Géométries et groupes classiques

T. Delcroix

Table des matières

Chapitre 1

Actions de groupes

1.1 Actions et représentations

Définition 1.1.1. Un *groupe* est un ensemble G muni d'une loi de composition interne *associative*

$$G \times G \to G, (g, h) \to gh$$

satisfaisant les propriétés suivantes.

1. Il existe un élément neutre e : $\forall g \in G, ge = eg = g$
2. tout élément admet un inverse : $\forall g \in G, \exists g^{-1} \in G, gg^{-1} = g^{-1}g = e$.

Exemple 1.1.2.

1. Le groupe $\mathfrak{S}_n$ des permutations de l'ensemble $\{1, \ldots, n\}$.
2. Les groupes sous-jacents à un corps $\mathbb{K}$: le groupe additif $(\mathbb{K}, +)$ et le groupe multiplicatif $(\mathbb{K}^*, \times)$.
3. Le groupe additif sous-jacent à un espace vectoriel, par exemple $(\mathbb{R}^n, +)$.

(À partir de maintenant, G désigne un groupe, la loi interne $(g, h) \mapsto gh$ est sous-entendue, et e désigne l'élément neutre.)

Définition 1.1.3. Une *action* du groupe G sur un ensemble E est une application

$$G \times E \to E, (g, x) \mapsto g \cdot x$$

satisfaisant les deux propriétés suivantes.

1. $\forall x \in E$, $e \cdot x = x$;
2. $\forall x \in E$, $\forall g, h \in G$, $g \cdot (h \cdot x) = (gh) \cdot x$.

Exemple 1.1.4.

1. L'action naturelle de $\mathfrak{S}_n$ sur $\{1, \ldots, n\}$.
2. Plus généralement, si E est un ensemble quelconque, l'ensemble $\mathrm{Bij}(E)$ des bijections de E dans lui-même est un groupe pour la loi de composition, et ce groupe agit sur E.
3. Le groupe multiplicatif $(\mathbb{R}^*, \times)$ agit sur $\mathbb{R}^n$ par homothéties :

$$\mathbb{R}^* \times \mathbb{R}^n \to \mathbb{R}^n, (t, x) \mapsto tx.$$

4. Plus généralement, si V est un espace vectoriel sur le corps $\mathbb{K}$, l'existence d'une action du groupe multiplicatif $(\mathbb{K}^*, \times)$ sur V par homothéties fait partie des axiomes définissant les espaces vectoriels.

5. Le groupe $(\mathbb{R}^n, +)$ agit sur $\mathbb{R}^n$ par translations :

$$\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n, (x, y) \mapsto x + y.$$

Plus généralement :

Proposition 1.1.5. *Pour tout groupe G, l'application*

$$G \times G \to G, (g, h) \mapsto gh$$

définit une action du groupe G sur l'ensemble G, appelée action par translations.

Démonstration. On vérifie les deux axiomes des actions de groupes :

1. $\forall h \in G$, $e \cdot h = e$ par définition de l'élément neutre.
2. $\forall g_1, g_2, h \in G$, $g_1(g_2 h) = (g_1 g_2)h$ par associativité de la loi de groupe.

□

Définition 1.1.6. Soient G et H deux groupes. Une application $\phi : G \to H$ est un morphisme de groupes si pour tous g_1 et g_2 dans G,

$$\phi(g_1 g_2) = \phi(g_1)\phi(g_2)$$

où le produit dans le terme de gauche est la loi de groupe de G, et celui dans le terme de droite est la loi de groupe de H.

Exemple 1.1.7. L'application signature $\varepsilon : \mathfrak{S}_n \to \{\pm 1\}$ est un morphisme de groupes. Rappelons deux définitions possibles de la signature :

1. si $\sigma \in \mathfrak{S}_n$,
$$\varepsilon(\sigma) = \prod_{1 \leq i < j \leq n} \text{signe}(\sigma(j) - \sigma(i))$$
2. si σ peut s'écrire comme un produit d'un nombre pair de transpositions, alors $\varepsilon(\sigma) = 1$, sinon $\varepsilon(\sigma) = -1$.

Définition 1.1.8. Si ϕ est un morphisme de groupe, on note son image par

$$\text{Im}(\phi) = \{\phi(g); g \in G\}$$

et on définit son noyau par

$$\ker(\phi) = \{g \in G; \phi(g) = e_H\}.$$

Exercice 1.1.9. Ce sont des sous-groupes.

Définition 1.1.10. Soit E un ensemble. On appelle représentation de G sur E (ou représentation de G dans $\text{Bij}(E)$) un morphisme de groupes de G dans $\text{Bij}(E)$.

Proposition 1.1.11. *On a la correspondance suivante entre actions et représentations :*

1. *Si G agit sur E, alors l'application*

$$G \to \text{Bij}(E), g \mapsto (x \mapsto g \cdot x)$$

définit une représentation de G sur E.

2. *Si $\phi : G \to \mathrm{Bij}(E)$ est une représentation de G sur E, alors l'application*

$$G \times E \to E, (g, x) \mapsto \phi(g)(x)$$

définit une action de G sur E.

Démonstration.

1. Pour $g \in G$, notons $\psi_g : E \to E, x \mapsto g \cdot x$. On a $\psi_{gh} = \psi_g \circ \psi_h$ puisque $\forall x$, $(gh) \cdot x = g \cdot (h \cdot x)$. Il reste à vérifier que ψ_g est bien dans $\mathrm{Bij}(E)$. Pour cela, on remarque que $\psi_g \circ \psi_{g^{-1}} = \psi_e = \mathrm{Id}_E$, donc chaque ψ_g est bien une bijection.
2. On vérifie les axiomes définissant une action de groupe :

$$\forall x \in E, \qquad e \cdot x = \phi(e)(x) = \mathrm{Id}_E(x) = x$$

(car un morphisme de groupe envoie élément neutre sur élément neutre), et $\forall x \in E$, $\forall g, h \in G$,

$$\begin{aligned} g \cdot (h \cdot x) &= \phi(g)\left(\phi(h)(x)\right) \\ &= \left(\phi(g) \circ \phi(h)\right)(x) \\ &= \phi_{gh}(x) \\ &= (gh) \cdot x. \end{aligned}$$

□

1.2 Description d'une action : orbites et stabilisateurs

Dans la suite, G désigne un groupe, E un ensemble, et on se donne une action de G sur E (donc aussi une représentation ϕ de G dans $\mathrm{Bij}(E)$).

Définition 1.2.1. Pour $x \in E$, on définit :

1. l'orbite de x
$$\mathrm{orb}(x) = \{y \in E; \exists g \in G, g \cdot x = y\} \subset E$$
2. le stabilisateur de x
$$\mathrm{Stab}(x) = \{g \in G; g \cdot x = x\} \subset G.$$

Exercice 1.2.2. La relation $x \sim y$ si $\mathrm{orb}(x) = \mathrm{orb}(y)$ définit une relation d'équivalence sur E.

Exercice 1.2.3. Pour tout $x \in E$, le stabilisateur $\mathrm{Stab}(x)$ est un sous-groupe de G.

Exemple 1.2.4.

— Considérons l'action naturelle de $\mathfrak{S}_3 = \{\mathrm{Id}, (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}$ sur l'ensemble $\{1, 2, 3\}$. Alors l'orbite de 1 est $\mathrm{orb}(1) = \{1, 2, 3\}$ et son stabilisateur est $\mathrm{Stab}(1) = \{\mathrm{Id}, (2\ 3)\}$.

— L'actions précédente induit une action naturelle de $\mathfrak{S}_3$ sur l'ensemble E des couples d'éléments de $\{1, 2, 3\}$. On a

$$E = \{(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3)\}$$

et, par exemple,

$$\mathrm{orb}((1,1)) = \{(1,1), (2,2), (3,3)\} \qquad \mathrm{Stab}((1,1)) = \{\mathrm{Id}, (2\ 3)\}$$

$$\mathrm{orb}((1,2)) = E \qquad \mathrm{Stab}((1,2)) = \{\mathrm{Id}\}$$

Définition 1.2.5.

1. L'action de G sur E est *transitive* s'il existe $x \in E$ tel que $\mathrm{orb}(x) = E$.
2. L'action de G sur E est *libre* si pour tout $x \in E$, $\mathrm{Stab}(x) = \{e\}$.
3. La représentation ϕ de G dans $\mathrm{Bij}(E)$ est *fidèle* si $\ker(\phi) = \{e\}$.

Exercice 1.2.6. Montrer que la représentation ϕ est fidèle si et seulement si l'intersection des stabilisateurs est triviale, c'est-à-dire

$$\bigcap_{x \in E} \mathrm{Stab}(x) = \{e\}.$$

Définition 1.2.7.

Si G est un groupe fini, on appelle *ordre* de G le cardinal de l'ensemble G.

Si $g \in G$, on appelle *ordre* de g le cardinal du sous-groupe engendré par g.

Exercice 1.2.8. Montrer que l'ordre d'un élément g est égal au plus petit entier $k \in \mathbb{N}^*$ tel que $g^k = e$.

Théorème 1.2.9. *Tout groupe d'ordre n admet une représentation fidèle dans $\mathfrak{S}_n$.*

Démonstration. Il suffit de considérer l'action de G par translation sur lui-même. Cette action est libre : pour tout $g \in G$,

$$\mathrm{Stab}(g) = \{h \in G; hg = g\} = \{e\}.$$

Une action libre est fidèle (par exemple en utilisant un exercice précédent, donc on a une représentation fidèle de G dans $\mathrm{Bij}(G)$. Enfin, $\#G = n$, donc $\mathrm{Bij}(G) \simeq \mathfrak{S}_n$. □

Exercice 1.2.10. Étudier les actions suivantes :

1. l'action triviale de G sur un ensemble quelconque, définie par $g \cdot x = x$ pour tout $g \in G$, $x \in E$;
2. l'action par de G par translation sur lui-même ;
3. l'action de $(\mathbb{R} \times \{0\}, +)$ par translation sur $\mathbb{R}^2$;
4. l'action de $(\mathbb{R}^*, \times)$ par homothéties sur $\mathbb{R}^2$.
5. l'action de $\mathfrak{S}_n$ sur les couples d'éléments de $\{1, \ldots, n\}$.

Exercice 1.2.11. On a vu un exemple d'action fidèle mais pas libre. Lequel (ou lesquels) ?

Un exemple fondamental d'action : l'action par conjugaison

Proposition 1.2.12. *L'application*

$$G \times G \to G, (g, h) \mapsto ghg^{-1}$$

définit une action de G sur G, appelée action par conjugaison.

Démonstration. Les vérifications sont immédiates, à écrire en exercice. □

Terminologie 1.2.13. Pour cette action, les orbites sont appelées les *classes de conjugaison* et les stabilisateurs sont appelés *centralisateurs*.

Remarque 1.2.14. En général, les orbites et stabilisateurs d'une action sont difficiles à déterminer. C'est particulièrement vrai pour l'action par conjugaison. On verra dans le cours plusieurs cas particuliers où on sait décrire les orbites ou des représentants de chaque orbite.

Un morphisme de groupe $\phi : G \to H$ est appelé un *automorphisme* de G si $H = G$ et il existe un morphisme $\phi^{-1} : G \to G$ tel que $\phi \circ \phi^{-1} = \phi^{-1} \circ \phi = \mathrm{Id}_G$. On note $\mathrm{Aut}(G)$ le groupe (pour la composition) des automorphismes de G.

Proposition 1.2.15. *La représentation $\phi : G \to \mathrm{Bij}(G)$ associée à l'action par conjugaison est à valeur dans les automorphismes de groupes :* $\mathrm{Im}(\phi) \subset \mathrm{Aut}(G)$.

Démonstration. Soit $g, h, k \in G$, on a

$$\begin{aligned}\phi(g)(hk) &= ghkg^{-1} \\ &= ghg^{-1}gkg^{-1} \\ &= \phi(g)(h)\phi(g)(k).\end{aligned}$$

□

Terminologie 1.2.16. Le noyau de ϕ est appelé le *centre* de G, et noté $Z(G)$. L'image de ϕ est appelé le groupe des *automorphismes intérieurs* de G, et noté $\mathrm{Int}(G)$.

1.2.0.1 Classes de conjugaison dans $\mathfrak{S}_n$

On note $(a_1\ a_2\ \cdots\ a_k)$ (avec $a_j \in \{1, \ldots, n\}$) le k-cycle qui envoie a_i sur a_{i+1} pour $1 \leq i < k$, qui envoie a_k sur a_1, et qui laisse fixe tous les éléments de $\{1, \ldots, n\} \setminus \{a_1, \ldots, a_k\}$.

Proposition 1.2.17. *Soient σ et τ deux éléments de $\mathfrak{S}_n$. Si τ est un k-cycle de la forme $(a_1\ a_2\ \cdots\ a_k)$, alors la permutation obtenue en conjuguant τ par σ est le k-cycle $\sigma\tau\sigma^{-1} = (\sigma(a_1)\ \sigma(a_2)\ \cdots\ \sigma(a_k))$.*

Démonstration. Si $x \in \{1, \ldots, n\} \backslash \{\sigma(a_1), \ldots, \sigma(a_k)\}$, alors on a $\sigma^{-1}(x) \in \{1, \ldots, n\} \backslash \{a_1, \ldots, a_k\}$, donc $\sigma\tau\sigma^{-1}(x) = \sigma\sigma^{-1}(x) = x$. Si $1 \leq i < k$, alors $\sigma\tau\sigma^{-1}(\sigma(a_i)) = \sigma\tau(a_i) = \sigma(a_{i+1})$. Enfin, puisqu'il ne reste plus qu'une possibilité, on a $\sigma\tau\sigma^{-1}(\sigma(a_k)) = \sigma(a_1)$. □

Rappelons que toute permutation de $\mathfrak{S}_n$ peut s'écrire comme un produit de cycles à supports disjoints. La proposition ci-dessus suffit donc à décrire la conjugaison de n'importe quelle permutation par une autre permutation. En travaillant un tout petit peu plus, on montre la proposition suivante.

Proposition 1.2.18. *Deux éléments de $\mathfrak{S}_n$ sont conjugués si et seulement si leurs écritures comme produit de cycles à supports disjoints compte le même nombre de k-cycles pour tout k.*

1.3 Espaces quotients, groupes quotients

Soit G un groupe, et H un sous-groupe de G.

Proposition 1.3.1. *La relation sur G définie par $x \sim y$ si $\exists h \in H$, $x = yh$ est une relation d'équivalence.*

Démonstration. Il faut vérifier chacune des propriétés des relations d'équivalence.

1. Elle est reflexive : en prenant $h = e$, on a $x \sim x$.
2. Elle est symétrique : si $x = yh$, alors $y = xh^{-1}$ (et $h^{-1} \in H$).
3. Elle est transitive : si $x = yh$ et $y = zh'$, alors $x = z(h'h)$ (et $h'h \in H$).

□

Terminologie 1.3.2. On appelle *espace quotient* (ou espace des classes à droite modulo H) et on note G/H l'ensemble des classes d'équivalence pour cette relation. On note la classe d'équivalence de g par gH ou $\bar{g}$.

Théorème 1.3.3 (Théorème de Lagrange)**.** *Soit G un groupe fini, H un sous-groupe de G. Alors*

$$\#G = (\#H)(\#G/H).$$

Démonstration. Il est clair que $G = \bigcup_{g\in G} gH$ car tout g est contenu dans gH (pour le même g). On a $gH \cap g'H \neq \emptyset$ si et seulement si $g \sim g'$, et si l'intersection est non vide, alors $gH = g'H$ (tout ceci grâce à la proposition précédente). Choisissons pour chacune des classes d'équivalence un représentant. En les ordonnant, on obtient m éléments g_1, g_2, ..., g_m, où $m = \#(G/H)$. Le groupe G est donc couvert par la réunion disjointe

$$G = \bigsqcup_{i=1}^{m} g_i H,$$

et chaque ensemble g_iH est de cardinal $\#H$. On en conclut l'égalité du théorème. □

Remarque 1.3.4. Plus généralement, les classes d'équivalence d'une relation d'équivalence sur un ensemble E forment une partition de E. Dans le cas particulier ici, toutes les classes d'équivalence ont même cardinal, ce qui permet d'obtenir la relation.

Corollaire 1.3.5. *L'ordre d'un élément de G divise l'ordre de G.*

Corollaire 1.3.6 (Formule des classes)**.** *Si un groupe fini G agit sur un ensemble E, et $x \in E$, alors*

$$\#G = (\#\operatorname{orb}(x))(\#\operatorname{Stab}(x)).$$

Démonstration. (à détailler chez vous) On a une bijection entre $\operatorname{orb}(x)$ et l'espace quotient $G/\operatorname{Stab}(x)$. □

Définition 1.3.7. Un sous-groupe H de G est *distingué* si $\forall g \in G$, $gH = Hg$. On note $H \triangleleft G$.

Ici, on utilise la notation $gH = \bar{g}$ la classe d'équivalence pour la relation précédemment introduite, et $Hg = \{hg, h \in H\}$, qui forme aussi une classe d'équivalence, pour la variante de relation d'équivalence que vous devinerez facilement.

Définition 1.3.8. Un groupe G est *simple* si ses seuls sous-groupes distingués sont $\{e\}$ et G.

Proposition 1.3.9. *Soit $H \triangleleft G$ un sous-groupe distingué de G. Alors l'application*

$$G/H \times G/H \to G/H, (\bar{g}_1, \bar{g}_2) \mapsto \overline{g_1 g_2}$$

définit une loi de groupe sur G/H, qui est alors appelé groupe quotient.

Démonstration. Il faut d'abord vérifier que c'est bien défini : $\overline{g_1 g_2}$ ne doit pas dépendre du choix de $g_1 \in \bar{g}_1$ et $g_2 \in \bar{g}_2$. Si $g_1' = g_1 h_1$ et $g_2' = g_2 h_2$, on a $g_1' g_2' = g_1 h_1 g_2 h_2$. Or $Hg_2 = g_2 H$ (car H est distingué), donc $\exists h_3 \in H$ tel que $h_1 g_2 = g_2 h_3$. Donc $g_1' g_2' = g_1 g_2 h_3 h_2$ définit la même classe que $g_1 g_2$.

Le reste des vérifications est élémentaire. On a par exemple pour élément neutre $\bar{e}$ et l'inverse de $\bar{g}$ est $\overline{g^{-1}}$. □

Exercice 1.3.10. Le noyau d'un morphisme de groupes est un sous-groupe distingué.

Exercice 1.3.11. Associer à toute représentation ϕ de G sur E une représentation ψ *fidèle* du groupe quotient $G/\ker(\phi)$ sur E, telle que, si $\pi : G \to G/\ker\phi$ est l'application quotient, on a $\phi = \psi \circ \pi$.

1.4 Produits semi-directs

1.4.1 La loi définissant un produit semi-direct

Soit N et H deux groupes. Soit $\theta : H \to \mathrm{Aut}(N)$ un morphisme de groupes. On considère la loi interne sur l'ensemble $N \times H$ définie par

$$(n, h)(n', h') = (n\theta(h)(n'), hh'). \tag{1.1}$$

On va montrer que cette loi définit une structure de groupe sur l'*ensemble* $N \times H$.

1.4.2 La loi est associative (TD1, exo12, question 1)

Soient n_1, n_2 et n_3 des éléments de N, soient h_1, h_2 et h_3 des éléments de H.

$$\begin{aligned}
(n_1, h_1)\left((n_2, h_2)(n_3, h_3)\right) &= (n_1, h_1)(n_2\theta(h_2)(n_3), h_2h_3) \\
&= (n_1\theta(h_1)(n_2\theta(h_2)(n_3)), h_1h_2h_3) \\
&= (n_1\theta(h_1)(n_2)\theta(h_1)(\theta(h_2)(n_3)), h_1h_2h_3) \\
&= (n_1\theta(h_1)(n_2)\theta(h_1h_2)(n_3), h_1h_2h_3) \\
&= (n_1\theta(h_1)(n_2), h_1h_2)(n_3, h_3) \\
&= \left((n_1, h_1)(n_2, h_2)\right)(n_3, h_3)
\end{aligned}$$

On a montré par le calcul ci-dessus que la loi définie par (1.1) est associative.

