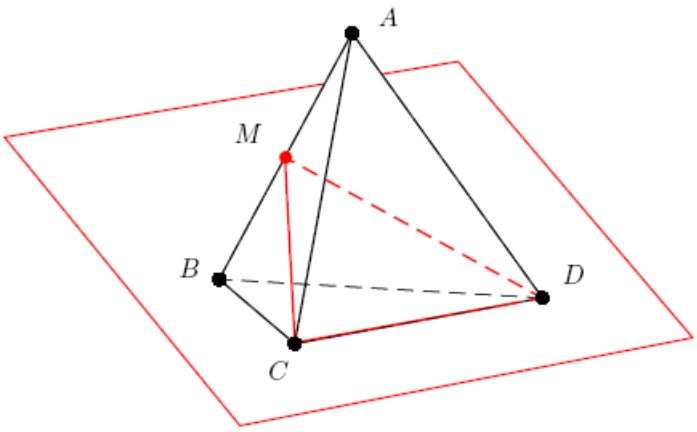
Université Claude Bernard-Lyon 1

Carnet de voyage en Algébrie



$\mathrm{Is}(\mathcal{T}) \simeq \mathfrak{S}_4$

Notes de TD de :
Chrystel Bouvier
Elisabeth Bruyère

Encadrées par :
Philippe Caldero

J'aimerais vivre en Théorie, parce qu'en Théorie, tout va bien...

Table des matières

Chapitre I

Algèbre linéaire

I.1 Dualité

I.1.1 Généralités

Exercice I.1.1. Soit E un espace de dimension n, $(e_i)_{1\leq i\leq n}$ une base de E et $(e_i^*)_{1\leq i\leq n}$ sa base duale dans E^*.

1. On considère un vecteur v dans E. Montrer que les coordonnées de v dans la base (e_i) sont données par
$$v = \sum_{i=1}^{n} e_i^*(v)e_i.$$
2. On considère une forme linéaire φ dans E^*. Montrer que les coordonnées de φ dans la base duale sont données par
$$\varphi = \sum_{i=1}^{n} \varphi(e_i)e_i^*.$$

Soluce.

1. Si on note $v = \sum_{i=1}^n v_i e_i$ la décomposition de v dans la base (e_i), on a par définition
$$e_j^*(v) = e_j^*\big(\sum_{i=1}^{n} v_i e_i\big) = \sum_{i=1}^{n} v_i e_j^*(e_i) = v_j.$$
D'où la formule qui décompose v dans la base (e_i).
2. On décompose $\varphi = \sum_{i=1}^n \varphi_i e_i^*$ dans la base duale et on évalue l'égalité en e_j :
$$\varphi(e_j) = \big(\sum_{i=1}^{n} \varphi_i e_i^*\big)(e_j) = \sum_{i=1}^{n} \varphi_i e_i^*(e_j) = \varphi_j.$$

Remarque. On vient de montrer deux égalités, mais en réalité, ces égalités sont non seulement équivalentes, mais ce sont les mêmes ! Effectivement, si l'on sait qu'un espace de dimension finie est isomorphe à son bidual en posant

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

on se rend compte que pour déduire la seconde formule de la première, il suffit de remplacer E par E^*, v par φ, et e_i par e_i^*.

C'est là que réside toute l'efficacité de la (belle) dualité. Non seulement elle met en relation les objets géométriques (ici, les vecteurs) et les fonctions sur ces objets, mais elle dédouble aussi les formules... et les théorèmes.

Exercice I.1.2 (Lagrange et Taylor)**.**
Soit $\mathbb{K}$ un corps de caractéristique nulle (donc infini, puisqu'il contient $\mathbb{Q}$) et $E := \mathbb{K}_n[X]$ l'algèbre des polynômes sur $\mathbb{K}$ de degré inférieur à n.

1. Soit a dans $\mathbb{K}$. Pour k de 0 à n, on pose $\varphi_k(P) = P^{(k)}(a)$.
   (a) Montrer que $(\varphi_k)_{0\leq k\leq n}$ est une base de E^* dont la base (anté)-duale est la base des polynômes $(P_k)_{0\leq k\leq n}$, avec $P_k = \frac{(X-a)^k}{k!}$.
   (b) En déduire la formule de Taylor polynomiale, pour $P \in \mathbb{K}_n[X]$
   $$P = \sum_{k=0}^{n} P^{(k)}(a)\frac{(X-a)^k}{k!}.$$
2. Soit a_k, $0 \leq k \leq n$ des éléments de $\mathbb{K}$ deux à deux distincts. Pour k de 0 à n, on pose $\psi_k(P) = P(a_k)$.
   (a) Montrer que $(\psi_k)_{0\leq k\leq n}$ est une base de E^* dont la base (anté)-duale est la base des polynômes $(Q_k)_{0\leq k\leq n}$, avec
   $$Q_k = \prod_{j\neq k} \frac{X-a_j}{a_k-a_j}.$$
   (b) En déduire la formule des polynômes interpolateurs de Lagrange, pour $P \in \mathbb{K}_n[X]$
   $$P = \sum_{k=0}^{n} P(a_k)Q_k.$$

Soluce. Avant de commencer, notons un principe général. Soit E un espace vectoriel et $(\varphi_i)_{i\in I}$ une famille de formes linéaires sur E. On suppose qu'il existe une famille $(e_i)_{i\in I}$ telle que $\varphi_i(e_j) = \delta_{ij}$, où δ_{ij} désigne le symbole de Kronecker. Alors, la famille $(\varphi_i)_{i\in I}$ est libre. Cela se voit immédiatement en écrivant une combinaison linéaire des φ_i, qui s'annule, et en l'appliquant à e_j. De façon symétrique, en appliquant φ_j à une combinaison linéaire nulle des e_i, on montre que la famille $(e_i)_{i\in I}$ est libre [1].

1. En fait, c'est la méthode « duale ».

1. On vérifie que $\varphi_i(P_j) = \delta_{ij}$, pour tout i, j. La famille $(\varphi_i)_{0\leq i\leq n}$ est libre, par le principe ci-dessus, ce qui implique que c'est une base, par un argument de cardinalité. La base (anté)-duale est donc la famille $(P_j)_{0\leq j\leq n}$, qui, du coup, constitue bien une base.
   La formule de Taylor, n'est rien d'autre que la formule de l'exercice précédent appliquée à cet exemple.
2. On vérifie que $\psi_i(Q_j) = \delta_{ij}$, pour tout i, j. La famille $(\psi_i)_{0\leq i\leq n}$ est libre, par le principe ci-dessus, ce qui implique que c'est une base, par un argument de cardinalité. La base (anté)-duale est donc la famille $(Q_j)_{0\leq j\leq n}$, qui, du coup, constitue bien une base.
   La formule de Lagrange n'est rien d'autre que la formule de l'exercice précédent appliquée à cet exemple.

Exercice I.1.3 (Deux formules pour le prix d'une)**.** Soit u un endomorphisme du $\mathbb{K}$-espace vectoriel E, P et Q deux polynômes de $\mathbb{K}[X]$. On notera $D = \operatorname{pgcd}(P, Q)$ et $M = \operatorname{ppcm}(P, Q)$.

1. Montrer que

   $$\ker D(u) = \ker P(u) \cap \ker Q(u) \text{ et que } \operatorname{Im} D(u) = \operatorname{Im} P(u) + \operatorname{Im} Q(u).$$

2. Montrer que

   $$\ker M(u) = \ker P(u) + \ker Q(u) \text{ et que } \operatorname{Im} M(u) = \operatorname{Im} P(u) \cap \operatorname{Im} Q(u).$$

Soluce. 1. On pose $DA = P$, $DB = Q$. On sait alors que A et B sont premiers entre eux.

• $\ker D(u) \subset \ker P(u) \cap \ker Q(u)$. Par symétrie, il suffit de montrer $\ker D(u) \subset \ker P(u)$. Si $x \in \ker D(u)$, alors $P(u)(x) = (AD)(u)(x) = A(u)D(u)(x) = A(u)(0) = 0$.

• $\ker P(u) \cap \ker Q(u) \subset \ker D(u)$. On a $D = UP + VQ$ par Bezout. Donc, si $x \in \ker P(u) \cap \ker Q(u)$, on a $D(u)(x) = U(u)P(u)(x) + V(u)Q(u)(x) = U(u)(0) + V(u)(0) = 0$.

La seconde égalité s'en déduit facilement par dualité en appliquant la première à ${}^t u$ et en prenant l'orthogonal. Précisons :

En appliquant l'égalité précédente à la transposée ${}^t u \in \mathcal{L}(E^*)$, il vient

$$\ker D({}^t u) = \ker P({}^t u) \cap \ker Q({}^t u).$$

Or, d'une part

$$\ker D({}^t u) = \ker \big({}^t D(u)\big) = \operatorname{Im} D(u)^\perp.$$

D'autre part,

$$\begin{aligned}\ker P({}^t u) \cap \ker Q({}^t u) &= \ker {}^t P(u) \cap \ker {}^t Q(u) = \operatorname{Im} P(u)^\perp \cap \operatorname{Im} Q(u)^\perp \\ &= \big(\operatorname{Im} P(u) + \operatorname{Im} Q(u)\big)^\perp.\end{aligned}$$

D'où l'égalité demandée.

2. On pose $M = RP$, avec R divise Q (puisque P divise M qui divise PQ).
• $\ker P(u) + \ker Q(u) \subset \ker M(u)$. Montrons que $\ker P(u) \subset \ker M(u)$, on aura de même $\ker Q(u) \subset \ker M(u)$, d'où l'inclusion cherchée. Soit x dans $\ker P(u)$, on a donc

$$M(u)(x) = (RP)(u)(x) = R(u)(P(u)(x)) = R(u)(0) = 0.$$

• $\ker M(u) \subset \ker P(u) + \ker Q(u)$. On calque la preuve du *lemme des noyaux*. On a A et B (comme ci-dessus) premiers entre eux et donc, $1 = TB + SA$ par Bezout. De plus, comme $PQ = DM$, il vient $M = ADB$. Soit x dans $\ker M(u)$. En évaluant en x l'égalité $\mathrm{Id} = T(u)B(u) + S(u)A(u)$, on a $x = Id(x) = T(u)B(u)(x) + S(u)A(u)(x)$, ce qui donne facilement la décomposition voulue, puisque, si x est dans $\ker M(u)$,

$$P(u)\big(T(u)B(u)\big)(x) = T(u)M(u)(x) = 0,$$

et de même

$$Q(u)\big(S(u)A(u)\big)(x) = S(u)(AQ)(u)(x) = S(u)M(u)(x) = 0.$$

L'autre égalité se prouve par dualité de la même manière que précédemment.

I.1.2 Trace et dualité

Exercice I.1.4. [Dualité et l'algèbre des matrices]
Soit $\mathbb{K}$ un corps, n un entier positif, et φ une forme linéaire sur l'espace $\mathcal{M}_n(\mathbb{K})$ des matrices carrées de taille n.

1. Montrer qu'il existe une unique matrice A de $\mathcal{M}_n(\mathbb{K})$ telle que
$$\varphi(X) = \mathrm{tr}(AX),$$
pour toute matrice X de $\mathcal{M}_n(\mathbb{K})$.
2. Application 1 : Montrer que si de plus φ vérifie $\varphi(XY) = \varphi(YX)$ pour tout X, Y dans $\mathcal{M}_n(\mathbb{K})$, alors $\varphi = \lambda\,\mathrm{tr}$ pour un λ dans $\mathbb{K}$.
3. Application 2 : Montrer que tout hyperplan de $\mathcal{M}_n(\mathbb{K})$ contient au moins une matrice inversible.
4. Application 3 : Montrer que tout hyperplan de $\mathcal{M}_n(\mathbb{R})$ contient au moins une matrice orthogonale.

Soluce. 1. Si on note $X = (x_{ij})$, on a par linéarité que

$$\varphi(X) = \sum_{1 \le i,j \le n} a_{i,j} x_{i,j},$$

pour des scalaires $(a_{i,j})$ déterminés de façon unique en fonction de φ. On voit facilement que la matrice $A = (a_{j,i})$ fait l'affaire : cela repose sur la formule

$$\mathrm{tr}(AE_{ij}) = a_{j,i},$$

où les E_{ij} sont les matrices élémentaires.

L'application de $\mathcal{M}_n(\mathbb{K})$ dans $\mathcal{M}_n(\mathbb{K})^*$ qui envoie A sur la forme linéaire $\varphi_A : X \mapsto \mathrm{tr}(AX)$ est donc une application linéaire surjective, donc injective par égalité des dimensions. A est donc bien unique.

2. On note A la matrice telle que $\varphi(X) = \mathrm{tr}(AX)$. Puisque φ vérifie $\varphi(XY) = \varphi(YX)$, il vient $\mathrm{tr}(AXY) = \mathrm{tr}(AYX) = \mathrm{tr}(XAY)$, par propriété de la trace. Ainsi, pour toute matrice X, on a

$$\mathrm{tr}\left((AX - XA)Y\right) = 0,$$

pour toute matrice Y. En utilisant l'unicité dans la question qui précède, on en déduit $AX - XA = 0$, et ce, pour tout X.

Donc, A est une matrice qui commute à toute matrice X. On sait que cela implique que A est une homothétie. Montrons cette assertion.

Soit v un vecteur non nul, et X une matrice telle que la droite engendrée par v est un sous-espace propre de X (ce qui signifie que la valeur propre correspondante est de multiplicité 1). Comme A commute avec X, A stabilise tout sous-espace propre de X, donc A stabilise la droite engendrée par v. On obtient donc un scalaire λ_v tel que $Av = \lambda_v v$.

Il ne reste plus qu'à montrer que λ_v ne dépend pas de v, c'est-à-dire que pour tout w non nul, $\lambda_w = \lambda_v$. Deux cas se présentent selon si w est proportionnel à v ou non :

• 1er Cas. Si $w = \mu v$:

Dans ce cas,

$$\lambda_w w = Aw = A(\mu v) = \mu Av = \mu\lambda_v v = \lambda_v w.$$

Donc, $\lambda_v = \lambda_w$, puisque w est non nul.

• 2nd Cas. Si $w \neq \mu v$:

Dans ce cas (v, w) forme une partie libre.

$$\lambda_{v+w}(v + w) = A(v + w) = Av + Aw = \lambda_v v + \lambda_w w.$$

Et donc, $\lambda_w = \lambda_{v+w} = \lambda_v$.

On sait maintenant que A est une matrice scalaire : $A = \lambda \mathrm{I}_n$. Et donc, $\varphi(X) = \lambda\,\mathrm{tr}(X)$.

3. Soit H un hyperplan de $\mathcal{M}_n(\mathbb{K})$. On sait qu'il existe une forme linéaire non nulle φ telle que $H = \ker(\varphi)$. Soit A la matrice telle que $\varphi(X) =$

$\mathrm{tr}(AX)$ pour toute matrice X. La matrice A est donc une matrice non nulle, de rang $r > 0$.

La matrice A est donc équivalente à la matrice J_r définie par

$$J_r := \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{pmatrix}$$

On a donc deux matrices inversibles P et Q telles que $A = PJ_rQ$. Ceci implique, par propriété de la trace, que

$$\mathrm{tr}(AX) = \mathrm{tr}(PJ_rQX) = \mathrm{tr}\big(J_r(QXP)\big).$$

Posons $Y = QXP$. On voit que X est inversible si et seulement si Y l'est. Trouver une matrice inversible dans H revient donc à trouver une matrice inversible X telle que $\mathrm{tr}(AX) = 0$, c'est-à-dire, trouver Y inversible telle que $\mathrm{tr}(J_rY) = 0$.

Il suffit de prendre

$$Y = \begin{pmatrix} 0 & 0 & \cdots & \cdots & 0 & 1 \\ 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \cdots & \cdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 1 & 0 \end{pmatrix}$$

Le déterminant de Y est $(-1)^{n-1}$ (on reconnaît une matrice de permutation correspondant à un n-cycle.) De plus, la diagonale de J_rY est nulle et donc sa trace est bien nulle comme désiré.

4. La méthode est similaire, mais utilise la décomposition polaire plutôt que le théorème du rang. Soit H un hyperplan de $\mathcal{M}_n(\mathbb{R})$. Il existe donc, comme ci-dessus, une forme linéaire φ et une matrice A non nulles telles que $H = \ker(\varphi)$ et $\varphi(X) = \mathrm{tr}(AX)$ pour toute matrice X.

   On peut donc décomposer A en $A = OS$, avec O orthogonale et S une matrice symétrique positive[2]. Il existe une matrice orthogonale P et une matrice diagonale D telles que $S = PDP^{-1}$.

   Ceci implique, par propriété de la trace, que

   $$\mathrm{tr}(AX) = \mathrm{tr}(OSX) = \mathrm{tr}(OPDP^{-1}X) = \mathrm{tr}(P^{-1}XOPD).$$

   Posons $Y = P^{-1}XOP$. Comme les matrice orthogonales forment un groupe, on voit que X est orthogonale si et seulement si Y l'est. Trouver

2. On notera ici que A n'étant pas supposée inversible, la matrice S n'est pas censée être définie positive. On perd aussi l'unicité de la décomposition.

une matrice inversible dans H revient donc à trouver une matrice orthogonale X telle que $\mathrm{tr}(AX) = 0$, c'est-à-dire, trouver Y orthogonale telle que $\mathrm{tr}(YD) = 0$.
Il suffit de prendre la même matrice Y que ci-dessus. Effectivement, toute matrice de permutation permute la base canonique, et donc transforme une base orthonormée (en l'occurence, la base canonique) de l'espace euclidien $\mathbb{R}^n$ en une base orthonormée (la base canonique permutée) ; la matrice Y est bien orthogonale.
De plus, la diagonale de YD est nulle et donc sa trace est bien nulle comme désiré.

I.1.3 Dualité et topologie

Exercice I.1.5. Soit E un espace vectoriel de dimension finie sur le corps $\mathbb{K}$, où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

1. Montrer que toute forme linéaire est continue.
2. En déduire que $\mathrm{GL}_n(\mathbb{K})$ est ouvert dans $\mathcal{M}_n(\mathbb{K})$.

Soluce.
1. On fixe une base $(e_i)_{i\in I}$ de E et une norme sur E. Le caractère continu d'une forme linéaire fixée φ ne dépend pas de la norme choisie puisque, en dimension finie, toutes les normes sont équivalentes. Soit $x = \sum_i x_i e_i$ sur la sphère de rayon 1 de E. On a
$$|\varphi(x)| = |\sum_i x_i \varphi(e_i)| \leq \sum_i |x_i||\varphi(e_i)| \leq \sum_i |\varphi(e_i)|.$$
Ainsi, φ est bornée, donc continue en 0, et donc continue partout par linéarité.
2. D'après la formule du déterminant, l'application det de $\mathcal{M}_n(\mathbb{K})$ dans $\mathbb{K}$ est une somme et produit de formes coordonnées. Elle est donc continue, et l'image inverse de l'ouvert $\mathbb{K}^*$ est un ouvert. On en déduit le résultat demandé.

Exercice I.1.6. Montrer que tout sous-espace vectoriel est fermé (en dim finie).

Soluce. Soit F un sous-espace vectoriel et $F^\perp$ son orthogonal dans le dual. On a $F = (F^\perp)^\perp$ et donc $F = \{v,\, \phi(v) = 0, \phi \in F^\perp\}$. Comme toute forme linéaire ϕ est continue, voir exercice I.1.5, $\phi^{-1}(0)$ est fermé. L'espace F est une intersection de fermés donc fermé.

I.2 Rang

Exercice I.2.1.
Comparer le rang d'une matrice à coefficients entiers sur $\mathbb{Q}$, sur $\mathbb{C}$, puis sur le corps $\mathbb{F}_p$.

Soluce. Le rang d'une matrice, sur un corps quelconque, est égal à la taille du plus grand mineur non nul.

Comme la matrice est à coefficients entiers, chaque mineur est un entier.

Qu'il soit vu dans $\mathbb{Z}$, dans $\mathbb{Q}$ ou dans $\mathbb{C}$, un entier non nul reste non nul (car nous sommes en caractéristique non nulle). Donc, les mêmes mineurs sont non nuls, qu'on les considères dans $\mathbb{Q}$ ou $\mathbb{C}$.

En revanche, si on considére un entier dans $\mathbb{F}_p$, c'est-à-dire, si l'on prend sa classe modulo p, il est possible qu'il s'annule.

Soit donc, $A = (a_{ij})$ une matrice à coefficients dans $\mathbb{Z}$. On peut réduire A modulo p pour obtenir $\overline{A} = (\bar{a}_{ij})$. Comme le déterminant est polynomial à coefficients entiers, on a $\det(\overline{A}) = \overline{\det(A)}$.

Ceci implique que si un mineur est nul, il reste nul après réduction de la matrice. Mais on voit facilement que la réciproque est fausse. Le rang devient plus petit après réduction modulo p.

Par exemple, la matrice $\begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix}$ a pour déterminant 15. Elle est de rang 2 sur $\mathbb{Q}$ et sur $\mathbb{C}$, mais de rang 1 sur $\mathbb{Z}/3\mathbb{Z}$ et sur $\mathbb{Z}/5\mathbb{Z}$.

Exercice I.2.2.
Montrer que deux projecteurs sont semblables si et seulement s'ils sont équivalents.

Soluce. L'implication est valable pour deux endomorphismes quelconques : si deux endomorphismes sont semblables, alors ils sont équivalents. La réciproque est fausse en général, mais vraie pour les projecteurs.

Rappel. Un projecteur est un endomorphisme p d'un espace E, tel que $p^2 = p$. Un projecteur annule donc le polynôme $P(X) = X^2 - X$, qui est scindé [3] à racines simples : $P(X) = X(X-1)$. Un projecteur est donc diagonalisable, et possède les deux valeurs propres 0 et 1 (on a aussi les cas extrêmes de l'endomorphisme nul, où 0 est seule valeur propre, et de l'endomorphisme identité, où 1 est seule valeur propre).

D'après le rappel, tout projecteur est semblable à une matrice diagonale comportant sur sa diagonale des 0 et des 1. Le nombre de 1 correspond donc à la fois à la trace et au rang du projecteur.

Ainsi, si deux projecteurs sont équivalents, alors ils ont même rang r et sont tous les deux semblables à la matrice $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. Ils sont donc semblables ; cela prouve l'assertion réciproque.

