

Mathématiques générales

Notes de cours

Arnaud Durand

Département de Mathématiques d'Orsay

arnaud.durand@universite-paris-saclay.fr

Avant-propos

Ce document a été rédigé pendant l'année universitaire 2002–2003, à partir d'un cours de Claudine Picaronny suivi à l'ENS Cachan. Les références ont été actualisées et le texte a été (un peu) révisé en 2026.

Conventions et notations

Dans l'ensemble de ces notes, on adopte les conventions suivantes. On note $\mathbb{N} = \{0, 1, 2, \dots\}$, $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$, et $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ les ensembles usuels de nombres. Le symbole i désigne l'unité imaginaire, et $\mathbb{U} = \{z \in \mathbb{C} : |z| = 1\}$ le cercle unité. Pour un ensemble E , $|E|$ désigne son cardinal. Le symbole $\subset$ autorise l'égalité ; une inclusion stricte est notée $\subsetneq$. Une somme vide vaut 0 et un produit vide vaut 1.

Les anneaux sont unitaires et leurs morphismes préservent l'unité ; un anneau intègre est non nul. On note A^* le groupe des éléments inversibles d'un anneau A , et $(a_1, \dots, a_r)$ l'idéal engendré par ces éléments. Les notations $a \wedge b$ et $a \vee b$ désignent respectivement un pgcd et un ppcm, lorsqu'ils existent, définis à un facteur inversible près ; pour les polynômes sur un corps, on choisit les représentants unitaires lorsqu'ils sont non nuls. La caractéristique $\text{Car}(A)$ est le plus petit entier $n \geq 1$ tel que $n1_A = 0$, s'il en existe un, et vaut 0 sinon.

Pour un ensemble E , $\mathfrak{S}(E)$ est le groupe de ses permutations ; on pose $\mathfrak{S}_n = \mathfrak{S}(\{1, \dots, n\})$ et on note $\mathfrak{A}_n$ son sous-groupe des permutations paires. Pour un groupe G , 1 désigne l'élément neutre, $\text{Aut}(G)$ son groupe d'automorphismes et $\langle S \rangle$ le sous-groupe engendré par S . Si H est un sous-groupe de G , G/H désigne l'ensemble des classes à gauche, $[G : H]$ leur nombre, et $H \triangleleft G$ signifie que H est distingué ; dans ce dernier cas, G/H est un groupe quotient. La notation $H < G$ signifie ici que H est un sous-groupe, sans imposer une inclusion stricte.

Pour une extension de corps $\mathbb{K} \subset \mathbb{L}$, $[\mathbb{L} : \mathbb{K}]$ désigne la dimension de $\mathbb{L}$ comme espace vectoriel sur $\mathbb{K}$. Pour q puissance d'un nombre premier, $\mathbb{F}_q$ désigne un corps à q éléments, unique à isomorphisme près.

Pour des A -modules M, N , $\text{Hom}_A(M, N)$ désigne l'ensemble des applications A -linéaires de M dans N . Pour un espace vectoriel E sur $\mathbb{K}$, $E^* = \text{Hom}_{\mathbb{K}}(E, \mathbb{K})$ est son dual, $\mathcal{L}(E)$ l'algèbre de ses endomorphismes et $\mathcal{GL}(E)$ le groupe de ses automorphismes. On note Id l'application identité, $\ker u$ et $\text{Im } u$ le noyau et l'image d'une application linéaire u , et $\text{rg } u = \dim \text{Im } u$ son rang. La notation $\langle S \rangle$ désigne aussi, selon le contexte, le sous-module ou le sous-espace vectoriel engendré par S ; dans ce dernier cas, on emploie également $\text{Vect}(S)$.

Pour un anneau commutatif A , $M_{m,n}(A)$ est l'ensemble des matrices à m lignes et n colonnes, $M_n(A) = M_{n,n}(A)$, I_n la matrice identité, $\text{GL}_n(A)$ le groupe des matrices inversibles et $\text{SL}_n(A)$ celui des matrices de déterminant 1. Le symbole de Kronecker $\delta_{i,j}$ vaut 1 si $i = j$ et 0 sinon. Pour une matrice M , ${}^t M$ désigne sa transposée ; si ses coefficients sont complexes, $\overline{M}$ est sa conjuguée coefficient par coefficient et $M^* = {}^t \overline{M}$ sa transposée conjuguée. La notation tr désigne la trace d'une matrice carrée ou d'un endomorphisme en dimension finie.

Table des matières

Avant-propos	i
1 Rappels sur les groupes	1
1.1 Opération d'un groupe sur un ensemble	1
1.2 Produit direct de deux groupes	3
1.3 Produit semi-direct de deux groupes	4
2 Corps finis	8
2.1 Rappels	8
2.2 Constructions de corps finis	10
2.3 Structure des corps finis	13
2.4 Prolongements	16
3 Réduction d'un endomorphisme	19
3.1 Polynôme minimal et polynôme caractéristique d'un endomorphisme	19
3.2 Sous-espaces stables	20
3.3 Sous-espaces cycliques et théorème de Cayley-Hamilton	21
3.4 Lemme de décomposition des noyaux	23
3.5 Invariants de similitude	25
3.6 Endomorphismes semi-simples	30
3.7 Réduction des endomorphismes normaux	33
4 Modules sur un anneau principal	39
4.1 Généralités sur les modules	39
4.2 Classes d'isomorphisme des modules de type fini sur un anneau principal	40
4.3 Applications	44
5 Formes bilinéaires, formes quadratiques	55
5.1 Généralités	55
5.2 Classification des formes quadratiques	59
5.3 Isotropie	63
6 Sous-groupes finis de $O_2(\mathbb{R})$ et $O_3(\mathbb{R})$	67
6.1 Sous-groupes finis de $O_2(\mathbb{R})$	67
6.2 Sous-groupes finis de $O_3(\mathbb{R})$	68
6.3 Problème de la coloration du cube	75
7 Combinatoire	80
7.1 Formule du crible	80
7.2 Séries génératrices	81
8 Introduction à la géométrie algébrique	84
8.1 Algèbres de polynômes à plusieurs indéterminées	84
8.2 Ensembles algébriques	87
Bibliographie	92

Rappels sur les groupes

Références : [4, 11]

1.1 Opération d'un groupe sur un ensemble

Cette section contient des rappels sur les actions de groupes. Nous donnerons leur définition et quelques-unes de leurs propriétés. Nous finirons par une application aux p -groupes.

1.1.1 Définition

Définition 1.1. Une *opération d'un groupe G sur un ensemble E* est la donnée d'une application

$$G \times E \rightarrow E, \quad (g, e) \mapsto g \cdot e$$

telle que pour tous $e \in E$, $g, g' \in G$, on ait

- (i). $(gg') \cdot e = g \cdot (g' \cdot e)$.
- (ii). $1 \cdot e = e$.

De manière équivalente, il s'agit aussi de la donnée d'un morphisme de groupes

$$\Psi : G \rightarrow \mathfrak{S}(E), \quad g \mapsto (e \mapsto g \cdot e).$$

Remarque. On a $g \cdot e = \Psi(g)(e)$.

1.1.2 Orbites et stabilisateurs

Proposition 1.1. Une opération du groupe G sur E définit une relation d'équivalence sur E :

$$e \sim e' \iff \exists g \in G : e = g \cdot e'.$$

Nous verrons bientôt que cette relation d'équivalence est primordiale dans l'étude des actions de groupes. Donnons quelques propriétés qui lui sont relatives.

Définition 1.2. Une classe d'équivalence par la relation $\sim$ est appelée *orbite*.

Proposition 1.2. L'orbite de e est

$$\omega(e) = \{g \cdot e, g \in G\}.$$

L'application

$$G \rightarrow \omega(e), \quad g \mapsto g \cdot e$$

est donc surjective et on a $g \cdot e = g' \cdot e$ si et seulement si $(g'^{-1}g) \cdot e = e$.

Définition 1.3. On appelle *stabilisateur* de e et on note $\text{Stab}_G(e)$ le sous-groupe de G défini par

$$\text{Stab}_G(e) = \{g : g \cdot e = e\}.$$

Avec cette définition, on peut donc écrire que $g \cdot e = g' \cdot e$ si et seulement si $g \text{Stab}_G(e) = g' \text{Stab}_G(e)$. On en déduit la proposition suivante et son corollaire immédiat.

Proposition 1.3. $G/\text{Stab}_G(e)$ est en bijection d'ensembles avec $\omega(e)$.

Dans la suite, nous noterons $|E|$ le cardinal d'un ensemble E .

Corollaire 1.4. Si G est fini, $|\omega(e)|$ divise $|G|$.

1.1.3 Equation aux classes

La dernière proposition concernant les actions de groupes que nous citerons est simple, mais pourtant très importante. Elle est connue sous le nom d'équation aux classes.

Proposition 1.5 (Equation aux classes). On note E_G l'ensemble des orbites de E sous l'action de G . Si E est fini, on a

$$|E| = \sum_{\omega \in E_G} |\omega|.$$

Démonstration. Il suffit de remarquer que les classes d'équivalence par la relation d'équivalence $\sim$ forment une partition de E . □

1.1.4 Une application à l'étude des p -groupes

Nous allons terminer cette section sur les opérations de groupes par une application à l'étude des p -groupes.

Définition 1.4. Soit p un nombre premier. On appelle p -groupe tout groupe fini dont l'ordre est une puissance de p .

Proposition 1.6. Le centre d'un p -groupe non trivial est non trivial.

Démonstration. Soit P un p -groupe non trivial. On fait opérer P sur lui-même par automorphisme intérieur.

$$P \times P \rightarrow P, \quad (g, x) \mapsto gxg^{-1}.$$

L'orbite de x est sa classe de conjugaison. Le stabilisateur de x est l'ensemble des éléments de P qui commutent avec x , c'est-à-dire le centralisateur de x

$$C_P(x) := \{g \in P : gx = xg\}.$$

Faisons quelques remarques.

- (i). Si $|\omega(x)| = 1$, $\text{Stab}_P(x) = P$ donc $x \in Z(P)$ (le centre de P , i.e. l'ensemble des éléments qui commutent avec tous les éléments de P) et réciproquement.
- (ii). Si $x \notin Z(P)$, $|\omega(x)|$ est une puissance de p non triviale, donc p divise $|\omega(x)|$.

On écrit alors l'équation aux classes.

$$|P| = \sum_{\omega} |\omega| = \underbrace{\sum_{|\omega|=1} |\omega|}_{=|Z(P)|} + \underbrace{\sum_{|\omega|>1} |\omega|}_{\in p\mathbb{N}}.$$

On conclut en utilisant le fait que p divise $|P|$, donc divise $|Z(P)|$. □

1.2 Produit direct de deux groupes

Cette section a pour vocation de fixer les idées sur le produit direct de deux groupes. Nous ne rentrerons pas dans les détails.

1.2.1 Définition

Définition 1.5. Le *produit direct* de H et K est l'ensemble $H \times K$ muni du produit

$$(h, k)(h', k') := (hh', kk').$$

1.2.2 Quelques propriétés

L'application

$$H \rightarrow H \times K, \quad h \mapsto (h, 1)$$

est un morphisme de groupes injectif. On note $\overline{H}$ son image. Il s'agit d'un sous-groupe de $H \times K$.

De même, l'application

$$K \rightarrow H \times K, \quad k \mapsto (1, k)$$

est aussi un morphisme de groupes injectif. On note $\overline{K}$ son image. Il s'agit également d'un sous-groupe de $H \times K$.

Proposition 1.7. On dispose des propriétés suivantes.

- (i). $\overline{H}$ et $\overline{K}$ commutent $((h, 1)(1, k) = (h, k) = (1, k)(h, 1))$.
- (ii). $(h, k) = (h, 1)(1, k)$ et $H \times K = \langle \overline{H}, \overline{K} \rangle$ (le sous-groupe engendré par $\overline{H}$ et $\overline{K}$).
- (iii). $\overline{H} \cap \overline{K} = \{(1, 1)\}$, ce qui est équivalent au fait que la décomposition précédente de (h, k) en produit d'un élément de $\overline{H}$ par un élément de $\overline{K}$ est unique.

Intéressons-nous maintenant au cas où le produit direct n'est pas construit à partir de deux groupes quelconques, mais au cas où il provient de deux sous-groupes d'un même groupe.

Définition 1.6. Soit G un groupe. Soient H et K deux sous-groupes de G . On dit que G est *produit direct* de H et K si l'application

$$H \times K \rightarrow G, \quad (h, k) \mapsto hk$$

est un isomorphisme de groupes.

Exemple. $H \times K$ est produit direct de $\overline{H}$ et $\overline{K}$.

La caractérisation du produit direct qui figure ci-dessous est certainement la proposition la plus importante de cette section.

Proposition 1.8. Soit G un groupe. Soient H et K deux sous-groupes de G . Alors G est produit direct de H et K si et seulement si

- (i). H et K commutent,
- (ii). $H \cap K = \{1\}$,
- (iii). $G = \langle H, K \rangle$.

Démonstration. On considère l'application

$$\varphi : H \times K \rightarrow G, \quad (h, k) \mapsto hk.$$

- φ est un morphisme de groupes si et seulement si

$$\varphi(h, k)\varphi(h', k') = \varphi((h, k)(h', k')) \iff hkh'k' = hh'kk',$$

soit encore

$$kh' = h'k,$$

ce qui signifie que H et K commutent.

- On suppose que φ est un morphisme. On a

$$\ker(\varphi) = \{(h, k) : hk = 1\} = \{(h, h^{-1}), h \in H \cap K\},$$

donc φ est injectif si et seulement si $H \cap K = \{1\}$.

- On suppose encore que φ est un morphisme. φ est surjectif si et seulement si $G = HK$, ce qui revient au fait que G est engendré par H et K , car ces deux groupes commutent.

□

Remarque. Supposons que l'on ait seulement

- (i). H et K commutent,
- (ii). $H \cap K = \{1\}$.

Alors $\langle H, K \rangle$ est produit direct de H et K . En particulier, si H et K sont finis, on a $|\langle H, K \rangle| = |H||K|$.

Exemple. $\mathbb{Z}/6\mathbb{Z}$ est produit direct de $\mathbb{Z}/3\mathbb{Z}$ et $\mathbb{Z}/2\mathbb{Z}$. On peut considérer l'application

$$\mathbb{Z}/6\mathbb{Z} \rightarrow \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}, \quad x \bmod 6 \mapsto (x \bmod 2, x \bmod 3).$$

1.3 Produit semi-direct de deux groupes

Après avoir traité du produit direct de deux groupes, nous nous intéressons à leur produit semi-direct. Nous verrons ses principales propriétés. Nous finirons en montrant que certains groupes bien connus peuvent s'exprimer comme un produit semi-direct de deux groupes.

1.3.1 Définition

Définition 1.7. Soient H et K deux groupes. Soit φ un morphisme de K dans $\text{Aut}(H)$ (l'ensemble des automorphismes de H). On définit le *produit semi-direct de H par K via φ* comme l'ensemble $H \times K$ muni de la loi interne

$$(h, k)(h', k') := (h\varphi(k)(h'), kk').$$

On le note $H \rtimes_{\varphi} K$.

Remarque. Si φ est triviale, c'est-à-dire que pour tout $k \in K$, $\varphi(k) = \text{Id}$, on retrouve le produit direct de H et K . Vérifions que nous obtenons bien un groupe.

Proposition 1.9. La loi définie ci-dessus est une loi de groupe.

Démonstration. La démonstration est fastidieuse, mais ne présente aucune difficulté.

- Associativité : puisque $\varphi(k)$ est un automorphisme et $\varphi(kk') = \varphi(k)\varphi(k')$,

$$\begin{aligned} (h, k)((h', k')(h'', k'')) &= (h\varphi(k)(h'\varphi(k')(h'')), kk'k'') \\ &= (h\varphi(k)(h')\varphi(kk')(h''), kk'k'') = ((h, k)(h', k'))(h'', k''). \end{aligned}$$

- Élément neutre : $(h, k)(1, 1) = (1, 1)(h, k) = (h, k)$.

- Inverse : en posant $a = \varphi(k^{-1})(h^{-1})$, on obtient

$$(h, k)(a, k^{-1}) = (h\varphi(k)(a), 1) = (1, 1), \quad (a, k^{-1})(h, k) = (a\varphi(k^{-1})(h), 1) = (1, 1).$$

Ainsi $(h, k)^{-1} = (\varphi(k^{-1})(h^{-1}), k^{-1})$.

□

1.3.2 Quelques propriétés

L'application

$$H \rightarrow H \rtimes_{\varphi} K, \quad h \mapsto (h, 1)$$

est un morphisme de groupes injectif car

$$(h, 1)(h', 1) = (h\varphi(1)(h'), 1) = (hh', 1).$$

On note $\overline{H}$ son image. Il s'agit d'un sous-groupe de $H \rtimes_{\varphi} K$.

De même, l'application

$$K \rightarrow H \rtimes_{\varphi} K, \quad k \mapsto (1, k)$$

est aussi un morphisme de groupes injectif. On note $\overline{K}$ son image. Il s'agit également d'un sous-groupe de $H \rtimes_{\varphi} K$.

Proposition 1.10. *On dispose des propriétés suivantes.*

(i). $\overline{H} \cap \overline{K} = \{(1, 1)\}.$

(ii). *la conjugaison permet de retrouver l'action de $\varphi(k)$ sur h :*

$$(1, k)(h, 1)(1, k^{-1}) = (1\varphi(k)(h), k)(1, k^{-1}) = (\varphi(k)(h)\varphi(k)(1), 1) = (\varphi(k)(h), 1).$$

(iii). $H \rtimes_{\varphi} K = \overline{H} \overline{K}$ car $(h, k) = (h, 1)(1, k)$ (attention : on ne peut pas commuter ces deux éléments).

(iv). $\overline{H}$ est distingué dans $H \rtimes_{\varphi} K$.

Nous allons maintenant nous intéresser au cas où un groupe est produit semi-direct de deux de ses sous-groupes.

Définition 1.8. Soit G un groupe. Soient H et K deux sous-groupes de G . On suppose H distingué dans G (ce que l'on note $H \triangleleft G$). On a alors un morphisme de groupes de G dans $\text{Aut}(H)$ induit par la conjugaison, que l'on peut restreindre à K . Plus précisément, on note

$$\varphi : K \rightarrow \text{Aut}(H), \quad g \mapsto (h \mapsto ghg^{-1}).$$

On dit que G est *produit semi-direct de H par K via φ* si et seulement si l'application

$$H \rtimes_{\varphi} K \rightarrow G, \quad (h, k) \mapsto hk$$

est un isomorphisme de groupes.

Ci-dessous figure une caractérisation très importante du produit semi-direct.

Proposition 1.11. *Soit G un groupe. Soit $H \triangleleft G$ et $K < G$. Alors G est produit semi-direct de H et K si et seulement si*

(i). $H \cap K = \{1\},$

(ii). $G = \langle H, K \rangle.$

Démonstration. On note

$$\pi : H \rtimes_{\varphi} K \rightarrow G, \quad (h, k) \mapsto hk.$$

- π est un morphisme si et seulement si φ est le morphisme de K dans $\text{Aut}(H)$ induit par la conjugaison que nous avons évoqué dans la définition 1.8. En effet,

$$\begin{aligned} \pi(h, k)\pi(h', k') &= \pi((h, k)(h', k')) \\ \iff hkh'k' &= h\varphi(k)(h')kk' \\ \iff kh'k^{-1} &= \varphi(k)(h') \end{aligned}$$

- On suppose que π est un morphisme. On a

$$\ker(\pi) = \{(h, k) : hk = 1\} = \{(h, h^{-1}), h \in H \cap K\},$$

donc π est injectif si et seulement si $H \cap K = \{1\}$.

- On suppose encore que π est un morphisme. π est surjectif si et seulement si $G = HK$, ce qui revient au fait que G est engendré par H et K , car K normalise H (i.e. pour tout $k \in K$, $kHk^{-1} = H$).

□

Remarques. Si H et K sont deux sous-groupes de G tels que

- (i). H est normalisé par K ,
- (ii). $H \cap K = \{1\}$.

Alors $\langle H, K \rangle$ est produit semi-direct de H et K . En particulier, si H et K sont finis, on a $|\langle H, K \rangle| = |H||K|$.

Si G est produit semi-direct de H par K , alors le quotient G/H est isomorphe à K .

$$\begin{array}{ccc} G & \xrightarrow{\sim} & H \rtimes_{\varphi} K \\ \downarrow & & \downarrow (h, k) \mapsto k \\ G/H & \xrightarrow{\sim} & K \end{array}$$

L'isomorphisme du haut envoie hk sur (h, k) . La composée est un morphisme surjectif de noyau H , d'où la factorisation. On peut relever G/H en un sous-groupe complémentaire de H dans G , à savoir K .

$$\begin{array}{ccccccc} 1 & \longrightarrow & H & \hookrightarrow & G & \twoheadrightarrow & G/H \longrightarrow 1 \\ & & & & \uparrow & \swarrow \sim & \\ & & & & K & & \end{array}$$

G est produit direct de H et K si et seulement si H et K sont distingués. En effet, si $G = H \times K$, H centralise K (i.e. chaque élément de H appartient au centralisateur de chaque élément de K) et K normalise K , donc G normalise K . Réciproquement, on suppose $H \triangleleft G$ et $K \triangleleft G$. On a donc

$$\underbrace{h \underbrace{(kh^{-1}k^{-1})}_{\in H \triangleleft G}}_{\in H} \text{ et } \underbrace{(hkh^{-1})}_{\in K \triangleleft G} k^{-1},$$

donc $hkh^{-1}k^{-1} \in H \cap K = \{1\}$. H centralise K , donc $G = H \times K$.

1.3.3 Exemples de produits semi-directs

Quelques groupes classiques peuvent être mis sous la forme d'un produit semi-direct. Pour un corps $\mathbb{K}$ et $n \geq 1$, $\text{GL}_n(\mathbb{K})$ est le groupe des matrices inversibles de taille n et $\text{SL}_n(\mathbb{K})$ son sous-groupe de déterminant 1. On note $\text{O}_n(\mathbb{R}) = \{M : {}^t M M = I_n\}$ le groupe orthogonal (où ${}^t M$ est la transposée et I_n l'identité), et $\text{SO}_n(\mathbb{R})$ son sous-groupe de déterminant 1.

Exemple. • le groupe des permutations, pour $n \geq 2$,

$$\mathfrak{S}_n = \mathfrak{A}_n \rtimes \langle (1\ 2) \rangle.$$

- le groupe linéaire

$$\mathrm{GL}_n(\mathbb{K}) = \mathrm{SL}_n(\mathbb{K}) \rtimes \left\{ \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & 1 & 0 \\ 0 & \cdots & 0 & \lambda \end{pmatrix}, \lambda \in \mathbb{K}^* \right\}.$$

- le groupe orthogonal

$$\mathrm{O}_n(\mathbb{R}) = \mathrm{SO}_n(\mathbb{R}) \rtimes \{1, \sigma\},$$

où σ est une réflexion

$$\sigma := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & 1 & 0 \\ 0 & \cdots & 0 & -1 \end{pmatrix}.$$

Définition 1.9. Pour $n \geq 3$, on appelle *n -ième groupe diédral* le groupe des isométries du plan qui conservent le polygone régulier à n sommets

$$\mathfrak{P}_n := \left\{ e^{i \frac{2k\pi}{n}}, k \in \{0, \dots, n-1\} \right\}.$$

Ce groupe sera noté D_{2n} .

Pour tout k , D_{2n} contient la rotation de centre 0 et d'angle $\frac{2k\pi}{n}$. Elle envoie le sommet 1 sur le sommet $e^{i \frac{2k\pi}{n}}$ et conserve $\mathfrak{P}_n$.

On pose $\mathrm{D}_{2n}^+ := \mathrm{D}_{2n} \cap \mathrm{SO}_2(\mathbb{R})$. C'est l'ensemble des rotations d'angle $\frac{2k\pi}{n}$ avec $k \in \{0, \dots, n-1\}$, car une rotation qui conserve $\mathfrak{P}_n$ envoie un sommet sur un autre et on peut montrer que donc elle est de centre 0 et d'angle $\frac{2k\pi}{n}$. Alors, si l'on note ρ la rotation de centre 0 et d'angle $\frac{2\pi}{n}$, on a $\mathrm{D}_{2n}^+ = \langle \rho \rangle$.

On dispose de la suite exacte suivante

$$1 \longrightarrow \mathrm{D}_{2n}^+ \hookrightarrow \mathrm{D}_{2n} \xrightarrow[\det]{s} \{-1, 1\} \longrightarrow 1$$

On exhibe une symétrie σ dans D_{2n} . Elle fournit la section s définie par $s(1) = \mathrm{Id}$ et $s(-1) = \sigma$. Alors

$$\mathrm{D}_{2n} = \{\mathrm{Id}, \rho, \dots, \rho^{n-1}\} \cup \{\sigma, \sigma\rho, \dots, \sigma\rho^{n-1}\}.$$

On pose $K := \{\mathrm{Id}, \sigma\}$. On a $\sigma\rho^i\sigma^{-1} = \rho^{-i}$ donc $\mathrm{D}_{2n}^+ \triangleleft \mathrm{D}_{2n}$. En outre, l'intersection de D_{2n}^+ et K est l'identité. On en déduit que $\langle \mathrm{D}_{2n}^+, K \rangle$ est produit semi-direct de D_{2n}^+ et K . Comme $\langle \mathrm{D}_{2n}^+, K \rangle$ est d'ordre $2n$, on en déduit que $\mathrm{D}_{2n} = \langle \mathrm{D}_{2n}^+, K \rangle$. Finalement, on a

$$\mathrm{D}_{2n} = \langle \rho \rangle \rtimes \{\mathrm{Id}, \sigma\}.$$

Remarque. On peut également définir le n -ième groupe diédral comme le produit semi-direct de $\mathbb{Z}/n\mathbb{Z}$ par $\mathbb{Z}/2\mathbb{Z}$ via l'application φ suivante

$$\varphi : \mathbb{Z}/2\mathbb{Z} \rightarrow \mathrm{Aut}(\mathbb{Z}/n\mathbb{Z}), \quad 0 \mapsto \mathrm{Id}, \quad 1 \mapsto (x \mapsto -x).$$

2.1 Rappels

2.1.1 Caractéristique d'un corps et sous-corps premier

Soit A un anneau. On a un morphisme d'anneaux

$$\varphi : \mathbb{Z} \rightarrow A$$

$$n \mapsto n \cdot 1 = \begin{cases} \underbrace{1 + \dots + 1}_{n \text{ termes}} & \text{si } n > 0 \\ 0 & \text{si } n = 0 \\ \underbrace{-1 - \dots - 1}_{-n \text{ termes}} & \text{si } n < 0. \end{cases}$$

L'image de φ est le plus petit sous-anneau de A contenant 1, i.e. le sous-anneau engendré par 1. Ce sous-anneau est isomorphe à $\mathbb{Z}/\ker \varphi$. Comme $\mathbb{Z}$ est principal, on a $\ker \varphi = n\mathbb{Z}$, avec $n \geq 0$.

Définition 2.1. L'unique entier $n \geq 0$ tel que $\ker \varphi = n\mathbb{Z}$ est appelé la *caractéristique* de A , notée $\text{Car}(A)$ (ou $\text{Car } A$). Ainsi $\text{Car}(A) = 0$ si aucun entier strictement positif n'annule 1_A ; sinon, c'est le plus petit entier strictement positif qui annule 1_A .

Dans le cas particulier où A est intègre (par exemple, A est un corps), $\mathbb{Z}/n\mathbb{Z}$ est isomorphe à un sous-anneau de A intègre donc $n = 0$ ou n premier.

Remarque. La caractéristique est une propriété "locale" autour de 1 de l'anneau : si A est un sous-anneau de B avec la même unité, alors A et B ont la même caractéristique.

Si $A = \mathbb{K}$ est un corps, sa caractéristique est 0 ou un nombre premier p .

- On suppose $\text{Car } \mathbb{K} = 0$. Le sous-anneau engendré par 1 est isomorphe à $\mathbb{Z}$ (ce n'est pas un corps).

$$\begin{array}{ccc} \mathbb{Z} & \xhookrightarrow{\varphi} & \mathbb{K} \\ \downarrow & \nearrow \hat{\varphi} & \\ \mathbb{Q} & & \end{array}$$

$\hat{\varphi}$ est défini par $\hat{\varphi}(\frac{a}{b}) = \varphi(a)\varphi(b)^{-1}$. On peut faire le rapprochement avec la définition du corps des fractions d'un anneau comme solution d'un problème universel : pour tout $A \xhookrightarrow{\varphi} \mathbb{L}$, il existe un prolongement de φ à un corps $\mathbb{K}$ (qui sera alors $\text{Fr } A$) qui soit un morphisme de corps.

$$\begin{array}{ccc} \mathbb{Z} & \xrightarrow{\sim} & \text{Im } \varphi \\ \downarrow & & \downarrow \\ \mathbb{Q} & \xrightarrow{\sim} & \text{Im } \hat{\varphi} \hookrightarrow \mathbb{K} \end{array}$$

$\text{Im } \hat{\varphi}$ est le plus petit sous-corps de $\mathbb{K}$: c'est le sous-corps premier de $\mathbb{K}$. Dans ce cas, il est isomorphe à $\mathbb{Q}$.

- On suppose $\text{Car } \mathbb{K} = p$. Le sous-anneau engendré par 1 est isomorphe à $\mathbb{Z}/p\mathbb{Z}$, c'est aussi le sous-corps premier de $\mathbb{K}$.

$$\begin{array}{ccc} \mathbb{Z} & \xrightarrow{\varphi} & \mathbb{K} \\ \downarrow & \nearrow & \\ \mathbb{Z}/p\mathbb{Z} & & \end{array}$$

Définition 2.2. Le plus petit sous-corps de $\mathbb{K}$ (celui engendré par 1) est appelé *sous-corps premier* de $\mathbb{K}$.

Remarques. Il est isomorphe à $\mathbb{Q}$ si $\text{Car } \mathbb{K} = 0$ et à $\mathbb{Z}/p\mathbb{Z}$ si $\text{Car } \mathbb{K} = p$.

Le sous-corps premier de $\mathbb{C}$ est $\mathbb{Q}$: on ne peut pas trouver de sous-corps fini de $\mathbb{C}$.

2.1.2 Séparabilité

Dans toute cette partie, $\mathbb{K}$ désigne un corps *commutatif*.

Définition 2.3. Soit P un polynôme irréductible sur $\mathbb{K}$. On dit que P est *séparable* si, dans un corps de décomposition, ses racines sont simples.

On considère l'opération de dérivation

$$D : \mathbb{K}[X] \rightarrow \mathbb{K}[X], \quad \sum_{i=0}^d a_i X^i \mapsto \sum_{i=1}^d i a_i X^{i-1}.$$

Lemme 2.1. Soit $P \in \mathbb{K}[X]$. Soit $a \in \mathbb{K}$. Alors $P = P(a) + DP(a)(X - a) + R(X) \cdot (X - a)^2$.

Démonstration. On écrit $P = b_0 + b_1(X - a) + \underbrace{b_2(X - a)^2 + \dots}_{R(X) \cdot (X - a)^2}$. En spécialisant en a , on obtient $b_0 = P(a)$. En dérivant et en spécialisant en a , on a $b_1 = DP(a)$. □

Corollaire 2.2. Si a est racine du polynôme P , alors a est racine simple si et seulement si $DP(a) \neq 0$.

Corollaire 2.3. Supposons P irréductible sur $\mathbb{K}$. P est non séparable si et seulement si $DP = 0$.

Démonstration. Dans un corps de décomposition de P , le polynôme minimal sur $\mathbb{K}$ de toute racine de P est égal à P à un facteur scalaire non nul près. Ainsi, P admet une racine multiple si et seulement si P et DP ont une racine commune, ce qui équivaut à $P \mid DP$. Comme $\deg DP < \deg P$ lorsque $DP \neq 0$, cette dernière condition équivaut à $DP = 0$. □

Deux cas peuvent alors se présenter

- On suppose $\text{Car } \mathbb{K} = 0$. On note $P = \sum_i a_i X^i$. $DP = 0$ si et seulement si pour tout $i \geq 1$, $a_i = 0$, c'est-à-dire si P est constant. On en déduit que *tous les polynômes irréductibles sont séparables*.
- On suppose $\text{Car } \mathbb{K} = p$. On note encore $P = \sum_i a_i X^i$. $DP = 0$ si et seulement si pour tout $i \geq 1$, $i a_i = 0$, c'est-à-dire si pour tout i premier avec p , $a_i = 0$, soit si $P = Q(X^p)$ avec $Q \in \mathbb{K}[X]$.

Lemme 2.4. On note $\mathbb{K}^p = \{x^p : x \in \mathbb{K}\}$. Si $\mathbb{K} = \mathbb{K}^p$, il existe Q tel que $P = Q(X^p)$ si et seulement s'il existe R tel que $P = R(X)^p$.

Démonstration. Si A est un anneau commutatif de caractéristique p , on a dans $A[X, Y]$, $(X + Y)^p = X^p + Y^p$. Donc

$$P(X) = R(X)^p = \left(\sum_i b_i X^i \right)^p = \sum_i b_i^p X^{ip} = Q(X^p),$$

et

$$P(X) = Q(X^p) = \sum_i a_i X^{ip} = \sum_i b_i^p X^{ip} = \left(\sum_i b_i X^i \right)^p = R(X)^p.$$

□

Corollaire 2.5. Si $\mathbb{K} = \mathbb{K}^p$, tout polynôme irréductible est séparable.

Démonstration. Si P est non séparable, on a $DP = 0$ donc il existe Q tel que $P(X) = Q(X^p)$, donc il existe R tel que $P(X) = R(X)^p$, donc P n'est pas irréductible. □

Définition 2.4. Un corps commutatif $\mathbb{K}$ est dit *parfait* si et seulement si soit $\text{Car } \mathbb{K} = 0$, soit $\text{Car } \mathbb{K} = p$ et $\mathbb{K} = \mathbb{K}^p$.

Si $\mathbb{K}$ est parfait, tous les polynômes irréductibles de $\mathbb{K}[X]$ sont séparables. On dit que *toutes les extensions algébriques de $\mathbb{K}$ sont séparables*.

Pour une extension finie $\mathbb{L}/\mathbb{K}$, on note $[\mathbb{L} : \mathbb{K}]$ sa dimension comme espace vectoriel sur $\mathbb{K}$. Pour finir, donnons un exemple d'extension non séparable. Il s'agit de

$$\mathbb{Z}/p\mathbb{Z}(T^p) \subset \mathbb{Z}/p\mathbb{Z}(T).$$

En effet, $P = X^p - T^p$ est irréductible sur $\mathbb{Z}/p\mathbb{Z}(T^p)$ et non séparable.

On note $D_{\mathbb{K}}(P)$ un corps de décomposition de P sur $\mathbb{K}$, c'est-à-dire une extension engendrée par toutes ses racines. Pour P irréductible, $R_{\mathbb{K}}(P)$ désigne un corps de rupture, isomorphe à $\mathbb{K}[X]/(P)$.

2.2 Constructions de corps finis

2.2.1 Structure d'espace vectoriel

Soit $\mathbb{K}$ un corps fini (*a priori* éventuellement non commutatif...). Soit $\mathbb{K}_0$ son sous-corps premier. Comme $\mathbb{K}_0$ est fini, il existe p premier tel que $\mathbb{K}_0 \simeq \mathbb{Z}/p\mathbb{Z}$. On a aussi $\text{Car}(\mathbb{K}) = p$.

On munit $\mathbb{K}$ d'une structure de $\mathbb{K}_0$ -espace vectoriel

$$\mathbb{K}_0 \times \mathbb{K} \rightarrow \mathbb{K}, \quad (\lambda, x) \mapsto \lambda x.$$

Comme $\mathbb{K}$ est fini, c'est un espace vectoriel de dimension finie. On note $n = [\mathbb{K} : \mathbb{K}_0]$. Alors $\mathbb{K}$ est de cardinal p^n .

2.2.2 Théorème de Wedderburn

Théorème 2.6 (Wedderburn). Tout corps fini est commutatif.

Démonstration. Soit $\mathbb{K}$ un corps fini. On note $Z(\mathbb{K})$ son centre. C'est un sous-corps commutatif de $\mathbb{K}$. On note $q = |Z(\mathbb{K})|$. On peut munir $\mathbb{K}$ d'une structure de $Z(\mathbb{K})$ -espace vectoriel

$$Z(\mathbb{K}) \times \mathbb{K} \rightarrow \mathbb{K}, \quad (\lambda, x) \mapsto \lambda x.$$

Comme $\mathbb{K}$ est fini, c'est un $Z(\mathbb{K})$ -espace vectoriel de dimension finie, notée n . Alors $|\mathbb{K}| = q^n$.

Le groupe $\mathbb{K}^*$ opère sur $\mathbb{K}$ par automorphisme intérieur

$$\mathbb{K}^* \times \mathbb{K} \rightarrow \mathbb{K}, \quad (x, y) \mapsto xyx^{-1}.$$

Si $y \in \mathbb{K}$, son orbite est sa classe de conjugaison $\{xyx^{-1}, x \in \mathbb{K}^*\}$ et son stabilisateur est $\{x \in \mathbb{K}^* : xy = yx\}$.

On note $\mathbb{K}_y = \{x \in \mathbb{K} : xy = yx\}$. C'est un sous-corps de $\mathbb{K}$ qui contient $Z(\mathbb{K})$, donc qui peut être vu comme un $Z(\mathbb{K})$ -espace vectoriel de dimension finie. On en déduit l'existence de n_y tel que $|\mathbb{K}_y| = q^{n_y}$.

Le stabilisateur de y est $\mathbb{K}_y^*$.

Lemme 2.7. $n_y | n$.

Démonstration. Nous proposerons deux démonstrations pour ce lemme. La première est la suivante. On munit $\mathbb{K}$ d'une structure de $\mathbb{K}_y$ -espace vectoriel à gauche

$$\mathbb{K}_y \times \mathbb{K} \rightarrow \mathbb{K}, \quad (\lambda, x) \mapsto \lambda x.$$

Comme $\mathbb{K}$ est fini, il est de dimension finie d_y sur $\mathbb{K}_y$. On a donc $q^n = |\mathbb{K}| = |\mathbb{K}_y|^{d_y} = q^{n_y d_y}$, puis $n = n_y d_y$.

Ci-après figure la seconde démonstration du lemme. $\mathbb{K}_y^*$ est un sous-groupe de $\mathbb{K}^*$, d'après le théorème de Lagrange, on a $q^{n_y} - 1 | q^n - 1$. Soient s, r tels que $n = sn_y + r$ et $0 \leq r < n_y$. On a

$$q^n - 1 = q^{sn_y + r} - 1 = q^r(q^{sn_y} - 1) + q^r - 1 = q^r(q^{n_y} - 1)(q^{(s-1)n_y} + \dots + 1) + q^r - 1 = B(q^{n_y} - 1) + q^r - 1.$$

On a $0 \leq r < n_y$. Donc $0 \leq q^r - 1 < q^{n_y} - 1$. Or $q^{n_y} - 1 | q^n - 1$. Donc $q^r - 1 = 0$, soit $r = 0$ (car $q \geq 2$). $\square$

Reprenons la démonstration du théorème. On écrit l'équation aux classes

$$|\mathbb{K}| = \sum_{y \in \mathbb{K} \setminus \mathbb{K}^*} |\omega_y| = \sum_{y \in \mathbb{K} \setminus \mathbb{K}^*} \frac{|\mathbb{K}^*|}{|\text{Stab}_{\mathbb{K}^*} y|},$$

c'est-à-dire

$$q^n = \sum_{y \in \mathbb{K} \setminus \mathbb{K}^*} \frac{q^n - 1}{q^{n_y} - 1}.$$

Or $|\omega_y| = 1$ si et seulement si $y \in Z(\mathbb{K})$ si et seulement si $n_y = n$. On peut donc écrire

$$q^n = \underbrace{q}_{\sum_{y \in Z(\mathbb{K})} 1} + \sum_{\substack{y \in \mathbb{K} \setminus \mathbb{K}^* \\ y \notin Z(\mathbb{K})}} \frac{q^n - 1}{q^{n_y} - 1}.$$

On note $\Phi_d(X) \in \mathbb{Z}[X]$ le d -ième polynôme cyclotomique (il s'agit du produit de $X - \xi$, où ξ décrit les racines d -ièmes primitives de l'unité). Si n_y divise strictement n , on a

$$q^n - 1 = \prod_{d|n} \Phi_d(q) = \prod_{d|n_y} \Phi_d(q) \prod_{\substack{d|n \\ d \nmid n_y}} \Phi_d(q) = (q^{n_y} - 1) \prod_{\substack{d|n \\ d \nmid n_y}} \Phi_d(q).$$

On en déduit que $\Phi_n(q)$ divise $\frac{q^n - 1}{q^{n_y} - 1}$. En s'aidant de l'équation aux classes, on a directement $\Phi_n(q)$ divise $q - 1$.

On remarque que pour tout $\alpha \in \mathbb{C}$ différent de 1 tel que $|\alpha| = 1$, on a $|q - \alpha| > |q - 1|$ (un dessin permet de se convaincre de cette propriété). Donc, si $n > 1$, on a $|\Phi_n(q)| > q - 1$. D'où $n = 1$, i.e. $\mathbb{K} = Z(\mathbb{K})$. $\square$

2.2.3 Le corps fini à q éléments $\mathbb{F}_q$

Théorème 2.8. Soit p un nombre premier. Soit $q = p^r$, avec $r \geq 1$. Il existe un corps de cardinal q unique à isomorphisme près. Ce corps est le corps de décomposition du polynôme $X^q - X$ sur son sous-corps premier (lequel est isomorphe à $\mathbb{Z}/p\mathbb{Z}$). C'est aussi l'ensemble des racines de $X^q - X$ dans tout corps qui le contient.

“Le” corps à q éléments est noté $\mathbb{F}_q$. En particulier, $\mathbb{Z}/p\mathbb{Z}$ en tant que corps est noté $\mathbb{F}_p$ (la notation $\mathbb{Z}/p\mathbb{Z}$ étant plutôt réservée à l'anneau).

On note $\mathbb{K}$ un corps commutatif de caractéristique p .

Lemme 2.9. Le polynôme $X^q - X$ n'a que des racines simples dans $\mathbb{K}$.

Démonstration. Posons $P = X^q - X$. On a $DP = qX^{q-1} - 1 = -1$. Aucune racine de P ne peut être racine de DP donc les racines de P sont simples. $\square$

Lemme 2.10. L'application

$$\sigma : \mathbb{K} \rightarrow \mathbb{K}, \quad x \mapsto x^q$$

est un morphisme de corps injectif appelé morphisme de Frobenius.

Démonstration. D'abord, $\sigma(xy) = \sigma(x)\sigma(y)$ car $\mathbb{K}$ est supposé commutatif. Ensuite, $\sigma(x + y) = (x + y)^{p^r} = (x^p + y^p)^{p^{r-1}} = \dots = x^{p^r} + y^{p^r} = \sigma(x) + \sigma(y)$. Enfin, σ est injectif car non nul ($\sigma(1) = 1$). $\square$

Corollaire 2.11. $\{x \in \mathbb{K} : x^q = x\}$ est un sous-corps de $\mathbb{K}$. C'est l'ensemble des racines de $X^q - X$ dans $\mathbb{K}$.

Démonstration. (théorème 2.8) Démontrons l'existence. Soit $\mathbb{K} = D_{\mathbb{F}_p}(X^q - X)$ “le” corps de décomposition de $X^q - X$ sur $\mathbb{F}_p$. Soit $S = \{x \in \mathbb{K} : x^q = x\}$. Le polynôme $X^q - X$ est scindé sur $\mathbb{K}$ et a toutes ses racines simples, d'après le lemme 2.9. Donc $|S| = q$. D'après le corollaire 2.11, S est un sous-corps de $\mathbb{K}$. C'est donc un corps de cardinal q .

On remarque que, $\mathbb{K}$ étant le plus petit sous-corps de $\mathbb{K}$ contenant les racines de $X^q - X$, c'est-à-dire contenant S , on a en fait $\mathbb{K} = S$.

Intéressons-nous à l'unicité. Soit $\mathbb{K}$ un corps à q éléments. Soit $\mathbb{K}_0$ son sous-corps premier. On a $|\mathbb{K}_0| = p$. Soit $x \in \mathbb{K}$. On a $x = 0$ ou $x^{q-1} = 1$, d'après le théorème de Lagrange. Donc $x^q = x$, puis $\mathbb{K} \subset \{x \in \mathbb{K} : x^q = x\}$. D'où l'égalité de ces deux derniers ensembles.

