## Notes de travail pour l'agrégation

Ces notes sont regroupées en trois parties approximatives : des « développements », des « thèmes » et des « exercices ». Les premiers sont en principe assez rapidement transformables en un développement d'oral ; les seconds sont plutôt des réflexions, regroupements ou approfondissement de points importants du programme (ou voisins du programme) ; les troisièmes sont, comme leur nom l'indique, des exercices. Les thèmes et les exercices peuvent parfois (en fait, très souvent) donner lieu à des développements, après une éventuelle adaptation de forme.

J'insiste très lourdement sur le fait que le contenu de ces notes n'est profitable qu'avec un (très) important travail personnel. En particulier, en dépit de la présence d'un certain nombre de réponses et corrigés aux exercices, il vous sera presque tout le temps beaucoup plus profitable de chercher longuement sur une question, même sans trouver de réponse, et même sans jamais n'avoir la réponse de toute votre vie, que de lire le corrigé trop rapidement. Je vous demande donc avec insistance de bien vouloir, systématiquement, oublier qu'un corrigé est fourni avec l'exercice (lorsque c'est le cas) et de ne vous y reporter éventuellement que pour vérifier que votre réponse est la bonne, ou alors, après avoir séché suffisamment longuement. On peut considérer qu'il est tout à fait raisonnable de sécher pendant plusieurs heures cumulées sur un exercice, et donc, comme la phase de recherche s'étale dans le temps au fur et à mesure de votre emploi du temps, de laisser l'exercice en chantier pendant plusieurs jours ou plusieurs semaines. Pensez aussi à partager vos réflexions et interrogations avec les autres étudiants.

## DÉVELOPPEMENTS

# Contenu et automorphismes de $k(X)$

Dans le premier paragraphe je fais quelques rappels sur le théorème qui affirme que si $A$ est un anneau factoriel, alors $A[X]$ (et par récurrence, aussi $A[X_1, \ldots, X_n]$) est un anneau factoriel. On peut considérer cela comme un résultat bien connu, et l'utiliser sans démonstration pour proposer en développement le calcul du groupe des $k$-automorphismes du corps $k(X)$. C'est l'objet du deuxième paragraphe.

## 1 Factorialité des anneaux de polynômes

La démonstration du théorème qui affirme que *si $A$ est factoriel, alors $A[X]$ est factoriel* utilise de façon cruciale la notion de contenu : pour $P \in A[X]$, son *contenu* noté $c(P)$ est le pgcd de ses coefficients. Comme pour le pgcd, ce contenu est bien défini à multiplication près par un inversible de $A$. La propriété essentielle est connue sous le nom de *lemme de Gauus* :

**Lemme 1 (Gauss) :** *Soit $A$ un anneau factoriel et $P, Q$ deux polynômes à coefficients dans $A$. Alors $c(PQ) = c(P)c(Q)$.*

Si $c(P) = 1$ on dit que $P$ est un polynôme *primitif*. On a l'énoncé précisément :

**Théorème 2 :** *Soit $A$ factoriel et $\{a_p\}_{p\in\mathcal{P}}$ une famille d'irréductibles. Alors $A[X]$ est un anneau factoriel, avec pour famille explicite d'irréductibles la réunion de :*

(1) *la famille $\{a_p\}_{p\in\mathcal{P}}$, et*

(2) *la famille des polynômes non constants et primitifs de $A[X]$, irréductibles dans $K[X]$.*

Par exemple la famille des nombres premiers et des polynômes non constants de $\mathbb{Z}[X]$, à coefficients premiers entre eux, et irréductibles dans $\mathbb{Q}[X]$, est une famille d'irréductibles de $\mathbb{Z}[X]$.

On trouve les démonstrations correspondantes dans les livres d'algèbre commutative.

## 2 Application aux automorphismes de $k(X)$

On peut faire un développement sur les automorphismes de $k(X)$, soit en prenant pour acquis le théorème, soit en n'admettant que le lemme de Gauss et en prouvant le corollaire suivant :

**Corollaire 3 :** *Soit $A$ un anneau factoriel, $K$ son corps de fractions, $P \in A[X]$ tel que $c(P) = 1$. Alors, $P$ est irréductible dans $A[X]$ ssi il est irréductible dans $K[X]$.*

**Preuve :** Supposons $P$ irréductible dans $K[X]$. Si $P = QR$ dans $A[X]$, cette égalité étant valable aussi dans $K[X]$, alors soit $Q$ soit $R$ est une constante non nulle de $K$, disons $Q$. Donc d'après le lemme de Gauss $1 = c(P) = c(Q)c(R) = ac(R)$, et on déduit que $Q = a$ est en fait inversible dans $A$, donc $P$ est irréductible dans $A[X]$.

Réciproquement, supposons $P$ irréductible dans $A[X]$. Supposons que $P = QR$ avec $Q$ et $R$ dans $K[X]$. Il est clair qu'il existe des éléments $\lambda, \mu \in K^\times$ tels que $Q = \lambda Q'$ et $R = \mu R'$ avec $Q', R'$ dans $A[X]$ et de contenu 1. On a donc $P = \lambda\mu Q'R'$ et en utilisant encore le lemme de Gauss, $1 = \lambda\mu$. Il s'ensuit que $P = Q'R'$ et comme $P$ irréductible dans $A[X]$ par hypothèse, $Q'$ ou $R'$ est un inversible de $A$. Donc $Q$ ou $R$ est un inversible de $K$. ▫

**Théorème 4 :** *Le morphisme* $\mathrm{GL}_2(k) \to \mathrm{Aut}_k(k(x))$ *donné par*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \left(X \mapsto \frac{aX+b}{cX+d}\right)$$

*induit un isomorphisme entre* $\mathrm{PGL}_2(k)$ *et le groupe des $k$-automorphismes du corps des fractions rationnelles $k(X)$.*

**Preuve :** Soit et montrons comment on en déduit le théorème.

Il est facile de vérifier que l'application de l'énoncé du théorème est bien un morphisme de groupes et que de plus, une matrice $(2, 2)$

induit l'identité de $k(X)$ ssi c'est une homothétie. On a donc une injection de $\mathrm{PGL}_2(k)$ dans $\mathrm{Aut}_k(k(x))$ et il ne reste qu'à montrer qu'elle est surjective. Soit $\varphi\colon k(X) \to k(X)$ un automorphisme, notons $F = \varphi(X) = U(X)/V(X)$. L'image de $\varphi$ est le corps $k(F)$, donc si c'est un automorphisme on a $[k(X) : k(F)] = 1$. Écrivons $F = U(X)/V(X)$ sous forme irréductible, et posons $n = \max\{\deg(U), \deg(V)\}$. Si on montre que $[k(X) : k(F)] = n$, on obtient $n = 1$ et le résultat en découle immédiatement.

Il nous reste à montrer que $[k(X) : k(F)] = n$. D'abord, il est clair que $X$ est algébrique sur $k(F)$ car annulé par le polynôme de degré $n$ :

$$P(T) \overset{\text{déf}}{=} V(T)F - U(T)\,.$$

Il nous suffit maintenant de montrer que $P$ est irréductible dans $k(F)[T]$. L'égalité

$$\underbrace{[k(X) : k]}_{\infty} = \underbrace{[k(X) : k(F)]}_{\text{fini}}\,[k(F) : k]$$

montre que $F$ est transcendant sur $k$, donc $k[F]$ est un anneau de polynômes. La clé de la preuve est que $P(T) = V(T)F - U(T)$ est irréductible dans $k[T][F]$ :

- si on a admis le théorème 2, cela provient du fait qu'il est de degré 1 en $F$ et primitif ;
- si on n'a admis que le corollaire 3, on dit que $P$ est de degré 1 en $F$ donc irréductible dans $k(X)[F]$, et comme il est primitif, il est irréductible dans $k[X][F]$ d'après ce corollaire.

Comme $k[T][F] = k[F][T]$, on en déduit que le polynôme $P(T)$ est irréductible dans $k[F][T]$, donc dans $k(F)[T]$, cqfd. □

# Déterminant des matrices à coefficients dans un anneau

Cette note présente une relecture (complétée) d'un résultat que l'on trouve dans Leichtnam, Schauer, Exercices corrigés de Mathématiques posés aux oraux X-ENS, Algèbre 1, *Ellipses*. Il peut servir de développement pour la leçon :

- Déterminant. Exemples et applications.

Par ailleurs, comme les anneaux les plus importants au programme (mis à part les corps) sont les anneaux principaux, en insistant plus sur l'aspect « matrices à coefficients dans un anneau principal » on peut aussi imaginer une utilisation dans les leçons :

- Anneaux $\mathbb{Z}/n\mathbb{Z}$. Applications.
- Anneaux principaux. Applications.

Soit donc $A$ un anneau commutatif avec un élément unité noté 1.

**Théorème :** *Soit $M$ une matrice de taille $n$ à coefficients dans $A$, et $f : A^n \to A^n$ l'endomorphisme $A$-linéaire associé. Alors :*

(1) *$f$ est surjectif ssi $f$ est bijectif ssi* $\det(f)$ *est inversible dans $A$.*
(2) *$f$ est injectif ssi* $\det(f)$ *est non diviseur de zéro dans $A$.*

*De plus, dans le cas injectif,*

(3) *Si $A = \mathbb{Z}$, le conoyau de $f$ est fini de cardinal* $|\det(f)|$.
(4) *Si $A = k[X]$, le conoyau de $f$ est un $k$-e.v. de dimension finie égale à* $\deg(\det(f))$.

Rappelons que le *conoyau* est le quotient de l'ensemble but par l'image, i.e. $\mathrm{coker}(f) = A^n/f(A^n)$. Le conoyau est une mesure du défaut de surjectivité de la même façon que le noyau est une mesure du défaut d'injectivité. Ainsi, $f$ est surjectif si et seulement si $\mathrm{coker}(f) = 0$.

**Démonstration :** On notera $e_1, \ldots, e_n$ la base canonique de $A^n$.

(1) Si $f$ est surjectif, pour tout $i$ il existe un vecteur $\epsilon_i$ tel que $f(\epsilon_i) = e_i$. Si l'on pose $g(e_i) = \epsilon_i$ pour tout $i$, on définit un unique morphisme $g : A^n \to A^n$. De plus, on a $f \circ g = \mathrm{Id}$ car ceci est vrai pour tous les $e_i$, qui forment une partie génératrice. On en déduit que $\det(f)\det(g) = 1$ et donc $\det(f)$ est inversible. Alors, la formule de la comatrice :

$$M\,{}^t\widetilde{M} = {}^t\widetilde{M}M = \det(M)\,\mathrm{Id}$$

montre que $f$ est bijectif. Comme enfin bijectif implique surjectif, on a tout démontré.

(2) Posons $d = \det(f)$. Si $d$ est non diviseur de zéro, supposons que $f(x) = 0$. Matriciellement, on a $Mx = 0$ et en appliquant la transposée de la comatrice, on trouve $dx = 0$. En regardant les coordonnées de $x$, l'hypothèse sur $d$ implique que $x = 0$ donc $f$ est injectif.

Réciproquement, si $d$ est diviseur de zéro, on va montrer que $f$ n'est pas injectif. Soit $u \in A$ non nul tel que $ud = 0$.

Si pour tout mineur $\mu$ de $M$ on a $u\mu = 0$, alors en particulier ceci est vrai pour les mineurs de taille 1, i.e. les coefficients de la matrice $M$. On a donc $f(ue_1) = 0$, or $ue_1 \neq 0$, donc $f$ n'est pas injectif.

Sinon, il existe une matrice extraite $N$ de $M$ telle que $u\det(N) \neq 0$. Choisissons une telle matrice de taille $r$ maximale ; on a $r < n$ puisque $ud = 0$. Quitte à réordonner les vecteurs de base à la source et au but, c'est-à-dire à multiplier $M$ à gauche et à droite par des matrices de permutation, on peut supposer que $N$ est la matrice de taille $r$ située en haut à gauche. Maintenant, pour chaque $i \in \{1, \dots, n\}$, bordons les $r$ premières lignes de $M$ inférieurement avec la $i$-ème ligne, et appelons $P_i$ la matrice de taille $r+1$ située à gauche :

$$P_i = \begin{pmatrix} m_{1,1} & \dots & m_{1,r+1} \\ \vdots & & \vdots \\ m_{r,1} & \dots & m_{r,r+1} \\ m_{i,1} & \dots & m_{i,r+1} \end{pmatrix}.$$

Pour $i \leqslant r$ la matrice $P_i$ a deux lignes égales donc son déterminant est nul, et pour $i \geqslant r+1$ c'est une matrice extraite de $M$ de taille $r+1$, donc son déterminant est annulé par $u$ compte tenu de l'hypothèse sur $r$. Dans les deux cas $u\det(P_i) = 0$, et si on développe par rapport à la dernière ligne, on trouve $u\sum_{j=1}^{r+1}(-1)^j m_{i,j}\mu_j = 0$ où $\mu_j$ est le mineur du coefficient de position $(r+1, j)$. Pour $i$ variant, ces égalités disent exactement que $M(ux) = 0$ où $x$ est le vecteur de coordonnées $(-\mu_1, \dots, (-1)^{r+1}\mu_{r+1}, 0, \dots, 0)$. Comme $u\mu_{r+1} = u\det(N) \neq 0$, on a $ux \neq 0$, donc $f$ n'est pas injectif.

(3) D'après les résultats sur les classes de congruence de matrices à coefficients dans un anneau principal, il existe des matrices $R, S$ inversibles à coefficients dans $\mathbb{Z}$ telles que $D := RMS$ est diagonale d'éléments diagonaux égaux aux facteurs invariants $d_1, \dots, d_n$ tels que $d_i | d_{i+1}$ pour tout $i$. On en déduit que

$$\mathrm{coker}(f) \simeq \mathbb{Z}^n / D(\mathbb{Z}^n) \simeq \mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_n\mathbb{Z}$$

de sorte que $|\mathrm{coker}(f)| = d_1 \dots d_n = |\det(f)|$.

(4) Le raisonnement est le même : il existe des matrices $R, S$ inversibles à coefficients dans $k[X]$ telles que $D := RMS$ est diagonale d'éléments diagonaux égaux aux facteurs invariants $P_1, \dots, P_n$ tels que $P_i | P_{i+1}$ pour tout $i$. On en déduit que

$$\mathrm{coker}(f) \simeq \frac{k[X]}{(P_1)} \times \cdots \times \frac{k[X]}{(P_n)}$$

puis $\dim_k(\mathrm{coker}(f)) = \deg(P_1) + \cdots + \deg(P_n) = \deg(P_1 \dots P_n) = \deg(\det(f))$. ▫

# Endomorphismes cycliques

La notion d'endomorphisme cyclique peut être évoquée entre autres dans les leçons :

Exemples d'applications des idéaux d'un anneau commutatif unitaire.
Anneaux principaux.
Sous-espaces stables d'un endom. d'un espace vectoriel de dim. finie. Applications.
Polynômes d'endomorphismes. Applications.
Formes linéaires et hyperplans en dimension finie. Exemples et applications.

Pour la leçon sur les idéaux et celle sur les anneaux principaux, on parlera de $\mu$ et $\mu_x$, générateurs d'idéaux de $k[X]$ ayant une signification géométrique. Pour la leçon sur les formes linéaires, on parlera du lemme 2 ci-dessous. Les endomorphismes cycliques sont aussi un peu présents dans la leçon sur l'exponentielle de matrices, si on évoque la non surjectivité de l'exponentielle de $\mathfrak{sl}_n(\mathbb{C})$ (cf le développement correspondant).

$\therefore$

Soit $E$ un espace vectoriel de dimension finie sur un corps $k$. Pour $u \in \mathrm{L}(E)$ et $P \in k[X]$ on notera parfois simplement $Pu$ au lieu de $P(u)$. On considère le morphisme d'algèbres

$$\begin{array}{rcl} \varphi\colon k[X] & \to & \mathrm{L}(E) \\ P & \mapsto & Pu \end{array}$$

Étant donné en plus un $x \in E$ on peut considérer le morphisme de $k$-espaces vectoriels

$$\begin{array}{rcl} \varphi_x\colon k[X] & \to & E \\ P & \mapsto & Pu(x) \end{array}$$

On note $\mu$ resp. $\mu_x$ le générateur unitaire du noyau de $\varphi$, resp. $\varphi_x$. On note $k[u]$ resp. $E_x$ l'image de $\varphi$, resp. de $\varphi_x$. Il est clair que pour tout $x \in E$, on a $\mu_x | \mu$.

**Définition :** $u$ est *cyclique* ssi il existe $x \in E$ tel que $E_x = E$.

**Théorème :** *Soit $u \in \mathrm{L}(E)$, $\mu$ son polynôme minimal, $\chi$ son polynôme caractéristique. Les conditions suivantes sont équivalentes :*

(1) *$u$ est cyclique.*

(2) *$\mu = \chi$.*

(3) *l'ensemble des endomorphismes qui commutent avec $u$ est égal à $k[u]$.*

L'ensemble des endomorphismes qui commutent avec $u$ est appelé *commutant* de $u$ et noté parfois $Z(u)$. Il contient toujours $k[u]$. On montrera (1)$\Leftrightarrow$(2) après d'un premier lemme et (1)$\Leftrightarrow$(3) après un second. Dans le lemme 2 on exhibe un certain sous-espace vectoriel stable qui a un supplémentaire stable : notez le lien avec la semi-simplicité, où le phénomène étudié est précisément l'existence de supplémentaires stables. Si on ne démontre pas tout, on peut admettre le lemme 1, mais pas le lemme 2 qui est le cœur de la preuve !

**Lemme 1 :** *Il existe $a \in E$ tel que $\mu_a = \mu$.*

**Preuve :** Soit $\mu = P_1^{\alpha_1} \dots P_r^{\alpha_r}$ la décomposition en facteurs irréductibles, et $E_i = \ker(P_i^{\alpha_i} u)$. Choisissons $x_i \in E_i$ tel que $P_i^{\alpha_i - 1} u(x) \neq 0$, il est alors clair que $\mu_{x_i} = P_i^{\alpha_i}$. Posons $a = x_1 + \dots + x_r$, alors par définition de $\mu_a$ on a

$$0 = \mu_a u(a) = \underbrace{\mu_a u(x_1)}_{\in E_1} + \dots + \underbrace{\mu_a u(x_r)}_{\in E_r}$$

Comme les $E_i$ sont en somme directe il vient $\mu_a u(x_i) = 0$ pour tout $i$. Par définition de $\mu_{x_i}$ on obtient $\mu_{x_i} | \mu_a$ et comme les $\mu_{x_i}$ sont premiers entre eux deux à deux, leur produit divise $\mu_a$. Ceci montre que $\mu | \mu_a$, et comme $\mu_a | \mu$ on a fini. ▫

Cela suffit à montrer que (1)⇔(2). En effet par définition de $\mu_a$ et de $E_a$, l'application $\varphi_a$ donne un isomorphisme $k[X]/(\mu_a) \simeq E_a$. Donc si on choisit $a$ comme dans le lemme, $\mu = \chi$ équivaut à $\mu_a = \chi$, ou encore $\deg(\mu_a) = n$, i.e. $\dim(E_a) = n$.

**Lemme 2 :** *Soit $a \in E$ tel que $\mu_a = \mu$. Alors $E_a$ est un sous-espace $u$-stable pour lequel il existe un supplémentaire $u$-stable.*

**Preuve :** Soit $k = \deg(\mu_a) = \deg(\mu)$. Par hypothèse les vecteurs $e_1 = a$, $e_2 = u(a)$,... $e_k = u^{k-1}(a)$ forment une base de $E_a$. Complétons-la avec des vecteurs $e_{k+1}, \ldots, e_n$ en une base de $E$ et soit $\{e_i^*\}$ la base duale. Considérons le sous-espace

$$\begin{aligned} G = & \{ x \in E,\ u^i(x) \text{ n'a pas de composante sur } e_k, \text{ pour tout } i \geqslant 0 \} \\ = & \bigcap_{i \geqslant 0} \ker(e_k^* \circ u^i) \\ = & \bigcap_{i=0}^{k-1} \ker(e_k^* \circ u^i) \\ & \text{car } u^k \text{ est combinaison linéaire des } u^i \text{ pour } i \leqslant k-1. \end{aligned}$$

Il est clair que $G$ est $u$-stable et que $E_a \cap G = \{0\}$. Si on montre que $\dim(G) = n - k$ on aura $E = E_a \oplus G$ et ce sera terminé. Pour cela montrons que les $k$ équations d'hyperplans qui définissent $G$ sont linéairement indépendantes :

$$\begin{aligned} \textstyle\sum_{j=0}^{k-1} \lambda_j (e_k^* \circ u^j) = 0 \ & \Rightarrow \ \textstyle\sum_{j=0}^{k-1} \lambda_j (e_k^* \circ u^{j+i}) = 0 \ (\forall i) \\ & \Rightarrow \ e_k^*(u^i(P(u)) = 0 \ (\forall i) \text{ où } P(X) = \textstyle\sum \lambda_j X^j \\ & \Rightarrow \ P(u)(a) \in G \end{aligned}$$

Comme par ailleurs $P(u)(a) \in E_a$ par définition, on obtient $P(u)(a) = 0$. Il en découle que $\mu_a | P$, et comme $P$ est de degré au plus $k-1$ il doit être nul. Donc tous les $\lambda_i$ sont nuls, ce qui montre que les équations sont linéairement indépendantes comme on le voulait. En conséquence, $\dim(G) = n - k$ et on a le résultat attendu. □

Il reste à montrer (1)⇔(3). Pour le sens direct, il suffit de montrer qu'un endomorphisme $v$ qui commute avec $u$ est un polynôme en $u$. Or par hypothèse il existe $x$ tel que $E_x = E$, en particulier $v(x) = Pu(x)$ pour un certain polynôme $P$. On va montrer que $v = Pu$. Pour cela soit $y \in E = E_x$, il s'écrit $y = Qu(x)$. Comme $v$ commute avec $u$ il commute avec tout polynôme en $u$, donc

$$v(y) = v(Qu(x)) = Qu(v(x)) = Qu(Pu(x)) = Pu(Qu(x)) = Pu(y)$$

Pour (3)⇒(1) on reprend les notations de la preuve du lemme 1, $\mu = P_1^{\alpha_1} \ldots P_r^{\alpha_r}$, $E_i = \ker(P_i^{\alpha_i} u)$, $a = x_1 + \cdots + x_r$ vérifiant $\mu_a = \mu$. On a une décomposition $E = E_a \oplus G$ en sous-espaces stables. Soit $\pi$ le projecteur sur $G$ parallèlement à $E_a$. Comme $E_a$ et $G$ sont $u$-stables, $\pi$ commute avec $u$, donc par hypothèse, on a $\pi = P(u)$ pour un certain polynôme $P$. On en déduit que $P(u_{|E_a}) = \pi_{|E_a} = 0$, donc $\mu_u = \mu_{u_{|E_a}}$ divise $P$, donc $\pi = P(u) = 0$. Ainsi $G = 0$ et $E_a = E$, c'est-à-dire, $u$ est cyclique.

**Bibliographie :**

[Gou] GOURDON, Les Maths en tête, Mathématiques pour M', *Ellipses*.

# Endomorphismes partiellement isométriques et norme

Leçons concernées :

Exemples d'applications des idéaux d'un anneau commutatif unitaire.
Anneaux principaux.
Sous-espaces stables d'un endom. d'un espace vectoriel de dim. finie. Applications.
Polynômes d'endomorphismes. Applications.
Formes linéaires et hyperplans en dimension finie. Exemples et applications.
Dualité en dimension finie.

Soit $E$ un espace vectoriel euclidien (de dimension finie $n$), de produit scalaire noté $\langle , \rangle$.

On dit qu'un endomorphisme $u \in \mathrm{L}(E)$ est *partiellement isométrique* (en abrégé PI) ssi $\| u(x) \| = \| x \|$, pour tout $x \in (\ker(u))^\perp$. On rappelle que cela équivaut à dire que $\langle u(x), u(y) \rangle = \langle x, y \rangle$, pour tous $x, y \in (\ker(u))^\perp$, ou encore que $u^*u$ est un projecteur orthogonal.

Enfin, on rappelle que si $s \in \mathrm{L}(E)$ est un endomorphisme symétrique positif, il existe un unique endomorphisme symétrique positif $r$ tel que $s = r^2$, appelé la racine carrée de $s$. Si $f \in \mathrm{L}(E)$ est quelconque, on note $|f|$ la racine carrée de $f^*f$. C'est donc un endomorphisme symétrique positif. Nous allons montrer :

**Théorème :** Soit $f \in \mathrm{L}(E)$.

(1) *pour tout $v$ PI on a* $\mathrm{tr}(vf) \leqslant \mathrm{tr}(|f|)$.

(2) $N(f) = \mathrm{tr}(|f|)$ *est une norme sur* $\mathrm{L}(E)$ *avec* $N(f) = \mathrm{tr}(f)$ *si $f$ est symétrique positif.*

**Lemme 1 :** $\ker f = \ker |f|$.

**Preuve :** On montre d'abord que $\ker f = \ker f^*f$, pour cela seule $\supset$ est à montrer. Or $f^*fx = 0$ implique $\| fx \|^2 = \langle f^*fx, x \rangle = 0$.

Ceci se réécrit $\ker f = \ker |f|^2$. On en déduit le lemme en appliquant cela à $f$ et $|f|$, observant que $|f|$ a même module que $f$. □

**Lemme 2 :** $\exists\, ! \, u \in \mathrm{L}(E)$ *qui soit PI avec* $|f| = uf$, $(\ker(u))^\perp = \mathrm{im}(f)$ *et* $u^*uf = f$.

**Preuve :** Sur $\mathrm{im}(f)$ on définit $u$ par $u(f(y)) = |f|(y)$. Ceci est légitime car d'après le lemme 1, si $f(y) = 0$ alors $|f|(y) = 0$. Sur $(\mathrm{im}(f))^\perp$ on définit $u$ par $u(x) = 0$. On a ainsi défini un endomorphisme $u$ sur $E$, tel que $uf = |f|$ et $(\ker(u))^\perp = \mathrm{im}(f)$, manifestement unique avec ces propriétés. Il reste à voir qu'il est PI et que $u^*uf = f$.

Sur $(\ker(u))^\perp = \mathrm{im}(f)$, $u$ est isométrique car pour $x = f(y)$ on a

$$\begin{aligned} \| ux \|^2 &= \| \, |f|(y) \, \|^2 = \langle |f|(y), |f|(y) \rangle = \langle |f|^2(y), (y) \rangle \\ &= \langle f^*f(y), y \rangle = \| \, f(y) \, \|^2 = \| \, x \, \|^2 \quad . \end{aligned}$$

Alors $u^*u$ est le projecteur orthogonal sur $(\ker(u))^\perp = \mathrm{im}(f)$ et il en découle immédiatement que $u^*uf = f$. □

**Preuve du théorème :** On utilisera le fait suivant noté $(*)$ : si $p$ est un projecteur orthogonal et $f$ un endormophisme symétrique positif, alors $\mathrm{tr}(pf) \leqslant \mathrm{tr}(f)$. Pour voir cela on choisit une base orthonormale $\{e_i\}$ qui diagonalise $p$, de sorte que $p(e_i) = e_i$ sur $\mathrm{im}(p)$ et $p(e_i) = 0$ sur $\ker(p)$. Alors

$$\begin{aligned} \mathrm{tr}(pf) &= \textstyle\sum_{i=1}^n \langle pfe_i, e_i \rangle = \sum \langle fe_i, p^*e_i \rangle = \sum \langle fe_i, pe_i \rangle \\ &= \sum_{e_i \in \mathrm{im}(p)} \langle fe_i, e_i \rangle \leqslant \sum_{e_i \in \mathrm{im}(p)} + \sum_{e_i \in \ker(p)} = \mathrm{tr}(f) \end{aligned}$$

où l'inégalité provient du fait que $f$ est positif.

On va montrer l'énoncé un peu plus fort : si $f$ est symétrique positif, pour tous $u, v$ PI on a $\mathrm{tr}(uvf) \leqslant \mathrm{tr}(f)$. Soit $g$ la racine

carrée de $f$ et fixons une base orthonormée $\{e_i\}$. On a

$$\begin{aligned}\mathrm{tr}(uvf) &= \sum_{i=1}^n \langle uvfe_i, e_i\rangle = \sum\langle g^2e_i, v^*u^*e_i\rangle = \sum\langle ge_i, gv^*u^*e_i\rangle \\ &\leqslant \sum \|ge_i\|\,\|gv^*u^*e_i\| \text{ par Cauchy-Schwarz} \\ &\leqslant \sqrt{\sum\|ge_i\|^2}\,\sqrt{\sum\|gv^*u^*e_i\|^2} \text{ par Cauchy-Schwarz encore.}\end{aligned}$$

Or d'une part

$$\sum \|ge_i\|^2 = \sum\langle ge_i, ge_i\rangle = \sum\langle g^2e_i, e_i\rangle = \mathrm{tr}(g^2) = \mathrm{tr}(f)$$

D'autre part,

$$\begin{aligned}\sum\|gv^*u^*e_i\|^2 &= \sum\langle uvg^2v^*u^*e_i, e_i\rangle = \mathrm{tr}(uvfv^*u^*) \\ &= \mathrm{tr}((u^*u)vfv^*) \overset{(*)}{\leqslant} \mathrm{tr}(vfv^*) = \mathrm{tr}(v^*vf) \overset{(*)}{\leqslant} \mathrm{tr}(f)\end{aligned}$$

où les deux inégalités proviennent du fait $(*)$ ci-dessus. Ceci conclut notre calcul, et permet de montrer (1).

Passons au point (2), i.e. à la vérification du fait qu'on a une norme.

D'abord $|f|$ est symétrique positif, donc diagonalisable à valeurs propres positives. Il s'ensuit que $N(f) = \mathrm{tr}(|f|) \geqslant 0$ et aussi que si $N(f) = 0$ alors toutes ses valeurs propres sont nulles, donc $|f|$ qui est diagonalisable est nul, donc $f$ aussi puisque $\ker f = \ker|f|$ (lemme 1).

Ensuite clairement $N(\lambda f) = |\lambda|N(f)$.

Enfin pour $f, g \in \mathrm{L}(E)$ quelconques, par le lemme 2 on peut choisir $u$ PI tel que $|f+g| = u(f+g)$. On a alors, utilisant le point (1),

$$N(f+g) = \mathrm{tr}(uf) + \mathrm{tr}(ug) \leqslant \mathrm{tr}(|f|) + \mathrm{tr}(|g|) = N(f) + N(g)$$

Le théorème est prouvé. □

### Bibliographie

[Gug] Gugger, Problèmes corrigés de mathématiques posés au concours de Polytechnique, Tome 6, *Ellipses*.

## Endomorphismes semi-simples

Soit $E$ un espace vectoriel sur un corps $k$, de dimension finie $n$. On dit qu'un endomorphisme $u \in \mathrm{L}(E)$ est *semi-simple* si et seulement si tout sous-espace $u$-stable $F \subset E$ possède un supplémentaire $u$-stable. Nous utiliserons cette notion surtout lorsque le corps de base $k$ est *parfait*, ce qui veut dire par définition que $k$ est soit de caractéristique 0, soit de caractéristique $p > 0$ avec un endomorphisme de Frobenius surjectif.

Exemples de corps parfaits : les corps de caractéristique 0, les corps finis, les corps algébriquement clos.

Exemples de corps non parfaits : corps de fractions rationnelles en une ou plusieurs indéterminées sur un corps de carcactéristique $p > 0$, typiquement, $\mathbb{F}_p(X)$.

**Théorème :** *Soient les conditions :*

(i) *$u$ est semi-simple.*

(ii) *le polynôme minimal de $u$ est produit de polynômes irréductibles distincts.*

(iii) *$u$ est diagonalisable sur une clôture algébrique de $k$.*

*Alors (i) $\iff$ (ii) $\Leftarrow$ (iii), et si $k$ est parfait les trois conditions sont équivalentes.*

**Lemme :** *Soit $u$ un endomorphisme et $\mu_u = P_1^{\alpha_1}\dots P_r^{\alpha_r}$ la décomposition de son polynôme minimal en facteurs irréductibles. Soit $E_i = \ker(P_i^{\alpha_i}(u))$. Pour tout sous-espace vectoriel $F \subset E$ qui est $u$-stable, on a $F = \oplus F \cap E_i$.*

**Preuve :** Il est clair que les espaces $F \cap E_i$ sont en somme directe. Il suffit de voir qu'ils engendrent $F$. Or pour $x \in F$, on peut écrire $x = x_1 + \dots + x_r$ avec $x_i \in E_i$. On utilise le fait que les projecteurs $\pi_i : E \to E_i$, $x \mapsto x_i$, sont des polynômes en $u$. Alors si $F$ est stable par $u$, il est stable par $\pi_i$, donc $x_i = \pi_i(x) \in F$ et c'est gagné. □

**Preuve du théorème :**

i ⇒ ii Soit $u$ semi-simple. Supposons que la décomposition du polynôme minimal $\mu_u$ contient un facteur carré : $\mu_u = P^2Q$. On va montrer que $(PQ)(u) = 0$ ce qui contredira le fait que le polynôme minimal de $u$ est $P^2Q$. Soit $F = \ker(P(u))$ et $S$ un supplémentaire $u$-stable de $F$. Soit $a = (PQ)(u)$, alors :

- $a$ est nul sur $F$ puisque $a = (QP)(u) = Q(u) \circ P(u)$.
- $a$ est nul sur $S$. En effet si $y \in S$, on a $a(y) \in F$ puisque $P(u)[a(y)] = (P^2Q)(u)(y) = 0$ et $a(y) \in S$ puisque $S$ est stable par $u$, donc par $a$ qui est un polynôme en $u$. Donc $a(y) \in F \cap S$, donc $a(y) = 0$ car $F \cap S = 0$, cqfd.

En conclusion $a = 0$, d'où la contradiction cherchée, donc il n'y pas de facteur carré dans $\mu_u$.

ii ⇐ i Réciproquement supposons que $\mu_u = P_1 \dots P_r$ avec tous les $P_i$ irréductibles distincts. Soit $F$ un sous-espace stable, on va lui construire un supplémentaire stable. Soit $E_i = \ker(P_i(u))$. D'après le lemme on a $F = \oplus F \cap E_i$ de sorte que si pour chaque $i$ on construit un supplémentaire stable pour $F \cap E_i$ dans $E_i$, par somme on aura un supplémentaire pour $F$ dans $E$. Comme $\mu_{u_{|E_i}} = P_i$, on se ramène ainsi au cas où $\mu_u = P$ est irréductible.

Si $F = E$ on a un supplémentaire stable $G = 0$ est c'est fini. Sinon, il existe $x \in E - F$. Considérons le morphisme de $k$-algèbres $\varphi\colon k[X] \to E$ défini par

$$Q \mapsto Q(u)(x)$$

On note $G_x = \{Q(u)(x),\ Q \in k[X]\}$ son image, et $P_x$ le polynôme unitaire générateur de son noyau. On va montrer que $F \cap G_x = 0$. Ceci fait, en itérant on construira $G_{x'}, G_{x''}, \dots$ et le supplémentaire cherché sera $G_x \oplus G_{x'} \oplus G_{x''} \dots$.

Par définition de $\mu_u = P$ on a $P_x | P$ donc ils sont égaux puisque $P$ est irréductible. Soit $y \in F \cap G_x$, que l'on peut écrire sous la forme $y = Q(u)(x)$. J'affirme que $P | Q$ de sorte que $y = 0$, ce qui conclura à $F \cap G_x = 0$. En effet, si $P$ ne divise pas $Q$ alors ces polynômes sont premiers entre eux, choisissons une relation de Bézout $UP + VQ = 1$. L'image par $\varphi$ de cette relation de Bézout donne, dans $E$ :

$$U(u)[\underbrace{P(u)(x)}_{=0}] + V(u)[\underbrace{Q(u)(x)}_{y}] = x$$

Or $V(u)(y) \in F$ car $F$ est stable sous $u$. Ceci contredit le choix de $x \in E - F$.

iii ⇐ ii Le polynôme minimal est inchangé par extension du corps de base, donc si $u$ est diagonalisable sur une clôture algébrique $\overline{k}$ de $k$, alors $\mu_u$ est scindé dans $\overline{k}$ à racines simples et distinctes. A fortiori, comme polynôme à coefficients dans $k$, il est sans facteur carré.

ii ⇐ iii lorsque $k$ est parfait. Montrons d'abord que le polynôme dérivé $\mu'$ est non nul. Dans le cas contraire, ceci veut dire que c'est un polynôme en $X^p$ i.e. $\mu_u(X) = F(X^p)$. Comme $k$ est parfait, tous les coefficients de $P$ sont des puissances $p$-èmes et donc $F(X^p) = (G(X))^p$. Ceci contredit le fait que $\mu_u$ est sans facteur carré. Il en résulte que $\mu'_u \neq 0$, et donc le pgcd de $\mu_u$ et $\mu'_u$ comme polynômes à coefficients dans $\overline{k}$ est égal à 1. Le pgcd est inchangé par extension du corps de base (ce fait est, par exemple, un corollaire du calcul du pgcd par l'algorithme d'Euclide), donc finalement $\mu_u$ est sans facteur carré, c'est-à-dire produit de polynômes irréductibles distincts de $k[X]$. ▫

**Contre-exemple 1 :** Soit le corps non parfait $k = \mathbb{F}_2(T)$, corps des fractions rationnelles en l'indéterminée $T$ sur $\mathbb{F}_2$. Considérons l'espace vectoriel $E = k^2$ et l'endomorphisme

$$u = \begin{pmatrix} 1 & T+1 \\ 1 & 1 \end{pmatrix}.$$

Le polynôme caractéristique de $u$ est $\chi_u(X) = X^2 + T$ (attention : $1 = -1$ dans $k$). Ce polynôme est irréductible, car $T$ n'est pas un carré dans $k$, donc $u$ est un endomorphisme semi-simple. Supposons que $u$ est diagonalisable sur une clôture algébrique $\overline{k}$ de $k$. Soit $\alpha$ une racine de $\chi_u$ dans $\overline{k}$, on a $\chi_u(X) = (X + \alpha)^2$. Donc $u$ est semblable dans $\overline{k}$ à l'homothétie $\alpha\,\mathrm{Id}$, et domme les homothéties commutent à

toutes les matrices, il s'ensuit qu'en fait $u = \alpha\,\mathrm{Id}$. Ceci n'est pas le cas, donc $u$ n'est pas diagonalisable sur $\overline{k}$. ▫

**Contre-exemple 2 :** Voici une métode plus facile, et plus conceptuelle aussi, pour donner un contre-exemple. Soit $A$ une algèbre unitaire et associative sur un corps $k$, et $\mathrm{End}_k(A)$ l'anneau des endomorphismes de $k$-espace vectoriel. Pour tout $a \in A$, on note $G_a : A \to A$ l'endomorphisme de multiplication à gauche par $a$, tel que $G_a(x) = ax$. On vérifie alors facilement qu'en associant à $a$ le morphisme $G_a$ on définit un morphisme injectif de $k$-algèbres $A \hookrightarrow \mathrm{End}_k(A)$. Si $A$ est de dimension finie $n$, l'algèbre $\mathrm{End}_k(A)$ est isomorphe à l'algèbre des matrices carrées $(n, n)$.

Soit le corps des fractions rationnelles $k = \mathbb{F}_p(T)$, soit le corps $A = k[U]/(U^p - T)$ et $u$ l'image de l'indéterminée $U$ dans $A$. Le polynôme minimal de $u \in \mathrm{End}_k(A) \simeq \mathrm{M}_p(k)$ est $X^p - T$, qui est irréductible, donc $u$ est semi-simple. En revanche, il n'est pas diagonalisable sur une clôture algébrique $\overline{k}$, car son polynôme minimal a une seule racine $\alpha$ dans $\overline{k}$ et $u$ n'est pas une homothétie. ▫

La décomposition $u = d + n$ dite de Jordan-Dunford, valable pour un endomorphisme dont le polynôme caractéristique est scindé, s'étend comme suit.

**Proposition :** *Soit $k$ un corps parfait et soit $u \in \mathrm{L}(E)$ un endomorphisme quelconque. Alors il existe un couple $(s, n)$ unique avec*

(1) $u = s + n$,

(2) *$s$ semi-simple et $n$ nilpotent,*

(3) $sn = ns$.

Le cas particulier $k = \mathbb{R}$ est le plus important pour nous. Démontrons le résultat dans ce cas particulier très simple. On peut plonger $\mathrm{M}_n(\mathbb{R})$ dans $\mathrm{M}_n(\mathbb{C})$ et pour tout endomorphisme $a$, représenté par une matrice complexe dans une base fixée, notons $\overline{a}$ l'endomorphisme représenté par la matrice dont les coefficients sont les complexes conjugués. La proposition dit juste ceci : on peut écrire la décomposition $u = d + n$ dans $\mathbb{C}$. On a $\overline{u} = u$ et comme $\overline{u} = \overline{d} + \overline{n}$, par unicité de la décomposition de Dunford on a $\overline{d} = d$, $\overline{n} = n$. Donc $d$ et $n$ sont en fait à coefficients dans $\mathbb{R}$. Clairement $d$ est semi-simple, on a donc la décomposition cherchée.

**Preuve :** La démonstration utilise un peu de théorie de Galois. Soit $K$ le corps de décomposition du polynôme caractéristique $\mu_u$. Comme $k$ est parfait, c'est une extension galoisienne de $k$. Soit $G$ le groupe de Galois de $K$ sur $k$. Si on choisit une base de $E$ alors $\mathrm{L}(E \otimes_k K)$ s'identifie à l'anneau des matrices $(n, n)$ à coefficients dans $K$. Via cette identification, le groupe $G$ agit sur $\mathrm{L}(E \otimes_k K)$ en agissant sur les coefficients des matrices. La théorie de Galois nous dit que $k = K^G$, et donc les éléments de $\mathrm{L}(E \otimes_k K)$ fixés par $G$ sont les éléments de $\mathrm{L}(E)$.

Sur $K$, on peut écrire la décomposition $u = d + n$ où $d$ et $n$ sont dans $\mathrm{L}(E \otimes_k K)$. Pour tout $\sigma \in G$, on a $u^\sigma = u$ car $u \in \mathrm{L}(E)$. Or on peut écrire $u^\sigma = d^\sigma + n^\sigma$. Il est facile (immédiat !) de voir que $d^\sigma$ est diagonalisable et $n^\sigma$ est nilpotent, donc par unicité de la décomposition $d + n$ on doit avoir $d^\sigma = d$ et $n^\sigma = n$. Ainsi $d$ et $n$ sont fixes sous $G$, donc dans $\mathrm{L}(E)$. On pose $s = d$ qui est bien semi-simple (puisque diagonalisable lorsqu'on passe sur $K$). Sur $k$, on a la décomposition $u = s + n$ souhaitée. ▫

**Contre-exemple 3 :** Nous reprenons la méthode du contre-exemple 2. Soit $k = \mathbb{F}_p(T)$ et l'algèbre $A = k[U, V]/(U^p - T, V^p)$ qui n'est pas un corps. Soient $u, v$ les images de $U, V$ dans $A$. On vérifie que $u$ est semi-simple de polynôme minimal irréductible $X^p - T$, $u + v$ est semi-simple de polynôme minimal $X^p - T$ également, et $v$ est nilpotent de polynôme minimal $X^p$. Ainsi on a $u = (u + v) - v$ ce qui met en défaut l'unicité de la décomposition $s + n$.

**Bibliographie :**

[BMP] Beck, Malick, Peyré, Objectif Agrégation, *H & K*. Précisément : Application 4.32 p. 160, exercice 4.23 p. 229 et exercice 6.8 p. 324.

[FGN2] Francinou, Gianella, Nicolas, Exercices de mathématiques d'oraux X-ENS : Algèbre, Tome 2, p. 122, *Cassini*.

[Gou] Gourdon, Les Maths en tête, Mathématiques pour M', *Ellipses*.

## Une équation différentielle quaternionique

Ce résultat peut faire l'objet d'un développement dans les leçons : Endomorphismes diagonalisables, et Exponentielle de matrices. Applications. Il faut aimer les calculs...

**Théorème :** *Soit $\mathbb{H}$ l'algèbre des quaternions, et $q$ un quaternion pur de norme 1. Les fonctions $f\colon \mathbb{R} \to \mathbb{H}$ de classe $C^1$ et vérifiant $f'(t) = qf(t)$ sont les fonctions de la forme :*

$$f(t) = \big(\cos(t) + q\sin(t)\big)f(0)$$

**Preuve :** Le calcul usuel pour montrer que les solutions de l'équation proposée sont de la forme $f(t) = \exp(qt)f(0)$ est de poser $g(t) = \exp(-qt)f(t)$ et de vérifier que $g'(t) = 0$. Cela fonctionne ici, mais comme $\mathbb{H}$ n'est pas commutatif il faut prendre soin de préciser que $q$ et $\exp(-qt)$ commutent :

$$g'(t) = -q\exp(-qt)f(t) + \exp(-qt)qf(t) = 0 \qquad \text{car } f' = qf.$$

Pour calculer l'exponentielle on utilise le plongement $\mathbb{H} \subset M_2(\mathbb{C})$ donné dans [Perrin] :

$$\phi\colon \mathbb{H} \xrightarrow{\sim} \left\{ M(a,b) = \begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix} \text{ avec } a,b \in \mathbb{C} \right\}$$

via $q = \alpha + \beta i + \gamma j + \delta k \mapsto M(\alpha + \beta i, \gamma - \delta i)$. Ce plongement est continu puisqu'en dimension finie, toutes les applications linéaires sont continues (ici entre deux $\mathbb{R}$-EV). Il en résulte que $\exp(\phi(qt)) = \phi(\exp(qt))$. Notons $q = \beta i + \gamma j + \delta k$ notre quaternion pur, et

$$M = \phi(q) = \begin{pmatrix} \beta i & -\gamma - \delta i \\ \gamma - \delta i & -\beta i \end{pmatrix}$$

son image dans $M_2(\mathbb{C})$. Pour calculer $\exp(tM)$, étudions d'abord $M$. On a $\mathrm{tr}(M) = 0$ et $\det(M) = \beta^2 - (-\delta^2 - \gamma^2) = |q|^2 = 1$ donc le

polynôme caractéristique est $\chi_M = X^2+1$. D'après Cayley-Hamilton on a $(M - i\,\mathrm{Id})(M + i\,\mathrm{Id}) = 0$ et donc

$$(M + i\,\mathrm{Id}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} (\beta+1)i \\ \gamma - \delta i \end{pmatrix}$$

resp.

$$(M - i\,\mathrm{Id}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} (\beta-1)i \\ \gamma - \delta i \end{pmatrix}$$

est un vecteur propre pour la valeur propre $i$ resp. $-i$. Observons que l'un de ces vecteurs est nul si $\gamma - \delta i = 0$, car alors $\beta = \pm 1$. Mais dans ce cas $q = \pm i$ et le calcul de exp est connu ; on suppose donc désormais $\omega := \gamma - \delta i \neq 0$. On prend la matrice de passage

$$P = \begin{pmatrix} (\beta+1)i & (\beta-1)i \\ \omega & \omega \end{pmatrix}$$

et pour calculer $P^{-1}$ on va utiliser Cayley-Hamilton là encore. On a $\mathrm{tr}(P) = (\beta+1)i + \omega$ et $\det(P) = 2i\omega$ donc $\chi_P = X^2 - ((\beta+1)i + \omega)X + 2i\omega$. On en tire $P(P - ((\beta+1)i + \omega)\,\mathrm{Id}) = -2i\omega$ donc

$$P^{-1} = \frac{-1}{2i\omega}\big(P - ((\beta+1)i + \omega)\,\mathrm{Id}\big) = \frac{-1}{2i\omega} \begin{pmatrix} -\omega & (\beta-1)i \\ \omega & -(\beta+1)i \end{pmatrix}$$

On a $P^{-1}MP = \mathrm{diag}(i, -i)$ donc $\exp(tM) = P\,\mathrm{diag}(e^{it}, e^{-it})P^{-1}$, le calcul donne

$$\exp(tM) = \frac{-1}{2i\omega} \begin{pmatrix} \omega[-(\beta+1)ie^{it} + (\beta-1)ie^{-it}] & * \\ \omega^2[-e^{it} + e^{-it}] & * \end{pmatrix}$$

Le calcul des coefficients $*$ n'est pas nécessaire puisqu'on sait que $\exp(tM) \in \mathbb{H}$, donc ces coefficients sont les conjugués ad hoc. En simplifiant ces expressions et en repassant dans $\mathbb{H}$ on trouve $\exp(tq) = \cos(t) + q\sin(t)$ d'où le résultat. □

# L'exponentielle de $\mathrm{SL}_n(\mathbb{C})$ n'est pas surjective

Soit $G$ un sous-groupe fermé de $\mathrm{GL}_n(\mathbb{C})$. On définit son algèbre de Lie (voir [MT]) par

$$\mathfrak{g} = \{\, M \in \mathrm{M}_n(\mathbb{C}) \text{ t.q. pour tout } t \in \mathbb{R},\ \exp(tM) \in G \,\}$$

L'algèbre de Lie de $\mathrm{GL}_n(\mathbb{C})$ est évidemment $\mathfrak{gl}_n(\mathbb{C}) = M_n(\mathbb{C})$. Par ailleurs, à partir de la formule $\det(\exp(M)) = \exp(\mathrm{tr}(M))$ on voit que l'algèbre de Lie de $\mathrm{SL}_n(\mathbb{C})$ est

$$\mathfrak{sl}_n(\mathbb{C}) = \{\, M \in \mathrm{M}_n(\mathbb{C}) \text{ t.q. } \mathrm{tr}(M) = 0 \,\}$$

En utilisant la forme de Jordan d'une matrice on démontre que $\exp\colon \mathfrak{gl}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ est surjective, voir [Gou]. Nous allons montrer que ce n'est pas un fait général :

**Théorème :** *L'exponentielle $\mathfrak{sl}_n(\mathbb{C}) \to \mathrm{SL}_n(\mathbb{C})$ n'est pas surjective.*

La démonstration nécessite deux petits lemmes intéressants pour eux-mêmes. On utilisera la décomposition $DU$ des matrices inversibles (diagonalisable × unipotent) que l'on obtient à partir de la décomposition de Dunford $D+N$ en mettant $D$ en facteur. Enfin rappelons que l'indice d'unipotence d'une matrice unipotente $U$ est défini comme étant égal à l'indice de nilpotence de $U - \mathrm{Id}$. L'exponentielle a un bon comportement vis-à-vis de cette décomposition :

**Lemme 1 :** *L'exponentielle d'une matrice nilpotente $N$ d'indice de nilpotence $n$ est unipotente d'indice d'unipotence $n$. L'exponentielle envoie la décomposition $D + N$ sur la décomposition $DU$ (mais le $D$ n'est pas le même !).*

**Preuve :** Posons $U = \exp(N)$. La première assertion découle de :

$$U - \mathrm{Id} = N + \frac{1}{2}N^2 + \cdots = N\,(\underbrace{\mathrm{Id} + \frac{1}{2}N + \ldots}_{\text{inversible}})$$

Ensuite, si $P = D + N$ avec $DN = ND$ alors $\exp(P) = \exp(D)\exp(N)$. On pose $D' = \exp(D)$ qui est diagonalisable, et $U = \exp(N)$ qui est unipotente. Il est clair que $D'U = UD'$. ▫

**Lemme 2 :** *Soit $N$ une matrice nilpotente d'indice de nilpotence $n$. Alors, toute matrice commutant avec $N$ est un polynôme en $N$. En particulier une telle matrice n'a qu'une valeur propre.*

**Preuve :** Puisque $N^{n-1} \neq 0$, il existe un vecteur $x \in \mathbb{C}^n$ tel que $N^{n-1}(x) \neq 0$. On vérifie que $\{1, x, N(x), \ldots, N^{n-1}(x)\}$ est une famille libre, donc une base de $\mathbb{C}^n$. Il en découle que tout vecteur $y \in \mathbb{C}^n$ s'écrit $y = \sum a_i N^i(x)$ et est donc l'image de $x$ par un polynôme en $N$, à savoir $P(N) = \sum a_i N^i$.

Soit $C$ une matrice commutant avec $N$. Par ce qui précède, il existe un polynôme $Q$ tel que $C(x) = Q(N)(x)$. Pour montrer que $C = Q(N)$, il suffit que montrer que $C(y) = Q(N)(y)$ pour tout vecteur $y$. Or on peut écrire $y = P(N)(x)$ pour un certain polynôme $P$, donc

$$C(y) = (C \circ P(N))(x) = (P(N) \circ C)(x)$$

(puisque $C$ commute avec $N$ et donc avec tout polynôme en $N$)

$$\cdots = P(N)(C(x)) = P(N)(Q(N)(x)) = Q(N)(P(N)(x)) = Q(N)(y)$$

On a obtenu $C = Q(N) = q_0 + q_1 N + \cdots + q_d N^d$. La somme des termes de degré $\geqslant 1$ est un endomorphisme nilpotent, donc en se plaçant dans une base de trigonalisation de $N$ on voit que $C$ a pour seule valeur propre $q_0$. ▫

**Preuve du théorème :** Soit $U \in \mathrm{M}_n(\mathbb{C})$ unipotente d'indice d'unipotence égal à $n$. Soit $\lambda$ une racine $n$-ème de l'unité, de sorte que $M := \lambda U \in \mathrm{SL}_n(\mathbb{C})$. Nous allons montrer que si $\lambda \neq 1$ alors $M$ n'est pas l'exponentielle d'une matrice de trace nulle.

Supposons avoir une matrice de trace nulle $P$ telle que $\exp(P) = M$. Soit la décomposition $P = D + N$ avec $D$ diagonalisable, $N$ nilpotente, et $DN = ND$. On a alors $\exp(P) = \exp(D)\exp(N)$ qui est la décomposition « diagonalisable × unipotent » (lemme 1). Or $M = \lambda U$ donc par unicité de la décomposition, on obtient $\exp(D) = \lambda$ et $\exp(N) = U$.

L'indice de nilpotence de $N$ est égal à l'indice d'unipotence de $U$ (lemme 1) donc $n$. D'après le lemme 2, la matrice $D$ n'a qu'une valeur propre $\mu$, donc $\mathrm{tr}(D) = n\mu$. Or $\mathrm{tr}(P) = \mathrm{tr}(D) = 0$ car $N$ est de trace nulle. Donc $\mu = 0$ et finalement $D = 0$. En conclusion $\exp(P) = \exp(N)$ ce qui contredit $M = \lambda \exp(N)$. ▫

**Remarques :** (1) Le résultat classique qui dit que l'exponentielle réalise un difféomorphisme

$$\{ \text{ matrices nilpotentes } \} \to \{ \text{ matrices unipotentes } \}$$

n'est pas formellement nécessaire pour ce qui précède, mais il apporte un éclairage intéressant sur le lemme 1 et c'est une bonne idée de l'avoir en tête. (Pour la preuve, lire [MT], chap. 3.)