3. Cette fois-ci, c'est vrai sur tout corps, car $0 \neq 1$ est la seule chose dont on soit sûr.

Exercice I.2.3. *[Propriétés topologiques du rang. [H2G2], Chap. I Proposition 4.1] On se propose de montrer que dans $\mathcal{M}_n(\mathbb{C})$, l'ensemble $\mathcal{O}_r$ des matrices de rang r est connexe et que son adhérence est l'ensemble des matrices de rang inférieur à r.

1. Montrer à l'aide du théorème du rang que $\mathcal{O}_r$ est l'image du groupe $\mathrm{GL}_n(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C})$ par une application continue et en déduire que $\mathcal{O}_r$ est connexe.
2. En utilisant la caractérisation du rang par des mineurs, montrer que l'ensemble des matrices de rang strictement inférieur à r est un fermé.
3. Conclure que $\overline{\mathcal{O}_r} = \bigcup_{k=0}^{r} \mathcal{O}_k$.

Soluce. 1. On sait qu'une matrice est de rang r si et seulement si elle est équivalente à la matrice $J_r := \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{pmatrix}$.

Considérons l'application :

$$\begin{aligned} \Phi\colon \mathrm{GL}_n(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C}) &\to \mathcal{O}_r \\ (P, Q) &\mapsto P J_r Q^{-1} \end{aligned}$$

L'application Φ est donc bien définie et surjective.

Montrons que cette application est continue. La multiplication des matrices est une application polynomiale (et même bilinéaire !) de $\mathcal{M}_n(\mathbb{C}) \times \mathcal{M}_n(\mathbb{C})$ dans $\mathcal{M}_n(\mathbb{C})$. De plus, la formule

$$Q^{-1} = \frac{1}{\det(Q)}\, {}^t\mathrm{com}(Q)$$

prouve que $Q \mapsto Q^{-1}$ est continue de $\mathrm{GL}_n(\mathbb{C})$ dans lui-même. Il en résulte bien que Φ est continue.

Comme $\mathrm{GL}_n(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C})$ est connexe par l'exercice I.5.5, on a $\mathcal{O}_r = \Phi(GL_n(\mathbb{C}) \times GL_n(\mathbb{C}))$ est connexe (et même connexe par arcs).

2. Considérons l'application :

$$f\colon \mathcal{M}_n(\mathbb{C}) \to \mathbb{R},\ A \mapsto \sum |\Delta_r|$$

qui, à une matrice A, fait correspondre la somme des modules de tous ses mineurs de taille r. On a clairement que la fonction f est continue. De plus, $f(A) = 0$ si et seulement si tous les mineurs de taille r sont nuls, si et seulement si $rg(A) < r$ car le rang d'une matrice A est égal à la taille maximale d'un mineur non nul de A.

L'ensemble des matrices de rang strictement inférieur à r est donc $f^{-1}(0)$. C'est donc l'image réciproque d'un fermé par une application continue, c'est donc bien un fermé.

3. Notons N_r l'ensemble des matrices de rang inférieur ou égal à r. On a $\mathcal{O}_r \subset N_r$. En prenant l'adhérence, on obtient $\overline{\mathcal{O}_r} \subset \overline{N_r} = N_r$ car N_r est un fermé.
   Montrons l'inclusion inverse. Soit B une matrice de rang k, $k \leq r$. Alors, B est équivalente à la matrice I_k, donc $B = P\mathrm{I}_k Q^{-1}$ avec P et Q dans $\mathrm{GL}_n(\mathbb{C})$. On pose $\mathrm{I}_{k,m} := \begin{pmatrix} \mathrm{I}_k & 0 & 0 \\ 0 & \frac{1}{m}\mathrm{I}_{r-k} & 0 \\ 0 & 0 & 0 \end{pmatrix}$ et $A_m := P\mathrm{I}_{k,m}Q^{-1}$.
   La suite A_m est donc clairement une suite de matrices de rang r.
   Montrons qu'elle converge vers B. En utilisant la norme[4] $||\,||_\infty$ des matrices, on voit que $\lim_m \mathrm{I}_{k,m} = \mathrm{I}_k$. De plus, comme la fonction $M \mapsto PMQ^{-1}$ est continue (car linéaire, par exemple), il vient que $\lim_m P\mathrm{I}_{k,m}Q^{-1} = P \lim_m \mathrm{I}_{k,m}Q^{-1} = P\mathrm{I}_k Q^{-1} = B$. Donc B est bien dans $\overline{\mathcal{O}_r}$.

Remarque. On note que l'on aurait pu un peu simplifier la preuve en posant tout simplement $\Phi(P, Q) = PJ_rQ$.

Exercice I.2.4. *[Actions de groupes et topologie. [H2G2] Chap. II Proposition 2.3.1]
Montrer que sur $\mathbb{C}$, l'ensemble des matrices de projection de rang r est connexe. Montrer qu'il est fermé.
On pourra utiliser la méthode de l'exercice précédent pour la connexité et utiliser la trace pour la fermeture.

Soluce. • Connexité.

Rappel. On rappelle le fait suivant, voir exercice I.2.2 : la matrice A est la matrice d'une projection de rang r si et seulement si A est semblable à J_r, si et seulement s'il existe P dans $\mathrm{GL}_n(\mathbb{C})$ tel que $A = PJ_rP^{-1}$, avec J_r comme dans l'exercice I.1.4.

Notons $\mathcal{P}$ l'ensemble des matrices de projections et $\mathcal{P}_r$ l'ensemble des matrices de projections de rang r. Considérons l'application :

$$\begin{aligned} f\colon \mathrm{GL}_n(\mathbb{C}) &\to \mathcal{P}_r \\ P &\mapsto PJ_rP^{-1} \end{aligned}$$

On a par ce qui précède que f est bien définie et surjective. De plus, f est continue (comme dans l'exercice précédent) et $\mathrm{GL}_n(\mathbb{C})$ est connexe, par l'exercice I.5.5, donc l'ensemble $\mathcal{P}_r$ est connexe.

• Fermeture : Considérons l'application :

$$\begin{aligned} \Phi\colon \mathcal{M}_n(\mathbb{C}) &\to \mathcal{M}_n(\mathbb{C}) \\ A &\mapsto A^2 - A \end{aligned}$$

4. On peut prendre celle qui nous fait plaisir puisqu'elles sont toutes équivalentes.

Φ est continue, car polynomiale, et de plus, $\mathcal{P} = \Phi^{-1}(\{0\})$. Donc $\mathcal{P}$ est fermé.

Attention, le rang n'est pas continu sur $\mathcal{M}_n(\mathbb{C})$, voir l'exercice I.2.3, avec les matrices $\mathrm{I}_{k,m}$. Toutefois, curieusement, il l'est lorsque l'on se restreint aux matrices de projections. En effet, comme on l'a vu dans l'exercice I.2.2, pour une matrice de projection A, $\mathrm{tr}(A)$ est la somme des valeurs propres, et donc c'est le nombre de 1 dans le spectre. La trace de A est donc égale à son rang puisque $\mathrm{Spec}(A) = \{1; 0\}$.

Or, l'application trace est continue sur $\mathcal{M}_n(\mathbb{C})$ et en particulier sur $\mathcal{P}$, et $\mathcal{P}_r = \mathrm{tr}^{-1}(\{r\})$ est donc un fermé dans $\mathcal{P}$ qui est fermé dans $\mathcal{M}_n(\mathbb{C})$. Donc $\mathcal{P}_r$ est fermé.

I.3 Valeurs propres

Exercice I.3.1. [Valeurs propres, diagonalisation, trace]
Soit A une matrice carrée de rang 1 sur un corps $\mathbb{K}$. Montrer que A est diagonalisable si et seulement si $\mathrm{tr}(A)$ est non nulle.

Soluce. Soit n la taille de A.

Comme A est de rang 1, il vient que $\dim\ker A = n-1$, par la formule du rang. Ainsi, 0 est valeur propre, de multiplicité géométrique $n-1$. Cela implique que 0 est valeur propre de multiplicité algébrique $m_a(0) = n-1$ ou n, puisque l'on a toujours $m_a(\lambda) \geq m_g(\lambda)$..

Le spectre de A est donc constitué de 0 avec multiplicité $n-1$ et, une autre valeur propre, disons, λ. On remarque que λ peut encore être nulle si $m_a(0) = n$. Puisque la trace de A est la somme des valeurs propres avec multiplicité, on a alors $\mathrm{tr}(A) = \lambda$. Deux cas se présentent :

• 1er Cas. Si $\mathrm{tr}(A) \neq 0$: alors $\lambda \neq 0$. Dans ce cas, on a donc que la multiplicité algébrique $m_a(0)$ vaut $n-1$. Celle de λ vaut 1, ce qui implique que $m_g(\lambda)$ vaut également 1. On a donc les égalités

$$m_g(0) = n-1 = m_a(0),\ m_g(\lambda) = 1 = m_a(\lambda).$$

Ceci implique (c'est même équivalent) que A est diagonalisable.

• 2nd Cas. Si $\mathrm{tr}(A) = 0$: alors $\lambda = 0$. Dans ce cas, A n' est pas diagonalisable, puisque $m_a(0) > m_g(0)$.

Ce qui prouve l'assertion.

Exercice I.3.2. Montrer qu'une matrice carrée dont tous les coefficients sont égaux à 1 est diagonalisable.

Soluce. Soit n la taille de A.

La matrice A est de rang 1 puisqu'elle est non nulle et que tous les vecteurs colonnes sont colinéaires (et même carrément égaux).

Par la formule du rang, on en déduit que $\dim\ker A = n-1$; 0 est donc valeur propre de multiplicité géométrique $n-1$. Ceci implique que sa multiplicité algébrique est au moins $n-1$.

Or, la trace de A est la somme des valeurs propres, et comme $\mathrm{tr}(A) = n$, il vient, comme dans l'exercice I.3.1 que n est valeur propre de multiplicité géométrique 1, et 0 est valeur propre de multiplicité géométrique $n-1$. Ainsi, A est diagonalisable.

Remarque : Cet exercice peut être présenté comme une application de l'exercice précédent, il est alors possible de rédiger un énoncé avec deux questions enchaînées.

Exercice I.3.3.
Quel est le déterminant de l'endomorphisme transposition : $M \mapsto {}^tM$ de l'espace des matrices carrées sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$?

Soluce. Soit u l'endomorphisme transposé en question

$$u : \mathcal{M}_n(\mathbb{K}) \to \mathcal{M}_n(\mathbb{K}),\ M \mapsto {}^tM.$$

L'endomorphisme u est involutif, *i. e.* $u^2 = \mathrm{Id}$, donc il annule le polynôme $P = X^2 - 1$.

Celui-ci est scindé à racines simples, puisque[5] $P(X) = (X-1)(X+1)$. Il vient que u est diagonalisable, et son spectre est inclus[6] dans $\{1, -1\}$. Cherchons maintenant les sous-espaces propres.

On a $\ker(u - \mathrm{Id}) = \{M \in \mathcal{M}_n(\mathbb{K}), {}^t M = M\} = S_n(\mathbb{K})$, le sous-espace des matrices symétriques de $\mathcal{M}_n(\mathbb{K})$, de dimension $\dim S_n(\mathbb{K}) = n(n+1)/2$.

De même, $\ker(u + \mathrm{Id}) = \{M \in \mathcal{M}_n(\mathbb{K}), {}^t M = -M\} = A_n(\mathbb{K})$, le sous-espace des matrices antisymétriques de $\mathcal{M}_n(\mathbb{K})$, de dimension donnée par $\dim A_n(\mathbb{K}) = n(n-1)/2$.

Ainsi, 1 et -1 sont bien valeurs propres, de multiplicité géométrique (et donc algébrique, car u est diagonalisable) respective $n(n+1)/2$ et $n(n-1)/2$. Comme le déterminant est le produit des valeurs propres avec leur multiplicité, on en déduit que $\det(u) = (1)^{n(n+1)/2}(-1)^{n(n-1)/2}$.

Le déterminant de la transposée est donc $(-1)^{n(n-1)/2}$.

I.4 Réduction et arithmétique des polynômes

Exercice I.4.1. *[[H2G2] Chap. III Exercice 6.1]
Soit M la matrice diagonale par blocs :

$$M = \begin{pmatrix} A & 0 \\ 0 & A' \end{pmatrix},$$

où A et A' sont des matrices carrées.

1. Montrer que le polynôme minimal de M est égal au ppcm des polynômes minimaux de A et A'.
2. En déduire que M est diagonalisable si et seulement si A et A' le sont.

Soluce. Soit μ_M, μ_A, $\mu_{A'}$ les polynômes minimaux des matrices concernées.

1. • Montrons que μ_M divise $\mathrm{ppcm}(\mu_A, \mu_{A'})$:

5. On est sur un corps de caractéristique différente de 2, sinon, on aurait $1 = -1$.

6. Attention ! P est seulement annulateur, et donc, les valeurs propres sont racines de P mais la réciproque n'est pas claire. D'où l'inclusion, et non l'inégalité.

Rappelons que si P est un polynôme alors $P(M)$ est diagonale par blocs :

$$P(M) = \begin{pmatrix} P(A) & 0 \\ 0 & P(A') \end{pmatrix}$$

Cette égalité résulte facilement de la multiplication par blocs des matrices. On peut commencer par la montrer par récurrence sur k pour le polynôme $P = X^k$, puis achever la preuve pour tout P par linéarité.

En particulier, pour $P = \text{ppcm}(\mu_A, \mu_{A'})$, il vient

$$\text{ppcm}(\mu_A, \mu_{A'})(M) = \begin{pmatrix} \text{ppcm}(\mu_A, \mu_{A'})(A) & 0 \\ 0 & \text{ppcm}(\mu_A, \mu_{A'})(A') \end{pmatrix}$$

Or, μ_A divise $\text{ppcm}(\mu_A, \mu_{A'})$, donc $\text{ppcm}(\mu_A, \mu_{A'})(A) = 0$. Et de même, $\text{ppcm}(\mu_A, \mu_{A'})(A') = 0$.

On en déduit $\text{ppcm}(\mu_A, \mu_{A'})(M) = 0$. Le polynôme $\text{ppcm}(\mu_A, \mu_{A'})$ est donc annulateur de M, et donc, μ_M divise $\text{ppcm}(\mu_A, \mu_{A'})$.

- Montrons que $\text{ppcm}(\mu_A, \mu_{A'})$ divise μ_M.

On a

$$\begin{pmatrix} \mu_M(A) & 0 \\ 0 & \mu_M(A') \end{pmatrix} = \mu_M(M) = 0.$$

On en déduit que $\mu_M(A) = \mu_M(A') = 0$, puis que μ_A divise μ_M et μ'_A divise μ_M.

Il vient donc, par la propriété fondamentale du ppcm, que $\text{ppcm}(\mu_A, \mu_{A'})$ divise μ_M.

Conclusion, comme $\text{ppcm}(\mu_A, \mu_{A'})$ et μ_M sont par construction deux polynômes unitaires qui se divisent l'un l'autre, ils sont égaux.

2. Supposons maintenant A et A' diagonalisables. Alors, il existe deux matrices inversibles P et P' ainsi que deux matrices diagonales D et D' telles que $PAP^{-1} = D$ et $P'A'P'^{-1} = D'$. Un calcul par blocs donne

$$\begin{aligned} \begin{pmatrix} P & 0 \\ 0 & P' \end{pmatrix} M \begin{pmatrix} P & 0 \\ 0 & P' \end{pmatrix}^{-1} &= \begin{pmatrix} P & 0 \\ 0 & P' \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & A' \end{pmatrix} \begin{pmatrix} P^{-1} & 0 \\ 0 & P'^{-1} \end{pmatrix} \\ &= \begin{pmatrix} D & 0 \\ 0 & D' \end{pmatrix}, \end{aligned}$$

Ce qui prouve que la matrice M est diagonalisable.

Réciproquement, si M est diagonalisable, alors le polynôme minimal μ_M est scindé simple. Or, μ_A et $\mu_{A'}$ divisent μ_M, par la question qui précède. Ainsi, μ_A et $\mu_{A'}$ sont également scindés simples et donc A et A' sont diagonalisables.

Exercice I.4.2.
Soit E un espace vectoriel de dimension finie sur un corps $\mathbb{K}$ et u un endomorphisme de E. On considère un sous-espace F de E stable par u.

1. Montrer qu'il existe un unique polynôme unitaire D_F de $\mathbb{K}[X]$ tel que
$$\forall P \in \mathbb{K}[X],\ D_F | P \Leftrightarrow \mathrm{Im}(P(u)) \subset F.$$
2. Montrer que D_F divise le polynôme minimal μ_u de u.

Soluce.

Idée clef. Dans cet exercice, on construit un polynôme qui généralise le polynôme minimal ; le polynôme minimal étant un cas particulier lorsque le sous-espace F est nul.

1. Soit I l'ensemble des P tels que $\mathrm{Im}\big(P(u)\big) \subset F$. Montrons que I est un idéal (il est clairement non vide).
   • Montrons que I est stable par $+$:
   On suppose P et Q dans I. Par définition de I, on a donc
   $$\mathrm{Im}\big(P(u)\big) \subset F \text{ et } \mathrm{Im}\big(Q(u)\big) \subset F.$$
   Soit $x \in E$, on a
   $$\big(P(u) + Q(u)\big)(x) = P(u)(x) + Q(u)(x),$$
   où $P(u)(x) \in F$ et $Q(u)(x) \in F$.
   Comme F est un sous-espace vectoriel, on a finalement
   $$\big(P(u) + Q(u)\big)(x) \in F.$$
   • Soit $A \in \mathbb{K}[X]$ et $P \in I$. Montrons que $AP \in I$.
   Soit $x \in E$, on a
   $$(AP)(u)(x) = A(u)\big(P(u)(x)\big) \in A(u)(F) \subset F,$$
   car F est stable par u.
   Ainsi, I est un idéal de $\mathbb{K}[X]$. Comme l'anneau $\mathbb{K}[X]$ est principal, I s'écrit $I = D_F\mathbb{K}[X]$, pour un polynôme D_F. De plus, D_F est unique s'il est choisi unitaire. D'où l'assertion.
2. Montrons que D_F divise μ_u.
   On a $\mathrm{Im}\big(\mu_u(u)\big) = \mathrm{Im}(0) = 0 \subset F$. Donc $\mu_u \in I$ et D_F divise μ_u.

Exercice I.4.3. On considère une matrice A de $\mathcal{M}_n(\mathbb{C})$ et un polynôme P de $\mathbb{C}[X]$.

1. Montrer que λ est valeur propre de A si et seulement si $P(\lambda)$ est valeur propre de $P(A)$.
2. Montrer que $P(A)$ est inversible si et seulement si le polynôme P est premier avec le polynôme minimal μ_A de A.
3. On décompose le polynôme minimal de A

$$\mu_A = \prod_{i=1}^{k} (X - \lambda_i)^{m_i},$$

où les λ_i sont deux à deux distincts. Montrer que $P(A)$ est nilpotent si et seulement si P est divisible par $\prod_{i=1}^{k}(X - \lambda_i)$.

Soluce. 1. Comme on travaille sur le corps des complexes, A est trigonalisable : $A = QTQ^{-1}$ avec T triangulaire supérieure et $Q \in \mathrm{GL}_n(\mathbb{C})$. Il vient que $P(A) = QP(T)Q^{-1}$.

Les valeurs propres de A sont sur la diagonale de T et de même, les valeurs propres de $P(A)$ sont sur la diagonale de $P(T)$. Or, comme T est triangulaire, la diagonale de $P(T)$ est constituée des $P(\lambda)$ où λ parcourt la diagonale de T. Ceci prouve l'équivalence demandée.

2. On sait qu'une matrice est inversible si et seulement si 0 n'appartient pas à son spectre. On en conclut, par la question qui précède, que $P(A)$ est inversible si et seulement si $P(\lambda) \neq 0$ pour tout λ dans le spectre de A. Mais ceci est équivalent à dire que P et μ_A n'ont pas de racine commune, et donc sont premiers entre eux.

3. On sait (par Cayley-Hamilton) qu'une matrice est nilpotente si et seulement si son spectre est réduit à 0. D'après la première question, $P(A)$ est nilpotente si et seulement si $P(\lambda) = 0$ pour tout λ dans le spectre de A. Comme on choisit les λ distincts, cela donne bien que P est multiple de $\prod_{i=1}^{k}(X - \lambda_i)$.

Exercice I.4.4.
On considère deux matrices A et B de $\mathcal{M}_n(\mathbb{C})$. Soit $\Phi_{A,B}$ l'endomorphisme

$$\Phi_{A,B} : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_n(\mathbb{C}),\ M \mapsto AM + MB.$$

1. Soit $L_A = \Phi_{A,0}$ l'endomorphisme de multiplication à gauche par A. Montrer que $(L_A)^k = L_{A^k}$ et en déduire que pour tout polynôme P de $\mathbb{C}[X]$, on a $P(L_A) = L_{P(A)}$.
2. Montrer que L_A est nul si et seulement si A est nulle.
3. Montrer que A et L_A ont même polynôme minimal, puis, qu'ils ont même spectre.
4. Soit $R_B = \Phi_{0,B}$ l'endomorphisme de multiplication à droite par B. Montrer que L_A et R_B commutent et que $\Phi_{A,B} = L_A + R_B$.
5. En déduire que les valeurs propres de $\Phi_{A,B}$ sont de la forme $\lambda_i + \rho_j$, avec λ_i, resp. ρ_j, valeur propre de A, resp. de B.

Soluce. 1. Il suffit de vérifier que pour toute matrice M, on a

$$(L_A)^k(M) = L_A\big(\cdots(L_A(M)\cdots\big) = A^kM = L_{A^k}(M).$$

Par linéarité, on en déduit facilement $P(L_A) = L_{P(A)}$.

2. Tout d'abord, si A est nulle, alors L_A l'est également. Réciproquement, si L_A est nulle, on a $AM = 0$ pour tout M dans $\mathcal{M}_n(\mathbb{C})$. En particulier, pour $M = A^*$, il vient $\mathrm{tr}(AA^*) = 0$, et comme $\mathrm{tr}(AA^*)$ est une norme, on obtient $A = 0$.

3. D'après les questions qui précèdent,

$$P(A) = 0 \Leftrightarrow L_{P(A)} = 0 \Leftrightarrow P(L_A) = 0.$$

En particulier, A et L_A ont même polynôme minimal.

Comme le spectre est l'ensemble des racines du polynôme minimal, ils ont même spectre.