Le polynôme $X^q - X$ a au plus q racines et tous les éléments de $\mathbb{K}$ (il y en a q) sont racines. Donc il a exactement ses racines dans $\mathbb{K}$, i.e. il est scindé sur $\mathbb{K}$. Le plus petit sous-corps de $\mathbb{K}$ contenant les racines de $X^q - X$ est lui-même (tout élément de $\mathbb{K}$ est un de ses racines).

Donc $\mathbb{K}$ est corps de décomposition de $X^q - X$ sur $\mathbb{K}_0$. On en déduit l'unicité de $\mathbb{K}$ à isomorphisme près.

Soit maintenant $\mathbb{K}$ un corps à q éléments et $\mathbb{L}$ un sur-corps de $\mathbb{K}$. On a

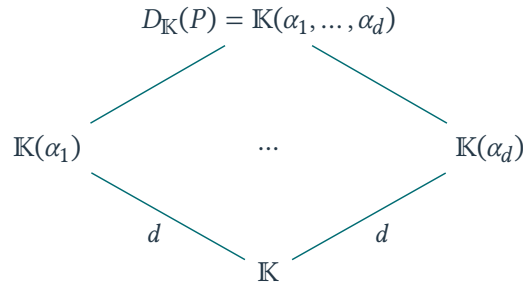
$$\mathbb{K} \subset \{x \in \mathbb{L} : x^q = x\} \subset \mathbb{L}.$$

Mais $X^q - X$ a au plus q racines. Donc $\mathbb{K} = \{x \in \mathbb{L} : x^q = x\}$. Ceci implique que $\mathbb{K}$ est l'unique sous-corps à q éléments de $\mathbb{L}$. $\square$

Donnons une application classique de ce résultat.

Proposition 2.12. Soit P un polynôme irréductible sur un corps fini $\mathbb{K}$. On note $\mathbb{L}$ un corps de rupture de P . Alors $\mathbb{L}$ est un corps de décomposition de P (i.e. P est scindé sur $\mathbb{L}$).

Démonstration. Soit M un corps de décomposition de P sur K . Posons $q = |K|$ et $d = \deg P$. Les corps finis sont parfaits, car le morphisme $x \mapsto x^p$ y est injectif, donc surjectif. Soient $\alpha_1, \dots, \alpha_d$ les racines distinctes de P . Chacun des sous-corps $K(\alpha_i)$ de M a pour cardinal q^d . Or un corps contient au plus un sous-corps de ce cardinal, puisque celui-ci est nécessairement l'ensemble des racines de $X^{q^d} - X$. Donc les corps de rupture $K(\alpha_i)$, qui sont de cardinal q^d , sont tous égaux : $K(\alpha_1) = \dots = K(\alpha_d)$. On en déduit que $M = K(\alpha_1)$.



□

On dit alors que *toutes les extensions finies des corps finis sont normales*. Ces extensions sont aussi séparables : si K est fini de caractéristique p , le morphisme de K qui à x associe x^p est bijectif, donc $K = K^p$ et K est parfait. Alors, *toutes les extensions finies des corps finis sont séparables*.

Exemple. Traitons le cas de $\mathbb{F}_4$. Ses éléments sont $0, 1, x$ et x^2 , où x est une racine de $X^2 + X + 1$. En effet, on a $X^4 - X = X(X-1)(X^2 + X + 1)$, c'est-à-dire, puisque nous travaillons en caractéristique 2, $X^4 + X = X(X+1)(X^2 + X + 1)$. La table d'addition est la suivante

0	1	x	x^2
1	0	x^2	x
x	x^2	0	1
x^2	x	1	0

La table de multiplication est

1	x	x^2
x	x^2	1
x^2	1	x

2.3 Structure des corps finis

2.3.1 Structure du groupe multiplicatif

Théorème 2.13. $\mathbb{F}_q^*$ est cyclique.

On obtient ce résultat comme corollaire du théorème suivant.

Théorème 2.14. Soit K un corps commutatif. Soit G un sous-groupe fini de K^* . Alors G est cyclique.

Démonstration. Soit G un sous-groupe fini de K^* . On pose $n = |G|$. Si d est un diviseur de n , on note G_d le sous-ensemble de G formé des éléments d'ordre d . On a

$$G = \bigsqcup_{d|n} G_d.$$

Lemme 2.15. $|G_d| \leq \varphi(d)$ (où $\varphi(d) = |\{1 \leq a \leq d : a \wedge d = 1\}|$ est l'indicatrice d'Euler).

Démonstration. Plus précisément, nous allons démontrer que soit G_d est vide, soit $|G_d| = \varphi(d)$. Supposons G_d non vide et fixons $g_0 \in G_d$. Soit $g \in G_d$. $\langle g \rangle$ est de cardinal d . $\langle g \rangle$ est aussi inclus dans l'ensemble des racines du polynôme $X^d - 1$ dans $\mathbb{K}$ qui est de cardinal au plus d . D'où l'égalité de $\langle g \rangle$ et de cet ensemble. g_0 est une racine de $X^d - 1$. Donc $g_0 \in \langle g \rangle$, puis $\langle g_0 \rangle \subset \langle g \rangle$. Comme g_0 est d'ordre d , on en déduit que $\langle g \rangle = \langle g_0 \rangle$. g est donc un générateur du groupe cyclique $\langle g_0 \rangle$ qui a pour ordre d . On en déduit que $|G_d| = \varphi(d)$. $\square$

Dans le groupe cyclique $\mathbb{Z}/n\mathbb{Z}$, comptons directement les éléments selon leur ordre. On sait que pour tout d diviseur de n , il existe un élément d'ordre d (il s'agit de $\frac{n}{d}$). Alors, pour tout $d|n$, on a $\varphi(d)$ éléments d'ordre d . On en déduit que

$$n = |\mathbb{Z}/n\mathbb{Z}| = \sum_{d|n} |(\mathbb{Z}/n\mathbb{Z})_d| = \sum_{d|n} \varphi(d).$$

Revenons à G . On peut écrire

$$n = |G| = \sum_{d|n} |G_d| \leq \sum_{d|n} \varphi(d) = n.$$

Donc toutes les inégalités sont des égalités. En particulier, on a $|G_n| = \varphi(n)$ et donc G_n n'est pas vide. $\square$

Nous en déduisons de manière triviale un résultat à mettre en relation avec le théorème de l'élément primitif.

Corollaire 2.16. *Toute extension finie d'un corps fini est monogène.*

Démonstration. Soient $\mathbb{K} \subset \mathbb{L}$, avec $\mathbb{L}$ fini. On note ξ un générateur de $\mathbb{L}^*$. Alors, on a $\mathbb{L} = \mathbb{K}(\xi)$. $\square$

2.3.2 Sous-corps d'un corps fini

Soient $\mathbb{K} \subset \mathbb{M}$ deux corps finis. On cherche à décrire les sous-corps de $\mathbb{M}$ contenant $\mathbb{K}$. Dans le cas particulier où $\mathbb{K}$ est le sous-corps premier de $\mathbb{M}$, on obtient les sous-corps de $\mathbb{M}$.

On pose $q = |\mathbb{K}|$. En notant $n = [\mathbb{M} : \mathbb{K}]$, on a $|\mathbb{M}| = q^n$.

Proposition 2.17. *Soit $\mathbb{L}$ un sous-corps de $\mathbb{M}$ contenant $\mathbb{K}$. Alors*

- (i). $\mathbb{L}$ est de cardinal q^d , avec $d|n$,
- (ii). $\mathbb{L}$ est l'ensemble des racines du polynôme $X^{q^d} - X$ dans $\mathbb{M}$.

Démonstration. La multiplicativité du degré s'écrit ici :

$$\underbrace{[\mathbb{M} : \mathbb{K}]}_n = [\mathbb{M} : \mathbb{L}] \cdot \underbrace{[\mathbb{L} : \mathbb{K}]}_d.$$

La seconde partie de la proposition se déduit du fait que $\mathbb{L}$ est un sous-corps de $\mathbb{M}$ de cardinal q^d et du théorème 2.8. $\square$

Lemme 2.18. *Si $d|n$, alors $X^{q^d} - X | X^{q^n} - X$.*

Démonstration. $d|n$ donc $n = \ell d$. On a $q^n - 1 = q^{\ell d} - 1 = (q^d - 1)(q^{(\ell-1)d} + \dots + 1)$, donc $q^d - 1 | q^n - 1$. En utilisant la même identité remarquable, on montre que $X^{q^{d-1}} - 1 | X^{q^{n-1}} - 1$. En multipliant par X , on a le résultat voulu. $\square$

Proposition 2.19. Soit $d|n$. Alors $\mathbb{M}$ contient un sous-corps de cardinal q^d contenant $\mathbb{K}$.

Démonstration. On note S l'ensemble des racines de $X^{q^d} - X$ dans $\mathbb{M}$. $X^{q^d} - X | X^{q^n} - X$ et $X^{q^n} - X$ est scindé à racines simples sur $\mathbb{M}$, donc $X^{q^d} - X$ l'est aussi. On en déduit que $|S| = q^d$. En outre, on sait que S est un sous-corps de $\mathbb{M}$ (voir le corollaire 2.11). Enfin, si $x^q = x$, $x^{q^d} = x$, donc S est un sur-corps de $\mathbb{K}$. $\square$

Récapitulons ces résultats dans le théorème suivant.

Théorème 2.20. L'application

$$\begin{aligned} \{\mathbb{L} \text{ corps} : \mathbb{K} \subset \mathbb{L} \subset \mathbb{M}\} &\rightarrow \{d \in \mathbb{N} : d|n\} \\ \mathbb{L} &\mapsto [\mathbb{L} : \mathbb{K}] \end{aligned}$$

est une bijection croissante d'ensembles partiellement ordonnés (le premier par l'inclusion, le second par la divisibilité). Sa réciproque est

$$\begin{aligned} \{d \in \mathbb{N} : d|n\} &\rightarrow \{\mathbb{L} \text{ corps} : \mathbb{K} \subset \mathbb{L} \subset \mathbb{M}\} \\ d &\mapsto \{\text{racines de } X^{q^d} - X\}. \end{aligned}$$

On peut écrire

$$\mathbb{K} \subset \mathbb{L} \subset \mathbb{L}' \subset \mathbb{M} \iff [\mathbb{L} : \mathbb{K}] | [\mathbb{L}' : \mathbb{K}].$$

2.3.3 Polynômes irréductibles sur un corps fini

Soit $\mathbb{K}$ un corps fini de cardinal q .

Lemme 2.21. Soit $P \in \mathbb{K}[X]$ irréductible. On note d le degré de P . Alors $P | X^{q^d} - X$.

Démonstration. Un corps de rupture de P est de degré d sur $\mathbb{K}$, donc de cardinal q^d . Soit $\mathbb{L}$ un tel corps. On a $\mathbb{L} = \mathbb{K}(\alpha)$, où α est une racine de P . En outre, $|\mathbb{L}| = q^d$. Donc α est racine de $X^{q^d} - X$. Or P est égal au polynôme minimal de α sur $\mathbb{K}$ à un facteur scalaire non nul près. Donc $P | X^{q^d} - X$ (on retrouve que P est scindé sur $\mathbb{L}$). $\square$

Remarque. $P | X^{q^n} - X$ pour tout n multiple de d .

Proposition 2.22. Soit P un diviseur irréductible de $X^{q^n} - X$ dans $\mathbb{K}[X]$. Alors le degré de P divise n .

Démonstration. Comme $P | X^{q^n} - X$, $\mathbb{F}_{q^n}$ contient un corps de rupture de P de degré $\deg P$ sur $\mathbb{K}$. Donc $\deg P | n$.

$$\left. \begin{array}{c} D_{\mathbb{K}}(X^{q^n} - X) \\ | \\ D_{\mathbb{K}}(P) = \mathbb{K}(\alpha) \\ | \\ d \\ | \\ \mathbb{K} \end{array} \right\} n$$

$\square$

Théorème 2.23. Soit $n \in \mathbb{N} \setminus \{0\}$. On note $\mathcal{J}_d(q)$ l'ensemble des polynômes de $\mathbb{F}_q[X]$ irréductibles unitaires et de degré d . Alors "la" décomposition en produit d'irréductibles de $X^{q^n} - X$ dans $\mathbb{F}_q[X]$ est

$$X^{q^n} - X = \prod_{d|n} \prod_{P \in \mathcal{J}_d(q)} P.$$

Démonstration. Nous savons déjà que :

- $\forall d|n, \forall P \in \mathcal{J}_d(q), P | X^{q^n} - X$,
- si P est un diviseur irréductible de $X^{q^n} - X$ supposé unitaire, alors il existe $d|n$ tel que $P \in \mathcal{J}_d(q)$.

Donc

$$X^{q^n} - X = \prod_{d|n} \prod_{P \in \mathcal{J}_d(q)} P^{m_P},$$

avec $m_P \geq 1$. Or $X^{q^n} - X$ n'a que des racines simples dans une extension suffisamment grande, donc ne peut être divisible par le carré d'un irréductible. Donc $m_P = 1$. $\square$

Remarque. Cette décomposition de $X^{q^n} - X$ permet de classer les éléments de $\mathbb{F}_{q^n}$ de la manière suivante.

$$\underbrace{X^{q^n} - X}_{\text{éléments de } \mathbb{F}_{q^n}} = \prod_{d|n} \underbrace{\prod_{P \in \mathcal{J}_d(q)} P}_{\substack{\text{éléments de } \mathbb{F}_{q^d} \text{ qui ne sont} \\ \text{pas dans } \mathbb{F}_{q^e} \text{ pour } e|d, e \neq d}}.$$

Par exemple, pour $p = 2$ et $n = 2$,

$$\underbrace{X^4 + X}_{\mathbb{F}_4} = \underbrace{X(X+1)}_{\mathbb{F}_2} \underbrace{(X^2 + X + 1)}_{\mathbb{F}_4 \setminus \mathbb{F}_2}.$$

2.4 Prolongements

2.4.1 Représentation d'un corps fini comme quotient d'un anneau de polynômes

Soit P un polynôme de degré n irréductible sur $\mathbb{F}_q$. Alors $\mathbb{F}_q[X]/(P)$ est de cardinal q^n .

Soit un entier $n \geq 1$. On choisit un polynôme P irréductible de degré n sur $\mathbb{F}_p$. Alors $\mathbb{F}_p[X]/(P)$ est un corps à p^n éléments.

Il existe un tel polynôme irréductible de degré n . En effet, si ξ est un générateur de $\mathbb{F}_{p^n}^*$, le corps engendré par ξ sur $\mathbb{F}_p$ ne peut être que $\mathbb{F}_{p^n}$, donc son polynôme minimal est de degré n .

Il peut exister plusieurs tels polynômes : cette représentation de $\mathbb{F}_{p^n}$ n'est en général pas unique.

Exemple. Pour commencer, considérons l'extension $\mathbb{F}_2 \subset \mathbb{F}_8$. On a

$$\begin{aligned} X^8 + X &= X(X+1)(X^6 + X^5 + X^4 + X^3 + X^2 + X + 1) \\ &= X(X+1) \underbrace{(X^3 + X + 1)(X^3 + X^2 + 1)}_{\substack{2 \text{ irréductibles de degré } 3}}. \end{aligned}$$

Donc $\mathbb{F}_8 \simeq \mathbb{F}_2[X]/(X^3 + X + 1)$ ou encore $\mathbb{F}_8 \simeq \mathbb{F}_2[X]/(X^3 + X^2 + 1)$.

Continuons avec l'extension $\mathbb{F}_2 \subset \mathbb{F}_{16}$. On a

$$\begin{aligned} X^{16} + X &= \underbrace{X(X+1)(X^2 + X + 1)}_{\substack{\text{le seul irréductible} \\ \text{de degré } 2}} \underbrace{(X^4 + X^3 + X^2 + X + 1)(X^4 + X^3 + 1)(X^4 + X + 1)}_{\substack{3 \text{ irréductibles de degré } 4}}. \end{aligned}$$

En effet, si $P = X^4 + aX^3 + bX^2 + cX + d$ est irréductible, on a $1 = P(0) = d$ et $1 = P(1) = a + b + c$. Les possibilités pour P sont donc $X^4 + X^3 + X^2 + X + 1$, $X^4 + X^3 + 1$, $X^4 + X + 1$, mais pas $(X^2 + X + 1)^2$ qui n'est pas irréductible. On a finalement par exemple $\mathbb{F}_{16} \simeq \mathbb{F}_2[X]/(X^4 + X^3 + X^2 + X + 1)$.

Dans la décomposition précédente de $X^{16} + X$, on peut également lire des informations sur le groupe multiplicatif $\mathbb{F}_{16}^*$.

$$X^{16} + X = \underbrace{X(X+1)}_1 \underbrace{(X^2+X+1)}_{\substack{2 \text{ éléments d'ordre } 3 \\ (X^2+X+1|X^3+1)}} \underbrace{(X^4+X^3+X^2+X+1)}_{\substack{4 \text{ éléments d'ordre } 5 \\ (X^4+X^3+X^2+X+1|X^5+1)}} \\ \underbrace{(X^4+X^3+1)(X^4+X+1)}_{\varphi(15)=8 \text{ éléments d'ordre } 15}.$$

2.4.2 Critères d'irréductibilité de polynômes

Soit $P \in \mathbb{F}_q[X]$. On se pose la question de savoir si P est irréductible.

On suppose $n = \deg P \geq 1$. P irréductible sur $\mathbb{F}_q$ si et seulement si P n'a pas de racine dans une extension de $\mathbb{F}_q$ de degré $< n$ et même $\leq \frac{n}{2}$ (si l'on s'intéresse au plus petit diviseur).

Remarques. Soient $P, Q \in \mathbb{K}[X]$ et une extension $\mathbb{K} \subset \mathbb{L}$. On a $P \wedge_{\mathbb{K}} Q = P \wedge_{\mathbb{L}} Q$.

P a une racine dans $\mathbb{F}_{q^d}$ si et seulement si P a une racine commune avec $X^{q^d} - X$ si et seulement si $P \wedge (X^{q^d} - X) \neq 1$.

On en déduit le critère suivant :

$$P \text{ irréductible} \iff P \wedge (X^{q^d} - X) = 1 \quad \text{pour tout } 1 \leq d \leq \lfloor n/2 \rfloor.$$

Les informations que nous venons de donner quant à l'irréductibilité d'un polynôme de $\mathbb{F}_q[X]$ permettent de s'intéresser au problème de l'irréductibilité d'un polynôme de $\mathbb{Z}[X]$.

Pour étudier ce problème, on dispose d'abord du critère d'Eisenstein.

En outre, on peut utiliser le critère de réduction présenté ci-après. On remarque que si $P = QR$ dans $\mathbb{Z}[X]$, alors $\bar{P} = \bar{Q}\bar{R}$ dans $\mathbb{F}_p[X]$. Si P est primitif et si sa réduction modulo p conserve son degré et est irréductible, alors P est irréductible dans $\mathbb{Z}[X]$ (et dans $\mathbb{Q}[X]$). (On peut être amené à conduire des raisonnements plus compliqués, mais qui utilisent le même argument : imaginons par exemple que P soit primitif de degré 5, que ses réductions modulo 2 et modulo 3 conservent ce degré, que modulo 2, P soit produit d'un irréductible de degré 1 et d'un irréductible de degré 4, et que modulo 3, P soit produit d'un irréductible de degré 2 et d'un irréductible de degré 3, alors P est irréductible).

Il est à noter que ce critère ne fonctionne pas toujours. Par exemple, le polynôme $X^4 + 1$ est irréductible dans $\mathbb{Z}[X]$ (c'est un polynôme cyclotomique), alors qu'il est réductible modulo p pour tout nombre premier p .

2.4.3 Algorithme de Berlekamp

C'est un algorithme qui permet de décomposer un polynôme $P \in \mathbb{F}_q[X]$ en produit d'irréductibles (voir [7]).

On commence par une décomposition sans facteurs carrés, en conservant les multiplicités. En caractéristique p , le quotient $P/(P \wedge P')$ ne suffit pas toujours : par exemple, pour $P = X^p$, ce quotient vaut 1. Les facteurs de multiplicité divisible par p sont traités en extrayant une racine p -ième, possible coefficient par coefficient puisque $\mathbb{F}_q$ est parfait, puis en procédant récursivement. On applique ensuite Berlekamp à chaque facteur unitaire sans facteur carré. On suppose donc ici $P = P_1 \dots P_k$, les P_i étant unitaires, irréductibles et deux à deux distincts. D'après le lemme chinois, il existe un isomorphisme d'anneaux

$$\varphi : \mathbb{F}_q[X]/(P) \rightarrow \mathbb{F}_q[X]/(P_1) \times \dots \times \mathbb{F}_q[X]/(P_k).$$

Considérons $H = \{x \in \mathbb{F}_q[X]/(P) : x^q = x\}$. Alors on a

$$\varphi(H) = \left\{ (x_1, \dots, x_k) \in \prod_{i=1}^k \mathbb{F}_q[X]/(P_i) : \forall i, x_i^q = x_i \right\} \simeq \prod_{i=1}^k \underbrace{\{x_i \in \mathbb{F}_q[X]/(P_i) : x_i^q = x_i\}}_{\text{sous-corps à } q \text{ éléments de } \mathbb{F}_q[X]/(P_i)}.$$

Donc H est en bijection avec $(\mathbb{F}_q)^k$.

Lemme 2.24. Soit

$$\eta : \mathbb{F}_q[X]/(P) \rightarrow \mathbb{F}_q[X]/(P), \quad x \mapsto x^q - x.$$

η est un endomorphisme du $\mathbb{F}_q$ -espace vectoriel $\mathbb{F}_q[X]/(P)$ de noyau H .

Démonstration. $\eta(x + y) = (x + y)^q - (x + y) = (x^q - x) + (y^q - y) = \eta(x) + \eta(y)$ et $\eta(\lambda x) = (\lambda x)^q - \lambda x = \lambda^q x^q - \lambda x \stackrel{\text{car } \lambda \in \mathbb{F}_q}{=} \lambda(x^q - x) = \lambda \eta(x)$. $\square$

Corollaire 2.25. H est un sous-espace vectoriel de dimension k .

On peut ainsi calculer k : on commence par calculer la matrice de η sur la base $(1, X, \dots, X^{\deg P - 1})$. On calcule le rang de cette matrice et on en déduit la dimension de H , i.e. la valeur de k .

On suppose $k > 1$ (si $k = 1$, P est irréductible). On peut déterminer $h \in H \setminus \mathbb{F}_q \cdot 1$ (i.e. h non constant). Soit $\hat{h}$ relevant h dans $\mathbb{F}_q[X]$, avec $\deg \hat{h} < \deg P$. On a $\eta(h) = 0$ si et seulement si $P | \hat{h}^q - \hat{h} = \prod_{\lambda \in \mathbb{F}_q} (\hat{h} - \lambda)$. On en déduit que

$$P = \prod_{\lambda \in \mathbb{F}_q} (P \wedge (\hat{h} - \lambda)).$$

On remarque que si $\lambda \neq \mu$, $(\hat{h} - \lambda) - (\hat{h} - \mu) = \mu - \lambda \in \mathbb{F}_q^*$, donc $(\hat{h} - \lambda) \wedge (\hat{h} - \mu) = 1$. En outre, $\deg(P \wedge (\hat{h} - \lambda)) < \deg P$. On obtient donc une vraie factorisation de P . On réitère le processus et on cherche les multiplicités si nécessaire.

2.4.4 Algèbres de polynômes sur un corps fini

Soit $\mathbb{K}$ un corps et soit $n \geq 1$. On note $\mathcal{F}(\mathbb{K}^n, \mathbb{K})$ l'ensemble des applications de $\mathbb{K}^n$ dans $\mathbb{K}$. C'est une $\mathbb{K}$ -algèbre. On a un morphisme de $\mathbb{K}$ -algèbres

$$\begin{aligned} \mathbb{K}[X_1, \dots, X_n] &\rightarrow \mathcal{F}(\mathbb{K}^n, \mathbb{K}) \\ P &\mapsto (\tilde{P} : (x_1, \dots, x_n) \mapsto P(x_1, \dots, x_n)). \end{aligned}$$

Proposition 2.26. Si $\mathbb{K}$ est infini, ce morphisme est injectif.

Démonstration. Le résultat se démontre par récurrence sur n . Supposons $\tilde{P} = 0$. On note $P = \sum_i Q_i(X_1, \dots, X_{n-1})X_n^i$. Pour tous $a_1, \dots, a_{n-1}$, $\sum_i Q_i(a_1, \dots, a_{n-1})X_n^i$ est nul sur $\mathbb{K}$. Le seul polynôme ayant une infinité de racines est 0. Donc le polynôme est nul, i.e. pour tout i , $Q_i(a_1, \dots, a_{n-1}) = 0$. On applique alors l'hypothèse de récurrence à chaque Q_i : $\tilde{Q}_i = 0 \Rightarrow Q_i = 0$. Donc $P = 0$. $\square$

Remarque. Le morphisme n'est pas surjectif : pour s'en convaincre, il suffit de considérer une fonction nulle partout sauf en un point.

Proposition 2.27. Si $\mathbb{K}$ est un corps fini de cardinal q , ce morphisme n'est pas injectif. Son noyau est l'idéal $(X_1^q - X_1, \dots, X_n^q - X_n)$. En revanche, il est surjectif.

Démonstration. On calcule le noyau par récurrence sur n . Soit P tel que $\tilde{P} = 0$. On a $P = \sum_i Q'_i(X_1, \dots, X_{n-1})X_n^i$. On écrit P sous la forme $P = (X_n^q - X_n)R + \sum_{i < q} Q_i(X_1, \dots, X_{n-1})X_n^i$. Pour tous $a_1, \dots, a_n \in \mathbb{K}$, on a $\tilde{P}(a_1, \dots, a_n) = 0$ donc $\sum_{i < q} Q_i(a_1, \dots, a_{n-1})a_n^i = 0$. Le polynôme $\sum_{i < q} Q_i(a_1, \dots, a_{n-1})X_n^i$ admet au moins q racines et est de degré $< q$, donc il est nul. Alors, pour tout $i < q$, $Q_i(a_1, \dots, a_{n-1}) = 0$. Donc, pour tout $i < q$, $\tilde{Q}_i = 0$. On applique aux Q_i l'hypothèse de récurrence. Donc $Q_i \in (X_1^q - X_1, \dots, X_{n-1}^q - X_{n-1})$, puis $P \in (X_1^q - X_1, \dots, X_n^q - X_n)$.

Étudions la surjectivité. Toute fonction $f \in \mathcal{F}(\mathbb{K}^n, \mathbb{K})$ peut s'écrire $f = \sum_{(x_1, \dots, x_n)} f(x_1, \dots, x_n) \mathbf{1}_{(x_1, \dots, x_n)}$. On remarque que $x \mapsto 1 - x^{q-1}$ est la fonction indicatrice de 0 sur $\mathbb{K}$. Donc $(x_1, \dots, x_n) \mapsto \prod_{i=1}^n (1 - x_i^{q-1})$ est la fonction indicatrice de $(0, \dots, 0)$ sur $\mathbb{K}^n$. Alors, $(x_1, \dots, x_n) \mapsto \prod_{i=1}^n (1 - (x_i - y_i)^{q-1})$ est la fonction indicatrice du point $(y_1, \dots, y_n)$ sur $\mathbb{K}^n$. $\square$

Réduction d'un endomorphisme

Références : [8, 10]

Commençons par fixer quelques notations. Soit $\mathbb{K}$ un corps commutatif, E un $\mathbb{K}$ -espace vectoriel de dimension finie $n \geq 1$. On note $\mathcal{L}(E)$ l'algèbre des endomorphismes de E . On note $\mathcal{GL}(E)$ le groupe des automorphismes de E . $M_n(\mathbb{K})$ désignera l'algèbre des matrices $n \times n$ à coefficients dans $\mathbb{K}$ et $GL_n(\mathbb{K})$ le groupe linéaire des matrices inversibles.

Le problème que nous étudierons est le suivant. $\mathcal{GL}(E)$ opère sur $\mathcal{L}(E)$ par conjugaison :

$$\mathcal{GL}(E) \times \mathcal{L}(E) \rightarrow \mathcal{L}(E), \quad (w, u) \mapsto wuw^{-1}.$$

Deux endomorphismes d'une même orbite sont dits semblables. Une orbite est appelée classe de similitude.

Le but est de caractériser de façon effective les classes de similitude. On remarque que deux endomorphismes u et v sont semblables si et seulement si u et v ont même matrice dans des bases éventuellement différentes. Nous serons amenés à chercher pour telle matrice une matrice simple, d'où le terme réduction.

On note $E^* = \text{Hom}_{\mathbb{K}}(E, \mathbb{K})$ le dual de E et, pour $V \subset E^*$, $V^\circ = \{x \in E : \forall \lambda \in V, \lambda(x) = 0\}$ son annulateur dans E . On note Id l'identité et $\text{Vect}(S)$ le sous-espace engendré par une partie S . La notation $M_{m,n}(\mathbb{K})$ désigne les matrices à m lignes et n colonnes, $M_n(\mathbb{K}) = M_{n,n}(\mathbb{K})$, I_n la matrice identité, $GL_n(\mathbb{K})$ le groupe des matrices inversibles et $SL_n(\mathbb{K})$ celui des matrices de déterminant 1. On note $\ker u$ et $\text{Im } u$ le noyau et l'image de u , $\text{rg } u = \dim \text{Im } u$ son rang et $\text{tr } u$ sa trace. Pour une matrice A , ${}^t A$ est sa transposée, $\bar{A}$ sa conjuguée coefficient par coefficient, $A^* = {}^t \bar{A}$ sa transposée conjuguée et $\text{adj}(A)$ sa comatrice transposée.

3.1 Polynôme minimal et polynôme caractéristique d'un endomorphisme

Soit $u \in \mathcal{L}(E)$.

3.1.1 Polynôme minimal

On définit un morphisme de $\mathbb{K}$ -algèbres

$$\varphi_u : \mathbb{K}[X] \rightarrow \mathcal{L}(E), \quad P \mapsto P(u).$$

Son image est la sous-algèbre commutative $\mathbb{K}[u] = \{P(u), P \in \mathbb{K}[X]\}$.

Son noyau est un idéal de $\mathbb{K}[X]$. Comme $\mathcal{L}(E)$ est de dimension finie, φ_u ne peut être injectif. Comme $\mathbb{K}[X]$ est principal, $\ker \varphi_u$ également. Il admet un unique générateur unitaire appelé polynôme minimal de u et noté π_u .

Par convention, le polynôme minimal de l'endomorphisme de l'espace nul vaut 1 ; cette convention s'appliquera aux restrictions ci-dessous.

On a $\mathbb{K}[u] \simeq \mathbb{K}[X]/(\pi_u)$. En particulier, $\dim \mathbb{K}[u] = \deg \pi_u$.

Remarques. Ceci est faux en général si E est de dimension infinie. Par exemple,

$$\mathbb{K}[X] \rightarrow \mathbb{K}[X], \quad P \mapsto XP$$

n'a pas de polynôme minimal, ni même de polynôme annulateur non nul.

On a, pour $u \in \mathcal{L}(E)$ et $w \in \mathcal{GL}(E)$, pour tout polynôme P , $P(w^{-1}uw) = w^{-1}P(u)w$. Donc l'idéal annulateur de u est celui de $w^{-1}uw$. Le polynôme minimal est un invariant de la classe de similitude.

3.1.2 Polynôme caractéristique

On commence par définir le polynôme caractéristique d'une matrice $A \in M_n(\mathbb{K})$ en posant

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

C'est le déterminant de la matrice $XI_n - A \in M_n(\mathbb{K}[X])$.

On remarque que, si $P \in GL_n(\mathbb{K})$,

$$\chi_{PAP^{-1}} = \det(XI_n - PAP^{-1}) = \det(P(XI_n - A)P^{-1}) = \det P \det(XI_n - A) \det P^{-1} = \det(XI_n - A) = \chi_A.$$

Définition 3.1. Toutes les matrices de l'endomorphisme u ont le même polynôme caractéristique que l'on appelle *polynôme caractéristique de u* et que l'on note χ_u .

Remarque. Le polynôme caractéristique est aussi un invariant de la classe de similitude.

3.2 Sous-espaces stables

Soit $u \in \mathcal{L}(E)$.

Définition 3.2. Soit F un sous-espace vectoriel de E . On dit que F est *stable* par u (ou u -stable) si $u(F) \subset F$.

Remarque. Si F est u -stable et G est un supplémentaire de F , alors la matrice de u dans une base associée à la somme directe $F \oplus G$ s'écrit

$$\begin{pmatrix} \star & \star \\ 0 & \star \end{pmatrix} \begin{matrix} \} F \\ \} G \end{matrix},$$

où $\star$ désigne des coefficients sur lesquels on ne connaît aucune information particulière *a priori*.

Exemple. • Si F est u -stable de dimension 1, F est une droite propre : $F = \mathbb{K}e$ et $u(e) = \lambda e$.

- Un sous-espace d'un sous-espace propre est stable : $F \subset \ker(u - \lambda \text{Id})$.
- $\ker v$ et $\text{Im } v$ sont u -stables, lorsque v et u commutent.
- $\{0\}$ et E sont u -stables.

Remarque. Si F est un sous-espace vectoriel u -stable, alors u induit sur F un endomorphisme

$$u_F : F \rightarrow F, \quad x \mapsto u(x).$$

En prenant G un supplémentaire de F , $\mathcal{B}_F$ une base de F et $\mathcal{B}_G$ une base de G , la matrice de u dans la base $\mathcal{B}_F \cup \mathcal{B}_G$ est

$$\begin{pmatrix} A & \star \\ 0 & \star \end{pmatrix} \begin{matrix} \} F \\ \} G \end{matrix},$$

où A est la matrice de u_F dans la base $\mathcal{B}_F$.

On en déduit que le polynôme caractéristique de u_F divise celui de u : $\chi_{u_F} | \chi_u$.

On a aussi $\pi_u(u) = 0$ donc $\pi_u(u_F) = 0$. Alors $\pi_{u_F} | \pi_u$.

Proposition 3.1. Si $E = F \oplus G$, avec F et G u -stables, alors $\chi_u = \chi_{u_F} \chi_{u_G}$ et $\pi_u = \pi_{u_F} \vee \pi_{u_G}$.

La démonstration de ce résultat est laissée en exercice.

Proposition 3.2. Soit $Q \in \mathbb{K}[X]$. Posons $F = \ker Q(u)$. (F est donc u -stable.) Alors $\pi_{u_F} = \pi_u \wedge Q$.

Démonstration. On sait d'une part que π_{u_F} divise π_u . D'autre part, $Q(u_F) = Q(u)_F = 0$. Donc $\pi_{u_F}|Q$. On en déduit que $\pi_{u_F}|\pi_u \wedge Q$.

Réciproquement, posons $R = \pi_u \wedge Q$. Comme $R|Q$, on a $\ker R(u) \subset \ker Q(u)$. En écrivant, d'après le théorème de Bézout, $R = A\pi_u + BQ$, on a $R(u) = B(u)Q(u)$, donc $\ker Q(u) \subset \ker R(u)$.

Supposons $R = \pi_{u_F}S$. On sait que $R|\pi_u$, donc $\pi_u = RT$. Alors $R(u)T(u) = 0$ donc $\text{Im } T(u) \subset \ker R(u)$. On a $\ker R(u) = F \subset \ker \pi_{u_F}(u)$. Donc $\text{Im } T(u) \subset \ker \pi_{u_F}(u)$, puis $\pi_{u_F}(u)T(u) = 0$.

On a $\pi_u = RT = \underbrace{\pi_{u_F}T}_{\substack{\text{annule } u \\ \text{et divise } \pi_u}} S$. Par minimalité de π_u , on a $S \in \mathbb{K}[X]^*$, donc $R \in \mathbb{K}^*\pi_{u_F}$. $\square$

3.3 Sous-espaces cycliques et théorème de Cayley-Hamilton

Définition 3.3. Soit $x \in E$. On appelle *sous-espace u -cyclique engendré par x* le plus petit sous-espace vectoriel u -stable qui contient x . On le note $\langle x \rangle_u$.

Un tel sous-espace existe car une intersection de sous-espaces u -stables est u -stable :

$$\langle x \rangle_u = \bigcap_{\substack{F \text{ } u\text{-stable} \\ x \in F}} F.$$

Proposition 3.3. On a $\langle x \rangle_u = \{P(u)(x), P \in \mathbb{K}[X]\}$.

La démonstration de cette proposition est laissée en exercice.

Proposition 3.4. Soit $x \in E$. Soit $F = \langle x \rangle_u$. Soit $P \in \mathbb{K}[X]$. On a l'équivalence suivante

$$P(u_F) = 0 \iff P(u)(x) = 0.$$

Démonstration. L'implication directe est triviale, car $x \in F$. Traitons l'implication réciproque. Soit $y \in F$. Il existe $Q \in \mathbb{K}[X]$ tel que $y = Q(u)(x)$. On a $P(u)(y) = P(u)(Q(u)(x)) = Q(u)(P(u)(x)) = 0$, donc $P(u_F) = 0$. $\square$

Proposition 3.5. Soit $x \in E \setminus \{0\}$ et $F = \langle x \rangle_u$. Soit d l'entier défini par

$$\begin{cases} (x, u(x), \dots, u^{d-1}(x)) \text{ libre} \\ (x, u(x), \dots, u^d(x)) \text{ liée.} \end{cases}$$

Alors $(x, u(x), \dots, u^{d-1}(x))$ est une base de F et $\dim F = d$. En outre, la matrice de u_F dans cette base est de la forme

$$\begin{pmatrix} 0 & \cdots & \cdots & 0 & a_d \\ 1 & \ddots & & \vdots & \vdots \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & a_2 \\ 0 & \cdots & 0 & 1 & a_1 \end{pmatrix},$$

$$\text{et } \pi_{u_F} = X^d - a_1X^{d-1} - a_2X^{d-2} - \dots - a_d.$$

Démonstration. $(x, u(x), \dots, u^{d-1}(x))$ est libre si et seulement si pour tout $P \in \mathbb{K}[X]$ tel que $\deg P < d$ et $P(u)(x) = 0$, on a $P = 0$. $(x, u(x), \dots, u^d(x))$ est liée si et seulement si il existe un polynôme $P \in \mathbb{K}[X]$ non nul, de degré au plus d , tel que $P(u)(x) = 0$. On en déduit que le polynôme minimal de u_F est exactement de degré d .

Soit $(\lambda_0, \dots, \lambda_d) \in \mathbb{K}^{d+1} \setminus \{0\}$ tel que $\lambda_0 x + \lambda_1 u(x) + \dots + \lambda_d u^d(x) = 0$. On a $\lambda_d \neq 0$ sinon $(x, u(x), \dots, u^{d-1}(x))$ serait liée. En divisant par λ_d , on obtient une égalité de la forme

$$u^d(x) = a_d x + a_{d-1} u(x) + \dots + a_1 u^{d-1}(x).$$

Le polynôme $X^d - a_1 X^{d-1} - a_2 X^{d-2} - \dots - a_d$ est le polynôme P de plus petit degré parmi les non nuls vérifiant $P(u)(x) = 0$. C'est donc π_{u_F} .

Par hypothèse, $(x, u(x), \dots, u^{d-1}(x))$ est une famille libre de F . Montrons qu'elle est génératrice. Soit $Q \in \mathbb{K}[X]$. On a $Q = R\pi_{u_F} + S$ avec $\deg S < \deg \pi_{u_F} = d$. $Q(u)(x) = S(u)(x) \in \text{Vect}(x, u(x), \dots, u^{d-1}(x))$.

On écrit la matrice de u_F dans cette base

$$\begin{pmatrix} 0 & \cdots & \cdots & 0 & a_d \\ 1 & \ddots & & \vdots & \vdots \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & a_2 \\ 0 & \cdots & 0 & 1 & a_1 \end{pmatrix} \begin{matrix} \} x \\ \vdots \\ \vdots \\ \} u^{d-2}(x) \\ \} u^{d-1}(x) \end{matrix}.$$

□

Définition 3.4. Soit $P \in \mathbb{K}[X]$ unitaire non constant. On note $d = \deg P$. On pose $P = X^d - a_1 X^{d-1} - \dots - a_d$. On appelle *matrice compagnon* associée à P la matrice $d \times d$ suivante :

$$\begin{pmatrix} 0 & \cdots & \cdots & 0 & a_d \\ 1 & \ddots & & \vdots & \vdots \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & a_2 \\ 0 & \cdots & 0 & 1 & a_1 \end{pmatrix}.$$

On la note $\mathcal{C}(P)$.

Exemple. Soit $x \in E \setminus \{0\}$ et $F = \langle x \rangle_u$. u_F admet comme matrice la matrice compagnon associée à π_{u_F} .

Proposition 3.6. Soit u un endomorphisme de E qui admet comme matrice la matrice compagnon associée à un polynôme P . Alors E est u -cyclique (i.e. il existe $x \in E$ tel que $E = \langle x \rangle_u$) et $P = \chi_u = \pi_u$.

Démonstration. Soit $(e_1, \dots, e_d)$ une base de E sur laquelle la matrice de u est $\mathcal{C}(P)$. Alors $E = \langle e_1 \rangle_u$. En effet, $\text{Vect}(e_1, \dots, e_d) = \text{Vect}(e_1, \dots, u^{d-1}(e_1))$.

On sait alors que les coefficients de $u^d(e_1)$ sur la base $(e_1, \dots, u^{d-1}(e_1)) = (e_1, \dots, e_d)$ définissent ceux du polynôme minimal :

$$\pi_u = X^d - a_1 X^{d-1} - \dots - a_d.$$

Calculons χ_u en développant $\det(XI_d - \mathcal{C}(P))$ suivant sa dernière colonne. Les mineurs correspondants sont triangulaires par blocs, et l'on obtient

$$\chi_u = (X - a_1)X^{d-1} - \sum_{i=1}^{d-1} a_{d-i+1} X^{i-1} = X^d - a_1 X^{d-1} - \dots - a_d = P.$$

□

Corollaire 3.7. Soit F un sous-espace vectoriel u -cyclique. Alors $\pi_{u_F} = \chi_{u_F}$.

Théorème 3.8 (Cayley-Hamilton). π_u divise χ_u .

Démonstration. Soit $x \in E$. On pose $F = \langle x \rangle_u$. D'après le corollaire 3.7, on a $\pi_{u_F} = \chi_{u_F}$. Donc, $\chi_{u_F}(u)(x) = 0$. Or $\chi_{u_F} | \chi_u$. Donc $\chi_u(u)(x) = 0$, puis $\chi_u(u) = 0$. On en déduit que $\pi_u | \chi_u$.

Nota bene : Une autre preuve part de l'identité $(XI_n - A) \operatorname{adj}(XI_n - A) = \chi_A(X)I_n$. En comparant les coefficients des puissances de X , puis en multipliant les identités obtenues par les puissances appropriées de A , une somme télescopique donne $\chi_A(A) = 0$. $\square$

3.4 Lemme de décomposition des noyaux

3.4.1 Enoncé et premières conséquences

Soit $u \in \mathcal{L}(E)$.

Lemme 3.9 (Décomposition des noyaux). Soit P un polynôme de $\mathbb{K}[X]$ tel que $P(u) = 0$. On suppose que $P = P_1 \dots P_r$ où les P_i sont premiers entre eux deux à deux. Alors E admet la décomposition en somme directe de sous-espaces stables suivante

$$E = \bigoplus_{i=1}^r \ker P_i(u).$$

En outre, les projecteurs associés à cette décomposition sont des polynômes en u .

Démonstration. Si $P = 0$, un facteur P_i est nul et tous les autres sont constants non nuls, puisque les facteurs sont premiers entre eux deux à deux. La décomposition se réduit alors à $E = \ker P_i(u)$, les autres noyaux étant nuls; les projecteurs sont Id et 0 . On peut donc supposer $P \neq 0$.

On a $P = P_1 \dots P_r$. Si $i \in \{1, \dots, r\}$, on pose $Q_i = \prod_{j \neq i} P_j = \frac{P}{P_i}$. Comme les P_i sont premiers entre eux deux à deux, les Q_i sont premiers entre eux dans leur ensemble. (En effet, si A est un diviseur irréductible de Q_i , il existe $j \neq i$ tel que A divise P_j et aucun autre des P_k , donc $A \nmid Q_j$.) D'après le théorème de Bézout, il existe des polynômes $A_1, \dots, A_r$ tels que $\sum_i A_i Q_i = 1$.