(2) Le lemme 2 est un cas particulier d'un résultat concernant les endomorphismes *cycliques* (endomorphismes dont le polynôme minimal est de degré $n$) :

« $u$ est cyclique ssi les seuls endomorphismes commutant avec $u$ sont les polynômes en $u$ »

**Bibliographie :**

Je ne connais pas de référence pour ce développement, tel quel. On a utilisé :

[Gou] Gourdon, Les Maths en tête, Mathématiques pour M', *Ellipses.*

[MT] Mneimné, Testard, Introduction à la théorie des groupes de Lie classiques, *Hermann.*

# L'exponentielle de $\mathrm{SO}_n(\mathbb{R})$ est surjective

**Théorème :** *L'exponentielle* $\mathfrak{so}_n(\mathbb{R}) \to \mathrm{SO}_n(\mathbb{R})$ *est surjective.*

Nous allons d'abord décrire l'algèbre de Lie $\mathfrak{so}_n(\mathbb{R})$, puis démontrer le théorème après deux lemmes. La démonstration repose sur deux choses : on traite d'abord le cas $n = 2$, puis on s'y ramène en utilisant la réduction d'une isométrie en une matrice diagonale par blocs avec pour blocs des matrices de rotations planes.

**Lemme 1 :** $\mathfrak{so}_n(\mathbb{R})$ *est la sous-algèbre de Lie de* $\mathfrak{gl}_n(\mathbb{R}) = \mathrm{M}_n(\mathbb{R})$ *formée des matrices antisymétriques.*

**Preuve :** L'algèbre de Lie de $\mathrm{SO}_n(\mathbb{R})$ est la même que celle de $\mathrm{O}_n(\mathbb{R})$. Soient $S_n \subset \mathrm{M}_n(\mathbb{R})$ l'espace vectoriel des matrices symétriques, et $f : \mathrm{M}_n(\mathbb{R}) \to S_n$ définie par $f(M) = {}^tMM - \mathrm{Id}$. On a $\mathrm{O}_n(\mathbb{R}) = \{M \in \mathrm{M}_n(\mathbb{R}),\ f(M) = 0\}$. Donnons le calcul avec les deux descriptions de l'algèbre de Lie.

Dans la première méthode on montre que $f$ est une submersion en l'identité puis $\mathfrak{so}_n(\mathbb{R}) = \ker(d_{\mathrm{Id}}f)$. Or $f(\mathrm{Id}+H) = (\mathrm{Id}+{}^tH)(\mathrm{Id}+H) - \mathrm{Id} = {}^tH + H + {}^tHH$ de sorte que $d_{\mathrm{Id}}f(H) = {}^tH + H$. Ainsi $d_{\mathrm{Id}}f : \mathrm{M}_n(\mathbb{R}) \to S_n$ est surjective, puisque $A \in S_n$ est l'image de $(1/2)A$, donc $f$ est une submersion au voisinage de l'identité. On trouve bien $\mathfrak{so}_n(\mathbb{R}) = \ker(d_{\mathrm{Id}}f)$ = l'ensemble des matrices antisymétriques.

Dans la deuxième méthode on dit que $\mathfrak{so}_n(\mathbb{R})$ est l'ensemble des $H \in \mathrm{M}_n(\mathbb{R})$ telles que $\exp(uH) \in \mathrm{O}_n(\mathbb{R})$ pour tout $u \in \mathbb{R}$ (je note $u$ au lieu de $t$ pour ne pas risquer de confusion avec la transposition). Ceci s'exprime par

$$ {}^t\exp(uH)\exp(uH) = \exp(u\,{}^tH)\exp(uH) = \mathrm{Id}\ ,\ \forall u \in \mathbb{R}\ . $$

Les DL à l'ordre 1 en $u$ sont donc égaux de part et d'autre, d'où $\mathrm{Id} + u({}^tH + H) = \mathrm{Id}$. Il s'ensuit que ${}^tH + H = 0$, cqfd. □

Pour toute $\mathbb{R}$-algèbre unitaire, associative, de dimension finie $A$, et $x \in A$, on note $\exp_A(x)$ la somme de la série normalement convergente $\sum_{n\geqslant 0} x^n/n!$.

**Lemme 2 :** *Soit* $f : A \to B$ *un morphisme d'algèbres unitaires, associatives, de dimension finie. Alors pour tout* $x \in A$ *on a* $f(\exp_A(x)) = \exp_B(f(x))$.

**Preuve :** Comme $f$ est un morphisme d'algèbres on a $f(\sum_{n=0}^N x^n/n!) = \sum_{n=0}^N f(x)^n/n!$. De plus $f$ est continue comme toute application linéaire entre espaces vectoriels de dimension finie, donc en passant à la limite on trouve le résultat. □

**Preuve du théorème :** Nous prouvons d'abord le cas $n = 2$. Toute matrice de $\mathrm{SO}_2(\mathbb{R})$ est une matrice de rotation de la forme

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

Introduisons la matrice :

$$ I = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathfrak{so}_2(\mathbb{R})\ . $$

Comme $I^2 = -1$, l'expression $f(a + bi) = a + bI$ (on note $a$ au lieu de $a\,\mathrm{Id}$) définit un morphisme de $\mathbb{R}$-algèbres $f : \mathbb{C} \to \mathrm{M}_2(\mathbb{R})$. Le résultat en découle puisque d'après le lemme 2,

$$ \begin{aligned} \exp(\theta I) &= \exp(f(\theta i)) = f(\exp_{\mathbb{C}}(\theta i)) = f(\cos(\theta) + i\sin(\theta)) \\ &= \cos(\theta) + \sin(\theta)I = R_\theta\ . \end{aligned} $$

Pour $n$ quelconque, on utilise la réduction des matrices orthogonales. Pour toute matrice $M \in \mathrm{SO}_n(\mathbb{R})$ il existe une matrice orthogonale $P \in \mathrm{O}_n(\mathbb{R})$ telle que $PMP^{-1}$ soit de la forme diagonale par blocs suivante :

$$ \mathrm{diag}(\mathrm{Id}_r, R_{\theta_1}, \ldots, R_{\theta_s})\ . $$

D'après le cas $n = 2$ c'est donc l'exponentielle de la matrice diagonale par blocs

$$ \mathrm{diag}(0_r, \theta_1 I, \ldots, \theta_s I)\ , $$

qui est antisymétrique. Comme l'exponentielle respecte la conjugaison, $M$ est donc l'exponentielle de la matrice

$$P^{-1}\,\mathrm{diag}(0_r, \theta_1 I, \ldots, \theta_s I)\,P\ .$$

Comme $P$ est orthogonale, il est immédiat de vérifier que cette matrice est encore antisymétrique. □

**Corollaire :** $\mathrm{SO}_n(\mathbb{R})$ est connexe par arcs.

**Preuve :** C'est l'image par une application continue de $\mathfrak{so}_n(\mathbb{R})$ qui est un espace vectoriel, donc connexe par arcs. □

**Remarque :** Dans le calcul de $\mathfrak{so}_n(\mathbb{R})$ utilisant la submersion $f : \mathrm{M}_n(\mathbb{R}) \to S_n$ (lemme 1), il faut bien prendre garde que pour avoir $d_{\mathrm{Id}}f$ surjective, l'espace d'arrivée doit être $S_n$ et non $\mathrm{M}_n(\mathbb{R})$.

**Bibliographie :**

[MT] Mneimné, Testard, Introduction à la théorie des groupes de Lie classiques, *Hermann*.

# Sous-groupe de Frattini

Références : Calais, *Eléments de théorie des groupes* ou Rotman, *An Introduction to the theory of groups*.

Soit $G$ un groupe. On définit le *sous-groupe de Frattini de* $G$, noté $\Phi(G)$, comme étant l'intersection des sous-groupes maximaux de $G$. (Un sous-groupe strict $M \subset G$ est dit *maximal* s'il n'y a pas de sous-groupe compris strictement entre $M$ et $G$.)

Le but de ce qui suit est de montrer que si $G$ est un $p$-groupe fini, $\Phi(G)$ est engendré par les puissances $p$-èmes et les commutateurs. La preuve montre aussi que $G/\Phi(G)$ est un $\mathbb{F}_p$-espace vectoriel, et fournit sa dimension.

Pour un développement d'oral, on peut élaguer un peu ; par exemple, on peut sauter la question (1) ; le point (4) peut être considéré comme « connu » ; la question (3) ne sert qu'à montrer (4) ; la question (7) (n')est (qu')un bonus.

Leçons concernées :

Groupes opérant sur un ensemble. Exemples et applications.
Exemples de sous-groupes distingués et de groupes quotients. Applications.
Groupes finis. Exemples et applications.
Exemples de parties génératrices d'un groupe.
Dimension d'un espace vectoriel. Rang. Exemples et applications.

**L'exercice.**

(1) On dit qu'une partie $X$ d'un groupe $G$ est *superflue* si pour toute partie $A \subset G$, on a :

$$\langle A, X\rangle = G \Rightarrow \langle A\rangle = G\ .$$

Montrez que si $X$ est superflu, alors tout élément $x \in X$ est superflu. Montrez que si $G$ est de type fini, la réciproque est vraie. Donnez un contre-exemple lorsque $G$ n'est pas de type fini.

(2) Montrez que l'ensemble des éléments superflus de $G$ est égal à $\Phi(G)$.

(3) Dans toute la suite, $G$ est un $p$-groupe fini. Soit $H \subset G$ un sous-groupe et $N = N_G(H)$ son normalisateur. En faisant agir $H$ par conjugaison sur $G/H$, montrez que $H \neq G \Rightarrow H \neq N$.

(4) Déduisez-en que les sous-groupes maximaux de $G$ sont distingués et d'indice $p$.

(5) Notons $G^p[G,G]$ le sous-groupe engendré par les puissances $p$-èmes et les commutateurs. Montrez que $G^p[G,G] \subset \Phi(G)$.

(6) Soit $i\colon G/G^p[G,G] \to G/\Phi(G)$ le morphisme surjectif induit par l'inclusion précédente. En justifiant que $G/G^p[G,G]$ est un $\mathbb{F}_p$-espace vectoriel et en utilisant (1), montrez que $i$ est un isomorphisme puis que $G^p[G,G] = \Phi(G)$.

(7) On note $r$ le *rang* de $G$, c'est-à-dire le nombre minimal d'éléments d'un système de générateurs de $G$. Montrez que la dimension de $G/\Phi(G)$ comme $\mathbb{F}_p$-espace vectoriel est égale à $r$.

**Correction.**

(1) Supposons $X$ superflu, soient $x \in X$, $A \subset G$. Si $\langle A, x\rangle = G$ alors $\langle A, X\rangle = G$ donc $\langle A\rangle = G$. Donc $x$ est superflu.

On suppose maintenant $G$ de type fini et on suppose que tout élément $x \in X$ est superflu. Soit $A \subset G$ tel que $\langle A, X\rangle = G$. Choisissons un système générateur fini $g_1, \ldots, g_n$ de $G$, alors chaque $g_i$ s'écrit comme un produit d'éléments de $A \cup X$ et de leurs inverses ; chaque tel produit fait intervenir un nombre fini d'éléments de $X$ ; comme il y a un nombre fini de $g_i$, au total les éléments de $X$ qui sont en jeu forment un ensemble fini $X_0 \subset X$. Puisque les $g_i$ engendrent $G$, on a ainsi $\langle A, X_0\rangle = G$. Comme chaque $x \in X_0$ est superflu par hypothèse, on peut les enlever un à un, et comme $X_0$ est fini, après un nombre fini d'étapes on trouve $\langle A\rangle = G$. On a montré que $X$ est superflu.

Si $G$ n'est pas de type fini, il n'est pas vrai en général qu'une partie est superflue dès que tous ses éléments le sont. Par exemple, soit $G = \mathbb{Z}[1/n]$ le groupe additif des rationnels dont le dénominateur est une puissance de $n$. Posons $x_k = 1/n^k$ et $X = \{x_k\}_{k\geqslant 1}$. On voit que les sous-groupes monogènes $\langle x_k\rangle$ sont emboîtés et leur union est $G$. On en déduit que $X$ engendre $G$, alors que les $x_k$ sont tous superflus.

(2) Soit $x \in \Phi(G)$ et $A \subset G$ une partie. Si $A$ n'engendre pas $G$, alors par le lemme de Zorn, il existe un sous-groupe maximal $M$ contenant $\langle A\rangle$. Comme $x \in M$ par définition, on a donc $\langle A, x\rangle \subset M$ donc $\langle A, x\rangle \neq G$.

Réciproquement, si $x$ est un éléments superflu, soit $M$ un sous-groupe maximal. Comme $M$ n'engendre pas $G$ (c'est un sous-groupe strict), et que $x$ est superflu, alors $\langle M, x\rangle \neq G$. En particulier, $x \in M$, par maximalité de $M$. Donc $x \in \Phi(G)$.

(3) Faisons agir $H$ sur $G/H$ par $h.gH = hgh^{-1}H$. Ceci a bien un sens, car si on change le représentant $g$ pour un représentant $g' = gk$ de la même classe ($k \in H$), alors du fait que $hkh^{-1} \in H$, on a $h.g'H = hgkh^{-1}H = hgh^{-1}hkh^{-1}H = hgh^{-1}H = h.gH$.

Les orbites ponctuelles pour cette action sont les $\{gH\}$ tels que pour tout $h \in H$, on a $hgh^{-1}H = gH$. Ceci équivaut à $g^{-1}hgh^{-1}H = H$, ou encore $g^{-1}hg \in H$ pour tout $h$, i.e. $g^{-1}Hg = H$. Ce sont donc

les orbites $\{gH\}$ avec $gH \in N/H$. La formule des classes donne :

$$|G/H| = |N/H| + \sum |\omega|$$

où le premier terme compte les orbites de cardinal 1 (orbites ponctuelles), et le second, les orbites de cardinal $> 1$ (donc divisible par $p$). Modulo $p$, ceci donne $0 \equiv |N/H|$ $(p)$ et il en découle que $|N/H| \neq 1$, donc $H \neq N$.

(4) Soit $M$ un sous-groupe maximal de $G$. D'après (2), le normalisateur $N$ de $M$ contient strictement $M$, donc $N = G$ par maximalité, c'est-à-dire, $M$ est distingué. On regarde ensuite le quotient $\pi\colon G \to G/M$ et on choisit, par le lemme de Cauchy, un sous-groupe $H \subset G/M$ cyclique d'ordre $p$. La préimage de $H$ par $\pi$ est un sous-groupe qui contient strictement $M$, donc c'est $G$ par maximalité. Ainsi $G/M \simeq H$ est cyclique d'ordre $p$, cqfd.

(5) Soit $x \in G$ un élément qui est soit une puissance $p$-ème, soit un commutateur. Soit $M$ un sous-groupe maximal de $G$. Comme $G/M \simeq \mathbb{Z}/p\mathbb{Z}$ d'après (3), en particulier $G/M$ est abélien et d'exposant $p$ (c'est-à-dire que la puissance $p$-ème de tout élément est égale à 1). Donc l'image de $x$ dans $G/M$ est 1. Il en découle que $x \in \Phi(G)$. Ainsi, le sous-groupe $G^p[G,G]$, qui est engendré par les puissances $p$-èmes et les commutateurs, est inclus dans $\Phi(G)$.

(6) Par le procédé de quotient, dans le groupe $E = G/G^p[G,G]$, les puissances $p$-èmes et les commutateurs sont nuls. Donc ce groupe est abélien (on le notera donc additivement) et d'exposant $p$. Ainsi l'application $\mathbb{F}_p \times E \to E$, donnée par $(n,g) \mapsto ng$, est bien définie et fait de $E$ un $\mathbb{F}_p$-espace vectoriel. Notons $s$ sa dimension.

Il reste à voir que $i$ est injectif (donc un isomorphisme). Supposons qu'il existe un élément non nul $e_1$ dans le noyau de $i$. On peut compléter $e_1$ en une base $e_1, e_2, \ldots, e_s$ de $E$. Ce système engendre $E$, donc $i(e_2), \ldots, i(e_s)$ engendrent $G/\Phi(G)$ (rappelons-nous que $i(e_1) = 0$). Notons $x_i$ un antécédent de $e_i$ dans $G$. Il s'ensuit que la partie formée de $x_2, \ldots, x_s$ et $\Phi(G)$ engendre $G$. Comme $\Phi(G)$ est fini et ses éléments sont superflus d'après (2), on peut les enlever un à un tout en conservant une partie génératrice ; ainsi $x_2, \ldots, x_s$ engendre $G$. Ceci est impossible, car il en découlerait que $e_2, \ldots, e_s$ engendrent $E$, en contradiction avec le fait que $e_1, e_2, \ldots, e_s$ est une base. Donc $i$ est injectif, c'est un isomorphisme, donc $G^p[G,G] = \Phi(G)$.

(7) Soit $r$ le rang de $G$. En choisissant des éléments $x_1, \ldots, x_s$ dont les images dans $E = G/\Phi(G)$ forment une base, on voit que $x_1, \ldots, x_s$ et $\Phi(G)$ engendrent $G$. En utilisant le point (2), comme dans (6), on voit que $x_1, \ldots, x_s$ engendrent $G$. Donc $s \geqslant r$. Maintenant, soit $y_1, \ldots, y_r$ un système de générateurs de $G$, de cardinal minimal égal au rang. Alors les images des $y_i$ engendrent $E$, donc $r \geqslant s$. En conclusion $r = s$.

# Compter les mauvais prix

Supposons être dans un pays où il n'y a que deux types de pièces de monnaie, disons des pièces de $a$€ et $b$€, avec $a$ et $b$ premiers entre eux. Alors, d'après le théorème de Bézout, on peut payer en espèces tous les montants entiers, en demandant au besoin aux commerçants de rendre la monnaie. Si les commerçants ne rendent pas la monnaie, il y a des *mauvais prix* qu'on ne peut pas payer exactement.

**Théorème** *Le nombre de mauvais prix est fini et égal à*

$$\frac{1}{2}(a-1)(b-1)\ .$$

Pour la preuve, nous utiliserons l'*écriture de Bézout unique* suivante : pour tout entier $n \in \mathbb{Z}$ il existe un couple unique $(\lambda, \mu) \in \mathbb{Z}^2$ tel que $n = \lambda a + \mu b$ avec $0 \leqslant \lambda \leqslant b-1$. On l'obtient immédiatement en prenant une écriture de Bézout quelconque et en faisant la division euclidienne de $\lambda$ par $b$.

**Lemme 1** *Le plus grand mauvais prix est $m = ab - (a+b)$.*

**Preuve :** Démontrons que $ab-(a+b)$ est un mauvais prix. Supposons que $ab-(a+b)$ soit un bon prix, c'est-à-dire qu'il existe $(\lambda, \mu) \in \mathbb{N}^2$ tels que $ab-(a+b) = \lambda a + \mu b$. On en déduit $ab = (\lambda+1)a + (\mu+1)b$. Ainsi $a$ divise $(\mu+1)b$ donc, comme $a$ et $b$ sont premiers entre eux, par le lemme de Gauß on obtient que $a$ divise $\mu+1$. De même, on montre que $b$ divise $\lambda+1$ :

$$\exists \mu' \in \mathbb{N},\ \mu+1 = \mu' a$$
$$\exists \lambda' \in \mathbb{N},\ \lambda+1 = \lambda' b$$

On a donc $ab = \lambda' ab + \mu' ab$ d'où en divisant par $ab$ : $1 = \lambda' + \mu'$. Nécessairement, $\lambda' = 0$ ou $\mu' = 0$ (sinon leur somme serait au moins égale à 2). Cela implique alors que $\lambda + 1 = \lambda' b = 0$ (si $\lambda' = 0$) ou que $\mu + 1 = \mu' a = 0$ (si $\mu' = 0$). C'est impossible puisque $\lambda \geqslant 0$ et $\mu \geqslant 0$. Cela signifie donc que $ab-(a+b)$ est un mauvais prix.

Démontrons ensuite que tout mauvais prix $n$ vérifie $n \leqslant ab-(a+b)$. Considérons l'écriture de Bézout unique $n = \lambda a + \mu b$ avec $0 \leqslant \lambda \leqslant b-1$. Comme c'est un mauvais prix on doit avoir $\mu \leqslant -1$ et donc $n = \lambda a + \mu b \leqslant (b-1)a - b = ab - a - b$. ▫

**Lemme 2** *Soit $E = \{0, 1, \ldots, m\}$. Alors, l'involution $i\colon E \to E$ définie par $i(x) = m - x$ échange les mauvais prix et les bons prix.*

**Preuve :** Supposons que $x$ est un bon prix. Si $m-x$ était aussi un bon prix, alors la somme $x + (m-x) = m$ serait un bon prix. Ceci n'est pas vrai. Donc, $m-x$ est un mauvais prix.

Supposons maintenant que $x$ est un mauvais prix et montrons que $m-x$ est un bon prix. Considérons l'écriture de Bézout unique de $m-x$, c'est-à-dire $m-x = \lambda a + \mu b$ avec $(\lambda, \mu) \in \mathbb{Z}^2$ et $0 \leqslant \lambda \leqslant b-1$. Il suffit de montrer qu'alors $\mu \geqslant 0$. Or, on a

$$\begin{aligned} x &= m - (m-x) = m - (\lambda a + \mu b) = ab - (a+b) - (\lambda a + \mu b) \\ &= (b - (\lambda+1))a - (\mu+1)b\ . \end{aligned}$$

Posons $\lambda' := b - (\lambda+1)$ et $\mu' := -(\mu+1)$, de sorte que $x = \lambda' a + \mu' b$. Vu qu'on a choisi l'écriture de Bézout unique on a $\lambda' \geqslant 0$, si de plus $\mu' \geqslant 0$ alors $x$ serait un bon prix, contrairement à l'hypothèse. Ainsi, $\mu' < 0$ i.e. $\mu \geqslant 0$ qui est ce qu'on voulait montrer. Donc $m-x$ est un bon prix. ▫

La conclusion du théorème est maintenant évidente. ▫

# Nombre d'automorphismes diagonalisables sur un corps fini

Dans le développement proposé ici, on dénombre les matrices inversibles à coefficients dans un corps fini $\mathbb{F}_q$ qui sont diagonalisables. Ce développement peut être utilisé dans les leçons suivantes :

Groupes opérant sur un ensemble. Exemples et applications.
Groupes finis. Exemples et applications.
Groupe linéaire d'un espace vectoriel de dimension finie $E$, sous-groupes de $\mathrm{GL}(E)$.
Applications.
Nombres premiers. Applications.
Corps finis. Applications.
Endomorphismes diagonalisables.
Méthodes combinatoires, problèmes de dénombrement.

**Théorème :** *Soit $n \geqslant 1$ un entier. Alors le nombre de matrices diagonalisables dans le groupe linéaire $\mathrm{GL}_n(\mathbb{F}_q)$ sur le corps fini $\mathbb{F}_q$ est égal à*

$$\sum_{\substack{(n_1,\dots,n_{q-1}) \\ \text{t.q. } n_1+\cdots+n_{q-1}=n}} \frac{|\mathrm{GL}_n(\mathbb{F}_q)|}{|\mathrm{GL}_{n_1}(\mathbb{F}_q)|\dots|\mathrm{GL}_{n_{q-1}}(\mathbb{F}_q)|}\,.$$

La preuve utilise d'abord un exercice de [Gourdon] (page 176), puis adapte un argument que l'on trouve dans [FGN1] (*Nombre d'involutions*, page 17).

**Preuve :** On commence par observer qu'une matrice $A \in \mathrm{M}_n(\mathbb{F}_q)$ est diagonalisable si et seulement si $A^q - A = 0$. En effet si $A$ est diagonalisable, on peut écrire $A = PDP^{-1}$ avec $D$ diagonale. Comme les coefficients de $D$ sont dans $\mathbb{F}_q$ on a $D^q = D$ dont on déduit $A^q = A$. Réciproquement, si $A^q = A$, alors $A$ est annulé par le polynôme $X^q - X$, qui est à racines simples. Donc le polynôme minimal de $A$, diviseur de $X^q - X$, est à racines simples, donc $A$ est diagonalisable.

Si $A \in \mathrm{GL}_n(\mathbb{F}_q)$, alors $A$ est diagonalisable ssi $A^{q-1} = \mathrm{Id}$. Or on sait que le groupe multiplicatif $\mathbb{F}_q^\times$ est cyclique. Choisissons un générateur $\zeta$ : c'est donc une racine primitive $(q-1)$-ième de l'unité. Dès lors, on a la factorisation

$$X^{q-1} - 1 = (X-1)(X-\zeta)\dots(X-\zeta^{q-2})$$

On a donc $(A - \mathrm{Id})(A - \zeta\,\mathrm{Id})\dots(A - \zeta^{q-2}\,\mathrm{Id}) = 0$. Comme les polynômes $X - \zeta^i$ sont premiers entre eux, on en déduit que $E = \oplus E_i$ où $E_i = \ker(A - \zeta^i\,\mathrm{Id})$ pour $i = 0, \dots, q-2$. (On peut faire courrir $i$ de 1 à $q-1$, ce qui ne change rien et donne une notation plus agréable.) Soit $n_i = \dim(E_i)$, on a $n_1 + \cdots + n_{q-1} = n$. Réciproquement, étant donné un $(q-1)$-uplet de sous-espaces vectoriels qui décomposent $E$ en somme directe, l'automorphisme $A$ est complètement déterminé puisque sa restriction à $E_i$ est la multiplication par $\zeta^i$. On a donc une bijection entre l'ensemble des matrices diagonalisables et l'ensemble des tels uplets, pour $(n_1, \dots, n_{q-1})$ variable.

Pour chaque $N = (n_1, \dots, n_{q-1})$ fixé, notons $Z_N$ l'ensemble des $(q-1)$-uplets de sous-espaces vectoriels comme ci-dessus ; nous allons dénombrer $Z_N$. Il y a une action de $G = \mathrm{GL}_n(\mathbb{F}_q)$ sur $Z_N$, qui à $(E_i)$ associe $(g(E_i))$. Étant donnés des uplets $(E_i)$ et $(E'_i)$, on peut choisir des bases $(e_{i,j})$, $(e_{i,j})$ de $E_i$ resp. $E'_i$ (avec le même nombre d'éléments). On définit un automorphisme linéaire $g$ qui envoie $e_{i,j}$ sur $e'_{i,j}$, de sorte que $g.(E_i)) = (E'_i)$. Il en résulte que l'action de $G$ sur $Z_N$ n'a qu'une orbite. Par ailleurs, le stabilisateur de $(E_i)$ est constitué des automorphismes qui stabilisent chaque $E_i$, donc c'est le produit des $\mathrm{GL}_{n_i}(\mathbb{F}_q)$. Il s'ensuit que le cardinal de $Z_N$ est égal à

$$\frac{|\mathrm{GL}_n(\mathbb{F}_q)|}{|\mathrm{GL}_{n_1}(\mathbb{F}_q)|\dots|\mathrm{GL}_{n_{q-1}}(\mathbb{F}_q)|}$$

Le nombre de matrices diagonalisables dans $\mathrm{GL}_n(\mathbb{F}_q)$ est la somme des cardinaux des $Z_N$, ce qui donne le résultat. □

**Bibliographie**

[FGN1] Francinou, Gianella, Nicolas, Exercices de mathématiques des oraux de l'Ecole polytechnique et des Ecoles normales supérieures : Algèbre, Tome I, *Cassini*.
[Gourdon] Gourdon, Algèbre, *Ellipses*.

# Nombre d'endomorphismes nilpotents sur un corps fini

On dénombre les matrices carrées à coefficients dans un corps fini $\mathbb{F}_q$ qui sont nilpotentes d'indice maximal. Ce développement présente un lien avec les leçons suivantes :

Groupes opérant sur un ensemble. Exemples et applications.
Groupes finis. Exemples et applications.
Groupe linéaire d'un espace vectoriel de dimension finie $E$, sous-groupes de $\mathrm{GL}(E)$. Applications.
Corps finis. Applications.
Endomorphismes nilpotents.
Méthodes combinatoires, problèmes de dénombrement.

**Théorème :** *Soit $n \geqslant 1$ un entier et $\mathbb{F}_q$ un corps fini. Alors le nombre de matrices nilpotentes d'indice maximal $n$ dans l'anneau $\mathrm{M}_n(\mathbb{F}_q)$ est égal à $(q^n-1)(q^n-q)\dots(q^n-q^{n-2})$.*

La preuve utilise une propriété importante des endomorphismes *cycliques* (voir [Gourdon] p. 279 et suivantes, notamment p. 282). Une définition simple est que $f$ est cyclique si et seulement si son polynôme minimal est de degré $n$. Une définition équivalente est qu'il existe un vecteur $x$ tel que tout vecteur $y$ est l'image de $x$ par un polynôme en $f$. La propriété importante que nous utiliserons est qu'un endomorphisme $f$ est cyclique si et seulement si son commutant ([1]) se réduit à l'ensemble des polynômes en $f$.

On utilise ici un cas particulier puisqu'une matrice nilpotente d'indice $n$ a pour polynôme minimal $X^n$. Sur cet exemple, on peut démontrer directement la propriété sur le commutant ; c'est ce que je fais dans la preuve qui suit. Ce cas particulier est aussi utilisé dans la note sur la non-surjectivité de l'exponentielle de $\mathrm{SL}_n(\mathbb{C})$. Voir aussi la note sur les endomorphismes cycliques mise sur la page web.

[1] Le commutant de $f$ est l'ensemble des endomophismes qui commutent avec $f$.

**Preuve :** Considérons l'action par conjugaison de $\mathrm{GL}_n(\mathbb{F}_q)$ sur $\mathrm{M}_n(\mathbb{F}_q)$. Alors l'ensemble des endomorphismes nilpotents d'indice $n$ est égal à l'orbite de la matrice nilpotente

$$N_0 = \begin{pmatrix} 0 & \dots & \dots & \dots & 0 \\ 1 & 0 & & & \vdots \\ 0 & 1 & 0 & & \vdots \\ \vdots & & \ddots & 0 & \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}$$

En effet, si $N$ est nilpotente d'indice $n$, c'est-à-dire $N^n = 0$ mais $N^{n-1} \neq 0$, alors on peut choisir un vecteur $x$ qui ne soit pas dans $\ker N^{n-1}$. On vérifie que $\{1, x, N(x), \dots, N^{n-1}(x)\}$ est une famille libre, donc une base de $(\mathbb{F}_q)^n$. La matrice de $N$ dans cette base est la matrice ci-dessus, ce qui démontre l'assertion.

Il reste à calculer le cardinal du stabilisateur de cette matrice, puis à utiliser la formule $|\mathcal{O}(x)| = |G|/|G_x|$. On va montrer que le stabilisateur en question, qui est le commutant de $N_0$ dans $\mathrm{GL}_n(\mathbb{F}_q)$, est égal à l'ensemble des polynômes en $N_0$. (En fait tout sera vrai pour une matrice nilpotente $N$ d'indice $n$, je note donc $N$ au lieu de $N_0$).

On choisit encore $x \notin \ker N^{n-1}$ et on observe que comme $\{1, x, N(x), \dots, N^{n-1}(x)\}$ est une base, tout vecteur $y$ s'écrit $y = \sum a_i N^i(x)$ et est donc l'image de $x$ par un polynôme en $N$, à savoir $P(N) = \sum a_i N^i$.

Soit $C$ une matrice commutant avec $N$. Par ce qui précède pour $y = C(x)$, il existe un polynôme $P$ tel que $C(x) = P(N)(x)$. Pour montrer que $C = P(N)$, il suffit de montrer que pour tout vecteur $y$ on a $C(y) = P(N)(y)$. Or $y = Q(N)(x)$ pour un certain polynôme $Q$, donc

$$C(y) = (C \circ Q(N))(x) = (Q(N) \circ C)(x)$$

(puisque $C$ commute avec $N$ et donc avec tout polynôme en $N$)

$$\cdots = Q(N)(C(x)) = Q(N)(P(N)(x)) = P(N)(Q(N)(x)) = P(N)(y) \ .$$

On a obtenu $C = P(N) = a_0 + a_1 N + \cdots + a_d N^d$, un polynôme en $N$.

Pour finir considérons le morphisme $\mathbb{F}_q[X] \to \mathrm{M}_n(\mathbb{F}_q)$ qui à l'indéterminée $X$ associe $N$. Son noyau est l'idéal engendré par le polynôme minimal de $N$ i.e. $X^n$. Donc le sous-espace $\mathbb{F}_q[N] \subset \mathrm{M}_n(\mathbb{F}_q)$ des polynômes en $N$ est isomorphe à $\mathbb{F}_q[X]/X^n$, de dimension $n$ sur $\mathbb{F}_q$. Pour une matrice $C = a_0 + a_1 N + \cdots + a_{n-1} N^{n-1}$, si $a_0 = 0$, alors $C$ est nilpotente, donc non inversible. Si $a_0 \neq 0$, $C$ est la somme de la matrice d'homothétie $a_0 \,\mathrm{Id}$, inversible, et d'une matrice nilpotente, les deux commutant entre elles ; $C$ est donc inversible. En conclusion l'ensemble des matrices inversibles commutant avec $N$ est déterminé par la seule condition $a_0 \neq 0$, il possède $(q-1)q^{n-1}$ éléments. Donc le cardinal recherché est

$$\begin{aligned} \tfrac{|\mathrm{GL}_n(\mathbb{F}_q)|}{(q-1)q^{n-1}} &= \tfrac{(q^n-1)(q^n-q)\dots(q^n-q^{n-1})}{(q-1)q^{n-1}} \\ &= (q^n-1)(q^n-q)\dots(q^n-q^{n-2}) \ . \end{aligned}$$

□

## Bibliographie

[Gourdon] Gourdon, Algèbre, *Ellipses*.

# Polynômes invariants sous le groupe alterné

Soient $A$ un anneau commutatif unitaire et $n \geqslant 2$ un entier. Le groupe symétrique $\mathfrak{S}_n$ agit sur l'anneau de polynômes $A[X_1, \ldots, X_n]$ en permutant les variables, et un polynôme invariant pour cette action est dit *symétrique*. Un polynôme qui est invariant pour l'action restreinte du groupe alterné $\mathfrak{A}_n$ est dit *alterné* ; donc tout polynôme symétrique est alterné. Le théorème fondamental des fonctions symétriques dit que l'anneau des polynômes symétriques est engendré par les fonctions symétriques élémentaires $S_1^n, \ldots, S_n^n$. Qu'en est-il pour l'anneau des polynômes alternés ? Son calcul est classique lorsque $A$ est le corps des nombres complexes, ou d'ailleurs n'importe quel anneau dans lequel 2 est inversible ; mais il est plus original sur un anneau de base quelconque. On peut trouver ce développement dans le livre [S], page 602.

Ce développement peut être utilisé dans les leçons :

Groupes opérant sur un ensemble. Exemples et applications.
Groupes finis. Exemples et applications.
Groupe des permutations d'un ensemble fini. Applications.
Algèbre des polynômes à $n$ indéterminées ($n \geqslant 2$).
Polynômes symétriques. Applications.

## 1 Cas où 2 est inversible dans $A$

Rappelons que la fonction symétrique élémentaire de degré $i$ en $n$ variables est définie par

$$S_i = \sum_{\{\alpha_1 < \cdots < \alpha_i\}} X_{\alpha_1} \ldots X_{\alpha_i} \ ,$$

où la somme est étendue à toutes les parties ordonnées de $i$ éléments de $\{1, \ldots, n\}$. Nous la notons $S_i$ plutôt que $S_i^n$, car dans ce complément le nombre de variables sera toujours $n$. Nous introduisons aussi le polynôme de Vandermonde défini par

$$V_n = \prod_{i>j} (X_i - X_j) \ .$$

Ce polynôme est invariant par les permutations paires, et une permutation impaire change $V_n$ en $-V_n$. Ainsi, sur un corps de caractéristique différente de 2, c'est un polynôme alterné, non symétrique. Pour déterminer tous les polynômes alternés, nous aurons besoin de quelques remarques sur les calculs dans $A[X_1, \ldots, X_n]$ et plus particulièrement sur la divisibilité par les polynômes $X_i - X_j$. Lorsque $A$ est un anneau factoriel, par exemple un corps, il est clair que ces polynômes sont des irréductibles distincts. En conséquence, si un polynôme en $n$ variables est divisible par tous les $X_i - X_j$ il est divisible par le polynôme de Vandermonde. En fait, si $A$ est quelconque (pas nécessairement factoriel ni même intègre), il est facile de vérifier par des calculs directs que cette affirmation reste vraie :

**Lemme 1** *Soit $A$ un anneau commutatif et unitaire, et $F \in A[X_1, \ldots, X_n]$. Alors,*
(i) *$X_i - X_j$ est régulier dans $A[X_1, \ldots, X_n]$, pour tous $i, j$. De manière équivalente, le polynôme de Vandermonde $V_n$ est régulier.*
(ii) *Si $F(..., X, ..., X, ...) = 0$, l'indéterminée $X$ étant aux places $i$ et $j$, alors $F$ est divisible par $X_i - X_j$.*
(iii) *Si $F$ est divisible par tous les polynômes $X_i - X_j$ pour $i > j$, il est divisible par le polynôme de Vandermonde.*

Nous utiliserons fréquemment le fait général et élémentaire suivant : un polynôme à coefficients dans un anneau $R$, de coefficient dominant régulier dans $R$, est régulier.

**Preuve :** (i) La remarque que l'on vient de faire, avec $R = A[X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n]$ et $X = X_i$, nous dit que le polynôme $X_i - X_j$ est polynôme en $X_i$ de coefficient dominant égal à 1, donc il est régulier dans $A[X_1, \ldots, X_n]$. Le résultat en découle pour $V_n$, car un produit d'éléments réguliers est régulier.

(ii) En faisant agir une permutation $\sigma$ qui envoie $i$ sur 1 et $j$ sur 2, on se ramène au cas $i = 1$, $j = 2$, ce qui simplifie les notations. Effectuons la division euclidienne de $F$, vu comme un polynôme en $X_1$ à coefficients des polynômes en les autres variables, par $X_1 - X_2$. On trouve $F = (X_1 - X_2)Q + R$ où $R$ est un polynôme constant en $X_1$, c'est-à-dire $R = R(X_2, X_3, \dots)$. Lorsqu'on spécialise $X_1$ et $X_2$ en $X$ comme indiqué, on trouve $0 = 0 + R(X, X_3, \dots)$. Comme $X$ est une indéterminée, ceci donne $R = 0$ de sorte que $X_1 - X_2$ divise $F$.
(iii) Les couples d'entiers $(i, j)$ tels que $1 \leqslant j < i \leqslant n$ sont en nombre fini égal à $N = n(n-1)/2$. Munissons l'ensemble de ces couples de l'ordre lexicographique, et pour tout $k \leqslant N$, notons $E_k$ l'ensemble des $k$ premiers couples. On va montrer par récurrence sur $k$ qu'il existe un polynôme $Q_k \in A[X_1, \dots, X_n]$ tel que

$$F = \left( \prod_{(i,j) \in E_k} (X_i - X_j) \right) Q_k \ .$$

Pour $k = 1$, c'est une simple conséquence de l'hypothèse. Supposons maintenant que cette propriété est vraie pour un entier $k \leqslant N - 1$. Notons $(u, v)$ le $k+1$-ème couple d'entiers et $\Pi$ le produit des $X_i - X_j$ avec $(i, j) \in E_k$. On sait que $X_u - X_v$ ne divise pas $\Pi$, mais $X_u - X_v$ divise $F$ par hypothèse. Donc lorsqu'on fait $X_u = X_v$, le produit $\Pi$ ne s'annule pas, et est régulier d'après le point (i), mais $F$ s'annule. Il s'ensuit que $Q_k$ s'annule. D'après le point (ii), il existe un polynôme $Q_{k+1}$ tel que $Q_k = (X_u - X_v)Q_{k+1}$ et la propriété est vraie pour $k + 1$. L'assertion à démontrer est le cas $k = N$. □

**Lemme 2** *Soit $A$ un anneau commutatif et unitaire, $B$ la $A$-algèbre des polynômes symétriques en $X_1, \dots, X_n$ et $C$ la $B$-algèbre des polynômes alternés. Alors il existe un $B$-automorphisme d'algèbres involutif*

$$\begin{array}{ccc} C & \to & C \\ F & \mapsto & \overline{F} \end{array}$$

*défini en choisissant une permutation impaire quelconque $\sigma$ et en posant $\overline{F} = \sigma F$. Le polynôme $\overline{F}$ est appelé le* conjugué *de $F$. De plus, $F$ est symétrique si et seulement si $\overline{F} = F$.*

En particulier, on a $\overline{V}_n = -V_n$.

**Preuve :** Soit $F$ un polynôme alterné, $\sigma$ une permutation impaire et $\overline{F} = \sigma F$. Si $\sigma'$ est une autre permutation impaire, la permutation $\sigma^{-1}\sigma'$ est paire donc $(\sigma^{-1}\sigma')F = F$, de sorte que $\sigma' F = \sigma(\sigma^{-1}\sigma')F = \sigma$. Ceci montre que $\overline{F}$ est indépendant du choix de la permutation impaire $\sigma$. Par ailleurs, si $\tau$ est une permutation paire, alors $\sigma^{-1}\tau\sigma$ l'est aussi et donc

$$\tau\overline{F} = \sigma(\sigma^{-1}\tau\sigma F) = \sigma F = \overline{F} \ .$$

Ceci montre que $\overline{F}$ est alterné. Il est immédiat de vérifier que $\overline{\overline{F}} = F$ et que $F$ est symétrique si et seulement si $\overline{F} = F$. □

Nous avons en main les outils pour démontrer le résultat suivant :

**Théorème 1** *Soit $A$ un anneau commutatif et unitaire tel que 2 est inversible dans $A$. Alors l'anneau des polynômes alternés est engendré par les fonctions symétriques élémentaires et le polynôme de Vandermonde. Plus précisément, notons $B = A[S_1, ..., S_n]$ l'algèbre des polynômes symétriques et $\Delta_n \in B$ le polynôme égal au carré du polynôme de Vandermonde. Alors, l'anneau des polynômes alternés est isomorphe comme $B$-algèbre à $B[T]/(T^2 - \Delta_n)$, par un isomorphisme qui envoie $V_n$ sur la classe de $T$.*

L'expression du polynôme $\Delta_n$ nous est familière : ce n'est rien d'autre que le discriminant du polynôme $\prod_{i=1}^{n}(T - X_i)$, vu comme polynôme en $T$.

**Preuve :** Si $F$ est un élément de $C$, on note $B \cdot F$ le sous-$B$-module qu'il engendre. Si $F$ est régulier, le morphisme surjectif de $B$-modules $B \to B \cdot F$, $b \mapsto bF$ est injectif, donc c'est un isomorphisme. Dans

$C$, les éléments 1 et $V_n$ sont réguliers (voir lemme 1(i)), donc les sous-modules $B \cdot 1$ et $B \cdot V_n$ sont isomorphes à $B$.

Ces sous-modules sont en somme directe, car si $P + V_n R = 0$, alors en passant au conjugué on trouve $P - V_n R = 0$, dont on déduit $P = R = 0$.

Nous allons vérifier qu'ils engendrent $C$. Soit $F$ un polynôme alterné et $\overline{F}$ son conjugué, comme défini dans le lemme 1. On considère les deux polynômes alternés $P = F + \overline{F}$ et $Q = F - \overline{F}$. Pour calculer $\overline{F}$, on peut utiliser n'importe quelle permutation impaire $\sigma$. Si on choisit $\sigma = (ij)$, on obtient

$$Q = F - \sigma F = F(..., X_i, ..., X_j, ...) - F(..., X_j, ..., X_i, ...)$$

donc $Q(..., X, ..., X, ...) = 0$, l'indéterminée $X$ étant aux places $i$ et $j$. D'après le point (ii) du lemme ci-dessus, on en déduit que $Q$ est divisible par $X_i - X_j$. Comme ceci est vrai pour tous $i, j$, d'après le point (ii) du lemme on trouve que $Q$ est divisible par le polynôme de Vandermonde : il existe un polynôme $R$ tel que $Q = V_n R$. Voyons maintenant que $P$ et $R$ sont symétriques. Pour $P$ c'est clair, car $\overline{P} = \overline{F} + F = P$. Par ailleurs, on a $\overline{Q} = \overline{F} - F = -Q$ et comme $\overline{V}_n = -V_n$, on trouve $V_n \overline{R} = V_n R$. D'après le point (i) du lemme, $V_n$ est régulier, donc $\overline{R} = R$, et $R$ est symétrique. Comme $F = \frac{1}{2}(P+Q) = \frac{1}{2}P + \frac{1}{2}V_n R$, on obtient bien que $B \cdot 1$ et $B \cdot V_n$ engendrent $C$. On a montré que l'algèbre $C$ est engendrée par les fonctions symétriques élémentaires et le polynôme $V_n$.

Il nous reste à définir un isomorphisme $B[T]/(T^2 - \Delta_n) \simeq C$. Définissons un morphisme de $B$-algèbres $f : B[T] \to C$ par $f(T) = V_n$. Comme l'image de $T^2 - \Delta_n$ par $f$ est $(V_n)^2 - \Delta_n = 0$, alors $f$ induit un morphisme $\widetilde{f} : B[T]/(T^2 - \Delta_n) \to C$. Le $B$-module $B[T]/(T^2 - \Delta_n)$ est libre de rang 2 avec pour base 1 et l'image de $T$, et $\widetilde{f}$ envoie cette base sur la base $\{1, V_n\}$ donc c'est un isomorphisme. □

## 2 Cas général

Le 2 qui pose problème s'explique bien sûr par le fait que c'est l'indice de $\mathfrak{A}_n$ dans $\mathfrak{S}_n$. Introduisons le polynôme symétrique

$$\Theta_n = \prod_{i>j} (X_i + X_j) \ .$$

Lorsque 2 n'est pas inversible, il n'est plus vrai que $V_n$ engendre l'anneau des polynômes alternés : en fait, si $A$ est de caractéristique 2, on a $+1 = -1$ donc $V_n = \Theta_n$ est symétrique. On va maintenant s'affranchir de l'hypothèse $2 \in A^\times$.

Pour n'importe quel anneau unitaire $A$, il y a un morphisme $\mathbb{Z} \to A$, qui à $n$ associe $n1$. En prenant les images des coefficients par ce morphisme, on peut voir tout polynôme $P$ à coefficients dans l'anneau des entiers $\mathbb{Z}$ comme un polynôme à coefficients dans $A$. Pour simplifier, on le note encore par la même lettre $P$, lorsque l'anneau dans lequel on considère les coefficients de $P$ est clair d'après le contexte.

Considérons d'abord $V_n$ et $\Theta_n$ à coefficients dans $\mathbb{Z}$. D'après notre remarque introductive, lorsqu'on réduit modulo 2 pour obtenir des polynômes à coefficients dans le corps fini $\mathbb{F}_2$ on a :

$$V_n + \Theta_n = 2\Theta_n = 0 \ .$$

Il s'ensuit qu'il existe un polynôme $W_n$ à coefficients dans $\mathbb{Z}$ tel que $2W_n = V_n + \Theta_n$. Nous appellerons ce polynôme $W_n$ le *polynôme de Vandermonde modifié*. Pour tout anneau $A$, on peut le voir comme un polynôme à coefficients dans $A$. Nous allons démontrer un analogue du théorème 1, valable pour n'importe quel anneau commutatif, en remplaçant $V_n$ par $W_n$.

**Exemple 1** *Lorsque $n = 2$, on trouve :*

$$W_2 = \frac{1}{2}\big((Y - X) + (Y + X)\big) = Y \ .$$

*Lorsque $n = 3$, on trouve après un bref calcul :*

$$\begin{aligned} W_3 &= \tfrac{1}{2}\big((Z - Y)(Z - X)(Y - X) + (Z + Y)(Z + X)(Y + X)\big) \\ &= YZ^2 + XY^2 + ZX^2 + XYZ \ . \end{aligned}$$

Il sera utile de collecter deux petits renseignements sur $W_n$ :

**Lemme 3** *Soit $A$ un anneau commutatif unitaire, et*

$$W_n = \frac{1}{2}\left(\prod_{i>j}(X_i - X_j) + \prod_{i>j}(X_i + X_j)\right)$$

*le polynôme de Vandermonde modifié. Alors,*
(i) *$W_n$ est régulier dans $A[X_1, \ldots, X_n]$.*
(ii) *Le conjugué de $W_n$ est $\overline{W}_n = \Theta_n - W_n = W_n - V_n$.*

**Preuve :** On démontre le point (i) par récurrence sur $n \geqslant 2$. Pour $n = 2$ le résultat est clair. Pour $n \geqslant 3$, on calcule le coefficient dominant de $W_n$ vu comme polynôme en $X_n$. Pour obtenir des termes de degré maximal en $X_n$, lorsqu'on développe le produit qui définit $V_n$, on doit retenir $X_n$ dans chacun des facteurs $(X_n - X_i)$ pour $n > i$. Le coefficient devant $X_n$ est donc

$$\prod_{n>i>j}(X_i - X_j) \ ,$$

c'est-à-dire $V_{n-1}$. De même, le coefficient dominant de $\Theta_n$ est $\Theta_{n-1}$, de sorte que le coefficient dominant de $W_n$ est $W_{n-1}$. D'après l'hypothèse de récurrence, $W_{n-1}$ est régulier dans $A[X_1, \ldots, X_{n-1}]$, et ceci implique que $W_n$ est régulier dans $A[X_1, \ldots, X_n]$.

Pour démontrer le point (ii), on raisonne d'abord dans l'anneau des polynômes à coefficients dans $\mathbb{Z}$. Une permutation impaire $\sigma$ envoie $2W_n = V_n + \Theta_n$ sur $-V_n + \Theta_n = 2(\Theta_n - W_n)$, en d'autres termes,

$$\overline{W}_n = \Theta_n - W_n = W_n - V_n \ .$$

En prenant les images par le morphisme $\mathbb{Z} \to A$, ces expressions restent valables dans l'anneau $A$. □

Nous introduisons enfin le polynôme $\Gamma_n := W_n\overline{W}_n$. Ce polynôme est clairement symétrique, et en utilisant le fait que $\overline{V}_n = -V_n$ et $\overline{\Theta}_n = \Theta_n$, on trouve $\Theta_n = W_n + \overline{W}_n$. En remplaçant $\Gamma_n$ par sa valeur de définition, il vient l'identité :

$$W_n^2 - \Theta_n W_n + \Gamma_n = 0 \ .$$

Les polynômes $\Theta_n$ et $\Gamma_n$ sont appelés la *trace* et la *norme* de $W_n$.

**Théorème 2** *Soit $A$ un anneau commutatif et unitaire. Alors l'anneau des polynômes alternés est engendré par les fonctions symétriques élémentaires et le polynôme de Vandermonde modifié $W_n$. Plus précisément, notons $B = A[S_1, ..., S_n]$ l'algèbre des polynômes symétriques et $\Theta_n, \Gamma_n$ les éléments de $B$ introduits ci-dessus. Alors, l'anneau des polynômes alternés est isomorphe comme $B$-algèbre à $B[T]/(T^2 - \Theta_n T + \Gamma_n)$, par un isomorphisme qui envoie $W_n$ sur la classe de $T$.*

**Preuve :** Notons $B \cdot 1$ et $B \cdot W_n$ les sous-$B$-modules de $C$ engendrés par 1 et $W_n$. Comme 1 et $W_n$ sont réguliers (lemme 3), ces deux sous-modules sont isomorphes à $B$. De plus ils sont en somme directe : si $P + W_n R = 0$, alors en passant au conjugué on trouve $P + (W_n - V_n)R = 0$. En soustrayant ces égalités, on trouve $-V_n R = 0$. Comme $V_n$ est régulier d'après le point (i) du lemme 1, il vient $R = 0$, puis $P = 0$.

Nous allons vérifier que $B \cdot 1$ et $B \cdot W_n$ engendrent $C$. Soit $F$ un polynôme alterné et posons $Q = F - \overline{F}$. En procédant comme dans la preuve du théorème 1, on montre que $Q$ est divisible par $V_n$, donc il existe un polynôme $R$ tel que $Q = V_n R$. Ce polynôme est symétrique. Le polynôme $P = F - W_n R$ vérifie

$$\overline{P} = \overline{F} - \overline{W}_n R = (F - V_n R) - (W_n - V_n)R = F - W_n R = P \ ,$$

donc il est symétrique également. Finalement, on a obtenu $F = P + W_n R$ avec $P$ et $R$ symétriques, d'où le résultat d'engendrement annoncé.

Pour finir, définissons un morphisme de $B$-algèbres $f : B[T] \to C$ par $f(T) = W_n$. D'après la définition de $\Gamma_n$, l'image de $T^2 - \Theta_n T + \Gamma_n$

par $f$ est $(W_n)^2 - \Theta_n W_n + \Gamma_n = 0$. Ainsi $f$ induit un morphisme $\widetilde{f} : B[T]/(T^2 - \Theta_n T + \Gamma_n) \to C$. Le $B$-module $B[T]/(T^2 - \Theta_n T + \Gamma_n)$ est libre de rang 2 avec pour base 1 et l'image de $T$, et $\widetilde{f}$ envoie cette base sur la base $\{1, W_n\}$ donc c'est un isomorphisme. ▫

**Bibliographie :**


[AB] Arnaudiès, Bertin, Groupes, Algèbres et Géométrie, tome II, *Ellipses*.

Calcul des polynômes invariants sous le groupe alterné sur le corps des complexes :

[Gob] Goblot, Algèbre commutative, *Masson*.

Preuve du théorème des fonctions symétriques avec une récurrence *simple* (on le trouve partout avec une récurrence double) :

[LS] Leichtnam, Schauer, Exercices corrigés de Mathématiques posés aux oraux X-ENS, Algèbre I, p. 53, *Ellipses*.

[S] A. Szpirglas, Mathématiques algèbre L3, Cours complet avec 400 tests et exercices corrigés, *Pearson*.


# $\mathrm{SL}(E)$ est engendré par les transvections

Je propose ici une (re-)lecture de la démonstration qui est dans [Perrin] du fait que les transvections engendrent $\mathrm{SL}(E)$. Je vous conseille d'accompagner les résultats qui suivent de dessins pour bien les comprendre !

Soit $E$ un espace vectoriel de dimension finie sur un corps $k$. Une *transvection* est un endomorphisme $u \in \mathrm{SL}(E)$ tel que $u(t) = t + f(t)a$ où $f$ est une forme linéaire sur $E$ et $a \in \ker(f)$, $a \neq 0$. L'hyperplan $H = \ker(f)$ est déterminé par $u$ de manière unique.