4. Cela provient des égalités suivantes, valables pour tout M

$$(AM)B = A(MB),\ \Phi_{A,B} = AM + MB = L_A(M) + R_B(M).$$

5. Soit $\{\lambda_i, i \in I\}$, resp. $\{\rho_j, j \in J\}$, le spectre de A, resp. B. Alors, $\{\lambda_i, i \in I\}$ est le spectre de L_A, et de même, on montre que $\{\rho_j, j \in J\}$ est le spectre de R_B.

De plus, comme L_A et R_B commutent, ils sont trigonalisables dans une base commune. Donc, les valeurs propres de $L_A + R_B$ sont les $\lambda_i + \rho_j$, $i \in I$, $j \in J$. C'est donc le spectre de $\Phi_{A,B}$.

Exercice I.4.5.
On considère deux matrices A et B de $\mathcal{M}_n(\mathbb{C})$. On note μ_A le polynôme minimal de A. Montrer que les conditions suivantes sont équivalentes :
(i) A et B possèdent une valeur propre commune,
(ii) il existe une matrice M non nulle de $\mathcal{M}_n(\mathbb{C})$ telle que $AM = MB$,
(iii) la matrice $\mu_A(B)$ est non inversible.

Soluce. Etudions de plus près l'assertion (ii). On note que $M \mapsto AM - MB$ définit un endomorphisme $\varphi_{A,B}$ de $\mathcal{M}_n(\mathbb{C})$.

L'assertion (ii) est équivalente au fait que $\varphi_{A,B}$ est non injective, c'est-à-dire que 0 est valeur propre de $\varphi_{A,B}$. Or, d'après l'exercice I.4.4, les valeurs propres de $\varphi_{A,B}$ sont de la forme $\lambda_i - \rho_j$, où les λ_i, resp. les ρ_j, sont les valeurs propres de A, resp. de B.

Il devient clair que (i) est équivalent à (ii).

Maintenant, par l'exercice I.4.3, $\mu_A(B)$ est non inversible si et seulement si μ_A et μ_B ont un facteur commun, et donc, sur $\mathbb{C}$, une racine commune. Il en résulte que A et B ont une valeur propre commune.

Exercice I.4.6.
Soit A une matrice de $\mathcal{M}_n(\mathbb{C})$ et P un polynôme de $\mathbb{C}[X]$. On note

$$\mu_A = \prod_{i=1}^{s} (X - \lambda_i)^{m_i}$$

la décomposition du polynôme minimal μ_A de A, avec les λ_i deux à deux distincts.
On veut montrer que les conditions suivantes sont équivalentes :
(i) $P(A)$ est diagonalisable
(ii) $P^{(k)}(\lambda_i) = 0$, pour tout i, $1 \leq i \leq s$, et tout k, $1 \leq k \leq m_i - 1$.

1. On suppose ici que A est nilpotente d'indice m. Montrer que la famille $(A, A^2, \cdots, A^{m-1})$ est linéairement indépendante, et en déduire l'équivalence dans ce cas.
   On pourra penser à l'unicité dans la décomposition de Dunford.
2. On suppose ici que $A - \lambda \mathrm{I}_n$ est nilpotente d'indice m pour un λ dans $\mathbb{C}$. Montrer l'équivalence dans ce cas.
3. En utilisant le lemme des noyaux, montrer l'équivalence dans le cas général.

Soluce. 1. On suppose

$$\sum_{i=1}^{m-1} \beta_i A^i = 0,\ \alpha_i \in \mathbb{C}$$

Il en résulte que le polynôme $Q = \sum_{i=1}^{m-1} \beta_i X^i$ annule A. Il est donc

divisible par $\mu_A = X^m$. Par un argument de degré, on a forcément $Q = 0$ et donc $\beta_i = 0$ pour tout i. Ceci prouve l'indépendance des A^i.

Montrons maintenant l'équivalence. On pose $P = \sum_{i=0}^{r} \alpha_i X^i$.

(ii)⇒(i) Le spectre de A est réduit à 0. Donc, l'assertion (ii) s'écrit $P^{(k)}(0) = 0$, pour tout k, $1 \leq k \leq m-1$. Par Taylor, cela revient à dire que les coefficients de P en X^k, $1 \leq k \leq m-1$, sont tous nuls. Compte tenu du fait que $A^k = 0$ pour $k \geq m$, on a $P(A) = \alpha_0 \mathrm{I}_n$. Ce qui implique (i).

Réciproquement, on suppose $P(A)$ diagonalisable. Alors, la matrice $\alpha_0 \mathrm{I}_n + \sum_{k=1}^{m-1} \alpha_k A^k$ est diagonalisable, donc elle s'écrit $D + 0$ avec D diagonalisable et 0 nilpotente. Or, la matrice $\sum_{k=1}^{m-1} A^k$ est nilpotente (on sait que la somme de matrices nilpotentes qui commutent est encore nilpotente), et elle commute bien sûr à $\alpha_0 \mathrm{I}_n$. Par l'unicité de la décomposition de Dunford, il vient $\sum_{k=1}^{m-1} \alpha_k A^k = 0$ et donc tous les α_k sont nuls, pour tout k, $1 \leq k \leq m-1$. Ceci implique (ii).

2. Dans ce cas, en appliquant le résultat précédent à $B = A - \lambda \mathrm{I}_n$, qui est nilpotente d'indice m, il vient que $P(A)$ est diagonalisable si et seulement si $Q(B)$ est diagonalisable, avec $Q(X) = P(X + \lambda)$, si et seulement si $Q^{(k)}(0) = 0$, pour tout k, $1 \leq k \leq m-1$. Ceci est équivalent à $P^{(k)}(\lambda) = 0$, pour tout k, $1 \leq k \leq m-1$.

3. Afin de pouvoir utiliser les restrictions, on va montrer l'équivalence dans le cadre des endomorphismes plutôt que celui des matrices. Soit donc u un endomorphisme de l'espace $\mathbb{C}^n$ dont la matrice dans la base canonique est donnée par A.

   Rappel. Rappelons le lemme des noyaux : on sait que l'on a une décomposition de $\mathbb{C}^n$ en somme directe de sous-espaces caractéristiques K_i, avec $K_i = \ker(u - \lambda_i \mathrm{Id})^{m_i}$. Les sous-espaces K_i sont stables par u, et les restrictions $(u - \lambda_i \mathrm{Id})_{K_i}$ sont nilpotentes d'indice m_i.

   En particulier, les K_i sont stables par $P(u)$. Par l'exercice I.4.1, on sait que $P(u)$ est diagonalisable si et seulement si toutes ses restrictions $P(u)_{K_i}$ le sont [7].

   Or, on a $P(u)_{K_i} = P(u_{K_i})$, et par construction, $u_{K_i} - \lambda_i \mathrm{Id}$ est nilpotente d'indice m_i. Il n'y a plus qu'à appliquer le point précédent.

   Remarque. L'ensemble des polynômes P tels que $P(A)$ est nilpotent forme un idéal de $\mathbb{C}[X]$, voir exercice I.4.3. Il n'en est pas de même pour l'ensemble des polynômes P tels que $P(A)$ est diagonalisable.

7. Pour être plus précis, il faut juste ajouter une petite récurrence à l'exercice.

Exercice I.4.7 (La réduction au secours des polynômes)**.**
Après un cursus de licence en mathématiques, on pourrait croire que la réduction doit beaucoup à la connaissance des polynômes. C'est un peu vrai, mais souvent, c'est l'inverse qui se passe : dans l'étude des racines d'un polynôme, la réduction a su devenir indispensable. Voici une histoire qui illustre cet adage.
Soit $P = X^n + \sum_{i=0}^{n-1} a_i X^i$ un polynôme unitaire de degré n dans $\mathbb{C}[X]$. On pose

$$R = \mathrm{Max}\{|a_0|, 1+|a_1|, 1+|a_2|, \cdots, 1+|a_{n-1}|\}.$$

Le but de l'exercice est de montrer que si λ est une racine de P, alors $\lambda \leq R$.

1. On considère la matrice compagnon C_P donnée par

$$C_P = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & 0 & -a_1 \\ 0 & 1 & \ddots & 0 & 0 & -a_2 \\ \vdots & 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 1 & 0 & -a_{n-2} \\ 0 & 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix}.$$

   Montrer que son polynôme caractéristique est $(-1)^n P$. En déduire que toute racine de P est valeur propre de C_P.

2. Soit λ une racine de P, donc une valeur propre de C_P, et soit $v = (v_i)_{1\leq i\leq n} \in \mathbb{C}^n$ un vecteur propre associé à λ. Montrer, en considérant $\mathrm{Max}_i\{|v_i|\}$ que $|\lambda| \leq R$.

Soluce. 1. Il suffit de calculer $\det(C_P - X\mathrm{I}_n)$. En développant ce déterminant par rapport à la première colonne, on obtient

$$\det(C_P - X\mathrm{I}_n) = -X\det(A_1 - X\mathrm{I}_{n-1}) - \det A_2,$$

où A_1 est la matrice compagnon du polynôme $X^{n-1} + a_{n-1}X^{n-2} + \cdots + a_1$ et A_2 est la matrice

$$A_2 = \begin{pmatrix} 0 & 0 & \cdots & \cdots & -a_0 \\ 1 & -X & 0 & \cdots & -a_2 \\ 0 & 1 & -X & \cdots & -a_3 \\ \vdots & \cdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 & -X - a_{n-1} \end{pmatrix}$$

On trouve $\det(A_2) = (-1)^{n-2}(-a_0) = (-1)^{n-1}a_0$ en développant par rapport à la première ligne.

A partir de ces identités et du cas évident $n = 1$, on prouve par récurrence sur n que le polynôme caractéristique de C_P est

$$-X(-1)^{n-1}(X^{n-1} + a_{n-1}X^{n-2} + \cdots + a_1) - (-1)^{n-1}a_0 = (-1)^n P.$$

La dernière assertion est claire.

2. Soit v_k tel que le maximum des $|v_i|$ est atteint pour i de 1 à n. Comme v est non nul (c'est un vecteur propre !), on a $v_k \neq 0$.
En regardant l'égalité $C_P v = \lambda v$ sur la k-ième coordonnée, on trouve

$$v_{k-1} - a_{k-1}v_n = \lambda v_k, \text{ si } k > 1, \text{ et } -a_0 v_n = \lambda v_k, \text{ si } k = 1.$$

Supposons $k > 1$. Il vient, en divisant par v_k et en utilisant l'inégalité triangulaire

$$|\lambda| \leq |\frac{v_{k-1}}{v_k}| + |a_{k-1}| \cdot |\frac{v_n}{v_k}| \leq 1 + |a_{k-1}| \leq R.$$

Le cas $k = 1$ est analogue.

I.5 Topologie matricielle

Exercice I.5.1.
On pose ici $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Montrer que $\mathrm{GL}_n(\mathbb{K})$ est un ouvert dense de $\mathcal{M}_n(\mathbb{K})$.

Soluce. On a l'égalité $\mathrm{GL}_n(K) = \det^{-1}(\mathbb{K}^*)$, c'est donc l'image réciproque d'un ouvert par le déterminant (de $\mathcal{M}_n(\mathbb{K})$ dans $\mathbb{K}$) qui est continu (voir le rappel dans l'exercice I.5.2), donc c'est un ouvert.

Densité : Soit $A \in \mathcal{M}_n(\mathbb{K})$. Montrons qu'il existe une suite A_k de matrices de $\mathrm{GL}_n(\mathbb{K})$ qui converge vers A.
Si A est inversible, la suite constante $A_k = A, \forall k \in \mathbb{N}$ convient.
Si A n'est pas inversible, alors, 0 est valeur propre de A. Comme l'ensemble des valeurs propres de A est fini, donc discret, il existe un réel non nul s, tel que $0 < |\lambda| < 1/s$ implique que λ n'est pas valeur propre de A. La suite $(A - (1/k)\mathrm{I}_n)_{k \geq s}$ est donc bien une suite de $\mathrm{GL}_n(\mathbb{K})$ (sinon $1/k$ serait valeur propre). Cette suite tend vers A, car $(1/k)\mathrm{I}_n$ tend vers 0, par exemple pour la norme sup.

Exercice I.5.2 (Continuité de l'application $A \mapsto \chi_A$)**.**
Montrer que l'application de $\mathcal{M}_n(\mathbb{C})$ dans le $\mathbb{C}$-espace vectoriel $\mathbb{C}[X]_n$ des polynômes de degré inférieur à n, qui, à la matrice A, associe son polynôme caractéristique χ_A, est continue. En déduire que $\chi_{AB} = \chi_{BA}$, pour tout A, B, dans $\mathcal{M}_n(\mathbb{C})$.

Soluce.

Rappel. Tout d'abord, il faut bien comprendre ce que signifie la continuité dans ce contexte. On rappelle que les espaces de dimension finie $\mathcal{M}_n(\mathbb{C})$ et $\mathbb{C}[X]_n$ sont dotés d'une topologie dont une base d'ouverts est donnée par l'ensemble des boules pour une norme. Comme toutes les normes sont équivalentes, on sait que la topologie ne dépend pas de la norme choisie. Comme les formes linéaires sont continues, il vient que toute fonction polynomiale de $\mathcal{M}_n(\mathbb{C})$ dans $\mathbb{C}$, c'est-à-dire polynomiale en les coordonnées de la matrice, est continue. Pour montrer qu'une fonction de $\mathcal{M}_n(\mathbb{C})$ dans $\mathbb{C}[X]_n$ est continue, il suffit de montrer que toute coordonnée (dans la base canonique de l'espace des polynômes) de cette fonction est continue, et cela sera vrai si ces coordonnées sont polynomiales.

Fixons une matrice $A = (a_{ij})$ dans $\mathcal{M}_n(\mathbb{C})$. L'application qui, à t dans $\mathbb{C}$, associe $\det(A - t\mathrm{I}_n)$ est polynomiale en t et a_{ij}, $1 \leq i, j \leq n$.

Les coefficients en les t^k du polynôme sont alors des polynômes en les a_{ij} :

$$\det(A - t\mathrm{I}_n) = \sum_{k=0}^{n} P_k(a_{ij}) t^k$$

La fonction χ, donnée par $A \mapsto \chi_A$, s'écrit dans les bases canoniques respectives de $\mathcal{M}_n(\mathbb{C})$ et de $\mathbb{C}[X]_n$ comme

$$(a_{ij})_{ij} \mapsto \big(P_k(a_{ij})\big)_k.$$

C'est une fonction polynomiale donc continue.

Montrons maintenant la dernière assertion. On commence par un cas particulier.

• Si A est inversible.

Alors dans ce cas, l'invariance du polynôme caractéristique par conjugaison donne :

$$\chi_{AB} = \chi A^{-1}(AB)A = \chi_{BA}.$$

• Cas général.

Considérons l'application f qui, à M dans $\mathcal{M}_n(\mathbb{C})$, associe $\chi_{MB} - \chi_{BM}$ dans $\mathbb{C}[X]_n$. Elle est continue, cela provient du fait que χ est continue, et que la multiplication à gauche, ou à droite, par B est continue (elle est linéaire !). De plus, d'après ce qui précède, f est nulle sur $\mathrm{GL}_n(\mathbb{C})$.

Or, par l'exercice I.5.1, $\mathrm{GL}_n(\mathbb{C})$ est dense dans $\mathcal{M}_n(\mathbb{C})$. Il en résulte, par densité, que f est nulle sur tout $\mathcal{M}_n(\mathbb{C})$.

D'où l'assertion.

Remarque. On peut se demander quels sont les polynômes P_k pour tout nombre k. Pour $k = 0$, c'est le déterminant. Pour $k = n - 1$, c'est la forme linéaire $(-1)^{n-1}\operatorname{tr}(A)$. Pour le cas général, on montre que P_k est la somme des mineurs diagonaux de taille k, multipliée par $(-1)^k$.

Exercice I.5.3 (Discriminant du polynôme caractéristique)**.**
Trouver une matrice de $\mathcal{M}_2(\mathbb{R})$ qui n'est pas dans l'adhérence de l'ensemble des matrices diagonalisables.

Soluce.

Idée clef. Contrairement au cas complexe, l'ensemble des matrices réelles diagonalisables n'est pas dense dans l'espace des matrices réelles. Le fait d'avoir un polynôme caractéristique scindé sur $\mathbb{R}$ est trop restrictif. Comme on est en dimension 2, les polynômes caractéristiques sont de degré 2, et le fait qu'ils soient non scindé se voit grâce à l'aide d'une fonction continue : le discriminant.

La matrice que l'on cherche n'est pas limite d'une suite de matrices diagonalisables.

Pour commencer, notons que toute matrice dont le polynôme caractéristique est scindé sur $\mathbb{R}$ est soit diagonalisable, soit limite d'une suite de matrices diagonalisables.

En effet : si une matrice A a un polynôme caractéristique scindé sur $\mathbb{R}$, mais n'est pas diagonalisable, alors elle a une valeur propre double λ et peut s'écrire $A = P\begin{pmatrix}\lambda & a\\ 0 & \lambda\end{pmatrix}P^{-1}$. On considère alors la suite de matrices diagonalisables $(M_n)_n$ avec $M_n = P\begin{pmatrix}\lambda & a\\ 0 & \lambda+\frac{1}{n}\end{pmatrix}P^{-1}$. La suite $(M_n)_n$ tend vers A et ainsi, A est limite d'une suite de matrices diagonalisables.

Nous allons alors montrer que toute matrice de $\mathcal{M}_2(\mathbb{R})$ dont le polynôme caractéristique n'est pas scindé ne peut être vue comme limite de matrices diagonalisables.

On construit l'application ρ qui à une matrice 2×2 associe le discriminant Δ de son polynôme caractéristique :

$$\rho : \begin{pmatrix}a & b\\ c & d\end{pmatrix} \mapsto \Delta(X^2 - (a+d)X + (ad-bc)) = (a+d)^2 - 4(ad-bc).$$

L'application ρ est polynomiale en les coefficients de la matrice $\begin{pmatrix}a & b\\ c & d\end{pmatrix}$, donc continue de $\mathcal{M}_2(\mathbb{R})$ dans $\mathbb{R}$.

Supposons qu'il existe une suite $(M_n)_{n\in\mathbb{N}}$ de matrices diagonalisables dans $\mathbb{R}$ qui tende vers une matrice M de $M_2(\mathbb{R})$ non diagonalisable. Par continuité de ρ, $(\rho(M_n))_{n\in\mathbb{N}}$ tend alors vers $\rho(M)$. Or , pour tout n, M_n est

diagonalisable ; son polynôme caractéristique est donc scindé de degré égal à 2. Cela implique que le discriminant de M_n vérifie $\rho(M_n) \geqslant 0$, pour tout n, et, $\mathbb{R}^+$ étant fermé, on a alors $\rho(M) \geqslant 0$, et M est diagonalisable. On obtient une contradiction avec l'hypothèse de départ.

Finalement, on peut donc choisir n'importe quelle matrice de $\mathcal{M}_2(\mathbb{R})$ dont le polynôme caractéristique n'est pas scindé sur $\mathbb{R}$.

Par exemple $M = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, dont le polynôme caractéristique $\chi_M(X)$ est $X^2 + 1$.

Exercice I.5.4 (Une application non continue issue de Dunford)**.**
Soit $n > 1$. On considère l'application

$$\varphi : \ \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_n(\mathbb{C}), \ M \mapsto D,$$

où D est la composante diagonalisable de M dans la décomposition de Dunford. Montrer que φ n'est pas continue.

Soluce. On sait que l'ensemble $\mathcal{D}_n(\mathbb{C})$ des matrices diagonalisables dans $\mathbb{C}$ est dense dans $\mathcal{M}_n(\mathbb{C})$.

Comme l'application φ se restreint en l'identité sur $\mathcal{D}_n(\mathbb{C})$, si elle était continue, alors par densité, ce serait l'application identité sur $\mathcal{M}_n(\mathbb{C})$. Ce qui est absurde, puisque cela signifierait que toute matrice est diagonalisable, ça se saurait ! On note le garde-fou $n > 1$, qui permet d'affirmer cette dernière assertion.

Exercice I.5.5. *[$\mathrm{GL}_n(\mathbb{C})$ est connexe. [H2G2] Chapitre II Proposition 1.4] On se propose de montrer que $\mathrm{GL}_n(\mathbb{C})$ est connexe par arc. Pour cela, on se donne deux matrices A et B de $\mathrm{GL}_n(\mathbb{C})$ et on veut construire un chemin de $\mathrm{GL}_n(\mathbb{C})$ qui relie A à B.

1. Montrer que l'application qui, à un complexe t, associe le complexe $\det(A(1-t)+tB)$, définit un polynôme non nul.
2. En déduire qu'il existe une courbe continue γ dans $\mathbb{C}$ qui relie 0 à 1 et telle que pour tout t de γ, $\det(A(1-t)+tB) \neq 0$.
3. Conclure.

Soluce.

1. Par définition, $\det : A \mapsto \sum_{\sigma \in \mathcal{S}_n} \epsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n}$ est polynomiale en les coefficients de A.
   Par conséquent, $P : t \mapsto \det(A(1-t)+tB)$ est polynomiale en t. De plus, pour $t = 0$, on a $P(0) = \det(A) \neq 0$ car $A \in \mathrm{GL}_n(\mathbb{C})$.
   Donc, la fonction polynomiale P n'est pas nulle. Le polynôme P associé est donc non nul.
2. Le polynôme P étant non nul, il possède un nombre fini de zéros dans $\mathbb{C}$. On peut donc tracer dans $\mathbb{C}$ un chemin γ reliant 0 à 1 qui évite les zéros de P.

A l'oral, un dessin peut suffire. Sinon : on peut dire la chose suivante : dans le plan complexe, l'ensemble $Z = \{z_i\}$ constitué des zéros de P est fini, donc discret pour la topologie normique de $\mathbb{C}$. On peut donc dessiner des disques fermés D_i disjoints autour des z_i. Il suffit de tracer un segment allant de 0 à 1 qui « contourne » les D_i lorsque ce segment rencontre D_i. Le contournement se fait par le cercle qui forme la frontière de D_i.

Comme $Z = \{z, P(z) = 0\}$ est fini dans $\mathbb{C}$, c'est un ensemble discret pour la topologie de $\mathbb{C}$. Donc, pour chaque $t_i \in Z \cap [0;1]$, il existe un disque ouvert D_i tel que $D_i \cap Z = \{t_i\}$. Soit C_i le cercle frontière de D_i, il intersecte $[0;1]$ en t_i^- et t_i^+. On construit alors un chemin continu γ par concaténation de : $\gamma' = [0, t_1^-]$; $\gamma'' = \overset{\frown}{(t_1^-; t_1^+)}$, ...