On pose pour tout i , $\pi_i = (A_i Q_i)(u) \in \mathbb{K}[u]$. On a $\sum_i \pi_i = \operatorname{Id}$. En outre, si $i \neq j$, $P | Q_i Q_j$, donc $P | (A_i Q_i)(A_j Q_j)$, puis $\pi_i \pi_j = 0$. On sait alors que les π_i sont des projecteurs et donc que

$$E = \bigoplus_i \operatorname{Im} \pi_i, \quad \ker \pi_i = \bigoplus_{j \neq i} \operatorname{Im} \pi_j.$$

$P_i(u)\pi_i = 0$ car $P | P_i A_i Q_i$. Donc $\operatorname{Im} \pi_i \subset \ker P_i(u)$.

Réciproquement, soit $x \in \ker P_i(u)$. Soit $j \neq i$. P_i divise $A_j Q_j$, donc $\pi_j(x) = 0$. On a $x = \sum_j \pi_j(x) = \pi_i(x) \in \operatorname{Im} \pi_i$. D'où $\ker P_i(u) \subset \operatorname{Im} \pi_i$. $\square$

On décompose le polynôme minimal de u en produit de puissances d'irréductibles deux à deux non associés :

$$\pi_u = P_1^{\alpha_1} \dots P_r^{\alpha_r}.$$

Alors le lemme de décomposition des noyaux assure que

$$E = \bigoplus_{i=1}^r \ker P_i^{\alpha_i}(u).$$

Sur $\ker P_i^{\alpha_i}(u)$, le polynôme minimal de la restriction de u est $P_i^{\alpha_i}$ (voir la proposition 3.2).

Définition 3.5. Les sous-espaces vectoriels $\ker P_i^{\alpha_i}(u)$ sont appelés les *sous-espaces caractéristiques* de u .

Proposition 3.10. Soit P un polynôme. On a

$$\{0\} \subset \ker P(u) \subset \ker P^2(u) \subset \dots \subset \ker P^i(u) \subset \dots$$

Si $\ker P^i(u) = \ker P^{i+1}(u)$, alors $\ker P^{i+1}(u) = \ker P^{i+2}(u)$. La suite est donc strictement croissante, puis stationnaire : il existe un entier α tel que

$$\begin{cases} \forall i < \alpha, \ker P^i(u) \subsetneq \ker P^{i+1}(u) \\ \forall i \geq \alpha, \ker P^i(u) = \ker P^{i+1}(u). \end{cases}$$

On a alors

$$\ker P^\alpha(u) = \bigcup_{i \in \mathbb{N}} \ker P^i(u).$$

La démonstration est laissée au lecteur.

Proposition 3.11. Lorsque P est un diviseur irréductible du polynôme minimal de u ,

$$\bigcup_{i \in \mathbb{N}} \ker P^i(u) = \ker P^\alpha(u),$$

où α est la multiplicité de P dans π_u .

Démonstration. Pour $i < \alpha$, on a $\pi_{\ker P^i(u)} = P^i$ et $\pi_{\ker P^{i+1}(u)} = P^{i+1}$, d'après la proposition 3.2. Nécessairement, on a $\ker P^i(u) \subsetneq \ker P^{i+1}(u)$. En outre, on a $\pi_{\ker P^{\alpha+1}(u)} = P^\alpha$, donc $P^\alpha(u)$ est nul sur $\ker P^{\alpha+1}(u)$, d'où l'égalité $\ker P^{\alpha+1}(u) = \ker P^\alpha(u)$. $\square$

3.4.2 Application à la décomposition de Dunford

Théorème 3.12 (Décomposition de Dunford). On suppose π_u scindé. Alors il existe un unique couple $(d, n) \in \mathcal{L}(E)^2$ tel que d est diagonalisable, n est nilpotent, $dn = nd$ et $u = d + n$. En outre, d et n sont des polynômes en u .

Remarque. En étudiant une réduction commune à d et n (d et n commutent), on obtient la réduction de Jordan de u .

Démonstration. On a $\pi_u = \prod_{i=1}^r (X - \lambda_i)^{\alpha_i}$. Alors $E = \bigoplus_{i=1}^r \ker(u - \lambda_i \text{Id})^{\alpha_i}$. On introduit $\pi_1, \dots, \pi_r$ les projecteurs associés à cette décomposition. Pour tout i , on définit d_i et n_i de la sorte :

$$u_{\ker(u - \lambda_i \text{Id})^{\alpha_i}} = \underbrace{\lambda_i \text{Id}_{\ker(u - \lambda_i \text{Id})^{\alpha_i}}}_{d_i} + \underbrace{(u - \lambda_i \text{Id})_{\ker(u - \lambda_i \text{Id})^{\alpha_i}}}_{n_i}.$$

On remarque que n_i est bien nilpotent, puisque $n_i^{\alpha_i} = 0$. On a finalement

$$u = \underbrace{\sum_{i=1}^r \lambda_i \pi_i}_d + \underbrace{\sum_{i=1}^r (u - \lambda_i \text{Id}) \pi_i}_n.$$

On peut aisément montrer que d est diagonalisable, que n est nilpotent et que d et n commutent. Comme pour tout i , $\pi_i \in \mathbb{K}[u]$, alors d et n appartiennent à $\mathbb{K}[u]$.

Traitons l'unicité. On suppose avoir aussi $u = d' + n'$ avec d' diagonalisable, n' nilpotent, $d'n' = n'd'$. On sait donc que d' et n' commutent avec u , donc avec d et n qui sont des polynômes en u .

On a $d' - d = n - n'$. D'une part, $d' - d$ est diagonalisable, car deux endomorphismes diagonalisables qui commutent ont une base commune de diagonalisation. D'autre part, $n - n'$ est nilpotent, car n et n' sont nilpotents et commutent. On a donc un morphisme à la fois diagonalisable et nilpotent, il est nécessairement nul. $\square$

Corollaire 3.13. Soit $M \in M_n(\mathbb{R})$. M admet une décomposition de Dunford (D, N) dans $M_n(\mathbb{C})$. On a en outre $D, N \in M_n(\mathbb{R})$.

Démonstration. $\mathbb{C}$ est algébriquement clos, donc le polynôme minimal de M est scindé. Donc, il existe un unique couple $(D, N) \in M_n(\mathbb{C})^2$ tel que $M = D + N$, D diagonalisable, N nilpotente et $DN = ND$. M est réelle, donc $M = \overline{M}$. Donc $M = \overline{D} + \overline{N}$, avec $\overline{D}$ diagonalisable, $\overline{N}$ nilpotente et $\overline{D}\overline{N} = \overline{N}\overline{D}$. Par unicité, on a $D = \overline{D}$ et $N = \overline{N}$. Donc D et N sont réelles.

Attention ! D n'est pas nécessairement diagonalisable dans $M_n(\mathbb{R})$. □

3.5 Invariants de similitude

3.5.1 Résultats préliminaires

Proposition 3.14. Il existe un sous-espace u -cyclique F de E tel que $\pi_{u_F} = \pi_u$.

La démonstration de cette proposition nécessite plusieurs lemmes. Le lecteur remarquera qu'elle est très similaire à la démonstration du fait que, dans un groupe abélien, il existe un élément d'ordre l'exposant du groupe.

Lemme 3.15. Supposons $\pi_u = P^\alpha$, avec P irréductible. Alors il existe un sous-espace u -cyclique F tel que $\pi_{u_F} = \pi_u$.

Démonstration. On a $\{0\} \subsetneq \ker P(u) \subsetneq \dots \subsetneq \ker P^\alpha(u) = E$. On prend $x \in E \setminus \ker P^{\alpha-1}(u)$. Alors, si $F = \langle x \rangle_u$, $\pi_{u_F} = P^\alpha$ (car π_{u_F} divise P^α et car $P^{\alpha-1}(u)(x) \neq 0$). □

Lemme 3.16. Soient x et y tels que les polynômes $\pi_{u_{\langle x \rangle_u}}$ et $\pi_{u_{\langle y \rangle_u}}$ soient premiers entre eux. Alors $\pi_{u_{\langle x+y \rangle_u}} = \pi_{u_{\langle x \rangle_u}} \pi_{u_{\langle y \rangle_u}}$.

Démonstration. Posons $P = \pi_{u_{\langle x \rangle_u}}$ et $Q = \pi_{u_{\langle y \rangle_u}}$. On a $P(u)(x) = 0$ et $Q(u)(y) = 0$, donc $PQ(u)(x) = 0$ et $PQ(u)(y) = 0$, puis $PQ(u)(x+y) = 0$. On pose $R = \pi_{u_{\langle x+y \rangle_u}}$. Alors $R|PQ$.

En écrivant $x = (x+y) - y$, on peut montrer que $P|QR$. On a $P \wedge Q = 1$. D'après le lemme de Gauss, on a $P|R$. De même, on peut montrer que $Q|R$. Donc $PQ = P \vee Q|R$. □

Démonstration. (proposition 3.14) On écrit $\pi_u = \prod_{i=1}^r P_i^{\alpha_i}$, où les P_i sont irréductibles. Pour tout i , il existe $x_i \in E$ tel que $\pi_{u_{\langle x_i \rangle_u}} = P_i^{\alpha_i}$. On a $\pi_{u_{\langle \sum_i x_i \rangle_u}} = \pi_u$. □

3.5.2 Invariants de similitude et réduction de Frobenius

Théorème 3.17 (Réduction de Frobenius). Soit $u \in \mathcal{L}(E)$. Il existe un unique entier r et une unique famille de polynômes unitaires non constants $Q_1, \dots, Q_r$ soumis à la condition $Q_r | \dots | Q_1$ tels que u admette pour matrice

$$\begin{pmatrix} \mathcal{C}(Q_1) & & 0 \\ & \ddots & \\ 0 & & \mathcal{C}(Q_r) \end{pmatrix}.$$

Remarques. La matrice par blocs ainsi ordonnée est unique ; la base qui la réalise ne l'est pas nécessairement. Deux endomorphismes u et v sont semblables si et seulement si ils ont le même ensemble de matrices. Donc u et v sont semblables si et seulement si ils correspondent, via ce théorème, à la même famille de polynômes.

Définition 3.6. Les polynômes $Q_1, \dots, Q_r$ sont appelés les *invariants de similitude* de u .

Proposition 3.18. On a $\pi_u = Q_1$ et $\chi_u = \prod_{i=1}^r Q_i$.

Démonstration. D'après le théorème 3.17, on peut écrire

$$E = \bigoplus_{i=1}^r F_i,$$

avec F_i u -stable et u_{F_i} admet $\mathcal{C}(Q_i)$ comme matrice. D'après la proposition 3.6, F_i est u -cyclique et $\chi_{u_{F_i}} = \pi_{u_{F_i}} = Q_i$. On a alors

$$\pi_u = \bigvee_{i=1}^r \pi_{u_{F_i}} = \bigvee_{i=1}^r Q_i = Q_1, \quad \chi_u = \prod_{i=1}^r \pi_{u_{F_i}} = \prod_{i=1}^r Q_i.$$

□

Corollaire 3.19. $\pi_u | \chi_u | (\pi_u)^r$, donc π_u et χ_u ont les mêmes diviseurs irréductibles.

Démonstration. (théorème 3.17) L'existence se démontre par récurrence sur n . D'abord, la proposition 3.14 assure qu'il existe un sous-espace u -cyclique F_1 tel que $\pi_{u_{F_1}} = \pi_u$.

La preuve consiste en le fait d'exhiber un supplémentaire u -stable de F_1 . Soit x un u -générateur de F_1 . Soit $d_1 = \dim F_1$. Alors $(x, u(x), \dots, u^{d_1-1}(x))$ est une base de F_1 . Il existe donc une forme linéaire λ telle que

$$\begin{cases} \lambda(x) = \dots = \lambda(u^{d_1-2}(x)) = 0 \\ \lambda(u^{d_1-1}(x)) = 1. \end{cases}$$

Posons $G = \langle \lambda \rangle_{t_u}^\circ$, l'orthogonal de $\langle \lambda \rangle_{t_u}$, où ${}^t u$ est l'endomorphisme transposé de u défini par

$${}^t u : E^* \rightarrow E^*, \quad \lambda \mapsto \lambda \circ u.$$

Montrons que G est un supplémentaire u -stable de F_1 .

Montrons que G est u -stable. Soit $y \in G$. Pour tout i , on a $\lambda \circ u^i(y) = 0$. Donc, pour tout i , on a $\lambda \circ u^i(u(y)) = 0$. D'où $u(y) \in G$.

Montrons que $F_1 \cap G = \{0\}$. Soit $y \in F_1 \cap G$. On peut écrire $y = \alpha_0 x + \alpha_1 u(x) + \dots + \alpha_{d_1-1} u^{d_1-1}(x)$. On suppose que $\forall k > j, \alpha_k = 0$. On a $0 = \lambda \circ u^{d_1-1-j}(y) = \lambda(\alpha_0 u^{d_1-1-j}(x) + \alpha_1 u^{d_1-j}(x) + \dots + \alpha_j u^{d_1-1}(x)) = \alpha_j$. On en déduit que $\alpha_0 = \dots = \alpha_{d_1-1} = 0$, i.e. $y = 0$.

Montrons que $\dim G = \text{codim } F_1$. On sait déjà que $\dim G \leq \text{codim } F_1$. On a $G = \langle \lambda \rangle_{t_u}^\circ$, donc $\dim G = \text{codim } \langle \lambda \rangle_{t_u}$ (car $\dim F + \dim F^\circ = \dim E$). Or $\dim \langle \lambda \rangle_{t_u} = \deg \pi_{t_u|_{\langle \lambda \rangle_{t_u}}} \leq \deg \pi_{t_u} = \deg \pi_u = \dim F_1$, car $\pi_{t_u} = \pi_u$.

Nota bene : L'application

$$\mathcal{L}(E) \rightarrow \mathcal{L}(E^*), \quad u \mapsto {}^t u$$

est un anti-isomorphisme d'algèbres (${}^t(v \circ u) = {}^t u \circ {}^t v$). Pour tout $P \in \mathbb{K}[X]$, on a $P({}^t u) = {}^t P(u)$. On en déduit que u et ${}^t u$ ont le même idéal annulateur, donc le même polynôme minimal.

On en déduit que $\dim G \geq \text{codim } F_1$.

Une fois G construit, on applique l'hypothèse de récurrence à u_G : on a $Q_r | \dots | Q_2$ et u_G admet pour matrice

$$\begin{pmatrix} \mathcal{C}(Q_2) & & 0 \\ & \ddots & \\ 0 & & \mathcal{C}(Q_r) \end{pmatrix}.$$

Comme $E = F_1 \oplus G$, u admet pour matrice

$$\begin{pmatrix} \mathcal{C}(Q_1) & & 0 \\ & \ddots & \\ 0 & & \mathcal{C}(Q_r) \end{pmatrix}.$$

Comme $Q_2 = \pi_{u_G}$ et $Q_1 = \pi_u$, on a aussi $Q_2|Q_1$.

Reste à démontrer l'unicité de la famille $Q_1, \dots, Q_r$ (voir [8]). On suppose avoir deux familles $Q_r| \dots | Q_1$ et $R_s| \dots | R_1$ telles que $E = F_1 \oplus \dots \oplus F_r$ et $E = G_1 \oplus \dots \oplus G_s$, où chaque F_i est un sous-espace u -cyclique tel que $\pi_{u_{F_i}} = Q_i$ et chaque G_j est un sous-espace u -cyclique tel que $\pi_{u_{G_j}} = R_j$.

On sait déjà que $\pi_u = Q_1$ et $\pi_u = R_1$. Donc $Q_1 = R_1$.

La suite de notre démonstration nécessite un lemme.

Lemme 3.20. *Soit F un sous-espace u -cyclique. Soit $Q \in \mathbb{K}[X]$. Alors $Q(u)(F)$ est encore un sous-espace u -cyclique et*

$$\pi_{u_{Q(u)(F)}} = \frac{\pi_{u_F}}{\pi_{u_F} \wedge Q}.$$

Démonstration. Soit x tel que $F = \langle x \rangle_u$. On a $F = \{P(u)(x), P \in \mathbb{K}[X]\}$, donc $Q(u)(F) = \{Q(u)(P(u)(x)), P \in \mathbb{K}[X]\} = \{P(u)(Q(u)(x)), P \in \mathbb{K}[X]\}$. D'où $Q(u)(F) = \langle Q(u)(x) \rangle_u$. $Q(u)(F)$ est donc u -cyclique.

Soit $P \in \mathbb{K}[X]$. Comme x engendre F , on a

$$P(u_{Q(u)(F)}) = 0 \iff (PQ)(u)(x) = 0 \iff \pi_{u_F} \mid PQ \iff \frac{\pi_{u_F}}{\pi_{u_F} \wedge Q} \mid P.$$

La dernière équivalence découle du lemme de Gauss. Le générateur unitaire de cet idéal annulateur est donc le quotient annoncé. $\square$

Reprenons la preuve du théorème. On démontre l'unicité par récurrence sur i . L'hypothèse de récurrence sera $\forall j < i, Q_j = R_j$.

On applique $Q_i(u)$ aux deux décompositions $E = F_1 \oplus \dots \oplus F_r$ et $E = G_1 \oplus \dots \oplus G_s$. Comme tous les sous-espaces sont u -stables, on a $Q_i(u)(E) = Q_i(u)(F_1) \oplus \dots \oplus Q_i(u)(F_{i-1})$ (car pour $j \geq i$, $Q_j|Q_i$ donc $Q_i(u)(F_j) = 0$) et $Q_i(u)(E) = Q_i(u)(G_1) \oplus \dots \oplus Q_i(u)(G_s)$.

D'après le lemme et l'hypothèse de récurrence, pour tout $\ell < i$, le polynôme minimal des restrictions de u à $Q_i(u)(F_\ell)$ et $Q_i(u)(G_\ell)$ est le même (c'est $\frac{Q_\ell}{Q_i}$). Ce sont des sous-espaces u -cycliques, donc en particulier, ils ont la même dimension.

Alors $\dim Q_i(u)(E) = \sum_{\ell=1}^{i-1} \dim Q_i(u)(F_\ell) = \sum_{\ell=1}^{i-1} \dim Q_i(u)(G_\ell)$, ce qui impose $\bigoplus_{j \geq i} Q_i(u)(G_j) = \{0\}$. Donc $Q_i(u)(G_i) = \{0\}$, puis $R_i|Q_i$. Par symétrie, on obtient l'égalité. Si une décomposition s'arrêtait avant l'autre, l'égalité des dimensions des blocs déjà identifiés laisserait une somme de blocs non nuls de dimension totale nulle, ce qui est impossible. Ainsi $r = s$. $\square$

3.5.3 Réduction de Jordan et autres applications

Cas particulier d'un endomorphisme nilpotent

u est nilpotent si et seulement si $\pi_u = X^d$, où d est l'indice de nilpotence de u . Le théorème 3.17 assure qu'il existe un unique entier r et une unique famille de polynômes $Q_r| \dots | Q_1$ tels que u admette pour matrice

$$\begin{pmatrix} \mathcal{C}(Q_1) & & 0 \\ & \ddots & \\ 0 & & \mathcal{C}(Q_r) \end{pmatrix}.$$

On a $Q_1 = \pi_u = X^d$ donc pour tout i , $Q_i = X^{d_i}$, avec $d_1 = d$ et $d_1 \geq \dots \geq d_r$. Or la matrice compagnon associée au polynôme X^{d_i} est

$$\begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ 1 & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix} \begin{matrix} \uparrow \\ \vdots \\ d_i \\ \vdots \\ \downarrow \end{matrix}.$$

On appelle une telle matrice bloc de Jordan et on la note $J(0, d_i)$. Réciproquement, un bloc de Jordan de valeur propre nulle est une matrice compagnon associée à une puissance de X .

Pour un endomorphisme u nilpotent, le théorème 3.17 est équivalent au théorème donnant la réduction de Jordan de u . On en déduit, grâce au lemme des noyaux, le théorème de jordanisation dans le cas où π_u est scindé.

Théorème 3.21 (Réduction de Jordan). *On suppose π_u scindé. Alors, dans une certaine base, u a pour matrice une matrice diagonale par blocs formée de blocs de Jordan*

$$J(\lambda, d) = \begin{pmatrix} \lambda & 0 & \cdots & \cdots & 0 \\ 1 & \ddots & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & \lambda \end{pmatrix} \begin{matrix} \uparrow \\ \vdots \\ d \\ \vdots \\ \downarrow \end{matrix}.$$

Démonstration. Donnons une démonstration rapide de ce résultat. Notons $\pi_u = \prod_{i=1}^r (X - \lambda_i)^{\alpha_i}$. Alors $E = \bigoplus_{i=1}^r \ker(u - \lambda_i \text{Id})^{\alpha_i}$. Sur chaque $F_i = \ker(u - \lambda_i \text{Id})^{\alpha_i}$, l'endomorphisme $u_{F_i} - \lambda_i \text{Id}_{F_i}$ est nilpotent. Il existe donc une base de F_i dans laquelle il a pour matrice une matrice diagonale par blocs formée de blocs de Jordan $J(0, d_{i,j})$. Dans cette même base, la matrice de u est donc la matrice diagonale par blocs formée des blocs de Jordan $J(\lambda_i, d_{i,j})$. D'où le résultat. $\square$

En guise de réciproque

Supposons que l'on connaisse la réduction de Jordan d'un endomorphisme. Comment calculer ses invariants de similitude ?

On note $\lambda_1, \dots, \lambda_s$ les valeurs propres de u . Pour chaque λ_i , on note $d_{i,1} \geq \dots \geq d_{i,r_i}$ les tailles des r_i blocs de Jordan de diagonale λ_i . La matrice considérée est la matrice diagonale par blocs suivante

$$\begin{pmatrix} J(\lambda_1, d_{1,1}) & & & & & \\ & \ddots & & & & 0 \\ & & J(\lambda_1, d_{1,r_1}) & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & J(\lambda_s, d_{s,1}) \\ & 0 & & & & \ddots \\ & & & & & & J(\lambda_s, d_{s,r_s}) \end{pmatrix}.$$

On remarque alors que

$$Q_1 = \pi_u = \bigvee_{i=1}^s \pi_{u_{\ker(u - \lambda_i \text{Id})^{d_{i,1}}}} = \prod_{i=1}^s (X - \lambda_i)^{d_{i,1}}.$$

Un bloc de Jordan correspond à un sous-espace cyclique engendré par le premier vecteur de la base correspondant à ce bloc. Donc, pour tout i , il existe x_i tel que le bloc de Jordan $J(\lambda_i, d_{i,1})$ corresponde au sous-espace $\langle x_i \rangle_u$. Alors, en posant $F = \langle x_1 + \dots + x_s \rangle_u$, F est un sous-espace cyclique correspondant aux blocs de Jordan $J(\lambda_1, d_{1,1}), \dots, J(\lambda_s, d_{s,1})$. On a $\pi_{u_F} = Q_1$, donc F correspond aussi à la matrice compagnon associée au polynôme Q_1 (voir la démonstration de la proposition 3.14). On note G le supplémentaire de F correspondant aux blocs restants. On recommence le processus sur u_G . On obtient

$$Q_2 = \pi_{u_G} = \prod_{r_i \geq 2} (X - \lambda_i)^{d_{i,2}}, \quad Q_3 = \prod_{r_i \geq 3} (X - \lambda_i)^{d_{i,3}}.$$

etc. On pose $r = \max(r_1, \dots, r_s)$. La famille construite convient. Par unicité, les polynômes $Q_1, \dots, Q_r$ sont les invariants de similitude de u .

Exemple. Sur un corps de caractéristique différente de 2, traitons le cas de la matrice

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

En utilisant la méthode exposée précédemment, on montre que les invariants de similitude de cette matrice sont

$$Q_1 = X^2(X-1)(X-2)^2, \quad Q_2 = (X-1)(X-2), \quad Q_3 = X-2.$$

Extension du corps de base

Soit $K \subset L$ une extension de corps. Soit $M \in M_n(K) \subset M_n(L)$.

Proposition 3.22. *Les invariants de similitude de M sont les mêmes dans $M_n(K)$ et dans $M_n(L)$.*

Démonstration. On note $Q_1, \dots, Q_r$ les invariants de similitude de M dans $M_n(K)$. Il existe alors $P \in GL_n(K)$ vérifiant

$$M = P \begin{pmatrix} \mathcal{C}(Q_1) & & 0 \\ & \ddots & \\ 0 & & \mathcal{C}(Q_r) \end{pmatrix} P^{-1}.$$

Cette égalité reste vraie dans $M_n(L)$. Par unicité, on en déduit que $Q_1, \dots, Q_r$ sont les invariants de M dans $M_n(L)$. $\square$

Corollaire 3.23. *Soient $M, N \in M_n(K)$. M et N sont semblables dans $M_n(K)$ si et seulement si elles sont semblables dans $M_n(L)$.*

Ce corollaire permet de démontrer un résultat classique.

Proposition 3.24. *Soit $M \in M_n(K)$. Alors M et ${}^tM$ sont semblables.*

Nous donnerons deux preuves de ce résultat.

Démonstration. (1) Commençons par supposer que M est un bloc de Jordan

$$M = \begin{pmatrix} \lambda & 0 & \cdots & \cdots & 0 \\ 1 & \ddots & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & \lambda \end{pmatrix} \begin{matrix} \uparrow \\ \vdots \\ n \\ \vdots \\ \downarrow \end{matrix}.$$

Alors

$${}^tM = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}.$$

Si l'on note $(e_1, \dots, e_n)$ une base et l'on appelle u l'endomorphisme dont la matrice dans la base $(e_1, \dots, e_n)$ est M . Dans la base $(e_n, \dots, e_1)$, la matrice de u est ${}^tM$. M et ${}^tM$ sont donc semblables.

Supposons π_M scindé. La matrice M admet une réduite de Jordan : il existe $P \in \text{GL}_n(\mathbb{K})$ telle que

$$M = P \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_s \end{pmatrix} P^{-1}.$$

On a

$${}^tM = {}^tP^{-1} \begin{pmatrix} {}^tJ_1 & & 0 \\ & \ddots & \\ 0 & & {}^tJ_s \end{pmatrix} {}^tP.$$

Or pour tout i , il existe Q_i inversible telle que ${}^tJ_i = Q_i J_i Q_i^{-1}$. Donc, on a

$${}^tM = {}^tP^{-1} \underbrace{\begin{pmatrix} Q_1 & & 0 \\ & \ddots & \\ 0 & & Q_s \end{pmatrix}}_Q \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_s \end{pmatrix} \underbrace{\begin{pmatrix} Q_1^{-1} & & 0 \\ & \ddots & \\ 0 & & Q_s^{-1} \end{pmatrix}}_{Q^{-1}} {}^tP.$$

Finalement,

$${}^tM = \underbrace{{}^tP^{-1} Q P^{-1}}_{S^{-1}} \underbrace{P \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_s \end{pmatrix} P^{-1}}_M \underbrace{P Q^{-1} {}^tP}_S.$$

Plaçons-nous dans le cas général. On considère une extension de $\mathbb{K}$ dans laquelle π_M est scindé. On est ramené au cas précédent. $\square$

Démonstration. (2) Tout part de la remarque suivante : E est u -cyclique si et seulement si $\pi_u = \chi_u$. Soit Q un polynôme unitaire de degré n . La matrice compagnon $\mathcal{C}(Q)$ vérifie alors $\pi_{\mathcal{C}(Q)} = \chi_{\mathcal{C}(Q)}$. Comme $\mathcal{C}(Q)$ et sa transposée ont des polynômes caractéristiques et minimaux égaux, on a $\pi_{{}^t\mathcal{C}(Q)} = \chi_{{}^t\mathcal{C}(Q)}$. Donc E est ${}^t\mathcal{C}(Q)$ -cyclique. ${}^t\mathcal{C}(Q)$ est donc semblable à la matrice compagnon associée à son polynôme minimal, c'est-à-dire Q . De cette propriété découle le résultat. $\square$

3.6 Endomorphismes semi-simples

3.6.1 Généralités

Définition 3.7. Soit $u \in \mathcal{L}(E)$. On dit que u est *simple* si $E \neq \{0\}$ et si les seuls sous-espaces vectoriels u -stables sont triviaux (i.e. sont soit $\{0\}$, soit E tout entier).

Exemple. Si $\dim E = 1$, tout endomorphisme de E est simple.

Proposition 3.25. Soit $u \in \mathcal{L}(E)$. u est simple si et seulement si E est u -cyclique et π_u est irréductible.

Démonstration. $(\Rightarrow)$ Soit $y \in E \setminus \{0\}$. $\langle y \rangle_u \not\supseteq \{0\}$ donc $\langle y \rangle_u = E$. E est donc u -cyclique.

Soit P un diviseur irréductible de π_u . Alors $\ker P(u)$ est un sous-espace vectoriel u -stable non nul. C'est donc E . Donc $\pi_u | P$, puis $P = \pi_u$ à un inversible près.

$(\Leftarrow)$ Soit F un sous-espace vectoriel u -stable non nul. Soit $y \in F \setminus \{0\}$. Soit $G = \langle y \rangle_u$. On a $\pi_{u_G} | \pi_u$. Or $G \not\supseteq \{0\}$, donc $\deg \pi_{u_G} \geq 1$. Par irréductibilité de π_u , on a $\pi_{u_G} = \pi_u$. Comme G est u -cyclique, on a $\dim G = \deg \pi_{u_G} = \deg \pi_u = \dim E$, E étant u -cyclique lui aussi. On a donc $E = G \subset F$. $\square$

Corollaire 3.26. Dans le cas où $\mathbb{K}$ est algébriquement clos, s'il existe un endomorphisme simple, E est une droite.

Définition 3.8. Soit $u \in \mathcal{L}(E)$. Un sous-espace vectoriel F est dit *u-simple* si F est non nul, F est u -stable et u_F est simple.

On déduit de la proposition 3.25 que F est u -simple si et seulement si F est u -cyclique et π_{u_F} est irréductible.

Définition 3.9. Soit $u \in \mathcal{L}(E)$. On dit que u est *semi-simple* si E se décompose comme une somme directe de sous-espaces u -simples.

Remarque. Si $\mathbb{K}$ est algébriquement clos, u est semi-simple si et seulement si il est diagonalisable. En effet, un sous-espace u -simple est une droite propre.

Théorème 3.27. Soit $u \in \mathcal{L}(E)$. u est semi-simple si et seulement si tout sous-espace u -stable admet un supplémentaire u -stable.

Démonstration. ($\Rightarrow$) On a $E = F_1 \oplus \dots \oplus F_r$, avec F_i u -simple. On remarque que si G est un sous-espace stable, soit $F_i \subset G$, soit $F_i + G = F_i \oplus G$ (puisque $F_i \cap G$ est un sous-espace stable inclus dans F_i).

On montre, par récurrence descendante sur $\dim F$ que tout sous-espace vectoriel u -stable F admet un supplémentaire u -stable. Soit F u -stable. Si $\dim F < n$, il existe i tel que F_i ne soit pas inclus dans F . Alors $F_i + F = F_i \oplus F$. On applique l'hypothèse de récurrence à $F_i \oplus F$.

($\Leftarrow$) Cette partie de la démonstration nécessite un lemme.

Lemme 3.28. Tout sous-espace vectoriel u -stable non nul contient un sous-espace vectoriel u -simple.

Démonstration. Soit $F \supsetneq \{0\}$ u -stable. Soit P un diviseur irréductible de π_{u_F} . Soit y un élément non nul de $\ker P(u) \cap F$. Alors $\langle y \rangle_u$ est un sous-espace u -simple contenu dans F (car $\pi_{u_{\langle y \rangle_u}} = P$). $\square$

Finissons la preuve du théorème. Soit F un sous-espace vectoriel de dimension maximale parmi ceux qui se décomposent en somme directe de sous-espaces u -simples. Si $F \subsetneq E$, il existe un supplémentaire u -stable G non nul de F . D'après le lemme, G contient un sous-espace u -simple, noté G_0 . Alors $F \oplus G_0$ est somme directe de sous-espaces u -simples et $F \oplus G_0 \supsetneq F$, ce qui contredit la définition de F . Donc $F = E$. $\square$

3.6.2 Polynôme minimal d'un endomorphisme semi-simple

Théorème 3.29. Soit $u \in \mathcal{L}(E)$. u est semi-simple si et seulement si le polynôme minimal de u n'est pas divisible par le carré d'un polynôme non inversible (i.e. la décomposition en irréductibles de π_u est de la forme $\prod_i P_i$, où les P_i sont irréductibles et deux à deux non associés).

Démonstration. ($\Rightarrow$) On suppose u semi-simple. Alors $E = \bigoplus_{i=1}^r F_i$, où F_i est u -simple. En particulier, $\pi_{u_{F_i}}$ est irréductible. Alors $\pi_u = \bigvee_{i=1}^r \pi_{u_{F_i}} = \prod_{i=1}^q P_i$, où les P_i sont distincts et choisis parmi les $\pi_{u_{F_i}}$.

($\Leftarrow$) On suppose que l'on a $\pi_u = \prod_{i=1}^q P_i$, où les P_i sont irréductibles et deux à deux non associés. D'après le lemme de décomposition des noyaux, on peut écrire $E = \bigoplus_{i=1}^q \ker P_i(u)$. Le polynôme minimal de la restriction de u à $\ker P_i(u)$ est P_i . D'après le théorème 3.17, $\ker P_i(u)$ se décompose en somme directe de sous-espaces u -cycliques : $\ker P_i(u) = \bigoplus_{j=1}^{r_i} F_{i,j}$. Comme P_i est irréductible, $\pi_{u_{F_{i,j}}} = P_i$. Donc $F_{i,j}$ est u -simple.

Remarque. Si l'on connaît la théorie des modules sur les anneaux principaux, on peut montrer que $\ker P_i(u)$ est un $\mathbb{K}[X]/(P_i)$ -espace vectoriel. En choisissant une base, on montre qu'il existe une décomposition $\ker P_i(u) = \bigoplus_j F_{i,j}$, avec les $F_{i,j}$ u -simples, i.e. les $F_{i,j}$ sont des $\mathbb{K}[X]/(P_i)$ -droites. $\square$

Corollaire 3.30. Soit $M \in M_n(\mathbb{K})$. Soit $\mathbb{L}$ une extension de $\mathbb{K}$ dans laquelle π_M est scindé. Si M est diagonalisable dans $M_n(\mathbb{L})$, alors M est semi-simple sur $\mathbb{K}$. La réciproque est vraie si l'extension $\mathbb{K} \subset \mathbb{L}$ est séparable.

Démonstration. Si M est diagonalisable dans $M_n(\mathbb{L})$, alors son polynôme minimal n'a que des racines simples. Donc ce polynôme ne peut être divisible par un carré non inversible dans $\mathbb{K}[X]$ (on utilise ici le fait que π_M est le même sur $\mathbb{K}$ et sur $\mathbb{L}$).

Réciproquement, si M est semi-simple sur $\mathbb{K}$, on a $\pi_M = P_1 \dots P_q$, où les P_i sont des irréductibles non associés deux à deux. Dans une extension $\mathbb{L}$ de $\mathbb{K}$ où π_M est scindé, P_i et P_j ne peuvent avoir une racine commune si $i \neq j$ (on peut invoquer le fait que P_i et P_j ont le même pgcd dans $\mathbb{K}$ et dans $\mathbb{L}$). Si l'extension est séparable, on a de plus que chaque P_i n'a que des racines simples. $\square$

Corollaire 3.31. Si u est semi-simple, pour tout sous-espace vectoriel F u -stable, u_F est semi-simple.

Démonstration. $\pi_{u_F} \mid \pi_u$. $\square$

Remarque. Si u est semi-simple, on décompose π_u en produit d'irréductibles $\pi_u = P_1 \dots P_q$. Alors $E = \bigoplus_{i=1}^q \ker P_i(u)$. Si F est u -stable, le lemme de décomposition des noyaux appliqué à u_F donne

$$F = \bigoplus_{i=1}^q \ker P_i(u_F) = \bigoplus_{i=1}^q (\ker P_i(u) \cap F).$$

Si en outre u est diagonalisable, pour tout i , on a $P_i = X - \lambda_i$. Tout sous-espace vectoriel de $\ker P_i(u)$ étant u -stable, on en déduit que F est u -stable si et seulement si $F = \bigoplus_{i=1}^q F_i$, où les F_i sont des sous-espaces vectoriels de $\ker P_i(u)$.

Proposition 3.32. Soient u et v deux endomorphismes semi-simples. Alors u et v sont semblables si et seulement si $\chi_u = \chi_v$.

Démonstration. $(\Rightarrow)$ C'est vrai même si u et v ne sont pas semi-simples.

$(\Leftarrow)$ Si on a $\chi_u = \chi_v = \prod_{i=1}^q P_i^{\alpha_i}$, où les P_i sont irréductibles deux à deux non associés, alors u et v admettent comme matrice

$$\begin{pmatrix} \mathcal{C}(P_1) & & & & \\ & \ddots & & & \\ & & \mathcal{C}(P_1) & & \\ & & & \ddots & \\ & & & & \mathcal{C}(P_2) & & \\ & & & & & \ddots & \\ & & & & & & \mathcal{C}(P_2) & & \\ & & & & & & & \ddots & \end{pmatrix}.$$

$\uparrow$
 α_1 blocs
 $\downarrow$
 $\uparrow$
 α_2 blocs
 $\downarrow$
 $\ddots$

Ils sont donc semblables. $\square$

3.6.3 Lien avec la décomposition de Dunford

Théorème 3.33. On note $\lambda_1, \dots, \lambda_r$ les racines de π_u dans une clôture algébrique de $\mathbb{K}$. On suppose l'extension $\mathbb{K} \subset \mathbb{K}(\lambda_1, \dots, \lambda_r)$ séparable. Même si π_u n'est pas scindé et si l'on ne dispose pas de $\lambda_1, \dots, \lambda_r$, on peut donner d semi-simple et n nilpotent dans $\mathbb{K}[u]$ tels que $u = d + n$ et $dn = nd$.

Démonstration. On note $\pi_u = P_1^{\alpha_1} \dots P_r^{\alpha_r}$, où les P_i sont des polynômes irréductibles unitaires distincts. On pose $Q = P_1 \dots P_r$. En caractéristique nulle, $Q = \pi_u / (\pi_u \wedge \pi'_u)$. En caractéristique positive, on conserve chaque facteur irréductible une seule fois; le simple quotient par le pgcd avec la dérivée peut perdre des facteurs. On pose $s = \max(\alpha_1, \dots, \alpha_r)$. Comme $\pi_u \mid Q^s$, on a $Q(u)^s = 0$. $Q'(u)$ est inversible, puisque que $Q' \wedge \pi_u = 1$ (par séparabilité de l'extension considérée).

On pose (ce qui n'est pas sans rappeler la méthode de Newton)

$$x_0 = u, \quad x_{n+1} = x_n - \frac{Q(x_n)}{Q'(x_n)} \quad (n \geq 0).$$

On va montrer par récurrence que cette suite est bien définie et que pour tout n , $Q(x_n) \in (Q(u)^{2^n})$. Le résultat est immédiat au rang 0. On le suppose vrai au rang n . Q' étant un polynôme, on a $Q'(x_{n+1}) - Q'(x_n) \in (x_{n+1} - x_n)$. Par définition de x_{n+1} , on a $x_{n+1} - x_n \in (Q(x_n))$. Donc $Q'(x_{n+1}) = Q'(x_n) + Q(x_n)v$, avec $v \in \mathbb{K}[u]$, est inversible, puisque somme d'un élément inversible et d'un élément nilpotent qui commutent. En outre,

$$Q(x_{n+1}) = Q\left(x_n - \frac{Q(x_n)}{Q'(x_n)}\right) = \underbrace{Q(x_n) - \frac{Q(x_n)}{Q'(x_n)}Q'(x_n)}_{=0} + \left(\frac{Q(x_n)}{Q'(x_n)}\right)^2 w \in (Q(x_n)^2).$$

D'où le résultat au rang $n + 1$.

$Q(x_n) \in (Q(u)^{2^n})$ et $Q(u)^s = 0$, donc pour n tel que $2^n \geq s$, $Q(x_n) = 0$. On note $d = x_{[\log_2 s]} \in \mathbb{K}[u]$. On a $Q(d) = 0$, donc $\pi_d \mid Q$ est sans facteur carré, donc d est semi-simple. On pose $n = u - d \in \mathbb{K}[u]$. n est nilpotent :

$$n = x_0 - x_{[\log_2 s]} = \sum_{i=0}^{[\log_2 s]-1} \underbrace{(x_i - x_{i+1})}_{\in (Q(x_i)) \Rightarrow \text{nilpotent}}.$$

Enfin, comme d et n sont des polynômes en u , ils commutent. □

3.7 Réduction des endomorphismes normaux

3.7.1 Définitions

Dans cette partie, $\mathbb{K}$ est le corps $\mathbb{R}$ ou $\mathbb{C}$. E désigne un $\mathbb{R}$ -espace vectoriel euclidien ou un $\mathbb{C}$ -espace vectoriel hermitien. On note $\langle \cdot, \cdot \rangle$ le produit scalaire, linéaire en la première variable dans le cas complexe.

Soit $u \in \mathcal{L}(E)$. On rappelle qu'il existe un unique endomorphisme de E , appelé adjoint de u et noté u^* vérifiant $\forall x, y \in E, \langle u(x), y \rangle = \langle x, u^*(y) \rangle$. Si $\mathcal{B} = (e_1, \dots, e_n)$ désigne une base orthonormée de E et si l'on définit la famille $(a_{i,j})$ par $u(e_j) = \sum_i a_{i,j} e_i$, alors $a_{i,j} = \langle u(e_j), e_i \rangle = \langle e_j, u^*(e_i) \rangle$. On en déduit que la matrice de u^* dans la base $\mathcal{B}$, i.e. l'adjointe de la matrice M de u sur $\mathcal{B}$, est $M^* = {}^t \bar{M}$.

Définition 3.10. Un endomorphisme u est dit *normal* si $uu^* = u^*u$.

Exemples. • Si $u = u^*$, on dit que u est auto-adjoint. Dans le cas réel, il est dit symétrique. Dans le cas complexe, il est dit hermitien.

- Si $u = -u^*$, on dit, dans le cas réel, que u est antisymétrique et, dans le cas complexe, que u est antihermitien.
- Si $u^* = u^{-1}$, on dit, dans le cas réel, que u est orthogonal et, dans le cas complexe, que u est unitaire.

3.7.2 Réduction

Proposition 3.34. Soit u un endomorphisme normal. Soit F un sous-espace u -stable. Alors $F^\perp$ est u -stable. En outre, u_F est un endomorphisme normal de F et $(u_F)^* = (u^*)_F$.

Remarque. On a $E = F \oplus F^\perp$. Donc, F admet un supplémentaire u -stable : son orthogonal. En particulier, u est semi-simple.

Démonstration. $E = F \oplus F^\perp$. Sur une base orthonormée adaptée à cette décomposition, la matrice de u s'écrit

$$M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \begin{matrix} \} F \\ \} F^\perp \end{matrix}.$$

Comme cette base est orthonormée, la matrice de u^* est

$$M^* = \begin{pmatrix} A^* & 0 \\ B^* & C^* \end{pmatrix}.$$

Comme $uu^* = u^*u$, on a $MM^* = M^*M$. Or

$$MM^* = \begin{pmatrix} AA^* + BB^* & BC^* \\ CB^* & CC^* \end{pmatrix}, \quad M^*M = \begin{pmatrix} A^*A & A^*B \\ B^*A & B^*B + C^*C \end{pmatrix}.$$

On a donc $AA^* + BB^* = A^*A$, puis en prenant la trace, $\text{Tr}(BB^*) = 0$, d'où $B = 0$. On a aussi $A^*A = AA^*$ et $C^*C = CC^*$. On en déduit donc que $F^\perp$ est u -stable (car $B = 0$), que $(u_F)^* = (u^*)_F$ et que u_F est normal (car $A^*A = AA^*$). $\square$

Théorème 3.35. Soit $u \in \mathcal{L}(E)$. u est normal si et seulement si il admet sur une base orthonormée une matrice de la forme :

- si $\mathbb{K} = \mathbb{C}$,

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

- si $\mathbb{K} = \mathbb{R}$,

$$\begin{pmatrix} \lambda_1 & & & & & & 0 \\ & \ddots & & & & & \\ & & \lambda_r & & & & \\ & & & a_1 & -b_1 & & \\ & & & b_1 & a_1 & & \\ & & & & & \ddots & \\ 0 & & & & & & a_s & -b_s \\ & & & & & & b_s & a_s \end{pmatrix}.$$

Démonstration. On suppose $\mathbb{K} = \mathbb{C}$. ($\Leftarrow$) u est normal si et seulement si sa matrice dans une base orthonormée est normale. On constate qu'une matrice diagonale est normale. On en déduit que u est normal.

($\Rightarrow$) Soit e un vecteur propre normé de u . On a $E = \mathbb{C}e \oplus (\mathbb{C}e)^\perp$ et on sait que $u_{(\mathbb{C}e)^\perp}$ est un endomorphisme normal. On conclut par récurrence sur $\dim E$.