**Théorème :** *Les transvections engendrent* $\mathrm{SL}(E)$.

**Lemme 1 :** *Soit $x \in E - \{0\}$ et $H_1, H_2$ hyperplans distincts tels que $x \notin H_1 \cup H_2$. Alors il existe une transvection $u$ telle que $u(x) = x$ et $u(H_1) = H_2$.*

**Preuve :** L'idée est de chercher une transvection $u$ qui fixe $x$ et $H_1 \cap H_2$, c'est-à-dire une transvection d'hyperplan $H = H_1 \cap H_2 + kx$. Il suffit alors de trouver des droites vectorielles $ky$ resp. $kz$, supplémentaires de $H_1 \cap H_2$ dans $H_1$ resp. $H_2$, avec $u(y) = z$, pour avoir le résultat. De l'hypothèse sur $x$ il résulte que $H + H_1 = H + H_2 = E$. Soit $z \in H_2 - H$, on peut donc l'écrire

$$z = a + y \qquad \text{avec} \qquad a \in H \text{ et } y \in H_1.$$

Soit $f$ la forme linéaire équation de $H$ telle que $f(y) = 1$ : elle existe car $y \notin H$. Soit $u$ la transvection définie par $u(t) = t + f(t)a$. Alors $u(y) = y + a = z$ et c'est gagné. ▫

**Lemme 2 :** *Supposons* $\dim(E) \geqslant 2$ *et soient $x, y \in E - \{0\}$. Alors il existe $u$, produit de une ou deux transvections, tel que $u(x) = y$.*

**Preuve :** Si $x$ et $y$ ne sont pas colinéaires on peut choisir un hyperplan $H$ contenant $y - x$ mais pas $x$. On pose $a = y - x$ et $u(t) = t + f(t)a$ où $f$ est une équation de $H$ telle que $f(x) = 1$. On a alors $u(x) = x + y - x = y$.

Si $x$ et $y$ ne sont pas colinéaires, on choisit un $z$ qui ne leur est pas colinéaire, et d'après ce qui précède il existe deux transvections $u_1, u_2$ telles que $u_1(x) = z$ et $u_2(z) = y$ donc $u = u_2 \circ u_1$ convient. (C'est ici qu'on utilise $\dim(E) \geqslant 2$.) ▫

**Preuve du théorème :** On fait une récurrence sur $n = \dim(E)$. Si $n = 1$ il n'y a rien à démontrer. Si $n \geqslant 2$, soit $v \in \mathrm{SL}(E)$. Soit $x \in E - \{0\}$, par le lemme 2 quitte à composer avec une ou deux transvections on peut supposer que $v(x) = x$. Soit ensuite $H \subset E$ un hyperplan tel que $x \notin H$. Alors $x = v(x) \notin v(H)$ donc d'après le lemme 1, quitte à composer avec une autre transvection on peut supposer que $v(H) = H$. L'hypothèse de récurrence appliquée à $v_{|H}$ nous dit que $v_{|H}$ est un produit de transvections $u_{i,H}$. Chacune de ces transvections s'étend en une unique transvection $u_i$ de $E$ qui fixe $x$. Comme $v(x) = x$ on obtient que $v$ est le produit des $u_i$. ▫

**Bibliographie :**

[Perrin] Perrin, Cours d'Algèbre, *Ellipses*.

**THÈMES**

# Factorisation des polynômes sur les corps finis

## 1 L'algorithme de Berlekamp

Soit $K = \mathbb{F}_q$ un corps fini, et soit $P \in K[X]$ un polynôme. On souhaite factoriser $P$ ; l'algorithme de Berlekamp prend $P$ en entrée, et ressort soit $P$ si celui-ci est irréductible, soit un diviseur non trivial de $P$. Les étapes sont les suivantes :

(1) si $P' = 0$ alors il existe $Q$ tel que $P(X) = Q(X^p)$. Soit $R$ le polynôme dont les coefficients sont les racines des coefficients de $Q$ (bien déterminés car le Frobenius $F$ est un isomorphisme de $K$), alors $P(X) = (R(X))^p$ et l'algorithme renvoie $R$ et s'arrête.

(2) si $P \wedge P' \neq 1$ alors c'est un facteur non trivial de $P$ donc l'algorithme renvoie $P \wedge P'$ et s'arrête.

(3) sinon, $P$ a ses facteurs irréductibles distincts. Soit $A = K[X]/P$ qui est une $K$-algèbre de dimension $n$, et considérons l'endomorphisme de $K$-ev $F - \mathrm{Id}$. Soit $r$ la dimension de $N = \ker(F - \mathrm{Id})$ ; on a $r \geqslant 1$ car $K \subset N$. Si $r = 1$ alors $P$ est irréductible, l'algorithme le dit et s'arrête. Si $r \geqslant 2$ on prend un élément de $N \backslash K$, classe d'un polynôme $Q \in K[X]$. On décrit alors tous les $\alpha \in K$ et on calcule le pgcd de $P$ et $Q - \alpha$. Un lemme montre que pour un certain $\alpha$ on trouve un truc non trivial, l'algorithme le renvoie et s'arrête.

## 2 Mise en pratique

Voici une question (hyper classique) posée à l'oral 2005 dans la leçon sur les racines des polynômes : factoriser $P = X^9 + X^6 - X + 1$ sur $K = \mathbb{F}_3$.

La première chose à faire est de voir s'il a une racine dans $K$. C'est vite fait car $K$ n'a que 3 éléments, et on trouve qu'il n'y a pas de racine. Appliquons maintenant l'algorithme :

(1) $P' = -1$.

(2) $P \wedge P' = 1$ car $P' = -1$.

(3) On est donc dans le vif du sujet. L'algèbre qui nous intéresse est

$$A = K[X]/(X^9 + X^6 - X + 1)\,.$$

C'est un $K$-ev de dimension 9 dont une base est $\{1, X, X^2, \ldots, X^8\}$. Pour calculer la matrice de l'endomorphisme $F - \mathrm{Id}$ on aura besoin des puissances de $X^3$ jusqu'à $X^{24}$. Je vais calculer aussi $X^{10}$ et $X^{11}$, vous verrez tout de suite pourquoi en suivant le calcul :

$$\begin{aligned}
X^9 &= -X^6 + X - 1 \\
X^{10} &= -X^7 + X^2 - X \\
X^{11} &= -X^8 + X^3 - X^2 \\
X^{12} &= -(-X^6 + X - 1) + X^4 - X^3 = X^6 + X^4 - X^3 - X + 1 \\
X^{15} &= (-X^6 + X - 1) + X^7 - X^6 - X^4 + X^3 \\
&= X^7 + X^6 - X^4 + X^3 + X - 1 \\
X^{18} &= (-X^7 + X^2 - X) + (-X^6 + X - 1) - X^7 + X^6 + X^4 - X^3 \\
&= X^7 + X^4 - X^3 + X^2 - 1 \\
X^{21} &= (-X^7 + X^2 - X) + X^7 - X^6 + X^5 - X^3 \\
&= -X^6 + X^5 - X^3 + X^2 - X \\
X^{24} &= -(-X^6 + X - 1) + X^8 - X^6 + X^5 - X^4 \\
&= X^8 + X^5 - X^4 - X + 1
\end{aligned}$$

On peut alors écrire la matrice de $F - \mathrm{Id}$ :

$$\begin{pmatrix}
0 & 0 & 0 & -1 & 1 & -1 & -1 & 0 & 1 \\
0 & -1 & 0 & 1 & -1 & 1 & 0 & -1 & -1 \\
0 & 0 & -1 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & -1 & -1 & 1 & -1 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 1 \\
0 & 0 & 1 & -1 & 1 & 1 & -1 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\begin{matrix}
1 \\ X \\ X^2 \\ X^3 \\ X^4 \\ X^5 \\ X^6 \\ X^7 \\ X^8
\end{matrix}$$

Le noyau contient la droite engendrée par la première colonne ; il faut décider s'il est plus gros. Utilisons le pivot de Gauss ; c'est long

mais on trouve le vecteur $(0,1,-1,-1,1,1,-1,0,1)$ (par exemple). Un représentant polynôme est

$$Q(X) = X^8 - X^6 + X^5 + X^4 - X^3 - X^2 + X$$

On calcule ensuite $P \wedge (Q - \alpha)$. Gardons $\alpha$ général au début, on fera $\alpha = 0$, 1 ou 2 en temps utile. On utilise l'algorithme d'Euclide : la première division est

$$\begin{aligned} X^9 + X^6 - X + 1 = \\ (X^8 - X^6 + X^5 + X^4 - X^3 - X^2 + X - \alpha)X \\ +(X^7 - X^5 + X^4 + X^3 - X^2 + (\alpha - 1)X + 1). \end{aligned}$$

La seconde est

$$\begin{aligned} X^8 - X^6 + X^5 + X^4 - X^3 - X^2 + X - \alpha = \\ (X^7 - X^5 + X^4 + X^3 - X^2 + (\alpha - 1)X + 1)X - \alpha(X^2 + 1). \end{aligned}$$

C'est gagné car si $\alpha = 0$, on a un reste nul, donc $X^7 - X^5 + X^4 + X^3 - X^2 - X + 1$ divise $P$ :

$$P(X) = (X^7 - X^5 + X^4 + X^3 - X^2 - X + 1)(X^2 + 1)$$

On pourrait imaginer de continuer l'algorithme avec $\alpha \neq 0$, mais on ne trouverait rien d'autre (je ne sais pas si c'est un fait général de l'algorithme de Berlekamp), car on poursuivrait avec $X^2 + 1$ comme reste, or on sait déjà qu'il divise $P$.

Le polynôme $X^2 + 1$ est irréductible sur $\mathbb{F}_3$. Il faut ensuite recommencer avec le facteur $P_1 = X^7 - X^5 + X^4 + X^3 - X^2 - X + 1$. Avec moins de détail cette fois :

$$\begin{aligned} X^7 &= X^5 - X^4 - X^3 + X^2 + X - 1 \\ X^9 &= -X^6 + X - 1 \\ X^{12} &= X^6 + X^4 - X^3 - X + 1 \\ X^{15} &= X^6 + X^5 + X^4 + X^2 - X + 1 \\ X^{18} &= X^5 + X^3 - X^2 + X + 1 \end{aligned}$$

Matrice de $F - \mathrm{Id}$

$$\begin{pmatrix} 0 & 0 & 0 & -1 & 1 & 1 & 1 \\ 0 & -1 & 0 & 1 & -1 & -1 & 1 \\ 0 & 0 & -1 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 1 & 1 & -1 \end{pmatrix} \begin{matrix} 1 \\ X \\ X^2 \\ X^3 \\ X^4 \\ X^5 \\ X^6 \end{matrix}$$

Ici on voit très rapidement que le système formé par les six colonnes de droite est de rang maximal, donc la dimension du noyau est 1 et $P$ est irréductible. En conclusion la décomposition en facteurs irréductibles de $X^9 + X^6 - X + 1$ est

$$(X^7 - X^5 + X^4 + X^3 - X^2 - X + 1)(X^2 + 1)\ .$$

## 3 Décomposition sur les exemples simples

Dans les cas simples, soit on aura une racine dans le corps de base, soit il y aura suffisamment de facteurs irréductibles de petit degré (i.e. 2 ou 3) qui sont facilement détectables et permettent de conclure. Dans ce cas-là, on "va à la pêche" ce qui nécessite tout de même de la méthode.

Rappelons tout d'abord la remarque utile :

**Lemme :** *Soit $K$ un corps et $P \in K[X]$ de degré $n$. Alors $P$ est réductible ssi il est divisible par un polynôme de degré $\leqslant n/2$.*

D'abord notons qu'il est facile de tester si un polynôme de degré $\leqslant 3$ est irréductible, car cela équivaut à dire qu'il n'a pas de racine. On voit alors, par exhaustion, que sur $\mathbb{F}_3$, les polynômes irréductibles unitaires de degré 2 sont : $X^2 + 1$, $X^2 + X - 1$, $X^2 - X - 1$.

**Remarque :** comme on est en caractéristique différente de 2, à translation près sur la variable, dans tout polynôme unitaire de degré 2 on peut camoufler le terme en $X$ puisque $X^2 + aX = (X + \frac{1}{2}a)^2 + \ldots$ Par ailleurs il est clair que $P$ est irréductible ssi $P(X + a)$ l'est. Donc partant des polynômes irréductibles de degré 2 et sans terme en $X$

(il n'y en a qu'un : $X^2+1$), on les trouve tous en susbtituant $X+a$ ($a \in \mathbb{F}_3$) à $X$. On trouve ainsi $X^2+X-1$ et $X^2-X-1$ sans calcul.

Pour trouver les polynômes irréductibles de degré 3 on peut utiliser le fait que $X^3-X$ induit la fonctions nulle sur $\mathbb{F}_3$. Si $P$ est un polynôme irréductible unitaire de degré 3, alors $P-(X^3-X)$ est un polynôme de degré $\leqslant 2$ dont la fonction associée ne s'annulle pas. Il est donc de la forme $aQ$ avec $a \in \mathbb{F}_3^\times$ et $Q$ unitaire sans racine, i.e. égal à 1 ou irréductible de degré 2. En sens inverse, partant de $Q=1$ ou $Q$ de degré 2 unitaire irréductible, le polynôme $P=X^3-X+aQ$ est irréductible unitaire de degré 3. On trouve ainsi :

$X^3-X+1$ ,
$X^3-X-1$ ,

$$(X^3-X)+(X^2+X-1) = X^3+X^2-1 ,$$
$$(X^3-X)-(X^2+X-1) = X^3-X^2+X+1 ,$$

$(X^3-X)+(X^2+1)$ ,
$(X^3-X)-(X^2+1)$ ,

$$(X^3-X)+(X^2-X-1) = X^3+X^2+X-1 ,$$
$$(X^3-X)-(X^2-X-1) = X^3-X^2+1 .$$

**Remarque :** vérifions que les polynômes trouvés sont en nombre correct à l'aide de la formule

$$\operatorname{card} I(n,q) = \frac{1}{n}\sum_{d|n} \mu(d) q^{n/d} .$$

On trouve $\operatorname{card} I(2,3) = \frac{1\times 3^2+(-1)\times 3}{2} = 3$ et $\operatorname{card} I(3,3) = \frac{1\times 3^3+(-1)\times 3}{3} = 8$.

# Que veut dire être canonique ?

Alors que le mot « canonique » est très présent en mathématiques, on n'en trouve pas de définition précise. Ceci est assez surprenant dans une discipline qui s'enorgueillit avec raison d'avoir pour principe de base de commencer par définir les termes qu'elle emploie. L'absence d'une définition officielle est d'ailleurs l'occasion de discussions dans la communauté au sujet du concept de canonicité, que les uns utilisent abusivement comme fourre-tout, et sur lequel d'autres ont au contraire des idées personnelles très précises.

$\therefore$

Il y a un sens du mot canonique qui ne pose pas de problème. Dans certains énoncés de définitions ou de théorèmes, on définit explicitement un objet qui, souvent, vérifie telle ou telle propriété et on le *nomme* canonique. C'est le cas pour la *surjection canonique* d'un ensemble (resp. d'un groupe, d'un anneau) vers son quotient par une relation d'équivalence (resp. par un sous-groupe distingué, par un idéal), la *base canonique* du $k$-espace vectoriel $k^n$, l'*injection canonique* d'un anneau commutatif $A$ dans l'anneau de polynômes $A[X]$, etc. Dans cette note, nous voulons parler de l'autre sens qu'a le mot « canonique », celui qui exprime le fait qu'un objet est unique, ou naturel ; il s'agit d'une qualité de l'objet, pas de son nom. Plus précisément, nous proposons de discuter et illustrer l'idée suivante : *on peut qualifier un objet de « canonique » si son existence ne dépend pas de divers choix.* Un point intéressant est que l'on peut discuter de la nature et du nombre de ces choix, comme nous allons le voir.

$\therefore$

Un bon exemple introductif est celui de l'isomorphisme de bidualité pour les espaces vectoriels de dimension finie fixée. Rappelons qu'il s'agit de l'isomorphisme $\mathrm{ev} : E \to E^{**}$ qui envoie un vecteur $x$ sur l'application d'évaluation $\mathrm{ev}_x : \varphi \mapsto \varphi(x)$. On voit que la définition

de ev ne dépend pas d'un choix de base pour $E$. De plus $\mathrm{ev} = \mathrm{ev}_E$ ne dépend pas de l'espace vectoriel $E$, au sens où si $u : E \to F$ est un isomorphisme, alors l'isomorphisme

$$F \xrightarrow{u^{-1}} E \xrightarrow{\mathrm{ev}_E} E^{**} \xrightarrow{u^{**}} F^{**}$$

obtenu en transformant $\mathrm{ev}_E$ à l'aide de $u$, est égal à $\mathrm{ev}_F$. Ceci légitime le fait que l'on note ev au lieu de $\mathrm{ev}_E$. Par ailleurs, pour tout scalaire non nul $\lambda$ dans le corps de base, l'application $\lambda\,\mathrm{ev}$ vérifie les mêmes propriétés que ev ; mais le simple fait de nommer $\lambda\,\mathrm{ev}$ nécessite de choisir $\lambda$, alors que l'existence de ev est un fait indépendant de notre intervention, et qui ne dépend donc même pas du corps des scalaires, en un sens évident. On pourrait tout de même avancer l'argument que $-\,\mathrm{ev}$ est très naturel aussi, et il y aurait un contre-argument selon lequel introduire $-1$ relève encore d'un petit choix. Cet exemple montre bien certains choix qu'on peut être amené à faire ou ne pas faire, et les discussions sur leur « naturalité » que cela amène.

$$\therefore$$

Nous allons entrer dans le vif du sujet en parlant des isomorphismes entre un espace vectoriel et son dual. Soient $n \geqslant 1$ un entier et $k$ un corps. Dans la suite, tous les espaces vectoriels sont des $k$-espaces vectoriels de dimension $n$.

**Définitions.** Un isomorphisme $i : E \to E^*$ est *indépendant de la base* si pour tout $u \in \mathrm{GL}(E)$, on a $u^* \circ i \circ u = i$. Une collection d'isomorphismes $i_E : E \to E^*$, pour $E$ variable, est *indépendante de $E$ à isomorphisme près* si pour tout isomorphisme $u : E \to F$, on a $u^* \circ i_F \circ u = i_E$.

Ces conditions s'expriment par la commutativité des diagrammes naturels

$$\begin{array}{ccc} E & \xrightarrow{i_E} & E^* \\ {\scriptstyle u}\downarrow & & \uparrow{\scriptstyle u^*} \\ F & \xrightarrow{i_F} & F^* \end{array}$$

avec $F = E$ dans le cas d'un isomorphisme indépendant de la base. Il est clair que si $\{i_E\}$ est une collection d'isomorphismes indépendante de $E$ à isomorphisme près, alors chaque $i_E$ est indépendant de la base. En guise de note culturelle, signalons que le langage adapté pour formuler la notion de « collection $\{i_E\}$ indépendante de $E$ à isomorphisme près » est celui des *catégories* ; on parle alors d'isomorphismes $i_E$ *fonctoriels* par rapport aux isomorphismes $u : E \to F$. Nous n'aurons pas besoin de ce langage.

On rappelle qu'il y a une correspondance entre isomorphismes $i : E \to E^*$ et formes bilinéaires non dégénérées $\langle -, - \rangle$ sur $E$, donnée par la formule $i(x) = \langle x, - \rangle$. Il sera plus agréable d'utiliser les définitions ci-dessus en termes de formes, comme suit.

**Définitions.** Une forme bilinéaire non dégénérée $\langle -, - \rangle$ sur $E$ est *indépendante de la base* si l'on a $\langle u(x), u(y) \rangle = \langle x, y \rangle$ pour tout $u \in GL(E)$ et tous $x, y \in E$. Une collection de formes bilinéaires non dégénérées $\langle -, - \rangle_E$ sur des $E$ variables est *indépendante de $E$ à isomorphisme près* si l'on a $\langle u(x), u(y) \rangle_F = \langle x, y \rangle_E$ pour tout isomorphisme $u : E \to F$ et tous $x, y \in E$.

**Exemples.** (1) Les deux seuls corps $k$ dont tout élément $\lambda \neq 0$ vérifie $\lambda^2 = 1$ sont $\mathbb{F}_2$ et $\mathbb{F}_3$. Si $E$ est un espace vectoriel de dimension 1 sur l'un de ces corps, on définit une application $E \times E \to k$, $(x, y) \mapsto xy$ comme suit :

$$xy = \begin{cases} 0 & \text{si } x = 0 \text{ ou } y = 0, \\ 1 & \text{si } x, y \neq 0 \text{ et } x = y, \\ -1 & \text{si } x, y \neq 0 \text{ et } x \neq y. \end{cases}$$

On montre facilement que cette application est une forme bilinéaire en observant par exemple que si $e$ est un vecteur de base de $E$ et $x = ae$, $y = be$, on a $xy = ab$. Si $E = k$, cette forme est simplement le produit du corps $k$.

(2) Si $E$ est un $k$-espace vectoriel de dimension 2 sur le corps $k = \mathbb{F}_2$, l'espace des formes bilinéaires alternées est de dimension 1. Il possède donc un unique élément non nul : le déterminant dans une quelconque base, qui dans le contexte présent ne dépend pas de la base. Pour

$x, y \in E$, nous noterons $\det(x, y)$ la valeur correspondante de cette forme bilinéaire alternée.

**Théorème.** *Soit $E$ un espace vectoriel de dimension $n \geqslant 1$ sur un corps $k$. Il existe un isomorphisme $i : E \to E^*$ indépendant de la base si et seulement si $n = 1$ et $k = \mathbb{F}_2$, ou $n = 1$ et $k = \mathbb{F}_3$, ou $n = 2$ et $k = \mathbb{F}_2$. Lorsque c'est le cas, il existe même une collection d'isomorphismes $i_E : E \to E^*$ indépendante de $E$ à isomorphisme près. Les formes bilinéaires correspondantes sont les suivantes :*

(1) *$n = 1$ et $k = \mathbb{F}_2$ : la forme symétrique $\langle x, y\rangle_E = xy$,*

(2) *$n = 1$ et $k = \mathbb{F}_3$ : l'une des formes symétriques $\langle x, y\rangle_E = xy$ ou $\langle x, y\rangle_E = -xy$,*

(3) *$n = 2$ et $k = \mathbb{F}_2$ : la forme alternée $\langle x, y\rangle_E = \det(x, y)$.*

*Dans chacun de ces cas, le morphisme composé $E \xrightarrow{i_E} E^* \xrightarrow{i_{E^*}} E^{**}$ est égal à l'isomorphisme de bidualité* $\mathrm{ev} : E \to E^{**}$.

**Commentaires.** (a) Le cas (1) n'est pas surprenant car on a en fait un résultat beaucoup plus fort : si $E$ et $F$ sont deux espaces vectoriels de dimension 1 sur $\mathbb{F}_2$, il existe un unique isomorphisme $i_{E,F} : E \to F$. Ainsi, non seulement $E$ et son dual $E^*$, mais $E$ et n'importe quel autre espace vectoriel $F$ sont isomorphes canoniquement (ici au sens le plus fort possible).

(b) Dans le cas (2), pour chaque espace vectoriel $E$ on a deux choix naturels opposés de forme bilinéaire. Le théorème affirme que chacune des deux formes s'étend en une collection indépendante de $E$ à isomorphisme près (et donc est vraiment « canonique » au sens le plus fort). Ceci est remarquable car il aurait pu se passer qu'il n'y ait pas de possibilité de choisir entre une forme et son opposée de manière compatible pour tous les espaces vectoriels.

(c) Dans tous les cas exceptionnels, comme il existe des collections indépendantes de $E$, on note $i$ au lieu de $i_E$ et $\langle x, y\rangle$ au lieu de $\langle x, y\rangle_E$. La dernière affirmation du théorème est donc que $\mathrm{ev} = i^2$, c'est-à-dire que les isomorphismes de dualité canoniques exceptionnels fournissent une racine carrée de l'isomorphisme de bidualité (voire deux, si $n = 1$ et $k = \mathbb{F}_3$).

**Preuve.** Cas $n \geqslant 3$. Soient $E$ un espace vectoriel de dimension $n$. Supposons qu'il existe un isomorphisme $i : E \to E^*$ indépendant de la base. Soit $\langle -, -\rangle$ la forme bilinéaire non dégénérée correspondante, invariante sous $\mathrm{GL}(E)$. Choisissons $x \in E$ non nul. Comme le noyau de la forme $\langle x, -\rangle$ est de dimension au moins 2, il existe $y \in E$ tel que $\langle x, y\rangle = 0$ et $\{x, y\}$ est une famille libre. Par ailleurs, comme $x \neq 0$ il existe $z \in E$ tel que $\langle x, z\rangle \neq 0$. Quitte à remplacer $z$ par $z + y$, on peut supposer que $\{x, z\}$ est une famille libre de $E$. Soit $u \in \mathrm{GL}(E)$ un automorphisme qui fixe $x$ et qui envoie $y$ sur $z$. On a alors $0 = \langle x, y\rangle = \langle u(x), u(y)\rangle = \langle x, z\rangle \neq 0$, une contradiction.

Cas $n = 2$. Dans une base fixée de $E$, on note $B$ la matrice de la forme bilinéaire $\langle -, -\rangle$ et $M$ la matrice d'un automorphisme variable $u \in \mathrm{GL}(E)$. Alors la propriété d'invariance $\langle u(x), u(y)\rangle = \langle x, y\rangle$ est équivalente à l'égalité matricielle ${}^tMBM = B$, pour toute matrice inversible $M$ de transposée ${}^tM$. Notons $B = \left(\begin{smallmatrix} r & s \\ t & u \end{smallmatrix}\right)$. En prenant $M = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ puis $M = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, on trouve $r = u = s+t = 0$ donc $B = s\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. En prenant maintenant $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ quelconque, l'égalité ${}^tMBM = B$ est équivalente à l'égalité $ad - bc = 1$. Ceci a lieu pour tout $M$ si et seulement si le seul élément non nul de $k$ est 1, c'est-à-dire $k = \mathbb{F}_2$. On a alors $B = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ dont on vérifie que c'est la matrice de la forme $(x, y) \mapsto \det(x, y)$ introduite dans l'exemple (2).

Cas $n = 1$. Dans ce cas, tout automorphisme $u$ est une homothétie de rapport $\lambda \neq 0$. L'égalité d'invariance s'écrit $\langle x, y\rangle = \langle u(x), u(y)\rangle = \lambda^2\langle x, y\rangle$. Pour que cette égalité soit vérifiée pour tous $\lambda, x, y$, il faut et il suffit que $\lambda^2 = 1$ pour tout scalaire $\lambda$ non nul dans $k$. Ceci signifie que $k^* = \{\pm 1\}$, c'est-à-dire $k = \mathbb{F}_2$ ou $k = \mathbb{F}_3$. Dans ce cas, n'importe quelle forme bilinéaire non dégénérée $\langle -, -\rangle$ est invariante. Dans une base fixée, une telle forme est déterminée par sa matrice composée d'un seul scalaire non nul. Si $k = \mathbb{F}_2$ il y a une seule possibilité, et si $k = \mathbb{F}_3$ il y a deux possibilités. Il est facile de vérifier que les formes en question sont la forme de l'exemple (1) et son opposée.

Le fait que les formes $\pm\langle x, y\rangle$ et $\det(x, y)$ soient définies en des termes qui ne font intervenir que la structure des espaces vectoriels en jeu montrent qu'elles sont invariantes par tout isomorphisme $u : E \to F$. Ceci montre que, dans chacun des cas exceptionnels, on définit ainsi

une collection d'isomorphismes $i_E : E \to E^*$ indépendante de $E$ à isomorphisme près.

Il ne reste qu'à montrer que la composée $i_{E^*} \circ i_E$ est égale à ev. Le plus simple est de le prouver matriciellement. On fixe une base $\mathscr{B}$ de $E$ et on note $\mathscr{B}^*, \mathscr{B}^{**}$ les bases correspondantes de $E^*, E^{**}$. Dans les cas (1) et (2), il existe $\lambda \neq 0$ tel que

$$\mathrm{Mat}_{\mathscr{B},\mathscr{B}^*}(i_E) = \mathrm{Mat}_{\mathscr{B}^*,\mathscr{B}^{**}}(i_{E^*}) = (\lambda) \quad \text{et} \quad \mathrm{Mat}_{\mathscr{B},\mathscr{B}^{**}}(\mathrm{ev}) = (1).$$

Comme $\lambda^2 = 1$, on a bien $i_{E^*} \circ i_E = \mathrm{ev}$. Dans le cas (3), on a

$$\mathrm{Mat}_{\mathscr{B},\mathscr{B}^*}(i_E) = \mathrm{Mat}_{\mathscr{B}^*,\mathscr{B}^{**}}(i_{E^*}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

et

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

Ici encore on trouve que $i_{E^*} \circ i_E = \mathrm{ev}$. □

Pour revenir au thème de cette note, on peut conclure que dans le cas des isomorphismes entre un espace vectoriel de dimension finie $n$ et son dual, la situation intermédiaire où l'on aurait un isomorphisme $i : E \to E^*$ qui est invariant par changement de base, mais pas par isomorphisme $u : E \to F$, ne se produit pas. C'est tout blanc ou tout noir et on conclut :

- si $n = 1$ et $\mathrm{card}(k) \leqslant 3$, ou si $n = 2$ et $k = \mathbb{F}_2$, il existe un isomorphisme canonique entre un $k$-espace vectoriel de dimension $n$ et son dual ;

- dans tous les autres cas, il n'existe pas d'isomorphisme canonique entre un $k$-espace vectoriel de dimension $n$ et son dual ;

Il est intéressant de noter que lorsque $n = 1$ et $k = \mathbb{F}_3$, on doit décider : la forme bilinéaire $\langle x, y \rangle_E = -xy$ est-elle canonique ?

# Le dual en dimension infinie

Soient $k$ un corps et $E$ un $k$-espace vectoriel de dimension *infinie*. Nous voulons démontrer dans cette note le résultat suivant : *l'espace vectoriel dual $E^*$ n'est pas isomorphe à $E$.*

Ce résultat est « bien connu » mais il est difficile d'en trouver une démonstration dans la littérature. On le trouve démontré aux pages 244-248 de Jacobson, *Lectures in Abstract Algebra II*, édité par Springer-Verlag. La preuve de Jacobson, assez longue, est aussi valable pour les espaces vectoriels sur un corps gauche (i.e. non commutatif). La preuve simple donnée ci-dessous est due à Andrea Ferretti.

Rappelons que tout espace vectoriel possède une base $\{e_i\}_{i \in I}$ (c'est un corollaire du lemme de Zorn), ce qui veut dire que $E$ est isomorphe à l'espace vectoriel $k^{(I)}$ des fonctions à support fini de $I$ dans $k$. La *dimension* de $E$ est par définition égale au *cardinal* de $I$, et le résultat énoncé ci-dessus résultera de l'affirmation plus précise que $\dim(E) < \dim(E^*)$.

## 1 Un peu de théorie des ensembles

On dispose de deux bonnes références qui donnent un bref aperçu des bases de la théorie des ensembles : l'annexe B de Lang, *Algèbre* (Dunod) et le livre de Halmos, *Introduction à la théorie des ensembles* (Gabay) pour un exposé un peu plus complet.

**1.1 L'ensemble de tous les ensembles.** Commençons par rappeler le célèbre *paradoxe de Russell* qui dit qu'*il n'existe pas d'ensemble de tous les ensembles.* En effet, supposons qu'il existe un tel ensemble $\mathscr{E}$, et notons $A = \{X \in \mathscr{E},\ X \notin X\}$ l'ensemble des éléments de $\mathscr{E}$ qui n'appartiennent pas à eux-mêmes. Posons-nous la question : est-ce que $A \in A$ ? Si oui, par définition cela veut dire que $A \notin A$ ce qui est contradictoire. Si non, par définition cela veut dire que $A \in A$ ce qui est contradictoire. On obtient donc un objet $A$ qui ne peut exister.

**1.2 Les cardinaux.** Soient $A$ et $B$ deux ensembles. S'il existe une bijection $A \to B$, on dit que $A$ *et* $B$ *ont même cardinal* et on écrit $\mathrm{card}(A) = \mathrm{card}(B)$. S'il existe une injection $A \to B$, ou de manière équivalente (c'est facile) s'il existe une surjection $B \to A$, on écrit $\mathrm{card}(A) \leqslant \mathrm{card}(B)$, ou $\mathrm{card}(B) \geqslant \mathrm{card}(A)$. S'il existe une injection mais pas de bijection entre $A$ et $B$, on écrit $\mathrm{card}(A) < \mathrm{card}(B)$. Il est clair que $\mathrm{card}(A) \leqslant \mathrm{card}(B)$ et $\mathrm{card}(B) \leqslant \mathrm{card}(C)$ implique $\mathrm{card}(A) \leqslant \mathrm{card}(C)$.

**1.3 Quelques « calculs » de cardinaux.**

**Théorème (Cantor-Schröder-Bernstein).** *Si* $\mathrm{card}(A) \leqslant \mathrm{card}(B)$ *et* $\mathrm{card}(B) \leqslant \mathrm{card}(A)$ *alors* $\mathrm{card}(A) = \mathrm{card}(B)$.

*Preuve :* voir Lang, Annexe B, Th. B.3.1.

**Lemme 1.** *Notons* $\mathcal{P}(A)$ *l'ensemble des parties de* $A$. *Alors* $\mathrm{card}(\mathcal{P}(A)) > \mathrm{card}(A)$.

*Preuve :* L'application $A \to \mathcal{P}(A)$ qui envoie $a$ sur $\{a\}$ est injective donc $\mathrm{card}(\mathcal{P}(A)) \geqslant \mathrm{card}(A)$. Supposons qu'il existe une bijection $f : A \to \mathcal{P}(A)$, notons $B = \{a \in A, a \notin f(a)\}$ et soit $a_0 \in A$ dont l'image par la bijection $f$ est $B$. Si $a_0 \in B$, par définition on a $a_0 \notin f(a_0) = B$ ce qui est impossible. Si $a_0 \notin B$, par définition on a $a_0 \in f(a_0) = B$ ce qui est impossible. Donc $f$ n'existe pas.

**Lemme 2.** *Soit* $n \geqslant 1$ *un entier naturel et* $A^n$ *le produit cartésien* $n$*-uple de* $A$. *Si* $A$ *est infini, alors* $\mathrm{card}(A^n) = \mathrm{card}(A)$.

*Preuve :* voir Lang, Annexe B, Cor. B.3.7.

**Lemme 3.** *Notons* $\mathcal{P}_*(A)$ *l'ensemble des parties finies de* $A$. *Si* $A$ *est infini alors* $\mathrm{card}(\mathcal{P}_*(A)) = \mathrm{card}(A)$.

*Preuve :* voir Lang, Annexe B, Cor. B.3.9.

## 2 Dual en dimension infinie

Nous utiliserons les résultats classiques rappelés dans la partie précédente sous la forme suivante.

**Lemme 4.** *Soient* $A, B$ *deux ensembles et* $z \in B$. *Notons* $B^{(A)}$ *l'ensemble des applications* $f : A \to B$ *telles que* $f^{-1}(B\backslash\{z\})$ *est fini. Si* $A$ *est infini et* $\mathrm{card}(A) \geqslant \mathrm{card}(B) \geqslant 2$, *alors* $\mathrm{card}(B^{(A)}) = \mathrm{card}(A)$.

Dans notre application, le point $z$ sera le zéro d'un espace vectoriel et les éléments de $B^{(A)}$ seront donc les fonctions à support fini.

*Preuve :* Comme $\mathrm{card}(B) \geqslant 2$, il existe $b \in B\backslash\{z\}$. L'application $A \to B^{(A)}$ qui envoie $a$ sur l'indicatrice de $\{a\}$, définie par $f(a) = b$ et $f(x) = z$ si $x \neq a$, est injective. D'après le théorème de Cantor-Schröder-Bernstein, il suffit donc de construire une injection $B^{(A)} \hookrightarrow A$. Pour tout $f \in B^{(A)}$, notons $S_f = f^{-1}(B\backslash\{z\})$. Comme chaque $f$ est déterminée par sa restriction à $S_f$, et compte tenu de l'hypothèse $\mathrm{card}(A) \geqslant \mathrm{card}(B)$, on conclut avec la suite d'injections :

$$B^{(A)} \simeq \coprod_{S \in \mathcal{P}_*(A)} B^S \overset{\text{Lm 3}}{\simeq} \coprod_{a \in A} B = A \times B \overset{\text{Hyp.}}{\hookrightarrow} A \times A \overset{\text{Lm 2}}{\simeq} A.$$

Nous arrivons maintenant au résultat qui nous intéresse.

**Théorème.** *Soient* $k$ *un corps,* $E$ *un* $k$*-espace vectoriel de dimension* infinie *et* $E^*$ *l'espace vectoriel dual. Alors* $\dim(E) < \dim(E^*)$. *En particulier,* $E^*$ *n'est pas isomorphe à* $E$.

**Démonstration :** Rappelons que tout espace vectoriel possède une base : c'est un corollaire du lemme de Zorn. La donnée d'une base $\{e_i\}_{i \in I}$ de $E$ équivaut à celle d'un isomorphisme entre $E$ et l'espace vectoriel $k^{(I)}$ des fonctions à support fini de $I$ dans $k$ : l'isomorphisme en question envoie $x = \sum x_i e_i$ sur la fonction à support fini $f$ telle que $f(i) = x_i$. La dimension de $E$ est égale par définition au cardinal de $I$. Enfin, il est facile de voir que le dual $E^*$ s'identifie alors à l'espace vectoriel $k^I$ de *toutes* les fonctions de $I$ dans $k$, car une forme $\varphi : E \to k$ est déterminée par sa valeur sur chaque $e_i$ et réciproquement.

Notons $\ell$ le sous-corps premier de $k$, c'est-à-dire son plus petit sous-corps ; à isomorphisme près, c'est $\mathbb{Q}$ ou un corps fini $\mathbb{F}_p$. Considérons le $\ell$-espace vectoriel $F = \ell^{(I)}$. On note $F^*$ le dual $\ell$-linéaire,

isomorphe à $\ell^I$. On a :

$$\mathrm{card}(F^*) = \mathrm{card}(\ell^I) \geqslant \mathrm{card}(\{0,1\}^I) = \mathrm{card}(\mathcal{P}(I)) > \mathrm{card}(I)$$

d'après la bijection bien connue $\{0,1\}^I \simeq \mathcal{P}(I)$ et le Lemme 1. Par ailleurs, comme $\ell$ est au plus dénombrable (i.e. de cardinal inférieur ou égal à celui de $\mathbb{N}$) et $I$ est infini, on a $\mathrm{card}(I) \geqslant \mathrm{card}(\ell) \geqslant 2$. Utilisant le Lemme 4, on trouve $\mathrm{card}(F) = \mathrm{card}(I) < \mathrm{card}(F^*)$. En particulier $\dim_\ell(F) < \dim_\ell(F^*)$.

Montrons maintenant que $\dim_\ell F^* \leqslant \dim_k E^*$. Il suffit de démontrer que pour toute famille $\ell$-libre d'applications $\varphi_s : I \to \ell$, $s \in S$, la famille $\varphi'_s = \varphi_s : I \to \ell \subset k$ est $k$-libre. Considérons une combinaison linéaire (finie) nulle $\sum_s \alpha_s \varphi'_s = 0$ avec $\alpha_s \in k$. Notons $\{f_j\}_{j\in J}$ une base de $k$ comme $\ell$-espace vectoriel, et écrivons $\alpha_s = \sum_j \alpha_{sj} f_j$ sur cette base, avec $\alpha_{sj} \in \ell$. Pour tout $i$, on a :

$$0 = \sum_{s,j} \alpha_{sj} f_j \varphi_s(i) = \sum_j \left( \sum_s \alpha_{sj} \varphi_s(i) \right) f_j$$

donc $\sum_s \alpha_{sj} \varphi_s(i) = 0$ pour tout $j$, puisque $\{f_j\}$ est une base. Ceci montre que $\sum_s \alpha_{sj} \varphi_s = 0$ comme applications de $I$ dans $\ell$, et comme les $\varphi_s$ sont $\ell$-libres, on trouve $\alpha_{sj} = 0$ pour tous $s, j$. Finalement $\alpha_s = 0$ donc la famille des $\varphi'_s$ est libre.

On conclut en alignant ces inégalités : $\dim_k(E) = \dim_\ell(F) < \dim_\ell(F^*) \leqslant \dim_k(E^*)$. □

**Références**

Lang, *Algèbre*, Annexe B, Dunod.
Halmos, *Introduction à la théorie des ensembles*, Gabay.

# Produit semi-direct

## 1 Définitions générales

Soit $G$ un groupe que l'on souhaite étudier. S'il n'est pas simple, il a un sous-groupe distingué $1 \subsetneq N \subsetneq G$ et on peut commencer par étudier $N$ et le groupe quotient $Q := G/N$. Ceci conduit aux définitions abstraites suivantes :

**Définition :** Une *suite exacte* de groupes est une suite de groupes

$$1 \longrightarrow N \xrightarrow{i} G \xrightarrow{\pi} Q \longrightarrow 1$$

(on écrit 1 au lieu de $\{1\}$) dont les flèches sont des morphismes de groupes tels que l'image de chacun est égale au noyau du suivant. Plus précisément cela signifie que :

(1) $i$ est injectif,

(2) $\pi$ est surjectif,

(3) l'image de $i$ est égale au noyau de $\pi$.

Lorsqu'on a une telle suite exacte on dit que $G$ est *extension de $N$ par $Q$* (ou de $Q$ par $N$ chez certains auteurs ; peu importe). Dans ce cas $N$ est distingué dans $G$.

**Exemples :** Les groupes $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$ et $\mathbb{Z}/8\mathbb{Z}$ ou encore le groupe diédral $\mathbb{D}_4$ sont tous les trois extensions

$$1 \longrightarrow \mathbb{Z}/4\mathbb{Z} \longrightarrow G \longrightarrow \mathbb{Z}/2\mathbb{Z} \longrightarrow 1$$

Les groupes $(\mathbb{Z}/2\mathbb{Z})^3$ et le groupe des quaternions $\mathbb{H}_8$ sont tous les deux extensions

$$1 \longrightarrow \mathbb{Z}/2\mathbb{Z} \longrightarrow G \longrightarrow \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \longrightarrow 1$$

Rappelons que le groupe des quaternions est le groupe $\mathbb{H}_8 = \{\pm 1, \pm i, \pm j, \pm k\}$. C'est un sous-groupe de l'algèbre des quaternions, ce qui explique son nom et sa notation. ▫

Étant donnée une extension $1 \to N \to G \xrightarrow{\pi} Q \to 1$, il existe une situation particulièrement simple :

**Proposition :** Les conditions suivantes sont équivalentes :

(1) Il existe un sous-groupe $H \subset G$ tel que $H \cap N = 1$ et $HN = G$ ([2]).

(2) $\pi\colon G \to Q$ a une section, c'est-à-dire un morphisme $s\colon Q \to G$ tel que $\pi \circ s = \mathrm{Id}_Q$.

Lorsqu'elles sont vérifiées on dit que $G$ est *produit semi-direct* de $N$ par $H$ (en abrégé PSD) et on note $G = N \rtimes H$. On dit aussi que la suite exacte a une section, ou qu'elle est scindée. Enfin, on dit que $H$ est un *complément* pour $N$.

**Preuve :** Sous les hypothèses de (1) on montre facilement que $\pi_{|H}$ est un isomorphisme, donc son inverse produit une section $s\colon Q \to G$. Ceci prouve que (1) $\Rightarrow$ (2). Réciproquement si on a une section il est facile de vérifier que le sous-groupe $H = s(Q)$ vérifie les conditions de (1). Donc (2) $\Rightarrow$ (1). ▫

Si $N$ et $Q$ sont fixés on appelle parfois *problème de l'extension* la question de retrouver tous les groupes $G$ qui sont extensions de $N$ par $Q$. C'est un problème difficile, et les seules extensions que l'on sait décrire de manière générale sont les PSD.

Soit $G = N \rtimes H$ un PSD, on observe que $H$ agit sur $N$ par conjugaison d'où un morphisme de groupes $H \to \mathrm{Aut}(N)$. Ceci mène à une autre manière de décrire les PSD.

En effet supposons avoir des groupes abstraits $N$ et $H$ ainsi qu'un morphisme $\theta\colon H \to \mathrm{Aut}(N)$ qu'on note $\theta(h)(n) = h \cdot n$ comme une action. On peut définir un groupe noté $N \rtimes_\theta H$ de la façon suivante : comme ensemble il s'agit simplement du produit $N \times H$, et la loi de multiplication est

$$(n, h).(n', h') = (n(h \cdot n'), hh')$$

**Exercice :** vérifiez que ceci définit bien une loi de groupe. Soit $N^*$ (resp. $H^*$) l'ensemble des éléments de la forme $(n, 1)$ (resp. $(1, h)$). Vérifiez que $N^* \simeq N$, $H^* \simeq H$ et $N \rtimes_\theta H$ est PSD de $N^*$ par $H^*$. Il est intéressant aussi de noter que l'action initiale de $H$ sur $N$ n'est autre que la conjugaison dans le groupe $N \rtimes_\theta H$. Enfin, montrez que les quatre conditions suivantes sont équivalentes : (i) $H^*$ est distingué dans $N \rtimes_\theta H$, (ii) $\theta$ est trivial, (iii) $N^*$ et $H^*$ commutent, (iv) le produit est un produit direct. On retiendra :

**Proposition :** Soient $N$ et $H$ des groupes. Alors c'est équivalent de se donner :

(1) Un groupe $G$ avec des injections $N \hookrightarrow G$, $H \hookrightarrow G$ qui font de $G$ le PSD de $N$ et $H$.

(2) Un morphisme de groupes $\theta\colon H \to \mathrm{Aut}(N)$.

Dans le cas (1) on parle plutôt de *PSD interne* car on prend le point de vue du groupe $G$, engendré par deux sous-groupes. Dans le cas (2) on parle de *PSD externe* car on part de $H$ et $N$ et on construit un groupe qui les contient.

## 2 Exemples

**Extensions d'ordre** 8. À isomorphisme près il y a 5 groupes finis d'ordre 8 : c'est un exercice aussi classique qu'instructif. Ce sont $(\mathbb{Z}/2\mathbb{Z})^3$, $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$, $\mathbb{Z}/8\mathbb{Z}$, le groupe diédral $\mathbb{D}_4$ et le groupe des quaternions $\mathbb{H}_8$. Voici toutes les extensions entre groupes d'ordres 2 et 4 :

[2] Si $N \lhd G$, c'est un fait général que pour tout sous-groupe $H \subset G$ l'ensemble des produits $hn$ (resp. des produits $nh$) avec $h \in H$ et $n \in N$ est un sous-groupe noté $HN$ (resp. $NH$). En fait $HN = NH =$ le sous-groupe engendré par $H$ et $N$.

| | $\mathbb{Z}/2\mathbb{Z} \to ? \to \mathbb{Z}/4\mathbb{Z}$ | $\mathbb{Z}/4\mathbb{Z} \to ? \to \mathbb{Z}/2\mathbb{Z}$ | $\mathbb{Z}/2\mathbb{Z} \to ? \to (\mathbb{Z}/2\mathbb{Z})^2$ | $(\mathbb{Z}/2\mathbb{Z})^2 \to ? \to \mathbb{Z}/2\mathbb{Z}$ |
|---|---|---|---|---|
| $(\mathbb{Z}/2\mathbb{Z})^3$ | | | × | × |
| $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$ | × | × | | |
| $\mathbb{Z}/8\mathbb{Z}$ | × | × | | |
| $\mathbb{D}_4$ | | × | | |
| $\mathbb{H}_8$ | | × | × | |

La colonne des extensions de $\mathbb{Z}/4\mathbb{Z}$ par $\mathbb{Z}/2\mathbb{Z}$ montre qu'un groupe extension de deux groupes assez gentils peut être assez compliqué. On y trouve une extension qui est produit direct : $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$, une extension qui est produit semi-direct non trivial : $\mathbb{D}_4$, et deux extensions qui ne sont même pas produits semi-directs : $\mathbb{Z}/8\mathbb{Z}$ et $\mathbb{H}_8$.

On voit aussi que les groupes non isomorphes $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$ et $\mathbb{Z}/8\mathbb{Z}$ sont extensions tous deux de de $\mathbb{Z}/4\mathbb{Z}$ par $\mathbb{Z}/2\mathbb{Z}$, et de $\mathbb{Z}/2\mathbb{Z}$ par $\mathbb{Z}/4\mathbb{Z}$. Donc ils sont « indistinguables » du point de vue des extensions.

**Contre-exemples**. Voici des exemples d'extensions non scindées. On a vu

$$1 \longrightarrow \mathbb{Z}/2\mathbb{Z} \longrightarrow \mathbb{H}_8 \longrightarrow \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \longrightarrow 1$$

Plus généralement on obtient une extension non scindée si on a un groupe $G$ de centre $Z$ tel que $Q = G/Z$ est un groupe non trivial de centre non trivial :

$$1 \longrightarrow Z \longrightarrow G \longrightarrow Q = G/Z \longrightarrow 1$$

En effet si l'extension était scindée, comme $Z$ est central le PSD devrait être un produit direct. Donc $G \simeq Z \times Q$ mais comme $Z \times Z(Q)$ est inclus dans le centre de $Z \times Q$, ceci contredit le fait que $Z$ est le centre de $G$. Voici une autre extension non scindée :

$$1 \longrightarrow \mathbb{C}^\times \longrightarrow \mathrm{GL}_n(\mathbb{C}) \longrightarrow \mathrm{PGL}_n(\mathbb{C}) \longrightarrow 1$$

En effet ici encore on quotiente par le centre, donc si l'extension était scindée ce serait un produit direct $G = \mathbb{C}^\times \times H$ avec $H \simeq \mathrm{PGL}_n(\mathbb{C})$. En particulier $H \lhd G$ donc $H \cap \mathrm{SL}_n(\mathbb{C}) \lhd \mathrm{SL}_n(\mathbb{C})$. Comme $\mathrm{PSL}_n(\mathbb{C})$ est simple on en déduit facilement que $H = \mathrm{SL}_n(\mathbb{C})$. Ceci est impossible car $\mathrm{SL}_n(\mathbb{C})$ a un centre alors que $\mathrm{PGL}_n(\mathbb{C})$ n'en a pas. (Voir les trois premiers exercices dans les Exercices du chapitre IV de [Per], p. 108). Pour finir mentionnons l'extension non scindée

$$1 \longrightarrow \{\pm 1\} \longrightarrow G \longrightarrow \mathrm{SO}_3(\mathbb{R}) \longrightarrow 1$$

où $G$ est le groupe des quaternions de norme 1. En effet supposons qu'il existe un sous-groupe $H \subset G$ tel que $G = \{\pm 1\} \rtimes H$ (un complément). Comme $i \in H$ ou $i \in -H$, on déduit dans tous les cas que $-1 = i^2 \in H$. Ceci est impossible. (Lire [Per] chap. VII remarque 2.2.)

**Différents PSD**. Un groupe qui est PSD de deux sous-groupes peut l'être de différentes façons. Considérons par exemple le groupe $G = \mathfrak{S}_3 \times \mathbb{Z}/2\mathbb{Z}$ qui est produit de ses deux sous-groupes $N = \mathfrak{S}_3$ et $H = \mathbb{Z}/2\mathbb{Z}$. Soit $\tau = (12)$ la transposition dans $N$ et $x$ l'élément non trivial de $H$, alors il est facile de voir que si on choisit pour $H'$ le sous-groupe d'ordre 2 engendré par $(\tau, x)$ on obtient une expression comme PSD non trivial $G = N \rtimes H'$.

**Déterminant**. Le morphisme qui à $\lambda \in k^\times$ associe la matrice diagonale $(\lambda, 1, \ldots, 1)$ donne une section du déterminant de sorte que

$$\mathrm{GL}_n(k) = \mathrm{SL}_n(k) \rtimes k^\times$$

**Symétries des polygones réguliers**. On a

$$\mathbb{D}_n = \mathbb{Z}/n\mathbb{Z} \rtimes \mathbb{Z}/2\mathbb{Z}$$

**Signature**. N'importe quelle transvection de $\mathfrak{S}_n$ donne un section de la signature d'où

$$\mathfrak{S}_n = \mathfrak{A}_n \rtimes \mathbb{Z}/2\mathbb{Z}$$

**Groupe affine**. Soit $\mathcal{E}$ un espace affine de direction $E$, alors

$$\mathrm{GA}(\mathcal{E}) = \mathrm{T}(\mathcal{E}) \rtimes \mathrm{GL}(E)$$

**Groupe de Galois de $X^n - a$.** Quand ce polynôme est-il irréductible ? Si $a$ est une puissance $d$-ème pour un $d|n$ alors il est réductible. Lorsque $n = 4k$ il peut aussi se produire un gag si $a = -4b^4$ :

$$X^{4k} + 4b^4 = (X^{2k} + 2bX^k + 2b^2)(X^{2k} - 2bX^k + 2b^2)$$

Le théorème 6.9.1 de [Lan] affirme que dans tous les autres cas le polynôme est irréductible : *soit $k$ un corps et $n \geqslant 2$ un entier, $a$ un élément non nul de $k$, alors si $a$ n'est une puissance $d^{\text{ème}}$ pour aucun $d|n$ et si $a$ n'est pas de la forme $-4b^4$ lorsque $4|n$, alors $X^n - a$ est irréductible.* C'est donc une CNS.

Soit $a \in \mathbb{Q}$ choisi comme dans ce théorème, de sorte que le polynôme $P(X) = X^n - a$ est irréductible. Soit $K \subset \mathbb{C}$ le corps de décomposition de $P$ et $G = \mathrm{Gal}(K/\mathbb{Q})$. Soit $\alpha$ une racine de $P$ et $\zeta$ une racine primitive $n$-ème de l'unité, alors

$$P(X) = (X - \alpha)(X - \zeta\alpha)\ldots(X - \zeta^{n-1}\alpha)$$

On voit ainsi que $K = \mathbb{Q}(\alpha, \zeta)$. Un automorphisme $\sigma \in G$ doit envoyer $\alpha$ sur une autre racine $\zeta^i\alpha$ pour un $i \in \mathbb{Z}/n\mathbb{Z}$ et $\zeta$ sur une autre racine primitive $n$-ème de l'unité $\zeta^j$ pour un $j \in (\mathbb{Z}/n\mathbb{Z})^\times$. Soit $\sigma_{i,j}$ le $\mathbb{Q}$-automorphisme de $K$ ainsi défini :

$$\sigma_{i,j}(\alpha) = \zeta^i\alpha \quad \text{et} \quad \sigma_{i,j}(\zeta) = \zeta^j$$

Il est immédiat de vérifier que $\sigma_{i,j} \circ \sigma_{k,l} = \sigma_{i+kj,jl}$ On a donc un isomorphisme $\sigma_{i,j} \mapsto (i, j)$ :

$$G \simeq \mathbb{Z}/n\mathbb{Z} \rtimes (\mathbb{Z}/n\mathbb{Z})^\times$$

**Théorème de Schur-Zassenhaus**. Terminons par un très beau résultat de théorie des groupes, dont la preuve est difficile. Ce théorème dit que toute extension de groupes finis

$$1 \longrightarrow A \longrightarrow G \xrightarrow{\pi} Q \longrightarrow 1$$

avec $|A|$ et $|Q|$ premiers entre eux est scindée. Notons que pour démontrer ce théorème il suffit de trouver un sous-groupe $H \subset G$ d'ordre égal à $[G : A] = |Q|$. En effet si $H$ est un tel sous-groupe alors $\ker(\pi_{|H}) = A \cap H$ est trivial car $|A|$ et $|H|$ sont premiers entre eux, donc $\pi_{|H}$ est un isomorphisme, d'où le résultat.