3. On veut montrer que $\mathrm{GL}_n(\mathbb{C})$ est connexe (et même connexe par arcs) en construisant un arc continu dans $\mathrm{GL}_n(\mathbb{C})$ et reliant A à B. Soit α de $[0;1]$ dans $\mathrm{GL}_n(\mathbb{C})$ qui envoie t sur $\big(1-\gamma(t)\big)A + \gamma(t)B$. L'application est visiblement continue puique γ l'est et, par construction, est bien à valeur dans $\mathrm{GL}_n(\mathbb{C})$, puisque l'on a pris soin de faire en sorte que le déterminant de la matrice associée ne s'annule pas. On a donc construit un arc connexe reliant A à B, dans $\mathrm{GL}_n(\mathbb{C})$.

Exercice I.5.6 (Application de la décomposition polaire)**.**
On considère la norme quadratique n_2 de l'espace vectoriel euclidien $\mathbb{R}^n$ et sa norme subordonnée N_2 sur les matrices réelles de $\mathcal{M}_n(\mathbb{C})$. Le but de l'exercice est de montrer que pour tout matrice A de $\mathrm{GL}_n(\mathbb{C})$,

$$N_2(A)^2 = \rho({}^tAA),$$

où ρ désigne le rayon spectral, c'est-à-dire le maximum des modules des valeurs propres.

Rappel. On rappelle que pour toute matrice inversible A, on a la décomposition polaire $A = OS$, avec O orthogonale et S symétrique définie positive. La matrice S étant l'unique « racine carrée »de la matrice symétrique définie positive ${}^tAA$, *i. e.* $S^2 = {}^tAA$.

1. Montrer que pour toute matrice M de $\mathcal{M}_n(\mathbb{C})$ et toute matrice orthogonale O, $N_2(OM) = N_2(M)$.
2. Soit S une matrice symétrique définie positive.
   (a) Montrer que pour tout x de norme 1 dans $\mathbb{R}^n$, $n_2(Sx) \leq \rho(S)$.
   (b) En déduire que $N_2(S) = \rho(S)$.
3. En déduire que $N_2(A) = \rho(S)$.
4. Soit T une matrice symétrique définie positive et R sa racine carrée (définie positive). Montrer $\rho(T) = \rho(R)^2$.
5. Conclure le résultat proposé.

Soluce. 1. Par définition, n_2 est invariante par O. Il vient

$$N_2(OM) = \sup_{n_2(x)=1} n_2(OMx) = \sup_{n_2(x)=1} n_2(Mx) = N_2(M).$$

2. (a) Soit x de norme 1 dans $\mathbb{R}^n$. Comme S est sympétrique réelle, on peut décomposer $x = \sum_i x_i v_i$ dans une base orthonormale (v_i) de vecteurs propres pour S, avec v_i associé à la valeur propre réelle positive λ_i :

$$n_2(Sx)^2 = n_2(\sum_i \lambda_i x_i v_i)^2 = \sum_i \lambda_i^2 x_i^2 \leq \rho(S)^2 \sum_i x_i^2 = \rho(S)^2.$$

   Ce qui prouve l'assertion demandée.

   (b) Il suffit de montrer que la borne est atteinte. Soit k telle que v_k est de valeur propre $\lambda_k := \max_i \lambda_i$, c'est-à-dire $\lambda_k = \rho(S)$, puisque les λ_i sont positifs. On a bien

$$n_2(Sv_k) = n_2(\lambda_k v_k) = \lambda_k n_2(v_k) = \lambda_k.$$

3. On a donc, d'après ce qui précède

$$N_2(A) = N_2(OS) = N_2(S) = \rho(S).$$

4. Comme R est symétrique définie positive, ses valeurs propres λ_i sont réelles strictement positives. Les valeurs propres de $T = R^2$ sont donc les λ_i^2. Comme la fonction $x \mapsto x^2$ est croissante sur $\mathbb{R}^+$, on a donc $\mathrm{Max}_i\{\lambda_i^2\} = \left(\mathrm{Max}_i\{\lambda_i\}\right)^2$, et donc $\rho(T) = \rho(R)^2$, comme désiré.
5. D'après ce qui précède,

$$N_2(A)^2 = N_2(S)^2 = N_2(S^2) = N_2({}^tAA) = \rho({}^tAA).$$

Exercice I.5.7 (Une preuve topologique de Cayley-Hamilton)**.**
Montrer que l'ensemble $\mathcal{D}$ des matrices diagonalisables de $\mathcal{M}_n(\mathbb{C})$ est dense dans $\mathcal{M}_n(\mathbb{C})$. En déduire une preuve express du théorème de Cayley-Hamilton sur $\mathbb{C}$.

Soluce. Prenons une matrice complexe quelconque A et montrons que l'on peut approcher A par une suite de matrices diagonalisables.
Comme on travaille sur $\mathbb{C}$, A est trigonalisable, c'est à dire que l'on a $A = PTP^{-1}$, avec

$$T = \mathrm{diag}(\lambda_1, \cdots, \lambda_1, \lambda_2, \cdots, \lambda_2, \cdots, \lambda_k, \cdots, \lambda_k) + N,$$

où N est strictement triangulaire supérieure et où les λ_i sont distincts [8].

On peut trouver ϵ assez petit pour que

$$T_\epsilon := T + \mathrm{diag}(\frac{\epsilon}{1}, \frac{\epsilon}{2}, \ldots, \frac{\epsilon}{n})$$

n'ait que des entrées distinctes sur la diagonale principale. Montrons par exemple que $\epsilon = \min_{i \neq j} |\lambda_i - \lambda_j|$ convient.

Effectivement, si, par l'absurde, on a $\lambda_i + \frac{\epsilon}{s} = \lambda_j + \frac{\epsilon}{r}$, avec $1 \leq i \leq j \leq k$ et $1 \leq s < r \leq n$. Si $i = j$, c'est impossible car $s < r$. Si $i \neq j$, on a

$$|\lambda_i - \lambda_j| = \epsilon|\frac{1}{s} - \frac{1}{r}| \leq \epsilon(1 - \frac{1}{r}) < \epsilon,$$

d'où la contradiction.

Ainsi, T_ϵ est une matrice triangulaire supérieure avec des valeurs propres distinctes, donc diagonalisable et qui tend vers T quand ϵ tend vers zéro. Ainsi, $PT_\epsilon P^{-1}$ est diagonalisable et tend vers A, par continuité du produit matriciel.

Preuve express de Cayley-Hamilton :

L'application

$$\begin{aligned} \psi \colon \mathcal{M}_n(\mathbb{C}) &\to \mathcal{M}_n(\mathbb{C}) \\ A &\mapsto \chi_A(A) \end{aligned}$$

8. Attention, cela ne résulte pas tout à fait de théorème de trigonalisation, car on a assemblé les λ_i entre eux. C'est en fait une conséquence du lemme des noyaux : on décompose l'espace en somme directe de sous-espaces sur lesquels on n'a qu'une valeur propre λ_i.

est continue. Montrons cette assertion. On sait d'une part, que l'application $A \mapsto \chi_A$ est continue [9] de $\mathcal{M}_n(\mathbb{C})$ dans l'espace vectoriel (normé) $\mathbb{C}_n[X]$ des polynômes sur $\mathbb{C}$ de degré inférieur à n. D'autre part, comme $A \mapsto A^k$ est continue de $\mathcal{M}_n(\mathbb{C})$ dans lui-même, il vient facilement que $(P, A) \mapsto P(A)$ est continue de $\mathbb{C}_n[X] \times \mathcal{M}_n(\mathbb{C})$ dans $\mathcal{M}_n(\mathbb{C})$. L'application ψ est donc continue comme composée de deux applications continues.

Montrons que si A est diagonalisable, alors $\chi_A(A) = 0$.
On suppose d'abord que $A = D = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$. Dans ce cas, $\chi_D = \prod_{i=1}^{n}(X - \lambda_i)$. Il vient par le calcul par blocs que $\chi_D(D)$ est une matrice diagonale et ses coefficients diagonaux sont sous la forme $\chi_D(\lambda_i)$. Ils sont donc tous nuls, ce qui prouve que $\chi_D(D)$ est la matrice nulle.
Dans le cas général, si A est diagonalisable, alors $A = PDP^{-1}$ pour une matrice inversible P. On a

$$\chi_A(A) = \chi_D(A) = \chi_D(PDP^{-1}) = P\chi_D(D)P^{-1} = P0_nP^{-1} = 0_n.$$

Ainsi, ψ est nulle sur l'ensemble dense des matrices diagonalisables. Par continuité, elle est donc nulle partout sur $\mathcal{M}_n(\mathbb{C})$. Donc $\chi_A(A) = 0$ pour tout A.

9. Les coefficients des X^k sont, au signe près, les sommes des mineurs diagonaux de taille $n - k$. La continuité provient donc de la continuité du déterminant.

Exercice I.5.8. Montrer que les groupes $\mathrm{O}_n(\mathbb{R})$ et $\mathrm{U}_n(\mathbb{C})$ sont compacts.

Soluce. On va montrer que $\mathrm{O}_n(\mathbb{R})$ et $\mathrm{U}_n(\mathbb{C})$ sont des fermés bornés.

Les applications $f : M \mapsto {}^tMM$ et $g : M \mapsto M^*M$ sont continues. Donc, les ensembles $\mathrm{O}_n(\mathbb{R}) = f^{-1}(\mathrm{I}_n)$ et $\mathrm{U}_n(\mathbb{C}) = g^{-1}(\mathrm{I}_n)$ sont fermés.

Pour montrer qu'ils sont bornés, il suffit de voir que $M \mapsto \mathrm{tr}({}^tMM)$ est une norme euclidienne, et que $M \mapsto \mathrm{tr}(M^*M)$ est une norme hermitienne [10].

On a $\mathrm{tr}({}^tMM) = \mathrm{tr}(\mathrm{I}_n) = n$, pour toute matrice M de $\mathrm{O}_n(\mathbb{R})$. Donc, $\mathrm{O}_n(\mathbb{R})$ est borné.

De même pour $\mathrm{U}_n(\mathbb{C})$ en utilisant la norme hermitienne.

Remarque. On vient de montrer que $\mathrm{O}_n(\mathbb{R})$ était sur la sphère de $\mathcal{M}_n(\mathbb{R})$ de rayon $\sqrt{n}$. La réciproque est clairement fausse et il est facile d'en trouver des contre-exemples : par exemple la matrice $D := \mathrm{diag}(\sqrt{n}, 0, \cdots, 0)$ vérifie $tr(\ {}^tDD) = n$, mais elle n'est pas dans $\mathrm{O}_n(\mathbb{R})$.

On montre, en revanche, que $\mathrm{O}_n(\mathbb{R})$ est l'ensemble des *points extrémaux* de la sphère de rayon 1 pour une autre norme : la norme subordonnée à la norme euclidienne de $\mathbb{R}^n$, voir [Szpirglas, Corollaire 7.110, p. 345].

10. En fait un calcul simple montre que $\mathrm{tr}({}^tMM) = \sum_{i,j} m_{ij}^2$, $tr(M^*M) = \sum_{ij} |m_{ij}|^2$. On reconnaît, respectivement, la norme euclidienne de $\mathbb{R}^{n^2}$, et la norme hermitienne de $\mathbb{C}^{n^2}$.

I.6 Dunford, exponentielle et topologie

Exercice I.6.1.
Calculer l'exponentielle de la matrice $\begin{pmatrix} 0 & x \\ x & 0 \end{pmatrix}$.

Soluce. On montre facilement par récurrence sur $n \geq 0$ que :

$$A^{2n} = \begin{pmatrix} x^{2n} & 0 \\ 0 & x^{2n} \end{pmatrix} = x^{2n}\mathrm{I}_2,\ A^{2n+1} = \begin{pmatrix} 0 & x^{2n+1} \\ x^{2n+1} & 0 \end{pmatrix} = x^{2n+1}J,$$

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

On a donc

$$\begin{aligned} \exp(A) &= \sum_{n\geqslant 0} \frac{1}{n!}A^n \\ &= \sum_{n\geqslant 0} \frac{1}{(2n)!}A^{2n} + \sum_{n\geqslant 0} \frac{1}{(2n+1)!}A^{2n+1} \\ &= \sum_{n\geqslant 0} \frac{x^{2n}}{(2n)!}\mathrm{I}_2 + \sum_{n\geqslant 0} \frac{x^{2n+1}}{(2n+1)!}J \end{aligned}$$

Et donc au final,

$$\exp(A) = \cosh(x)\mathrm{I}_2 + \sinh(x)J.$$

Remarque. On peut aussi diagonaliser A, histoire de rester dans le bain ! Car $A = PDP^{-1}$, avec

$$D = \begin{pmatrix} x & 0 \\ 0 & -x \end{pmatrix},\ P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},\ P^{-1} = \frac{1}{2}P.$$

Il vient :

$$\exp(A) = P\begin{pmatrix} x & 0 \\ 0 & -x \end{pmatrix}P^{-1} = \begin{pmatrix} \mathrm{ch}(x) & \mathrm{sh}(x) \\ \mathrm{sh}(x) & \mathrm{ch}(x) \end{pmatrix}.$$

Exercice I.6.2 (Dunford et exponentielle de matrices)**.**
Résoudre dans $\mathcal{M}_n(\mathbb{C})$ l'équation $\exp(A) = \mathrm{I}_n$.

Soluce. Donnons tout de suite une idée générale de la soluce. On conseille de lire [H2G2, Chap. V et VI] pour les informations utilisées concernant l'exponentielle et de la décomposition de Dunford.

Idée clef. Nous allons utiliser l'unicité de la décomposition de Dunford appliquée :

• d'une part, à la matrice I_n : I_n étant diagonalisable (car diagonale !), la décomposition de Dunford de I_n est : $\mathrm{I}_n = \mathrm{I}_n + 0_n$ où 0_n est la matrice nulle

• d'autre part, à la matrice $\exp(A)$: on va donc écrire $\exp(A)$ comme la somme d'une matrice diagonalisable et d'une matrice nilpotente

On cherche $A \in \mathcal{M}_n(\mathbb{C})$ telle que : $\exp(A) = \mathrm{I}_n$. Sur $\mathbb{C}$, le polynôme caractéristique de A est toujours scindé (d'où la résolution sur $\mathcal{M}_n(\mathbb{C})$ et non $\mathcal{M}_n(\mathbb{R})$). Ceci nous permet de considérer la décomposition de Dunford de la matrice A, qui nous donne l'existence d'un couple *unique* (D, N) tel que : $A = D + N$, D diagonalisable, N nilpotente, avec D et N qui commutent.

Puisque D et N commutent, on peut écrire : $\exp(D+N) = \exp(D)\exp(N)$, [11].

L'équation devient alors : $\exp(D)\exp(N) = \mathrm{I}_n$.

On écrit :

$$\exp(D)\exp(N) = \exp(D)+\exp(D)\exp(N)-\exp(D) = \exp(D)+\exp(D)(\exp(N)-\mathrm{I}_n)$$

avec :

• $\exp(D)$ diagonalisable, car D diagonalisable [12].

• $\exp(D)(\exp(N)-\mathrm{I}_n)$ nilpotente, puisque $(\exp(N)-\mathrm{I}_n)$ est nilpotente [13] et que $\exp(D)$ et $(\exp(N) - \mathrm{I}_n)$ commutent. En effet :

$$\begin{aligned}\exp(D)(\exp(N) - \mathrm{I}_n) &= \exp(D)\exp(N) - \exp(D) = \exp(D + N) - \exp(D)\\ &= \exp(N)\exp(D) - \exp(D) = (\exp(N) - \mathrm{I}_n)\exp(D).\end{aligned}$$

• $\exp(D)$ et $\exp(D)(\exp(N) - \mathrm{I}_n)$ commutent, d'après ce qui précède.

Ainsi, la décomposition de Dunford pour $\exp(A)$ est :

$$\exp(A) = \exp(D) + \exp(D)(\exp(N) - \mathrm{I}_n).$$

On est donc amené à résoudre :

$$\exp(D) + \exp(D)(\exp(N) - \mathrm{I}_n) = \mathrm{I}_n + 0_n.$$

Par *unicité* de la décomposition de Dunford, on a nécessairement :

$$\exp(D) = \mathrm{I}_n,\ \exp(D)(\exp(N) - \mathrm{I}_n) = 0_n.$$

Or $\exp(D)$ est inversible, d'inverse $\exp(-D)$, puisque $\exp(D)\exp(-D) = \exp(0_n) = \mathrm{I}_n$.

Donc, $\exp(D)(\exp(N) - \mathrm{I}_n) = 0_n$ implique $\exp(N) - \mathrm{I}_n = 0_n$.

Posons p l'indice de nilpotence de N et supposons, par l'absurde $p > 1$. Le développement de $\exp(N)$ donne

$$0_n = \exp(N) - \mathrm{I}_n = N + \frac{N^2}{2!} + \cdots + \frac{N^{p-1}}{(p-1)!}.$$

Multiplions l'égalité par N^{p-2} on obtient $N^{p-1} + 0_n + .. + 0_n = 0_n$, ce qui contredit la minimalité de p. Donc $N = 0_n$.

11. Voir par exemple [H2G2, VI.2]

12. Comme D est conjuguée à une matrice diagonale, et comme $\exp(PMP^{-1}) = P\exp(M)P^{-1}$, on obtient facilement que $\exp(D)$ est conjuguée à une matrice diagonale.

13. Pour le voir, on triangularise N et on voit que $\exp(N) - \mathrm{I}_n$ est semblable à une matrice triangulaire avec des zéros sur la diagonale.

Attention. Le fait que $\exp(N) = \mathrm{I}_n$ n'implique pas immédiatement que $N = 0_n$ car exp n'est pas injective.

Finalement : $A = D$, donc A est diagonalisable, c'est-à-dire semblable à une matrice diagonale, disons, $\mathrm{diag}(\lambda_1, .., \lambda_n)$. La matrice $\exp(A)$ est alors semblable à la matrice diagonale $\mathrm{diag}(e^{\lambda_1}, .., e^{\lambda_n})$. Comme $\exp(A) = \mathrm{I}_n$, on a nécessairement :

$$\lambda_j = 2ik_j\pi,\ k_j \in \mathbb{Z}.$$

Réciproquement, toute matrice de la forme $P\,\mathrm{diag}(\lambda_1, .., \lambda_n)P^{-1}$, avec λ_j comme ci-dessus, vérifie l'équation $\exp(A) = \mathrm{I}_n$.

Conclusion : l'ensemble des solutions est

$$\{P\Delta P^{-1},\ P \in GL_n(\mathbb{C}), \Delta \text{ diagonale à coefficients : } 2ik\pi, k \in \mathbb{Z}\}.$$

Remarques :

On a montré au passage que si $A = D + N$ est la décomposition de Dunford de A alors la décomposition de Dunford de l'exponentielle de A est : $\exp(A) = \exp(D) + \exp(D)(\exp(N) - \mathrm{I}_n)$

De plus, dans cette correction, on démontre que A est diagonalisable, il est donc possible de proposer en exercice la variante suivante :

Variante :

1- Montrer que A est diagonalisable si et seulement si $\exp(A)$ est diagonalisable.

2- Résoudre $\exp(A) = \mathrm{I}_n$.

Exercice I.6.3.
Soit A une matrice de $\mathcal{M}_n(\mathbb{C})$. Montrer qu'il existe un polynôme P_A de $\mathbb{C}[X]$ tel que $P_A(A) = \exp(A)$.

Soluce. Soit $\mathcal{P}_A$ l'ensemble des polynômes en A, et soit n la taille de A.

Il est clair que $\mathcal{P}_A$ est un sous-espace vectoriel de l'espace vectoriel $\mathcal{M}_n(\mathbb{C})$. Or, tout sous-espace vectoriel de dimension finie est fermé. Comme $\exp(A)$ est la limite d'une suite d'éléments de $\mathcal{P}_A$ (c'est la limite de la suite $\sum_{k=0}^{n} \frac{1}{k!}A^k$), cette suite converge donc dans $\mathcal{P}_A$. Cela signifie qu'il existe donc P_A dans $\mathcal{P}_A$ tel que $P_A(A) = \exp(A)$.

Chapitre II

Formes quadratiques

II.1 Fondamentaux

Exercice II.1.1.
Soit E un espace vectoriel (de dimension finie) et b une forme bilinéaire sur E. On considère l'application ϕ_b définie par

$$\phi_b : E \to E^*,\ x \mapsto \phi_b(x) = b(x, ?).$$

1. Montrer que ϕ_b est bien définie et qu'elle est linéaire.
2. On suppose dans cette question que b est une forme non dégénérée, c'est-à-dire que $\ker \phi_b$ est trivial.
   (a) Soit F un sous-espace de E et

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

   Montrer que $\phi_b(F^{\perp_b}) = F^{\perp}$ et en déduire

   $$\dim F^{\perp_b} = n - \dim F.$$

   (b) Soit u dans $\mathrm{End}(E)$ et u^* son adjoint pour la forme b. Soit ${}^t u$ la transposée de u dans $\mathrm{End}(E^*)$.
   Montrer que $u^* = \phi_b^{-1} \circ {}^t u \circ \phi_b$.
3. Montrer que si ψ est une application linéaire de E dans E' et si F' est un sous-espace vectoriel de E', alors $\dim \psi^{-1}(F') \geq \dim F'$. En déduire

   $$\dim F^{\perp_b} \geq n - \dim F.$$

Soluce. 1. Pour tout x de E, $\phi_b(x)$ est l'application de E dans $\mathbb{K}$ définie par $\phi_b(x)(y) = b(x, y)$. Ainsi, $\phi_b(x)$ est bien linéaire, puisque b est bilinéaire (et donc linéaire à droite). On a donc $\phi_b(x) \in E^*$, ce qui prouve que ϕ_b est bien définie.

On montre que $\phi_b(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 \phi_b(x_1) + \lambda_2 \phi(x_2)$ en appliquant cette égalité à y dans E et en utilisant le fait que b est bilinéaire (et donc linéaire à gauche).
(Attention, ce raisonnement ne marche plus pour les formes hermitiennes car elles sont antilinéaires à gauche).