1.4.3 Inverse et élément neutre (TD1, exo12, questions 2-3)

Vérifier que l'élément (e_N, e_H) fournit un élément neutre pour la loi (1.1) est essentiellement immédiat. Pour vérifier que chaque élément admet un inverse, on commence par deviner celui-ci. D'après l'équation définissant la loi interne, si (n', h') est l'inverse de (n, h), alors $hh' = e_H$ et $n\theta(h)(n') = e_N$, donc $h' = h^{-1}$ et $n' = (\theta(h))^{-1}(n^{-1})$. Ce choix fonctionne en effet : on a aussi

$$(\theta(h))^{-1}(n^{-1}), h^{-1})(n, h) = (\theta(h))^{-1}(n^{-1})\theta(h^{-1})(n), h^{-1}h) = (e_N, e_H)$$

car $\theta(h^{-1}) : N \to N$ est un morphisme de groupes.

L'ensemble $N \times H$ munit de la loi interne (1.1) forme donc bien un groupe.

Terminologie 1.4.1. Le groupe obtenu de cette manière est noté $N \rtimes_\theta H$, et on l'appelle produit semi-direct de N et H (relativement à θ).

1.4.4 Produit semi-direct et produit direct (TD1, exo12, question 4)

Le choix du morphisme trivial $\theta : H \to \mathrm{Aut}(N), h \mapsto \mathrm{Id}$ fournit la structure de produit direct sur le produit $N \times H$. En effet, dans ce cas la formule de la loi interne est simplement

$$(n, h)(n', h') = (nn', hh').$$

Réciproquement, si θ est non trivial, soit $h \in H$ tel que $\theta(h) \neq \mathrm{Id}$. Puisque $\theta(h) \neq \mathrm{Id}$, il existe $n' \in N$ tel que $\theta(h)(n') \neq n'$. On a donc

$$(e_N, h)(n', e_H) = (\theta(h)(n'), h) \neq (n', h)$$

et donc $N \rtimes_\theta H$ n'est pas le groupe produit.

1.4.5 Quotient (TD1, exo12, questions 5-6)

Notons $N' = N \times \{e_H\} \subset N \rtimes_\theta H$. Montrons que c'est un sous-groupe distingué : il suffit de montrer que pour tout $(n,h) \in N \rtimes_\theta H$, et $n' \in N$, il existe $n'' \in N$ tel que

$$(n,h)(n',e_H) = (n'',e_H)(n,h).$$

En utilisant la définition du produit, on aurait donc

$$(n\theta(h)(n'),h) = (n''n,h)$$

L'unique possibilité est de poser $n'' = n\theta(h)(n')n^{-1}$, qui est bien un élément de N. On vérifie par les mêmes calculs que ce choix fonctionne toujours.

Enfin, montrons que $N \rtimes_\theta H/N' \simeq H$. Considérons l'application $N \rtimes_\theta H \to H, (n,h) \to h$. Étant donné la définition de la loi de groupe, cette application est un morphisme de groupes. Son noyau est égal à N' par définition, donc le sous-groupe image est isomorphe à $N \rtimes_\theta H/N'$ (il s'agit d'un résultat de base de théorie des groupes, à revoir dans vos cours des années précédentes si besoin : pour tout morphisme de groupe $\phi : G \to H$, les groupes $\mathrm{Im}(\phi)$ et $G/\ker(\phi)$ sont isomorphes). Le morphisme est évidemment surjectif, donc H est isomorphe à $N \rtimes_\theta H/N'$.

1.5 Un exemple de géométrie discrète : les automorphismes de graphes

Définition 1.5.1. On appelle **graphe** un couple (S,A) formé d'un ensemble S et d'un sous-ensemble $A \subset (S \times S \setminus \{(s,s), s \in S\})$.

Terminologie 1.5.2. Si (S,A) est un graphe, on appelle *sommets* les éléments de S et *arêtes orientées* les éléments de A. Si $(o,e) \in A$, on dit que o est l'*origine* de l'arête orientée et que e est son *extrémité*.

On s'intéresse aux graphes finis, c'est-à-dire que S est un ensemble fini, et on suppose pour fixer les notations que $S = \{1,\dots,n\}$. Notons $\mathcal{G}$ l'ensemble des graphes sur S. Le groupe $\mathfrak{S}_n$ agit sur S par son action naturelle, et ceci induit une action de $\mathfrak{S}_n$ sur $\mathcal{G}$: si $\sigma \in \mathfrak{S}_n$, $\sigma \cdot (S,A) = (S,A')$ où

$$A' = \{(\sigma(o),\sigma(e)) \mid (o,e) \in A\}.$$

Terminologie 1.5.3. On note $\mathrm{Aut}(S,A) \subset \mathfrak{S}_n$ le stabilisateur de (S,A) pour l'action définie ci-dessus, et on appelle ses éléments les **automorphismes** du graphe (S,A).

Remarque 1.5.4. On a une action naturelle de $\mathrm{Aut}(S,A)$ sur les sommets S, mais aussi une action induite sur les arêtes orientées : l'action $\sigma \cdot (o,e) = (\sigma(o),\sigma(e))$ est bien définie sur A pour $\sigma \in \mathrm{Aut}(S,A)$.

Exemple 1.5.5. Pour $S = \{1,2,3\}$, considérons le graphe donné par l'ensemble d'arêtes orientées $A = \{(1,2),(2,3),(3,1)\}$. On a

$$\mathrm{Aut}(S,A) = \{\mathrm{Id}, (1\ 2\ 3), (1\ 3\ 2)\}$$

donc par la formule des classes, l'orbite de (S,A) est de cardinal 2. Plus précisément,

$$\mathrm{orb}(S,A) = \{(S,A), (S,\{(2,1),(3,2),(1,3)\})\}.$$

Exercice 1.5.6. Donner une représentation graphique des deux graphes impliqués dans l'exemple précédent. Montrer que l'action induite par $\mathrm{Aut}(S,A)$ sur les sommets est transitive. Montrer que l'action induite par $\mathrm{Aut}(S,A)$ sur les arêtes orientées est transitive.

Chapitre 2

Matrices équivalentes, matrices semblables

Le but de ce chapitre est d'étudier les actions :

I. de $\mathrm{GL}_n(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ sur $\mathrm{M}_n(\mathbb{K})$ par $(g,h) \cdot A = gAh^{-1}$,

II. de $\mathrm{GL}_n(\mathbb{K})$ sur $\mathrm{M}_n(\mathbb{K})$ par $g \cdot A = gAg^{-1}$.

Le but est comme toujours de comprendre les orbites et les stabilisateurs. Ces actions commencent à être trop compliquées pour que l'on décrive chaque orbite et chaque stabilisateurs en détails. L'objectif principal sera ici de déterminer un représentant de chaque orbite. Pour la première action on travaillera sur un corps quelconque $\mathbb{K}$. Pour l'action par conjugaison, la description d'un représentant de chaque orbite ne sera donnée que sur $\mathbb{C}$ et $\mathbb{R}$.

2.1 Matrices équivalentes

2.1.1 Rappels sur les matrices

Soit $\mathbb{K}$ un corps. On note $\mathrm{M}_n(\mathbb{K})$ l'ensemble des matrices $n \times n$ à coefficients dans $\mathbb{K}$. On notera régulièrement $A = (a_{i,j})_{1 \leq i,j \leq n}$ et on omettra l'information $1 \leq i,j \leq n$ lorsqu'elle est évidente par le contexte. Le nombre $a_{i,j} \in \mathbb{K}$ est appelé le coefficient de A à la i-ième ligne et la j-ième colonne.

Cet ensemble est en fait muni d'une structure très riche.

L'ensemble $\mathrm{M}_n(\mathbb{K})$, muni de l'addition $A + B = (a_{i,j} + b_{i,j})_{1 \leq i,j \leq n}$ et de la multiplication par les élément de $\mathbb{K}$: $\lambda A = (\lambda a_{i,j})$, forme un espace vectoriel de dimension n^2. Il est muni d'une base standard, définie par les *matrices élémentaires* $E_{k,l} \in \mathrm{M}_n(\mathbb{K})$. Ces dernières sont les matrices définies par

$$E_{k,l} = (\delta_{i,k}\delta_{j,l})_{1 \leq i,j \leq n}$$

c'est-à-dire que ses coefficients sont tous nuls excepté un 1 à la k-ième ligne et l-ième colonne.

On pensera aussi à $\mathrm{M}_n(\mathbb{K})$ comme à l'espace vectoriel formé des applications linéaires de $\mathbb{K}^n$ dans $\mathbb{K}^n$. Plus précisément, à une matrice A correspond l'application linéaire :

$$A : \underline{x} = (x_1, \ldots, x_n) \in \mathbb{K}^n \longmapsto A\underline{x} = \left(\sum_{j=1}^{n} a_{1,j}x_j, \sum_{j=1}^{n} a_{2,j}x_j, \ldots, \sum_{j=1}^{n} a_{n,j}x_j \right) \in \mathbb{K}^n.$$

On peut ainsi définir l'image et le noyau de A :

$$\mathrm{Im}(A) = A(\mathbb{K}^n) \qquad \ker(A) = \{\underline{x} \in \mathbb{K}^n \mid A\underline{x} = 0\}.$$

Rappelons le théorème du rang, qui relie les dimensions de ces deux espaces. On note $\mathrm{rg}(A)$ la dimension de l'image de A.

Théorème 2.1.1. *Pour toute matrice $A \in \mathrm{M}_n(\mathbb{K})$, on a*

$$\dim \ker(A) + \mathrm{rg}(A) = n.$$

Enfin, $\mathrm{M}_n(\mathbb{K})$ est muni d'une multiplication

$$AB = \left(\sum_{k=1}^{n} a_{i,k} b_{k,j} \right)_{1 \leq i,j \leq n}$$

d'élément neutre $I_n = (\delta_{i,j})_{1 \leq i,j \leq n} = \sum_{i=1}^{n} E_{i,i}$. Cette multiplication fait de $\mathrm{M}_n(\mathbb{K})$ un anneau. Rappelons la définition de $\mathrm{GL}_n(\mathbb{K})$.

Définition 2.1.2. Le *groupe général linéaire* $\mathrm{GL}_n(\mathbb{K})$ est le groupe des éléments inversibles de l'anneau $\mathrm{M}_n(\mathbb{K})$. Il s'agit donc des éléments $g \in \mathrm{M}_n(\mathbb{K})$ tels qu'il existe un élément $g^{-1} \in \mathrm{M}_n(\mathbb{K})$ avec $gg^{-1} = g^{-1}g = I_n$. Il s'agit également des éléments tels que l'application linéaire $\mathbb{K}^n \to \mathbb{K}^n$ associée soit un isomorphisme.

Pour terminer cette section de rappels, rappelons la définition de deux applications fondamentales.

Définition 2.1.3.

1. L'application *trace* est définie par

$$\mathrm{tr} : \mathrm{M}_n(\mathbb{K}) \to \mathbb{K}, A \mapsto \sum_{i=1}^{n} a_{i,i}.$$

2. L'application *déterminant* est définie par

$$\det : \mathrm{M}_n(\mathbb{K}) \to \mathbb{K}, A \mapsto \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i}.$$

Proposition 2.1.4.

1. *On a $\mathrm{GL}_n(\mathbb{K}) = \det^{-1}(\mathbb{K}^*)$.*
2. *Le déterminant, restreint à $\mathrm{GL}_n(\mathbb{K})$, est un morphisme de groupe.*
3. *Pour deux matrices quelconques A et B de $\mathrm{M}_n(\mathbb{K})$, on a $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.*

Démonstration. En exercice (ou revoir les cours des années précédentes). □

2.1.2 Les orbites de l'action de $\mathrm{GL}_n \times \mathrm{GL}_n$ sur M_n

Théorème 2.1.5. *Le groupe $\mathrm{GL}_n(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ agit sur $\mathrm{M}_n(\mathbb{K})$ avec exactement $n+1$ orbites. Ces orbites sont les $\mathrm{rg}^{-1}(r) = \{A \in \mathrm{M}_n(\mathbb{K}) \mid \mathrm{rg}(A) = r\}$ pour $r \in \{0, 1, \ldots, n\}$.*

Définition 2.1.6. On dit que deux matrices A et B de $\mathrm{M}_n(\mathbb{K})$ sont *équivalentes* si elles sont dans la même orbite, c'est-à-dire s'il existe $(g, h) \in \mathrm{GL}_n(\mathbb{K})$ tel que $B = gAh^{-1}$.

On en déduit une autre manière d'énoncer le théorème.

Théorème 2.1.7. *Deux matrices sont équivalents si et seulement si elles ont même rang.*

En réalité, l'énoncé ci-dessus n'est pas formellement équivalent au précédent : on ne précise pas qu'il y a exactement $n+1$ orbites. Ceci est facilement corrigé en exhibant une matrice de rang r pour tout $0 \leq r \leq n$. Pour les utiliser dans la preuve, on introduit les matrices suivantes :

$$I_{r,n} := \sum_{i=1}^{r}$$

il s'agit de matrices diagonales ayant des coefficients diagonaux égaux à 1 sur les r premières lignes, et égaux à 0 sur les autres lignes.

On montre le théorème en prouvant les deux directions.

Lemme 2.1.8. *Si A et B ont même rang, alors elles sont équivalentes.*

On prouvera ce lemme dans la section suivante.

Lemme 2.1.9. *Si A et B sont équivalentes, alors elles ont même rang.*

Démonstration. On écrit $B = gAh^{-1}$ avec $g, h \in \mathrm{GL}_n(\mathbb{K})$, ou de manière équivalente, $Bh = gA$. Il est facile de vérifier que B et Bh ont même image : h est un isomorphisme de $\mathbb{K}^n$ dans $\mathbb{K}^n$, donc $h(\mathbb{K}^n) = \mathbb{K}^n$, et $Bh(\mathbb{K}^n) = B(\mathbb{K}^n)$. D'autre part, puisque g est un isomorphisme, gA et A ont même noyau. Par le Théorème du rang, $\mathrm{rg}(A) = n - \dim\ker(A)$ et $\mathrm{rg}(B) = n - \dim\ker(B)$, on a donc

$$\begin{aligned}
\mathrm{rg}(A) &= n - \dim\ker(A) \\
&= n - \dim\ker(gA) \\
&= n - \dim\ker(Bh) \\
&= \mathrm{rg}(Bh) \\
&= \mathrm{rg}(B).
\end{aligned}$$

□

2.1.3 Pivot de Gauss

Pour démontrer que deux matrices de même rang sont équivalentes, on applique l'algorithme du pivot de Gauss pour montrer que toute matrice de rang r est équivalente à $I_{r,n}$.

Initialement, le pivot de Gauss consiste en des opérations sur les lignes et les colonnes :

1. échange des lignes i et j : $L_i \leftrightarrow L_j$,
2. échange des colonnes i et j : $C_i \leftrightarrow C_j$,
3. dilatation de la ligne i par un $\lambda \in \mathbb{K}^*$: $L_i \leftarrow \lambda L_i$,
4. dilatation de la colonne i par un $\lambda \in \mathbb{K}^*$: $C_i \leftarrow \lambda C_i$,
5. transvection sur les lignes : $L_i \leftarrow L_i + \lambda L_j$,
6. transvection sur les colonnes : $C_i \leftarrow C_i + \lambda C_j$.

Par ces opérations, on a une procédure qui transforme une matrice A en une matrice de type $I_{r,n}$.

Étape 0 (cas d'arrêt) si $A = 0_n$ la matrice nulle, alors on s'arrête.

Étape 1 si $a_{1,1} = 0$, on échange des lignes et/ou des colonnes pour avoir $a_{1,1} \neq 0$.

Étape 2 si $a_{1,1} \neq 1$, on utilise une dilatation sur la ligne 1 pour avoir $a_{1,1} = 1$.

Étape 3 par une suite de transvections sur les lignes et les colonnes, on s'assure $a_{1,j} = 0$ pour $j \neq 1$ et $a_{i,1} = 0$ pour $i \neq 1$.

Étape 4 on peut maintenant écrire

$$A = \begin{pmatrix} 1 & 0_{1\times n-1} \\ 0_{n-1\times 1} & A' \end{pmatrix}$$

on reprend à l'étape 0 avec A' à la place de A.

Exercice 2.1.10. Décrire les transvections utilisées dans l'étape 3 en détails.

Le lien avec l'action de $\mathrm{GL}_n(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ est obtenu de la manière suivante.

Définition 2.1.11.

1. Soit $\sigma \in \mathfrak{S}_n$. La *matrice de permutation* associée à σ est la matrice

$$P_\sigma := \sum_{i=1}^{n} E_{\sigma(i),i}.$$

   Dans le cas de la transposition (k, l), on note aussi $P_{k,l}$ la transposition associée.
2. Les *matrices de dilatation* sont les matrices $D_i(\lambda) = I_n + (\lambda - 1)E_{i,i}$ pour $1 \leq i \leq n$ et $\lambda \in \mathbb{K}^*$.
3. Les *matrices de transvections* sont les matrices $T_{i,j}(\lambda) = I_n + \lambda E_{i,j}$ pour $1 \leq i \neq j \leq n$.

Exercice 2.1.12. Dessiner des schémas décrivant chacune de ces matrices.

Proposition 2.1.13. *La multiplication d'une matrice A par une matrice de transposition, de dilatation, de transvection,* à gauche *réalise l'opération correspondante sur* les lignes *de la matrice A. La multiplication d'une matrice A par une matrice de transposition, de dilatation, de transvection,* à droite *réalise l'opération correspondante sur* les colonnes *de la matrice A.*

Démonstration. S'en convaincre. □

Ainsi l'algorithme décrit plus tôt fournit une matrice g et une matrice k telles que $gAk = I_{r,n}$, où g et k sont des produits de matrices de dilatation, transposition, transvection. Pour conclure, il nous reste à remarquer que ces matrices sont toutes inversibles (exercice), donc que $g \in \mathrm{GL}_n(\mathbb{K})$ et $h = k^{-1} \in \mathrm{GL}_n(\mathbb{K})$ sont telles que $gAh^{-1} = I_{r,n}$. Enfin, $r = \mathrm{rg}(A)$ par le premier lemme démontré.

2.1.4 Exemples de stabilisateurs

Exercice 2.1.14. Déterminer les stabilisateurs de $I_{r,2}$ $(0 \leq r \leq 2)$ pour l'action de $\mathrm{GL}_2(\mathbb{K}) \times \mathrm{GL}_2(\mathbb{K})$ sur $\mathrm{M}_2(\mathbb{K})$.

Réponses :

$$\begin{aligned} \mathrm{Stab}(0_2) &= \mathrm{GL}_2(\mathbb{K}) \times \mathrm{GL}_2(\mathbb{K}) \\ \mathrm{Stab}(I_{1,2}) &= \left\{ \left(\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}, \begin{pmatrix} a & 0 \\ c & e \end{pmatrix} \right) \mid a, d, e \in \mathbb{K}^*, b, c \in \mathbb{K} \right\} \\ \mathrm{Stab}(I_2) &= \mathrm{diag}(\mathrm{GL}_2(\mathbb{K})) = \{(g, g) \mid g \in \mathrm{GL}_2(\mathbb{K})\} \end{aligned}$$

2.2 Matrices semblables

2.2.1 Forme de Jordan sur $\mathbb{R}$ ou $\mathbb{C}$

Soit $\mathbb{K}$ un corps, quelconque pour l'instant.

Définition 2.2.1. Deux matrices A et B dans $\mathrm{M}_n(\mathbb{K})$ sont *semblables (dans* $\mathrm{M}_n(\mathbb{K})$*)* s'il existe $g \in \mathrm{GL}_n(\mathbb{K})$ tel que

$$B = gAg^{-1}.$$

Définition 2.2.2. Soit $A \in \mathrm{M}_n(\mathbb{K})$. L'ensemble des matrices semblables à A dans $\mathrm{M}_n(\mathbb{K})$ est appelé *la classe de similitude de A.*

On appelle ainsi *classes de similitudes* les orbites de l'action de $\mathrm{GL}_n(\mathbb{K})$ sur $\mathrm{M}_n(\mathbb{K})$ par conjugaison $g \cdot A = gAg^{-1}$.

Définition 2.2.3. Un *bloc de Jordan* de taille $k \in \mathbb{N}^*$ et de valeur propre $\lambda \in \mathbb{K}$ est la matrice

$$J_{k,\lambda} = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix} = \lambda I_k + \sum_{i=1}^{k-1} E_{i,i+1} \in \mathrm{M}_k(\mathbb{K})$$

Attention pour l'énoncé suivant on se place dans le cas $\mathbb{K} = \mathbb{C}$.

