la suite orthonormée $e_{i_1},\dots,e_{i_k}$ peut se compléter en une base orthonormée directe de E. Donc les conditions demandées dans l'assertion (3), avec la condition de linéarité, déterminent uniquement l'application linéaire $*$ sur $\Lambda^k E$

Par la question (1) ii), la famille $(e_{i_1}\wedge\dots\wedge e_{i_k})_{0\le k\le n,\ 1\le i_1<\dots<i_k\le n}$ est une base de ΛE. Il existe donc un unique endomorphisme linéaire $*$ de ΛE, dépendant à priori du choix de $\mathscr{B}$, telle que pour tous les $k, i_1,\dots,i_\ell$, nous ayons

$$*(e_{i_1}\wedge\dots\wedge e_{i_k}) = \varepsilon_{\{i_1,\dots,i_k\}}\ e_1\wedge\dots\wedge\widehat{e_{i_1}}\wedge\dots\wedge\widehat{e_{i_k}}\wedge\dots\wedge e_{i_k}$$

(avec la convention que $e_{i_1}\wedge\dots\wedge e_{i_k}=\mathbf{1}$ si $k=0$, et que la permutation $\{i_1,\dots,i_k\}$ est l'identité si $k=0$). Par construction et linéarité, nous avons $*(\Lambda^k E)\subset\Lambda^{n-k}E$ pour tout $k=0,\dots,n$. Nous avons $*\circ *=(-1)^{k(n-k)}\,\mathrm{id}$ sur $\Lambda^k E$, en remarquant que la signature de la permutation $(k+1,\dots,n,1,\dots,k)$ est égale à $(-1)^{k(n-k)}$, ainsi que celle de tous ses conjugués. De plus, les conditions demandées dans l'assertion (3) sont satisfaites pour la base $\mathscr{B}$, ainsi que pour toutes ses permutations qui restent directe, c'est-à-dire dont la signature est $+1$.[59] Montrons que $*$ ne dépend en fait pas du choix de $\mathscr{B}$.

59. Pour toute permutation $\sigma\in\mathscr{S}_n$, pour toute base $\mathscr{B}=(e_1,\dots,e_n)$ d'un espace vectoriel E de dimension n sur un corps commutatif, le déterminant de l'unique isomorphisme linéaire de E envoyant e_i sur $e_{\sigma^{-1}(i)}$ pour tout $i=1,\dots,n$ est égal à la signature $\epsilon(\sigma)$.

Soient $\mathscr{C} = (f_1, \dots, f_n)$ une autre base orthonormée directe de E, et P la matrice de passage de $\mathscr{B}$ à $\mathscr{C}$, qui est orthogonale et vérifie $\det P > 0$, donc appartient à $\mathrm{SO}(n)$.

Puisque la matrice de passage $\breve{P}$ de $\mathscr{B}^*$ à $\mathscr{C}^*$ est égale à $({}^tP)^{-1}$ par la question A (3), nous avons $\det(\breve{P}) = \frac{1}{\det P} = 1$. Par la question B (1) et la formule (22), nous avons donc

$$e_1 \wedge \cdots \wedge e_n = \det{}_{\mathscr{B}^*} = \det(\breve{P}) \ \det{}_{\mathscr{C}^*} = \det{}_{\mathscr{C}^*} = f_1 \wedge \cdots \wedge f_n \ .$$

Il suffit donc de montrer que pour tout $k \in \{1, \dots, n-1\}$, la restriction de $*$ à $\Lambda^k E$ ne dépend pas du choix de $\mathscr{B}$.

Si $n = 2$ et $P = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, alors $f_i = Pe_i$ pour $i = 1, 2$, donc

$$\begin{aligned} *(f_1) &= *((\cos\theta)\ e_1 + (\sin\theta)\ e_2) = (\cos\theta)\ *(e_1) + (\sin\theta)\ *(e_2) \\ &= (\cos\theta)\ e_2 + (\sin\theta)\ (-e_1) = f_2 \end{aligned}$$

et de même $*(f_2) = -f_1$.

Comme $\mathrm{SO}(n)$ est engendré par les rotations sur les plans de coordonnées de la base $\mathscr{B}$ et les permutations paires de la base $\mathscr{B}$ (voir l'exercice E.22), l'assertion (3) en découle.

(5) Ces affirmations découlent de la définition du produit mixte et du produit vectoriel.

2 Groupes orthogonaux définis positifs

Dans cette partie, le corps de base est $\mathbb{R}$. Soient $n \in \mathbb{N}$ et E un espace euclidien de dimension n, de forme quadratique, norme et produit scalaire notés respectivement q, $x \mapsto \|x\|$ et $(x, y) \mapsto \langle x, y\rangle$. Une référence globale pour cette partie est [Per1, Chap. VI].

2.1 Propriétés de transitivité

Appelons *suite orthonormée* dans E une suite finie $(x_1, \dots, x_k)$, où $k \in \mathbb{N}$, d'éléments de E, de norme 1 et deux à deux orthogonaux.

Proposition 2.1. *Pour tout $k = 0, \dots, n$, l'action diagonale de $\mathrm{O}(q)$ sur l'ensemble des suites orthonormées à k éléments de E est transitive*[60]*, et simplement transitive si $k = n$.*

Si $k < n$, l'action de $\mathrm{SO}(q)$ sur l'ensemble des suites orthonormées à k éléments de E est transitive.

Si E est orienté, l'action de $\mathrm{SO}(q)$ sur l'ensemble des bases orthonormées directes de E est simplement transitive.

Pour tout $k = 0, \dots, n$, le groupe $\mathrm{SO}(q)$ agit transitivement sur l'ensemble des sous-espaces vectoriels de E de dimension k.

En particulier, l'action de $\mathrm{O}(n)$ sur la sphère unité $\mathbb{S}_{n-1}$ de l'espace euclidien standard $\mathbb{R}^n$ est transitive, ainsi que celle de $\mathrm{SO}(n)$ si $n \geq 2$.

La dernière propriété de la proposition 2.1 est en général complètement fausse lorsque q est une forme quadratique quelconque sur un corps commutatif quelconque, voir par exemple [Per1, Chap. VIII].

Démonstration. La dernière affirmation découle de l'antépénultiemme, en prenant des bases orthonormées.

Toute suite orthonormée à k éléments se complète en une base orthonormée de E, en lui concaténant une base orthonormée de son orthogonal. Et si $k < n$, quitte à changer le dernier vecteur par son opposé, nous pouvons supposer que la base est dans n'importe quelle orientation choisie. Nous nous ramenons donc au cas $k = n$.

Il est immédiat que l'image d'une base orthonormée (respectivement base orthonormée positive) de E par un élément de $\mathrm{O}(q)$ (respectivement $\mathrm{SO}(q)$) est encore une base orthonormée (respectivement base orthonormée positive). Puisque l'action de $\mathrm{GL}(E)$ sur l'ensemble des bases de E est simplement transitive, et puisqu'être une isométrie de q se vérifie sur les couples de vecteurs de n'importe quelle base de E, l'action de $\mathrm{O}(q)$ sur l'ensemble des bases orthonormées est simplement transitive. Comme un élément de $\mathrm{O}(q)$ appartient à $\mathrm{SO}(q)$ si et seulement s'il envoie une base orthonormée sur une base orthonormée dans la même orientation, le résultat en découle. □

60. Rappelons qu'une action d'un groupe G sur un ensemble X est *transitive* si

$$\forall\, x, y \in X,\ \exists\, g \in G \quad y = gx$$

et *simplement transitive* si

$$\forall\, x, y \in X,\ \exists!\, g \in G \quad y = gx\ .$$

2.2 Centre et partie génératrice des groupes orthogonaux

L'énoncé suivant regroupe les résultats sur le calcul du centre[61] et d'une partie génératrice de $\mathrm{O}(q)$ et $\mathrm{SO}(q)$.

Théorème 2.2. (1) *Le centre de* $\mathrm{O}(q)$ *est* $Z(\mathrm{O}(q)) = \{-\mathrm{id}, +\mathrm{id}\}$.

(2) *Si* $n \geq 3$, *le centre de* $\mathrm{SO}(q)$ *est* $\mathrm{SO}(q) \cap \{-\mathrm{id}, +\mathrm{id}\}$.

(3) *Tout élément de* $\mathrm{O}(q)$ *est produit d'au plus* n *réflexions orthogonales.*

(4) *Tout élément de* $\mathrm{SO}(q)$ *est produit d'au plus* n *renversements orthogonaux.*

En particulier, $\mathrm{O}(q)$ n'est pas commutatif si $n \geq 2$, car il contient d'autres éléments que $-\mathrm{id}, +\mathrm{id}$.

Il est facile de vérifier, en utilisant le fait que $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SO}(2)$ si et seulement si $\det A = 1$ et ${}^tA = A^{-1}$, c'est-à-dire $A = \begin{pmatrix} a & -c \\ c & a \end{pmatrix}$ avec $a^2 + c^2 = 1$, que l'application de $\mathbb{R}/2\pi\mathbb{Z}$ dans $\mathrm{SO}(2)$ qui envoie $\theta \mod 2\pi$ sur $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ est un isomorphisme de groupes (et un homéomorphisme pour les topologies évidentes). Donc $\mathrm{SO}(2)$ est abélien, et en particulier égal à son centre.

Si $n \geq 3$, nous avons donc $Z(\mathrm{SO}(q)) = \{\mathrm{id}\}$ si n est impair et $Z(\mathrm{SO}(q)) = \{-\mathrm{id}, +\mathrm{id}\}$ si n est pair.

Le groupe $\mathrm{O}(q)$ est donc engendré par l'ensemble des réflexions orthogonales.

Le groupe $\mathrm{SO}(q)$ est donc engendré par l'ensemble des renversements orthogonaux.

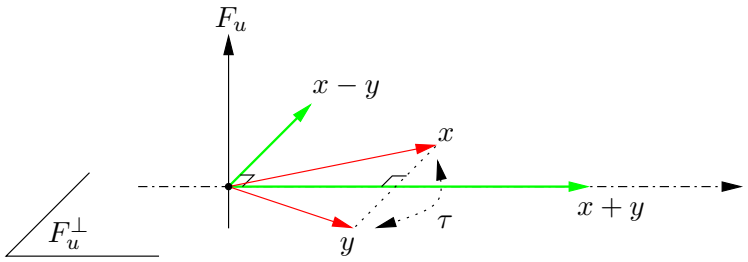
Remarque. Ce théorème est encore valable si k est un corps quelconque de caractéristique différente de 2 et si q est une forme quadratique non dégénérée quelconque, sauf le point (1) quand $k = \mathbb{F}_3$ est un corps fini à 3 éléments, $n = 2$ et q a pour matrice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ dans au moins une base de E, voir [Per1, Chap. VIII].

Démonstration. (1) Il est immédiat que $\{-\mathrm{id}, +\mathrm{id}\} \subset Z(\mathrm{O}(q))$. Soit $u \in Z(\mathrm{O}(q))$. Pour toute droite vectorielle D de E, puisque $u \circ s_D \circ u^{-1} = s_{u(D)}$ par la proposition 1.9, nous avons $u(D) = D$. Un automorphisme linéaire qui préserve toutes les droites étant une homothétie, et le déterminant d'un élément de $\mathrm{O}(q)$ valant ± 1, nous avons $u \in \{-\mathrm{id}, +\mathrm{id}\}$.

(2) Le même argument que ci-dessus en remplaçant réflexions orthogonales par renversements orthogonaux montre que tout élément u du centre de $\mathrm{SO}(q)$ préserve tout plan vectoriel de E. Si $n \geq 3$, alors toute droite vectorielle est l'intersection de deux plans vectoriels, donc u préserve toutes les droites vectorielles, et nous concluons comme ci-dessus.

(3) Soit $u \in \mathrm{O}(q)$. Notons $F_u = \{x \in E \ : \ u(x) = x\}$ le sous-espace vectoriel des points fixes de u. Montrons par récurrence sur la codimension $p_u = n - \dim F_u$ de F_u que u est une composition d'au plus p_u réflexions orthogonales. Si $p_u = 0$, alors $u = \mathrm{id}$ et la convention sur les produits vides conclut. Sinon, soient $x \in F_u^\perp - \{0\}$ et $y = u(x)$. Alors $y \neq x$ et $y \in F_u^\perp$ car u est orthogonale et F_u est stable par u (voir la remarque (1) de la partie 1.5).

61. Rappelons que le *centre* d'un groupe G est son sous-groupe distingué (formé des éléments qui commutent avec tous les autres)
$$Z(G) = \{g \in G \ : \ \forall\, h \in G, \ \ gh = hg\}\,.$$



Notons τ la symétrie orthogonale par rapport à l'hyperplan orthogonal à $x-y$. Cet hyperplan contient F_u car $x, y \in F_u^\perp$, et il contient $x+y$ car $\langle x-y, x+y\rangle = \|x\|^2 - \|y\|^2 = 0$. La réflexion orthogonale τ envoie y sur x, en prenant la différence des deux égalités $\tau(x+y) = x+y$ et $\tau(x-y) = -(x-y)$. Donc $\tau \circ u$ fixe x. D'où $F_{\tau\circ u}$, qui contient F_u et $x \notin F_u$, est de codimension $p_{\tau\circ u}$ strictement inférieure à celle de F_u. Par récurrence, il existe donc des réflexions orthogonales $\rho_1, \dots, \rho_k$ avec $k \leq p_{\tau\circ u}$ telles que $\tau\circ u = \rho_1\circ\cdots\circ\rho_k$. Donc $u = \tau \circ \rho_1 \circ \cdots \circ \rho_k$ est un produit d'au plus $p_{\tau\circ u} + 1 \leq p_u$ réflexions orthogonales.

Puisque $p_u \leq n$, le résultat en découle.

(4) Le lemme suivant va nous permettre de passer des réflexions à des renversements.

Lemme 2.3. *Si $n \geq 3$, et si ρ_1 et ρ_2 sont des réflexions orthogonales, alors il existe des renversements orthogonaux σ_1 et σ_2 tels que $\rho_1 \circ \rho_2 = \sigma_1 \circ \sigma_2$.*

Démonstration. Si $n = 3$, le résultat est immédiat, car $\sigma_i = -\rho_i$ est un renversement orthogonal pour $i = 1, 2$. Soient H_1 et H_2 les hyperplans fixes de ρ_1 et ρ_2. Soit V un sous-espace vectoriel de $H_1 \cap H_2$ de dimension $n-3$. Puisque $\rho_1 \circ \rho_2$ vaut l'identité sur V, il préserve son orthogonal $V^\perp$, qui est de dimension 3. Il existe donc des renversements orthogonaux σ_1 et σ_2 de $V^\perp$ tels que $(\rho_1 \circ \rho_2)\,|_{V^\perp} = \sigma_1 \circ \sigma_2$. En prolongeant σ_1 et σ_2 par l'identité sur V et par linéarité à E, le résultat en découle. □

Démontrons maintenant l'assertion (4). Si $u \in \mathrm{SO}(q)$, alors puisque le déterminant d'une réflexion orthogonale est -1, il existe des réflexions orthogonales $\rho_1, \rho_2, \dots, \rho_{2k-1}, \rho_{2k}$ avec $2k \leq n$ telles que $u = \rho_1 \circ \rho_2 \circ \cdots \circ \rho_{2k-1} \circ \rho_{2k}$. Le lemme précédant conclut. □

2.3 Classification à conjugaison près des transformations orthogonales

Le but de cette partie est de donner une « forme normale » pour tout élément de $\mathrm{O}(q)$.

Supposons tout d'abord que E est de dimension 2. Tout élément de $\mathrm{O}(q)$ est produit d'au plus 2 réflexions orthogonales d'après la partie précédente, et toute réflexion orthogonale de E est de déterminant -1. Donc tout élément de $\mathrm{O}(q) - \mathrm{SO}(q)$ est une réflexion orthogonale, c'est-à-dire une symétrie orthogonale par rapport à une droite vectorielle D de E. Dans une base orthonormée adaptée à la décomposition en somme directe orthogonale $E = D \overset{\perp}{\oplus} D^\perp$, la matrice de u est donc

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

Puisque $\mathrm{O}(q)$ agit transitivement sur l'ensemble des droites vectorielles de E, deux éléments de $\mathrm{O}(q) - \mathrm{SO}(q)$ sont conjugués.

Soit maintenant $u \in \mathrm{SO}(2)$. Si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est la matrice de u dans une base orthonormée de E, alors l'équation $A^{-1} = {}^tA$ équivaut à $c = -b$ et $d = a$ puisque $\det A = 1$. Et l'équation $\det A = 1$ devient alors à $a^2 + b^2 = 1$.

Ceci montre au passage que le groupe $\mathrm{SO}(2)$ est isomorphe au groupe multiplicatif $\mathbb{U} = (\mathbb{S}_1, \times)$ des nombres complexes de module 1 par (l'inverse de) $a + ib \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, lui-même isomorphe à $\mathbb{R}/\mathbb{Z}$ par (l'inverse de) $t \mod \mathbb{Z} \mapsto e^{2i\pi t}$.

Il existe donc un unique $\theta \in \mathbb{R}/2\pi\mathbb{Z}$ tel que $a = \cos\theta$ et $b = \sin\theta$, de sorte que la matrice de u est

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} .$$

Cet élément θ ne dépend pas au signe près du choix de la base orthonormée, car les matrices de u dans deux bases orthonormées sont conjuguées [62] et les valeurs propres complexes de A sont $e^{-i\theta}, e^{i\theta}$. Il est appelé l'*angle* de la rotation u. Deux éléments de $\mathrm{O}(q)$ sont conjugués dans $\mathrm{O}(q)$ si et seulement si ou bien ils sont tous les deux de déterminant -1, ou bien ils sont tous les deux des rotations de même angle modulo signe. Deux éléments du groupe abélien $\mathrm{SO}(q)$ sont conjugués dans $\mathrm{SO}(q)$ si et seulement s'ils sont égaux.

Le résultat suivant généralise ceci en toute dimension.

Théorème 2.4. *Pour tout élément $u \in \mathrm{O}(q)$, il existe une suite $(p_{-1}, p_1, \theta_1, \dots, \theta_r)$ où $p_{-1}, p_1, r \in \mathbb{N}$ vérifient $p_{-1} + p_1 + 2r = n$ et $\theta_1, \dots, \theta_r \in (\mathbb{R} - \pi\mathbb{Z})/2\pi\mathbb{Z}$, cette suite étant unique modulo permutation et changement de signe de $\theta_1, \dots, \theta_r$, telle qu'il existe une décomposition en somme directe orthogonale*

$$E = E_{-1} \overset{\perp}{\oplus} E_1 \overset{\perp}{\oplus} P_1 \overset{\perp}{\oplus} \dots \overset{\perp}{\oplus} P_r$$

avec $E_{\pm 1}$ l'espace propre de u pour la valeur propre ± 1 de dimension $p_{\pm 1}$, et $P_1 \dots, P_r$ des plans vectoriels invariants par u, tels que $u|_{P_i}$ soit la rotation plane d'angle θ_i. Dans une base orthonormée adaptée à cette décomposition, la matrice de u est

$$D(p_{-1}, p_1, \theta_1, \dots, \theta_r) = \begin{pmatrix} -1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & -1 & & & & & & & & \\ & & & 1 & & & & & & & \\ & & & & \ddots & & & & & & \\ & & & & & 1 & & & & & \\ & & & & & & \cos\theta_1 & -\sin\theta_1 & & & \\ & & & & & & \sin\theta_1 & \cos\theta_1 & & & \\ & & & & & & & & \ddots & & \\ & & & & & & & & & \cos\theta_r & -\sin\theta_r \\ & & & & & & & & & \sin\theta_r & \cos\theta_r \end{pmatrix} .$$

Deux éléments de $\mathrm{O}(q)$ sont conjugués dans $\mathrm{O}(q)$ si et seulement s'ils ont la même suite $(p_{-1}, p_1, \theta_1, \dots, \theta_r)$ modulo permutation et changement de signe de $\theta_1, \dots, \theta_r$.

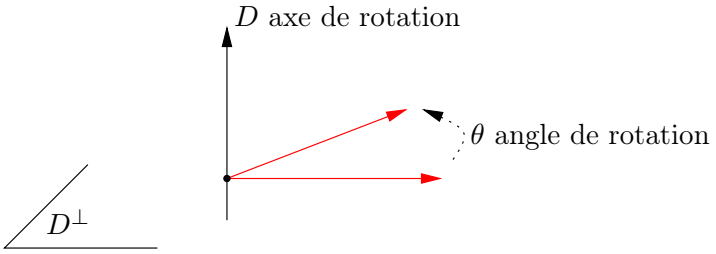
62. En effet, si $A' = {}^tPAP$ et $P^{-1} = {}^tP$, alors $A' = P^{-1}AP$.

Remarques. (1) La trace de cet élément $u \in \mathrm{O}(q)$ est

$$\mathtt{tr}\, u = -p_{-1} + p_1 + \sum_{i=1}^{r} 2\cos\theta_i \ .$$

(2) Un tel élément u appartient $\mathrm{SO}(q)$ si et seulement si p_{-1} est pair. Si n est impair et si $u \in \mathrm{SO}(q)$, alors nécessairement p_1 est non nul : u admet toujours au moins un vecteur fixe non nul.

(3) En dimension 3, tout élément non trivial u de $\mathrm{SO}(q)$ admet donc une et une seule droite invariante D, appelée son *axe de rotation*, et u induit une rotation plane non triviale sur le plan orthogonal $D^\perp$ à son axe de rotation, dont l'angle de rotation est appelé l'*angle de rotation* de u. Il est bien défini modulo 2π et changement de signe, et bien défini modulo 2π si on a fixé une orientation de l'axe de rotation et de l'espace ambiant, voir l'exercice E.21 ci-dessous.



(4) Fixons une orientation de E et $u \in \mathrm{SO}(q)$. Quitte à changer le signe d'un vecteur de base en appliquant le théorème 2.4 à u (qui appartient à $\mathrm{O}(q)$), il existe une base orthonormée directe $\mathscr{B} = (e_1, \ldots, e_n)$ de E dans laquelle la matrice de u est égale à la matrice $D(p_{-1}, p_1, \theta_1, \ldots, \theta_r)$, où $p_{-1}, p_1, r \in \mathbb{N}$ vérifient $p_{-1} + p_1 + 2r = n$ et $\theta_1, \ldots, \theta_r \in (\mathbb{R} - \pi\mathbb{Z})/2\pi\mathbb{Z}$. Une telle suite $(p_{-1}, p_1, \theta_1, \ldots, \theta_r)$ est uniquement déterminée modulo permutation et changement de signe de $\theta_1, \ldots, \theta_r$, avec changement d'un nombre pair de signes si $p_{-1} = p_1 = 0$. De plus, deux éléments de $\mathrm{SO}(q)$ sont conjugués dans $\mathrm{SO}(q)$ si et seulement s'ils ont la même suite $(p_{-1}, p_1, \theta_1, \ldots, \theta_r)$ modulo permutation et changement de signe de $\theta_1, \ldots, \theta_r$, avec changement d'un nombre pair de signes si $p_{-1} = p_1 = 0$.[63]

Démonstration. Le résultat découle du corollaire 1.17 (4) et de la description des éléments de $\mathrm{O}(2)$ précédant l'énoncé. Nous en donnons une démonstration résumée pour information.

63. En effet, soit $u' \in \mathrm{SO}(q)$ ayant pour matrice $D(p'_{-1}, p'_1, \theta'_1, \ldots, \theta'_r)$ dans une base orthonormée directe $\mathscr{B}' = (e'_1, \ldots, e'_n)$ de E.

Supposons que $p'_{-1} = p_{-1}$, $p'_1 = p_1$ (et donc $r = r'$), et qu'il existe une permutation $\sigma \in \mathscr{S}_r$ et des signes $(\epsilon_i)_{1\leq i\leq r} \in \{\pm 1\}^r$ tels que $\theta'_i = \epsilon_i\theta_i$ pour tout $i = 1, \ldots, r$ et $\prod_{i=1}^r \epsilon_i = 1$ si $p_{-1} = p_1 = 0$. L'unique application $v : E \to E$ qui est linéaire et envoie respectivement $e_1, e_2, \ldots, e_{p_{-1}+p_1}$ sur $\prod_{i=1}^r \epsilon_i\, e'_1, e'_2, \ldots, e'_{p_{-1}+p_1}$, ainsi que respectivement $e_{p_{-1}+p_1+2i}$ et $e_{p_{-1}+p_1+2i-1}$ sur $\epsilon_i e'_{p_{-1}+p_1+2i}$ et $e'_{p_{-1}+p_1+2i-1}$ pour tout $i = 1, \ldots, r$ conjuge u et u'. C'est un élément de $\mathrm{O}(q)$, car elle envoie une base orthonormée sur une base orthonormée. De plus, puisque les bases $\mathscr{B}$ et $\mathscr{B}'$ sont toutes les deux directes, et puisque $\prod_{i=1}^r \epsilon_i = 1$ si $p_{-1} = p_1 = 0$, son déterminant est égal à 1. Donc v est un élément de $\mathrm{SO}(q)$ qui conjuge u et u'.

Réciproquement, supposons que u' soit conjuguée à u dans $\mathrm{SO}(q)$ (ce qui est en particulier vérifié si $u' = u$). Alors les dimensions des espaces propres de valeurs propres -1 et $+1$ de u et u' sont égales. Par conséquent $p'_{-1} = p_{-1}$, $p'_1 = p_1$ (et donc $r = r'$). De plus, u et u' ont les mêmes valeurs propres non réelles à permutation près. Donc, comme l'échange de deux valeurs propres conjuguées $e^{i\theta}$ et $e^{-i\theta}$ revient à changer de signe θ, il existe une permutation $\sigma \in \mathscr{S}_r$ et des signes $(\epsilon_i)_{1\leq i\leq r} \in \{\pm 1\}^r$ tels que $\theta'_i = \epsilon_i\theta_i$ pour tout $i = 1, \ldots, r$. Enfin, un calcul de déterminant (et le fait que $\mathscr{B}$ et $\mathscr{B}'$ sont toutes les deux directes) montre que $\prod_{i=1}^r \epsilon_i = 1$ si $p_{-1} = p_1 = 0$.

Montrons par récurrence sur n l'existence d'une telle décomposition en somme directe orthogonale. Si $n = 1$, c'est immédiat, et le cas $n = 2$ a déjà été traité avant l'énoncé. Nous pouvons donc supposer que $n \geq 3$. En fixant une base orthonormée de E, nous pouvons supposer que E est l'espace euclidien usuel $\mathbb{R}^n$, voir le théorème 1.12 (2) et la formule (16).

Si u admet une valeur propre réelle λ, alors celle-ci vaut ± 1, car si x est un vecteur propre non nul associé, alors $\lambda^2\langle x, x\rangle = \langle u(x), u(x)\rangle = \langle x, x\rangle$, donc $\lambda^2 = 1$. En appliquant l'hypothèse de récurrence à l'hyperplan $x^\perp$ (qui est stable par u par la remarque (1) de la partie 1.5), le résultat en découle.

Sinon, soit λ une valeur propre complexe non réelle de u et $x \in \mathbb{C}^n$ un vecteur propre complexe non nul associé. Puisque u est réel, le vecteur $\overline{x}$ (dont les composantes sont les conjuguées des composantes de x) est vecteur propre de u pour la valeur propre $\overline{\lambda}$. Puisque $\lambda \neq \overline{\lambda}$, les vecteurs x et $\overline{x}$ sont linéairement indépendants sur $\mathbb{C}$. Les deux vecteurs $e_1 = x + \overline{x}$ et $e_2 = i(x - \overline{x})$ sont réels, et engendrent sur $\mathbb{C}$ le plan complexe $\mathbb{C}\,x + \mathbb{C}\,\overline{x}$, donc sont linéairement indépendants sur $\mathbb{R}$. Puisque

$$\begin{cases} u(e_1) = u(x) + u(\overline{x}) = \lambda\, x + \overline{\lambda}\,\overline{x} = (\mathrm{Re}\ \lambda)\, e_1 + (\mathrm{Im}\ \lambda)\, e_2 \\ u(e_2) = i\, u(x) - i\, u(\overline{x}) = i\,\lambda\, x - i\,\overline{\lambda}\,\overline{x} = -(\mathrm{Im}\ \lambda)\, e_1 + (\mathrm{Re}\ \lambda)\, e_2\ , \end{cases}$$

le plan vectoriel réel P dans $\mathbb{R}^n$ engendré sur $\mathbb{R}$ par e_1 et e_2 est invariant par u. Par le cas $n = 2$ et par application de l'hypothèse de récurrence sur $P^\perp$, le résultat en découle.

Les autres affirmations en découlent aisément. □

Exercice E.21. *Dans l'espace euclidien orienté usuel $\mathbb{R}^3$, pour tout vecteur unitaire $\vec{n}$ de $\mathbb{R}^3$ et pour tout $\psi \in \mathbb{R}$, notons $R_{\vec{n},\,\psi}$ l'élément de $\mathrm{SO}(3)$ d'axe de rotation la droite vectorielle engendrée par $\vec{n}$, induisant sur le plan vectoriel $\vec{n}^\perp$ orthogonal à $\vec{n}$ orienté*[64] *par $\vec{n}$ la rotation plane d'angle ψ.*

(1) *Montrer que*

$$R_{\vec{n},\,\psi} = R_{-\vec{n},\,-\psi}$$

et que pour tout $x \in \mathbb{R}^3$,

$$R_{\vec{n},\,\psi}(x) = x + (1 - \cos\psi)\ \vec{n} \wedge (\vec{n} \wedge x) + (\sin\psi)\ \vec{n} \wedge x\ . \tag{27}$$

(2) *Montrer que les rotations d'angle ψ autour des axes de coordonnées de $\mathbb{R}^3$, identifiées à leur matrice dans la base canonique (e_1, e_2, e_3) de $\mathbb{R}^3$, sont les éléments de $\mathrm{SO}(3)$ suivants*

$$R_{e_1,\,\psi} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & -\sin\psi \\ 0 & \sin\psi & \cos\psi \end{pmatrix}$$

$$R_{e_2,\,\psi} = \begin{pmatrix} \cos\psi & 0 & \sin\psi \\ 0 & 1 & 0 \\ -\sin\psi & 0 & \cos\psi \end{pmatrix}$$

$$R_{e_3,\,\psi} = \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}\ .$$

64. c'est-à-dire que (u, v) est une base orthonormée positive de $\vec{n}^\perp$ si et seulement si $(u, v, \vec{n})$ est une base orthonormée positive de $\mathbb{R}^3$

(3) *Notons*

$$\eta_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \eta_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \eta_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Montrer, en utilisant l'assertion (6) de la proposition 3.4, que pour tout vecteur unitaire $\vec{n} = (n_1, n_2, n_3) \in \mathbb{S}_2$ *et pour tout* $\psi \in \mathbb{R}$*, nous avons*

$$R_{\vec{n},\,\psi} = \exp\Big(\psi \sum_{k=1}^{3} n_k\, \eta_k\Big)\,.$$

2.4 Groupe dérivé du groupe orthogonal

Rappelons que

- le *groupe dérivé* $D(G) = [G, G]$ d'un groupe G est le sous-groupe de G engendré par les commutateurs $[a, b] = a\, b\, a^{-1}\, b^{-1}$,
- ses éléments, qui sont des produits de commutateurs car $[a, b]^{-1} = [b, a]$, ne sont pas forcément réduits à un seul commutateur,
- il est distingué dans G car $c\,[a, b]\,c^{-1} = [c\,a\,c^{-1}, c\,b\,c^{-1}]$, et
- le groupe quotient $G/[G, G]$ est abélien. C'est en fait le plus gros groupe quotient abélien de G, au sens que si G' est un groupe abélien, alors tout morphisme de groupes $f : G \to G'$ *factorise* par $G/[G, G]$. Ceci signifie qu'il existe un unique morphisme de groupes $\overline{f} : G/[G, G] \to G'$ tel que le diagrame suivant, dont la flèche verticale est la projection canonique $p : G \to G/[G, G]$, commute (c'est-à-dire que $f = \overline{f} \circ p$) :

$$\begin{array}{ccc} G & & \\ {\scriptstyle p}\downarrow & \searrow^{f} & \\ G/[G,G] & \xrightarrow[\overline{f}]{} & G'\,. \end{array}$$

Proposition 2.5.

(1) *Si* s *et* s' *sont des symétries orthogonales par rapports à des sous-espaces vectoriels de même dimension, alors* s *et* s' *sont conjuguées dans* $\mathrm{SO}(q)$*, c'est-à-dire qu'il existe* $u \in \mathrm{SO}(q)$ *tel que* $s' = u\, s\, u^{-1}$*.*

(2) *Si* $n \geq 2$*, le groupe dérivé de* $\mathrm{O}(q)$ *est* $\mathrm{SO}(q)$*.*

(3) *Si* $n \geq 3$*, le groupe* $\mathrm{SO}(q)$ *est* parfait*, c'est-à-dire égal à son groupe dérivé.*

Démonstration. (1) Si A et A' sont les sous-espaces vectoriels fixes de s et s' respectivement, par la dernière affirmation de la proposition 2.1, il existe $u \in \mathrm{SO}(q)$ tel que $u(A') = A$, et alors $u\, s\, u^{-1} = s'$ par la formule (17) dans la proposition 1.9.

(2) Puisque $\det(a\, b\, a^{-1}\, b^{-1}) = 1$, nous avons $[\mathrm{O}(q), \mathrm{O}(q)] \subset \mathrm{SO}(q)$. Réciproquement, si ρ, ρ' sont des réflexions orthogonales, alors par l'assertion (1), il existe $u \in \mathrm{SO}(q)$ tel que $\rho' = u\ \rho\ u^{-1}$, donc $\rho'\rho = \rho'\rho^{-1} = u\ \rho\ u^{-1}\rho^{-1} \in [\mathrm{O}(q), \mathrm{O}(q)]$. Comme tout élément de $\mathrm{SO}(q)$ est produit d'un nombre pair de réflexions orthogonales, l'inclusion réciproque en découle.

(3) L'inclusion $[\mathrm{SO}(q), \mathrm{SO}(q)] \subset \mathrm{SO}(q)$ est triviale. Réciproquement, puisque tout élément de $\mathrm{SO}(q)$ est produit de renversements, il suffit de montrer que tout renversement est un commutateur. Puisque tous les renversements sont conjugués par l'assertion (1), il suffit de montrer qu'il existe un renversement qui est un commutateur. Puisque $n \geq 3$, il existe une suite orthonormée (e_1, e_2, e_3) dans E. Pour $i = 1, 2, 3$, soit σ_i le renversement de sous-espace vectoriel fixe l'orthogonal du plan vectoriel $\mathrm{Vect}(\{e_1, e_2, e_3\} - \{e_i\})$. Par l'assertion (1), soit $u \in \mathrm{SO}(q)$ tel que $\sigma_1 = u\ \sigma_2\ u^{-1}$. Alors $\sigma_3 = \sigma_1\ \sigma_2 = \sigma_1\ \sigma_2^{-1} = u\ \sigma_2\ u^{-1}\ \sigma_2^{-1}$ est un commutateur. □

2.5 Simplicité des groupes spéciaux orthogonaux modulo leur centre

Rappelons que

- un groupe G est *simple* s'il est non trivial et s'il ne contient pas d'autre sous-groupe distingué que son sous-groupe trivial et lui-même ;
- les groupes $\mathbb{Z}/p\mathbb{Z}$ avec p premier sont les seuls groupes simples abéliens ;
- les groupes alternés $\mathfrak{A}_n$ (des permutations de $\{1, \ldots, n\}$ de signature 1) pour $n \geq 5$ sont simples, voir [Per1, §I.8] ;
- les groupes projectifs spéciaux linéaires $\mathrm{PSL}_n(k) = \mathrm{SL}_n(k)/Z(\mathrm{SL}_n(k))$ sont simples pour tout corps commutatif k et tout $n \geq 2$ sauf si $(n, k) = (2, \mathbb{F}_2), (2, \mathbb{F}_3)$, voir [Per1, §IV.4] ;
- un groupe simple non abélien est parfait, mais la réciproque n'est pas vraie. Par exemple, $\mathrm{SO}(6)$ est parfait par la proposition 2.5, mais n'est pas simple, car son centre $\{-\mathrm{id}, \mathrm{id}\}$ est non trivial.

Théorème 2.6. *Si $n \geq 3$ et $n \neq 4$, alors le groupe $\mathrm{PSO}(n) = \mathrm{SO}(n)/Z(\mathrm{SO}(n))$ est simple.*

Notons que si n est impair, alors le centre $Z(\mathrm{SO}(n))$ est trivial, donc le groupe $\mathrm{SO}(n)$ est simple. En particulier, $\mathrm{SO}(3)$ est simple, et la démonstration ci-dessous commence en fait par montrer ceci. Il existe plusieurs démonstrations possibles de la simplicité de $\mathrm{SO}(3)$. Nous utilisons par choix personnel de l'auteur (c'est beau la topologie !) celle ci-dessous (voir par exemple [FGN, page 67]). Voir par exemple [Per1, §VI.6] ou [Ber1, §8.5] pour une autre démonstration.

Remarque. Le groupe $\mathrm{PSO}(4)$ n'est pas simple, il est isomorphe à $\mathrm{SO}(3) \times \mathrm{SO}(3)$, voir par exemple [Per1, §VII.3].

Démonstration. Étape 1. Montrons que $\mathrm{SO}(3)$ est simple.

Soit G un sous-groupe distingué non trivial de $\mathrm{SO}(3)$. Montrons qu'il est égal à $\mathrm{SO}(3)$. Comme les renversements engendrent $\mathrm{SO}(3)$, et que deux renversements sont conjugués, il suffit de montrer que G contient un renversement.

Montrons tout d'abord que nous pouvons supposer que G est connexe par arcs. Soit G_0 la composante connexe par arcs de l'élément neutre e dans G.

Celle-ci est un sous-groupe de G, car si $g, h \in G_0$, si $\alpha, \beta : [0, 1] \to G$ sont des chemins continus dans G de e à respectivement g, h, alors $t \mapsto \alpha(t)(\beta(t))^{-1}$ est un chemin continu dans G de e à gh^{-1}. Le sous-groupe G_0 est distingué dans $\mathrm{SO}(3)$, car le conjugué par $h \in \mathrm{SO}(3)$ de tout chemin continu de g à e dans G est un chemin continu dans G (car G est distingué) de hgh^{-1} à $heh^{-1} = e$.

Supposons par l'absurde que G_0 soit trivial. Alors toutes les composantes connexes par arcs de G le sont, car l'action (par homéomorphismes) de G par translations à gauche sur

lui-même est transitive sur ses composantes connexes par arcs. Soient g un élément non trivial de G et A son axe de rotation. Soit h une rotation d'angle θ non nul modulo π, d'axe de rotation B orthogonal à A, donc ne préservant pas A. L'application de $[0,1]$ dans $\mathrm{SO}(3)$ qui à t associe la rotation h_t d'axe de rotation B et d'angle de rotation $t\theta$ est continue. Donc $t \mapsto h_t g (h_t)^{-1}$ est un chemin continu dans G (puisque G est distingué) de g à hgh^{-1}. Il est donc constant et $g = hgh^{-1}$. Mais l'axe de rotation de hgh^{-1}, qui est $h(A)$, n'est pas égal à A, une contradiction. Donc quitte à remplacer G par G_0, nous pouvons supposer que G est connexe par arcs.

Soit g un élément non trivial de G. Quitte à le remplacer par une puissance entière (puisque G est un sous-groupe), nous pouvons supposer que son angle de rotation θ, qui n'est pas nul, appartient à $[\frac{\pi}{2}, \frac{3\pi}{2}]$. Puisque G est connexe par arcs, il existe un chemin continu α de e à g dans G. L'application $f : t \mapsto \mathtt{tr}(\alpha(t)) - 1$ est continue, vaut $2 \geq 0$ en $t = 0$ et $2\cos\theta \leq 0$ en $t = 1$. Par le théorème des valeurs intermédiaires, il existe donc t_0 tel que $f(t_0) = 0$. Donc $\alpha(t_0) \in G$ est une rotation d'angle $\pm\frac{\pi}{2}$, et $\alpha(t_0)^2$ est un élément de G qui est un renversement. Ceci termine la démonstration de l'étape 1.

Étape 2. Pour $n \geq 5$, montrons que $\mathrm{PSO}(n)$ est simple.

Soit $\overline{G}$ un sous-groupe distingué non trivial de $\mathrm{PSO}(n)$. Notons G son image réciproque par la projection canonique $\mathrm{SO}(n) \to \mathrm{PSO}(n)$, qui est un sous-groupe distingué de $\mathrm{SO}(n)$ contenant proprement le centre de $\mathrm{SO}(n)$. Montrons que $G = \mathrm{SO}(n)$. Comme dans l'étape 1, il suffit de montrer que G contient un renversement.