En particulier la démonstration est simple si $Q$ est un $p$-groupe, car alors il suffit de choisir pour $H$ un $p$-Sylow de $G$. Ainsi dans $\mathfrak{A}_4$ le sous-groupe $V = \{1, (12)(34), (13)(24), (14)(23)\}$ est un sous-groupe distingué isomorphe à $(\mathbb{Z}/2\mathbb{Z})^2$ et le quotient est $\mathbb{Z}/3\mathbb{Z}$. Il en découle que $\mathfrak{A}_4 \simeq V \rtimes \mathbb{Z}/3\mathbb{Z}$.

Pour démontrer le théorème de Schur-Zassenhaus on observe d'abord que de manière générale le nombre de conjugués $g^{-1}Hg$ d'un sous-groupe $H \subset G$ est égal à l'indice $[G : N]$ du normalisateur $N = N_G(H)$. En effet le groupe $G$ agit transitivement sur l'ensemble des conjugués de $H$, et le stabilisateur de $H$ est $N$.

**Preuve (esquisse) :** On fait une récurrence sur $|G|$ et $|A|$. Dans le cas $|G| = |A| = 1$ il n'y a rien à dire. Rappelons qu'il suffit de trouver un sous-groupe $H \subset G$ d'ordre égal à $[G : A]$.

Soit $p$ un facteur premier de $A$, $S$ un $p$-Sylow de $A$, $N = N_G(S)$. Alors $S$ est aussi un $p$-Sylow de $G$ et $N_A(S) = N \cap A$. Les $p$-Sylow de $G$ (qui sont conjugués à $S$) sont tous dans $A$ donc d'après l'observation précédente :

$$[G : N] = [A : N \cap A]$$

On en déduit $[N : N \cap A] = [G : A]$.

Si $N \cap A \subsetneq A$, par l'hypothèse de récurrence appliquée à $(N, N \cap A)$ le groupe $N$ contient un sous-groupe d'ordre $[N : N \cap A] = [G : A]$. Cela fournit un sous-groupe de $G$ qui répond à la question.

Si $N \cap A = A$ alors $[G : N] = 1$ donc $S$ est distingué dans $G$. Par l'hypothèse de récurrence appliquée à $(G/S, A/S)$ il existe un sous-groupe $H \subset G$ contenant $S$ avec $[H : S] = [G : A]$. Soit $Z \neq 1$ le centre de $S$, un petit calcul montre que $Z \lhd H$ : si $z \in Z$ alors $h^{-1}zh \in S$ et

$$h^{-1}zhs = h^{-1}z(hsh^{-1})h = h^{-1}(hsh^{-1})zh = sh^{-1}zh$$

donc $h^{-1}zh$ commute avec $S$ donc est dans $Z$. Par l'hypothèse de récurrence appliquée à $(H/Z, S/Z)$ il existe un sous-groupe $K \subset H$ contenant $Z$ avec $[K : Z] = [G : A]$.

On s'est ramené à résoudre le problème pour $(K, Z)$ i.e. au cas où $A$ est un $p$-groupe central abélien. La conclusion de la preuve dans ce cas particulier, fait appel à la *cohomologie des groupes*, théorie qui va au-delà du programme de l'Agrégation.

### Bibliographie


[Lan] Lang, Algèbre, *Dunod*.
[Per] Perrin, Cours d'Algèbre, *Ellipses* ou *ENSJF*.


# L'algèbre des quaternions

## 1 Définition

Pour définir l'algèbre des quaternions $\mathbb{H}$, on fixe un $\mathbb{R}$-espace vectoriel de dimension 4 et une base que l'on note $\{1, i, j, k\}$. On définit une application $\mathbb{R}$-bilinéaire $\mathbb{H} \times \mathbb{H} \to \mathbb{H}$ pour laquelle 1 est élément neutre, en posant :

$$i^2 = j^2 = k^2 = -1 \quad ;$$

$$ij = -ji = k \quad ; \quad ik = -ki = -j \quad ; \quad jk = -kj = i \ .$$

Il est facile de vérifier que cette multiplication munit $\mathbb{H}$ d'une structure de $\mathbb{R}$-algèbre associative.

Le sous-espace vectoriel de $\mathbb{H}$ engendré par 1 est noté simplement $\mathbb{R}$ ; c'est une sous-algèbre. Le sous-espace vectoriel de $\mathbb{H}$ engendré par $i, j, k$ est noté $P$ ; il n'est pas stable par multiplication. Un quaternion de $\mathbb{R}$ est dit *réel* et un quaternion de $P$ est dit *imaginaire pur*. On a évidemment $\mathbb{H} = \mathbb{R} \oplus P$ donc on peut parler de la *partie réelle* et de la *partie imaginaire* d'un quaternion.

L'algèbre $\mathbb{H}$ n'est pas commutative ; son centre noté $Z(\mathbb{H})$ est la sous-algèbre engendrée par 1, c'est-à-dire $Z(\mathbb{H}) = \mathbb{R}$.

Il y a sur $\mathbb{H}$ une *conjugaison* qui est définie ainsi : si $q = a + bi + cj + dk$ alors son conjugué est $\overline{q} = a - bi - cj - dk$. On vérifie aisément que c'est un anti-automorphisme, c'est-à-dire que c'est un automorphisme d'espace vectoriel et que $\overline{q_1 q_2} = \overline{q_2 q_1}$. Dit autrement, la conjugaison est un isomorphisme entre $\mathbb{H}$ et l'algèbre *opposée* à $\mathbb{H}$, c'est-à-dire l'algèbre $\mathbb{H}^\circ$ dans laquelle le produit de deux éléments $q_1$ et $q_2$ est par définition $q_2 q_1$ (produit dans $\mathbb{H}$).

Soit un quaternion $q \in \mathbb{H}$. Il est facile de voir que $q \in \mathbb{R}$ ssi $\overline{q} = q$, et $q \in P$ ssi $\overline{q} = -q$. Mais il existe une autre caractérisation des quaternions réels et imaginaires purs, un peu plus surprenante et purement algébrique. Précisément, un simple calcul montre que $q \in \mathbb{R}$ ssi $q^2 \in \mathbb{R}_{\geqslant 0}$, et $q \in P$ ssi $q^2 \in \mathbb{R}_{\leqslant 0}$. Notons à ce propos que la

relation d'ordre total de $\mathbb{R}$ ne s'étend pas à $\mathbb{H}$. De ce fait, la notation « $q \geqslant 0$ » est ambigüe et c'est pourquoi nous préférons écrire $q \in \mathbb{R}_{\geqslant 0}$.

On définit ensuite la norme d'un quaternion $q \in \mathbb{H}$ par $N(q) = q\overline{q}$. On vérifie que si $q = a + bi + cj + dk$ alors $N(q) = a^2 + b^2 + c^2 + d^2$. En particulier $N(q) \in \mathbb{R}$, $N(\overline{q}) = N(q)$, et $N(q) = 0$ si et seulement si $q = 0$. Il s'ensuit que si $q \neq 0$, alors $N(q)^{-1}\overline{q}$ est inverse à gauche et à droite pour $q$. Donc $\mathbb{H}$ est ce que l'on appelle une *algèbre à division* ou encore parfois un *corps gauche*. On a donc une application multiplicative $N : \mathbb{H} \to \mathbb{R}$ et un morphisme de groupes induit $\mathbb{H}^* \to \mathbb{R}^*$. Le fait que le corps $\mathbb{H}$ soit non commutatif introduit parfois certaines subtilités par rapport à la théorie des corps commutatifs. Nous essaierons de les mettre en lumière au fur et à mesure.

# 2 Sous-corps de $\mathbb{H}$

## 2.1 L'équation $X^2 + 1 = 0$

Dans ce qui suit, on notera $G$ le sous-groupe de $\mathbb{H}^*$ formé des quaternions de norme 1. On note que $q^2 = -1$ ssi on a simultanément $q^2 \in \mathbb{R}_{\leqslant 0}$ et $q \in G$. D'après ce que l'on a dit plus haut, cela veut donc dire que $q \in P \cap G$. L'espace vectoriel euclidien $P$ est de dimension 3 et $P \cap G$ est sa sphère unité. On obtient donc que l'équation $X^2 + 1 = 0$ possède dans $\mathbb{H}$ un ensemble de solutions isomorphe à la sphère $S^2$. En particulier, on constate que la finitude du nombre de racines d'un polynôme, phénomène classique de la théorie des corps commutatifs, est prise en défaut ici.

## 2.2 Formes polaires

Tout quaternion non nul peut donc s'écrire $q = tp$ où $t = N(q)$ et $p \in G$. Cette écriture sera appelée la *forme polaire* de $q$, en analogie avec le cas complexe.

Pour les quaternions non réels, une autre écriture nous sera aussi utile : partant de la décomposition en partie réelle et partie imaginaire $q = r + q'$, on peut considérer la forme polaire de $q'$ et obtenir

$$q = r + tp \quad \text{où} \quad r \in \mathbb{R},\ p \in P \cap G,\ t = N(q - r)\ .$$

## 2.3 Commutants

Étant donné $q \in \mathbb{H}$, on définit le commutant de $q$, noté $Z(q)$, comme étant l'ensemble des quaternions $x \in \mathbb{H}$ tels que $qx = xq$. C'est un sous-corps de $\mathbb{H}$ contenant $\mathbb{R}$, que nous allons calculer. Notons $K$ le sous-corps de $\mathbb{H}$ engendré par $q$, c'est le corps des polynômes en $q$, qui est monogène donc évidemment commutatif. On a les formules de produit des dimensions d'espaces vectoriels $[\mathbb{H} : \mathbb{R}] = [\mathbb{H} : K]\,[K : \mathbb{R}]$ et $[Z(q) : \mathbb{R}] = [Z(q) : K]\,[K : \mathbb{R}]$. Il s'ensuit que $[K : \mathbb{R}]$ vaut 1 ou 2, puis que $[Z(q) : \mathbb{R}]$ vaut aussi 1, 2 ou 4.

Si $q \in \mathbb{R}$, on a $K = \mathbb{R}$ et $Z(q) = \mathbb{H}$. Sinon, $q$ n'est pas central, donc $Z(q) \neq \mathbb{H}$. On a $[Z(q) : \mathbb{R}] = [K : \mathbb{R}] = 2$. Dans ce cas $Z(q)$ est un corps commutatif isomorphe à $\mathbb{C}$. Si on écrit $q = r + tp$ avec $r$ la partie réelle de $q$ et $t = N(q - r)$, on a $p^2 = -1$ donc on définit un isomorphisme de $\mathbb{R}$-algèbres explicite $f : \mathbb{C} \to Z(q)$ en posant $f(i) = p$.

Notons qu'en particulier, la $\mathbb{R}$-algèbre $\mathbb{H}$ elle-même n'est pas monogène. Ceci est à comparer au fait que toute $\mathbb{R}$-algèbre de dimension finie qui est un corps commutatif est monogène, d'après le théorème de l'élément primitif.

## 2.4 Structures de $\mathbb{C}$-algèbre sur $\mathbb{H}$

Une structure de $\mathbb{C}$-algèbre sur $\mathbb{H}$ est donnée par un morphisme $f : \mathbb{C} \to \mathbb{H}$. Un tel morphisme est déterminé par l'image de $i$ par $f$, qui doit être un élément de carré $-1$. On voit donc que les structures de $\mathbb{C}$-algèbre sur $\mathbb{H}$ sont en bijection avec $P \cap G$.

## Une remarque sur la transposée d'une matrice

Soit $k$ un corps, $n \geqslant 1$ un entier, $E$ un $k$-espace vectoriel de dimension $n$. C'est un fait classique que toute matrice $M \in \mathrm{M}_n(k)$ est semblable à sa transposée. Dans cette note, nous tentons d'expliquer ce fait et nous le démontrons.

Il n'est pas évident de comprendre ce que cet énoncé matriciel signifie pour les endomorphismes. En effet, l'énoncé analogue mène à considérer un endomorphisme $f : E \to E$ et son dual $f^* : E^* \to E^*$. Comme $f$ et $f^*$ n'agissent pas sur les mêmes espaces, on ne peut les conjuguer l'un en l'autre... Sauf à utiliser, pour conjuguer, un isomorphisme $g : E \xrightarrow{\sim} E^*$ de sorte qu'on ait $f^* = gfg^{-1}$. On a alors le diagramme commutatif :

$$\begin{array}{ccc} E & \xrightarrow{f} & E \\ {\scriptstyle g}\downarrow & & \downarrow{\scriptstyle g} \\ E^* & \xrightarrow{f^*} & E^* \end{array}$$

Dans ce cas, l'isomorphisme $g$ ne représente pas un changement de base dans $E$. Plutôt, il faut penser à $g$ comme à une forme bilinéaire non dégénérée sur $E$, en se souvenant que formes bilinéaires $b : E \times E \to k$ et morphismes $g : E \to E^*$ se correspondent via $g(x) = b(x, \cdot)$ et $b(x, y) = (g(x))(y)$. En termes de la forme bilinéaire non dégénérée $b$ associée à $g$, la relation $f^* = gfg^{-1}$ fournit :

$$\text{pour tous } x, y \text{ dans } E,\ b(f(x), y) = b(x, f(y)) \ .$$

Dit autrement, $f$ est autoadjoint pour $b$. Il y a même un petit bonus, car il se trouve que non seulement toute matrice $M$ est semblable à sa transposée, mais en plus, la matrice de passage peut être choisie symétrique. Nous pouvons finalement exprimer ceci sous la forme d'un bel énoncé et le démontrer :

**Proposition** *Soit $f$ un endomorphisme de $E$. Alors, il existe une forme bilinéaire symétrique non dégénérée sur $E$ telle que $f$ est autoadjoint pour cette forme.*

**Démonstration :** bien que l'on ait fait des efforts pour obtenir un énoncé sur les endomorphismes, la démonstration est matricielle. Vérifions qu'on a bien traduit notre problème initial. Dire que $f$ est autoadjoint pour une forme bilinéaire symétrique non dégénérée $\varphi$ signifie, si l'on fixe une base de $E$ et que l'on désigne par $M$ la matrice de $f$ et par $P$ la matrice de $\varphi$, qu'on a

$$^t(MX)PY = {}^tXPMY$$

pour tous vecteurs colonnes $X, Y$. Ceci signifie que $PM = {}^tMP$, donc $M$ et sa transposée sont semblables et conjuguées par une matrice symétrique. Il s'agit donc de trouver une matrice inversible et symétrique $P$ satisfaisant cette égalité. Commençons par le cas où $M$ est une matrice compagnon :

$$M = \begin{pmatrix} 0 & \dots & 0 & b_1 \\ 1 & \ddots & \vdots & b_2 \\ & \ddots & 0 & \vdots \\ 0 & & 1 & b_n \end{pmatrix}$$

c'est-à-dire $m_{i,j} = \delta_{i\geqslant 2}\delta_{j\leqslant n-1}\delta_{i,j+1} + \delta_{j,n}b_i$. L'égalité des coefficients $(PM)_{i,j} = ({}^tMP)_{i,j}$ donne

$$\delta_{j\leqslant n-1}p_{i,j+1} + \delta_{j,n}\sum_{k=1}^{n} p_{i,k}b_k = \delta_{i\leqslant n-1}p_{i+1,j} + \delta_{i,n}\sum_{k=1}^{n} p_{k,j}b_k \ .$$

On en déduit que $p_{i,j+1} = p_{i+1,j}$ à chaque fois que ceci a un sens, c'est-à-dire qu'il existe des scalaires $q_2, \dots, q_{2n}$ tels que $p_{i,j} = q_{i+j}$ pour tous $i, j$, et que $q_{n+i+1} = \sum_{k=1}^{n} q_{k+i}b_k$ pour $1 \leqslant i \leqslant n-1$. On voit donc que la donnée de $q_2, \dots, q_{n+1}$ quelconques détermine $P$, et que toutes les matrices $P$ ainsi obtenues sont symétriques. Enfin il est clair que l'une de ces matrices est inversible, car si l'on prend $q_2 = \cdots = q_n = 0$ et $q_{n+1} = 1$, on trouve une matrice de déterminant $\pm 1$.

Dans le cas général, $M = A^{-1}NA$ où $N$ est diagonale par blocs avec pour blocs des matrices compagnon $N_i$ (forme normale de Frobenius). Pour chaque $i$ choisissons $Q_i$ inversible symétrique telle

$Q_iN_i = {}^tN_iQ_i$. Soit $Q$ la matrice diagonale par blocs de blocs $Q_i$, on vérifie que $P = {}^tAQA$ est inversible symétrique et que $PM = {}^tMP$. □

N.B. On peut s'en sortir aussi avec la réduction de Jordan, mais il faut alors passer par une clôture algébrique de $k$ et ensuite faire quelques contorsions pour justifier que le résultat vaut tout de même sur $k$. C'est donc un exemple où la réduction de Frobenius est plus efficace, car, contrairement à la réduction de Jordan, elle ne nécessite aucune hypothèse sur le corps de base.

**Remarque** Si $k = \mathbb{R}$, il n'existe pas en général de produit scalaire euclidien pour lequel $f$ est autoadjoint (et idem si $k = \mathbb{C}$, pas de produit scalaire hermitien). Dit autrement, la forme bilinéaire symétrique non dégénérée fournie par le résultat ci-dessus ne peut pas toujours être choisie positive. Il suffit de prendre pour contre-exemple un endomorphisme non diagonalisable, puisque tout endomorphisme autoadjoint pour un produit scalaire (ou produit hermitien) est diagonalisable. Par exemple soient $E = \mathbb{R}^2$ et l'endomorphisme (semis-simple et non diagonalisable) de matrice

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

(notez que $I^2 = -1$). Les matrices $P$ vérifiant $PI = {}^tIP$ sont les matrices de la forme

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

La forme correspondante est non dégénérée ssi $a^2 + b^2 \neq 0$, et on vérifie que sa signature est $(1, -1)$ pour toute valeur de $(a, b)$. Donc la forme associée n'est jamais un produit scalaire.

# Quelques remarques sur les anneaux $\mathbb{Z}/n\mathbb{Z}$

Le contenu de cette note peut servir dans les leçons :

- Groupes finis. Exemples et applications.
- Groupe linéaire d'un e. v. de dimension finie $E$, sous-groupes de $\mathrm{GL}(E)$. Applications.
- Anneaux $\mathbb{Z}/n\mathbb{Z}$. Applications.
- Anneaux principaux. Applications.
- Opérations élémentaires sur les lignes et les colonnes d'une matrice. Exemples et applications.
- Méthodes combinatoires, problèmes de dénombrement.



## 1 Structure de $\mathbb{Z}/n\mathbb{Z}$

### 1.1 Généralités

Sur le groupe abélien $\mathbb{Z}/n\mathbb{Z}$, il n'y a qu'une structure d'anneau possible. En effet, un produit $ab$ est une somme $a + \cdots + a$ avec $b$ termes, donc toute multiplication est une itération finie d'additions et la multiplication est déterminée par l'addition.

Dit autrement, la structure de l'anneau $\mathbb{Z}/n\mathbb{Z}$ est déterminée par sa structure de groupe additif. Ceci se reflète aussi sur les éléments

inversibles de l'anneau, qui sont les générateurs du groupe additif, et sur les idéaux, qui sont les sous-groupes additifs.

Un autre fait notable est que dans l'anneau $\mathbb{Z}/n\mathbb{Z}$, comme d'ailleurs dans tout anneau fini $A$, les éléments non nuls sont soit inversibles, soit diviseurs de zéro. En effet, pour $x \in A\backslash\{0\}$, la multiplication par $x$ donne un endomorphisme de groupe abélien $m_x : A \to A$. Si $m_x$ est surjectif, alors 1 est dans l'image, donc il existe $y \in A$ tel que $xy = 1$ et $x$ est inversible. Si $m_x$ n'est pas surjectif, alors il n'est pas injectif (car $A$ est fini) et donc il y a un élément non nul $y$ dans le noyau. Alors, $xy = 0$ et $x$ est un diviseur de zéro.

Soit $n = p_1^{\alpha_1} \dots p_r^{\alpha_r}$ la décomposition en facteurs premiers de $n$. La structure algébrique de l'anneau $\mathbb{Z}/n\mathbb{Z}$ est pour l'essentiel gouvernée par la décomposition en produit donnée par l'isomorphisme du théorème des restes chinois :

$$\mathbb{Z}/n\mathbb{Z} \simeq \mathbb{Z}/p_1^{\alpha_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_r^{\alpha_r}\mathbb{Z}$$

et par la structure particulière des facteurs $\mathbb{Z}/p^\alpha\mathbb{Z}$.

### 1.2 Structure de $\mathbb{Z}/p^\alpha\mathbb{Z}$

Décrivons donc plus en détail l'anneau $A = \mathbb{Z}/p^\alpha\mathbb{Z}$. Une très bonne manière de se représenter les éléments de $A$ est d'utiliser l'écriture en base $p$ : pour tout $x \in A$, il existe des entiers uniques $0 \leqslant x_i \leqslant p-1$ tels que $x = x_0 + x_1 p + x_2 p^2 + \cdots + x_{\alpha-1} p^{\alpha-1}$. Si un élément de $A$ est écrit ainsi, on a $x \in A^*$ ssi $x_0 \neq 0$, ou encore, ssi $x \notin (p)$. En particulier, on voit que si on note $\pi : \mathbb{Z}/p^\alpha\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ la surjection canonique, alors $x$ est inversible dans $A$ si et seulement si $\pi(x)$ est inversible dans $\mathbb{Z}/p\mathbb{Z}$. (En général, si $f : R \to S$ est un morphisme d'anneaux commutatifs et unitaires, l'image d'un inversible est inversible, mais la réciproque n'est pas vraie.) De plus, l'idéal $(p)$ est égal à l'idéal des éléments nilpotents, et on a $A = A^* \sqcup (p)$.

Par ailleurs, les idéaux de $A$ forment une chaîne :

$$0 \subset (p^{\alpha-1}) \subset \cdots \subset (p^2) \subset (p) \subset A \ .$$

Ceci permet de définir la *valuation p-adique* d'un élément non nul $x \in A$ comme étant le plus grand entier $k \leqslant \alpha - 1$ tel que $x \in (p^k)$. On peut alors écrire $x = p^k u$, où $u$ est inversible dans $A$, et cette écriture est unique.

## 2 Puissances dans $\mathbb{Z}/n\mathbb{Z}$

### 2.1 Puissances $k$-ièmes

**Proposition :** *Soient $k \geqslant 2$ et $n \geqslant 2$ deux entiers. Alors, l'application $\mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ d'élévation à la puissance $k$ est bijective si et seulement si tous les facteurs premiers $p$ de $n$ sont de multiplicité 1 et tels que $p-1$ est premier avec $k$.*

**Preuve :** Notons $\varphi : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ telle que $\varphi(x) = x^k$. Cette application est multiplicative. Soit $n = p_1^{\alpha_1} \dots p_r^{\alpha_r}$ la décomposition en facteurs premiers de $n$. Par le théorème des restes chinois, on a un isomorphisme $\mathbb{Z}/n\mathbb{Z} \simeq \mathbb{Z}/p_1^{\alpha_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_r^{\alpha_r}\mathbb{Z}$. Si l'on décrit $\varphi$ via cet isomorphisme, il est clair que $\varphi(x_1, \dots, x_r) = (x_1^k, \dots, x_r^k)$, de sorte que $\varphi$ est bijective si et seulement si pour tout $i$, l'application d'élévation à la puissance $k$ dans $\mathbb{Z}/p_i^{\alpha_i}\mathbb{Z}$ est bijective. Ceci nous ramène au cas où $n = p^\alpha$.

Soit $x \in \mathbb{Z}/p^\alpha\mathbb{Z}$. Si $x$ est inversible, alors $\varphi(x)$ est inversible, c'est-à-dire qu'il n'est pas dans $(p)$. Si $x$ n'est pas inversible, il est dans $(p)$ et donc $\varphi(x) = x^k$ est dans $(p^k)$. (Pour le dire autrement, l'application $\varphi$ multiplie la valuation $p$-adique par $k$ de sorte que les éléments de l'image ont une valuation multiple de $k$.) On voit donc que si $\alpha \geqslant 2$, l'élément $p \in A$ n'est pas dans l'image de $\varphi$. Donc $\alpha = 1$ si $\varphi$ est bijectif.

L'application $\varphi$ envoie 0 sur 0 et sa restriction à $(\mathbb{Z}/p\mathbb{Z})^*$ est un morphisme de groupes multiplicatifs. Il reste à voir quand celui-ci est bijectif. Or $(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}$ et $\varphi$ s'identifie, comme endomorphisme du groupe additif $\mathbb{Z}/(p-1)\mathbb{Z}$, à la multiplication par $k$. Celle-ci est bijective ssi $k$ est premier à $p-1$. □

## 2.2 Carrés

Le résultat précédent dit que l'application d'élévation au carré ($k = 2$) dans $\mathbb{Z}/n\mathbb{Z}$ n'est bijective que lorsque $n = 2$. Donc en général, les carrés forment un sous-ensemble strict, que l'on va dénombrer, généralisant le résultat correspondant pour $\mathbb{Z}/p\mathbb{Z}$ avec $p$ premier. Ici encore, en utilisant le théorème chinois, le nombre de carrés est le produit des nombres de carrés dans des anneaux $\mathbb{Z}/p_i^{\alpha_i}\mathbb{Z}$. Ceci nous ramène au cas où $n = p^\alpha$ et nous citerons le résultat dans ce cas. Nous ne traiterons que le cas où $p \geqslant 3$, mais le cas où $p = 2$ se traite de la même manière (la seule différence provenant de la structure du groupe des inversibles).

**Lemme :** *Soit $p$ un nombre premier impair et $\alpha \geqslant 1$ un entier.*
*(i) Le cardinal de $(\mathbb{Z}/p^\alpha\mathbb{Z})^{*2}$, ensemble des carrés des éléments inversibles de $\mathbb{Z}/p^\alpha\mathbb{Z}$, est égal à $p^{\alpha-1}\frac{p-1}{2}$.*
*(ii) Soit $i$ un entier tel que $i \leqslant \alpha$. Alors, la multiplication par $p^i$ induit une injection de groupes abéliens $\mathbb{Z}/p^{\alpha-i}\mathbb{Z} \hookrightarrow \mathbb{Z}/p^\alpha\mathbb{Z}$ et l'image de $(\mathbb{Z}/p^{\alpha-i}\mathbb{Z})^{*2}$ est égale à $p^i(\mathbb{Z}/p^\alpha\mathbb{Z})^{*2}$.*
*(iii) Le cardinal de $p^i(\mathbb{Z}/p^\alpha\mathbb{Z})^{*2}$ est égal à $p^{\alpha-i-1}\frac{p-1}{2}$.*

**Preuve :** (i) Comme $p$ est impair, $(\mathbb{Z}/p^\alpha\mathbb{Z})^* \simeq \mathbb{Z}/p^{\alpha-1}(p-1)\mathbb{Z}$ et l'élévation au carré s'identifie à la multiplication par 2 dans $\mathbb{Z}/p^{\alpha-1}(p-1)\mathbb{Z}$. Comme 2 divise $p-1$, l'image est donc le sous-groupe strict engendré par 2, d'indice 2. Donc le cardinal de $(\mathbb{Z}/p^\alpha\mathbb{Z})^{*2}$ est $p^{\alpha-1}\frac{p-1}{2}$.

(ii) Il est immédiat de voir que le noyau de la multiplication par $p^i$ de $\mathbb{Z}/p^\alpha\mathbb{Z}$ dans lui-même est égal à l'idéal engendré par $p^{\alpha-i}$, d'où la première assertion. On peut décrire cette application ainsi : à $x = x_0 + x_1p + \cdots + x_{\alpha-i-1}p^{\alpha-i-1}$ on associe $p^ix = p^i(x_0 + x_1p + \cdots + x_{\alpha-i-1}p^{\alpha-i-1})$. L'image de $(\mathbb{Z}/p^{\alpha-i}\mathbb{Z})^{*2}$ est $p^i(\mathbb{Z}/p^{\alpha-i}\mathbb{Z})^{*2}$. C'est aussi $p^i(\mathbb{Z}/p^\alpha\mathbb{Z})^{*2}$, puisque dans l'écriture $p^i(x_0 + x_1p + \cdots + x_{\alpha-1}p^{\alpha-1})^2$ les termes $x_jp^j$ avec $j \geqslant \alpha - i$ sont annulés par $p^i$.

(iii) D'après (ii) le cardinal de $p^i(\mathbb{Z}/p^\alpha\mathbb{Z})^{*2}$ est égal à celui de $(\mathbb{Z}/p^{\alpha-i}\mathbb{Z})^{*2}$, d'où le résultat d'après (i). □

**Proposition :** *Si $p$ est un nombre premier impair, le nombre de carrés dans $A = \mathbb{Z}/p^\alpha\mathbb{Z}$ est égal à*

$$1 + \frac{p-1}{2}(p + p^3 + \cdots + p^{2\beta-1})$$

*si $\alpha = 2\beta$ est pair, et*

$$1 + \frac{p-1}{2}(1 + p^2 + \cdots + p^{2\beta})$$

*si $\alpha = 2\beta + 1$ est impair.*

**Preuve :** Si $x \in A$ est non nul, il s'écrit de manière unique sous la forme $x = p^iu$ avec $0 \leqslant i \leqslant \alpha - 1$ et $u \notin (p)$, c'est-à-dire $u$ inversible dans $A$. Il est clair que $x$ est un carré si et seulement si $i$ est pair et $u$ est un carré. En d'autres termes, l'ensemble des carrés non nuls dans $A$ est

$$A^{*2} \sqcup p^2A^{*2} \sqcup p^4A^{*2} \sqcup \cdots \sqcup p^{2\beta-2}A^{*2} \quad \text{si } \alpha = 2\beta \text{ est pair,}$$

$$A^{*2} \sqcup p^2A^{*2} \sqcup p^4A^{*2} \sqcup \cdots \sqcup p^{2\beta}A^{*2} \quad \text{si } \alpha = 2\beta + 1 \text{ est impair.}$$

En utilisant le lemme qui donne le cardinal de $p^{2k}A^{*2}$ et en tenant compte du fait que $0 \in A$ est un carré, on trouve que le nombre de carrés dans $A$ est

$$1 + p^{2\beta-1}\frac{p-1}{2} + p^{2\beta-3}\frac{p-1}{2} + \cdots + p\frac{p-1}{2}$$

si $\alpha = 2\beta$, et l'expression similaire si $\alpha$ est impair. □

**Références :** Je ne connais de référence ni pour la description de l'application d'élévation à la puissance $k$-ième, ni pour le calcul du nombre de carrés de $\mathbb{Z}/n\mathbb{Z}$.

# 3 Matrices à coefficients dans $\mathbb{Z}/n\mathbb{Z}$

Soit $r \geqslant 1$ un entier. Dans ce paragraphe, nous nous intéressons aux groupes linéaires $\mathrm{GL}_r(\mathbb{Z}/n\mathbb{Z})$ et $\mathrm{SL}_r(\mathbb{Z}/n\mathbb{Z})$. Nous utiliserons les remarques simples qui suivent.

Si $f : A \to B$ est un morphisme d'anneaux commutatifs et unitaires, il y a une application $\mathrm{M}_r(A) \to \mathrm{M}_r(B)$ entre les ensembles de matrices carrées de taille $r$, notée encore $f$ pour simplifier, obtenue en associant à une matrice $M = (m_{i,j})$ la matrice $f(M) = (f(m_{i,j}))$. Puisque $f$ est un morphisme d'anneaux et que l'addition et la multiplication des matrices s'expriment par des additions et des multiplications entre les coefficients des matrices, cette application $f : \mathrm{M}_r(A) \to \mathrm{M}_r(B)$ est un morphisme d'anneaux. Puisque le déterminant d'une matrice est lui aussi un polynôme en les coefficients de la matrice, on a $\det(f(M)) = f(\det(M))$. Il en découle que $f$ induit des morphismes de groupes $\mathrm{GL}_r(A) \to \mathrm{GL}_r(B)$ et $\mathrm{SL}_r(A) \to \mathrm{SL}_r(B)$.

Par ailleurs, dans le cas où l'anneau des coefficients des matrices est un anneau produit, il est clair que $\mathrm{M}_r(A \times B) \simeq \mathrm{M}_r(A) \times \mathrm{M}_r(B)$, $\mathrm{GL}_r(A \times B) \simeq \mathrm{GL}_r(A) \times \mathrm{GL}_r(B)$ et $\mathrm{SL}_r(A \times B) \simeq \mathrm{SL}_r(A) \times \mathrm{SL}_r(B)$. Pour étudier $\mathrm{GL}_r(\mathbb{Z}/n\mathbb{Z})$ et $\mathrm{SL}_r(\mathbb{Z}/n\mathbb{Z})$, utilisant le théorème des restes chinois on est ramené au cas où $n = p^\alpha$.

## 3.1 Nombre d'éléments de $\mathrm{GL}_r(\mathbb{Z}/n\mathbb{Z})$ et $\mathrm{SL}_r(\mathbb{Z}/n\mathbb{Z})$

**Proposition :** *On a*

$$|\mathrm{GL}_r(\mathbb{Z}/p^\alpha\mathbb{Z})| = p^{(\alpha-1)r^2}(p^r-1)(p^r-p)\dots(p^r-p^{r-1}).$$

**Preuve :** Commençons par le cas $\alpha = 1$. Dans ce cas, l'anneau de coefficients est le corps $k = \mathbb{Z}/p\mathbb{Z}$. Une matrice est dans $\mathrm{GL}_r(\mathbb{Z}/p\mathbb{Z})$ ssi ses vecteurs colonnes forment une base. Le premier vecteur doit être non nul, il y a donc $p^r - 1$ façons de le choisir. Le deuxième vecteur ne doit pas être dans la droite engendrée par le premier, il y a donc $p^r - p$ façons de le choisir. En continuant ainsi, on trouve $|\mathrm{GL}_r(\mathbb{Z}/p\mathbb{Z})| = (p^r-1)(p^r-p)\dots(p^r-p^{r-1})$.

Passons au cas général. On notera $\mathbb{Z}/p^\alpha\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$, $x \mapsto \overline{x}$ le morphisme de réduction. Rappelons-nous que $x$ est inversible ssi $\overline{x}$ est inversible (voir 1.2). Nous allons voir que le morphisme induit $\nu : \mathrm{GL}_r(\mathbb{Z}/p^\alpha\mathbb{Z}) \to \mathrm{GL}_r(\mathbb{Z}/p\mathbb{Z})$ est surjectif. En effet, si $M \in \mathrm{GL}_r(\mathbb{Z}/p\mathbb{Z})$ et qu'on considère une matrice $N \in \mathrm{M}_r(\mathbb{Z}/p^\alpha\mathbb{Z})$ obtenue en relevant de manière arbitraire les coefficients de $M$, on a $\overline{\det(N)} = \det(M)$ qui est inversible. Donc $\det(N)$ est inversible, i.e. $N \in \mathrm{GL}_r(\mathbb{Z}/p^\alpha\mathbb{Z})$ et $N$ est un antécédent pour $M$. On regarde maintenant le noyau $H = \ker(\nu)$. C'est l'ensemble des matrices $\mathrm{Id} + N$ où $N$ est à coefficients dans $p(\mathbb{Z}/p^\alpha\mathbb{Z}) \simeq \mathbb{Z}/p^{\alpha-1}\mathbb{Z}$. Comme les matrices ont $r^2$ coefficients, on trouve $|H| = (p^{\alpha-1})^{r^2}$. Finalement $|\mathrm{GL}_r(\mathbb{Z}/p^\alpha\mathbb{Z})| = |\mathrm{GL}_r(\mathbb{Z}/p\mathbb{Z})|.|H|$ et ceci donne le résultat annoncé. □

**Proposition :** *On a* $|\mathrm{SL}_r(\mathbb{Z}/p^\alpha\mathbb{Z})| = p^{(\alpha-1)(r^2-1)}(p^r-1)(p^r-p)\dots(p^r-p^{r-2})p^{r-1}$.

**Preuve :** On considère le morphisme déterminant det : $\mathrm{GL}_r(\mathbb{Z}/p^\alpha\mathbb{Z}) \to (\mathbb{Z}/p^\alpha\mathbb{Z})^*$. Il est surjectif, car tout $x \in (\mathbb{Z}/p^\alpha\mathbb{Z})^*$ est le déterminant d'une matrice de dilatation diagonale $(x, 1, \dots, 1)$. Il s'ensuit que le cardinal du noyau, le groupe spécial linéaire, est

$$|\mathrm{SL}_r(\mathbb{Z}/p^\alpha\mathbb{Z})| = \frac{|\mathrm{GL}_r(\mathbb{Z}/p^\alpha\mathbb{Z})|}{p^{\alpha-1}(p-1)}\ .$$

Compte tenu de la proposition précédente, ceci mène au résultat annoncé. □

**Références :** Le calcul du cardinal de $\mathrm{GL}_r(\mathbb{Z}/p\mathbb{Z})$ et d'autres groupes linéaires sur les corps finis est fait dans Perrin [P]. Le calcul du cardinal de $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$ peut être trouvé dans [FGN2], exercice 3.23 (lire la fin de la correction).

## 3.2 Surjection $\mathrm{SL}_r(\mathbb{Z}) \to \mathrm{SL}_r(\mathbb{Z}/n\mathbb{Z})$

On peut se poser la question de savoir si toute matrice inversible à coefficiens dans $\mathbb{Z}/n\mathbb{Z}$ peut être relevée en une matrice inversible à co-

efficients dans $\mathbb{Z}$. Mais ceci est presque tout le temps faux, pour la raison que le groupe des inversibles de $\mathbb{Z}/n\mathbb{Z}$ est plus gros que le groupe des inversibles de $\mathbb{Z}$ : il est clair qu'une matrice de $\mathrm{GL}_r(\mathbb{Z}/n\mathbb{Z})$ de déterminant inversible, mais distinct de $\pm 1$, ne peut pas être relevée dans $\mathrm{GL}_r(\mathbb{Z})$. Le théorème suivant est donc assez surprenant :

**Théorème :** *Le morphisme de réduction* $\mathrm{SL}_r(\mathbb{Z}) \to \mathrm{SL}_r(\mathbb{Z}/n\mathbb{Z})$ *est surjectif.*

**Preuve :** On fait une récurrence sur $r$. Comme $\mathrm{SL}_1(\mathbb{Z}) \simeq \mathrm{SL}_1(\mathbb{Z}/n\mathbb{Z}) \simeq 1$, le résultat est clair pour $r = 1$. Supposons-le vrai pour l'entier $r-1$, et soit $A \in \mathrm{M}_r(\mathbb{Z})$ une matrice carrée telle que $\det(A) \equiv 1\ (n)$. D'après le théorème des invariants de similitude, il existe deux matrices $U, V$ dans $\mathrm{GL}_r(\mathbb{Z})$ telles que $UAV$ est une matrice diagonale, d'éléments $a_1, \ldots, a_m$. Posons $b = a_2 \ldots a_m$ et considérons les matrices

$$W = \begin{pmatrix} b & 1 & & & \\ b-1 & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}, \; X = \begin{pmatrix} 1 & -a_2 & & & \\ 0 & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix},$$

$$A' = \begin{pmatrix} 1 & 0 & & & \\ 1-a_2 & a_1 a_2 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}.$$

Comme $a_1 b = \det(A) \equiv 1\ (n)$, on voit que $WUAVX \equiv A'\ (n)$. Par l'hypothèse de récurrence, la matrice carrée de taille $(r-1, r-1)$ en bas à droite de $A'$ se relève en une matrice $C \in \mathrm{SL}_{r-1}(\mathbb{Z})$. On vérifie alors facilement que

$$B = U^{-1}W^{-1} \left(\begin{array}{c|c} 1 & 0 \\ \hline 1-a_1 & \\ & C \\ 0 & \end{array}\right) X^{-1}V^{-1}$$

est une matrice dans $\mathrm{SL}_r(\mathbb{Z})$ qui relève $A$. ▫

**Remarques et références :** Ce théorème est une jolie application du théorème des invariants de similitude, sous forme matricielle. La démonstration donnée ici est la reproduction fidèle de celle que l'on trouve en pages 20-21 du livre de Shimura [Shi]. Dans le cas $r = 2$, la preuve du théorème se trouve aussi dans [FGN2], exercice 3.23, p. 204, et dans le livre d'Hellegouarch [H], chapitre 5, § 3, p. 295.

Une des raisons de l'importance de ce théorème provient de l'étude des *groupes fuchsiens* et des *sous-groupes de congruence de* $\mathrm{SL}_2(\mathbb{Z})$ tels que le sous-groupe :

$$\Gamma(n) = \ker\left(\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})\right).$$

Ce groupe intervient dans l'étude des formes modulaires, qui sont l'un des ingrédients de la preuve du théorème de Fermat. Vous trouverez plus de détails sur tout cela dans le livre d'Hellegouarch [H].

## Références


[FGN2] S. Francinou, H. Gianella, S. Nicolas, *Exercices de Mathématiques Oraux X-ENS, Algèbre 2*, Cassini.

[H] Y. Hellegouarch, *Invitation aux Mathématiques de Fermat-Wiles*, Masson, 1997.

[P] D. Perrin, *Cours d'Algèbre*, Ellipses.

[Shi] G. Shimura, *Introduction to the Arithmetic Theory of Automorphic Forms*, Princeton University Press, 1971.

# L'exponentielle complexe

D'un point de vue historique, les concepts familiers d'angle, cosinus, sinus, exponentielle, et même le nombre $\pi$ qui est au départ de cette aventure, sont apparus de manière plus chaotique que ce que l'enseignement de collège et lycée peut laisser croire. Dans ce petit texte, nous expliquons comment on peut présenter ces concepts de manière tout à fait différente, en tirant profit du développement de l'analyse et de ses fondations rigoureuses au $18^{\text{ème}}$, $19^{\text{ème}}$ et $20^{\text{ème}}$ siècles. Cette présentation met l'exponentielle sur le devant de la scène. Le cosinus, le sinus et le nombre $\pi$ ne sont définis qu'ensuite, à partir de celle-ci.

Pour que ce texte soit lisible avec des connaissances légères, un grand soin est mis pour éviter tant que possible la théorie des fonctions holomorphes ou le calcul différentiel avancé (théorème des fonctions implicites notamment). Cependant, un peu de topologie est inévitable pour démontrer que l'exponentielle induit un paramétrage périodique du cercle $\mathbb{U}$ des nombres complexes de module 1. Ceci nécessite quelques arguments plus sophistiqués (abordables tout de même au niveau L3) qui ont été réunis dans la dernière partie, qu'il est possible de ne pas lire.

## 1 Définition et principales propriétés

Dans le corps des complexes $\mathbb{C}$, une partie est dite *ouverte* si elle est réunion de disques ouverts $D(z_0, r) = \{z \in \mathbb{C},\ |z - z_0| < r\}$. La donnée de ces ouverts définit une topologie sur $\mathbb{C}$.

**Proposition 1** *La série de fonctions $\sum_{n\geqslant 0} z^n/n!$ est normalement convergente sur tout compact $K \subset \mathbb{C}$. Sa somme est notée $\exp(z)$ ou $e^z$.*

**Démonstration :** Soit $M$ tel que $|z| \leqslant M$ pour tout $z \in K$. On a alors

$$\sum_{n\geqslant 0} \sup_{z\in K} \frac{|z|^n}{n!} \leqslant \sum_{n\geqslant 0} \frac{M^n}{n!} \ .$$

On applique à $u_n = M^n/n!$ le critère de d'Alembert pour les séries. Ici $u_{n+1}/u_n = M/n + 1$ tend vers 0, donc on a convergence de la série $\sum M^n/n!$ et il en découle que la série de fonctions converge normalement sur $K$. □

La propriété fondamentale de l'exponentielle est la suivante :

**Théorème 1** *Pour tous nombres complexes $s, t$ on a $e^s e^t = e^{s+t}$. En particulier l'exponentielle définit un morphisme de groupes* $\exp : \mathbb{C} \to \mathbb{C}^*$.

**Démonstration :** Soit $K$ un compact qui contient $s$ et $t$. Pour $s$ fixé, la série qui définit $e^s$ est absolument convergente, et idem pour $e^t$. D'après le théorème sur le produit de Cauchy on a

$$e^s e^t = \sum_{n\geqslant 0} \frac{s^n}{n!} \sum_{n\geqslant 0} \frac{t^n}{n!} = \sum_{n\geqslant 0} \sum_{p+q=n} \frac{s^p}{p!}\frac{t^q}{q!} \ .$$

Par ailleurs, en utilisant la formule du binôme,

$$\sum_{p+q=n} \frac{s^p}{p!}\frac{t^q}{q!} = \sum_{p=0}^{n} \frac{s^p}{p!}\frac{t^{n-p}}{(n-p)!} = \frac{1}{n!}\sum_{p=0}^{n} \binom{n}{p} s^p t^{n-p} = \frac{1}{n!}(s+t)^n \ ,$$

donc finalement $e^s e^t = \sum_{n\geqslant 0} (s+t)^n/n! = e^{s+t}$. □

Voici une propriété élémentaire de l'exponentielle, utile pour la suite :

**Lemme 1** *Pour tout $z \in \mathbb{C}$, on a $\exp(\overline{z}) = \overline{\exp(z)}$.*

**Démonstration :** La conjugaison complexe est une application $\mathbb{R}$-linéaire, donc continue. Il en découle que le conjugé de $\exp(z)$, c'est-à-dire le conjugué de la limite des sommes partielles $\sum_{n\leqslant N} z^n/n!$, est égal à la limite des sommes partielles $\sum_{n\leqslant N} \overline{z}^n/n!$, c'est-à-dire à $\exp(\overline{z})$. □

Le théorème 1 exprime une propriété algébrique, mais nous allons voir que celle-ci a un pendant analytique tout aussi fondamental. Commençons pas le cas réel :

**Proposition 2** *La restriction de l'exponentielle à $\mathbb{R}$ est de classe $C^\infty$, de dérivée $(\exp_{\mathbb{R}})' = \exp_{\mathbb{R}}$, et définit un isomorphisme de groupes $\exp_{\mathbb{R}} : \mathbb{R} \simeq \mathbb{R}^{+*}$.*

**Démonstration :** Par définition $\exp_{\mathbb{R}}$ est développable en série entière au voisinage de tout point, donc elle est de classe $C^\infty$. D'après le théorème 1, pour $x, h$ réels on a

$$e^{x+h} = e^x e^h = e^x(1 + h + h^2/2 + \dots) = e^x + e^x h + h^2\varphi(x,h)$$

pour une certaine fonction $\varphi$ dépendant de manière $C^\infty$ de $h$. Alors le calcul de la dérivée par taux d'accroissement fournit immédiatement que la valeur de la dérivée en $x$ est $e^x$.

On considère ensuite la fonction $\ln : \mathbb{R}^{+*} \to \mathbb{R}$ définie par $\ln(x) = \int_1^x \frac{dt}{t}$. Cette fonction est dérivable et par définition $\ln'(x) = 1/x$ pour tout $x > 0$. Soient les fonctions dérivables

$$\begin{array}{lll} f : \mathbb{R}^{+*} \to \mathbb{R}^{+*} & , & f(x) = x\exp(-\ln(x)) \ , \\ g : \mathbb{R} \to \mathbb{R} & , & g(x) = \ln(\exp(x)) - x \ . \end{array}$$

Deux calculs de dérivée montrent que $f'(x) = g'(x) = 0$ pour tout $x$. Donc ces fonctions sont constantes, $f(x) = f(1) = 1$ et $g(x) = g(0) = 0$. Ceci signifie que $\exp_{\mathbb{R}}$ et $\ln$ sont des bijections réciproques l'une de l'autre, donc $\exp_{\mathbb{R}}$ est un isomorphisme de groupes d'inverse $\ln$. ▫

Avant de passer au cas complexe, prenons le temps d'un commentaire. Une fonction $f : \mathbb{C} \to \mathbb{C}$ est une fonction de deux variables réelles, à valeurs dans un espace à deux dimensions réelles. Pour se ramener à la situation plus simple des applications de $\mathbb{R}$ dans $\mathbb{R}$, il est naturel d'essayer de composer $f$ avec des applications $\mathbb{R} \to \mathbb{C}$ à la source et $\mathbb{C} \to \mathbb{R}$ au but.

Il existe deux applications $\mathbb{C} \to \mathbb{R}$ privilégiées : la partie réelle et la partie imaginaire, et nous les utilisons toujours sans même y penser. D'ailleurs l'étude de $f$ est équivalente à l'étude simultanée de $\Re(f)$ et $\Im(f)$. Quant à la composition à la source par des applications $\mathbb{R} \to \mathbb{C}$, elle correspond à tracer des chemins dans $\mathbb{C}$ et regarder le comportement de $f$ le long de ces chemins. On fera grand usage de cette technique de chemins dans la suite, et dès maintenant pour donner la caractérisation différentielle de l'exponentielle :

**Théorème 2** *L'exponentielle est l'unique fonction $f : \mathbb{C} \to \mathbb{C}$ satisfaisant la propriété suivante : pour tout $z \in \mathbb{C}$, la fonction $g : \mathbb{R} \to \mathbb{C}$ telle que $g(t) = f(tz)$ est dérivable et vérifie*

$$\left\{ \begin{array}{l} g'(t) = zg(t) \\ g(0) = 1 \ . \end{array} \right.$$

**Démonstration :** D'après le théorème 1, pour $t, h$ réels on a

$$e^{(t+h)z} = e^{tz}e^{hz} = e^{tz}(1+hz+h^2z^2/2+\dots) = e^{tz}+ze^{tz}h+h^2\varphi(t,h,z)$$

pour une certaine fonction $\varphi$ dépendant de manière $C^\infty$ de $h$. Utilisant cela, on vérifie que $f = \exp$ satisfait la propriété annoncée puisque le calcul de la dérivée par taux d'accroissement fournit $g'(t) = ze^{tz} = zg(t)$ et de plus $g(0) = 1$.

Réciproquement soit $f$ une fonction vérifiant cette propriété. Posons $h(t) = e^{-tz}g(t)$, il vient immédiatement $h'(t) = -ze^{-tz}g(t) + e^{-tz}g'(t) = 0$ pour tout $t \in \mathbb{R}$. Donc $h$ est constante : $h(t) = a$, et $g(t) = ae^{tz}$. Comme $g(0) = 1$, on trouve $g(t) = e^{tz}$ puis $f(z) = g(1) = e^z$. ▫

**Théorème 3** *L'application $\exp : \mathbb{C} \to \mathbb{C}^*$ est surjective.*

Une raison pour cela est que l'ensemble d'arrivée $\mathbb{C}^*$ est connexe. À titre de comparaison, il est intéressant de constater que $\mathbb{R}^*$ n'est pas connexe, et que l'exponentielle des nombres réels a pour image $\mathbb{R}^{+*}$, qui est la composante connexe de l'élément neutre.

**Démonstration :** Non seulement $\mathbb{C}^*$ est connexe, mais il est même connexe par arcs. Étant donné $z \in \mathbb{C}^*$, on peut l'atteindre depuis l'élément neutre 1 par un chemin tracé dans $\mathbb{C}^*$ :

$$\zeta : [0,1] \to \mathbb{C}^* \text{ continu et tel que } \zeta(0) = 1 \text{ et } \zeta(1) = z.$$

L'idée intuitive est d'utiliser ce chemin pour construire un « petit bout de logarithme », c'est-à-dire une fonction $\ell : [0,1] \to \mathbb{C}$ telle que pour tout $t \in [0,1]$ on a $\exp(\ell(t)) = \zeta(t)$. Pour guider ce choix, on s'inspire de l'observation que formellement, si $\ell(t) = \ln(\zeta(t))$,

alors $\ell'(t) = \zeta'(t)/\zeta(t)$. Rigoureusement maintenant, définissons $\ell : [0,1] \to \mathbb{C}$ par

$$\ell(t) = \int_0^t \frac{\zeta'(s)}{\zeta(s)}\, ds \ .$$

Comme $\zeta$ est continue et ne s'annule pas, alors $\zeta'/\zeta$ est continue donc $\ell$ est dérivable de dérivée $\ell'(t) = \zeta'(t)/\zeta(t)$. Posons $g(t) = \zeta(t)\exp(-\ell(t))$. Utilisant le théorème 2, on trouve

$$g'(t) = \zeta'(t)\exp(-\ell(t)) - \frac{\zeta'(t)}{\zeta(t)}\zeta(t)\exp(-\ell(t)) = 0$$

donc $g$ est constante : $g(t) = g(0) = 1$. Il s'ensuit que $\exp(\ell(t)) = \zeta(t)$ pour tout $t$. En particulier $\exp(\ell(1)) = \zeta(1) = z$, donc l'exponentielle est surjective. ▫

J'insiste sur le fait qu'il n'existe pas de fonction réciproque globale pour l'exponentielle. On ne sait construire que des « petits bouts de logarithme » sur des ouverts $V \subset \mathbb{C}^*$, mais ceci ne donne pas lieu à une construction globale car ces petits bouts ne coïncident pas sur les différents ouverts.

Le résultat le plus difficile concernant l'exponentielle est le suivant :

**Théorème 4** *L'application* $\exp : \mathbb{C} \to \mathbb{C}^*$ *est ouverte, c'est-à-dire que l'image de tout ouvert est un ouvert.*

Pour ne pas alourdir l'exposition, nous l'admettrons pour l'instant. Nous proposons au lecteur de lire la partie 3 pour la démonstration.

## 2 L'exponentielle et le cercle

Dans cette partie, on utilise l'exponentielle pour décrire le groupe multiplicatif $\mathbb{U}$ des nombres complexes de module 1. Ce groupe est muni de la topologie induite de celle de $\mathbb{C}$, c'est-à-dire que ses ouverts sont les $U \cap \mathbb{U}$ avec $U$ ouvert dans $\mathbb{C}$. On note que si $x \in \mathbb{R}$, on a

$$|e^{ix}|^2 = e^{ix}\overline{e^{ix}} = e^{ix}e^{\overline{ix}} = e^{ix}e^{-ix} = 1$$

d'après le lemme 1. Ainsi, on peut définir un morphisme de groupes

$$\varphi : \mathbb{R} \to \mathbb{U} \quad , \quad \varphi(x) = e^{ix} \ .$$

On va justifier que, tout comme l'exponentielle, $\varphi$ est surjectif et ouvert.