Théorème 2.2.4 (Décomposition de Jordan complexe)**.** *Toute matrice $A \in \mathrm{M}_n(\mathbb{C})$ est semblable dans $\mathrm{M}_n(\mathbb{C})$ à une matrice B diagonale par blocs, dont les blocs sont des blocs de Jordan :*

$$\exists g \in \mathrm{GL}_n(\mathbb{C}), \qquad B = gAg^{-1} = \begin{pmatrix} J_{k_1,\lambda_1} & 0 & \cdots & 0 \\ 0 & J_{k_2,\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & J_{k_m,\lambda_m}. \end{pmatrix}$$

De plus, la matrice B de cette forme est unique à permutation des blocs diagonaux près, et on l'appelle la forme de Jordan *de A.*

Notation 2.2.5. Étant donné un angle $\theta \in]0, \pi[$, on note

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

la *matrice de rotation* associée.

Si on se donne de plus un entier $k \in \mathbb{N}^*$ et un réel non-nul $\lambda \in \mathbb{R}^*$, on note $K_{k,\lambda,\theta}$ la matrice définie par blocs par

$$K_{k,\lambda,\theta} = \begin{pmatrix} \lambda R_\theta & I_2 & 0_2 & \cdots & 0_2 \\ 0_2 & \lambda R_\theta & I_2 & \cdots & 0_2 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0_2 & 0_2 & 0_2 & \cdots & \lambda R_\theta \end{pmatrix} \in \mathrm{M}_{2k}(\mathbb{R}).$$

Théorème 2.2.6 (Décomposition de Jordan réelle)**.** *Toute matrice $A \in \mathrm{M}_n(\mathbb{R})$ est semblable dans $\mathrm{M}_n(\mathbb{R})$ à une matrice B diagonale par blocs, dont chaque bloc est soit de la forme $K_{k,\lambda,\theta}$ pour $k \in \mathbb{N}^*$, $\lambda \in \mathbb{R}^*$, $\theta \in]0, \pi[$, soit de la forme $J_{k,\lambda}$ pour $k \in \mathbb{N}^*$ et $\lambda \in \mathbb{R}$. On appelle B la* forme de Jordan réelle *de A, et elle est unique à permutation des blocs diagonaux près.*

Exemple 2.2.7. Si $\lambda \in \mathbb{C}$, la matrice diagonale λI_n est sa propre forme de Jordan. En fait, l'orbite de λI_n par conjugaison est restreinte à $\{\lambda I_n\}$.

Exercice 2.2.8. La matrice R_θ est sa propre forme de Jordan réelle. Quelle est sa forme de Jordan complexe ?

2.2.2 Invariants de conjugaison

Rappelons la définition des polynômes caractéristiques et minimaux d'une matrice, impliqués par exemple dans la diagonalisation des matrices.

Remarque 2.2.9. Une matrice est diagonalisable si et seulement si elle appartient à la classe de similitude d'une matrice diagonale.

Définition 2.2.10. Soit $A \in \mathrm{M}_n(\mathbb{K})$. On appelle *polynôme caractéristique* de A et on note χ_A le polynôme en la variable X défini par

$$\chi_A(X) = \det(XI_n - A).$$

Définition 2.2.11. Soit $A \in \mathrm{M}_n(\mathbb{K})$. On appelle *polynôme minimal* de A et on note P_A le polynôme en la variable X, de plus petit degré et de coefficient directeur égal à un, tel que, si on remplace X par A dans l'expression d, on ait

$$P_A(A) = 0.$$

Proposition 2.2.12. *Si A et B sont semblables, alors $\chi_A = \chi_B$.*

Démonstration. Écrivons $B = gAg^{-1}$, pour un $g \in \mathrm{GL}_n(\mathbb{K})$. On a

$$\begin{aligned}
\chi_B(X) &= \det(XI_n - B) \\
&= \det\left(Xgg^{-1} - gAg^{-1}\right) \\
&= \det\left(g(XI_n - A)g^{-1}\right) \\
&= \det(g)\det(XI_n - A)\det(g)^{-1} && \text{car det est un morphisme} \\
&= \det(XI_n - A) && \text{car } \mathbb{K} \text{ est commutatif} \\
&= \chi_A(X),
\end{aligned}$$

ce qu'il fallait démontrer. □

Corollaire 2.2.13. *Si A et B sont semblables, alors* $\det(A) = \det(B)$ *et* $\mathrm{tr}(A) = \mathrm{tr}(B)$.

Démonstration. Le déterminant est (au signe près), le coefficient constant du polynôme caractéristique. La trace est l'opposé de son coefficient sous-directeur (le coefficient de X^{n-1}). □

Proposition 2.2.14. *Si A et B sont semblables, alors $P_A = P_B$.*

Démonstration. Il suffit de montrer que pour tout polynôme Q, $Q(A) = 0$ si et seulement si $Q(B) = 0$. Par symétrie des rôles de A et B, il suffit de montrer que pour tout polynôme Q,

$$Q(A) = 0 \Longrightarrow Q(B) = 0.$$

Supposons $Q(A) = 0$, notons $Q(X) = a_0 + a_1X + \cdots + a_kX^k$, et $B = gAg^{-1}$ pour un $g \in \mathrm{GL}_n(\mathbb{K})$. Alors

$$\begin{aligned} Q(B) &= a_0 + a_1 gAg^{-1} + a_2 gAg^{-1}gAg^{-1} + \cdots + a_k \left(gAg^{-1}\right)^k \\ &= g\left(a_0 + a_1A + \cdots + a_kA^k\right)g^{-1} \\ &= gQ(A)g^{-1} \\ &= 0 \end{aligned}$$

□

Proposition 2.2.15. *Si A et B sont semblables, $\lambda \in \mathbb{K}$ et $k \in \mathbb{N}$, alors*

$$\dim \ker\left((A - \lambda I_n)^k\right) = \dim \ker\left((A - \lambda I_n)^k\right).$$

Démonstration. Notons $B = gAg^{-1}$. On a aussi $(B - \lambda I_n)^k = g(A - \lambda I_n)^k g^{-1}$ par la simplification habituelle $g^{-1}g = I_n$. Alors on voit en particulier que les matrices $(B - \lambda I_n)^k$ et $(A - \lambda I_n)^k$ sont équivalentes. Elles ont par conséquent même rang. La conclusion du théorème suit par application du Théorème du rang. □

Proposition 2.2.16 (Application pour déterminer les blocs de Jordan (sur $\mathbb{C}$))**.** *Soit λ une valeur propre de $A \in \mathrm{M}_n(\mathbb{C})$, alors*

1. *la* somme des tailles des blocs *de Jordan de valeur propre λ dans la forme de Jordan de A est égale à la multiplicité de λ comme racine du polynôme caractéristique χ_A ;*
2. *la* taille du plus grand bloc *de valeur propre λ est égale à la multiplicité de λ comme racine du polynôme minimal P_A ;*
3. *le* nombre de blocs *de Jordan de valeur propre λ est égal à la dimension du sous-espace propre de A pour la valeur propre λ.*

Démonstration. En exercice : déterminer ces quantités sur la forme de Jordan de A et utiliser les propositions précédentes. □

Remarque 2.2.17. En raffinant un peu le troisième point ci-dessus, on peut complètement déterminer la forme de Jordan en appliquant plusieurs algorithmes de type pivot de Gauss...

2.2.3 Exemples d'orbites et stabilisateurs

Exercice 2.2.18. Dans $\mathrm{M}_2(\mathbb{C})$, *décrire* les orbites de l'action de $\mathrm{GL}_2(\mathbb{C})$ par conjugaison, et les stabilisateurs (à conjugaison près).

Réponses : Il y a deux types d'orbites :
- les orbites formées des matrices diagonalisables de valeurs propres (λ_1, λ_2), pour chaque choix de $(\lambda_1, \lambda_2) \in \mathbb{C}^2$;
- les orbites formées des matrices non-diagonalisables de valeur propre double λ, pour chaque choix de $\lambda \in \mathbb{C}$.

Pour les stabilisateurs à conjugaison près, il suffit de déterminer les stabilisateurs pour les formes de Jordan.
- Si $A = \mathrm{diag}(\lambda_1, \lambda_2)$ avec $\lambda_1 \neq \lambda_2$, on a $\mathrm{Stab}(A) = \{\text{matrices diagonalisables inversibles}\}$.
- Si $A = \mathrm{diag}(\lambda, \lambda)$, on a $\mathrm{Stab}(A) = \mathrm{GL}_2(\mathbb{C})$.
- Si $A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$, on a $\mathrm{Stab}(A) = \left\{ \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} \mid a \in \mathbb{C}^*, b \in \mathbb{C} \right\}$.

Chapitre 3

Groupes classiques

3.1 Notion de groupe topologique, exemple de GL_n

3.1.1 Espaces topologiques

Si E est un ensemble, on note $P(E)$ l'ensemble des parties (on dit aussi sous-ensembles) de E.

Définition 3.1.1. Un *espace topologique* est un ensemble E muni d'un ensemble $\mathcal{T} \subset P(E)$ de parties de E tel que :

1. $\{\emptyset, E\} \subset \mathcal{T}$
2. $\forall I \subset \mathcal{T}$, $\bigcup_{U \in I} U \subset E$ est un élément de $\mathcal{T}$
3. $\forall n \in \mathbb{N}$, $\forall U_1, \ldots, U_n \in \mathcal{T}$, $U_1 \cap \cdots \cap U_n \subset E$ est un élément de $\mathcal{T}$.

On omettra souvent de préciser l'ensemble $\mathcal{T}$ en parlant d'un espace topologique $(E, \mathcal{T})$, et ceci en particulier quand la topologie est évidente par le contexte.

Terminologie 3.1.2. Les éléments de $\mathcal{T}$ sont appelés les *ouvert* de l'espace topologique E, la première propriété signifie que l'ensemble vide et E sont des ouverts de E, la deuxième propriété s'énonce *une réunion quelconque d'ouverts et ouverte* et la troisième propriété s'énonce *une intersection finie d'ouverts est ouverte.*

Le complémentaire d'un ouvert est appelé un fermé de E. On a les propriétés équivalentes : une intersection quelconque de fermés est fermée et une réunion finie de fermés est fermée.

Exemple 3.1.3. Sur n'importe quel ensemble on peut mettre au moins deux topologies :

1. la topologie *grossière* $\mathcal{T} = \{\emptyset, E, \}$,
2. la topologie *discrète* $\mathcal{T} = P(E)$.

Exemple 3.1.4. Si on se donne $\mathcal{S} \subset P(E)$ qui ne satisfait pas les hypothèses de la définition, on peut définir la *topologie engendrée* par $\mathcal{S}$: il s'agit de l'ensemble des réunions quelconques d'intersections finies d'éléments de $\mathcal{S}$, et cela fournit une structure d'espace topologique sur E.

Exemple 3.1.5. La topologie usuelle sur $\mathbb{R}$ est la topologie engendrée par les intervalles ouverts.

Rappelons deux propositions élémentaires qui permettent de construire des espaces topologiques à partir d'autres espaces topologiques.

Proposition 3.1.6. *Si $(E, \mathcal{T})$ est un espace topologique, et $F \subset E$ est une partie de E, alors $\{U \cap F \mid U \in \mathcal{T}\}$ fournit une structure d'espace topologique sur F, appelée la* topologie induite.

Proposition 3.1.7. *Si $(E_1, \mathcal{T}_1)$ et $(E_2, \mathcal{T}_2)$ sont deux espaces topologiques, alors on appelle* topologie produit *sur $E_1 \times E_2$ la topologie engendrée par les produits $U_1 \times U_2$ où $U_1 \in \mathcal{T}_1$ et $U_2 \in \mathcal{T}_2$.*

Exemple 3.1.8.
- Comme ensemble, $\mathbb{C} \simeq \mathbb{R}^2$, et la topologie usuelle de $\mathbb{C}$ est la topologie produit issue de cet isomorphisme.
- La topologie induite sur $\mathbb{R}^2 \simeq \mathbb{R}^2 \times \{0\} \subset (\mathbb{R}^3$, topol. produit) est la même que la topologie produit.

Si on ne le précise pas, E désigne un espace topologique dans la suite, et $\mathcal{T}$ désigne sa topologie.

Définition 3.1.9. Soit D une partie de E. L'*adhérence* $\overline{D}$ de D est la plus petit fermé de E qui contient D.

Remarque 3.1.10. L'adhérence d'un fermé est ce fermé lui-même.

Définition 3.1.11. On dit que D est dense dans E si $\overline{D} = E$.

En particulier, une partie D est toujours dense dans son adhérence !

Exemple 3.1.12. L'intervalle ouvert $]0, 1[$ est ouvert dans l'intervalle fermé $[0, 1]$. Les rationnels $\mathbb{Q}$ sont denses dans les réels $\mathbb{R}$.

Définition 3.1.13. Un espace topologique E *n'est pas connexe* s'il existe deux ouvert U et V de E, non-vides, disjoints, tels que $E = U \sqcup V$. Il est dit *connexe* dans le cas contraire.

Exemple 3.1.14. Un intervalle est connexe dans $\mathbb{R}$. La réunion de deux points distincts dans $\mathbb{R}$ n'est pas connexe.

Définition 3.1.15.
1. Un espace topologique E est *séparé* si $\forall x_1, x_2 \in E$, $\exists U_1, U_2$ ouverts de E tels que $x_1 \in U_1$, $x_2 \in U_2$ et $U_1 \cap U_2 = \emptyset$.
2. Un espace topologique $(E, \mathcal{T})$ est compact si il séparé et si pour tout $I \subset \mathcal{T}$ tel que $\bigcup_{U \in I} U = E$, il existe un sous-ensemble fini $\{U_1, \ldots, U_n\}$ de I tel que $U_1 \cup \cdots U_n = E$.

Exemple 3.1.16. Une réunion finie de points, un intervalle fermé, sont compacts dans $\mathbb{R}$. Un intervalle ouvert n'est pas compact.

On va rappeler à la troisième section une caractérisation pratique de la compacité dans les espaces vectoriels normés.

3.1.2 Fonctions continues

Définition 3.1.17. Soient E_1 et E_2 deux espaces topologiques. Une application $f : E_1 \to E_2$ est *continue* si pour tout ouvert U de E_2, son image réciproque $f^{-1}(U)$ est un ouvert de E_1.

Remarque 3.1.18. Par définition, on a aussi la caractérisation alternative : l'application f est continue si pour tout fermé F de E_2, $f^{-1}(F)$ est fermé dans E_1.

Exemple 3.1.19. Supposons $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.
- Les fonctions monomiales $\mathbb{K}^n \to \mathbb{K}, (x_1, \ldots, x_n) \to x_1^{a_1} \cdots x_n^{a_n}$, pour $(a_1, \ldots, a_n) \in \mathbb{N}^n$, sont continues.
- Les fonctions polynomiales, combinaisaon linéaires de fonctions monomiales, sont également continues.

— L'inverse d'une fonction polynomiale est continue sur l'ouvert où la fonction ne s'annule pas, donc les fonctions rationnelles sont continues sur leurs ensembles de définition.
— la fonction exponentielle $\exp : \mathbb{K} \to \mathbb{K}$ est continue.

Proposition 3.1.20. *Un espace topologique E est connexe si et seulement si toute application continue $E \to (\{0,1\},$ topol. discrète$)$ est constante.*

Définition 3.1.21. Un espace topologique E est *connexe par arcs* si pour tout couple $(x, y) \in E^2$ de points de E, il existe une application continue $\gamma : [0,1] \to E$ telle que $\gamma(0) = x$ et $\gamma(1) = y$.

Terminologie 3.1.22. On appelle γ un chemin (continu) entre x et y.

Proposition 3.1.23. *Si E est connexe par arcs, alors E est connexe.*

Remarque 3.1.24. La réciproque est fausse : se rappeler le contre-exemple donné par l'adhérence du graphe de $x \mapsto \sin(1/x)$.

Proposition 3.1.25. *Soit $f : E_1 \to E_2$ une application continue entre deux espaces topologiques séparés. Alors l'image par f d'une partie compacte de E_1 est une partie compacte de E_2.*

Remarque 3.1.26. Attention, ici, on parle de l'image directe par f et as de l'image réciproque.

3.1.3 Espaces vectoriels normés

Pour cette section, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

Définition 3.1.27. Un $\mathbb{K}$-*espace vectoriel normé* est un $\mathbb{K}$-espace vectoriel E muni d'une norme, c'est-à-dire d'une application $\|\cdot\| : E \to \mathbb{R}_+$ telle que

1. pour $x \in E$, $\|x\| = 0$ si et seulement si $x = 0$,
2. pour $x \in E$ et $t \in \mathbb{K}$, $\|tx\| = |t|\|x\|$,
3. pour x, $y \in E$, $\|x + y\| \leq \|x\| + \|y\|$.

Dans cette section, E désigne un espace vectoriel normé.

Définition 3.1.28. On appelle *boule ouverte* dans E les ensembles $B(x, r) = \{y \in E \mid \|y - x\| < r\}$ pour $x \in E$ et $r \in \mathbb{R}_+^*$.

Un $\mathbb{K}$-espace vectoriel normé est muni d'une structure d'espace topologique associée à la norme : la topologie engendrée par les boules ouvertes.

Définition 3.1.29. Deux normes $\|\cdot\|_1$ et $\|\cdot\|_2$ sont *équivalentes* s'il existe deux réels a, $b \in \mathbb{R}_+^*$ tels que pour tout $x \in E$,

$$a\|x\|_1 \leq \|x\|_2 \leq b\|x\|_1.$$

Proposition 3.1.30. *Deux normes sont équivalentes si et seulement si elles définissent la même topologie.*

Théorème 3.1.31. *Sur un espace vectoriel de dimension finie E, toutes les normes sont équivalentes, et la topologie commune qu'elles définissent coïncide avec la topologie produit pour tout isomorphisme $E \simeq \mathbb{K}^n$.*

En vue de ce théorème, il parait stupide de s'intéresser aux normes. Cependant, la caractérisation de la compacité dans les espaces vectoriels normés justifie cette digression. Rappelons qu'une partie K de E est *bornée* si l'application norme est bornée sur K.

Proposition 3.1.32. *Une partie K de E est compacte si et seulement si elle est fermée et bornée.*

Notons qu'en dépit de l'équivalence des normes, différentes normes peuvent s'avérer utiles dans des contextes différents.

Exemple 3.1.33. Rappelons les normes les plus classiques sur l'espace vectoriel $\mathbb{R}^n$:
— $\|(x_1, \ldots, x_n)\|_1 = |x_1| + \cdots + |x_n|$
— $\|(x_1, \ldots, x_n)\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$
— $\|(x_1, \ldots, x_n)\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$

Exemple 3.1.34. Si $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ sont deux $\mathbb{K}$-espaces vectoriels normés de dimension finies, l'espace vectoriel $\mathcal{L}(E, F)$ des applications linéaires entre E et F est muni d'une *norme d'opérateurs*, définie par

$$\|f\| = \sup_{x \in E \setminus \{0\}} \frac{\|f(x)\|_F}{\|x\|_E} \qquad \text{pour } f \in \mathcal{L}(E, F)$$

Enfin, sans réécrire tout en détail, on a :

Proposition 3.1.35. *Un espace vectoriel normé est un espace métrique (pour la distance $d(x, y) = \|y - x\|$), ce qui permet d'utiliser des critères séquentiels pour caractériser les ensembles fermés, les adhérences, la continuité d'une fonction, etc.*

3.1.4 Notion de groupe topologique

Définition 3.1.36. Un *groupe topologique* est un groupe G tel que :

1. l'ensemble sous-jacent est muni d'une structure d'espace topologique,
2. le produit de groupe $G \times G \to G, (g, h) \mapsto gh$ est une application continue,
3. l'inverse $G \to G, g \mapsto g^{-1}$ est une application continue.

Remarque 3.1.37. Tout sous-groupe d'un groupe topologique est un groupe topologique pour la topologie induite.

Exemple 3.1.38.
— $(\mathbb{R}, +)$
— $(\mathbb{R}^n, +)$
— $(\mathbb{R}^*, \times)$
— $(\mathbb{C}, \times)$

Pour la culture, même si vous ne savez pas ce qu'est une variété, on a :

Définition 3.1.39. Un *groupe de Lie* est un groupe G tel que :

1. l'ensemble sous-jacent est muni d'une structure de variété différentiable,
2. le produit de groupe $G \times G \to G, (g, h) \mapsto gh$ est une application différentiable,
3. l'inverse $G \to G, g \mapsto g^{-1}$ est une application différentiable.

Un ouvert dans $\mathbb{R}^n$ est un exemple de variété différentiable. En particulier, les exemples précédents sont des groupes de Lie.

Définition 3.1.40. Un *groupe algébrique* est un groupe G tel que :

1. l'ensemble sous-jacent est muni d'une structure de variété algébrique,
2. le produit de groupe $G \times G \to G, (g, h) \mapsto gh$ est une application polynomiale,

3. l'inverse $G \to G, g \mapsto g^{-1}$ est une application polynomiale.