Soit g_0 un élément de G différent de $\pm\,\mathrm{id}$. Nous allons le modifier par étapes successives jusqu'à obtenir un renversement dans G, en augmentant la dimension de son sous-espace vectoriel fixe.

Il existe au moins un plan vectoriel P qui n'est pas invariant par g_0, sinon g_0 préserverait toutes les droites vectorielles (qui sont intersections de deux plans vectoriels car $n \geq 3$), donc serait une homothétie, donc serait égale à $\pm\,\mathrm{id}$ car orthogonale. Soit τ le renversement par rapport à l'orthogonal de P. Alors $g_1 = g_0 \tau g_0^{-1} \tau^{-1}$ appartient à G (car G est un sous-groupe distingué) et est le produit de deux renversements τ^{-1} fixant $P^\perp$ et $g_0 \tau g_0^{-1}$ fixant $g_0(P^\perp) = g_0(P)^\perp$. Donc g_1 est non trivial car $P \neq g_0(P)$ et il laisse fixe l'intersection $P^\perp \cap g_0(P)^\perp$, qui est de dimension au moins $n-4$. Puisque $n \geq 5$, l'élément g_1 fixe au moins un vecteur non nul a, et en particulier n'est pas $-\,\mathrm{id}$.

Soit $b \in E$ un vecteur (non nul) tel que $g_1(b)$ et b soient non colinéaires, ce qui est possible car g_1 n'est pas une homothétie. Notons ρ_x la réflexion orthogonale par rapport à l'hyperplan orthogonal à un vecteur non nul x de E. Posons $g_2 = g_1(\rho_b\rho_a)g_1^{-1}(\rho_b\rho_a)^{-1}$, qui est un élément de G. Nous avons, puisque $g_1(a) = a$,

$$g_2 = (g_1\rho_b g_1^{-1})(g_1\rho_a g_1^{-1})\rho_a\rho_b = \rho_{g_1(b)}\rho_{g_1(a)}\rho_a\rho_b = \rho_{g_1(b)}\rho_b\ .$$

Puisque $g_1(b)$ et b sont non colinéaires, ce produit de deux réflexions orthogonales g_2 est non trivial, et l'ensemble de ses point fixes $F = g_1(b)^\perp \cap b^\perp$ est de dimension $n-2$.

Soit E' un sous-espace vectoriel de E de dimension 3 contenant le plan vectoriel $F^\perp$. Notons G' le sous-groupe de $\mathrm{SO}(E')$ formé des restrictions à E' des éléments de G valant l'identité sur $(E')^\perp$. Il est non trivial car il contient $g_2|_{E'}$, et il est distingué dans $\mathrm{SO}(E')$. Donc par l'étape 1, il est égal à $\mathrm{SO}(E')$. Il contient donc un renversement de E', qui, en le prolongeant par l'identité sur $(E')^\perp$, donne un renversement dans G. Ceci conclut la démonstration. □

2.6 Exercices

Exercice E.22. Soit E un espace euclidien de dimension $n \geq 2$, muni d'une base orthonormée $\mathscr{B}$. Montrer que $\mathrm{SO}(E)$ est engendré par les rotations sur les plans de coordonnées de la base $\mathscr{B}$ (valant l'identité sur l'orthogonal de ces plans). Montrer que $\mathrm{SO}(E)$ est engendré par les rotations sur le premier plan de coordonnées de la base $\mathscr{B}$ et les permutations paires de la base $\mathscr{B}$.

Exercice E.23. Soit E un espace euclidien de dimension $n \geq 2$.

(1) Montrer que le groupe spécial orthogonal $\mathrm{SO}(E)$ de E agit transitivement sur l'ensemble des *croix orthogonales*, c'est-à-dire des n-uplets de droites deux à deux orthogonales.

(2) Supposons E orienté, muni de son produit vectoriel. Montrer que

$$G = \left\{ u \in \mathrm{GL}(E) \ : \ \forall\, x_1, \ldots, x_{n-1} \in E, \quad u(x_1 \wedge \cdots \wedge x_{n-1}) = u(x_1) \wedge \cdots \wedge u(x_{n-1}) \right\} .$$

est un sous-groupe de $\mathrm{GL}(E)$ et que $G = \mathrm{SO}(E)$.

Exercice E.24. Montrer qu'une matrice antisymétrique réelle S ne peut avoir pour valeur propre -1. Montrer que $(I-S)(I+S)^{-1}$ est une matrice spéciale orthogonale. En déduire que $\mathrm{SO}(n)$ est connexe.

Exercice E.25. Dans l'espace euclidien usuel $\mathbb{R}^n$, identifions tout vecteur au vecteur colonne de ses coordonnées dans la base canonique. Soient A une matrice réelle inversible $n \times n$, dont toutes les valeurs propres sont réelles, notées $\lambda_1, \ldots, \lambda_n$ de sorte que

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n| > 0 \, .$$

(1) Montrer que

$$\lambda_1^2 \, \lambda_2^2 \leq \sup_{\|x\| \leq 1, \|y\| \leq 1} \begin{vmatrix} \langle Ax, Ax \rangle & \langle Ax, Ay \rangle \\ \langle Ax, Ay \rangle & \langle Ay, Ay \rangle \end{vmatrix}$$

(2) Soient $\mu_1, \ldots, \mu_n$ les racines carrées des valeurs propres de ${}^tAA$ de sorte que[65]

$$\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n > 0 \, .$$

Montrer que $|\, \lambda_1 \, \lambda_2 \,| \leq \mu_1 \, \mu_2$.

65. Notons que $\mu_1, \ldots, \mu_n$ sont les valeurs singulières de la matrice inversible A, voir la partie 1.10.

2.7 Indications pour la résolution des exercices

Correction de l'exercice E.21. (1) En écrivant $x \in \mathbb{R}^3$ comme somme $x = x' + x''$ d'un vecteur x' colinéaire à $\vec{n}$ et d'un vecteur x'' orthogonal à $\vec{n}$, nous avons par définition

$$R_{\vec{n},\,\psi}(x) = x' + (\cos\psi)\; x'' + (\sin\psi)\; \vec{n} \wedge x''$$

La première formule en découle.

En utilisant les relations $\vec{n} \wedge x' = 0$ et, pour tous les $a, b, c \in \mathbb{R}^3$,

$$a \wedge (b \wedge c) = \langle a, c\rangle\, b - \langle a, b\rangle\, c\ ,$$

(voir l'exercice E.13), de sorte que

$$\vec{n} \wedge (\vec{n} \wedge x) = \vec{n} \wedge (\vec{n} \wedge x'') = \langle \vec{n}, x''\rangle\, \vec{n} - \langle \vec{n}, \vec{n}\rangle\, x'' = -x''\ ,$$

la dernière formule (27) en découle.

(2) Cette assertion découle de la formule (27) en prenant pour $(\vec{n}, x)$ les divers couples d'éléments de $\{e_1, e_2, e_3\}$.

(3) Par l'assertion (1), l'application linéaire $\frac{d}{d\psi}_{|\psi=0} R_{\vec{n},\,\psi}$ est égale à $x \mapsto \vec{n} \wedge x$. Un petit calcul [66] montre que cette application a pour matrice $\sum_{k=1}^{3} n_k\; \eta_k$. Puisque $\psi \mapsto R_{\vec{n},\,\psi}$ est un sous-groupe à un paramètre de $\mathrm{GL}_3(\mathbb{R})$, la proposition 3.4 (6) montre que

$$R_{\vec{n},\,\psi} = \exp\Big(\psi \sum_{k=1}^{3} n_k\; \eta_k\Big)\ .$$

Ceci montre le résultat. [67] □

Correction de l'exercice E.22. Soit $u \in \mathrm{SO}(E)$, montrons qu'il est produit de rotations sur les plans de coordonnées de la base $\mathscr{B}$. Raisonnons par récurrence sur $n \geq 2$, le résultat étant immédiat pour $n = 2$.

Notons $\mathscr{B} = (e_1, \ldots, e_n)$. Soient $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tels que $u(e_1) = \sum_{i=1}^{n} \lambda_i\; e_i$. Soit $k \in \{1, \ldots, n\}$ tel que $\lambda_k \neq 0$ et $\lambda_i = 0$ si $k < i \leq n$.

Montrons par récurrence sur k que nous pouvons supposer que $k = 1$ quitte à post-composer u par des rotations sur des plans de coordonnées. En effet, si $k \geq 2$, soit r une

---

66. Rappelons (voir l'exercice E.13) que $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \wedge \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} \begin{vmatrix} y_2 & z_2 \\ y_3 & z_3 \end{vmatrix} \\ \begin{vmatrix} y_3 & z_3 \\ y_1 & z_1 \end{vmatrix} \\ \begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} y_2 z_3 - y_3 z_2 \\ y_3 z_1 - y_1 z_3 \\ y_1 z_2 - y_2 z_1 \end{pmatrix}$ pour tous les $y_1, y_2, y_3, z_1, z_2, z_3 \in \mathbb{R}$.

67. Il possible d'utiliser les matrices $J_1 = i\,\eta_1, J_2 = i\,\eta_2, J_3 = i\,\eta_3$, qui sont hermitiennes, au lieu des matrices η_1, η_2, η_3. Cette utilisation est due aux physiciens, qui préfèrent travailler avec des matrices hermitiennes plutôt qu'avec des matrices antihermitiennes. Ceci fait aussi apparaître une analogie avec l'expression des rotations dans le plan euclidien orienté usuel, qui sont les applications linéaires $z \mapsto e^{i\theta}\, z$ de multiplication par $e^{i\theta}$ pour $\theta \in \mathbb{R}$: nous avons

$$R_{\vec{n},\,\psi} = \exp\Big(-i\,\psi \sum_{k=1}^{3} n_k\; J_k\Big)\ .$$

rotation du plan de coordonnées $\mathrm{Vect}\{e_{k-1}, e_k\}$ telle que $r(\lambda_{k-1}\ e_{k-1} + \lambda_k\ e_k) \in \mathbb{R}e_{k-1}$, qui existe puisque les rotations d'un plan euclidien agissent transitivement sur les droites vectorielles de ce plan. Alors

$$r \circ u(e_1) = \sum_{i=1}^{k-2} \lambda_i\ e_i + r(\lambda_{k-1}\ e_{k-1} + \lambda_k\ e_k) \in \mathrm{Vect}\{e_1, \dots, e_{k-1}\}\ ,$$

Ceci conclut par récurrence.

Donc $u(e_1) = \lambda_1\ e_1$ et puisque u préserve la norme, nous avons $\lambda_1 = \pm 1$. Quitte à postcomposer par la rotation d'angle π sur le premier plan de coordonnée, nous pouvons donc supposer que $u(e_1) = e_1$. Comme u est orthogonal, il préserve alors l'hyperplan $e_1^\perp$ orthogonal à e_1 (voir la remarque (1) de la partie 1.5), et la restriction v de u à $e_1^\perp$ appartient à $\mathrm{SO}(e_1^\perp)$. Par récurrence, v est une composition de rotations sur les plans de coordonnées de la base $(e_2, \dots, e_n)$ de $e_1^\perp$. En les étendant par l'identité sur $\mathbb{R}\,e_1$, nous obtenons donc que u est une composition de rotations sur les plans de coordonnées de $\mathscr{B}$.

La dernière affirmation découle du fait que le groupe des permutations paires de $\mathscr{B}$ agit transitivement sur les plans de coordonnées, et donc que toute rotation dans un plan de coordonnées est conjugué par une permutation paire de $\mathscr{B}$ à une rotation sur le premier plan de coordonnées de $\mathscr{B}$.

Correction de l'exercice E.23. (1) Puisque le groupe orthogonal $\mathrm{O}(E)$ agit transitivement sur l'ensemble des bases orthonormées, il agit transitivement sur l'ensemble des croix orthogonales. Comme il est possible de changer un vecteur de base par son opposé, l'action de $\mathrm{SO}(E)$ est aussi transitive.

(2) Il est immédiat de vérifier que G est un sous-groupe de $\mathrm{GL}(E)$.

Soit $u \in \mathrm{SO}(E)$. Puisque $\det u = 1$, nous avons, pour tous les $x_1, \dots, x_{n-1}, z \in E$,

$$\begin{aligned}\langle u(x_1) \wedge \cdots \wedge u(x_{n-1}), z\rangle &= \det(u(x_1), \dots, u(x_{n-1}), z) = (\det u)\det(x_1, \dots, x_{n-1}, u^{-1}(z)) \\ &= \langle x_1 \wedge \cdots \wedge x_{n-1}, u^{-1}(z)\rangle = \langle u(x_1 \wedge \cdots \wedge x_{n-1}), z\rangle\ .\end{aligned}$$

Puisque ceci est vrai pour tout z et par la définition du produit vectoriel (voir la formule (23)), nous avons donc $u(x_1) \wedge \cdots \wedge u(x_{n-1}) = u(x_1 \wedge \cdots \wedge x_{n-1})$ pour tous les éléments $x_1, \dots, x_{n-1} \in E$, et $u \in \mathrm{SO}(E)$.

Réciproquement, soit $u \in G$. Pour toute droite (vectorielle) D de E, puisque u appartient à $\mathrm{GL}(E)$, nous avons que $u(D)$ et $u(D^\perp)^\perp$ sont des droites de E. De plus, tout élément non nul de $u(D)$ s'écrit comme $u(x_1 \wedge \cdots \wedge x_{n-1})$ avec $x_1, \dots, x_{n-1}$ des éléments linéairement indépendants de l'hyperplan $D^\perp$. Donc il s'écrit de la forme $u(x_1) \wedge \cdots \wedge u(x_{n-1})$ avec $u(x_1), \dots, u(x_{n-1})$ des éléments de $u(D^\perp)$. Par conséquent (voir la proposition 1.26 (iii)), $u(D)$ est contenu dans $(u(D^\perp))^\perp$, avec égalité par argument de dimension. Fixons-nous une croix orthogonale C_0. Comme toute droite d'une croix orthogonale est l'intersection des $n-1$ hyperplans orthogonaux aux autres droites de cette croix, le groupe G envoie la croix orthogonale C_0 sur une autre croix orthogonale. Puisque le groupe $\mathrm{SO}(E)$ agit transitivement sur les croix orthogonales par l'assertion (1), et puisque G est un groupe qui contient $\mathrm{SO}(E)$, quitte à précomposer u par un élément de $\mathrm{SO}(E)$, nous pouvons supposer que u fixe la croix orthogonale C_0. Dans une base orthonormée positive $(e_1, \dots, e_n)$ de vecteurs de la croix orthogonale C_0, la matrice de u est donc diagonale. Soient $\lambda_1, \dots, \lambda_n \in \mathbb{R} - \{0\}$ ses coefficients diagonaux. Nous avons l'égalité $e_1 \wedge \cdots \wedge e_{n-1} = e_n$ (voir la proposition

1.26 (iv)), ainsi que ses permutations cycliques. Nous avons donc

$$\begin{aligned}\lambda_n e_n = u(e_n) = u(e_1 \wedge \cdots \wedge e_{n-1}) = u(e_1) \wedge \cdots \wedge u(e_{n-1}) \\ = (\lambda_1 \dots \lambda_{n-1})\, e_1 \wedge \cdots \wedge e_{n-1} = (\lambda_1 \dots \lambda_{n-1}) e_n \,.\end{aligned}$$

Donc $\lambda_1 \dots \lambda_{n-1} = \lambda_n$ et de même par permutations cycliques. En prenant les rapports des équations consécutives, ceci montre que $\lambda_1^2 = \cdots = \lambda_n^2$, donc que les λ_i sont des racines de l'unité. Puisqu'elles sont réelles, elles ne peuvent être que ± 1. Donc $u \in \mathrm{O}(E)$ et $\det u = \lambda_1 \dots \lambda_{n-1}\lambda_n = \lambda_n^2 = 1$. Ceci montre que $u \in \mathrm{SO}(E)$.

3 Groupes unitaires définis positifs

Dans cette partie, le corps de base est $\mathbb{C}$, muni de son automorphisme involutif de corps $z \mapsto \overline{z}$. Soient $n \in \mathbb{N}$ et E un espace hermitien de dimension n, de forme quadratique hermitienne (de signature $(0, n)$), de norme et de produit scalaire notés respectivement q, $x \mapsto \|x\|$ et $(x, y) \mapsto \langle x, y\rangle$. Une référence générale pour cette partie est [MT].

3.1 Propriétés de transitivité

Appelons *suite orthonormée* dans E une suite finie $(x_1, \ldots, x_k)$, où $k \in \mathbb{N}$, d'éléments de E, de norme 1 et deux à deux orthogonaux.

Proposition 3.1. *Pour tout $k = 0, \ldots, n$, l'action diagonale de $\mathrm{U}(q)$ sur l'ensemble des suites orthonormées à k éléments de E est transitive, et simplement transitive si $k = n$.*

Si $k < n$, l'action de $\mathrm{SU}(q)$ sur l'ensemble des suites orthonormées à k éléments de E est transitive.

Pour tout $k = 0, \ldots, n$, le groupe $\mathrm{SU}(q)$ agit transitivement sur l'ensemble des sous-espaces vectoriels de E de dimension k.

En particulier, l'action de $\mathrm{U}(n)$ sur la sphère unité

$$\mathbb{S}_{2n-1} = \{(z_1, \ldots, z_n) \in \mathbb{C}^n : |z_1|^2 + \ldots |z_n|^2 = 1\}$$

de l'espace hermitien standard $\mathbb{C}^n$ est transitive, ainsi que celle de $\mathrm{SU}(n)$ si $n \geq 2$. Par contre $\mathrm{SU}(1)$ est trivial, donc n'agit pas transitivement sur le cercle $\mathbb{S}_1$.

Démonstration. La dernière affirmation lorsque $k < n$ découle de l'avant-dernière, en prenant des bases orthonormées. Elle est immédiate lorsque $k = n$.

Il est immédiat que l'image d'une suite orthonormée de E par un élément de $\mathrm{U}(n)$ est encore une suite orthonormée. Puisque l'action de $\mathrm{GL}(E)$ sur l'ensemble des bases de E est simplement transitive, et puisqu'être une isométrie de q se vérifie sur les couples de vecteurs de n'importe quelle base de E, l'action de $\mathrm{U}(q)$ sur l'ensemble des bases orthonormées est simplement transitive.

Toute suite orthonormée à k éléments se complète en une base orthonormée de E, en lui concaténant une base orthonormée de son orthogonal. Donc l'action de $\mathrm{U}(q)$ sur l'ensemble des suites orthonormées à k éléments est transitive. Si $k < n$, il est possible de modifier l'action d'un élément $g \in \mathrm{U}(q)$ sur le dernier élément d'une base orthonormée par un nombre complexe de module 1 pour obtenir un élément de $\mathrm{SU}(q)$ qui coïncide avec g sur les $n-1$ premiers éléments. Ceci montre que l'action de $\mathrm{SU}(q)$ sur l'ensemble des suites orthonormées à k éléments est transitive si $k < n$. □

3.2 Centre et partie génératrice des groupes unitaires

Notons $\mathbb{U} = (\mathbb{S}_1, \times)$ le groupe multiplicatif des nombres complexes de module 1, et

$$\mu_n = \{e^{2i\pi k/n} \ : \ k = 0, \ldots, n-1\}$$

le sous-groupe (fini et isomorphe à $\mathbb{Z}/n\mathbb{Z}$) de $\mathbb{U}$ des racines n-èmes de l'unité.

Pour tout $\zeta \in \mathbb{U}$ et pour toute droite vectorielle D de E, il existe un et un seul élément $s_{D,\zeta}$ de $\mathrm{U}(q)$ tel que la restriction de $s_{D,\zeta}$ à l'orthogonal de D soit l'identité, et la restriction

de $s_{D,\zeta}$ à D soit $\zeta\,\mathrm{id}_D$. Un tel élément est appelé une *réflexion complexe*[68]. Elle est *triviale* (c'est-à-dire égale à l'identité) si et seulement si $\zeta = 1$. Son déterminant est égal à ζ. Notons que l'inverse d'une réflexion complexe est encore une réflexion complexe :

$$s_{D,\zeta}^{-1} = s_{D,\zeta^{-1}} .$$

Exercice E.26. *Montrer qu'une réflexion complexe non triviale de E est une application de la forme $\rho : x \mapsto x - \check{r}(x)r$ où $\check{r}$ est une forme linéaire non nulle et r un vecteur non nul tel que $1 - \check{r}(r)$ soit un élément de $\mathbb{U} - \{1\}$. Le couple $(r, \check{r})$, bien déterminé modulo multiplication du premier élément par un scalaire non nul et division du second élément par le même scalaire, est appelé le couple des* racines et coracines *de ρ.*

Théorème 3.2. (1) *Le centre de* $\mathrm{U}(q)$ *est* $Z(\mathrm{U}(q)) = \{\lambda\,\mathrm{id} \ : \ \lambda \in \mathbb{U}\}$.

(2) *Le centre de* $\mathrm{SU}(q)$ *est* $Z(\mathrm{SU}(q)) = \{\lambda\,\mathrm{id} \ : \ \lambda \in \mu_n\}$.

(3) *Tout élément de* $\mathrm{U}(q)$ *est produit d'au plus* $2n$ *réflexions complexes.*

Démonstration. (1) L'inclusion $\{\lambda\,\mathrm{id} \ : \ \lambda \in \mathbb{U}\} \subset Z(\mathrm{U}(q))$ est immédiate, car nous avons $(\lambda\,\mathrm{id})^*(\lambda\,\mathrm{id}) = |\lambda|^2\,\mathrm{id} = \mathrm{id}$ si $\lambda \in \mathbb{U}$.

Soit $u \in Z(\mathrm{U}(q))$. Pour toute droite vectorielle D de E, notons $s_D = s_{D,-1}$ l'unique endomorphisme linéaire de E valant id sur l'orthogonal de D, et $-\mathrm{id}$ sur D. Alors nous avons $s_D \in \mathrm{U}(q)$, et par unicité, nous avons $v \circ s_D \circ v^{-1} = s_{v(D)}$ pour tout $v \in \mathrm{U}(q)$. Donc $u(D) = D$. Un automorphisme linéaire qui préserve toutes les droites étant une homothétie, et le déterminant d'un élément de $\mathrm{U}(q)$ étant de module 1 (voir la ligne qui suit la formule (15)), nous avons $u \in \{\lambda\,\mathrm{id} \ : \ \lambda \in \mathbb{U}\}$.

(2) Montrons que $Z(\mathrm{SU}(q)) = Z(\mathrm{U}(q)) \cap \mathrm{SU}(q)$. Comme nous avons $Z(\mathrm{U}(q)) \cap \mathrm{SU}(q) = \{\lambda\,\mathrm{id} \ : \ \lambda \in \mu_n\}$ par l'assertion (1), le résultat en découle.

L'inclusion $Z(\mathrm{U}(q)) \cap \mathrm{SU}(q) \subset Z(\mathrm{SU}(q))$ est immédiate. Si $v \in \mathrm{U}(q)$, soit $\lambda \in \mathbb{U}$ une racine n-ème de son déterminant (qui est non nul). Alors $v = (\lambda\,\mathrm{id})v'$ avec $v' = \frac{1}{\lambda}v \in \mathrm{SU}(q)$. Donc un élément qui commute avec tous les éléments de $\mathrm{SU}(q)$, puisqu'il commute aussi avec les homothéties, commute avec tous les éléments de $\mathrm{U}(q)$, ce qu'il fallait démontrer.

(3) Soit $u \in \mathrm{U}(q)$. Notons $F_u = \{x \in E \ : \ u(x) = x\}$ le sous-espace vectoriel des points fixes de u. Montrons par récurrence sur la codimension $p_u = n - \dim F_u$ de F_u que u est une composition d'au plus $2p_u$ réflexions complexes. Comme $p_u \leq n$, cela montre le résultat.

Si $p_u = 0$, alors $u = \mathrm{id}$ et la convention sur les produits vides conclut. Sinon, soient $x \in F_u^\perp - \{0\}$ et $y = u(x)$. Alors $y \in F_u^\perp$ car u est unitaire et F_u est stable par u (voir la remarque (1) de la partie 1.5).

Notons $\zeta \in \mathbb{U}$ un élément tel que $\langle x, y\rangle = \zeta|\langle x, y\rangle|$. Il est unique si x et y ne sont pas orthogonaux, et n'importe quel élément de $\mathbb{U}$ convient sinon.

Si $x = \zeta y$, alors en notant τ la réflexion complexe $s_{\mathbb{C}x,\zeta}$ fixant l'hyperplan orthogonal à x et valant ζ id sur la droite engendrée par x, nous avons $u \circ \tau(x) = u(\zeta x) = \zeta y = x$. Donc $F_{u\circ\tau}$ contient x et F_u car $x \in F_u^\perp$. Donc $p_{u\circ\tau} < p_u$. Par récurrence, il existe donc des réflexions complexes $\rho_1, \ldots, \rho_k$ avec $k \leq 2p_{\tau\circ u}$ telles que $u \circ \tau = \rho_1 \circ \cdots \circ \rho_k$. Donc $u = \rho_1 \circ \cdots \circ \rho_k \circ \tau^{-1}$ est un produit d'au plus $2p_{\tau\circ u} + 1 \leq 2p_u$ réflexions complexes.

68. La terminologie est litigieuse car c'est une involution si et seulement si $\zeta \in \{-1, +1\}$.

Sinon, $x-\zeta y \neq 0$. Notons $\tau' = s_{\mathbb{C}(x-\zeta y),-1}$ la réflexion complexe fixant l'hyperplan orthogonal à $x-\zeta y$ et valant $-\mathrm{id}$ sur la droite engendrée par $x-\zeta y$. Cet hyperplan contient F_u car $x, y \in F_u^\perp$, et il contient $x+\zeta y$ car

$$\langle x+\zeta y, x-\zeta y\rangle = \|x\|^2 + \zeta\,\overline{\langle x,y\rangle} - \overline{\zeta}\,\langle x,y\rangle - \|y\|^2 = |\zeta|^2\,|\langle x,y\rangle| - |\zeta|^2\,|\langle x,y\rangle| = 0\ .$$

La réflexion orthogonale τ' envoie y sur $\zeta^{-1}x$, en prenant la différence des deux égalités $\tau'(x+\zeta y) = x+\zeta y$ et $\tau'(x-\zeta y) = -(x-\zeta y)$. Donc $\tau'\circ u$ envoie x sur $\zeta^{-1}x$. Notons $\tau = s_{\mathbb{C}x,\zeta}$, qui envoie $\zeta^{-1}x$ sur x. D'où $F_{\tau\circ\tau'\circ u}$, qui contient F_u et $x \notin F_u$, est de codimension $p_{\tau\circ\tau'\circ u}$ strictement inférieure à celle de F_u. Par récurrence, il existe donc des réflexions orthogonales $\rho_1, \ldots, \rho_k$ avec $k \leq 2p_{\tau\circ u}$ telles que $\tau\circ\tau'\circ u = \rho_1\circ\cdots\circ\rho_k$. Donc $u = (\tau')^{-1}\circ\tau^{-1}\circ\rho_1\circ\cdots\circ\rho_k$ est un produit d'au plus $2p_{\tau\circ\tau'\circ u} + 2 \leq 2p_u$ réflexions orthogonales.

Puisque $p_u \leq n$, l'assertion (3) en découle. □

3.3 Classification à conjugaison près des transformations unitaires

Le résultat suivant est l'analogue du théorème 2.4. Il dit en particulier que tout élément de $\mathrm{U}(n)$ a ses valeurs propres de module 1 et est diagonalisable en base orthonormée. Ceci est un cas particulier du fait que $u \in \mathrm{U}(n)$ commute avec $u^* = u^{-1}$, donc est normal (voir le théorème 1.13).

Proposition 3.3. *Tout élément de* $\mathrm{U}(n)$ *est conjugué dans* $\mathrm{U}(n)$ *à un élément de la forme*

$$D(\theta_1, \cdots, \theta_n) = \begin{pmatrix} e^{i\theta_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & e^{i\theta_n} \end{pmatrix}$$

où $\theta_1, \cdots, \theta_n \in \mathbb{R}$.

Remarque. Un tel élément appartient à $\mathrm{SU}(n)$ si et seulement si $\theta_1+\cdots+\theta_n = 0 \mod 2\pi$. Deux éléments de $\mathrm{U}(n)$ (respectivement $\mathrm{SU}(n)$) sont conjugués dans $\mathrm{U}(n)$ (respectivement $\mathrm{SU}(n)$) si et seulement s'ils sont les mêmes $\theta_1, \cdots, \theta_n$ modulo 2π et permutations.

Démonstration. Nous ne donnons une démonstration que parce que la répétition est une méthode pédagogique éprouvée.

Soient $\lambda \in \mathbb{C}$ une valeur propre d'un élément $u \in \mathrm{U}(n)$ et x un vecteur propre non nul associé. Nous avons

$$\lambda\,\|x\|^2 = \langle u(x), x\rangle = \langle x, u^*(x)\rangle = \langle x, u^{-1}(x)\rangle = \left\langle x, \frac{1}{\lambda}\,x\right\rangle = \frac{1}{\overline{\lambda}}\,\|x\|^2\ .$$

Donc $|\lambda| = 1$.

Montrons par récurrence sur n que u est diagonalisable en base orthonormée. Si $n = 1$, le résultat est vrai. Si $n \geq 2$, pour tout vecteur propre non nul x de u, l'hyperplan $x^\perp$ est stable par u (voir la remarque (1) de la partie 1.5), et la restriction de u à $x^\perp$ est encore unitaire. Le résultat en découle par récurrence, et par concaténation de bases orthonormées. □

3.4 Sur l'application exponentielle des matrices

Notons $\mathbb{K}$ le corps des nombres réels $\mathbb{R}$ ou le corps des nombres complexes $\mathbb{C}$, muni de sa valeur absolue usuelle $|\cdot|$. Un outil essentiel pour la partie 3.5 est celui de l'exponentielle des matrices réelles ou complexes, dont nous rappelons les propriétés.

Soit $n \in \mathbb{N} - \{0\}$. L'application *exponentielle* des matrices $\exp : \mathscr{M}_n(\mathbb{K}) \to \mathrm{GL}_n(\mathbb{K})$ est définie (voir juste ci-dessous l'assertion (1) de la proposition 3.4) par

$$X \mapsto \exp X = \sum_{k\in\mathbb{N}} \frac{1}{k!}\, X^k \,.$$

Nous résumons ses propriétés dans le résultat suivant.

Proposition 3.4.

(1) *La série* $\sum_{k\in\mathbb{N}} \frac{1}{k!}\, X^k$ *converge normalement*[69] *dans* $\mathscr{M}_n(\mathbb{K})$ *(pour n'importe quelle norme d'algèbre sur* $\mathscr{M}_n(\mathbb{K})$*).*

(2) *Pour tous les* $X, Y \in \mathscr{M}_n(\mathbb{K})$*, si* $XY = YX$*, alors*

$$\exp(X+Y) = \exp X \ \exp Y$$

et en particulier $\exp X$ *est inversible, d'inverse* $(\exp X)^{-1} = \exp(-X)$*.*

(3) *Pour tous les* $\lambda_1, \dots, \lambda_n \in \mathbb{K}$*, nous avons*

$$\exp \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \begin{pmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{pmatrix} .$$

Plus généralement, si $T \in \mathscr{M}_n(\mathbb{K})$ *est triangulaire supérieure de coefficients diagonaux* $\lambda_1, \dots, \lambda_n$*, alors* $\exp T$ *est triangulaire supérieure de coefficients diagonaux* $e^{\lambda_1}, \dots, e^{\lambda_n}$*.*

69. Rappelons les propriétés suivantes (voir par exemple [Car]). Une série $\sum x_n$ de terme général x_n pour $n \in \mathbb{N}$ dans un espace vectoriel normé réel ou complexe E (de norme notée $\| \ \|$) est dite *normalement convergente* si la série à termes positifs ou nuls $\sum \|x_n\|$ converge. Si E est complet et si $\sum x_n$ est une série normalement convergente, alors la suite $(\sum_{k=1}^n x_k)_{n\in\mathbb{N}}$ est convergente dans E, vers une limite appelée la *somme* de la série et notée $\sum_{n\in\mathbb{N}} x_n$, qui vérifie

$$\Big\| \sum_{n\in\mathbb{N}} x_n \Big\| \leq \sum_{n\in\mathbb{N}} \|x_n\| \,.$$

Une *algèbre normée* A sur $\mathbb{K}$ (réelle ou complexe) est une algèbre sur $\mathbb{K}$ munie d'une norme $\| \ \|$ telle que pour tous les $x, y \in A$, nous ayons

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

Une *série entière* dans A est une série $\sum a_n\, x^n$ de terme général $a_n\, x^n$ pour $n \in \mathbb{N}$, avec $a_n \in \mathbb{K}$ et $x \in A$. Son *rayon de convergence (normale)* $r \in [0, +\infty]$ (avec les conventions usuelles pour 0 et ∞) est la borne supérieure des $\rho > 0$ telle que la série $\sum a_n\, x^n$ converge normalement pour tout x dans la boule ouverte $B(0,\rho)$ de centre 0 et de rayon $B(0,\rho)$. Si A est une *algèbre de Banach* (c'est-à-dire une algèbre normée complète, par exemple une algèbre normée de dimension finie, auquel cas toutes les normes (d'algèbre) sont équivalentes) et si $\sum a_n\, x^n$ est une série entière de rayon de convergence r strictement positif, alors l'application f de $B(0,r)$ dans E définie par $x \mapsto \sum_{n\in\mathbb{N}} a_n\, x^n$ est de classe C^∞. De plus, pour tout $k \in \mathbb{N}$, cette application admet un développement limité à l'ordre k en 0 donné par $f(x) = \sum_{n=1}^k a_n\, x^n + \mathrm{O}(x^{k+1})$.

(4) *Pour tout* $X \in \mathscr{M}_n(\mathbb{K})$*, nous avons*

$$^t(\exp X) = \exp(^t X)$$

et si $\mathbb{K} = \mathbb{C}$*, alors*

$$\overline{\exp X} = \exp \overline{X} \quad \text{et} \quad (\exp X)^* = \exp(X^*)$$

De plus, si $P \in \mathrm{GL}_n(\mathbb{K})$*, alors*

$$P(\exp X)P^{-1} = \exp(PXP^{-1}) \ .$$

(5) *Pour tout* $X \in \mathscr{M}_n(\mathbb{K})$*, en notant* $\mathtt{tr}\, X = \sum_{i=1}^n a_{i,i}$ *la* trace *de* $X = (a_{i,j})_{1 \leq i,j \leq n}$*, nous avons*

$$\det(\exp X) = e^{\mathtt{tr}\, X} \ .$$

(6) *L'application* exp *est de classe* C^∞ *(entre deux ouverts de* $\mathbb{K}^{n^2}$*). De plus la différentielle de* exp *en la matrice nulle* 0 *de* $\mathscr{M}_n(\mathbb{K})$ *est l'application identité de* $\mathscr{M}_n(\mathbb{K})$ *:*

$$d\exp_0 = \mathrm{id}_{\mathscr{M}_n(\mathbb{K})} \ .$$

Donc exp *est un* C^∞*-difféomorphisme d'un voisinage ouvert de* 0 *dans* $\mathscr{M}_n(\mathbb{K})$ *dans un voisinage ouvert de la matrice identité* I_n *dans* $\mathscr{M}_n(\mathbb{K})$*.*

(7) *Appelons* sous-groupe à un paramètre *de* $\mathrm{GL}_n(\mathbb{K})$ *toute application de classe* C^∞ *de* $\mathbb{R}$ *dans l'ouvert* $\mathrm{GL}_n(\mathbb{K})$ *qui est aussi un morphisme de groupes.*[70] *L'application de* $\mathscr{M}_n(\mathbb{K})$ *dans l'ensemble des sous-groupes à un paramètre de* $\mathrm{GL}_n(\mathbb{K})$*, qui à* $X \in \mathscr{M}_n(\mathbb{K})$ *associe l'application*

$$t \mapsto \exp(tX)$$

est une bijection.

(8) *Soit* $\mathrm{Ad} : \mathrm{GL}_n(\mathbb{K}) \to \mathrm{GL}(\mathscr{M}_n(\mathbb{K}))$ *le morphisme de groupes défini par*

$$\mathrm{Ad} : x \mapsto \left\{ \mathrm{Ad}\, x : X \mapsto x\, X\, x^{-1} \right\} .$$

Soit $\mathrm{ad} : \mathscr{M}_n(\mathbb{K}) \to \mathrm{End}(\mathscr{M}_n(\mathbb{K}))$ *l'application linéaire définie par*

$$\mathrm{ad} : X \mapsto \left\{ \mathrm{ad}\, X : Y \mapsto XY - YX \right\} .$$

Alors l'application Ad *est* C^∞*, et sa différentielle en la matrice identité est*

$$d\,\mathrm{Ad}_{I_n} = \mathrm{ad} \ .$$

De plus, pour tout $X \in \mathscr{M}_n(\mathbb{K})$*, nous avons*

$$\mathrm{Ad}(\exp X) = \exp(\mathrm{ad}\, X) \ .$$

(9) *Pour tout* $X \in \mathscr{M}_n(\mathbb{K})$*, la différentielle* $d\exp_X$ *de l'application exponentielle en* X *est inversible si et seulement si* X *n'admet pas de valeurs propres complexes* λ *et* μ *telles que* $\lambda - \mu \in 2i\pi\mathbb{Z} - \{0\}$*.*

70. En fait, tout morphisme de groupes continu de $\mathbb{R}$ dans $\mathrm{GL}_n(\mathbb{K})$ est automatiquement de classe C^∞, voir par exemple [Pau2, Coro. 5.22].

Démonstration. (1) Rappelons que $\mathscr{M}_n(\mathbb{K})$ est une algèbre sur $\mathbb{K}$ de dimension finie. En notant $\|\cdot\|$ la norme usuelle sur $\mathbb{K}^n$ (de sorte que $\|v\|^2 = \sum_{k=1}^n |v_k|^2$ pour tout vecteur $v = (v_1, \dots, v_n)$ dans $\mathbb{K}^n$), soit $\|\cdot\|$ la *norme d'opérateur* sur $\mathscr{M}_n(\mathbb{K})$, définie par

$$\|X\| = \sup_{v \in \mathbb{K}^n,\ \|v\| \leq 1} \|X\, v\|$$

pour tout $X \in \mathscr{M}_n(\mathbb{K})$, en identifiant tout $v \in \mathbb{K}^n$ avec le vecteur colonne de ses coordonnées. La propriété cruciale de sous-multiplicativité ($\|XY\| \leq \|X\|\ \|Y\|$ pour tous les $X, Y \in \mathscr{M}_n(\mathbb{K})$) de la norme d'opérateur montre la convergence normale de la série entière $\exp X$ dans toute l'algèbre de Banach $\mathscr{M}_n(\mathbb{K})$, avec de plus

$$\|\exp(X)\| \leq e^{\|X\|}\ .$$

Les propriétés (2) à (4) sont élémentaires. La démonstration de la première est formellement la même que dans le cas de l'application exponentielle de $\mathbb{R}$ dans $\mathbb{R}$, et justifiée par la commutativité des matrices. Mais on prendra bien garde que la formule $\exp(X+Y) = \exp X \exp Y$ n'est en général pas correcte sans l'hypothèse que X et Y commutent. [71] Une démontration soigneuse de la propriété (4) pour justifier le passage à la limite utilise la continuité (par la linéarité) de l'application de $\mathscr{M}_n(\mathbb{K})$ dans $\mathscr{M}_n(\mathbb{K})$ définie par $X \mapsto P\, X\, P^{-1}$, pour tout $P \in \mathrm{GL}_n(\mathbb{K})$.

Notons que le déterminant et la trace sont invariants par conjugaison, puisque det est un morphisme de groupes (nous avons $\det(xy) = \det(x)\det(y)$ pour tous les $x, y \in \mathscr{M}_n(\mathbb{K})$) à valeurs dans un groupe abélien et que $\mathtt{tr}(xy) = \mathtt{tr}(yx)$ pour tous les $x, y \in \mathscr{M}_n(\mathbb{K})$. L'assertion (5) se montre alors en trigonalisant [72] (sur $\mathbb{C}$) la matrice X et en utilisant les assertions (3) et (4). En effet, puisque $\mathbb{C}$ est algébriquement clos, il existe $P \in \mathrm{GL}_n(\mathbb{C})$ et $T \in \mathscr{M}_n(\mathbb{C})$ triangulaire supérieure telle que $X = P\,T\,P^{-1}$, de sorte que

$$\det(\exp(X)) = \det(\exp(P\,T\,P^{-1})) = \det(P\,\exp(T)\,P^{-1}) = \det(\exp(T)) = e^{\mathtt{tr}\,T} = e^{\mathtt{tr}\,X}\ .$$

Les deux premières affirmations de l'assertion (6) sont élémentaires par l'expression en tant que série entière de l'application exponentielle, et la dernière découle du théorème d'inversion locale. [73]

Montrons l'assertion (7), qui dit en particulier que les sous-groupes à un paramètre de $\mathrm{GL}_n(\mathbb{K})$ sont exactement les applications $t \mapsto \exp(tX)$ pour $X \in \mathscr{M}_n(\mathbb{K})$.