**Lemme 2** *Le morphisme* $\varphi$ *est surjectif.*

**Démonstration :** Soit $z \in \mathbb{U}$. D'après le théorème 3 il existe un complexe $x + iy$ tel que $e^{x+iy} = z$. Comme $e^{iy} \in \mathbb{U}$ on trouve $e^x = ze^{-iy} \in \mathbb{U}$. Comme par ailleurs $e^x$ est un réel strictement positif, on trouve $e^x = 1$ donc $x = 0$ par la proposition 2. Donc $z = e^{iy}$ est dans l'image de $\varphi$. ▫

**Lemme 3** *Le morphisme* $\varphi$ *est ouvert.*

**Démonstration :** Les ouverts de $\mathbb{R}$ sont les réunions d'intervalles ouverts $]a,b[$ donc ce sont exactement les parties de la forme $U \cap \mathbb{R}$ avec $U$ ouvert de $\mathbb{C}$. Il s'agit de montrer que l'image de $U \cap \mathbb{R}$ est ouverte dans $\mathbb{U}$. La partie $iU := \{iz, z \in U\}$ est un ouvert de $\mathbb{C}$, donc par le théorème 4 il en va de même de $V := \exp(iU)$. Par conséquent $\exp(iU) \cap \mathbb{U}$ est un ouvert de $\mathbb{U}$, donc il suffit de démontrer que $\varphi(U \cap \mathbb{R}) = \exp(iU) \cap \mathbb{U}$. L'inclusion directe est claire. Pour la réciproque, soit $z = e^{i(x+iy)} \in \mathbb{U}$ avec $x + iy \in U$. Alors $e^{-y} = ze^{-ix} \in \mathbb{U}$. Par le même argument que dans la preuve du lemme 2, on trouve $y = 0$ donc $z = \varphi(x) \in \varphi(U \cap \mathbb{R})$. ▫

**Théorème 5** *Il existe un unique nombre réel strictement positif noté* $\pi$ *tel que le noyau de* $\varphi : \mathbb{R} \to \mathbb{U}$ *est égal à* $2\pi\mathbb{Z}$. *Le morphisme* $\varphi$ *induit un isomorphisme de groupes*

$$\overline{\varphi} : \mathbb{R}/2\pi\mathbb{Z} \simeq \mathbb{U} \ .$$

**Démonstration :** D'après le résultat classique sur les sous-groupes additifs de $\mathbb{R}$, le noyau $N := \ker(\varphi)$ est soit dense dans $\mathbb{R}$, soit de la forme $a\mathbb{Z}$ pour un certain $a \in \mathbb{R}$, $a \geqslant 0$.

Supposons $N$ dense dans $\mathbb{R}$. Comme $\varphi(x) = 1$ pour tout $x \in N$ et comme $\varphi$ est continu, il s'ensuit que $\varphi$ est constant égal à 1. Ceci est impossible car $\varphi$ est surjectif, donc $N = a\mathbb{Z}$.

Supposons $a = 0$. Alors $N = 0$ donc $\varphi$ est un isomorphisme de groupes. Par ailleurs $\varphi$ est continu. De plus, son inverse $\psi = \varphi^{-1} : \mathbb{U} \to \mathbb{R}$ doit l'être aussi puisque si $U \subset \mathbb{R}$ est un ouvert alors $\psi^{-1}(U) = \varphi(U)$ est ouvert étant donné que $\varphi$ est ouvert par 3. Il s'ensuit que $\varphi$ est un homéomorphisme. Ceci est impossible, puisque $\mathbb{R}$ n'est pas compact alors que $\mathbb{U}$ l'est.

Donc $a > 0$, et en posant $\pi := a/2$ on obtient le résultat annoncé. ▫

Si $z$ est un nombre complexe non nul, son module $\rho = |z|$ est non nul et il est clair que $\rho^{-1}z$ est un nombre complexe de module 1. D'après la proposition ci-dessus, il existe un nombre réel $\theta$ tel que $\rho^{-1}z = e^{i\theta}$. De plus, deux tels nombres réels $\theta$, $\theta'$ diffèrent d'un multiple de $2\pi$, c'est-à-dire que la classe $\overline{\theta} \in \mathbb{R}/2\pi\mathbb{Z}$ est uniquement déterminée.

**Définition 1** *Pour tout nombre complexe $z \neq 0$, l'écriture $z = \rho e^{i\theta}$ est appelée la* forme polaire *de $z$ et l'élément $\overline{\theta} \in \mathbb{R}/2\pi\mathbb{Z}$ est appelé l'*argument *de $z$ et noté* $\mathrm{Arg}(z)$.

**Corollaire 1** *On a un isomorphisme de groupes commutatifs*

$$\begin{aligned} &\mathbb{C}^* \to \mathbb{R}^{+*} \times \mathbb{R}/2\pi\mathbb{Z} \\ &z \mapsto (|z|, \mathrm{Arg}(z)) \ . \end{aligned}$$

**Démonstration :** Il est clair que c'est un isomorphisme. Son inverse associe à $(\rho, \overline{\theta})$ le complexe $z = \rho e^{i\theta}$, où $\theta$ est un représentant quelconque de $\overline{\theta}$. ▫

Pour manipuler l'argument d'un nombre complexe, on en choisira systématiquement un représentant $\theta \in \mathbb{R}$. Une des difficultés de la notion d'argument est la confusion fréquente entre l'argument et ses représentants réels. Par exemple, dans l'écriture polaire on appelle souvent $\theta$ l'argument alors que ce n'est que l'un de ses représentants. Cette confusion n'est évidemment pas souhaitable.

Nous continuons le déroulement des propriétés de l'exponentielle et des diverses quantités qui lui sont reliées.

**Théorème 6 (Euler)** *On a $e^{i\pi} = -1$.*

**Démonstration :** Par définition $\pi$ est le plus petit réel strictement positif tel que $e^{2i\pi} = 1$. Posons $z = e^{i\pi}$, on a donc $z \neq 1$ et $z^2 = e^{2i\pi} = 1$. Dans le corps $\mathbb{C}$ l'équation $X^2 = 1$ a pour seules racines 1 et $-1$. Donc $z = -1$. ▫

On peut ensuite développer la trigonométrie sur cette lancée. On introduit les fonctions trigonométriques cos, sin... et leurs propriétés principales se déduisent de celles de l'exponentielle. Nous n'indiquons que le début de cette belle histoire.

**Définitions 1** *Pour tout nombre complexe $z$, on définit :*

(1) *le* cosinus $\cos(z) = \frac{e^{iz}+e^{-iz}}{2}$,

(2) *le* sinus $\sin(z) = \frac{e^{iz}-e^{-iz}}{2i}$,

(3) *le* cosinus hyperbolique $\mathrm{ch}(z) = \frac{e^{z}+e^{-z}}{2} = \cos(iz)$,

(4) *le* sinus hyperbolique $\mathrm{sh}(z) = \frac{e^{z}-e^{-z}}{2} = -i\sin(iz)$.

Nous laissons au lecteur le soin d'ajouter à cette liste la définition des tangente, cotangente, tangente hyperbolique et cotangente hyperbolique.

**Proposition 3** *On a les propriétés :*

(1) *si $z$ est réel, alors* $\cos(z)$, $\sin(z)$, $\mathrm{ch}(z)$ *et* $\mathrm{sh}(z)$ *sont réels.*

(2) $e^{iz} = \cos(z) + i\sin(z)$ *et* $e^z = \mathrm{ch}(z) + \mathrm{sh}(z)$.

(3) $\cos^2(z) + \sin^2(z) = 1$.

(4) $\mathrm{ch}^2(z) - \mathrm{sh}^2(z) = 1$.

**Démonstration :** La vérification de ces propriétés est facile et laissée au lecteur. □

Une autre observation parfois utile est que pour $x$ réel, $\cos(x)$ et $\sin(x)$ sont la partie réelle et la partie imaginaire de $e^{ix}$, et $\mathrm{ch}(x)$ et $\mathrm{sh}(x)$ sont la partie paire et la partie impaire de $e^x$.

**Proposition 4** *Pour tous nombres complexes $a$ et $b$, on a :*

(1) $\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b)$.

(2) $\sin(a+b) = \sin(a)\cos(b) - \sin(b)\cos(a)$.

**Démonstration :** Ces formules découlent de la multiplicativité de l'exponentielle, et nous laissons la vérification au lecteur. □

Pour faire le lien avec la présentation habituelle du nombre $\pi$, nous devons le relier au périmètre du cercle. Ce lien résulte de la *définition de la longueur* qui est rappelée dans la preuve du résultat que voici :

**Proposition 5** *Le rapport entre le périmètre d'un cercle et son diamètre est égal à $\pi$.*

**Démonstration :** Rappelons que si $c : [a;b] \to \mathbb{R}^2$ est un arc de classe $C^1$ qui est injectif sauf en un nombre fini de points, alors cet arc possède une *longueur* définie par

$$L(c) = \int_a^b \parallel c'(t) \parallel \, dt \ .$$

Revenant à notre cercle $\mathcal{C}$, on peut le supposer centré en l'origine 0 et c'est alors l'ensemble des points de coordonnées $(x,y)$ vérifiant $x^2+y^2 = R^2$ où $R>0$ est le rayon. À un tel point $(x,y)$ correspond le nombre complexe $z = x+iy$ qui est alors de module $R$. D'après le théorème 5 et la définition de $\pi$, le paramétrage

$$\begin{array}{rcl} c : [0;2\pi] & \to & \mathcal{C} \\ t & \mapsto & Re^{it} \end{array}$$

est injectif sauf en $O$ et $2\pi$. On a $c'(t) = Rie^{it}$ qui est de module $R$ donc

$$L(\mathcal{C}) = \int_0^{2\pi} \parallel Rie^{it} \parallel \, dt = \int_0^{2\pi} R\, dt = 2\pi R \ .$$

C'est bien le produit du diamètre par $\pi$. □

## 3 L'exponentielle est ouverte

Cette partie est consacrée à la preuve du théorème mentionné plus haut :

**Théorème 4** *L'application* $\exp : \mathbb{C} \to \mathbb{C}^*$ *est ouverte.*

Pour démontrer que l'exponentielle est surjective (théorème 3), on a utilisé un chemin tracé dans $\mathbb{C}^*$ et joignant 1 à $z$. Il est tout à fait remarquable que le raisonnement fonctionnait pour n'importe quel chemin $\zeta : [0,1] \to \mathbb{C}^*$ joignant 1 à $z$.

A priori, un changement de chemin $\zeta$ change l'antécédent $\ell(1)$, et une meilleure notation serait donc $\ell(1,\zeta)$. Le fait que l'exponentielle est ouverte va résulter de la possibilité de choisir le chemin $\zeta$ de sorte que l'antécédent $\ell(1,\zeta)$ varie continuement avec $z$, au moins localement au voisinage de 1. Tout simplement, on joint 1 à $z$ par la ligne droite :

$$\zeta(t) = 1 + t(z-1) \ .$$

La contrainte que le chemin soit tracé tout entier dans $\mathbb{C}^*$ nous force à éliminer les $z$ qui sont sur le demi-axe réel négatif. C'est une illustration du fait qu'on ne sait construire que des « petits bouts de logarithme » sur des ouverts de $\mathbb{C}^*$. Cependant, il se trouve qu'il sera largement suffisant pour nous de construire $\ell(1,\zeta)$ sur le disque ouvert centré en 1 et de rayon 1, noté $D$. On considère donc la fonction $\ell : [0,1] \times D \to \mathbb{C}$ définie par

$$\ell(t,z) = \int_0^t \frac{z-1}{1+s(z-1)}\, ds \ .$$

**Lemme 4** *La fonction $\lambda : D \to \mathbb{C}$ définie par $\lambda(z) = \ell(1,z)$ est continue.*

**Démonstration :** On peut invoquer le théorème de continuité dans les intégrales à paramètre. Une autre possibilité est de passer par les séries entières, comme suit. Pour tout $(s,z) \in [0,1] \times D$ on a $|s(z-1)| < 1$ donc on a le développement en série convergente

$$\frac{1}{1+s(z-1)} = \sum_{n \geqslant 0} (-s(z-1))^n \ .$$

Vérifions que cette série de fonctions converge uniformément sur $[0,1]$. D'abord notons que la somme de cette série est une fonction de $s$ continue sur $[0,1]$, donc son module est majoré par un certain $M > 0$. Le reste au rang $N$ de la série est

$$R_N(s) = \sum_{n \geqslant N} (-s(z-1))^n = \frac{(-s(z-1))^N}{1+s(z-1)} \ ,$$

donc $\sup_{s \in [0,1]} |R_N(s)| \leqslant M\,|z-1|^N$. Puisque $|z-1| < 1$, ce sup tend vers 0 lorsque $N \to \infty$, d'où convergence uniforme. On sait qu'on peut alors intervertir la somme infinie et l'intégrale :

$$\lambda(z) = (z-1) \int_0^1 \sum_{n \geqslant 0} (-s(z-1))^n \, ds = \sum_{n \geqslant 0} \int_0^1 (-1)^n (z-1)^{n+1} s^n \, ds$$

$$= \sum_{n \geqslant 0} (-1)^n \frac{(z-1)^{n+1}}{n+1} \ .$$

Ainsi $\lambda$ est développable en série entière sur $D$, donc continue sur $D$. □

**Remarque 1** Dans le développement en série entière de $\lambda$, on reconnaît le développement habituel de $\ln(z)$ en $z = 1$, ce qui n'est pas un hasard...

Dans la dernière partie de la preuve, nous utiliserons constamment les morphismes de translation, qui sont des homéomorphismes :

- pour $z_0 \in \mathbb{C}$, la translation additive de $\mathbb{C}$ dans $\mathbb{C}$ définie par $z \mapsto z_0 + z$. L'image d'une partie $A \subset \mathbb{C}$ par cette translation est notée $z_0 + A$.

- pour $z_0 \in \mathbb{C}^*$, la translation multiplicative de $\mathbb{C}^*$ dans $\mathbb{C}^*$ définie par $z \mapsto z_0 z$. L'image d'une partie $A \subset \mathbb{C}^*$ par cette translation est notée $z_0 A$.

Soit $U \subset \mathbb{C}$ un ouvert, il s'agit de démontrer que $V = \exp(U)$ est un ouvert. Soit $y \in V$, donc $y = \exp(x)$, $x \in U$, et vérifions que la partie

$$\Omega = y\lambda^{-1}(-x+U)$$

est un voisinage ouvert de $y$ dans $V$. Dans cette notation, $-x+U$ est le translaté additif de $U$, qui contient 0. Puis $\lambda^{-1}(-x+U)$ est sa préimage par $\lambda$, qui contient donc 1. Enfin $\Omega$ est le translaté multiplicatif par $y$, il contient donc $y$.

Comme $-x+U$ est ouvert et $\lambda$ continue d'après le lemme 4, alors $\lambda^{-1}(-x+U)$ est ouvert et donc $\Omega$ aussi.

Enfin vérifions que $\Omega \subset V$. Par définition de $\Omega$, pour tout $z \in \Omega$ il existe $v \in D$ tel que $z = yv$ et $\lambda(v) + x \in U$. Or $\exp(\lambda(v)) = v$ par la preuve du théorème 3, et $\exp(x) = y$. Partant de $\lambda(v) + x \in U$ et prenant l'exponentielle, on trouve

$$z = vy = \exp(\lambda(v)) \exp(x) \in \exp(U) = V \ ,$$

ce que l'on voulait. □

# Un peu de culture mathématique sur les groupes de Lie et l'exponentielle

Voici quelques commentaires sur l'importance de l'exponentielle dans la théorie des groupes de Lie. Un *groupe de Lie* est un groupe qui est muni d'une structure différentiable qu'on appelle une structure de *variété*. Son espace tangent en l'élément neutre $e \in G$ est un espace vectoriel, qui se trouve muni d'une application bilinéaire appelée *crochet* qui en fait une *algèbre de Lie*. Voyons comment présenter ces objets de la manière la plus simple possible.

**1. La notion de variété avec les mains.** Pour passer sous silence la définition précise de ce qu'est une variété, disons que l'idée grossière est que c'est un espace topologique $M$ sur lequel on sait définir la notion d'espace tangent en chacun de ses points. Il semble clair que si $M \subset \mathbb{R}^n$ est un $k$-plan affine ($k \leqslant n$), alors l'espace tangent de $M$ en chacun de ses points est $M$ lui-même, donc $M$ est une variété. Par exemple $\mathbb{R}^n$ (affine) ou n'importe quel ouvert $U \subset \mathbb{R}^n$ est une variété, et pour tout $x \in U$ on a $T_xU = \mathbb{R}^n$ (vectoriel). Plus généralement, une façon naturelle de définir des objets géométriques est de considérer une fonction $f : \mathbb{R}^n \to \mathbb{R}^p$ de classe $C^\infty$ et de regarder l'ensemble $Z(f) = \{x \in \mathbb{R}^n\,,\, f(x) = 0\}$. Alors, si la différentielle en tout point $d_xf : \mathbb{R}^n \to \mathbb{R}^p$ est surjective, $M = Z(f)$ est une variété (une sous-variété de $\mathbb{R}^n$) ([3]). Cette condition imposée sur la différentielle de $f$ provient du théorème des fonctions implicites ; illustrons tout de suite sa necessité. Dans le cas $n = 2$, $p = 1$, on sait qu'en un point $(a, b) \in Z(f)$, donc $f(a, b) = 0$, la tangente à la courbe $Z(f)$ est l'ensemble des $(v, w) \in \mathbb{R}^2$ tels que

$$\frac{\partial f}{\partial x}(a,b)(v-a) + \frac{\partial f}{\partial y}(a,b)(w-b) = 0\,,$$

sauf lorsque $(\partial f/\partial x)(a, b) = (\partial f/\partial y)(a, b) = 0$ où il y a un problème. Dans l'exemple de la fonction $f_1(x, y) = x^2 + y^2 - 1$, en tout point $x_0 = (a, b) \in Z(f_1)$ on a $d_{x_0}f_1 = (\partial f_1/\partial x)(a, b) = (\partial f_1/\partial y)(a, b) \neq (0, 0)$ donc $M_1 = Z(f_1)$ est bien une variété. En revanche pour $f_2(x, y) = x^2 - y^2$, si $x = (a, b) \in Z(f_2)$ on a $d_xf_2 = (2a, -2b)$ donc pour $x_0 = (0, 0) \in Z(f_2)$ on a $d_{x_0}f_2 = 0$. Ainsi $M_2 = Z(f_2)$ n'est pas une variété et on dit que $x_0$ est un point *singulier* (mais $M_2\backslash\{x_0\}$ est une variété).

**2. L'espace tangent avec les mains.** L'espace tangent en un point $x$ d'une variété $M$ est une approximation linéaire de $M$ en $x$, de la même façon que la tangente à une courbe est une approximation de la courbe. Sans entrer dans une définition trop précise, disons qu'on peut définir l'espace tangent en $x \in M$ comme étant l'ensemble des vecteurs tangents $\gamma'(0)$ aux courbes $t \mapsto \gamma(t)$ de classe $C^1$ tracées sur $M$ et telles que $\gamma(0) = x$ (avec $t$ dans un petit intervalle $[-\epsilon, \epsilon]$). Il est noté $T_xM$ et ses points sont des vecteurs. Prenons l'exemple de $M = \{x \in \mathbb{R}^n\,,\, f(x) = 0\}$ où $f : \mathbb{R}^n \to \mathbb{R}^p$ est une submersion, et $x_0 \in M$, alors on peut montrer que le plan tangent $T_{x_0}M$ est le plan de $\mathbb{R}^n$ d'équations $d_{x_0}f(v) = 0$. En d'autres termes, notant $f = (f_1, \ldots, f_p)$ et $\mathrm{Jac}(f)$ la matrice des $(\partial f_i/\partial x_j)(x_0)$,

$$T_{x_0}M = \{v = (v_1, \ldots, v_n) \text{ t.q. } \mathrm{Jac}(f)v = 0\}\,.$$

Une petite variation de $x$ dans son espace tangent peut être pensée comme un élément de la forme $x + v$ où $v \in T_xM$ est un vecteur. Si $f : M \to N$ est une application différentiable entre deux variétés (ce que l'on n'a pas défini mais on fait comme si), avec $x \in M$ et $y = f(x)$, alors il y a une application linéaire induite qui est la différentielle $d_xf : T_xM \to T_yN$. Si $v \in T_xM$ et $w = d_xf(v)$ alors la petite variation $x + v$ s'envoie sur la petite variation $y + w$, en d'autres termes on retrouve une forme familière :

$$f(x + v) = f(x) + d_xf(v) + (\text{termes d'ordre supérieur en } v)\,.$$

Ces remarques sont pratiques pour faire des calculs simples, comme nous le ferons ci-dessous.

[3] Une telle $f$ est appelée une *submersion*. À vrai dire, pour avoir une bonne notion de plans tangents il suffit de prendre $f$ de classe $C^1$, et on obtient alors une notion de variété de classe $C^1$. Nous ne considérons que des variétés de classe $C^\infty$ pour simplifier.

**3. Groupes de Lie avec les mains.** Un groupe de Lie est un ensemble $G$ qui est muni d'une structure de groupe et d'une structure de variété, les deux étant compatibles au sens où la multiplication $G \times G \to G$ et l'inversion $G \to G$ sont des application différentiables. Par exemple il est clair que le groupe des nombres complexes de module 1, identifié au cercle $S^1$, est un groupe de Lie. Voici d'autres exemples : $GL_n(\mathbb{R})$, $GL_n(\mathbb{C})$ (ce sont des ouverts de certains $\mathbb{R}^n$), $SL_n(\mathbb{R})$, $SL_n(\mathbb{C})$, $O_n(\mathbb{R})$, $U_n(\mathbb{C})$ (ici c'est moins évident que ce sont des variétés, il faut le démontrer en écrivant ces ensembles sous la forme $M = Z(f)$ et vérifier que $f$ est à différentielles $d_x f$ surjectives ; exercice : faites-le !).

**4. L'algèbre de Lie d'un groupe de Lie, avec les mains.** Pour tout groupe de Lie $G$, on note $\mathfrak{g} = T_eG$ l'espace tangent en l'élément neutre $e \in G$. Cet espace vectoriel a une importante capitale car il se trouve qu'il existe une application bilinéaire $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, appelée le *crochet de Lie*, qui renferme presque toute la structure de $G$. Pour le trouver on considère l'application de conjugaison $c : G \times G \to G$, $(g, h) \mapsto ghg^{-1}$, qui renferme de l'information sur la multiplication et sur l'inversion de $G$. Si on différentie $c$ en l'élément $(e, e)$ on obtient une application

$$d_{(e,e)}c : T_{(e,e)}(G \times G) \to T_eG \ .$$

Or on a un isomorphisme naturel $T_{(e,e)}(G \times G) \simeq T_eG \times T_eG$ (c'est facile à montrer ; croyez-moi) donc on a $d_{(e,e)}c : T_eG \times T_eG \to T_eG$. Malheureusement, comme toute différentielle, $d_{(e,e)}c$ est une application linéaire, or on veut une application bilinéaire. Il va falloir ruser un peu. Ce que l'on fait c'est que pour tout $g \in G$ on considère

$$c_g : G \to G \quad , \quad h \mapsto ghg^{-1}$$

et sa différentielle $d_e(c_g) : T_eG \to T_eG$. Ceci est une application linéaire, autrement dit

$$d_e(c_g) \in \mathscr{L}(T_eG, T_eG) \ .$$

On a obtenu une application

$$\varphi : G \to \mathscr{L}(T_eG, T_eG) \quad , \quad g \mapsto \varphi(g) = d_e(c_g) \ .$$

Comme $\varphi(e) = \mathrm{Id}$, si on différentie de nouveau en $e$ on trouve

$$d_e\varphi : T_eG \to T_{\mathrm{Id}}(\mathscr{L}(T_eG, T_eG)) \simeq \mathscr{L}(T_eG, T_eG)$$

puisque l'espace tangent d'un espace vectoriel en un point s'identifie à l'espace vectoriel lui-même. Ainsi on a obtenu l'application recherchée :

$$[\cdot, \cdot] : T_eG \times T_eG \to T_eG$$

qui envoie $(v, w)$ sur $[v, w] := [(d_e\varphi)(v)](w)$. Voici donc défini le crochet en général. On peut montrer qu'il satisfait une propriété très importante appelée *identité de Jacobi* : $[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0$.

**5. L'exemple des groupes de matrices.** Il est assez facile de calculer le crochet de Lie pour tous les groupes de Lie qui sont des sous-groupes fermés de $GL_n(\mathbb{R})$ ou $GL_n(\mathbb{C})$. Posons, disons, $G = GL_n(\mathbb{R})$. C'est un ouvert de $\mathbb{R}^2 = M_n(\mathbb{R})$, défini par la non-annulation du déterminant, donc c'est bien une variété. Les formules qui donnent la multiplication et l'inversion montrent que ces applications sont bien différentiables (exercice : vérifiez que c'est clair pour vous). Donc $G$ est un groupe de Lie. De plus comme c'est un ouvert de l'espace vectoriel $M_n(\mathbb{R})$, l'espace tangent en chacun de ses points s'identifie à $M_n(\mathbb{R})$. On va voir que le crochet de Lie de deux vecteurs, c'est-à-dire deux matrices $V, W \in M_n(\mathbb{R})$, est donné par la formule familière $[V, W] = VW - WV$, valable pour les algèbres de Lie de $GL_n(\mathbb{R})$, $GL_n(\mathbb{C})$ ou de tous leurs sous-groupes fermés.

Pour différentier en l'identité on considèrera des petites variations $M = \mathrm{Id} + V$, $N = \mathrm{Id} + W$ (voir ci-dessus). Ici $c_M : GL_n \to GL_n$ est l'application $N \mapsto MNM^{-1}$. (Exercice : tout à l'heure on a obervé que $d_{(e,e)}c$ est une application linéaire : calculez-la !). Comme

$$M(\mathrm{Id} + W)M^{-1} = \mathrm{Id} + MWM^{-1} + (\text{termes d'ordre supérieur en } W)$$

(en fait ici les termes d'ordre supérieur sont nuls), on trouve

$$d_{\mathrm{Id}}c_M : M_n \to M_n \quad , \quad W \mapsto MWM^{-1} \ .$$

D'où l'application

$$\varphi : GL_n \to \mathscr{L}(M_n, M_n) \quad , \quad M \mapsto d_{\mathrm{Id}}c_M \ .$$

Cette application envoie l'identité $\mathrm{Id} = \mathrm{Id}_n \in GL_n$ sur l'identité $\mathrm{Id} = \mathrm{Id}_{n^2} \in \mathscr{L}(M_n, M_n)$. On les notes toutes deux Id sans distinction. Il s'agit maintenant de calculer

$$d_{\mathrm{Id}}\varphi : T_{\mathrm{Id}}GL_n = M_n \to T_{\mathrm{Id}}\mathscr{L}(M_n, M_n) = \mathscr{L}(M_n, M_n)$$

Notons $\psi_V : M_n \to M_n$ l'endomorphisme $\psi_V(W) = VW - WV$, formule linéaire en $V$. On va montrer que

$$\varphi(\mathrm{Id} + V) = \mathrm{Id} + \psi_V + (\text{termes d'ordre supérieur en } V)$$

Notons $\varphi_{\mathrm{Id}+V}$ au lieu de $\varphi(\mathrm{Id}+V)$, on a

$$\begin{aligned}\varphi_{\mathrm{Id}+V}(W) &= (\mathrm{Id}+V)W(\mathrm{Id}+V)^{-1} = (\mathrm{Id}+V)W(\mathrm{Id}-V + (\text{t.o.s. en } V)\\ &= W + VW - WV + (\dots) = \mathrm{Id}(W) + \psi_V(W) + (\text{t.o.s.})\end{aligned}$$

et on obtient bien $\varphi(\mathrm{Id}+V) = \mathrm{Id} + \psi_V + (\text{t.o.s.})$. En conclusion

$$[V, W] = [(d_{\mathrm{Id}}\varphi)(V)](W) = \psi_V(W) = VW - WV\ .$$

Voici fini ce calcul. (Exercice : montrez que ce crochet vérifie bien l'identité de Jacobi.)

**6. L'algèbre de Lie et l'exponentielle.** Soit $G$ un groupe de Lie qui est un sous-groupe fermé de $GL_n(\mathbb{R})$. Nous allons voir que son algèbre de Lie $\mathfrak{g} = T_eG$ est intimement liée à l'exponentielle via :

$$T_eG = \{A \in M_n(\mathbb{R})\ ,\ \exp(tA) \in G \text{ pour tout } t \in \mathbb{R}\}\ .$$

La propriété capitale pour voir ceci est le fait que exp: $\mathrm{M}_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$ est analytique (donc de classe $C^\infty$) et $d(\exp)_0 = \mathrm{Id}$, donc elle réalise un difféomorphisme entre un voisinage $\Omega_0$ de $0 \in \mathrm{M}_n(k)$ et un voisinage $\Omega_{\mathrm{Id}}$ de $\mathrm{Id} \in \mathrm{GL}_n(k)$. Notons $\log : \Omega_{\mathrm{Id}} \to \Omega_0$ son inverse. On va utiliser la définition de $T_eG$ comme ensemble des vecteurs tangents $\gamma'(0)$ aux courbes $t \mapsto \gamma(t) \in G$, $t \in [-\epsilon, \epsilon]$, de classe $C^1$ et telles que $\gamma(0) = \mathrm{Id}$.

Il est clair que $\{A \in M_n(\mathbb{R}), \exp(tA) \in G, \forall t \in \mathbb{R}\} \subset T_eG$ car si $\exp(tA) \in G$ pour tout $t \in \mathbb{R}$, alors la courbe $\gamma(t) = \exp(tA)$ vérifie $\gamma'(0) = A$ et donc $A \in T_eG$. Réciproquement soit $A = \gamma'(0)$ pour une certaine courbe $\gamma : [-\epsilon, \epsilon] \to G$ centrée en $A$, et $t \in \mathbb{R}$. Pour tout $n$ assez grand on a :

- $t/n < \epsilon$ et $\gamma(t/n) = \mathrm{Id} + tA/n + O(1/n^2)$,
- $\gamma(t/n) \in \Omega_{\mathrm{Id}}$, de sorte que
- $\gamma(t/n) = \exp(\log(\gamma(t/n))) = \exp(\log(\mathrm{Id} + tA/n + O(1/n^2)))$.

On en déduit que la limite de $\gamma(t/n)^n$ est égale à $\exp(tA)$, or comme $G$ est fermé, cette limite est dans $G$. Donc $T_eG \subset \{A \in M_n(\mathbb{R}), \exp(tA) \in G, \forall t \in \mathbb{R}\}$.

**7. Le dictionnaire groupes ↔ algèbres.** Il y a un analogue à l'exponentielle pour les groupes de Lie généraux, c'est un morphisme $\exp_G : \mathfrak{g} \to G$ que l'on appelle encore l'exponentielle. Ces concepts sont fondamentaux dans la théorie car on montre qu'on peut établir une sorte de dictionnaire entre les groupes de Lie et leurs algèbres de Lie. Voici quelques exemples :

- soit $G$ un groupe de Lie connexe, alors il y a une bijection entre l'ensemble des sous-groupes de Lie connexes $H \subset G$ et l'ensemble des sous-algèbres de Lie $\mathfrak{h} \subset \mathfrak{g}$.

- soient $G_1$, $G_2$ des groupes de Lie connexes, alors il y a une bijection entre l'ensemble des morphismes de groupes de Lie connexes $G_1 \to G_2$ et l'ensemble des morphismes d'algèbres de Lie $\mathfrak{g}_1 \subset \mathfrak{g}_2$ (en fait il faut faire une hypothèse supplémentaire qui est que $G_1$ est *simplement connexe*).

Ceci est magique, car on ramène dans une certaine mesure l'étude des groupes de Lie, des objets très compliqués (pleins de topologie et de géométrie différentielle) à celles de leurs algèbres, bien plus simple (c'est de l'algèbre linéaire). (Question : pourquoi s'est-on limité ci-dessus à des groupes *connexes* ?)

**8. Exemples.** On note les algèbres de Lie par des lettres gothiques. Soit $k = \mathbb{R}$ ou $\mathbb{C}$. On dispose de deux techniques pour calculer l'algèbre de Lie d'un sous-groupe fermé $G \subset GL_n(k)$ : soit comme espace tangent $T_eG$, soit comme $\{A \in M_n(\mathbb{R}), \exp(tA) \in G, \forall t \in \mathbb{R}\}$.

- $\mathfrak{gl}_n(k) = M_n(k)$ (clair ; exercice : dites pourquoi !).

- $\mathfrak{sl}_n(k)$ est l'ensemble des matrices de trace nulle. En effet faisons-le par la première méthode (espace tangent). $SL_n(k) = Z(f)$ où $f(M) = \det(M) - 1$, donc l'espace tangent est défini par l'équation $d_{\mathrm{Id}}f(V) = 0$. Or il est facile de voir que $\det(\mathrm{Id} + V) = 1 + \mathrm{tr}(V) + (\text{t.o.s. en } V)$ donc $d_{\mathrm{Id}}f(V) = \mathrm{tr}(V)$. On trouve donc l'équation $\mathrm{tr}(V) = 0$. (Exercice : vérifiez le calcul $\det(\mathrm{Id} + V) = 1 + \mathrm{tr}(V) + (\text{t.o.s. en } V)$ et plus généralement calculez la différentielle du déterminant en une matrice $M$.) Faisons-le maintenant par la deuxième méthode. Soit $V$ telle que $\det(\exp(tV)) = 1$ pour tout $t \in \mathbb{R}$. Alors $\det(\exp(tV)) = \exp(\mathrm{tr}(tV)) = 1$ (car les valeurs propres de $\exp(tV)$ sont les exponentielles des valeurs propres de $tV$ ; exercice : vérifiez-le et vérifiez la formule $\det(\exp(A)) = \exp(\mathrm{tr}(A))$). Il s'ensuit que la trace de $tV$ est nulle (si $k = \mathbb{R}$) ou multiple de $2i\pi$ (si $k = \mathbb{C}$), et ceci pour tout $t \in \mathbb{R}$, donc la trace est nulle.

- $\mathfrak{so}_n(\mathbb{R})$ est l'ensemble des matrices antisymétriques de $M_n(\mathbb{R})$. (Exercice.)

- $\mathfrak{su}_n(\mathbb{R})$ est l'ensemble des matrices antihermitiennes de $M_n(\mathbb{C})$. (Exercice ; dites au passage ce qu'est une matrice antihermitienne.)

- l'algèbre de Lie du groupe de Lie réel des quaternions de norme 1 est égale à l'ensemble des quaternions purs. (Exercice.)

- l'algèbre de Lie du groupe de Lie réel des nombres complexes non nuls est... l'algèbre de Lie du groupe de Lie réel des nombres complexes de module 1 est...

**Bibliographie**


Mneimné, Testard, Introduction à la théorie des groupes de Lie classiques, *Hermann*
Rouvière, Petit guide de Calcul différentiel (...), *Cassini*


# Représentations linéaires des groupes finis

Réf : Serre, *Représentations linéaires des groupes finis*, Hermann.

On considère un groupe fini $G$ et des $\mathbb{C}$-espaces vectoriels de dimension finie $V, V_1, V_2$, etc. On note $\mathrm{Hom}(V_1, V_2)$ l'espace vectoriel des application $\mathbb{C}$-linéaires de $V_1$ dans $V_2$ et $\mathrm{End}(V)$ la $\mathbb{C}$-algèbre $\mathrm{Hom}(V, V)$.

## 1 Introduction

L'étude de la réduction d'un endomorphisme $f$ d'un espace vectoriel $V$ est le tout premier pas de la *théorie des représentations*. L'étude classique de familles d'endomorphismes $f_i$ qui commutent entre eux (codiagonalisabilité, cotrigonalisabilité) amène au second pas, qui est l'étude de familles quelconques.

Bien sûr, on gagne en structure et en souplesse en prenant pour objet d'étude la sous-$k$-algèbre $A$ de $\mathrm{End}(V)$ engendrée par les $f_i$. Le bon cadre abstrait pour faire cela est de considérer une $k$-algèbre abstraite $A$ et d'étudier les morphismes $A \to \mathrm{End}(V)$, appelés *représentations linéaires de $A$*.

Si les $f_i$ sont inversibles, il est naturel de considérer, plutôt que $A$, le sous-groupe $G$ de $\mathrm{GL}(V)$ qu'ils engendrent. Le bon cadre abstrait pour faire cela est de considérer un groupe abstrait $G$ et d'étudier les morphismes $G \to \mathrm{GL}(V)$, appelés *représentations linéaires de $G$*.

## 2 Notions de base

Les notions de cette partie font sens pour un groupe $G$ pas nécessairement fini, un corps de base $k$ pas nécessairement isomorphe à $\mathbb{C}$, et des espaces vectoriels pas nécessairement de dimension finie.

**Définitions 1** *Soient $G$ un groupe et $V, V_1, V_2$ des $k$-espaces vectoriels.*

(1) *Une* représentation *de $G$ dans un $k$-espace vectoriel $V$ est un morphisme de groupes $\rho : G \to \mathrm{GL}(V)$, c'est-à-dire une action de $G$ sur $V$ par automorphismes $k$-linéaires.*

(2) *Un* morphisme *de $\rho^1 : G \to \mathrm{GL}(V_1)$ dans $\rho^2 : G \to \mathrm{GL}(V_2)$ est une application linéaire $f : V_1 \to V_2$ telle que $f \circ \rho^1_s = \rho^2_s \circ f$, pour tout $s \in G$.*

(3) *Une* sous-représentation *$W$ de $V$ est un sous-espace vectoriel qui est $G$-stable, c'est-à-dire tel que $\rho_s(W) \subset W$ pour tout $s \in G$. La* représentation quotient *de $V$ par $W$ est la représentation $\overline{\rho} : G \to \mathrm{GL}(V/W)$ telle que $\overline{\rho}_s : V/W \to V/W$ est le morphisme induit de $\rho_s : V \to V$.*

(4) *Soit $R$ un espace vectoriel de base $(e_t)_{t\in G}$, par exemple $R = k^G$ et $e_t$ égal à l'indicatrice de $\{t\}$. La* représentation régulière *est la représentation $\rho : G \to \mathrm{GL}(R)$ définie par $\rho_s(e_t) = e_{st}$.*

**Remarque 2** On utilise souvent des notations et une terminologie simplifiées : une représentation est désignée par la seule lettre $V$, l'action $\rho$ étant sous-entendue ; on note $s(v)$, $s.v$ ou $sv$ au lieu de $\rho_s(v)$ ; un morphisme de représentations $f : V_1 \to V_2$ est souvent appelé $G$-morphisme de $V_1$ dans $V_2$ et on note $\mathrm{Hom}_G(V_1, V_2)$ l'ensemble de ces morphismes.

**Exemples 1** *Les définitions de base de la théorie comprennent d'autres notions importantes, mais que nous utiliserons moins ici (nous utilisons les notations simplifiées de 2) :*

(1) *La* représentation duale *d'une représentation $V$ est la représentation d'espace vectoriel sous-jacent le dual $V'$ définie par $s.\varphi = \varphi \circ s^{-1}$.*

(2) *Étant données deux représentations $V_1$ et $V_2$, leur* somme directe *est la représentation d'espace sous-jacent $V_1 \oplus V_2$ définie par $s.(v_1 \oplus v_2) = s(v_1) \oplus s(v_2)$. La* représentation Hom *est la représentation d'espace sous-jacent l'espace des applications linéaires* $\mathrm{Hom}(V_1, V_2)$ *définie par $s.\varphi = s \circ \varphi \circ s^{-1}$.*

**Exercice 1** *On reprend les notations de 1. Si $V$ est un espace vectoriel quelconque, on appelle* représentation triviale d'espace $V$ *et on note $V^{\mathrm{triv}}$ la représentation de $G$ sur $V$ définie par $s.v = v$, pour tout $s \in G$. Si $W$ est une représentation de $G$, on note $W^G = \{v \in V, \forall s \in G, s.v = v\}$ l'espace vectoriel des points fixes. Montrez que :*

(1) *il existe une bijection canonique*

$$\mathrm{Hom}_G(V^{\mathrm{triv}}, W) \xrightarrow{\sim} \mathrm{Hom}(V, W^G),$$

(2) *on a* $\mathrm{Hom}_G(V_1, V_2) = \mathrm{Hom}(V_1, V_2)^G$.

**Lemme 1** *Soit $f : V_1 \to V_2$ un morphisme de représentations de $G$. Alors, le noyau* $\ker(f)$ *est une sous-représentation de $V_1$, l'image* $\mathrm{im}(f)$ *est une sous-représentation de $V_2$, et l'isomorphisme* $V_1/\ker(f) \simeq \mathrm{im}(f)$ *est un isomorphisme de représentations.*

**Démonstration :** C'est une vérification facile. □

## 3 Semi-simplicité

À partir de maintenant, les hypothèses faites au début du texte deviennent véritablement importantes : *$G$ est fini, le corps de base est $k = \mathbb{C}$, et les espaces vectoriels sont de dimension finie.*

Dans toute la suite, on dira simplement *représentation* au lieu de « représentation linéaire de dimension finie ». La dimension d'une représentation est aussi appelée son *degré*.

**Théorème 1** *Soit $V$ une représentation de $G$. Alors, toute sous-représentation $W \subset V$ admet un supplémentaire stable, c'est-à-dire un supplémentaire en tant que représentation.*

**Démonstration :** Rappelons la correspondance entre supplémentaires de $W$ et projecteurs sur $W$, donnée ainsi : à $W'$ on associe le projecteur sur $W$ parallèlement à $W'$, et à $p$ on associe son noyau.

Partons d'un projecteur $p : V \to V$ sur $W$. Définissons une nouvelle application linéaire par

$$p^0 = \frac{1}{|G|} \sum_{t\in G} tpt^{-1}.$$

Comme $t(W) \subset W$ pour tout $t$, on voit que $p^0(V) \subset W$. De plus, si $x \in W$ on a

$$p^0(x) = \frac{1}{|G|} \sum_{t \in G} t(p(\underbrace{t^{-1}x}_{\in W})) = \frac{1}{|G|} \sum_{t \in G} t(t^{-1}x) = x$$

donc $p^0$ est un projecteur sur $W$. Enfin $p^0$ est un $G$-morphisme car

$$sp^0s^{-1} = \frac{1}{|G|} \sum_{t \in G} stpt^{-1}s^{-1} = \frac{1}{|G|} \sum_{u \in G} upu^{-1} = p^0.$$

Donc $W^0 = \ker(p^0)$ est un supplémentaire stable de $W$. □

**Définition 1** *Une représentation $V$ d'un groupe $G$ est dite* irréductible *ou* simple *si elle est non nulle et si ses seules sous-représentations sont $0$ et $V$. Une représentation est dite* semi-simple *si elle est somme directe de sous-représentations simples.*

Bien sûr, il y a une analogie avec le concept de simplicité pour les groupes : on cherche à décomposer les représentations en morceaux plus petits. Par ailleurs, il y a un lien direct avec le concept d'endomorphisme semi-simple d'un espace vectoriel : il est équivalent de dire que $u \in \mathrm{End}(V)$ est semi-simple au sens habituel, ou que $V$ est semi-simple comme représentation de l'algèbre $\mathbb{C}[u]$ des polynômes en $u$, ou encore, que $V$ est semi-simple comme représentation du groupe (infini) $G = \mathbb{C}[u] \cap \mathrm{GL}(V)$.

**Théorème 2** *Toute représentation est semi-simple.*

On prendra garde que ce résultat n'est plus vrai sur corps de caractéristique divisant l'ordre de $G$, ou si $G$ n'est pas fini.

**Démonstration :** Soit $V = V_1 \oplus \cdots \oplus V_s$, avec $s \leqslant \dim(V)$, une décomposition maximale (i.e. $s$ maximal) en somme directe de sous-représentations non nulles. Par le théorème précédent, s'il existe un indice $i$ et une sous-représentation stricte $W_i \subset V_i$, celle-ci possède un supplémentaire stable $W_i'$. On a alors $V_i = W_i \oplus W_i'$ en contradiction avec la maximalité de $s$. □

Cette décomposition n'est pas unique : par exemple si $V$ est triviale (voir 1), une décomposition en sous-représentations n'est rien d'autre qu'une décomposition en somme directe de droites et il y a une infinité de façons de faire cela. Mais on verra (théorème 4) que le nombre de sous-représentations isomorphes à une représentation irréductible donnée est, lui, unique.

# 4 Caractères et fonctions centrales

**Définition 2** *Soit $\rho : G \to \mathrm{GL}(V)$ une représentation de $G$. Le* caractère *de $V$ est la fonction $\chi_\rho : G \to \mathbb{C}$ définie par $\chi_\rho(s) = \mathrm{tr}(\rho_s)$. On dit qu'un caractère est* irréductible *si la représentation l'est.*

L'intérêt des caractères vient du fait remarquable qu'ils déterminent les représentations à isomorphisme près, comme nous le verrons (corollaire 4).

Voyons maintenant quelques propriétés élémentaires des caractères. Dans la suite, on notera parfois $s_V$ au lieu de $\rho_s$.

**Proposition 1** *Si $\dim(V) = n$ et $\chi = \chi_\rho$, alors :*

(1) $\chi(1) = n$,

(2) $\chi(s^{-1}) = \chi(s)^*$ *où $z^*$ est le complexe conjugué de $z$,*

(3) $\chi(tst^{-1}) = \chi(s)$.

**Démonstration :** Le point (1) vient de ce que $\chi(1) = \mathrm{tr}(\mathrm{Id}) = n$. Le point (2) vient du fait que si l'on note $d = |G|$, on a $(s_V)^d = \mathrm{Id}$. Ainsi $s_V$ est diagonalisable à valeurs propres $\lambda_i$ racines $d$-ièmes de l'unité, de sorte que $\lambda_i^{-1} = \lambda_i^*$. Alors $(s_V)^{-1}$ a pour valeurs propres les $\lambda_i^*$, donc

$$\chi(s^{-1}) = \mathrm{tr}((s_V)^{-1}) = \sum \lambda_i^* = (\sum \lambda_i)^* = \mathrm{tr}(s_V)^* = \chi(s)^*.$$

Enfin (3) est une propriété élémentaire de la trace. □

**Proposition 2** *Soient $V_1, V_2$ deux représentations de $G$ de caractères $\chi_1, \chi_2$. Soit $V = V_1 \oplus V_2$ la représentation somme directe, de caractère $\chi$. Alors, on a $\chi = \chi_1 + \chi_2$.*

**Démonstration :** En effet, pour tout $s \in G$ on a $s_V = s_{V_1} \oplus s_{V_2}$. Matriciellement, après choix d'une base adaptée à la décomposition $V_1 \oplus V_2$, cet endomorphisme se représente par une matrice diagonale par blocs. On trouve alors $\chi(s) = \mathrm{tr}(s_V) = \mathrm{tr}(s_{V_1}) + \mathrm{tr}(s_{V_2}) = \chi_1(s) + \chi_2(s)$. □

On voit que l'ensemble des caractères est un sous-ensemble de l'espace vectoriel de fonctions $\mathcal{F}(G, \mathbb{C})$ qui est stable par somme, mais n'est stable ni par passage à l'opposé, ni par multiplication par un scalaire $\lambda \in \mathbb{C}$ (à cause de la propriété $\chi(1) = n$ de 1). Mais il y a un candidat naturel à être un plus petit sous-espace vectoriel contenant les caractères, c'est l'espace des fonctions centrales :

**Définition 3** *Une fonction $f : G \to \mathbb{C}$ est* centrale *si elle vérifie $f(tst^{-1}) = f(s)$ pour tous $s, t \in G$, ou de manière équivalente, $f(uv) = f(vu)$ pour tous $u, v \in G$.*

Ainsi, une fonction centrale est une fonction qui est constante sur les classes de conjugaison de $G$. Elle passe donc au quotient en une fonction sur $G/\sim$, où $\sim$ désigne la conjugaison, et réciproquement. En conclusion, l'ensemble des fonctions centrales sur $G$ est un sous-espace vectoriel de $\mathcal{F}(G, \mathbb{C})$ isomorphe à $\mathcal{F}(G/\sim, \mathbb{C})$. En particulier, sa dimension est égale au cardinal de $G/\sim$, le nombre de classes de conjugaison de $G$.

Nous allons étudier les caractères et les fonctions centrales à l'aide de deux formes binaires naturelles : si $\varphi, \psi$ sont des fonctions sur $G$, on définit

(1) une forme bilinéaire symétrique par $\langle \varphi, \psi \rangle = \frac{1}{|G|} \sum_{t \in G} \varphi(t)\psi(t^{-1})$,

(2) un produit scalaire hermitien par $(\varphi, \psi) = \frac{1}{|G|} \sum_{t \in G} \varphi(t)\psi(t)^*$.

Noter que $\langle -, - \rangle$ est bilinéaire alors que $(-, -)$ est linéaire en la première variable et semi-linéaire en la seconde variable. Par ailleurs, si $\chi$ est un caractère, alors on a $\langle \varphi, \chi \rangle = (\varphi, \chi)$.

# 5 Lemme de Schur

**Lemme 2 (Lemme de Schur)** *Soient $V_1, V_2$ deux représentations irréductibles de $G$. Alors :*

$$\mathrm{Hom}_G(V_1, V_2) \simeq \begin{cases} 0 & \text{si } V_1 \not\simeq V_2, \\ \mathbb{C} & \text{si } V_1 \simeq V_2. \end{cases}$$

**Démonstration :** Soit $f : V_1 \to V_2$ un morphisme de $G$-représentations. Si $f \neq 0$, alors $\ker(f) \neq V_1$ donc $\ker(f) = 0$ par irréductibilité de $V_1$. Dans ce cas, $\mathrm{im}(f) \neq 0$ donc $\mathrm{im}(f) = V_2$ par irréductibilité de $V_2$. Alors $f$ induit un isomorphisme de $V_1$ sur $V_2$. Par contraposée, ceci donne $\mathrm{Hom}_G(V_1, V_2) = 0$ si $V_1 \not\simeq V_2$.

Si $V_1 \simeq V_2 \simeq V$, alors $f : V \to V$ possède au moins une valeur propre $\lambda$. Le morphisme $f - \lambda\,\mathrm{Id} : V \to V$ est encore un $G$-morphisme ; son noyau est non nul par choix de $\lambda$, donc égal à $V$ par irréductibilité. Ainsi $f = \lambda\,\mathrm{Id}$. □

**Corollaire 1** *Soient $V_1, V_2$ deux représentations irréductibles de $G$ et $h : V_1 \to V_2$ une application linéaire quelconque. Posons $h^0 = \frac{1}{|G|} \sum_{t \in G} t^{-1}ht$. Alors :*

(1) *$h^0 = 0$ si $V_1 \not\simeq V_2$,*

(2) *$h^0 = \frac{1}{n}\mathrm{tr}(h)$ si $V_1 = V_2 = V$, où $n = \dim(V)$.*

**Démonstration :** On a tout fait pour que $h^0$ soit un $G$-morphisme. Par le lemme de Schur, dans le cas (1) on a $h^0 = 0$. Dans le cas (2), il existe un $\lambda$ tel que $h^0 = \lambda\,\mathrm{Id}$, et en prenant la trace on trouve $n\lambda = \mathrm{tr}(h^0) = \mathrm{tr}(h)$. □

**Corollaire 2** *Soient $V_1, V_2$ deux représentations irréductibles de $G$ données, après choix de bases sur $V_1$ et $V_2$, par les matrices $\rho_t^1 = t_{V_1}$ et $\rho_t^2 = t_{V_2}$ :*

$$\rho_t^1 = (r_{i_1 j_1}(t)) \qquad \text{et} \qquad \rho_t^2 = (r_{i_2 j_2}(t))$$

*où les $r_{i_1 j_1}, r_{i_2 j_2}$ sont des fonctions sur $G$. On a, pour tous $i_1, j_1, i_2, j_2$ :*

(1) $\langle r_{i_2j_2}, r_{j_1i_1}\rangle = 0$ *si* $V_1 \not\simeq V_2$,

(2) $\langle r_{i_2j_2}, r_{j_1i_1}\rangle = \frac{1}{n}\delta_{i_2i_1}\delta_{j_2j_1}$ *si* $V_1 = V_2 = V$, *où* $n = \dim(V)$.

**Démonstration :** Dans les bases $(e_{i_1})$ et $(\epsilon_{i_2})$ choisies sur $V_1$ et $V_2$, une application linéaire $h : V_1 \to V_2$ est déterminée par $h(e_{i_1}) = \sum_{i_2} x_{i_2i_1}\epsilon_{i_2}$. Sa matrice est alors $M_h = (x_{i_2i_1})$. De même $M_{h^0} = (x^0_{i_2i_1})$ où

$$\begin{aligned} x^0_{i_2i_1} &= \frac{1}{|G|}\sum_{t\in G, j_2, j_1} r_{i_2j_2}(t^{-1})x_{j_2j_1}r_{j_1i_1}(t) \\ &= \frac{1}{|G|}\sum_{j_2,j_1}\left(\sum_{t\in G} r_{i_2j_2}(t^{-1})r_{j_1i_1}(t)\right)x_{j_2j_1}. \end{aligned}$$

Le corollaire 1 au lemme de Schur dit que si $V_1 \not\simeq V_2$, on doit avoir $x^0_{i_2i_1} = 0$ identiquement en les $x_{j_2j_1}$. Il s'ensuit que chaque terme de la somme ci-dessus doit s'annuler, c'est-à-dire

$$\frac{1}{|G|}\sum_{t\in G} r_{i_2j_2}(t^{-1})r_{j_1i_1}(t) = 0.$$

C'est le contenu du point (1). On déduit le point (2) du corollaire 1 de manière identique. □

## 6 Orthogonalité des caractères

On rappelle (*cf* 2) qu'un caractère irréductible est le caractère d'une représentation irréductible.

**Théorème 3** *Les caractères irréductibles de* $G$ *sont*

- *de norme 1 :* $(\chi, \chi) = 1$,
- *orthogonaux :* $(\chi_1, \chi_2) = 0$ *si* $V_1 \not\simeq V_2$.

**Démonstration :** On reprend les notations du corollaire 2. On a :

$$\begin{aligned} (\chi_1, \chi_2) = \langle \chi_1, \chi_2\rangle &= \frac{1}{|G|}\sum_{t\in G}\chi_1(t)\chi_2(t^{-1}) \\ &= \frac{1}{|G|}\sum_{t\in G, i_1, i_2} r_{i_1i_1}(t)r_{i_1i_2}(t^{-1}) = \sum_{i_1,i_2}\langle r_{i_1i_1}, r_{i_2i_2}\rangle. \end{aligned}$$

Si $V_1 \not\simeq V_2$, ceci vaut 0. Si $V_1 = V_2 = V$, ceci vaut $\sum_{i_1,i_2}\delta_{i_1,i_2} = 1$. □

**Corollaire 3** *Le nombre de caractères irréductibles est fini, inférieur ou égal au nombre de classes de conjugaison de* $G$.

**Démonstration :** En effet, les caractères irréductibles distincts sont orthogonaux dont linéairement indépendants dans l'espace des fonctions centrales, qui est de dimension égale au nombre de classes de conjugaison de $G$. □

**Notation 1** *Dans toute la suite, on note* $\chi_1, \ldots, \chi_h$ *les différents caractères irréductibles de* $G$. *Pour chaque* $i$, *on choisit une représentation* $W_i$ *de caractère* $\chi_i$ *et on note* $n_i = \dim(W_i)$.

**Théorème 4** *Soit* $V$ *une représentation de* $G$, *de caractère* $\varphi$, *et soit* $V = w_1 \oplus \cdots \oplus W_k$ *une décomposition en somme directe de représentations irréductibles. Si* $W$ *est une décomposition irréductible de* $G$ *de caractère* $\chi$, *le nombre des* $W_i$ *isomorphes à* $W$ *est égal au produit scalaire* $(\varphi, \chi) = \langle\varphi, \chi\rangle$. *En particulier, il est indépendant de la décomposition choisie ; on l'appelle la* multiplicité de $W$ dans $V$.