Le lieu des zéros d'un polynôme dans $\mathbb{R}^n$, ou son complémentaire sont des exemples de variétés algébriques. En particulier, les exemples précédents sont des groupes algébriques.

Un exemple fondamental : $\mathrm{GL}_n(\mathbb{K})$ pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$

On se rappelle que $\mathrm{M}_n(\mathbb{K})$ est un espace vectoriel de dimension finie. Il est donc muni de sa topologie canonique. Le groupe $\mathrm{GL}_n(\mathbb{K})$ est muni de la topologie induite. La multiplication de matrices $\mathrm{M}_n(\mathbb{K}) \times \mathrm{M}_n(\mathbb{K}) \to \mathrm{M}_n(\mathbb{K})$ est continue car ses composantes sont polynomiales, donc le produit de groupe dans $\mathrm{GL}_n(\mathbb{K})$ est continu. L'inverse de matrice est continu également : il suffit d'utiliser l'expression de l'inverse d'une matrice grâce à sa comatrice pour s'en convaincre.

3.1.5 Propriétés topologiques de $\mathrm{GL}_n(\mathbb{K})$ pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$

Théorème 3.1.41. *Le groupe $\mathrm{GL}_n(\mathbb{C})$ est*

1. *un ouvert de $\mathrm{M}_n(\mathbb{C})$*
2. *dense dans $\mathrm{M}_n(\mathbb{C})$*
3. *non-compact*
4. *connexe.*

On veut prouver le résultat pour tout n. Il est toujours instructif de commencer par traiter les petits cas. Dans le cas $n = 1$, on a $\mathrm{GL}_1(\mathbb{C}) = \mathbb{C}^*$.
Propriété 1 : C'est un ouvert de $\mathbb{C}$, on peut par exemple le prouver en montrant que son complémentaire, $\{0\}$, est fermé par critère séquentiel.
Propriété 2 : C'est un sous-ensemble dense dans $\mathbb{C}^*$: si on considère la suite $1/k$ pour $k \in \mathbb{N}^*$, on a une suite d'éléments de $\mathbb{C}^*$ qui converge vers 0. On a donc $\mathbb{C} = \mathbb{C}^* \cap \{0\} \subset \overline{\mathbb{C}^*} \subset \mathbb{C}$ donc tous ces ensembles sont égaux.
Propriété 3 : La preuve précédente montre que $\mathbb{C}^*$ n'est pas fermé dans $\mathbb{C}$, donc ce n'est pas un compact de $\mathbb{C}$.
Propriété 4 : Pour montrer que $\mathbb{C}^*$ est connexe, on montre que $\mathbb{C}^*$ est connexe par arcs. Les segments dans $\mathbb{C}$ forment des exemples basiques de chemins continus, pour peu qu'on les paramètre par $[0, 1]$: les composantes de l'application correspondante sont linéaires. Ils permettent de relier deux points quelconques dans $\mathbb{C}$. On peut de plus concaténer des segments : pour aller d'un point x à un point z, on peut d'abord emprunter le segment entre x et y puis le segments entre y et z. Le tout donne un chemin continu entre x et z si on paramètre le premier segment par $[0, 1/2]$ et le second par $[1/2, 1]$. Dans le cas de $\mathbb{C}^*$, il est interdit de passer par 0. Si x et y sont deux points de $\mathbb{C}^*$ et $y \notin \mathbb{R}^*_- x$, alors le segment entre x et y fournit un chemin continu *dans* $\mathbb{C}^*$ entre x et y. Si $y \in \mathbb{R}^*_- x$, alors on peut emprunter d'abord le segment entre y et ix, puis le segment entre ix et x, qui sont tous deux dans $\mathbb{C}^*$.

Démonstration. On considère maintenant $n \in \mathbb{N}^*$ quelconque.
Propriété 1 : On utilise l'application déterminant, qui est polynomiale donc continue. On a $\mathrm{GL}_n(\mathbb{C}^*) = \det^{-1}(\mathbb{C}^*)$ et $\mathbb{C}^* = \mathrm{GL}_1(\mathbb{C})$ est ouvert comme vu ci-dessus, donc $\mathrm{GL}_n(\mathbb{C})$ est ouvert.
Propriété 2 : Il existe de nombreuses façon de montrer que $\mathrm{GL}_n(\mathbb{C})$ est dense dans $\mathrm{M}_n(\mathbb{C})$, la clef étant de construire une suite (g_k) de matrices dans $\mathrm{GL}_n(\mathbb{C})$ convergeant vers A pour tout $A \in \mathrm{M}_n(\mathbb{C})$. On peut ici se servir du Chapitre 2, et imiter le cas $n = 1$. Supposons que A est une matrice de rang r, alors il existe g et h dans $\mathrm{GL}_n(\mathbb{C})$ tels que $A = gI_{r,n}h^{-1}$, où $I_{r,n} = \mathrm{diag}(1, \dots, 1, 0, \dots, 0)$ est la matrice diagonale avec des 1 comme r premiers coefficients diagonaux, et des 0 ensuite. On considère la suite définie par $g_k = g\,\mathrm{diag}(1, \dots, 1, 1/k, \dots, 1/k)h^{-1}$. Ses éléments sont clairement dans $\mathrm{GL}_n(\mathbb{C})$, et elle converge vers A, ce qu'il fallait démontrer.

Propriété 3 : Comme dans le cas $n = 1$, la preuve précédente montre que $\mathrm{GL}_n(\mathbb{C})$ n'est pas fermé.
Propriété 4 : Là aussi, on va montrer que $\mathrm{GL}_n(\mathbb{C})$ est connexe par arcs. Par concaténation des chemins, on sait qu'il suffit de relier toutes les matrices inversibles à I_n *via* un chemin continu de matrices inversibles. Pour ça aussi on peut utiliser le Chapitre 2 et la décomposition de Jordan (En réalité, on utilise surtout le fait que dans $\mathrm{GL}_n(\mathbb{C})$, toutes les matrices sont trigonalisables). Soit $A \in \mathrm{GL}_n(\mathbb{C})$ quelconque, et écrivons $A = gBg^{-1}$ où B est la forme de Jordan de A. Pour toute valeur propre λ de B, on choisit un chemin continu γ_λ dans $\mathbb{C}^*$ reliant λ à 1, en utilisant la connexité par arcs de $\mathbb{C}^*$. Plus précisément, disons que $\gamma_\lambda(0) = 1$ et $\gamma_\lambda(1) = \lambda$. Considérons la matrice B_t pour $t \in [0, 1]$ où, dans chaque bloc de Jordan $J_{\lambda,k}$, les coefficients diagonaux sont remplacés par $\gamma_\lambda(t)$, et les coefficients hors diagonale sont multipliés par t. Alors le chemin $\gamma(t) = gB_tg^{-1}$ est un chemin continu dans $\mathrm{GL}_n(\mathbb{C})$, tel que $\gamma(0) = gI_ng^{-1} = I_n$ et $\gamma(1) = gBg^{-1} = A$. □

Théorème 3.1.42. *Le groupe* $\mathrm{GL}_n(\mathbb{R})$ *est*

1. *un ouvert de* $\mathrm{M}_n(\mathbb{R})$
2. *dense dans* $\mathrm{M}_n(\mathbb{R})$
3. *non-compact*
4. *non-connexe.*

Démonstration. La même preuve fonctionne pour les trois premières propriétés. Pour la quatrième, il suffit de trouver deux ouverts non-vides disjoints qui recouvrent $\mathrm{GL}_n(\mathbb{R})$. On utilise le fait que $\mathbb{R}^* = \mathbb{R}^*_- \sqcup \mathbb{R}^*_+$ où les $\mathbb{R}^*_\pm$ sont ouverts, non-vides et disjoints. Alors $\mathrm{GL}_n(\mathbb{R}) = \det^{-1}(\mathbb{R}^*_-) \sqcup \det^{-1}(\mathbb{R}^*_+)$ et les $\det^{-1}(\mathbb{R}^*_\pm)$ sont ouverts, non-vides et disjoints. □

3.1.6 Cas de $\mathrm{SL}_n(\mathbb{C})$

Théorème 3.1.43. *Le groupe* $\mathrm{SL}_n(\mathbb{C})$ *est*

1. *fermé dans* $\mathrm{M}_n(\mathbb{C})$
2. *non-compact (si* $n \geq 2$*)*
3. *connexe.*

Démonstration. **Propriété 1 :** On a $\mathrm{SL}_n(\mathbb{C}) = \det^{-1}(\{1\})$, et $\{1\}$ est fermé dans $\mathbb{C}$, donc $\mathrm{SL}_n(\mathbb{C})$ est fermé dans $\mathrm{M}_n(\mathbb{C})$.
Propriété 2 : Il suffit de montrer que $\mathrm{SL}_n(\mathbb{C})$ n'est pas borné. On peut considérer pour cela les éléments $A_k = \mathrm{diag}(k, 1/k, 1, \ldots, 1)$ et la norme $\|\cdot\|_\infty$ sur $\mathrm{M}_n(\mathbb{C}) \simeq \mathbb{C}^{n^2}$. On a alors $\|A_k\|_\infty = k \to \infty$.
Propriété 3 : On utilise la connexité par arcs de $\mathrm{GL}_n(\mathbb{C})$. Soient A et B deux éléments de $\mathrm{SL}_n(\mathbb{C})$. Alors il existe un chemin continu $\tilde{\gamma}(t)$ dans $\mathrm{GL}_n(\mathbb{C})$ tel que $\tilde{\gamma}(0) = A$ et $\tilde{\gamma}(1) = B$. On considère le chemin $\gamma(t) = \mathrm{diag}(1/\det(\gamma(t)), 1 \ldots, 1)\gamma(t)$. C'est un chemin continu dans $\mathrm{SL}_n(\mathbb{C})$ reliant A et B. □

3.2 Les groupes $\mathrm{O}_n(\mathbb{R})$ et $\mathrm{SO}_n(\mathbb{R})$

3.2.1 Les groupes orthogonaux en général

Soit $\mathbb{K}$ un corps quelconque, et V un $\mathbb{K}$-espace vectoriel. On note $\mathrm{GL}(V)$ le groupe des applications linéaires inversibles de V dans V.

Définition 3.2.1. Une *forme quadratique* sur V est une application $q : V \to \mathbb{K}$ telle qu'il existe une forme bilinéaire symétrique $B : V \times V \to \mathbb{K}$ avec pour tout $u \in V$, $q(u) = B(u, u)$.

Le groupe $\mathrm{GL}(V)$ agit sur les formes bilinéaires symétriques par

$$(g \cdot B)(u, v) = B(g^{-1}u, g^{-1}v),$$

donc sur les formes quadratiques par pré-composition :

$$(g \cdot q)(u) = q(g^{-1}u).$$

Définition 3.2.2. Le groupe orthogonal de la forme quadratique q, noté $\mathrm{O}(q)$, est le stabilisateur de q sous l'action de $\mathrm{GL}(V)$ par pré-composition. Autrement dit,

$$\mathrm{O}(q) = \left\{ g \in \mathrm{GL}(V) \mid \forall u \in V \quad q(g^{-1}u) = q(u) \right\}$$

Définition 3.2.3. Deux formes quadratiques q et q' sont dites *équivalentes* si elles sont dans la même orbite, c'est-à-dire $\exists g \in \mathrm{GL}(V)$ tel que $q' = g \cdot q$.

Deux formes quadratiques équivalentes ont des groupes orthogonaux *conjugués*.

3.2.2 Exemple principal du cours : $\mathrm{O}_n(\mathbb{R})$

On considère l'espace vectoriel $\mathbb{R}^n$, muni de son produit scalaire $\langle \cdot, \cdot \rangle$ usuel : si les éléments de $\mathbb{R}^n$ sont pensés comme des vecteurs colonnes, on a pour x et $y \in \mathbb{R}^n$,

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

C'est une forme bilinéaire symétrique (définie positive), qui définit la forme quadratique

$$\|x\|_2^2 = \langle x, x \rangle = \sum_{i=1}^{n} x_i^2.$$

Le groupe orthogonal associé est noté $\mathrm{O}_n(\mathbb{R})$ (ou plus simplement O_n ou $\mathrm{O}(n)$) et est appelé *groupe orthogonal* tant qu'aucune confusion n'est possible.

Il y a plusieurs autre descriptions possibles du groupe orthogonal, en utilisant les différents points de vue sur $\mathrm{GL}(\mathbb{R}^n) = \mathrm{GL}_n(\mathbb{R})$: comme groupe de matrices, etc. Les éléments de $\mathrm{O}_n(\mathbb{R})$, vu comme matrices, sont appelés les *matrices orthogonales*.

On a

$$\begin{aligned}
\mathrm{O}_n(\mathbb{R}) &= \left\{ M \in \mathrm{M}_n(\mathbb{R}) \mid {}^t M M = I_n \right\} \\
&= \left\{ M \in \mathrm{M}_n(\mathbb{R}) \mid M {}^t M = I_n \right\} \\
&= \left\{ M \in \mathrm{GL}_n(\mathbb{R}) \mid M^{-1} = {}^t M \right\} \\
&= \{ M \in \mathrm{M}_n(\mathbb{R}) \mid \text{les colonnes de } M \text{ forment une b.o.n. de } \mathbb{R}^n \} \\
&= \{ M \in \mathrm{M}_n(\mathbb{R}) \mid \mathbb{R}^n \to \mathbb{R}^n, x \mapsto Mx \text{ est une isométrie linéaire de } \mathbb{R}^n \} \\
&= \{ M \in \mathrm{M}_n(\mathbb{R}) \mid \forall x, y \in \mathbb{R}^n \quad \langle Mx, My \rangle = \langle x, y \rangle \} \\
&= \{ M \in \mathrm{M}_n(\mathbb{R}) \mid \forall x \in \mathbb{R}^n \quad \|Mx\|_2 = \|x\|_2 \}
\end{aligned}$$

Proposition 3.2.4. *Si* $M \in \mathrm{O}_n(\mathbb{R})$, *alors* $\det(M) = \pm 1$

Démonstration. On

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

□

Définition 3.2.5. On appelle *groupe spécial orthogonal* le groupe

$$\mathrm{SO}_n(\mathbb{R}) := \{M \in \mathrm{O}_n(\mathbb{R}) \mid \det(M) = 1\}.$$

En conséquence directe de la définition, $\mathrm{SO}_n(\mathbb{R})$ est distingué dans $\mathrm{O}_n(\mathbb{R})$.

3.2.3 Classes de conjugaison dans $\mathrm{O}_n(\mathbb{R})$

Le théorème suivant, dit de *réduction des matrices orthogonales*, montre essentiellement que la forme de Jordan d'une matrice de $\mathrm{O}_n(\mathbb{R})$ est dans $\mathrm{O}_n(\mathbb{R})$, et que l'on peut utiliser une matrice de $\mathrm{O}_n(\mathbb{R})$ comme matrice de passage.

Théorème 3.2.6. *Pour toute matrice orthogonale $M \in \mathrm{O}_n(\mathbb{R})$, il existe : une matrice orthogonale $P \in \mathrm{O}_n(\mathbb{R})$, deux entiers naturels p et $q \in \mathbb{N}$, et un nombre fini d'angles $\theta_1, \ldots, \theta_k \in \mathbb{R}\backslash\pi\mathbb{Z}$) tels que*

$$\begin{aligned} PM^tP = PMP^{-1} &= \mathrm{diag}(I_p, -I_q, R_{\theta_1}, \ldots, R_{\theta_k}) \\ &= \begin{pmatrix} I_p & 0 & 0 & \cdots & 0 \\ 0 & -I_q & 0 & \cdots & 0 \\ 0 & 0 & R_{\theta_1} & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & R_{\theta_k} \end{pmatrix} \end{aligned}$$

Remarque 3.2.7. L'action de $\mathrm{GL}_n(\mathbb{K})$ sur $\mathrm{M}_n(\mathbb{K})$ par $g \cdot A = gA^tg$ est appelée action par congruences. En restriction à $\mathrm{O}_n(\mathbb{R})$, c'est la même action que l'action par conjugaison puisque $g^{-1} = {}^tg$ pour $g \in \mathrm{O}_n(\mathbb{R})$.

Remarque 3.2.8. En voyant la matrice P comme une matrice de changement de base, on a la formulation suivante du théorème : si f est une isométrie linéaire de $\mathbb{R}^n$, alors il existe une base orthonormée de $\mathbb{R}^n$ telle que la matrice de f dans cette base soit $\mathrm{diag}(I_p, -I_q, R_{\theta_1}, \ldots, R_{\theta_k})$.

Exercice 3.2.9. Montrer que $M \in \mathrm{SO}_n(\mathbb{R})$ si et seulement si q est pair.

3.2.4 Propriétés topologiques de $\mathrm{O}_n(\mathbb{R})$ et $\mathrm{SO}_n(\mathbb{R})$

Théorème 3.2.10.

1. *Les groupes $\mathrm{O}_n(\mathbb{R})$ et $\mathrm{SO}_n(\mathbb{R})$ sont compacts.*
2. *Le groupe $\mathrm{O}_n(\mathbb{R})$ n'est pas connexe.*
3. *Le groupe $\mathrm{SO}_n(\mathbb{R})$ est connexe par arcs.*

Démonstration. Pour la première propriété, comme on travaille dans l'espace vectoriel de dimension finie $\mathrm{M}_n(\mathbb{R})$, il suffit de montrer que ces deux groupes sont fermés et bornés, pour un choix arbitraire de norme. Pour la fermeture, on remarque que l'application $f : \mathrm{M}_n(\mathbb{R}) \to \mathrm{M}_n(\mathbb{R}), M \mapsto {}^tMM$ est continue (car toutes ses composantes sont des applications polynomiales) donc $\mathrm{O}_n(\mathbb{R}) = f^{-1}(\{I_n\})$ est un fermé. Le groupe spécial orthogonal est l'intersection des deux fermés $\mathrm{SL}_n(\mathbb{R})$ et $\mathrm{O}_n(\mathbb{R})$, il est donc fermé lui aussi.

Pour montrer que $\mathrm{O}_n(\mathbb{R})$ est borné, le plus simple est d'utiliser la norme d'opérateur $\|\cdot\|$ subordonnée à la norme Euclidienne $\|\cdot\|_2$ sur $\mathbb{R}^n$. En effet, si $M \in \mathrm{O}_n(\mathbb{R})$, alors l'application linéaire $\mathbb{R}^n \to \mathbb{R}^n$ définie par M est une isométrie, et donc $\|Mx\|_2 = \|x\|_2$ pour tout $x \in \mathbb{R}^n$. On a alors $\|M\| = 1$ par définition, donc $\mathrm{O}_n(\mathbb{R})$ est borné. Puisque $\mathrm{SO}_n(\mathbb{R})$ est contenu dans $\mathrm{O}_n(\mathbb{R})$, le même énoncé est vrai pour $\mathrm{SO}_n(\mathbb{R})$.

Pour la deuxième propriété, on utilise encore que det est une application continue, et que

$$\mathrm{O}_n(\mathbb{R}) = (\mathrm{O}_n(\mathbb{R}) \cap \overset{-1}{\det}(\mathbb{R}_-^*)) \cup (\mathrm{O}_n(\mathbb{R}) \cap \overset{-1}{\det}(\mathbb{R}_+^*))$$

fournit donc une décomposition en ouverts disjoints. Il reste à vérifier que les deux ouverts sont non vides : on a $\det(I_n) = 1$ et $\det(\mathrm{diag}(-1, I_{n-1})) = -1$.

Enfin, pour la troisième propriété, on construit un chemin continu explicite entre $M \in \mathrm{SO}_n(\mathbb{R})$ et I_n, en utilisant la réduction des matrices orthogonales. On écrit, en utilisant le théorème,

$$M = P\,\mathrm{diag}(I_p, -I_q, R_{\theta_1}, \dots, R_{\theta_k})P^{-1}$$

avec $P \in \mathrm{O}_n(\mathbb{R})$. Par l'exercice 131, q est pair. Or, on a $-I_2 = R_\pi$, donc

$$M = P\,\mathrm{diag}(I_p, R_\pi, \dots, R_\pi, R_{\theta_1}, \dots, R_{\theta_k})P^{-1}.$$

Le chemin

$$\gamma(t) := P\,\mathrm{diag}(I_p, R_{t\pi}, \dots, R_{t\pi}, R_{t\theta_1}, \dots, R_{t\theta_k})P^{-1}$$

fournit un chemin continu entre $\gamma(0) = I_n$ et $\gamma(1) = M$. □

3.2.5 Les éléments de $\mathrm{O}_2(\mathbb{R})$ et $\mathrm{SO}_2(\mathbb{R})$

Proposition 3.2.11. *Toute matrice de* $\mathrm{SO}_2(\mathbb{R})$ *est de la forme* R_θ *avec* $\theta \in [0, 2\pi[$.