71. Par exemple, si $X = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ et $Y = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ (qui ne commutent pas), alors nous avons

$$\exp(X+Y) = \begin{pmatrix} 1 & e \\ 0 & e \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 0 & e \end{pmatrix} = \exp(X)\,\exp(Y)\ .$$

72. Une base d'un espace vectoriel V *trigonalise* un endomorphisme de V si la matrice de cet endomorphisme dans cette base est triangulaire supérieure.

73. Nous renvoyons par exemple à [Ave, Car, Cha] pour une démonstration du résultat suivant.

Théorème 3.5. (Théorème d'inversion locale) *Soient E et F deux espaces vectoriels normés réels de dimension finie, U un ouvert de E, k un élément de $(\mathbb{N}-\{0\})\cup\{\infty\}$ et $f : U \to F$ une application de classe C^k. Si, en un point x de U, la différentielle $df_x : E \to F$ est bijective, alors il existe un voisinage ouvert V de x contenu dans U, et un voisinage ouvert W de $f(x)$, tels que $f : V \to W$ soit un C^k-difféomorphisme.*

Par l'assertion (2) pour la propriété de morphisme de groupes, par le théorème de dérivation des applications composées, appliqué à l'application $t \mapsto tX$ linéaire donc C^∞ et l'application exp de classe C^∞ par l'assertion (6), il est immédiat que $t \mapsto \exp(tX)$ est un sous-groupe à un paramètre de $GL_n(\mathbb{K})$ pour tout $X \in \mathscr{M}_n(\mathbb{K})$. Réciproquement, soit φ un sous-groupe à un paramètre de $GL_n(\mathbb{K})$. Posons $X = \frac{d}{dt}_{|t=0}\varphi(t)$, qui appartient à $\mathscr{M}_n(\mathbb{K})$. En dérivant en $s = 0$ l'équation $\varphi(t+s) = \varphi(t)\varphi(s)$, nous obtenons $\frac{d\varphi}{dt}(t) = \varphi(t)X$. L'unique solution de cette équation différentielle linéaire du premier ordre telle que $\varphi(0) = I_n$ est $t \mapsto \exp(tX)$.

Montrons l'assertion (8). L'application $\iota : x \mapsto x^{-1}$ de $GL_n(\mathbb{K})$ dans $GL_n(\mathbb{K})$ est C^∞, car par la formule donnant l'inverse, les coefficients de x^{-1} sont des fractions rationnelles en les coefficients de x de dénominateur ne s'annulant pas. Comme l'application Φ de $\mathscr{M}_n(\mathbb{K}) \times \mathscr{M}_n(\mathbb{K})$ dans $\mathscr{M}_n(\mathbb{K})$ définie par $(x, y) \mapsto xy$ est bilinéaire, donc de différentielle en (I_n, I_n) égale à $(X, Y) \mapsto X + Y$, et puisque $\Phi(x, \iota(x)) = e$ pour tout $x \in \mathscr{M}_n(\mathbb{K})$, la différentielle de ι en I_n est

$$d\iota_{I_n} = -\operatorname{id}_{\mathscr{M}_n(\mathbb{K})} .$$

L'application de $\mathscr{M}_n(\mathbb{K}) \times \mathscr{M}_n(\mathbb{K})$ dans $\operatorname{End}(\mathscr{M}_n(\mathbb{K}))$ définie par $(x, y) \mapsto \{Z \mapsto xZy\}$ est bilinéaire donc C^∞. En particulier, sa différentielle en (I_n, I_n) est égale à l'application $(X, Y) \mapsto \{Z \mapsto XZ + ZY\}$. Le calcul de la différentielle $d\operatorname{Ad}_{I_n}$, qui est une application de $\mathscr{M}_n(\mathbb{K})$ dans $\operatorname{End}(\mathscr{M}_n(\mathbb{K}))$, s'en déduit, par le théorème de dérivation des fonctions composées :

$$d\operatorname{Ad}_{I_n}(X) = \{Z \mapsto XZ + Zd\iota_{I_n}(X)\} = \{Z \mapsto XZ - ZX\} .$$

L'application $\operatorname{Ad} : GL_n(\mathbb{K}) \to GL(\mathscr{M}_n(\mathbb{K}))$ est donc un morphisme de groupes et une application C^∞. Donc les applications $t \mapsto \operatorname{Ad}(\exp(tX))$ et $t \mapsto \exp(t \operatorname{ad} X)$ sont deux sous-groupes à un paramètre de $GL(\mathscr{M}_n(\mathbb{K}))$. Leur dérivée en $t = 0$ est respectivement $d\operatorname{Ad}_{I_n} \circ d\exp_0(X)$ et $\operatorname{ad} X$. Nous avons $d\exp_0 = \operatorname{id}_{\mathscr{M}_n(\mathbb{K})}$ par l'assertion (6) et $d\operatorname{Ad}_{I_n} = \operatorname{ad}$ par ce qui précède. Donc ces deux sous-groupes à un paramètre ont même dérivée en $t = 0$. Par l'assertion (7), ils sont donc égaux, ce qui montre le résultat en les évaluant en $t = 1$.

Pour démontrer l'assertion (9), nous commençons par un lemme calculant la différentielle de l'exponentielle en un élément quelconque X de $\mathscr{M}_n(\mathbb{K})$. Pour tout endomorphisme u d'un espace vectoriel réel V de dimension finie, notons $\Theta(u) = \frac{\operatorname{id} - \exp u}{u}$ la valeur de la série (normalement convergente pour n'importe quel choix de norme sur $\operatorname{End}(V)$, par exemple la norme d'opérateur)

$$\frac{\operatorname{id} - \exp u}{u} = -\sum_{n \in \mathbb{N}} \frac{1}{(n+1)!}\, u^n .$$

Pour tout $Y \in \mathscr{M}_n(\mathbb{K})$, notons L_Y l'élément de $\operatorname{End}(\mathscr{M}_n(\mathbb{K}))$ défini par $Z \mapsto YZ$.

Lemme 3.6. *Pour tout $X \in \mathscr{M}_n(\mathbb{K})$, nous avons*

$$d\exp_X = L_{\exp X} \circ \frac{\operatorname{id} - \exp(-\operatorname{ad} X)}{-\operatorname{ad} X} .$$

Démonstration. Remarquons que l'application $Y \mapsto L_Y$ de $\mathscr{M}_n(\mathbb{K})$ dans $\operatorname{End}(\mathscr{M}_n(\mathbb{K}))$ est linéaire et que si Y est inversible, alors L_Y est inversible, d'inverse $L_{Y^{-1}}$.

Soit $X \in \mathscr{M}_n(\mathbb{K})$. Puisque $(L_{\exp X})^{-1} = L_{(\exp X)^{-1}} = L_{\exp(-X)}$, il s'agit de montrer que pour tout Y dans $\mathscr{M}_n(\mathbb{K})$, nous avons

$$\exp(-X)\, d\exp_X(Y) = \Theta(-\operatorname{ad} X)(Y)\ .$$

Notons $f_X : \mathbb{R} \to \operatorname{End}(\mathscr{M}_n(\mathbb{K}))$ l'application C^∞ définie par

$$f_X : s \mapsto \left\{Y \mapsto s \exp(-sX)\, d\exp_{sX}(Y)\right\}\ .$$

Soient $s, t \in \mathbb{R}$. En dérivant par rapport à X l'équation $\exp((s+t)X) = \exp(sX)\, \exp(tX)$, nous obtenons par bilinéarité, pour tout Y dans $\mathscr{M}_n(\mathbb{K})$,

$$(s+t) d\exp_{(s+t)X}(Y) = s\, d\exp_{sX}(Y) \exp(tX) + t \exp(sX)\, d\exp_{tX}(Y)\ .$$

En multipliant à gauche par la matrice $\exp(-(s+t)X)$, nous avons, en utilisant les assertions (2) et (8),

$$\begin{aligned} f_X(s+t) &= \exp(-tX)\, f_X(s) \exp(tX) + f_X(t) = \operatorname{Ad}(\exp(-tX)) \circ f_X\ (s) + f_X(t) \\ &= \exp(\operatorname{ad}(-tX)) \circ f_X\ (s) + f_X(t) = \exp(-t \operatorname{ad} X) \circ f_X\ (s) + f_X(t)\ . \end{aligned}$$

En dérivant cette équation par rapport à t en $t = 0$, et comme ${f_X}'(0) = d\exp_0 = \mathrm{id}$ par l'assertion (6), nous obtenons

$${f_X}'(s) = \mathrm{id} - \operatorname{ad} X \circ f_X(s)\ .$$

Il est facile de vérifier que l'application $s \mapsto s\, \Theta(-s \operatorname{ad} X)$ est aussi une solution de cette équation différentielle linéaire du premier ordre, dont la valeur en $s = 0$ est $0 = f_X(0)$. Par unicité, nous avons donc $f_X(s) = s\, \Theta(-s \operatorname{ad} X)$ pour tout $s \in \mathbb{R}$. En prenant $s = 1$, le résultat en découle. □

Revenons maintenant à la démonstration de l'assertion (9). Si $(e_1, \dots, e_n)$ est une base de $\mathbb{C}^n$ trigonalisant X, alors la base $(E_{i,j})_{1 \le i,j \le n}$ (ordonnée en mettant bout à bout les lignes de cette matrice) de $\operatorname{End}(\mathbb{C}^n)$ associée à $(e_1, \dots, e_n)$ trigonalise $\operatorname{ad} X$. Donc si $\lambda_1, \dots, \lambda_n$ sont les valeurs propres complexes de X, alors les $\lambda_i - \lambda_j$ pour $1 \le i, j \le n$ sont les valeurs propres complexes de $\operatorname{ad} X$, et les valeurs propres complexes de $\frac{1-\exp(-\operatorname{ad} X)}{-\operatorname{ad} X}$ sont 1 si X admet une valeur propre complexe de multiplicité au moins 2 et les $\frac{1-e^{\lambda_i-\lambda_j}}{\lambda_i-\lambda_j}$ pour les valeurs propres distinctes λ_i et λ_j. En particulier, $\frac{1-\exp(-\operatorname{ad} X)}{-\operatorname{ad} X}$ est inversible si et seulement si la différence de deux valeurs propres distinctes de X n'est pas un multiple non nul de $2i\pi$. Le résultat découle alors du lemme 3.6. □

Donnons quelques notations qui seront utiles dans cette sous-partie et la suivante. Soit $n \in \mathbb{N} - \{0\}$.

Rappelons que $\mathrm{Sym}_n = \mathrm{Sym}_n(\mathbb{R}) = \{X \in \mathscr{M}_n(\mathbb{R}) \ :\ {}^tX = X\}$ est le sous-espace vectoriel réel de $\mathscr{M}_n(\mathbb{R})$ (muni de n'importe quelle norme) formé des matrices symétriques. Notons Sym_n^{++} l'ouvert [74] de Sym_n formé des matrices symétriques définies positives (c'est-à-dire dont la forme quadratique associée est définie positive).

74. L'application qui à une forme quadratique q sur $\mathbb{R}^n$ associe $\min_{x \in \mathbb{S}_{n-1}} q(x)$ (ou de manière équivalente, qui à une matrice symétrique A, associe $\min\{\, {}^tXAX \ :\ X \in \mathbb{S}_{n-1}\}$) est bien définie et continue par compacité de $\mathbb{S}_{n-1}$. De plus, q ou A sont définies positives si et seulement si cette fonction est strictement positive, ce qui est une condition ouverte. Dans la littérature, en particulier agrégative, cet ouvert Sym_n^{++} est souvent noté $\mathscr{S}_n^{++}(\mathbb{R})$, voire $\mathscr{S}_n^{++}$.

Rappelons que $\mathrm{Herm}_n = \mathrm{Herm}_n(\mathbb{C}) = \{X \in \mathscr{M}_n(\mathbb{C}) \ : \ X^* = X\}$ est le sous-espace vectoriel réel de $\mathscr{M}_n(\mathbb{C})$ (muni de n'importe quelle norme) formé des matrices hermitiennes. Notons Herm_n^{++} l'ouvert [75] de Herm_n formé des matrices hermitiennes définies positives (c'est-à-dire dont la forme hermitienne associée est définie positive).

L'exercice suivant montre d'une autre manière que les sous-espaces topologiques Sym_n^{++} et Herm_n^{++} sont ouverts dans Sym_n et Herm_n réciproquement.

Exercice E.27. *(**Critère de Sylvester**) Soit $A = (a_{i,j})_{1\leq i,j\leq n}$ une matrice $n \times n$ symétrique réelle ou hermitienne complexe. Montrer que A est définie positive si et seulement si ses* mineurs principaux $\det\big((a_{i,j})_{1\leq i,j\leq k}\big)$ *pour $k = 1, \ldots, n$ sont strictement positifs.*

Proposition 3.7. *Soit $n \in \mathbb{N} - \{0\}$.*

(1) *L'application* $\exp : \mathrm{Sym}_n \to \mathrm{Sym}_n^{++}$ *est un C^∞-difféomorphisme.*

(2) *L'application* $\exp : \mathrm{Herm}_n \to \mathrm{Herm}_n^{++}$ *est un C^∞-difféomorphisme.*

(3) *Il existe un et un seul C^∞-difféomorphisme noté $x \mapsto \sqrt{x}$ de Sym_n^{++} (respectivement Herm_n^{++}) dans lui-même tel que $(\sqrt{x}\,)^2 = x$ pour tout x dans Sym_n^{++} (respectivement Herm_n^{++}).*

(4) *Notons* $\mathrm{Nil}_n = \left\{ \begin{pmatrix} 0 & * & * \\ 0 & \ddots & * \\ 0 & 0 & 0 \end{pmatrix} \right\}$ *le sous-espace vectoriel sur $\mathbb{K}$ de $\mathscr{M}_n(\mathbb{K})$ formé des matrices* nilpotentes [76] *triangulaires supérieures. Notons* $\mathrm{Uni}_n = \left\{ \begin{pmatrix} 1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & 1 \end{pmatrix} \right\}$ *le sous-espace affine* [77] *de $\mathscr{M}_n(\mathbb{K})$ formé des matrices* unipotentes [78] *triangulaires supérieures. L'application* $\exp : \mathrm{Nil}_n \to \mathrm{Uni}_n$ *est un C^∞-difféomorphisme.*

(5) *Notons $\mathscr{N}_n$ la partie de $\mathscr{M}_n(\mathbb{K})$ formée des matrices nilpotentes, et $\mathscr{U}_n$ la partie de $\mathscr{M}_n(\mathbb{K})$ formée des matrices unipotentes. Alors l'application $E_n : \mathscr{N}_n \to \mathscr{U}_n$ définie*

75. L'argument est presque le même que celui de la note de bas de page précédente. Cet ouvert Herm_n^{++} est parfois noté $\mathscr{H}_n^{++}$.

76. c'est-à-dire dont une puissance est nulle, voir l'exercice ci-dessous pour des caractérisations.

Exercice E.28. *Soient $n \in \mathbb{N} - \{0\}$, k un corps de caractéristique nulle et $N \in \mathscr{M}_n(k)$. Montrer que les assertions suivantes sont équivalentes.*

- *La matrice N est nilpotente.*
- *Toutes les valeurs propres de N dans une clôture algébrique de k sont égales à 0.*
- *Le polynôme caractéristique $\det(XI_n - N)$ de N est égal à X^n.*
- *Il existe $m \in \{1, \ldots, n\}$ tel que le polynôme minimal unitaire de N soit égal à X^m.*

77. car $\mathrm{Uni}_n = I_n + \mathrm{Nil}_n$

78. c'est-à-dire dont la différence avec l'identité est nilpotente, voir l'exercice ci-dessous pour des caractérisations.

Exercice E.29. *Soient $n \in \mathbb{N} - \{0\}$, k un corps de caractéristique nulle et $U \in \mathscr{M}_n(k)$. Montrer que les assertions suivantes sont équivalentes.*

- *La matrice U est unipotente.*
- *Toutes les valeurs propres de U dans une clôture algébrique de k sont égales à 1.*
- *Le polynôme caractéristique $\det(XI_n - U)$ de U est égal à $(X-1)^n$.*
- *Il existe $m \in \{1, \ldots, n\}$ tel que le polynôme minimal unitaire de U soit égal à $(X-1)^m$.*

par $X \mapsto \sum_{k=0}^{n-1} \frac{1}{k!}\, X^k$ *est un homéomorphisme, d'inverse l'application* $L_n : \mathscr{U}_n \to \mathscr{N}_n$ *définie par* $X \mapsto \sum_{k=1}^{n-1} \frac{(-1)^{k+1}}{k}\, (X - I_n)^k$.

Démonstration. (1) Il suffit de remplacer $(\mathbb{C}, z \mapsto \overline{z})$ par $(\mathbb{R}, \mathrm{id})$, ainsi que le corollaire 1.15 (2) par le corollaire 1.17 (2), dans la démonstration de l'assertion (2) ci-dessous.

(2) L'application $\exp : \mathrm{Herm}_n \to \mathrm{Herm}_n^{++}$ est bien définie, car $(\exp x)^* = \exp(x^*)$ et si $x \in \mathrm{Herm}_n$, alors x est à valeurs propres réelles et diagonalisable en base orthonormée (voir le corollaire 1.15 (2)), donc il existe un élément $P \in U(n)$ et une matrice diagonale réelle D tels que $x = PDP^{-1}$, donc $\exp x = P(\exp D)P^{-1} \in \mathrm{Herm}_n^{++}$.

L'application $\exp : \mathrm{Herm}_n \to \mathrm{Herm}_n^{++}$ est surjective, car si $y \in \mathrm{Herm}_n^{++}$, alors par le corollaire 1.15 (2), il existe un élément $P \in U(n)$ et une matrice diagonale à coefficients diagonaux strictement positifs D_+ tels que $y = PD_+P^{-1}$, et si $x = P(\ln D_+)P^{-1}$ où $\ln D_+$ est la matrice diagonale de coefficients diagonaux les logarithmes des coefficients diagonaux de D_+, alors $y = \exp x$.

L'application $\exp : \mathrm{Herm}_n \to \mathrm{Herm}_n^{++}$ est injective, car si $z \in \mathrm{Herm}_n$ alors toute base qui diagonalise $\exp z$ diagonalise z et réciproquement. Donc si $x, x' \in \mathrm{Herm}_n$ vérifient $\exp x = \exp(x')$, alors x, x' sont simultanément diagonalisables en base orthonormée donc $x = PDP^{-1}$ et $x' = PD'P^{-1}$ avec $P \in U(n)$ et D, D' diagonales réelles. Comme $D = \ln \exp D = \ln \exp D' = D'$, nous avons $x = x'$.

L'application $\exp : \mathrm{Herm}_n \to \mathrm{Herm}_n^{++}$ est donc une bijection de classe C^∞ entre deux ouverts de l'espace vectoriel réel Herm_n de dimension finie. Les valeurs propres des éléments X de Herm_n sont réelles (et donc la différence de deux d'entre elles ne peut être un multiple entier non nul de $2i\pi$). Donc la différentielle $d\exp_X$ de exp en tout élément X de Herm_n est injective par la proposition 3.4 (9). Par conséquent, l'application linéaire $d\exp_X$ est bijective de Herm_n dans Herm_n. Par le théorème d'inversion locale 3.5, l'application $\exp : \mathrm{Herm}_n \to \mathrm{Herm}_n^{++}$ est donc un C^∞-difféomorphisme local bijectif, donc un C^∞-difféomorphisme.

(3) Il suffit de poser

$$\sqrt{x} = \exp\left(\frac{1}{2}\exp^{-1}(x)\right).$$

(4) et (5) Notons que Nil_n est un sous-espace vectoriel de $\mathscr{M}_n(\mathbb{K})$ stable par multiplication et que $\mathrm{Uni}_n = I_n + \mathrm{Nil}_n$ est un sous-espace affine de $\mathscr{M}_n(\mathbb{K})$, qui est de plus un sous-groupe de $\mathrm{GL}_n(\mathbb{K})$. En particulier, il est possible d'effectuer du calcul différentiel sur Nil_n et Uni_n.

Notons en remarque préliminaire que $\mathscr{N}_n$ est stable par combinaisons linéaires de ses matrices commutantes (si X et Y commutent et sont nilpotentes d'ordre p et q, alors [79]

79. Soient E un espace vectoriel (pas forcément de dimension finie) sur un corps commutatif k (quelconque), et $u, v \in \mathscr{L}(E)$. Si u et v sont nilpotents, soient $m, n \in \mathbb{N}$ tels que $u^m = 0$ et $v^m = 0$. Si u et v commutent, alors la formule du binôme montre que

$$(u+v)^{n+m} = \sum_{k=0}^{n+m} a_k\, u^k v^{n+m-k}$$

où $a_k \in \mathbb{N}$. Pour tout $k = 1, \ldots, n+m$, nous avons $k \geq m$ (auquel cas $u^k = 0$) ou $n+m-k \geq n$ (auquel cas $v^{n+m-k} = 0$), et donc dans les deux cas $u^k v^{n+m-k} = 0$. Donc $u+v$ est nilpotent, d'ordre de nilpotence au plus $n+m$.

Voici un contre-exemple sans l'hypothèse de commutation. La matrice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ est somme des matrices

leur somme $X+Y$ est nilpotente d'ordre au plus $p+q$), et par multiplication de ses matrices commutantes (si X et Y commutent et sont nilpotentes d'ordre p et q alors XY est nilpotente d'ordre au plus $\min\{p,q\}$).

Nous avons $y \in \mathrm{Uni}_n$ (respectivement $y \in \mathscr{U}_n$) si et seulement si $y - I_n \in \mathrm{Nil}_n$ (respectivement $y - I_n \in \mathscr{N}_n$). Pour tout $x \in \mathrm{Nil}_n$ (respectivement $x \in \mathscr{N}_n$), nous avons $x^k = 0$ pour tout $k \geq n$, donc $\exp x = \sum_{k=0}^{n-1} \frac{1}{k!}\, x^k = I_n + \sum_{k=1}^{n-1} \frac{1}{k!}\, x^k$ appartient à Uni_n (respectivement $\mathscr{U}_n$) car $\sum_{k=1}^{n-1} \frac{1}{k!}\, x^k$ appartient à Nil_n (respectivement $\mathscr{N}_n$) par la remarque préliminaire.

Notons $\ln : \mathscr{U}_n \to \mathscr{M}_n(\mathbb{K})$ l'application $y \mapsto \sum_{k=1}^{n-1} \frac{(-1)^{k+1}}{k}\, (y - I_n)^k$, qui est à valeurs dans $\mathscr{N}_n$ par la remarque préliminaire, en remarquant qu'elle envoie Uni_n dans Nil_n. Alors ln et exp sont continues, et de classe C^∞ en restriction à Uni_n et Nil_n respectivement, car polynomiales en les coefficients.

Considérons les polynômes réels

$$E_n(X) = \sum_{k=0}^{n-1} \frac{1}{k!}\, X^k \quad \text{et} \quad L_n(X) = \sum_{k=1}^{n-1} \frac{(-1)^{k+1}}{k}\, (X-1)^k$$

en la variable X. Puisque les applications réelles $\exp : \mathbb{R} \to]0, +\infty[$ et $\ln :]0, +\infty[\to \mathbb{R}$ sont inverses l'une de l'autre et ont les développements limités bien connus en 0 et en 1 à l'ordre n, les applications polynomiales réelles associées à E_n et L_n vérifient les propriétés asymptotiques $E_n(L_n(t)) - t = \mathrm{O}((t-1)^n)$ quand $t \to 1$ et $L_n(E_n(t)) - t = \mathrm{O}(t^n)$ quand $t \to 0$. Donc les polynômes réels $E_n(L_n(X+1)) - (X+1)$ et $L_n(E_n(X)) - X$, qui ont un zéro d'ordre n en 0, sont divisibles par X^n. En passant aux polynômes d'endomorphismes, il en découle que $\ln \circ \exp$ vaut l'identité sur Nil_n (respectivement $\mathscr{N}_n$) et que $\exp \circ \ln$ vaut l'identité sur Uni_n (respectivement $\mathscr{U}_n$). □

Remarques. (1) Notons que $\sqrt{PxP^{-1}} = P\sqrt{x}P^{-1}$ pour tous les $x \in \mathrm{Herm}_n^{++}$ et $P \in \mathrm{U}(n)$ (respectivement $x \in \mathrm{Sym}_n^{++}$ et $P \in \mathrm{O}(n)$), et que

$$\sqrt{\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}} = \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix}$$

pour tous les $\lambda_1, \ldots, \lambda_n > 0$.

(2) De même,

$$\exp : \{x \in \mathrm{Sym}_n \ : \ \mathtt{tr}\ x = 0\} \to \{x \in \mathrm{Sym}_n^{++} \ : \ \det x = 1\}$$

et

$$\exp : \{x \in \mathrm{Herm}_n \ : \ \mathtt{tr}\ x = 0\} \to \{x \in \mathrm{Herm}_n^{++} \ : \ \det x = 1\}$$

sont des homéomorphismes. [80]

Exercice E.30. Soit $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{R}$, et soit $n \in \mathbb{N} - \{0\}$.

---

nilpotentes $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ et $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, mais n'est pas nilpotente (ses valeurs propres sont ± 1).

80. En fait, ce sont des C^∞-difféomorphismes, lorsque l'on munit les images de leur structure de sous-variété différentielle des ouverts Sym_n^{++} et Herm_n^{++} (voir par exemple [Pau2], [Cha, Chap. 6], [Laf]).

(1) Soient $k, n_1, \dots, n_k \in \mathbb{N} - \{0\}$ et $A_1 \in \mathscr{M}_{n_1}(\mathbb{K}), \dots, A_k \in \mathscr{M}_{n_k}(\mathbb{K})$. Montrer que

$$\exp \begin{pmatrix} A_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & A_k \end{pmatrix} = \begin{pmatrix} \exp(A_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \exp(A_k) \end{pmatrix}.$$

(2) Pour tous les $\lambda, t \in \mathbb{K}$, calculer $\exp J_{0,t}$ puis $\exp J_{\lambda,t}$ où

$$J_{0,t} = \begin{pmatrix} 0 & t & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & t \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix} \quad \text{et} \quad J_{\lambda,t} = \begin{pmatrix} \lambda & t & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & t \\ 0 & \dots & \dots & \dots & \lambda \end{pmatrix}.$$

(3) Est-ce que l'application $\exp : \mathscr{M}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ est injective ?

(4) Pour tout $x \in \mathscr{M}_n(\mathbb{C})$, montrer que $\exp(x) = \mathrm{id}$ si et seulement si x est diagonalisable de valeurs propres dans $2i\pi\mathbb{Z}$.

(5) Est-ce que l'application $\exp : \mathscr{M}_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$ est injective ?

(6) Pour tout $A \in \mathscr{M}_n(\mathbb{K})$, montrer que la matrice $\exp(A)$ appartient à l'anneau $\mathbb{K}[A]$ des polynômes en la matrice A.

Exercice E.31. Montrer que $\mathrm{GL}_n(\mathbb{K})$ est sans sous-groupe arbitrairement petit. [81]

Exercice E.32. (Propriétés de surjectivité de l'application exponentielle) Soit $n \in \mathbb{N} - \{0\}$.

(1) Montrer que l'application $\exp : \mathscr{M}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ est surjective. Pour ceci, pour tout $A \in \mathscr{M}_n(\mathbb{C})$, noter $\mathbb{C}[A]$ l'anneau des polynômes en la matrice A, noter $\mathbb{C}[A]^\times$ le groupe multiplicatif des polynômes en A inversibles dans l'anneau $\mathbb{C}[A]$, et montrer successivement que
- $\mathbb{C}[A]^\times = \mathbb{C}[A] \cap \mathrm{GL}_n(\mathbb{C})$,
- $\exp(\mathbb{C}[A]) \subset \mathbb{C}[A]^\times$,
- $\mathbb{C}[A]^\times$ est connexe par arcs,
- $\exp(\mathbb{C}[A])$ est ouvert dans $\mathbb{C}[A]^\times$,
- $\exp(\mathbb{C}[A])$ est fermé dans $\mathbb{C}[A]^\times$,
- $\exp(\mathbb{C}[A]) = \mathbb{C}[A]^\times$.

(2) Montrer que l'application $\exp : \mathscr{M}_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$ n'est pas surjective, que son image est strictement contenue dans $\mathrm{GL}_n^+(\mathbb{R}) = \{x \in \mathrm{GL}_n(\mathbb{R}) \ : \ \det x > 0\}$ si $n \geq 2$, et que son image est exactement $\{x^2 \ : \ x \in \mathrm{GL}_n(\mathbb{R})\}$.

81. Un groupe topologique est dit *sans sous-groupe arbitrairement petit* s'il existe un voisinage U de l'identité tel que le seul sous-groupe contenu dans U soit le sous-groupe trivial. Un célèbre théorème de Gleason-Yamabe dit qu'un groupe topologique localement compact sans sous-groupe arbitrairement petit est un groupe de Lie réel. Par exemple, pour le lecteur ou la lectrice qui connaît le corps des p-adiques $\mathbb{Q}_p$ pour $p \in \mathbb{N} - \{0, 1\}$ premier (voir par exemple [Ser]), le groupe topologique additif localement compact $\mathbb{Q}_p$ contient une base de voisinages de l'élément neutre formée de sous-groupes compacts ouverts : il suffit de prendre $(p^k\mathbb{Z}_p)_{k\in\mathbb{N}}$ (et de même pour $\mathrm{GL}_n(\mathbb{Q}_p)$ avec $\big(\ker \varphi_k\big)_{k\in\mathbb{N}}$) où $\varphi_k : \mathrm{GL}_n(\mathbb{Z}_p) \to \mathrm{GL}_n(\mathbb{Z}_p/(p^k\mathbb{Z}_p))$ est le morphisme de groupes de réduction modulo $p^k\mathbb{Z}_p$ des coefficients des matrices de $\mathrm{GL}_n(\mathbb{Z}_p)$.

(3) Montrer qu'un élément A de $\mathrm{GL}_n(\mathbb{R})$ appartient à l'image de l'application exponentielle $\exp : \mathscr{M}_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$ si et seulement si, pour tous les $\lambda \in]-\infty, 0[$ et $k \in \{1, \dots, n\}$, le nombre de blocs de Jordan $J_k(\lambda) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$ de A de taille $k \times k$ et de valeur propre λ, dans la décomposition de Jordan de A sur $\mathbb{C}$, est pair.

(4) Montrer qu'un élément A de $\mathrm{SL}_n(\mathbb{R})$ appartient à l'image de l'application exponentielle

$$\exp : \mathfrak{sl}_n(\mathbb{R}) = \{x \in \mathscr{M}_n(\mathbb{R}) \ : \ \mathtt{tr}\, x = 0\} \to \mathrm{GL}_n(\mathbb{R})$$

si et seulement s'il appartient à l'image de $\exp : \mathscr{M}_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$.

(5) Montrer que l'application

$$\exp : \mathfrak{so}(n) = \{x \in \mathscr{M}_n(\mathbb{R}) \ : \ {}^t x + x = 0\} \ \to \ \mathrm{SO}(n)$$

est surjective.

(6) Montrer que l'application

$$\exp : \mathfrak{u}(n) = \{x \in \mathscr{M}_n(\mathbb{C}) \ : \ x^* + x = 0\} \ \to \ \mathrm{U}(n)$$

est surjective.

(7) Montrer que l'application

$$\exp : \mathfrak{su}(n) = \{x \in \mathscr{M}_n(\mathbb{C}) \ : \ x^* + x = 0, \ \mathtt{tr}\, x = 0\} \ \to \ \mathrm{SU}(n)$$

est surjective.

(8) Montrer que l'application

$$\exp : \mathfrak{sl}_2(\mathbb{C}) = \{x \in \mathscr{M}_2(\mathbb{C}) \ : \ \mathtt{tr}\, x = 0\} \to \mathrm{SL}_2(\mathbb{C})$$

n'est pas surjective, et que son image est exactement $\{x \in \mathrm{SL}_2(\mathbb{C}) \ : \ (x+I_2)^2 \neq 0\}$. On pourra montrer que si $X \in \mathfrak{sl}_2(\mathbb{C})$ et si ω est une des deux racines carrées complexes de $-\det X$, alors $\exp X = (\cosh \omega)\, I_2 + \frac{\sinh \omega}{\omega}\, X$.

(9) Montrer que l'application

$$p \circ \exp : \mathfrak{sl}_n(\mathbb{C}) = \{x \in \mathscr{M}_n(\mathbb{C}) \ : \ \mathtt{tr}\, x = 0\} \to \mathrm{PSL}_n(\mathbb{C})$$

où $p : \mathrm{SL}_n(\mathbb{C}) \to \mathrm{PSL}_n(\mathbb{C}) = \mathrm{SL}_n(\mathbb{C})/Z(\mathrm{SL}_n(\mathbb{C}))$ est la projection canonique, est surjective.

Exercice E.33. Soit $A \in \mathscr{M}_n(\mathbb{C})$. Montrer que A est normale [82] inversible si et seulement s'il existe une matrice $B \in \mathscr{M}_n(\mathbb{C})$ normale telle que $A = \exp B$.

82. Rappelons qu'une matrice est *normale* si elle commute avec son adjointe.

3.5 Décomposition polaire et topologie des groupes classiques

Une référence globale pour cette partie est [MT].

Notons $\mathbb{K}$ le corps des nombres réels $\mathbb{R}$ ou le corps des nombres complexes $\mathbb{C}$, muni de sa valeur absolue usuelle $|\cdot|$. Nous munissons tout espace vectoriel sur $\mathbb{K}$ de dimension finie de la topologie définie par n'importe laquelle de ses normes (elles sont toutes équivalentes, et définissent aussi les mêmes bornés).

Un *groupe topologique* est un groupe G muni d'une topologie rendant continues les opérations de groupes, c'est-à-dire telle que l'application $(x, y) \mapsto xy^{-1}$ de $G \times G$ dans G soit continue.

Par exemple, le groupe $\mathbb{U}$ des nombres complexes de module 1, muni de la topologie induite de celle de $\mathbb{C}$, est un groupe topologique compact connexe par arcs. Tout sous-groupe de $\mathrm{GL}_n(\mathbb{K})$, munie de la topologie induite de celle de $\mathbb{K}^{n^2}$, est un groupe topologique, car la multiplication des matrices est polynomiale en les coefficients, et le passage à l'inverse des matrices est une fraction rationnelle de dénominateur ne s'annulant pas en les coefficients.

Proposition 3.8. *Pour tout $n \in \mathbb{N}$, les groupes topologiques*

$$\mathrm{O}(n) = \{x \in \mathscr{M}_n(\mathbb{R}) : \ {}^t x\, x = I_n\} \quad \text{et} \quad \mathrm{U}(n) = \{x \in \mathscr{M}_n(\mathbb{C}) : \ x^*\, x = I_n\}$$

sont compacts.

Démonstration. Ils sont fermés, car définis par des équations continues (car polynomiales sur $\mathbb{R}$), et bornés, car leurs vecteurs colonnes sont de norme 1, dans les espaces vectoriels de dimension finie $\mathscr{M}_n(\mathbb{R})$ et $\mathscr{M}_n(\mathbb{C})$ respectivement. □

Décomposition polaire de $\mathrm{GL}_n(\mathbb{R})$ et de $\mathrm{GL}_n(\mathbb{C})$

Soit $n \in \mathbb{N} - \{0\}$. Le résultat suivant permet de ramener l'étude des propriétés topologiques des groupes linéaires à ceux des groupes orthogonaux et unitaires. Il dit en particulier que pour tout $A \in \mathrm{GL}_n(\mathbb{C})$, il existe un unique couple de matrices (U, H) où $U \in \mathrm{U}(n)$ est unitaire et $H \in \mathrm{Herm}_n^{++}$ est hermitienne définie positive tel que $A = U\,H$, et pour tout $A \in \mathrm{GL}_n(\mathbb{R})$, il existe un unique couple de matrices (O, S) où $O \in \mathrm{O}(n)$ est orthogonale et $S \in \mathrm{Sym}_n^{++}$ est symétrique définie positive tel que $A = O\,S$.

Théorème 3.9. (Décomposition polaire) *L'application de $\mathrm{Herm}_n^{++} \times \mathrm{U}(n)$ dans $\mathrm{GL}_n(\mathbb{C})$ définie par $(x, y) \mapsto xy$ est un homéomorphisme*[83] *(appelé la* décomposition polaire *de $\mathrm{GL}_n(\mathbb{C})$), d'inverse $x \mapsto (\sqrt{xx^*}, \sqrt{xx^*}^{-1} x)$, où $x^* = {}^t\overline{x}$ est la matrice adjointe de x et $\sqrt{\ }$ est l'application définie dans la proposition 3.7 (3).*

L'application de $\mathrm{Sym}_n^{++} \times \mathrm{O}(n)$ dans $\mathrm{GL}_n(\mathbb{R})$ définie par $(x, y) \mapsto xy$ est un homéomorphisme[84] *(appelé la* décomposition polaire *de $\mathrm{GL}_n(\mathbb{R})$), d'inverse l'application définie par $x \mapsto (\sqrt{x\,{}^t x}, \sqrt{x\,{}^t x}^{-1} x)$.*

83. En fait, lorsque l'on munit (voir par exemple [Pau2], [Cha, Chap. 6], [Laf]) $\mathrm{U}(n)$ de sa structure de sous-variété différentielle de l'ouvert $\mathrm{GL}_n(\mathbb{C})$ de $\mathbb{R}^{2n^2}$, la démonstration ci-dessous montre que la décomposition polaire de $\mathrm{GL}_n(\mathbb{C})$ est un C^∞-difféomorphisme.

84. En fait, lorsque l'on munit (voir par exemple [Pau2], [Cha, Chap. 6], [Laf]) $\mathrm{O}(n)$ de sa structure de sous-variété différentielle de l'ouvert $\mathrm{GL}_n(\mathbb{R})$ de $\mathbb{R}^{n^2}$, la démonstration ci-dessous montre que la décomposition polaire de $\mathrm{GL}_n(\mathbb{R})$ est un C^∞-difféomorphisme.

De même, les applications de $\mathrm{U}(n) \times \mathrm{Herm}_n^{++}$ dans $\mathrm{GL}_n(\mathbb{C})$ et de $\mathrm{O}(n) \times \mathrm{Sym}_n^{++}$ dans $\mathrm{GL}_n(\mathbb{R})$, définies par $(x, y) \mapsto xy$, sont des homéomorphismes, dont les inverses sont les applications $x \mapsto (x\sqrt{x^*x}^{-1}, \sqrt{x^*x})$ et $x \mapsto (x\sqrt{{}^t x\, x}^{-1}, \sqrt{{}^t x\, x})$ respectivement.

Démonstration. Nous démontrons le premier résultat, le second étant analogue.

Les applications $\phi : (x, y) \mapsto xy$ et $\psi : x \mapsto (\sqrt{xx^*}, \sqrt{xx^*}^{-1} x)$ sont bien définies (en effet xx^* est hermitienne[85] définie positive[86] pour tout $x \in \mathrm{GL}_n(\mathbb{C})$), et continues (car polynomiale en les coefficients pour la première et composée d'applications continues pour la seconde, d'après la proposition 3.7 (3)). Il est immédiat que $\phi \circ \psi$ vaut l'identité de $\mathrm{GL}_n(\mathbb{C})$. Si $x \in \mathrm{Herm}_n^{++}$ et $y \in U(n)$, alors

$$\sqrt{(xy)(xy)^*} = \sqrt{x(yy^*)x^*} = \sqrt{x^2} = x \ .$$

Donc $\psi \circ \phi$ vaut l'identité de $\mathrm{Herm}_n^{++} \times \mathrm{U}(n)$. □

Corollaire 3.10. *L'application de* $\mathrm{GL}_n(\mathbb{R})$ *dans* Sym_n^{++} *définie par* $x \mapsto x\,{}^t x$ *induit par passage au quotient un homéomorphisme*[87]

$$\mathrm{GL}_n(\mathbb{R})/\,\mathrm{O}(n) \simeq \mathrm{Sym}_n^{++} \ .$$

L'application de $\mathrm{GL}_n(\mathbb{C})$ *dans* Herm_n^{++} *définie par* $x \mapsto x\,x^*$ *induit par passage au quotient un homéomorphisme*[88]

$$\mathrm{GL}_n(\mathbb{C})/\,\mathrm{U}(n) \simeq \mathrm{Herm}_n^{++} \ .$$

Démonstration. Nous montrons le second résultat, le premier se démontre de manière analogue. Nous munissons l'ensemble quotient $\mathrm{GL}_n(\mathbb{C})/\,\mathrm{U}(n)$ pour l'action par translations à gauche de $\mathrm{U}(n)$ sur $\mathrm{GL}_n(\mathbb{C})$ (définie par $x \mapsto \{y \mapsto yx^{-1}\}$ pour $x \in \mathrm{U}(n)$ et $y \in \mathrm{GL}_n(\mathbb{C})$) de la topologie quotient.