**Démonstration :** Soit $\chi_i$ le caractère de $W_i$. D'après la proposition 2, on a $\varphi = \chi_1 + \cdots + \chi_k$. Ainsi $(\varphi, \chi) = (\chi_1, \chi) + \cdots + (\chi_k, \chi)$. D'après le théorème précédent, le $i$-ième terme de cette somme vaut 0 si $W_i \not\simeq W$ et 1 si $W_i \simeq W$. Le résultat en découle. □

**Corollaire 4** *Deux représentations de même caractère sont isomorphes.*

**Démonstration :** En effet, d'après le résultat précédent, elles contiennent le même nombre de fois toute représentation irréductible $W$ donnée. □

**Théorème 5** *Si $\varphi$ est le caractère d'une représentation, alors $(\varphi, \varphi)$ est un entier positif, et $(\varphi, \varphi) = 1$ si et seulement si ce caractère est irréductible.*

Ceci donne un critère pratique d'irréductibilité.

**Démonstration :** Soit $V$ une représentation dont $\varphi$ est le caractère, et soit $m_i = (\varphi, \chi_i)$ le nombre de fois (multiplicité) que $W_i$ apparaît dans $V$. La représentation $V$ est isomorphe à la somme directe des $m_i W_i$, où l'on note $mW$ la somme directe de $m$ fois une représentation $W$. D'après les relations d'orthogonalité des caractères, on a

$$(\varphi, \varphi) = \left( \sum_i m_i \chi_i, \sum_j m_j \chi_j \right) = \sum_i (m_i)^2.$$

Par ailleurs, ce nombre vaut 1 ssi l'un des $m_i$ est égal à 1 et les autres sont nuls, ssi $V$ est isomorphe à l'une des $W_i$. □

## 7 Décomposition de la représentation régulière

Rappelons que la représentation régulière est la représentation $\rho : G \to \mathrm{GL}(R)$ dans l'espace vectoriel $R = \mathrm{Vect}(e_t, t \in G)$ avec $\rho_s(e_t) = e_{st}$.

**Proposition 3** *Le caractère $r = r_G$ de la représentation régulière est donné par*

$$r(1) = |G|$$
$$r(s) = 0 \quad \text{si } s \neq 1.$$

**Démonstration :** Si $s = 1$, on a $\rho_s = \mathrm{Id}$ donc $r(1) = \dim(R) = |G|$. Si $s \neq 1$, on a $st \neq t$ pour tout $t$, donc les termes diagonaux de la matrice de $\rho_s$ dans la base $(e_t)$ sont tous nuls et $r(s) = \mathrm{tr}(\rho_s) = 0$. □

**Corollaire 5** *Avec les notations de 1 pour les caractères irréductibles distincts de $G$, on a :*

(1) *la multiplicité de $W_i$ dans la représentation régulière est égale à $n_i$.*

(2) *les dimensions $n_i$ vérifient $(n_1)^2 + \cdots + (n_h)^2 = |G|$.*

(3) *pour tout $s \in G$ distinct de 1, on a $n_1\chi_1(s) + \cdots + n_h\chi_h(s) = 0$.*

**Démonstration :** D'après le théorème 4, la multiplicité de $W_i$ dans $R$ est égale à

$$\langle r, \chi_i \rangle = \frac{1}{|G|} \sum_{t \in G} r(t^{-1})\chi_i(t) = \chi_i(1) = n_i.$$

On en déduit que $r = n_1\chi_1 + \cdots + n_h\chi_h$. Compte tenu de la proposition 3, en évaluant en $s = 1$, on trouve $|G| = (n_1)^2 + \cdots + (n_h)^2$. En évaluant en $s \neq 1$, on trouve $n_1\chi_1(s) + \cdots + n_h\chi_h(s) = 0$. □

## 8 Nombre des représentations irréductibles

**Proposition 4** *Soit $f$ une fonction centrale sur $G$, et soit $\rho : G \to \mathrm{GL}(V)$ une représentation de $G$. Soit l'endomorphisme $\rho_f = \sum_{t \in G} f(t)t$ où $t = t_V$ désigne l'image de $t$ dans $\mathrm{GL}(V)$. Si $V$ est irréductible, de dimension $n$ et de caractère $\chi$, l'endomorphisme $\rho_f$ est une homothétie de rapport*

$$\lambda = \frac{|G|}{n}(f, \chi^*).$$

**Démonstration :** Pour tout $s \in G$, on a

$$s\rho_f s^{-1} = \sum_{t\in G} f(t)sts^{-1} = \sum_{u\in G} f(s^{-1}us)u = \sum_{u\in G} f(u)u = \rho_f$$

donc $\rho_f$ est un morphisme de représentations. D'après le lemme de Schur, $\rho_f$ est une homothétie. Soit $\lambda$ son rapport, de sorte que $\lambda \,\mathrm{Id} = \rho_f = \sum_{t\in G} f(t)t$. En prenant les images par $\chi$ on trouve

$$\lambda n = \chi(\lambda \,\mathrm{Id}) = \sum_{t\in G} f(t)\chi(t) = |G|\,(f, \chi^*),$$

comme annoncé. □

**Théorème 6** *Les caractères irréductibles $\chi_1, \ldots, \chi_h$ forment une base de l'espace des fonctions centrales sur $G$.*

**Démonstration :** On sait déjà par le théorème 3 que les $\chi_i$ sont linéairement indépendants. Soit $H$ l'espace qu'ils engendrent. Il suffit de montrer que $H^\perp = 0$, où l'orthogonal s'entend pour le produit scalaire hermitien $(-,-)$. Soit $\varphi$ une fonction centrale orthogonale aux $\chi_i$, et $f = \varphi^*$ qui est orthogonale aux $\chi_i^*$. Pour une représentation $\rho$ variable, posons $\rho_f = \sum_{t\in G} f(t)t$. La proposition 4 montre que $\rho_f = 0$ lorsque $\rho$ est irréductible. Comme toute représentation est somme directe d'irréductibles, on conclut qu'on a toujours $\rho_f = 0$. En particulier, si $\rho$ est la représentation régulière, on trouve

$$0 = \rho_f(e_1) = \sum_{t\in G} f(t)t(e_1) = \sum_{t\in G} f(t)e_t$$

donc $f(t) = 0$ pour tout $t \in G$. Ainsi $f = 0$, donc $\varphi = 0$, cqfd. □

**Corollaire 6** *Le nombre des classes d'isomorphisme de représentations irréductibles de $G$ est égal au nombre de classes de conjugaison de $G$.* □

Pour finir, nous allons donner de toute représentation $V$ une décomposition moins fine qu'une décomposition en représentations irréductibles, mais qui a l'avantage d'être unique ; on l'appelle la *décomposition canonique*. Pour l'obtenir, on part d'une décomposition en irréductibles $V = U_1 \oplus \cdots \oplus U_m$ et pour chaque $i = 1, \ldots, h$ on appelle $V_i$ la somme directe des $U_j$ qui sont isomorphes à $W_i$.

**Théorème 7** *La décomposition $V = V_1 \oplus \cdots \oplus V_h$ ne dépend pas de la décomposition initiale de $V$ en somme de représentations irréductibles. Le projecteur sur $V_i$ associé à cette décomposition est*

$$p_i = \frac{n_i}{|G|} \sum_{t\in G} \chi_i(t)^* t.$$

**Démonstration :** Il suffit de démontrer la deuxième assertion. Posons $q_i = \frac{n_i}{|G|}\sum_{t\in G} \chi_i(t)^* t$. La proposition 4 montre que si $W \subset V$ est une sous-représentation irréductible de caractère $\chi$ et dimension $n$, la restriction $q_i|_W$ est une homothétie de rapport $\frac{n_i}{n}(\chi_i, \chi)$. C'est donc 0 si $\chi \neq \chi_i$ et 1 si $\chi_i = \chi$. En conséquence $q_i|_{V_j}$ est égale à 0 si $j \neq i$ et à $\mathrm{Id}_{V_i}$ si $j = i$, donc $q_i$ est bien le projecteur $p_i$. □

# Décomposition de Bruhat

Soient $k$ un corps, $n \geqslant 1$ un entier, et $\mathrm{GL}_n(k)$ le groupe linéaire. Le théorème de décomposition dont il est question s'énonce ainsi : pour toute matrice carrée $A \in \mathrm{GL}_n(k)$, il existe une permutation $\sigma$, une matrice unipotente supérieure $U$ et une matrice triangulaire supérieure $T$ telles que $A = UM_\sigma T$ (où $M_\sigma$ est la matrice de permutation associée à $\sigma$), et de plus la permutation $\sigma$ est unique. Rappelons qu'une matrice unipotente supérieure est une matrice triangulaire supérieure dont les coefficients diagonaux sont tous égaux à 1.

Nous allons établir ce résultat comme conséquence d'un énoncé plus géométrique d'algèbre linéaire. Si l'on omet tous les commentaires vaseux, la preuve est assez concise. Notons $E$ un espace vectoriel de dimension $n$ sur $k$. On appelle *drapeau (complet)* de $E$ une suite strictement croissante de sous-espaces vectoriels $F_0 \subsetneq \cdots \subsetneq F_n$. On notera simplement $F$ un tel drapeau ; observez que $\dim(F_i) = i$ pour tout $i$.

**Théorème.** *Soient $F$ et $G$ deux drapeaux complets de $E$. Alors il existe une permutation $\sigma \in \mathfrak{S}_n$ et une base ordonnée $(e_1, \ldots, e_n)$ de $E$ telles que $e_i \in F_i \cap G_{\sigma(i)}$ pour tout $i \in \{1, \ldots, n\}$. De plus, la permutation $\sigma_{F,G} := \sigma$ est unique.*

**Commentaires.** (1) Cet énoncé signifie simplement qu'il existe une base de $E$ dont chaque vecteur appartient à l'un des $F_i$ et à l'un des $G_j$.

(2) L'unicité de $\sigma = \sigma_{F,G}$ implique que $\sigma_{G,F} = {\sigma_{F,G}}^{-1}$, car si l'on pose $\tau = \sigma^{-1}$ et $e'_j = e_{\tau(j)}$ on a $e'_j \in G_j \cap F_{\tau(j)}$ d'où $\tau = \sigma_{G,F}$. Cependant, la preuve procèdera différemment, en montrant *d'abord* que $\sigma_{G,F} = {\sigma_{F,G}}^{-1}$ *puis* que $\sigma_{F,G}$ est unique.

(3) Il y a dans la littérature mathématique une imprécision de langage extrêmement (trop) présente, qui consiste à confondre *base* et *base ordonnée*. Rappelons qu'une base est par définition une famille libre et génératrice, et qu'elle n'est pas ordonnée a priori. Par exemple, contrairement à ce qu'on lit d'habitude, c'est pour une base *ordonnée* $\mathscr{B}$ qu'il existe une unique forme $n$-linéaire alternée $\varphi = \det_{\mathscr{B}}$ telle que $\varphi(\mathscr{B}) = 1$. De même, c'est avec l'ensemble des bases *ordonnées* de $E$ que le groupe linéaire $\mathrm{GL}(E)$ est en bijection ; notez que le stabilisateur d'une base *non ordonnée* $\mathscr{B}$ pour l'action du groupe $\mathrm{GL}(E)$ s'identifie au groupe de permutations $\mathfrak{S}_{\mathscr{B}}$. Nous utiliserons ce fait ci-dessous (preuve du corollaire 1). Un dernier exemple est donné par la matrice d'une application linéaire dans des bases fixées : c'est un tableau de scalaires dont les lignes et les colonnes *ne* sont *pas* ordonnées tant que les bases ne le sont pas, et cela n'a pas de sens de parler par exemple de sa *première ligne*. Ce sont les contraintes de notre mode de représentation en dimension 2 sur une feuille ou au tableau qui font qu'on est obligé d'ordonner les lignes et les colonnes pour écrire une matrice. Pour dissiper cette imprécision, on prend soin de noter $(e_1, \ldots, e_n)$ une base ordonnée et $\{e_1, \ldots, e_n\}$ une base non ordonnée.

**Preuve.** Fixons un entier $i \geqslant 1$. On observe que si pour un certain $j = j_0 \geqslant 1$ l'inclusion $F_{i-1} + G_j \subset F_i + G_j$ est une égalité, alors elle l'est encore pour tout $j \geqslant j_0$ puisque $F_{i-1} + G_j = F_{i-1} + G_{j_0} + G_j = F_i + G_{j_0} + G_j = F_i + G_j$. Il s'ensuit que lorsque $j$ croît de 0 à $n$, cette inclusion qui est stricte pour $j = 0$ devient une égalité pour un certain $j \geqslant 1$, puis le reste. Notons $\sigma_{F,G}(i) = j$ cet entier minimal, qui est donc caractérisé par les relations $F_{i-1} + G_{j-1} \subsetneq F_i + G_{j-1}$ et $F_{i-1} + G_j = F_i + G_j$.

Pour montrer que $\sigma_{F,G}$ est une permutation, il suffit de montrer que $\sigma_{G,F} \circ \sigma_{F,G} = \mathrm{id}$. Pour cela, notons $j = \sigma_{F,G}(i)$ et montrons qu'alors $i = \sigma_{G,F}(j)$, c'est-à-dire $F_{i-1} + G_{j-1} \subsetneq F_{i-1} + G_j$ et $F_i + G_{j-1} = F_i + G_j$. Or supposant que $F_{i-1} + G_{j-1} = F_{i-1} + G_j$, on déduit $F_i + G_{j-1} = F_i + G_j = F_{i-1} + G_j = F_{i-1} + G_{j-1}$ ce qui est une contradiction ; ceci établit l'inclusion stricte désirée. Par ailleurs, on a :

$$\begin{aligned} \dim(F_i + G_{j-1}) &= \dim(F_{i-1} + G_{j-1}) + 1 \\ &= \dim(F_{i-1} + G_j) = \dim(F_i + G_j), \end{aligned}$$

ce qui montre que l'inclusion $F_i + G_{j-1} \subset F_i + G_j$ est une égalité.

Pour $i$ et $j$ quelconques, utilisant la formule reliant la somme de deux sous-espaces à leur intersection, on voit que $F_{i-1}+G_j = F_i+G_j$ équivaut à $\dim(F_i \cap G_j) = \dim(F_{i-1} \cap G_j) + 1$. Si $j = \sigma(i)$ avec $\sigma = \sigma_{F,G}$, cette égalité a lieu donc on peut choisir $e_i \in F_i \cap G_{\sigma(i)} \backslash F_{i-1} \cap G_{\sigma(i)}$. En particulier, on a $e_i \in F_i \backslash F_{i-1}$ ce qui montre que $(e_1, \ldots, e_n)$ est une base (ordonnée) de $E$.

Montrons enfin que $\sigma_{F,G}$ est unique. Supposons qu'il existe une permutation $\varphi \in \mathfrak{S}_n$ et des vecteurs $e_i \in F_i \cap G_{\varphi(i)}$ formant une base de $E$. Alors $F_i = \mathrm{Vect}(e_1, \ldots, e_i)$ donc $e_i \notin F_{i-1}$. Posant $j = \varphi(i)$, on a donc $\dim(F_i \cap G_j) = \dim(F_{i-1} \cap G_j) + 1$ puis $F_{i-1} + G_j = F_i + G_j$ d'après l'argument sur la somme et l'intersection utilisé quelques lignes plus haut. Il s'ensuit que $\varphi(i)$ est inférieur ou égal à $\sigma_{F,G}(i)$ tel que celui-ci a été défini en début de preuve. Comme ceci vaut pour tout $i$ et que $\sigma_{F,G}$ est une bijection, ceci implique que $\varphi = \sigma_{F,G}$. ▫

**Corollaire 1.** *Soit $X$ l'ensemble des paires de drapeaux de $E$, muni de l'action naturelle de $G = \mathrm{GL}(E)$ définie par $u.(F,G) = (u(F), u(G))$. Alors l'application $(F,G) \mapsto \sigma_{F,G}$ passe au quotient en une bijection canonique entre l'ensemble d'orbites $X/G$ et le groupe symétrique $\mathfrak{S}_n$.*

Cet énoncé dit qu'après transport par un automorphisme de $E$, le drapeau $G$ est obtenu par (unique) permutation des vecteurs d'une base de $F$. Autrement dit, à automorphisme linéaire et à permutation près, il n'y a qu'un drapeau dans $E$. Si $E = k^n$, on peut choisir par exemple le drapeau canonique $F$ tel que $F_i$ est engendré par les $i$ premiers vecteurs de la base canonique.

**Preuve.** Soient $\sigma = \sigma_{F,G}$ et $(e_i)$ une base satisfaisant les conclusions du théorème. On a alors $G_j = \mathrm{Vect}(e_{\sigma^{-1}(1)}, \ldots, e_{\sigma^{-1}(j)})$ puisque pour chaque $k \leqslant j$ le vecteur $e_{\sigma^{-1}(k)}$ appartient à $G_k$ qui est inclus dans $G_j$, et ces vecteurs sont en nombre $j = \dim(G_j)$. Ainsi $G$ est entièrement déterminé par la base ordonnée $(e_i)$ et la permutation $\sigma$. Si l'on fixe une base ordonnée $\mathscr{B}$ de $E$ et qu'on note $u \in \mathrm{GL}(E)$ l'unique automorphisme linéaire qui envoie $(e_i)$ sur $\mathscr{B}$, on voit que la paire $(u(F), u(G))$ est entièrement déterminée par $\sigma$. Notez qu'on a utilisé la base $\mathscr{B}$ non pas pour définir l'application $X/G \to \mathfrak{S}_n$, qui est donc bien canonique, mais seulement pour vérifier que c'est une bijection. ▫

Revenons aux notations de départ pour en déduire la décomposition de Bruhat sous sa forme matricielle. On observera que la preuve du corollaire 2 utilise seulement le théorème, pas le corollaire 1. On rappelle que $M_\sigma$ désigne la matrice de permutation associée à $\sigma$.

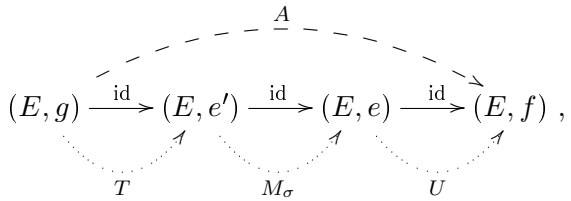
**Corollaire 2.** *Pour tout $A \in \mathrm{GL}_n(k)$, il existe une permutation $\sigma$, une matrice $U$ unipotente supérieure et une matrice $T$ triangulaire supérieure, telles que $A = UM_\sigma T$. De plus la permutation $\sigma$ est unique.*

**Preuve.** Notons $f = (f_1, \ldots, f_n)$ la base (ordonnée) canonique de $E = k^n$ et $g = (g_1, \ldots, g_n)$ la base (ordonnée) formée par les vecteurs colonnes de la matrice $A$. Ainsi $A$ n'est autre que la matrice de l'identité exprimée dans les bases $g$ à la source et $f$ au but, c'est-à-dire, en symboles $A = \mathrm{Mat}_{g,f}(\mathrm{id})$.

Soient $F, G$ les deux drapeaux de $E$ définis par $F_i = \mathrm{Vect}(f_1, \ldots, f_i)$ et $G_j = \mathrm{Vect}(g_1, \ldots, g_j)$. Notons $\sigma = \sigma_{F,G}$ et $e = (e_1, \ldots, e_n)$ la base ordonnée fournie par le théorème, qui vérifie $e_i \in F_i \cap G_{\sigma(i)}$. Puisque $e_i \in F_i \backslash F_{i-1}$, lorsqu'on exprime $e_i$ sur la base $f$ sa composante sur $f_i$ est non nulle ; quitte à normaliser $e_i$ on peut donc supposer que cette composante est 1. Ceci signifie que la matrice de passage $U = \mathrm{Mat}_{e,f}(\mathrm{id})$ est unipotente supérieure.

Notons maintenant $\tau = \sigma^{-1}$ et $e'_i = e_{\tau(i)}$ ; clairement la matrice de passage $\mathrm{Mat}_{e',e}(\mathrm{id})$ est la matrice de permutation $M_\sigma$. Les vecteurs $e'_i$ vérifient $e'_i \in F_{\tau(i)} \cap G_i$ et le raisonnement fait précédemment montre que la matrice de passage $T = \mathrm{Mat}_{g,e'}(\mathrm{id})$ est triangulaire supérieure (mais on ne peut plus normaliser de manière à ce qu'elle soit unipotente, car cela changerait la normalisation de $e_i$). On conclut en

disant que la matrice de l'application composée




écrite dans les bases indiquées, est le produit $A = UM_\sigma T$. □

## EXERCICES

# Anneaux factoriels et non factoriels

Vous trouverez des détails sur certains des points qui suivent (mais pas tous) dans le Cours d'Algèbre de Daniel Perrin.

## 1 Quelques rappels sur les anneaux factoriels

**Cadre général.** L'arithmétique, c'est l'étude des relations de divisibilité. Le cadre habituel pour cette étude est celui des anneaux commutatifs, unitaires. On s'intéresse la plupart du temps à des anneaux intègres, à la fois car la plupart des exemples historiques qui ont motivé les mathématiciens (dont les anneaux d'entiers dans les corps de nombres) sont des anneaux intègres, et aussi car l'hypothèse d'intégrité est très commode, voire souvent nécessaire pour mener à bien les raisonnements que l'on a en vue. En résumé, disons qu'on considère, dans cette feuille au moins, la famille $\mathcal{F}$ des anneaux commutatifs unitaires et intègres. Il faut tout de même noter qu'on est souvent amené à réduire modulo un idéal et que dans ce cas, on tombe sur un anneau qui possède toutes les caractéristiques précédentes, sauf peut-être l'intégrité.

**Anneaux factoriels.** Dans la famille $\mathcal{F}$, les anneaux les plus sympathiques pour faire de l'arithmétique sont les anneaux *factoriels*, qui sont par définition les anneaux $A$ tels que :

(I) $A$ est intègre,
(E) tout élément non nul s'écrit comme un produit $a = up_1 \dots p_r$ avec $u$ inversible et les $p_i$ irréductibles (non distincts)
(U) une décomposition comme dans (E) est unique, à permutation près des facteurs et aux éléments inversibles près.

On notera qu'un anneau factoriel n'est pas nécessairement noethérien (par exemple, l'anneau $\mathbb{Q}[X_1, X_2, X_3, \dots]$ de polynômes en une infinité de variables est factoriel non noethérien). Néanmoins, les anneaux que l'on étudie en arithmétique sont presque tout le temps noethériens (extensions de $\mathbb{Z}$ ou d'un corps engendrées par un nombre fini d'éléments). Or, un anneau noethérien vérifie toujours la condition (E). Finalement, retenons que la propriété véritablement distinctive des anneaux factoriels est la validité de la condition (U).

**Pgcd et ppcm.** Dans cette note constituée d'exercices, on s'intéresse en particulier au comportement des pgcd et ppcm, et notamment à leur (non-)existence éventuelle. Rappelons que dans un anneau intègre $A$, l'ensemble $A\backslash\{0\}$ des éléments non nuls est un monoïde multiplicatif, l'ensemble $A^*$ des inversibles est un sous-monoïde qui est un groupe et qui donne naissance à la relation d'*association* : $a$ et $b$ sont associés s'il existe $u \in A^*$ tel que $a = ub$. On note parfois $a \sim b$. La relation de divisibilité :

$$a \geqslant b \quad \text{ssi} \quad a \,|\, b \,,$$

est une relation d'ordre sur $A\backslash\{0\}$, ou sur son quotient $(A\backslash\{0\})/\sim$, compatible à la structure de monoïde. Si deux éléments $a$ et $b$ de $A\backslash\{0\}$ possèdent un majorant, on l'appelle « un » pgcd (dans $A\backslash\{0\}$) ou « le » pgcd (dans $(A\backslash\{0\})/\sim$). S'ils possèdent un minorant, on l'appelle « un » ou « le » ppcm. (Attention au fait que si l'on choisit la convention inverse pour la relation d'ordre, c'est-à-dire si l'on écrit que $a$ divise $b$ se note $a \leqslant b$ au lieu de $a \geqslant b$, alors les notions de sup et d'inf sont renversées.)

## 2 Exercices

Chaque exercice s'autorise à utiliser le résultat de ceux qui le précèdent, et il est donc nécessaire de les faire dans l'ordre.

**Exercice 1** *Soit $A$ un anneau intègre et $a, b, x$ non nuls dans $A$. Montrez que si $xa$ et $xb$ possèdent un pgcd, alors $a$ et $b$ possèdent un pgcd et on a la formule* $\operatorname{pgcd}(xa, xb) = x \operatorname{pgcd}(a, b)$.

*Remarque : il n'est pas toujours vrai que si $a$ et $b$ possèdent un pgcd, alors $xa$ et $xb$ en possèdent un. On donnera un contre-exemple dans l'exercice 4.*

**Exercice 2** *Soit $A$ un anneau intègre noethérien. Montrez que les conditions suivantes sont équivalentes :*

*(i) $A$ est factoriel,*

*(ii) tout couple d'éléments de $A$ possède un pgcd.*

**Exercice 3** *Soient $A$ un anneau intègre et $a, b$ deux éléments de $A$. Montrez que si $a$ et $b$ possèdent un ppcm $m$, alors ils possèdent un pgcd $d$ et on a la formule $ab = md$.*

*Remarque : il n'est pas toujours vrai que si $a$ et $b$ possèdent un pgcd, ils possèdent un ppcm. On donnera un contre-exemple dans l'exercice 4.*

Terminons avec un exemple d'anneau (intègre, noethérien) non factoriel. Pour fabriquer cet exemple, on suit une idée simplissime et extrêment importante qui utilise à fond la possibilité donnée par les quotients d'anneaux de polynômes de construire des exemples d'anneaux possédant un certain nombre fixé (fini) d'éléments satisfaisant un certain nombre (fini) de relations données. Précisément, l'idée est que pour trouver un anneau non factoriel, on va chercher un anneau contenant un élément qui possède deux décompositions en irréductibles distinctes de la forme $xy = zt$ où $x, y, z, t$ sont des irréductibles tous distincts. On peut construire un tel anneau de manière « universelle » en considérant l'anneau de l'exercice.

**Exercice 4** *On considère l'anneau $A = k[X, Y, Z, T]/(XY - ZT)$, quotient d'un anneau de polynômes en quatre indéterminées. On note $x, y, z, t$ les images de $X, Y, Z, T$ dans $A$.*

(1) *On considère l'anneau $R = k[z, t] \subset A$ isomorphe à $k[Z, T]$. Montrez que tout élément de $A$ possède une écriture unique $a = a_0 + xa_1(x) + ya_2(y)$ où $a_0 \in R$, $a_1 \in R[x]$ et $a_2 \in R[y]$.*

(2) *Montrez que $A$ est intègre en construisant une injection de $A$ dans $B = k(X)[Z, T]$.*

(3) *Montrez que $z$ et $t$ sont des irréductibles de $A$. Indiquez pourquoi $x$ et $y$ le sont aussi.*

(4) *Justifiez qu'aucun de ces quatre irréductibles n'est multiple d'un autre ; en particulier, ils sont non associés deux à deux.*

(5) *Montrez que $A$ n'est pas factoriel.*

(6) *Montrez que $xz$ et $xy$ n'ont pas de pgcd.*

(7) *Montrez que $x$ et $z$ ont un pgcd mais n'ont pas de ppcm (utiliser l'exercice 3).*

# 3 Corrigés

**Corrigé exercice 1.** Supposons que $xa$ et $xb$ possèdent un pgcd et notons-le $e$. Comme $x$ divise $xa$ et $xb$, on a $x$ divise $e$. Soit $d$ tel que $e = xd$. Il suffit de montrer que $d$ est un pgcd pour $a$ et $b$ pour résoudre l'exercice. D'abord, du fait que $e = xd$ divise $xa$ et $xb$, on tire que $d$ divise $a$ et $b$. Ensuite, si $d'$ divise $a$ et $b$, alors $xd'$ divise $xa$ et $xb$, donc $xd'$ divise $\mathrm{pgcd}(xa, xb) = e = xd$. Ainsi $d'$ divise $d$, ce qui conclut la démonstration. □

**Corrigé exercice 2.** On sait que (i) implique (ii), et il faut donc prouver la réciproque. Supposons pour cela que $A$ n'est pas factoriel, c'est-à-dire qu'il ne vérifie pas la condition (U). Alors il existe un élément $a \in A$ non nul qui possède deux décompositions en irréductibles distinctes : $a = up_1^{\alpha_1} \dots p_r^{\alpha_r} = vq_1^{\beta_1} \dots q_s^{\beta_s}$ (où l'on suppose que les $\alpha_i$ et les $\beta_j$ sont $> 0$).

Notons $P = \{p_1, \dots, p_r\}$ et $Q = \{q_1, \dots, q_s\}$. Pour tout $a \in A$, notons $\mu(a)$ la plus petite des valeurs $\alpha_1 + \cdots + \alpha_r$, prise parmi toutes les décompositions en irréductibles possibles $a = up_1^{\alpha_1} \dots p_r^{\alpha_r}$. Si on choisit $a$ tel que $\mu(a)$ est minimal, on aura $P \cap Q = \varnothing$. En effet, sinon $P$ et $Q$ ont un élément commun, que l'on peut supposer être $p_1 = q_1$ quitte à renuméroter. Alors, en divisant par $p_1 = q_1$ de part et d'autre, on obtient un élément $a'$ qui possède deux décompositions en irréductibles distinctes :

$$a' = up_1^{\alpha_1 - 1} p_2^{\alpha_2} \dots p_r^{\alpha_r} = vq_1^{\beta_1 - 1} \dots q_s^{\beta_s}$$

et $\mu(a') < \mu(a)$, en contradiction avec le choix de $a$.

On a donc un élément $a$ décomposé de deux manières, comme ci-dessus, avec $P$ et $Q$ disjoints. Soit $b = p_1 q_1$. Si $a$ et $b$ possèdent un pgcd, alors en utilisant l'exercice 1 on trouve

$$\mathrm{pgcd}(a,b) = \mathrm{pgcd}(up_1^{\alpha_1}\dots p_r^{\alpha_r}, p_1 q_1) = p_1 \,\mathrm{pgcd}(up_1^{\alpha_1 - 1}\dots p_r^{\alpha_r}, q_1) = p_1$$

puisque $q_1$ ne divise aucun des $p_i$. Par ailleurs

$$\mathrm{pgcd}(a,b) = \mathrm{pgcd}(vq_1^{\beta_1}\dots q_s^{\beta_s}, p_1 q_1) = q_1 \,\mathrm{pgcd}(vq_1^{\beta_1 - 1}\dots q_s^{\beta_s}, p_1) = q_1 \,.$$

On obtient $p_1 = q_1$ ce qui n'est pas possible. Donc $a$ et $b$ ne possèdent pas de pgcd, ce que l'on voulait démontrer. □

**Corrigé exercice 3.** Notons $m$ le ppcm de $a$ et $b$. Comme $ab$ est un multiple commun de $a$ et $b$, c'est un multiple de $m$ : il existe $d$ tel que $ab = md$. Montrons que $d$ est un pgcd pour $a$ et $b$.

Comme $m$ est multiple de $a$ et $b$, il existe $u, v$ tels que $m = ua = vb$. On en déduit que $ab = uad$ donc $b = ud$, et $ab = vbd$ donc $a = vd$. Ainsi $d$ est diviseur commun de $a$ et $b$.

Soit $e$ un diviseur quelconque de $a$ et $b$. Alors il existe $\alpha, \beta$ tels que $a = e\alpha$ et $b = e\beta$, donc $e\alpha\beta$ qui vaut à la fois $a\beta$ et $b\alpha$ est un multiple commun de $a$ et $b$. Ainsi, il existe $f$ tel que $e\alpha\beta = mf$. Alors, $b\alpha = e\alpha\beta = mf = vbf$ donc $\alpha = vf$. On en tire $vd = a = e\alpha = evf$ donc $d = ef$, i.e. $e$ divise $d$ et $d$ est le pgcd de $a$ et $b$. □

Précédons la correction de l'exercice 4 par un commentaire sur les calculs dans les anneaux quotients. Comme toujours avec les quotients (quotient $X/R$ d'un ensemble par une relation d'équivalence, quotient $G/H$ d'un groupe par un sous-groupe distingué ou non, quotient $A/I$ d'un anneau par un idéal), la plupart du temps la seule manière de faire des calculs avec les éléments d'un quotient est d'en choisir des préimages (dans $X$, resp. $G$, resp. $A$) et de raisonner « en haut » mais évidemment modulo ($R$, $H$, $I$). Avec un anneau $A$, une seconde possibilité se présente parfois : il s'agit grosso modo d'utiliser un analogue *ad hoc* de la division euclidienne pour obtenir une écriture *unique* agréable pour les éléments de $A$. L'exemple modèle est celui des quotients d'anneaux de polynôme en une variable : dans $A = k[X]/(P)$ avec $P$ unitaire de degré $n$, par division euclidienne tout élément de $A$ s'écrit de manière unique sous la forme d'un polynôme de degré $\leqslant n-1$ en $x$, l'image de $X$ dans $A$.

**Corrigé exercice 4.** (1) Choisissons un polynôme $P \in k[X,Y,Z,T]$ d'image $a \in A$, et raisonnons modulo $(XY - ZT)$. On écrit $P$ comme polynôme en $X$ et $Y$ à coefficients dans $k[Z,T]$. Il est donc somme de monômes $X^i Y^j$, et à chaque fois que $i \geqslant 1$ et $j \geqslant 1$, on peut mettre en facteur $XY$ dans ce monôme et le remplacer par $ZT$, vu comme une « constante » de notre sous-anneau $k[Z,T]$. En itérant ce procédé, on arrive à une écriture dans laquelle n'apparait plus de produit $XY$. C'est la forme demandée par l'énoncé, et l'unicité est claire (je vous la laisse).

(2) Je commence par dire que souvent, on écrit indifféremment $x$ ou $X$ lorsque l'anneau dans lequel on le considère est clair. Je ferai ce genre d'abus ci-dessous, en particulier, je note plus volontiers $B = k(x)[z,t]$. Soit le morphisme $f' : k[x,y,z,t] \to B$ qui envoie $x, z, t$ sur eux-mêmes et $y$ sur $zy/x$. Le polynôme $xy - zt$ (ici, il faudrait vraiment écrire $XY - ZT$) est envoyé sur 0 donc $f'$ se factorise en un morphisme $f : A \to B$. Montrons que $f$ est injectif : si $f(a) = 0$ avec $a = a_0 + xa_1(x) + ya_2(y)$ comme dans (1), alors $a_0 + xa_1(x) + \frac{zt}{x}a_2(\frac{zt}{x}) = 0$. Regardons cela comme une égalité dans $R(x)$, qui possède une base formée des puissances positives ou négatives de $x$. On voit que $a_0$ est constant, $xa_1(x)$ ne possède que des monômes de degré (en $x$) strictement positif et $\frac{zt}{x}a_2(\frac{zt}{x})$ ne possède que des monômes de degré (en $x$) strictement négatif. Donc $a_0 = a_1 = a_2 = 0$ puis $a = 0$.

(3) On suppose que l'on a une écriture $z = ab$ dans $A$, avec $a = a_0 + xa_1(x) + ya_2(y)$ et $b = b_0 + xb_1(x) + yb_2(y)$. On utilise le morphisme injectif de la question (2), que l'on voit comme une injection, pour plonger cette égalité dans $B = k(x)[z,t]$, qui est un anneau de polynômes en deux variables sur un corps. On sait que $z$ est un irréductible de $B$, donc $a$ est inversible dans $B$ i.e. $a \in k(x)$, et $b = a^{-1}z$ (ou l'inverse). Ceci implique que ni $a_0 + xa_1(x) + \frac{zt}{x}a_2(\frac{zt}{x})$ ni $b_0 + xb_1(x) + \frac{zt}{x}b_2(\frac{zt}{x})$ ne possèdent de monôme contenant $t$, donc :

$$a_0 \in k\ ,\ a_1 \in k[x]\ ,\ a_2 = 0 \qquad \text{et} \qquad b_0 \in k.z\ ,\ b_1 \in k[x].z\ ,\ b_2 = 0\ .$$

On voit ainsi que $a$ est en fait un polynôme en $x$ et $b$ est produit de $z$ par un polynôme en $x$. Pour avoir $z = ab$ on doit avoir $a \in k^*$ et $b = a^{-1}z$. Ceci montre que $z$ est irréductible dans $A$. Le raisonnement pour $t$ est symétrique. Pour montrer que $x$ et $y$ sont irréductibles, on fait pareil en utilisant des écritures $a_0 + za_1(z) + ta_2(t)$ avec les $a_i$ à coefficients dans $k[x, y]$.

(4) Montrons par exemple que $x$ n'est multiple ni de $y$ ni de $z$, les autres cas sont semblables. Le fait que $x$ n'est pas multiple de $y$ est équivalent au fait que la classe de $x$ est non nulle dans $A/(y)$. Or on a $A/(y) \simeq k[x, y, z, t]/(xy - zt, y) \simeq k[x, z, t]/(zt)$ et il est clair que la classe de $x$ est non nulle. Le fait que $x$ n'est pas multiple de $z$ est équivalent au fait que la classe de $x$ est non nulle dans $A/(z)$, or $A/(z) \simeq k[x, y, z, t]/(xy - zt, z) \simeq k[x, y, t]/(xy)$ et il est clair que la classe de $x$ est non nulle.

(5) L'élément $xy$ possède deux écritures distinctes comme produit d'irréductibles : $xy = zt$.

(6) On commence par une petite observation : deux irréductibles distincts (c'est-à-dire, non associés) $p$ et $q$ ont 1 pour pgcd, car sinon ce pgcd est associé à $p$ et à $q$, ce qui n'est pas possible. D'après l'exercice 1, si $xz$ et $xy$ ont un pgcd alors $z$ et $y$ en ont un et on a $\mathrm{pgcd}(xz, xy) = x\,\mathrm{pgcd}(z, y) = x$. Par ailleurs, $xy = zt$, de sorte que par le même raisonnement $\mathrm{pgcd}(xz, xy) = \mathrm{pgcd}(xz, zt) = z\,\mathrm{pgcd}(x, t) = z$. Comme $z$ et $t$ ne sont pas associés, c'est une contradiction, donc $xz$ et $xy$ n'ont pas de pgcd.

(7) Les éléments $x$ et $z$ sont des irréductibles distincts donc ils ont un pgcd qui est $d = 1$. S'ils ont un ppcm $m$, d'après le résultat de l'exercice 3 on a la relation $xz = md = m$. Or $xy = zt$ est manifestement un multiple de $x$ et de $z$, il doit donc être multiple de $m = xz$. Or $xy$ multiple de $xz$ implique $y$ multiple de $z$, ce qui n'est pas le cas d'après la question (4). Donc $x$ et $z$ n'ont pas de ppcm. □

# La droite projective

La droite projective complexe $\mathbb{P}^1(\mathbb{C})$ est un espace topologique qui possède de nombreuses structures supplémentaires. Ceci en fait un objet mathématique extrêmement important, situé au croisement de nombreuses branches des mathématiques, dédiées à l'étude de ces différentes structures. Nous donnerons quatre constructions et vérifierons, sous forme d'exercices, qu'elles donnent bien naissance au même objet, à homéomorphisme près.

a) La droite projective complexe $\mathbb{P}^1(\mathbb{C})$ est l'ensemble des droites vectorielles (i.e. des droites passant par l'origine) du $\mathbb{C}$-espace vectoriel $\mathbb{C}^2$. Il y a une application surjective $\pi : \mathbb{C}^2\backslash\{0\} \to \mathbb{P}^1(\mathbb{C})$ qui envoie un vecteur $x \neq 0$ sur la droite qu'il engendre. L'ensemble $\mathbb{C}^2\backslash\{0\}$ est un ouvert de $\mathbb{C}^2$, et on munit $\mathbb{P}^1(\mathbb{C})$ de la topologie quotient via $\pi$ (voir exercice ci-dessous sur la topologie quotient).

b) La sphère euclidienne $S^2 \subset \mathbb{R}^3$ est la sphère unité dans l'espace euclidien de dimension 3. Elle est munie de la métrique induite de celle de $\mathbb{R}^3$, ce qui en fait un espace topologique.

c) Le compactifié d'Alexandroff de $\mathbb{C}$ est l'ensemble $\widehat{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$ réunion de $\mathbb{C}$ et d'un point qui n'est pas dans $\mathbb{C}$, qu'on appelle le *point à l'infini*. Sa topologie est celle dont les ouverts sont de deux types : soit les ouverts de $\mathbb{C}$, soit la réunion de $\infty$ et du complémentaire dans $\mathbb{C}$ d'un compact de $\mathbb{C}$ (voir exercice ci-dessous sur le compactifié d'Alexandroff).

d) L'espace $X$ est l'espace topologique obtenu en recollant les deux ensembles $Y = Z = \mathbb{C}$ le long des ouverts $U = Y\backslash\{0\}$ et $V = Z\backslash\{0\}$, au moyen de l'homéomorphisme $\varphi : U \to V$ défini par $\varphi(x) = 1/x$. Précisément, $X$ est le quotient de l'ensemble $Y \amalg Z$ par la relation d'équivalence telle que $y \sim z$ si et seulement si $y \in Y$, $z \in Z$ et $y = \varphi(z)$. Il est muni de la topologie quotient de celle de $Y \amalg Z$.

La présentation de la droite projective donnée dans a) relève de la géométrie projective (géométrie qui étudie les propriétés d'incidence),

de la géométrie algébrique (géométrie qui étudie les variétés dont les fonctions locales sont des polynômes), ou de la géométrie complexe (géométrie qui étudie les variétés dont les fonctions locales sont holomorphes). La présentation donnée dans b) relève de la géométrie riemannienne (géométrie qui étudie les variétés munies d'une métrique). Les deux dernières présentations sont des constructions classiques en topologie. La présentation d) est typique du procédé de construction par *recollement* pour fabriquer des variétés (topologiques, différentiables, holomorphes, algébriques, ou autres).

**Exercice 1** *Soit $X$ un espace topologique, $Y$ un ensemble, et $f : X \to Y$ une application surjective. Montrez que la famille des parties $A \subset Y$ telles que $f^{-1}(A)$ est un ouvert de $X$ définit une topologie sur $Y$. On appelle cette topologie la* topologie quotient de $X$ (via $f$). *Montrez que pour cette topologie, l'application $f : X \to Y$ vérifie la propriété suivante : pour tout espace topologique $Z$, une application $g : Y \to Z$ est continue si et seulement si $g \circ f : X \to Z$ est continue.*

**Corrigé.** Il est facile de montrer qu'on a bien défini une topologie sur $Y$, et je laisse les détails. Maintenant, soit $Z$ un espace topologique et $g : Y \to Z$ une application. Si $g$ est continue, alors $g \circ f$ est continue. C'est la réciproque qui demande une démonstration. Supposons donc que $g \circ f$ est continue, on veut montrer que $g$ l'est. Soit $W \subset Z$ un ouvert, il faut montrer que $V := g^{-1}(W)$ est ouvert. Par définition de la topologie sur $Y$, il est équivalent de montrer que $f^{-1}(V)$ est ouvert, ce qui est vrai puisque $f^{-1}(V) = (g \circ f)^{-1}(W)$ et $g \circ f$ est continue par hypothèse.

**Exercice 2** *Soit $X$ un espace topologique séparé et localement compact. Le* compactifié d'Alexandroff de $X$ *est l'ensemble égal à la réunion de $X$ et d'un point qui n'est pas dans $X$, noté $\infty$ . Montrez que la famille de parties suivantes définit une topologie sur $X$ : les ouverts de $X$ (ce sont les ouverts de $\widehat{X}$ ne contenant pas $\infty$), et la réunion de $\infty$ et du complémentaire d'un compact de $X$ (ce sont les ouverts de $\widehat{X}$ contenant $\infty$). Montrez que $\widehat{X}$ est compact.*

**Corrigé.** Il est facile de montrer qu'on a bien défini une topologie sur $\widehat{X}$, et je laisse les détails. Pour montrer que $\widehat{X}$ est compact, il suffit de montrer qu'il vérifie la propriété de Borel-Lebesgue. Soit donc $\{U_i\}$ un recouvrement de $\widehat{X}$ par des ouverts. Comme il s'agit d'un recouvrement, l'un de ces ouverts, disons $U_{i_0}$, contient $\infty$. Le complémentaire de $U_{i_0}$ est un compact $K$ de $X$, il est donc recouvert par un nombre fini d'ouverts relatifs $U_{i_1} \cap K, \ldots, U_{i_n} \cap K$, où les $U_{i_k}$ sont extraits du recouvrement initial. Donc les $U_{i_k}$ pour $0 \leqslant k \leqslant n$ recouvrent $\widehat{X}$.

**Exercice 3** *(*Lemme : une bijection qui échange des bases d'ouverts est un homéo.*) Soient $X, Y$ des espaces topologiques et $f : X \to Y$ une bijection. Montrez que $f$ est un homéomorphisme si et seulement si c'est une bijection et pour tout $x \in X$, il existe $\{U_i\}_{i \in I}$ une base de voisinages ouverts de $x$ telle que $\{f(U_i)\}$ est une base de voisinages ouverts de $f(x)$.*

**Corrigé.** Seule la partie *si* n'est pas évidente. Nous supposons donc que $f$ est une bijection satisfaisant la propriété indiquée, et nous voulons vérifier que c'est un homéomorphisme. Les hypothèses étant symétriques en $x$ et $y$, il suffit de démontrer que $f$ est continue, ou ouverte, au choix (une application entre espaces topologiques est ouverte, par définition, si l'image d'un ouvert est un ouvert ; donc $f$ est un homéomorphisme ssi $f$ est continue et ouverte). Montrons que $f$ est ouverte. Soit $U$ un ouvert de $X$. Par hypothèse, tout $x \in U$ possède un voisinage ouvert $U_x$ inclus dans $U$, et tel que $f(U_x)$ est ouvert. Il s'ensuit que $f(U) = \cup_{x \in U} f(U_x)$ est ouvert.

Dans les deux exercices qui suivent, on construit des applications $X \xrightarrow{a} \mathbb{P}^1(\mathbb{C}) \xrightarrow{b} S^2 \xrightarrow{\sigma} \widehat{\mathbb{C}}$ et on montre que ce sont des homéomorphismes en utilisant le lemme ci-dessus, c'est-à-dire en montrant que ces applications échangent des bases de voisinages ouverts.

**Exercice 4** *Dans l'espace euclidien $E = \mathbb{R}^3$, on considère la sphère unité $S^2$ et le plan équatorial $\mathcal{P}$. On appelle* projection stéréographique *notée $\sigma : S^2 \backslash \{N\} \to \mathcal{P}$ la projection depuis le pôle nord $N$, définie par $\sigma(M) = (NM) \cap \mathcal{P}$.*

*(1) Faites un dessin.*
*(2) On identifie $E$ à $\mathbb{C}\oplus\mathbb{R}$, de sorte que le plan équatorial s'identifie à $\mathbb{C}$. On note $(z,t)\in\mathbb{C}\oplus\mathbb{R}$ les coordonnées d'un point $M\in E$. Donnez l'expression de $\sigma$ dans ces coordonnées.*
*(3) Montrez que $\sigma$ est une bijection en donnant l'expression de $\sigma^{-1}$. On l'étend en une bijection de $S^2$ dans $\widehat{\mathbb{C}}=\mathcal{P}\cup\{\infty\}$, encore notée $\sigma$, en posant $\sigma(N)=\infty$.*

**Corrigé.** (2) La coordonnée $z$ de $M$ est l'affixe complexe de la projection de $M$ sur le plan $\mathcal{P}$, notée $m$. Notons $O$ l'origine de l'espace, on voit que le point $M'=\sigma(M)$ est sur le plan contenant $O$, $M$ et $N$. En particulier, en tant que point de $\mathcal{P}$ il appartient à la demi-droite $[Om)$, donc son affixe est de la forme $\lambda z$ avec $\lambda\in\mathbb{R}^+$. En appliquant le théorème de Thalès dans le triangle $ONM'$ on trouve

$$\frac{t}{1}=\frac{\lambda-1}{\lambda}$$

donc $\lambda=\frac{1}{1-t}$. Finalement $\sigma(z,t)=\frac{z}{1-t}$.
(3) Notons $z'$ l'affixe du point $M'\in\mathcal{P}$. Les mêmes considérations que précédemment dans le plan $(OMN)$ montrent que $M=\sigma^{-1}(M')$ a des coordonnées de la forme $(\mu z',t)$ pour un certain $\mu\in\mathbb{R}^+$. On a les deux contraintes supplémentaires

(i) $\mu^2|z'|^2+t^2=1$ car $M\in S^2$, et
(ii) $\frac{\mu z'}{1-t}=z'$ car $\sigma(M)=M'$.

De (ii) in tire $\mu=1-t$. En remplaçant $t=1-\mu$ dans (i) on trouve

$$\mu^2|z|^2+\mu^2-2\mu=0\ .$$

Le cas $\mu=0$ correspond à $z'=0$, et dans le cas $\mu\neq0$ on trouve

$$\mu=\frac{2}{|z'|^2+1}\qquad\text{et enfin}\qquad\sigma^{-1}(z')=\left(\frac{2z'}{|z'|^2+1},\frac{|z'|^2-1}{|z'|^2+1}\right)\ .$$

Cette dernière expression est valable sans restriction sur $z'$, le traitement à part du cas $\mu=0$ n'était que temporaire.

**Exercice 5** *On conserve les notations utilisées ci-dessus. Pour tout $v=(a,b)\in\mathbb{C}^2\backslash\{0\}$, on note $\pi(x)=(a:b)$ image dans la droite projective $\mathbb{P}^1(\mathbb{C})$, et on appelle $a,b$ ses* coordonnées projectives. *(Ce ne sont pas des coordonnées au sens propre du terme, puisqu'elles ne sont définies qu'à multiplication près par un scalaire inversible $\lambda\in\mathbb{C}^*$.)*

*On définit $a:X\to\mathbb{P}^1(\mathbb{C})$ par $a_{|Y}$ qui envoie $y\in Y$ sur $(y:1)$ et $a_{|Z}$ qui envoie $z\in Z$ sur $(1:z)$. On définit $b:\mathbb{P}^1(\mathbb{C})\to S^2$ par*

$$b(u:v)=\left(\frac{2u\overline{v}}{|u|^2+|v|^2},\frac{|u|^2-|v|^2}{|u|^2+|v|^2}\right)\in S^2\subset\mathbb{C}\oplus\mathbb{R}\simeq\mathbb{R}^3$$

*(voir notations de l'exercice précédent).*

*(1) Vérifiez que $a$ et $b$ sont bien définies.*

*(2) On considère le point $0\in\mathbb{C}=Z\subset X$ et ses voisinages $B(0,\epsilon)=\{z\in Z,\,|z|<\epsilon\}\subset X$, pour $\epsilon>0$. Montrez que via $a,b,\sigma$, cette base d'ouverts s'envoie sur des bases de voisinages ouverts des points images :*

$$\begin{array}{ccccccc} X & \xrightarrow{a} & \mathbb{P}^1(\mathbb{C}) & \xrightarrow{b} & S^2 & \xrightarrow{\sigma} & \widehat{\mathbb{C}} \\ 0\in Z & \mapsto & (1:0) & \mapsto & (z,t)=(0,1) & \mapsto & \infty \end{array}$$

*Les calculs étant similaires en les points autres que $\infty$, on admettra que ceci démontre que $a,b,\sigma$ sont des homéomorphismes, compte tenu du lemme (« une bijection qui échange des bases d'ouverts est un homéo »).*

**Corrigé.** (1) Pour vérifier que $a$ est bien définie il suffit de voir que l'application $\tilde{a}:Y\amalg Z\to\mathbb{P}^1(\mathbb{C})$ définie par $a_{|Y}$ et $a_{|Z}$, passe au quotient par la relation d'équivalence qui définit $X$ (on dit que $a_{|Y}$ et $a_{|Z}$ *se recollent*). Rappelons-nous que pour tout $\lambda\in\mathbb{C}^*$ on a $(\lambda a:\lambda b)=(a:b)$. Alors l'assertion à voir est claire, car si $y=\varphi(z)=1/z$, on a

$$a_{|Y}(y)=(y:1)=(1/z:1)=(1:z)=a_{|Z}(z)\ .$$

Pour $b$, il n'y a qu'à voir que l'image $b(u:v)$ est bien dans $S^2$, ce qui est facile.

(2) On vérifie que l'image de la boule $B(0,\epsilon) \subset X$ est comme suit :
- dans $\mathbb{P}^1(\mathbb{C})$, sur l'ensemble $\{(a:b),\ \epsilon|a| > |b|\}$,
- dans $S^2$, sur l'ensemble $\{(z,t) \in \mathbb{C} \oplus \mathbb{R},\ t > \frac{1-\epsilon^2}{1+\epsilon^2}\}$.
- dans $\widehat{\mathbb{C}}$, sur le complémentaire de la boule fermée centrée en $0 \in \mathbb{C}$ de rayon $1/\epsilon$.

Ce sont des bases de voisinages d'ouverts des points respectifs.

# Dualité et sous-réseaux

Soit $i\colon A \to B$ un morphisme injectif entre deux $\mathbb{Z}$-modules libres de même rang fini $n$. Le conoyau $E = B/A$ est donc un groupe fini, et on a une suite exacte

$$0 \to A \to B \to E \to 0\ .$$

Précisément, la théorie des invariants de similitude dit qu'on peut choisir des bases (distinctes) dans $A$ et $B$ dans lesquelles la matrice de $i$ s'écrit

$$\begin{pmatrix} a_1 & & & 0 \\ & a_2 & & \\ & & \ddots & \\ 0 & & & a_n \end{pmatrix}$$

avec $a_i | a_{i+1}$ pour tout $i$. On en déduit que $E \simeq \mathbb{Z}/a_1\mathbb{Z} \times \cdots \times \mathbb{Z}/a_n\mathbb{Z}$, en particulier l'ordre de $E$ est le produit des $|a_i|$, c'est-à-dire encore, la valeur absolue du déterminant de $i$. Si on dualise, au sens du dual $A^* = \mathrm{Hom}(A,\mathbb{Z})$ ([4]), on obtient un morphisme transposé $i^*\colon B^* \to A^*$ qui est encore une injection de $\mathbb{Z}$-modules libres de rang $n$. La question que je pose est de déterminer le conoyau dans la suite ci-dessous :

$$0 \to B^* \to A^* \to ?\ \to 0$$

La réponse n'est certainement pas $E^*$, car $E^* = 0$ : vérifiez-le. Le but de l'exercice qui suit est de montrer que le conoyau recherché est le dual de Pontryagin $E^\dagger = \mathrm{Hom}(E, \mathbb{Q}/\mathbb{Z})$ ([5]) ([6]).

[4] Le Hom désigne les homomorphismes de groupes abéliens.

[5] Ici encore.