L'application linéaire correspondante à une telle matrice est exactement la rotation d'angle θ centrée en l'origine.

Démonstration. Le théorème général de réduction ne donne pas directement ce résultat a priori, puisqu'il st seulement à conjugaison près. Un calcul direct donne l'énoncé de la proposition très rapidement. On écrit

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

On a $M \in \mathrm{SO}_2(\mathbb{R})$ si et seulement si ${}^tMM = I_2$ et $\det(M) = 1$, c'est à dire si et seulement si les quatre équations suivantes sont vérifiées

$$\begin{aligned} a^2 + b^2 &= 1 \\ ac + bd &= 0 \\ c^2 + d^2 &= 1 \\ ad - bc &= 1 \end{aligned}$$

Par la première équation, on peut écrire $a = \cos(\theta)$ et $b = \sin(\theta)$ avec $\theta \in [0, 2\pi[$. La seconde équation montre que le vecteur (c, d) est orthogonal au vecteur (a, b), donc il existe $t \in \mathbb{R}$ tel que $(c, d) = t(-b, a)$. La troisième équation montre que le vecteur (c, d) est de norme 1, donc que $t = \pm 1$. Enfin la dernière équation permet de vérifier que $t = 1$ est la seule possibilité, donc $M = R_\theta$. □

Dans le cas $M \in \mathrm{O}_2(\mathbb{R}) \setminus \mathrm{SO}_2(\mathbb{R})$, le calcul ci-dessus s'applique excepté la dernière ligne, où $t = -1$, donc M est de la forme

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

Le théorème général donne une meilleure information géométrique. Formellement, il assure que M est conjugué, via une matrice orthogonale, à la matrice

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

L'isométrie de $\mathbb{R}^2$ associé à cette dernière matrice est la réflexion orthogonale par rapport à l'axe des abscisses. Remarquons que l'axe des abscisses est exactement l'ensemble des points invariants, donc le sous-espace propre pour la valeur propre 1. En général, l'isométrie associée à M est la réflexion orthogonale par rapport à son sous-espace propre pour la valeur propre 1.

On finit ce paragraphe sur les isométries linéaires de $\mathbb{R}^2$ par une quasi-tautologie :

Théorème 3.2.12. *Le groupe* $\mathrm{SO}_2(\mathbb{R})$ *est isomorphe* comme groupe topologique *au cercle unité* $S^1 \subset \mathbb{R}^2$ *(où la loi de groupe sur* S^1 *est induite par la multiplication dans* $\mathbb{C} \simeq \mathbb{R}^2$*).*

Démonstration. Il suffit d'exhiber les bijections réciproques, qui sont facilement vérifiées à la fois morphismes de groupes et continues :

$$S^1 \to \mathrm{SO}_2(\mathbb{R}) \qquad (x, y) \mapsto \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$$

$$\mathrm{SO}_2(\mathbb{R}) \to S^1 \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (a, b)$$

□

3.2.6 Les éléments de $\mathrm{O}_3(\mathbb{R})$ et $\mathrm{SO}_3(\mathbb{R})$

Le théorème de réduction des matrices orthogonales permet de classer les éléments de $\mathrm{O}_3(\mathbb{R})$ en quatre familles selon leur réduction.

- Il convient de mettre à part l'*identité* I_3.
- Si $M \in \mathrm{O}_3(\mathbb{R})$ est conjuguée à $\mathrm{diag}(1, R_\theta)$ avec $\theta \in]0, 2\pi[$, alors M est une *rotation* dans $\mathbb{R}^3$ d'axe le sous-espace propre pour la valeur propre 1 et d'angle θ.
- Si $M \in \mathrm{O}_3(\mathbb{R})$ est conjuguée à $\mathrm{diag}(1, 1, -1)$, alors M est une *réflexion* orthogonale par rapport au plan des vecteurs propres pour la valeur propre 1.
- Enfin, si $M \in \mathrm{O}_3(\mathbb{R})$ est conjuguée à $\mathrm{diag}(-1, R_\theta)$ avec $\theta \in]0, 2\pi[$, alors M est une *roto-réflexion* dans $\mathbb{R}^3$ d'axe le sous-espace propre pour la valeur propre -1 et d'angle θ.

Remarque 3.2.13. Les rotations et réflexions orthogonales correspondent à l'intuition comme transformations de $\mathbb{R}^3$. Une roto-réflexion d'axe D et d'angle θ est la composée (dans n'importe quel ordre) de la rotation d'axe D et d'angle θ avec la réflexion par rapport au plan orthogonal à D.

Exercice 3.2.14. Soit $M \in \mathrm{O}_3(\mathbb{R})$. Alors on peut reconnaître facilement le type de M :

- Si $\det(M) = 1$, alors M est soit l'identité, soit une rotation (et on peut considérer l'identité comme une rotation d'angle 0).
- Si $\det(M) = -1$ et 1 est une valeur propre de M, alors M est une réflexion.
- Sinon, M est une roto-réflexion.

3.3 Les groupes U_n et SU_n

3.3.1 Les groupes unitaires en général (sur $\mathbb{C}$)

Soit V un espace vectoriel complexe de dimension finie.

Définition 3.3.1. Une *forme Hermitienne* sur V est une application $\phi : E \times E \to \mathbb{C}$ telle que pour tout $\lambda \in \mathbb{C}$, u, v, $w \in V$, on a :

1. $\phi(\lambda u + v, w) = \bar{\lambda}\phi(u, v) + \phi(v, w)$
2. $\phi(u, \lambda v + w) = \lambda\phi(u, v) + \phi(u, w)$
3. $\phi(u, v) = \overline{\phi(v, u)}$.

Remarque 3.3.2. — Les deux dernières propriétés impliquent la première.
— Si on a seulement les deux premières propriétés, on parle de *forme sesquilinéaire*.

On dit que la forme Hermitienne ϕ définit un *produit scalaire Hermitien* si elle est *définie positive*, c'est-à-dire

$$\forall u \in V, \phi(u, u) \in \mathbb{R}_+ \qquad \text{et} \qquad \phi(u, u) = 0 \text{ ssi } u = 0.$$

Dans ce cas, on appelle (V, ϕ) un *espace Hermitien*, et on peut parler d'isométries de l'espace Hermitien, de bases othonormées, etc.

Exercice 3.3.3. Le groupe $\mathrm{GL}(V)$ agit sur l'ensemble des formes Hermitiennes par précomposition : $(g \cdot \phi)(u, v) = \phi(g^{-1}u, g^{-1}v)$.

Définition 3.3.4. Le *groupe unitaire* $\mathrm{U}(\phi)$ *associé à la forme Hermitienne* ϕ est son stabilisateur pour cette action :

$$\mathrm{U}(\phi) := \{g \in \mathrm{GL}(V) \mid g \cdot \phi = \phi\}.$$

3.3.2 La forme Hermitienne usuelle sur $\mathbb{C}^n$

L'application $\mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$, $(Z, W) \mapsto \langle Z, W\rangle := Z^*W = \sum_{k=1}^n \bar{z}_k w_k$ définit un produit scalaire Hermitien sur $\mathbb{C}^n$. Le groupe unitaire associé est noté $\mathrm{U}_n(\mathbb{C})$ (ou plus simplement U_n ou $\mathrm{U}(n)$) et il est généralement appelé *groupe unitaire*.

Comme d'habitude, on peut voir les éléments de U_n comme des applications linéaires inversibles $\mathbb{C}^n \to \mathbb{C}^n$, ou comme des matrices, ou comme des bases de $\mathbb{C}^n$ (données par les vecteurs colonnes), etc. On en déduit différentes descriptions possibles de $\mathrm{U}_n(\mathbb{C})$:

$$\begin{aligned}
\mathrm{U}_n &= \{M \in \mathrm{M}_n(\mathbb{C}) \mid M^*M = I_n\} \\
&= \{M \in \mathrm{M}_n(\mathbb{C}) \mid MM^* = I_n\} \\
&= \left\{M \in \mathrm{GL}_n(\mathbb{C}) \mid M^{-1} = M^*\right\} \\
&= \{M \in \mathrm{M}_n(\mathbb{C}) \mid \text{les colonnes de } M \text{ forment une b.o.n. pour } \langle\cdot,\cdot\rangle\} \\
&= \{M \in \mathrm{M}_n(\mathbb{C}) \mid \text{l'application } x \mapsto Mx \text{ est une isométrie pour } \langle\cdot,\cdot\rangle\}
\end{aligned}$$

On appelle *matrices unitaires* les éléments de U_n.

Proposition 3.3.5. *Si* $M \in \mathrm{U}_n$, *alors* $|\det(M)| = 1$.

Démonstration. On a

$$\begin{aligned} 1 &= \det(I_n) \\ &= \det(M^*M) \\ &= \det(M^*)\det(M) \\ &= \overline{\det(M)}\det(M) \\ &= |\det(M)|^2. \end{aligned}$$

□

Définition 3.3.6. Le *groupe spécial unitaire* SU_n est le groupe

$$\mathrm{SU}_n = \mathrm{SL}_n(\mathbb{C}) \cap \mathrm{U}_n\,.$$

Remarque 3.3.7. C'est un sous-groupe distingué de U_n.

3.3.3 Réduction des matrices unitaires

Théorème 3.3.8. *Soit $U \in \mathrm{U}_n$ une matrice unitaire. Il existe $P \in \mathrm{U}_n$ et $\lambda_1, \lambda_2, \ldots, \lambda_n$ des nombres complexes de module 1 tels que $U = P\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)P^{-1}$.*

Remarque 3.3.9.
- Ce théorème donne un représentant diagonal unique (à permutation des coefficients diagonaux près) de chaque classe de conjugaison dans $\mathrm{U}_n(\mathbb{C})$.
- Si on voit U comme définissant une application linéaire $f : \mathbb{C}^n \to \mathbb{C}^n$ et P comme une matrice de passage, alors l'énoncé peut se traduire en :

Théorème 3.3.10. *Si E est un espace Hermitien et $f : E \to E$ est une isométrie alors il existe une base orthonormée formée de vecteurs propres pour f, et toutes les valeurs propres ont module 1.*

Lemme 3.3.11. *Si $f : E \to E$ est une isométrie et λ est une valeur propre de f, alors $|\lambda| = 1$.*

Démonstration. Soir x un vecteur propre pour la valeur propre λ. On a

$$\begin{aligned} 1 &= \frac{\langle f(x), f(x)\rangle}{\langle x, x\rangle} \\ &= \frac{\langle \lambda x, \lambda x\rangle}{\langle x, x\rangle} \\ &= \bar{\lambda}\lambda\frac{\langle x, x\rangle}{\langle x, x\rangle} \\ &= |\lambda|^2 \end{aligned}$$

□

Lemme 3.3.12. *Si $f : E \to E$ est une isométrie, u est un vecteur propre de f pour une valeur propre λ quelconque, alors $u^\perp = \{v \in E \mid \langle v, u\rangle = 0\}$ est stable par f.*

Démonstration. Soit $x \in u^\perp$. On a

$$\begin{aligned} \langle f(x), u\rangle &= \langle f(x), \frac{1}{\lambda}f(u)\rangle \\ &= \frac{1}{\lambda}\langle f(x), f(u)\rangle \\ &= \frac{1}{\lambda}\langle x, u\rangle \\ &= 0 \end{aligned}$$

□

Preuve du Théorème. Par récurrence sur $n = \dim(E)$:

— Si $n = 1$, alors f est la multiplication par un λ avec $|\lambda| = 1$.

— Supposons le résultat prouvé pour la dimension n, et soit E un espace Hermitien de dimension $n+1$. Puisqu'on travaille sur un espace vectoriel complexe, f admet au moins une valeur propre λ_{n+1}, qui a module 1 par le premier lemme. Choisissons e_{n+1} l'un des vecteurs propres correspondants, de norme 1. Le sous-espace $F = e_{n+1}^{\perp}$ est un sous-espace Hermitien de dimension n, et il est stable par F. On peut donc appliquer l'hypothèse de récurrence à $f|_F$ et obtenir une b.o.n. $(e_1, \ldots, e_n)$ de F formée de vecteurs propres pour f. La base $(e_1, \ldots, e_{n+1})$ fournit la b.o.n. recherchée dans l'énoncé du théorème.

□

3.3.4 Propriétés topologiques de U_n et SU_n

Théorème 3.3.13. *Les groupes* U_n *et* SU_n *sont compacts et connexes par arcs.*

Démonstration. Les preuves sont essentiellement identiques au cas de O_n et SO_n. Pour la connexité : soit $U \in \mathrm{U}_n$, qu'on écrit $U = P\,\mathrm{diag}(e^{i\theta_1}, \ldots, e^{i\theta_n})P^{-1}$ avec $U \in \mathrm{U}_n$ et $\theta_1, \ldots, \theta_n \in \mathbb{R}$ grâce au théorème précédent.

On utilise le chemin continu $\gamma(t) = P\,\mathrm{diag}(e^{it\theta_1}, \ldots, e^{it\theta_n})P^{-1}$ qui relie $\gamma(0) = I_n$ et $\gamma(1) = U$, et qui reste dans U_n puisque toute matrice diagonale dont les coefficients diagonaux sont de module 1 est une matrice unitaire.

Considérons maintenant le cas où $U \in \mathrm{SU}_n$. Il n'est pas assuré que le chemin ci-dessus reste dans SU_n : $\det(\gamma(t)) = e^{it(\theta_1+\cdots+\theta_n)}$ et qu'on peut par exemple avoir $\theta_1 + \cdots + \theta_n = 2\pi$. Pour résoudre ce problème, il suffit de choisir dès le départ les représentants $\theta_1, \ldots, \theta_n$ de sorte que $\theta_1 + \cdots + \theta_n = 0$. □

Chapitre 4

Application exponentielle et décomposition polaire

4.1 Matrices Hermitiennes

4.1.1 Définition

On rappelle que si $M \in \mathrm{M}_n(\mathbb{C})$, la matrice *adjointe* ou *transconjuguée*, notée M^*, est la matrice dont les coefficients sont les $\bar{m}_{j,i}$ si les coefficients de M sont les $m_{i,j}$.

Définition 4.1.1. Une matrice $M \in \mathrm{M}_n(\mathbb{C})$ est *Hermitienne* si $M^* = M$.

Exemple 4.1.2.

1. Si $n = 1$, les matrices Hermitiennes coïncident avec les nombres réels : $\bar{z} = z$ ssi $z \in \mathbb{R}$.
2. Si $n = 2$, une matrice $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ satisfait $M^* = M$ si et seulement si a, $d \in \mathbb{R}$ et $b = \bar{c}$. On peut donc l'écrire en termes de quatre coefficients réels α, β, γ, δ sous la forme
$$M = \begin{pmatrix} \alpha & \beta + i\gamma \\ \beta - i\gamma & \delta \end{pmatrix}$$
3. Une matrice diagonale est Hermitienne si et seulement si ses coefficients sont réels. En fait pour toute matrice Hermitienne, les coefficients diagonaux sont réels.

Définition 4.1.3. Une matrice $M \in \mathrm{M}_n(\mathbb{C})$ est *anti-Hermitienne* si $M^* = -M$.

Exemple 4.1.4.

1. Si $n = 1$, les matrices anti-Hermitiennes coïncident avec les nombres imaginaires purs : $\bar{z} = -z$ ssi $z \in i\mathbb{R}$.
2. Si $n = 2$, une matrice $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ satisfait $M^* = -M$ si et seulement si a, d sont imaginaires purs et $b = -\bar{c}$. On peut donc l'écrire en termes de quatre coefficients réels α, β, γ, δ sous la forme
$$M = \begin{pmatrix} i\alpha & \beta + i\gamma \\ -\beta + i\gamma & i\delta \end{pmatrix}$$
3. Une matrice diagonale est anti-Hermitienne si et seulement si ses coefficients diagonaux sont imaginaires purs. En fait pour toute matrice anti-Hermitienne, les coefficients diagonaux sont imaginaires purs.

On note H_n l'ensemble des matrices Hermitiennes, et AH_n l'ensemble des matrices anti-Hermitiennes.

Proposition 4.1.5.

1. *Les sous-ensembles H_n et AH_n sont des sous-espaces vectoriels* réels *de l'espace vectoriel* $\mathrm{M}_n(\mathbb{C}) \simeq \mathbb{R}^{2n^2}$.
2. *Ce ne sont pas des sous-espaces vectoriels complexes de* $\mathrm{M}_n(\mathbb{C})$.
3. *Ils fournissent une décomposition en somme directe* $\mathrm{M}_n(\mathbb{C}) = H_n \oplus AH_n$.
4. *La dimension de l'espace vectoriel H_n est n^2, de même que celle de AH_n.*

Démonstration. Soit $t \in \mathbb{C}$, M, $N \in \mathrm{M}_n(\mathbb{C})$. Alors on a

$$(tM + N)^* = \bar{t}M^* + N^*.$$

On en déduit que, si $t \in \mathbb{R}$, M et $N \in H_n$, alors $tM + N \in H_n$, c'est-à-dire que H_n est un sous-espace vectoriel réel de $\mathrm{M}_n(\mathbb{C})$. De même pour AH_n.

Si $M \in H_n \cap AH_n$, alors $M = M^* = -M$, donc $M = 0$. On en déduit que H_n et AH_n sont en somme directe $H_n \cap AH_n = \{0\}$. On en déduit aussi que H_n n'est pas un sous-espace vectoriel complexe : si $i \in \mathbb{C}$ et $M \in H_n \setminus \{0\}$, on a $iM \in AH_n$ donc $iM \notin H_n$. Plus généralement, on remarque que la multiplication par n'importe quel nombre imaginaire pur envoie H_n dans AH_n et AH_n dans H_n.

Plus précisément que la multiplication par i fournit un isomorphisme entre H_n et AH_n. En effet, c'est une application $\mathbb{R}$-linéaire car elle est $\mathbb{C}$-linéaire sur $\mathrm{M}_n(\mathbb{C})$, et c'est un isomorphisme entre H_n et AH_n d'inverse la multiplication par $-i$ par la remarque précédente. On a en particulier $\dim(H_n) = \dim(AH_n)$. On pourra déduire de cette égalité que $\dim(H_n) = n^2$ une fois prouvée la décomposition $\mathrm{M}_n(\mathbb{C}) = H_n \oplus AH_n$.

Pour prouver ce dernier résultat, on écrit, si $M \in \mathrm{M}_n(\mathbb{C})$,

$$M = \frac{M + M^*}{2} + \frac{M - M^*}{2}.$$

Le premier terme de cette décomposition est dans H_n alors que le second est dans AH_n, on a donc montré de cette manière $\mathrm{M}_n(\mathbb{C}) = H_n + AH_n$. On avait déjà prouvé que ces deux sous-espaces étaient en somme directe, donc on a terminé la preuve. □

Remarque 4.1.6. Une autre méthode pour calculer la dimension de H_n serait de déterminer explicitement une base, comme on l'a fait implicitement pour l'exemple de H_2 : les matrices suivantes fournissent une base de H_2.

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

4.1.2 Réduction des matrices Hermitiennes

On rappelle que $\langle X, Y \rangle = X^*Y$ désigne le produit scalaire Hermitien standard sur $\mathbb{C}^n$.

Proposition 4.1.7. *Si $H \in \mathrm{M}_n(\mathbb{C})$ est Hermitienne, alors pour tout X, $Y \in \mathbb{C}^n$,*

$$\langle HX, Y \rangle = \langle X, HY \rangle$$

Démonstration. On a

$$\begin{aligned}\langle HX, Y\rangle &= (HX)^* Y \\ &= X^* H^* Y \\ &= X^* HY \\ &= \langle X, HY\rangle\end{aligned}$$

□

Théorème 4.1.8. *Soit $H \in H_n$ une matrice Hermitienne. Alors il existe une matrice unitaire $P \in U(n)$ et une matrice diagonale réelle D telles que*

$$H = P^* DP = P^{-1} DP.$$

Exercice 4.1.9. La preuve de ce théorème est obtenue par récurrence de manière identique à celle pour les matrices unitaires, une fois que les deux lemmes suivants sont démontrés.

Attention, pour être complétement rigoureux dans la démonstration, il faut être capable de manipuler les changements de bases efficacement, ou traduire la définition de matrices Hermitiennes du monde des matrices au monde des applications linéaires entre deux espaces Hermitiens sans base fixée.