2. (a) On commence par noter que, puisque $\ker \phi_b$ est trivial, et que $\dim E = \dim E^*$, la formule du rang implique que ϕ_b est un isomorphisme.
On va montrer directement l'égalité $F^{\perp_b} = \phi_b^{-1}(F^\perp)$.

$$x \in F^{\perp_b} \Leftrightarrow \forall y \in F,\ b(x,y) = 0 \Leftrightarrow \forall y \in F,\ \phi_b(x)(y) = 0 \quad \text{(II.1)}$$
$$\Leftrightarrow \phi_b(x) \in F^\perp \Leftrightarrow x \in \phi_b^{-1}(F^\perp). \quad \text{(II.2)}$$

On a donc finalement $\dim F^{\perp_b} = \dim F^\perp = n - \dim F$.
Attention ! En général, on n'a pas $F^{\perp_b}$ en somme directe avec F. Par exemple, si on prend la forme quadratique q sur $\mathbb{R}^2$ donnée par $q(x,y) = x^2 - y^2$, on vérifie qu'elle est non dégénérée car sa matrice dans la base canonique est $\mathrm{diag}(1,-1)$.
On a $b\big((1,1),(1,1)\big) = q(1,1) = 0$. Soit D la droite engendrée par $(1,1)$, alors $D^{\perp_b}$ est de dimension $2-1=1$ et contient D. Donc $D^{\perp_b} = D$.

(b) Par unicité de l'adjoint, il suffit de vérifier l'égalité :

$$b\big(\phi_b^{-1} \circ {}^t u \circ \phi_b(x), y\big) = b(x, u(y)).$$

On va utiliser pour cela les égalités de définition :

$$\phi_b(x)(y) = b(x,y),\ l\big(u(x)\big) = ({}^t u(l))(x),\ x,y \in E,\ l \in E^*.$$

$$b\big(\phi_b^{-1} \circ {}^t u \circ \phi_b(x), y\big) = \phi_b\big(\phi_b^{-1} \circ {}^t u \circ \phi_b(x)\big)(y) = {}^t u(\phi_b(x))(y) \quad \text{(II.3)}$$
$$= \phi_b(x)\big(u(y)\big) = b(x, u(y)). \quad \text{(II.4)}$$

Remarque. La notion d'adjoint pour une forme est essentiel dans la réduction. Tout vient du fait que si un endomorphisme u stabilise un sous-espace F alors u^* stabilise son orthogonal. Il est aussi important quand on veut montrer que les valeurs propres d'un endomorphisme hermitien, resp. antihermitien, resp. unitaire, sont réelles, resp. imaginaires pures, resp. de norme 1.

3. On considère l'application linéaire $\psi_0 : \psi^{-1}(F') \to F'$, $x \mapsto \psi(x)$.
Comme $\ker \psi \subset \psi^{-1}(F')$, on a $\ker \psi_0 = \ker \psi$. Et comme $\mathrm{Im}\, \psi_0 = F' \cap \mathrm{Im}(\psi)$, on a par la formule de Grassmann,

$$\begin{aligned}\dim \mathrm{Im}\, \psi_0 &= \dim F' + \dim \mathrm{Im}\, \psi - \dim(F' + \mathrm{Im}\, \psi) \\ &\geq \dim F' + \dim \mathrm{Im}\, \psi - \dim E.\end{aligned}$$

Si on applique la formule du rang, on obtient donc

$$\begin{aligned}\dim \psi^{-1}(F') &= \dim \ker \psi_0 + \dim \operatorname{Im} \psi_0 = \dim \ker \psi + \dim \operatorname{Im} \psi_0 \\ &= \dim E - \dim \operatorname{Im} \psi + \dim \operatorname{Im} \psi_0 \geq \dim F'\end{aligned}$$

. Du coup, $\dim F^{\perp_b} = \dim \phi_b^{-1}(F^\perp) \geq \dim F^\perp$.

II.2 Matrices symétriques réelles et réduction

Exercice II.2.1 (Version matricielle du théorème de réduction simultanée)**.** Soit $A \in \mathcal{S}_n^{++}(\mathbb{R})$, $B \in \mathcal{S}_n(\mathbb{R})$. Il existe $P \in \mathrm{GL}_n(\mathbb{R})$ et D diagonale telles que : ${}^tPAP = \mathrm{I}_n$ et ${}^tPBP = D$.

Soluce. La matrice A étant symétrique définie positive, elle est la matrice, dans la base canonique de $\mathbb{R}^n$, d'une forme symétrique définie positive q. Soit Q la matrice de passage de la base canonique à une base *orthonormée pour* q : on a ${}^tQAQ = \mathrm{I}_n$.

Remarque. Attention, Q est inversible mais n'est pas un élément de $\mathrm{O}_n(\mathbb{R})$. Effectivement, la nouvelle base est *orthonormée pour la forme* q, mais ne l'est pas pour la base canonique. Pour se fixer les idées, notons que l'on peut trouver cette base grâce au procédé d'orthonormalisation de Gram-Schmidt. Dans ce cas, Q est triangulaire supérieure avec des réels strictement positifs sur la diagonale, et il y a très peu de chance qu'une telle matrice (lorsque ce n'est pas $\pm\mathrm{I}_n$) soit dans $\mathrm{O}_n(\mathbb{R})$.

Comme B est symétrique, la matrice ${}^tQBQ$ l'est également (il suffit de la transposer pour s'en convaincre). Ainsi, en appliquant le théorème de réduction des matrices symétriques réelles, on obtient qu'il existe une matrice Ω dans $\mathrm{O}_n(\mathbb{R})$, et une matrice D diagonale, telles que

$${}^tQBQ = \Omega D \Omega^{-1} = \Omega D\, {}^t\Omega$$

Pour la trouver en pratique, on cherche une base orthonormale de vecteurs propres pour ${}^tQBQ$.

En multipliant à gauche par ${}^t\Omega$ et à droite par Ω les deux égalités ${}^tQAQ = \mathrm{I}_n$ et ${}^tQBQ = \Omega D\, {}^t\Omega$, on obtient ${}^t(Q\Omega)A(Q\Omega) = \mathrm{I}_n$ et ${}^t(Q\Omega)B(Q\Omega) = D$. On pose $P = Q\Omega$ pour obtenir ${}^tPAP = \mathrm{I}_n$ et ${}^tPBP = D$. Attention, encore une fois, P est inversible, mais non orthogonale !

Exercice II.2.2.
Soit A dans S_n^{++}, l'ensemble des matrices symétriques réelles définies positives, et soit B une matrice symétrique réelle. On munit $\mathbb{R}^n$ de sa forme euclidienne canonique $(\ ,\ )$, et d'une forme notée $\langle\ ,\ \rangle$, et donnée par

$$\langle X, Y\rangle = (A^{-1}X, Y),\ X, Y \in \mathbb{R}^n.$$

Montrer que AB est symétrique pour la forme $\langle\ ,\ \rangle$. En déduire que AB est diagonalisable sur $\mathbb{R}$.

Soluce. Avant de commencer, notons que A étant symétrique, A^{-1} l'est également. Effectivement, pour toute matrice carrée M, on a

$$^tM\,{}^t(M^{-1}) = {}^t(M^{-1}M) = {}^t\mathrm{I}_n = \mathrm{I}_n,$$

et il en résulte que ${}^t(M^{-1}) = ({}^tM)^{-1}$, puis que, ${}^t(A^{-1}) = ({}^tA)^{-1} = A^{-1}$.

La matrice A est symétrique définie positive. Cela implique que ses valeurs propres sont toutes strictement positives, et donc, les valeurs propres de A^{-1}, qui sont leurs inverses, sont également strictement positives. Donc, la matrice A^{-1} est symétrique définie positive.

Or, la forme $\langle\ ,\ \rangle$ a pour matrice A^{-1} dans la base canonique, puisque

$$\langle X, Y\rangle = (A^{-1}X, Y) = {}^t(A^{-1}X)Y = {}^tX\,{}^tA^{-1}Y = {}^tXA^{-1}Y.$$

On en déduit que la forme $\langle\ ,\ \rangle$ est une forme symétrique définie positive.

Remarque. On peut toutefois montrer ceci sans invoquer ces pauvres valeurs propres, si souvent accaparées ! Si X est non nul, alors $X' := A^{-1}X$ est également non nul, et

$$\langle X, X\rangle = (A^{-1}X, X) = (X', AX') = (AX', X') > 0,$$

puisque A est définie positive.

Montrons donc que AB vérifie $\langle ABX, Y\rangle = \langle X, ABY\rangle$. On a

$$\begin{aligned}\langle ABX, Y\rangle &= (A^{-1}ABX, Y) = (BX, Y) = (X, BY)\\ &= (AA^{-1}X, BY) = (A^{-1}X, ABY) = \langle X, ABY\rangle.\end{aligned}$$

On en déduit que l'endomorphisme de $\mathbb{R}^n$ ayant pour matrice AB dans la base canonique, est égal à son propre adjoint pour le produit scalaire $\langle\ ,\ \rangle$. La matrice AB se trouve donc être diagonalisable.

Exercice II.2.3 (Théorème d'Apollonius)**.**
On considère deux formes quadratiques q et q' sur $\mathbb{R}^n$, avec q' définie positive. On choisit une base (u_i) de $\mathbb{R}^n$ orthonormée pour la forme q' (Gram et Schmidt nous y autorisent !). Montrer que $\sum_{i=1}^n q(u_i)$ ne dépend pas du choix de la base orthonormée choisie (pour q'). Voyez-vous le lien avec la formule d'Apollonius ? On la rappelle : si OAB est un triangle du plan euclidien, avec I milieu de $[AB]$, alors $OA^2 + OB^2 = 2OI^2 + \frac{1}{2}AB^2$.

Soluce. Si (v_i) est une autre base orthonormée pour q' et P la matrice de passage de (u_i) à (v_i), on a

$$ {}^tP \operatorname{mat}_{(u_i)}(q) P = \operatorname{mat}_{(v_i)}(q),\ {}^tPP = {}^tP \operatorname{mat}_{(u_i)}(q') P = \operatorname{mat}_{(v_i)}(q') = \mathrm{I}_n $$

Donc, ${}^tP = P^{-1}$ et il en résulte que $\operatorname{mat}_{(u_i)}(q)$ et $\operatorname{mat}_{(v_i)}(q)$ sont semblables. Elles ont donc même trace : $\sum_{i=1}^n q(u_i) = \sum_{i=1}^n q(v_i)$, par définition de la matrice d'une forme quadratique dans une base.

Pour faire un lien avec la formule d'Appolonius, on prend pour q la norme euclidienne canonique de $\mathbb{R}^2$ et pour q' la forme dont la matrice est I_2 dans la base $\big(\overrightarrow{OA}, \overrightarrow{OB}\big)$. Soit φ' la forme bilinéaire symétrique associée à q'.

On a

$$ \varphi'(\sqrt{2}\overrightarrow{OI}, \sqrt{2}\overrightarrow{OI}) = 2\varphi'(\frac{\overrightarrow{OA}+\overrightarrow{OB}}{2}, \frac{\overrightarrow{OA}+\overrightarrow{OB}}{2}) = \frac{1}{2}\big(q(\overrightarrow{OA}) + q(\overrightarrow{OB})\big) = 1, $$

$$ \varphi'(\frac{1}{\sqrt{2}}\overrightarrow{AB}, \frac{1}{\sqrt{2}}\overrightarrow{AB}) = \frac{1}{2}\varphi'(\overrightarrow{OB}-\overrightarrow{OA}, \overrightarrow{OB}-\overrightarrow{OA}) = \frac{1}{2}\big(q(\overrightarrow{OB}) + q(\overrightarrow{OA})\big) = 1, $$

$$ \varphi'(\sqrt{2}\overrightarrow{OI}, \frac{1}{\sqrt{2}}\overrightarrow{AB}) = \varphi'(\frac{\overrightarrow{OA}+\overrightarrow{OB}}{2}, \overrightarrow{OB}-\overrightarrow{OA}) = \frac{1}{2}\big(q(\overrightarrow{OB}) - q(\overrightarrow{OA})\big) = 0. $$

Donc, la base $(\sqrt{2}\overrightarrow{OI}, \frac{1}{\sqrt{2}}\overrightarrow{AB})$ est également orthonormée pour q' et on a bien par ce qui précède

$$ q(\overrightarrow{OA}) + q(\overrightarrow{OB}) = 2q(\overrightarrow{OI}) + \frac{1}{2}q(\overrightarrow{AB}). $$

Remarque. On a utilisé l'invariance de la trace. Mais, à l'aide de l'invariance du déterminant, on obtient que le q-volume du parallélotope engendré par la base (u_i) est indépendant du choix effectué. Cela provient du fait que ce q-volume est égal à $\sqrt{|\det\big(\operatorname{mat}_{(u_i)}(q)\big)|}$.

II.3 Formes quadratiques et polynômes réels

Exercice II.3.1 (Formes de Hankel)**.**
Le but de l'exercice est de construire une forme quadratique réelle s associée à un polynôme P de $\mathbb{R}[X]$ et telle que la signature de s permette de calculer le nombre de racines distinctes ainsi que le nombre de racines réelles distinctes de P.
Soit P un polynôme réel de degré n, $n \geq 1$, et soit $x_1, \ldots, x_t$ ses racines (distinctes). Supposons que x_k est de multiplicité m_k. On note

$$s_k = m_1 x_1^k + \cdots + m_t x_t^k$$

les sommes de Newton (et l'on pose $s_0 = n$).

1. Montrer que $\sum_{0 \leq i,j \leq n-1} s_{i+j} X_i X_j$ définit une forme quadratique s, appelée forme de Hankel ou forme camarade [a], sur $\mathbb{C}^n$ et qu'elle définit une forme quadratique $s_{\mathbb{R}}$ sur $\mathbb{R}^n$.
2. Pour k dans $[1, t]$, soit φ_k la forme linéaire sur $\mathbb{C}^n$ donnée par

   $$\varphi_k(X_0, \ldots, X_{n-1}) = X_0 + x_k X_1 + x_k^2 X_2 + \cdots + x_k^{n-1} X_{n-1}$$

   et soit (p, q) la signature de $s_{\mathbb{R}}$.
   (a) Montrer par un déterminant de Vandermonde que les φ_k sont indépendants.
   (b) Montrer que $s = \sum_{k=1}^{t} m_k \varphi_k^2$.
   (c) En déduire que le nombre t de racines distinctes de P, est $p + q$.
3. Soit $n \geq 2$. Montrer que la signature de $\varphi_k^2 + \bar{\varphi}_k^2$ sur $\mathbb{R}^n$ vaut $(1, 0)$ si x_k est réel et $(1, 1)$ sinon.
4. En déduire que le nombre de racines réelles distinctes de P est égal à $p - q$.

a. C'est Denis Roussillon qui nous a proposé cette terminologie, en référence à la matrice compagnon qui permet de calculer les racines d'un polynôme.

Soluce. 1. On a clairement un polynôme homogène de degré 2 sur $\mathbb{C}$. Comme le polynôme P est réel et que les s_k sont des polynômes symétriques des racines, ceux-ci sont réels ; on a bien une forme quadratique réelle.

2. (a) Pour montrer qu'ils sont indépendants, on écrit leurs coordonnées en colonnes dans la base duale de la base canonique. On obtient

une matrice M de $\mathcal{M}_{n,t}(\mathbb{C})$ donnée par

$$M = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_t \\ x_1^2 & x_2^2 & \cdots & x_t^2 \\ \vdots & \vdots & \cdots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_t^{n-1} \end{pmatrix}$$

La matrice M est de rang maximal t : comme les x_k sont deux à deux distincts, on voit, par un déterminant de Vandermonde, que tous ses mineurs de taille t sont non nuls.

(b) Le coefficient de X_iX_j dans $\sum_{k=1}^t m_k\varphi_k^2$ vaut

$$\sum_{k=1}^t m_k x_k^i x_k^j = \sum_{k=1}^t m_k x_k^{i+j} = s_{i+j}.$$

Cela donne l'égalité.

(c) Le rang de s est t car les φ_k sont indépendants. C'est aussi $p+q$.

3. Si x_k est réel, $\varphi_k^2 + \bar{\varphi}_k^2 = 2\varphi_k^2$, de signature $(1,0)$ car φ_k est non nulle. Sinon, on vérifie que $\varphi_k^2 + \bar{\varphi}_k^2 = 2\,\mathrm{Re}(\varphi_k)^2 - 2\,\mathrm{Im}(\varphi_k)^2$. C'est donc bien une forme quadratique réelle. On pourra remarquer ensuite que, comme x_k n'est pas réel, $\varphi_k^2 + \bar{\varphi}_k^2$ est de rang 2 car φ_k et $\bar{\varphi}_k$ sont indépendants. Donc, la signature est bien $(1,1)$ dans ce cas.

4. La forme s est somme de $m_k\varphi_k^2$, avec les φ_k indépendants. On regroupe les formes φ_k conjuguées entre elles, lorsqu'elles ne sont pas réelles. Soit r le nombre de racines réelles distinctes, alors, par la question qui précède, la signature de s est égale à $(r,0) + (\frac{t-r}{2}, \frac{t-r}{2}) = (\frac{t+r}{2}, \frac{t-r}{2})$. D'où le résultat, puisque $\frac{t+r}{2} - \frac{t-r}{2} = r$.

Remarque. D'expérience, toute la difficulté dans la présentation de la forme de Hankel réside dans l'ambiguïté du corps sur lequel est définie la forme. Il faut bien avoir conscience qu'il y a une forme, vue sur $\mathbb{C}$ et une, vue sur $\mathbb{R}$.

Exercice II.3.2 (Déterminant et équation de cercle)**.**
On considère trois points non alignés, $A_i := (x_i, y_i)$, $1 \le i \le 3$, du plan affine euclidien $\mathbb{A}^2$. Montrer que l'équation du cercle passant par les points A_i (x_i, y_i), $1 \le i \le 3$ est donnée par

$$\begin{vmatrix} 1 & x & y & x^2+y^2 \\ 1 & x_1 & y_1 & x_1^2+y_1^2 \\ 1 & x_2 & y_2 & x_2^2+y_2^2 \\ 1 & x_3 & y_3 & x_3^2+y_3^2 \end{vmatrix} = 0$$

Soluce. Pour tout i, les coordonnées des points A_i vérifient l'équation proposée. En effet, si on remplace (x, y) par (x_i, y_i), le déterminant est celui d'une matrice dont deux lignes sont égales ; il est donc nul.

De plus, en développant le déterminant par rapport à la première ligne, on obtient bien une équation de cercle, puisque l'on reconnaît la forme $ax^2 + ay^2 + bx + cy + d = 0$. Attention au piège, tout de même : il faut prouver que $a \neq 0$.

Or, le développement par rapport à la première ligne donne

$$-a = \begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix}.$$

En substituant la ligne L_2, resp. L_3, par $L_2 - L_1$, resp. $L_3 - L_1$, et en développant par rapport à la première colonne, il vient

$$-a = \begin{vmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{vmatrix}.$$

Or, comme les A_i ne sont pas alignés, $\overrightarrow{A_1A_2}$ et $\overrightarrow{A_1A_3}$ ne sont pas proportionnels, et donc $a \neq 0$.

Conclusion : comme il n'existe qu'un unique cercle passant par trois points non alignés (« le » cercle circonscrit du triangle), on obtient bien l'équation voulue.

Remarque. Si les trois points avaient été alignés (mais deux à deux distincts), les termes en degré 2 en x et y auraient disparu et on aurait obtenu l'équation de la droite passant par ces trois points.

En manipulant légèrement le déterminant 3×3 ci-dessus, on obtient au passage que les trois points A_i sont alignés si et seulement si

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{vmatrix} = 0.$$

Ceci est facilement visible en plaçant les points A_i de $\mathbb{A}^2$ sur le plan affine d'équation $z = 1$, plongé dans l'espace vectoriel $\mathbb{R}^3$, en identifiant $A = (x, y)$ à $\overrightarrow{OA} = (x, y, 1)$. On voit alors que A_1, A_2, A_3 sont alignés si et seulement si $\overrightarrow{OA_1}$, $\overrightarrow{OA_2}$, $\overrightarrow{OA_3}$ sont coplanaires.

Chapitre III

Actions de groupes par isométries

III.1 Le tétraèdre

Exercice III.1.1. *[Le tétraèdre. [H2G2], Chap. XII proposition 3.12]
Le but de l'exercice est de décrire le groupe des isométries du tétraèdre régulier. On précise que le tétraèdre est un tétraèdre plein.

1. Montrer que si une isométrie de l'espace euclidien $\mathbb{R}^3$ stabilise un tétraèdre régulier, alors elle permute ses quatre sommets.

2. En déduire que le groupe des isométries du tétraèdre est isomorphe à $\mathfrak{S}_4$.

3. Déterminer le cardinal de $\mathrm{Is}^+(\mathcal{T})$ et dresser la liste de ses éléments.

4. Déterminer le cardinal de $\mathrm{Is}^-(\mathcal{T})$ et dresser la liste de ses éléments.

5. Exhiber un morphisme naturel de $\mathfrak{S}_4$ sur $\mathfrak{S}_3$.

Soluce. 1. On notera $\mathcal{T}$ le tétraèdre. On sait qu'il possède quatre sommets S_1, S_2, S_3 et S_4. De plus, il est convexe, c'est-à-dire que si M et N sont deux points de $\mathcal{T}$, alors, le segment $[MN]$ est entièrement dans $\mathcal{T}$.

Définition : $S \in \mathcal{T}$ est dit extrémal si la condition suivante est vérifiée

$$\forall M,\, N \in \mathcal{T},\, S \in [MN] \Longrightarrow S = M \text{ ou } S = N.$$

Nous allons tout d'abord montrer que l'ensemble des sommets du tétraèdre est exactement égal à l'ensemble de ses points extrémaux, puis

qu'une isométrie du tétraèdre stabilise l'ensemble de ses points extrémaux.

On sait qu'une application affine envoie un segment sur un segment. La proposition suivante s'en déduit :

Proposition 0 : Sur un segment $[AB]$, les points extrémaux sont A et B. Soit $g \in \mathrm{Is}(\mathcal{T})$, alors g envoie un point extrémal de $\mathcal{T}$ sur un point extrémal de $\mathcal{T}$.

Preuve :

La première assertion découle de la définition. Montrons la seconde.

Soient S un point extrémal de $\mathcal{T}$ et $g \in \mathrm{Is}(\mathcal{T})$. Notons $S' = g(S)$. Montrons que S' est extrémal.