On suppose $\mathbb{K} = \mathbb{R}$. ($\Leftarrow$) Une matrice diagonale est normale et une matrice de similitude directe aussi, puisque

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = (a^2 + b^2)I_2.$$

($\Rightarrow$) Cette partie de la démonstration nécessite quelques lemmes.

Lemme 3.36. Soit u un endomorphisme d'un $\mathbb{R}$ -espace vectoriel euclidien E . Alors u admet une droite stable ou un plan stable.

Démonstration. Si π_u admet une racine réelle, alors u admet une droite stable.

Sinon, π_u admet un diviseur irréductible de degré 2. Soit P un tel diviseur. Alors pour tous $x \in \ker P(u) \setminus \{0\}$, $\langle x \rangle_u$ est un sous-espace u -cyclique sur lequel le polynôme minimal de u est P , c'est donc un plan. $\square$

On remarque que si l'on est capable de majorer le degré des irréductibles, on peut majorer la dimension des sous-espaces stables.

Lemme 3.37. Soit $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$. M est normale si et seulement si soit $b = c$ (dans ce cas, M est diagonalisable), soit $b = -c$ et $a = d$.

Démonstration. Calculons

$$M^t M = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix}, \quad {}^t M M = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix}.$$

Donc M est normale si et seulement si

$$(b^2 = c^2 \text{ et } (a - d)(b - c) = 0) \iff (b = c \text{ ou } (b = -c \text{ et } a = d)).$$

Dans le cas où $b = c$, on a $\chi_M = X^2 - (a + d)X + ad - b^2$. Le discriminant s'écrit $\Delta = (a - d)^2 + 4b^2 > 0$, sauf si $a = d$ et $b = 0$, qui constitue un cas trivial. M est donc diagonalisable. $\square$

Remarque. Un endomorphisme symétrique admet donc une valeur propre. On peut en déduire, par récurrence, qu'un tel endomorphisme est diagonalisable.

Montrons le résultat par récurrence sur $\dim E$.

Si u admet une valeur propre réelle, on note e un vecteur propre normé de u associé à cette valeur propre. On a $E = \mathbb{R}e \oplus (\mathbb{R}e)^\perp$. On sait alors que $u|_{(\mathbb{R}e)^\perp}$ est normal et on lui applique l'hypothèse de récurrence.

Sinon, u admet un plan stable $\mathcal{P}$. $u|_{\mathcal{P}}$ est normal et sans valeur propre, donc il admet sur une base orthonormée de $\mathcal{P}$ une matrice de la forme

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

On a $E = \mathcal{P} \oplus \mathcal{P}^\perp$. $u|_{\mathcal{P}^\perp}$ est normal et on lui applique l'hypothèse de récurrence. $\square$

Corollaire 3.38 (Etude de cas particuliers). Du théorème de réduction des endomorphismes normaux, on peut déduire que :

- u est hermitien si et seulement si il admet sur une base orthonormée une matrice de la forme

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \text{ avec } \lambda_j \in \mathbb{R}.$$

- u est antihermitien si et seulement si il admet sur une base orthonormée une matrice de la forme

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \text{ avec } \lambda_j \in i\mathbb{R}.$$

- u est unitaire si et seulement si il admet sur une base orthonormée une matrice de la forme

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \text{ avec } |\lambda_j| = 1.$$

- u est symétrique si et seulement si il admet sur une base orthonormée une matrice de la forme

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \text{ avec } \lambda_j \in \mathbb{R}.$$

- u est antisymétrique si et seulement si il admet sur une base orthonormée une matrice de la forme

$$\begin{pmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & & 0 & -b_1 & \\ & & & b_1 & 0 & \\ & & & & & \ddots \\ & 0 & & & & 0 & -b_s \\ & & & & & b_s & 0 \end{pmatrix}, \text{ avec } b_j \neq 0.$$

- u est orthogonal si et seulement si il admet sur une base orthonormée une matrice de la forme

$$\begin{pmatrix} \pm 1 & & & & & \\ & \ddots & & & & \\ & & \pm 1 & & & \\ & & & a_1 & -b_1 & \\ & & & b_1 & a_1 & \\ & & & & & \ddots \\ & 0 & & & & a_s & -b_s \\ & & & & & b_s & a_s \end{pmatrix}, \text{ avec } a_j^2 + b_j^2 = 1.$$

Remarque. Les endomorphismes semi-simples sont tous des endomorphismes normaux pour un certain produit scalaire.

3.7.3 Applications en topologie

On note $U_n(\mathbb{C}) = \{M : M^*M = I_n\}$ le groupe unitaire, $O_n(\mathbb{R}) = \{M : {}^tMM = I_n\}$ le groupe orthogonal et $SO_n(\mathbb{R})$ son sous-groupe de déterminant 1.

Connexité de $U_n(\mathbb{C})$ et $SO_n(\mathbb{R})$

Les résultats énoncés dans cette partie peuvent se retrouver dans [10].

Proposition 3.39. $U_n(\mathbb{C})$ est connexe. $O_n(\mathbb{R})$ a deux composantes connexes : $SO_n(\mathbb{R})$ et $O_n(\mathbb{R}) \setminus SO_n(\mathbb{R})$.

Démonstration. On note diag la matrice diagonale, ou diagonale par blocs, dont les arguments sont les blocs successifs. Si $M \in U_n(\mathbb{C})$, le théorème spectral donne

$$M = P^* \text{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})P, \quad P \in U_n(\mathbb{C}).$$

Le chemin $t \mapsto P^* \text{diag}(e^{it\theta_1}, \dots, e^{it\theta_n})P$, pour $0 \leq t \leq 1$, relie I_n à M dans $U_n(\mathbb{C})$.

Posons $R(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$. Pour $M \in SO_n(\mathbb{R})$, la réduction orthogonale réelle donne des blocs 1, -1 et $R(\theta)$. Le nombre de blocs -1 est pair puisque $\det M = 1$; on les regroupe deux par deux en blocs $R(\pi)$. Ainsi,

$$M = {}^tP \text{diag}(I_p, R(\theta_1), \dots, R(\theta_r))P, \quad P \in O_n(\mathbb{R}), \quad p + 2r = n.$$

Le chemin $t \mapsto {}^tP \text{diag}(I_p, R(t\theta_1), \dots, R(t\theta_r))P$ relie I_n à M dans $SO_n(\mathbb{R})$. Ces deux groupes sont donc connexes par arcs.

Si $\sigma \in O_n(\mathbb{R}) \setminus SO_n(\mathbb{R})$, alors

$$SO_n(\mathbb{R}) \rightarrow O_n(\mathbb{R}) \setminus SO_n(\mathbb{R}), \quad M \mapsto \sigma M$$

est un homéomorphisme donc $O_n(\mathbb{R}) \setminus SO_n(\mathbb{R})$ est aussi connexe. Comme le déterminant est continu et que $\det SO_n(\mathbb{R}) = \{1\}$ et $\det O_n(\mathbb{R}) \setminus SO_n(\mathbb{R}) = \{-1\}$, $SO_n(\mathbb{R})$ et $O_n(\mathbb{R}) \setminus SO_n(\mathbb{R})$ ne peuvent être dans un même composante connexe. $\square$

Remarque. Dans un groupe topologique, la composante connexe à laquelle appartient 1 est un sous-groupe distingué. Les autres composantes connexes sont les classes modulo ce sous-groupe.

Proposition 3.40. $U_n(\mathbb{C})$ et $O_n(\mathbb{R})$ sont compacts.

Démonstration. Celle-ci est laissée en exercice. □

Décomposition polaire

Proposition 3.41. On note $S_n^+(\mathbb{R})$ l'ensemble des matrices symétriques positives (au sens large), et $S_n^{++}(\mathbb{R})$ celui des matrices symétriques définies positives. On a un homéomorphisme

$$\varphi : O_n(\mathbb{R}) \times S_n^{++}(\mathbb{R}) \rightarrow GL_n(\mathbb{R}), \quad (O, S) \mapsto OS.$$

Démonstration. φ est continue (en tant que restriction de l'application bilinéaire continue $M_n(\mathbb{R}) \times M_n(\mathbb{R}) \rightarrow M_n(\mathbb{R}) : (A, B) \mapsto AB$).

Soit $M \in GL_n(\mathbb{R})$. Supposons que l'on puisse écrire $M = OS$. Alors ${}^tMM = {}^tS^tOOS = S^2$.

Lemme 3.42. ${}^tMM \in S_n^{++}(\mathbb{R})$.

Proposition 3.43. Soit $N \in S_n^{++}(\mathbb{R})$. Il existe un unique $S \in S_n^{++}(\mathbb{R})$ tel que $N = S^2$.

Démonstration. Traitons l'existence de S . Il existe $P \in O_n(\mathbb{R})$ et $\lambda_1, \dots, \lambda_n > 0$ tels que

$$N = {}^tP \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} P.$$

On peut écrire $N = S^2$ avec

$$S = {}^tP \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} P.$$

Intéressons-nous maintenant à l'unicité de S .

- *Première méthode.* Si $N = S^2$, on note (λ_i) la famille des valeurs propres distinctes de N . D'après le lemme de décomposition des noyaux, on a

$$\ker(N - \lambda_i I_n) = \ker(S - \sqrt{\lambda_i} I_n) \oplus \underbrace{\ker(S + \sqrt{\lambda_i} I_n)}_{=\{0\}, \text{ sinon } S \text{ aurait une valeur propre négative}}.$$

Or $E = \bigoplus_i \ker(N - \lambda_i I_n)$. Donc S est unique.

- *Deuxième méthode.* Il existe un polynôme Q (polynôme d'interpolation de Lagrange) tel que pour tout i , $Q(\lambda_i) = \sqrt{\lambda_i}$. On en déduit que la matrice S construite précédemment vérifie $S = Q(N)$. Soit T une autre solution : $T^2 = N = S^2$. Ainsi S est un polynôme en $N = T^2$, donc S est un polynôme en T . Donc $ST = TS$. On peut alors écrire $(S - T)(S + T) = 0$. $S + T$ est la somme de deux matrices définies positives, donc est définie positive. $S + T$ est donc inversible. Finalement, $S = T$. □

Revenons à la preuve du théorème. Si $M = OS$, on a $S^2 = {}^tMM$. S est nécessairement l'unique racine carrée définie positive de ${}^tMM$, donc est unique. On pose alors $O = MS^{-1}$. On a ${}^tOO = {}^tS^{-1} {}^tMMS^{-1} = S^{-1}S^2S^{-1} = I_n$. Donc $O \in O_n(\mathbb{R})$. On a ainsi la bijectivité de φ .

Montrons que φ^{-1} est continue. Soit (M_i) une suite de matrices de $GL_n(\mathbb{R})$ qui converge vers une matrice $M \in GL_n(\mathbb{R})$. Soient $M_i = O_i S_i$ et $M = OS$ les décompositions polaires correspondantes.

Lemme 3.44. O est l'unique valeur d'adhérence de la suite (O_i) .

Démonstration. Soit U une valeur d'adhérence de (O_i) . Soit ψ telle que $O_{\psi(i)} \rightarrow U$. Comme $M = \lim_i M_i = \lim_i O_i S_i$, on en déduit que $S_{\psi(i)}$ converge vers ${}^t U O S$. Comme $S_{\psi(i)} \in S_n^{++}(\mathbb{R})$, on a ${}^t U O S \in S_n^+(\mathbb{R})$. Or ${}^t U O S \in \text{GL}_n(\mathbb{R})$, donc ${}^t U O S \in S_n^{++}(\mathbb{R})$. L'unicité de la décomposition polaire d'une telle matrice montre que ${}^t U O = I_n$, i.e. $U = O$. $\square$

Comme $O_n(\mathbb{R})$ est compact, on en déduit que la suite (O_i) converge vers O et donc (S_i) converge vers S . $\square$

Corollaire 3.45. $\text{GL}_n(\mathbb{R})$ a deux composantes connexes : $\text{GL}_n^+(\mathbb{R}) = \{M : \det M > 0\}$ et $\text{GL}_n^-(\mathbb{R}) = \{M : \det M < 0\}$.

Démonstration. $S_n^{++}(\mathbb{R})$ est convexe, donc connexe. (On peut même montrer, grâce à l'exponentielle, que $S_n^{++}(\mathbb{R})$ est homéomorphe à $\mathbb{R}^{\frac{n(n+1)}{2}}$.) $\square$

Corollaire 3.46. $O_n(\mathbb{R})$ est un sous-groupe compact maximal de $\text{GL}_n(\mathbb{R})$.

Démonstration. Soit $G \supset O_n(\mathbb{R})$ un sous-groupe compact de $\text{GL}_n(\mathbb{R})$. Soit $M \in G$. Soit $M = OS$ sa décomposition polaire. On a $S \in G$. Donc, pour tout i , $S^i \in G$. G est compact donc la suite $(\|S^i\|)$ est bornée et les valeurs propres de S sont ≤ 1 . S'il existait une valeur propre < 1 , la limite d'une sous-suite convergente de S^i ne serait pas inversible. Donc $S = I_n$. $\square$

Modules sur un anneau principal

Références : [5, 8, 11]

4.1 Généralités sur les modules

Soit A un anneau intègre commutatif. On note $\text{Hom}_A(M, N)$ l'ensemble des applications A -linéaires entre deux A -modules M, N , et $\delta_{i,j}$ le symbole de Kronecker, égal à 1 si $i = j$ et à 0 sinon.

Définition 4.1. Un A -module est la donnée de $(M, +, \cdot)$ tel que

- $(M, +)$ soit un groupe abélien additif,
 - $A \times M \rightarrow M$ soit une loi externe,
- vérifiant
- $\forall a \in A, \forall m, m' \in M, a(m + m') = am + am'$.
 - $\forall a, a' \in A, \forall m \in M, (aa')m = a(a'm)$.
 - $\forall a, a' \in A, \forall m \in M, (a + a')m = am + a'm$.
 - $\forall m \in M, 1m = m$.

Il est élémentaire de définir les notions de sous-module, de A -morphisme, de quotient, de somme directe, de noyau et d'image à partir de ce que l'on connaît des espaces vectoriels.

Définition 4.2. Soit M un A -module et X une partie de M . Il existe un plus petit sous-module de M qui contient X . Il s'agit de l'intersection de tous les sous-modules de M qui contiennent X . On l'appelle *sous-module engendré par X* et on le note $\langle X \rangle$.

Proposition 4.1.

$$\langle X \rangle = \left\{ \sum_{i=1}^k a_i x_i, k \in \mathbb{N}, a_i \in A, x_i \in X \right\}.$$

Définition 4.3. Si X vérifie $\langle X \rangle = M$, on dit que X est une *partie génératrice* de M .

Remarque. Un module est dit *de type fini* s'il admet une partie génératrice finie. Un A -module de type fini peut avoir des sous-modules qui ne sont pas de type fini.

Exemple. On considère le module $A = A \cdot 1$. Les sous-modules de A sont les idéaux de A et réciproquement.

Définition 4.4. Un anneau A est dit *noetherien* si tous ses idéaux sont de type fini.

Théorème 4.2. Si A est noetherien, un A -module de type fini a tous ses sous-modules de type fini.

Exemple. Si $\mathbb{K}$ est un corps, $\mathbb{K}[X_i, i \in \mathbb{N}]$ est un anneau non noetherien, puisque l'idéal engendré par $(X_i, i \in \mathbb{N})$ n'est pas de type fini.

Définition 4.5. • Une partie X de M est dite *libre* si pour toute famille presque nulle (λ_x) , telle que $\sum_{x \in X} \lambda_x x = 0$, on a pour tout x , $\lambda_x = 0$.

- Si une partie X de M est à la fois libre et génératrice, on dit que c'est une *base* de M .
- On dit que M est un *module libre* si il admet une base. Dans ce cas, si X est une base de M , on obtient un isomorphisme de A -modules

$$A^{(X)} \rightarrow M, \quad (\lambda_x)_{x \in X} \mapsto \sum_{x \in X} \lambda_x x.$$

Ici $A^{(X)}$ est l'ensemble des familles à support fini (tous les coefficients sauf un nombre fini sont nuls). Dans le cas où X est finie, on démontre que toutes les bases ont le même cardinal, que l'on appelle *rang* du module.

Démonstration. La démonstration de l'existence du rang est basée sur le théorème suivant.

Théorème 4.3 (Krull). *Tout anneau commutatif unitaire non nul admet un idéal maximal (donc un quotient qui est un corps).*

On suppose avoir deux bases : une de cardinal n et l'autre de cardinal m . On a donc $A^n \simeq A^m$. On note $\mathcal{M}$ un idéal maximal de A . On a donc $A^n / \mathcal{M} A^n \simeq A^m / \mathcal{M} A^m$, puis $(A / \mathcal{M})^n \simeq (A / \mathcal{M})^m$, ce qui impose $n = m$. $\square$

Proposition 4.4. *Soit X une base de A^n . Les bases de A^n sont les PX avec $P \in \text{GL}_n(A)$, où $\text{GL}_n(A) = \{P \in M_n(A) : \exists Q \in M_n(A) : PQ = I_n\} = \{P \in M_n(A) : \det P \in A^*\}$.*

Démonstration. $(\Rightarrow)$ Si P admet une inverse Q dans $M_n(A)$, $PQ = I_n$, donc $\det P \det Q = 1$, puis $\det P \in A^*$.

$(\Leftarrow)$ $P \text{ adj}(P) = (\det P) I_n$. Si $\det P \in A^*$, P a pour inverse $(\det P)^{-1} \text{adj}(P)$ qui appartient à $M_n(A)$. $\square$

Définition 4.6. On suppose A intègre. Un élément $m \in M$ est de *torsion* s'il existe $a \neq 0$ tel que $am = 0$, i.e. $\{m\}$ n'est pas libre.

On appelle *sous-module de torsion* de M le sous-ensemble formé par ses éléments de torsion. On le note M_T ou $\text{Tor } M$.

Démonstration. Soient $m, m' \in M_T$. Il existe $a \neq 0$ et $a' \neq 0$ tels que $am = 0$ et $a'm' = 0$. On a $aa'(m + m') = 0$ et $aa' \neq 0$. Soit maintenant $b \in A$. On a $a(bm) = bam = 0$ avec $a \neq 0$, donc $bm \in M_T$. $\square$

Proposition 4.5. • Si M est un A -module libre, $M_T = \{0\}$.

- $(M/M_T)_T = \{0\}$.

Démonstration. Soit X une base de M . Soit $m \in M_T$. Il existe une famille $(\lambda_x) \in A^{(X)}$ telle que $m = \sum_x \lambda_x x$. Il existe $a \neq 0$ tel que $am = 0$. On a $\sum_x a\lambda_x x = 0$. Donc, pour tout x , $a\lambda_x = 0$, puis $\lambda_x = 0$.

Soit $m \in M$ et $\bar{m}$ sa classe dans M/M_T . Si $\bar{m} \in (M/M_T)_T$, il existe $a \neq 0$ tel que $a\bar{m} = 0$, donc $am \in M_T$. Il existe $b \neq 0$ tel que $b(am) = 0$. On a donc $m \in M_T$, puis $\bar{m} = 0$. $\square$

4.2 Classes d'isomorphisme des modules de type fini sur un anneau principal

Soit A un anneau principal.

Définition 4.7. Soit M un A -module. On appelle *idéal annulateur* de M l'idéal $\text{Ann } M = \{a \in A : aM = 0\}$.

Proposition 4.6. Si M est de type fini, on a l'équivalence $M = M_T \iff \text{Ann } M \neq \{0\}$.

Théorème 4.7 (Classes d'isomorphisme de A -modules). Soit M un A -module de type fini. Alors il existe deux entiers naturels uniques r et s et une suite d'éléments de $A \setminus (A^* \cup \{0\})$ $a_s \mid \dots \mid a_1$ unique aux inversibles près tels que

$$M \simeq A/(a_1) \oplus \dots \oplus A/(a_s) \oplus A^r.$$

Ceci décrit les classes d'isomorphisme de A -modules.

Théorème 4.8 (De la base adaptée). Soit M un A -module libre de rang m et soit N un sous-module de M . Alors N est un A -module libre de rang $n \leq m$ et, plus précisément, il existe une base $\{e_1, \dots, e_m\}$ de M et des éléments non nuls $a_n \mid \dots \mid a_1$ de A tels que $(a_1 e_1, \dots, a_n e_n)$ soit une base de N . Les $a_1, \dots, a_n$ sont alors uniques aux inversibles près.

Remarques. Avec les notations du théorème, on a $M/N \simeq A/(a_1) \oplus \dots \oplus A/(a_n) \oplus A^{m-n}$. On a le diagramme suivant

$$\begin{array}{ccc} A^m & \xrightarrow{\quad} & A/(a_1) \oplus \dots \oplus A/(a_n) \oplus A^{m-n} \\ \uparrow \sim & \nearrow \varphi & \uparrow \sim \\ M & \xrightarrow{\quad} & M/N \end{array}$$

φ est un morphisme surjectif de noyau N . Par factorisation de ce morphisme, on a le résultat.

Faire le quotient de M par N revient à quotienter composante par composante, d'où le nom de "base adaptée".

On obtient que l'unicité dans le théorème 4.7 implique l'unicité dans le théorème 4.8. En effet supposons avoir des éléments non nuls $a_n \mid \dots \mid a_1$ de A tels que $(a_1 e_1, \dots, a_n e_n)$ soit une base de N . Soit ℓ défini par a_i inversible pour tout $i \geq \ell$ et a_i non inversible pour tout $i < \ell$. On a, d'après ce qui précède, $M/N \simeq A/(a_1) \oplus \dots \oplus A/(a_{\ell-1}) \oplus A^{m-n}$. L'unicité dans le théorème 4.7 assure que $m - n$ est unique (donc n est unique) et l'unicité aux inversibles près de $a_1, \dots, a_{\ell-1}$. En outre, pour $i \geq \ell$, $a_i = 1$ aux inversibles près. On en déduit l'unicité dans le théorème 4.8.

Soit M un A -module de type fini et $\{m_1, \dots, m_n\}$ une partie génératrice de M . Alors on définit un morphisme surjectif

$$\varphi : A^n \twoheadrightarrow M, \quad (a_1, \dots, a_n) \mapsto \sum_{i=1}^n a_i m_i.$$

On a donc $M \simeq A^n / \ker \varphi$.

Supposons montrée l'existence dans le théorème 4.8 : il existe $(e_1, \dots, e_n)$ base de A^n , il existe $m \leq n$ et il existe des éléments non nuls $a_m \mid \dots \mid a_1$ de A tels que $(a_1 e_1, \dots, a_m e_m)$ soit une base de $\ker \varphi$. On a, d'après notre remarque précédente, $A^n / \ker \varphi \simeq A/(a_1) \oplus \dots \oplus A/(a_m) \oplus A^{n-m}$, d'où $M \simeq A/(a_1) \oplus \dots \oplus A/(a_m) \oplus A^{n-m}$. On retire les a_i inversibles et on obtient une décomposition de M telle qu'elle est décrite dans le théorème 4.7.

Démonstration. Nous allons commencer par démontrer l'existence dans le théorème 4.8.

Lemme 4.9. Soit M un A -module de type fini. Soit α une forme linéaire (i.e. un morphisme $M \rightarrow A$) non nulle. Alors $\alpha(M)$ est un idéal principal non nul de A . Si $m_0 \in M$ vérifie $\alpha(M) = (\alpha(m_0))$, on a $M = Am_0 \oplus \ker \alpha$.

Démonstration. Soit $m \in M$. $\alpha(m) \in (\alpha(m_0))$, donc il existe $a \in A$ tel que $\alpha(m) = a\alpha(m_0)$. On a $m = \underbrace{am_0}_{\in Am_0} + \underbrace{(m - am_0)}_{\in \ker \alpha}$. Donc $M = Am_0 + \ker \alpha$.

Supposons $am_0 \in \ker \alpha$. On a $0 = \alpha(am_0) = \underbrace{a}_{\neq 0} \alpha(m_0)$, donc $a = 0$, puis $am_0 = 0$. Finalement, $M = Am_0 \oplus \ker \alpha$. $\square$

Lemme 4.10. Soit M un A -module libre de rang m . Soit α une forme linéaire non nulle de M . Alors $\ker \alpha$ est un A -module libre de rang $m - 1$.

Démonstration. Le résultat se démontre par récurrence sur $m \geq 1$ (une forme non nulle ne peut exister sur le module nul). Si $m = 1$, écrivons $M = Ae_1$. Comme $\alpha(e_1) \neq 0$ et A est intègre, $\alpha(ae_1) = 0$ impose $a = 0$: le noyau est donc libre de rang 0. Supposons maintenant $m \geq 2$ et le résultat établi aux rangs inférieurs. On note $(e_1, \dots, e_m)$ une base de M et $(e_1^*, \dots, e_m^*)$ sa base duale. Alors $\ker e_i^* = \bigoplus_{j \neq i} Ae_j$.

On prend m_0 dans M tel que $\alpha(M) = (\alpha(m_0))$, si bien que l'on a $M = Am_0 \oplus \ker \alpha$.

Si $\ker \alpha = \{0\}$, la décomposition précédente donnerait $M = Am_0 \simeq A$, contrairement à $m \geq 2$ et à l'invariance du rang. Le noyau est donc non nul, et il existe i tel que $e_i^*(\ker \alpha) \neq \{0\}$. $e_{i|_{\ker \alpha}}^*$ est donc une forme linéaire non nulle. On lui applique le lemme précédent : il existe $m_1 \in \ker \alpha$ tel que $\ker \alpha = Am_1 \oplus \ker e_{i|_{\ker \alpha}}^*$. Ainsi, on a $M = Am_0 \oplus Am_1 \oplus (\ker e_i^* \cap \ker \alpha)$.

$\ker e_i^* \cap \ker \alpha$ est le noyau de la forme linéaire $\alpha|_{\ker e_i^*}$ définie sur $\ker e_i^*$, qui est un module libre de rang $m - 1$. Si cette restriction est nulle, son noyau est $\ker e_i^*$, libre de rang $m - 1$; sinon, l'hypothèse de récurrence donne un noyau libre de rang $m - 2$. Le premier cas est exclu par l'égalité des rangs dans la somme directe suivante. Or $M = Am_0 \oplus Am_1 \oplus (\ker e_i^* \cap \ker \alpha)$. $\ker \alpha = Am_1 \oplus (\ker e_i^* \cap \ker \alpha)$ est donc un module libre de rang $m - 1$.

On remarque que $Am_1 \simeq A$, car m_1 n'est pas de torsion, puisque $(\ker \alpha)_T \subset M_T = \{0\}$, car M libre. $\square$

Revenons à la démonstration de l'existence dans le théorème 4.8. Soit M un module libre de rang m et N un sous-module de M . On pose $\mathcal{F} = \{\alpha(N), \alpha \in \text{Hom}_A(M, A)\}$. $\mathcal{F}$ est un ensemble d'idéaux de A . On en cherche un maximal (c'est possible, car A est noethérien, puisque principal ; en outre, on peut identifier un tel idéal maximal : on associe à tout idéal I non nul de $\mathcal{F}$ un entier n_I qui est le nombre d'irréductibles comptés avec leurs multiplicités d'une décomposition en produit d'irréductibles d'un générateur de I , on choisit alors I tel que n_I soit minimal). Soit donc I un idéal maximal pour l'inclusion dans $\mathcal{F}$. Soient $\alpha \in \text{Hom}_A(M, A)$ tel que $I = \alpha(N)$, $b \in A$ tel que $I = (b)$ et $f \in N$ tel que $b = \alpha(f)$. On peut supposer $N \neq \{0\}$. Dans ce cas, $\mathcal{F} \neq \{0\}$ et $b \neq 0$.

Lemme 4.11. Pour tout $\beta \in \text{Hom}_A(M, A)$, $b \mid \beta(f)$.

Démonstration. Soit $d = b \wedge \beta(f)$. Comme A est principal, on a une relation de Bézout : il existe $x, y \in A$ tels que $xb + y\beta(f) = d$. On a $(x\alpha + y\beta)(f) = d$. Alors $(x\alpha + y\beta)(N) \supset (d) \supset (b) = \alpha(N)$. Comme $\alpha(N)$ est maximal dans $\mathcal{F}$, on en déduit qu'il y a égalité partout, donc $(d) = (b)$, puis $b \mid \beta(f)$. $\square$

Lemme 4.12. Il existe $e \in M$ tel que $\alpha(e) = 1$ et $f = be$.

Démonstration. Soit $(e_1, \dots, e_m)$ une base de M et $(e_1^*, \dots, e_m^*)$ sa base duale. On a $f = \sum_{i=1}^m e_i^*(f)e_i$ et pour tout i , $b \mid e_i^*(f)$. Il existe donc $c_1, \dots, c_m$ dans A tels que pour tout i , $bc_i = e_i^*(f)$. Alors on a $f = b \sum_{i=1}^m c_i e_i$, puis $f = be$ en posant $e = \sum_{i=1}^m c_i e_i$. On en déduit $\alpha(e) = 1$. $\square$

Reprenons la démonstration du théorème. On a $\alpha(N) = (b) = (\alpha(f))$. On en déduit que $N = Af \oplus (\ker \alpha \cap N)$. Comme $\alpha(e) = 1$, on a $\alpha(M) \supset (\alpha(e)) = A$, d'où $\alpha(M) = (\alpha(e))$, puis $M = Ae \oplus \ker \alpha$.

On sait que $\ker \alpha$ est un A -module libre de rang $m - 1$. On suppose par récurrence qu'il existe une base adaptée à son sous-module $\ker \alpha \cap N$: il existe $(e'_1, \dots, e'_{m-1})$ une base de $\ker \alpha$, il existe $n \leq m$ et des éléments non nuls $a_{n-1} \mid \dots \mid a_1$ tels que $(a_1 e'_1, \dots, a_{n-1} e'_{n-1})$ soit une base de $\ker \alpha \cap N$.

On pose $e_1 = e'_1, \dots, e_{n-1} = e'_{n-1}, e_n = e$ et $e_i = e'_{i-1}$ pour tout $i > n$. Comme $M = Ae \oplus \ker \alpha$, $(e_1, \dots, e_m)$ est une base de M .

On pose $a_n = b$. Comme $N = Abe \oplus (\ker \alpha \cap N)$, $(a_1 e_1, \dots, a_n e_n)$ est une base de N . En particulier, N est libre, de rang n .

Si $n = 1$, la condition de divisibilité est vide. Pour $n \geq 2$, on remarque que $\alpha(a_{n-1} e_{n-1} + a_n e_n) = a_n = b$ et $a_{n-1} e_{n-1} + a_n e_n \in N$. Le lemme de divisibilité s'applique à tout élément f de N vérifiant $\alpha(f) = b$. On en déduit que $b \mid \beta(a_{n-1} e_{n-1} + a_n e_n)$, pour toute forme linéaire β . En particulier, pour $\beta = e_{n-1}^*$, on a $a_n \mid a_{n-1}$.

Démontrons maintenant l'unicité dans le théorème 4.7. Supposons pouvoir écrire $M \simeq A/(a_1) \oplus \dots \oplus A/(a_s) \oplus A^r$, avec $a_s \mid \dots \mid a_1$, et $M \simeq A/(b_1) \oplus \dots \oplus A/(b_t) \oplus A^q$, avec $b_t \mid \dots \mid b_1$. Montrons que $r = q$, $s = t$ et pour tout i , $a_i = b_i$ aux inversibles près.

Lemme 4.13. $(A/(a_1) \oplus \dots \oplus A/(a_s) \oplus A^r)_T = A/(a_1) \oplus \dots \oplus A/(a_s)$.

Si le sous-module de torsion est nul, les deux familles de facteurs sont vides. Dans le cas contraire, $s, t \geq 1$. Remarquons que $\text{Ann}(A/(a)) = (a)$. On a donc $\text{Ann}(\bigoplus_i A/(a_i)) = \bigcap_i (a_i) = (a_1)$. Du lemme, on déduit $A/(a_1) \oplus \dots \oplus A/(a_s) \simeq A/(b_1) \oplus \dots \oplus A/(b_t)$. En considérant l'idéal annulateur, on a $a_1 = b_1$.

Lemme 4.14. Si $M \simeq A/(a_1) \oplus \dots \oplus A/(a_s) \oplus A^r$, alors $M/M_T \simeq A^r$.

Démonstration. On a un isomorphisme $\varphi : M \xrightarrow{\sim} A/(a_1) \oplus \dots \oplus A/(a_s) \oplus A^r$. Donc $\varphi(M_T) = \varphi(M)_T = A/(a_1) \oplus \dots \oplus A/(a_s)$. Posons $T = A/(a_1) \oplus \dots \oplus A/(a_s)$. Par passage au quotient, on a

$$\begin{array}{ccc} M & \xrightarrow[\sim]{\varphi} & T \oplus A^r \\ \downarrow & & \downarrow \\ M/M_T & \xrightarrow[\sim]{\bar{\varphi}} & (T \oplus A^r)/T \simeq A^r \end{array}$$

□

On déduit directement de ce lemme que $r = q$ (c'est le rang du module libre M/M_T).

Remarque. On a obtenu que M est libre si et seulement $M_T = \{0\}$.

Lemme 4.15. Soit $b \in A \setminus \{0\}$. Soit $a \in A$. On a $bA/(a) \simeq A/(a')$, où $a' = \frac{a}{a \wedge b}$.

Démonstration. On a le diagramme suivant

$$\begin{array}{ccc} A & \twoheadrightarrow & A/(a) \\ & \searrow \varphi & \downarrow \bar{x} \mapsto b\bar{x} \\ & & bA/(a) \end{array}$$

φ est un morphisme surjectif, déterminons son noyau.

Soit $y \in A$. $y \in \ker \varphi \iff by \in (a) \iff a \mid by \iff \frac{a}{a \wedge b} \mid \frac{b}{a \wedge b} y \iff a' \mid y \iff y \in (a')$. Donc $\ker \varphi = (a')$. □

Lemme 4.16. Soit $b \in A \setminus \{0\}$. Soit $a \in A$. On a $(A/(a))/(bA/(a)) \simeq A/(a \wedge b)$.

Démonstration. On a le diagramme suivant

$$\begin{array}{ccccc} & & \varphi & & \\ & \searrow & \curvearrowright & \searrow & \\ A & \twoheadrightarrow & A/(a) & \twoheadrightarrow & (A/(a))/(bA/(a)) \end{array}$$

φ est un morphisme surjectif, déterminons son noyau.

Soit $y \in A$. On a $y \in \ker \varphi \iff y + (a) \in bA/(a) \iff \exists x \in A : y + (a) = bx + (a) \iff \exists x, z \in A : y = bx + za \iff y \in (a) + (b) \iff y \in (a \wedge b)$. □

Pour achever l'unicité des facteurs de torsion, remplaçons désormais M par $M_T \neq 0$. On a $A/(a_1) \oplus \dots \oplus A/(a_s) \simeq M \simeq A/(b_1) \oplus \dots \oplus A/(b_t)$.

Soit p un diviseur irréductible de a_s . On a $pM \simeq A/(a'_1) \oplus \dots \oplus A/(a'_s)$, où $a'_i = \frac{a_i}{p}$. On a aussi $M/pM \simeq (A/(p))^s$.

Comme $a_1 = b_1$, il existe ℓ tel que pour tout $i \leq \ell$, $p \mid b_i$, et pour tout $i > \ell$, $p \nmid b_i$. On a $pM \simeq A/(b'_1) \oplus \dots \oplus A/(b'_\ell) \oplus A/(b_{\ell+1}) \oplus \dots \oplus A/(b_t)$, où $b'_i = \frac{b_i}{p}$, pour $i \leq \ell$. On a aussi $M/pM \simeq (A/(p))^\ell$.

La structure de A -module de M/pM provient d'une structure de $A/(p)$ -espace vectoriel ($A/(p)$ est un corps). On en déduit que $\ell = s$ (c'est la dimension de M/pM vu comme $A/(p)$ -espace vectoriel).

On en déduit que $s \leq t$. Par symétrie, on a $s = t$.

D'où $\ell = t$. On en déduit que $A/(a'_1) \oplus \dots \oplus A/(a'_s) \simeq pM \simeq A/(b'_1) \oplus \dots \oplus A/(b'_t)$. On omet les facteurs inversibles, puis on complète les deux listes par des unités à droite pour les comparer avec la même longueur. On conclut par récurrence sur le nombre de diviseurs irréductibles de a_1 comptés avec leurs multiplicités. On en déduit que $a_i = b_i$ aux inversibles près.

L'initialisation de la récurrence correspond à a_1 irréductible, qui est un cas trivial (dans ce cas, tous les a_i et b_i sont égaux).

□

4.3 Applications

4.3.1 Cas particulier des groupes abéliens

Théorème de décomposition des groupes abéliens

Un groupe abélien est naturellement muni d'une structure de $\mathbb{Z}$ -module :

$$\mathbb{Z} \times G \rightarrow G$$

$$(n, g) \mapsto n \cdot g = \begin{cases} \underbrace{g + \dots + g}_{n \text{ termes}} & \text{si } n > 0 \\ 0 & \text{si } n = 0 \\ \underbrace{-g - \dots - g}_{-n \text{ termes}} & \text{si } n < 0. \end{cases}$$

Remarque. • Un morphisme de groupes abéliens est aussi un morphisme de $\mathbb{Z}$ -modules.

- Un sous-groupe d'un groupe abélien est un sous- $\mathbb{Z}$ -module.

Soit G un groupe abélien. Soit $g \in G$. g est de torsion si et seulement si $\exists n \neq 0 : ng = 0$ si et seulement si $\exists n > 0 : ng = 0$ si et seulement si g est d'ordre fini. On en déduit la proposition suivante :

Proposition 4.17. *Le sous-module de torsion de G est le sous-groupe formé par les éléments d'ordre fini dans G .*

Le théorème 4.7 s'écrit dans ce contexte de la manière suivante.

Théorème 4.18 (Décomposition des groupes abéliens). *Soit G un groupe abélien de type fini. Alors il existe deux uniques entiers r et s et une unique famille d'entiers naturels > 1 vérifiant $n_s \mid \dots \mid n_1$ tels que $G \simeq \mathbb{Z}/n_1\mathbb{Z} \oplus \dots \oplus \mathbb{Z}/n_s\mathbb{Z} \oplus \mathbb{Z}^r$.*

Dans le cas particulier où G est fini non trivial, on a :

- $r = 0$,
- $n_1\mathbb{Z} = \text{Ann } G$. Donc n_1 est le ppcm des ordres des éléments de G , c'est-à-dire l'exposant du groupe G . n_1 est également l'ordre d'un élément de G , c'est donc le plus grand ordre possible.
- $|G| = \prod_{i=1}^s n_i$.

Dans le cas d'un groupe abélien, on peut retrouver assez facilement l'unicité des n_i . Prenons en effet $n_s \mid \dots \mid n_1$ et $m_t \mid \dots \mid m_1$ tels que $\bigoplus_{i=1}^s \mathbb{Z}/n_i\mathbb{Z} = \bigoplus_{i=1}^t \mathbb{Z}/m_i\mathbb{Z}$. Le lecteur montrera les résultats suivants :

- (i). $\forall d \in \mathbb{N}^*, \forall n \in \mathbb{N}^*, |d\mathbb{Z}/n\mathbb{Z}| = \frac{n}{d \wedge n}$.
- (ii). $\forall d \in \mathbb{N}^*, \prod_{i=1}^s (d \wedge n_i) = \prod_{i=1}^t (d \wedge m_i)$.
- (iii). $s = t$ et $\forall i, n_i = m_i$.

Le lecteur montrera également que $\mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z} \simeq \mathbb{Z}/a \wedge b\mathbb{Z} \times \mathbb{Z}/a \vee b\mathbb{Z}$, en

- utilisant le théorème 4.18,
- décomposant a et b en puissances de nombres premiers et en utilisant le lemme chinois,
- exhibant un isomorphisme.

Exemples d'application

Structure du groupe multiplicatif d'un corps commutatif Donnons une autre démonstration du théorème 2.14 qui assure que tout sous-groupe fini G du groupe multiplicatif d'un corps commutatif est cyclique.

Démonstration. Si $G = \{1\}$, il est cyclique. Supposons désormais $G \neq \{1\}$. D'après le théorème 4.18, on a $G \simeq G_1 \oplus \dots \oplus G_s$, où $G_i \simeq \mathbb{Z}/n_i\mathbb{Z}$ et $n_i \mid n_{i-1}$. Donc tous les éléments de G sont racines du polynôme $X^{n_1} - 1$. Or ce polynôme a au plus n_1 racines. Donc $|G| \leq n_1$, ce qui impose $s = 1$. $\square$

Le problème de Galois inverse pour les groupes abéliens L'extension galoisienne

$$\begin{array}{c} \mathbb{K} = D_{\mathbb{Q}}(P) \\ | \\ \mathbb{Q} \end{array}$$

correspond au groupe de Galois $\text{Gal}_{\mathbb{Q}}(\mathbb{K})$. On s'intéresse à la question suivante : quels sont les groupes finis qui sont des groupes de Galois d'extensions galoisiennes de $\mathbb{Q}$?

Cette question est très difficile et le problème est encore ouvert. Cependant, on connaît des groupes finis qui sont des groupes de Galois d'extensions galoisiennes de $\mathbb{Q}$: par exemple, les groupes abéliens finis.

Théorème 4.19. *Les groupes abéliens finis sont des groupes de Galois d'extensions galoisiennes de $\mathbb{Q}$.*

La démonstration repose sur les propositions suivantes que nous ne démontrerons pas.

Proposition 4.20. *On note $\text{Gal}_{\mathbb{K}}(\mathbb{L})$ le groupe des automorphismes de $\mathbb{L}$ fixant $\mathbb{K}$ point par point. On considère l'extension cyclotomique galoisienne de degré $\deg \Phi_n = \varphi(n)$*

$$\begin{array}{c} \mathbb{Q}(e^{\frac{2i\pi}{n}}) = D_{\mathbb{Q}}(\Phi_n) \\ | \\ \mathbb{Q} \end{array}.$$

On a

$$\text{Gal}_{\mathbb{Q}}(\mathbb{Q}(e^{\frac{2i\pi}{n}})) = \{ \sigma : e^{\frac{2i\pi}{n}} \mapsto e^{\frac{2ik\pi}{n}}, k \wedge n = 1 \} \simeq (\mathbb{Z}/n\mathbb{Z})^*.$$

Proposition 4.21. *On considère une extension galoisienne finie $\mathbb{Q} - \mathbb{K}$. On note $G = \text{Gal}_{\mathbb{Q}}(\mathbb{K})$. Soit H un sous-groupe de G . On pose $\mathbb{K}^H = \{x \in \mathbb{K} : \forall \sigma \in H, \sigma(x) = x\}$. On a alors*

$$\begin{array}{c} \mathbb{K} \\ | \\ \mathbb{K}^H \\ | \\ \mathbb{Q} \end{array}.$$

Alors l'extension $\mathbb{Q} - \mathbb{K}^H$ est galoisienne si et seulement si $H \triangleleft G$. Dans ce cas, $\text{Gal}_{\mathbb{Q}}(\mathbb{K}^H) \simeq G/H$.

Montrons maintenant le lemme suivant.

Lemme 4.22. Soit G un groupe abélien fini. Il existe un entier n tel que G est isomorphe à un quotient de $(\mathbb{Z}/n\mathbb{Z})^*$.

Démonstration. Si G est trivial, $n = 2$ convient. Sinon, on cherche un n convenable sous la forme $n = p_1 \dots p_t$ où les p_i sont premiers deux à deux distincts. Alors

$$(\mathbb{Z}/n\mathbb{Z})^* \simeq \prod_{i=1}^t (\mathbb{Z}/p_i\mathbb{Z})^* \simeq \prod_{i=1}^t \mathbb{Z}/(p_i - 1)\mathbb{Z}.$$

D'après le théorème 4.18, il existe s et $n_s \mid \dots \mid n_1$ tels que $G \simeq \mathbb{Z}/n_1\mathbb{Z} \times \dots \times \mathbb{Z}/n_s\mathbb{Z}$.

Pour que G soit isomorphe à un quotient de $(\mathbb{Z}/n\mathbb{Z})^*$, il suffit que $s = t$ et que pour tout i , $\mathbb{Z}/n_i\mathbb{Z}$ soit isomorphe à un quotient de $(\mathbb{Z}/p_i\mathbb{Z})^*$, c'est-à-dire $s = t$ et pour tout i , $n_i \mid p_i - 1$ (car les quotients d'un groupe cyclique $\mathbb{Z}/k\mathbb{Z}$ sont, à isomorphisme près, les groupes cycliques $\mathbb{Z}/\ell\mathbb{Z}$, avec $\ell \mid k$).