Lemme 4.1.10. *Si H est une matrice Hermitienne, alors les valeurs propres de H sont réelles.*

Démonstration. Soit λ une valeur propre de H, et X un vecteur propre pour la valeur propre λ. Alors

$$\langle HX, X\rangle = \langle \lambda X, X\rangle = \bar{\lambda}\langle X, X\rangle$$

et

$$\langle X, HX\rangle = \langle X, \lambda X\rangle = \lambda\langle X, X\rangle$$

or les deux sont égaux par la proposition, donc $\lambda = \bar{\lambda}$. □

Lemme 4.1.11. *Soit X un vecteur propre pour H. Alors $X^\perp$ est stable par H.*

Démonstration. Soit $Y \in X^\perp$. On a

$$\langle X, HY\rangle = \langle HX, Y\rangle = \langle \lambda X, Y\rangle = \bar{\lambda}\langle X, Y\rangle = 0$$

□

En particulier, le théorème précédent assure que les matrices Hermitiennes sont diagonalisables. Si deux matrices Hermitiennes commutent, elles sont donc co-diagonalisables par le théorème de diagonalisation simultanée. En suivant la preuve de ce théorème, on vérifie de plus que la matrice de changement de base communes peut être choisie unitaire. C'est le résultat suivant.

Théorème 4.1.12. *Soient A_1 et A_2 deux matrices Hermitiennes qui commutent. Alors il existe une matrice unitaire P et deux matrices diagonales réelles D_1 et D_2 telles que $A_1 = P^* D_1 P$ et $A_2 = P^* D_2 P$.*

Démonstration. Le théorème précédent appliqué à A_1 implique : $\mathbb{C}^n$ est la somme directe orthogonale (pour le produit scalaire Hermitien) des sous-espaces propres E_1, ..., E_r de A_1. De même, pour A_2, $\mathbb{C}^n$ est la somme orthogonale des sous-espaces propres F_1, ..., F_s de A_2. Si on

montre que A_2 laisse stable les sous-espaces propres de A_1, alors pour obtenir le théorème il suffit de choisir une base orthonormée adaptée à la décomposition en somme directe orthogonale

$$\mathbb{C}^n = \bigoplus_{1\leq i\leq r, 1\leq j\leq s}^{\perp} E_i \cap F_j.$$

Le fait que A_2 laisse stable les sous-espaces propres de A_1 est standard. On rappelle la preuve rapidement. Soit λ une valeur propre de A_1. Comme A_2 commute avec A_1, on a

$$A_2(A_1 - \lambda I_n) = (A_1 - \lambda I_n)A_2.$$

Si X est un vecteur propre de A_1 pour la valeur propre λ, on a donc

$$(A_1 - \lambda I_n)(A_2 X) = A_2((A_1 - \lambda I_n)X) = 0$$

donc A_2X est dans le sous-espace propre $\ker(A_1 - \lambda I_n)$ de A_1 pour la valeur propre λ. □

4.2 Décomposition polaire pour $\mathrm{GL}_n(\mathbb{C})$

4.2.1 Matrices Hermitiennes positives

Définition 4.2.1. Soit $H \in H_n$ une matrice Hermiienne. On dit que H est *positive* si pour tout $X \in \mathbb{C}^n$, $X^*HX \geq 0$. On dit que H est *définie positive* si pour tout $X \in \mathbb{C}^n \setminus \{0\}$, $X^*HX > 0$.

On note H_n^+ l'ensemble des matrices Hermitiennes positives, et H_n^{++} l'ensemble des matrices Hermitiennes définies positives. On a bien sûr $H_n^{++} \subset H_n^+$.

Exemple 4.2.2.

- La matrice I_n est définie positive.
- La matrice $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ est positive mais pas définie positive.
- La matrice $\begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$ n'est pas positive.
- La matrice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ n'est pas positive.

Proposition 4.2.3. *Soit H une matrice Hermitienne. Alors H est positive si et seulement si toutes ses valeurs propres sont positives, et H est définie positive si et seulement si toutes ses valeurs propres sont strictement positives.*

Démonstration. Soit X est un vecteur propre de H pour la valeur propre λ. Alors $X^*HX = \lambda X^*X$, donc les implications dans le sens direct sont évidentes.

Pour l'autre sens, on utilise la réduction des matrices Hermitiennes. On écrit $H = P^*DP$ où les coefficients diagonaux de la matrice diagonale D sont les valeurs propres de H. Si toutes les valeurs propres de H sont positives, alors pour tout $Y \in \mathbb{C}^n$, $Y^*DY \geq 0$. Comme P est inversible, pour $X \in \mathbb{C}^n$ on peut considérer $Y = PX$ et avoir ainsi $X^*HX = Y^*DY \geq 0$.

Le cas où H est définie positive se traite exactement de la même manière. □

Corollaire 4.2.4.

- *Si H est Hermitienne positive, alors* $\mathrm{tr}(H) \geq 0$ *et* $\det(H) \geq 0$.
- *Si $H \in H_n^{++}$, alors* $\mathrm{tr}(H) > 0$ *et* $\det(H) > 0$.
- *Soit $H \in H_n^+$. Alors H est définie positive si et seulement si le déterminant de H est non nul. En d'autre termes, $H_n^{++} = H_n^+ \cap \mathrm{GL}_n(\mathbb{C})$.*

4.2.2 Racines carrées

Théorème 4.2.5. *Soit $H \in H_n^+$. Alors il existe une unique matrice Hermitienne positive P telle que $P^2 = H$. De plus, si $H \in H_n^{++}$, alors $P \in H_n^{++}$.*

On notera $P = \sqrt{H}$ cette racine carrée dans H_n^+.

Attention, en général il peut exister beaucoup de matrices R telles que $H = R^2$. Le théorème dit qu'il y en a exactement un dans H_n^+. Par exemple, $\sqrt{I_2} = I_2$ mais toutes les matrices R suivantes vérifient $R^2 = I_2$:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}, \dots$$

Démonstration. Supposons que $P \in H_n^+$ est une racine carrée de H comme dans l'énoncé. On a alors $PH = P^3 = HP$, c'est-à-dire que les deux matrices P et H commutent. Les sous-espaces propres de H sont donc stables par P. Par le théorème de diagonalisation simultanée, on peut trouver une base de vecteurs propres commune à H et P. Si e_i est un élément de cette base et λ, μ les valeurs propres respectives de H et P correspondantes, alors on a $\lambda e_i = P^2 e_i = \mu^2 e_i$, donc $\mu = \sqrt{\lambda}$ qui est la racine carrée positive de λ. On en déduit que sur l'ensemble du sous-espace propre de H associé à la valeur propre λ, P agit comme $\sqrt{\lambda}\mathrm{Id}$. Or $\mathbb{C}^n$ est la somme directe des sous-espaces propres de H, donc ceci détermine complètement P. Réciproquement, si P est définie par $PX = \sqrt{\lambda}X$ pour X vecteur propre de H pour la valeur propre λ, alors P définit bien une racine carrée de H, Hermitienne positive car on peut choisir une base orthonormée de $\mathbb{C}^n$ qui rend H et P diagonales à coefficients réels positifs. □

4.2.3 Décomposition polaire

Théorème 4.2.6. *Soit $A \in \mathrm{GL}_n(\mathbb{C})$. Il existe un unique couple (U, P) tel que $U \in \mathrm{U}(n)$, $P \in H_n^{++}$ et $A = UP$.*

Remarque 4.2.7. Pour $n = 1$, on a $\mathrm{GL}_1(\mathbb{C}) = \mathbb{C}^*$, $H_1^{++} = \mathbb{R}_+^*$ et $\mathrm{U}(1) = \{z \in \mathbb{C} \mid |z| = 1\}$. L'énoncé ci-dessus est bien entendu vrai dans ce cas : il existe une unique manière d'écrire un nombre complexe non-nul a comme $a = up$ avec $|u| = 1$ et $p \in \mathbb{R}_+^*$. On a en effet $p = |a| = \sqrt{\bar{a}a}$ et $u = \frac{a}{p}$. La preuve dans le cas général suit la même trame.

Démonstration. Commençons par montrer l'unicité. Soient U et P comme dans l'énoncé. Alors $A^*A = (UP)^*UP = P^*U^*UP = P^2$ car $P \in H_n$ et $U \in U(n)$. Or $(A^*A)^* = A^*A$ et $X^*(A^*A)X = (AX)^*(AX) > 0$ pour $X \in \mathbb{C}^* \setminus \{0\}$, donc $A^*A \in H_n^{++}$. On en déduit que P est l'unique racine carrée Hermitienne positive de A^*A, donc est déterminée uniquement par A. La relation $U = AP^{-1}$ détermine uniquement U étant données A et P, donc on a prouvé l'unicité.

Pour l'existence, on vérifie simplement que les choix imposés ci-dessus fonctionnent. En posant $P = \sqrt{A^*A}$ et $U = AP^{-1}$, on a bien $P \in H_n^+ +$ et $UP = A$. Il reste à vérifier que U est unitaire. On calcule $U^*U = (AP^{-1})^*AP^{-1} = P^{-1}A^*AP^{-1} = P^{-1}P^2P^{-1} = I_n$, ce qui montre que $U \in U(n)$ et finit la preuve. □

Théorème 4.2.8. *L'application $\phi : U(n) \times H_n^{++} \to \mathrm{GL}_n(\mathbb{C})$ définie par $\phi(U, P) = UP$ est un homéomorphisme.*

Ici les topologies impliquées sont bien sûr les topologies induites sur ces sous-ensembles de $\mathrm{M}_n(\mathbb{C})$.

Démonstration. La décomposition polaire démontrée juste avant montre que ϕ est une bijection. Cette application est continue car le produit de matrices est continu. Il reste donc uniquement à montrer que l'inverse ϕ^{-1} de ϕ est continu.

Par la preuve de la décomposition polaire, on a $\phi^{-1}(A) = (U, P)$ où $P = \sqrt{A^*A}$ et $U = AP^{-1}$. Si on montre que la racine carrée est continue, alors on aura montré que ϕ^{-1} est continue, et on aura fini la preuve.

On va utiliser le critère séquentiel de continuité. Supposons que (B_k) est une suite dans H_n^{++}, qui converge vers $C \in H_n^{++}$. En particulier, (B_k) est bornée et (en exercice), $(\sqrt{B_k})$ est également bornée. La suite $(\sqrt{B_k})$ vit dans un compact, on veut montrer qu'elle converge vers $\sqrt{C}$. Pour cela on utilisera le fait classique de topologie suivant : dans un espace compact, une suite converge si et seulement si elle a une unique valeur d'adhérence.

Considérons donc une sous-suite convergente $(\sqrt{B_{k_j}})$, et notons Q sa limite. Par continuité du produit de matrices, la suite (B_{k_j}) converge vers Q^2, mais on sait déjà qu'elle converge vers C, donc $Q^2 = C$. De plus, H_n^+ est fermé (en exercice), donc Q est l'unique racine carrée Hermitienne positive $\sqrt{C}$ de C. Par le fait de topologie rappelé ci-dessus, la suite $(\sqrt{B_k})$ est convergente et sa limite est $\sqrt{C}$, ce qu'on voulait démontrer. □

4.3 Décomposition polaire réelle

4.3.1 Passer du complexe au réel

Les matrices à coefficients réels sont les matrices à coefficients complexes qui sont égales à leur conjuguées :

$$\mathrm{M}_n(\mathbb{R}) = \{M \in \mathrm{M}_n(\mathbb{C}) \mid \bar{M} = M\}.$$

On obtient de cette manière diverses correspondances entre des groupes ou sous-ensembles de matrices étudiés plus tôt en considérant uniquement des matrices à coefficients réels, par exemple $\mathrm{GL}_n(\mathbb{R}) = \mathrm{GL}_n(\mathbb{C}) \cap \mathrm{M}_n(\mathbb{R})$ et $\mathrm{O}(n) = \mathrm{U}(n) \cap \mathrm{M}_n(\mathbb{R})$. D'autre part, l'ensemble des matrices symétriques réelles $\S_n$ s'obtient à partir des matrices Hermitiennes par $S_n = H_n \cap \mathrm{M}_n(\mathbb{R})$, de même que les matrices symétriques réelles positives et définies positives s'obtiennent respectivement par $S_n^+ = H_n^+ \cap \mathrm{M}_n(\mathbb{R})$ et $S_n^{++} = H_n^{++} \cap \mathrm{M}_n(\mathbb{R})$. De même l'ensemble des matrices anti-symétriques réelles AS_n est égal à $AH_n \cap \mathrm{M}_n(\mathbb{R})$.

L'analogue de la première proposition qu'on a démontré sur les matrices Hermitiennes est :

Proposition 4.3.1. *Les sous-ensembles S_n et AS_n sont des sous-espaces vectoriels réels de $\mathrm{M}_n(\mathbb{R})$. Ils sont en somme directe et $\mathrm{M}_n(\mathbb{R}) = S_n \oplus AS_n$. La dimension de S_n est $\frac{n(n+1)}{2}$ et celle de AS_n est $\frac{n(n-1)}{2}$.*

Excepté pour les dimensions, c'est une conséquence directe du cas des matrices Hermitiennes. Pour les dimensions, le plus simple est de décrire une base de ces espaces vectoriels. Notons que contrairement à H_n et AH_n, les espaces vectoriels S_n et AS_n ne sont pas isomorphes.

4.3.2 Réduction des matrices réelles

Théorème 4.3.2. *Soit $M \in S_n$. Il existe une matrice orthogonale $P \in O(n)$ et une matrice diagonale réelle D, telles que $M = {}^tPDP = P^{-1}DP$.*

Démonstration. La preuve est une copie parfaite de celle pour les matrices Hermitiennes, en remplaçant le produit scalaire Hermitien de $\mathbb{C}^n$ par le produit scalaire Euclidien de $\mathbb{R}^n$. □

Attention, on ne peut pas déduire directement l'énoncé précédent du cas des matrices Hermitiennes : si on écrit $M = U^*DU$ avec $U \in U(n)$, on peut très bien avoir $U \in U(n) \setminus O(n)$. Par exemple, en dimension 1, on a toujours $e^{-i\theta} \cdot 1 \cdot e^{i\theta} = 1$.

4.3.3 Racines carrées des matrices symétriques définies positives

Théorème 4.3.3. *Soit $M \in S_n^+$. Alors il existe une unique matrice symétrique réelle positive P telle que $P^2 = M$. De plus, si $M \in S_n^{++}$, alors $P \in S_n^{++}$.*

Démonstration. Puisque $M \in S_n^+ \subset H_n^+$, il existe une unique matrice $P \in H_n^+$ telle que $P^2 = M$. Considérons $\bar{P}$. On a $(\bar{P})^* = \overline{P^*} = \bar{P}$, c'est-à-dire $\bar{P} \in H_n$. Puisque les valeurs propre de P sont réelles, elles sont égales à celles de $\bar{P}$, donc $\bar{P} \in H_n^+$. On a enfin $\bar{P}^2 = \bar{M} = M$, donc par unicité $\bar{P} = P$ et $P \in H_n^+ \cap \mathrm{M}_n(\mathbb{R}) = S_n^+$. □

4.3.4 Décomposition polaire réelle

Théorème 4.3.4. *L'application $\phi : O(n) \times S_n^{++} \to \mathrm{GL}_n(\mathbb{R})$ définie par $\phi(U, P) = UP$ est un homéomorphisme.*

Démonstration. Par définition, on est en train de considérer la restriction à $O(n) \times S_n^{++}$ de l'homéomorphisme donné par la décomposition polaire complexe. Il s'agit d'un homéomorphisme sur son image, mais il faut vérifier que cette image est exactement $\mathrm{GL}_n(\mathbb{R})$. L'inclusion $\phi(O(n) \times S_n^{++}) \subset \mathrm{GL}_n(\mathbb{R})$ est évidente. L'autre inclusion est encore une conséquence de la décomposition polaire complexe : soit $A \in \mathrm{GL}_n(\mathbb{R}) \subset \mathrm{GL}_n(\mathbb{C})$, alors $P := \sqrt{A^*A} \in S_n^{++}$ car $A^*A = {}^tAA \in S_n^{++}$, et $U = AP^{-1} \in U(n) \cap \mathrm{GL}_n(\mathbb{R}) = O(n)$, donc $A \in \phi(O(n) \times S_n^{++})$. □

4.4 Exponentielle de matrices

4.4.1 Rappels (définition et principales propriétés)

Soit $A \in \mathrm{M}_n(\mathbb{C})$. On considère la série à valeur dans $\mathrm{M}_n(\mathbb{C})$ de terme général $\frac{A^k}{k!}$, et on note sa somme $\exp(A)$ si la série converge.

Proposition 4.4.1. *La série converge pour tout choix de $A \in \mathrm{M}_n(\mathbb{C})$. De plus, en restriction à n'importe quel compact de $\mathrm{M}_n(\mathbb{C})$, la fonction* exp *ainsi définie est limite uniforme des fonctions $A \mapsto \sum_{k=0}^m \frac{A^k}{k!}$.*

On montrera ces propriétés en comparant avec le cas de l'exponentielle réelle.

Démonstration. Fixons une norme sous-multiplicative $|||\cdot|||$ sur $\mathrm{M}_n(\mathbb{C})$, par exemple la norme associée à la norme Hermitienne sur $\mathbb{C}^n$. La propriété de sous-multiplicativité signifie que pour tous A, $B \in \mathrm{M}_n(\mathbb{C})$,

$$|||AB||| \leq |||A|||\,|||B|||.$$

Montrons d'abord que la série converge. On a, pour tout m, $l \in \mathbb{N}$,

$$\left|\left|\left| \sum_{k=m}^{l} \frac{A^k}{k!} \right|\right|\right| \leq \sum_{k=m}^{l} \frac{|||A|||^k}{k!}$$

par inégalité triangulaire et sous-multiplicativité. Puisque la série de terme général $\frac{|||A|||^k}{k!}$ converge vers $\exp(|||A|||)$, la suite de ses sommes partielles est de Cauchy. L'inégalité ci-dessus montre donc que la suite des sommes partielles de terme général $\frac{A^k}{k!}$ est de Cauchy, donc convergente car un espace vectoriel de dimension finie est complet.

Pour la convergence uniforme, plaçons nous sur un compact $K \subset \mathrm{M}_n(\mathbb{C})$. Un tel ensemble est borné. Soit $M \in \mathbb{R}$ tel que $|||A||| \leq M$ pour tout $A \in K$. On a, pour les mêmes raisons que précédemment,

$$\begin{aligned}\left|\left|\left|\exp(A) - \sum_{k=0}^{m} \frac{A^k}{k!}\right|\right|\right| &\leq \sum_{k=m+1}^{\infty} \frac{|||A|||^k}{k!} \\ &\leq \sum_{k=m+1}^{\infty} \frac{M^k}{k!}.\end{aligned}$$

La dernière expression converge vers zéro uniformément en $A \in K$. □

Corollaire 4.4.2. *La fonction* $\exp : \mathrm{M}_n(\mathbb{C}) \to \mathrm{M}_n(\mathbb{C})$ *est continue.*

Démonstration. Une limite uniforme de fonctions continues est continue. Or les fonctions $A \mapsto \sum_{k=0}^{m} \frac{A^k}{k!}$ sont continues, donc exp est continue sur tout compact de $\mathrm{M}_n(\mathbb{C})$. Cela suffit pour conclure, car $\mathrm{M}_n(\mathbb{C})$ peut être recouvert par des compacts (par exemple les boules fermées centrées en zéro). □

Proposition 4.4.3. *Si* A *et* B *commutent, alors* $\exp(A+B) = \exp(A)\exp(B)$.

Attention, c'est faux en général.

Démonstration. On considère le produit de Cauchy des deux séries définissant $\exp(A)$ et $\exp(B)$. Sa somme partielle est

$$\begin{aligned}U_m &:= \left(\sum_{k=0}^{m} \frac{A^k}{k!}\right)\left(\sum_{l=0}^{m} \frac{B^l}{l!}\right) \\ &= \sum_{k,l=0}^{m} \frac{A^k B^l}{k!l!}.\end{aligned}$$

La somme partielle de la série définissant $\exp(A+B)$ est

$$\begin{aligned}V_m &:= \sum_{q=0}^{m} \frac{(A+B)^q}{q!} \\ &= \sum_{q=0}^{m}\sum_{p=0}^{q} \frac{1}{q!}\binom{q}{p} A^p B^{q-p} \\ &= \sum_{0 \leq k,l,k+l \leq m} \frac{A^k B^l}{k!l!}.\end{aligned}$$

On a

$$\begin{aligned}|||U_m - V_m||| &= \left|\left|\left| \sum_{k,l \geq 0, k+l > m} \frac{A^k B^l}{k!l!} \right|\right|\right| \\ &\leq \sum_{k,l \geq 0, k+l > m} \frac{|||A|||^k |||B|||^l}{k!l!} \\ &\leq u_m - v_m\end{aligned}$$

où u_m est la somme partielle du produit de Cauchy des séries définissant $\exp(|||A|||)$ et $\exp(|||B|||)$ et v_m est la somme partielle de la série définissant $\exp(|||A|||+|||B|||)$. Puisque $\exp(|||A|||+|||B|||) = \exp(|||A|||)\exp(|||B|||)$ (et la convergence est absolue), $u_m - v_m$ converge vers zéro.