[6] Le dual de Pontryagin d'un groupe topologique commutatif localement compact $G$ est l'ensemble des homomorphismes continus de $G$ dans le groupe $\mathbb{U}$ des complexes de module 1. Or on a une injection $\mathbb{Q}/\mathbb{Z} \hookrightarrow \mathbb{U}$ donnée par $r \mapsto e^{2i\pi r}$. Si $G$ est fini discret, tout morphisme $G \to \mathbb{U}$ est continu, et il est clair qu'il a son image dans $\mathbb{Q}/\mathbb{Z}$, de sorte que $\mathrm{Hom}_{cont}(G,\mathbb{U}) = \mathrm{Hom}(E,\mathbb{Q}/\mathbb{Z})$.

**Exercice 1** (1) *Démontrez qu'il existe un entier $m$ tel que $mB \subset A$.*

(2) *Notons $F$ le conoyau de $i^*\colon B^* \to A^*$. Définissez un accouplement*

$$\langle \cdot, \cdot \rangle \colon E \times F \to \mathbb{Q}/\mathbb{Z}\,,$$

*montrez que cet accouplement est non-dégénéré, et déduisez-en le résultat.* (Indications : si $(e, f) \in E \times F$, alors $e$ est la classe d'un élément $b \in B$ et $f$ est la classe d'une « forme » $\varphi\colon A \to \mathbb{Z}$. Utilisez alors (1) pour définir $\langle e, f \rangle$.)

# Éléments d'ordre fini dans un groupe

*Soit $G$ un groupe et $E$ l'ensemble de ses éléments d'ordre fini. Montrer que si $E$ est fini, c'est un sous-groupe de $G$.*

**Corrigé.** Quitte à remplacer $G$ par le sous-groupe engendré par $E$, on peut supposer que $E$ engendre $G$. La question revient alors à montrer que $G$ est fini.

Écrivons $E = \{e_1, \dots, e_r\}$. Pour chaque $(i, j)$, l'élément $e_i e_j e_i^{-1}$ est d'ordre fini, donc de la forme $e_{\sigma(i,j)}$. On a donc $e_i e_j = e_{\sigma(i,j)} e_i$. De plus, clairement, si $i \neq j$ alors $\sigma(i, j) \neq i$.

On va montrer que tout $g \in G$ a une écriture de la forme $g = (e_1)^{\epsilon_1} \dots (e_r)^{\epsilon_r}$, avec $\epsilon_i \in \{0, 1\}$, ce qui entraînera clairement le résultat. Comme $E$ engendre $G$, on peut écrire $g = e_{i_1} \dots e_{i_s}$ où on suppose la longueur $s$ choisie minimale. Utilisant $e_r e_j = e_{\sigma(r,j)} e_r$, on voit que si l'un des $i_k$ est égal à $r$ on peut le faire passer à droite pour avoir $g = e_{i'_1} \dots e_{i'_{s-1}} e_r$. Alors, aucun des indices restants n'est égal à $r$, car sinon on le passerait de nouveau à droite et on ferait diminuer la longueur puisque $(e_r)^2 \in E$. Pour la même raison, dans la suite, si on utilise les relations $e_i e_j = e_{\sigma(i,j)} e_i$ pour transformer $e_{i'_1} \dots e_{i'_t}$, l'élément $e_r$ ne pourra jamais apparaître. On itère alors ce procédé avec $e_{i'_1} \dots e_{i'_t}$ pour faire passer $e_{r-1}$ à droite, etc.

# Espaces propres et dualité

Les trois exercices qui suivent portent sur le thème suivant : étant donné un endomorphisme $f$ d'un espace vectoriel de dimension finie, le dual d'un sous-espace propre (resp. caractéristique) est-il le sous-espace propre (resp. caractéristique) du dual ? (Ici le dual de $f$ est à comprendre comme le transposé.)

Ceux qui souhaitent faire quelques révisions peuvent faire les trois exercices dans l'ordre, et ceux qui se sentent plus à l'aise peuvent passer directement au troisième.

**Exercice 1** *Soient $k$ un corps, $E$ un $k$-espace vectoriel de dimension finie $n$, et $f \in \mathcal{L}(E)$. Pour toute valeur propre $\lambda$ de $f$, soient $\alpha$ sa multiplicité dans le polynôme minimal de $f$ et $\beta$ sa multiplicité dans le polynôme caractéristique de $f$. Montrez qu'on peut définir le sous-espace caractéristique $E(\lambda)$ par l'une quelconque des égalités :*

(1) $E(\lambda) = \ker (f - \lambda)^{\alpha}$.

(2) $E(\lambda) = \ker (f - \lambda)^{\beta}$.

(3) $E(\lambda) = \ker (f - \lambda)^{n}$.

**Exercice 2** *Soient $k$ un corps, $E$ un $k$-espace vectoriel de dimension finie $n$, et $f \in \mathcal{L}(E)$. Pour toute valeur propre $\lambda$ de $f$, on note $E(\lambda)$, resp. $E[\lambda]$ le sous-espace caractéristique, resp. le sous-espace propre correspondant. Soit $f^* \in \mathcal{L}(E^*)$ le transposé, et $E^*(\lambda)$, $E^*[\lambda]$ les sous-espaces correspondants. Montrez que* $\dim E^*(\lambda) = \dim E(\lambda)$ *et* $\dim E^*[\lambda] = \dim E[\lambda]$.

**Exercice 3** *Soient $k$ un corps, $E$ un $k$-espace vectoriel de dimension finie $n$, et $f \in \mathcal{L}(E)$.*

(1) *Montrez que* $E = \mathrm{im}(f^n) \oplus \ker(f^n)$.

(2) *Montrez que les morphismes $E^* \to E(\lambda)^*$ et $E^* \to E[\lambda]^*$, obtenus par dualité à partir des inclusions $E(\lambda) \subset E$ et $E[\lambda] \subset E$, sont surjectifs.*

*On considère les composés $\varphi(\lambda)\colon E^*(\lambda) \to E^* \to E(\lambda)^*$ et $\varphi[\lambda]\colon E^*[\lambda] \to E^* \to E[\lambda]^*$.*

(3) *Montrez que $\varphi(\lambda)$ est un isomorphisme.*

(4) *Montrez que $\varphi[\lambda]$ est un isomorphisme si et seulement si $f$ est semi-simple en $\lambda$ (ce qui signifie que la multiplicité de $\lambda$ dans le polynôme minimal de $f$ est égale à 1, ou encore, que $E(\lambda) = E[\lambda]$).*

# Le $n$ de $\mathrm{GL}_n$

Soient $K$ un corps et $m, n$ deux entiers. On suppose qu'on a un isomorphisme de groupes $\mathrm{GL}_n(K) \simeq \mathrm{GL}_m(K)$. A-t-on $m = n$ ?

*Si* $\mathrm{car}(K) \neq 2$, il est classique que $\mathrm{GL}_n(K) \simeq \mathrm{GL}_m(K)$ implique $m = n$. Pour le voir, on considère les sous-groupes finis $G \subset \mathrm{GL}_n(K)$ qui sont d'exposant 2, c'est-à-dire tels que $M^2 = \mathrm{Id}$ pour tout $M \in G$. Un tel groupe est abélien. De plus toutes les matrices de $G$ sont annulées par le polynôme à racines simples $X^2 - 1$, donc elles sont diagonalisables (cet argument échoue si $\mathrm{car}(K) = 2$). Comme $G$ est abélien, ces matrices sont simultanément diagonalisables, et donc il existe un conjugué de $G$ dont tous les éléments sont des matrices diagonales, avec des coefficients $\pm 1$. Donc $|G| \leqslant 2^n$ et ce cardinal est atteint. Donc si $\mathrm{GL}_n(K) \simeq \mathrm{GL}_m(K)$, alors $2^n = 2^m$ et donc $n = m$.

La question reste posée si $\mathrm{car}(K) = 2$, et la réponse est encore positive, comme on va le voir dans l'exercice 5. Un regard rétrospectif sur les raisonnements faits permettra de voir qu'ils montrent en fait que $\mathrm{GL}_n(K)$ *détermine* $n$*, indépendamment de* $K$. Ainsi, si $K$ et $L$ sont deux corps et si $\mathrm{GL}_n(K) \simeq \mathrm{GL}_m(L)$, alors on a $m = n$.

Plus généralement, si $K$ et $L$ sont deux corps et $\mathrm{GL}_n(K) \simeq \mathrm{GL}_m(L)$, on peut se demander si on a $K \simeq L$, ou au moins $\mathrm{car}(K) = \mathrm{car}(L)$. Il est possible de montrer (nous ne le ferons pas ici : voir le recueil d'exercices en référence ci-dessous) que si $m = n \geqslant 2$, alors $\mathrm{car}(K) = \mathrm{car}(L)$. Il se trouve que c'est faux si $n = 1$, et nous nous contenterons de donner un contre-exemple (exercice 6), qui fait un exercice très intéressant par ailleurs.

Avant de passer aux choses sérieuses, donnons des références :

J. Fresnel, M. Matignon, *Algèbre et Géométrie, un recueil d'exercices*, École Mathématique et Informatique Bordeaux 1. Peut se commander auprès de l'E.M.I. de Bordeaux 1, prix 10 €. J'en ai un exemplaire et je peux scanner la page s'il y a des personnes intéressées.

J. Fresnel, *Algèbre des matrices*, Hermann. C'est l'exercice A.4.7.21.3.

$\therefore$

**Définition 1** *Soit $G$ un groupe (pas nécessairement fini). On dit que c'est un* $p$-groupe *si et seulement si tout élément est d'ordre fini égal à une puissance de $p$.*

**Définition 2** *Soit $G \subset \mathrm{GL}_n(K)$ un sous-groupe. On dit que $G$ est* unipotent *si et seulement si tous ses éléments sont des matrices unipotentes, c'est-à-dire sommes de la matrice identité et d'une matrice nilpotente.*

**Exercice 1** *Soient $K$ un corps de caractéristique $p$ et $T \subset \mathrm{GL}_n(K)$ le sous-groupe des matrices triangulaires supérieures avec des 1 sur la diagonale.*

(1) *Montrez que $M \in \mathrm{GL}_n(K)$ est d'ordre une puissance de $p$ si et seulement si $M$ est unipotente. (Rappelez-vous qu'en caractéristique $p$, $(M+N)^p = M^p + N^p \ldots$).*

(2) *Soit $H \subset \mathrm{GL}_n(K)$ un sous-groupe. Montrez que les conditions suivantes sont équivalentes :*

(i) *$H$ est unipotent.*

(ii) *$H$ est un $p$-sous-groupe.*

(iii) *$H$ est conjugué à un sous-groupe de $T$.*

(3) *Montrez que les conditions suivantes sont équivalentes :*

(i) *$H$ est unipotent maximal.*

(ii) *$H$ est un $p$-sous-groupe maximal.*

(iii) *$H$ est conjugué à $T$.*

**Définition 3** *Soit $G$ un groupe et $H, K$ deux sous-groupes. On note $[H, K]$ le sous-groupe engendré par les commutateurs $hkh^{-1}k^{-1}$ avec $h \in H$, $k \in K$.*

*On appelle* suite centrale descendante de $G$ *la suite $C^iG$ définie par $C^1G = G$ et $C^{i+1}G = [G, C^iG]$. S'il existe $i$ tel que $C^iG = \{1\}$, on dit que $G$ est un groupe* nilpotent*, et on appelle* indice de nilpotence *le plus petit tel indice $i$.*

**Exercice 2** *Mêmes notations que dans l'exercice précédent.*

(1) *Si $H$ est un $p$-sous-groupe maximal, montrez que $C^n H = \{1\}$ et $C^{n-1} H \neq \{1\}$.*

(2) *Déduisez-en que si $\mathrm{GL}_n(K) \simeq \mathrm{GL}_m(K)$, alors $n = m$.*

Dans le cas $m = n = 1$, voici le contre-exemple au fait que

$$\text{« } \mathrm{GL}_n(K) \simeq \mathrm{GL}_m(K) \Rightarrow \mathrm{car}(K) = \mathrm{car}(L) \text{ ».}$$

**Exercice 3** *Montrez que $\mathrm{GL}_1(\mathbb{Q}) \simeq \mathrm{GL}_1(\mathbb{F}_3(t))$, où $\mathbb{F}_3(t)$ est le corps des fractions rationnelles en une variable.*

*(Indications : décrivez $\mathbb{Q}^\times$ et $\mathbb{F}_3(t)^\times$ en termes des irréductibles de $\mathbb{Z}$ et de $\mathbb{F}_3[t]$...)*

# Groupe d'exposant 3 non abélien

On dit qu'un groupe est d'exposant fini s'il existe un entier $k \geqslant 1$ tel que pour tout élément $x$ de $G$, on a $x^k = 1$. Dans ce cas, on appelle exposant de $G$ le plus petit tel entier.

(1) Montrer qu'un groupe d'exposant 2 est abélien.

(2) Donner un exemple de groupe d'exposant 3 non abélien.

**Corrigé.** (1) On suppose que pour tout $x$ dans $G$, on a $x^2 = 1$. Ceci veut dire que $x^{-1} = x$. Par ailleurs, pour $x, y$ dans $G$ on a $(xy)^2 = xyxy = 1$. En multipliant à gauche par $x^{-1} = x$ et à droite par $y^{-1} = y$, on trouve $yx = xy$, donc $G$ est abélien.

(2) On applique une stratégie de base pour trouver des contre-exemples : on essaie de trouver le plus simple contre-exemple possible. Dans notre cas, on va chercher un contre-exemple parmi les groupes finis, de cardinal le plus petit possible. D'après le lemme de Cauchy, le groupe $G$ doit être un 3-groupe, car si un nombre premier $p$ divise l'ordre de $G$, alors il existe un élément $x$ d'ordre $p$. Mais comme $x^3 = 1$, on trouve que $p = 3$.

Donc l'ordre de $G$ est de la forme $3^n$ pour un certain $n$. Supposons que $n$ est choisi minimal, c'est-à-dire qu'il existe un groupe d'ordre $3^n$ non abélien, mais tous les groupes d'ordre $3^{n-1}$ sont abéliens. Par une propriété classique des $p$-groupes, il existe dans $G$ un sous-groupe distingué $N$ d'ordre $3^{n-1}$.

Posons $H = G/N$, c'est un groupe d'ordre 3 donc isomorphe à $\mathbb{Z}/3\mathbb{Z}$. Comme $G$ est d'exposant 3, n'importe quel antécédent d'un générateur de $H$ est d'ordre 3 donc engendre un sous-groupe de $G$ isomorphe à $H$. Ceci montre que la suite exacte

$$1 \to N \to G \to H \to 1$$

est scindée, ou dit autrement, que $G$ est produit semi-direct de $N$ par $H$. Ce produit semi-direct est décrit par un morphisme

$H \to \mathrm{Aut}(N)$, et $G$ est non abélien ssi ce morphisme est non trivial. Comme $H$ n'a pas de sous-groupe non trivial, ce morphisme est non trivial ssi il est injectif.

Par ailleurs, d'après l'hypothèse faite sur $G$, le groupe $N$ est d'exposant 3, et par choix de $n$ il est abélien. Donc $N$ est isomorphe à $(\mathbb{Z}/3\mathbb{Z})^{n-1}$. Ainsi $\mathrm{Aut}(N) \simeq \mathrm{GL}_{n-1}(\mathbb{Z}/3\mathbb{Z})$. Il s'agit donc de trouver un sous-groupe d'ordre 3 dans $\mathrm{GL}_{n-1}(\mathbb{Z}/3\mathbb{Z})$. Le plus petit $n$ pour lequel c'est possible est $n = 3$, et un exemple de tel sous-groupe est le sous-groupe des matrices unipotentes de la forme

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$$

avec $a$ appartenant à $\mathbb{Z}/3\mathbb{Z}$.

Pour conclure, on pose $N = (\mathbb{Z}/3\mathbb{Z})^2$ et $H = \mathbb{Z}/3\mathbb{Z}$. On appelle $\theta : H \to \mathrm{Aut}(N) = \mathrm{GL}_2(\mathbb{Z}/3\mathbb{Z})$ le morphisme qui envoie $a$ sur la matrice unipotente ci-dessus. Alors le produit semi-direct de $N$ par $H$ correspondant à $\theta$ est un groupe d'ordre 27, non abélien. On doit faire attention tout de même que, dans le raisonnement qui précède, rien ne montre que ce groupe est d'exposant 3. Il faut le vérifier, ce qui est un exercice facile de manipulation sur le produit semi-direct.

Une fois cette réponse trouvée, on peut regarder cet exemple droit dans les yeux et reconnaître un groupe connu qui s'est caché... On voit que la clef de l'exercice est donnée par les matrices unipotentes et qu'en fait le groupe qu'on a construit n'est autre que le groupe $U_3(\mathbb{F}_3)$ des matrices unipotentes triangulaires supérieures

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}$$

à coefficients dans le corps à trois éléments $\mathbb{F}_3$. Notez que sur cette description, il est facile de voir que ce groupe est d'exposant 3.

# Groupes nilpotents

Références : Calais, *Eléments de théorie des groupes* (th. 7.60 page 254) ou Rotman, *An Introduction to the theory of groups* (th. 5.39).

**Définition 1** *Soit $G$ un groupe. On définit la* suite centrale descendante *de $G$ par récurrence, par $C^1G = G$ et $C^{i+1}G = [G, C^iG]$ (groupe engendré par les commutateurs d'un élément de $G$ et un élément de $C^iG$). C'est une suite décroissante de sous-groupes distingués de $G$. (Vérifiez-le.)*

*On dit que $G$ est* nilpotent *s'il existe $n$ tel que $C^{n+1}G = \{1\}$. Si c'est le cas, le plus petit tel $n$ est appelé la* classe de nilpotence *de $G$.*

**Remarque 3** $\star\star\star$ Si $G$ est nilpotent, alors pour un certain $n$ on a $C^nG \neq 1$ et $C^{n+1}G = \{1\}$. Ceci veut dire que $C^nG$ est central, donc le centre de $G$ est non trivial. (On connaît une autre classe de groupes qui ont cette propriété : les $p$-groupes finis.)

**Exercice 1** *(1) Montrez qu'un groupe $G$ est nilpotent si et seulement s'il existe une suite de sous-groupes distingués de $G$,*

$$G^1 = G \supset G^2 \supset \cdots \supset G^n \supset G^{n+1} = 1$$

*telle que $G^i/G^{i+1}$ est central dans $G/G^{i+1}$ pour tout $i$ (ou encore, de manière équivalente, $[G, G^i] \subset G^{i+1}$).*

(2) *Montrez que tout sous-groupe et tout quotient d'un groupe nilpotent sont nilpotents.*

(3) *Montrez que tout produit* fini *de groupes nilpotents est nilpotent.*

**Exercice 2** *Soit $p$ un nombre premier, montrez qu'un $p$-groupe fini est nilpotent.*
*(*Indication : utilisez le (1) de l'exercice précédent en construisant une suite

$$\cdots \supsetneq G_3 \supsetneq G_2 \supsetneq G_1$$

Il est plus simple de construire cette suite à l'envers, c'est pourquoi la numérotation est inversée. Commencez par choisir un sous-groupe $G_1 (= G^n)$ avec les propriétés requises, puis $G_2 (= G^{n-1})$,...*)*

D'après les derniers exercices, tout produit fini de $p$-groupes finis est nilpotent. En fait, tout les groupes finis nilpotents sont comme ça :

**Exercice 3** Tout groupe fini nilpotent $G$ est produit direct de ses sous-groupes de Sylow.

*On note $|G| = (p_1)^{\alpha_1} \dots (p_n)^{\alpha_n}$ l'ordre de $G$ et $Z \subset G$ son centre. Pour tout $i$ on note $Z_i$ le $p_i$-Sylow de $Z$ (il n'y en a qu'un, car les $p_i$-Sylow sont conjugués or $Z$ est abélien), et on choisit un $p_i$-Sylow $P_i$ de $G$. Montrez successivement que :*

(1) *$Z_i \subset P_i$.*

(2) *$P_i$ est distingué dans $G$.*

(3) *Pour $i \neq j$, on a $[P_i, P_j] = \{1\}$, c'est-à-dire, $P_i$ et $P_j$ commutent.*

(4) *Conclure.*

**Corrigé du dernier exercice.**

(1) Comme $G$ est nilpotent, son centre $Z$ n'est pas réduit à $\{1\}$. Considérons le morphisme de quotient $\pi\colon G \to G/Z$. L'image de $P_i$ par $\pi$ est un sous-$p_i$-groupe de $G/Z$, isomorphe à $P_i/P_i \cap Z_i$ (en effet, il est clair que le noyau de $P_i \hookrightarrow G \to G/Z$ est $P_i \cap Z_i$). Comme $P_i \cap Z_i \subset Z_i$, on a

$$|\pi(P_i)| = |P_i/P_i \cap Z_i| = \frac{|P_i|}{|P_i \cap Z_i|} \geqslant \frac{|P_i|}{|Z_i|} .$$

Or si on note $(p_i)^{\beta_i}$ l'ordre de $Z_i$, on a $|G/Z| = (p_1)^{\alpha_1-\beta_1} \dots (p_n)^{\alpha_n-\beta_n}$. Donc $|P_i|/|Z_i| = (p_i)^{\alpha_i-\beta_i}$ est l'ordre d'un $p_i$-Sylow de $G/Z$. Un sous-$p_i$-groupe de $G/Z$ ne peut avoir un ordre strictement plus grand, donc $|P_i/P_i \cap Z_i| = (p_i)^{\alpha_i-\beta_i}$. Ainsi $|P_i \cap Z_i| = (p_i)^{\beta_i} = |Z_i|$, donc $P_i \cap Z_i = Z_i$ puis $Z_i \subset P_i$.

(2) Soit $g \in G$, on doit montrer que $gP_ig^{-1} = P_i$. Or on a là deux $p_i$-Sylow de $G$, qui d'après le (1) contiennent $Z_i$. On considère le morphisme de quotient $\pi\colon G \to G/Z$. Par une récurrence sur l'ordre de $G$, les images de $gP_ig^{-1}$ et $P_i$ par $\pi$, qui sont des $p_i$-Sylow de $G/Z$ (voir (1)), sont égales. Ceci s'écrit, d'après (1) encore, $P_i/Z_i = gP_ig^{-1}/Z_i$. On en déduit que $P_i = gP_ig^{-1}$.

(3) Soient $a \in P_i$ et $b \in P_j$, utilisant le fait que $P_i$ et $P_j$ sont distingués, on a :

$$aba^{-1}b^{-1} = \underbrace{(\underbrace{aba^{-1}}_{\in P_j})b^{-1}}_{\in P_j} = \underbrace{a(\underbrace{ba^{-1}b^{-1}}_{\in P_i})}_{\in P_i} \in P_i \cap P_j$$

Comme $P_i$ et $P_j$ sont d'ordres premiers entre eux, $P_i \cap P_j = \{1\}$. Donc tous les commutateurs sont triviaux, le sous-groupe de commutateurs engendré est trivial.

(4) Comme les $P_i$ commutent entre eux, pour tout $m \leqslant n$, l'application suivante

$$\begin{aligned} \varphi_m\colon \quad & P_1 \times P_2 \times \dots \times P_m \to G \\ & (a_1, a_2, \dots, a_m) \mapsto a_1a_2 \dots a_m \end{aligned}$$

est un morphisme de groupes. (Il est utile de considérer tous les $\varphi_m$ et pas seulement $\varphi_n$ pour pouvoir faire la récurrence qui va suivre.) Montrons par récurrence sur $m$ que $\varphi_m$ est injectif. En effet, pour $m = 1$ c'est clair, et pour $m \geqslant 2$, si $a_1a_2 \dots a_m = 1$ alors $a_1a_2 \dots a_{m-1} = (a_m)^{-1}$ est dans l'intersection de $P_1 \times P_2 \times \dots \times P_{m-1}$, vu comme sous-groupe de $G$ par l'hypothèse de récurrence, et de $P_m$. Ces deux groupes ont des ordres premiers entre eux, donc leur intersection est réduite à $\{1\}$, donc $a_1a_2 \dots a_{m-1} = (a_m)^{-1} = 1$. Par l'hypothèse de récurrence on a donc $a_1 = a_2 = \dots = a_m = 1$, et $\varphi_m$ est injectif. Pour $m = n$, on a $\varphi_n$ injectif entre deux groupes de même ordre, c'est un isomorphisme.

## Groupes sans automorphismes

*Trouver tous les groupes $G$ tels que* $\mathrm{Aut}(G) = \{1\}$.

**Corrigé.** Parmi les automorphismes d'un groupe, il y a toujours les automorphismes intérieurs. Plus précisément, l'ensemble des automorphismes intérieurs forme un sous-groupe distingué de $\mathrm{Aut}(G)$, isomorphe à $G/Z(G)$ où $Z(G)$ est le centre de $G$. Donc si $\mathrm{Aut}(G) = \{1\}$ alors $G = Z(G)$, c'est-à-dire, $G$ est abélien. On écrit donc $G$ en notation additive.

On observe ensuite que l'application $x \mapsto -x$ est un automorphisme (si $G$ n'est pas abélien, l'application $x \mapsto x^{-1}$ n'est pas un morphisme.) Si $\mathrm{Aut}(G) = \{1\}$, on a donc $-x = x$ pour tout $x \in G$, c'est-à-dire $2x = 0$, donc $G$ est un $\mathbb{F}_2$-espace vectoriel (en détails : on a une application bien définie $\mathbb{F}_2 \times G \to G$ qui à $(n, x)$ associe $nx$.) De plus les automorphismes de $G$ comme groupe sont exactement les automorphismes de $G$ comme $\mathbb{F}_2$-ev (c'est immédiat à vérifier). Si $\dim_{\mathbb{F}_2}(G) \geqslant 2$, alors $G$ possède une base ayant deux éléments distincts $e_1, e_2$, et on peut définir un automorphisme non trivial qui échange $e_1$ et $e_2$, et fixe tous les autres éléments de la base. Donc si $\mathrm{Aut}(G) = \{1\}$, on doit avoir $\dim_{\mathbb{F}_2}(G) \leqslant 1$. Si la dimension est 0 on a $G \simeq \{1\}$ et si la dimension est 1 on a $G \simeq \mathbb{Z}/2\mathbb{Z}$. Il est clair que ces deux groupes vérifient $\mathrm{Aut}(G) = \{1\}$.

# Homographies et birapport

**Exercice 1** *(Le groupe des homographies.) On appelle* homographie *une application du plan complexe dans lui-même de la forme $h(z) = \frac{az+b}{cz+d}$, où $a, b, c, d$ sont quatre complexes tels que $ad - bc \neq 0$.*

*(1) Indiquez l'ensemble de définition et l'image d'une homographie.*
*(2) Montrez qu'une homographie se prolonge de manière naturelle en une bijection de $\mathbb{P}^1(\mathbb{C})$ dans $\mathbb{P}^1(\mathbb{C})$.*

*Dorénavant, par le terme d'*homographie*, on entendra une transformation $h : \mathbb{P}^1(\mathbb{C}) \to \mathbb{P}^1(\mathbb{C})$ de la forme précédente.*

*(3) Montrez que l'ensemble $\mathcal{H}$ des homographies possède une structure naturelle de groupe et que ce groupe est engendré par les similitudes directes et par l'application $z \mapsto 1/z$.*
*(4) À la matrice $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ on associe l'homographie $h(z) = \frac{az+b}{cz+d}$. Montrez que ceci définit un morphisme de groupes $\mathrm{GL}_2(\mathbb{C}) \to \mathcal{H}$. Déterminez son image et son noyau.*

**Corrigé.** (1) On voit qu'il y a deux cas à distinguer, selon que $c = 0$ ou $c \neq 0$. Dans le premier cas, $h$ est une similitude directe, son ensemble de définition est $\mathbb{C}$ et son image est $\mathbb{C}$ également.

Si $c \neq 0$ l'ensemble de définition de $h$ est $\mathbb{C}\backslash\{-d/c\}$. On voit ici que $h$ induit une bijection de $\mathbb{C}\backslash\{-d/c\}$ sur $\mathbb{C}\backslash\{a/c\}$ d'inverse $g(y) = \frac{-dy+b}{cy-a}$. Donc $h$ a pour image $\mathbb{C}\backslash\{a/c\}$.
(2) On va garder la lettre $h$ pour l'extension à $\mathbb{P}^1(\mathbb{C})$. Si $c = 0$, c'est-à-dire que $h$ est une similitude directe, alors c'est une bijection de $\mathbb{C}$ et il suffit de poser $h(\infty) = \infty$ pour obtenir une bijection de $\mathbb{P}^1(\mathbb{C})$ qui fixe $\infty$.

Si $c \neq 0$ on a vu que $h$ induit une bijection de $\mathbb{C}\backslash\{-d/c\}$ sur $\mathbb{C}\backslash\{a/c\}$. Si on pose $h(-d/c) = \infty$ et $h(\infty) = a/c$, on obtient une bijection de $\mathbb{P}^1(\mathbb{C})$, et l'inverse est $g(y) = \frac{-dy+b}{cy-a}$ avec $g(\infty) = -d/c$ et $g(a/c) = \infty$.
(3) On va montrer que $h$ est un sous-groupe du groupe des bijections de $\mathbb{P}^1(\mathbb{C})$. L'identité appartient à $\mathcal{H}$, car c'est l'homographie correspondant à $a = d = 1$ et $b = c = 0$. On a déja vu que l'inverse

(pour la composition) d'une homographie est une homographie, et on a donné sa formule. Il reste à voir que la composée de deux homographies $h, h'$ est une homographie. Or

$$h(h'(z)) = \frac{a\frac{a'z+b'}{c'z+d'} + b}{c\frac{a'z+b'}{c'z+d'} + d} = \frac{(aa'+bc')z + (ab'+bd')}{(ca'+dc')z + (cb'+dd')} .$$

Montrons que $\mathcal{H}$ est engendré par les similitudes directes et par $i(z) = 1/z$. Soit $h$ une homographie. Si c'est une similitude directe il n'y a rien à démontrer. Sinon, on a $c \neq 0$ et on écrit

$$h(z) = \frac{az+b}{cz+d} = \frac{\frac{a}{c}(cz+d) + b - \frac{ad}{c}}{cz+d} = \frac{a}{c} + \frac{bc-ad}{c}\frac{1}{cz+d} .$$

On voit donc que $h$ est composée de $z \mapsto cz+d$, puis $i$, puis $z \mapsto \frac{bc-ad}{c}z + \frac{a}{c}$. C'est bien une composée de similitudes directes et de $i$.
(4) L'application $A \mapsto f(A) = h$ de l'énoncé envoie la matrice identité sur la transformation identité de $\mathbb{P}^1(\mathbb{C})$. De plus, en comparant l'expression du produit matriciel de deux matrices $A$ et $A'$ avec l'expression de la composée des homographies $h = f(A)$ et $h' = f(A')$, on voit que $f$ est un morphisme de groupes. Par définition des homographies, ce morphisme est surjectif. Calculons son noyau. Soit $A \in \mathrm{GL}_2(\mathbb{C})$ et $h = f(A)$ telle que $h(z) = \frac{az+b}{cz+d} = z$ pour tout $z \in \mathbb{P}^1(\mathbb{C})$. Alors on a $az+b = cz^2+dz$ pour tout $z$, et cette égalité de polynômes en $z$ implique $c = 0$, $a = d$, $b = 0$. Ceci signifie que $A$ est une matrice d'homothétie. Le noyau de $f$ est le sous-groupe des homothéties.

**Exercice 2** *(Le birapport.)*
*(1) Soient $a, b, c$ dans $\mathbb{P}^1(\mathbb{C})$ distincts. Montrez qu'il existe une unique homographie $h$ envoyant $a, b, c$ sur $0, 1, \infty$. On supposera pour simplifier qu'aucun des points $a, b, c$ n'est égal au point à l'infini.*

*Soient $a, b, c, d$ quatre éléments de $\mathbb{P}^1(\mathbb{C})$ dont les trois premiers sont distincts. La valeur $h(d) \in \mathbb{P}^1(\mathbb{C})$, où $h$ est l'homographie de la question précédente, est appelée le* birapport *de $a, b, c, d$ et notée $[a, b, c, d]$.*

*(2) Soit $z \in \mathbb{P}^1(\mathbb{C})$. Calculez $[0, 1, \infty, z]$.*

*(3) Montrez que le birapport est invariant par homographie, c'est-à-dire que si $f$ est une homographie, alors $[f(a), f(b), f(c), f(d)] = [a, b, c, d]$. (*Considérez l'homographie $hf^{-1}$ où $h$ est l'homographie de la première question.*)*

*(4) Donnez une formule pour le birapport $[a, b, c, d]$.*

**Corrigé.** (1) Une homographie qui possède $a$ pour 0 et $c$ comme pôle est de la forme $h(z) = \lambda\frac{z-a}{z-c}$. Si on veut de plus $h(b) = 1$ alors nécessairement $1 = \lambda\frac{b-a}{b-c}$ donc $\lambda = \frac{b-c}{b-a}$. Finalement l'homographie $h$, si elle existe, est uniquement déterminée par $h(z) = \frac{b-c}{b-a}\frac{z-a}{z-c}$, et cette homographie envoie $a, b, c$ sur $0, 1, \infty$ comme demandé, donc elle convient.
(2) L'unique homographie $h$ qui envoie $0, 1, \infty$ sur $0, 1, \infty$ est l'identité, donc $[0, 1, \infty, z] = h(z) = z$.
(3) L'homographie $g = hf^{-1}$ envoie $f(a)$ sur $h(a) = 0$, $f(b)$ sur $h(b) = 1$ et $f(c)$ sur $h(c) = \infty$. Donc par définition du birapport, on a $[f(a), f(b), f(c), z] = g(z)$ pour tout $z$, en particulier $[f(a), f(b), f(c), f(d)] = (hf^{-1})(f(d)) = h(d) = [a, b, c, d]$.
(4) On l'a déjà vue dans la première question : l'unique homographie qui envoie $a, b, c$ sur $0, 1, \infty$ est $h(z) = \frac{b-c}{b-a}\frac{z-a}{z-c}$ donc $[a, b, c, d] = h(d) = \frac{b-c}{b-a}\frac{d-a}{d-c}$.

**Exercice 3** *(Symétries du birapport.) Soient $a, b, c, d$ quatre éléments de $\mathbb{P}^1(\mathbb{C})$ dont les trois premiers sont distincts.*
*(1) Montrez que $[c, d, a, b] = [a, b, c, d]$. (*Considérez l'homographie $\frac{[a,b,c,d]}{h}$ où $h$ est l'unique homographie qui envoie $a, b, c$ sur $0, 1, \infty$.*)*

*(2) Montrez que $[a, d, c, b] = [a, b, c, d]^{-1}$. (*Considérez l'homographie $\frac{1}{[a,b,c,d]}\,h$.*)*

*(3) Montrez que $[b, a, c, d] = 1 - [a, b, c, d]$. (*Considérez l'homographie...*)*

**Corrigé.** (1) L'homographie $g = \frac{[a,b,c,d]}{h}$ envoie $c$ sur 0, $d$ sur 1 et $a$ sur $\infty$ donc $[c, d, a, b] = g(b) = \frac{[a,b,c,d]}{h(b)} = [a, b, c, d]$ puisque $h(b) = 1$.
(2) L'homographie $g = \frac{1}{[a,b,c,d]}\,h$ envoie $a$ sur 0, $d$ sur 1 et $c$ sur $\infty$ donc $[a, d, c, b] = g(b) = \frac{1}{[a,b,c,d]}$.

(3) L'homographie $g = 1 - h$ envoie $b$ sur 0, $a$ sur 1 et $c$ sur $\infty$ donc $[b, a, c, d] = g(d) = 1 - h(d) = 1 - [a, b, c, d]$.

**Exercice 4** *Montrez qu'une application $f : \mathbb{P}^1(\mathbb{C}) \to \mathbb{P}^1(\mathbb{C})$ qui laisse invariant le birapport de quatre points est une homographie.*

**Corrigé.** L'hypothèse signifie que pour tout quadruplet $(a, b, c, d)$ de complexes dont les trois premiers sont distincts, alors $f(a)$, $f(b)$, $f(c)$ sont distincts et

$$[f(a), f(b), f(c), f(d)] = [a, b, c, d] \ .$$

En particulier, pour tout $z \in \mathbb{P}^1(\mathbb{C})$ on doit avoir

$$[f(0), f(1), f(\infty), f(z)] = [0, 1, \infty, z] = z \ .$$

On sait qu'il existe une unique homographie $h$ qui envoie 0 sur $f(0)$, 1 sur $f(1)$ et $\infty$ sur $f(\infty)$. Cette homographie préserve le birapport, d'après la question (3) de l'exercice précédent, donc

$$\begin{aligned} z &= [f(0), f(1), f(\infty), f(z)] = [h(0), h(1), h(\infty), f(z)] \\ &= [h(0), h(1), h(\infty), h(h^{-1}(f(z)))] = [0, 1, \infty, h^{-1}(f(z))] \\ &= h^{-1}(f(z)) \ . \end{aligned}$$

En composant par $h$ on trouve $h(z) = f(z)$ pour tout $z$, donc $f$ est égale à l'homographie $h$.

**Exercice 5** *(Cercles de $\mathbb{P}^1(\mathbb{C})$.) Dans $\mathbb{P}^1(\mathbb{C}) = \mathbb{C} \cup \{\infty\}$, la réunion d'une droite de $\mathbb{C}$ et du point $\infty$ est appelée un* cercle de $\mathbb{P}^1(\mathbb{C})$ passant par $\infty$. *La famille des cercles de $\mathbb{P}^1(\mathbb{C})$ est donc constituée des cercles passant par $\infty$, que l'on vient de définir, et des cercles ne passant pas par l'infini, qui sont les cercles ordinaires dans $\mathbb{C}$. Cette terminologie est justifiée par le fait que la projection stéréographique envoie les cercles tracés sur la sphère $S^2$ sur les cercles de $\mathbb{P}^1(\mathbb{C})$ au sens où l'on vient de les définir. Dans cet exercice, on souhaite démontrer que les homographies préservent la famille des cercles de $\mathbb{P}^1(\mathbb{C})$, et en application, donner une condition de cocyclicité.*

*(1) On veut montrer que $h(\mathcal{C})$ est un cercle de $\mathbb{P}^1(\mathbb{C})$, pour tout cercle $\mathcal{C}$ de $\mathbb{P}^1(\mathbb{C})$ et pour toute homographie $h$. Montrez qu'il suffit d'établir ce résultat pour $h(z) = 1/z$, ce que l'on supposera dans la suite.*
*(2) Montrez qu'on peut se ramener au cas où $\mathcal{C}$ est symétrique par rapport à l'axe réel. (*Poser $g(z) = e^{i\theta}z$ et calculer $ghg$.*)*
*(3) Montrez que $h(\mathcal{C})$ est un cercle, en distinguant quatre cas :*

$$\begin{aligned} &\infty \notin \mathcal{C} \text{ et } 0 \in \mathcal{C} \quad ; \quad \infty \in \mathcal{C} \text{ et } 0 \notin \mathcal{C} \quad ; \\ &\infty \notin \mathcal{C} \text{ et } 0 \notin \mathcal{C} \quad ; \quad \infty \in \mathcal{C} \text{ et } 0 \in \mathcal{C} \quad . \end{aligned}$$

*(*Méfiez-vous car le centre de $h(\mathcal{C})$ n'est pas forcément l'image par $h$ du centre de $\mathcal{C}$.*)*
*(4) Montrez que quatre nombres complexes $a, b, c, d$ sont cocycliques ou alignés ssi leur birapport est réel.*

**Corrigé.** (1) On sait que les similitudes directes envoient les cercles de $\mathbb{C}$ sur des cercles, et les droites sur des droites, donc globalement elles envoient les cercles de $\mathbb{P}^1(\mathbb{C})$ sur des cercles de $\mathbb{P}^1(\mathbb{C})$. Comme le groupe des homographies est engendré par les similitudes directes et l'homographie $h(z) = 1/z$, si cette dernière préserve la famille des cercles de $\mathbb{P}^1(\mathbb{C})$, il en sera de même de toute homographie.

(2) Le calcul direct donne $(ghg)(z) = h(z)$. On peut toujours choisir $\theta$ de telle sorte que $\mathcal{C}' = g(\mathcal{C})$ soit symétrique par rapport à l'axe réel : si $\mathcal{C}$ est un cercle passant par l'infini, c'est-à-dire une droite de $\mathbb{C}$, on la fait tourner de façon à la rendre orthogonale à l'axe réel, et si $\mathcal{C}$ est un cercle ne passant pas par l'infini, c'est-à-dire un cercle de $\mathbb{C}$, on le fait tourner de façon à mettre son centre sur l'axe réel. Notre calcul initial dit que $h(\mathcal{C}) = g(h(\mathcal{C}'))$. Donc si on sait démontrer que $h(\mathcal{C}')$ est un cercle, alors comme $g$ est une similitude directe, $g(h(\mathcal{C}'))$ est aussi un cercle, donc $h(\mathcal{C})$ est un cercle. Donc il suffit de démontrer le résultat demandé pour $\mathcal{C}'$, en d'autres termes on peut supposer que $\mathcal{C}$ est symétrique par rapport à l'axe réel. De plus, si $\mathcal{C}$ est un cercle on peut supposer que son centre est sur le demi-axe $\mathbb{R}^+$ et si c'est une droite on peut supposer qu'elle coupe le demi-axe $\mathbb{R}^+$.

(3) On traite les quatre cas successivement.

$\boxed{\infty \notin \mathcal{C} \text{ et } 0 \in \mathcal{C}}$ ici $\mathcal{C}$ est un cercle dont on note $a \in \mathbb{R}^+$ l'affixe du centre et $r$ le rayon. L'hypothèse $0 \in \mathcal{C}$ signifie que $a = r$. Les points de $\mathcal{C}$ sont de la forme $z = r + re^{i\theta}$ et on a

$$\begin{aligned} h(z) &= \tfrac{1}{r(1+e^{i\theta})} = \tfrac{1}{2r\cos(\theta/2)e^{i\theta/2}} \\ &= \tfrac{1}{2r\cos(\theta/2)} e^{-i\theta/2} = \tfrac{1}{2r}(1 - i\tan(\theta/2)) \ . \end{aligned}$$

La quantité $\tan(\theta/2)$ décrit $\mathbb{R}$, et les points $h(z)$ décrivent la droite d'équation $\Re(z) = 1/2r$.

$\boxed{\infty \in \mathcal{C} \text{ et } 0 \notin \mathcal{C}}$ ici $\mathcal{C}$ est une droite ne passant pas par l'origine, donc d'équation $\Re(z) = t$ pour un $t \in \mathbb{R}^+$. Notons $\mathcal{C}'$ le cercle de centre $a = 1/2t$ et de rayon $a = 1/2t$. Le cas précédent montre que $h(\mathcal{C}')$ est la droite d'équation $\Re(z) = t$, donc $h(\mathcal{C}') = \mathcal{C}$, donc $h(\mathcal{C}) = \mathcal{C}'$. (Noter que $h^2$ est l'identité !)

$\boxed{\infty \notin \mathcal{C} \text{ et } 0 \notin \mathcal{C}}$ ici $\mathcal{C}$ est un cercle dont on note $a \in \mathbb{R}^+$ l'affixe du centre et $r$ le rayon, avec $a \neq r$. On doit préparer le calcul par une petite réflexion. L'image de $\mathcal{C}$ va être un cercle, et la difficulté principale est de trouver son centre. Pour cela on regarde l'image du diamètre situé sur l'axe réel, qui est le segment $[a - r; a + r]$. On a $h(a - r) = 1/(a - r)$ et $h(a + r) = 1/(a + r)$, et le segment $[h(a - r); h(a + r)]$ a pour milieu :

$$\alpha := \frac{\frac{1}{a-r} + \frac{1}{a+r}}{2} = \frac{a}{a^2 - r^2} \ .$$

Les points de $\mathcal{C}$ sont de la forme $z = a + re^{i\theta}$. Pour vérifier que $h(z)$ est sur un cercle de centre $\alpha$, on calcule

$$\begin{aligned} \left|\frac{1}{z} - \frac{a}{a^2 - r^2}\right| &= \left|\frac{1}{a + re^{i\theta}} - \frac{a}{a^2 - r^2}\right| = \left|\frac{\cancel{a^2} - r^2 - \cancel{a^2} - are^{i\theta}}{(a + re^{i\theta})(a^2 - r^2)}\right| \\ &= \frac{r}{|a^2 - r^2|} \left|\frac{r + ae^{i\theta}}{a + re^{i\theta}}\right| = \frac{r}{|a^2 - r^2|} \left|e^{i\theta} \frac{a + re^{-i\theta}}{a + re^{i\theta}}\right| \\ &= \tfrac{r}{|a^2 - r^2|} \end{aligned}$$

car $a + re^{-i\theta}$ et $a + re^{i\theta}$, étant complexes conjugués, ont même module. On a démontré que $h(\mathcal{C})$ est le cercle de centre $a/(a^2 - r^2)$ et de rayon $r/|a^2 - r^2|$.

$\boxed{\infty \in \mathcal{C} \text{ et } 0 \in \mathcal{C}}$ ici $\mathcal{C}$ est la réunion de l'axe des imaginaires purs et de $\{\infty\}$. Ses éléments s'écrivent $z = it$ pour $t \in \mathbb{R}$ (sauf $\infty$). On a $h(z) = 1/(it) = (-1/t)i$ et $h(\infty) = 0$. Ces points décrivent l'axe imaginaire pur donc $h(\mathcal{C}) = \mathcal{C}$.

Dans chacun des quatre cas considérés, $h(\mathcal{C})$ est un cercle (au sens des cercles de $\mathbb{P}^1(\mathbb{C})$).

(4) L'énoncé est un peu imprécis car si deux ou plus des points sont confondus, le birapport n'est pas défini. Mais dans ce cas, les points sont cocycliques (au fait, pourquoi ?), donc on laisse de côté ce cas particulier. L'observation clé est que trois points distincts $a, b, c$ sont sur un cercle $\mathcal{C}$ de $\mathbb{P}^1(\mathbb{C})$ et un seul : s'ils sont alignés, $\mathcal{C}$ est la droite en question, et sinon, $\mathcal{C}$ est le cercle circonscrit au triangle $abc$. Alors, $a, b, c, d$ sont cocycliques ou alignés ssi $d \in \mathcal{C}$.

Appelons $h$ l'homographie qui envoie $a, b, c$ sur $0, 1, \infty$, de sorte que $h(d) = [a, b, c, d]$. Alors $d \in \mathcal{C}$ ssi $h(d) \in h(\mathcal{C})$. Or $h(\mathcal{C})$ est un cercle de $\mathbb{P}^1(\mathbb{C})$ d'après les questions précédentes, contenant $0, 1, \infty$, donc c'est en fait l'axe réel. On a donc trouvé :

$a, b, c, d$ cocycliques ou alignés ssi $[a, b, c, d] = h(d) \in h(\mathcal{C}) = \mathbb{R}$.

## Matrices réelles qui sont des exponentielles

La décomposition $D+N$ usuelle a un analogue multiplicatif, pour les matrices inversibles, qui est une décomposition « diagonalisable $\times$ unipotent ». Rappelons qu'une matrice $U$ est dite *unipotente* si sa seule valeur propre est 1, ou dit autrement, si $U-\mathrm{Id}$ est nilpotente. L'*indice d'unipotence* de $U$ est défini comme étant égal à l'indice de nilpotence de $U-\mathrm{Id}$. Voici en exercice cette décomposition :

**Exercice 1 :** Montrez que pour tout $G \in \mathrm{GL}_n(\mathbb{C})$, il existe un couple unique $(D,U)$ tel que $G=DU$ avec $D$ diagonalisable, $U$ unipotent, et $DU=UD$.

**Exercice 2 :** On dit qu'une matrice $M \in \mathrm{M}_n(\mathbb{R})$ est *semi-simple* si elle est diagonalisable sur $\mathbb{C}$. (Plus généralement une matrice carrée à coefficients dans un corps $k$ est dite semi-simple si elle est diagonalisable sur une clôture algébrique de $k$).

(1) Soit $M \in \mathrm{M}_n(\mathbb{R})$, montrez qu'il existe un unique couple $(S,N)$ composé d'une matrice semi-simple $S$, une matrice nilpotente $N$, telles que $M=S+N$ et $SN=NS$.

*Indication : commencez par écrire la décomposition $D+N$ de $M$ dans $\mathbb{C}$.*

(2) Soit $M \in \mathrm{GL}_n(\mathbb{R})$, montrez qu'il existe un unique couple $(S,U) \in \mathrm{M}_n(\mathbb{R})^2$ composé d'une matrice semi-simple $S$, une matrice unipotente $U$, telles que $M=SU$ et $SU=US$.

(*Indication : commencez par écrire la décomposition $S+N$ de $M$.*)

**Exercice 3 :** Soit $S$ une matrice semi-simple réelle, et $S=PDP^{-1}$ avec $D$ diagonale complexe. Montrez que $D$ et la matrice conjuguée $\overline{D}$ sont semblables. Quelle matrice de similitude $M$ peut-on choisir ? Montrez que $S$ est réelle si et seulement si $P^{-1}\overline{P}M^{-1}$ est diagonale par blocs, les blocs ayant des tailles que l'on précisera.

**Exercice 4 :** La résolution de cet exercice nécessite d'avoir fait le précédent. On veut montrer que $X \in \mathrm{GL}_n(\mathbb{R})$ est l'exponentielle d'une matrice de $\mathrm{M}_n(\mathbb{R})$ si et seulement s'il existe une matrice $Y \in \mathrm{GL}_n(\mathbb{R})$ telle que $X=Y^2$.

(1) Montrez que si $X$ est l'exponentielle d'une matrice de $\mathrm{M}_n(\mathbb{R})$ alors c'est le carré d'une matrice de $\mathrm{GL}_n(\mathbb{R})$.

Réciproquement, on suppose que $X=Y^2$ et on écrit la décomposition $Y=SU$ en produit commutatif semi-simple par unipotent.

(2) La matrice $U^2$ est-elle l'exponentielle d'une matrice réelle $N$ ?

(3) Utilisez l'exercice précédent pour montrer que $S^2$ est l'exponentielle d'une matrice semi-simple $T$.
(*Indic : On écrit $S=PDP^{-1}$. Définissez $\log(D^2)$ puis montrez que $T=P(\log(D^2))P^{-1}$ est réelle. Attention aux valeurs propres qui ont même carré !*)

(4) Montrez que $T$ et $N=\log(U^2)$ commutent. Concluez.

# Corrigé

**Exercice 1 :** On écrit la décomposition $G = D+N$ avec $DN = ND$. Comme $G$ est inversible, ses valeurs propres qui sont aussi celles de $D$ sont non nulles, donc $D$ est inversible. Alors $G = D + N = D(\mathrm{Id} + D^{-1}N)$. Comme $D$ et $N$ commutent, $D^{-1}N$ est nilpotente donc $U := \mathrm{Id} + D^{-1}N$ est unipotente. Il est facile de vérifier que $DU = UD$. Noter qu'on retrouve la décomposition de départ grâce à $N = D(U - \mathrm{Id})$, ce qui permet de montrer l'unicité (faites-le).

**Exercice 2 :**

**Exercice 3 :** En séparant les valeurs propres réelles $\lambda_j$ et les valeurs propres complexes conjuguées $z_k$, $\overline{z}_k$, on a :

$$D = \mathrm{diag}\left((\lambda_1, a_1), \ldots, (\lambda_r, a_r), (z_1, b_1), (\overline{z}_1, b_1), \ldots, (z_s, b_s), (\overline{z}_s, b_s)\right)$$

C'est une matrice diagonale par blocs, les blocs étant des homothéties de rapports tous distincts. Leurs tailles sont données par les multiplicités : $a_1, \ldots, a_r, b_1, b_1, \ldots, b_s, b_s$ ce que l'on note en abrégé $(\underline{a}, \underline{b}, \underline{b})$. Pour conjuguer on permute juste les valeurs propres complexes :

$$D = \mathrm{diag}\left((\lambda_1, a_1), \ldots, (\lambda_r, a_r), (\overline{z}_1, b_1), (z_1, b_1), \ldots, (\overline{z}_s, b_s), (z_s, b_s)\right)$$

Notons $\sigma$ la permutation qui est l'identité sur les $r$ premiers entiers et qui échange les suivants par paires, et $M = M_\sigma$ la matrice de permutation associée. On a donc $\overline{D} = M^{-1}DM$. Or $S$ est réelle ssi $\overline{S} = S$ càd $\overline{P}\,\overline{D}\,\overline{P}^{-1} = PDP^{-1}$. Comme $\overline{D} = M^{-1}DM$ cela donne

$$\overline{P}M^{-1}DM\overline{P}^{-1} = PDP^{-1}$$

c'est-à-dire que $P^{-1}\overline{P}M^{-1}$ commute avec $D$. Vue la forme de $D$, ceci équivaut à dire que $P^{-1}\overline{P}M^{-1}$ est diagonale par blocs de type $(\underline{a}, \underline{b}, \underline{b})$.