Il suffit maintenant de prouver l'existence de s nombres premiers $p_1, \dots, p_s$ tels que pour tout i , $n_i \mid p_i - 1$. Le théorème de la progression arithmétique de Dirichlet assure qu'il existe une infinité de nombres premiers $p \equiv 1 \pmod{n_1}$. On prend les s premiers : $p_1, \dots, p_s$. On a $n_i \mid n_1 \mid p_i - 1$. $\square$

Démonstration. (théorème 4.19) Il existe n et $H \triangleleft (\mathbb{Z}/n\mathbb{Z})^*$ tels que $G \simeq (\mathbb{Z}/n\mathbb{Z})^*/H$. On a alors

$$\left. \begin{array}{c} \mathbb{K} = \mathbb{Q}(e^{\frac{2i\pi}{n}}) \\ H \mid \\ \mathbb{K}^H \\ (\mathbb{Z}/n\mathbb{Z})^*/H \simeq G \mid \\ \mathbb{Q} \end{array} \right\} (\mathbb{Z}/n\mathbb{Z})^*$$

Le groupe G est bien groupe de Galois d'une extension galoisienne : $\mathbb{Q} - \mathbb{K}^H$. $\square$

4.3.2 Structure de $\mathbb{K}[X]$ -module induite par un endomorphisme sur un espace vectoriel de dimension finie

Soit $\mathbb{K}$ un corps commutatif. Soit E un $\mathbb{K}$ -espace vectoriel de dimension finie. Soit $u \in \mathcal{L}(E)$. On munit E de la loi externe

$$\mathbb{K}[X] \times E \rightarrow E, \quad (P, x) \mapsto P(u)(x).$$

On vérifie qu'on munit ainsi E d'une structure de $\mathbb{K}[X]$ -module. On le note E_u .

Remarques. • Un sous-module de E_u est un sous-espace u -stable de E .

- Un endomorphisme de E_u est un endomorphisme de E qui commute avec u .
- Un sous-espace vectoriel u -cyclique de E est un sous-module monogène de E_u (d'où la notation $\langle x \rangle_u$).
- E_u est de torsion et $\text{Ann } E_u = (\pi_u)$.

Lemme 4.23. Soit $Q \in \mathbb{K}[X]$ unitaire non constant. Soit F un sous-espace vectoriel u -stable de E (i.e. un sous- $\mathbb{K}[X]$ -module de E_u). Alors F est isomorphe en tant que $\mathbb{K}[X]$ -module à $\mathbb{K}[X]/(Q)$ si et seulement si F est u -cyclique et $\pi_{u_F} = Q$.

Démonstration. ($\Rightarrow$) Il existe $\psi : \mathbb{K}[X]/(Q) \rightarrow F$ un isomorphisme de $\mathbb{K}[X]$ -module. Soit $x = \psi(1)$. Alors on a $\psi(P(X)) = P \cdot \psi(1) = P \cdot x = P(u)(x)$. Comme ψ est surjectif, F est u -cyclique. L'injectivité de ψ , quant à elle, assure que $P(u)(x) = 0$ si et seulement si $Q \mid P$, c'est-à-dire que $Q = \pi_{u_F}$.

($\Leftarrow$) Réciproquement, si F est u -cyclique de polynôme minimal Q , on a un morphisme surjectif

$$\mathbb{K}[X] \rightarrow F, \quad P \mapsto P(u)(x)$$

de noyau $(\pi_{u_F}) = (Q)$. $\square$

Appliquons à E_u le théorème 4.7. On obtient l'existence d'un unique entier s et d'une famille de polynômes $Q_s \mid \dots \mid Q_1$ unique aux inversibles près tels que $E_u \simeq \bigoplus_{i=1}^s \mathbb{K}[X]/(Q_i)$. Pour tout i , on appelle F_i le sous- $\mathbb{K}[X]$ -module de E_u qui est l'image de $\mathbb{K}[X]/(Q_i)$ par l'isomorphisme $\bigoplus_{i=1}^s \mathbb{K}[X]/(Q_i) \rightarrow E_u$. On a $E_u = \bigoplus_{i=1}^s F_i$ (somme directe de sous- $\mathbb{K}[X]$ -modules de E_u), donc $E = \bigoplus_{i=1}^s F_i$ (somme directe de sous-espaces vectoriels u -stables de E). En outre, pour tout i , $F_i \simeq \mathbb{K}[X]/(Q_i)$. D'après le lemme, pour tout i , F_i est un sous-espace u -cyclique de E et $\pi_{u_{F_i}} = Q_i$. On retrouve le théorème 3.17 sur les invariants de similitude.

Proposition 4.24. Soient $u, v \in \mathcal{L}(E)$. Alors u et v sont semblables si et seulement si les $\mathbb{K}[X]$ -modules E_u et E_v sont isomorphes.

Remarque. La classe d'isomorphisme du $\mathbb{K}[X]$ -module E_u caractérise la classe de similitude de u dans $\mathcal{L}(E)$.

Démonstration. Soit w un automorphisme de l'espace vectoriel E . w est un isomorphisme de $\mathbb{K}[X]$ -modules de E_u dans E_v si et seulement si

$$\begin{aligned} & \forall P \in \mathbb{K}[X], \forall x \in E, w(P \cdot x) = P \cdot w(x) \\ \iff & \forall P \in \mathbb{K}[X], \forall x \in E, w(P(u)(x)) = P(v)w(x) \\ \iff & \forall P \in \mathbb{K}[X], wP(u) = P(v)w \\ \iff & wu = vw. \end{aligned}$$

□

Remarque. On appelle module simple un module non nul qui n'a aucun sous-module non trivial. Ainsi, un endomorphisme u est simple si et seulement si E_u est un module simple. Certaines notions vues dans le cadre des endomorphismes d'un espace vectoriel apparaissent comme des cas particuliers de notions plus générales concernant les modules sur les anneaux principaux. Nous n'entrerons pas dans les détails ici.

4.3.3 Classes d'équivalence de matrices à coefficients dans un anneau principal

Cas général

Soit A un anneau principal. Soient $M, N \in M_{m,n}(A)$. Le rang d'une telle matrice désigne son rang sur le corps des fractions de A .

Définition 4.8. M et N sont dites *équivalentes* si et seulement si il existe des matrices $P \in GL_m(A)$ et $Q \in GL_n(A)$ telles que $M = PNQ$.

Proposition 4.25. La relation "être équivalente" est une relation d'équivalence.

Remarque. Soit $M \in M_{m,n}(A)$. M définit canoniquement un morphisme $A^n \rightarrow A^m$, noté abusivement M . Deux matrices M et N sont équivalentes si et seulement si l'ensemble des matrices de M pour des bases quelconques de A^n et A^m est égal à l'ensemble des matrices de N pour des bases quelconques de A^n et A^m si et seulement si M et N ont une matrice en commun.

Théorème 4.26. Soit $M \in M_{m,n}(A)$. Alors il existe un unique entier $r = \text{rg } M$ et une famille d'éléments non nuls de A $a_r \mid \dots \mid a_1$ unique aux inversibles près tels que M soit équivalente à la matrice

$$\begin{pmatrix} a_1 & & & & 0 \\ & \ddots & & & \\ & & a_r & & \\ & & & 0 & \\ 0 & & & & \ddots \\ & & & & & 0 \end{pmatrix}.$$

Remarque. $r \leq \min(m, n)$ est le rang de la matrice M .

Démonstration. On a $M : A^n \rightarrow A^m$. On montre que $\ker M$ est facteur direct de A^n . En effet, si $x \in A^n$ est tel qu'il existe $\alpha \neq 0$ vérifiant $\alpha x \in \ker M$, alors $x \in \ker M$ (on a ainsi $(A^n / \ker M)_T = \{0\}$).

D'après le théorème 4.8, il existe $(e_1, \dots, e_n)$ une base de A^n , il existe $s \leq n$ et il existe $b_s \mid \dots \mid b_1$ non nuls tels que $(b_1 e_1, \dots, b_s e_s)$ est une base de $\ker M$. Comme $b_i e_i \in \ker M$, d'après la remarque ci-dessus, $e_i \in \ker M$. Les b_i sont donc inversibles et $\ker M = \bigoplus_{i=1}^s A e_i$. On a donc $A^n = \ker M \oplus \bigoplus_{i>s} A e_i$.

Montrons maintenant que $\text{Im } M = \bigoplus_{i>s} A M e_i$. On sait déjà que l'image de M est égale à $\sum_{i=1}^n A M e_i = \sum_{i>s} A M e_i$, car pour tout $i \leq s$, $M e_i = 0$. Soit $(x_i)_{i>s} \in A^{n-s}$ tels que $\sum_{i>s} x_i M e_i = 0$. On a alors $\sum_{i>s} x_i e_i \in \ker M$. Comme $\ker M \cap \bigoplus_{i>s} A e_i = \{0\}$, on a $\sum_{i>s} x_i e_i = 0$, donc pour tout $i > s$, $x_i = 0$.

On applique le théorème 4.8 au sous-module $\text{Im } M$ de A^m : il existe $(f_1, \dots, f_m)$ une base de A^m , il existe $r \leq m$ et il existe $a_r \mid \dots \mid a_1$ non nuls tels que $(a_1 f_1, \dots, a_r f_r)$ est une base de $\text{Im } M$. On a clairement $r = n - s$.

Il existe $g_1, \dots, g_r \in A^n$ tels que $M g_i = a_i f_i$. Montrons que

$$A^n = \underbrace{\bigoplus_{i=1}^s A e_i}_{\ker M} \oplus \bigoplus_{i=1}^r A g_i.$$

Soit $e \in A^n$. On a $M e = \sum_{i=1}^r x_i a_i f_i$, car $\text{Im } M = \bigoplus_{i>s} A M e_i$. Donc $M e = M \sum_{i=1}^r x_i g_i$, puis $e - \sum_{i=1}^r x_i g_i \in \ker M$. On a donc $A^n = \ker M + \sum_{i=1}^r A g_i$. Supposons avoir $f \in \ker M$ et $x_i \in A$ tels que $f + \sum_{i=1}^r x_i g_i = 0$. On a $0 = M(f + \sum_{i=1}^r x_i g_i) = \sum_{i=1}^r x_i a_i f_i$. Donc pour tout i , $x_i = 0$. On en déduit aussi que $f = 0$.

La matrice de M dans la base $(g_1, \dots, g_r, e_1, \dots, e_s)$ de A^n et dans la base $(f_1, \dots, f_m)$ de A^m s'écrit

$$D = \begin{pmatrix} a_1 & & & & 0 \\ & \ddots & & & \\ & & a_r & & \\ & & & 0 & \\ 0 & & & & \ddots \\ & & & & & 0 \end{pmatrix}.$$

En appelant P et Q les matrices de passage convenables, on a $M = PDQ$.

Montrons que les a_i sont uniques aux inversibles près. On a le diagramme commutatif suivant

$$\begin{array}{ccc} A^n & \xrightarrow{M} & A^m \\ Q \downarrow \sim & & \sim \uparrow P \\ A^n & \xrightarrow{D} & A^m \end{array}$$

Il existe un automorphisme de A^m (c'est P) qui envoie $\text{Im } D$ sur $\text{Im } M$. Soit $(h_1, \dots, h_m)$ la base canonique de A^m . On a $\text{Im } D = \bigoplus_{i=1}^r A a_i h_i$. Sur la base $(P h_1, \dots, P h_m)$, on a $\text{Im } M = \bigoplus_{i=1}^r A a_i P h_i$. $(P h_1, \dots, P h_m)$ est une base adaptée à l'image de M . Les a_i sont alors uniques aux inversibles près (on invoque l'unicité du théorème 4.8). $\square$

La proposition suivante donne l'expression des a_i en fonction des coefficients de la matrice M .

Proposition 4.27. Soit $M = PDQ \in M_{m,n}(A)$, avec $P \in GL_m(A)$, $Q \in GL_n(A)$ et $D \in M_{m,n}(A)$ diagonale, de coefficients diagonaux non nuls $a_1, \dots, a_r$ vérifiant $a_r \mid \dots \mid a_1$, tous ses autres coefficients étant nuls. Alors

- r est le plus grand entier s (avec la convention que le mineur de taille 0 vaut 1) tel qu'il existe un mineur de taille s extrait de M et non nul,
- si $r \geq 1$, a_r est, à multiplication par une unité près, le pgcd des coefficients de M ,
- pour $1 \leq i < r$, a_i est, à multiplication par une unité près, le quotient du pgcd des mineurs de taille $r + 1 - i$ extraits de M et du pgcd des mineurs de taille $r - i$ extraits de M .

Remarque. On redémontre ainsi l'unicité des coefficients a_i aux inversibles près (puisque'un pgcd est unique aux inversibles près).

Démonstration. Soit $M \in M_{m,n}(A)$. On note $\mathcal{J}_j(M)$ l'idéal engendré par les mineurs de taille j extraits de M , pour $1 \leq j \leq \min(m, n)$.

On note

$$D = \begin{pmatrix} a_1 & & & & \\ & \ddots & & & \\ & & a_r & & \\ & & & 0 & \\ & 0 & & & \ddots \\ & & & & & 0 \end{pmatrix}.$$

Alors, pour $j > r$, on a $\mathcal{J}_j(D) = (0)$ et, pour $j \leq r$, $\mathcal{J}_j(D) = (a_{i_1} \dots a_{i_j}, 1 \leq i_1 < \dots < i_j \leq r) = (a_r \dots a_{r+1-j})$.

Montrons que $\mathcal{J}_j(M)$ reste constant dans la classe d'équivalence de M .

Remarque. On a $\mathcal{J}_j(M) = \mathcal{J}_j({}^t M)$. Il suffit donc de montrer que pour toute matrice $P \in GL_m(A)$, pour toute matrice $M \in M_{m,n}(A)$, $\mathcal{J}_j(PM) = \mathcal{J}_j(M)$. En effet, on aurait alors $\mathcal{J}_j(PMQ) = \mathcal{J}_j(MQ) = \mathcal{J}_j({}^t(MQ)) = \mathcal{J}_j({}^t Q {}^t M) = \mathcal{J}_j({}^t M) = \mathcal{J}_j(M)$.

Il suffit même de démontrer que pour toute matrice $P \in GL_m(A)$, pour toute matrice $M \in M_{m,n}(A)$, $\mathcal{J}_j(PM) \subset \mathcal{J}_j(M)$, puisqu'alors $\mathcal{J}_j(M) = \mathcal{J}_j(P^{-1}PM) \subset \mathcal{J}_j(PM)$ et on a égalité des deux idéaux.

Montrons donc que pour toute matrice $P \in GL_m(A)$, pour toute matrice $M \in M_{m,n}(A)$, $\mathcal{J}_j(PM) \subset \mathcal{J}_j(M)$.

Soit Δ un mineur de taille j extrait de PM . On note $PM = (x_{i,k})_{\substack{1 \leq i \leq m \\ 1 \leq k \leq n}}$, $P = (p_{i,k})_{\substack{1 \leq i \leq m \\ 1 \leq k \leq m}}$ et $M = (m_{i,k})_{\substack{1 \leq i \leq m \\ 1 \leq k \leq n}}$. On a $x_{i,k} = \sum_{\ell=1}^m p_{i,\ell} m_{\ell,k}$, donc

$$\begin{aligned} \Delta &= \det(x_{i_a, k_b})_{1 \leq a, b \leq j} = \det\left(\sum_{\ell=1}^m p_{i_a, \ell} m_{\ell, k_b}\right)_{1 \leq a, b \leq j} \\ &= \sum_{\ell_1, \dots, \ell_j=1}^m p_{i_1, \ell_1} \dots p_{i_j, \ell_j} \underbrace{\det(m_{\ell_a, k_b})_{1 \leq a, b \leq j}}_{\text{nul ou égal, au signe près, à un mineur de } M} \in \mathcal{J}_j(M). \end{aligned}$$

D'où $\mathcal{J}_j(PM) \subset \mathcal{J}_j(M)$. □

Cas particulier où l'anneau A est euclidien

On suppose A euclidien. Dans ce cas, on dispose d'un algorithme qui prend en entrée $M \in M_{m,n}(A)$ et sort $a_r \mid \dots \mid a_1$ tels que M est équivalente à

$$\begin{pmatrix} a_1 & & & & \\ & \ddots & & & \\ & & a_r & & \\ & & & 0 & \\ & 0 & & & \ddots \\ & & & & & 0 \end{pmatrix}.$$

On redémontre ainsi la partie existence du théorème 4.26.

Faisons quelques rappels utiles à la compréhension de cet algorithme.

Définition 4.9. Pour $n \geq 1$, notons I_n la matrice identité et E_{ij} la matrice dont le seul coefficient non nul est un 1 en position (i, j) . Les *matrices d'opérations élémentaires* sont les suivantes :

- *Transpositions*, pour $i \neq j$:

$$\tau_{i,j} = I_n - E_{ii} - E_{jj} + E_{ij} + E_{ji}, \quad \det \tau_{i,j} = -1.$$

Elles échangent les vecteurs e_i et e_j de la base canonique et engendrent les matrices de permutation $(\delta_{i,\sigma(j)})_{1 \leq i,j \leq n}$, pour $\sigma \in \mathfrak{S}_n$.

- *Transvections*, pour $a \in A$ et $i \neq j$:

$$T_{i,j,a} = I_n + aE_{ij}, \quad \det T_{i,j,a} = 1.$$

Puisque A est euclidien, elles engendrent $SL_n(A)$ (voir [11]).

- *Dilatations*, pour $a \in A^*$:

$$D_a = I_n + (a - 1)E_{nn} = \text{diag}(1, \dots, 1, a), \quad \det D_a = a.$$

Ces matrices appartiennent donc toutes à $GL_n(A)$.

Intéressons-nous à ce que l'on obtient en multipliant une matrice $M \in M_{m,n}(A)$ par une matrice d'opération élémentaire.

Proposition 4.28. Soit $M \in M_{m,n}(A)$. On note

$$M = \begin{pmatrix} L_1 \\ \vdots \\ L_m \end{pmatrix} = (C_1 \cdots C_n).$$

- *Multiplication par une matrice de permutation.*

- Soit $\sigma \in \mathfrak{S}_n$. On pose $\Sigma = (\delta_{i,\sigma(j)})_{1 \leq i,j \leq n}$. Alors

$$M\Sigma = \begin{pmatrix} C_{\sigma(1)} & \cdots & C_{\sigma(n)} \end{pmatrix},$$

et, en particulier,

$$M\tau_{i,j} = (C_1 \cdots C_{i-1} C_j C_{i+1} \cdots C_{j-1} C_i C_{j+1} \cdots C_n),$$

où $\tau_{i,j}$ désigne la matrice associée à la transposition (i, j) .

- Soit $\sigma \in \mathfrak{S}_m$. On pose $\Sigma = (\delta_{i,\sigma(j)})_{1 \leq i,j \leq m}$. Alors

$$\Sigma M = \begin{pmatrix} L_{\sigma^{-1}(1)} \\ \vdots \\ L_{\sigma^{-1}(m)} \end{pmatrix},$$

et, en particulier,

$$\tau_{i,j}M = \begin{pmatrix} L_1 \\ \vdots \\ L_{i-1} \\ L_j \\ L_{i+1} \\ \vdots \\ L_{j-1} \\ L_i \\ L_{j+1} \\ \vdots \\ L_m \end{pmatrix}.$$

- Multiplication par une matrice de transvection.

- Soit une matrice de transvection $T_{i,j,a} \in \text{GL}_n(A)$. Alors

$$MT_{i,j,a} = \begin{pmatrix} C_1 & \cdots & C_{j-1} & C_j + aC_i & C_{j+1} & \cdots & C_n \end{pmatrix}.$$

- Soit une matrice de transvection $T_{i,j,a} \in \text{GL}_m(A)$. Alors

$$T_{i,j,a}M = \begin{pmatrix} L_1 \\ \vdots \\ L_{i-1} \\ L_i + aL_j \\ L_{i+1} \\ \vdots \\ L_m \end{pmatrix}.$$

- Multiplication par une matrice de dilatation.

- Soit une matrice de dilatation $D_a \in \text{GL}_n(A)$. Alors

$$MD_a = \begin{pmatrix} C_1 & \cdots & C_{n-1} & aC_n \end{pmatrix}.$$

- Soit une matrice de dilatation $D_a \in \text{GL}_m(A)$. Alors

$$D_aM = \begin{pmatrix} L_1 \\ \vdots \\ L_{m-1} \\ aL_m \end{pmatrix}.$$

On appelle G_m (respectivement G_n) le sous-groupe de $\text{GL}_m(A)$ (respectivement $\text{GL}_n(A)$) engendré par les matrices d'opérations élémentaires.

En effectuant une suite d'opérations élémentaires sur M , la matrice obtenue est de la forme PMQ , où $P \in G_m$ et $Q \in G_n$.

On note $v : A \setminus \{0\} \rightarrow \mathbb{N}$ un stathme euclidien : pour $b \neq 0$, une division $a = qb + r$ est possible avec $r = 0$ ou $v(r) < v(b)$. Si $M = (m_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in M_{m,n}(A) \setminus \{0\}$, on pose $v(M) = \min(v(m_{i,j}), m_{i,j} \neq 0)$.

Proposition 4.29. *L'algorithme suivant permet d'obtenir à partir de M la matrice*

$$D = \begin{pmatrix} a_1 & & & & \\ & \ddots & & & 0 \\ & & a_r & & \\ & & & 0 & \\ 0 & & & & \ddots \\ & & & & & 0 \end{pmatrix}$$

telle que $a_r \mid \dots \mid a_1$ et qu'il existe $P \in \text{GL}_m(A)$ et $Q \in \text{GL}_n(A)$ vérifiant $M = PDQ$.

Démonstration. On montre aisément ce résultat par induction sur $(m+n, v(M))$. □

Voici l'algorithme. Si la matrice, ou le bloc restant, est nul ou de taille nulle, on s'arrête sur ce bloc. Après chaque diminution de $v(M)$, on reprend à la première étape. Les opérations et permutations sont mémorisées pour retrouver les matrices de changement de bases.

- Première étape :

$$M \leftarrow \Sigma M \Sigma'.$$

Σ et Σ' sont des matrices de permutation choisies de sorte que $v(m_{1,1}) = v(M)$ après modification de M .

- Deuxième étape :

- On suppose qu'il existe $i > 1$ tel que $m_{1,1} \nmid m_{1,i}$. Alors il existe $q, r \in A$ tels que $m_{1,i} = qm_{1,1} + r$ avec $r \neq 0$ et $v(r) < v(m_{1,1})$.

$$M \leftarrow MT_{1,i,-q}.$$

Ceci revient en fait à $C_i \leftarrow C_i - qC_1$. $v(M)$ a strictement diminué.

- On suppose que pour tout $i > 1$, $m_{1,1} \mid m_{1,i}$. Alors pour tout i , il existe $q_i \in A$ tel que $m_{1,i} = q_i m_{1,1}$.

$$M \leftarrow MT_{1,2,-q_2} \dots T_{1,n,-q_n}.$$

$$M \text{ devient alors } \begin{pmatrix} m_{1,1} & 0 & \dots & \dots & 0 \\ & & \star & & \end{pmatrix}.$$

- On suppose qu'il existe $j > 1$ tel que $m_{1,1} \nmid m_{j,1}$. Alors il existe $q, r \in A$ tels que $m_{j,1} = qm_{1,1} + r$ avec $r \neq 0$ et $v(r) < v(m_{1,1})$.

$$M \leftarrow T_{j,1,-q}M.$$

Ceci revient en fait à $L_j \leftarrow L_j - qL_1$. $v(M)$ a strictement diminué.

- On suppose que pour tout $j > 1$, $m_{1,1} \mid m_{j,1}$. Alors pour tout j , il existe $q_j \in A$ tel que $m_{j,1} = q_j m_{1,1}$.

$$M \leftarrow T_{m,1,-q_m} \dots T_{2,1,-q_2}M.$$

$$M \text{ devient alors } \begin{pmatrix} m_{1,1} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \star & \\ 0 & & & \end{pmatrix}.$$

- On suppose que pour tous $i, j > 1$, $m_{1,1} \mid m_{i,j}$. On conserve le pivot $m_{1,1}$ et on applique récursivement l'algorithme au bloc $B = (m_{i,j})_{2 \leq i \leq m, 2 \leq j \leq n}$, sans le diviser par le pivot. Tous ses coefficients, puis tous ses invariants, restent divisibles par $m_{1,1}$. La taille diminue. Enfin, on inverse l'ordre des facteurs non nuls pour obtenir la convention $a_r \mid \dots \mid a_1$.

- On suppose qu'il existe $i, j > 1$ tels que $m_{1,1} \nmid m_{i,j}$.

$$M \leftarrow T_{1,i,1}M.$$

$$M \text{ devient alors } \begin{pmatrix} m_{1,1} & m_{1,2} & \cdots & m_{1,n} \\ 0 & & & \\ \vdots & & m_{i,j} & \\ 0 & & & \end{pmatrix}.$$

Il existe $q, r \in A$ tels que $m_{i,j} = qm_{1,1} + r$ avec $r \neq 0$ et $v(r) < v(m_{1,1})$.

$$M \leftarrow MT_{1,j,-q}.$$

$v(M)$ a strictement diminué.

4.3.4 Calcul effectif des invariants de similitude et autres applications

Calcul des invariants de similitude d'une matrice

Proposition 4.30. Soit E un $\mathbb{K}$ -espace vectoriel de dimension finie. Soit $u \in \mathcal{L}(E)$. Soit M une matrice de u . Soient $Q_r \mid \dots \mid Q_1$ les invariants de similitude de u . Alors les invariants de la classe d'équivalence dans $M_n(\mathbb{K}[X])$ de la matrice $XI_n - M$ sont, dans l'ordre décroissant de divisibilité adopté ici, $Q_1, \dots, Q_r, \underbrace{1, \dots, 1}_{n-r \text{ facteurs}}$.

Corollaire 4.31. On dispose d'un algorithme pour calculer les invariants de similitude de u .

Démonstration. $\mathbb{K}[X]$ est euclidien. □

Preuve de la proposition 4.30. Identifions E à $\mathbb{K}^n$ dans la base choisie. L'application

$$\Phi : \mathbb{K}[X]^n \longrightarrow E_u, \quad \sum_{k=0}^d X^k v_k \longmapsto \sum_{k=0}^d M^k v_k$$

est un morphisme surjectif de $\mathbb{K}[X]$ -modules. Son noyau est $(XI_n - M)\mathbb{K}[X]^n$: en effet, pour $k \geq 1$,

$$X^k v - M^k v = (XI_n - M) \sum_{j=0}^{k-1} X^{k-1-j} M^j v.$$

Tout vecteur polynomial est donc congru à sa valeur par Φ modulo ce sous-module. Ainsi,

$$\mathbb{K}[X]^n / (XI_n - M)\mathbb{K}[X]^n \simeq E_u \simeq \bigoplus_{i=1}^r \mathbb{K}[X] / (Q_i).$$

Le théorème de la base adaptée et l'unicité des facteurs invariants donnent les facteurs non inversibles $Q_1, \dots, Q_r$; les $n - r$ autres facteurs sont des unités, que l'on normalise à 1. □

Résolution d'équations diophantiennes du premier degré

On considère l'équation

$$a_1 x_1 + \cdots + a_n x_n = b, \quad a_1, \dots, a_n, b \in \mathbb{Z}.$$

Si tous les a_i sont nuls, toute famille d'entiers convient lorsque $b = 0$, et il n'y a aucune solution sinon. Dans les autres cas, on place en première position un coefficient non nul de valeur absolue minimale. Pour $j \geq 2$, écrivons $a_j = b_j a_1 + r_j$, avec $0 \leq r_j < |a_1|$. L'équation devient

$$a_1 \underbrace{(x_1 + b_2 x_2 + \cdots + b_n x_n)}_{X_1} + r_2 x_2 + \cdots + r_n x_n = b.$$

Le changement de variables est inversible sur $\mathbb{Z}^n$. Si un reste est non nul, la valeur absolue du plus petit coefficient non nul diminue ; on recommence avec ce coefficient comme pivot. Si tous les restes sont nuls, il existe des solutions si et seulement si $a_1 \mid b$, et, dans les coordonnées de cette dernière étape, elles sont

$$\left(\frac{b}{a_1} - \sum_{i=2}^n b_i x_i, x_2, \dots, x_n \right), \quad (x_2, \dots, x_n) \in \mathbb{Z}^{n-1}.$$

Pour retrouver les solutions de l'équation initiale, on remonte ensuite, en ordre inverse, les changements de variables et les permutations des étapes précédentes.

De manière plus générale, on peut résoudre en utilisant l'algorithme présenté précédemment un système d'équations du type $AX = B$, avec $A \in M_{m,n}(\mathbb{Z})$ et $B \in \mathbb{Z}^m$.

Génération de $GL_n(A)$

A est toujours supposé euclidien.

Toute matrice de $GL_n(A)$ est équivalente à l'identité : pour tout $A \in GL_n(A)$, il existe deux matrices U et V , produits de matrices d'opérations élémentaires, vérifiant $UAV = I_n$, i.e. $A = U^{-1}V^{-1}$.

$GL_n(A)$ est donc engendré par les matrices d'opérations élémentaires (voir [11]).

Méthode du pivot de Gauss

On travaille sur un corps. On part de l'équation $MX = B$ et, en multipliant à gauche par des matrices d'opérations élémentaires, on obtient l'équation $TX = B'$, où la matrice T est triangulaire supérieure. On peut donc plus facilement trouver X .

On peut utiliser la même méthode pour inverser une matrice.

Citons également deux propositions que le lecteur démontrera en exercice ou pourra trouver dans [5].

Proposition 4.32 (Décomposition LU). Soit $M \in GL_n(K)$. Il existe une matrice L triangulaire inférieure et une matrice U triangulaire supérieure telles que $M = LU$ si et seulement si pour tout i , le mineur fondamental extrait de M de taille i (i.e. le déterminant de la sous-matrice contenant les i premières lignes et les i premières colonnes de M) est non nul.

Si M admet une décomposition LU , il y a unicité de L et U à leur diagonale près : plus précisément, si $M = LU = L'U'$, alors il existe D diagonale telle que $L' = LD$ et $U' = D^{-1}U$.

Proposition 4.33 (Décomposition de Cholesky). Soit M une matrice réelle symétrique. Alors M est définie positive si et seulement si il existe une matrice triangulaire supérieure inversible T telle que $M = {}^tTT$ si et seulement si pour tout i , le mineur fondamental extrait de M de taille i est strictement positif.

La matrice T ainsi obtenue est unique si l'on impose à ses coefficients diagonaux d'être strictement positifs.

Ces méthodes interviennent aussi dans le calcul du rang et l'étude de $GL_2(\mathbb{Z})$.

Formes bilinéaires, formes quadratiques

Références : [1, 3, 11, 12]

5.1 Généralités

Soit $\mathbb{K}$ un corps commutatif. Soient E et F des $\mathbb{K}$ -espaces vectoriels de dimension finie. On note $E^* = \text{Hom}_{\mathbb{K}}(E, \mathbb{K})$ le dual de E .

Définition 5.1. Une application $B : E \times F \rightarrow \mathbb{K}$ est appelée *forme bilinéaire* si elle est linéaire en chacune de ses variables.

Une application $B : E \times E \rightarrow \mathbb{K}$ est appelée *forme bilinéaire symétrique* si elle est bilinéaire et si pour tous $x, y \in E$, on a $B(x, y) = B(y, x)$.

Soit B une forme bilinéaire symétrique.

Définition 5.2. Des éléments x et y de E sont dits *orthogonaux* si $B(x, y) = 0$.

Deux parties X et Y de E sont dites *orthogonales* si pour tout $x \in X$ et tout $y \in Y$, $B(x, y) = 0$.

Soit X une partie de E . On appelle *orthogonal* de X le sous-espace vectoriel $X^\perp = \{y \in E : \forall x \in X, B(x, y) = 0\}$.

Proposition 5.1. (i). $(X \cup Y)^\perp = X^\perp \cap Y^\perp$.

(ii). $(X + Y)^\perp = X^\perp \cap Y^\perp$, quand X et Y sont des sous-espaces vectoriels.

(iii). $X \subset X^{\perp\perp}$.

(iv). $(X \cap Y)^\perp \supset X^\perp + Y^\perp$.

Définition 5.3. On appelle *noyau* de la forme bilinéaire symétrique B le sous-espace vectoriel $\ker B = E^\perp$.

Proposition 5.2. B définit une application linéaire

$$E \rightarrow E^*, \quad x \mapsto (y \mapsto B(x, y)).$$

Le noyau de B est le noyau de cette application linéaire.

Définition 5.4. On dit que la forme bilinéaire symétrique B est *non dégénérée* quand $\ker B = \{0\}$.

Pour le lemme et son corollaire ci-dessous, on suppose B non dégénérée. Alors B induit un isomorphisme entre E et son dual. On a alors le lemme suivant.

Lemme 5.3. Soit F un sous-espace vectoriel de E . Alors $\dim F + \dim F^\perp = \dim E$.

Démonstration. On a le diagramme commutatif suivant

$$\begin{array}{ccc} E & \xrightarrow{\sim} & E^* \\ & \searrow \psi_F & \downarrow \text{restriction} \\ & & F^* \end{array}$$

On a $\ker \psi_F = \{x \in E : \forall y \in F, B(x, y) = 0\} = F^\perp$. Le théorème du rang assure que $\dim E = \dim F^* + \dim F^\perp$, d'où le résultat. $\square$

Corollaire 5.4. Soient F et G des sous-espaces vectoriels de E . On a

- (i). $F = F^{\perp\perp}$.
- (ii). $(F \cap G)^\perp = F^\perp + G^\perp$.

Remarque. Soit G un supplémentaire de $\ker B$. Alors $B : G \times G \rightarrow \mathbb{K}$ est non dégénérée. En effet, si $x \in G$ est tel que pour tout $y \in G$, $B(x, y) = 0$, alors pour tout $y \in E = G + \ker B$, $B(x, y) = 0$, donc $x \in \ker B$, puis $x = 0$.

Définition 5.5. Un vecteur $x \in E \setminus \{0\}$ est dit *isotrope* si $B(x, x) = 0$.

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

Un sous-espace vectoriel F est dit *totalement isotrope* si $F \subset F^\perp$.

Remarques. F est un sous-espace vectoriel totalement isotrope si et seulement si $B|_{F \times F} = 0$.

Si $D = \mathbb{K}e$ est une droite, on a e isotrope si et seulement si D isotrope si et seulement si D totalement isotrope.

Proposition 5.5. Soit B une forme bilinéaire symétrique non dégénérée et F un sous-espace vectoriel de E . Alors $\ker B|_{F \times F} = F \cap F^\perp$. Les propositions suivantes sont équivalentes :

- (i). $B|_{F \times F}$ est non dégénérée.
- (ii). F est non isotrope.
- (iii). $F^\perp$ est non isotrope.
- (iv). $B|_{F^\perp \times F^\perp}$ est non dégénérée.
- (v). $E = F \oplus F^\perp$.

Démonstration. $\ker B|_{F \times F} = \{x \in F : \forall y \in F, B(x, y) = 0\} = F \cap F^\perp$. Donc ((i).) $\iff$ ((ii).).

$F^\perp \cap F^{\perp\perp} = F \cap F^\perp$. Donc ((ii).) $\iff$ ((iii).).

$\ker B|_{F^\perp \times F^\perp} = F^\perp \cap F^{\perp\perp}$. Donc ((iii).) $\iff$ ((iv).).

$\dim E = \dim F + \dim F^\perp$. Donc ((iv).) $\iff$ ((v).). □

Définition 5.6. On considère $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . On a

$$B\left(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j\right) = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} (B(e_i, e_j))_{1 \leq i, j \leq n} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

$(B(e_i, e_j))_{1 \leq i, j \leq n}$ est appelée *matrice de la forme bilinéaire symétrique B dans la base $\mathcal{B}$* . On la note $M_{\mathcal{B}}(B)$. Elle est symétrique.

Remarques. L'ensemble des matrices de B est

$$\left\{ {}^t P (B(e_i, e_j))_{1 \leq i, j \leq n} P, P \in \mathrm{GL}_n(\mathbb{K}) \right\}.$$

$\ker B = \left\{ \sum_{i=1}^n x_i e_i, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \ker M_{\mathcal{B}}(B) \right\}$. En particulier, B est non dégénérée si et seulement si $M_{\mathcal{B}}(B)$ est inversible.

$\mathrm{rg} M_{\mathcal{B}}(B)$ est la dimension maximale d'un sous-espace vectoriel sur lequel B est non dégénérée. On appelle aussi cette dimension le rang de B .

Définition 5.7. On note $\mathbb{K}^2 = \{a^2 : a \in \mathbb{K}\}$ l'ensemble des carrés et $\mathbb{K}^{*2} = \{a^2 : a \in \mathbb{K}^*\}$ le sous-groupe des carrés non nuls. On suppose B non dégénérée. Soient M et M' deux matrices de B . Alors elles sont inversibles et il existe $P \in \text{GL}_n(\mathbb{K})$ telle que $M = {}^t P M' P$. On a donc $\det M = \det P^2 \det M'$. On appelle *discriminant* de B la classe de $\det M$ modulo $\mathbb{K}^{*2}$. On le note $\Delta(B) \in \mathbb{K}^*/\mathbb{K}^{*2}$.

Définition 5.8. La base $\mathcal{B} = (e_1, \dots, e_n)$ est dite *orthogonale* si pour tous $i \neq j$, on a $B(e_i, e_j) = 0$.

Remarques. Ceci équivaut à dire que $M_{\mathcal{B}}(B)$ est diagonale.

On suppose $\mathbb{K}$ de caractéristique 2. Sur $\mathbb{K}^2$, on considère la forme bilinéaire symétrique

$$B : ((x, y), (x', y')) \mapsto xy' + yx'.$$

Elle a pour matrice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ et est non dégénérée. On a, pour tous $x, y \in \mathbb{K}$, $B((x, y), (x, y)) = 2xy = 0$. La forme quadratique associée à B est nulle. Pour une telle forme bilinéaire symétrique, il ne peut y avoir de base orthogonale.

Désormais, on suppose que $\mathbb{K}$ n'est pas de caractéristique 2. On associe à B sa forme quadratique

$$q : E \rightarrow \mathbb{K}, \quad x \mapsto B(x, x).$$

Proposition 5.6. Si $B \neq 0$, alors $q \neq 0$.

Démonstration. Supposons $q = 0$. Alors, pour tout x , $B(x, x) = 0$. Soient $x, y \in E$. On a $0 = B(x + y, x + y) = B(x, x) + 2B(x, y) + B(y, y)$, donc $B(x, y) = 0$ ($\mathbb{K}$ n'est pas de caractéristique 2). $\square$

Corollaire 5.7. B admet une base orthogonale.

Démonstration. Si $B = 0$, toute base convient. Sinon, on écrit $E = \ker B \oplus G$. On a $B|_{G \times G}$ non dégénérée, donc non nulle. Alors q_G est non nulle et on a un vecteur non isotrope $g \in G$. On écrit $G = \mathbb{K}g \oplus H$, avec $H = (\mathbb{K}g)^\perp \cap G$. $B|_{H \times H}$ est non dégénérée. Une récurrence élémentaire permet de conclure. $\square$

Remarque. Une forme quadratique peut être définie comme un polynôme homogène de degré 2 en n variables.

$$q(x_1, \dots, x_n) = \sum_{i=1}^n a_i x_i^2 + \sum_{1 \leq i < j \leq n} a_{i,j} x_i x_j.$$

q admet pour matrice

$$\begin{pmatrix} a_1 & & \frac{1}{2}a_{1,j} \\ & \ddots & \\ \frac{1}{2}a_{i,j} & & a_n \end{pmatrix}.$$

Chercher une base orthogonale pour la forme bilinéaire symétrique associée à q , c'est chercher une base sur laquelle q s'exprime uniquement avec des carrés :

$$q\left(\sum_{i=1}^n y_i f_i\right) = \sum_{i=1}^n b_i y_i^2.$$

L'algorithme de mise sous carrés de Gauss permet d'obtenir une telle base orthogonale. On considère une forme quadratique

$$q : (x_1, \dots, x_n) \mapsto \sum_{i=1}^n a_i x_i^2 + \sum_{1 \leq i < j \leq n} a_{i,j} x_i x_j.$$

- Premier cas : il existe i tel que $a_i \neq 0$. On suppose $i = 1$. Alors

$$\begin{aligned}
 q(x_1, \dots, x_n) &= a_1 x_1^2 + x_1 \sum_{j=2}^n a_{1,j} x_j + \underbrace{\sum_{i=2}^n a_i x_i^2 + \sum_{2 \leq i < j \leq n} a_{i,j} x_i x_j}_{r(x_2, \dots, x_n)} \\
 &= a_1 \left(x_1 + \frac{1}{2} \sum_{j=2}^n \frac{a_{1,j}}{a_1} x_j \right)^2 - \frac{1}{4a_1} \left(\sum_{j=2}^n a_{1,j} x_j \right)^2 + r(x_2, \dots, x_n).
 \end{aligned}$$

$y_1(x_1, \dots, x_n)$ $s(x_2, \dots, x_n)$

On obtient

$$q(x_1, \dots, x_n) = a_1 y_1(x_1, \dots, x_n)^2 + s(x_2, \dots, x_n).$$

On répète le processus sur s . Le changement de coordonnées $y_1 = x_1 + \frac{1}{2a_1} \sum_{j=2}^n a_{1,j} x_j$, $y_j = x_j$ pour $j \geq 2$, est inversible : sa matrice est triangulaire supérieure de diagonale égale à 1.

- Second cas : pour tout i , $a_i = 0$. Il existe $j > i$ tel que $a_{i,j} \neq 0$. On suppose $i = 1$ et $j = 2$. Alors

$$\begin{aligned}
 q(x_1, \dots, x_n) &= a_{1,2} x_1 x_2 + x_1 \sum_{j=3}^n a_{1,j} x_j + x_2 \sum_{j=3}^n a_{2,j} x_j + \underbrace{\sum_{3 \leq i < j \leq n} a_{i,j} x_i x_j}_{r(x_3, \dots, x_n)} \\
 &= a_{1,2} \left(x_1 + \sum_{j=3}^n \frac{a_{2,j}}{a_{1,2}} x_j \right) \left(x_2 + \sum_{j=3}^n \frac{a_{1,j}}{a_{1,2}} x_j \right) \\
 &\quad - \frac{1}{a_{1,2}} \underbrace{\left(\sum_{j=3}^n a_{2,j} x_j \right) \left(\sum_{j=3}^n a_{1,j} x_j \right)}_{s(x_3, \dots, x_n)} + r(x_3, \dots, x_n) \\
 &= \frac{a_{1,2}}{4} \left(\left(x_1 + x_2 + \sum_{j=3}^n \frac{a_{2,j} + a_{1,j}}{a_{1,2}} x_j \right)^2 - \left(x_1 - x_2 + \sum_{j=3}^n \frac{a_{2,j} - a_{1,j}}{a_{1,2}} x_j \right)^2 \right) + s(x_3, \dots, x_n).
 \end{aligned}$$

$y_1(x_1, \dots, x_n)$ $y_2(x_1, \dots, x_n)$

On obtient

$$q(x_1, \dots, x_n) = \frac{a_{1,2}}{4} (y_1(x_1, \dots, x_n)^2 - y_2(x_1, \dots, x_n)^2) + s(x_3, \dots, x_n).$$

On répète le processus sur s . Le changement de coordonnées $(x_1, \dots, x_n) \mapsto (y_1, y_2, x_3, \dots, x_n)$ est inversible : son bloc supérieur gauche est $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, de déterminant $-2 \neq 0$.

Exemple.

$$\begin{aligned}
 q(x, y, z, t) &= xy + xz + yz + zt = (x + z)(y + z) - z^2 + zt \\
 &= \frac{1}{4}(x + y + 2z)^2 - \frac{1}{4}(x - y)^2 - \left(z - \frac{t}{2} \right)^2 + \frac{t^2}{4}.
 \end{aligned}$$

Remarque. L'algorithme s'arrête quand la forme quadratique à réduire est nulle.

A ce stade, on a $q = b_1 y_1^2 + \dots + b_r y_r^2$, où les y_i sont des formes linéaires linéairement indépendantes. On a ainsi une base sur laquelle la matrice de q est

$$\begin{pmatrix} b_1 & & & & \\ & \ddots & & & \\ & & b_r & & \\ & & & 0 & \\ & 0 & & & \ddots \\ & & & & & 0 \end{pmatrix}.$$

r est le rang de q . La méthode de Gauss permet ainsi de déterminer le rang d'une forme quadratique.