La décomposition polaire montre que l'application $\widetilde{\theta} : x \mapsto x\,x^*$ est bien définie, continue et surjective[89]. Pour tous les $x, y \in \mathrm{GL}_n(\mathbb{C})$ nous avons $xx^* = yy^*$ si et seulement si $x^{-1}y = y^*(x^*)^{-1} = (x^{-1}y)^*$, c'est-à-dire si et seulement si $x^{-1}y$ appartient à $\mathrm{U}(n)$, ou encore si et seulement s'il existe $z \in \mathrm{U}(n)$ tel que $x = yz$. Donc $\widetilde{\theta}$ passe au quotient en une application θ bijective de $\mathrm{GL}_n(\mathbb{C})/\,\mathrm{U}(n)$ dans Herm_n^{++}, qui est continue par les propriétés de la topologie quotient.

Notons $\pi : \mathrm{GL}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})/\,\mathrm{U}(n)$ la projection canonique, qui est continue par définition de la topologie quotient. L'application réciproque de θ est continue, car c'est l'application $x \mapsto \pi(\sqrt{x})$, où $x \mapsto \sqrt{x}$ est l'application continue définie dans la proposition 3.7 (3). Donc θ est un homéomorphisme. □

Application à la topologie des groupes classiques

Soit G est un groupe topologique (par exemple un sous-groupe de $\mathrm{GL}_n(\mathbb{K})$, avec la topologie induite).

85. car $(xx^*)^* = (x^*)^*x^* = xx^*$

86. car pour tout $v \in \mathbb{C}^n$ (identifié au vecteur colonne de ses coordonnées), si $\langle \cdot, \cdot \rangle$ est le produit hermitien standard de $\mathbb{C}^n$, alors $\langle xx^*v, v\rangle = \langle x^*v, x^*v\rangle = \|x^*v\|^2$, qui est positif ou nul, et qui est nul si et seulement si $x^*v = 0$, c'est-à-dire si v est nul car x est inversible

87. En fait, lorsque l'on munit (voir par exemple [Pau3]) $\mathrm{GL}_n(\mathbb{R})/\,\mathrm{O}(n)$ de sa structure de variété différentielle quotient, cette application est un C^∞-difféomorphisme.

88. En fait, lorsque l'on munit (voir par exemple [Pau3]) $\mathrm{GL}_n(\mathbb{C})/\,\mathrm{U}(n)$ de sa structure de variété différentielle quotient, cette application est un C^∞-difféomorphisme.

89. Ceci revient à dire que l'action de $\mathrm{GL}_n(\mathbb{C})$ sur Herm_n^{++} définie par $x \mapsto \{y \mapsto xyx^*\}$ où $x \in \mathrm{GL}_n(\mathbb{C})$ et $y \in \mathrm{Herm}_n^{++}$ est transitive.

Proposition 3.11. *L'ensemble $\pi_0 G$ des composantes connexes par arcs de G, muni de la loi de composition interne qui aux composantes connexes par arcs de $x \in G$ et de $y \in G$ associe la composante connexe par arcs de xy, est un groupe.*

Démonstration. Cette loi est en effet bien définie, car pour tous les x, x', y, y' dans G, si $\alpha : [0,1] \to G$ est un chemin continu de x à x' et $\beta : [0,1] \to G$ est un chemin continu de y à y', alors $t \mapsto \alpha(t)\,\beta(t)$ est chemin continu de xy à $x'y'$.

Pour tout $x \in G$, notons $[x]$ la composante connexe par arcs de x dans G. Posons $[x]^{-1} = [x^{-1}]$, ce qui ne dépend pas du choix des représentants, car pour tous les x, y dans G, si $\alpha : [0,1] \to G$ est un chemin continu de x à y, alors $t \mapsto \alpha(t)^{-1}$ est chemin continu de x^{-1} à y^{-1}. La composante connexe $[e]$ de l'élément neutre e de G est élément neutre dans $\pi_0 G$ car $[x][e] = [xe] = [x] = [ex] = [e][x]$. L'associativité et le fait que $[x]^{-1}$ soit l'inverse de x sont aussi immédiats par construction. □

Si G_0 est la composante connexe par arcs de l'élément neutre de G, alors l'application de G dans $\pi_0 G$, qui à un point associe sa composante connexe par arcs, est un morphisme de groupes surjectif par construction, de noyau G_0. Il induit donc par passage au quotient un isomorphisme de groupes $G/G_0 \simeq \pi_0 G$.

Corollaire 3.12. *Soient $n, p, q \in \mathbb{N} - \{0\}$. Dans chaque ligne du tableau suivant, le groupe matriciel de la colonne de gauche est homéomorphe à l'espace topologique produit de la colonne centrale lorsqu'il est indiqué, et a pour groupe des composantes connexes par arcs celui de la colonne de droite.*

G	décomposition polaire	$\pi_0 G$
$\mathrm{GL}_n(\mathbb{R})$	$\mathrm{O}(n) \times \mathbb{R}^{\frac{n(n+1)}{2}}$	$\mathbb{Z}/2\mathbb{Z}$
$\mathrm{GL}_n(\mathbb{C})$	$\mathrm{U}(n) \times \mathbb{R}^{n^2}$	1
$\mathrm{SL}_n(\mathbb{R})$	$\mathrm{SO}(n) \times \mathbb{R}^{\frac{n(n+1)}{2}-1}$	1
$\mathrm{SL}_n(\mathbb{C})$	$\mathrm{SU}(n) \times \mathbb{R}^{n^2-1}$	1
$\mathrm{O}(n)$		$\mathbb{Z}/2\mathbb{Z}$
$\mathrm{SO}(n)$		1
$\mathrm{O}(p,q)$	$\mathrm{O}(p) \times \mathrm{O}(q) \times \mathbb{R}^{pq}$	$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$
$\mathrm{SO}(p,q)$	$S(\mathrm{O}(p) \times \mathrm{O}(q)) \times \mathbb{R}^{pq}$	$\mathbb{Z}/2\mathbb{Z}$
$\mathrm{U}(n)$		1
$\mathrm{SU}(n)$		1
$\mathrm{U}(p,q)$	$\mathrm{U}(p) \times \mathrm{U}(q) \times \mathbb{R}^{2pq}$	1
$\mathrm{SU}(p,q)$	$S(\mathrm{U}(p) \times \mathrm{U}(q)) \times \mathbb{R}^{2pq}$	1

Dans le tableau ci-dessus, nous identifions $\mathrm{O}(p) \times \mathrm{O}(q)$ avec le sous-groupe de $\mathrm{GL}_{p+q}(\mathbb{R})$ diagonal par blocs

$$\left\{ \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} \,:\, x \in \mathrm{O}(p),\ \ y \in \mathrm{O}(q) \right\},$$

dont l'intersection avec $\mathrm{SL}_{p+q}(\mathbb{R})$ est notée $S(\mathrm{O}(p) \times \mathrm{O}(q))$. Nous identifions $\mathrm{U}(p) \times \mathrm{U}(q)$ avec le sous-groupe de $\mathrm{GL}_{p+q}(\mathbb{C})$ diagonal par blocs

$$\left\{ \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} \,:\, x \in \mathrm{U}(p),\ \ y \in \mathrm{U}(q) \right\},$$

dont l'intersection avec $\mathrm{SL}_{p+q}(\mathbb{C})$ est notée $S(\mathrm{U}(p)\times\mathrm{U}(q))$.

Démonstration. Les deux premiers homéomorphismes

$$\mathrm{GL}_n(\mathbb{R})\simeq \mathrm{O}(n)\times\mathbb{R}^{\frac{n(n+1)}{2}} \quad\text{et}\quad \mathrm{GL}_n(\mathbb{C})\simeq \mathrm{U}(n)\times\mathbb{R}^{n^2}$$

s'obtiennent en composant les homéomorphismes

$$\mathrm{GL}_n(\mathbb{R})\to \mathrm{Sym}_n^{++}\times\mathrm{O}(n) \quad\text{et}\quad \mathrm{GL}_n(\mathbb{C})\to \mathrm{Herm}_n^{++}\times\mathrm{U}(n)$$

donnés par le théorème 3.9 et les homéomorphismes

$$\exp^{-1}:\mathrm{Sym}_n^{++}\to\mathrm{Sym}_n \quad\text{et}\quad \exp^{-1}:\mathrm{Herm}_n^{++}\to\mathrm{Herm}_n$$

respectivement donnés par les assertions (1) et (2) de la proposition 3.7, avec un calcul de dimension élémentaire des espaces vectoriels Sym_n et Herm_n (déjà effectué dans la partie 1.2).

Rappelons que $I_{p,q}=\begin{pmatrix}-I_p & 0\\ 0 & I_q\end{pmatrix}$ où I_k est la matrice identité $k\times k$. Pour chacun des groupes matriciels G de la colonne de gauche ci-dessous, notons T_eG l'espace vectoriel réel [90] de la colonne du milieu (dont la dimension sur $\mathbb{R}$ est donnée par la colonne de droite).

G	T_eG	dimension sur $\mathbb{R}$
$\mathrm{GL}_n(\mathbb{R})$	$\mathscr{M}_n(\mathbb{R})$	n^2
$\mathrm{GL}_n(\mathbb{C})$	$\mathscr{M}_n(\mathbb{C})$	$2n^2$
$\mathrm{SL}_n(\mathbb{R})$	$\{X\in\mathscr{M}_n(\mathbb{R}) \,:\, \mathtt{tr}\ X=0\}$	n^2-1
$\mathrm{SL}_n(\mathbb{C})$	$\{X\in\mathscr{M}_n(\mathbb{C}) \,:\, \mathtt{tr}\ X=0\}$	$2n^2-2$
$\mathrm{O}(n)$	$\{X\in\mathscr{M}_n(\mathbb{R}) \,:\, {}^tX=-X\}$	$\frac{n(n-1)}{2}$
$\mathrm{SO}(n)$	$\{X\in\mathscr{M}_n(\mathbb{R}) \,:\, {}^tX=-X\}$	$\frac{n(n-1)}{2}$
$\mathrm{O}(p,q)$	$\{X\in\mathscr{M}_{p+q}(\mathbb{R}) \,:\, {}^tX\,I_{p,q}+I_{p,q}\,X=0\}$	$\frac{(p+q)(p+q-1)}{2}$
$\mathrm{SO}(p,q)$	$\{X\in\mathscr{M}_{p+q}(\mathbb{R}) \,:\, {}^tX\,I_{p,q}+I_{p,q}\,X=0\}$	$\frac{(p+q)(p+q-1)}{2}$
$\mathrm{U}(n)$	$\{X\in\mathscr{M}_n(\mathbb{C}) \,:\, X^*=-X\}$	n^2
$\mathrm{SU}(n)$	$\{X\in\mathscr{M}_n(\mathbb{C}) \,:\, X^*=-X,\ \ \mathtt{tr}\ X=0\}$	n^2-1
$\mathrm{U}(p,q)$	$\{X\in\mathscr{M}_{p+q}(\mathbb{C}) \,:\, X^*\,I_{p,q}+I_{p,q}\,X=0\}$	$\frac{(p+q)(p+q-1)}{2}$
$\mathrm{SO}(p,q)$	$\{X\in\mathscr{M}_{p+q}(\mathbb{C}) \,:\, X^*\,I_{p,q}+I_{p,q}\,X=0,\ \ \mathtt{tr}\,X=0\}$	$\frac{(p+q)(p+q-1)}{2}$

Nous renvoyons par exemple [91] à [MT, Chap. 3] ou [Deh, §VII.7] pour une démonstration de l'affirmation que si $G=\mathrm{SL}_n(\mathbb{C}),\mathrm{SL}_n(\mathbb{R}),\mathrm{O}(p,q),\mathrm{SO}(p,q),\mathrm{U}(p,q),\mathrm{SU}(p,q)$ (avec $n=p+q$ dans ces quatre derniers cas), et si $G=\mathrm{Sp}_{\frac{n}{2}}(\mathbb{C}),\mathrm{Sp}_{\frac{n}{2}}(\mathbb{R})$ lorsque n est pair, alors l'application $(x,y)\mapsto xy$ induit un homéomorphisme

$$(\mathrm{U}(n)\cap G)\times(\mathrm{Herm}_n^{++}\cap G)\simeq G \tag{28}$$

90. Il s'interprète comme l'espace tangent en l'élément neutre de la sous-variété différentielle G de $\mathrm{GL}_N(\mathbb{K})$ pour $(\mathbb{K},N)$ approprié, voir par exemple [Pau2], [Cha, Chap. 6], [Laf]. Mais nous n'aurons pas besoin de ce fait.

91. Voir aussi le problème 11 page 63 de [Zav] pour une démonstration directe de la décomposition polaire de $\mathrm{U}(p,q)$ et de $\mathrm{O}(p,q)$.

(donc $(\mathrm{O}(n)\cap G)\times(\mathrm{Sym}_n^{++}\cap G)\simeq G$ lorsque $G\subset\mathrm{GL}_n(\mathbb{R})$), et que l'application exponentielle induit un homéomorphisme

$$(\mathrm{Herm}_n\cap T_eG)\simeq(\mathrm{Herm}_n^{++}\cap G)$$

(donc $(\mathrm{Sym}_n\cap T_eG)\simeq(\mathrm{Sym}_n^{++}\cap G)$ lorsque $G\subset\mathrm{GL}_n(\mathbb{R})$).

Un calcul par blocs montre que $\begin{pmatrix}A&B\\C&D\end{pmatrix}\in\mathrm{O}(p+q)\cap\mathrm{O}(p,q)$ si et seulement si $A\in\mathrm{O}(p)$, $B=0$, $C=0$ et $D\in\mathrm{O}(q)$. De même,

$$\begin{pmatrix}A&B\\C&D\end{pmatrix}\in\mathrm{Sym}_{p+q}\cap\{X\in\mathscr{M}_n(\mathbb{R})\ :\ {}^tXI_{p,q}+I_{p,q}X=0\}$$

si et seulement si $A=0$, $D=0$ et $C={}^tB$ (et B est une matrice réelle $p\times q$ quelconque).

Un calcul par blocs montre que $\begin{pmatrix}A&B\\C&D\end{pmatrix}\in\mathrm{U}(p+q)\cap\mathrm{U}(p,q)$ si et seulement si $A\in\mathrm{U}(p)$, $B=0$, $C=0$ et $D\in\mathrm{U}(q)$. De même,

$$\begin{pmatrix}A&B\\C&D\end{pmatrix}\in\mathrm{Herm}_{p+q}\cap\{X\in\mathscr{M}_n(\mathbb{C})\ :\ X^*I_{p,q}+I_{p,q}X=0\}$$

si et seulement si $A=0$, $D=0$ et $C=B^*$ (et B est une matrice complexe $p\times q$ quelconque). La colonne centrale du tableau du corollaire 3.12 s'en déduit.

Pour tous les groupes topologiques G et les $k\in\mathbb{N}$, l'application de π_0G dans $\pi_0(G\times\mathbb{R}^k)$ qui à une composante connexe par arcs C de G associe $C\times\mathbb{R}^k$ est une bijection de π_0G dans $\pi_0(G\times\mathbb{R}^k)$. Par la décomposition polaire, l'inclusion des groupes topologiques

$$\mathrm{O}(n),\ \ \mathrm{U}(n),\ \ \mathrm{SO}(n),\ \ \mathrm{SU}(n),\ \ \mathrm{O}(p)\times\mathrm{O}(p),\ \ S(\mathrm{O}(p)\times\mathrm{O}(p))$$

dans

$$\mathrm{GL}_n(\mathbb{R}),\ \ \mathrm{GL}_n(\mathbb{C}),\ \ \mathrm{SL}_n(\mathbb{R}),\ \ \mathrm{SL}_n(\mathbb{C}),\ \ \mathrm{O}(p,q),\ \ \mathrm{SO}(p,q)$$

respectivement est donc un morphisme de groupes qui préserve les composantes connexes par arcs. Pour montrer la colonne de droite du tableau ci-dessus, il suffit donc de montrer que $\pi_0\,\mathrm{O}(n)\simeq\mathbb{Z}/2\mathbb{Z}$, $\pi_0\,\mathrm{SO}(n)=1$, $\pi_0\,\mathrm{U}(n)=1$, et $\pi_0\,\mathrm{SU}(n)=1$.

Montrons que $\mathrm{U}(n)$ est connexe par arcs. Nous avons vu dans la partie 3.3 que pour tout $x\in\mathrm{U}(n)$, il existe $P\in\mathrm{U}(n)$ et $\theta_1,\cdots,\theta_n\in\mathbb{R}$ tels que $x=PD(\theta_1,\cdots,\theta_n)P^{-1}$, où

$$D(\theta_1,\cdots,\theta_n)=\begin{pmatrix}e^{i\theta_1}&0&0\\0&\ddots&0\\0&0&e^{i\theta_n}\end{pmatrix}.$$

L'application $c:[0,1]\to\mathrm{U}(n)$ définie par $t\mapsto PD(t\theta_1,\cdots,t\theta_n)P^{-1}$ est alors un chemin continu dans $\mathrm{U}(n)$ partant de l'élément neutre, et arrivant en x.

Comme $PD(\theta_1,\cdots,\theta_n)P^{-1}$ appartient à $\mathrm{SU}(n)$ si et seulement si $\theta_1+\cdots+\theta_n=0$, et puisqu'alors $t\theta_1+\cdots+t\theta_n=0$ pour tout $t\in[0,1]$, l'image de ce chemin c est contenue dans $\mathrm{SU}(n)$ si son extrémité $c(1)$ appartient à $\mathrm{SU}(n)$. Donc $\mathrm{SU}(n)$ est connexe par arcs.

Montrons que $\mathrm{SO}(n)$ est connexe par arcs. Nous avons vu dans la partie 2.3 que tout élément de $\mathrm{O}(n)$ est conjugué dans $\mathrm{O}(n)$ à un élément de la forme

$$D'(p,q,\theta_1,\cdots,\theta_r) = \begin{pmatrix} -I_p & 0 & 0 & 0 & 0 \\ 0 & I_q & 0 & 0 & 0 \\ 0 & 0 & \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix} & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \begin{pmatrix} \cos\theta_r & -\sin\theta_r \\ \sin\theta_r & \cos\theta_r \end{pmatrix} \end{pmatrix}$$

où $p,q,r \in \mathbb{N}$ vérifient $p+q+2r=n$ et $\theta_1,\cdots,\theta_r \in \mathbb{R}$, en notant I_k la matrice identité $k\times k$ pour tout $k\in\mathbb{N}$. De plus, nous avons vu que $D'(p,q,\theta_1,\cdots,\theta_r)$ appartient à $\mathrm{SO}(n)$ si et seulement si p est pair. En écrivant $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \cos\pi & -\sin\pi \\ \sin\pi & \cos\pi \end{pmatrix}$, ceci montre que pour tout élément x de $\mathrm{SO}(n)$, il existe $P\in\mathrm{O}(n)$, $p,r\in\mathbb{N}$ et $\theta_1,\cdots,\theta_r \in \mathbb{R}$ tels que $x = PD'(0,p,\theta_1,\cdots,\theta_r)P^{-1}$. Donc l'application $c:[0,1]\to\mathrm{SO}(n)$ définie par $t\mapsto PD'(0,p,t\theta_1,\cdots,t\theta_r)P^{-1}$ est un chemin continu dans $\mathrm{SO}(n)$ entre l'élément neutre et x. Donc $\mathrm{SO}(n)$ est connexe par arcs.

Si $x_0 = \begin{pmatrix} -1 & 0 \\ 0 & I_{n-1} \end{pmatrix}$, alors $x_0\in\mathrm{O}(n)$. Comme l'application det est continue, à valeurs dans l'espace discret $\{\pm1\}$ sur $\mathrm{O}(n)$, et puisque $\det(x_0)=-1$, les composantes connexes par arcs de $\mathrm{O}(n)$ contenant x_0 et I_n sont distinctes. Puisque $\mathrm{SO}(n)$ est connexe par arcs et puisque $\mathrm{O}(n) = \mathrm{SO}(n) \coprod x_0\,\mathrm{SO}(n)$, il y a exactement deux composantes connexes par arcs dans $\mathrm{O}(n)$. Donc l'application qui à la composante connexe par arc de x associe $\det x$ est un isomorphisme de groupes de $\pi_0\,\mathrm{O}(n)$ dans $\mathbb{Z}/2\mathbb{Z}$. □

Exercice E.34. *(1) Soit E un espace euclidien ou hermitien de dimension finie, de norme $\|\ \|$. Soient K un sous-groupe compact de $\mathrm{GL}(E)$ et C un convexe compact non vide de E invariant par K. Le but de cette question est de montrer le* théorème du point fixe de Kakutani *qui dit qu'il existe un point fixe par K dans C.*

- *Montrer que l'application de E dans $\mathbb{R}$ définie par $x\mapsto \|x\|_K = \max_{g\in K}\|gx\|$ est une norme, et que $\|x+y\|_K = \|x\|_K + \|x\|_K$ si et seulement si x et y sont positivement dépendants (c'est-à-dire s'il existe $\lambda_1,\lambda_2\in[0,+\infty[$ non tous deux nuls tels que $\lambda_1 x+\lambda_2 y = 0$).*
- *Montrer que tout point $x_0\in C$ réalisant le minimum de l'application de C dans $\mathbb{R}$ définie par $x\mapsto\|x\|_K$ est fixe par K.*

(2) Montrer que tout sous-groupe compact K_0 de $\mathrm{GL}_n(\mathbb{R})$ ou $\mathrm{GL}_n(\mathbb{C})$ préserve un produit scalaire euclidien sur $\mathbb{R}^n$ ou hermitien sur $\mathbb{C}^n$, respectivement.[92] *En déduire que $\mathrm{O}(n)$ est l'unique à conjugaison près sous-groupe compact maximal (pour l'inclusion) dans $\mathrm{GL}_n(\mathbb{R})$, et que $\mathrm{U}(n)$ est l'unique à conjugaison près sous-groupe compact maximal (pour l'inclusion) dans $\mathrm{GL}_n(\mathbb{C})$.*

92. Il existe aussi une démonstration de cette affirmation qui utilise le théorème de l'ellipsoïde de John, voir par exemple [FGN], et une démonstration qui utilise l'existence d'une *mesure de Haar* sur le groupe topologique compact K_0, c'est-à-dire d'une mesure borélienne de probabilité sur K_0 invariante par translations à gauche par K_0, mais il faudrait alors la construire.

3.6 Exercices supplémentaires

Exercice E.35. Montrer que tout élément $A \in \mathrm{GL}_n(\mathbb{C})$ s'écrit de manière unique $A = DU$ avec $D, U \in \mathrm{GL}_n(\mathbb{C})$, D diagonalisable, U unipotente et D et U commutantes. Montrer que de plus D et U sont des polynômes en A. Ce résultat s'appelle le théorème de la décomposition de Dunford multiplicative. Calculer la décomposition de Dunford multiplicative de $\exp A$.

Exercice E.36. Montrer que tout sous-groupe fini de $\mathrm{GL}_n(\mathbb{R})$ (respectivement $\mathrm{GL}_n(\mathbb{C})$) est conjugué à un sous-groupe de $\mathrm{O}(n)$ (respectivement $\mathrm{U}(n)$).

Exercice E.37. Soit E un espace vectoriel réel de dimension n, muni de la topologie définie par n'importe laquelle de ses normes. Montrer que l'ensemble des bases de E est un ouvert de l'espace topologique produit E^n qui est homéomorphe à $\mathbb{R}^{n(n+1)/2} \times \mathrm{O}(n)$. Si E est orienté, montrer que l'ensemble des bases positives de E est un ouvert de E^n qui est homéomorphe à $\mathbb{R}^{n(n+1)/2} \times \mathrm{SO}(n)$. Si E est un espace euclidien, montrer que l'ensemble de ses bases orthonormées est une partie de E^n homéomorphe à $\mathrm{O}(n)$. Si E est un espace euclidien orienté, montrer que l'ensemble de ses bases orthonormées positives est une partie de E^n homéomorphe à $\mathrm{SO}(n)$.

Exercice E.38. Soit $A \in \mathscr{M}_n(\mathbb{C})$ une matrice hermitienne définie positive. Montrer qu'il existe une matrice triangulaire supérieure P telle que $A = P^*P$.

Exercice E.39. (L'inégalité d'Hadamard) Soient $n \in \mathbb{N} - \{0\}$ et $B \in \mathrm{GL}_n(\mathbb{C})$. Munissons $\mathbb{C}^n$ de la norme hermitienne usuelle, et identifions les éléments de $\mathbb{C}^n$ avec les vecteurs colonnes de leurs coordonnées dans la base canonique de $\mathbb{C}^n$. Montrer *l'inégalité d'Hadamard*

$$|\det B| \leq \prod_{i=1}^{n} \|C_i\|$$

où les C_i sont les colonnes de B. Quand y a-t-il égalité ?

Exercice E.40. (Matrices de Householder) Soit $n \in \mathbb{N} - \{0\}$. Le but de cet exercice est de donner un algorithme pour écrire tout élément de $\mathrm{O}(n)$ et $\mathrm{U}(n)$ comme produit de réflexions orthogonales.

Si X est un vecteur unitaire de l'espace euclidien (respectivement hermitien) usuel $\mathbb{R}^n$ (respectivement $\mathbb{C}^n$), identifié au vecteur colonne de ses coordonnées dans la base canonique, notons $Q_X = I_n - 2XX^*$ (en remarquant que $A^* = {}^tA$ si A est une matrice réelle). Une *matrice (de réflexion) de Householder* est une matrice M dans $\mathscr{M}_n(\mathbb{R})$ (respectivement $\mathscr{M}_n(\mathbb{C})$) telle qu'il existe un vecteur unitaire X tel que $M = Q_X$.

a) Montrer qu'un tel vecteur est unique modulo multiplication par un scalaire de valeur absolue 1. Montrer qu'une matrice de Householder est symétrique et orthogonale (respectivement hermitienne et unitaire). Montrer que si $P \in \mathrm{O}(n)$ (respectivement $P \in \mathrm{U}(n)$), alors $PQ_XP^{-1} = Q_{PX}$. Calculer le déterminant d'une matrice de Householder.

b) Montrer que si X est le premier vecteur colonne d'une matrice $A \in \mathrm{O}(n)$ (respectivement $A \in \mathrm{U}(n)$), si e_1 est le premier vecteur de la base canonique, alors il existe un élément $\alpha \in \mathbb{U} = \{z \in \mathbb{C} : |z| = 1\}$ tel que si $X' = \frac{X - \alpha e_1}{\|X - \alpha e_1\|}$, alors $Q_{X'}A$ est triangulaire supérieure en blocs $(1, n-1)$.

c) Montrer que les matrices de Householder engendrent $\mathrm{O}(n)$, et que tout élément de $\mathrm{SO}(n)$ est le produit d'un nombre pair inférieur à n de matrices de Householder.

d) Montrer que les matrices de Householder et les matrices diagonales de $\mathrm{U}(n)$ engendrent $\mathrm{U}(n)$.

Exercice E.41. (Introduction aux groupes de réflexions complexes) Considérons des entiers $n, d, e \in \mathbb{N} - \{0\}$. Pour toute bijection $\sigma \in \mathfrak{S}_n$, notons P_σ la matrice de la permutation de la base canonique de $\mathbb{C}^n$ définie par σ, et identifions $\mathfrak{S}_n$ avec le sous-groupe de $\mathrm{GL}_n(\mathbb{C})$ constitué des matrices de permutations par l'application $\sigma \mapsto P_\sigma$. Notons D le groupe des matrices diagonales dont les coefficients diagonaux $\xi_1, \dots, \xi_n$ vérifient

$$(\xi_1 \dots \xi_n)^d = 1 \quad \text{et} \quad \forall\, i = 1, \dots, n, \quad (\xi_i)^{de} = 1\ .$$

Notons G le sous-groupe de $\mathrm{GL}_n(\mathbb{C})$ engendré par $\mathfrak{S}_n$ et D.

(1) Pour $a_1, \dots, a_n \in \mathbb{C}^\times$, chercher les valeurs propres et vecteurs propres de

$$\begin{pmatrix} 0 & \dots & 0 & a_n \\ a_1 & \ddots & & 0 \\ & \ddots & \ddots & \vdots \\ 0 & & a_{n-1} & 0 \end{pmatrix}.$$

(2) Pour tout $A = (a_{i,j})_{1 \le i,j \le n} \in \mathscr{M}_n(\mathbb{C})$, calculer $P_\sigma A (P_\sigma)^{-1}$.

(3) Montrer que le groupe G est isomorphe au produit semi-direct de $\mathfrak{S}_n$ et de D, et en déduire son ordre.

(4) Montrer que G est engendré par des réflexions complexes, que l'on déterminera. [93]

Exercice E.42. (Le revêtement universel $\mathrm{SU}(2) \to \mathrm{SO}(3)$) Notons

$$\mathfrak{su}(2) = \{X \in \mathscr{M}_2(\mathbb{C}) \ :\ X^* + X = 0,\ \ \mathtt{tr}\, X = 0\}$$

l'espace vectoriel réel des matrices anti-hermitiennes de trace nulle de taille 2.

(1) Montrer que le groupe $\mathrm{SU}(2)$ agit linéairement sur $\mathfrak{su}(2)$ par l'application définie par $A \mapsto \{X \mapsto AXA^{-1}\}$.

(2) Montrer que le triplet

$$\left(\xi_1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix},\quad \xi_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},\quad \xi_3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\right)$$

est une base de l'espace vectoriel réel $\mathfrak{su}(2)$, et que l'application $\sqrt{\det}$ est une norme euclidienne sur $\mathfrak{su}(2)$, rendant orthonormée la base (ξ_1, ξ_2, ξ_3), invariante par l'action de $\mathrm{SU}(2)$.

(3) Montrer que l'application Φ de $\mathrm{SU}(2)$ dans $\mathrm{GL}_3(\mathbb{R})$ qui à A associe la matrice de l'application linéaire $X \mapsto AXA^{-1}$ dans la base (ξ_1, ξ_2, ξ_3) est un morphisme de groupes continu de $\mathrm{SU}(2)$ dans $\mathrm{SO}(3)$, de noyau $\{\pm \mathrm{id}\}$.

93. Un théorème de classification de Coxeter (voir par exemple [Bou, Hum]) classe à conjugaison près les sous-groupes finis de $\mathrm{GL}_n(\mathbb{R})$ engendrés par des réflexions orthogonales, dont le groupe des permutations de la base canonique est un exemple typique. Un résultat analogue de Shephard-Todd [ST] en 1954 classe à conjugaison près les sous-groupes finis de $\mathrm{GL}_n(\mathbb{C})$ engendrés par des réflexions complexes. Les groupes ci-dessus pour $d, e \in \mathbb{N} - \{0\}$ sont des exemples typiques.

(4) Montrer que l'application

$$\exp : \mathfrak{so}(n) = \{x \in \mathscr{M}_n(\mathbb{R}) \ : \ {}^t x + x = 0\} \ \to \ \mathrm{SO}(n)$$

est surjective.

(5) Considérons l'application $T\Phi$ de $\mathfrak{su}(2)$ dans $\mathscr{M}_3(\mathbb{R})$, qui à $X \in \mathfrak{su}(2)$ associe la matrice dans la base (ξ_1, ξ_2, ξ_3) de l'endomorphisme linéaire $Y \mapsto XY - YX$ de $\mathfrak{su}(2)$. Calculer $T\Phi(x_1\xi_1 + x_2\xi_2 + x_3\xi_3)$ pour tous les $x_1, x_2, x_3 \in \mathbb{R}$. Montrer que $T\Phi$ est un isomorphisme linéaire de $\mathfrak{su}(2)$ dans $\mathfrak{so}(3)$, et que le diagramme suivant est commutatif

$$\begin{array}{ccc} \mathfrak{su}(2) & \xrightarrow{T\Phi} & \mathfrak{so}(3) \\ \exp \downarrow & & \downarrow \exp \\ \mathrm{SU}(2) & \xrightarrow{\Phi} & \mathrm{SO}(3) \end{array} .$$

On pourra utiliser le fait (voir la Proposition 3.4 (7)) que si E est un espace vectoriel réel de dimension finie, alors l'application de $\mathscr{L}(E)$ dans l'ensemble des sous-groupes différentiables à un paramètre de $\mathrm{GL}(E)$ définie par $X \mapsto (\exp(tX))_{t\in\mathbb{R}}$ est une bijection d'inverse l'application $(g_t)_{t\in\mathbb{R}} \mapsto \frac{d}{dt}_{|t=0} g_t$. En déduire que Φ est surjectif.

3.7 Indications pour la résolution des exercices

Correction de l'exercice E.26. Si $\rho = s_{D,\zeta}$ est une réflexion complexe non triviale, notons $\check{r}$ une forme linéaire non nulle de noyau l'hyperplan vectoriel $D^\perp$ orthogonal à D, et r l'unique élément (non nul car $\zeta \neq 1$) de D tel que $\check{r}(r) = 1 - \zeta$. Alors l'application $x \mapsto x - \check{r}(x)r$ vaut l'identité sur $D^\perp$. Pour tout $\lambda \in \mathbb{C}$, elle envoie le vecteur λr sur $\lambda r - \check{r}(\lambda r)r = \lambda(1 - \check{r}(r))r = \zeta\lambda r$, donc vaut $\zeta\,\mathrm{id}$ sur D. Par conséquent, elle coïncide avec ρ.

Correction de l'exercice E.27. Notons $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{R}$, et E l'espace euclidien ou hermitien usuel $\mathbb{K}^n$, muni de sa base canonique $\mathscr{B} = (e_1, \dots, e_n)$. Pour tout $k = 1, \dots, n$, notons E_k le sous-espace vectoriel de E engendré par les k premiers vecteurs de la base $\mathscr{B}$. Notons q la forme quadratique ou quadratique hermitienne de E dont la matrice dans la base $\mathscr{B}$ est égale à A.

Si q est définie positive, alors pour tout $k = 1, \dots, n$, la restriction de q au sous-espace vectoriel E_k est encore définie positive sur E_k. Donc le mineur principal d'ordre k de A, qui est le déterminant de la matrice A_k de $q|_{E_k}$ dans la base $(e_1, \dots, e_k)$, est strictement positif.

Montrons la réciproque par récurrence sur la dimension n, le cas $n = 1$ étant immédiat (il est aussi possible de commencer à $n = 0$). Soit $n \geq 2$. Par récurrence, la restriction de la forme q à E_{n-1} est définie positive. Puisque le déterminant de A est non nul, la forme q est non dégénérée. Donc l'orthogonal $(E_{n-1})^\perp$ de l'hyperplan vectoriel E_{n-1} est une droite vectorielle supplémentaire de E_{n-1}, engendrée par un vecteur e'_n. Soit P la matrice de passage de la base $\mathscr{B}$ à la base $\mathscr{B}' = (e_1, \dots, e_{n-1}, e'_n)$. Par la formule (2), nous avons $\begin{pmatrix} A_{n-1} & 0 \\ 0 & q(e'_n) \end{pmatrix} = {}^tP\,A\,\overline{P}$. Donc en prenant le déterminant, nous obtenons que $q(e'_n)\det A_{n-1} = |\det P|^2 \det A$. Puisque $q(e'_n)$ a le même signe que $\frac{\det A}{\det A_{n-1}} > 0$, la restriction de q à la droite $\mathbb{K}e'_n$ est aussi définie positive. Par la décomposition en somme directe orthogonale $E = E_{n-1} \oplus \mathbb{K}e'_n$, la forme quadratique ou quadratique hermitienne q est donc définie positive sur E.

Correction de l'exercice E.30. (1) Ceci découle, par combinaison linéaire et passage à la limite, du fait que pour tout $\ell \in \mathbb{N}$, nous avons

$$\begin{pmatrix} A_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & A_k \end{pmatrix}^\ell = \begin{pmatrix} (A_1)^\ell & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & (A_k)^\ell \end{pmatrix}.$$

(2) Comme les matrices d'homothétie commutent avec toutes les matrices dans $\mathscr{M}_n(\mathbb{K})$, et puisque $J_{\lambda,t} = \lambda I_n + J_{0,t}$, nous avons

$$\exp(J_{\lambda,t}) = \exp(\lambda I_n + J_{0,t}) = \exp(\lambda I_n)\exp(J_{0,t}) = e^\lambda \exp(J_{0,t})\,.$$

En calculant les puissances de $J_{0,t}$ d'ordre $k \leq n-1$, dont les coefficients sont nuls sauf

ceux sur la k-ème sur-diagonale, égaux à t^k, nous avons

$$\exp(J_{0,t}) = \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{n-1}}{(n-1)!} \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \frac{t^2}{2!} \\ \vdots & & & \ddots & t \\ 0 & \cdots & \cdots & \cdots & 1 \end{pmatrix}.$$

(3) Pour tout $n \geq 1$, l'application $\exp : \mathscr{M}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ n'est pas injective car $\exp I_n = \exp \begin{pmatrix} 2i\pi & 0 \\ 0 & I_{n-1} \end{pmatrix}$ par la question (1) et puisque $\exp(2ik\pi) = 1$ pour tout $k \in \mathbb{Z}$.

(4) Puisque $\exp(PyP^{-1}) = P\exp(y)P^{-1}$ et puisque la classe de conjugaison de la matrice identité est réduite à l'identité, le problème est invariant à conjugaison près.

Puisque $e^{2i\pi k} = 1$ pour tout $k \in \mathbb{Z}$, il est immédiat par la question (1) appliquée avec $n_1 = \cdots = n_k = 1$ que si x est diagonalisable de valeurs propres dans $2i\pi\mathbb{Z}$, alors $\exp x = \mathrm{id}$.

Si x est diagonalisable et admet une valeur propre qui n'appartient pas à $2i\pi\mathbb{Z}$, alors $\exp x$ admet une valeur propre différente de 1 par la question (1) appliquée avec $n_1 = \cdots = n_k = 1$, donc $\exp x \neq \mathrm{id}$.

Enfin, si x n'est pas diagonalisable, puisque l'exponentielle d'une matrice diagonale par blocs $\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_k \end{pmatrix}$ est la matrice diagonale par blocs $\begin{pmatrix} \exp(A_1) & & 0 \\ & \ddots & \\ 0 & & \exp(A_k) \end{pmatrix}$ par la question (1), nous pouvons supposer que $x = \lambda\,\mathrm{id} + J$ est un bloc de Jordan de taille $n \geq 2$ et de valeur propre λ, avec $J = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}$. Auquel cas (voir la question (2)), puisque $\lambda\,\mathrm{id}$ et J commutent, nous avons $\exp x = e^{\lambda} \sum_{k=0}^{n-1} \frac{1}{p!} J^p$, qui n'est pas diagonale car $n \geq 2$, donc $\exp x \neq \mathrm{id}$.

(5) Si $n = 1$, alors l'application $\exp : \mathbb{R} \to]0, +\infty[$ est injective, car c'est un C^∞-difféomorphisme (d'inverse le logarithme). Si $n = 2$, et donc si $n \geq 2$ en considérant des matrices diagonales par blocs et en utilisant la question (1), pour tout $k \in \mathbb{Z}$, la matrice réelle $2k\pi \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, qui est diagonalisable sur $\mathbb{C}$ de valeurs propres $\pm 2ik\pi$, a pour image par l'exponentielle la matrice identité, par la question (4).

(6) Ceci découle du fait que $\mathbb{K}[A]$ est un sous-espace vectoriel de l'espace vectoriel normé de dimension finie $\mathscr{M}_n(\mathbb{K})$, donc est fermé, et que $\exp A$ est une limite dans $\mathscr{M}_n(\mathbb{K})$ d'éléments de $\mathbb{K}[A]$. Montrons plus précisément que $\exp(A) \in \mathbb{K}_{n-1}[A]$.