**Exercice 4 :** (1) Si $X = \exp(Z)$ alors $X = \exp(\frac{1}{2}Z + \frac{1}{2}Z) = (\exp(\frac{1}{2}Z))^2$.

(2) On sait que l'exponentielle réalise un homéomorphisme entre l'ensemble des matrices nilpotentes et l'ensemble des matrices unipotentes. Donc il existe une unique matrice nilpotente $N$ telle que $U = \exp(N)$. La question qui se pose est de savoir si $N$ est réelle. En fait, l'homéomorphisme inverse est explicite et montre bien que $N$ est réelle :

$$N = \log(U) = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k} (U - \mathrm{Id})^k$$

(3) Soit $S = PDP^{-1}$ et $(\lambda_j, a_j)$ et $(z_k, b_k)$, $(\overline{z}_k, b_k)$ les valeurs propres pondérées. D'après l'exercice précédent, il existe une matrice de permutation $M$ (adaptée aux multiplicités) telle que $P^{-1}\overline{P}M^{-1}$ commute avec $D$, i.e. est diagonale par blocs de type $(\underline{a}, \underline{b}, \underline{b})$. On a :

$$D^2 = \mathrm{diag}\left((\lambda_1^2, a_1), \ldots, (\lambda_r^2, a_r), (z_1^2, b_1), (\overline{z}_1^2, b_1), \ldots, (z_s^2, b_s), (\overline{z}_s^2, b_s)\right)$$

Il faut prendre garde que l'écriture ci-dessus ne reflète plus forcément ni la séparation entre valeurs propres réelles et non réelles, ni les multiplicités (qu'il est important de compter pour identifier les matrices qui commutent). Notons $z_k = \rho_k e^{i\theta_k}$ avec $\theta_k \in ]0, \pi[ \cup ]\pi, 2\pi[$. Définissons la matrice diagonale

$$\log(D^2) = \mathrm{diag}\Big((\log\lambda_1^2, a_1), \ldots, (\log\lambda_r^2, a_r), (\log\rho_1^2 + 2i\theta_1, b_1), (\log\rho_1^2 - 2i\theta_1, b_1), \ldots\Big)\,.$$

L'existence éventuelle de deux valeurs propres opposées fait augmenter les multiplicités lors du passage au carré. Quoi qu'il en soit, $\log(D^2)$ est encore diagonale par blocs homothétiques selon le découpage $(\underline{a}, \underline{b}, \underline{b})$. (Le type $\underline{a}'$ est éventuellement plus grossier que $\underline{a}$.) Comme $P^{-1}\overline{P}M^{-1}$ est diagonale par blocs de type $(\underline{a}, \underline{b}, \underline{b})$ elle commute avec $\log(D^2)$. De plus $M$ échange les valeurs propres complexes de $\log(D^2)$ et leurs conjuguées, donc

$$\overline{\log(D^2)} = M^{-1}\log(D^2)M$$

Il s'ensuit que $P(\log(D^2))P^{-1}$ est réelle (cf exercice précédent). On pose $T = P(\log(D^2))P^{-1}$. On a

$$\exp(T) = P\exp(\log(D^2))P^{-1} = PD^2P^{-1} = S^2$$

(4) Dans la décomposition $Y = SU$, $U$ commute avec $S$. Donc $P^{-1}U^2P$ commute avec $P^{-1}S^2P = D^2$, i.e. elle est diagonale par blocs de type $(\underline{a}, \underline{b}, \underline{b})$. En prenant le logarithme comme dans la question (2) on obtient que $P^{-1}NP$ est diagonale par blocs de même type. Donc elle commute avec $P^{-1}(\log D^2)P$ c'est-à-dire avec $T$. Il s'ensuit que

$$\begin{aligned}\exp(T+N) &= \exp(T)\exp(N)\\ &= P\exp(\log D^2)P^{-1}\exp(N) = S^2U^2 = M\ .\end{aligned}$$

**Bibliographie :**


[Gou] Gourdon, Les Maths en tête, Mathématiques pour M', *Ellipses.*

[MT] Mneimné, Testard, Introduction à la théorie des groupes de Lie classiques, *Hermann.*


# Rotations et homographies

Le but de l'exercice qui suit est de décrire un morphisme injectif de groupes $\mathrm{SO}_3(\mathbb{R}) \hookrightarrow \mathrm{PGL}_2(\mathbb{C})$.

**Exercice 1** *On considère l'espace euclidien $E = \mathbb{R}^3$ identifié à $\mathbb{C}\oplus\mathbb{R}$, sa sphère unité $S^2$, et le plan équatorial $\mathbb{C}$. La* projection stéréographique *$\sigma : S^2 \to \mathbb{C}\cup\{\infty\} = \widehat{\mathbb{C}}$ est la projection depuis le pôle nord $N$, définie par $\sigma(M) = (NM)\cap\mathbb{C}$ et $\sigma(N) = \infty$.*

*(1) On note $(z,t) \in \mathbb{C}\oplus\mathbb{R}$ les coordonnées d'un point $M \in E$. Donnez l'expression de $\sigma$ et $\sigma^{-1}$ dans ces coordonnées.*

*(2) Le plan tangent en un point de $S^2$ étant orienté par la normale sortante en ce point, on note $f_\theta$ la rotation de $E$ d'axe $[ON)$ et d'angle $\theta$, et $g_\varphi$ la rotation de $E$ d'axe $[Ox)$ et d'angle $\varphi$. Justifiez que $\mathrm{SO}_3(\mathbb{R})$ est engendré par les rotations $f_\theta$ et $g_\varphi$.*

*(3) Une rotation $r \in \mathrm{SO}_3(\mathbb{R})$ induit une bijection $\widetilde{r} := \sigma^{-1}\circ r\circ\sigma$ de $\mathbb{P}^1(\mathbb{C})$. Calculez $\widetilde{f}_\theta$ et montrez que*

$$\widetilde{g}_\varphi(z) = -i\,\frac{z\cos(\frac{\varphi}{2}) + i\sin(\frac{\varphi}{2})}{z\sin(\frac{\varphi}{2}) - i\cos(\frac{\varphi}{2})}\ .$$

*Déduisez-en que pour tout $r \in \mathrm{SO}_3(\mathbb{R})$, la bijection $\widetilde{r}$ est une homographie. Montrez que $r \mapsto \widetilde{r}$ définit un morphisme de groupes injectif $\mathrm{SO}_3(\mathbb{R}) \hookrightarrow \mathrm{PGL}_2(\mathbb{C})$.*

**Corrigé.** (1) Cf feuille d'exercices sur la droite projective : on avait trouvé

$$\sigma(z,t) = \frac{z}{1-t} \quad \text{et} \quad \sigma^{-1}(z) = \left(\frac{2z}{|z|^2+1}, \frac{|z|^2-1}{|z|^2+1}\right)\ .$$

(2) Soit $r$ une rotation, $x$ un vecteur directeur de son axe, et $\psi$ son angle. Il existe $\theta$ et $\varphi$ tels que $x = (f_\theta g_\varphi)(N)$. Donc $(f_\theta g_\varphi)^{-1}\circ r\circ(f_\theta g_\varphi)$ est la rotation d'axe $[ON)$ et d'angle $\psi$, autrement dit, c'est $f_\psi$. Il s'ensuit que $r = (f_\theta g_\varphi)\circ f_\psi\circ(f_\theta g_\varphi)^{-1}$, ce qui démontre que $\mathrm{SO}_3(\mathbb{R})$ est engendré par les rotations $f_\theta$ et $g_\varphi$.

(3) On a $f_\theta(z,t) = (e^{i\theta}z, t)$ d'où $\widetilde{f}_\theta(z) = e^{i\theta}z$. C'est une homographie. Passons à $g_\varphi$ : matriciellement, on a

$$Mat(g_\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) \\ 0 & \sin(\varphi) & \cos(\varphi) \end{pmatrix}.$$

Soit $z = a + ib \in \mathbb{C}$, on a

$$g_\varphi(\sigma^{-1}(z)) = g_\varphi\left(\frac{2a}{a^2+b^2+1}, \frac{2b}{a^2+b^2+1}, \frac{a^2+b^2-1}{a^2+b^2+1}\right)$$

$$= \left(\frac{2a}{a^2+b^2+1}, \frac{2b\cos(\varphi) - (a^2+b^2-1)\sin(\varphi)}{a^2+b^2+1}, \frac{2b\sin(\varphi) + (a^2+b^2-1)\cos(\varphi)}{a^2+b^2+1}\right)$$

et on en déduit

$$\begin{aligned} \widetilde{g}_\varphi(z) &= \frac{2a + i(2b\cos(\varphi) - (a^2+b^2-1)\sin(\varphi)}{a^2+b^2+1-2b\sin(\varphi) - (a^2+b^2-1)\cos(\varphi)} \\ &= \frac{-2i\big(a\sin(\frac{\varphi}{2}) + i(-b\sin(\frac{\varphi}{2}) + \cos(\frac{\varphi}{2})\big)\big(a\cos(\frac{\varphi}{2}) + i(b\cos(\frac{\varphi}{2}) + \sin(\frac{\varphi}{2})\big)}{2\ |a\sin(\frac{\varphi}{2}) + i(b\sin(\frac{\varphi}{2}) - \cos(\frac{\varphi}{2})|^2} \\ &= -i\ \frac{z\cos(\frac{\varphi}{2}) + i\sin(\frac{\varphi}{2})}{z\sin(\frac{\varphi}{2}) - i\cos(\frac{\varphi}{2})}\ . \end{aligned}$$

Les rotations $f_\theta$ et $g_\varphi$ engendrent $\mathrm{SO}_3(\mathbb{R})$, elles induisent des homographies de $\mathbb{P}^1(\mathbb{C})$ via $\sigma$, donc toute rotation $r \in \mathrm{SO}_3(\mathbb{R})$ induit une homographie. Ceci donne un morphisme de groupes injectif $\mathrm{SO}_3(\mathbb{R}) \hookrightarrow \mathrm{PGL}_2(\mathbb{C})$.

**Exercice 2** À venir... Pour la construction de $\mathrm{SO}_3(\mathbb{R}) \hookrightarrow \mathrm{PGL}_2(\mathbb{C})$ via les quaternions, voir la note sur les sous-groupes finis de $\mathrm{PGL}_2(\mathbb{C})$.

**Corrigé.**

# Un sous-espace vectoriel de fonctions

*Soit $k$ un corps et $V$ un sous-$k$-espace vectoriel de dimension finie de l'espace des fonctions de $k$ dans $k$. Montrez qu'il existe $x_1, \dots, x_n$ dans $k$ tels que l'application $f \mapsto (f(x_1), \dots, f(x_n))$ réalise un isomorphisme $V \simeq k^n$.*

Pour $x \in k$, notons $e_x$ la forme linéaire d'évaluation des fonctions en $x$. Il suffit de montrer que les $e_x$ engendrent $V^*$, car alors on pourra en extraire une base $e_{x_1}, \dots, e_{x_n}$ de $V^*$ et ceci répondra à la question. (Pourquoi ?)

Il s'agit de montrer que pour toute forme $\varphi \in V^*$, il existe $x_1, \dots, x_n$ et $\lambda_1, \dots, \lambda_n$ tels que $\varphi = \sum_{j=1}^n \lambda_j e_{x_j}$. Si on fixe une base $f_1, \dots, f_n$ de $V$, cette égalité est équivalente aux $n$ égalités $\varphi(f_i) = \sum_{j=1}^n \lambda_j f_i(x_j)$. À leur tour, ces $n$ égalités sont équivalentes à l'égalité matricielle $\Phi = A\Lambda$ où $\Phi$ est le vecteur colonne des $\varphi(f_j)$, $A$ est la matrice carrée des $f_i(x_j)$, et $\Lambda$ est le vecteur colonne des $\lambda_j$. Il suffit donc de trouver $x_1, \dots, x_n$ tels que $A$ est inversible, car alors les $\lambda_j$ seront uniquement déterminés et le problème sera résolu.

Montrons par récurrence que pour tout $k \leqslant n$ il existe $x_1, \dots, x_k$ tels que la matrice carrée de taille $k$ en haut à gauche de $A$ est inversible. Pour $k = 1$ c'est clair ; supposons l'hypothèse vraie au rang $k-1$. Si pour tout $x_k \in k$ le déterminant de taille $k$ est nul, on développe ce déterminant par rapport à la dernière colonne, il vient :

$$m_{k,k} f_k(x_k) - m_{k-1,k} f_{k-1}(x_k) + \cdots + (-1)^{k-1} m_{1,k} f_1(x_k) = 0 \,,$$

où $m_{i,j}$ est le mineur adéquat. Comme $m_{k,k} \neq 0$ par hypothèse, ceci montre que $f_k$ est combinaison linéaire de $f_1, \dots, f_{k-1}$, ce qui est impossible puisque $f_1, \dots, f_n$ est une base de $V$. Par contraposée, il existe $x_k$ tel que la matrice carrée de taille $k$ en haut à gauche de $A$ est inversible.

Pour $k = n$, ceci montre que $A$ est inversible. On peut donc poser $\Lambda = A^{-1}\Phi$ et on a alors $\varphi = \sum_{j=1}^n \lambda_j e_{x_j}$.

# Sous-groupes finis de $\mathrm{PGL}_2(\mathbb{C})$

Dans cette note j'énonce (sans démonstration, mais avec une référence) le théorème de Dickson sur les sous-groupes finis de $\mathrm{PGL}_2(\mathbb{C})$, et j'explique (avec démonstration, mais sans référence) comment identifier précisément ces sous-groupes en donnant des générateurs sous forme d'homographies. Les leçons concernées sont :

> Groupes finis. Exemples et applications.
> Groupe linéaire d'un espace vectoriel de dimension finie $E$, sous-groupes de $\mathrm{GL}(E)$. Applications.
> Sous-groupes finis de $\mathrm{O}(2,\mathbb{R})$, de $\mathrm{O}(3,\mathbb{R})$. Applications.
> Exemples de parties génératrices d'un groupe.
> Corps des fractions rationnelles à une indéterminée sur un corps commutatif. Applications.
> Homographies de la droite complexe. Applications.

*Notations.* Le groupe $\mathrm{PGL}_2(\mathbb{C})$ est le quotient de $\mathrm{GL}_2(\mathbb{C})$ par le sous-groupe des homo-théties. On désignera par la même lettre une matrice inversible $(2,2)$ et son image dans $\mathrm{PGL}_2(\mathbb{C})$. Nous notons $\mathbb{D}_n$ le groupe diédral d'ordre $2n$.

## 1 Le théorème de Dickson

### 1.1 Énoncé

Rappelons que $\mathrm{PGL}_2(\mathbb{C})$ est un groupe d'une grande importance géométrique car il s'identifie au groupe des homographies de la droite projective $\mathbb{P}^1(\mathbb{C})$ via l'application

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \left( x \mapsto \frac{ax+b}{cx+d} \right)$$

Il se trouve qu'il a exactement les mêmes sous-groupes finis qu'un autre groupe important en géométrie, à savoir le groupe $\mathrm{SO}_3(\mathbb{R})$ des isométries positives de l'espace euclidien de dimension trois. Ce n'est pas un hasard, comme nous l'expliquerons ci-dessous. Le théorème de Dickson ([Suzuki], theorem 6.17 du chapter 3, § 2) est le suivant :

**Théorème 1** *Les sous-groupes finis de* $\mathrm{PGL}_2(\mathbb{C})$ *sont isomorphes à l'un des 5 groupes suivants : un groupe cyclique, un groupe diédral, ou l'un des groupes* $\mathfrak{A}_4$, $\mathfrak{S}_4$, $\mathfrak{A}_5$.

**Remarque 4** Voici comment lire cet énoncé dans le livre de Suzuki - mais lisez ensuite la remarque 2 ci-dessous. Le théorème 6.17 de [Suzuki] concerne les sous-groupes de $\mathrm{SL}_2(k)$ avec $k$ algébriquement clos de caractéristique quelconque. La démonstration est longue et assez technique et je ne vous conseille pas de la lire. Pour retrouver l'énoncé ci-dessus, notez que si $k$ est algébriquement clos on a $\mathrm{PSL}_2(k) \simeq \mathrm{PGL}_2(k)$ (exercice), donc il suffit de lire la partie I du théorème (caractéristique 0) et de quotienter par $\{\pm 1\}$ pour obtenir les sous-groupes de $\mathrm{PSL}_2(k)$. On trouve, dans l'ordre donné par Suzuki : les groupes cycliques, les groupes diédraux, puis $\mathfrak{A}_4$ (vu que $\mathrm{PSL}(2,\mathbb{F}_3) \simeq \mathfrak{A}_4$), $\mathfrak{S}_4$ (en effet le quotient du groupe $\widetilde{\Sigma}_4$ donné par Suzuki est $\mathfrak{S}_4$) et $\mathfrak{A}_5$ (vu que $\mathrm{PSL}(2,\mathbb{F}_5) \simeq \mathfrak{A}_5$).

**Remarque 5** Vous n'avez pas besoin de la remarque 1, car il vous sera très facile de retenir le théorème puisque vous connaissez bien la liste des sous-groupes finis de $\mathrm{SO}_3(\mathbb{R})$. Si on vous demande une référence, vous n'aurez qu'à citer le livre de Suzuki... on ne vous en demandera pas plus.

## 1.2 Lien avec $\mathrm{SO}_3(\mathbb{R})$

Ce paragraphe est inclus pour assouvir votre curiosité, mais il n'est pas nécessaire pour lire la partie suivante.

Le lien entre sous-groupes finis de $\mathrm{SO}_3(\mathbb{R})$ et de $\mathrm{PGL}_2(\mathbb{C})$ a une explication simple :

**Affirmation :** *Il existe une injection* $\mathrm{SO}_3(\mathbb{R}) \hookrightarrow \mathrm{PGL}_2(\mathbb{C})$.

Admettant cela, le théorème de Dickson nous redonne les seuls sous-groupes finis possibles de $\mathrm{SO}_3(\mathbb{R})$. En effet, si $H$ est un sous-groupe fini de $\mathrm{SO}_3(\mathbb{R})$, c'est un sous-groupe de $\mathrm{PGL}_2(\mathbb{C})$ via notre injection, et d'après le théorème de Dickson il est isomorphe à l'un des 5 que l'on sait.

Nous allons maintenant construire cette injection à l'aide de l'algèbre des quaternions $\mathbb{H}$ (j'utilise [Perrin], chapitre VII comme référence sur $\mathbb{H}$). Rappelons que $\mathbb{H}$ est la $\mathbb{R}$-algèbre composée des éléments $a + bi + cj + dk$ où $a, b, c, d \in \mathbb{R}$, avec les règles de multiplication bien connues. Le groupe $G \subset \mathbb{H}$ des quaternions de norme 1 est homéomorphe à la sphère $S^3$.

**Preuve :** Il y a deux points importants :

(1) On a une injection $\mathbb{H} \hookrightarrow \mathrm{M}_2(\mathbb{C})$. En effet $\mathbb{H}$ a une structure de $\mathbb{C}$-ev de dimension 2 : on peut voir $\mathbb{C}$ comme sous-corps de $\mathbb{H}$ comme l'ensemble des $a + bi$, et puisque $k = ij$ dans $\mathbb{H}$, on peut écrire $a + bi + cj + dk = (a + bi) + (c + di)j$ de sorte que $\{1, j\}$ est une base de $\mathbb{H}$ comme $\mathbb{C}$-ev. (Attention : la loi extérieure est ici $(\lambda, q) = \lambda q$ alors que Perrin (chap. VII, § 4) utilise la loi extérieure $(\lambda, q) = q\lambda$.) Si on fixe un quaternion $q$, il est facile devoir que la multiplication à droite par $q$, $D_q(x) = xq$ est $\mathbb{C}$-linéaire. (Ici, Perrin considère la multiplication à gauche à la place.) On obtient donc un morphisme injectif $\mathbb{H} \to \mathrm{End}_{\mathbb{C}\text{-ev}}(\mathbb{H}) \simeq \mathrm{M}_2(\mathbb{C})$, qui à $q$ associe $D_q$.

(2) On en déduit un morphisme $f\colon G \to \mathrm{PGL}_2(\mathbb{C})$. En effet on peut faire agir un élément $g \in G$ par la conjugaison par son image dans $\mathrm{M}_2(\mathbb{C})$. Il s'agit d'une action par automorphismes de $\mathbb{C}$-algèbre d'où un morphisme de groupes $G \to \mathrm{Aut}_{\mathbb{C}\text{-alg}}(\mathrm{M}_2(\mathbb{C}))$. Or on connaît ce groupe d'automorphismes : le morphisme $\mathrm{GL}_2(\mathbb{C}) \to \mathrm{Aut}_{\mathbb{C}\text{-alg}}(\mathrm{M}_2(\mathbb{C}))$ qui à $P$ associe la conjugaison par $P$ a pour noyau les homothéties, donc il se factorise en un morphisme injectif $\mathrm{PGL}_2(\mathbb{C}) \to \mathrm{Aut}_{\mathbb{C}\text{-alg}}(\mathrm{M}_2(\mathbb{C}))$. Ce morphisme est aussi surjectif (donc un isomorphisme) car on sait que les automorphismes de $\mathbb{C}$-algèbre de $\mathrm{M}_2(\mathbb{C})$ sont tous intérieurs (voir par exemple [FGN] p. 273).

On conclut en cherchant le noyau de $f$. Si $g \in \ker(f)$, il agit trivialement sur $\mathrm{M}_2(\mathbb{C})$ et en particulier sur la sous-algèbre $\mathbb{H}$. Or la conjugaison par $g$ est triviale sur $\mathbb{H}$ ssi $g$ est central, donc $g \in \mathbb{R}$, or $G \cap \mathbb{R} = \{\pm 1\}$. Donc $\ker(f) = \{\pm 1\}$, or on sait que $G/\{\pm 1\} \simeq \mathrm{SO}_3(\mathbb{R})$ : cela s'obtient en disant que la restriction de la conjugaison par $g \in G$ à l'ensemble des quaternions purs se fait par isométries, voir Perrin chap. VII, § 2. Donc $f\colon G \to \mathrm{PGL}_2(\mathbb{C})$ se factorise en l'injection qu'on voulait $\mathrm{SO}_3(\mathbb{R}) \hookrightarrow \mathrm{PGL}_2(\mathbb{C})$. ▫

**Remarque 6** Il y a une autre façon de voir cette injection. On peut identifier la sphère unité $S^2 \subset \mathbb{R}^3$ à la droite projective complexe (voir cours d'E. Halberstadt sur la droite projective). Les isométries de $\mathbb{R}^3$ agissent sur la sphère et donc sur $\mathbb{P}^1(\mathbb{C})$, et tout revient alors à vérifier que cette action se fait par homographies.

## 2 Expliciter les sous-groupes finis de $\mathrm{PGL}_2(\mathbb{C})$

On peut résumer les résultats qui vont suivre en un énoncé un peu imprécis :

**Théorème 2** *Soit $G$ l'un des 5 groupes $\mathbb{Z}/n\mathbb{Z}$, $\mathbb{D}_n$, $\mathfrak{A}_4$, $\mathfrak{S}_4$, $\mathfrak{A}_5$. Alors les sous-groupes de $\mathrm{PGL}_2(\mathbb{C})$ qui sont isomorphes à $G$ sont tous conjugués entre eux. En particulier ils sont tous conjugués à un sous-groupe de $\mathrm{SO}_3(\mathbb{R})$. De plus, on sait produire un sous-groupe explicite de chacun des 5 types.*

### 2.1 Un résultat clé

Notons $\mu_n^* \subset \mathbb{C}$ le groupe des racines primitives $n$-èmes de l'unité. On définit une relation d'équivalence dans $\mu_n^*$ par $\zeta \sim \zeta'$ ssi $\zeta' \in \{\zeta, \zeta^{-1}\}$. La classe de $\zeta$ est notée $[\zeta]$, et l'application $[\zeta] \mapsto \zeta + \zeta^{-1}$ définit une injection $\mu_n^*/\sim \hookrightarrow \mathbb{C}$.

La clé de la démonstration du théorème 2 est le résultat suivant.

**Proposition 1** *Soit $A \in \mathrm{PGL}_2(\mathbb{C})$ et $n > 1$. Alors LCSSE ([7]) :*

(i) *$A$ est d'ordre $n$.*

(ii) *Comme homographie de $\mathbb{P}^1(\mathbb{C})$, $A$ est conjugué à $x \mapsto \zeta x$, pour un $[\zeta] \in \mu_n^*/\sim$.*

(iii) *Il existe $[\zeta] \in \mu_n^*/\sim$ tel que $(\zeta + \zeta^{-1} + 2)\det(A) - \mathrm{tr}(A)^2 = 0$.*

*En particulier l'ensemble $\mu_n^*/\sim$ classifie les classes de conjugaison d'éléments d'ordre $n$.*

[7] Les conditions suivantes sont équivalentes

Pour justifier (ii), observez que $x \mapsto \zeta x$ et $x \mapsto \zeta^{-1}x$ sont conjuguées (par l'homographie $x \mapsto 1/x$), de sorte que le fait d'être conjugué à $x \mapsto \zeta x$ ne dépend pas de la classe de $\zeta$ pour la relation $\sim$. Quant à (iii), notez que $\det(A)$ et $\mathrm{tr}(A)$ ne sont pas bien définies car $A$ n'est que la classe d'une matrice modulo homothéties. Cependant, si on multiplie $A$ par un scalaire $\lambda \neq 0$, alors $\det(A)$ et $\mathrm{tr}(A)^2$ sont tous deux mutlipliés par $\lambda^2$, donc la condition (iii) a un sens indépendant du représentant matriciel choisi pour $A$.

**Preuve :** On travaille avec un représentant de $A$ dans $\mathrm{GL}_2(\mathbb{C})$. Le polynôme caractéristique est $\chi(X) = X^2 - \mathrm{tr}(A)X + \det(A)$ et les valeurs propres sont $\lambda^{\pm} = \frac{1}{2}(\mathrm{tr}(A) \pm \delta)$ où $\delta^2 = \mathrm{tr}(A)^2 - 4\det(A)$. Posons $\zeta_A := \lambda^+/\lambda^- \in \mathbb{C}$ de sorte que

$$\zeta_A + \zeta_A^{-1} = \frac{\mathrm{tr}(A) + \delta}{\mathrm{tr}(A) - \delta} + \frac{\mathrm{tr}(A) - \delta}{\mathrm{tr}(A) + \delta} = \frac{\mathrm{tr}(A)^2}{\det(A)} - 2$$

Il est clair que $A$ est d'ordre $n$ (dans $\mathrm{PGL}_2(\mathbb{C})$) ssi $\zeta_A = \zeta \in \mu_n^*$. Ceci équivaut à dire que $A$ est semblable à la matrice diagonale $(\zeta, 1)$, donc (i) $\Leftrightarrow$ (ii). Par ailleurs on a vu que (i) $\Rightarrow$ (iii). Montrons la réciproque. Sous la condition (iii) on a

$$\zeta_A + \zeta_A^{-1} = \frac{\mathrm{tr}(A)^2}{\det(A)} - 2 = \zeta + \zeta^{-1}$$

Ceci entraîne que $\zeta_A = \zeta$ ou $\zeta^{-1}$, d'où on déduit (i). □

Les cas particuliers $n = 2, 3, 4, 6$ sont les seuls cas où $\mu_n^*/\sim$ est réduit à un point : pour $n = 2$ c'est clair, et les autres cas sont ceux pour lesquels $\varphi(n) = 2$, donc $\zeta + \zeta^{-1}$ prend alors une seule valeur. (Les cas $n = 2$ et $n = 3$ seront fondamentaux dans la suite.)

**Corollaire 1** *$A$ est d'ordre $n = 2, 3, 4, 6$ ssi $\mathrm{tr}(A)^2 = i\det(A)$ pour $i = 0, 1, 2, 3$ resp.* □

**Preuve :** On utilise le (iii) de la proposition 1. Les racines primitives de l'unité correspondantes sont $-1$, $j = e^{2i\pi/3}$, $i$ et $u = e^{i\pi/3}$. Les polynômes minimaux de $j$ et $u$ sur $\mathbb{Q}$ sont les polynômes cyclotomiques $\Phi_3 = (X-j)(X-\bar{j}) = X^2+X+1$ et $\Phi_6 = (X-u)(X-\bar{u}) =$

$X^2-X+1$. On en déduit les valeurs correspondantes pour $\zeta+\zeta^{-1}+2$. □

**Corollaire 2** *Soit $A \in \mathrm{PGL}_2(\mathbb{C})$ distinct de l'identité et d'ordre fini. Notons $A$ comme une homographie $A(x) = \frac{ax+b}{cx+d}$. Alors $A$ a deux points fixes sur $\mathbb{P}^1(\mathbb{C})$.*

**Preuve :** D'après le (iii) de la proposition 1, il existe une racine de l'unité $\zeta$ telle que $(\zeta + \zeta^{-1} + 2)\det(A) - \mathrm{tr}(A)^2 = 0$. Si $c = 0$ alors $a \neq d$ car $A$ est d'ordre fini, et alors $A$ a pour points fixes le point $\infty$ et le point $b/(d-a)$. Si $c \neq 0$, il n'y a pas de point fixe à l'infini, et l'équation donnant les points fixes est $cx^2 + (d-a)x - b = 0$. Le discriminant est $\Delta = \mathrm{tr}(A)^2 - 4\det(A) = (\zeta + \zeta^{-1} - 2)\det(A)$. Or $\zeta + \zeta^{-1} - 2 \neq 0$ pour toute racine $n$-ième de l'unité, donc il y a deux solutions distinctes. □

## 2.2 Démonstration du théorème de Dickson

Dans ce paragraphe *qui n'est pas terminé*, je voudrais démontrer le théorème de Dickson en suivant la stratégie de la preuve du théorème de classification des sous-groupes finis de $\mathrm{SO}_3(\mathbb{R})$.

Soit maintenant $G$ un sous-groupe de $\mathrm{PGL}_2(\mathbb{C})$, fini de cardinal $n$. Notons $\mathcal{F}$ l'ensemble des points fixes des éléments de $G$ distincts de l'identité : d'après le corollaire ci-dessus c'est un ensemble fini de cardinal compris entre 2 et $2(n-1)$. Le groupe $G$ agit sur $\mathcal{F}$ car si $x$ est un point fixe de $g \in G$, et $h \in G$, alors $h(x)$ est un point fixe de $hgh^{-1}$. D'après la formule de Burnside, le nombre d'orbites pour cette action est

$$k = \frac{1}{n}\sum_{g\in G} |\,\mathrm{Fix}(g)| = \frac{1}{n}(|\mathcal{F}| + 2(n-1))$$

Rappelons ensuite que $\mathrm{PGL}_2(\mathbb{C})$ est simplement 3-transitif sur $\mathbb{P}^1(\mathbb{C})$, ce qui veut dire qu'étant donnés deux triplets $(r,s,t)$ et $(r',s',t')$ d'éléments de $\mathbb{P}^1(\mathbb{C})$ avec $r,s,t$ tous distincts et $r',s',t'$ tous distincts, il existe un unique $F \in \mathrm{PGL}_2(\mathbb{C})$ tel que $F(r) = r'$, $F(s) = s'$, $F(t) = t'$. Maintenant soit... *(à suivre)*

## 2.3 Démonstration du théorème 2

On prouve le théorème 2 en cherchant un représentant explicite de chaque classe de conjugaison de sous-groupe fini dans $\mathrm{PGL}_2(\mathbb{C})$. Pour cela, on choisit des présentations des groupes finis en question par générateurs et relations, et on exprime ces relations dans $\mathrm{PGL}_2(\mathbb{C})$ à l'aide de la proposition 1 et du corollaire. Pour simplifier les écritures, nous noterons parfois les homographies sous la forme $\frac{ax+b}{cx+d}$.

**Proposition 2** *Les sous-groupes cycliques de $\mathrm{PGL}_2(\mathbb{C})$ sont tous conjugués à*

$$C_n := \langle\, \zeta x \,\rangle$$

Notez que ce sous-groupe ne dépend pas du choix de $\zeta$ qui apparaît. La preuve de cette proposition est simplement le (ii) de la proposition 1.

**Proposition 3** *Les sous-groupes diédraux de $\mathrm{PGL}_2(\mathbb{C})$ sont tous conjugués à*

$$D_n := \langle\, \zeta x, \frac{1}{x} \,\rangle$$

**Preuve :** Soit $\mathbb{D}_n = \langle\, r,s \,|\, r^n = s^2 = (rs)^2 = 1 \,\rangle$. Soit $\nu\colon \mathbb{D}_n \to \mathrm{PGL}_2(\mathbb{C})$ un morphisme injectif, $a = \nu(r)$ et $b = \nu(s)$. D'après la proposition 1(ii), quitte à appliquer une première conjugaison dans $\mathrm{PGL}_2(\mathbb{C})$, on peut supposer que $a(x) = \zeta^{-1}x$ pour un $\zeta \in \mu_n^*$. D'après le corollaire, comme $b$ est d'ordre 2 on peut écrire $b(x) = \frac{rx+s}{tx-r}$. En écrivant le fait que $ab = \nu(rs)$ est aussi d'ordre 2 on trouve que $r = 0$, donc $b(x) = \frac{s}{tx}$. Alors, la conjugaison par $\lambda/x$ (où $\lambda := \sqrt{s/t}$) envoie $a$ sur $x \mapsto \zeta x$ et $b$ sur $x \mapsto 1/x$. On a ainsi montré que l'image de $\nu$ est conjuguée au groupe $D_n = \langle\, \zeta x, 1/x \,\rangle$. □

**Proposition 4** *Les sous-groupes de $\mathrm{PGL}_2(\mathbb{C})$ isomorphes à $\mathfrak{A}_4$ sont tous conjugués à*

$$A_4 := \langle\, jx, \frac{x+2}{x-1} \,\rangle \simeq \langle\, (123), (12)(34) \,\rangle$$

*avec $j$ racine primitive troisième de l'unité.*

L'écriture ci-dessus signifie qu'on a choisi pour $\mathfrak{A}_4$ les générateurs (123) et (12)(34), et que l'homographie $x \mapsto jx$ correspond à (123), et $x \mapsto \frac{x+2}{x-1}$ correspond à (12)(34).

**Preuve :** Soit $\nu\colon \mathfrak{A}_4 \to \mathrm{PGL}_2(\mathbb{C})$ un morphisme injectif. Soit $a = \nu(123)$ et $b = \nu((12)(34))$. Écrivons $b(x) = \frac{rx+s}{tx-r}$. Quitte à appliquer une première conjugaison dans $\mathrm{PGL}_2(\mathbb{C})$, on peut supposer que $a(x) = jx$. En écrivant le fait que $ab = \nu(134)$ est d'ordre 3 on trouve $2r^2 = st$. Donc $t \neq 0$, car sinon on aurait $r = 0$ ce qui est impossible. Alors la conjugation par $\phi(x) = rx/t$ laisse $a$ inchangé et envoie $b$ sur $x \mapsto \frac{x+2}{x-1}$. Donc l'image de $\nu$ est conjuguée au groupe annoncé. On n'a utilisé que certaines relations, et on pourrait les utiliser toutes pour montrer que le groupe obtenu est en effet isomorphe à $\mathfrak{A}_4$. Cependant le fait d'admettre le théorème de Dickson nous assure qu'il existe des sous-groupes isomorphes à $\mathfrak{A}_4$, de sorte que le sous-groupe produit ci-dessus *doit* être isomorphe à $\mathfrak{A}_4$. Nous ferons de même ci-dessous. □

**Proposition 5** *Les sous-groupes de* $\mathrm{PGL}_2(\mathbb{C})$ *isomorphes à* $\mathfrak{S}_4$ *sont tous conjugués à*

$$S_4 := \langle\, ix, \frac{x+1}{x-1} \,\rangle \simeq \langle\, (1234), (12) \,\rangle$$

**Preuve :** Soit $\nu\colon \mathfrak{S}_4 \to \mathrm{PGL}_2(\mathbb{C})$ un morphisme injectif. Soit $a = \nu(1234)$ et $b = \nu(12)$, on a $ab = \nu(134)$. Écrivons $b(x) = \frac{rx+s}{tx-r}$. Quitte à appliquer une première conjugaison on peut supposer que $a(x) = ix$. En écrivant le fait que $ab$ est d'ordre 3 on trouve $r^2 = st$. La conjugation par $\phi(x) = rx/t$ laisse $a$ inchangé et envoie $b$ sur $x \mapsto \frac{x+1}{x-1}$. □

**Proposition 6** *Les sous-groupes de* $\mathrm{PGL}_2(\mathbb{C})$ *isomorphes à* $\mathfrak{A}_5$ *sont tous conjugués à*

$$A_5 := \langle\, \delta x, \frac{x+1}{\Delta x - 1} \,\rangle \simeq \langle\, (12345), (12)(34) \,\rangle$$

*où* $\delta$ *est une quelconque racine primitive cinquième de l'unité et* $\Delta := 1 - \delta - \delta^{-1}$.

**Preuve :** Soit $\nu\colon \mathfrak{A}_5 \to \mathrm{PGL}_2(\mathbb{C})$ un morphisme injectif. Soit $a = \nu(12345)$, $b = \nu((12)(34))$, on a $ab = \nu((12345)(12)(34)) = \nu(135)$. Écrivons $b(x) = \frac{rx+s}{tx-r}$. Il existe une racine primitive cinquième de l'unité $\delta$ telle qu'après une première conjugaison on ait $a(x) = \delta x$. On écrit que $ab$ est d'ordre 3, ce qui donne $\Delta r^2 = st$. Alors, la conjugaison par $x \mapsto \frac{r\Delta}{t}x$ laisse $a$ inchangé et transforme $b$ en $\frac{x+1}{\Delta x-1}$. □

## Bibliographie


[FGN] Francinou, Gianella, Nicolas, Exercices de mathématiques des oraux de l'Ecole polytechnique et des Ecoles normales supérieures : Algèbre, Tome I, *Cassini*.

[Perrin] Perrin, Cours d'Algèbre, *Ellipses*.

[Suzuki] Suzuki, Group Theory I, *Springer*. Le livre n'est pas à la BU, mais il est à la bibliothèque de Chevaleret à la cote 20 SUZ 82.

# Suites exactes

Considérons une suite de groupes et de morphismes de groupes

$$\ldots \longrightarrow G_{n-1} \xrightarrow{f_{n-1}} G_n \xrightarrow{f_n} G_{n+1} \longrightarrow \ldots$$

On dit que cette suite est *exacte en* $G_n$ si l'image de $f_{n-1}$ est égale au noyau de $f_n$. On dit que la suite est *exacte* si elle est exacte en chaque groupe de la suite.

L'exemple le plus important de suite exacte est celui des *suites exactes à trois termes* appelées aussi *suites exactes courtes*, qui sont les suites de la forme

$$1 \longrightarrow H \xrightarrow{i} G \xrightarrow{p} Q \longrightarrow 1.$$

Une telle suite est exacte si et seulement si $i$ est injectif, $p$ est surjectif et $\mathrm{im}(i) = \ker(p)$. Dit autrement, cela veut dire que $i$ identifie $H$ au noyau de $p$, et que $p$ s'identifie au morphisme de quotient. Les suites exactes courtes donnent donc un moyen compact d'écrire un groupe, un sous-groupe distingué et le quotient. Cette terminologie est beaucoup utilisée par exemple pour étudier le problème dit de l'*extension* : on dispose de deux groupes $H$ et $Q$, et on cherche tous les groupes $G$ qui possèdent un sous-groupe distingué isomorphe à $H$ avec quotient $G/H$ isomorphe à $Q$. Un tel groupe $G$ est alors appelé une *extension de $Q$ par $H$*, ce qui explique la terminologie.

Lorsqu'on se limite aux suites exactes à trois termes, l'avantage de présenter les quotients sous forme de suite exacte peut paraître discutable. La pratique, et l'intuition que l'on acquiert par la manipulation des suites exactes, montrent qu'en fait ce formalisme est extrêmement efficace même dans ce cas simple. Mais les suites exactes avec plus de termes apparaissent très naturellement en algèbre, et l'utilité du concept est alors encore plus visible. Le but des exercices qui suivent est de donner un exemple simple pour illustrer ce fait.

**Exercice 1** *Soit* $1 \longrightarrow G_1 \longrightarrow G_2 \longrightarrow \ldots \longrightarrow G_n \longrightarrow 1$ *une suite exacte de groupes finis. Montrez que le produit alterné des cardinaux* $n_i$ *des groupes* $G_i$*, c'est-à-dire le produit*

$$n_1 n_2^{-1} n_3 n_4^{-1} \ldots$$

*est égal à* 1.

*Soit* $0 \longrightarrow E_1 \longrightarrow E_2 \longrightarrow \ldots \longrightarrow E_n \longrightarrow 0$ *une suite exacte d'espaces vectoriels de dimension finie sur un corps* $k$ *(c'est-à-dire, une suite exacte des groupes commutatifs sous-jacents). Montrez que la somme alternée des dimensions* $\dim(E_i)$ *est égale à* 0.

**Exercice 2** *Soit* $\varphi(n)$ *l'indicateur d'Euler de* $n$*, c'est-à-dire le nombre d'entiers* $1 \leqslant k \leqslant n$ *premiers avec* $n$*. On sait que si* $a$ *et* $b$ *sont premiers entre eux, on a* $\varphi(ab) = \varphi(a)\varphi(b)$*. On va démontrer une généralisation de cette formule au cas où* $a$ *et* $b$ *ne sont pas nécessairement premiers entre eux. On note alors* $d$ *leur pgcd et* $m$ *leur ppcm.*

*(1) On considère la suite d'applications :*

$$0 \longrightarrow \mathbb{Z}/d\mathbb{Z} \xrightarrow{i} (\mathbb{Z}/ab\mathbb{Z})^* \xrightarrow{\Delta} (\mathbb{Z}/a\mathbb{Z})^* \times (\mathbb{Z}/b\mathbb{Z})^* \xrightarrow{p} (\mathbb{Z}/d\mathbb{Z})^* \longrightarrow 1$$

*avec* $i(r) = 1 + mr$*,* $\Delta(s) = (s, s)$ *et* $p(t, u) = tu^{-1}$ *où l'on s'est autorisé à désigner par une même lettre un élément et sa classe résiduelle, lorsque le modulo est clair d'après le contexte. Montrez qu'il s'agit d'une suite exacte de groupes finis.*

*(2) Déduire de la question précédente et de l'exercice précédent une formule pour* $\varphi(ab)$*.*

## Automorphismes du groupe des quaternions

On note $\mathbb{H}$ le groupe des quaternions. C'est un groupe d'ordre 8, engendré par deux éléments $i$ et $j$ dont on note $k$ le produit, possédant un seul élément central non trivial noté $-1$, avec une multiplication déterminée par les formules

$$i^2 = j^2 = k^2 = -1 \quad ; \quad ij = -ji = k \quad ;$$

$$ik = -ki = -j \quad ; \quad jk = -kj = i.$$

On peut construire facilement ce groupe comme sous-groupe de $\mathrm{GL}_4(\mathbb{R})$ ou du groupe symétrique $\mathfrak{S}_8$.

**Exercice 1** *Pour être bien à l'aise avec $\mathbb{H}$, montrez que :*

(1) *$\mathbb{H}$ possède un unique élément d'ordre 2, qui est $-1$.*

(2) *Le centre de $\mathbb{H}$ est $Z = \{1, -1\}$.*

(3) *Tout $x \in \mathbb{H} - Z$ est d'ordre 4 et vérifie $x^2 = -1$.*

(4) *Le quotient $\mathbb{H}/Z$ est isomorphe à $(\mathbb{Z}/2\mathbb{Z})^2$.*

(5) *Tous les sous-groupes de $\mathbb{H}$ sont distingués.*

**Exercice 2** *Dans cet exercice, on constate que pour le groupe $V = (\mathbb{Z}/2\mathbb{Z})^2$, les trois structures ensembliste, de groupe, de $\mathbb{F}_2$-espace vectoriel, ont essentiellement les mêmes automorphismes.*

(1) *Soit $p$ un nombre premier et $G$ un groupe abélien tel que pour tout $x \in G$, on a $px = 0$. Montrez qu'il existe une unique structure de $\mathbb{F}_p$-espace vectoriel sur $G$ compatible avec sa loi de groupe abélien. Déduisez-en que $\mathrm{Aut}(G)$ est le groupe linéaire $\mathrm{GL}(G)$ des automorphismes de $G$ vu comme $\mathbb{F}_p$-espace vectoriel.*

(2) *Notons 0 l'élément neutre de $V = (\mathbb{Z}/2\mathbb{Z})^2$. Pour tout automorphisme $f : V \to V$, on note $f'$ la bijection induite par $f$ sur $V - \{0\}$. Montrez que le morphisme*

$$\mathrm{Aut}(V) \to \mathfrak{S}_{V-\{0\}} \simeq \mathfrak{S}_3 \quad , \quad f \mapsto f'$$

*est un isomorphisme.*

**Exercice 3** *Soit $G$ un groupe, $Z$ son centre, $c : G \to \mathrm{Aut}(G)$ le morphisme qui à $g$ associe la conjugaison $c_g : x \mapsto gxg^{-1}$. On rappelle que l'image de $c$ est le sous-groupe distingué $\mathrm{Int}(G) \lhd \mathrm{Aut}(G)$ des automorphismes intérieurs et que $c$ induit un isomorphisme $G/Z \simeq \mathrm{Int}(G)$.*

*On note $V = \mathbb{H}/Z \simeq \mathrm{Int}(\mathbb{H})$.*

(1) *Décrivez les automorphismes intérieurs de $\mathbb{H}$.*

(2) *Montrez que tout automorphisme $f : \mathbb{H} \to \mathbb{H}$ vaut l'identité sur le centre. Montrez que le morphisme induit $\overline{f} : V \to V$ est l'identité si et seulement si $f$ est intérieur.*

(3) *Soit $\sigma$ une permutation de l'ensemble $\{i, j, k\}$. Montrez qu'il existe un unique automorphisme $f_\sigma$ de $\mathbb{H}$ qui envoie $i$ sur $\sigma(i)$ et $j$ sur $\sigma(j)$. En considérant les permutations $\sigma = (ij)$ et $\tau = (ijk)$ et les automorphismes $f_\sigma$ et $f_\tau$, montrez que la suite*

$$1 \longrightarrow V \xrightarrow{c} \mathrm{Aut}(\mathbb{H}) \xrightarrow{f \mapsto \overline{f}} \mathrm{Aut}(V) \longrightarrow 1$$

*est exacte et scindée, c'est-à-dire que $\mathrm{Aut}(\mathbb{H})$ est produit semi-direct $V \rtimes \mathrm{Aut}(V)$.*

(4) *Montrez que le groupe symétrique $\mathfrak{S}_4$ est un produit semi-direct $(\mathbb{Z}/2\mathbb{Z})^2 \rtimes \mathfrak{S}_3$ et que c'est le seul avec un 3-Sylow non distingué. Déduisez-en que $\mathrm{Aut}(\mathbb{H}) \simeq \mathfrak{S}_4$.*

**Exercice 4** *Soit $f \in \mathrm{Aut}(\mathbb{H})$. Montrez que $f$ est déterminé par $f(i)$ et $f(j)$. Montrez qu'on dispose d'au plus 6 choix pour $f(i)$, puis d'au plus 4 choix pour $f(j)$. Déduisez-en que n'importe quelle façon de faire ces choix définit un automorphisme.*

**Exercice 5** *Voici une présentation un peu plus conceptuelle de l'isomorphisme $\mathrm{Aut}(\mathbb{H}) \simeq \mathfrak{S}_4$. On considère l'ensemble des parties à trois éléments $x = \{\epsilon_1 i, \epsilon_2 j, \epsilon_3 k\}$ où les signes $\epsilon_i \in \{\pm 1\}$ sont variables ; cet ensemble est de cardinal 8. On considère ensuite l'ensemble $X$ des paires $y = \{x, -x\}$, de cardinal 4.*

(1) *Montrez qu'un automorphisme de $\mathbb{H}$ permute $X$.*

(2) *Montrez que le morphisme $\mathrm{Aut}(\mathbb{H}) \to \mathfrak{S}_X$ est injectif.*

(3) *Déduisez-en que $\mathrm{Aut}(\mathbb{H}) \simeq \mathfrak{S}_4$.*

# Symétries du Sudoku

Les grilles de Sudoku standard sont composées de 9 blocs de taille 3×3, dont on doit remplir les cases avec les chiffres de 1 à 9 de telle sorte qu'un chiffre n'apparaisse qu'une fois dans chaque ligne, chaque colonne, et chaque bloc. Voici un exemple :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 6 | 9 | 7 | 4 | 5 | 3 | 2 | 1 | 8 |
| 8 | 2 | 4 | 6 | 1 | 9 | 7 | 5 | 3 |
| 3 | 5 | 1 | 2 | 8 | 7 | 6 | 9 | 4 |
| 7 | 3 | 5 | 8 | 6 | 1 | 9 | 4 | 2 |
| 9 | 4 | 8 | 7 | 3 | 2 | 1 | 6 | 5 |
| 1 | 6 | 2 | 5 | 9 | 4 | 3 | 8 | 7 |
| 5 | 1 | 9 | 3 | 7 | 8 | 4 | 2 | 6 |
| 4 | 7 | 6 | 9 | 2 | 5 | 8 | 3 | 1 |
| 2 | 8 | 3 | 1 | 4 | 6 | 5 | 7 | 9 |

Le but de cette feuille d'exercices est de décrire les grilles de Sudoku à « symétrie » près. La définition mathématique d'une grille de Sudoku s'obtient en mettant en valeur les différents éléments constitutifs du Sudoku, tels qu'introduits ci-dessus : les *cases*, *blocs*, *lignes*, *colonnes* et *chiffres*.

**Définitions.** Soit $\mathcal{C}$ l'ensemble $\{1,\dots,9\} \times \{1,\dots,9\}$.

- une *case* est un élément de $\mathcal{C}$ i.e. un couple $(i,j)$,
- une *ligne* est une partie de la forme $L_i := \{i\} \times \{1,\dots,9\}$,
- une *colonne* est une partie de la forme $C_j := \{1,\dots,9\} \times \{j\}$,
- une *bande* est une réunion de trois lignes, de la forme $B_i := L_{3i-2} \cup L_{3i-1} \cup L_{3i}$,
- une *pile* est une réunion de trois colonnes, de la forme $P_j := C_{3j-2} \cup C_{3j-1} \cup C_{3j}$,
- un *bloc* est l'intersection d'une bande et d'une pile.
- une *grille de Sudoku* est une application $f : \mathcal{C} \to \{1,\dots,9\}$ dont la restriction aux lignes, aux colonnes et aux blocs est injective. On appelle $f(i,j)$ le *chiffre* de la case $(i,j)$.

On peut permuter les cases (en respectant la structure ligne-colonne-bloc) ou les chiffres d'une grille de Sudoku. Ceci revient à agir à la source ou au but d'une application $f : \mathcal{C} \to \{1,\dots,9\}$. On arrive ainsi à la définition suivante.

**Définition.** Soit $X$ l'ensemble des grilles de Sudoku. Une *transformation de Sudoku* (ou simplement *transformation*) est une paire de bijections $(g : \mathcal{C} \to \mathcal{C}, \sigma : \{1,\dots,9\} \to \{1,\dots,9\})$ telle que si $x$ est une ligne ou une colonne, alors $g(x)$ est une ligne ou une colonne. Une telle transformation donne lieu à une bijection $(g,\sigma) : X \to X$, $f \mapsto \sigma \circ f \circ g^{-1}$.

On notera $G$ le groupe des transformations de Sudoku, agissant sur $X$ par la formule précédente. Le thème de cette feuille est donc de décrire l'ensemble des orbites de $X$ sous $G$, ou encore, l'ensemble $X/G$.

(1) Donnez une liste aussi longue que possible de transformations de Sudoku.

(2) Soit $g$ une transformation de Sudoku. Montrez que s'il existe une ligne qui est envoyée par $g$ sur une ligne, alors toutes les lignes le sont. Montrez que si $g$ vaut l'identité sur $L_1 \cup C_1$, alors c'est l'identité.

(3) Déduisez de la question (2) que le groupe $G$ est engendré par les transformations suivantes :

- les renumérotations de chiffres i.e. les éléments de la forme $(1,\sigma)$,
- la réflexion par rapport à la diagonale (ensemble des cases $(i,i)$),
- les échanges de deux bandes,

- les échanges de deux lignes dans la même bande,
- les échanges de deux piles,
- les échanges de deux colonnes dans la même pile.

On s'intéresse maintenant au nombre d'orbites, i.e. au nombre de grilles à transformation près. Pour simplifier le problème, on regarde désormais les grilles composées de 4 blocs de taille 2×2, dont on remplit les cases avec les chiffres de 1 à 4 de telle sorte qu'un chiffre n'apparaisse qu'une fois dans chaque ligne, chaque colonne, et chaque bloc. Par exemple :

| | | | |
|---|---|---|---|
| 3 | 2 | 4 | 1 |
| 1 | 4 | 2 | 3 |
| 4 | 3 | 1 | 2 |
| 2 | 1 | 3 | 4 |

On attribue des symboles aux transformations suivantes :

| | renumérotation des symboles 1,2,3,4 |
|---|---|
| $\overrightarrow{\square}$ | échange des deux bandes |
| $\overrightarrow{\square}1$ | échange des lignes dans la bande 1 |
| $\overrightarrow{\square}2$ | échange des lignes dans la bande 2 |
| $\downarrow\!\square\!\downarrow$ | échange des piles |
| $\downarrow\!\square\!\downarrow 1$ | échange des colonnes dans la pile 1 |
| $\downarrow\!\square\!\downarrow 2$ | échange des colonnes dans la pile 2 |
| $\diagdown$ | réflexion autour de la diagonale NO-SE |
| $\diagup$ | réflexion autour de la diagonale NE-SO |
| — | réflexion autour de l'axe de symétrie horizontal |
| $\vert$ | réflexion autour de l'axe de symétrie vertical |
| $\circlearrowleft$ | rotation de 1/4 tour dans le sens indiqué |

(4) On note $\langle - \rangle$ le sous-groupe engendré. Montrez que :

- $\overrightarrow{\square}1 \in \langle \overrightarrow{\square}, \overrightarrow{\square}2 \rangle$,
- $\downarrow\!\square\!\downarrow 1 \in \langle \downarrow\!\square\!\downarrow, \downarrow\!\square\!\downarrow 2 \rangle$.

Déduisez-en que $G$ est engendré par les renumérotations et l'ensemble $\{\overrightarrow{\square}, \overrightarrow{\square}2, \downarrow\!\square\!\downarrow, \downarrow\!\square\!\downarrow 2, \diagdown\}$.

(5) (Cette question n'est pas indispensable pour la suite.) Montrez que :

- $— \in \langle \overrightarrow{\square}, \overrightarrow{\square}2 \rangle$,
- $\vert \in \langle \downarrow\!\square\!\downarrow, \downarrow\!\square\!\downarrow 2 \rangle$,
- $\diagup \in \langle —, \vert, \diagdown \rangle$,
- $\circlearrowleft \in \langle \overrightarrow{\square}, \diagup, \downarrow\!\square\!\downarrow 1, \downarrow\!\square\!\downarrow 2 \rangle$.

(6) Soit $H$ le sous-groupe de $G$ engendré par les renumérotations et l'ensemble $\{\overrightarrow{\square}2, \downarrow\!\square\!\downarrow 2, \diagdown\}$. On part d'une grille de Sudoku quelconque. Montrez qu'en utilisant des transformations de $H$, on peut successivement ramener cette grille aux grilles suivantes :

① 

| | | | |
|---|---|---|---|
| 1 | 2 | * | * |
| 3 | 4 | * | * |
| * | * | * | * |
| * | * | * | * |

② 

| | | | |
|---|---|---|---|
| 1 | 2 | 3 | 4 |
| 3 | 4 | * | * |
| * | * | * | * |
| * | * | * | * |

③ 

| | | | |
|---|---|---|---|
| 1 | 2 | 3 | 4 |
| 3 | 4 | * | * |
| 2 | * | * | * |
| 4 | * | * | * |

④ 

| | | | |
|---|---|---|---|
| 1 | 2 | 3 | 4 |
| 3 | 4 | 1 | * |
| 2 | * | * | * |
| 4 | * | * | * |

(7) Complétez autant que possible cette dernière grille et montrez qu'on arrive finalement à deux grilles possibles. Déduisez-en que $H$ agit librement sur $X$ et que $X/H$ contient deux éléments.

(8) Dans un cadre général, considérons un groupe $G$ agissant sur un ensemble $X$. Pour tout sous-groupe $H$ de $G$, définissez une relation d'équivalence naturelle $\mathcal{R}$ sur $X/H$, induite par l'action de $G$, telle que la surjection $X/H \to X/G$ se factorise en une bijection $(X/H)/\mathcal{R} \simeq X/G$.

(9) En utilisant la question précédente dans le contexte du Sudoku, montrez que la relation d'équivalence induite par $G$ sur $X/H$ est triviale (c'est-à-dire que $x\mathcal{R}y$ ssi $x = y$). Déduisez-en que $X/H \simeq X/G$ et qu'il y a deux grilles de Sudoku distinctes à transformation près.