5.2 Classification des formes quadratiques

Le groupe $\mathcal{GL}(E)$ opère sur l'ensemble $\mathcal{Q}(E)$ des formes quadratiques sur E de la manière suivante :

$$\mathcal{GL}(E) \times \mathcal{Q}(E) \rightarrow \mathcal{Q}(E), \quad (u, q) \mapsto q \circ u^{-1}.$$

Intéressons-nous aux orbites de cette action.

Définition 5.9. Deux formes quadratiques q et q' sur E sont dites *équivalentes* si elles appartiennent à la même orbite.

5.2.1 Importance de $\mathbb{K}^*/\mathbb{K}^{*2}$

Commençons par traiter le cas de la dimension 1. Si $q : \mathbb{K} \rightarrow \mathbb{K}$ est non nulle, elle s'écrit $q(xe) = ax^2$, avec $a \neq 0$. Un changement de base conduit à $q(u(xe)) = q(xbe) = ab^2 x^2$. Alors deux formes quadratiques non nulles sont équivalentes si et seulement si elles ont même discriminant. Essayons de généraliser ce résultat.

Proposition 5.8. Soit q une forme quadratique non dégénérée sur E . Soit $\mathfrak{R}$ un système de représentants de $\mathbb{K}^*/\mathbb{K}^{*2}$. Alors q admet une matrice de la forme

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

où pour tout i , $\lambda_i \in \mathfrak{R}$.

Démonstration. Soit q une forme quadratique non dégénérée qui admet pour matrice dans une base orthogonale

$$\begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}.$$

Pour tout i , il existe $\beta_i \in \mathfrak{R}$ tel que $\alpha_i \in \beta_i \mathbb{K}^{*2}$. Alors pour tout i , il existe $\gamma_i \in \mathbb{K}^*$ tel que $\alpha_i = \beta_i \gamma_i^2$. On a

$$\begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix} = \begin{pmatrix} \gamma_1 & & 0 \\ & \ddots & \\ 0 & & \gamma_n \end{pmatrix} \begin{pmatrix} \beta_1 & & 0 \\ & \ddots & \\ 0 & & \beta_n \end{pmatrix} \begin{pmatrix} \gamma_1 & & 0 \\ & \ddots & \\ 0 & & \gamma_n \end{pmatrix}.$$

Ainsi $\begin{pmatrix} \beta_1 & & 0 \\ & \ddots & \\ 0 & & \beta_n \end{pmatrix}$ est une matrice de q .

□

Corollaire 5.9. Pour $\dim E \geq 1$, l'ensemble des classes d'équivalence des formes quadratiques non dégénérées sur E est fini si et seulement si $\mathbb{K}^*/\mathbb{K}^{*2}$ est fini.

Démonstration. Si $\mathbb{K}^*/\mathbb{K}^{*2}$ est fini de cardinal m , la proposition précédente montre qu'il y a au plus m^n classes d'équivalence de formes quadratiques non dégénérées.

Réciproquement, le discriminant Δ est une application surjective de l'ensemble des classes d'équivalence de

formes quadratiques non dégénérées sur $\mathbb{K}^*/\mathbb{K}^{*2}$, car $\Delta \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & \alpha \end{pmatrix} = \alpha \mathbb{K}^{*2}$. Dans le diagramme suivant, on

fixe une base $\mathcal{B}$ de E et on pose $\det_{\mathcal{B}}(q) = \det(\text{Mat}_{\mathcal{B}}(q))$:

$$\begin{array}{ccc} \mathcal{Q}(E)_{\text{non dégénérées}} & \xrightarrow{\det_{\mathcal{B}}} & \mathbb{K}^* \\ & \searrow \Delta & \downarrow \\ & & \mathbb{K}^*/\mathbb{K}^{*2} \end{array}$$

□

5.2.2 Cas où $\mathbb{K}^* = \mathbb{K}^{*2}$

Ce cas se produit par exemple quand $\mathbb{K}$ est algébriquement clos (e.g. $\mathbb{K} = \mathbb{C}$). Le corps des nombres complexes constructibles à la règle et au compas fournit un autre exemple. Dans le corps des réels constructibles, en revanche, -1 n'est pas un carré.

On peut prendre $\mathfrak{R} = \{1\}$. La proposition 5.8 donne :

Proposition 5.10. Toute forme quadratique non dégénérée sur E admet l'identité comme matrice. Il y a une unique classe de formes quadratiques non dégénérées.

5.2.3 Cas où $\mathbb{K} = \mathbb{R}$

Dans ce cas, $\mathbb{K}^{*2} = \mathbb{R}^{++}$. On prend $\mathfrak{R} = \{1, -1\}$. La proposition 5.8 conduit à :

Proposition 5.11. Toute forme quadratique non dégénérée sur E admet une matrice de la forme

$$\begin{pmatrix} 1 & & & & \\ & \ddots & & & 0 \\ & & 1 & & \\ & & & -1 & \\ 0 & & & & \ddots \\ & & & & & -1 \end{pmatrix} \begin{matrix} \uparrow \\ r \\ \downarrow \\ \uparrow \\ s = n - r \\ \downarrow \end{matrix}.$$

Définition 5.10. Une forme quadratique q est dite *définie positive* si pour tout x non nul de E , $q(x) > 0$ (on notera $q > 0$). Elle est dite *définie négative* si pour tout x non nul de E , $q(x) < 0$ (on notera $q < 0$).

Proposition 5.12. Supposons que q admette pour matrice $\begin{pmatrix} I_r & 0 \\ 0 & -I_s \end{pmatrix}$. Alors

$$r = \max\{\dim F : q|_F > 0\}, \quad s = \max\{\dim F : q|_F < 0\}.$$

Démonstration. On note $E = F_0 \oplus G_0$, avec $\dim F_0 = r$ (respectivement $\dim G_0 = s$) et $q|_{F_0}$ a pour matrice I_r (respectivement $q|_{G_0}$ a pour matrice $-I_s$). On remarque que $q|_{F_0}$ est définie positive, tandis que $q|_{G_0}$ est définie négative.

Soit F un sous-espace vectoriel tel que $q|_F > 0$. Soit $x \in F \cap G_0$. Alors $q(x) \geq 0$ (car $x \in F$) et $q(x) \leq 0$ (car $x \in G_0$). Donc $q(x) = 0$, puis $x = 0$, car $q|_F$ est définie positive. On a donc $F + G_0 = F \oplus G_0$, ce qui impose $\dim F \leq r$. Par symétrie, on montre que tout sous-espace vectoriel G tel que $q|_G < 0$ vérifie $\dim G \leq s$. $\square$

Définition 5.11. (r, s) est unique. Ce couple est appelé la *signature* de q .

Corollaire 5.13. Deux formes quadratiques non dégénérées sont équivalentes si et seulement si elles ont la même signature.

Corollaire 5.14. Il y a $n + 1$ classes d'équivalences de formes quadratiques non dégénérées sur $\mathbb{R}^n$.

Si les mineurs principaux initiaux $\Delta_i = \det M_{\{1, \dots, i\}, \{1, \dots, i\}}$ sont tous non nuls, posons $\Delta_0 = 1$. La réduction de Gauss donne des coefficients diagonaux Δ_i / Δ_{i-1} . Ainsi, s est le nombre de changements de signe dans $(\Delta_0, \dots, \Delta_n)$ et $r = n - s$. Cette règle ne s'applique pas directement lorsqu'un de ces mineurs s'annule.

5.2.4 Cas où $\mathbb{K} = \mathbb{F}_\ell$, ℓ étant une puissance d'un nombre premier impair

On rappelle que

$$\mathbb{F}_\ell^* \rightarrow \mathbb{F}_\ell^{*2}, \quad x \mapsto x^2$$

est un morphisme surjectif de noyau $\{1, -1\}$. Donc $|\mathbb{F}_\ell^{*2}| = \frac{\ell-1}{2}$ et $\mathbb{F}_\ell^* / \mathbb{F}_\ell^{*2}$ est d'ordre 2. On choisit $\alpha \in \mathbb{F}_\ell^* \setminus \mathbb{F}_\ell^{*2}$ et on prend $\mathfrak{R} = \{1, \alpha\}$. La proposition 5.8 assure que :

Proposition 5.15. Toute forme quadratique non dégénérée admet une matrice de la forme

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \alpha & & \\ & 0 & & & \ddots & \\ & & & & & \alpha \end{pmatrix}.$$

Lemme 5.16. Soit q une forme quadratique de rang ≥ 2 . Alors pour tout $c \in \mathbb{F}_\ell^*$, il existe $e \in E$ tel que $q(e) = c$.

Démonstration. A partir d'une base orthogonale de E pour q , on peut trouver une partie libre $\{u, v\}$ telle que $q(xu + yv) = ax^2 + by^2$ avec a et b non nuls. On considère les deux sous-ensembles de $\mathbb{F}_\ell$ suivants : $A = \{ax^2, x \in \mathbb{F}_\ell\}$ et $B = \{c - by^2, y \in \mathbb{F}_\ell\}$. Comme a et b sont non nuls, ces ensembles sont en bijection avec l'ensemble des carrés de $\mathbb{F}_\ell$, qui est de cardinal $\frac{\ell+1}{2}$. Comme $\frac{\ell+1}{2} + \frac{\ell+1}{2} > \ell$, on a $A \cap B \neq \emptyset$: il existe $x, y \in \mathbb{F}_\ell$ tels que $ax^2 = c - by^2$. $\square$

Lemme 5.17. Toute forme quadratique de rang ≥ 3 admet un vecteur isotrope.

Démonstration. En effet, à partir d'une base orthogonale de E pour q , on peut trouver une partie libre $\{u, v, w\}$ telle que $q(xu + yv + zw) = ax^2 + by^2 + cz^2$, avec a, b et c non nuls. En reprenant le raisonnement mené pour démontrer le lemme, on obtient qu'il existe x et y tels que $ax^2 + by^2 + c = 0$, ce qui implique que le vecteur $xu + yv + w$ est isotrope. $\square$

Corollaire 5.18. Une forme quadratique de matrice $\begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}$ admet pour matrice l'identité.

Démonstration. Soit q de matrice $\begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}$. q est une forme quadratique non dégénérée sur un plan. D'après le lemme, il existe un vecteur e tel que $q(e) = 1$. e est non isotrope, donc $\mathbb{F}_\ell^2 = \langle e \rangle \oplus \langle e \rangle^\perp$. Soit f un vecteur non nul de $\langle e \rangle^\perp$. Sur la base (e, f) , la matrice de q est $\begin{pmatrix} 1 & 0 \\ 0 & q(f) \end{pmatrix}$. En considérant le discriminant de q , on remarque que $q(f) \in \alpha^2 \mathbb{F}_\ell^{*2} = \mathbb{F}_\ell^{*2}$. Il existe donc $\beta \in \mathbb{F}_\ell^*$ tel que $q(f) = \beta^2$. q admet l'identité pour matrice sur la base $(e, \frac{f}{\beta})$. $\square$

Corollaire 5.19. Toute forme quadratique non dégénérée admet comme matrice :

- si $\Delta(q) = \mathbb{F}_\ell^{*2}$,

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

- si $\Delta(q) \neq \mathbb{F}_\ell^{*2}$,

$$\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 & \\ & & & \alpha \end{pmatrix}.$$

Corollaire 5.20. Deux formes quadratiques non dégénérées de même dimension sont équivalentes si et seulement si elles ont le même discriminant.

Corollaire 5.21. Pour $n \geq 1$, il y a deux classes de formes quadratiques non dégénérées sur $\mathbb{F}_\ell^n$.

5.2.5 Quelques mots du cas où $\mathbb{K} = \mathbb{Q}$

On commence par remarquer que $\mathbb{Q}^*/\mathbb{Q}^{*2}$ est infini. On peut prendre $\mathfrak{R} = \{\varepsilon p_1 \dots p_k, \varepsilon = \pm 1, k \in \mathbb{N}, p_1, \dots, p_k \text{ premiers distincts}\}$.

On montre qu'à équivalence près, les valeurs absolues non triviales sur $\mathbb{Q}$ sont la valeur absolue usuelle $|\cdot|$ et les valeurs absolues ultramétriques $|\cdot|_p$ définies par $|n|_p = p^{-v_p(n)}$, pour $n \in \mathbb{Q}^*$, avec $|0|_p = 0$, où $v_p(n)$ est l'exposant de p dans la décomposition du rationnel non nul n (théorème d'Ostrowski). Les complétés de $\mathbb{Q}$ pour ces différentes valeurs absolues conduisent à $\mathbb{R}$ (pour la valeur absolue usuelle) d'une part, et aux corps p -adiques $\mathbb{Q}_p$ (pour la valeur absolue p -adique) d'autre part (voir [1]).

Sur $\mathbb{Q}_p$, les formes quadratiques non dégénérées sont classifiées par leur dimension, leur discriminant et leur invariant de Hasse (défini à l'aide du symbole de Hilbert; voir [12]). Le seul groupe des classes de carrés ne suffit pas à cette classification.

Le théorème suivant permet la classification des formes quadratiques sur un $\mathbb{Q}$ -espace vectoriel.

Théorème 5.22 (Hasse–Minkowski). Deux formes quadratiques sur $\mathbb{Q}$ sont équivalentes si et seulement si elles sont équivalentes sur $\mathbb{R}$ et sur $\mathbb{Q}_p$, pour tout p .

On consultera avec profit [12].

5.3 Isotropie

5.3.1 Plan hyperbolique

Définition 5.12. Soit P un plan et q une forme quadratique sur P . On dit que (P, q) est un *plan hyperbolique* s'il existe une base de P sur laquelle q est $(x, y) \mapsto x^2 - y^2$.

Les vecteurs isotropes sont les vecteurs $(x, y) \neq (0, 0)$ tels que $x = \pm y$. On a exactement deux droites isotropes : $y = x$ et $y = -x$.

Proposition 5.23. (P, q) est hyperbolique si et seulement si l'une des assertions équivalentes suivantes est vérifiée :

(i). q admet pour matrice

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

(ii). q admet pour matrice

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

(iii). q est non dégénérée et admet au moins une droite isotrope.

(iv). q est non nulle et admet deux droites isotropes distinctes.

Démonstration. ((i).) $\Leftrightarrow$ ((ii).). On remarque simplement que

$$x^2 - y^2 = 2 \underbrace{\frac{x-y}{2}}_X \underbrace{(x+y)}_Y.$$

((ii).) $\Rightarrow$ ((iii).). $\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1 \neq 0$, donc q est non dégénérée. Le premier vecteur de base engendre une droite isotrope.

((iii).) $\Rightarrow$ ((iv).) et ((ii).). Soit e un vecteur isotrope. Comme q est non dégénérée, il existe $g \in P$ tel que $B(e, g) \neq 0$. Quitte à diviser g par $B(e, g)$, on peut supposer $B(e, g) = 1$. Comme $q(e) = B(e, e) = 0$, (e, g) est une base de P . Comme pour tout λ , $q(g + \lambda e) = q(g) + 2\lambda$, il existe λ tel que $q(g + \lambda e) = 0$. La matrice de q sur la base $(e, g + \lambda e)$ est alors $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

((iv).) $\Rightarrow$ ((iii).). Soient e et f deux vecteurs non nuls respectivement sur chaque droite isotrope. (e, f) est une base de P sur laquelle la matrice de q est $\begin{pmatrix} 0 & \lambda \\ \lambda & 0 \end{pmatrix}$, avec $\lambda = B(e, f)$. Comme $q \neq 0$, on a $\lambda \neq 0$. Le déterminant de cette matrice vaut $-\lambda^2 \neq 0$, donc q est non dégénérée. $\square$

5.3.2 Espaces hyperboliques, sous-espaces totalement isotropes, indice

Définition 5.13. (E, q) est dit *hyperbolique* si q admet comme matrice

$$\begin{pmatrix} I_r & 0 \\ 0 & -I_r \end{pmatrix},$$

avec $\dim E = 2r$.

Remarques. Un espace hyperbolique est de dimension paire.
Une forme quadratique hyperbolique est non dégénérée.

Proposition 5.24. Les assertions suivantes sont équivalentes :

- (i). (E, q) est hyperbolique.
- (ii). $E = P_1 \oplus \dots \oplus P_r$, où les P_i sont des plans hyperboliques orthogonaux deux à deux.
- (iii). q admet la matrice

$$\begin{pmatrix} 0 & I_r \\ I_r & 0 \end{pmatrix}.$$

- (iv). q est non dégénérée et il existe un sous-espace totalement isotrope de dimension $\frac{\dim E}{2}$.

Démonstration. ((i).) $\Leftrightarrow$ ((ii).). Il suffit de changer l'ordre des vecteurs d'une base dans laquelle q a pour matrice $\begin{pmatrix} I_r & 0 \\ 0 & -I_r \end{pmatrix}$ pour obtenir une base dans laquelle q a pour matrice $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^{\oplus r}$, et réciproquement.

((ii).) $\Leftrightarrow$ ((iii).). Sur chaque plan hyperbolique, une forme quadratique qui admet pour matrice $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ admet aussi $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, et réciproquement. Quitte à changer l'ordre des vecteurs de base, on a l'équivalence.

((iii).) $\Rightarrow$ ((iv).). Si q admet la matrice $\begin{pmatrix} 0 & I_r \\ I_r & 0 \end{pmatrix}$, les r premiers vecteurs de base engendrent un sous-espace totalement isotrope de dimension $r = \frac{\dim E}{2}$. En outre, le déterminant de cette matrice est non nul, donc q est non dégénérée.

((iv).) $\Rightarrow$ ((iii).). La démonstration nécessite le lemme suivant.

Lemme 5.25. Soit q une forme quadratique non dégénérée. Soit F un sous-espace totalement isotrope. Alors il existe un sous-espace vectoriel H tel que

- (i). $F \subset H$.
- (ii). $\dim H = 2 \dim F$.
- (iii). (H, q) est hyperbolique.

Démonstration. On raisonne par récurrence sur $k = \dim F$. Si $k = 0$, le sous-espace $H = \{0\}$ convient. Supposons maintenant $k \geq 1$ et le résultat établi pour tout sous-espace totalement isotrope de dimension strictement inférieure à k , dans tout espace quadratique non dégénéré.

Soit $(e_1, \dots, e_k)$ une base de F . Comme F est totalement isotrope, on a $B(e_i, e_j) = 0$ pour tous i, j . Nous allons construire un plan hyperbolique contenant e_k et orthogonal aux vecteurs $e_1, \dots, e_{k-1}$.

Construisons f_k . Comme q est non dégénérée, tout sous-espace vectoriel G de E vérifie $\dim G + \dim G^\perp = \dim E$. Donc $\langle e_1, \dots, e_k \rangle^\perp \subsetneq \langle e_1, \dots, e_{k-1} \rangle^\perp$. Il existe alors $g \in \langle e_1, \dots, e_{k-1} \rangle^\perp \setminus \langle e_1, \dots, e_k \rangle^\perp$. On obtient ainsi, pour tout $i < k$, $B(g, e_i) = 0$ et $B(g, e_k) \neq 0$. Quitte à diviser g par $B(g, e_k)$, on peut supposer en outre $B(g, e_k) = 1$. Pour tout λ , on a $q(g + \lambda e_k) = q(g) + 2\lambda$. Il existe donc λ_0 tel que $q(g + \lambda_0 e_k) = 0$. On pose $f_k = g + \lambda_0 e_k$. On a ainsi $B(e_i, f_k) = \delta_{i,k}$ et $q(f_k) = 0$.

On remarque alors que le plan $P = \langle e_k, f_k \rangle$ est hyperbolique. Donc $q|_P$ est non dégénérée et $E = P \oplus P^\perp$. $\langle e_1, \dots, e_{k-1} \rangle$ est un sous-espace totalement isotrope de $P^\perp$ de dimension $k - 1$. $q|_{P^\perp}$ est non dégénérée, si bien que l'on peut appliquer l'hypothèse de récurrence. On obtient un sous-espace hyperbolique $H' \subset P^\perp$ contenant $\langle e_1, \dots, e_{k-1} \rangle$ et de dimension $2(k - 1)$. La somme orthogonale $H = P \oplus H'$ est alors hyperbolique, contient F et a pour dimension $2k$, ce qui achève la récurrence. $\square$

Reprenons la démonstration de la proposition. Dans le cas particulier où $\dim F = \frac{\dim E}{2}$, le sous-espace vectoriel H dont l'existence est assurée par le lemme est égal à E pour des raisons de dimension. $\square$

Définition 5.14. Un sous-espace totalement isotrope de dimension $\frac{\dim E}{2}$ est appelé un *Lagrangien*.

Définition 5.15. Soit q une forme quadratique non dégénérée. On appelle *indice* de q et on note $v(q)$ le maximum des dimensions des sous-espaces totalement isotropes de E .

Remarque. On a $v(q) \leq \lfloor \frac{\dim E}{2} \rfloor$. En outre, $v(q) = \frac{\dim E}{2}$ si et seulement si (E, q) est hyperbolique.

Proposition 5.26. Soit q une forme quadratique non dégénérée. Alors on peut écrire $E = H \oplus A$, la somme étant orthogonale, où H et A vérifient

- (i). H est hyperbolique de dimension $2v(q)$.
- (ii). $q|_A$ est anisotrope (i.e. $q|_A(x) = 0 \Rightarrow x = 0$).

Démonstration. Soit F un sous-espace totalement isotrope de dimension $v(q)$. Soit H un sous-espace vectoriel hyperbolique contenant F et de dimension $2 \dim F$ (voir lemme 5.25). Comme q est non dégénérée, on a $E = H \oplus H^\perp$. On pose $A = H^\perp$. Reste à montrer que $q|_A$ est anisotrope. Soit $x \in A$ tel que $q(x) = 0$. On a $B(x, F) = 0$. En effet, $x \in H^\perp \subset F^\perp$. Comme $q(x) = 0$, $q|_{\langle x, F \rangle} = 0$. Par maximalité de F , on a $x \in F$. Or $F \cap A \subset H \cap A = \{0\}$. Donc $x = 0$. $\square$

Remarque. On a en fait un résultat d'unicité (voir [11]) :

- (i). Tous les sous-espaces vectoriels totalement isotropes maximaux sont de dimension $v(q)$.
- (ii). Si on a deux décompositions $E = H \oplus A$ et $E = H' \oplus A'$, alors il existe un automorphisme orthogonal u tel que $u(H) = H'$ et $u(A) = A'$.
- (iii). La classe d'équivalence de $q|_A$ est unique.

5.3.3 Calcul de l'indice d'une forme quadratique dans les cas usuels

Dans cette partie, nous montrons comment calculer l'indice d'une forme quadratique non dégénérée dans les cas de corps déjà traités précédemment. On peut ainsi obtenir les classes d'équivalence de formes quadratiques non dégénérées puisque deux formes quadratiques non dégénérées sont équivalentes si et seulement si elles ont même indice et même forme anisotrope à équivalence près.

Cas où $\mathbb{K} = \mathbb{K}^2$

Lemme 5.27. Toute forme quadratique de rang ≥ 2 admet un vecteur isotrope.

Démonstration. On se ramène à $q(x, y) = ax^2 + by^2$, avec $a \neq 0$ et $b \neq 0$. On sait que $a + by^2 = 0$ admet une solution y_0 . Donc $q(1, y_0) = 0$. $\square$

Soit q une forme quadratique non dégénérée sur E . Soit F un sous-espace totalement isotrope maximal de E . Soit H un sous-espace vectoriel hyperbolique contenant F . Soit $A = H^\perp$. Alors, comme $q|_A$ est anisotrope, soit $A = \{0\}$, soit $\dim A = 1$. La parité de la dimension permet de savoir dans quel cas on se situe : plus précisément, $\dim A = \dim E \pmod{2}$.

- Si $\dim E$ est paire, $A = \{0\}$, E est hyperbolique et $\dim F = \frac{\dim E}{2}$.
- Si $\dim E$ est impaire, $\dim A = 1$ et $\dim F = \frac{\dim E - 1}{2} = \lfloor \frac{\dim E}{2} \rfloor$. On montre alors qu'il y a une unique classe de formes quadratiques de rang 1 ($ax^2 = (bx)^2$, où $a \neq 0$ et b est tel que $a = b^2$).

Cas où $\mathbb{K} = \mathbb{R}$

Soit F un sous-espace totalement isotrope maximal dont la dimension est notée ℓ . Soit $H \subset E$ un sous-espace hyperbolique de dimension 2ℓ contenant F , fourni par le lemme 5.25. Soit $A = H^\perp$.

Remarque. Soit q une forme quadratique non dégénérée. Comme q est continue sur E , qui est un $\mathbb{R}$ -espace vectoriel de dimension finie, si q est anisotrope, q est soit définie positive, soit définie négative. En chaque dimension strictement positive fixée, il y a deux classes de formes quadratiques anisotropes.

$q|_A$ admet donc pour matrice εI_m , où $\varepsilon = \pm 1$ et $m = \dim A$. On en déduit que q admet pour matrice

$$\begin{pmatrix} I_\ell & & 0 \\ & -I_\ell & \\ 0 & & \varepsilon I_m \end{pmatrix},$$

soit encore la matrice

$$\begin{pmatrix} I_r & \\ & -I_s \end{pmatrix}.$$

En convenant que $\varepsilon = 1$ lorsque $m = 0$, ℓ , m et ε sont propres à la classe d'équivalence de formes quadratiques considérée. On peut établir une correspondance bijective entre (ℓ, m, ε) et (r, s) , ce qui implique l'unicité de r et s . r est également caractérisé par $r = \max(\dim G, q|_G > 0)$.

Cas où $\mathbb{K} = \mathbb{F}_\ell$

D'après le lemme 5.17, on sait que toute forme quadratique de rang ≥ 3 admet un vecteur isotrope.

Traisons le cas de la dimension 2. On prend q une forme quadratique non dégénérée de rang 2. D'après le lemme 5.16, elle atteint la valeur 1. Donc elle admet $\begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix}$ comme matrice, où α est défini modulo $\mathbb{F}_\ell^{*2}$.

Il existe $(x, y) \neq (0, 0)$ tel que $x^2 + \alpha y^2 = 0$ si et seulement s'il existe $x \neq 0$ et $y \neq 0$ tels que $\alpha = -\left(\frac{x}{y}\right)^2$ si et seulement si $-\alpha \in \mathbb{F}_\ell^{*2}$. Donc q admet un vecteur isotrope si et seulement si $-\alpha \in \mathbb{F}_\ell^{*2}$. Si $-\alpha \in \mathbb{F}_\ell^{*2}$, alors q admet un vecteur isotrope et on a un plan hyperbolique et $\alpha = -1 \pmod{\mathbb{F}_\ell^{*2}}$. Sinon, q est anisotrope et $\alpha \neq -1 \pmod{\mathbb{F}_\ell^{*2}}$.

Revenons au cas général. Soit F un sous-espace totalement isotrope maximal de dimension d . Soit H un hyperbolique de dimension $2d$ contenant F . Soit $A = H^\perp$. $q|_A$ est anisotrope.

- Si $\dim E$ est impaire, $\dim A = 1$. On obtient ainsi $\dim F = \frac{\dim E - 1}{2} = \lfloor \frac{\dim E}{2} \rfloor$. On classifie ensuite les formes quadratiques en dimension 1. Elles sont classifiées par $\mathbb{F}_\ell^*/\mathbb{F}_\ell^{*2}$, via leur discriminant. On a donc deux classes. Le discriminant permet de savoir dans quelle classe on se trouve.
- Si $\dim E$ est paire, $\dim A \in \{0, 2\}$. Si $\dim A = 2$, $q|_A$ est dans la classe anisotrope ; si $\dim A = 0$, q est hyperbolique.
 - Si q est hyperbolique, $\dim A = 0$ et $\dim F = \frac{\dim E}{2}$.
 - Sinon $\dim A = 2$ et $\dim F = \frac{\dim E}{2} - 1$.

Sous-groupes finis de $O_2(\mathbb{R})$ et $O_3(\mathbb{R})$

Références : [2, 3, 4]

Pour une forme quadratique q sur E , $O(E, q)$ désigne le groupe des automorphismes préservant q . Pour la norme euclidienne usuelle, on écrit $O_n(\mathbb{R}) = \{M : {}^tMM = I_n\}$ et $SO_n(\mathbb{R}) = \{M \in O_n(\mathbb{R}) : \det M = 1\}$. Les signes $+$ dans G^+ désignent les isométries de déterminant 1.

Soit G un sous-groupe fini de $GL_n(\mathbb{R})$. Soit q une forme quadratique définie positive. On introduit une forme quadratique définie positive Q en posant pour tout $x \in \mathbb{R}^n$,

$$Q(x) = \sum_{g \in G} q(g(x)).$$

Alors $G \subset O(\mathbb{R}^n, Q)$. Tout élément de G est en effet une isométrie pour Q : pour tout $h \in G$ et tout $x \in \mathbb{R}^n$,

$$Q(h(x)) = \sum_{g \in G} q(gh(x)) = \sum_{k \in G} q(k(x)) = Q(x).$$

Tout sous-groupe fini de $GL_2(\mathbb{R})$ (*respectivement* $GL_3(\mathbb{R})$) est donc conjugué à un sous-groupe de $O_2(\mathbb{R})$ (*respectivement* $O_3(\mathbb{R})$).

6.1 Sous-groupes finis de $O_2(\mathbb{R})$

On rappelle que les éléments de $O_2(\mathbb{R})$ sont :

- les rotations, qui constituent exactement les éléments du groupe $SO_2(\mathbb{R}) = \{U \in O_2(\mathbb{R}) : \det U = 1\}$ et qui sont

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \text{ avec } \theta \in]-\pi, \pi].$$

- les symétries orthogonales par rapport à une droite, qui sont

$$\begin{pmatrix} a & b \\ b & -a \end{pmatrix}, \text{ avec } a^2 + b^2 = 1.$$

On a un isomorphisme de groupes

$$\mathbb{U} \rightarrow SO_2(\mathbb{R}), \quad e^{i\theta} \mapsto \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

En particulier, $SO_2(\mathbb{R})$ est un groupe abélien.

6.1.1 Sous-groupes finis de $SO_2(\mathbb{R})$

Proposition 6.1. *Les sous-groupes finis de $SO_2(\mathbb{R})$ sont cycliques.*

Pour chaque entier naturel $n \geq 1$, il y a exactement un sous-groupe fini d'ordre n .

Démonstration. $SO_2(\mathbb{R})$ est isomorphe à $\mathbb{U}$. Comme $\mathbb{U}$ est un sous-groupe du groupe multiplicatif du corps $\mathbb{C}$, ses sous-groupes finis sont cycliques.

Si $n \geq 1$, l'unique sous-groupe fini d'ordre n est alors formé par l'ensemble des rotations d'angle $\frac{2k\pi}{n}$, qui est engendré par la rotation d'angle $\frac{2\pi}{n}$. □

6.1.2 Sous-groupes finis de $O_2(\mathbb{R})$ non contenus dans $SO_2(\mathbb{R})$

Proposition 6.2. *Un sous-groupe fini de $O_2(\mathbb{R})$ non contenu dans $SO_2(\mathbb{R})$ est diédral. Pour chaque entier naturel $n \geq 1$, $O_2(\mathbb{R})$ contient, à conjugaison près, un unique groupe diédral d'ordre $2n$. Pour $n \geq 3$, il est réalisé par les isométries d'un polygone régulier à n sommets. Pour $n = 1$, il est engendré par une réflexion ; pour $n = 2$, par deux réflexions d'axes perpendiculaires.*

Démonstration. Soit G un sous-groupe fini de $O_2(\mathbb{R})$ non contenu dans $SO_2(\mathbb{R})$. On note G^+ le sous-groupe de G formé par les isométries positives : $G^+ = G \cap SO_2(\mathbb{R})$. Le morphisme $\det : G \rightarrow \{\pm 1\}$ est surjectif et son noyau est G^+ . Donc $[G : G^+] = 2$ et $G^+ \triangleleft G$. Soit $\sigma \in G \setminus G^+$. On sait que G^+ est cyclique. Soit ρ un de ses générateurs et m son ordre. Alors $G = \langle \rho \rangle \rtimes \langle \sigma \rangle$. On obtient ainsi le groupe diédral D_{2m} (voir partie 1.3.3).

Les réflexions du plan sont conjuguées par une rotation, et le sous-groupe des rotations d'ordre m est unique : cela donne l'unicité à conjugaison près. Si $m \geq 3$, choisissons ρ d'angle $2\pi/m$ et un vecteur unitaire u sur l'axe de σ . Le polygone

$$\mathfrak{P}_m = \bigcup_{i=0}^{m-1} [\rho^i(u), \rho^{i+1}(u)]$$

est conservé par G , et son groupe d'isométries est exactement G . □

6.2 Sous-groupes finis de $O_3(\mathbb{R})$

Dans toute la suite, les polyèdres réguliers sont centrés à l'origine. On rappelle que toute isométrie positive admet une matrice de la forme

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

Si cette isométrie ρ n'est pas l'identité, c'est une rotation d'axe $\ker(\rho - \text{Id})$ et d'angle θ dans la base orthonormée orientée choisie. Nous conseillons au lecteur de se reporter à [2, 3, 4].

6.2.1 Sous-groupes finis de $SO_3(\mathbb{R})$

Proposition 6.3. *$SO_3(\mathbb{R})$ admet des sous-groupes isomorphes à $\mathbb{Z}/n\mathbb{Z}$, D_{2n} , $\mathfrak{A}_4$, $\mathfrak{S}_4$ et $\mathfrak{A}_5$.*

Démonstration. • $\mathbb{Z}/n\mathbb{Z}$: Soit P un plan. On construit un groupe isomorphe à $\mathbb{Z}/n\mathbb{Z}$ en considérant les rotations d'axe $P^\perp$ et d'angle un multiple de $\frac{2\pi}{n}$.

• D_{2n} : Soit P un plan. On sait que $O(P) \simeq O_2(\mathbb{R})$ contient un sous-groupe G isomorphe à D_{2n} . On prolonge chaque élément g de ce groupe à $\mathbb{R}^3$ en posant

$$\hat{g}(x) = \begin{cases} (\det g)x & \text{si } x \in P^\perp \\ g(x) & \text{si } x \in P. \end{cases}$$

On pose $\hat{G} = \{\hat{g}, g \in G\}$. Alors $\hat{G} \simeq G$ et $\hat{G} < SO_3(\mathbb{R})$.

• $\mathfrak{A}_4$: On dispose de la proposition suivante.

Proposition 6.4. *Le groupe des isométries du tétraèdre régulier est isomorphe à $\mathfrak{S}_4$.*

Démonstration. Soit $\mathfrak{T} = \{A, B, C, D\}$ un tétraèdre régulier. On note O son isobarycentre. Soit $\mathcal{G} = \{g \in O_3(\mathbb{R}) : g(\mathfrak{T}) = \mathfrak{T}\}$. $\mathcal{G}$ opère sur $\mathfrak{T}$, donc il existe un morphisme de groupes $\varphi : \mathcal{G} \rightarrow \mathfrak{S}_{\mathfrak{T}} \simeq \mathfrak{S}_4$.

Montrons que φ est un isomorphisme :

• φ est injectif. Soit en effet $f \in O_3(\mathbb{R})$ tel que $f(A) = A$, $f(B) = B$, $f(C) = C$ et $f(D) = D$. f admet quatre points fixes non coplanaires, il s'agit donc de l'identité.

- φ est surjectif. La transposition (A, B) et le 3-cycle (B, C, D) engendrent $\mathfrak{S}_4$. Il reste à remarquer que (A, B) est l'image par φ de la réflexion par rapport au plan (OCD) , et que (B, C, D) est l'image par φ de la rotation d'axe (OA) et d'angle $\frac{2\pi}{3}$.

□

Corollaire 6.5. *Le groupe des isométries positives du tétraèdre régulier est isomorphe à $\mathfrak{A}_4$.*

- $\mathfrak{S}_4$: Nous allons montrer la proposition suivante.

Proposition 6.6. *Le groupe des isométries positives du cube (voir figure 6.1) est isomorphe à $\mathfrak{S}_4$.*

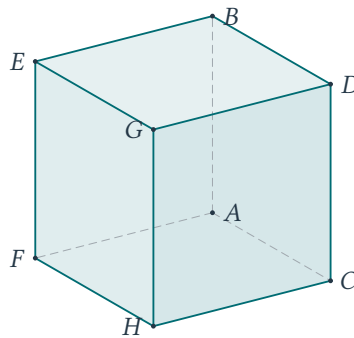


FIGURE 6.1 : Le cube

Démonstration. Soit $\mathfrak{C} = \{A, B, C, D, E, F, G, H\}$ un cube de côté de longueur α . Toute isométrie affine conservant $\mathfrak{C}$ fixe son isobarycentre O . Soit $\mathcal{G} = \{g \in O_3(\mathbb{R}) : g(\mathfrak{C}) = \mathfrak{C}\}$.

Un segment reliant deux sommets distincts peut être :

- une arête (les deux sommets sont consécutifs) (de longueur α),
- la diagonale d'une face (de longueur $\alpha\sqrt{2}$),
- une grande diagonale (de longueur $\alpha\sqrt{3}$).

Comme une isométrie conserve les distances, on remarque que tout élément de $\mathcal{G}$ envoie une grande diagonale sur une grande diagonale. $\mathcal{G}$ opère donc sur l'ensemble des grandes diagonales $\Delta = \{(AG), (BH), (CE), (DF)\}$. On en déduit un morphisme de groupes $\mathcal{G} \rightarrow \mathfrak{S}_\Delta \simeq \mathfrak{S}_4$. Celui-ci induit par restriction un morphisme de $\mathcal{G}^+ = \mathcal{G} \cap SO_3(\mathbb{R})$ dans $\mathfrak{S}_\Delta$, noté φ .

Montrons que φ est un isomorphisme :

- φ est injectif. Prenons $g \in \ker \varphi$. Pour tout $\delta \in \Delta$, $g(\delta) = \delta$, donc il existe $\varepsilon_\delta = \pm 1$ tel que $g_\delta = \varepsilon_\delta \text{Id}$. On suppose qu'il existe deux droites distinctes δ_1 et δ_2 de Δ telles que $\varepsilon_{\delta_1} = \varepsilon_{\delta_2} = 1$ et qu'il existe deux droites distinctes δ_3 et δ_4 de Δ telles que $\varepsilon_{\delta_3} = \varepsilon_{\delta_4} = -1$. $\ker(g - \text{Id})$ (respectivement $\ker(g + \text{Id})$) contient le plan (δ_1, δ_2) (respectivement (δ_3, δ_4)), donc est de dimension au moins 2. Comme $\ker(g - \text{Id}) \cap \ker(g + \text{Id}) = \{0\}$, $\ker(g - \text{Id}) + \ker(g + \text{Id})$ est au moins de dimension 4, ce qui est absurde. On en déduit qu'il existe trois droites distinctes δ_1, δ_2 et δ_3 de Δ telles que $\varepsilon_{\delta_1} = \varepsilon_{\delta_2} = \varepsilon_{\delta_3} = 1$ ou qu'il existe trois droites distinctes δ_1, δ_2 et δ_3 de Δ telles que $\varepsilon_{\delta_1} = \varepsilon_{\delta_2} = \varepsilon_{\delta_3} = -1$. Les trois droites δ_1, δ_2 et δ_3 étant non coplanaires, on en déduit que $g = \pm \text{Id}$. Comme $\det g = 1$ en dimension 3, on a $g = \text{Id}$.
- φ est surjectif. Dans des coordonnées où les sommets sont $(\pm 1, \pm 1, \pm 1)$, les matrices de permutation signées de déterminant 1 donnent $3! 2^2 = 24$ rotations distinctes du cube. L'injection précédente est donc une bijection sur $\mathfrak{S}_4$.

□

Corollaire 6.7. *Le groupe des isométries du cube est isomorphe à $\mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}$.*

Démonstration. $-\text{Id} \in G \setminus G^+$ et $-\text{Id}$ commute avec G^+ . Donc $\langle G^+, -\text{Id} \rangle \simeq G^+ \times \{\pm \text{Id}\}$. Pour des raisons de cardinalité, $G = \langle G^+, -\text{Id} \rangle$. D'où le résultat. $\square$

Définition 6.1. Soit $\mathfrak{P}$ un polyèdre régulier convexe centré à l'origine. On appelle ici *polyèdre dual* (à homothétie près) de $\mathfrak{P}$ le polyèdre dont les sommets sont les isobarycentres respectifs des faces de $\mathfrak{P}$.

Exemples. • Le tétraèdre est autodual.

- Le dual du cube est l'octaèdre, et réciproquement.
- Le dual de l'icosaèdre (voir figure 6.2) est le dodécaèdre (voir figure 6.3), et réciproquement.

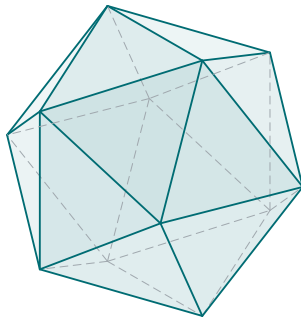


FIGURE 6.2 : L'icosaèdre

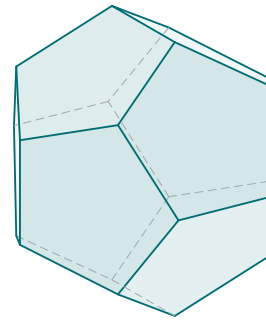


FIGURE 6.3 : Le dodécaèdre

Proposition 6.8. *Soit $\mathfrak{P}$ un polyèdre régulier convexe centré à l'origine. Le groupe des isométries de $\mathfrak{P}$ est isomorphe au groupe des isométries du dual de $\mathfrak{P}$.*

- $\mathfrak{A}_5$: Nous allons montrer la proposition suivante.

Proposition 6.9. *Soit G le groupe des isométries conservant un dodécaèdre (respectivement un icosaèdre), alors $G \simeq \mathfrak{A}_5 \times \mathbb{Z}/2\mathbb{Z}$. En particulier, $G^+ \simeq \mathfrak{A}_5$.*

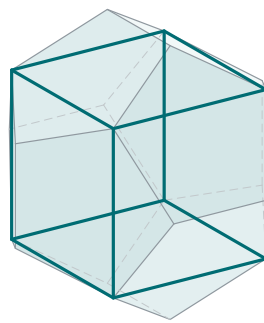


FIGURE 6.4 : Un cube inscrit dans un dodécaèdre

Démonstration. Le dodécaèdre contient cinq cubes dont les sommets sont des sommets du polyèdre (voir figure 6.4). Pour identifier son groupe, nous utiliserons le comptage des rotations et l'argument par conjugaison du cas V ci-dessous, qui ne dépend pas de la présente proposition.

Cherchons les éléments de G^+ :

- il y a l'identité Id ,
- il y a quinze éléments d'ordre 2 : les rotations d'angle π et d'axe une droite passant par les milieux respectifs de deux arêtes opposées,

- il y a vingt éléments d'ordre 3 : les rotations d'angle $\frac{2k\pi}{3}$, avec $k \in \{1, 2\}$, et d'axe une droite passant par deux sommets opposés,
- il y a vingt-quatre éléments d'ordre 5 : les rotations d'angle $\frac{2k\pi}{5}$, avec $k \in \{1, 2, 3, 4\}$, et d'axe une droite passant par les centres respectifs de deux faces opposées.

Ces rotations donnent 60 éléments distincts. Il n'y en a pas davantage : une rotation est déterminée par l'image d'une face (douze choix) et d'un de ses sommets (cinq choix), ce qui donne au plus 60 possibilités. Les axes de chaque type forment une orbite, par régularité du polyèdre. L'argument par classes de conjugaison et sous-groupes de Sylow donné au cas V ci-dessous s'applique donc et montre $G^+ \simeq \mathfrak{A}_5$. Comme $-\text{Id} \in G$ est central, $G \simeq G^+ \times \{\pm \text{Id}\}$. $\square$

 $\square$

La proposition 6.3 assure que $SO_3(\mathbb{R})$ admet des sous-groupes isomorphes à $\mathfrak{A}_4$, $\mathfrak{S}_4$, $\mathfrak{A}_5$, $\mathbb{Z}/n\mathbb{Z}$ et D_{2n} , n désignant un entier naturel non nul. Le théorème suivant montre que ce sont les seuls sous-groupes finis possibles.

Théorème 6.10. *Soit G un sous-groupe fini de $SO_3(\mathbb{R})$. Alors G est isomorphe à $\mathfrak{A}_4$, $\mathfrak{S}_4$, $\mathfrak{A}_5$, $\mathbb{Z}/n\mathbb{Z}$ ou D_{2n} , pour un certain entier naturel non nul n .*

Démonstration. Le groupe trivial est cyclique. Supposons donc $|G| > 1$ et commençons par une définition.

Définition 6.2. Soit $x \in \mathbb{R}^3$. On dit que x est un *pôle* de G si $\|x\| = 1$ et il existe $g \in G \setminus \{\text{Id}\}$ tel que $g(x) = x$.