Par le théorème de Cayley-Hamilton, si χ_A est le polynôme caractéristique de A, alors l'égalité $\chi_A(A) = 0$ permet d'exprimer A^n comme polynôme de degré au plus $n-1$ en A à coefficients dans $\mathbb{K}$. Donc par récurrence et linéarité, pour tout $k \in \mathbb{N}$, la somme $S_k = \sum_{i=0}^{k} \frac{1}{i!} A^i$ appartient au sous-espace vectoriel $\mathbb{K}_{n-1}[A]$ des polynômes de degré au

plus $n-1$ en A. Puisque le sous-espace vectoriel $\mathbb{K}_{n-1}[A]$ de l'espace vectoriel normé $\mathscr{M}_n(\mathbb{K})$ de dimension finie est fermé, la limite $\exp(A) = \lim_{k\to+\infty} S_n$ appartient aussi à $\mathbb{K}_{n-1}[A]$.

Correction de l'exercice E.31. Munissons $\mathscr{M}_n(\mathbb{R})$ de n'importe quelle norme matricielle $\|\ \|$. Par la proposition 3.4 (6), l'application exponentielle est un homéomorphisme local d'un voisinage ouvert V' de 0 dans $\mathscr{M}_n(\mathbb{K})$ à valeurs dans un voisinage ouvert U' de l'élément neutre dans $\mathrm{GL}_n(\mathbb{K})$. Soit $\epsilon > 0$ tel que $B(0,2\epsilon) \subset V'$, et posons $U = \exp(B(0,\epsilon))$. Si par l'absurde il existe un élément y non trivial de U dont toutes les puissances restent dans U, soit $x \in B(0,\epsilon)$ tel que $y = \exp x$. Soit $k \in \mathbb{N}$ tel que $kx \in B(0,2\epsilon) - B(0,\epsilon)$ qui existe car $0 < \|x\| < \epsilon$ et $\|kx\| = k\|x\|$. Alors $y^k = \exp(kx) \notin U$, une contradiction.

Correction de l'exercice E.32. La correction des questions (1) et (2) est issue du livre [Zav, page 48]. Voir [Djo] pour des développements plus poussés.

(1) Soit $A \in \mathscr{M}_n(\mathbb{C})$.

- Il est immédiat que $\mathbb{C}[A]^\times$ est contenu dans $\mathbb{C}[A] \cap \mathrm{GL}_n(\mathbb{R})$. Réciproquement, soit $B \in \mathbb{C}[A] \cap \mathrm{GL}_n(\mathbb{R})$. Par le théorème de Cayley-Hamilton, si $P = \sum_{i=0}^n a_i\ X^i$ est le polynôme caractéristique de B, alors $a_0 \neq 0$ car B est inversible, et si $B' = -\sum_{i=1}^n \frac{a_i}{a_0}\ B^{i-1}$, alors $B' \in \mathbb{C}[A]$ et $BB' = B'B = I_n$, donc $B \in \mathbb{C}[A]^\times$.

- Soit $N \in \mathbb{C}[A]$. Puisque $\mathbb{C}[A]$ est un sous-espace vectoriel de l'espace vectoriel complexe de dimension finie $\mathscr{M}_n(\mathbb{C})$, il est fermé. Donc $\exp(N) = \lim_{n\in\mathbb{N}} \sum_{i=0}^n \frac{1}{i!}\ A^i \in \mathbb{C}[A]$, et puisque $\exp(N) \in \mathrm{GL}_N(\mathbb{C})$, l'inclusion $\exp(\mathbb{C}[A]) \subset \mathbb{C}[A]^\times$ découle du premier point.

- Pour tous les $M_0, M_1 \in \mathbb{C}[A]^\times$ et $z \in \mathbb{C}$, la matrice $M(z) = zM_0 + (1-z)M_1$ appartient à $\mathbb{C}[A]$ et son déterminant, qui est un polynôme d'une variable en z, n'a qu'un nombre fini de racines, qui ne sont pas 0 et 1. Par connexité par arc de $\mathbb{C}$ privé d'un nombre fini de points, il existe donc un chemin $c : [0,1] \to \mathbb{C}$ entre 0 et 1 évitant les racines de $\det M(z)$. Alors $t \mapsto M(c(t))$ est un chemin continu entre M_0 et M_1 contenu dans $\mathbb{C}[A] \cap \mathrm{GL}_n(\mathbb{R}) = \mathbb{C}[A]^\times$. Ceci montre que $\mathbb{C}[A]^\times$ est connexe par arcs.

- L'application $\exp : \mathbb{C}[A] \to \mathbb{C}[A]$ est de classe C^∞ dans l'espace vectoriel complexe de dimension finie $\mathbb{C}[A]$, et sa différentielle en 0 est inversible (c'est l'identité de $\mathbb{C}[A]$). Par le théorème d'inversion locale, il existe un voisinage ouvert V de 0 dans $\mathbb{C}[A]$ et un voisinage ouvert U de I_n dans $\mathbb{C}[A]$ tel que $\exp : V \to U$ soit un homéomorphisme. Puisque deux éléments de $\mathbb{C}[A]$ commutent, nous avons donc pour tout $B \in \mathbb{C}[A]$,

$$\exp(B+V) = \exp(B)\exp(V) = \exp(B)\ U\ ,$$

qui, par la continuité de la multiplication à gauche par $\exp(B)$, est un voisinage ouvert de $\exp(B)$ dans $\mathbb{C}[A]$. Donc $\exp(\mathbb{C}[A])$ est ouvert dans $\mathbb{C}[A]$, et contenu dans $\mathbb{C}[A]^\times$, donc ouvert dans $\mathbb{C}[A]^\times$.

Remarque. Puisque l'application déterminant est polynomiale, donc continue, la partie $\mathbb{C}[A]^\times = \{B \in \mathbb{C}[A]\ :\ \det B \neq 0\}$ est ouverte dans $\mathbb{C}[A]$.

- Montrons que le complémentaire X de $\exp(\mathbb{C}[A])$ dans $\mathbb{C}[A]^\times$ est ouvert dans $\mathbb{C}[A]^\times$. Soit $B \in X$, montrons que $B\exp(\mathbb{C}[A])$, qui est un voisinage ouvert de B dans $\mathbb{C}[A]^\times$, est contenu dans X. Sinon, il existe $C, D \in \mathbb{C}[A])$ tel que $B\exp(C) = \exp(D)$. Puisque C et D commutent, nous avons donc $B = \exp(D-C)$, ce qui contredit le fait que $B \in X$.

- La partie $\exp(\mathbb{C}[A])$ est ouverte, fermée et non vide dans l'espace topologique connexe $\mathbb{C}[A]^\times$. Elle est donc égale à $\mathbb{C}[A]^\times$.

La conclusion de la question (1) vient du fait que si $A \in \mathrm{GL}_n(\mathbb{C})$, alors $A \in \mathbb{C}[A]^\times$, donc A appartient à l'image de l'exponentielle des matrices complexes.

(2) Les éléments $A = \begin{pmatrix} -1 & 0 \\ 0 & I_{n-1} \end{pmatrix}$ et $A' = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & I_{n-2} \end{pmatrix}$ ne sont pas dans l'image de l'exponentielle des matrices réelles, comme nous le verrons dans la question (3). Voici une démonstration directe pour A. Par l'absurde, s'il existe $B \in \mathscr{M}_n(\mathbb{R})$ telle que $A = \exp(B)$, soient $\lambda_1, \dots, \lambda_n$ les valeurs propres complexes avec multiplicités de B. Quitte à permuter, nous avons $\exp \lambda_1 = -1$ et $\exp \lambda_2 = \cdots = \exp \lambda_n = 1$. Donc $\lambda_1 \in \pi i + 2i\pi\mathbb{Z}$ et $\lambda_k \in 2i\pi\mathbb{Z}$ pour $k = 2, \dots, n$, ce qui implique en particulier que $\lambda_k \neq \overline{\lambda_1}$. Ceci contredit le fait que le spectre (l'ensemble des valeurs propres complexes) d'une matrice réelle est stable par la conjugaison complexe.

Soit $A \in \mathrm{GL}_n(\mathbb{R})$. S'il existe $B \in \mathscr{M}_n(\mathbb{R})$ tel que $A = \exp(B)$, alors $A = (\exp \frac{B}{2})^2$. Réciproquement, s'il existe $B \in \mathrm{GL}_n(\mathbb{R})$ tel que $A = B^2$, alors il existe $C \in M_n(\mathbb{C})$ tel que $B = \exp(C)$ par la question (1). Puisque la matrice B est réelle, nous avons aussi $B = \overline{\exp(C)} = \exp(\overline{C})$. D'où, puisque C et $\overline{C}$ commutent et comme $C + \overline{C}$ est une matrice réelle,

$$A = B^2 = \exp(C)\exp(\overline{C}) = \exp(C + \overline{C}) \in \exp(\mathscr{M}_n(\mathbb{R}))\ .$$

(3) Pour toute matrice $\Lambda = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ où $a, b \in \mathbb{R}$ avec $b \neq 0$ et $k \in \{1, \dots, n\}$, appelons encore *bloc de Jordan* la matrice $J'_k(\Lambda) = \begin{pmatrix} \Lambda & I_2 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & I_2 \\ 0 & & & \Lambda \end{pmatrix} \in \mathscr{M}_{2k}(\mathbb{R})$. Rappelons (*décomposition de Jordan réelle*) que toute matrice $A \in \mathrm{GL}_n(\mathbb{R})$ est conjuguée dans $\mathrm{GL}_n(\mathbb{R})$ à une matrice diagonale par blocs dont les blocs diagonaux sont des blocs de Jordan $J_k(\lambda)$ avec $\lambda \in \mathbb{R}^*$ ou $J'_k(\Lambda)$ comme ci-dessus.

Si A est dans l'image de l'exponentielle, par la question (2), la matrice A est le carré d'une matrice B de $\mathrm{GL}_n(\mathbb{R})$. Si λ est une valeur propre réelle strictement négative de A, alors c'est le carré d'une valeur propre imaginaire pure (non nulle) λ' de B, et comme $\overline{\lambda'} = -\lambda$ est aussi une valeur propre de B, de même carré que λ', ceci montre que les blocs de Jordan $J_k(\lambda)$ de A de taille k donnée sont en nombre pairs.

Réciproquement, par le comportement de l'élévation au carré des matrices vis-à-vis de la conjugaison et des matrices diagonales par blocs, et par la question (2), il suffit de montrer qu'un bloc de Jordan $J_k(\lambda)$ avec $\lambda > 0$ ou un bloc de Jordan $J'_k(\Lambda)$ comme ci-dessus ou une matrice $\begin{pmatrix} J_k(\lambda) & 0 \\ 0 & J_k(\lambda) \end{pmatrix}$ avec $\lambda < 0$ est un carré. Toujours par l'invariance de l'élévation au carré par la conjugaison, ceci découle du fait que

- la décomposition de Jordan réelle de $J_k(\lambda')^2 = \begin{pmatrix} \lambda'^2 & 2\lambda' & 1 & & 0 \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & \ddots & 2\lambda' \\ 0 & & & & \lambda'^2 \end{pmatrix}$ est $J_k(\lambda'^2)$

si $\lambda' > 0$,

- toute matrice réelle $\Lambda = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ avec $b \neq 0$ s'écrit

$$\Lambda = \rho \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \Lambda'^2$$

avec $\rho = \sqrt{a^2+b^2} > 0$, $\theta = \arcsin\frac{b}{\rho}$ et $\Lambda' = \sqrt{\rho}\begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}$ et la décomposition de Jordan réelle de $J_k(\Lambda')^2$ est $J_k(\Lambda'^2)$,

- nous avons $\begin{pmatrix} 0 & -J_k(\lambda') \\ J_k(\lambda') & 0 \end{pmatrix}^2 = \begin{pmatrix} -J_k(\lambda')^2 & 0 \\ 0 & -J_k(\lambda')^2 \end{pmatrix}$ et la décomposition de Jordan réelle de $-J_k(\lambda')^2$ est $J_k(-\lambda'^2)$ si $\lambda' > 0$,

(4) Soit $A \in \mathrm{SL}_n(\mathbb{R})$. Si A appartient à l'image de $\exp_{|\mathfrak{sl}_n(\mathbb{R})}$, alors A appartient à l'image de exp. Réciproquement, s'il existe $B \in \mathscr{M}_n(\mathbb{R})$ tel que $\exp B = A$, alors par la proposition 3.4 (5), nous avons $e^{\mathtt{tr}\, B} = \det(\exp B) = \det A = 1$. Donc $\mathtt{tr}\, B = 0$ et $B \in \mathfrak{sl}_n(\mathbb{R})$, ce qui montre le résultat.

(5) Un petit calcul par diagonalisation (ou, de manière analytique, le fait de remarquer que l'application $t \mapsto A(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$, qui vérifie l'équation différentielle linéaire à coefficients constants $A'(t) = B\ A(t)$ où $B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ avec condition initiale $A(0) = I_2$, a pour solution $A(t) = \exp(t\, B)$) montre que

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \exp\begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}\ .$$

Donc $\exp : \mathfrak{so}(2) \to \mathrm{SO}(2)$ est surjective.

L'application exponentielle commute avec les conjugaisons. La matrice exponentielle d'une matrice diagonale par blocs de blocs diagonaux $A_1, \ldots, A_k$ est la matrice diagonale par blocs de blocs diagonaux $\exp(A_1), \ldots, \exp(A_k)$. Le conjugué d'une matrice antisymétrique par un élément de $\mathrm{SO}(n)$ est encore antisymétrique. Le résultat découle alors du théorème 2.4, qui dit que tout élément de $\mathrm{SO}(n)$ est conjuguée dans $\mathrm{SO}(n)$ à une matrice diagonale par blocs avec des blocs diagonaux 1×1 valant $1 = e^0$ et des blocs diagonaux 2×2 qui sont des rotations (y compris d'angle π).

(6) et (7) L'application exponentielle commute avec les conjugaisons. La matrice exponentielle d'une matrice diagonale de coefficients diagonaux $a_1, \ldots, a_n$ est la matrice diagonale de coefficients diagonaux $e^{a_1}, \ldots, e^{a_n}$. Une matrice diagonale à coefficients diagonaux imaginaires purs est anti-hermitienne. Le conjugué d'une matrice anti-hermitienne (respectivement anti-hermitienne de trace nulle) par un élément de $\mathrm{U}(n)$ (respectivement $\mathrm{SU}(n)$) est encore anti-hermitienne (respectivement anti-hermitienne de trace nulle).

Le résultat découle alors de la proposition 3.3 qui dit que tout élément de $\mathrm{U}(n)$ (respectivement $\mathrm{SU}(n)$) est conjugué dans $\mathrm{U}(n)$ (respectivement $\mathrm{SU}(n)$) à une matrice diagonale de coefficients diagonaux $e^{i\theta_1}, \ldots, e^{i\theta_n}$ où $\theta_1, \ldots, \theta_n \in \mathbb{R}$ (respectivement $\theta_1, \ldots, \theta_n \in \mathbb{R}$ et $\theta_1 + \cdots + \theta_n = 0$).

(8) Pour tout $X = \begin{pmatrix} x & y \\ z & -x \end{pmatrix} \in \mathfrak{sl}_2(\mathbb{C})$, notons ω (qui dépend de x, y, z) une des deux

racines carrées complexes de x^2+yz, de sorte que

$$\det X = -x^2 - yz = -\,\omega^2\ .$$

Nous prolongerons par continuité la fonction $t \mapsto \frac{\sinh t}{t}$ définie sur $\mathbb{C}-\{0\}$, en lui donnant la valeur limite 1 en $t=0$. Par le théorème de Cayley-Hamilton qui dit que

$$X^2 - (\mathtt{tr}\ X)X + (\det\ X)I_2 = 0$$

(ou tout simplement en calculant le carré de X), nous avons

$$X^2 = \omega^2 I_2\ .$$

Donc $X^{2n} = \omega^{2n} I_2$ et $X^{2n+1} = \omega^{2n} X$ pour tout $n \in \mathbb{N}$. En développant les séries, nous avons donc

$$\exp X = (\cosh\omega)\, I_2 + \frac{\sinh\omega}{\omega}\, X = \begin{pmatrix} \cosh\omega + \frac{\sinh\omega}{\omega}\, x & \frac{\sinh\omega}{\omega}\, y \\ \frac{\sinh\omega}{\omega}\, z & \cosh\omega - \frac{\sinh\omega}{\omega}\, x \end{pmatrix}\ .$$

Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{C})$ tel que $A + I_2$ ne soit pas nilpotente, et montrons que A appartient à l'image de $\exp : \mathfrak{sl}_2(\mathbb{C}) \to \mathrm{SL}_2(\mathbb{C})$.

Par la formule centrée précédente, nous avons $A = \exp X$ si et seulement si

$$\cosh\omega = \frac{a+d}{2},\quad \frac{\sinh\omega}{\omega}\, x = \frac{a-d}{2},\quad \frac{\sinh\omega}{\omega}\, y = b,\quad \frac{\sinh\omega}{\omega}\, z = c\ .$$

Rappelons que $\cosh : \mathbb{C} \to \mathbb{C}$ est surjective[94] et que $\sinh\omega = -\sin(i\omega)$ est nul si et seulement s'il existe $n \in \mathbb{Z}$ tel que $\omega = i\pi n$. En particulier, $\frac{\sinh\omega}{\omega}$ est nul si et seulement s'il existe $n \in \mathbb{Z}-\{0\}$ tel que $\omega = i\pi n$. Fixons donc $\omega \in \mathbb{C}$ tel que $\cosh\omega = \frac{a+d}{2}$.

Supposons tout d'abord que $\omega \notin i\pi\mathbb{Z}-\{0\}$. Posons alors

$$x = \frac{\omega}{\sinh\omega}\,\frac{a-d}{2},\quad y = \frac{\omega}{\sinh\omega}\, b,\quad z = \frac{\omega}{\sinh\omega}\, c\ .$$

Pour montrer que $A = \exp X$, il suffit donc de montrer que $\omega^2 = x^2 + yz$. Or, en utilisant le fait que $bc = ad - 1$, nous avons

$$\begin{aligned} x^2 + yz &= \frac{\omega^2(a-d)^2}{4\sinh^2\omega} + \frac{\omega^2 bc}{\sinh^2\omega} = \frac{\omega^2((a-d)^2 + 4(ad-1))}{4\sinh^2\omega} \\ &= \frac{\omega^2((a+d)^2-4)}{4\sinh^2\omega} = \frac{\omega^2(\cosh^2\omega - 1)}{\sinh^2\omega} = \omega^2\ . \end{aligned}$$

Supposons au contraire qu'il existe un élément $n \in \mathbb{Z}-\{0\}$ tel que $\omega = i\pi n$. En particulier $\cosh w = (-1)^n$, donc $\mathtt{tr}\, A = a + d = 2(-1)^n$. Si λ_1, λ_2 sont les valeurs propres (complexes) de A, alors $\lambda_1\lambda_2 = \det A = 1$ et $\lambda_1 + \lambda_2 = \mathtt{tr}\, A = 2(-1)^n$, ce qui implique que $\lambda_1 = \lambda_2 = (-1)^n$.

94. Pour tout $a \in \mathbb{C}$, l'équation $\frac{e^z + e^{-z}}{2} = a$ admet une solution $z \in \mathbb{C}$ si et seulement si l'équation quadratique $X^2 - 2aX + 1 = 0$ (obtenue en posant $X = e^z$, en utilisant la surjectivité de $\exp : \mathbb{C} \to \mathbb{C}^\times$) a une solution non nulle. Or $X = a + \sqrt{a^2-1}$, qui ne s'annule pour aucune valeur de $a \in \mathbb{C}$, est une telle solution.

Si n est pair, alors la matrice $N = A - I_2$ est nilpotente. En effet, ses deux valeurs propres sont égales à 0, donc elle appartient à $\mathfrak{sl}_2(\mathbb{C})$ et vérifie (par trigonalisation) $N^2 = 0$. D'où

$$\exp N = I_2 + N = A \ ,$$

et A appartient bien à l'image de $\exp : \mathfrak{sl}_2(\mathbb{C}) \to \mathrm{SL}_2(\mathbb{C})$.

Si n est impair, alors un raisonnement analogue montre que $N = A + I_2$ est nilpotente, une contradiction à l'hypothèse sur A.

Réciproquement, soit $A \in \mathrm{SL}_2(\mathbb{C})$ tel que $N = A + I_2$ soit nilpotente, montrons par l'absurde que A n'appartient pas à l'image de $\exp : \mathfrak{sl}_2(\mathbb{C}) \to \mathrm{SL}_2(\mathbb{C})$. En effet, supposons au contraire qu'il existe $X \in \mathfrak{sl}_2(\mathbb{C})$ tel que $\exp X = A$. Si $N \neq 0$ alors $A = I_2 + N$ n'est pas diagonalisable, donc X n'est pas diagonalisable (sinon $A = \exp X$ le serait), donc les valeurs propres de X sont égales, et elles sont nulles car $\mathtt{tr}\, X = 0$. Par trigonalisation, les valeurs propres de $A = \exp X$ sont donc égales à $e^0 = 1$, une contradiction car les valeurs propres de A sont toutes deux égales à -1.

(9) Soit $B \in \mathrm{SL}_n(\mathbb{C})$. Par la surjectivité de l'exponentielle des matrices complexes, soit $A \in \mathscr{M}_n(\mathbb{C})$ telle que $\exp A = B$. Puisque $e^{\mathtt{tr}\, A} = \det(\exp A) = \det B = 1$, il existe $k \in \mathbb{Z}$ tel que $\mathtt{tr}\, A = 2i\pi k$. Posons $A' = A - \frac{2i\pi k}{n} I_n$. Alors $\mathtt{tr}\, A' = 0$. Puisque A et tout multiple de l'identité commutent, nous avons $\exp(A') = e^{\frac{2i\pi k}{n}} \exp A = e^{\frac{2i\pi k}{n}} B$, qui diffère de B par un élément du centre $Z(\mathrm{SL}_n(\mathbb{C})) = \{e^{\frac{2i\pi k}{n}} I_n \ : \ k \in \mathbb{Z}\}$ de $\mathrm{SL}_n(\mathbb{C})$.

Correction de l'exercice E.33. Pour toute matrice normale B, la matrice $A = \exp B$ est inversible, d'adjointe $A^* = \exp(B^*)$, donc qui commute avec A.

Réciproquement, si $A \in \mathscr{M}_n(\mathbb{C})$ est normale inversible, alors par le théorème de diagonalisation 1.14, il existe une matrice diagonale $D \in \mathscr{M}_n(\mathbb{C})$ et une matrice unitaire $U \in \mathrm{U}(n)$ telles que $A = U\, D\, U^{-1}$. Puisque A est inversible, les coefficients diagonaux de D sont non nuls. Par la surjectivité de l'exponentielle complexe $\mathbb{C} \to \mathbb{C}^*$, il existe une matrice diagonale D' telle que $D = \exp D'$. Alors $B = U\, D'\, U^{-1}$ est normale (car diagonalisable en base orthonormée) et

$$A = U\, D\, U^{-1} = U\, (\exp D')\, U^{-1} = \exp(U\, D'\, U^{-1}) = \exp B \ .$$

Correction de l'exercice E.34. Cette correction est issue du livre [Ale].

(1) Puisque K est compact et par continuité de l'application $g \ \mapsto gx$, la borne supérieure $\sup_{g\in K} \|gx\|$ est atteinte. Par linéarité de l'action de g et la propriété de la norme, nous avons $\|\lambda x\|_K = |\lambda|\ \|x\|_K$ et

$$\|x+y\|_K = \max_{g\in K} \|gx+gy\| \leq \max_{g\in K}(\|gx\| + \|gy\|) \leq \max_{g\in K} \|gx\| + \max_{g\in K} \|gy\| = \|x\|_K + \|y\|_K \ .$$

Le cas d'égalité et la compacité de K forcent l'existence d'un élément $g \in G$ tel que $\|gx + gy\| = \|gx\| + \|gy\|$, ce qui, par la linéarité de g et par le cas d'égalité de la sous-additivité de la norme dans les espaces euclidiens ou hermitiens, montre que x et y sont positivement dépendants. Par changement de variable, nous avons $\|gx\|_K = \|x\|_K$ pour tous les $g \in K$ et $x \in E$.

Puisque C est compact, l'application continue de C dans $\mathbb{R}$ définie par $x \mapsto \|x\|_K$ atteint sa borne inférieure, disons en $x_0 \in C$. Alors pour tout $g \in K$, le point gx_0 appartient à C

par invariance, donc le point $\frac{1}{2}(x_0 + gx_0)$ appartient à C par convexité. Donc

$$\|x_0\|_K \leq \|\frac{1}{2}(x_0 + gx_0)\|_K \leq \|x_0\|_K \ .$$

L'analyse du cas d'égalité dans l'inégalité triangulaire de la norme $\| \ \|_K$ montre que x_0 et gx_0 sont positivement dépendants, ce qui, puisqu'ils ont même norme pour $\| \ \|_K$, implique qu'ils sont égaux.

(2) Montrons le résultat dans $\mathrm{GL}_n(\mathbb{C})$, la démonstration est la même dans $\mathrm{GL}_n(\mathbb{R})$. Soit K_0 un sous-groupe compact maximal de $\mathrm{GL}_n(\mathbb{C})$. Montrons que K_0 est conjugué à un sous-groupe du sous-groupe compact $\mathrm{U}(n)$, ce qui implique que K_0 est conjugué à $\mathrm{U}(n)$ par maximalité.

Rappelons que Herm_n est l'espace vectoriel réel des matrices hermitiennes $n \times n$, que nous munissons d'un produit scalaire euclidien quelconque. L'application de $\mathrm{GL}_n(\mathbb{C})$ dans $\mathrm{GL}(\mathrm{Herm}_n)$ définie par $x \mapsto \{y \mapsto xyx^*\}$ est un morphisme de groupes bien défini et continu (car polynomial en les coefficients). Notons K l'image de K_0 par cette application, qui est un sous-groupe compact de $\mathrm{GL}(\mathrm{Herm}_n)$. Remarquons que $\{z^*z \ : \ z \in K_0\}$ est un compact non vide, contenu dans le convexe Herm_n^{++}, et qui est invariant par K, car $x(yy^*)x^* = (xy)(xy)^*$ pour tous les x, y dans le groupe K_0. Notons C l'enveloppe convexe dans l'espace vectoriel réel Herm_n de $\{zz^* : \ z \in K_0\}$, qui est aussi compacte non vide, contenue dans Herm_n^{++} et invariante par K. Par l'assertion (1), il existe alors un point z de C fixe par tout élément de K. Donc la matrice z est hermitienne définie positive, et $xzx^* = z$ pour tout $x \in K_0$. En particulier, K_0 préserve le produit scalaire hermitien de matrice z.

Par conséquent, $(\sqrt{z}^{-1}x\sqrt{z})(\sqrt{z}^{-1}x\sqrt{z})^* = \sqrt{z}^{-1}(xzx^*)\sqrt{z}^{-1} = \mathrm{id}$, ce qui montre que le conjugué de K_0 par $\sqrt{z}^{-1}$ est contenu dans $\mathrm{U}(n)$.

Correction de l'exercice E.36. C'est un cas particulier de l'exercice E.34, mais il a une démonstration bien plus élémentaire. Traitons le cas complexe, le cas réel est analogue. Si G est un sous-groupe fini de $\mathrm{GL}_n(\mathbb{C})$, alors

$$(x, y) \mapsto \langle x, y\rangle_G = \sum_{g\in G} \langle gx, gy\rangle$$

est un produit scalaire hermitien sur $\mathbb{C}^n$, qui est invariant par G par changement de variable dans la somme. Si q est la forme quadratique hermitienne associée, alors $G \subset \mathrm{U}(q)$. Par leur classification (voir le théorème 1.12), il existe $h \in \mathrm{GL}_n(\mathbb{C})$ tel que $q \circ h$ soit la forme quadratique hermitienne usuelle de $\mathbb{C}^n$. Donc

$$h^{-1}Gh \subset h^{-1}\,\mathrm{U}(q)h = \mathrm{U}(q \circ h) = \mathrm{U}(n) \ ,$$

ce qu'il fallait démontrer.

Correction de l'exercice E.37. Notons Base_E l'ensemble des bases de E, qui est une partie de E^n. L'action diagonale de $\mathrm{GL}(E)$ sur E^n préserve Base_E. Fixons une base $\mathscr{B}$ de E et identifions tout élément de $\mathrm{GL}(E)$ avec sa matrice dans la base $\mathscr{B}$. Par la continuité du déterminant, la partie Base_E est un ouvert de E^n. De plus, l'application de $\mathrm{GL}(E)$ dans Base_E définie par $g \mapsto g\mathscr{B}$ est continue, et bijective car $\mathrm{GL}(E)$ agit simplement transitivement sur l'ensemble des bases de E. Son inverse, qui est l'application qui à une base $\mathscr{B}'$ associe la matrice de passage $P_{\mathscr{B},\mathscr{B}'}$, est aussi continue. Donc Base_E est homéomorphe

à $\mathrm{GL}(E)$, qui par la décomposition polaire (voir le théorème 3.9), est homéomorphe à $\mathbb{R}^{n(n+1)/2} \times \mathrm{O}(n)$. Les autres affirmations s'en déduisent, par restriction.

Correction de l'exercice E.39. La matrice B^*B est hermitienne définie positive, et pour tout $j = 1, \ldots, n$, son coefficient diagonal (j, j) est égal à ${}^t\overline{C_j}C_j = \|C_j\|^2$. Par l'exercice E.9, le résultat découle du fait que $\det(B^*B) = |\det(B)|^2$.

Par l'exercice E.9, le cas d'égalité se produit exactement quand il existe une matrice diagonale D' telle que $B^*B = D'$ est diagonale. En posant $D = \sqrt{D'}$ (qui existe car D' est à coefficients diagonaux strictement positifs), l'égalité $B^*B = D'$ est équivalente à $(BD^{-1})^*(BD^{-1}) = I_n$, c'est-à-dire au fait que BD^{-1} soit unitaire. L'égalité étudiée $|\det(B)| = \prod_{j=1}^n \|C_j\|^2$ se produit donc exactement quand $B = UD$ avec D diagonale à coefficients diagonaux strictement positifs et U unitaire.

Correction de l'exercice E.40. Nous renvoyons par exemple à [MT, page 186]. Sauf pour la question c), nous ne traitons que le cas complexe, le cas réel étant complètement analogue.

a) Si $|\lambda| = 1$, alors $Q_{\lambda X} = I_n - 2|\lambda|^2 XX^* = Q_X$. Puisque X est unitaire, nous avons $Q_X X = X - 2XX^*X = -X$ et si $Y \in X^\perp$, alors $Q_X Y = Y$. Donc Q_X est la symétrie orthogonale par rapport à l'hyperplan orthogonal à la droite vectorielle D engendrée par X (et, en particulier, Q_X est unitaire). Un vecteur directeur unitaire de D est uniquement déterminé modulo multiplication par un scalaire de valeur absolue 1.

Nous avons $(I_n - 2XX^*)^* = I_n - 2XX^*$, donc Q_X est hermitienne. Pour tout $P \in \mathrm{U}(n)$, nous avons, puisque $P^{-1} = P^*$,

$$PQ_XP^{-1} = I_n - 2PXX^*P^* = I_n - 2(PX)(PX)^* = Q_{PX} \ .$$

Puisque X est unitaire, comme $\mathrm{U}(n)$ agit transitivement sur la sphère unité de $\mathbb{C}^n$, il existe $P \in \mathrm{U}(n)$ tel que $X = Pe_1$, où e_1 est le premier vecteur de la base canonique de $\mathbb{C}^n$. Donc, puisque $P^* = P^{-1}$,

$$\begin{aligned}\det(I_n - 2XX^*) &= \det(I_n - 2(Pe_1)(Pe_1)^*) = \det\left(P(I_n - 2e_1e_1^*)P^{-1}\right) \\ &= \det(I_n - 2e_1e_1^*) = -1 \ .\end{aligned}$$

b) Soit $\alpha \in \mathbb{U}$ tel que $\overline{\alpha}\,\langle X, e_1\rangle \in \mathbb{R}$. Si $\langle X, e_1\rangle \in \mathbb{R}$, nous prenons $\alpha = 1$. Puisque $\|X\| = \|\alpha\, e_1\| = 1$, la première colonne de la matrice $Q_{X'}A$ est

$$\begin{aligned}Q_{X'}X &= X - \frac{2}{\|X - \alpha\, e_1\|^2}(X - \alpha\, e_1)(X - \alpha\, e_1)^*X \\ &= X - \frac{2}{2 - 2\mathrm{Re}(\langle X, \alpha\, e_1\rangle)}(X - \alpha\, e_1)(1 - \langle X, \alpha\, e_1\rangle) \\ &= X - (X - \alpha\, e_1) = \alpha\, e_1 \ .\end{aligned}$$

Donc $Q_{X'}A$ est triangulaire supérieure en blocs $(1, n-1)$, de coefficient $(1,1)$ égal à α.

c) et d) Montrons par récurrence sur n que tout élément de $\mathrm{O}(n)$ (respectivement $\mathrm{U}(n)$) est un produit d'au plus n matrices de Householder (respectivement d'au plus n matrices de Householder puis d'une matrice diagonale).

Le résultat est immédiat si $n = 1$. Soient $n \geq 2$ et $A \in \mathrm{O}(n)$ (respectivement $A \in \mathrm{U}(n)$). Soit X' comme dans la démonstration de la question b). Puisque $Q_{X'}A$ est orthogonale

(respectivement unitaire) et préserve la première droite de coordonnée, elle préserve aussi son orthogonal, donc $Q_{X'}A = \begin{pmatrix} \alpha & 0 \\ 0 & A' \end{pmatrix}$ avec $\alpha = 1$ et $A' \in \mathrm{O}(n-1)$ (respectivement $\alpha \in \mathbb{U}$ et $A' \in \mathrm{U}(n-1)$). Remarquons que si Y est un vecteur unitaire de $\mathbb{R}^{n-1}$ (respectivement $\mathbb{C}^{n-1}$), alors $\begin{pmatrix} 0 \\ Y \end{pmatrix}$ est un vecteur unitaire de $\mathbb{R}^n$ (respectivement $\mathbb{C}^n$) et

$$I_n - 2\begin{pmatrix} 0 \\ Y \end{pmatrix}\begin{pmatrix} 0 \\ Y \end{pmatrix}^* = \begin{pmatrix} 1 & 0 \\ 0 & I_{n-1} - 2\,YY^* \end{pmatrix} .$$

Donc si Q est une matrice de Householder de $\mathscr{M}_{n-1}(\mathbb{R})$ (respectivement $\mathscr{M}_{n-1}(\mathbb{C})$), alors $\begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}$ est une matrice de Householder de $\mathscr{M}_n(\mathbb{R})$ (respectivement $\mathscr{M}_n(\mathbb{C})$).

Par récurrence, il existe donc $k \in \{1, \dots, n-1\}$ et des matrices de Householder $Q_{X_1}, Q_{X_2}, \dots, Q_{X_k}$ telles que $Q_{X_1}Q_{X_2}\dots Q_{X_k}Q_{X'}A$ soit la matrice identité (respectivement une matrice diagonale). Remarquons que $Q_X^{-1} = Q_X$ pour tout vecteur unitaire X. Donc A est produit d'au plus $n-1+1 = n$ matrices de Householder (respectivement d'au plus n matrices de Householder et d'une matrice diagonale).

Comme le déterminant d'une matrice de Householder est -1 par la question a), un élément de $\mathrm{SO}(n)$ est produit d'un nombre pair (au plus n) de matrices de Householder.

Remarque. Ceci redémontre que les réflexions engendrent $\mathrm{O}(n)$, et que tout élément de $\mathrm{SO}(n)$ est le produit d'un nombre pair inférieur à n de réflexions.

Correction de l'exercice E.42. (1) Pour tous les $A \in \mathrm{SU}(2)$ et $X \in \mathfrak{su}(2)$, nous avons $(AXA^{-1})^* = (A^*)^{-1}X^*A^* = -AXA^{-1}$ et $\mathtt{tr}(AXA^{-1}) = \mathtt{tr}\,X = 0$. Par conséquent $AXA^{-1} \in \mathfrak{su}(2)$. L'application $\mathrm{Ad}\,A : X \mapsto AXA^{-1}$ est linéaire et inversible d'inverse $\mathrm{Ad}(A^{-1})$. Comme $\mathrm{Ad}(AB) = (\mathrm{Ad}\,A) \circ (\mathrm{Ad}\,B)$, l'application $A \mapsto \mathrm{Ad}\,A$ est un morphisme de groupes de $\mathrm{SU}(2)$ dans $\mathrm{GL}(\mathfrak{su}(2))$.

(2) Toute matrice X de $\mathfrak{su}(2)$, étant anti-hermitienne de trace nulle, s'écrit de manière unique

$$X = \begin{pmatrix} i\,x_3 & -x_2 + i\,x_1 \\ x_2 + i\,x_1 & -i\,x_3 \end{pmatrix} = x_1\,\xi_1 + x_2\,\xi_2 + x_3\,\xi_3$$

où x_1, x_2, x_3 sont des nombres réels. Donc (ξ_1, ξ_2, ξ_3) est une base de l'espace vectoriel réel $\mathfrak{su}(2)$. De plus, l'application $X \mapsto \sqrt{\det X} = (x_1^2 + x_2^2 + x_3^2)^{\frac{1}{2}}$ est une norme euclidienne sur $\mathfrak{su}(2)$, rendant orthonormée la base (ξ_1, ξ_2, ξ_3). Puisque $\det(AXA^{-1}) = \det X$, cette norme est préservée par l'action de $\mathrm{SU}(2)$.

(3) Considérons l'application Φ de $\mathrm{SU}(2)$ dans $\mathrm{GL}_3(\mathbb{R})$ qui à A associe la matrice de $\mathrm{Ad}\,A$ dans la base (ξ_1, ξ_2, ξ_3). Elle est polynomiale en les coefficients, donc continue. Par ce qui précède, l'image de Φ est contenue dans le groupe orthogonal $\mathrm{O}(3)$. Comme $\mathrm{SU}(2)$ est connexe et Φ continue, son image est en fait contenue dans la composante connexe de l'élément neutre dans $\mathrm{O}(3)$, qui est $\mathrm{SO}(3)$.

Le noyau $\mathrm{Ker}\,\mathrm{Ad}$ de Ad est

$$\{g \in \mathrm{SU}(2) \ : \ \forall\, X \in \mathfrak{su}(2),\ \ gXg^{-1} = X\} .$$

En prenant $X = \xi_3$, qui est diagonale à valeurs propres distinctes, nous obtenons que tout élément de $\operatorname{Ker}\operatorname{Ad}$ est diagonal, donc de la forme $\begin{pmatrix} a & 0 \\ 0 & \overline{a} \end{pmatrix}$ avec $|a| = 1$. En prenant $X = \xi_2$, nous obtenons que $a = \overline{a}$ est réel, donc $a = \pm 1$, d'où $\operatorname{Ker}\operatorname{Ad}$ est contenu dans $\{\pm \operatorname{id}\}$. L'inclusion réciproque étant immédiate, nous avons $\operatorname{Ker}\operatorname{Ad} = \{\pm \operatorname{id}\}$. Donc $\operatorname{Ker}\Phi = \{\pm \operatorname{id}\}$.

(4) Voir la solution de l'exercice E.32 (5).

(5) Pour tous les $X, Y \in \mathfrak{su}(2)$, nous avons

$$(XY - YX)^* = Y^*X^* - X^*Y^* = YX - XY = -(XY - YX)$$

et $\mathtt{tr}(XY - YX) = 0$. Par conséquent $XY - YX \in \mathfrak{su}(2)$. Pour tout $X \in \mathfrak{su}(2)$, l'application

$$\operatorname{ad} X : Y \mapsto XY - YX$$

est clairement linéaire, et l'application $\operatorname{ad} : \mathfrak{su}(2) \to \mathscr{L}(\mathfrak{su}(2))$ définie par $X \mapsto \operatorname{ad} X$ est aussi clairement linéaire.

Un petit calcul montre que $\operatorname{ad}\xi_1(\xi_2) = -\operatorname{ad}\xi_2(\xi_1) = 2\,\xi_3$, $\operatorname{ad}\xi_1(\xi_3) = -\operatorname{ad}\xi_3(\xi_1) = -2\,\xi_2$, $\operatorname{ad}\xi_2(\xi_3) = -\operatorname{ad}\xi_3(\xi_2) = 2\,\xi_1$, et $\operatorname{ad}\xi_i(\xi_i) = 0$ pour $i = 1, 2, 3$. Donc la matrice de $\operatorname{ad}(x_1\,\xi_1 + x_2\,\xi_2 + x_3\,\xi_3)$ dans la base (ξ_1, ξ_2, ξ_3) est

$$\begin{pmatrix} 0 & -2\,x_3 & 2\,x_2 \\ 2\,x_3 & 0 & -2\,x_1 \\ -2\,x_2 & 2x_1 & 0 \end{pmatrix} \in \mathfrak{so}(3)\ .$$

Donc l'application $T\Phi$ est bien à valeurs dans $\mathfrak{so}(3)$. Elle est injective par la formule centrée ci-dessus, et puisque les espaces vectoriesl réels $\mathfrak{su}(2)$ et $\mathfrak{so}(3)$ sont tous les deux de dimension 3, l'application linéaire $T\Phi$ est un isomorphisme.