Soient $M', N' \in \mathcal{T}$, $M' \neq N'$, tels que $S' \in [M'N']$. On pose

$$M := g^{-1}(M'),\ N := g^{-1}(N').$$

Comme g^{-1} est une application affine, elle envoie le segment $[M'N']$ vers le segment $[g^{-1}(M')g^{-1}(N')]$. De plus, g^{-1} stabilise $\mathcal{T}$. Donc $S = g^{-1}(S')$ est dans $[g^{-1}(M')g^{-1}(N')] = [MN] \subset \mathcal{T}$. Vu que S est extrémal, on a $S = M$ ou $S = N$, et donc $S' = M'$ ou $S' = N'$. Ce qui prouve que S' est extrémal.

Remarque : On s'est servi uniquement du fait que g est affine, et non du fait qu'elle conserve les longueurs.

Proposition 1 : S est un sommet du tétraèdre si et seulement si S est extrémal.

Preuve :

Soit S un point extrémal de $\mathcal{T}$. Montrons que S est un sommet.

Si S est sur une arête, il est en particulier extrémal sur l'arête. D'après la propriété 0, c'est donc un sommet.

Supposons que S est sur une face, disons F. Projetons S, à partir d'un sommet S' de F, sur l'arête opposée à S'. Soit P' le point ainsi obtenu. Le tétraèdre étant convexe, le segment $[S'P']$ ainsi obtenu reste dans le tétraèdre. Comme S est extrémal sur ce segment, puisqu'il l'est dans le tétraèdre, d'après la proposition 0, $S = S'$ ou P'. D'après ce qui précède, dans les deux cas, S est un sommet.

Si S est un point intérieur au tétraèdre, on projette S, à partir d'un sommet, disons S'', sur la face opposée, disons F'', à S''. On obtient un point P'' de F''. Par convexité, le segment $[S''P'']$ reste dans le tétraèdre. Puis on applique le raisonnement précédent pour affirmer que S est un sommet.

Montrons que tout sommet du tétraèdre est extrémal. On rappelle que S_1, S_2, S_3, S_4 sont les sommets de $\mathcal{T}$. Comme les 4 sommets de $\mathcal{T}$ ne sont pas coplanaires, (S_1, S_2, S_3, S_4) constitue un repère de

l'espace affine $\mathbb{R}^3$. Dans ce repère, le tétraèdre (plein) $\mathcal{T}$ est l'ensemble des points M de coordonnées (x_1, x_2, x_3), avec $0 \leq x_i$ et $\sum_i x_i \leq 1$. Supposons par l'absurde que l'origine S_1 soit dans $[MM']$, avec $M = (x_i)$, $M' = (x'_i)$ sur le tétraèdre. Alors S_1 est barycentre, à coefficients positifs, de (M, α), (M', α'). On peut de plus supposer que les coefficients sont strictement positifs (sinon, on aurait tout de suite $S_1 = M$ ou M'). Or, $\alpha(x_1, x_2, x_3) + \alpha'(x'_1, x'_2, x'_3) = (0, 0, 0)$, avec $0 < \alpha, \alpha'$ et $\alpha + \alpha' = 1$, avec les x_i, x'_i positifs, implique $(x_1, x_2, x_3) = (x'_1, x'_2, x'_3) = (0, 0, 0)$. Ceci implique $M = M' = S_1$. Ce qui prouve que S_1 est extrémal.

Notons que l'on peut envoyer un sommet sur un autre sommet par une application affine (par exemple par une rotation d'ordre 3). On en déduit par la proposition 0 que tous les S_i sont extrémaux.

Par la proposition 1, on a alors : si $g \in \mathrm{Is}(\mathcal{T})$, alors g permute les 4 sommets de $\mathcal{T}$.

2. Notons $\mathcal{E}$ l'ensemble des sommets (et donc des points extrémaux). Montrons que $\mathrm{Is}(\mathcal{T})$ est isomorphe au groupe des permutations $\mathfrak{S}(\mathcal{E})$:

   Considérons

$$\begin{aligned} \phi\colon \mathrm{Is}(\mathcal{T}) &\to \mathfrak{S}(\mathcal{E}) \simeq \mathfrak{S}_4 \\ g &\mapsto \phi(g)\colon \mathcal{E} \to \mathcal{E} \\ S &\mapsto g(S) \end{aligned}$$

   L'application ϕ est juste *l'application restriction*, qui envoie g, qui est une application de l'espace affine dans lui-même, sur sa restriction à l'ensemble des sommets de $\mathcal{T}$. A ce titre, il s'agit d'un morphisme de groupes (il conserve la loi $\circ$).

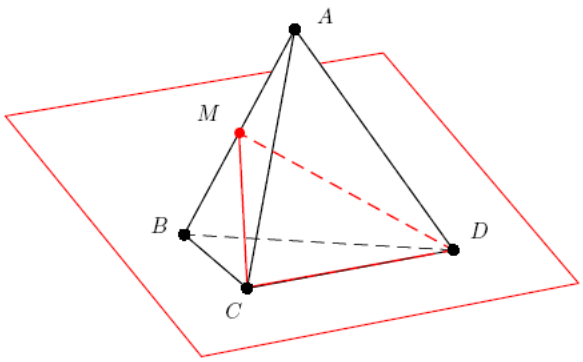
   Montrons que ϕ est injectif.

   Soit $g \in \mathrm{Is}(\mathcal{T}) \mid \phi(g) = \mathrm{Id}_{\mathcal{E}}$. Autrement dit, g laisse fixes les 4 sommets. Or, on a vu que (S_1, S_2, S_3, S_4) constitue un repère affine de $\mathbb{R}^3$. Donc g stabilise un repère de $\mathbb{R}^3$, c'est donc l'identité.
   Par suite, on a $\ker \phi = \{\mathrm{Id}_{\mathcal{T}}\}$ et donc ϕ est injectif.

   Montrons que ϕ est surjectif.

   On montre que $\mathrm{Im}\,\phi$ contient un système de générateurs de $\mathfrak{S}_4$: les transpositions.

   Pour cela, il suffit de montrer que $\mathrm{Im}\,\phi$ contient (12). Les autres transpositions s'y trouveront alors aussi par symétrie. On cherche donc une isométrie (forcément unique par ce qui précède) qui laisse fixes S_3 etS_4 et qui échange S_1 et S_2. Il s'agit de la symétrie orthogonale par rapport au plan médiateur de $[S_1S_2]$, c'est-à-dire le plan passant par S_3 et S_4 et le milieu de $[S_1S_2]$, voir figure ci-dessous



D'où l'isomorphisme annoncé.

3. Maintenant, on remarque que tout sous-groupe du groupe affine qui fixe un point peut être assimilé à un sous-groupe du groupe linéaire ; on peut par exemple lui appliquer le déterminant. Le sous-groupe $\mathrm{Is}^+(\mathcal{T})$ est le noyau du déterminant restreint à $\mathrm{Is}(\mathcal{T})$. Comme l'image du déterminant restreint à $\mathrm{Is}(\mathcal{T})$ prend exactement deux valeurs, 1 et -1, il vient que le sous-groupe $\mathrm{Is}^+(\mathcal{T})$ est d'indice 2 dans $\mathrm{Is}(\mathcal{T})$. Via l'isomorphisme entre $\mathrm{Is}(\mathcal{T})$ et $\mathfrak{S}_4$, le sous-groupe $\mathrm{Is}^+(\mathcal{T})$ est donc isomorphe à un sous-groupe d'indice 2 de $\mathfrak{S}_4$. Or, on sait qu'il n'y a qu'un seul tel sous-groupe : le groupe alterné $\mathfrak{A}_4$. On a donc, $\mathrm{Is}^+(\mathcal{T}) \simeq \mathfrak{A}_4$.

4. $|\mathrm{Is}^-(\mathcal{T})| = 24 - 12 = 12$.

   Remarque. Ce n'est pas un sous-groupe (le vilain !). En revanche, c'est une belle classe à gauche (ou à droite)...

5. Dans la suite, on appelle *bimédiane* du tétraèdre, une droite qui joint les milieux de deux arêtes opposées. On remarque que le tétraèdre possède 3 bimédianes, qu'il est plus facile de visualiser lorsque le tétraèdre est inscrit dans le cube, comme dans la figure III.3 : ces trois bimédianes sont alors les axes de symétrie du cube qui traversent par les milieux ses faces opposées. Par exemple, on voit parfaitement qu'elles sont deux à deux orthogonales.

   Le groupe $\mathrm{Is}(\mathcal{T}) \simeq \mathfrak{S}_4$ envoie une bimédiane sur une autre bimédiane. Il agit donc naturellement sur l'ensemble des bimédianes qui est de cardinal 3. On a donc construit, par définition d'une action, un morphisme de $\mathfrak{S}_4$ sur $\mathfrak{S}_3$.

   Ce morphisme est surjectif : un 4-cycle a pour image une transposition. On peut le voir encore une fois en inscrivant le tétraèdre dans le cube et en utilisant la réalisation du 4-cycle faite ci-dessus.

   Le noyau est le groupe de Klein : {e, composées de 3 transpositions à supports disjoints}. Ces transpositions correspondent aux 3 retournements (rotation d'angle π) autour des trois bimédianes.

III.2 Le cube

Exercice III.2.1 (Le groupe du cube [H2G2). , Chap. XII]

1. Montrer, en remarquant que l'on peut inscrire deux tétraèdres dans le cube (voir figure III.3), que le groupe G des isométries positives du cube contient un sous-groupe d'indice 2 isomorphe à $\mathfrak{A}_4$. On obtient, en particulier, que G est d'ordre 24.
2. En déduire, en faisant agir G sur les grandes diagonales du cube, que G est isomorphe à $\mathfrak{S}_4$.

Soluce. 1. Notons G le groupe des isométries positives du cube et $\mathcal{T} := \{T_1, T_2\}$ l'ensemble des deux tétraèdres inscrits dans le cube. Comme vu précédemment avec le tétraèdre, le groupe G respecte l'ensemble des sommets du cube. Il en résulte que G agit sur l'ensemble des tétraèdres inscrits dans le cube. Donc G s'envoie dans $\mathfrak{S}(\mathcal{T}) \simeq \mathfrak{S}_2$ par un morphisme d'action ϕ.

Par une rotation d'ordre 4 autour d'un axe passant par le milieu de faces opposées, on voit que l'on échange T_1 et T_2. Le morphisme ϕ est donc surjectif.

Montrons que $\ker\phi \simeq \mathfrak{A}_4$. Un élément de $\ker\phi$ est un élément qui laisse fixe le tétraèdre T_1 (du coup, il laissera automatiquement fixe l'autre tétraèdre). Il en résulte que $\ker\phi$ s'injecte dans le groupe des isométries positives du tétraèdre, c'est-à-dire $\mathfrak{A}_4$.

Montrons la surjectivité. Il suffit pour cela de montrer qu'un système de générateurs de $\mathfrak{A}_4$ est dans l'image, c'est-à-dire qu'un système de générateurs du groupe des isométries du tétraèdre se réalise comme système d'isométries positives du cube. Un bon système de générateurs [1] de $\mathfrak{A}_4$ est l'ensemble des 3-cycles. On voit que les 3-cycles se réalisent comme rotations autour des grandes diagonales du cube.

Conclusion, $H := \ker\phi \simeq \mathfrak{A}_4$ vérifie $G/H \simeq \operatorname{Im}\phi \simeq \mathfrak{S}_2$; il est bien d'indice 2.

2. On fait agir G sur les 4 grandes diagonales du cube. On vérifie que l'action est bien définie (une rotation respecte les distances, elle envoie bien une grande diagonale sur une grande diagonale). On obtient un morphisme d'action de G dans $\mathfrak{S}_4$. Il est surjectif, car les transpositions de $\mathfrak{S}_4$ sont réalisées par les rotations d'ordre 2 passant par les milieux d'arêtes opposées. Ce morphisme est injectif, puisqu'il est surjectif et que l'ordre des groupes est le même, 24.

1. Comment différencier un bon système d'un mauvais système de générateurs ? Alors, ben, le mauvais système de générateurs, il est là, il voit $\mathfrak{A}_4$, et il le génère quoi, pfouah ! Mais le bon système de générateurs, il voit $\mathfrak{A}_4$, il le génère... Mais c'est un bon système de générateurs.

III.3 Dictionnaires

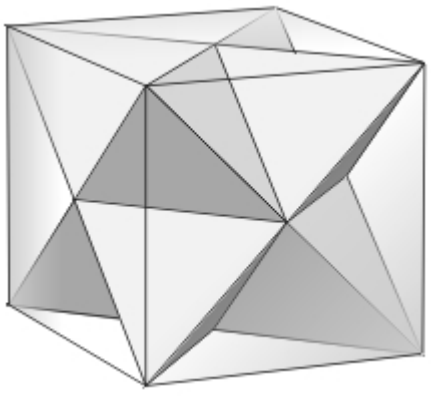
Nombre	Ordre	Isométries de $\mathcal{T}$	Permutations de $\mathfrak{S}_4$
1	1	Id	Id
8	3	rotation d'axe sommet-centre de face opposée	3-cycle
3	2	rotation d'angle π passant par le milieu de 2 arêtes opposées	produit de 2 transpositions à supports disjoints

Nombre	Ordre	Isométries de $\mathcal{T}$	Permutations de $\mathfrak{S}_4$
6	2	symétrie par rapport à un plan médiateur	transposition
6	4	composée d'une symétrie et d'une rotation (*)	4-cycle

(*) La composée d'une rotation et d'une symétrie mérite que l'on s'y attarde un peu.

Une première approche consiste à remarquer l'égalité $(123)(34) = (1234)$. Lorsque l'on bascule cette égalité dans $\mathrm{Is}(\mathcal{T})$ via l'isomorphisme obtenu, on voit que pour obtenir la permutation circulaire $(S_1S_2S_3S_4)$, il faut composer une rotation d'ordre 3 autour de la droite médiane passant par S_4 avec la symétrie par rapport au plan médiateur de S_3 et S_4. On pourrait reprocher à cette décomposition de ne pas être unique. Effectivement, un théorème donne une condition d'unicité pour la décomposition de l'antidéplacement $(S_1S_2S_3S_4)$ si la droite médiane et le plan médiateur sont orthogonaux. Ce qui n'est pas le cas.

Pour y remédier, on considère dans la figure III.3 qui suit les deux tétraèdres inscrits dans le cube :



Pour obtenir un 4-cycle, il suffit de composer une rotation d'ordre 4 autour de l'axe vertical du cube avec la symétrie par rapport au plan orthogonal

à cet axe. Le problème est que l'on n'a pas décomposé en deux isométries du même tétraèdre, mais en deux isométries qui échangent les deux tétraèdres inscrits dans le cube. Mais faut savoir ce qu'on veut !

Nombre	Ordre	$\mathrm{Is}^+(C_6)$	Permutations de $\mathfrak{S}_4$
1	1	Id	Id
8	3	rotation d'axe sommet-sommet opposé	3-cycle
3	2	rotation d'angle π passant par le centre de 2 faces opposées	produit de 2 transpositions à supports disjoints
6	2	rotation d'angle π passant par le centre de 2 arêtes opposées	transposition
6	4	rotation d'angle $\pm\pi/2$ passant par le centre de 2 faces opposées	4-cycle

FIGURE III.1 – Isomorphisme entre $\mathrm{Is}^+(C_6)$ et $\mathfrak{S}_4$

Chapitre IV

Combinatoire et arithmétique

IV.1 Combinatoire et fractions rationnelles

Exercice IV.1.1. *
Soit $n \geq m$ deux entiers positifs. On lance k dés à m faces numérotées de 1 à m. On veut connaître le nombre N_n de façons d'obtenir un total égal à n.

1. Montrer que N_n est égal au coefficient en z^n dans $P = (z + z^2 + \ldots + z^m)^k$.
2. Montrer que $P = z^k \frac{(1-z^m)^k}{(1-z)^k}$.
3. Montrer l'égalité $(1-z)^{-k} = \sum_{v \geq 0} \binom{v+k-1}{k-1} z^v$, avec $|z| < 1$.
4. En déduire
$$N_n = \sum_{0 \leq u \leq s} (-1)^u \binom{k}{u} \binom{n - um - 1}{k-1},$$
où $s = \text{Min}\{k, \frac{n-k}{m}\}$.

Soluce. 1. On voit en développant P que l'on a

$$P = (z+z^2+\ldots+z^m)^k = (z^1+z^2+\ldots+z^m)\cdots(z^1+z^2+\ldots+z^m) = \sum_{1 \leq i_1, \cdots, i_k \leq m} z^{i_1+\cdots+i_k}$$

Et donc,

$$P = \sum_{0 \leq n} N_n z^n$$

On pourra noter que cette somme est bien sûr finie puisque les N_n sont presque tous nuls.

2. On a :

$$\begin{aligned} P &= [z(1+\ldots+z^{m-1}]^k \\ P &= z^k(1+\ldots+z^{m-1})^k \\ P &= z^k\left(\frac{1-z^m}{1-z}\right)^k \end{aligned}$$

Ce qui donne le résultat voulu.

3. Montrons, par récurrence sur k que $(1-z)^{-k} = \sum_{v\geq 0}\binom{v+k-1}{k-1}z^v$.
Pour $k=1$, c'est le développement en série entière de $\frac{1}{1-z}$ sur le disque ouvert de rayon 1 :

$$(1-z)^{-1} = \sum_{v\geq 0} z^v$$

Comme $\binom{v+k-1}{k-1} = \binom{v}{0} = 1$, l'égalité en découle.

Supposons la propriété vraie au rang $k \geq 1$. Montrons qu'elle l'est au rang $k+1$. Par dérivation à l'intérieur du disque de convergence, on obtient :

$$\begin{aligned} k(1-z)^{-(k+1)} &= \sum_{v\geq 1} v\binom{v+k-1}{k-1}z^{v-1} \\ &= \sum_{u\geq 0}(u+1)\binom{u+k}{k-1}z^u \\ (1-z)^{-(k+1)} &= \sum_{u\geq 0}\frac{(u+1)}{k}\frac{(u+k)!}{(k-1)!(u+k-k+1)!}z^u \\ &= \sum_{u\geq 0}\frac{(u+1)}{k}\frac{(u+k)!}{(k-1)!(u+1)!}z^u \\ &= \sum_{u\geq 0}\frac{(u+k)!}{k!u!}z^u \end{aligned}$$

d'où : $(1-z)^{-k} = \sum_{u\geq 0}\binom{u+k-1}{k-1}z^u$

4. On sait que N_n est le coefficient de z^n dans $(z+z^2+\ldots+z^m)^k$

Or :

$$\begin{aligned}(z+z^2+\ldots+z^m)^k &= z^k(1+z+\ldots+z^{m-1})^k \\ &= z^k\left(\frac{1-z^m}{1-z}\right)^k \\ &= z^k(1-z^m)^k\sum_{v\geq 0}\binom{v+k-1}{k-1}z^v \\ &= z^k\left(\sum_{w=0}^{k}(-1)^w\binom{k}{w}z^{wm}\right)\left(\sum_{v\geq 0}\binom{v+k-1}{k-1}z^v\right)\end{aligned}$$

donc les termes en z^n sont ceux tels que $k+wm+v=n$ d'où :

$$N=\sum_{k+wm+v=n}(-1)^w\binom{k}{w}\binom{v+k-1}{k-1}$$

puis, en posant $u=w$, on a $v=n-k-um$ et enfin

$$N=\sum_{u\geq 0}(-1)^u\binom{k}{u}\binom{n-um-1}{k-1}.$$

Exercice IV.1.2 (Formule du crible et applications)**.**

1. Soit une famille finie A_i, $1 \leq i \leq m$, de sous-ensembles d'un ensemble fini E. Prouver la formule de cardinaux

$$\#\bigcup_{i\in I} A_i = \sum_{k=1}^{m}(-1)^{k+1} \sum_{1\leq i_1<i_2<\cdots<i_k\leq m} \#\bigcap_{j=1}^{k} A_{i_j}.$$

2. Première application : montrer que le nombre de surjections d'un ensemble de cardinal n vers un ensemble de cardinal m est égal à

$$S_{n,m} = \sum_{k=0}^{m}(-1)^k \binom{m}{k}(m-k)^n.$$

3. Deuxième application : montrer que le nombre de dérangements (c'est-à-dire des permutations qui ne fixent aucun élément) du groupe $\mathfrak{S}_n$ est égal à

$$D_n = n!\sum_{k=0}^{n}\frac{(-1)^k}{k!}.$$

4. Troisième application : montrer que l'indicatrice d'Euler $\varphi(n)$ d'un entier naturel n (nombre d'entiers de $[1, n]$ premiers avec n) vaut

$$\varphi(n) = n\prod_{i=1}^{l}(1-\frac{1}{p_i}),$$

où $n = \prod_{i=1}^{l} p_i^{\alpha_i}$ est la décomposition de n en facteurs premiers.

Soluce. 1. Nous allons utiliser pour cela les fonctions indicatrices. Dans la $\mathbb{Q}$-algèbre $\mathcal{F}$ des applications de E dans $\mathbb{Q}$ [1], on considère, pour tout sous-ensemble A de E, la fonction indicatrice $\mathbf{1}_A$ telle que $\mathbf{1}_A(x)$ vaut 1 ou 0 selon si x appartient à A ou non. On a bien évidemment

$$\#A = \sum_{x\in E}\mathbf{1}_A(x).$$

Comme l'application γ qui, à f dans $\mathcal{F}$, associe $\sum_{x\in E} f(x)$, est une *forme linéaire* sur $\mathcal{F}$, il suffit pour prouver la formule, de montrer

$$\mathbf{1}_{\bigcup_{i\in I} A_i} = \sum_{k=1}^{m}(-1)^{k+1} \sum_{1\leq i_1<i_2<\cdots<i_k\leq m} \mathbf{1}_{\bigcap_{j=1}^{k} A_{i_j}},$$

et d'appliquer γ.