On pose $n = |G|$. Soit X l'ensemble des pôles de G . On remarque que si $x \in X$, pour tout $g \in G$, $g(x) \in X$. En effet, comme $g \in SO_3(\mathbb{R})$, $\|g(x)\| = \|x\| = 1$. En outre, il existe $h \in G \setminus \{\text{Id}\}$ tel que $h(x) = x$, donc $ghg^{-1}(g(x)) = g(x)$. Observer que $ghg^{-1} \neq \text{Id}$ suffit pour conclure. On en déduit que G agit sur X de la manière suivante :

$$G \times X \rightarrow X, \quad (g, x) \mapsto g(x).$$

Soit $x \in X$. Comme x est un pôle, par définition, $\text{Stab}_G(x) \supsetneq \{\text{Id}\}$. Soit $g \in \text{Stab}_G(x)$. Comme $g(x) = x$, $g(\text{Vect } x) = \text{Vect } x$, d'où $g((\text{Vect } x)^\perp) = (\text{Vect } x)^\perp$. Considérons l'application de restriction

$$\pi_x : \text{Stab}_G(x) \rightarrow O((\text{Vect } x)^\perp), \quad g \mapsto g|_{(\text{Vect } x)^\perp}.$$

π_x est un morphisme :

- injectif : si $g|_{(\text{Vect } x)^\perp} = \text{Id}$, comme $g(x) = x$, alors $g = \text{Id}$.
- à valeurs dans $SO((\text{Vect } x)^\perp)$: dans une base adaptée à la décomposition $\mathbb{R}^3 = \text{Vect } x \oplus (\text{Vect } x)^\perp$, g a pour matrice

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & g|_{(\text{Vect } x)^\perp} & \\ 0 & & & \end{pmatrix},$$

donc $\det \pi_x(g) = \det g = 1$.

Comme $SO((\text{Vect } x)^\perp) \simeq SO_2(\mathbb{R})$, on en déduit que $\text{Stab}_G(x)$ est un groupe cyclique. On note n_x son ordre. Alors $n_x \mid n$ et $n_x \geq 2$.

Remarque. Si x et y sont dans la même orbite, alors $\text{Stab}_G(x)$ et $\text{Stab}_G(y)$ sont conjugués, donc $n_x = n_y$.

L'ensemble X est fini. En effet, si $g \in SO_3(\mathbb{R}) \setminus \{\text{Id}\}$, g admet exactement deux vecteurs unitaires fixes (les deux vecteurs unitaires de son axe). Donc $|X| \leq 2(n-1)$. On remarque également que si $x \in X$, $-x \in X$.

Posons $Y = \{(x, g) \in X \times G \setminus \{\text{Id}\} : g(x) = x\}$ et comptons le nombre d'éléments de Y :

- première méthode :

$$|Y| = \sum_{g \in G \setminus \{\text{Id}\}} |\{x \in X : g(x) = x\}| = 2(n-1).$$

- seconde méthode :

$$|Y| = \sum_{x \in X} |\{g \in G \setminus \{\text{Id}\} : g(x) = x\}| = \sum_{x \in X} (n_x - 1).$$

Soit k le nombre d'orbites de X modulo G . On numérote ces orbites $\omega_1, \dots, \omega_k$ par ordre de taille : $|\omega_1| \leq \dots \leq |\omega_k|$. Pour tout i , on appelle n_i l'ordre commun aux stabilisateurs respectifs des éléments de ω_i : pour tout $x \in \omega_i$, $n_x = n_i$. On a $n_1 \geq \dots \geq n_k$. Alors

$$|Y| = \sum_{i=1}^k \sum_{x \in \omega_i} (n_x - 1) = \sum_{i=1}^k (n_i - 1) \frac{n}{\widetilde{n_i}}.$$

On en déduit que

$$2\left(1 - \frac{1}{n}\right) = \sum_{i=1}^k \left(1 - \frac{1}{n_i}\right).$$

Or $n \geq n_1 \geq \dots \geq n_k \geq 2$, donc $1 > 1 - \frac{1}{n} \geq 1 - \frac{1}{n_1} \geq \dots \geq 1 - \frac{1}{n_k} \geq \frac{1}{2}$. Par conséquent, on obtient $\frac{k}{2} \leq \sum_{i=1}^k \left(1 - \frac{1}{n_i}\right) \leq k\left(1 - \frac{1}{n}\right)$, puis $\frac{k}{2} \leq 2\left(1 - \frac{1}{n}\right) \leq k\left(1 - \frac{1}{n}\right)$, ce qui impose $k \in \{2, 3\}$.

- On suppose $k = 2$. Alors $1 - \frac{1}{n} = \frac{1}{2} \left(\left(1 - \frac{1}{n_1}\right) + \left(1 - \frac{1}{n_2}\right) \right)$, d'où $n_1 = n_2 = n$.
- On suppose $k = 3$. Alors $3\left(1 - \frac{1}{n_3}\right) \leq \sum_{i=1}^3 \left(1 - \frac{1}{n_i}\right) = 2\left(1 - \frac{1}{n}\right) < 2$, donc $1 - \frac{1}{n_3} < \frac{2}{3}$, d'où $n_3 = 2$.
Ensuite, on a $\frac{1}{2} + 2\left(1 - \frac{1}{n_2}\right) \leq \sum_{i=1}^3 \left(1 - \frac{1}{n_i}\right) = 2\left(1 - \frac{1}{n}\right) < 2$, d'où $1 - \frac{1}{n_2} < \frac{3}{4}$, puis $n_2 \in \{2, 3\}$.
- On suppose $n_2 = 2$. Alors $1 + 1 - \frac{1}{n_1} = 2\left(1 - \frac{1}{n}\right)$, d'où $n_1 = \frac{n}{2}$, ce qui impose n pair.
- On suppose $n_2 = 3$. On a $\frac{1}{2} + \frac{2}{3} + 1 - \frac{1}{n_1} < 2$, d'où $n_1 < 6$. Or $n_1 \geq n_2$, donc $n_1 \in \{3, 4, 5\}$. Si $n_1 = 3$, $n = 12$. Si $n_1 = 4$, $n = 24$. Si $n_1 = 5$, $n = 60$.

Consignons les résultats que nous venons d'obtenir dans le tableau 6.1.

Cas	k	n	n_1	n_2	n_3	G
I	2		n	n	-	$\mathbb{Z}/n\mathbb{Z}$
II	3	pair	$\frac{n}{2}$	2	2	D_n
III	3	12	3	3	2	$\mathfrak{A}_4$
IV	3	24	4	3	2	$\mathfrak{S}_4$
V	3	60	5	3	2	$\mathfrak{A}_5$

TABLE 6.1 : Classification des sous-groupes finis de $SO_3(\mathbb{R})$

Le reste de la démonstration consiste à identifier à quel groupe G correspond chaque cas.

- Cas I : $k = 2$ et $n_1 = n_2 = n$. On a deux orbites de cardinal $\frac{n}{n_i} = 1$: $\omega_1 = \{x\}$ et $\omega_2 = \{y\}$. x et y sont fixes par G . Comme $-x \in X$, on a $y = -x$. Soit $D = \text{Vect } x$. Tous les éléments de G fixent D . G est donc isomorphe (par restriction) à un sous-groupe de $SO_2(\mathbb{R})$, donc à $\mathbb{Z}/n\mathbb{Z}$.
- Cas II : $k = 3$, n est pair, $n_1 = \frac{n}{2}$ et $n_2 = n_3 = 2$. On a trois orbites : ω_1 de cardinal 2, ω_2 et ω_3 de cardinal $\frac{n}{2}$. Nous allons proposer deux méthodes pour identifier le groupe diédral D_n .
 - (i). Commençons par la première méthode. Soit H le stabilisateur d'un élément de ω_1 : $H = \text{Stab}_G(x)$, avec $x \in \omega_1$. H est un sous-groupe cyclique de cardinal $\frac{n}{2}$. Soit y tel que $\omega_1 = \{x, y\}$. Comme y est dans la même orbite que x , $\text{Stab}_G(y)$ et H sont conjugués. Mais H étant d'indice 2, il est distingué dans G , ce qui implique $\text{Stab}_G(y) = H$. Tout élément de $G \setminus H$ est dans le stabilisateur d'un de ses pôles (lesquels pôles appartiennent à ω_2 ou ω_3), donc est d'ordre 2.

Soit $g \in G \setminus H$. On a $H \triangleleft G$ et $H \cap \langle g \rangle = \{\text{Id}\}$, car $g \notin H$. Donc $\langle H, g \rangle \simeq H \rtimes \langle g \rangle$. Pour des raisons de cardinalité, on a $G \simeq H \rtimes \langle g \rangle$.

Soit $h \in H$. Comme $g \notin H$, alors $gh \notin H$. Donc gh est d'ordre 2 : $ghgh = \text{Id}$. Sachant que $g^2 = \text{Id}$, on a $g^{-1}hg = h^{-1}$. Finalement $G \simeq \mathbb{Z}/(n/2)\mathbb{Z} \rtimes_{\varphi} \mathbb{Z}/2\mathbb{Z}$, où φ est le morphisme

$$\mathbb{Z}/2\mathbb{Z} \rightarrow \text{Aut}(\mathbb{Z}/(n/2)\mathbb{Z}), \quad 0 \mapsto \text{Id}, \quad 1 \mapsto (x \mapsto -x).$$

On reconnaît D_n .

(ii). Citons maintenant la seconde méthode.

- Supposons $n > 4$. On a $|\omega_1| \neq |\omega_2| = |\omega_3|$ et $n_1 \neq n_2 = n_3$. Soit $x \in \omega_1$. Alors $-x \in X$ et $|\text{Stab}_G(-x)| = |\text{Stab}_G(x)| = n_1$, donc $-x \in \omega_1$, puis $\omega_1 = \{-x, x\}$.

On pose $H = \text{Stab}_G(x)$. Soit $g \in G \setminus H$. $g(x) \in \omega_1$ et $g(x) \neq x$ (sinon, g appartiendrait à H), donc $g(x) = -x$. Donc, que g soit dans H ou non, $g(\text{Vect } x) = \text{Vect } x$, donc $g((\text{Vect } x)^\perp) = (\text{Vect } x)^\perp$. On définit le morphisme de restriction

$$\pi_x : G \rightarrow O((\text{Vect } x)^\perp), \quad g \mapsto g|_{(\text{Vect } x)^\perp}.$$

Soit $g \in G$. Dans une base adaptée à la décomposition $\mathbb{R}^3 = \text{Vect } x \oplus (\text{Vect } x)^\perp$, g a pour matrice

$$\begin{pmatrix} \varepsilon & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & g|_{(\text{Vect } x)^\perp} & \\ 0 & & & \end{pmatrix}.$$

Cette remarque permet de montrer que :

- π_x est injectif. Soit en effet $g \in G$ tel que $\pi_x(g) = \text{Id}$. On a $\det g = \varepsilon \det \pi_x(g) = \varepsilon$. Comme $g \in \text{SO}_3(\mathbb{R})$, on a $\varepsilon = 1$, d'où $g = \text{Id}$.
- $\text{Im } \pi_x$ n'est pas inclus dans $\text{SO}((\text{Vect } x)^\perp)$. On a $1 = \det g = \varepsilon \det \pi_x(g)$. Les seuls éléments dont l'image est dans $\text{SO}((\text{Vect } x)^\perp)$ sont ceux de H .

G est isomorphe à un sous-groupe fini de $O_2(\mathbb{R})$, non contenu dans $\text{SO}_2(\mathbb{R})$. Il est donc diédral.

- Supposons maintenant $n = 4$. G est un groupe d'ordre 4 dont tous les éléments non triviaux sont d'ordre 2, donc $G \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. On note $G = \{\text{Id}, g_1, g_2, g_3\}$. Chaque orbite ω_i s'écrit $\{x, y\}$, avec $y = h(x)$, pour tout $h \notin \text{Stab}_G(x)$. x et y sont dans une même orbite, donc $\text{Stab}_G(x)$ et $\text{Stab}_G(y)$ sont conjugués. Comme G est d'ordre 4, la conjugaison est triviale. Donc $\text{Stab}_G(x) = \text{Stab}_G(y)$.

On suppose, pour fixer les idées, $x \in \omega_1$ et $-x \in \omega_2$. Alors $\omega_1 = \{x, y\}$ et $\omega_2 = \{-x, -y\}$. Les éléments de ω_1 et ω_2 ont tous le même stabilisateur $\{\text{Id}, g_1\}$. On en déduit, en notant z un élément de ω_3 , que le stabilisateur de z contient g_2 et g_3 , ce qui implique qu'il est égal à G , ce qui est absurde. On en déduit que $y = -x$.

Les orbites sont de la forme $\omega_i = \{\pm x_i\}$. Si $g \neq \text{Id}$ fixe x_1 , alors $g(x_2) = -x_2$, donc $x_1 \perp x_2$. De même, $x_1 \perp x_3$ et $x_2 \perp x_3$. G est donc isomorphe de trois façons différentes à un sous-groupe de $O_2(\mathbb{R})$. Comme G n'est pas cyclique, il est diédral.

- Cas III : $k = 3$, $n = 12$, $n_1 = n_2 = 3$ et $n_3 = 2$. G est un groupe d'ordre 12. Ses éléments sont d'ordre 1, 2, 3.

Remarque. On pourrait déduire de la liste des groupes d'ordre 12 que $G \simeq \mathfrak{A}_4$.

On a trois orbites : ω_1 et ω_2 de cardinal 4 et ω_3 de cardinal 6.

Comme $x \in \omega_3 \iff |\text{Stab}_G(x)| = 2$, on sait que ω_3 est stable par passage à l'opposé. Donc $\omega_3 = \{\pm x, \pm y, \pm z\}$.

Soit $\omega_1 = \{t, u, v, w\}$. Soit $g \in \text{Stab}_G(t) \setminus \{\text{Id}\}$. Alors $g(t) = t$ et $g(\{u, v, w\}) = \{u, v, w\}$. Si g fixe u, v et w , alors g fixe au moins deux droites distinctes, ce qui contredit $g \neq \text{Id}$. Donc g agit non trivialement sur $\{u, v, w\}$ et comme $g^3 = \text{Id}$, il permute circulairement u, v et w . Ni u , ni v , ni w n'est un vecteur fixé par g .

En particulier, $-t \notin \omega_1$, car $-t$ est fixé par g . Par transitivité de G sur ω_1 , l'opposé d'aucun de ses éléments n'appartient à ω_1 . Ainsi $\omega_2 = -\omega_1$.

On fait opérer G sur ω_1 . On en déduit un morphisme $G \rightarrow \mathfrak{S}_{\omega_1} \simeq \mathfrak{S}_4$. Ce morphisme est injectif : on vient de voir que si g fixe chaque élément de ω_1 , alors $g = \text{Id}$. On en déduit que $G \simeq \mathfrak{A}_4$.

- Cas IV : $k = 3, n = 24, n_1 = 4, n_2 = 3$ et $n_3 = 2$. On a $x \in \omega_i \iff |\text{Stab}_G(x)| = n_i$. Donc chaque orbite est stable par passage à l'opposé. On en déduit que :

- ω_1 est formé de six vecteurs sur trois droites,
- ω_2 est formé de huit vecteurs sur quatre droites,
- ω_3 est formé de douze vecteurs sur six droites.

ω_2 définit quatre droites $\Delta_1, \dots, \Delta_4$. Ce sont les axes d'éléments d'ordre 3 dans G . G opère sur l'ensemble Δ de ces quatre droites. En effet, si $\Delta_i = \langle x \rangle$, avec $x \in \omega_2$, on a $g(\Delta_i) = \langle g(x) \rangle$ et $g(x) \in \omega_2$. On en déduit un morphisme $\varphi : G \rightarrow \mathfrak{S}_\Delta \simeq \mathfrak{S}_4$.

Montrons que φ est injectif. Soit $g \in G$ tel que $g|_\Delta = \text{Id}$, i.e. pour tout i , $g(\Delta_i) = \Delta_i$. On suppose $g \neq \text{Id}$.

On suppose que g fixe un élément de ω_2 , par exemple $x \in \Delta_1$. Comme g est d'ordre 3, il s'agit d'une rotation d'angle $\pm \frac{2\pi}{3}$. Alors g ne conserve aucune droite distincte de Δ_1 . Donc g permute circulairement les droites Δ_2, Δ_3 et Δ_4 . Absurde, alors g ne fixe aucun élément de ω_2 , d'où pour tout i , $g|_{\Delta_i} = -\text{Id}$.

Ces quatre droites engendrent $\mathbb{R}^3$: sinon, leur plan serait invariant par une rotation d'ordre 3 dont l'axe est dans ce plan, ce qui est impossible. Ainsi, on en déduit que $g = -\text{Id}$, ce qui contredit $g \in \text{SO}_3(\mathbb{R})$. On en déduit $g = \text{Id}$, d'où l'injectivité de φ . Finalement, $G \simeq \mathfrak{S}_4$.

- Cas V : $k = 3, n = 60, n_1 = 5, n_2 = 3$ et $n_3 = 2$. G est un groupe d'ordre 60. Ses éléments sont d'ordre 1, 2, 3, 5. On a de nouveau $x \in \omega_i \iff |\text{Stab}_G(x)| = n_i$. Donc chaque orbite est conservée par passage à l'opposé.
- Un élément d'ordre 2 est dans le stabilisateur de deux éléments opposés de ω_3 : ses deux pôles. Réciproquement, deux vecteurs opposés de ω_3 engendrent l'axe d'un unique élément d'ordre 2. On a donc $\frac{|\omega_3|}{2} = 15$ éléments d'ordre 2 dans G .
- Un élément d'ordre 3 est dans le stabilisateur de deux éléments opposés de ω_2 : ses deux pôles. Réciproquement, deux vecteurs opposés de ω_2 engendrent l'axe des deux rotations non triviales de leur stabilisateur commun. On a donc $2 \frac{|\omega_2|}{2} = 20$ éléments d'ordre 3 dans G .
- Un élément d'ordre 5 est dans le stabilisateur de deux éléments opposés de ω_1 : ses deux pôles. Réciproquement, deux vecteurs opposés de ω_1 engendrent l'axe des quatre rotations non triviales de leur stabilisateur commun. On a donc $4 \frac{|\omega_1|}{2} = 24$ éléments d'ordre 5 dans G .

Les quinze demi-tours forment une seule classe de conjugaison, puisque G agit transitivement sur leurs axes. Leur centralisateur a donc ordre $60/15 = 4$. Le centralisateur d'une rotation d'ordre 3 ou 5 est son stabilisateur axial, d'ordre 3 ou 5 : une rotation qui commute avec elle doit conserver son axe orienté. Les classes de conjugaison ont donc les tailles 1, 15, 20, 12, 12. Aucune somme contenant 1 d'une sous-famille propre non vide des quatre autres tailles ne divise 60. Par le théorème de Lagrange, G est simple.

Les sous-groupes de Sylow d'ordre 4 sont isomorphes à $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, car G ne contient aucun élément d'ordre 4. Chacun contient trois involutions. Une involution appartient à un seul de ces sous-groupes : tout tel sous-groupe est contenu dans son centralisateur, d'ordre 4. Il y a donc cinq sous-groupes de Sylow d'ordre 4. La conjugaison sur cet ensemble définit un morphisme non trivial $G \rightarrow \mathfrak{S}_5$, injectif par simplicité. Son image, d'ordre 60, est $\mathfrak{A}_5$ (la signature est triviale sur un groupe simple d'ordre 60). Ainsi $G \simeq \mathfrak{A}_5$.

Nous avons traité tous les cas répertoriés dans le tableau 6.1, ce qui achève la démonstration. □

6.2.2 Sous-groupes finis de $O_3(\mathbb{R})$ non contenus dans $SO_3(\mathbb{R})$

Proposition 6.11. $O_3(\mathbb{R})$ admet pour sous-groupes finis non inclus dans $SO_3(\mathbb{R})$ des sous-groupes isomorphes à $\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, $D_{2n} \times \mathbb{Z}/2\mathbb{Z}$, $\mathfrak{A}_4 \times \mathbb{Z}/2\mathbb{Z}$, $\mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}$, $\mathfrak{A}_5 \times \mathbb{Z}/2\mathbb{Z}$, $\mathfrak{S}_4$, D_{2n} ou $\mathbb{Z}/(2n)\mathbb{Z}$.

Démonstration. Soit H un sous-groupe fini de $SO_3(\mathbb{R})$. Alors $\langle H, -\text{Id} \rangle \simeq H \times \{\pm \text{Id}\}$ qui est un sous-groupe de $O_3(\mathbb{R})$ qui n'est pas inclus dans $SO_3(\mathbb{R})$. La proposition 6.3 assure que $O_3(\mathbb{R})$ admet pour sous-groupes finis non inclus dans $SO_3(\mathbb{R})$ des sous-groupes isomorphes à $\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, $D_{2n} \times \mathbb{Z}/2\mathbb{Z}$, $\mathfrak{A}_4 \times \mathbb{Z}/2\mathbb{Z}$, $\mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}$, $\mathfrak{A}_5 \times \mathbb{Z}/2\mathbb{Z}$. D'après la proposition 6.4, le groupe des isométries du tétraèdre régulier est un sous-groupe fini de $O_3(\mathbb{R})$ non contenu dans $SO_3(\mathbb{R})$ et est isomorphe à $\mathfrak{S}_4$.

On sait qu'il existe un sous-groupe de $O_2(\mathbb{R})$ non contenu dans $SO_2(\mathbb{R})$ isomorphe à D_{2n} (voir proposition 6.2). On peut construire à partir de celui-ci un sous-groupe de $O_3(\mathbb{R})$ non contenu dans $SO_3(\mathbb{R})$, toujours isomorphe à D_{2n} (en mettant Id sur l'axe orthogonal).

Enfin, le groupe engendré par l'élément

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos \frac{\pi}{n} & \sin \frac{\pi}{n} \\ 0 & -\sin \frac{\pi}{n} & \cos \frac{\pi}{n} \end{pmatrix}$$

est un sous-groupe de $O_3(\mathbb{R})$ non contenu dans $SO_3(\mathbb{R})$ qui est isomorphe à $\mathbb{Z}/(2n)\mathbb{Z}$. □

Le théorème suivant montre que les groupes évoqués dans la proposition précédente sont les seuls sous-groupes finis de $O_3(\mathbb{R})$ non contenus dans $SO_3(\mathbb{R})$.

Théorème 6.12. Soit G un sous-groupe fini de $O_3(\mathbb{R})$ non contenu dans $SO_3(\mathbb{R})$. Alors G est isomorphe à $\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, $D_{2n} \times \mathbb{Z}/2\mathbb{Z}$, $\mathfrak{A}_4 \times \mathbb{Z}/2\mathbb{Z}$, $\mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}$, $\mathfrak{A}_5 \times \mathbb{Z}/2\mathbb{Z}$, $\mathfrak{S}_4$, D_{2n} ou $\mathbb{Z}/(2n)\mathbb{Z}$.

Démonstration. On note $G^+ = G \cap SO_3(\mathbb{R})$. On a $[G : G^+] = 2$.

- (i). On suppose $-\text{Id} \in G$. Alors $G \simeq G^+ \times \{\pm \text{Id}\}$, donc $G \simeq G^+ \times \mathbb{Z}/2\mathbb{Z}$, où G^+ est un sous-groupe fini de $SO_3(\mathbb{R})$, qui est isomorphe à $\mathbb{Z}/n\mathbb{Z}$, D_{2n} , $\mathfrak{A}_4$, $\mathfrak{S}_4$ ou $\mathfrak{A}_5$, d'après le théorème 6.10.
- (ii). On suppose $-\text{Id} \notin G$. On pose $H = \langle G, -\text{Id} \rangle$. Comme $-\text{Id} \notin G$ et $-\text{Id}$ commute avec G , on a $H \simeq G \times \{\pm \text{Id}\}$. Donc $G \simeq H/\{\pm \text{Id}\} \simeq H^+$. Ainsi, $G \simeq H^+$. En outre, G^+ est un sous-groupe d'indice 2 de G , donc H^+ admet un sous-groupe d'indice 2. Nécessairement, d'après le théorème 6.10, H^+ est isomorphe à $\mathfrak{S}_4$, D_{2n} ou $\mathbb{Z}/(2n)\mathbb{Z}$ ($\mathbb{Z}/(2n+1)\mathbb{Z}$ n'admet pas de sous-groupe d'indice 2, car $2 \nmid 2n+1$; $\mathfrak{A}_4$ non plus, d'après ce que l'on connaît de sa structure; $\mathfrak{A}_5$ non plus, puisqu'il est simple).

□

6.3 Problème de la coloration du cube

6.3.1 Énoncé du problème

On dispose d'un cube et de q pots de peinture de couleurs distinctes. On cherche le nombre de colorations de ses six faces, chacune portant une seule couleur. Deux colorations sont identifiées si une rotation du cube transforme l'une en l'autre; les réflexions ne sont pas autorisées.

Traitions quelques cas simples :

- On suppose $q = 1$. Il y a alors une seule coloration possible du cube.
- On suppose $q = 2$. On dispose de deux couleurs A et B . Il y a a faces de couleur A et b faces de couleur B . On a $a + b = 6$. Cherchons le nombre de colorations possibles correspondant à chaque valeur de a .

a	colorations
0	1
1	1
2	2
3	2
4	2
5	1
6	1

Il y a donc dix colorations possibles du cube.

Revenons au cas général. Soit $\mathcal{C}$ l'ensemble des couleurs. Soit $(x_c)_{c \in \mathcal{C}} \in \mathbb{N}^{\mathcal{C}}$ tel que $\sum_c x_c = 6$. On note $N((x_c))$ le nombre de colorations du cube avec x_c faces de couleur c . On considère le polynôme

$$P(X_1, \dots, X_q) = \sum_{\substack{(x_c) \in \mathbb{N}^{\mathcal{C}} \\ \sum_c x_c = 6}} N((x_c)) X_1^{x_1} \dots X_q^{x_q}.$$

Remarques. Le nombre de colorations du cube avec q couleurs est alors $P(1, \dots, 1)$. En outre, le polynôme P est symétrique : le nombre de colorations du cube avec x_c faces de couleur c est le même que le nombre de colorations avec x_c faces de couleur $\sigma(c)$, pour toute permutation σ .

6.3.2 Théorème de Pólya

Soit X un ensemble fini de cardinal k . Soit G un groupe fini opérant sur X . Fixons quelques notations. Soit $g \in G$. Alors g définit une permutation sur X

$$\sigma_g : X \rightarrow X, \quad x \mapsto g \cdot x.$$

Pour tout $i \in \{1, \dots, k\}$, on note $c_i(g)$ le nombre de cycles de longueur i apparaissant dans la décomposition en produit de cycles disjoints de σ_g . On définit le polynôme

$$P_{G,X}(X_1, \dots, X_k) = \frac{1}{|G|} \sum_{g \in G} X_1^{c_1(g)} \dots X_k^{c_k(g)}.$$

Soit $\mathcal{C}$ un ensemble fini. L'opération de G sur X induit une action de G sur $\mathcal{F}(X, \mathcal{C}) = \{f : X \rightarrow \mathcal{C}\}$

$$G \times \mathcal{F}(X, \mathcal{C}) \rightarrow \mathcal{F}(X, \mathcal{C}), \quad (g, f) \mapsto (x \mapsto f(g^{-1} \cdot x)).$$

On définit une application

$$\omega : \mathcal{C} \rightarrow \mathbb{Q}[X_c, c \in \mathcal{C}], \quad c \mapsto X_c,$$

et on en déduit

$$w : \mathcal{F}(X, \mathcal{C}) \rightarrow \mathbb{Q}[X_c, c \in \mathcal{C}], \quad f \mapsto \prod_{x \in X} \omega(f(x)).$$

Remarque. Si $g \in G$ et $f \in \mathcal{F}(X, \mathcal{C})$, alors $w(g \cdot f) = w(f)$. En effet, $w(g \cdot f) = \prod_{x \in X} \omega(f(g^{-1} \cdot x)) = \prod_{y \in X} \omega(f(y)) = w(f)$.

Nous allons nous attacher à démontrer le théorème suivant.

Théorème 6.13 (Pólya).

$$\sum_{f \in \mathcal{F}(X, \mathcal{C})_G} w(f) = P_{G,X} \left(\sum_{c \in \mathcal{C}} X_c, \dots, \sum_{c \in \mathcal{C}} X_c^k \right).$$

Démonstration. On note Ω_f l'orbite de $f \in \mathcal{F}(X, \mathcal{C})$ sous l'action de G .

$$\sum_{f \in \mathcal{F}(X, \mathcal{C})_G} w(f) = \sum_{f \in \mathcal{F}(X, \mathcal{C})_G} \left(\frac{1}{|\Omega_f|} \sum_{f' \in \Omega_f} w(f') \right).$$

Mais $|\Omega_f| = \frac{|G|}{|\text{Stab}_G(f)|}$, donc on peut écrire

$$\begin{aligned} \sum_{f \in \mathcal{F}(X, \mathcal{C})_G} w(f) &= \frac{1}{|G|} \sum_{f \in \mathcal{F}(X, \mathcal{C})} |\text{Stab}_G(f)| w(f) \\ &= \frac{1}{|G|} \sum_{f \in \mathcal{F}(X, \mathcal{C})} \sum_{\substack{g \in G \\ g \cdot f = f}} w(f) = \frac{1}{|G|} \sum_{g \in G} \sum_{\substack{f \in \mathcal{F}(X, \mathcal{C}) \\ g \cdot f = f}} w(f). \end{aligned}$$

Le support $\text{supp } \tau$ d'un cycle τ est l'ensemble des éléments qui le composent (y compris son unique élément pour un cycle de longueur 1). Soit $g \in G$. Soit $f \in \mathcal{F}(X, \mathcal{C})$. Alors $g \cdot f = f$, si et seulement si pour tout $x \in X$, $f(g \cdot x) = f(x)$ (i.e. la couleur de la face x est la même que celle de la face $g(x)$), si et seulement si f est constante sur chaque orbite de X sous l'action de $\langle g \rangle$, si et seulement si f est constante sur chaque support de cycle dans la décomposition en produit de cycles disjoints de σ_g . Si $\tau_1 \dots \tau_\ell$ est la décomposition en produit de cycles disjoints de σ_g , on a ainsi $g \cdot f = f$, si et seulement si f est constante sur le support de τ_i , pour tout i . On a alors

$$w(f) = \prod_{x \in X} \omega(f(x)) = \prod_{i=1}^{\ell} \prod_{x \in \text{supp } \tau_i} \omega(f(x)).$$

Mais, en notant $c_i = f(x)$, pour tout $x \in \text{supp } \tau_i$ et $m_i = |\text{supp } \tau_i|$, on a

$$\prod_{x \in \text{supp } \tau_i} \omega(f(x)) = X_{c_i}^{m_i},$$

d'où $w(f) = X_{c_1}^{m_1} \dots X_{c_\ell}^{m_\ell}$. Considérer une fonction $f \in \mathcal{F}(X, \mathcal{C})$ telle que $g \cdot f = f$ revient à considérer le ℓ -uplet des valeurs prises par la fonction f sur le support de chacun des cycles $\tau_1, \dots, \tau_\ell$, et réciproquement. On en déduit que

$$\sum_{\substack{f \in \mathcal{F}(X, \mathcal{C}) \\ g \cdot f = f}} w(f) = \sum_{(c_1, \dots, c_\ell) \in \mathcal{C}^\ell} X_{c_1}^{m_1} \dots X_{c_\ell}^{m_\ell}.$$

Donc

$$\begin{aligned} \sum_{f \in \mathcal{F}(X, \mathcal{C})_G} w(f) &= \frac{1}{|G|} \sum_{g \in G} \sum_{\substack{f \in \mathcal{F}(X, \mathcal{C}) \\ g \cdot f = f}} w(f) \\ &= \frac{1}{|G|} \sum_{\substack{g \in G \\ \sigma_g = \tau_1 \dots \tau_\ell}} \sum_{(c_1, \dots, c_\ell) \in \mathcal{C}^\ell} \prod_{i=1}^{\ell} X_{c_i}^{|\text{supp } \tau_i|} \\ &= \frac{1}{|G|} \sum_{\substack{g \in G \\ \sigma_g = \tau_1 \dots \tau_\ell}} \prod_{i=1}^{\ell} \sum_{c \in \mathcal{C}} X_c^{|\text{supp } \tau_i|} \\ &= \frac{1}{|G|} \sum_{g \in G} \left(\sum_{c \in \mathcal{C}} X_c^1 \right)^{c_1(g)} \dots \left(\sum_{c \in \mathcal{C}} X_c^k \right)^{c_k(g)} = P_{G, X} \left(\sum_{c \in \mathcal{C}} X_c, \dots, \sum_{c \in \mathcal{C}} X_c^k \right). \end{aligned}$$

□

6.3.3 Application au problème de la coloration du cube

Liens avec le théorème de Pólya

On note X l'ensemble des six faces du cube, $\mathcal{C}$ l'ensemble des couleurs disponibles. On appelle G le groupe des isométries positives du cube, qui est isomorphe à $\mathfrak{S}_4$, d'après la proposition 6.6. $\mathcal{F}(X, \mathcal{C})$ est l'ensemble des colorations du cube dont on a numéroté les faces. Pour obtenir l'ensemble des colorations du cube, il suffit de considérer $\mathcal{F}(X, \mathcal{C})|_G$.

Pour $f \in \mathcal{F}(X, \mathcal{C})$, on a

$$w(f) = \prod_{c \in \mathcal{C}} X_c^{x_c(f)},$$

où $x_c(f)$ désigne le nombre de faces du cube portant la couleur c . Par conséquent,

$$\sum_{f \in \mathcal{F}(X, \mathcal{C})|_G} w(f) = \sum_{f \in \mathcal{F}(X, \mathcal{C})|_G} \prod_{c \in \mathcal{C}} X_c^{x_c(f)}.$$

En regroupant dans la somme les termes qui correspondent au même nombre de faces de chaque couleur, on obtient

$$\sum_{f \in \mathcal{F}(X, \mathcal{C})|_G} w(f) = \sum_{\substack{(x_c) \in \mathbb{N}^{\mathcal{C}} \\ \sum_c x_c = 6}} N((x_c)) \prod_{c \in \mathcal{C}} X_c^{x_c}.$$

Calcul de $P_{G,X}$

Rappelons que

$$P_{G,X}(X_1, \dots, X_k) = \frac{1}{|G|} \sum_{g \in G} X_1^{c_1(g)} \dots X_k^{c_k(g)}.$$

Afin de calculer $P_{G,X}$, nous allons passer en revue les éléments $g \in G$ et déterminer, pour tout i , le nombre de cycles de longueur i apparaissant dans la décomposition en produit de cycles à supports disjoints de σ_g .

	$\mathfrak{S}_4$	G	$c_1(g)$	$c_2(g)$	$c_3(g)$	$c_4(g)$	$c_5(g)$	$c_6(g)$
1	Id	Id	6	0	0	0	0	0
6	transposition	rotation d'axe passant par les milieux de deux arêtes opposées et d'angle π	0	3	0	0	0	0
8	3-cycle	rotation d'axe une grande diagonale et d'angle $\pm \frac{2\pi}{3}$	0	0	2	0	0	0
3	produit de deux transpositions	rotation d'axe une droite passant par les centres de deux faces opposées et d'angle π	2	2	0	0	0	0
6	4-cycle	rotation d'axe une droite passant par les centres de deux faces opposées et d'angle $\pm \frac{\pi}{2}$	2	0	0	1	0	0

On en déduit finalement

$$P_{G,X}(X_1, \dots, X_6) = \frac{1}{24} (X_1^6 + 6X_2^3 + 8X_3^2 + 3X_1^2 X_2^2 + 6X_1^2 X_4) .$$

Conséquences du théorème de Pólya

On applique le théorème de Pólya avec, par exemple, six couleurs.

$$\begin{aligned} \sum_{\substack{(x_1, \dots, x_6) \in \mathbb{N}^6 \\ \sum_i x_i = 6}} N((x_i)) X_1^{x_1} \dots X_6^{x_6} &= \frac{1}{24} \left((X_1 + \dots + X_6)^6 + 6(X_1^2 + \dots + X_6^2)^3 \right. \\ &\quad + 8(X_1^3 + \dots + X_6^3)^2 + 3(X_1 + \dots + X_6)^2 (X_1^2 + \dots + X_6^2)^2 \\ &\quad \left. + 6(X_1 + \dots + X_6)^2 (X_1^4 + \dots + X_6^4) \right). \end{aligned} \quad (6.1)$$

On en déduit le nombre de colorations possibles du cube avec six couleurs.

$$\frac{1}{24}(6^6 + 6 \cdot 6^3 + 8 \cdot 6^2 + 3 \cdot 6^2 \cdot 6^2 + 6 \cdot 6^2 \cdot 6) = 2226.$$

En développant le membre de droite de l'équation 6.1, ce qui est plutôt fastidieux, on pourrait en fait obtenir le nombre de colorations possibles avec x_i faces de couleur i , pour tout i .

Remarque. Le lecteur amusé pourra également s'intéresser au problème de coloration de la roulette : on considère une roulette comportant n secteurs angulaires identiques et on s'intéresse au nombre de façons de colorier ces secteurs angulaires avec un nombre donné de couleurs disponibles. On note $\mathbb{U}_n = \{z \in \mathbb{C} : z^n = 1\}$ le groupe des racines n -ièmes de l'unité. On remarque alors que le groupe $\mathbb{U}_n$ agit sur la roulette, de sorte que l'on peut encore une fois faire appel au théorème de Pólya.

Combinatoire

Références : [6, 9]

L'analyse combinatoire étudie les différentes façons de calculer le nombre d'éléments d'un certain ensemble. Commençons par un exemple simple.

Exemple. Quelle est la dimension de l'espace vectoriel des polynômes homogènes en n variables de degré d ?

Une base de cet espace est formée des polynômes $X_1^{d_1} \dots X_n^{d_n}$, avec $d_1 + \dots + d_n = d$. La dimension recherchée est donc le cardinal de l'ensemble $\{(d_1, \dots, d_n) \in \mathbb{N}^n : d_1 + \dots + d_n = d\}$. Une manière de calculer ce cardinal est de remarquer que choisir des entiers $d_1, \dots, d_n$ tels que $d_1 + \dots + d_n = d$ revient à choisir comment construire une suite de $d + n - 1$ zéros et uns comportant exactement $n - 1$ zéros : en effet, on peut représenter $(d_1, \dots, d_n)$ ainsi :

$$\underbrace{1 \dots 1}_{d_1} 0 \underbrace{1 \dots 1}_{d_2} 0 \dots 0 \underbrace{1 \dots 1}_{d_n}.$$

Dans cette séquence de chiffres, il y a $n - 1$ zéros et d uns. Il suffit donc de calculer le nombre de façons de choisir quels seront les zéros. Il y en a

$$\binom{d + n - 1}{n - 1}.$$

Le lecteur retrouvera les notions abordées dans ce chapitre dans le très complet [6], ainsi que dans [9]. Nous considérerons que certains résultats élémentaires, comme le principe de récurrence, le lemme des tiroirs ou encore la formule des classes (voir proposition 1.5), sont déjà connus.

7.1 Formule du crible

7.1.1 Énoncé

Proposition 7.1 (Formule du crible). Soit E un ensemble fini. Soient $A_1, \dots, A_n$ n parties de E . Alors

$$|A_1 \cup \dots \cup A_n| = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} |A_{i_1} \cap \dots \cap A_{i_k}|.$$

Démonstration. On note 1_A la fonction indicatrice de A , égale à 1 sur A et à 0 ailleurs. On pose, pour tout i , $B_i = E \setminus A_i$. Alors

$$1_{E \setminus \bigcup_{i=1}^n A_i} = 1_{\bigcap_{i=1}^n B_i} = \prod_{i=1}^n 1_{B_i} = \prod_{i=1}^n (1 - 1_{A_i}).$$

En développant le produit, on obtient

$$1 - 1_{\bigcup_{i=1}^n A_i} = 1 - \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} 1_{A_{i_1}} \dots 1_{A_{i_k}}.$$

On obtient finalement

$$1_{\bigcup_{i=1}^n A_i} = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} 1_{A_{i_1} \cap \dots \cap A_{i_k}}.$$

La formule du crible découle de l'intégration sur E de l'égalité précédente. □

7.1.2 Application au calcul du nombre de dérangements

Définition 7.1. Soit E un ensemble fini. On appelle *dérangement* de E tout élément de $\mathfrak{S}(E)$ qui n'a aucun point fixe.

Notre but est de calculer le nombre d_n de dérangements de l'ensemble $\{1, \dots, n\}$. Soit $X \subset \{1, \dots, n\}$. On note

$$F_X = \{\sigma \in \mathfrak{S}_n : \forall x \in X, \sigma(x) = x\}.$$

Remarque. F_X est un sous-groupe de $\mathfrak{S}_n$ isomorphe (par restriction) à $\mathfrak{S}(\{1, \dots, n\} \setminus X)$. En particulier, $|F_X| = (|\{1, \dots, n\} \setminus X|)!$.

Les éléments de $\mathfrak{S}_n$ qui ne sont pas des dérangements sont exactement les éléments qui ont au moins un point fixe, c'est-à-dire les éléments de

$$\bigcup_{i=1}^n F_{\{i\}}.$$

On applique la formule du crible.

$$\begin{aligned} \left| \bigcup_{i=1}^n F_{\{i\}} \right| &= \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} |F_{\{i_1\}} \cap \dots \cap F_{\{i_k\}}| = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} |F_{\{i_1, \dots, i_k\}}| \\ &= \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} (n-k)! = n! \sum_{k=1}^n \frac{(-1)^{k-1}}{k!}. \end{aligned}$$

On en déduit que le nombre de dérangements de $\{1, \dots, n\}$ est

$$d_n = n! \sum_{k=0}^n \frac{(-1)^k}{k!}, \quad \text{d'où} \quad \lim_{n \rightarrow +\infty} \frac{d_n}{n!} = \frac{1}{e}.$$

On munit $\mathfrak{S}_n$ d'une loi uniforme. Le résultat précédent indique que la probabilité qu'une permutation de $\{1, \dots, n\}$ soit un dérangement tend vers $\frac{1}{e}$ quand n tend vers l'infini.

7.2 Séries génératrices

Il y a deux façons d'aborder les séries génératrices :

- Considérer les séries entières, que l'on connaît déjà : le problème de leur rayon de convergence se pose alors. On est souvent amené à manipuler des séries entières dont on ne sait même pas si elles ont un rayon de convergence non nul !
- Considérer les séries formelles, qui fournissent un outil beaucoup plus satisfaisant que les séries entières dans le cadre de l'étude des séries génératrices.

Par conséquent, nous allons privilégier les séries formelles, que nous allons présenter maintenant.

7.2.1 Séries formelles

Soit A un anneau intègre. On munit $A[X]$ de la distance $d(P, Q) = e^{-\mathfrak{v}(P-Q)}$, où $\mathfrak{v}$ désigne la valuation d'un polynôme, que l'on définit de la manière suivante : si $P = \sum_i a_i X^i$, alors

$$\mathfrak{v}(P) = \begin{cases} \infty & \text{si } P = 0 \\ \inf\{i : a_i \neq 0\} & \text{sinon.} \end{cases}$$

On vérifie aisément les propriétés suivantes :

- $\mathfrak{v}(P) = \infty \iff P = 0$.

- $\mathfrak{b}(-P) = \mathfrak{b}(P)$.
- $\mathfrak{b}(P + Q) \geq \min(\mathfrak{b}(P), \mathfrak{b}(Q))$.

Ceci permet de montrer que d est une distance ultramétrique (i.e. $d(P, Q) \leq \max(d(P, R), d(Q, R))$). En particulier, la suite de terme général $Q_n = \sum_i a_i^{(n)} X^i$ est de Cauchy si et seulement si $d(Q_n, Q_{n+1})$ tend vers 0 quand n tend vers l'infini, c'est-à-dire que pour tout i , la suite $(a_i^{(n)})$ est stationnaire.

Définition 7.2. Le complété de $A[X]$ pour la distance d s'appelle l'algèbre des *séries formelles* à coefficients dans A et d'indéterminée X . On le note $A[[X]]$.

Le lecteur vérifiera aisément le résultat suivant, ainsi que son corollaire.

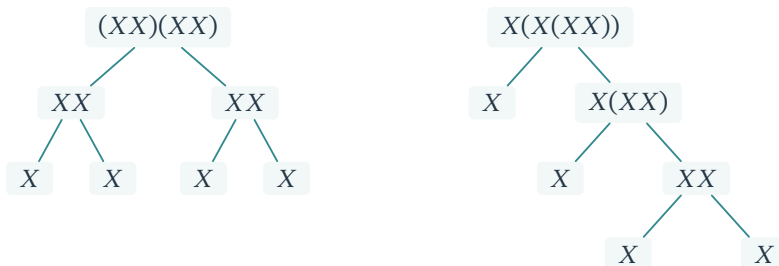
Proposition 7.2. La série formelle $\sum_i a_i X^i$ est inversible si et seulement si a_0 l'est.