Montrons que le diagramme suivant est commutatif

$$\begin{array}{ccc} \mathfrak{su}(2) & \xrightarrow{\operatorname{ad}} & \mathscr{L}(\mathfrak{su}(2)) \\ \exp\downarrow & & \downarrow\exp \\ \mathrm{SU}(2) & \xrightarrow{\mathrm{Ad}} & \mathrm{GL}(\mathfrak{su}(2)) \end{array}\ .$$

Pour tout $X \in \mathfrak{su}(2)$, le sous-groupe à un paramètre

$$t \mapsto \operatorname{Ad}(\exp(tX)) : Y \mapsto \exp(tX)Y\exp(-tX)$$

est différentiable, de dérivée en $t = 0$ égale à $\operatorname{ad} X : Y \mapsto XY - YX$. Par le rappel donné dansl'énoncé de la question (5), nous avons donc $\operatorname{Ad}(\exp(tX)) = \exp(t\operatorname{ad}(X))$ pour tout $t \in \mathbb{R}$. Le cas particulier $t = 1$ montre la commutativité cherchée.

Comme la matrice de l'exponentielle d'un endomorphisme est l'exponentielle de la matrice de cet endomorphisme, ceci montre la commutativité du diagramme de la question (5).

La surjectivité de Φ découle alors de la surjectivité de $\exp : \mathfrak{so}(3) \to \mathrm{SO}(3)$ (voir la question (4)) et de la commutativité du diagramme.

4 Sur quelques groupes d'isométries euclidiennes et affines euclidiennes

4.1 Sous-groupes finis de $\mathrm{SO}(3)$ et polyèdres réguliers de dimension 3

Le théorème suivant est un théorème de classification à conjugaison près (et pas seulement à isomorphisme de groupes près) des sous-groupes finis de rotations de l'espace euclidien de dimension 3. Par l'exercice E.47, il permet d'obtenir aussi la classification à conjugaison près des sous-groupes finis de $\mathrm{O}(3)$, $\mathrm{GL}_3(\mathbb{R})$, $\mathrm{O}(1,3)$, $\mathrm{SO}(1,3)$, $\mathrm{SO}_0(1,3)$, $\mathrm{GL}_2(\mathbb{C})$, $\mathrm{SL}_2(\mathbb{C})$, $\mathrm{PGL}_2(\mathbb{C})$ et $\mathrm{PSL}_2(\mathbb{C})$. Certains objets qui apparaissent dans l'énoncé du théorème 4.1 seront définis au cours de sa démonstration.

Théorème 4.1. *Tout sous-groupe fini de* $\mathrm{SO}(3)$ *est conjugué à un et un seul des groupes suivants :*

(1) *un groupe cyclique d'ordre* n *où* $n \in \mathbb{N} - \{0\}$*, engendré par une rotation d'angle* $\frac{2\pi}{n}$*,*

(2) *un groupe diédral d'ordre* $2n$ *où* $n \in \mathbb{N} - \{0,1\}$*, engendré par les renversements sur deux droites vectorielles faisant un angle* $\frac{\pi}{n}$*,*

(3) *le groupe des isométries directes d'un tétraèdre régulier de barycentre* 0*, isomorphe au groupe alterné* $\mathfrak{A}_4$*,*

(4) *le groupe des isométries directes d'un cube régulier de barycentre* 0*, ou, de manière équivalente, d'un octaèdre régulier de barycentre* 0*, isomorphe au groupe symétrique* $\mathfrak{S}_4$*,*

(5) *le groupe des isométries directes d'un dodécaèdre régulier de barycentre* 0*, ou, de manière équivalente, d'un icosaèdre régulier de barycentre* 0*, isomorphe au groupe alterné* $\mathfrak{A}_5$*.*

Démonstration. Nous nous inspirons par exemple des références [Aud, Exercice V.54], [Arn] et [AB, t. 1, Chap. IX]. Le résultat suivant permet de traiter, puis d'exclure par la suite, le premier cas.

Lemme 4.2. *Pour tout* $n \in \mathbb{N} - \{0\}$*, il existe un sous-groupe cyclique de* $\mathrm{SO}(3)$ *d'ordre* n*. Un tel sous-groupe est engendré par une rotation d'angle* $\frac{2\pi}{n}$*. Deux tels sous-groupes sont conjugués.*

Démonstration. Nous pouvons supposer que $n \geq 2$. Le résultat découle de la classification des éléments de $\mathrm{SO}(3)$ à conjugaison près (voir le théorème 2.4). Tout élément non trivial de $\mathrm{SO}(3)$ est une rotation autour d'un unique axe de rotation orienté d'angle de rotation dans le sens positif égal à $\theta \in]0, 2\pi[$. Cette rotation est d'ordre n si et seulement si $\theta = \frac{2\,k\pi}{n}$ où $k \in \{1, \ldots, n-1\}$ est tel que k et n soient premiers entre eux. Cette rotation engendre le même groupe cyclique que la rotation autour du même axe orienté d'angle $\frac{2\pi}{n}$. Par la transitivité de l'action de $\mathrm{SO}(3)$ sur $\mathbb{S}_2$, deux rotations $\rho_1, \rho_2 \in \mathrm{SO}(3)$ ayant le même angle de rotation dans le sens positif autour d'un même axe orienté sont conjuguées dans $\mathrm{SO}(3)$ (par tout élément de $\mathrm{SO}(3)$ envoyant l'axe orienté de ρ_1 sur celui de ρ_2). □

Notons $|E|$ le cardinal d'un ensemble fini E. Á partir de maintenant, nous considérons donc un sous-groupe fini non cyclique G de $\mathrm{SO}(3)$. Notons $n = |G| \in \mathbb{N}$ son ordre, qui vérifie $n \geq 4$, car les groupes d'ordre au plus 3 sont cycliques.

Notons X l'ensemble (non vide) des points fixes des éléments non triviaux de G dans la sphère unité $\mathbb{S}_2$ de l'espace euclidien standard $\mathbb{R}^3$. Il est fini car tout élément non trivial

de SO(3) admet deux points fixes seulement dans $\mathbb{S}_2$, les points d'intersection de son axe de rotation avec $\mathbb{S}_2$. L'ensemble X est invariant par G, car en notant $F_{g'}$ l'ensemble des points fixes dans $\mathbb{S}_2$ d'un élément $g' \in G$, pour tous les $g, h \in G$, nous avons $gF_h = F_{ghg^{-1}}$. Nous allons étudier les propriétés de l'action du groupe G sur l'ensemble X.

Notons $O_1, \dots, O_k$ les orbites de G dans X et $n_i \geq 2$ l'ordre commun des stabilisateurs des points de O_i dans G, de sorte que $X = \coprod_{i=1}^k O_i$ (union disjointe) et

$$|O_i| = \frac{n}{n_i}$$

par la formule des classes. Notons

$$\Gamma = \big\{(g,x) \in (G - \{\mathrm{id}\}) \times X \ : \ g(x) = x\big\} \ .$$

Tout élément de SO(3) différent de l'identité est une rotation d'angle non nul modulo 2π, donc admet exactement deux points fixes dans $\mathbb{S}_2$, les points d'intersection de son axe avec cette sphère. Donc

$$|\Gamma| = \sum_{g \in G - \{\mathrm{id}\}} \big|\{x \in X \ : \ g(x) = x\}\big| = 2(n-1) \ .$$

Tout élément de O_i est fixé par exactement $n_i - 1$ éléments non triviaux. Donc

$$|\Gamma| = \sum_{x \in X} \big|\{g \in G - \{\mathrm{id}\} \ : \ g(x) = x\}\big| = \sum_{i=1}^k \sum_{x \in O_i} (n_i - 1) = \sum_{i=1}^k n(1 - \frac{1}{n_i}) \ .$$

Les deux formules centrées précédentes impliquent que

$$2(1 - \frac{1}{n}) = \sum_{i=1}^k (1 - \frac{1}{n_i}) \ . \tag{29}$$

Nous pouvons supposer que $n_1 \leq \cdots \leq n_k$. Comme $n_1 \leq n$ (donc $1 - \frac{1}{n_1} \leq 1 - \frac{1}{n}$), nous ne pouvons pas avoir $k = 1$. Si nous avions $n_k = n$, alors G admettrait un point fixe global dans $\mathbb{S}_2$, donc fixerait une droite vectorielle orientée, donc serait isomorphe à un sous-groupe de SO(2), donc serait cyclique, ce que nous avons exclu. Par conséquent n_i, qui divise n, est inférieur ou égal à $\frac{n}{2}$. Si nous avions $k = 2$, la formule (29) donnerait l'égalité $\frac{1}{n_1} + \frac{1}{n_2} = \frac{2}{n}$. Cette égalité est impossible puisque son membre de gauche est au moins $\frac{4}{n}$. Donc $k \geq 3$. Puisque $n_i \geq 2$ et $\sum_{i=1}^k (1 - \frac{1}{n_i}) < 2$, nous avons $k \leq 3$. Par conséquent

$$k = 3$$

et la formule (29) s'écrit

$$\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} = 1 + \frac{2}{n} > 1 \ .$$

Si nous avions $n_1 \geq 3$, alors $n_2, n_3 \geq 3$ et le membre de gauche de l'inégalité ci-dessus serait inférieur ou égal à 1, ce qui contredirait cette inégalité. Donc

$$n_1 = 2 \ ,$$

ce qui implique que n est pair (car n_1 divise n), et la formule (29) s'écrit

$$\frac{1}{n_2}+\frac{1}{n_3}=\frac{1}{2}+\frac{2}{n}>\frac{1}{2}\,.$$

Donc de même $n_2\in\{2,3\}$. Si $n_2=2$, alors $n_3=\frac{n}{2}$, et nous décrirons la situation dans le **Cas 1** ci-dessous. Si $n_2=3$, alors $\frac{1}{n_3}=\frac{1}{6}+\frac{2}{n}>\frac{1}{6}$, donc n_3 peut prendre trois valeurs, $n_3=3$ (auquel cas $n=12$) ou $n_3=4$ (auquel cas $n=24$) ou $n_3=5$ (auquel cas $n=60$), et nous décrirons respectivement la situation dans les **Cas 2**, **Cas 3**, **Cas 4** ci-dessous.[95]

Nous décrivons maintenant les quatre situations ci-dessus, en faisant des disgressions afin de donner des informations complémentaires sur les groupes et les polyèdres (dont la construction d'iceux) qui apparaissent dans le théorème 4.1 ci-dessus.

Cas 1 : Supposons que $(n_1,n_2,n_3)=(2,2,\frac{n}{2})$. Montrons alors que G est diédral d'ordre $n=2n_3$, et que deux sous-groupes diédraux de $\mathrm{SO}(3)$ de même ordre sont conjugués.

Les hypothèses du cas 1 montrent que G admet un sous-groupe cyclique C d'ordre $\frac{n}{2}\geq 2$, le fixateur d'un point ξ de l'orbite O_3. Le cardinal de O_3 est donc égal à $\frac{n}{\frac{n}{2}}=2$. Puisque G préserve O_3 et puisqu'il contient un élément non trivial a fixant ξ (dont le seul autre point fixe est $-\xi$), nous avons par conséquent $O_3=\{\xi,-\xi\}$.

Comme les seules rotations envoyant un vecteur en son opposé sont les renversements, il existe donc un renversement $b\in G$ envoyant ξ sur $-\xi$, donc conjuguant a à son inverse a^{-1}. Par cardinalité, nous avons $G=C\cup bC$, donc G est engendré par une rotation a d'ordre $\frac{n}{2}$ et un renversement b valant $-\mathrm{id}$ sur l'axe de rotation de a. La description ci-dessous, bien plus longue que nécessaire, montre alors que G est un groupe diédral d'ordre n.

Rappelons tout d'abord que pour tous les groupes N et H, et pour tout morphisme de groupes $\phi:H\to\mathrm{Aut}(N)$, le *produit semidirect* de N et H par ϕ est le groupe G noté $N\rtimes_\phi H$ (ou simplement $N\rtimes H$ lorsque ϕ est sous-entendu, mais il vaut mieux préciser quand il y a plusieurs choix), d'ensemble sous-jacent $N\times H$ et de loi

$$(n,h)(n',h')=(n\,\phi(h)(n'),h\,h')\,.$$

Les groupes N et H sont identifiés à leur image dans G par les morphismes de groupes injectifs $i:n\mapsto(n,1)$ et $s:h\mapsto(1,h)$. Le sous-groupe N est distingué dans G. Le morphisme de groupes $p:(n,h)\mapsto h$ de G dans H est surjectif car $p\circ s=\mathrm{id}_H$. Son noyau est égal à $i(N)$ et identifié à N. Il induit donc par passage au quotient un isomorphisme de groupes de G/N dans H. Nous avons donc une suite exacte de groupes

$$1\longrightarrow N\xrightarrow{\ i\ }G\xrightarrow{\ p\ }H\longrightarrow 1\,,$$

qui est *scindée*, c'est-à-dire qu'il existe un morphisme de groupes injectif $s:H\to G$, appelé une *section* de p, tel que $p\circ s=\mathrm{id}_H$, dans le cas présent l'application $h\mapsto(1,h)$. Réciproquement, tout groupe G qui admet une telle suite exacte scindée est isomorphe au

95. Cette discussion, avec le cas 1 traité ci-dessous, montre le résultat partiel suivant, dont la démonstration, contrairement à celle complète du théorème 4.1, bien trop longue, rentre parfaitement dans le cadre d'un développement pour la leçon 101 sur les actions de groupes.

Théorème 4.3. *Tout sous-groupe fini de* $\mathrm{SO}(3)$ *non cyclique et non diédral est d'ordre 12, 24 ou 60.*

produit semi-direct de N par H pour le morphisme $\phi : H \to \mathrm{Aut}(N)$ défini (en identifiant N avec son image dans G par i) par $h \mapsto \{n \mapsto s(h)\, n\, s(h)^{-1}\}$.

Pour tout $m \in \mathbb{N} - \{0\}$, nous notons [96] D_{2m} le *groupe diédral* d'ordre $2m$, défini comme le produit semi-direct de $\mathbb{Z}/m\mathbb{Z}$ par le groupe $\mathbb{Z}/2\mathbb{Z}$ dont l'élément non trivial agit par l'automorphisme $x \mapsto -x$ de $\mathbb{Z}/m\mathbb{Z}$. Nous appellerons un *groupe diédral* tout groupe isomorphe à D_{2m} pour $m \in \mathbb{N} - \{0\}$, et nous allons en donner une liste partielle.

- Le groupe diédral D_{2m} admet les deux présentations suivantes [97] (la seconde étant une présentation de Coxeter, voir par exemple [Hum, Bou, dlH])

$$D_{2m} \simeq \langle a, b \,|a^m = b^2 = bab^{-1}a = 1\rangle \simeq \langle s, t \,|s^2 = t^2 = (st)^m = 1\rangle \ .$$

Il est possible de montrer que l'application envoyant a sur le générateur 1 de $\mathbb{Z}/m\mathbb{Z}$ et b sur le générateur 1 de $\mathbb{Z}/2\mathbb{Z}$ s'étend en un isomorphisme de groupes du groupe quotient $L(\{a,b\})/\langle\langle\{a^m, b^2, bab^{-1}a\}\rangle\rangle$ sur D_{2m}. Il est de même possible de montrer que l'application envoyant t sur le générateur 1 de $\mathbb{Z}/2\mathbb{Z}$, et s sur le conjugué de ce générateur par le générateur 1 de $\mathbb{Z}/m\mathbb{Z}$, s'étend en un isomorphisme de groupes du groupe quotient $L(\{s,t\})/\langle\langle\{s^2, t^2, (st)^m\}\rangle\rangle$ sur D_{2m}.

- Pour toutes les droites vectorielles D et D' d'un plan euclidien faisant un angle $\pm\frac{\pi}{m}$ entre elles (voir le dessin ci-dessous), le groupe diédral D_{2m} est isomorphe au sous-groupe $R_{D,D'}$ de $\mathrm{O}(2)$ engendré par les deux réflexions orthogonales σ_D et $\sigma_{D'}$ d'ensemble de points fixes D et D' respectivement, par l'unique morphisme de groupes envoyant le générateur 1 de $\mathbb{Z}/m\mathbb{Z}$ sur $s_{D'} \circ s_D$ (qui est une rotation d'angle $\pm\frac{2\pi}{m}$) et le générateur 1 de $\mathbb{Z}/2\mathbb{Z}$ sur s_D. Les présentations ci-dessus sont alors obtenues en posant $(t,s) = (\sigma_D, \sigma_{D'})$ et $(a,b) = (s_{D'} \circ s_D, \sigma_D)$.

- Si $m \geq 3$, le groupe diédral D_{2m} est isomorphe au groupe $G(P)$ des isométries d'un polygone régulier [98] P à m côtés de barycentre 0 dans un plan euclidien. En notant $b = s$ la symétrie orthogonale par rapport à une droite passant par un sommet A de P, et t la symétrie orthogonale par rapport à la médiatrice d'un côté de P adjacent à A, en posant $a = st$ la composition de s et t, alors ce groupe $G(P)$ est constitué de m rotations a^k et de m réflexions $a^{k'}b$ où $k, k' \in \{0, \ldots, m-1\}$.

En fait, tout sous-groupe fini de $\mathrm{O}(2)$ est conjugué ou bien au groupe engendré par la rotation d'angle $\frac{2\pi}{n}$ pour un entier $n \in \mathbb{N} - \{0\}$, ou bien à un groupe diédral $R_{D,D'}$ pour un entier $m \in \mathbb{N} - \{0,1\}$ comme ci-dessus. Deux tels sous-groupes diédraux de $\mathrm{O}(2)$ sont [99] conjugués dans $\mathrm{O}(2)$. Chaque sous-groupe diédral de $\mathrm{O}(2)$ laisse invariants, à homothétie

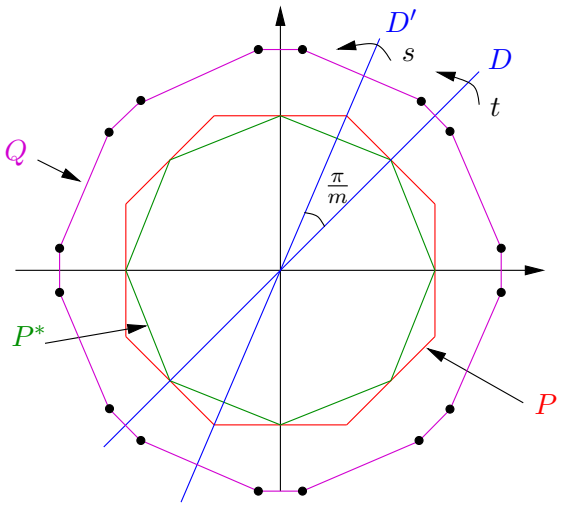
96. Certaines références le notent D_m.

97. Pour tout ensemble S, notons $L(S)$ le groupe libre sur S (voir par exemple [Pau4, Annexe B.2]). Pour toute partie R d'un groupe L, notons $\langle\langle R\rangle\rangle$ le sous-groupe distingué de L engendré par R, c'est-à-dire l'intersection de tous les sous-groupes distingués de L contenant R. Une *présentation* d'un groupe G est un triplet (S, R, ϕ) (ou un couple (S,R) quand ϕ est sous-entendu) où S est un ensemble, R une partie du groupe libre $L(S)$ et $\phi : L(S)/\langle\langle R\rangle\rangle \to G$ un isomorphisme de groupes. On dit aussi que G est *défini par générateurs et relations* par (S, R, ϕ). Si $S = \{s_1, \ldots, s_k\}$ et $R = \{r_1, \ldots, r_\ell\}$, on note parfois $\langle s_1, s_2, \ldots, s_k \mid r_1 = r_2 = \cdots = r_k = 1\rangle$ le groupe quotient $L(S)/\langle\langle R\rangle\rangle$.

98. Par exemple, le *polygone régulier standard* P_m est l'enveloppe convexe dans $\mathbb{R}^2$ de l'ensemble des racines m-èmes de l'unité $\{e^{2i\pi k/m} \ : \ k = 0, \ldots, m-1\}$, dont le groupe des isométries $G(P_m)$ est constitué des eléments $a^k : z \mapsto e^{2i\pi k/m} z$ et les $a^{k'}b$, où $b : z \mapsto \overline{z}$ et $k, k' \in \{0, \ldots, m-1\}$.

99. En effet, le groupe $\mathrm{O}(2)$ agit transitivement sur les couples de droites vectorielles faisant un angle $\pm\frac{\pi}{m}$.

près, exactement deux [100] polygones réguliers de barycentre 0 qui ont m côtés. Mais D_{2m} est isomorphe au groupe des isométries d'un ensemble non dénombrable (même à homothétie près) de polygones Q non réguliers du plan euclidien à homothétie près. Un tel polygone Q peut être défini comme l'enveloppe convexe de l'orbite par D_{2m} de tout point du cercle qui n'est pas fixe par l'une des réflexions de D_{2m} et n'est pas sur la médiatrice entre deux droites fixes par l'une des réflexions $ba^{k'}$ de D_{2m} et cycliquement consécutives.



Le groupe diédral D_{2m} est non abélien si et seulement si $m \geq 3$ (car nous avons alors $bab^{-1} = a^{-1} \neq a$). De plus, $D_2 \simeq \mathbb{Z}/2\mathbb{Z}$ et $D_4 = (\mathbb{Z}/2\mathbb{Z})^2$.

Lemme 4.4. *Pour tout $m \in \mathbb{N} - \{0, 1\}$, il existe un sous-groupe diédral, d'ordre $2m$, de* $\mathrm{SO}(3)$. *Tout tel sous-groupe est engendré par le groupe des isométries directes préservant un m-gone régulier de barycentre* 0, *par une rotation d'ordre m et un renversement valant* $-\mathrm{id}$ *sur l'axe de cette rotation, et par deux renversements par rapport à des droites vectorielles orientées faisant un angle $\frac{\pi}{m}$. Deux tels sous-groupes sont conjugués.*

Démonstration. Pour deux droites vectorielles orientées D et D' de $\mathbb{R}^3$ faisant dans cet ordre un angle $\frac{\pi}{m}$, le sous-groupe $G_{D,D'}$ de $\mathrm{SO}(3)$ engendré par les renversements respectivement t et s par rapport à ces droites est un groupe diédral. En effet, soit D'' la droite vectorielle orientée orthogonale à D et D', telle que, si les vecteurs directeurs unitaires définissant l'orientation de D, D', D'' sont notés v, v', v'' respectivement, alors la base (v, v', v'') de l'espace vectoriel orienté $\mathbb{R}^3$ soit directe. Alors la restriction de $G_{D,D'}$ au plan vectoriel $P = D \oplus D'$ orienté orthogonal à D'' est un sous-groupe diédral du groupe $\mathrm{O}(P)$, et st est une rotation d'axe D'' et d'ordre m, qui engendre $G_{D,D'}$ avec s. De plus, s induit $-\mathrm{id}$ sur l'axe D'' de st.

Comme le groupe $\mathrm{SO}(3)$ agit transitivement sur les couples de points de $\mathbb{S}_2$ faisant un angle donné, deux tels groupes $G_{D,D'}$ sont conjugués dans $\mathrm{SO}(3)$.

Le groupe $G_{D,D'}$ préserve le m-gone régulier du plan P qui est l'enveloppe convexe de l'orbite par $G_{D,D'}$ d'un vecteur non nul de D', et est exactement le groupe des rotations qui le préserve.

100. Les droites fixes des réflexions préservant un polygone régulier P doivent passer ou bien par un sommet de P ou bien par le milieu d'une arête de P. Les polygones réguliers à m côtés préservés par le groupe diédral $G(P)$ sont alors les homothétiques de P, ou ceux du polygone P^*, dit *dual* à P, dont les sommets sont les milieux des arêtes de P (voir la figure ci-dessous à homothétie près).

Soit G un sous-groupe diédral de $\mathrm{SO}(3)$ d'ordre $2m$. Les seuls éléments de $\mathrm{SO}(3)$ d'ordre 2 sont des renversements, donc en notant s et t des éléments d'ordre 2 de G tels que st soit d'ordre m, alors t et s sont des renversements, dont nous notons respectivement D et D' les droites vectorielles fixes. Ces droites sont distinctes car $m \geq 2$ donc $s \neq t$. Puisque le produit st est d'ordre m, et fixe la droite vectorielle D'' orthogonale à D et à D', c'est une rotation d'axe de rotation D'', et d'angle de rotation le double de l'angle entre D et D'. Donc $G = G_{D,D'}$. Le résultat en découle. □

Cas 2 : Supposons que $(n, n_1, n_2, n_3) = (12, 2, 3, 3)$.

Le groupe alterné $\mathfrak{A}_4$ des permutations de $\{1,2,3,4\}$ de signature paire est d'ordre $\frac{1}{2}\,4! = 12$. Notons (e_1, e_2, e_3, e_4) la base canonique de l'espace euclidien standard $\mathbb{R}^4$, et (x_1, x_2, x_3, x_4) ses coordonnées canoniques. La bijection $i \mapsto e_i$ de $\{1,2,3,4\}$ dans $\{e_1, e_2, e_3, e_4\}$ permet d'identifier[101] $\mathfrak{A}_4$ des permutations de $\{1,2,3,4\}$ de signature paire est d'ordre $\frac{1}{2}\,4! = 12$. Notons (e_1, e_2, e_3, e_4) la base canonique de l'espace euclidien standard $\mathbb{R}^4$, et (x_1, x_2, x_3, x_4) ses coordonnées canoniques. La bijection $i \mapsto e_i$ de $\{1,2,3,4\}$ dans $\{e_1, e_2, e_3, e_4\}$ permet d'identifier $\mathfrak{A}_4$ avec le sous-groupe du groupe linéaire $\mathrm{GL}_4(\mathbb{R})$ des permutations de déterminant 1 de la base canonique (e_1, e_2, e_3, e_4). Puisque le groupe $\mathfrak{A}_4$ envoie base orthonormée directe sur base orthonormée directe, il est contenu dans $\mathrm{SO}(4)$, et puisqu'il fixe la droite vectorielle $\mathbb{R}(1,1,1,1)$, il est contenu dans le fixateur dans $\mathrm{SO}(4)$ de $\mathbb{R}(1,1,1,1)$, et est égal au stabilisateur de $\{e_1, \ldots, e_4\}$ dans ce fixateur.

Notons $\mathscr{H}$ l'hyperplan euclidien affine d'équation $\sum_{i=1}^4 x_i = 1$, et Δ_3 l'enveloppe convexe de $\{e_1, \ldots, e_4\}$. Par l'unicité à isométrie près des espaces euclidiens de dimension donnée, il existe une isométrie de l'espace vectoriel $\vec{\mathscr{H}}$, obtenu en pointant $\mathscr{H}$ en le barycentre $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ de Δ_3, muni de la structure euclidienne induite de celle $\mathbb{R}^4$ sur l'espace euclidien standard $\mathbb{R}^3$. Appelons *tétraèdre régulier* de $\mathbb{R}^3$, et notons T l'image de Δ_3 par une telle isométrie et une homothétie de rapport $\frac{2}{\sqrt{3}}$ (l'inverse de la distance entre le barycentre $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ de Δ_3 et n'importe lequel de ses sommets). Ainsi $\mathfrak{A}_4$ s'identifie au sous-groupe de $\mathrm{SO}(3)$ préservant T. Le tétraèdre T, dont nous notons encore e_1, e_2, e_3, e_4 les sommets correspondants à ceux éponymes de Δ_3, est alors inscrit dans la sphère $\mathbb{S}_2$.

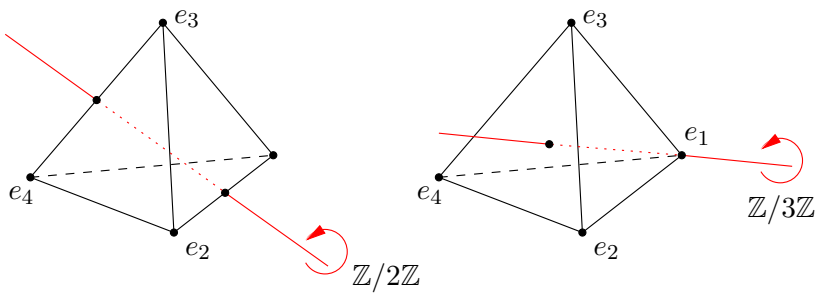
Le groupe $\mathfrak{A}_4$ contient, outre l'identité,

- trois doubles transpositions $(12)(34)$, $(13)(24)$ et $(14)(23)$, qui sont les rotations d'angles π par rapport aux médiatrices communes des trois paires d'arêtes opposées de T (voir la figure de gauche ci-dessous), et qui, avec l'identité, forment un sous-groupe distingué N de $\mathfrak{A}_4$ isomorphe à $(\mathbb{Z}/2\mathbb{Z})^2$, donc d'ordre 4,
- huit cycles de longueur 3, deux pour chacun des 4 sommets de T correspondant aux permutations cycliques non triviales des trois autres sommets (voir la figure de droite ci-dessous). Le sous-groupe engendré par le cycle (123) est noté H, et $\mathfrak{A}_4$ est isomorphe au produit semi-direct de N par H, pour l'action par conjugaison dans $\mathfrak{A}_4$ de H sur N. En effet, le quotient $H' = \mathfrak{A}_4/N$ est un groupe d'ordre 3 par cardinalité. Il est donc isomorphe au groupe cyclique $\mathbb{Z}/3\mathbb{Z}$. La suite exacte

$$1 \to N \xrightarrow{i} \mathfrak{A}_4 \xrightarrow{p} H' \to 1$$

101. par l'application de $\mathfrak{A}_4$ dans $\mathrm{GL}_4(\mathbb{R})$ qui à $\sigma \in \mathfrak{A}_4$ associe (la matrice dans la base canonique de) l'unique application linéaire u_σ qui envoie e_i sur $e_{\sigma^{-1}(i)}$ pour $i = 1, \ldots, 4$. Il est élémentaire de vérifier que $\sigma \mapsto u_\sigma$ est un morphisme de groupes injectif.

où $i : N \to \mathfrak{A}_4$ est l'inclusion et $p : \mathfrak{A}_4 \to H' = \mathfrak{A}_4/N$ est la projection canonique, est scindée, par la section $s : H' \to \mathfrak{A}_4$ d'image H, qui au générateur $p((123))$ de H' associe le générateur (123) de H.



Le groupe $\mathfrak{A}_4$ contient dix sous-groupes, qui sauf $\mathfrak{A}_4$ lui-même sont cycliques d'ordre 1, 2 ou 3 ou diédral d'ordre 4 : ce sont

- trois sous-groupes distingués, le sous-groupe trivial $\{\mathrm{id}\}$, le sous-groupe total $\mathfrak{A}_4$, et le sous-groupe $N \simeq (\mathbb{Z}/2\mathbb{Z})^2$, qui est l'unique 2-Sylow de $\mathfrak{A}_4$;
- trois sous-groupes cycliques d'ordre 2, les stabilisateurs des paires d'arêtes opposées, qui sont conjugués,
- quatre sous-groupes cycliques d'ordre 3, les fixateurs des 4 sommets de T, qui sont conjugués et sont les quatre 3-Sylow de $\mathfrak{A}_4$.

Le groupe $\mathfrak{A}_4$ admet pour présentation (voir par exemple [AB, Prop. IX.5])

$$\mathfrak{A}_4 \simeq \langle a, b \mid a^3 = b^3 = (ab)^2 = 1 \rangle$$

en prenant pour a le 3-cycle (123) et b le 3-cycle (234), qui vérifient clairement les égalités $a^3 = b^3 = (ab)^2 = 1$ (mais il faut encore montrer, en notant $G = L(\{a, b\})/\langle\langle\{a^3, b^3, (ab)^2\}\rangle\rangle$ le groupe quotient du groupe libre engendré par $\{a, b\}$ par le sous-groupe distingué engendré par $\{a^3, b^3, (ab)^2\}$, que l'unique morphisme de G dans $\mathfrak{A}_4$ envoyant l'image de a dans G sur a et l'image de b dans G sur b est un isomorphisme de groupes, voir loc. cit.).

Lemme 4.5. *Il existe un sous-groupe de* $\mathrm{SO}(3)$ *non cyclique et non diédral, d'ordre* 12. *Tout tel sous-groupe est isomorphe à* $\mathfrak{A}_4$*, et égal au groupe des isométries directes préservant un tétraèdre régulier de barycentre* 0. *Deux tels sous-groupes sont conjugués.*

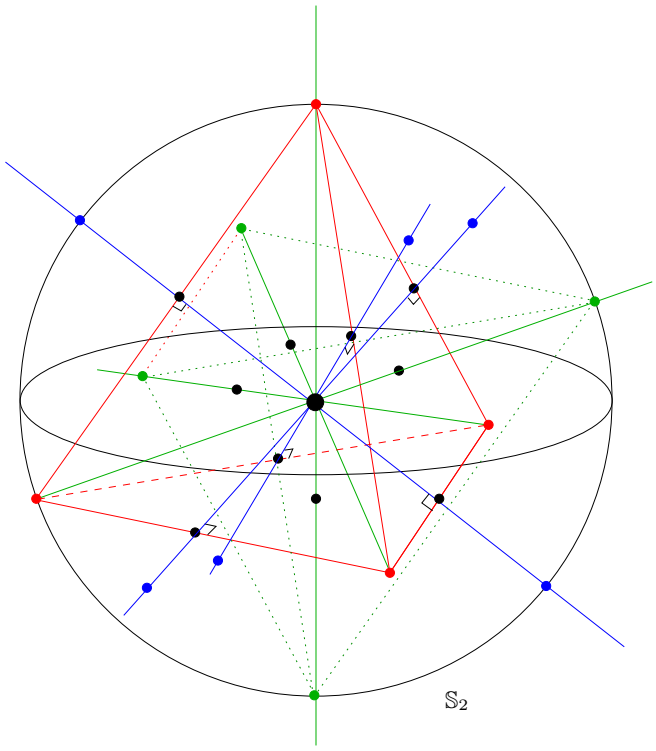
Démonstration. La première affirmation découle de la construction géométrique de $\mathfrak{A}_4$ ci-dessus.

Soit G un sous-groupe fini non cyclique non diedral de $\mathrm{SO}(3)$, d'ordre 12. Par l'analyse ayant conduit au **Cas 2**, G admet exactement trois orbites O_1, O_2, O_3 de points fixes d'élements non triviaux de G, dont le stabilisateur d'un élément est respectivement d'ordre 2, 3 et 3, et qui sont donc d'ordre respectivement 6, 4 et 4.

Dans le dessin ci-dessous,

- l'orbite O_2 est constituée des quatres points rouges, que nous montrerons former les sommets d'un tétraèdre régulier P (d'arêtes en traits rouges sur le dessin) dont G est le groupe des isométries directes,
- l'orbite O_1 est constituée des six points bleus (que nous montrerons être les points d'intersection avec la sphère des perpendiculaires communes (en trais bleus sur le dessin) aux paires d'arêtes opposées de P, et

• l'orbite O_3 est constituée des quatre points verts, que nous montrerons être les sommets du tétraèdre régulier P^* (d'arêtes en traits pointillés verts sur le dessin) dual (à homothétie près) de P, points d'intersection avec la sphère des hauteurs (en traits continus verts sur le dessin) du tétraèdre, dont le groupe des isométries directes est aussi égal à G.



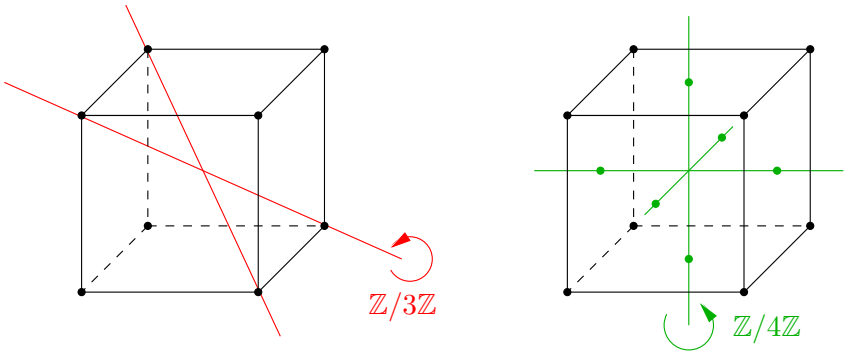
Notons P l'enveloppe convexe de O_2, qui est donc un tétraèdre. Le groupe G agit transitivement sur les sommets de ce tétraèdre, et le stabilisateur G_x d'un point x de O_2, qui est d'ordre 3 et préserve les trois autres sommets, est donc un groupe cyclique d'ordre 3, engendré par une rotation ρ d'axe contenant x et d'angle $\frac{2\pi}{3}$. Le groupe G_x agit donc simplement transitivement sur ces trois autres sommets, qui sont à même distance de x (puisque SO(3) préserve la distance euclidienne). Puisque G_x agit transitivement sur les trois arêtes ayant x pour sommet, toutes les arêtes de P ont la même longueur. Le groupe SO(3) agit transitivement sur la sphère $\mathbb{S}_2$, donc quitte à conjuguer G, nous pouvons supposer que $x = e_1$. Soit y un sommet de P différent de x. Par la transitivité de l'action du groupe des rotations d'axe passant par e_1 sur les points de la sphère $\mathbb{S}_2$ à distance $t \in [0, 2]$ donnée de e_1, nous pouvons supposer que y appartient au demi-cercle d'extrémité x passant par e_2. Comme la distance entre y et $\rho(y)$ est égale à la distance entre e_1 et y,

ceci n'est possible que si $y = e_2$. [102] Donc $P = T$. Comme G est d'ordre 12, contenu dans le groupe des isométries directes de T, qui est aussi d'ordre 12, nous avons $G = \mathfrak{A}_4$. Ceci montre le résultat. □

Cas 3 : Supposons que $(n, n_1, n_2, n_3) = (24, 2, 3, 4)$.

Le groupe symétrique $\mathfrak{S}_4$ des permutations de $\{1, 2, 3, 4\}$ est d'ordre $4! = 24$. Notons $C^3 = [-1, 1]^3$, appelé le *cube standard*, qui est un polyèdre convexe de l'espace euclidien standard $\mathbb{R}^3$, ayant

- 8 sommets $(\epsilon_1, \epsilon_2, \epsilon_3)$ où $\epsilon_i \in \{\pm 1\}$,
- 12 arêtes $[-1, 1] \times \{\epsilon_2\} \times \{\epsilon_3\}$, $\{\epsilon_1\} \times [-1, 1] \times \{\epsilon_3\}$ et $\{\epsilon_1\} \times \{\epsilon_2\} \times [-1, 1]$ où $\epsilon_i \in \{\pm 1\}$,
- 6 faces $[-1, 1]^2 \times \{\epsilon_3\}$, $[-1, 1] \times \{\epsilon_2\} \times [-1, 1]$ et $\{\epsilon_1\} \times [-1, 1]^2$ où $\epsilon_i \in \{\pm 1\}$.

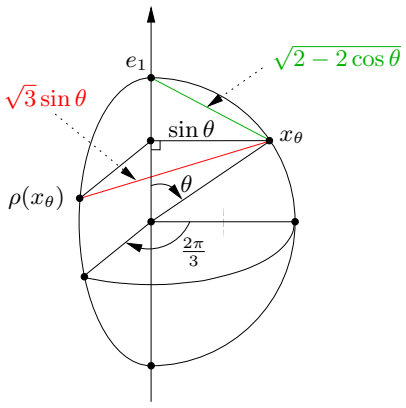


Notons G le groupe des isométries directes de l'espace affine euclidien orienté $\mathbb{R}^3$ préservant C^3. Elles fixent son barycentre, qui est le point $(0, 0, 0)$, donc G est un sous-groupe de $\mathrm{SO}(3)$. Appelons *grande diagonale* du cube C^3 toute droite vectorielle passant par deux sommets opposés de C^3. Elles sont au nombre de quatre, nous les numérotons (de manière indifférente) de 1 à 4. Le groupe G préserve l'ensemble des sommets de C^3, et envoie sommets opposés sur sommets opposés. Donc G préserve l'ensemble A des grandes diagonales. Les rotations d'angle $\pi/2$ autour des axes orientés de coordonnées appartiennent à G, donc le sous-groupe de $\mathrm{SO}(3)$ qu'elles engendrent est contenu dans G, et puisqu'elles agissent transitivement sur les sommets, l'action de G sur l'ensemble des sommets de C^3, donc sur A, est transitive. Ceci définit un morphisme de G dans $\mathfrak{S}_4$.