1. Pourquoi est-elle de dimension $\#E$?

Or, il est clair, en faisant un cas par cas, que $\mathbf{1}_{A\cap B} = \mathbf{1}_A\mathbf{1}_B$. Du coup, par récurrence,

$$\mathbf{1}_{\bigcap_k A_k} = \prod_k \mathbf{1}_{A_k}.$$

Il est encore plus clair que si $\overline{A}$ désigne le complémentaire de A dans E, on a $\mathbf{1}_{\overline{A}} = 1 - \mathbf{1}_A$ [2].

$$\begin{aligned}
\mathbf{1}_{\bigcup_{i\in I} A_i} &= 1 - \mathbf{1}_{\overline{\bigcup_{i\in I} A_i}} = 1 - \mathbf{1}_{\bigcap_{i\in I} \overline{A_i}} \\
&= 1 - \prod_{i\in I} \mathbf{1}_{\overline{A_i}} = 1 - \prod_{i\in I}(1 - \mathbf{1}_{A_i}) \\
&= 1 - \Big(\sum_{k=0}^{m}(-1)^k \sum_{1\le i_1<i_2<\cdots<i_k\le m} \prod_{j=1}^{k} \mathbf{1}_{A_{i_j}}\Big) \\
&= \sum_{k=1}^{m}(-1)^{k+1} \sum_{1\le i_1<i_2<\cdots<i_k\le m} \prod_{j=1}^{k} \mathbf{1}_{A_{i_j}} \\
&= \sum_{k=1}^{m}(-1)^{k+1} \sum_{1\le i_1<i_2<\cdots<i_k\le m} \mathbf{1}_{\bigcap_{j=1}^{k} A_{i_j}}.
\end{aligned}$$

2. On rappelle que le nombre d'applications d'un ensemble à n éléments vers un ensemble à m éléments est égal à m^n. Soit A_i, i de 1 à m, l'ensemble des applications f telles que i soit dans l'image directe de f. Alors, l'ensemble des surjections est égal à $\cap_{i=1}^m A_i$. Il vient donc, avec la formule du crible,

$$\begin{aligned}
S_{n,m} &= \#\cap_{i=1}^m A_i = m^n - \#\overline{\cap_{i=1}^m A_i} = m^n - \#\cup_{i=1}^m \overline{A_i} \\
&= m^n - \sum_{k=1}^{m}(-1)^{k+1} \sum_{1\le i_1<i_2<\cdots<i_k\le m} \#\bigcap_{j=1}^{k} \overline{A}_{i_j}.
\end{aligned}$$

Or, $\bigcap_{j=1}^k \overline{A}_{i_j}$ est l'ensemble des applications qui ne contiennent aucun des i_j, $1 \le j \le k$, dans leur image directe. Il y en a donc autant que d'applications d'un ensemble à n éléments dans un ensemble à $m-k$ éléments, c'est-à-dire $(m-k)^n$. Ce qui donne facilement la formule désirée.

3. Soit A_i, i de 1 à n, l'ensemble des permutations de $\mathfrak{S}_n$ qui fixent i. Alors, le cardinal de l'ensemble des dérangements est égal à

$$\begin{aligned}
D_n &= \#\cap_{i=1}^n \overline{A_i} = \#\overline{\cup_{i=1}^n A_i} = n! - \#\cup_{i=1}^n A_i \\
&= n! - \sum_{k=1}^{n}(-1)^{k+1} \sum_{1\le i_1<i_2<\cdots<i_k\le m} \#\bigcap_{j=1}^{k} A_{i_j}.
\end{aligned}$$

2. Ici, 1 désigne la fonction qui associe le nombre 1 à tout x de E.

Or, $\bigcap_{j=1}^{k} A_{i_j}$ est l'ensemble des permutations de $[1,n]$ qui fixent les éléments $i_1, \cdots, i_k$. Il y en a donc autant que de permutations d'un ensemble à $n-k$ éléments, c'est-à-dire $(n-k)!$. Ce qui donne facilement la formule désirée.

4. Soit A_i, i de 1 à l, l'ensemble des éléments m de $E := [1,n]$ tels que p_i divise m. Un nombre m est premier avec n si et seulement aucun p_i ne divise m. L'indicatrice d'Euler est donc égale à

$$\begin{aligned}\varphi(n) &= \#\cap_{i=1}^{l}\overline{A_i} = \#\overline{\cup_{i=1}^{n} A_i} = n - \#\cup_{i=1}^{n} A_i \\ &= n - \sum_{k=1}^{n}(-1)^{k+1} \sum_{1\le i_1<i_2<\cdots<i_k\le m} \#\bigcap_{j=1}^{k} A_{i_j}.\end{aligned}$$

Or, comme les p_{i_j} sont des premiers distincts, être divisible par tous les p_{i_j} revient à être divisible par leur produit. Envoyer x sur $\frac{x}{\prod_{j=1}^{k} p_{i_j}}$ fournit alors une bijection (à vérifier soigneusement) de $\bigcap_{j=1}^{k} A_{i_j}$ sur $[1,m]$, avec $m := \frac{n}{\prod_{j=1}^{k} p_{i_j}}$. Ainsi,

$$\#\bigcap_{j=1}^{k} A_{i_j} = \frac{n}{\prod_{j=1}^{k} p_{i_j}}$$

Il vient donc

$$\begin{aligned}\varphi(n) &= n - \sum_{k=1}^{n}(-1)^{k+1} \sum_{1\le i_1<i_2<\cdots<i_k\le m} \frac{n}{\prod_{j=1}^{k} p_{i_j}} \\ &= \sum_{k=0}^{n}(-1)^{k} \sum_{1\le i_1<i_2<\cdots<i_k\le m} \frac{n}{\prod_{j=1}^{k} p_{i_j}} \\ &= n\sum_{k=0}^{n}(-1)^{k} \sum_{1\le i_1<i_2<\cdots<i_k\le m} \prod_{j=1}^{k} \frac{1}{p_{i_j}} \\ &= n\prod_{i=1}^{l}(1-\frac{1}{p_i}).\end{aligned}$$

Exercice IV.1.3 (L'énigme des prisonniers)**.**
Afin de régler le problème de surpopulation des prisons, on décide de jouer avec le sort de 100 prisonniers [a]. On fournit aux prisonniers un matricule allant de 1 à 100. On leur indique une pièce, où se trouvent 100 boîtes numérotées de 1 à 100, contenant chacune le matricule d'un prisonnier, de sorte que chaque prisonnier soit représenté.
Lorsqu'un prisonnier rentre dans cette pièce, il va choisir 50 boîtes. Tous les prisonniers vont passer chacun à leur tour, et si un seul des prisonniers ne trouve pas son matricule dans une des boîtes, alors, tous les prisonniers sont fusillés [b].
Heureusement, un des prisonniers est mathématicien ; il va proposer à ses codétenus une stratégie pour améliorer leurs chances de survie [c]. Voici sa stratégie :
Chacun commence par ouvrir la boîte portant son numéro, tire le matricule qui est à l'intérieur, puis ouvre la boîte ayant ce matricule pour numéro, tire le matricule qui est à l'intérieur, et ainsi de suite jusqu'à ce qu'il trouve son matricule ou qu'il ait ouvert 50 boîtes.
Voici deux questions bien naturelles. Quelle est la probabilité pour que le groupe survive si on laisse chaque prisonnier agir au hasard ? Quelle est la probabilité pour que le groupe survive si l'on suit le conseil du mathématicien ?

a. Inutile de rappeler que l'histoire se déroule sous un régime totalitaire, et dans un pays totalement imaginaire.

b. Ce détail n'est pas vraiment utililisé dans la preuve. Il permet juste de créer de l'adrénaline et une motivation supplémentaire pour le candidat au concours.

c. Cette partie de l'énoncé a pour but non dissimulé de réintégrer une image positive du mathématicien dans la population carcérale. Ceci dit, on ne saura jamais pourquoi le mathématicien est sous les verrous. A cette question du jury, suggérez une erreur judiciaire.

Soluce. On est maintenant entre gens sérieux et nous allons remplacer le nombre 100 par un nombre n quelconque, que l'on peut supposer pair, le nombre 50 sera donc remplacé par $\frac{n}{2}$.

• Premier cas. Les risques du hasard.

Pour un prisonnier donné, la probabilité que son matricule apparaisse dans les tirages successifs est de $\frac{n/2}{n} = \frac{1}{2}$. L'indépendance de chaque prisonnier (l'anarchie, quoi !) fait que la chance de survie du groupe est de $(1/2)^n$.

• Second cas. Le risque sous contrôle.

Soit E_n l'ensemble des nombres entiers de 1 à n, P l'ensemble des prisonniers (ou si l'on préfère, l'ensemble de leurs matricules), et B l'ensemble des boîtes.

L'ensemble E_n est alors identifié à la fois à P, via les matricules, et à B, via la numérotation des boîtes. On pourra dire, sans ambiguïté, le prisonnier p, ou la boîte k, .

Le fait d'avoir placé un matricule dans chaque boîte fournit une bijec-

tion σ de E_n dans lui-même, telle que la boîte p contient le matricule $\sigma(p)$. L'identification fait donc de σ une permutation.

Tout le suspense réside maintenant dans cet élément σ de $\mathfrak{S}_n$.

La procédure du choix des boîtes proposée par le mathématicien se traduit de la façon suivante : soit p un prisonnier, il va tout d'abord se diriger vers la boîte p. Il va donc tirer le matricule $\sigma(p)$. La procédure consiste maintenant à ouvrir la boîte $\sigma(p)$, dont il tirera le matricule $\sigma\big(\sigma(p)\big) = \sigma^2(p)$. S'il suit convenablement la procédure, la k-ième boîte qu'il ouvrira contiendra le matricule $\sigma^k(p)$.

Il est primordial que le prisonnier p retrouve son propre matricule au bout de $\frac{n}{2}$ essais. Cela signifie que l'on veut $\sigma^k(p) = p$ pour $1 \leq k \leq \frac{n}{2}$. Et ceci doit être valable pour tout p.

On a donc résumé le problème en une propriété de la permutation σ. Etudions de plus près cette propriété.

Décomposons σ en cycles disjoints : $\sigma = c_1 c_2 \cdots c_m$. Soit p dans E_n. Si $\sigma(p) = p$, alors, le prisonnier p trouve son matricule dans la première boîte. On suppose $\sigma(p) \neq p$, alors p appartient à un seul des cycles c_l, puisque les cycles sont disjoints. De plus, $\sigma(p) = c_l(p)$. L'élément $c_l(p)$ appartient encore au cycle c_l par construction, et donc $\sigma^2(p) = c_l^2(p)$. Par récurrence, $\sigma^k(p) = c_l^k(p)$.

Supposons que tous les cycles de la décomposition de σ soient de longueur inférieure à $\frac{n}{2}$. Dans ce cas, chaque c_l est d'ordre $l \leq \frac{n}{2}$, et ainsi, $\sigma^k(p) = p$ pour un $k \leq \frac{n}{2}$. On a gagné (la vie, et l'agreg. Wouah, la chance !).

En revanche, supposons que σ contienne un cycle de longueur l, $l > \frac{n}{2}$, alors tout prisonnier p appartenant à ce cycle ne pourra trouver son matricule au bout de la procédure.

Reste à calculer la probabilité pour que σ soit une permutation dont les longueurs des cycles, dans la décomposition en cycles disjoints, soient toutes inférieures à $\frac{n}{2}$. Il vaut mieux calculer la probabilité de l'évènement complémentaire. Effectivement, s'il existe un cycle de longueur l, $l > \frac{n}{2}$, alors, celui-ci est unique (dans le sens où il ne peut pas y en avoir deux de longueur l, $l > \frac{n}{2}$, dans la décomposition de σ).

On fixe donc l, $l > \frac{n}{2}$, et on cherche le nombre de σ de $\mathfrak{S}_n$ telles que σ possède un cycle de longueur l dans sa décomposition. Pour une partie fixée de E_n de cardinal l, il y a $\frac{l!}{l} = (l-1)!$ l-cycles qui permutent cette partie. Effectivement, un cycle peut s'écrire exactement de l façons distinctes. Il ne reste plus qu'à permuter (de façon quelconque) les $n - l$ éléments restants. Il y a en tout, par unicité de la décomposition en cycles disjoints

$$\binom{n}{l} \times (n-l)!(l-1)! = \frac{n!}{l}$$

permutations possédant un cycle de longueur l.

Par les unicités indiquées ci-dessus, cela nous fait en tout

$$\sum_{l=\frac{n}{2}+1}^{n} \frac{n!}{l}$$

possibilités. C'est-à-dire que la probabilité de l'évènement complémentaire est

$$\frac{1}{n!}\sum_{l=\frac{n}{2}+1}^{n} \frac{n!}{l} = \sum_{l=\frac{n}{2}+1}^{n} \frac{1}{l} = \frac{1}{n}\sum_{l=\frac{n}{2}+1}^{n} \frac{1}{\frac{l}{n}}.$$

Quand n tend vers l'infini, cette probabilité tend donc vers

$$\int_{\frac{1}{2}}^{1} \frac{1}{t}\mathrm{d}t = \ln(2).$$

La probabilité cherchée est donc, pour n grand, proche de $1-\ln(2) \simeq 0,3068$.

Par exemple pour $n = 100$, on a

$$(\frac{1}{2})^{100} = 7,888 \times 10^{-31}, \;\; 1 - \sum_{l=51}^{100} \frac{1}{l} = 0,312$$

Y a pas photo !

Remarque. Si σ possède un cycle (forcément unique) de longueur $l > \frac{n}{2}$, alors les prisonniers du cycle correspondant n'auront aucune chance de retrouver leur matricule dans les boîtes.

IV.2 Arithmétique

IV.2.1 Arithmétique et théorème de Lagrange

Exercice IV.2.1 (Nombres de Fermat)**.** Soit $F_n = 2^{2^n}+1$ le n-ième nombre de Fermat. Montrer que si p est un premier qui divise F_n alors p congru à 1 mod 2^{n+1}.

Soluce. On travaille dans $(\mathbb{Z}/p\mathbb{Z})^*$, les classes seront donc des classes modulo p.

Puisque p divise F_n, $\overline{F_n} = \overline{0}$, et donc $\overline{2^{2^n}} = \overline{-1}$. En élevant au carré, on obtient $\overline{2^{2^{n+1}}} = \overline{1}$, et donc $\overline{2}^{2^{n+1}} = \overline{1}$.

L'ordre d de 2 dans $(\mathbb{Z}/p\mathbb{Z})^*$ divise donc 2^{n+1}. Montrons par l'absurde que $d = 2^{n+1}$: Si d divise strictement 2^{n+1} alors $d = 2^k$ avec $0 < k \leq n$. On a alors en particulier que d divise 2^n et par conséquent, $\overline{2^{2^n}} = \overline{1}$ ce qui contredit le fait que $2^{2^n} = -1$. Donc $d = 2^{n+1}$.

Maintenant, par Lagrange, d divise $p-1$ (qui est l'ordre du groupe $(\mathbb{Z}/p\mathbb{Z})^*$), donc $p-1 \equiv 0 \;\; (2^{n+1})$, d'où $p \equiv 1 \;\; (2^{n+1})$.

Exemple IV.2.2. Par exemple, $2^{32}+1$ est divisible par 641 (genre, de tête), et 641 est congru à 1 modulo 64.

Exercice IV.2.3.
Montrer qu'il existe un multiple de 2017 qui se termine par autant de 9 que l'on veut dans son écriture décimale. Même question avec n'importe quelle terminaison. Est-ce que cela marche avec 2013 ? avec 2015 ?

Soluce. On veut montrer que pour tout n, il existe k tel que $2017k = -1$ modulo 10^n. Or, 2017 est premier, il est donc premier avec 10, et donc avec 10^n. Le nombre 2017 possède donc un inverse a modulo 10^n, donc k existe, il suffit de prendre $-a$.

Cela marche encore pour 2013, même si celui-ci n'est pas un nombre premier, il est tout de même premier avec 10^n. Il a donc un inverse dans $(\mathbb{Z}/10^n\mathbb{Z})^*$

En revanche, cela ne marche plus pour 2015, puisque tous les multiples de 2015 se terminent par un 5 ou par un 0. On voit aussi que la preuve ne marche plus, puisque 2015 n'est pas premier avec 10.

Avec les nombres 2017 ou 2013, toute terminaison aurait convenu : une fois que l'on a trouvé l'inverse a modulo 10^n, il n'y a plus qu'à multiplier a par la terminaison en question.

Exercice IV.2.4. Les deux dates préférées de tout citoyen français qui se respecte sont 1789 et 1968. On peut remarquer qu'il existe un multiple de 1789 qui dans son écriture décimale se termine par 1968. Sauriez-vous expliquer ce phénomène ?

Soluce. Même chose : on veut montrer qu'il existe un entier k tel que $k\times 1789$ se termine par 1968. On travaille dans le groupe multiplicatif $(\mathbb{Z}/10000\mathbb{Z})^*$ [3]. En effet, 1789 est premier avec 10, donc, il l'est aussi avec 10000. Ainsi, $\overline{1789}$ est un élément du groupe $(\mathbb{Z}/10000\mathbb{Z})^*$. Notons $\overline{a}$ son inverse dans $(\mathbb{Z}/10000\mathbb{Z})^*$.

Une solution est de prendre k (entier positif !) tel que $\overline{k} = \overline{1968}\cdot\overline{a}$.

Exercice IV.2.5.
Quel est le dernier chiffre de l'écriture décimale de 7^{3^9} ?

Soluce. Il faut calculer 7^{3^9} modulo 10, puisqu'on ne s'intéresse qu'au chiffre des unités. Comme le groupe multiplicatif $(\mathbb{Z}/10\mathbb{Z})^*$ est d'ordre $\varphi(10) = 4$, d'après le théorème de Lagrange, $\overline{7}^4 = \overline{1}$.

Si on trouve le reste de 3^9 dans la division euclidienne par 4, on a (pratiquement) gagné. Effectivement, si $3^9 = 4k + r$, alors

$$7^{3^9} = 7^{4k+r} = (7^4)^k 7^r,$$

3. Quel est son ordre ?

qui est congru à 7^r modulo 10, d'après ce qui précède.

Or, $3 \equiv -1\ (4)$, donc l'ordre de 3 modulo 4 est 2, d'où : $3^9 \equiv 3\ (4)$, c'est-à-dire que le reste est 3. Le dernier chiffre est donc aussi le dernier chiffre de 7^3. On voit facilement que la réponse est 3.

Exercice IV.2.6. Trouver les trois derniers chiffres de 777^{401} dans son écriture décimale.

Soluce. Même chose. Il faut calculer 777^{401} modulo 1000. Mais cette fois-ci l'ordre de 777 est difficile à calculer dans $(\mathbb{Z}/1000\mathbb{Z})^*$. L'astuce consiste à voir que $(\mathbb{Z}/1000\mathbb{Z})^*$ est d'ordre $\phi(1000) = 1000 \times (1-1/2)(1-1/5) = 400$. Donc par Lagrange 777^{400} vaut 1 modulo 1000. Ainsi la réponse est 777.

IV.2.2 Théorème chinois

Notation. Pour tous x et m entiers, on notera $\bar{x}_m$ la réduction de x modulo m.

Exercice IV.2.7 (Théorème chinois, système de congruences)**.**
Soit m et n premiers entre eux. On considère le morphisme d'anneaux :

$$\tilde{\varphi}\colon \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$$

$$x \mapsto (\bar{x}_m, \bar{x}_n)$$

1. Déterminer $\ker \tilde{\varphi}$.
2. En déduire un isomorphisme d'anneau $\varphi : \mathbb{Z}/mn\mathbb{Z} \simeq \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$.
3. Soit $a, b \in \mathbb{Z}$, on considère le système :

$$(S) \begin{cases} x \equiv a \ \ [m] \\ x \equiv b \ \ [n] \end{cases}$$

   et le morphisme d'anneaux

$$\varphi\colon \mathbb{Z}/mn\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$$

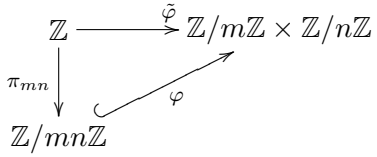
$$\bar{y}_{mn} \mapsto (\bar{y}_m, \bar{y}_n)$$

   (a) Déterminer $\varphi(\bar{x}_{mn})$ si x est solution de (S).
   (b) A l'aide du théorème de Bezout, déterminer $\varphi^{-1}(\bar{1}_m, \bar{0}_n)$ et $\varphi^{-1}(\bar{0}_m, \bar{1}_n)$
   (c) Résoudre (S) dans $\mathbb{Z}$.
4. Résoudre dans $\mathbb{Z}$ le système

$$(S) \begin{cases} x \equiv 6 \ \ [12] \\ x \equiv 13 \ \ [19] \end{cases}$$

Soluce. 1. L'entier x est dans $\ker(\tilde{\varphi})$ si et seulement si $(\bar{x}_m, \bar{x}_n) = (\bar{0}_m, \bar{0}_n)$. Ceci est vrai si et seulement si m divise x et n divise x. Par la propriété fondamentale du ppcm, c'est équivalent à dire que le ppcm de m et n divise x. Comme m et n sont premiers entre eux, cela revient à dire que mn divise x.
Nous avons donc montré l'égalité $\ker(\tilde{\varphi}) = mn\mathbb{Z}$.

2. On considère le passage au quotient par le noyau de $\tilde{\varphi}$. On obtient un morphisme injectif comme indiqué sur le diagramme :



où π_{mn} est la surjection canonique.
Or, $\#(\mathbb{Z}/mn\mathbb{Z}) = mn$ et $\#(\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}) = m \times n$, donc φ est bijective comme application injective entre deux ensembles de mêmes cardinaux.

3. (a) Si x est une solution de (S), alors $\varphi(\bar{x}_{mn}) = (\bar{a}_m, \bar{b}_n)$.
   (b) D'après l'identité de Bezout, avec m et n premiers entre eux, il existe u, v dans $\mathbb{Z}$ tels que : $um + vn = 1$.
   On a donc : $um \equiv 0[m], um \equiv 1[n]$, et ainsi, $\tilde{\varphi}(um) = (\bar{0}_m, \bar{0}_n)$.
   De même, $vn \equiv 1[m], vn \equiv 0[n]$, donc $\tilde{\varphi}(vn) = (\bar{1}_m, \bar{0}_n)$.
   Ceci donne la réponse à la question posée car φ est bijectif.
   (c) Comme
   $$(\bar{a}_m, \bar{b}_n) = a(\bar{1}_m, \bar{0}_n) + b(\bar{0}_m, \bar{1}_n) = a\tilde{\varphi}(vn) + b\tilde{\varphi}(um)$$
   et comme $\tilde{\varphi}$ est un morphisme, $x := a(vn) + b(um)$ vérifie $\tilde{\varphi}(x) = (\bar{a}_m, \bar{b}_n)$. On a donc, par passage au quotient : $\varphi(\bar{x}_{mn}) = (\bar{a}_m, \bar{b}_n)$.
   De plus, comme φ est bijectif, cette solution est unique.
   La solution de (S) est donc : $\{(avn + bum) + kmn, k \in \mathbb{Z}\}$.
4. Les entiers 12 et 19 sont bien premiers entre eux, leur ppcm est donc $12 \cdot 19 = 228$. Déterminons des coefficients de Bezout ; on trouve par exemple : $8 \times 12 - 5 \times 19 = 1$. On a par la suite $-6 \cdot 5 \cdot 19 + 13 \cdot 8 \cdot 12 = 678$. L'ensemble des solutions est donc
   $$\{678 + 228k, k \in \mathbb{Z}\} = \{222 + 228k, k \in \mathbb{Z}\}.$$

Remarques

1. Les deux premières questions démontrent le sens direct du théorème chinois. La réciproque est également vraie :

   Proposition. Si les anneaux $\mathbb{Z}/mn\mathbb{Z}$ et $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ sont isomorphes, alors m et n sont premiers entre eux.