Corollaire 7.3. Si $\mathbb{K}$ est un corps, alors $\mathbb{K}[[X]]^*$ est l'ensemble des séries qui ne sont pas multiples de X , c'est-à-dire $\mathbb{K}[[X]] \setminus (X)$.

On a alors $\mathbb{K}[[X]] = \mathbb{K}[[X]]^* \sqcup (X)$. Le spectre maximal de $\mathbb{K}[[X]]$ est alors réduit à un unique élément : (X) . $\mathbb{K}[[X]]$ est appelé anneau local. Il y a une unique classe d'irréductibles : celle de X .

7.2.2 Nombres de Catalan

On dispose d'un produit non associatif. On cherche le nombre C_n de façons de parenthéser un produit de n facteurs X . Par exemple, avec $n = 4$, on peut écrire $(XX)(XX)$ ou encore $X(X(XX))$. Ceci revient à chercher le nombre d'arbres binaires plans pleins à n feuilles (chaque nœud interne a deux fils ordonnés) : chaque feuille correspond à un X , si Y et Z sont les fils d'un même père, alors ce père est $(Y)(Z)$. Pour $n = 4$, voici deux arbres correspondants :



On remarque aisément que $C_1 = 1$ et, pour $n \geq 2$,

$$C_n = \sum_{i=1}^{n-1} C_i C_{n-i}.$$

On travaille désormais dans $\mathbb{Q}[[X]]$. On appelle S la série génératrice de (C_n) , c'est-à-dire la série formelle

$$S = \sum_{n=1}^{\infty} C_n X^n.$$

On a

$$S = X + \sum_{n=2}^{\infty} \sum_{i=1}^{n-1} C_i C_{n-i} X^n = X + \sum_{i=1}^{\infty} C_i X^i \sum_{n>i} C_{n-i} X^{n-i} = X + \sum_{i=1}^{\infty} C_i X^i \sum_{j=1}^{\infty} C_j X^j = X + S^2.$$

On en déduit que $S^2 - S + X = 0$. Le discriminant correspondant à ce trinôme est $\Delta = 1 - 4X$. On note T son unique racine carrée formelle de terme constant 1 (ses coefficients se déterminent successivement à partir de $T^2 = 1 - 4X$). Le développement binomial formel donne

$$\begin{aligned} T &= \sqrt{1 - 4X} = 1 + \sum_{k=1}^{\infty} \prod_{i=0}^{k-1} \left(\frac{1}{2} - i \right) \frac{(-4X)^k}{k!} \\ &= 1 + \sum_{k=1}^{\infty} \left(- \left(-\frac{1}{2} \right)^k \frac{(2k-2)!}{2^{k-1}(k-1)!} \right) \frac{(-1)^k 2^{2k}}{k!} X^k = 1 - \sum_{k=1}^{\infty} \frac{2(2k-2)!}{(k-1)!k!} X^k. \end{aligned}$$

Alors $S = \frac{1 \pm T}{2}$ et, comme S est sans terme constant,

$$S = \frac{1 - T}{2} = \sum_{k=1}^{\infty} \frac{(2k-2)!}{(k-1)!k!} X^k = \sum_{k=1}^{\infty} \frac{1}{k} \binom{2k-2}{k-1} X^k.$$

L'identification des coefficients donne $C_n = \frac{1}{n} \binom{2n-2}{n-1}$ pour $n \geq 1$.

7.2.3 Nombre de surjections

Pour $n, k \in \mathbb{N}$, on note $S_{n,k}$ le nombre de surjections d'un ensemble à n éléments dans un ensemble à k éléments. On convient que $S_{0,0} = 1$ et $S_{n,0} = 0$ pour $n \geq 1$. On remarque d'emblée que, si $n < k$, $S_{n,k} = 0$ et que, si $n = k$, $S_{n,k} = n!$.

Pour $n, k \geq 1$, en considérant les applications d'un ensemble à n éléments dans un ensemble à k éléments et en les regroupant selon le nombre d'éléments de leur image, il est facile de remarquer que

$$k^n = \sum_{i=1}^k \binom{k}{i} S_{n,i}.$$

La formule d'inversion de Pascal permettrait de conclure; nous allons utiliser les séries génératrices.

Nous partons d'une autre remarque : en distinguant si l'image du dernier élément de départ a un autre antécédent ou non, on a, pour $n, k \geq 1$,

$$S_{n,k} = k(S_{n-1,k} + S_{n-1,k-1}).$$

Pour $k \geq 1$, on appelle T_k la série génératrice de $(S_{n,k})_{n \geq k}$, i.e.

$$T_k = \sum_{n \geq k} S_{n,k} X^{n-k}.$$

Pour $k \geq 2$, on a alors

$$T_k = k \left(\sum_{n \geq k} S_{n-1,k} X^{n-k} + \sum_{n \geq k} S_{n-1,k-1} X^{n-k} \right) = k(XT_k + T_{k-1}).$$

Il en résulte que $(1 - kX)T_k = kT_{k-1}$. Comme $S_{n,1} = 1$ pour $n \geq 1$, on obtient

$$T_1 = \sum_{n \geq 1} X^{n-1} = \frac{1}{1-X}, \quad T_k = T_1 \prod_{i=2}^k \frac{i}{1-iX} = k! \prod_{i=1}^k \frac{1}{1-iX}.$$

On décompose cette fraction rationnelle en éléments simples :

$$\begin{aligned} T_k &= k! \sum_{i=1}^k \frac{1}{1-iX} \prod_{\substack{1 \leq j \leq k \\ j \neq i}} \frac{1}{1-\frac{j}{i}} = k! \sum_{i=1}^k \frac{(-1)^{k-i} i^{k-1}}{(k-i)!(i-1)!} \cdot \frac{1}{1-iX} \\ &= k! \sum_{i=1}^k \frac{(-1)^{k-i} i^{k-1}}{(k-i)!(i-1)!} \sum_{n=0}^{\infty} (iX)^n = \sum_{n \geq k} \left(k! \sum_{i=1}^k \frac{(-1)^{k-i} i^{(n-k)+(k-1)}}{(k-i)!(i-1)!} \right) X^{n-k}. \end{aligned}$$

Finalement, pour $n \geq k \geq 1$,

$$S_{n,k} = k! \sum_{i=1}^k \frac{(-1)^{k-i} i^{n-1}}{(k-i)!(i-1)!}.$$

Introduction à la géométrie algébrique

Référence : [13]

Dans ce chapitre, tous les anneaux considérés seront supposés commutatifs.

8.1 Algèbres de polynômes à plusieurs indéterminées

8.1.1 Polynômes homogènes

On s'intéresse à l'algèbre $A[X_1, \dots, X_n]$ des polynômes à n indéterminées et à coefficients dans un anneau A . On définit le degré du monôme $X_1^{i_1} \dots X_n^{i_n}$ par $i_1 + \dots + i_n$.

Définition 8.1. Un polynôme est dit *homogène* de degré d s'il est combinaison linéaire de monômes de degré d . On note $A[X_1, \dots, X_n]_d$ le A -module des polynômes homogènes de degré d (y compris le polynôme nul).

Proposition 8.1. On a

$$A[X_1, \dots, X_n] = \bigoplus_{d \in \mathbb{N}} A[X_1, \dots, X_n]_d.$$

Remarques. Le produit d'un polynôme homogène de degré d et d'un polynôme homogène de degré e est un polynôme homogène de degré $d + e$.

Soit $P \in A[X_1, \dots, X_n]$. Alors le polynôme P est homogène de degré d si et seulement si la relation $P(TX_1, \dots, TX_n) = T^d P(X_1, \dots, X_n)$ est vérifiée.

8.1.2 Théorèmes de transfert

Théorème 8.2. Soit A un anneau et X une indéterminée. Alors

- (i). si A est intègre, $A[X]$ l'est aussi,
- (ii). si A est noethérien, $A[X]$ l'est aussi (théorème de la base de Hilbert),
- (iii). si A est factoriel, $A[X]$ l'est aussi (théorème de Gauss).

Nous ne démontrerons que l'assertion (iii). du théorème précédent. On suppose A factoriel. On note K le corps des fractions de A . Commençons par quelques définitions.

Définition 8.2. Soit $P \in A[X]$ non nul. On appelle *contenu* de P et on note $c(P)$ le pgcd des coefficients de P . Il est défini à multiplication par un élément inversible de A près ; les égalités entre contenus ci-dessous sont entendues à cette réserve près.

Définition 8.3. $P \in A[X]$ est dit *primitif* si $c(P) = 1$.

Remarque. Pour tout $P \in A[X]$ non nul, il existe $\tilde{P}$ primitif tel que $P = c(P)\tilde{P}$.

Lemme 8.3 (Gauss). Soient $P, Q \in A[X]$ non nuls et p un irréductible de A . Alors si $p \mid c(PQ)$, $p \mid c(P)$ ou $p \mid c(Q)$.

Démonstration. Comme p est irréductible et A factoriel, (p) est premier, donc $A/(p)$ est intègre et $A/(p)[X]$ également. Donc $p \mid c(PQ)$ si et seulement si $PQ = 0 \pmod{(p)}$ si et seulement si $P = 0 \pmod{(p)}$ ou $Q = 0 \pmod{(p)}$ si et seulement si $p \mid c(P)$ ou $p \mid c(Q)$. $\square$

Lemme 8.4. Soient $P, Q \in A[X]$ non nuls. $c(PQ) = 1$ si et seulement si $c(P) = c(Q) = 1$.

Corollaire 8.5. Soient $P, Q \in A[X]$ non nuls. $c(PQ) = c(P)c(Q)$.

Démonstration. On écrit $P = c(P)\tilde{P}$ et $Q = c(Q)\tilde{Q}$, où $\tilde{P}$ et $\tilde{Q}$ sont primitifs. On a $PQ = c(P)c(Q)\tilde{P}\tilde{Q}$ et $\tilde{P}\tilde{Q}$ est primitif, d'après le lemme précédent. D'où le résultat. $\square$

Proposition 8.6. Soit $P \in A[X]$ non nul. Alors P est irréductible dans $A[X]$ si et seulement si P est une constante irréductible de A ou P est primitif et irréductible dans $\mathbb{K}[X]$.

Démonstration. Si P est dans A , l'ensemble de ses diviseurs dans A et l'ensemble de ses diviseurs dans $A[X]$ coïncident. Donc P est irréductible dans $A[X]$ si et seulement si il l'est dans A .

On suppose que P n'est pas dans A et qu'il est irréductible dans $A[X]$. On note $P = c(P)\tilde{P}$. Comme P est irréductible non constant, on a $c(P) = 1$. On note $P = QR$, avec Q et R dans $\mathbb{K}[X]$. Il existe a et b dans A non nuls tels que aQ et bR soient dans $A[X]$. Alors $abP = (aQ)(bR)$. Donc $ab = c(abP) = c(aQ)c(bR)$. Ainsi $P = \frac{aQ}{c(aQ)} \frac{bR}{c(bR)}$ dans $A[X]$. Comme P est irréductible, on a $\frac{aQ}{c(aQ)}$ ou $\frac{bR}{c(bR)}$ est inversible, ce qui signifie que P est irréductible dans $\mathbb{K}[X]$.

Réciproquement, on suppose P primitif et irréductible dans $\mathbb{K}[X]$. On écrit $P = QR$ dans $A[X]$. Alors cette décomposition est encore valable dans $\mathbb{K}[X]$ et par irréductibilité dans $\mathbb{K}[X]$, on a Q ou R inversible de $\mathbb{K}[X]$, donc Q ou R est un élément non nul de A . Mais $c(Q)c(R) = c(P) = 1$. Si par exemple Q est un élément non nul de A , il est égal à son contenu, qui est inversible. Donc P est irréductible dans $A[X]$. $\square$

Nous pouvons maintenant passer à la démonstration du théorème de Gauss.

Démonstration. (assertion (iii). du théorème 8.2) Commençons par l'existence. Soit $P \in A[X]$ non nul. Si P est constant, on utilise la factorialité de A . Sinon, on écrit $P = Q_1 \dots Q_r$, où les Q_i sont des irréductibles de $\mathbb{K}[X]$ (le facteur constant est incorporé à Q_1). Il existe $a_i \in A$ non nul tel que $a_i Q_i \in A[X]$. On a $(\prod_{i=1}^r a_i) P = (a_1 Q_1) \dots (a_r Q_r)$, donc $(\prod_{i=1}^r a_i) c(P) = c(a_1 Q_1) \dots c(a_r Q_r)$, puis $P = c(P) \prod_{i=1}^r \frac{a_i Q_i}{c(a_i Q_i)}$. Les $\frac{a_i Q_i}{c(a_i Q_i)}$ sont primitifs et irréductibles dans $\mathbb{K}[X]$, donc irréductibles dans $A[X]$. Il ne reste qu'à décomposer $c(P)$ en produit d'irréductibles de A .

Intéressons-nous à l'unicité. On suppose avoir $a_1 \dots a_p Q_1 \dots Q_r = b_1 \dots b_q R_1 \dots R_s$, où les a_i et b_i sont des irréductibles de A tandis que les Q_i et R_i sont des irréductibles non constants de $A[X]$. En prenant le contenu, on a $a_1 \dots a_p = b_1 \dots b_q$ et la factorialité de A donne $p = q$ et l'existence d'une permutation $\sigma \in \mathfrak{S}_p$ telle que $a_i = b_{\sigma(i)}$ aux inversibles de A près. La factorialité de $\mathbb{K}[X]$ donne ensuite $r = s$ et l'existence d'une permutation $\rho \in \mathfrak{S}_r$ telle que $Q_i = R_{\rho(i)}$ aux inversibles de $\mathbb{K}[X]$ près. Prendre le contenu montre que l'égalité a lieu aux inversibles de A près. $\square$

8.1.3 Résultant et discriminant

On suppose A factoriel et on note $\mathbb{K} = \text{Frac}(A)$ son corps des fractions. Soient $P, Q \in A[X]$ de degrés respectifs $m, n \geq 1$, et de coefficients dominants a_m, b_n .

Proposition 8.7. Les polynômes P et Q ont un facteur commun de degré strictement positif si et seulement si il existe $U, V \in A[X]$ non nuls tels que $\deg U < n$, $\deg V < m$ et $UP = VQ$.

Démonstration. Si $P = RV$ et $Q = RU$ avec $\deg R > 0$, ces quotients conviennent. Réciproquement, si P et Q étaient premiers entre eux dans $\mathbb{K}[X]$, l'égalité $UP = VQ$ impliquerait $Q \mid U$, contrairement à $0 \leq \deg U < n$. Le théorème de Gauss permet de passer de $\mathbb{K}[X]$ à $A[X]$. Un facteur constant non inversible de A ne suffit donc pas pour ce critère. $\square$

Définition 8.4. Le *résultant* $\text{Res}(P, Q)$ est le déterminant de la matrice de Sylvester : ses n premières lignes sont les coefficients de $X^{n-1}P, \dots, P$, et ses m dernières lignes ceux de $X^{m-1}Q, \dots, Q$, tous écrits dans l'ordre décroissant des puissances $X^{m+n-1}, \dots, 1$.

Proposition 8.8. Dans un corps de décomposition, écrivons $P = a_m \prod_{i=1}^m (X - \alpha_i)$ et $Q = b_n \prod_{j=1}^n (X - \beta_j)$, en comptant les racines avec multiplicité. Alors

$$\text{Res}(P, Q) = a_m^n \prod_{i=1}^m Q(\alpha_i) = a_m^n b_n^m \prod_{i=1}^m \prod_{j=1}^n (\alpha_i - \beta_j).$$

En particulier, le résultant est nul si et seulement si P et Q ont une racine commune dans une clôture algébrique, ou, de façon équivalente, un facteur commun de degré positif. On a aussi

$$\text{Res}(P, Q) = (-1)^{mn} \text{Res}(Q, P), \quad \text{Res}(PR, Q) = \text{Res}(P, Q) \text{Res}(R, Q).$$

Démonstration. Notons μ_Q la multiplication par Q dans $\mathbb{K}[X]/(P)$. Les divisions euclidiennes par P dans la matrice de Sylvester donnent $\text{Res}(P, Q) = a_m^n \det \mu_Q$. Dans un corps de décomposition, la multiplication par X est trigonalisable, avec valeurs propres $\alpha_1, \dots, \alpha_m$ comptées avec multiplicité. Ainsi $\det \mu_Q = \prod_i Q(\alpha_i)$, même lorsque les racines ne sont pas simples. Les autres identités en découlent. $\square$

On peut également calculer le résultant par un déterminant de taille n : si μ_P désigne la multiplication par P dans $\mathbb{K}[X]/(Q)$, alors

$$\text{Res}(P, Q) = (-1)^{mn} b_n^m \det \mu_P.$$

Définition 8.5. Pour $P = a_m \prod_{i=1}^m (X - \alpha_i)$, avec $m \geq 1$, on appelle *discriminant* le scalaire

$$\Delta(P) = a_m^{2m-2} \prod_{1 \leq i < j \leq m} (\alpha_i - \alpha_j)^2.$$

Il est nul si et seulement si P possède une racine multiple dans une clôture algébrique de $\mathbb{K}$, c'est-à-dire si le pgcd unitaire de P et P' dans $\mathbb{K}[X]$ est différent de 1. Si $\mathbb{K}$ est parfait, cela équivaut à l'existence d'un facteur irréductible de $\mathbb{K}[X]$ de multiplicité au moins deux.

On étend la définition du résultant aux constantes par $\text{Res}(P, c) = c^m$ pour $c \neq 0$. Lorsque $\deg P' = m - 1$ (en particulier en caractéristique nulle), la formule du résultant donne

$$\Delta(P) = (-1)^{m(m-1)/2} \frac{\text{Res}(P, P')}{a_m}.$$

Le discriminant ne doit donc pas être confondu avec le résultant de P et de sa dérivée. Par exemple,

$$\Delta(X^2 + bX + c) = b^2 - 4c, \quad \Delta(X^3 + pX + q) = -4p^3 - 27q^2.$$

8.2 Ensembles algébriques

8.2.1 Généralités

Soit K un corps.

Définition 8.6. Un sous-ensemble $V \subset K^n$ est un *ensemble algébrique affine* s'il existe un sous-ensemble $E \subset K[X_1, \dots, X_n]$ tel que $V = \{v \in K^n : \forall P \in E, P(v) = 0\}$. On note dans ce cas $V = \mathfrak{B}(E)$.

Exemples. $V = K^n$ est un ensemble algébrique affine (prendre $E = \{0\}$), tout comme $V = \emptyset$ ($E = K[X_1, \dots, X_n]$) et $V = \{(a_1, \dots, a_n)\}$ ($E = \{X_1 - a_1, \dots, X_n - a_n\}$).

Quand $E = \{P\}$ avec P non constant, on parle d'hypersurface et plus particulièrement d'hyperplan dans le cas où $\deg P = 1$ et de courbe algébrique si $n = 2$.

Proposition 8.9. (i). $\mathfrak{B}(E \cup E') = \mathfrak{B}(E) \cap \mathfrak{B}(E')$.

(ii). $\mathfrak{B}(\{P\}) \cup \mathfrak{B}(\{Q\}) = \mathfrak{B}(\{PQ\})$.

(iii). $\mathfrak{B}(E) = \mathfrak{B}((E))$.

(iv). Si $E \subset F$, $\mathfrak{B}(F) \subset \mathfrak{B}(E)$.

(v). Tout ensemble algébrique affine est défini par un ensemble fini de polynômes (ceci car $K[X_1, \dots, X_n]$ est noetherien).

(vi). Toute union finie d'ensembles algébriques affines est un ensemble algébrique affine. En particulier, tout sous-ensemble fini de K^n est un ensemble algébrique affine.

(vii). Les ensembles algébriques affines définissent une topologie, dite de Zariski, pour laquelle ce sont des fermés.

Exemples. • Si $n = 1$, un ensemble est algébrique si et seulement si il est fini (éventuellement vide) ou égal à K .
• L'ensemble $\{(x, \sin x), x \in \mathbb{R}\}$ n'est pas algébrique.

Définition 8.7. Soit V un ensemble algébrique affine. On lui associe l'idéal $\mathfrak{I}(V) = \{P \in K[X_1, \dots, X_n] : \forall v \in V, P(v) = 0\}$, appelé *idéal de définition* de V .

Remarque. Si $V \subset W$, on a $\mathfrak{I}(W) \subset \mathfrak{I}(V)$.

Proposition 8.10. Si V est un ensemble algébrique affine, on a $V = \mathfrak{B}(\mathfrak{I}(V))$ et $\mathfrak{I}(\mathfrak{B}(\mathfrak{I}(V))) = \mathfrak{I}(V)$.

Démonstration. Soit $E \subset K[X_1, \dots, X_n]$ tel que $V = \mathfrak{B}(E) = \mathfrak{B}((E))$. Par définition, $(E) \subset \mathfrak{I}(V)$, donc $V = \mathfrak{B}((E)) \supset \mathfrak{B}(\mathfrak{I}(V))$. Mais réciproquement il est évident que $V \subset \mathfrak{B}(\mathfrak{I}(V))$ par définition de $\mathfrak{I}(V)$. $\square$

Corollaire 8.11. Soient V et W deux ensembles algébriques. Alors $V = W$ si et seulement si $\mathfrak{I}(V) = \mathfrak{I}(W)$.

8.2.2 Composantes irréductibles

Définition 8.8. Un ensemble algébrique affine V est dit *irréductible* s'il est non vide et si toute décomposition $V = V_1 \cup V_2$ en deux fermés algébriques impose $V = V_1$ ou $V = V_2$.

Proposition 8.12. Tout ensemble algébrique affine non vide se décompose de façon unique comme réunion finie d'ensembles algébriques affines irréductibles non vides, appelés *composantes irréductibles*, dont aucun n'est contenu dans un autre. L'unicité est entendue à permutation près.

Démonstration. Considérons d'abord l'existence. On suppose, par l'absurde, que l'ensemble algébrique affine V ne se décompose pas au sens de la proposition. En particulier, V n'est pas irréductible, donc $V = V_1 \cup V_2$, avec $V_1 \subsetneq V$ et $V_2 \subsetneq V$ et au moins un ensemble algébrique affine parmi V_1 ou V_2 ne se décompose pas au sens de la proposition. On construit ainsi une suite infinie strictement décroissante d'ensembles algébriques, ce qui définit, en prenant les idéaux de définition, une suite strictement croissante d'idéaux, ce qui contredit le fait que $K[X_1, \dots, X_n]$ est noetherien.

Après avoir retiré les ensembles contenus dans un autre, on obtient une décomposition sans inclusion. Pour l'unicité, écrivons $V = \bigcup_i V_i = \bigcup_j W_j$ sous cette forme. Comme $V_i = \bigcup_j (V_i \cap W_j)$ est irréductible, il existe j tel que $V_i \subset W_j$. De même, il existe k tel que $W_j \subset V_k$. L'absence d'inclusion entre composantes impose $k = i$, puis $V_i = W_j$. Par symétrie, les deux familles coïncident. $\square$

Proposition 8.13. V est irréductible si et seulement si $\mathfrak{F}(V)$ est premier.

Démonstration. $(\Rightarrow)$ On suppose $\mathfrak{F}(V)$ non premier. Alors, il existe $P, Q \notin \mathfrak{F}(V)$ tels que $PQ \in \mathfrak{F}(V)$. On a $V = \underbrace{(V \cap \mathfrak{B}((P)))}_{V'} \cup \underbrace{(V \cap \mathfrak{B}((Q)))}_{V''}$. Comme V est irréductible, on a $V = V'$ ou $V = V''$. Dans tous les cas, on a $P \in \mathfrak{F}(V)$ ou $Q \in \mathfrak{F}(V)$, ce qui contredit leur définition.

$(\Leftarrow)$ Si $\mathfrak{F}(V)$ est premier, il est propre, donc V n'est pas vide. On suppose V non irréductible. On écrit $V = V' \cup V''$, avec $V', V'' \subsetneq V$. Alors $\mathfrak{F}(V) \subsetneq \mathfrak{F}(V'), \mathfrak{F}(V'')$, donc il existe $P \in \mathfrak{F}(V') \setminus \mathfrak{F}(V)$ et $Q \in \mathfrak{F}(V'') \setminus \mathfrak{F}(V)$. On a $(P) \subset \mathfrak{F}(V')$, donc $V' = \mathfrak{B}(\mathfrak{F}(V')) \subset \mathfrak{B}((P))$. De même, on a $V'' \subset \mathfrak{B}((Q))$. Donc $\mathfrak{B}((PQ)) = \mathfrak{B}((P)) \cup \mathfrak{B}((Q)) \supset V' \cup V'' = V$, d'où $PQ \in \mathfrak{F}(V)$, ce qui contredit la primalité de $\mathfrak{F}(V)$. $\square$

8.2.3 Idéal radical

Définition 8.9. Soit I un idéal d'un anneau A . On appelle *radical* de I et on note $\sqrt{I}$ l'idéal $\{a \in A : \exists m \geq 1 : a^m \in I\}$.

Définition 8.10. On dit qu'un idéal I est *radical* si $I = \sqrt{I}$.

Proposition 8.14. Si V est un ensemble algébrique affine, $\mathfrak{F}(V)$ est un idéal radical.

8.2.4 Théorème des zéros de Hilbert (Nullstellensatz)

Rappels sur les éléments entiers

Définition 8.11. Soient $A \subset B$ deux anneaux. Un élément $x \in B$ est dit *entier* sur A s'il existe $d \geq 1$ et $a_0, \dots, a_{d-1} \in A$ tels que $x^d + a_{d-1}x^{d-1} + \dots + a_0 = 0$.

Proposition 8.15. Les assertions suivantes sont équivalentes :

- (i). x est entier sur A .
- (ii). $A[x]$ est un A -module de type fini.
- (iii). Il existe un sous-anneau C de B contenant A et x qui est un A -module de type fini.

Démonstration. ((i).) $\Rightarrow$ ((ii).) $\Rightarrow$ ((iii).). C'est évident.

((iii).) $\Rightarrow$ ((i).). Soit $y_1, \dots, y_d$ une partie génératrice de C . Comme $y \mapsto xy$ est A -linéaire, on peut écrire $xy_i = \sum_j a_{i,j}y_j$, donc $0 = (x - a_{i,i})y_i - \sum_{j \neq i} a_{i,j}y_j$. On pose

$$M = \begin{pmatrix} x - a_{1,1} & & -a_{1,j} \\ & \ddots & \\ -a_{i,j} & & x - a_{d,d} \end{pmatrix}.$$

Alors $M \begin{pmatrix} y_1 \\ \vdots \\ y_d \end{pmatrix} = 0$, d'où ${}^t \text{Co} M M \begin{pmatrix} y_1 \\ \vdots \\ y_d \end{pmatrix} = 0$, puis $\det M \begin{pmatrix} y_1 \\ \vdots \\ y_d \end{pmatrix} = 0$. Comme $1 \in C$ est combinaison linéaire des y_i à coefficients dans A , on obtient finalement $\det M = 0$. Il suffit alors de remarquer que $\det M$ est un polynôme unitaire de degré d en x . $\square$

Quelques lemmes préliminaires

Lemme 8.16. Soit K un corps algébriquement clos et L une extension de K telle que $[L : K] < \infty$. Alors $L = K$.

Lemme 8.17. Soit $K \subset L$ une extension de corps telle que L est une K -algèbre de type fini. Alors $[L : K] < \infty$.

Démonstration. On pose $L = K[\alpha_1, \dots, \alpha_p]$ et on démontre le résultat par récurrence sur p . Pour $p = 0$, c'est immédiat; pour $p = 1$, un anneau de polynômes en une indéterminée n'étant pas un corps, α_1 est algébrique. Supposons $p \geq 2$. Si tous les α_i sont algébriques sur K , l'extension est finie. Sinon, après renumérotation, supposons α_1 transcendant. Puisque $L = K(\alpha_1)[\alpha_2, \dots, \alpha_p]$, l'hypothèse de récurrence appliquée sur $K(\alpha_1)$ donne $[L : K(\alpha_1)] < \infty$. Ainsi, pour $i \geq 2$, α_i est algébrique sur $K(\alpha_1)$. On peut donc écrire $\alpha_i^{n_i} + \beta_{i,1}\alpha_i^{n_i-1} + \dots + \beta_{i,n_i} = 0$, avec $\beta_{i,j} \in K(\alpha_1)$. Donc $\gamma_{i,0}\alpha_i^{n_i} + \gamma_{i,1}\alpha_i^{n_i-1} + \dots + \gamma_{i,n_i} = 0$, avec $\gamma_{i,j} \in K[\alpha_1]$ et $\gamma_{i,0} \neq 0$. En multipliant par $\gamma_{i,0}^{n_i-1}$, on en déduit que $\gamma_{i,0}\alpha_i$ est entier sur $K[\alpha_1]$. On pose $\gamma = \prod_{i=2}^p \gamma_{i,0} \in K[\alpha_1]$. Alors pour tout $i \geq 2$, $\gamma\alpha_i$ est entier sur $K[\alpha_1]$. L'anneau $C = K[\alpha_1][\gamma\alpha_2, \dots, \gamma\alpha_p]$ est un module de type fini sur $K[\alpha_1]$: les équations unitaires de ses générateurs permettent de borner leurs exposants. Comme tout $\alpha \in L$ est un polynôme en les α_i , une puissance assez grande de γ donne $\gamma^m \alpha \in C$. La caractérisation précédente des éléments entiers montre donc que $\gamma^m \alpha$ est entier sur $K[\alpha_1]$. En particulier, pour tout $Q \in K[X]$, tel que $Q(\alpha_1) \neq 0$, il existe m_Q tel que $\frac{\gamma^{m_Q}}{Q(\alpha_1)}$ est entier sur $K[\alpha_1]$. On a donc

$$\left(\frac{\gamma^{m_Q}}{Q(\alpha_1)} \right)^{d_Q} + \delta_{Q,1} \left(\frac{\gamma^{m_Q}}{Q(\alpha_1)} \right)^{d_Q-1} + \dots + \delta_{Q,d_Q} = 0,$$

ce qui donne encore $\gamma^{m_Q d_Q} + \delta_{Q,1} Q(\alpha_1) \gamma^{m_Q(d_Q-1)} + \dots + \delta_{Q,d_Q} Q(\alpha_1)^{d_Q} = 0$. Or γ s'écrit $R(\alpha_1)$. Choisissons Q irréductible non constant ne divisant pas R ; un tel polynôme existe car $K[X]$ possède une infinité d'irréductibles unitaires. L'égalité précédente est alors une relation polynomiale non nulle en α_1 : modulo Q , son membre de gauche vaut $R^{m_Q d_Q} \neq 0$. Cela contredit la transcendance de α_1 . $\square$

Nullstellensatz

Dans cette sous-section, le corps K est supposé algébriquement clos.

Théorème 8.18 (Nullstellensatz faible). Soit K un corps algébriquement clos. Tout idéal maximal de $K[X_1, \dots, X_n]$ est de la forme $(X_1 - \alpha_1, \dots, X_n - \alpha_n)$, avec $\alpha_1, \dots, \alpha_n \in K$.

Démonstration. Soit $\mathfrak{M}$ un idéal maximal de $K[X_1, \dots, X_n]$. On pose $L = K[X_1, \dots, X_n]/\mathfrak{M}$. L est un corps, puisque $\mathfrak{M}$ est maximal. Le lemme 8.17 assure que $[L : K] < \infty$, tandis que le lemme 8.16 assure que $L = K$. Par conséquent, pour tout i , il existe $\alpha_i \in K$ tel que $X_i - \alpha_i \in \mathfrak{M}$. Donc $(X_1 - \alpha_1, \dots, X_n - \alpha_n) \subset \mathfrak{M}$. Or $(X_1 - \alpha_1, \dots, X_n - \alpha_n)$ est maximal, donc on a égalité. $\square$

Corollaire 8.19. Soit I un idéal propre de $\mathbb{K}[X_1, \dots, X_n]$. Alors $\mathfrak{B}(I) \neq \emptyset$.

Démonstration. Il existe $\mathfrak{M}$ maximal qui contient I , d'après le théorème de Krull. Le théorème 8.18 assure qu'il existe $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ tels que $\mathfrak{M} = (X_1 - \alpha_1, \dots, X_n - \alpha_n)$. On a $\mathfrak{B}(I) \supset \mathfrak{B}(\mathfrak{M}) \ni (\alpha_1, \dots, \alpha_n)$, d'où le résultat. $\square$

Corollaire 8.20 (Nullstellensatz fort). Soit I un idéal de $\mathbb{K}[X_1, \dots, X_n]$. Alors $\mathfrak{I}(\mathfrak{B}(I)) = \sqrt{I}$.

Démonstration. Si $P \in \sqrt{I}$, il existe $m \geq 1$ tel que $P^m \in I$. Donc pour tout $v \in \mathfrak{B}(I)$, $P^m(v) = 0$. Donc $P(v) = 0$. D'où $\sqrt{I} \subset \mathfrak{I}(\mathfrak{B}(I))$.

La réciproque peut se démontrer avec une idée de Rabinowitsch. On note $I = (P_1, \dots, P_r)$, comme $\mathbb{K}[X_1, \dots, X_n]$ est noetherien. Soit $Q \in \mathfrak{I}(\mathfrak{B}(I))$. Si $Q = 0$, la conclusion est immédiate ; supposons donc $Q \neq 0$. On introduit J l'idéal de $\mathbb{K}[X_1, \dots, X_{n+1}]$ engendré par $P_1, \dots, P_r$ et $X_{n+1}Q - 1$. On a nécessairement $\mathfrak{B}(J) = \emptyset$. Le corollaire 8.19 assure que $J = \mathbb{K}[X_1, \dots, X_{n+1}]$. On peut donc écrire $1 = \sum_i G_i P_i + (X_{n+1}Q - 1)R$. En spécialisant X_{n+1} en $\frac{1}{Q}$, on obtient $1 = \sum_i G_i \left(X_1, \dots, X_n, \frac{1}{Q(X_1, \dots, X_n)} \right) P_i(X_1, \dots, X_n)$. Il s'agit d'une fraction rationnelle dont le numérateur est dans I et dont le dénominateur est une puissance de Q . Par conséquent, $Q \in \sqrt{I}$. $\square$

8.2.5 Courbes algébriques

Généralités

Définition 8.12 (Courbe algébrique). Une courbe algébrique plane affine est définie par une équation $P(X, Y) = 0$, où $P \in \mathbb{K}[X, Y]$ est non constant. Sur un corps algébriquement clos, on choisit une équation réduite (sans facteur irréductible répété) pour que les multiplicités ci-dessous dépendent de la courbe. Les dessins représentent les points réels des courbes considérées.

Proposition 8.21. Soient $P, Q \in \mathbb{K}[X, Y]$ sans facteur commun. Alors $\mathfrak{B}((P)) \cap \mathfrak{B}((Q))$ est fini.

Démonstration. P et Q sont premiers entre eux, donc ils vérifient une relation de Bézout dans $\mathbb{K}(X)[Y]$: $AP + BQ = 1$, avec $A, B \in \mathbb{K}(X)[Y]$. Donc on a $CP + DQ = E$, avec $C, D \in \mathbb{K}[X, Y]$ et $E \in \mathbb{K}[X] \setminus \{0\}$. Si $(a, b) \in \mathfrak{B}((P)) \cap \mathfrak{B}((Q))$, alors a est racine de E , donc il y a un nombre fini de possibilités pour a . Par symétrie, on a le résultat. $\square$

Définition 8.13. On considère $\mathcal{C} = \mathfrak{B}((P))$ une courbe algébrique. Soit $(a, b) \in \mathcal{C}$. On a $P(X + a, Y + b) = \underbrace{P(a, b)}_{=0} + Q_1 + \dots + Q_m$, où les Q_i sont les composantes homogènes de degré i de $P(X + a, Y + b)$. La multiplicité, relativement à l'équation P , du point (a, b) de $\mathcal{C}$ est $m_P(a, b) = \inf\{i : Q_i \neq 0\}$. Si $m_P(a, b) = 1$, le point (a, b) est régulier. On vérifie que $Q_1 = \frac{\partial P}{\partial X}(a, b)X + \frac{\partial P}{\partial Y}(a, b)Y$. Sinon, le point est dit singulier.

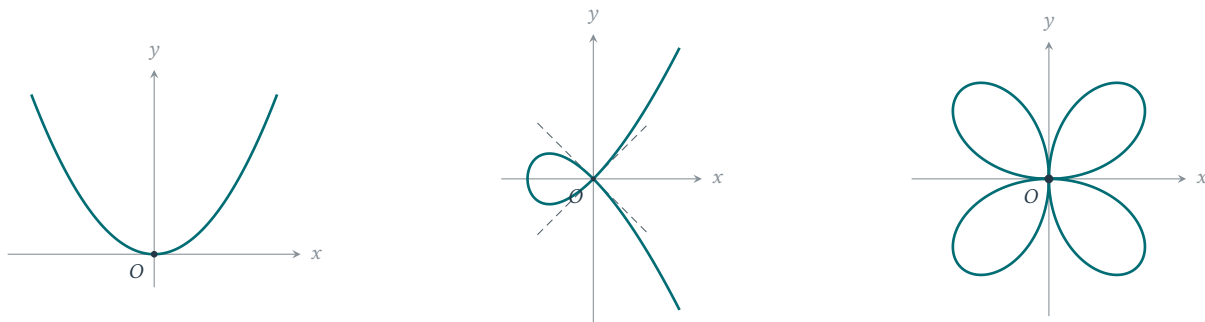


FIGURE 8.1 : Courbes réelles : point régulier, point double et point quadruple.

Exemples. Sur la figure 8.1, on a tracé de gauche à droite la courbe algébrique définie par le polynôme $P = Y - X^2$, celle définie par le polynôme $P = Y^2 - X^2 - X^3$ et celle définie par le polynôme $P = -4X^2Y^2 + (X^2 + Y^2)^3$. On considère le point $(0, 0)$ de ces courbes.

- Dans le premier cas, on a $Q_1 = Y$ et $Q_2 = -X^2$. La multiplicité est 1. Le point $(0, 0)$ est régulier.
- Dans le deuxième cas, on a $Q_1 = 0$, $Q_2 = Y^2 - X^2$ et $Q_3 = -X^3$. La multiplicité est 2. Le point $(0, 0)$ est singulier.
- Dans le troisième cas, on a $Q_1 = 0$, $Q_2 = 0$, $Q_3 = 0$, $Q_4 = -4X^2Y^2$, $Q_5 = 0$ et $Q_6 = (X^2 + Y^2)^3$. La multiplicité est 4. Le point $(0, 0)$ est singulier.

Définition 8.14. Soit $\mathcal{C} = \mathfrak{B}((P))$ une courbe algébrique. Soit $(\lambda, \mu) \in \mathbb{K}^2 \setminus \{(0, 0)\}$. La droite $(a, b) + \mathbb{K}(\lambda, \mu)$ est *tangente* à $\mathcal{C}$ en (a, b) si $P(a + T\lambda, b + T\mu) \in \mathbb{K}[T]$ est tel que la multiplicité de sa racine nulle est strictement supérieure à $m_P(a, b)$. Par convention, cette multiplicité vaut $+\infty$ lorsque le polynôme restreint est nul.

Exemple. Sur $\mathbb{R}$, on considère à nouveau $P = Y^2 - X^2 - X^3$. On a $P(\lambda T, \mu T) = T^2(\mu^2 - \lambda^2) - T^3\lambda^3$. Comme $m_P(0, 0) = 2$, la droite $\mathbb{K}(\lambda, \mu)$ est tangente si et seulement si $\mu^2 - \lambda^2 = 0$, c'est-à-dire $\lambda = \pm\mu$. Il y a donc deux tangentes, comme le suggère la figure 8.1.

Proposition 8.22. Si (a, b) est régulier, une tangente à $\mathcal{C}$ est la droite d'équation

$$\frac{\partial P}{\partial X}(a, b)(X - a) + \frac{\partial P}{\partial Y}(a, b)(Y - b) = 0.$$

Lien avec les ensembles algébriques projectifs

L'espace projectif $\mathbb{P}^n(\mathbb{K})$ est l'ensemble des droites vectorielles de $\mathbb{K}^{n+1}$; ses points sont notés $[x_0 : \dots : x_n]$, avec des coordonnées non toutes nulles, définies à un facteur non nul près. Un ensemble algébrique projectif correspond à un ensemble algébrique affine de $\mathbb{K}^{n+1}$ défini par un idéal engendré par des polynômes homogènes. On peut faire correspondre un polynôme $P = \sum_{i,j} a_{i,j} X^i Y^j \in \mathbb{K}[X, Y]$ avec un polynôme $\tilde{P} = \sum_{i,j} a_{i,j} X^i Y^j T^{\deg P - i - j}$ homogène de degré $\deg P$. On peut donc faire correspondre $\mathfrak{B}((P)) \subset \mathbb{K}^2$ avec $\mathfrak{B}((\tilde{P})) \subset \mathbb{P}^2(\mathbb{K})$. Si $v \in \mathfrak{B}((P))$, on a $(v, 1) \in \mathfrak{B}((\tilde{P}))$, ceci car $\tilde{P}(X, Y, 1) = P(X, Y)$. Réciproquement, si $w \in \mathfrak{B}((\tilde{P}))$, soit $(w_1/w_3, w_2/w_3) \in \mathfrak{B}((P))$ (si $w_3 \neq 0$), soit w est un point à l'infini.

Citons pour finir un théorème que nous ne démontrerons pas.

Théorème 8.23 (Bézout). Sur un corps $\mathbb{K}$ algébriquement clos, soient $\mathcal{C}$ et $\mathcal{C}'$ deux courbes du plan projectif, définies par des polynômes homogènes F et G sans facteur commun, de degrés d et e . Alors

$$\sum_{v \in \mathcal{C} \cap \mathcal{C}'} I_v(\mathcal{C}, \mathcal{C}') = de.$$

Ici I_v est la multiplicité locale d'intersection : dans une carte affine contenant v , si f, g sont les équations déshomogénéisées et $\mathfrak{m}_v$ l'idéal maximal du point, on pose

$$I_v(\mathcal{C}, \mathcal{C}') = \dim_{\mathbb{K}}(\mathbb{K}[x, y]_{\mathfrak{m}_v} / (f, g)).$$

L'anneau local $\mathbb{K}[x, y]_{\mathfrak{m}_v}$ est formé des fractions dont le dénominateur ne s'annule pas en v .

Pour deux courbes affines sans composante commune, la somme des multiplicités d'intersection dans le plan affine est donc au plus le produit des degrés : il faut ajouter les intersections à l'infini pour obtenir l'égalité. Par exemple, deux droites affines parallèles distinctes ont une unique intersection dans le plan projectif, située à l'infini.

Bibliographie

- [1] Y. Amice, *Les nombres p -adiques*, Presses Universitaires de France, Paris, 1992.
- [2] J.-M. Arnaudiès, *Les cinq polyèdres réguliers de $\mathbb{R}^3$ et leurs groupes*, Centre de documentation universitaire et S.E.D.E.S., Paris, 1969.
- [3] M. Berger, *Géométrie. Tomes 1 et 2*, Cassini, 2016.
- [4] A. Bouvier et D. Richard, *Groupes : observation, théorie, pratique*, 2^e édition revue et corrigée, Hermann, Paris, 1997.
- [5] P. G. Ciarlet, *Introduction à l'analyse numérique matricielle et à l'optimisation*, 5^e édition, Dunod, 2024.
- [6] L. Comtet, *Analyse combinatoire. Tomes I et II*, Presses Universitaires de France, Paris, 1970.
- [7] M. Demazure, *Cours d'algèbre*, 2^e édition, Cassini, 2008.
- [8] R. Goblot, *Algèbre linéaire*, Ellipses, 2005.
- [9] R. L. Graham, D. E. Knuth et O. Patashnik, *Mathématiques concrètes. Fondations pour l'informatique*, 2^e édition, Vuibert, 2003.
- [10] R. Mneimné et F. Testard, *Introduction à la théorie des groupes de Lie classiques*, Hermann, Paris, 1997.
- [11] D. Perrin, *Cours d'algèbre*, Ellipses, Paris, 1996.
- [12] J.-P. Serre, *Cours d'arithmétique*, 4^e édition, Presses Universitaires de France, Paris, 1994.
- [13] J. S. Milne, *Algebraic Geometry*, version 6.01, 2015 (chapitres 1, 2, 4, 6 et 7).