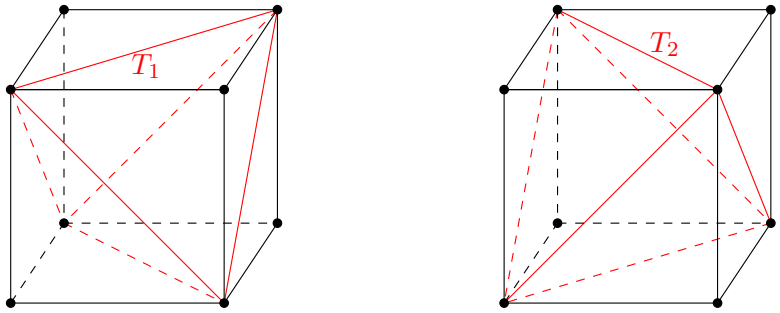
Comme la seule rotation envoyant un vecteur non nul sur son opposé est un renversement, et qu'un renversement ne vaut pas $-\mathrm{id}$ sur trois droites non coplanaires, tout

---

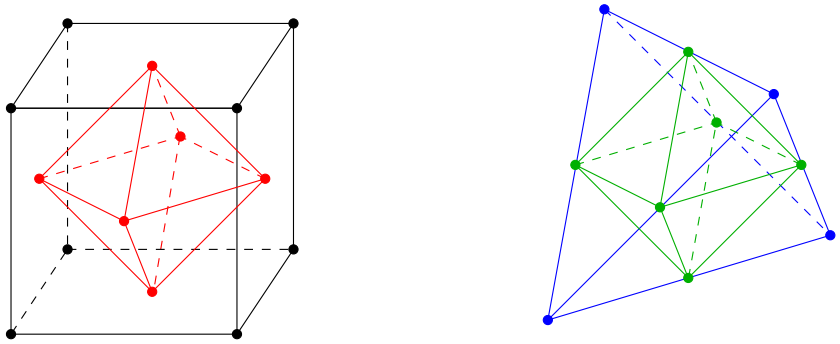
102. En effet, un calcul élémentaire montre que si x_θ est le point de ce demi-cercle faisant un angle $\theta \in]0, \pi[$ avec e_1, alors la distance entre x_θ et $\rho(x_\theta)$ est égale à $\sqrt{3}\sin\theta$, la distance entre x_θ et e_1 est égale à $\sqrt{2 - 2\cos\theta}$, et l'équation $\sqrt{3}\sin\theta = \sqrt{2 - 2\cos\theta}$ admet une et une seule solution $\theta = \arccos(-\frac{1}{3})$ dans $]0, \pi[$. Au passage, nous retrouvons que la longueur des arêtes du tétraèdre régulier inscrit dans la sphère $\mathbb{S}_2$ est $\sqrt{\frac{8}{3}}$.



élément de G qui préserve chaque grande diagonale doit valoir l'identité sur au moins une grande diagonale, et donc être une rotation d'angle $\frac{2\pi}{3}$, ce qui l'empêche de préserver les autres grandes diagonales. Donc le morphisme ci-dessus est injectif.



L'ensemble des sommets du cube est partitionné de manière unique en deux parties disjointes $A_1 \cup A_2$, chaque A_i ayant quatre éléments dont aucune paire n'est jointe par une arête de C^3. Par la transitivité sur les sommets, il existe une isométrie directe de C^3 qui envoie A_1 sur A_2. Le sous-groupe de G préservant A_1 est donc d'indice 2 dans G. Chaque A_i est l'ensemble des sommets d'un tétraèdre régulier T_i (voir le dessin ci-dessus), dont toute isométrie positive préserve C^3. Donc G est d'ordre 24, et le morphisme ci-dessus est un isomorphisme.



Le groupe $\mathfrak{S}_4$ a pour classes de conjugaison, outre celle de l'identité :

• la classe de conjugaison des six transpositions, qui sont les renversements du cube valant $-\mathrm{id}$ sur le plan vectoriel passant par l'une des six paires d'arêtes opposées, ou, autrement dit, la rotation d'angle π autour de la médiatrice commune de deux arêtes opposées,

• la classe de conjugaison des trois doubles transpositions, qui sont les rotations d'ordre 2 autour des droites vectorielles passant par les milieux des faces opposées du carré, donc par les sommets opposés de l'octaèdre dual du cube (voir le dessin ci-dessus à gauche),

• la classe de conjugaison des huit cycles de longueur 3, qui sont les deux rotations d'ordre 3 autour des quatre grandes diagonales, donc des droites vectorielles passant par les barycentres des faces opposées de l'octaèdre dual,

• la classe de conjugaison des six cycles de longueur 4, qui sont les deux rotations d'ordre 4 autour de chacune des trois droites vectorielles passant par deux sommets opposés de l'octaèdre dual, donc des droites vectorielles passant par les barycentres des faces opposées du cube.

Le groupe $\mathfrak{S}_4$ contient (voir par exemple [AB, §IX.1]) exactement trente sous-groupes, qui sauf $\mathfrak{S}_4$ et $\mathfrak{A}_4$ (qui est le stabilisateur de chacun des deux tétraèdres réguliers inscrits dans le cube, ou dans lesquels sont inscrits l'octaèdre dual, en tant qu'enveloppe convexe des milieux des arêtes du tétraèdre, voir les dessins ci-dessus) sont cycliques d'ordre 1, 2, 3 ou 4 ou diédraux d'ordre 4, 6 ou 8 :

- les quatre sous-groupes distingués, le sous-groupe trivial $\{\mathrm{id}\}$, le sous-groupe total $\mathfrak{S}_4$, le sous-groupe dérivé $\mathfrak{A}_4 = [\mathfrak{S}_4, \mathfrak{S}_4]$ et le sous-groupe $[\mathfrak{A}_4, \mathfrak{A}_4]$, qui comme vu dans le **Cas 2** est isomorphe à D_4 et constitué outre de l'identité des trois doubles transpositions ;
- les neuf sous-groupes cycliques d'ordre 2 répartis en deux classes de conjugais, engendrés par les transpositions ou les doubles transpositions, voir ci-dessus pour leur interprétation géométrique ;
- les trois 2-Sylow, qui sont diédraux d'ordre 8 ;
- les trois sous-groupes cycliques d'ordre 4 contenus dans les 2-Sylow ;
- les trois sous-groupes diédraux d'ordre 4 contenus dans les 2-Sylow ;
- les quatre 3-Sylow, qui sont cycliques d'ordre 3 (contenus dans $\mathfrak{A}_4$), voir ci-dessus pour leur interprétation géométrique ;
- les quatre stabilisateurs de points de $\{1, 2, 3, 4\}$, qui sont isomorphes à $\mathfrak{S}_3 \simeq D_6$.

Le groupe $\mathfrak{S}_4$ admet pour présentation (voir par exemple [AB, Prop. IX.6])

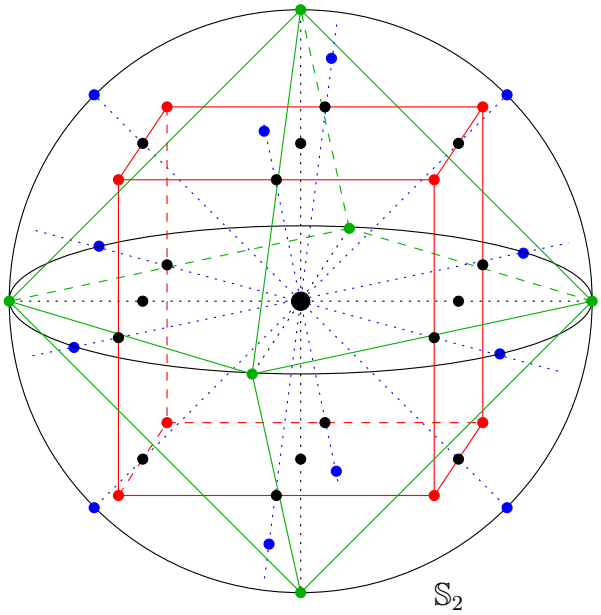
$$\mathfrak{A}_4 = \langle a, b \mid a^3 = b^4 = (ab)^2 = 1 \rangle$$

en prenant pour a le 3-cycle (234) et b le 4-cycle (1432), qui vérifient clairement $a^3 = b^4 = (ab)^2 = 1$ (mais il faut encore montrer, en notant $G = L(\{a, b\})/N(\{a^3, b^4, (ab)^2\})$ le groupe quotient du groupe libre engendré par $\{a, b\}$ par le sous-groupe distingué engendré par $\{a^3, b^4, (ab)^2\}$, que l'unique morphisme de G dans $\mathfrak{A}_4$ envoyant l'image de a dans G sur a et l'image de b dans G sur b est un isomorphisme de groupes, voir loc. cit.)

Lemme 4.6. *Il existe un sous-groupe de* $\mathrm{SO}(3)$ *non cyclique et non diédral, d'ordre* 24. *Tout tel sous-groupe est isomorphe à* $\mathfrak{S}_4$, *et est égal au groupe des isométries directes préservant un cube (ou, de manière équivalente, un octaèdre régulier) centré en* 0. *Deux tels sous-groupes sont conjugués.*

Démonstration. La première affirmation découle de la construction géométrique du groupe des isométries directes du cube standard ci-dessus.

Soit G un sous-groupe fini non cyclique non diedral de $\mathrm{SO}(3)$, d'ordre 24. Par l'analyse ayant conduit au **Cas 3**, G admet exactement trois orbites O_1, O_2, O_3 de points fixes de réflexions de G, qui sont d'ordre respectivement 12, 8 et 6, et dont le stabilisateur d'un élément est respectivement d'ordre 2, 3 et 4. Dans le dessin ci-dessous, O_2 est constitué des huits points rouges, que nous montrerons former les sommets d'un cube régulier P dont G est le groupe des isométries directes, O_1 est constitué des 12 points bleus (que nous montrerons être constitués des points d'intersection avec la sphère des perpendiculaires communes aux paires d'arêtes opposées de P, et O_3 des six points verts, que nous montrerons être les sommets de l'octaèdre régulier P^* dual (à homothétie près) de P, points d'intersection avec la sphère des médiatrices des faces opposées de P, dont le groupe des isométries directes est aussi G. Notons que contrairement au cas du tétraèdre, le polyèdre dual du cube n'est pas isométrique au cube.



Notons P l'enveloppe convexe de O_2, qui a donc 8 sommets. Le stabilisateur de chaque sommet, qui est d'ordre 3 est donc cyclique d'ordre 3, et G agit par cardinalité simplement transitivement sur les arêtes orientées. Comme chaque arête a deux sommets, il y a donc 12 arêtes. Notons que P est symétrique par rapport à l'origne (car si x est point fixe d'une rotation, alors $-x$ aussi), et par ce qui précède, tout sommet de P est trivalent, c'est-à-dire qu'exactement trois arêtes de P en partent. Puisque le stabilisateur G_y d'un point y de O_3 est d'ordre 4, il est cyclique d'ordre 4 car c'est un groupe de rotations ayant un point fixe. Donc les orbites de G_y dans O_2 sont deux carrés centrés sur le diamètre de $\mathbb{S}_2$ entre y et $-y$. Puisque les sommets sont trivalents, il doit y avoir une arête parallèle à cet axe entre les deux carrés, et par rotation, P est donc un produit d'un carré perpendiculaire à cet axe par un intervalle parallèle à cet axe. Puisque G agit transitivement sur le sommet, ce paralélotope est donc un cube. Puisque le groupe des isométries directe d'un cube est d'ordre 24, le groupe G est donc égal au groupe des isométries directes de P. Ceci montre la deuxième affirmation du lemme 4.6.

La dernière affirmation découle du fait que $\mathrm{SO}(3)$ agit transitivement sur les cubes centrés en 0 et inscrits dans la sphère, car il agit de manière transitive sur les points de la sphère, et le stabilisateur d'un point de la sphère agit transitivement sur les points de la sphère à distance donnée de ce point. □

Cas 4 : Supposons que $(n, n_1, n_2, n_3) = (60, 2, 3, 5)$.

Le groupe alterné $\mathfrak{A}_5$ des permutations de $\{1, 2, 3, 4, 5\}$ de signature paire est simple (voir par exemple [Per1]) d'ordre $\frac{1}{2}\,5! = 60$. [103]

Lemme 4.7. *Il existe un sous-groupe de* $\mathrm{SO}(3)$ *non cyclique et non diédral, d'ordre* 60. *Tout tel sous-groupe est isomorphe à* $\mathfrak{A}_5$*, engendré par le groupe des isométries directes préservant un dodécaèdre régulier (ou, de manière équivalente, un icosaèdre régulier). Deux tels sous-groupes sont conjugués.*

Démonstration. Nous renvoyons par exemple à [AB, Chap. IX]. □

103. En fait, tout groupe fini simple d'ordre 60 est isomorphe à $\mathfrak{A}_5$, voir [AB, t. 1, p. 25, théo. I.17], et le plus petit ordre d'un groupe fini simple non abélien est 60.

4.2 Groupes cristallographiques

Nous considérons maintenant quelques groupes d'isométries euclidiennes affines. Notons $\mathscr{E}$ un espace affine euclidien, et E son espace vectoriel associé. Pour simplifier les notations, nous fixons une fois pour toute un isomorphisme d'espace euclidien affine entre $\mathscr{E}$ et l'espace euclidien affine standard $\mathbb{R}^n$, associé à isomorphisme d'espace euclidien entre E et l'espace euclidien standard $\mathbb{R}^n$. Le lecteur saura remplacer quand nécessaire $\mathbb{R}^n$ par $\mathscr{E}$ ou E au besoin dans ce qui suit. Nous notons $\| \ \|$ et $\langle \ , \ \rangle$ la norme et le produit scalaire de $\mathbb{R}^n$.

Rappelons que le groupe $\mathrm{Aff}(\mathbb{R}^n)$ des bijections affines de l'espace affine standard $\mathbb{R}^n$ s'insère dans une suite exacte de groupes

$$1 \longrightarrow \mathbb{R}^n \xrightarrow{\tau} \mathrm{Aff}(\mathbb{R}^n) \xrightarrow{p} \mathrm{GL}(\mathbb{R}^n) = \mathrm{GL}_n(\mathbb{R}) \longrightarrow 1 \ ,$$

où $\tau : x \mapsto \tau_x$ associe à $x \in \mathbb{R}^n$ la *translation* $\tau_x : y \mapsto x + y$ de vecteur x, où $p : f \mapsto \overline{f}$ associe à une transformation affine f sa partie vectorielle $\overline{f}$, et où nous identifions comme d'habitude une bijection linéaire de $\mathbb{R}^n$ à sa matrice dans la base canonique de $\mathbb{R}^n$. Cette suite exacte est scindée [104], donc

$$\mathrm{Aff}(\mathbb{R}^n) \simeq \mathbb{R}^n \rtimes_\phi \mathrm{GL}_n(\mathbb{R}) \ ,$$

où $\phi : \mathrm{GL}_n(\mathbb{R}) \to \mathrm{Aut}(\mathbb{R}^n)$ est l'action linéaire de $\mathrm{GL}_n(\mathbb{R})$ sur $\mathbb{R}^n$. En particulier, pour tout $f \in \mathrm{Aff}(\mathbb{R}^n)$, il existe un unique $x \in \mathbb{R}^n$ et un unique $\rho \in \mathrm{GL}_n(\mathbb{R})$ tels que $f = \tau_x \circ \rho$.

Notons que l'application $(x, \rho) \mapsto \tau_x \circ \rho$ de $\mathbb{R}^n \times \mathrm{GL}_n(\mathbb{R})$ dans $\mathrm{Aff}(\mathbb{R}^n)$ est certes bijective, mais n'est pas un morphisme de groupes. En effet, la translation τ_x par $x \in \mathbb{R}^n$ et une transformation linéaire $\rho \in \mathrm{GL}_n(\mathbb{R})$ ne commutent en général pas :

$$\rho \circ \tau_x = \tau_{\rho(x)} \circ \rho \ . \tag{30}$$

Le groupe $\mathrm{Isom}(\mathbb{R}^n)$ des isométries euclidiennes affines de l'espace affine euclidien $\mathbb{R}^n$ est un sous-groupe de $\mathrm{Aff}(\mathbb{R}^n)$. Il s'insère dans la suite exacte de groupes

$$1 \longrightarrow \mathbb{R}^n \xrightarrow{\tau} \mathrm{Isom}(\mathbb{R}^n) \xrightarrow{p} \mathrm{O}(\mathbb{R}^n) \longrightarrow 1 \ ,$$

qui est aussi scindée. Donc est $\mathrm{Isom}(\mathbb{R}^n)$ est un produit semi-direct

$$\mathrm{Isom}(\mathbb{R}^n) \simeq \mathbb{R}^n \rtimes_\phi \mathrm{O}(n)$$

du groupe additif des translations affines de $\mathbb{R}^n$ et du groupe $\mathrm{O}(n)$ des isométries vectorielles de l'espace euclidien $\mathbb{R}^n$, pour l'action $\phi : \mathrm{O}(n) \to \mathrm{Aut}(\mathbb{R}^n, +)$ linéaire de $\mathrm{O}(n)$ sur $\mathbb{R}^n$. Nous munissons $\mathrm{Isom}(\mathbb{R}^n)$ de la topologie telle que l'application de produit soit un homéomorphisme. En particulier, $\mathrm{Isom}(\mathbb{R}^n)$ est métrisable complet, et localement compact (mais pas compact).

Nous noterons $\mathrm{Isom}^+(\mathbb{R}^n)$ le groupes des isométries affines euclidiennes directes (préservant l'orientation) de l'espace affine euclidien orienté $\mathbb{R}^n$, aussi appelé le groupe des *déplacements*. Tout élément $f \in \mathrm{Isom}^+(\mathbb{R}^n)$ s'écrit de manière unique $f = \tau_x \circ \rho$ où $x \in \mathbb{R}^n$ et $\rho \in \mathrm{SO}(n)$. Le groupe $\mathrm{Isom}^+(\mathbb{R}^n)$ est isomorphe au produit semi-direct $\mathbb{R}^n \rtimes_\phi \mathrm{SO}(n)$.

104. car toute transformation linéaire de l'espace vectoriel $\mathbb{R}^n$ est aussi une transformation affine de l'espace affine $\mathbb{R}^n$

Nous allons décrire dans cette partie quelques sous-groupes de $\mathrm{Isom}(\mathbb{R}^n)$, en commençant par des sous-groupes du groupe de translation $\mathbb{R}^n$.

Un *réseau* de l'espace euclidien $\mathbb{R}^n$ est un sous-groupe Λ du groupe additif de $\mathbb{R}^n$ tel qu'il existe une $\mathbb{R}$-base $(e_1, e_2, ..., e_n)$ de l'espace vectoriel réel $\mathbb{R}^n$ qui est une $\mathbb{Z}$-base du groupe abélien Λ (c'est-à-dire $\Lambda = \mathbb{Z}e_1 \oplus \cdots \oplus \mathbb{Z}e_n$). En particulier, un réseau de $\mathbb{R}^n$ est un groupe abélien libre de rang n. Nous renvoyons par exemple à [CS, Mar] pour une masse de renseignements sur les réseaux.

Nous appellerons *covolume* d'un réseau Λ de l'espace euclidien usuel $\mathbb{R}^n$, et noterons $\mathrm{vol}(\mathbb{R}^n)/\Lambda$ le volume euclidien du *parallélotope fondamental*

$$P_\Lambda = \{t_1 e_1 + \cdots + t_n e_n \ : \ (t_1, \ldots, t_n) \in [0,1]^n\}$$

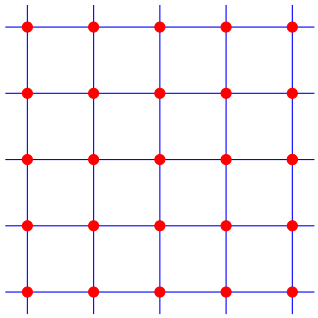
défini par une $\mathbb{Z}$-base $(e_1, e_2, ..., e_n)$ de Λ (et qui dépend de cette $\mathbb{Z}$-base). Ceci ne dépend pas du choix de cette base, car toute matrice de $\mathrm{GL}_n(\mathbb{Z})$ est de déterminant ± 1 (les seuls inversibles de l'anneau $\mathbb{Z}$), donc préserve le volume.

Comme le groupe $\mathrm{GL}_n(\mathbb{R})$ agit transitivement sur les $\mathbb{R}$-bases de l'espace vectoriel $\mathbb{R}^n$, comme le stabilisateur du *réseau standard* $\mathbb{Z}^n$ est $\mathrm{GL}_n(\mathbb{Z})$, l'application $g \mapsto g\mathbb{Z}^n$ de $\mathrm{GL}_n(\mathbb{R})$ dans l'ensemble $\mathscr{R}_n$ des réseaux de l'espace euclidien $\mathbb{R}^n$ induit par passage au quotient une bijection

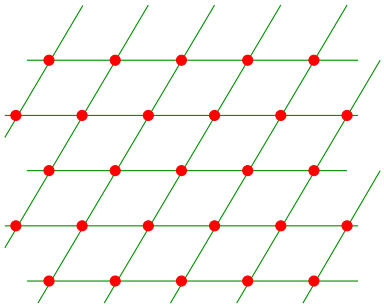
$$\mathrm{GL}_n(\mathbb{R})/\mathrm{GL}_n(\mathbb{Z}) \ \simeq \ \mathscr{R}_n \ .$$

Cette bijection identifie le sous-espace de $\mathscr{R}_n$ constitué des réseaux de covolume 1 avec $\mathrm{SL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z})$. Nous renvoyons par exemple à [Pau1] pour des informations élémentaires sur la topologie de l'espace de réseaux $\mathscr{R}_n$.

Les deux exemples les plus connus de réseaux du plan euclidien $\mathbb{R}^2 = \mathbb{C}$ sont le *réseau de Gauss* $\mathbb{Z} + i\mathbb{Z}$ et le *réseau d'Eisenstein* $\mathbb{Z} + e^{\frac{i\pi}{3}}\mathbb{Z} = \mathbb{Z} + e^{\frac{2i\pi}{3}}\mathbb{Z}$.



$\mathbb{Z} + i\mathbb{Z}$



$\mathbb{Z} + e^{\frac{2i\pi}{3}}\mathbb{Z}$

Nous donnons dans la proposition suivante une caractérisation topologique des réseaux.

Rappelons qu'une partie E d'un espace topologique X est *discrète* si la topologie induite sur E est la topologie discrète $\mathscr{P}(E)$ de E, ou, de manière équivalente, si tout point x de E est isolé dans E (c'est-à-dire admet un voisinage dans X qui ne rencontre E qu'en x). Rappelons que si X est un espace topologique muni d'une action (à gauche) d'un groupe G, et si $\pi : X \to G\backslash X$ est la projection canonique sur l'espace des orbites, alors la *topologie quotient* sur $G\backslash X$ est la topologie dont les ouverts sont les parties de $G\backslash X$ dont l'image réciproque par π est un ouvert de X, ou, de manière équivalente, la plus petite topologie rendant continue π.

Proposition 4.8. *Un sous-groupe Λ du groupe additif de $\mathbb{R}^n$ est un réseau si et seulement s'il vérifie l'une des conditions équivalentes suivantes :*

(i) Λ est discret, et le quotient $\mathbb{R}^n/\Lambda$ est compact ;

(ii) Λ est discret et engendre $\mathbb{R}^n$ en tant qu'espace vectoriel réel.

Démonstration. Montrons qu'être un réseau implique (i). Soit $(e_1, e_2, ..., e_n)$ une $\mathbb{R}$-base de $\mathbb{R}^n$. Il est immédiat que $\Lambda = \mathbb{Z}e_1 + \cdots + \mathbb{Z}e_n$ est discret dans $\mathbb{R}^n$. Soit K l'ensemble des $\lambda_1 e_1 + \cdots + \lambda_n e_n$ où $\lambda_1 \in [0, 1]$, qui est compact dans $\mathbb{R}^n$. La projection canonique $\pi : \mathbb{R}^n \to \mathbb{R}^n/\Lambda$ envoie ouvert sur ouvert, car si U est ouvert de $\mathbb{R}^n$, alors $\pi^{-1}(\pi(U)) = \bigcup_{x\in\Lambda} x + U$ est ouvert de $\mathbb{R}^n$, donc $\pi(U)$ est un ouvert de $\mathbb{R}^n/\Lambda$. Par conséquent, si deux éléments x et y dans $\mathbb{R}^n$ ont des images distinctes dans $\mathbb{R}^n/\Lambda$, si $\epsilon = \inf\{\|\lambda\| \ : \ \lambda \in \Lambda - \{0\}\}$, qui est strictement positif par discrétude, alors les images par π des boules ouvertes $B(x, \frac{\epsilon}{2})$ et $B(y, \frac{\epsilon}{2})$ sont des voisinages ouverts disjoints de $\pi(x)$ et $\pi(y)$.

Montrons que (i) implique (ii). Soit E' le sous-espace vectoriel réel de $\mathbb{R}^n$ engendré par Λ, et E'' un supplémentaire de E'. Alors l'isomorphisme naturel $\mathbb{R}^n \to E'\times''$ induit un homéomorphisme de $\mathbb{R}^n/\Lambda$ sur $(E'/\Lambda) \times E''$. Comme $E'/\Lambda \times E''$ est compact et E'/Λ non vide, l'image de la projection (continue)sur le second facteur, qui est E'', est compact. Donc E'' est réduit à $\{0\}$.

Montrons que (ii) implique être un réseau. Nous pouvons supposer que $n \geq 1$. Construisons par récurrence sur k des éléments linéairement indépendants e_k dans Λ avec $1 \leq k \leq n$. Soit e_1 un élément de $\Lambda - \{0\}$ de norme minimale, ce qui est possible car Λ est discret. Si $n > 2$, supposons construits $e_1, \ldots, e_k$ linéairement indépendants, avec $k \leq n - 1$, et posons $E_k = \mathbb{R}e_1 + \cdots + \mathbb{R}e_k$. Le sous-groupe $\Lambda_k = \mathbb{Z}e_1 + \cdots + \mathbb{Z}e_k$ est contenu dans Λ. Par l'implication être un réseau entraîne (ii), l'espace E_k/Λ_k est compact. Donc il existe $r_1 > 0$ tel que E_k soit recouvert par les translatés par les éléments de Λ_k de la boule $B(0, r_1)$. L'application $f : \mathbb{R}^n \to [0, +\infty[$ définie par $f : v \mapsto \min\{w \in \mathbb{R}^n \ : \ \|v - w\|\}$ est bien définie, car $f(v) = \|v - v^\perp\|$ où $v^\perp$ est la projection orthogonale de v sur E_k, et elle est continue. Comme Λ engendre l'espace vectoriel $\mathbb{R}^n$ et puisque la dimension de E_k est au plus $n-1$, il existe (au moins) un point v_k dans $\Lambda - E_k$. Poons $r_2 = f(v_k)$. Alors la borne inférieure de f sur $\Lambda - E_k$, qui est la borne inférieure de f sur le compact $B(0, r_1 + r_2)$, à cause de l'invariance de f par translations par Λ_k, est atteinte, en au moins un point, noté e_{k+1}. Ce point n'appartient pas à E_k, donc les vecteurs $e_1, \ldots, e_{k+1}$ sont linéairement indépendants, ce qui complète la construction par récurrence.

Alors la suite $(e_1, \ldots, e_n)$ est une $\mathbb{R}$-base de $\mathbb{R}^n$ (car libre), et une $\mathbb{Z}$-base de Λ. En effet, par l'absurde, s'il existe $x \in \Lambda - (\mathbb{Z}e_1 + \cdots + \mathbb{Z}e_n)$, alors, quitte à enlever à x un élément de $\mathbb{Z}e_1 + \cdots + \mathbb{Z}e_n$, il existe k avec $1 \leq k \leq n$ et $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ tels que $x = \lambda_1 e_1 + \cdots + \lambda_k e_k$, $\lambda_k \neq 0$ et $|\lambda_k| \leq \frac{1}{2}$. Mais alors $x \in \Lambda - E_{k-1}$ (en posant par convention $E_0 = \{0\}$) et $\min\{w \in E_{k-1}\|x - w\| \leq \|\lambda_k e_k\| < \|e_k\|$, ce qui contredit la propriété de minimalité. □

Nous allons nous intéreser à la structure des groupes discrets d'isométries des espaces affines euclidiens.

Un *groupe cristallographique* de dimension n est un sous-groupe Γ du groupe $\mathrm{Isom}(\mathbb{R}^n)$ des isométries euclidiennes affines de $\mathbb{R}^n$ tel que son *groupe des translations*

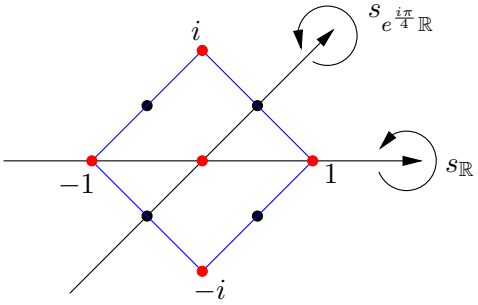
$$\Lambda_\Gamma = \{x \in \mathbb{R}^n \ : \ \tau_x \in \Gamma\}$$

soit un réseau de $\mathbb{R}^n$.

Pour tout réseau Λ de $\mathbb{R}^n$, notons $\mathrm{O}(\Lambda) = \{g \in \mathrm{O}(n) \ : \ g(\Lambda) = \Lambda\}$ le groupe des isométries vectorielles de l'espace euclidien $\mathbb{R}^n$ préservant le réseau Λ. Ce stabilisateur doit préserver l'ensemble $\{x \in \Lambda - \{0\} \ : \ \|x\| = \min_{y\in\Lambda-\{0\}} \|y\|\}$ des vecteurs de norme minimale. Notons que cet ensemble ne contient pas forcément de $\mathbb{Z}$-base de Λ.

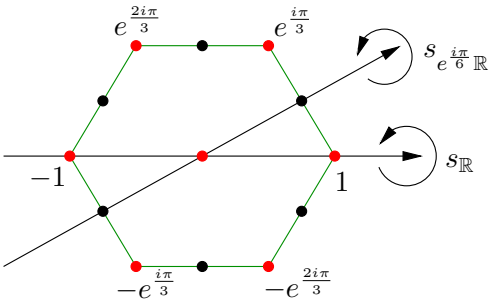
Par exemple si $n = 2$, le stabilisateur du réseau de Gauss $\mathbb{Z} + i\,\mathbb{Z}$, qui a exactement 4 vecteurs non nuls de norme minimale $\{\pm 1, \pm i\}$ qui sont les sommets du carré, est contenu dans le groupe des isométries de ce carré, et lui est en fait égal. C'est le groupe diédral D_8 d'ordre 8 engendré par la réflexion orthogonale $s_{\mathbb{R}}$ par rapport à l'axe réel et la réflexion orthogonale $s_{e^{\frac{i\pi}{4}}\mathbb{R}}$ par rapport à la droite vectorielle $e^{\frac{i\pi}{4}}\mathbb{R}$:

$$\mathrm{O}(\mathbb{Z} + i\,\mathbb{Z}) = D_8$$



De même, le stabilisateur du réseau d'Eisenstein $\mathbb{Z} + e^{\frac{i\pi}{3}}\,\mathbb{Z}$, qui a exactement 6 vecteurs non nuls de norme minimale $\{\pm 1, \pm e^{\frac{i\pi}{3}}, \pm e^{\frac{2i\pi}{3}}\}$ qui sont les sommets d'un hexagone régulier, est contenu (et en fait égal) dans le groupe des isométries de cet hexagone : c'est le groupe diédral D_{12} d'ordre 12 engendré par la réflexion orthogonale par rapport à l'axe réel et la réflexion orthogonale $s_{e^{\frac{i\pi}{6}}\mathbb{R}}$ par rapport à la droite $e^{\frac{i\pi}{6}}\mathbb{R}$:

$$\mathrm{O}(\mathbb{Z} + e^{\frac{i\pi}{3}}\,\mathbb{Z}) = D_{12}$$



Lemme 4.9. *Pour tout réseau Λ de $\mathbb{R}^n$, le groupe $\mathrm{O}(\Lambda)$ est fini.*

Démonstration. Si $r > 0$ est tel que la boule $B(0, r)$ de centre 0 et de rayon R de $\mathbb{R}^n$ contient une $\mathbb{Z}$-base de Λ, alors $\mathrm{O}(\Lambda)$ préserve $B(0, r) \cap \Lambda$. Cette intersection est finie car Λ est discret, et l'action de $\mathrm{O}(\Lambda)$ sur cette intersection est injective (car un élément de $\mathrm{GL}_n(\mathbb{R})$ est déterminé par son action sur une $\mathbb{R}$-base de $\mathbb{R}^n$). Donc $\mathrm{O}(\Lambda)$ est fini. □

Pour tout sous-groupe G de $\mathrm{O}(\Lambda)$, notons

$$\Lambda \rtimes G = \{\tau_x \circ g \ : \ x \in \Lambda, \ \ g \in G\} \ .$$

Par la formule (30) et puisque G préserve Λ, c'est un sous-groupe du groupe $\mathrm{Isom}(\mathbb{R}^n)$. De plus, c'est un groupe cristallographique, de groupe des translations égal à Λ par construction :

$$\Lambda_{\Lambda\rtimes G} = \Lambda \ .$$

La notation $\Lambda \rtimes G$ est cohérente avec le fait que ce groupe est en effet isomorphe au produit semi-direct de Λ et de G (pour l'action linéaire de G sur Λ). Notons que $\Lambda \rtimes G$ est un sous-groupe de $\Lambda \rtimes \mathrm{O}(\Lambda)$.

Mais tout groupe cristallographique n'est pas forcément de cette forme (seulement 13 sur 17 en dimension 2, voir le théorème 4.11). Le premier théorème de Bieberbach ci-dessous explique ce qui reste valable.

Nous allons nous intéresser à la classification des groupes cristallographiques modulo conjugaison dans $\mathrm{Aff}(\mathbb{R}^n)$, mais pas modulo conjugaison dans $\mathrm{Isom}(\mathbb{R}^n)$, ce qui serait trop peu restrictif).

Par exemple, les réseaux $\mathbb{Z}+\lambda\, i\, \mathbb{Z}$ où $\lambda > 1$ et $\mathbb{Z}+e^{i\theta}\,\mathbb{Z}$ pour $\theta \in]0, \frac{\pi}{2}[$ sont des groupes cristallographiques conjugués au réseau de Gauss $\mathbb{Z} + i\,\mathbb{Z}$ dans $\mathrm{Aff}(\mathbb{R}^2)$ (par $\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\lambda} \end{pmatrix}$ et $\begin{pmatrix} 1 & 0 \\ 0 & i\, e^{-i\theta} \end{pmatrix}$ respectivement), mais pas dans $\mathrm{Isom}(\mathbb{R}^2)$, comme précisé dans le résultat suivant.

Exercice E.43. *Montrer que deux réseaux de $\mathbb{R}^n$ sont conjugués dans $\mathrm{Aff}(\mathbb{R}^n)$. Montrer que tout réseau de covolume 1 de $\mathbb{R}^2$ est isométrique à un réseau de la forme $\mathbb{Z}+\tau\mathbb{Z}$ où $\mathrm{Im}\ \tau > 0$. Montrer que deux réseaux $\mathbb{Z}+\tau\mathbb{Z}$ et $\mathbb{Z}+\tau'\mathbb{Z}$, où $\mathrm{Im}\ \tau > 0$ et $\mathrm{Im}\ \tau' > 0$, sont isométriques si et seulement s'il existe une matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ telle que $\tau' = \frac{az+b}{cz+d}$.*

Le théorème suivant résume les propriétés structurelles des groupes cristallographiques. Sa première assertion est une caractérisation topologique des groupes cristallographiques, généralisant (et redémontrant) la proposition 4.8.

Théorème 4.10. (Théorèmes de Bieberbach)

(1) **(Premier théorème de Bieberbach)** *Un sous-groupe Γ de $\mathrm{Isom}(\mathbb{R}^n)$ est cristallographique si et seulement s'il est discret dans $\mathrm{Isom}(\mathbb{R}^n)$ et si l'espace topologique quotient $\Gamma\backslash\mathbb{R}^n$ est compact. Il existe un sous-groupe fini $\overline{\Gamma}$ de $\mathrm{O}(\Lambda_\Gamma)$ et une suite exacte (pas forcément scindée)*
$$1 \longrightarrow \Lambda_\Gamma \longrightarrow \Gamma \longrightarrow \overline{\Gamma} \longrightarrow 1\ .$$

(2) *Pour tout groupe cristallographique Γ de $\mathbb{R}^n$, le sous-groupe de translations $\Gamma \cap \tau(\mathbb{R}^n)$ est abélien maximal et distingué dans Γ. Tout sous-groupe abélien et distingué de Γ est contenu dans $\Gamma \cap \tau(\mathbb{R}^n)$.*

(3) **(Deuxième théorème de Bieberbach)** *Pour tout $n \in \mathbb{N} - \{0\}$, il existe un nombre fini de groupes cristallographiques de dimension n à isomorphisme près.*

(4) *Pour tout n, il existe $N(n)$ tel que pour tout groupe cristallographique Γ de $\mathbb{R}^n$, l'indice du sous-groupe de translations $\Gamma \cap \tau(\mathbb{R}^n)$ dans Γ est au plus $N(n)$.*

(5) **(Troisième théorème de Bieberbach)** *Deux sous-groupes cristallographiques de $\mathbb{R}^n$ sont isomorphes si et seulement s'ils sont conjugués dans $\mathrm{Aff}(\mathbb{R}^n)$.*

Démonstration. (1) Rappelons que l'application $p : \mathrm{Aff}(\mathbb{R}^n) \to \mathrm{GL}_n(\mathbb{R})$ qui à une transformation affine $f \in \mathrm{Aff}(\mathbb{R}^n)$ de l'espace affine $\mathbb{R}^n$ associe sa partie vectorielle $\overline{f} \in \mathrm{GL}_n(\mathbb{R})$ vérifie
$$f \circ \tau_x \circ f^{-1} = \tau_{\overline{f}(x)}\ .$$

Pour tout sous-groupe Γ de $\mathrm{Isom}(\mathbb{R}^n)$, notons $\overline{\Gamma} = p(\Gamma)$, qui est un sous-groupe du groupe compact $\mathrm{O}(n)$, et $\Lambda_\Gamma = \{\lambda \in \mathbb{R}^n \ : \ \tau_\lambda \in \Gamma\}$, qui est un sous-groupe du groupe abélien $\mathbb{R}^n$. Nous avons donc une suite exacte

$$1 \longrightarrow \Lambda_\Gamma \xrightarrow{\tau} \Gamma \xrightarrow{p} \overline{\Gamma} \longrightarrow 1 \ .$$

Pour tous les $\gamma \in \Gamma$ et $\lambda \in \Lambda_\Gamma$, nous avons $\tau_{\overline{\gamma}(\lambda)} = \gamma \circ \tau_\lambda \circ \gamma^{-1}$ appartient à Γ, donc $\gamma(\lambda) \in \Lambda_\Gamma$ et en particulier $\overline{\Gamma}$ est contenu dans $\mathrm{O}(\Lambda_\Gamma)$.

Supposons que Γ soit un groupe cristallographique. Montrons que $\overline{\Gamma}$ est fini. Ceci montrera que le groupe Γ est discret dans $\mathrm{Isom}(\mathbb{R}^n)$, car ayant un sous-groupe d'indice fini qui l'est. Ceci montrera que $\Gamma\backslash\mathbb{R}^n$ est compact, car séparé par une démonstration analogue à celle du point (i) de la proposition 4.8, et image d'un parallélogramme fondamental P_Λ (qui est compact) par la surjection canonique $\mathbb{R}^n \to \Gamma\backslash\mathbb{R}^n$, qui est continue par la définition de la topologie quotient.

Soit $(\gamma_k)_{k\in\mathbb{N}}$ une suite dans Γ et $g_k = \overline{\gamma_k}$. Pour tout $\lambda \in \Lambda_\Gamma$, comme vu ci-dessus, $g_k(\lambda) \in \Lambda_\Gamma$. Puisque Λ_Γ est discret et puisque la suite des $g_k(\lambda)$ est bornée (de norme inférieure à celle de λ), ceci implique que la suite $(g_k(\lambda))_{k\in\mathbb{N}}$ est constante à partir d'un certain rang. En faisant varier λ dans une $\mathbb{Z}$-base de Λ_Γ qui est une $\mathbb{R}$-base de $\mathbb{R}^n$, puisqu'un élément de $\mathrm{O}(n)$ est déterminé par les valeurs qu'elle prend sur une base, ceci dit que la suite $(g_k)_{k\in\mathbb{N}}$ est constante à partir d'un certain rang. Donc $\overline{\Gamma}$ est fini.