   Preuve : Supposons qu'il existe un isomorphisme d'anneaux
   $$g \colon \mathbb{Z}/mn\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}.$$
   Soit μ le ppcm de m et n. On voit que $\mu(\bar{a}_m, \bar{b}_n) = (\mu\bar{a}_m, \mu\bar{b}_n) = (\bar{0}_m, \bar{0}_n)$, pour tout $(\bar{a}_m, \bar{b}_n)$ de $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$. Par l'isomorphisme inverse g^{-1}, on obtient que $\mu\bar{x}_{mn} = \bar{0}_{mn}$ pour tout x dans $\mathbb{Z}$. En particulier, pour $x = 1$. On trouve donc $\mu = 0$ modulo mn. Or, μ divise mn, donc $\mu = mn$. Comme le pgcd de m et n est égal à $\frac{mn}{\mu}$, ceci implique que m et n sont premiers entre eux.

2. On généralise facilement la résolution des systèmes de congruences pour un nombre fini d'entiers deux à deux premiers entre eux. Par exemple, on cherche l'ensemble des x tels que x est congru à a, resp. b, resp. c, modulo l, resp. m, resp. n, avec l, m, n deux à deux premiers entre eux. On commence par chercher à résoudre le système constitué des deux premières congruences. La solution de ce premier système est de la forme $x \equiv x_0 \ \ [lm]$. Il ne reste plus qu'à résoudre le système donné par $x \equiv x_0 \ \ [lm]$ et $x \equiv c \ \ [n]$, en remarquant que, par hypothèse, lm et n sont premiers entre eux.

Exercice IV.2.8 (Système de congruences, cas général)**.**
On veut étudier un système de congruences comme dans l'exercice précédent, mais dans le cas où le pgcd de m et n est un entier δ. On notera μ le ppcm de m et n.
On considère le morphisme d'anneaux :

$$\tilde{\varphi}\colon \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$$

$$x \mapsto (\bar{x}_m, \bar{x}_n)$$

1. Déterminer $\ker \tilde{\varphi}$.
2. En déduire un morphisme injectif $\mathbb{Z}/\mu\mathbb{Z} \hookrightarrow \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$, que l'on notera φ dans la suite.
3. On suppose que r et s sont deux entiers tels que r divise s. Montrer que le morphisme de $\mathbb{Z}/s\mathbb{Z}$ dans $\mathbb{Z}/r\mathbb{Z}$ qui envoie $\bar{x}_s$ sur $\bar{x}_r$ est bien défini, c'est-à-dire, ne dépend pas du représentant x de $\bar{x}_s$ choisi dans $\mathbb{Z}$.
4. En déduire que l'on définit un morphisme ψ par

   $$\psi : \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/\delta\mathbb{Z},\ (\bar{x}_m, \bar{y}_n) \mapsto \bar{x}_\delta - \bar{y}_\delta,$$

   et que celui-ci est surjectif.
5. Montrer que $\mathrm{Im}(\varphi) \subset \ker(\psi)$ et en déduire l'égalité $\mathrm{Im}(\varphi) = \ker(\psi)$
6. Montrer que le système suivant n'a pas de solution :

   $$(S)\begin{cases} x \equiv 2 & [21] \\ x \equiv 11 & [35] \end{cases}$$

7. On suppose que a et b vérifient $\bar{a}_\delta = \bar{b}_\delta$. On veut résoudre le système

   $$(S)\begin{cases} x \equiv a & [m] \\ x \equiv b & [n] \end{cases}$$

   Soit u, v deux entiers tels que $um + vn = \delta$.
   (a) Montrer que $x_0 := \frac{1}{\delta}(avn + bun)$ est solution du système.
   (b) En déduire l'ensemble des solutions de (S).
8. Résoudre dans $\mathbb{Z}$

   $$(S)\begin{cases} x \equiv 3 & [21] \\ x \equiv 10 & [35] \end{cases}$$

Soluce. 1. Tout d'abord, $\ker\tilde{\varphi}$ a un sens puisque $\tilde{\varphi}$ est un morphisme de groupes (composante par composante, il s'agit de surjections canoniques). Le noyau de $\tilde{\varphi}$ est, par définition, constitué des x dans $\mathbb{Z}$ tels que $\bar{x}_m$ et $\bar{x}_n$ sont tous deux nuls. Ceci est équivalent au fait que m et n divisent x. Par la propriété fondamentale du ppcm, cela signifie que μ divise x. On a donc $\ker(\tilde{\varphi}) = \mu\mathbb{Z}$.

2. Par *passage au quotient*, on en déduit un morphisme injectif φ qui associe, à $\bar{x}_\mu$, l'élément $\varphi(\bar{x}_\mu) := \tilde{\varphi}(x)$. Le passage au quotient par le noyau assure que ce morphisme est *bien défini*, c'est-à-dire qu'il ne dépend pas du choix de x dans la classe $\bar{x}_\mu$. Notons au passage (mais pas au quotient !), que le fait d'avoir quotienté par le noyau, implique que l'image n'a pas été modifiée, c'est-à-dire $\mathrm{Im}(\varphi) = \mathrm{Im}(\tilde{\varphi})$.

3. On peut le montrer de deux façons, selon l'aisance que l'on a avec les structures algébriques. Une façon terre à terre consiste à écrire $s = kr$, avec k entier. Soit x et x' dans la classe $\bar{x}_s$. Alors, $x' = x+k's = x+k'kr$, et donc $\bar{x}'_r = \bar{x}_r$, ce qui prouve que l'image ne dépend pas du choix de x. On vérifie alors sans peine qu'il s'agit bien d'un morphisme.

   Sinon, une autre façon est de dire que, comme r divise s, on a l'inclusion $s\mathbb{Z} \subset r\mathbb{Z}$. Or, $r\mathbb{Z}$ est le noyau du morphisme canonique $\mathbb{Z} \to \mathbb{Z}/r\mathbb{Z}$. Le morphisme demandé est obtenu à partir du morphisme canonique, par passage au quotient par $s\mathbb{Z}$. L'avantage de cette méthode est que l'on a directement qu'il est bien défini, que c'est bien un morphisme, et qu'il est surjectif, puisque l'image ne change pas après passage au quotient.

4. Comme δ divise m et n, il vient que $\bar{x}_m \mapsto \bar{x}_\delta$ et $\bar{y}_n \mapsto \bar{y}_\delta$ définissent bien des morphismes, et donc ψ est un morphisme.

   Montrons qu'il est surjectif. Soit $\bar{x}_\delta$ dans $\mathbb{Z}/\delta\mathbb{Z}$ et x un représentant. Alors, $\psi(\bar{x}_m, \bar{0}_n) = \bar{x}_\delta$, ce qui prouve notre affaire.

5. Un élément de $\mathrm{Im}\,\varphi$ est de la forme $(\bar{x}_m, \bar{x}_n)$, pour un x dans $\mathbb{Z}$. On a alors
$$\psi(\bar{x}_m, \bar{x}_n) = \bar{x}_\delta - \bar{x}_\delta = \bar{0}_\delta,$$
   ce qui prouve la première assertion.

   L'égalité se déduit de l'inclusion par un argument de cardinalité. Le cardinal $|\,\mathrm{Im}\,\varphi|$ de $\mathrm{Im}\,\varphi$ est égal au cardinal de $\mathbb{Z}/\mu\mathbb{Z}$, puisque φ est injectif.

   On a donc
$$|\,\mathrm{Im}\,\varphi| = \mu = \frac{mn}{\delta} = \frac{|\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}|}{|\mathbb{Z}/\delta\mathbb{Z}|} = \frac{|\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}|}{|\,\mathrm{Im}(\psi)|} = |\ker(\psi)|.$$
   Ce qui achève la preuve.

6. Si, par l'absurde, une solution existait, on aurait un x dans $\mathbb{Z}$ tel que $\tilde{\varphi}(x) = (\bar{x}_{21}, \bar{x}_{35}) = (\bar{2}_{21}, \overline{11}_{35})$, et donc, $(\bar{2}_{21}, \overline{11}_{35})$ serait dans $\mathrm{Im}(\varphi)$, c'est-à-dire dans $\ker(\psi)$. Or, $\bar{2}_7 - \overline{11}_7 = \overline{-9}_7 \neq \bar{0}_7$. Absurde.

7. (a) On peut écrire $a = b + k\delta$, avec k entier. Ce qui donne

$$x_0 = \frac{1}{\delta}\big((b+k\delta)vn + bum\big) = b\frac{vn+um}{\delta} + kvn,$$

ce qui prouve $x_0 \equiv b \ \ [n]$. De la même manière, en remplaçant b par $a - k\delta$ dans l'expression de x_0, on trouve $x_0 \equiv a \ \ [m]$.

(b) Soit x une solution de (S). Alors, $\tilde{\varphi}(x) = \tilde{\varphi}(x_0)$, et donc $x - x_0 \in \ker(\tilde{\varphi}) = \mu\mathbb{Z}$.

L'ensemble des solutions est donc $\{x_0 + k\mu,\ k \in \mathbb{Z}\}$.

8. Cette fois-ci, nous sommes dans la situation où $(\bar{a}_m, \bar{b}_n) \in \ker(\psi)$, puisque $3 - 10$ est multiple de 7. On applique la méthode ci-dessus qui donne $2 \cdot 21 - 35 = 7$, puis $x_0 = \frac{1}{7}(-3\dot{3}5 + 10 \cdot 2 \cdot 21) = 45$.

L'ensemble des solutions est donc l'ensemble des $45 + 105k$, avec k parcourant $\mathbb{Z}$.

Remarque. On connaît tous la formule $mn = \mu\delta$, qui relie deux nombres entiers avec leur pgcd et leur ppcm. Ce qui est intéressant ici, est de voir que cette formule n'est que le reflet d'un phénomène qui se situe dans le monde des groupes, plus précisément, d'un morphisme (ici, j'ai nommé, ψ) qui part de $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$, dont l'image est $\mathbb{Z}/\delta\mathbb{Z}$ et dont le noyau est $\mathbb{Z}/\mu\mathbb{Z}$. C'est toute la poésie (et la puissance !) de l'algèbre moderne, de voir les nombres comme des ombres d'un monde solidement structuré (ici le monde des groupes et des morphismes), façon caverne de Platon.

Attention toutefois : dans la formule arithmétique, les nombres μ et δ commutent allègrement, alors que dans les morphismes, on ne peut pas inverser noyau et image. Le pgcd est lié à l'image, et le ppcm au noyau. C'est ce qu'il faut retenir de cet avatar du lemme chinois.

IV.2.3 Résidus quadratiques

Exercice IV.2.9. *[Résidus quadratiques, [H2G2] Chap. V, Proposition C1] Soit p un nombre premier impair.

1. Montrer que $\phi : (\mathbb{Z}/p\mathbb{Z})^* \to (\mathbb{Z}/p\mathbb{Z})^*$, $x \mapsto x^2$ est un morphisme de groupe.

2. Montrer que le noyau de ϕ est d'ordre 2. En déduire que $x^{\frac{p-1}{2}} = \pm 1$.

3. (a) Montrer que si x est un carré de $(\mathbb{Z}/p\mathbb{Z})^*$, alors $x^{\frac{p-1}{2}} = 1$.

   (b) Réciproquement, montrer en utilisant une propriété sur le nombre de racines d'un polynôme sur un corps, que si $x^{\frac{p-1}{2}} = 1$, alors x est un carré.

Soluce. 1. Rappelons que $(\mathbb{Z}/p\mathbb{Z})^*$ est l'ensemble des inversibles de l'anneau $\mathbb{Z}/p\mathbb{Z}$. Il est donc toujours muni d'une structure de groupe multiplicatif et donc, ϕ est bien définie.

Ici, $\mathbb{Z}/p\mathbb{Z}$ est un corps puisque p est premier. Donc, $(\mathbb{Z}/p\mathbb{Z})^*$ est tout simplement $\mathbb{Z}/p\mathbb{Z}$ privé de 0.

On a, par commutativité : $\phi(1) = 1$; $\phi(xy) = (xy)^2 = x^2y^2 = \phi(x)\phi(y)$ et $\phi(x^{-1}) = (x^{-1})^2 = x^{-2} = \phi(x)^{-1}$.

Donc ϕ est bien un morphisme de groupe.

2. Soit $a \in \ker\phi$, c'est-à-dire $\phi(a) = a^2 = 1$. D'où $(a+1)(a-1) = 0$.

Comme $\mathbb{Z}/p\mathbb{Z}$ est un corps, il est donc intègre. Par conséquent, $a = 1$ ou $a = -1$.

En outre, $p \neq 2$ donc $2 \neq 0$ et $1 \neq -1$. L'équation précédente a donc deux solutions disctinctes et $\ker\phi$ est d'ordre 2.

On rappelle que $(\mathbb{Z}/p\mathbb{Z})^*$ est d'ordre $p-1$. Soit $x \in (\mathbb{Z}/p\mathbb{Z})^*$. On a $(x^{\frac{p-1}{2}})^2 = x^{p-1} = 1$, d'après le théorème de Lagrange. D'où : $x^{\frac{p-1}{2}} \in \ker\phi$ et $x^{\frac{p-1}{2}} = \pm 1$.

3. (a) Supposons que x soit un carré de $(\mathbb{Z}/p\mathbb{Z})^*$. Il existe $y \in (\mathbb{Z}/p\mathbb{Z})^*$ tel que $x = y^2$. On a alors : $x^{\frac{p-1}{2}} = y^{p-1} = 1$ par Lagrange.

   (b) Le nombre de carrés (non nuls) est égal au cardinal de $\operatorname{Im}\phi$ par définition de ϕ. Il est donc égal à $\#(\mathbb{Z}/p\mathbb{Z})^*/\#\ker\phi$, c'est-à-dire $(p-1)/2$ d'après 2). Donc, d'après 3)a), le polynôme $X^{(p-1)/2} - 1$ possède au moins $(p-1)/2$ solutions (les carrés !). Or, comme on travaille sur un corps, le polynôme, qui est de degré $(p-1)/2$, n'en possède qu'au plus $(p-1)/2$. Donc les carrés sont les seules racines de ce polynôme, ce qui implique la réciproque voulue.

   Conclusion : l'application qui, à $x \in (\mathbb{Z}/p\mathbb{Z})^*$ associe $\left(\frac{x}{p}\right) := x^{\frac{p-1}{2}}$ est un morphisme surjectif sur $\{1; -1\}$ et son noyau est le sous-groupe des carrés de $\big((\mathbb{Z}/p\mathbb{Z})^*, \times\big)$

Exercice IV.2.10. Montrer que si p premier impair divise $n^2 + 1$, alors p est congru à 1 modulo 4.

Soluce. On travaille modulo p. Dans $(\mathbb{Z}/p\mathbb{Z})^*$, on a que $n^2 = -1$ (je devrais mettre des barres, mais j'en suis conscient !). Alors, -1 est un carré modulo p, ce qui implique par le symbole de Legendre que p est congru à 1 modulo 4.

Sans utiliser le symbole de Legendre :

Le fait que $n^2 = -1$ entraîne que $n^4 = 1$. Ainsi, l'ordre de n est un diviseur de 4. Or, cet ordre ne saurait [4] être égal à 1 ou 2 car, le cas échéant, on aurait $n^2 = 1$. Ainsi, l'ordre de n est égal à 4.

4. Y a qu'en maths qu'on peut encore parler comme ça, non ?

Par le théorème de Lagrange, on déduit alors que 4 divise $p-1$, d'où $p \equiv 1\ (4)$.

Remarque. On peut même se payer le luxe d'exhiber les racines de $x^2 = -1$ sur $\mathbb{Z}/p\mathbb{Z}$ dans ce cas. Comme $\mathbb{Z}/p\mathbb{Z}$ est un corps, une équation du second degré a au plus deux racines ; elles sont opposées, et il suffit donc d'en avoir une pour avoir l'autre. On a p congru à 1 modulo 4 et p premier, donc, si on pose $x = 1 \cdot 2 \cdots \frac{p-1}{2}$, il vient (dans $\mathbb{Z}/p\mathbb{Z}$) :

$$x^2 = \big(1 \cdot 2 \cdots \frac{p-1}{2}\big)\big(\frac{p-1}{2} \cdots 2 \cdot 1\big).$$

Or, $\frac{p-1}{2}$ est pair par hypothèses, il vient donc, en multipliant par $(-1)^{\frac{p-1}{2}}$:

$$x^2 = \big(1 \cdot 2 \cdots \frac{p-1}{2}\big)\big((-\frac{p-1}{2}) \cdots (-2) \cdot (-1)\big).$$

Par le théorème de Wilson, on obtient finalement :

$$x^2 = \big(1 \cdot 2 \cdots \frac{p-1}{2}\big)\big((\frac{p+1}{2}) \cdots (p-2) \cdot (p-1)\big) = (p-1)! = -1.$$

Le nombre x proposé est bien une racine de -1 modulo p.

IV.3 Codage RSA

Exercice IV.3.1. Le codage RSA. Principe, preuve et exemple.
Vous avez 3 heures.

Principe Bob transmet un message chiffré m à Alice qui va le décoder.

Celui qui code connaît la clé publique, celui qui décode doit connaître la clé privée.

1. Préparatifs : fabrication des clefs publique et privée

   *A*lice propose une clef publique pour qu'on puisse lui envoyer des messages codés. Elle possède un code privé sous la forme d'un nombre couple (n, d) de nombres entiers. Bob veut lui écrire un message codé, il a accès à la clef publique sous la forme d'un couple (n, e). L'idée de la méthode est qu'il est très difficile, voire quasiment impossible de trouver d à partir de e. Cela demande de factoriser n qui est composé de deux nombres premiers très grands, p et q.

   On choisit deux nombres premiers distincts très grands p et q, qu'on garde secrets. On calcule $n = pq$, puis $\varphi(n)$, où φ est la fonction indicatrice d'Euler. Comme $p \wedge q = 1$, alors $\varphi(pq) = (p-1)(q-1)$. Ensuite, on choisit e premier avec $(p-1)(q-1)$, et enfin on calcule d, l'inverse de e modulo $(p-1)(q-1)$, en utilisant les coefficients de Bezout. On a donc $ed \equiv 1 \ \ [(p-1)(q-1)]$.

   Les nombres p, q, d sont gardés secrets. Le couple (n, e) est transmis à celui qui code, c'est la clé publique.

2. Codage par Bob du message m. Bob calcule le message codé $M = m^e[n]$, puis, l'envoie à Alice.

3. Décodage par Alice. Alice reçoit le message M et calcule M^d modulo n (c'est-à-dire $(m^e)^d$ modulo n) ce qui lui permet de récupérer le message d'origine, puisque $(m^e)^d \equiv m[n]$.

Preuve Montrons que $m^{ed} \equiv m[n]$

- Tout d'abord, on a $ed \equiv 1 \ \ [(p-1)(q-1)]$. Donc il existe k dans $\mathbb{Z}$ tel que $ed = 1 + k(p-1)(q-1)$.
- Montrons ensuite que $m^e d \equiv m[p]$
  - Si m n'est pas premier avec p, alors p divise m, et donc p divise m^{ed} et donc
    $$m^{ed} \equiv 0 \equiv m[p].$$
  - Si $m \wedge p = 1$ alors d'après le petit théorème de Fermat, $m^{p-1} \equiv 1[p]$. Il vient alors
    $$m^{ed} \equiv m^{1+k(p-1)(q-1)} \equiv m \cdot (m^{p-1})^{k(q-1)} \equiv m \cdot 1 \equiv m[p].$$

  Dans les deux cas, $m^{ed} \equiv m[p]$.

- De même, on montre que $m^{ed} \equiv m[q]$
- Si $m^{ed} \equiv m[p]$ et $m^{ed} \equiv m[q]$ alors $m^{ed} \equiv m[pq]$. C'est le théorème chinois, comme dans l'exercice IV.2.7, qui dit qu'un nombre modulo pq est uniquement déterminé par ses congruences modulo p et modulo q, puisqu'ici, p et q sont premiers entre eux.
  On a donc bien $m^{ed} \equiv m[n]$.

Exemple

1. On choisit $p = 3, q = 11$ et on calcule $n = pq = 33$. On a donc $\varphi(33) = (p-1)(q-1) = 20$.
   Puis on choisit e premier avec 20, par exemple $e = 7$. La clé publique sera donc le couple $(33, 7)$ (transmise à Bob). Ensuite, on cherche d inverse de e modulo 20. Par exemple $d = 3$. La clé privée est donc $(33, 3)$ (possédée par Alice).
2. Bob veut envoyer le message chiffré $m = 9$. Il calcule $m^e = 9^7 = 4782969$. Or, $4782969 \equiv 15[33]$. Bob envoie donc le message chiffré $M = 15$ à Alice.
3. Alice reçoit le message $M = 15$, calcule $M^d = 15^3 = 3375$, puis $3375 \equiv 9[33]$. Donc le message décodé est donc bien $m = 9$.

Référence

[H2G2] Philippe Caldero et Jérôme Germoni : *Histoires hédonistes de groupes et de géométries,* Tome premier, Mathématiques en devenir, Calvage et Mounet, Paris, 2013.
[Szpirglas] Aviva Szpirglas : *Mathématiques L3 Algèbre,* Pearson Education, Paris, 2009.

« Travailler moins pour gagner plus » : ce fascicule d'exercices en algèbre vous amènera dans un premier temps sur les sentiers fleuris de l'agrégation interne. Dans le prochain volume, nous résoudrons, hop-hop, la crise mondiale.