On en déduit que U_m et V_m ont même limite. La limite de U_m est le produit des exponentielles car ces séries convergent normalement. □

Proposition 4.4.4. *L'application* exp *prends ses valeurs dans* $\mathrm{GL}_n(\mathbb{C})$. *Plus précisément, pour* $A \in \mathrm{M}_n(\mathbb{C})$, $\exp(A)$ *est inversible et son inverse est* $\exp(-A)$.

Démonstration. On commence par noter le cas particulier évident : $\exp(0) = I_n$. Soit $A \in \mathrm{M}_n(\mathbb{C})$. Alors puisque A et $-A$ commutent, on a $\exp(A)\exp(-A) = \exp(A-A) = \exp(0) = I_n$. □

Proposition 4.4.5. *Pour tout* $A \in \mathrm{M}_n(\mathbb{C})$, *on a* $\overline{\exp(A)} = \exp(\bar{A})$, ${}^t\exp(A) = \exp({}^tA)$ *et* $\exp(A)^* = \exp(A^*)$.

Démonstration. La preuve est la même pour les trois opérations. Faisons-le pour l'adjoint. Par définition, les sommes partielles $\sum_{k=0}^{m} \frac{(A^*)^k}{k!}$ convergent vers $\exp(A^*)$. Mais par linéarité de l'adjoint, et par son comportement sur les produits, la somme partielle est égale à $\left(\sum_{k=0}^{m} \frac{A^k}{k!}\right)^*$, et ceci converge vers $\exp(A)^*$ par continuité de l'adjoint. □

Proposition 4.4.6. *Soit* $A \in \mathrm{M}_n(\mathbb{C})$ *et* $P \in \mathrm{GL}_n(\mathbb{C})$. *Alors* $\exp(PAP^{-1}) = P\exp(A)P^{-1}$.

Exercice 4.4.7. Faire la preuve, c'est exactement le même principe que pour la proposition précédente.

4.4.2 Exponentielle et décomposition polaire

Théorème 4.4.8. *On a* $\exp(H_n) = H_n^{++}$.

Démonstration. Montrons d'abord que $\exp(H_n) \subset H_n^{++}$. Soit $A \in H_n$. On a $\exp(A)^* = \exp(A^*) = \exp(A)$ donc $\exp(A) \in H_n$. Soit $Z \in \mathbb{C}^n \setminus \{0\}$. Alors

$$\begin{aligned} Z^*\exp(A)Z &= Z^*\exp(A/2)\exp(A/2)Z \\ &= Z^*\exp(A^*/2)\exp(A/2)Z \\ &= \|\exp(A/2)Z\|^2 \\ &> 0 \end{aligned}$$

(notons que $\exp(A/2)Z \neq 0$ car $\exp(A/2)$ est inversible). Donc $\exp(A) \in H_n^{++}$.

Montrons maintenant que $H_n^{++} \subset \exp(H_n)$. Soit $B \in H_n^{++}$. Par réduction des matrices Hermitiennes, on peut écrire $B = U^*\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)U$ avec $U \in \mathrm{U}(n)$ et $\lambda_j \in \mathbb{R}$. De plus, puisque B est définie positive, les λ_j sont strictement positifs. Soit $A = U^*\,\mathrm{diag}(\ln\lambda_1, \ldots, \ln\lambda_n)U$. On a $A \in H_n$ et $\exp(A) = B$. □

Le vrai théorème est plutôt le suivant, dont on donnera des éléments de preuve en TD.

Théorème 4.4.9. *L'application* $\exp : H_n \to H_n^{++}$ *est un homéomorphisme.*

Remarque 4.4.10. On a démontré que exp est continue, et envoie H_n surjectivement sur H_n^{++}. Il reste à montrer que exp est injective, puis que son inverse $\exp^{-1} : H_n^{++} \to H_n$ est continue.

Corollaire 4.4.11. *Le groupe topologique* $\mathrm{GL}_n(\mathbb{C})$ *est homéomorphe à* $\mathrm{U}(n)\times H_n \simeq \mathrm{U}(n)\times\mathbb{R}^{n^2}$.

On en déduit par exemple une autre preuve que $\mathrm{GL}_n(\mathbb{C})$ est connexe par arcs : c'est une conséquence du corollaire précédent, de la connexité par arcs de $\mathbb{R}^{n^2}$ et de la connexité par arcs de $\mathrm{U}(n)$.

Théorème 4.4.12. *On a* $\exp(AH_n) = \mathrm{U}(n)$.

Remarque 4.4.13. Ici, $\exp : AH_n \to \mathrm{U}(n)$ n'est pas un homéomorphisme. Cette application n'est par exemple pas injective. En effet, déjà en rang un, $\exp : i\mathbb{R} \to S^1$ a pour noyau le sous-groupe $2\pi i\mathbb{Z}$.

Démonstration. Pour la première inclusion, si $A \in AH_n$, alors $\exp(A)^* = \exp(A^*) = \exp(-A) = \exp(A)^{-1}$ donc $\exp(A) \in \mathrm{U}(n)$.

Pour l'autre inclusion, on utilise la réduction des matrices unitaires. Soit $B \in \mathrm{U}(n)$, alors $B = U^* \operatorname{diag}(e^{i\theta_1}, \ldots, e^{i\theta_n})U$ avec $U \in \mathrm{U}(n)$. Posons $A = U^* \operatorname{diag}(i\theta_1, \ldots, i\theta_n)U$, alors on a $A \in AH_n$ et $\exp(A) = B$. □

Chapitre 5

Représentations linéaires des groupes finis

5.1 Représentations linéaires

5.1.1 Définition

Soit G un groupe et $\mathbb{K}$ un corps (commutatif).

Définition 5.1.1.
- Une *représentation linéaire de G à coefficients dans $\mathbb{K}$* est la donnée :
  - d'un $\mathbb{K}$-espace vectoriel V
  - et d'un morphisme de groupe $\rho : G \to \mathrm{GL}(V)$.
- Deux représentations linéaires (V_1, ρ_1) et (V_2, ρ_2) de G sont dites *équivalentes* s'il existe un isomorphisme d'espace vectoriels $\phi : V_1 \to V_2$ tel que, pour tout $g \in G$,

$$\rho_2(g) \circ \phi = \phi \circ \rho_1(g)$$

- La dimension de V est appelé le *degré* de la représentation.

Dans ce cours, on se concentrera sur le cas $\mathbb{K} = \mathbb{C}$, et on omettra souvent les termes *linéaire* et *à coefficients dans* $\mathbb{C}$ pour simplement dire représentation de G.

Remarque 5.1.2.
- Un morphisme $\rho : G \to \mathrm{GL}_n(\mathbb{C})$ peut être interprété comme une représentation linéaire complexe de G, de degré n, où l'espace vectoriel sous-jacent est $\mathbb{C}^n$.
- Réciproquement, si V est de dimension finie, comme on le supposera toujours dans ce cours, le choix d'une base identifie une représentation à un morphisme dans $\mathrm{GL}_n(\mathbb{C})$.
- Soient $\rho_1 : G \to \mathrm{GL}_n(\mathbb{C})$ et $\rho_2 : G \to \mathrm{GL}_n(\mathbb{C})$ deux morphismes. Alors les représentations associées sont équivalentes si elles sont conjuguées ,c'est-à-dire qu'il existe $h \in \mathrm{GL}_n(\mathbb{C})$ tel que pour tout $g \in Gh\rho_1(g)h^{-1} = \rho_2(g)$.

Exemple 5.1.3.
1. Pour tout groupe G et tout espace vectoriel V, on peut toujours considérer la représentation triviale $\rho : G \to \mathrm{GL}(V)$ dont l'image est le singleton $\{\mathrm{Id}_V\}$.
2. Pour $G = \mathbb{Z}/2\mathbb{Z}$, on construit facilement des exemples : $\rho_1 : G \to \mathrm{GL}_2(\mathbb{C}), \bar{k} \mapsto (-1)^k I_2$, $\rho_2 : G \to \mathrm{GL}_2(\mathbb{C}), \bar{k} \mapsto \begin{pmatrix} (-1)^k & 0 \\ 0 & 1 \end{pmatrix}$ ou $\rho_3 : G \to \mathrm{GL}_2(\mathbb{C}), \bar{k} \mapsto \begin{pmatrix} 1 & 0 \\ 0 & (-1)^k \end{pmatrix}$. Les représentations ρ_2 et ρ_3 sont équivalentes, mais pas ρ_2 et ρ_1.
3. Plus généralement, toujours pour $G = \mathbb{Z}/2\mathbb{Z}$, la donnée d'un morphisme $G \to \mathrm{GL}_n(\mathbb{C})$ est équivalente à la donnée d'une matrice $n \times n$ d'ordre 2 (l'image de $\bar{1}$).

4. Pour $G = \mathfrak{S}_n$, on a déjà rencontré un morphisme remarquable $\rho : \mathfrak{S}_n \to \mathrm{GL}_n(\mathbb{C})$ donné par l'isomorphisme entre $\mathfrak{S}_n$ et le groupe des matrices de permutations.
5. Plus généralement, si G agit sur un ensemble fini $\{1, \dots, n\}$, on peut construire une représentation de G dans $\mathrm{GL}_n(\mathbb{C})$ comme suit : si $(e_1, \dots, e_n)$ est la base standard de $\mathbb{C}^n$ on pose
$$\rho(g)\left(\sum_{i=1}^{n} z_i e_i\right) = \sum_{i=1}^{n} z_i e_{g\cdot i}.$$
Ceci définit une représentation et on appelle les représentations obtenues de cette façon des *représentations de permutations*.
6. Un cas particulier important est donné par l'action d'un groupe fini sur lui-même par multiplication à gauche. Ceci définit une représentation de degré $\#G$ de G appelée la *représentation régulière* de G.
7. Pour $G = \mathbb{Z}/2\mathbb{Z}$ par exemple, la représentation régulière est donnée par $\rho(\bar{1}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

5.1.2 Constructions élémentaires

5.1.2.1 Sous-représentations

Soit $\rho : G \to \mathrm{GL}(V)$ une représentation de G. Si $W \subset V$ est un sous-espace vectoriel de V stable par tous les $\rho(g)$ pour $g \in G$, alors on obtient une représentation $\rho|_W : G \to \mathrm{GL}(W), g \mapsto \rho(g)|_W$. Une telle représentation est appelée une *sous-représentation* de ρ.

Exemple 5.1.4. Considérons la représentation régulière de $\mathbb{Z}/2\mathbb{Z}$ donnée dans le dernier exemple précédent. Alors la droite $\mathbb{C}\binom{1}{1}$ est un sous-espace vectoriel de $\mathbb{C}^2$, stable par I_2 et $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, donc elle définit une sous-représentation de $\mathbb{Z}/2\mathbb{Z}$ de degré 1, qui est par ailleurs équivalente à la représentation triviale de degré 1.

5.1.2.2 Somme directe

Soient $\rho_1 : G \to \mathrm{GL}(V_1)$ et $\rho_2 : G \to \mathrm{GL}(V_2)$ deux représentations de G. La somme directe $\rho_1 \oplus \rho_2$ est la représentation linéaire $\rho_1 \oplus \rho_2 : G \to \mathrm{GL}(V_1 \oplus V_2)$ définie par $(\rho_1 \oplus \rho_2)(g)(v_1, v_2) = (\rho_1(g)(v_1), \rho_2(g)(v_2))$.

De la même manière, on peut définir la somme directe d'un nombre arbitraire de représentations.

En pratique, si $\rho_1 : G \to \mathrm{GL}_{n_1}(\mathbb{C})$ et $\rho_2 : G \to \mathrm{GL}_{n_2}(\mathbb{C})$ sont données par des matrices, la somme directe est définie par des matrices diagonales par blocs : $\rho_1 \oplus \rho_2 : G \to \mathrm{GL}_{n_1+n_2}(\mathbb{C})$, $(\rho_1 \oplus \rho_2)(g) = \begin{pmatrix} \rho_1(g) & 0 \\ 0 & \rho_2(g) \end{pmatrix}$.

Exemple 5.1.5.
— La représentation ρ de $\mathbb{Z}/2\mathbb{Z}$ définie par $\rho(\bar{1}) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ est la somme directe de $\rho_1 : \mathbb{Z}/2\mathbb{Z} \to \mathrm{GL}_1(\mathbb{C}) = \mathbb{C}^*$ et $\rho_2 : \mathbb{Z}/2\mathbb{Z} \to \mathbb{C}^*$ avec $\rho_1(\bar{1}) = -1$ et $\rho_2(\bar{1}) = 1$.
— Pour tout G, la représentation triviale de degré n de G est la somme directe de n copies de la représentation triviale de degré 1 de G.

Remarque 5.1.6.
— Si $\rho = \rho_1 \oplus \rho_2$, alors ρ_1 et ρ_2 sont des sous-représentations de ρ.
— L'un des buts de ce chapitre est de comprendre comment décomposer de manière optimale une représentation en somme directe de sous-représentations.

5.1.3 Représentations irréductibles

Définition 5.1.7. Une représentation $\rho : G \to \mathrm{GL}(V)$ est dite *irréductible* si ses seuls sous-espaces vectoriels stables sont $\{0\}$ et V.

Exemple 5.1.8. (Immédiat mais fondamental) Toute représentation linéaire de degré 1 est irréductible.

Exercice 5.1.9. Soit $n \in \mathbb{N}^*$ et $G = \mathbb{Z}/n\mathbb{Z}$. Déterminer tous les morphismes $\rho : G \to \mathbb{C}^*$.

Solution : Un tel morphisme est uniquement déterminé par l'image d'un générateur de G. Celle-ci doit être une racine n-ième de l'unité dans $\mathbb{C}$, et n'importe quel choix de racine n-ième de l'unité fournit un tel morphisme.

Chacun de ces morphismes donne un exemple de représentation irréductible de $\mathbb{Z}/n\mathbb{Z}$. En fait, ça les donne tous, comme le montre la proposition suivante.

Proposition 5.1.10. *Soit G un groupe fini abélien. Une représentation linéaire complexe $\rho : G \to \mathrm{GL}_n(\mathbb{C})$ est irréductible si et seulement si $n = 1$.*

La preuve de cette proposition repose sur le lemme suivant.

Lemme 5.1.11. *Soit G un groupe fini et $\rho : G \to \mathrm{GL}_n(\mathbb{C})$ une représentation de G. Alors pour tout $g \in G$, la matrice $\rho(g)$ est diagonalisable.*

Démonstration. Soit $g \in G$, et soit m l'ordre de g (qui est fini car G est fini). Comme ρ est un morphisme, on a $\rho(g)^m = I_n$. Donc le polynôme minimal de la matrice $\rho(g)$ divise le polynôme $X^m - 1$. Ce dernier est scindé à racines simples dans $\mathbb{C}$, donc le polynôme minimal de $\rho(g)$ l'est également, et $\rho(g)$ est diagonalisable. □

Preuve de la proposition. Par le lemme, chaque $\rho(g)$ est diagonalisable, pour $g \in G$. Puisque G est abélien, toutes ces matrices commutent. On peut donc co-diagonaliser la famille de matrices $\rho(G)$, dans une base de vecteurs propres communs $(e_1, \ldots, e_n)$. Chacune des droites $\mathbb{C}e_i$ est stable par $\rho(G)$, donc ρ est irréductible si et seulement si $n = 1$. □

Remarque 5.1.12. Attention, c'est faux sur un corps quelconque. Par exemple soit $\rho : \mathbb{Z}/3\mathbb{Z} \to \mathrm{GL}_2(\mathbb{R})$ le morphisme de groupe défini par $\rho(\bar{1}) = R_{2\pi/3}$ (la matrice de la rotation d'angle $2\pi/3$). Ce morphisme définit une représentation réelle de degré 2, qui est irréductible comme représentation linéaire réelle : la rotation d'angle $2\pi/3$ ne laisse stable que les sous-espaces vectoriels $\{0\}$ et $\mathbb{R}^2$ de $\mathbb{R}^2$.

Théorème 5.1.13 (Lemme de Schur)**.** *Soient $\rho_1 : G \to \mathrm{GL}(V_1)$ et $\rho_2 : G \to \mathrm{GL}(V_2)$ deux représentations linéaires complexes de dimension finie (comme toujours dans ce cours). Soit $\phi : V_1 \to V_2$ une application linéaire non identiquement nulle, telle que $\phi \circ \rho_1(g) = \rho_2(g) \circ \phi$ pour tout $g \in G$. Alors*

1. *si V_1 et V_2 sont irréductibles, alors ϕ est un isomorphisme,*
2. *si $V_1 = V_2$, alors il existe $\lambda \in \mathbb{C}^*$ tel que $\phi = \lambda Id_{V_1}$.*

Démonstration.

1. Le noyau de ϕ est un sous-espace de V_1 qui est stable par $\rho_1(G)$ par l'hypothèse d'équivariance de ϕ. Il n'est égal à V_1 car ϕ n'est pas identiquement nulle. Comme ρ_1 est irréductible, on en déduit que $\ker(\phi) = \{0\}$. De même, l'image de ϕ est un sous-espace de V_2 stable par $\rho_2(G)$, différent de $\{0\}$. Comme V_2 est irréductible, on en déduit $\mathrm{Im}(\phi) = V_2$ et donc ϕ est un isomorphisme entre V_1 et V_2.
2. Soit $\lambda \in \mathbb{C}$. En appliquant le premier point à l'application $\phi - \lambda \mathrm{Id}_{V_1}$, on a : soit $\phi - \lambda \mathrm{Id}_{V_1}$ est un isomorphisme, soit $\phi = \lambda \mathrm{Id}_{V_1}$. Puisque ϕ admet au moins une valeur propre, il existe un $\lambda \in \mathbb{C}$ tel que $\phi - \lambda \mathrm{Id}_{V_1}$ n'est *pas* un isomorphisme, d'où la conclusion.

□

5.1.4 Décomposition en représentations irréductibles

Théorème 5.1.14. *Soit G un groupe fini. Alors toute représentation linéaire complexe de degré fini de G est une somme directe d'un nombre fini de sous-représentations irréductibles.*

Remarque 5.1.15. On l'a déjà démontré pour les groupes finis abéliens.

Démonstration. Il suffit d'appliquer récursivement le résultat suivant. □

Théorème 5.1.16 (Lemme de Maschke)**.** *Soit $\rho : G \to \mathrm{GL}(W)$ une représentation de degré fini du groupe fini G, et V_1 un sous-espace $\rho(G)$-stable non trivial de W (ni $\{0\}$, ni W). Alors il existe un sous-espace $\rho(G)$-stable V_2 de W tel que $V_1 \oplus V_2 = W$ et $\rho = \rho|_{V_1} \oplus \rho|_{V_2}$.*

Démonstration. On choisit une base quelconque de W pour identifier ρ avec un morphisme $\rho : G \to \mathrm{GL}_n(\mathbb{C})$ où $n = \dim(W)$. On considère la matrice

$$H := \sum_{g \in G} \rho(g)^* \rho(g).$$

C'est un élément de H_n^{++} (pour un élément de la somme, on l'a déjà démontré, et la somme reste un élément de H_n^{++} car cet ensemble est un cône convexe). L'espace $W \simeq \mathbb{C}^n$ est donc équipé du produit scalaire Hermitien défini par $(X, Y) \mapsto X^* H Y$. Ce produit scalaire Hermitien est invariant par $\rho(G)$: pour $h \in G$, X, $Y \in \mathbb{C}^n$, on a

$$\begin{aligned}(\rho(h)X)^* H(\rho(h)Y) &= \sum_{g \in G} X^* \rho(h)^* \rho(g)^* \rho(g)\rho(h) Y \\ &= \sum_{g \in G} X^* \rho(gh)^* \rho(gh) Y \\ &= X^* H Y\end{aligned}$$

Donc si V_1 est un sous-espace de $\mathbb{C}^n = W$ stable par $\rho(G)$, alors son orthogonal par rapport à $(X, Y) \mapsto X^* H Y$ est aussi stable par $\rho(G)$. Si on le note V_2, on a bien $W = V_1 \oplus V_2$ et $\rho = \rho|_{V_1} \oplus \rho|_{V_2}$. □

Attention, le résultat est faux en général pour les groupes infinis. Par exemple, l'application

$$\rho : \mathbb{Z} \to \mathrm{GL}_2(\mathbb{C}), k \mapsto \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

est un morphisme de groupe, donc définit une représentation linéaire de degré 2 de $\mathbb{Z}$. La droite $V_1 = \mathbb{C}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ est un sous-espace stable $\rho(\mathbb{Z})$ de $\mathbb{C}^2$ (triviale), mais aucun sous-espace vectoriel de V_1 n'est stable par $\rho(\mathbb{Z})$: pour $\mathbb{C}\begin{pmatrix} z \\ 1 \end{pmatrix}$, $z \in \mathbb{C}$, on a

$$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} z \\ 1 \end{pmatrix} = \begin{pmatrix} z+k \\ 1 \end{pmatrix}$$

et $\begin{pmatrix} z+k \\ 1 \end{pmatrix} \in \mathbb{C}\begin{pmatrix} z+k \\ 1 \end{pmatrix}$ si et seulement si $k = 0$.