Réciproquement, supposons que Γ est discret dans $\mathrm{Isom}(\mathbb{R}^n)$ et que $\mathbb{R}^n/\Gamma$ est compact. Alors Λ_Γ, dont l'image par l'homéomorphisme $\lambda \mapsto \tau_\lambda$ est un sous-groupe de Γ, est discret dans $\mathbb{R}^n$. Nous renvoyons par exemple à [Bus] pour une démonstration courte du fait que $\overline{\Gamma}$ est discret dans $\mathrm{O}(n)$ donc fini par compacité. Ceci implique que l'image de Λ_Γ dans Γ est d'indice fini dans Γ, donc que $\mathbb{R}^n/\Lambda$ est compact, ce qui conclut par la proposition 4.8.

(2) Si $\gamma \in \Gamma$ commute avec τ_λ pour tout $\lambda \in \Lambda_\Gamma$, alors la relation $\tau_{\overline{\gamma}(\lambda)} = \gamma \circ \tau_\lambda \circ \gamma^{-1}$ implique que l'application linéaire $\overline{\gamma}$ fixe Λ_Γ, qui engendre $\mathbb{R}^n$ en tant qu'espace vectoriel. Donc $\overline{\gamma} = \mathrm{id}$ et γ appartient au noyau de la restriction de p à Γ, qui est égal à $\Gamma \cap \tau(\mathbb{R}^n)$. Donc celui-ci est un sous-groupe abélien maximal de Γ.

Si Γ' est un sous-groupe abélien distingué de Γ, alors pour touts les $\gamma \in \Gamma'$ et $\lambda \in \Lambda_\Gamma$, nous avons

$$\tau_{\overline{\gamma}(\lambda)-\lambda} = (\gamma\tau_\lambda\gamma^{-1})\tau_\lambda^{-1}) = \gamma(\tau_\lambda\gamma^{-1}\tau_\lambda^{-1}) = (\tau_\lambda\gamma^{-1}\tau_\lambda^{-1})\gamma = \tau_\lambda(\gamma^{-1}\tau_\lambda^{-1}\gamma) = \tau_{\lambda-\overline{\gamma}^{-1}(\lambda)} \ .$$

Donc l'application linéaire $(\overline{\gamma} - \mathrm{id})^2$ vaut 0 sur Λ_Γ, qui engendre $\mathbb{R}^n$ en tant qu'espace vectoriel. Par conséquent $(\overline{\gamma} - \mathrm{id})^2$ est nulle. Puisque $\overline{\gamma} \in \mathrm{O}(n)$, son polynôme minimal n'a pas de racine multiple (voir le théorème 2.4). Donc $\overline{\gamma} = \mathrm{id}$ et γ appartient au noyau de la restriction de p à Γ. Ceci montre que Γ' est contenu dans $\Gamma \cap \tau(\mathbb{R}^n)$

(3) Nous renvoyons par exemple à [Bus, §5] et [AB, Theo. XVII.23].

(4) Ceci découle de l'assertion précédente.

(5) Deux groupes cristallographiques conjugués dans $\mathrm{Aff}(\mathbb{R}^n)$ sont en particulier isomorphes. Pour la réciproque, nous renvoyons par exemple à [AB, Theo. XVII.20]. □

Le problème (une partie du 18ème problème de Hilbert, voir [Mil]) du calcul du nombre crist_n de classes d'isomorphisme (ou, de manière équivalente par le théorème de Bieberbach 4.10 (4), de classes de conjugaison dans $\mathrm{Isom}(\mathbb{R}^n)$) est largement ouvert. Il est immédiat

que $\mathrm{crist}_1 = 2$. Nous montrerons que $\mathrm{crist}_2 = 17$ dans le résultat suivant. Nous avons $\mathrm{crist}_3 = 219$ par un résultat de Fedorov, Schoenfliess et Barlow, et $\mathrm{crist}_4 = 4783$, voir [BBN+].

Le théorème suivant est un théorème de classification des groupes cristallographiques du plan affine euclidien. Notons $\langle P \rangle$ le sous-groupe de $\mathrm{Isom}(\mathbb{R}^n)$ engendré par une partie P de $\mathrm{Isom}(\mathbb{R}^n)$ et $\mathbb{Z}[e^{i\theta}] = \mathbb{Z} + e^{i\theta}\,\mathbb{Z} = \langle \tau_1, \tau_{e^{i\theta}} \rangle$ le groupe engendré par des translations planes τ_1 et $\tau_{e^{i\theta}}$.

Théorème 4.11. *Il existe exactement* 17 *classes d'isomorphismes de groupes cristallographiques du plan affine euclidien* $\mathbb{R}^2$*, donc* 5 *sont contenus dans* $\mathrm{Isom}^+(\mathbb{R}^n)$ *qui sont :*

(1) $\Gamma_1 = \mathbb{Z}[i] \in \mathrm{Isom}^+(\mathbb{R}^n)$,

(2) $\Gamma_2 = \mathbb{Z}[i] \rtimes G_2 \in \mathrm{Isom}^+(\mathbb{R}^n)$, *où* $G_2 = \langle -\mathrm{id} : z \mapsto -z \rangle \simeq \mathbb{Z}/2\mathbb{Z}$,

(3) $\Gamma_3 = \mathbb{Z}[i] \rtimes G_3$, *où* $G_3 = \langle s_{\mathbb{R}} : z \mapsto \overline{z} \rangle \simeq \mathbb{Z}/2\mathbb{Z}$,

(4) $\Gamma_4 = \mathbb{Z}[i] \rtimes G_4$, *où* $G_4 = \langle s_{\mathbb{R}e^{\frac{i\pi}{4}}} : z \mapsto i\,\overline{z} \rangle \simeq \mathbb{Z}/2\mathbb{Z}$,

(5) $\Gamma_5 = \mathbb{Z}[i] \rtimes G_5 \in \mathrm{Isom}^+(\mathbb{R}^n)$, *où* $G_5 = \langle s_{\mathbb{R}} : z \mapsto i\,z \rangle \simeq \mathbb{Z}/4\mathbb{Z}$,

(6) $\Gamma_6 = \mathbb{Z}[i] \rtimes G_6$, *où* $G_6 = \langle -\mathrm{id}, s_{\mathbb{R}} \rangle \simeq D_4$,

(7) $\Gamma_7 = \mathbb{Z}[i] \rtimes G_7$, *où* $G_7 = \langle -\mathrm{id}, s_{\mathbb{R}e^{\frac{i\pi}{4}}} \rangle \simeq D_4$,

(8) $\Gamma_8 = \mathbb{Z}[i] \rtimes G_8$, *où* $G_8 = \langle s_{\mathbb{R}}, s_{\mathbb{R}e^{\frac{i\pi}{4}}} \rangle \simeq D_8$,

(9) $\Gamma_9 = \mathbb{Z}[e^{\frac{i\pi}{3}}] \rtimes G_9 \in \mathrm{Isom}^+(\mathbb{R}^n)$, *où* $G_9 = \langle z \mapsto e^{\frac{2i\pi}{3}} z \rangle \simeq \mathbb{Z}/3\mathbb{Z}$,

(10) $\Gamma_{10} = \mathbb{Z}[e^{\frac{i\pi}{3}}] \rtimes G_{10} \in \mathrm{Isom}^+(\mathbb{R}^n)$, *où* $G_{10} = \langle z \mapsto e^{\frac{i\pi}{3}} z \rangle \simeq \mathbb{Z}/6\mathbb{Z}$,

(11) $\Gamma_{11} = \mathbb{Z}[e^{\frac{i\pi}{3}}] \rtimes G_{11}$, *où* $G_{11} = \langle s_{\mathbb{R}}, s_{\mathbb{R}e^{\frac{i\pi}{3}}} \rangle \simeq D_6$,

(12) $\Gamma_{12} = \mathbb{Z}[e^{\frac{i\pi}{3}}] \rtimes G_{12}$, *où* $G_{12} = \langle z \mapsto e^{\frac{2i\pi}{3}} z, s_{\mathbb{R}e^{\frac{i\pi}{6}}} \rangle \simeq D_6$,

(13) $\Gamma_{13} = \mathbb{Z}[e^{\frac{i\pi}{3}}] \rtimes G_{13}$, *où* $G_{13} = \langle s_{\mathbb{R}}, s_{\mathbb{R}e^{\frac{i\pi}{6}}} \rangle \simeq D_{12}$,

(14) $\Gamma_{14} = \langle \tau_i, \tau_{\frac{1}{2}} \circ s_{\mathbb{R}} \rangle$,

(15) $\Gamma_{15} = \langle -\mathrm{id}, z \mapsto \frac{i}{2} + \overline{z} \rangle$,

(16) $\Gamma_{16} = \langle -\mathrm{id}, z \mapsto \frac{1}{2} + \overline{z}, \tau_{\frac{i}{2}} \rangle$,

(17) $\Gamma_{17} = \langle z \mapsto i\,z, z \mapsto \frac{i}{2} + \overline{z} \rangle$.

Démonstration. Nous renvoyons par exemple à [AB, Chap. XVII] et [Bos]. □

4.3 Exercices

Exercice E.44. Montrer qu'à isomorphisme près, tout groupe fini d'ordre 12 est isomorphe à

- $\mathbb{Z}/12\mathbb{Z} \simeq \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$,
- $\mathbb{Z}/6\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \simeq \mathbb{Z}/3\mathbb{Z} \times (\mathbb{Z}/2\mathbb{Z})^2$,
- $\mathbb{Z}/3\mathbb{Z} \rtimes \mathbb{Z}/4\mathbb{Z}$, où le générateur de $\mathbb{Z}/4\mathbb{Z}$ agit sur $\mathbb{Z}/3\mathbb{Z}$ par l'unique automorphisme non trivial de $\mathbb{Z}/3\mathbb{Z}$,
- $D_{12} \simeq \mathbb{Z}/3\mathbb{Z} \rtimes (\mathbb{Z}/2\mathbb{Z})^2$, où les générateurs de chaque facteur de $(\mathbb{Z}/2\mathbb{Z})^2$ agissent sur $\mathbb{Z}/3\mathbb{Z}$ par l'unique automorphisme non trivial de $\mathbb{Z}/3\mathbb{Z}$,
- $\mathfrak{A}_4 \simeq (\mathbb{Z}/2\mathbb{Z})^2 \rtimes \mathbb{Z}/3\mathbb{Z}$, où le générateur de $\mathbb{Z}/3\mathbb{Z}$ agit sur $(\mathbb{Z}/2\mathbb{Z})^2$ par $\begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}$.

À quel groupe sont isomorphes $D_6 \times \mathbb{Z}/2\mathbb{Z}$ et $D_6 \rtimes \mathbb{Z}/2\mathbb{Z}$ où l'élément non trivial de $\mathbb{Z}/2\mathbb{Z}$ agit sur D_6 par l'inverse ?

Exercice E.45. Montrer que le centre du groupe diédral D_{2m} vérifie $Z(D_2) = D_2$, $Z(D_4) = D_4$, $Z(D_{2m}) = \{\mathrm{id}\}$ si m est impair au moins 3 et

$$Z(D_{2m}) = \{\mathrm{id}, b\, a^{\frac{m}{2}}\} \simeq \mathbb{Z}/2\mathbb{Z}$$

si m est pair au moins 4, où a est un élément de D_{2m} d'ordre m, et où b est un élément de D_{2m} d'ordre 2 qui conjugue a à son inverse.

Notons $G = (\mathbb{Z}/m\mathbb{Z}) \rtimes (\mathbb{Z}/m\mathbb{Z})^\times$ où le groupe multiplicatif $(\mathbb{Z}/m\mathbb{Z})^\times$ des éléments inversibles de l'anneau $\mathbb{Z}/m\mathbb{Z}$ agit sur $\mathbb{Z}/m\mathbb{Z}$ par multiplication. Montrer que l'ordre de G est $m\,\varphi(m)$, où $\varphi : m \mapsto \mathrm{Card}\big((\mathbb{Z}/m\mathbb{Z})^\times\big)$ est la fonction d'Euler. Montrer que le groupe des automorphismes du groupe diédral D_{2m} vérifie $\mathrm{Aut}(D_2) = \{\mathrm{id}\}$,

$$\mathrm{Aut}(D_4) = \mathrm{GL}_2(\mathbb{Z}/2\mathbb{Z}) \simeq \mathfrak{S}_3 \simeq D_6$$

et si $m \geq 3$, alors il existe un isomorphisme de groupes $G \simeq \mathrm{Aut}(D_{2m})$ qui envoie l'élément $(k, n) \in G$ sur l'unique automorphisme de D_{2m} tel que $a \mapsto a^n$ et $b \mapsto b\, a^k$.

Montrer que le groupe $\mathrm{Out}(D_{2m}) = \mathrm{Aut}(D_{2m})/\mathrm{Int}(D_{2m})$ des automorphismes extérieurs de D_{2m} vérifie $\mathrm{Out}(D_2) = \{\mathrm{id}\}$, $\mathrm{Out}(D_4) = \mathrm{Aut}(D_4) \simeq D_6$ et si $m \geq 3$, alors $\mathrm{Out}(D_{2m}) \simeq (\mathbb{Z}/m\mathbb{Z})^\times$ si m est pair et $\mathrm{Out}(D_{2m}) \simeq (\mathbb{Z}/m\mathbb{Z})^\times/\{\pm 1\}$ si m est impair.

Exercice E.46. Déterminer le centre $Z(G)$, le groupe dérivé $[G, G]$, le groupe des automorphismes $\mathrm{Aut}(G)$, le groupe des automorphismes intérieurs $\mathrm{Int}(G)$ et le groupe des automorphismes extérieurs $\mathrm{Out}(G)$ pour $G = \mathfrak{A}_4, \mathfrak{S}_4, \mathfrak{A}_5$.

Montrer que $\mathfrak{A}_4$ et $\mathfrak{S}_4$ sont résolubles. [105]

Exercice E.47. Notons $(e_0, \ldots, e_n)$ la base canonique de l'espace euclidien standard $\mathbb{R}^{n+1}$. Notons Δ_n l'enveloppe convexe de $\{e_0, \ldots, e_n\}$, appelée le *simplexe euclidien standard*, munie de la distance induite par celle de $\mathbb{R}^{n+1}$. Montrer que le groupe des isométries de Δ_n est le sous-groupe du groupe orthogonal $\mathrm{O}(n+1)$ préservant globalement Δ_n, et qu'il est isomorphe au groupe symétrique $\mathfrak{S}_{n+1}$.

Exercice E.48. (1) Soit $n \in \mathbb{N} - \{0\}$. Montrer que tout sous-groupe fini de $\mathrm{GL}_n(\mathbb{R})$ (respectivement $\mathrm{SL}_n(\mathbb{R})$) est conjugué à un sous-groupe de $\mathrm{O}(n)$ (respectivement $\mathrm{SO}(n)$).

(2) Soient $p \neq q$ dans $\mathbb{N} - \{0\}$. En utilisant les notations du corollaire 3.12, montrer que tout sous-groupe fini de $\mathrm{O}(p, q)$ (respectivement $\mathrm{SO}(p, q)$, $\mathrm{SO}_0(p, q)$) est conjugué à un sous-groupe de $\mathrm{O}(p) \times \mathrm{O}(q)$ (respectivement $S(\mathrm{O}(p) \times \mathrm{O}(q))$, $\mathrm{SO}(p) \times \mathrm{SO}(q)$).

(3) Le groupe $\mathrm{SL}_2(\mathbb{C})$ agit sur $\mathbb{P}_1(\mathbb{C}) = \mathbb{C} \cup \{\infty\}$ par homographies

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \left(z \mapsto \frac{az + b}{cz + d}\right)$$

(avec les conventions usuelles $\infty \mapsto \frac{a}{c}$ et $\frac{w}{0} = \infty$ si $w \in \mathbb{C} - \{0\}$). Le noyau de cette action est $\{\pm \mathrm{id}\}$, donc nous identifions tout élément de $\mathrm{PSL}_2(\mathbb{C})$ avec l'homographie qu'elle définit.

105. Si G est un groupe, sa suite dérivée est la suite décroissante de sous-groupes distingués $(G_n)_{n\in\mathbb{N}}$ de G définie par récurrence $G_0 = G$ et $G_{n+1} = [G_n, G_n]$. Un groupe G est dit *groupe résoluble* s'il existe $n \in \mathbb{N}$ tel que $G_n = \{\mathrm{id}\}$.

Notons $p : \mathbb{S}_2 \to \mathbb{P}_1(\mathbb{C}) = \mathbb{C} \cup \{\infty\}$ la projection stéréographique de pôle nord. Montrer que l'application de $\mathrm{SO}(3)$ dans $\mathrm{PSL}_2(\mathbb{C})$ qui à $g \in \mathrm{SO}(3)$ associe l'application de $\mathbb{P}_1(\mathbb{C})$ dans lui-même définie par $z \mapsto p \circ g \circ p^{-1}$ est un isomorphisme de groupes sur le sous-groupe K de $\mathrm{PSL}_2(\mathbb{C})$ défini par

$$K = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{PSL}_2(\mathbb{C}) \ : \ ad - bc = 1, \ .... \right\}.$$

(4) Montrer que les groupes $\mathrm{SO}_0(1,3)$ et $\mathrm{PSL}_2(\mathbb{C})$ sont isomorphes. Pour cela, considérer l'espace vectoriel réel $\mathrm{Herm}_2 = \{X \in \mathscr{M}_2(\mathbb{C}) \ : \ X^* = X\}$ formé des matrices hermitiennes de $\mathscr{M}_2(\mathbb{C})$, muni de la base $\mathscr{B} = (\sigma_0, \sigma_1, \sigma_2, \sigma_3)$ où

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

i) Montrer que l'application $-\det$ est une forme quadratique de signature $(1,3)$ sur Herm_2, dont la matrice dans la base $\mathscr{B}$ est la matrice $I_{1,3} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

ii) Montrer que l'application $\Phi' : \mathrm{SL}_2(\mathbb{C}) \to \mathrm{GL}(\mathrm{Herm}_2)$ définie par

$$A \mapsto \{\Phi'(A) : X \mapsto AXA^*\}$$

est un morphisme de groupes, de noyau $\{\pm \mathrm{id}\}$.

iii) Montrer que l'application $\Phi : \mathrm{SL}_2(\mathbb{C}) \to \mathrm{GL}_4(\mathbb{R})$ qui à $A \in \mathrm{SL}_2(\mathbb{C})$ associe la matrice dans la base $\mathscr{B}$ de $\Phi'(A)$, est un morphisme de groupes continu dont l'image est contenue dans $\mathrm{SO}_0(1,3)$.

iv) Montrer que toute rotation dans $\mathbb{R}^3$ est un produit de rotations autour des trois axes de coordonnées. Montrer en fait que toute rotation dans $\mathbb{R}^3$ est un produit de rotations autour de deux des trois axes de coordonnées.

v) Montrer que

$$\text{si } \ A = \begin{pmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{pmatrix} \quad \text{alors} \quad \Phi(A) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{31}$$

et

$$\text{si } \ A' = \begin{pmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \quad \text{alors} \quad \Phi(A') = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta & -\sin\theta \\ 0 & 0 & \sin\theta & \cos\theta \end{pmatrix}. \tag{32}$$

En déduire que l'image de Φ contient $\mathbf{K} = \mathrm{SO}(1) \times \mathrm{SO}(3)$, vu comme le sous-groupe diagonal par blocs $(1,3)$ de $\mathrm{SO}_0(1,3)$.

vi) Posons

$$\mathbf{A} = \left\{ \begin{pmatrix} \cosh\psi & \sinh\psi & 0 & 0 \\ \sinh\psi & \cosh\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \ : \ \psi \in \mathbb{R} \right\},$$

appelé le *sous-groupe de Cartan* de $\mathrm{SO}_0(1,3)$. Montrer que

$$\text{si } A'' = \begin{pmatrix} \cosh\frac{\psi}{2} & \sinh\frac{\psi}{2} \\ \sinh\frac{\psi}{2} & \cosh\frac{\psi}{2} \end{pmatrix} \quad \text{alors} \quad \Phi(A'') = \begin{pmatrix} \cosh\psi & \sinh\psi & 0 & 0 \\ \sinh\psi & \cosh\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{33}$$

En déduire que l'image de Φ contient $\mathbf{A}$.

vii) Montrer que le coefficient $(1,1)$ de tout élément de $\mathrm{SO}_0(1,3)$ est supérieur ou égal à 1. En écrivant les matrices par blocs 1-3 dans la décomposition $\mathbb{R}^4 = \mathbb{R} \times \mathbb{R}^3$, montrer que tout élément h de $\mathrm{SO}_0(1,3) \cap \mathrm{Sym}_4^{++}$ est conjugué par un élément k de $\mathbf{K}$ à un élément a de $\mathbf{A}$ (au sens que $h = kak^{-1}$). En utilisant la décomposition polaire (voir le corollaire 3.12) de $\mathrm{SO}_0(1,3)$, qui dit que l'application de $(\mathrm{SO}_0(1,3) \cap \mathrm{Sym}_4^{++}) \times \mathbf{K}$ dans $\mathrm{SO}_0(1,3)$ définie par $(h,k') \mapsto h\,k'$ est un homéomorphisme, montrer la *décomposition de Cartan*

$$\mathrm{SO}_0(1,3) = \mathbf{KAK} ,$$

c'est-à-dire que tout $g \in \mathrm{SO}_0(1,3)$, il existe des éléments $k, k'' \in \mathbf{K}$ et $a \in \mathbf{A}$ tels que $g = kak''$.

viii) En déduire le résultat.

(5) En déduire une classification à conjugaison près des sous-groupes finis de $\mathrm{O}(n)$, $\mathrm{GL}_3(\mathbb{R})$, $\mathrm{O}(1,3)$, $\mathrm{SO}(1,3)$, $\mathrm{SO}_0(1,3)$, $\mathrm{GL}_2(\mathbb{C})$, $\mathrm{SL}_2(\mathbb{C})$, $\mathrm{PGL}_2(\mathbb{C})$, $\mathrm{PSL}_2(\mathbb{C})$.

4.4 Indications pour la résolution des exercices

Correction de l'exercice E.44. Voir par exemple [Ale].

Correction de l'exercice E.45. Voir

https ://www.youtube.com/watch ?v=c-VlZK2GLJg

Correction de l'exercice E.46. Voir par exemple [Per2, AB].

Notons que $[\mathfrak{S}_4, \mathfrak{S}_4] = \mathfrak{A}_4$, que $[\mathfrak{A}_4, \mathfrak{A}_4]$ est le 2-Sylow de $\mathfrak{A}_4$,qui est le sous-groupe abélien N de $\mathfrak{A}_4$ (isomorphe à $(\mathbb{Z}/2\mathbb{Z})^2$) constitué, outre de l'identité, des doubles transpositions, et que $[N, N] = \{\mathrm{id}\}$ (car le groupe dérivé d'un groupe abélien est trivial). Donc $\mathfrak{A}_4$ et $\mathfrak{S}_4$ sont résolubles.

Correction de l'exercice E.47. Voir par exemple l'exercice 4.43 de [FGN].

Correction de l'exercice E.48.

(4) i) Un calcul immédiat donne

$$-\det\Big(\sum_{i=0}^{n} x_i\,\sigma_i\Big) = -(x_0)^2 + (x_1)^2 + (x_2)^2 + (x_3)^2\ . \tag{34}$$

ii) Notons que $\Phi'(A)$ envoie bien Herm_2 dans Herm_2, car

$$(AXA^*)^* = AX^*A^* = AXA^*$$

si $X \in \mathrm{Herm}_2$, et que $\Phi'(A)$ est linéaire et inversible d'inverse $\Phi'(A^{-1})$. C'est un morphisme de groupes, car pour tous les $A, B \in \mathrm{SL}_2(\mathbb{C})$ et $X \in \mathrm{Herm}_2$, nous avons

$$(AB)X(AB)^* = A(BXB^*)A^*\ ,$$

donc $\Phi'(AB) = \Phi'(A) \circ \Phi'(B)$.

Soit $A \in \ker\Phi$. Alors pour tout $X \in \mathrm{Herm}_2$, nous avons

$$AXA^* = X\ .$$

En prenant en particulier pour X la matrice identité, nous en déduisons que $AA^* = \mathrm{id}$, donc que A est unitaire, d'inverse égal à A^*. D'où $AX = XA$ pour tout $X \in \mathrm{Herm}_2$. En prenant pour X la matrice σ_3, qui est diagonale à valeurs propres distinctes, nous en déduisons donc que A doit être diagonale. En prenant $X = \sigma_1$, nous en déduisons que les coefficients diagonaux de A sont égaux. Comme $\det A = 1$, ceci implique que $A = \pm\,\mathrm{id}$. Donc $\ker\Phi' \subset \{\pm\,\mathrm{id}\}$ et l'inclusion réciproque étant immédiate,

$$\ker\Phi' = \{\pm\,\mathrm{id}\}\ .$$

iii) Puisque l'application de $\mathrm{GL}(\mathrm{Herm}_2)$ dans $\mathrm{GL}_4(\mathbb{R})$ qui à un élément de $\mathrm{GL}(\mathrm{Herm}_2)$ associe sa matrice dans la base $\mathscr{B}$ est un isomorphisme de groupes, l'application Φ est un morphisme de groupes. Il est continu, car polynomial en les coefficients. Puisque

$$\det(AXA^*) = |\det(A)|^2 \det X = \det X$$

pour tous les $A \in \mathrm{SL}_2(\mathbb{C})$ et $X \in \mathrm{Herm}_2$, l'application linéaire $\Phi'(A)$ préserve l'application $-\det$, et donc par la formule (34), l'image de Φ est contenue dans $\mathrm{O}(1,3)$.

Par la continuité de Φ et la connexité de $\mathrm{SL}_2(\mathbb{C})$, cette image est en fait contenue dans la composante connexe de l'élément neutre $\mathrm{SO}_0(1,3)$ de $\mathrm{O}(1,3)$.

iv) Si (e_1, e_2, e_3) est la base canonique de $\mathbb{R}^3$, et si $R \in \mathrm{SO}(3)$, soit D l'axe de rotation de R, soit R_1 une rotation d'axe $\mathbb{R}e_1$ telle que la droite vectorielle $R_1 D$ soit contenue dans le plan vectoriel $\mathbb{R}e_1 + \mathbb{R}e_3$, et soit R_2 une rotation d'axe $\mathbb{R}e_2$ telle que $R_2 R_1 D = \mathbb{R}e_3$. Alors $R_3 = R_2 R_1 R (R_2 R_1)^{-1}$ est une rotation d'axe $\mathbb{R}e_3$, donc $R = (R_1)^{-1}(R_2)^{-1} R_3 R_2 R_1$ est un produit de rotations autour des axes de coordonnées.

Notons R'_3 la rotation d'axe $\mathbb{R}e_3$ et d'angle π. Si R'_2 est une rotation d'axe $\mathbb{R}e_2$, alors $(R'_3)^{-1} \circ R'_2 \circ R'_3$ est une rotation d'axe $\mathbb{R}e_1$. Donc R'_2 est un produit de rotations d'axes $\mathbb{R}e_1$ ou $\mathbb{R}e_3$. En permutant la base, le résultat en découle.

v) L'application $X \mapsto AXA^*$ envoie les matrices $\sigma_0, \sigma_1, \sigma_2, \sigma_3$ sur respectivement σ_0, $\cos\theta\ \sigma_1 + \sin\theta\ \sigma_2$, $-\sin\theta\ \sigma_1 + \cos\theta\ \sigma_2$ et σ_3. D'où la formule (31).

L'application $X \mapsto A'XA'^*$ envoie les matrices $\sigma_0, \sigma_1, \sigma_2, \sigma_3$ sur respectivement σ_0, σ_1 $\cos\theta\ \sigma_2 + \sin\theta\ \sigma_3$ et $-\sin\theta\ \sigma_2 + \cos\theta\ \sigma_3$. D'où la formule (32).

Le résultat découle alors de la question iv).

vi) Un petit calcul montre que l'application linéaire $\Phi'(A) : X \mapsto AXA^*$ envoie $\sigma_0, \sigma_1, \sigma_2, \sigma_3$ sur respectivement

$$\cosh\psi\ \sigma_0 + \sinh\psi\ \sigma_1,\ \ \sinh\psi\ \sigma_0 + \cosh\psi\ \sigma_1,\ \ \sigma_2,\ \ \sigma_3\ .$$

Ceci montre la formule (33) et donc **A** est contenu dans l'image de Φ.

vii) Pour tout $g \in \mathrm{O}(1,3)$ de première colonne $\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$, la relation ${}^t g\, I_{1,3}\, g = I_{1,3}$ donne $-a^2 + b^2 + c^2 + d^2 = -1$. Donc $a^2 \geq 1$. D'où $a \geq 1$ ou $a \leq 1$. Par connexité, nous avons donc que le coefficient $(1,1)$ de tout élément de $\mathrm{SO}_0(1,3)$ est au moins 1.

Nous identifierons un vecteur de $\mathbb{R}^3$ avec son vecteur colonne. Soit donc $h = \begin{pmatrix} a & {}^t v \\ u & B \end{pmatrix}$, avec $a \in \mathbb{R}$, $u, v \in \mathbb{R}^3$ et $B \in \mathscr{M}_3(\mathbb{R})$, un élément de $\mathrm{SO}_0(1,3) \cap \mathrm{Sym}_4^{++}$. Par la symétrie de la matrice h, nous avons donc $v = u$ et $B \in \mathrm{Sym}_3$.

La condition d'appartenance de h à $\mathrm{O}(1,3)$, qui est ${}^t h\, I_{1,3}\, h = I_{1,3}$, s'écrit donc

$$\begin{pmatrix} a & {}^t u \\ u & B \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & I_3 \end{pmatrix} \begin{pmatrix} a & {}^t u \\ u & B \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & I_3 \end{pmatrix}\ .$$

Cette égalité est équivalente au système d'égalités

$$-a^2 + {}^t u\, u = -1, \quad -a\, u + B\, u = 0, \quad -u\, {}^t u + B^2 = I_3\ .$$

Puisque $h \in \mathrm{SO}_0(1,3)$, nous avons $a \geq 1$. La matrice symétrique $I_3 + u\ {}^t u$ est définie positive, donc admet une unique racine carrée définie positive $\sqrt{I_3 + u\, {}^t u}$ (voir la proposition 3.7 (3)). Puisque h est définie positive, B doit être définie positive. Donc le système ci-dessus équivaut à

$$a = \sqrt{1 + {}^t u\, u}, \quad B\, u = a\, u, \quad B = \sqrt{I_3 + u\, {}^t u}\ . \tag{35}$$

Pour tout $C' \in \mathrm{SO}(3)$, nous avons

$$\begin{pmatrix} 1 & 0 \\ 0 & C' \end{pmatrix} \begin{pmatrix} a & {}^t u \\ u & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & C' \end{pmatrix}^{-1} = \begin{pmatrix} a & {}^t u \\ C'u & C'\,B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & {}^t C' \end{pmatrix} = \begin{pmatrix} a & {}^t(C'u) \\ C'u & C'B\,{}^t C' \end{pmatrix} .$$

Puisque $\mathrm{SO}(3)$ agit transitivement sur les droites vectorielles de $\mathbb{R}^3$ (voir la proposition 2.1), il existe $C \in \mathrm{SO}(3)$ tel que Cu appartienne à la première droite de coordonnée, donc soit de la forme $Cu = (b, 0, 0)$. La matrice $CB\,{}^t C$ est alors de la forme

$$CB\,{}^t C = \sqrt{I_3 + Cu\,{}^t u\,{}^t C} = \sqrt{I_3 + Cu\,{}^t(Cu)} = \sqrt{I_3 + \begin{pmatrix} b \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} b & 0 & 0 \end{pmatrix}}$$

$$= \begin{pmatrix} \sqrt{1+b^2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

En posant $d = \sqrt{1+b^2}$, nous avons donc montré qu'il existe un élément $k = \begin{pmatrix} 1 & 0 \\ 0 & C \end{pmatrix}$ de $\mathbf{K}$ tel que

$$k\,h\,k^{-1} = \begin{pmatrix} a & b & 0 & 0 \\ b & d & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} .$$

En notant A cette matrice, la relation ${}^t A\, I_{1,3}\, A = I_{1,3}$, s'écrit

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} .$$

D'où $-a^2 + c^2 = -1$, $-ab + cd = 0$ et $-b^2 + d^2 = 1$. En multipliant par d la seconde équation, et en utilisant la troisième et $ad - bc = 1$, nous avons $-(1+bc)b + c(1+b^2) = 0$, donc $b = c$ et $a^2 - b^2 = 1$. Puisque $a \geq 0$, cette dernière équation implique qu'il existe $\psi \in \mathbb{R}$ tel que $a = \cosh\psi$ et $b = \sinh\psi$. Puisque $ad - bc = 1$, nous avons $ad = 1 + b^2 = a^2$, donc $a = d$ puisque $a \neq 0$. Donc $k\,h\,k^{-1}$ appartient au sous-groupe de Cartan $\mathbf{A}$.

Par la décomposition polaire, tout élément $g \in \mathrm{SO}_0(1,3)$ s'écrit sous la forme hk' avec $h \in \mathrm{SO}_0(1,3) \cap \mathrm{Sym}_4^{++}$ et $k' \in \mathbf{K}$. Par la première assertion, il existe $k \in \mathbf{K}$ et $a \in \mathbf{A}$ tels que $h = kak^{-1}$. Donc $g = ka(k^{-1}k')$, ce qui montre la décomposition de Cartan.

viii) Puisque Φ est un morphisme de groupes dont l'image contient des sous-groupes de $\mathrm{SO}_0(1,3)$ qui engendrent $\mathrm{SO}_0(1,3)$, l'image de Φ est égale à $\mathrm{SO}_0(1,3)$. Comme le noyau de Φ, qui coïncide avec le noyau de Φ', est égal à $\{\pm \mathrm{id}\}$, le résultat en découle.

Index

Références

[AB] J.-M. Arnaudiès et J. Bertin. *Groupes, algèbres et géométrie.* Ellipses, 1998.

[Ale] M. Alessandri. *Thèmes de géométrie. Groupes en situation géométrique.* Dunod, 1999.

[Arn] J.-M. Arnaudiès. *Les cinq polyèdres réguliers de $\mathbb{R}^3$ et leurs groupes.* Centre de documentation universitaire (Editions CEDES), 1969.

[Art] E. Artin. *Algèbre géométrique.* (Geometric algebra, Interscience Pub. 1957, Dover 2015), Ed. Gabay 1996.

[Aud] M. Audin. *Géométrie.* EDP Sciences, 2006.

[Ave] A. Avez. *Calcul différentiel.* Masson, 1983.

[BBN$^+$] H. Brown, R. Bülow, J. Neubüser, H. Wondratschek, et H. Zassenhaus. *Crystallographic groups of four-dimensional space.* Wiley Mono. Crystal. Wiley 1978.

[Ber1] M. Berger. *Géométrie. Vol. 2 : espaces euclidiens, triangles, cercles et sphères.* Cedic/Fernand Nathan, 2nd éd., 1979.

[Ber2] M. Berger. *Géométrie. Vol. 3 : convexes et polytopes, polyèdres réguliers, aires et volumes.* Cedic/Fernand Nathan, 2nd éd., 1979.

[Ber4] M. Berger. *Géométrie. Vol. 4 : formes quadratiques, coniques et quadriques.* Cedic/Fernand Nathan, 2nd éd., 1979. Cedic/Fernand Nathan, 2nd éd., 1979.

[Bor] A. Borel. *Introduction aux groupes arithmétiques.* Hermann, 1969.

[Bos] Y. Bossard. *Rosaces, frises et pavages, Vol 1 et Vol 2.* Cedic 1977 et 1979.

[Bou] N. Bourbaki. *Groupes et algèbres de Lie : chap. 4, 5 et 6.* Masson, 1981.

[BT] F. Bruhat et J. Tits. *Schémas en groupes et immeubles des groupes classiques sur un corps local. II : groupes unitaires.* Bull. Soc. Math. France **115** (1987) 141–195.

[Bus] P. Buser. *A geometric proof of Bieberbach theorems on crystallographic groups.* L'Ens. Math. **31** (1985) 137–145.

[Car] H. Cartan. *Cours de calcul différentiel.* Hermann, 2nde éd. 1977.

[Cha] M. Chaperon. *Calcul différentiel et calcul intégral 3e année : Cours et exercices avec solutions.* Dunod, 2008.

[CS] J. Conway et N. Sloane. *Sphere Packings, Lattices and Groups.* Grund. math. Wiss **290**, Springer Verlag, 1988.

[Deh] R. Deheuvels. *Formes quadratiques et groupes classiques.* PUF, 1981.

[Die] J. Dieudonné. *La géométrie des groupes classiques.* Erg. Math. Grenz., Springer Verlag, 1971.

[Djo] D. Djokovic. *On the exponential map in classical Lie groups.* J. Algebra **64** (1980) 76–88.

[dlH] P. de la Harpe. *An invitation to Coxeter groups.* In "Group theory from a geometrical viewpoint" (E. Ghys, A. Haefliger, A. Verjovsky eds.), 193–253, World Scientific, 1991.

[FGN] S. Francinou, H. Gianella, et S. Nicolas. *Exercices de Mathématiques. Algèbre Tome 3.* Cassini, 2008.

[GL] E. Ghys et J. Leys. *Un système triple orthogonal.* Images des Mathématiques, CNRS, 2008, https ://images.math.cnrs.fr/Un-systeme-triple-orthogonal.html.

[GL] G. H. Golub and C. F. Van Loan. *Matrix computations.* Fourth edition. Johns Hopkins Univ. Press, 2013.

[Gou] X. Gourdon. *Algèbre et probabilités.* Ellipses, 3ème éd, 2021.

[Hum] J. E. Humphreys. *Reflection groups and Coxeter groups.* Cambridge Univ. Press, 1990.

[Laf] J. Lafontaine. *Introduction aux variétés différentielles.* Press. Univ. Grenoble, 1996.

[Mar] J. Martinet. *Perfect lattices in Euclidean spaces.* Grund. math. Wissen. **327**, Springer Verlag 2003.

[Mil] J. Milnor. *Hilbert's problem 18 : On crystallographic groups, fundamental domains, and on sphere packing.* In "Mathematical Developments Arising from Hilbert Problems" (Proc. Sympos. Pure Math., Northern Illinois Univ., 1974), pp. 491–506. Proc. Sympos. Pure Math., Vol. XXVIII, Amer. Math. Soc. 1976.

[MT] N. Mneimné et F. Testard. *Introduction à la théorie des groupes de Lie classiques.* Hermann, 1986.

[Pau1] F. Paulin. *De la géométrie et de la dynamique de* $\mathrm{SL}_n(\mathbb{R})$ *et* $\mathrm{SL}_n(\mathbb{R})$. Dans "Sur la dynamique des groupes de matrices et applications arithmétiques" (avec Gilles Courtois et Françoise Dal'Bo), Journées X-UPS 2007, N. Berline, A. Plagne, C. Sabbagh eds., Editions Ecole Polytechnique, 2007.

[Pau2] F. Paulin. *Géométrie différentielle élémentaire.* Notes de cours de première année de master, Ecole Normale Supérieure, 2007, voir https://www.imo.universite-paris-saclay.fr/∼paulin/notescours/liste_notescours.html.

[Pau3] F. Paulin. *Groupes et géométries.* Notes de cours de seconde année de master, Université Paris-Sud, 2014, voir http://www.math.u-psud.fr/∼paulin/notescours/cours_georiem.pdf.

[Pau4] F. Paulin. *Introduction topologique à la géométrie.* Notes de cours de première année de master, Université Paris-Saclay, 2021, voir https://www.imo.universite-paris-saclay.fr/∼paulin/notescours/cours_GeometrieM1Orsay.pdf.

[Per1] D. Perrin. *Cours d'algèbre.* Ellipses, 1996.

[Per2] D. Perrin. *Géométrie algébrique : une introduction.* Intereditions/CNRS editions, 1995.

[Ser] J.-P. Serre. *Cours d'arithmétique.* Press. Univ. France, Paris, 1970.

[ST] G. Shephard et J. Todd. *Finite unitary reflection groups.* Canad. J. Math. **6** (1954) 274–304.

[Zav] M. Zavidovique. *Un Max de maths.* Calvage et Mounet, 2013.

Frédéric Paulin
Laboratoire de mathématique d'Orsay, UMR 8628 CNRS
Université Paris-Saclay, 91405 ORSAY Cedex, FRANCE
courriel : frederic.paulin@universite-paris-saclay.fr