5.2 Théorie des caractères

5.2.1 Définition, premières propriétés

Définition 5.2.1. Soit $\rho : G \to \mathrm{GL}(V)$ une représentation de G. On appelle *caractère* de ρ la fonction

$$\chi_\rho : G \to \mathbb{C}, g \mapsto \mathrm{tr}(\rho(g))$$

Exemple 5.2.2.

1. Si $\rho : G \to \mathrm{GL}_1(\mathbb{C}) = \mathbb{C}^*$ est de degré 1, alors $\chi_\rho = \rho$.
2. Pour $\rho : \mathbb{Z}/2\mathbb{Z} \to \mathrm{GL}_2(\mathbb{C}), \bar{k} \mapsto (-1)^k I_2$ on a $\chi_\rho(\bar{0}) = 2$ et $\chi_\rho(\bar{1}) = -2$.
3. Pour la représentation régulière de $\mathbb{Z}/2\mathbb{Z}$, on a $\chi_\rho(\bar{0}) = 2$ et $\chi_\rho(\bar{1}) = 0$.
4. L'image de l'élément neutre par χ_ρ est toujours égale au degré de la représentation ρ.
5. Si G agit sur $\{1, \dots, n\}$ et $\rho : G \to \mathrm{GL}_n(\mathbb{C})$ est la représentation de permutation associée, on a $\chi_\rho(g) =$ nombre de points fixes par g dans $\{1, \dots, n\}$.
6. En particulier, pour la représentation régulière de G, on a $\chi_\rho(g) = 0$ pour tout g différent de l'élément neutre.

Proposition 5.2.3. *Soit $\rho : G \to \mathrm{GL}(V)$ une représentation, et χ_ρ son caractère. La fonction $\chi_\rho : G \to \mathbb{C}$ est invariante par conjugaison :*

$$\forall g, h \in G, \chi_\rho(ghg^{-1}) = \chi_\rho(h).$$

Démonstration. C'est une conséquence directe de l'invariance de la trace par conjugaison. □

Remarque 5.2.4.

1. On appelle parfois *fonctions centrales* les fonctions $G \to \mathbb{C}$ invariantes par conjugaison.
2. Pour déterminer χ_ρ, si on connait les classes de conjugaison dans G, il suffit de déterminer ses valeurs en un représentant de chaque classe de conjugaison.

Corollaire 5.2.5. *Deux représentations équivalentes ont le même caractère.*

Notre but pour la suite du cours est de déterminer la décomposition en représentations irréductibles d'une représentation en fonction de son caractère. En particulier, on reconnaitra les représentations irréductibles par leurs caractères.

Voici un premier (petit) pas dans cette direction.

Proposition 5.2.6. *Soient $\rho_1 : G \to \mathrm{GL}_{n_1}(\mathbb{C})$ et $\rho_2 : G \to \mathrm{GL}_{n_2}(\mathbb{C})$ deux représentations. Alors $\chi_{\rho_1 \oplus \rho_2} = \chi_{\rho_1} + \chi_{\rho_2}$.*

Démonstration. Cela découle directement de l'interprétation de $\rho_1 \oplus \rho_2$ en termes de matrices diagonales par blocs :

$$\chi_{\rho_1 \oplus \rho_2}(g) = \mathrm{tr} \begin{pmatrix} \rho_1(g) & 0 \\ 0 & \rho_2(g) \end{pmatrix} = \mathrm{tr}(\rho_1(g)) + \mathrm{tr}(\rho_2(g)) = \chi_{\rho_1}(g) + \chi_{\rho_2}(g).$$

□

Une autre propriété sera utile dans la suite.

Proposition 5.2.7. *Soit ρ une représentation de G et χ_ρ son caractère. Alors pour tout $g \in G$,*

$$\chi_\rho(g^{-1}) = \overline{\chi_\rho(g)}$$

Démonstration. Soit $g \in G$. Notons $\lambda_1, \ldots, \lambda_n$ les valeurs propres de $\rho(g)$, répétées selon leur multiplicité. Alors $\chi_\rho(g) = \lambda_1 \cdots \lambda_n$. Comme G est fini, g est d'ordre fini, disons m. Puisque ρ est un morphisme de groupe, on a $\rho(g)^m = \mathrm{Id}_V$, donc le polynôme minimal de $\rho(g)$ divise $X^m - 1$. En particulier, chacune des valeurs propres est un nombre complexe de module 1. On a donc

$$\begin{aligned}\chi_\rho(g^{-1}) &= \lambda_1^{-1} + \cdots + \lambda_n^{-1} \\ &= \overline{\lambda_1} + \cdots + \overline{\lambda_n} \\ &= \overline{\chi_\rho(g)}\end{aligned}$$

□

5.2.2 Produit scalaire Hermitien de caractères

On munit l'espace vectoriel complexe des fonctions complexes $G \to \mathbb{C}$ du produit scalaire Hermitien :

$$\big(f \mid f'\big) := \frac{1}{\#G} \sum_{g \in G} f(g)\overline{f'(g)} \qquad \text{pour } f,\ f' : G \to \mathbb{C}$$

On va s'intéresser aux produits scalaires impliquant des caractères de représentations. Effectuons un premier calcul de produit scalaire, dont on verra les conséquences en termes de représentations dans les sections suivantes.

Théorème 5.2.8. *Soit $\rho : G \to \mathrm{GL}(V)$ une représentation, et χ_ρ son caractère. Soit $\mathbf{1} : G \to \mathbb{C}$ la fonction constante égale à 1. On note enfin V^G le sous-espace vectoriel de V formé par les vecteurs invariants par tous les $\rho(g)$ pour $g \in G$. Alors*

$$(\chi_\rho \mid \mathbf{1}) = \dim(V^G).$$

Démonstration. Soit $P : V \to V$ l'application linéaire définie par

$$P := \frac{1}{\#G} \sum_{g \in G} \rho(g).$$

On a

$$\begin{aligned}(\chi_\rho \mid \mathbf{1}) &= \frac{1}{\#G} \sum_{g \in G} \mathrm{tr}(\rho(g)) \\ &= \mathrm{tr}(P)\end{aligned}$$

On va maintenant déterminer la trace de P. Pour cela on calcule

$$\begin{aligned}P \circ P &= \frac{1}{\#G} \sum_{g \in G} \frac{1}{\#G} \sum_{h \in G} \rho(g)\rho(h) \\ &= \frac{1}{\#G} \sum_{g \in G} \frac{1}{\#G} \sum_{h \in G} \rho(gh) \\ &= \frac{1}{\#G} \sum_{g \in G} P\end{aligned}$$

car $g \mapsto gh$ est une bijection de G pour tout $h \in G$

$$P \circ P = P.$$

Ce calcul montre que P est un projecteur, c'est-à-dire que $\ker(P) \oplus \mathrm{Im}(P) = V$ et $P(x+y) = y$ pour $x \in \ker(P)$, $y \in \mathrm{Im}(P)$. Pour un projecteur, la trace est égale à la dimension de l'image. Il reste donc à montrer que $\mathrm{Im}(P) = V^G$.

Pour la première inclusion, si $y \in V^G$, on a

$$\begin{aligned} P(y) &= \frac{1}{\#G} \sum_{g \in G} \rho(g)(y) \\ &= \frac{1}{\#G} \sum_{g \in G} y \\ &= y \end{aligned}$$

donc $y \in \mathrm{Im}(P)$. Pour la seconde inclusion, si $y \in \mathrm{Im}(P)$, on choisit $\tilde{y} \in V$ tel que $P(\tilde{y}) = y$. Alors pour $h \in G$, on a

$$\begin{aligned} \rho(h)(y) &= \frac{1}{\#G} \sum_{g \in G} \rho(hg)(\tilde{y}) \\ &= P(\tilde{y}) \\ &= y. \end{aligned}$$

□

Remarque 5.2.9. La fonction **1** est le caractère de la représentation triviale de degré 1 de G.

5.2.3 Digression : représentations sur $L(V, W)$

Soient V et W deux espaces vectoriels complexes (de dimension finie toujours). On note $L(V, W)$ l'espace vectoriel formé par les applications linéaires $V \to W$.

Définition 5.2.10. Étant données deux représentations $\rho : G \to \mathrm{GL}(V)$ et $\pi : G \to \mathrm{GL}(W)$, on définit une représentation

$$\pi \otimes \rho^* : G \to \mathrm{GL}(L(V, W))$$

en posant, pour $\phi \in L(V, W)$,

$$\pi \otimes \rho^*(g)(\phi) := \pi(g) \circ \phi \circ \rho(g^{-1}).$$

Théorème 5.2.11. *Le caractère de $\pi \otimes \rho^*$ est donné par :*

$$\chi_{\pi \otimes \rho^*}(g) = \overline{\chi_\rho(g)} \chi_\pi(g)$$

Démonstration. Il s'agit d'un calcul de trace. Pour cela on va fixer une base de $L(V, W)$ issue de bases de V et W. Soit $(e_1, \ldots, e_n)$ une base de V et $(f_1, \ldots, f_m)$ une base de W. Alors les applications linéaires $e_i^* f_j : V \to W$ définies par

$$e_i^* f_j \left(\sum_{k=1}^{n} z_k e_k \right) = z_i f_j$$

pour $(i, j) \in \{1, \ldots, n\} \times \{1, \ldots, m\}$ forment une base de $L(V, W)$.

Pour $g \in G$, considérons l'application $\pi \otimes \rho^*(g)$ et calculons sa trace en utilisant cette base. Il s'agit de calculer la somme pour tout $(i, j) \in \{1, \ldots, n\} \times \{1, \ldots, m\}$ du coefficient de $e_i^* f_j$ dans $\pi \otimes \rho^*(g)(e_i^* f_j)$. Notons par ailleurs $(\rho(g^{-1}))_{k,l}$ les coefficients de la matrice de $\rho(g^{-1})$ dans

la base $(e_1, \ldots, e_n)$, et $(\pi(g))_{k,l}$ les coefficients de la matrice de $\pi(g)$ dans la base $(f_1, \ldots, f_m)$. On a

$$\pi \otimes \rho^*(g)(e_i^* f_j) = \pi(g) \circ e_i^* f_j \circ \rho(g^{-1}),$$

donc

$$\begin{aligned}\pi \otimes \rho^*(g)(e_i^* f_j)(e_i) &= \pi(g) \circ e_i^* f_j \circ \sum_{l=1^n} (\rho(g^{-1}))_{l,i} e_l \\ &= \pi(g)\left((\rho(g^{-1})_{i,i} f_j\right) \qquad = \sum_{k=1}^{m} (\pi(g))_{k,j} (\rho(g^{-1}))_{i,i} f_k\end{aligned}$$

Donc le coefficient de $e_i^* f_j$ dans $\pi \otimes \rho^*(g)(e_i^* f_j)$ est égal à $(\pi(g))_{j,j}(\rho(g^{-1}))_{i,i}$. On a pour conclure

$$\begin{aligned}\chi_{\pi\otimes\rho^*}(g) &= \sum_{i=1}^{n}\sum_{j=1}^{m} (\pi(g))_{j,j}(\rho(g^{-1}))_{i,i} \\ &= \left(\sum_{i=1}^{n} (\rho(g^{-1}))_{i,i}\right)\left(\sum_{j=1}^{m} (\pi(g))_{j,j}\right) \\ &= \chi_\rho(g^{-1})\chi_\pi(g) \\ &= \overline{\chi_\rho(g)}\chi_\pi(g)\end{aligned}$$

□

5.2.4 Résultats principaux de la théorie des caractères

5.2.4.1 Sur les représentations irréductibles

Théorème 5.2.12. *Soient ρ et π deux représentations irréductibles de G. Alors*

1. *si ρ n'est pas équivalente à π, alors*

$$(\chi_\pi \mid \chi_\rho) = 0$$

2. *si ρ est équivalente à π alors $\chi_\pi = \chi_\rho$ et*

$$(\chi_\rho \mid \chi_\rho) = 1.$$

La preuve de ce théorème consiste à mettre ensemble les deux derniers théorèmes démontrés, ainsi que le Lemme de Schur.

Démonstration. On a d'une part

$$(\chi_{\pi\otimes\rho^*} \mid \mathbf{1}) = (\chi_\pi \overline{\chi_\rho} \mid \mathbf{1}) = (\chi_\pi \mid \chi_\rho)$$

d'autre part,

$$(\chi_{\pi\otimes\rho^*} \mid \mathbf{1}) = \dim(L(V,W)^G).$$

Or par le Lemme de Schur, la dimension de $L(V,W)^G$ est égale à 0 si ρ et π ne sont pas équivalentes, et égale à 1 si elles le sont. □

5.2.5 Sur la décomposition en représentations irréductibles

Soit $\rho : G \to \mathrm{GL}(V)$ une représentation de G. Par les résultats de la première partie, on peut trouver des sous-espaces $V_1, \ldots, V_m$ de V, stables par $\rho(G)$, tels que

$$\rho = \rho|_{V_1} \oplus \cdots \oplus \rho|_{V_m}$$

et chacune des sous-représentations $\rho|_{V_i}$ est irréductible.

Soient $\rho_1, \ldots, \rho_k$ des représentations irréductibles non-équivalentes deux à deux de G, telles que pour tout $i \in \{1, \ldots, m\}$, il existe $j \in \{1, \ldots, k\}$ tel que $\rho|_{V_i}$ soit équivalente à ρ_j. Pour tout $j \in \{1, \ldots, k\}$, notons a_j le nombre de $i \in \{1, \ldots, m\}$ tels que $\rho|_{V_i}$ soit équivalente à ρ_j. La représentation ρ est donc équivalente à la représentation

$$\rho_1 \oplus \cdots \oplus \rho_1 \oplus \rho_2 \oplus \cdots \oplus \rho_k$$

où la représentation ρ_1 apparaît exactement a_1 fois, la représentation ρ_2 apparaît a_2 fois, etc. On note

$$\rho \sim \rho_1^{\oplus a_1} \oplus \rho_2^{\oplus a_2} \oplus \cdots \oplus \rho_k^{\oplus a_k}.$$

Théorème 5.2.13. *Soit ρ une représentation de G, et $\rho \sim \rho_1^{\oplus a_1} \oplus \rho_2^{\oplus a_2} \oplus \cdots \oplus \rho_k^{\oplus a_k}$ une décomposition en représentations irréductibles comme ci-dessus. Alors*

1. *Pour tout i, la multiplicité a_i est égale à $(\chi_\rho \mid \chi_{\rho_i})$. En particulier, si deux représentations ont même caractère, alors elles sont équivalentes.*
2. *On a $(\chi_\rho \mid \chi_\rho) = \sum_{i=1}^k a_i^2$. En particulier, ρ est irréductible si et seulement si $(\chi_\rho \mid \chi_\rho) = 1$.*

Démonstration. Maintenant qu'on a rappelé l'existence d'une décomposition en représentations irréductibles (prouvée dans la première partie du chapitre), c'est une conséquence directe du théorème précédent. □

5.2.5.1 Sur la représentation régulière

Théorème 5.2.14. *Toute classe d'équivalence de représentations irréductibles de G a un représentant qui apparaît dans la décomposition en représentations irréductibles de G, avec une multiplicité égale à son degré. En particulier, l'ordre de G est égal à la somme des degré aux carrés des classes d'équivalences de représentations irréductibles de G.*

Démonstration. On a déjà calculé le caractère de la représentation régulière de G. Notons r ce caractère. On rappelle que $r(e) = \#G$ et $r(g) = 0$ pour $g \neq e$.

Pour toute représentation π de G, de degré $\deg(\pi)$, on a

$$(\chi_\pi \mid r) = \frac{r(e)\overline{\chi_\pi(e)}}{\#G} = \deg(\pi)$$

Notons $\rho_1^{\oplus a_1} \oplus \rho_2^{\oplus a_2} \oplus \cdots \oplus \rho_k^{\oplus a_k}$ la décomposition en représentations irréductibles de la représentation régulière comme au paragraphe précédent. Si π est irréductible, on a aussi

$$(\chi_\pi \mid r) = \sum_{i=1}^k a_i \, (\chi_\pi \mid \rho_i)$$

où chaque produit scalaire apparaissant à droite est égal à 0 ou 1, et au plus un d'entre eux est égal à 1. Comme $\deg(\pi) > 0$, exactement un de ces produits scalaires $(\chi_\pi \mid \rho_i)$ est égal à un, π est équivalente ρ_i et le degré de π est égal à a_i. □

Corollaire 5.2.15. *Tout groupe fini n'a qu'un nombre fini de représentations irréductibles (à équivalence près).*

On peut en fait être plus précis : il y a autant de classes d'équivalences de représentations irréductibles de G que de classes de conjugaisons dans G. La preuve de ce résultat fait l'objet du dernier exercice du TD.

5.2.6 Table de caractères

À la vue des résultats principaux de la théorie des caractères, on peut résumer toute l'information sur les représentations d'un groupe fini de la manière suivante. Soit G un groupe fini. On associe à G un tableau, dit *table des caractères* de G, dont les lignes sont indexées par les classes d'équivalences de représentations irréductibles de G (ou de manière équivalente, les caractères de ces représentations) et les colonnes sont indexées par les classes de conjugaison dans G. À la case correspondant à une ligne donnée par un caractère χ et une colonne donnée par une classe de conjugaison $\mathcal{C}$, on place le coefficient $\chi(x)$, où x est un élément quelconque de $\mathcal{C}$. On inclue également le cardinal de chaque classe de conjugaison dans l'en-tête de la colonne correspondante.

Voir en TD les exemples de tables de caractères qu'on a déjà rempli.

Les résultats énoncés dans ce chapitre donnent beaucoup d'informations sur la table des caractères, et permettent notamment de compléter une table de caractères si elle est déjà presque entièrement remplie.

Par exemple, on peut utiliser les propriétés suivantes des tables de caractères :

- Par convention, on commence toujours par mettre le caractère **1** de la représentation triviale de degré 1 sur la première ligne. Tous les coefficients de la première ligne sont donc des 1.
- Par convention, la première colonne correspond toujours à la classe de conjugaison de l'élément neutre, de cardinal 1, et les coefficients de la première colonne correspondent donc aux degrés des représentations irréductibles associées.
- Si le groupe G est abélien, chaque classe de conjugaison est de cardinal 1, chaque représentation irréductible est de degré 1, et les lignes correspondent exactement aux différents morphismes de groupes $G \to \mathbb{C}^*$.
- Les caractères forment une famille **orthonormale** de fonctions centrales. Mais **attention**, si le groupe n'est pas abélien, il ne suffit pas de faire le produit scalaire des vecteurs lignes pour avoir le produit scalaire des caractères correspondants : chaque colonne doit être comptée autant de fois que le cardinal de la classe de conjugaison correspondante, et le résultat final doit être divisé par l'ordre du groupe.
- La somme des coefficients d'une colonne, pondérée par les coefficients de la première colonne doit donner la valeur du caractère de la représentation régulière sur la classe de conjugaison associée à cette colonne.
- En d'autres termes, le produit scalaire du vecteur i-ème colonne avec le vecteur première colonne doit être égal à $\#G$ si $i = 1$, et à 0 sinon.
- Si un caractère prend des valeurs non réelles, alors son conjugué doit apparaître également dans la table de caractères (**Exercice** : pourquoi ?).