UNIVERSITÉ DE MONTPELLIER

ALGÈBRE 3

COURS DE L2 - HLMA301

JÉRÉMIE BRIEUSSEL

ANNÉE 2019-2020

Ce document regroupe les notes du cours d'algèbre donné au premier semestre de L2 à l'université de Montpellier lors de l'année universitaire 2019-2020.

À l'issue du cours, on attend de l'étudiant.e qu'il.elle connaisse parfaitement (c'est-à-dire qu'il.elle soit capable d'énoncer instantanément) toutes les définitions, toutes les propositions, tous les lemmes, tous les théorèmes et tous les corollaires exposés. De plus, on attend qu'il.elle soit capable de redémontrer rapidement un certain nombre de propositions et théorèmes. Il ne s'agit pas ici d'apprendre par coeur comme on récite une poésie, mais au contraire de comprendre suffisamment bien pour pouvoir réénoncer, retrouver le résultat tout.e seul.e.

Le cours est parsemé de nombreux exercices, souvent sous forme d'exemples, et complété par les exercices des feuilles de TD. La connaissance et la compréhension du cours, ainsi que la capacité à résoudre ces exercices constituent les objectifs principaux du semestre. Un.e étudiant.e à l'aise se doit de chercher à en résoudre un maximum. En fonction de la difficulté, il est évidemment préférable d'en discuter avec le chargé de TD, qui pourra suggérer des pistes de résolutions, et vérifier le cas échéant que les solutions de l'étudiant.e sont correctes.

La lecture de ces notes ne peut pas remplacer le cours donné par l'enseignant, qui reste la base du travail de l'étudiant.e. La prise de notes complètes et détaillées lors du cours d'amphithéatre est essentielle.

Ce document est amené à évoluer. Merci d'écrire à

jeremie.brieussel@umontpellier.fr

pour signaler des erreurs, poser une question ou pour tout commentaire.

Chapitre 0

Petit panorama des structures algèbriques

Dans son sens moderne, l'algèbre désigne l'étude des structures algébriques, c'est-à-dire la structure d'espaces (ensembles) munis d'opérations. Dans ce chapitre introductif, on définit brièvement les structures algèbriques les plus usuelles.

Les définitions de ces structures sont regroupées en début de polycopié dans ce chapitre zéro afin d'en donner une vue d'ensemble, mais ne seront abordées en amphithéatre qu'au fur et à mesure de leur apparition dans le cours.

Une *opération* $\cdot$ sur un ensemble E est une application

$$\begin{array}{rccc} \cdot : & E \times E & \to & E \\ & (a,b) & \mapsto & a \cdot b \end{array}$$

Exemples 0.0.1. On connait déjà beaucoup d'exemples d'opérations :

(a) l'addition $+$ sur des ensembles de nombres $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ ou sur des ensembles de fonctions comme $\mathcal{F}(\mathbb{R}, \mathbb{R})$, l'ensemble des applications de $\mathbb{R}$ dans $\mathbb{R}$, ou $\mathcal{F}(X, \mathbb{C})$ l'ensemble des applications d'un ensemble X quelconque vers $\mathbb{C}$.
Dans ce dernier espace, si $f_1, f_2 \in \mathcal{F}(X, \mathbb{C})$, alors $f_1 + f_2$ est définie par :

$$\forall x \in X, (f_1 + f_2)(x) = f_1(x) + f_2(x),$$

où le $+$ de gauche est l'addition dans $\mathcal{F}(X, \mathbb{R})$ et le $+$ de droite l'addition dans $\mathbb{C}$.

(b) la multiplication $\times$ dans les mêmes espaces.

(c) la composition $\circ$ sur l'ensemble $\mathcal{F}(\mathbb{R}, \mathbb{R})$ ou plus généralement sur l'ensemble $\mathcal{F}(X, X)$ des applications d'un ensemble X dans lui-même.

(d) On peut aussi restreindre un opération à un sous-espace F (i.e. une partie) de E stable par l'opération, par exemple l'opération d'addition sur l'ensemble $2\mathbb{Z}$ des entiers relatifs pairs, la composition $\circ$ sur l'ensemble $C_0(\mathbb{R}, \mathbb{R})$ des applications continues de $\mathbb{R}$ dans $\mathbb{R}$ ou encore la composition $\circ$ sur l'ensemble $\mathcal{L}(V)$ des applications linéaires de l'espace vectoriel V.

(e) Soit X un ensemble quelconque, et $E = \mathcal{P}(X)$. L'union $\cup$, l'intersection $\cap$ et la différence symétrique Δ sont des opérations sur E.

0.1 Groupes

Définition 0.1.1. *Un ensemble muni d'une opération $(G, \cdot)$ est un groupe si les trois conditions suivantes sont satisfaites*

(a) $\forall a, b, c \in G, (a \cdot b) \cdot c = a \cdot (b \cdot c)$ *appelée associativité,*

(b) $\exists e \in G, \forall a \in G, a \cdot e = e \cdot a = a$*, cet élément* e *est dit élément neutre du groupe* G*,*

(c) $\forall a \in G, \exists b \in G, a \cdot b = b \cdot a = e$*, cet élément* b *est alors appelé inverse de* a*.*

On vérifiera en exercice que l'élément neutre est unique, et que chaque élément a possède un unique inverse b. Cet inverse sera souvent noté a^{-1} (notation dite *multiplicative*). Toutefois, si la loi $\cdot$ est une addition $+$, on notera plutôt $-a$ pour l'inverse.

Définition 0.1.2. *Un groupe $(G, \cdot)$ est dit commutatif (ou aussi abélien) si de plus*

(d) $\forall a, b \in G, a \cdot b = b \cdot a$*.*

Exemples 0.1.3. (a) $(\mathbb{Z}, +), (\mathbb{R}, +), (\mathbb{C}, +)$ sont des groupes commutatifs. Leur élément neutre est 0. C'est aussi le cas de $(\mathbb{Z}/n\mathbb{Z}, +)$.

(b) $(\mathbb{R} \setminus \{0\}, \times), (\mathbb{R}^+ \setminus \{0\}, \times), (\mathbb{C} \setminus \{0\}, \times)$ sont des groupes commutatifs. Leur élément neutre est 1.

(c) $(M_n(\mathbb{C}), +)$ est un groupe. Quel est son élément neutre ?

(d) $(\mathcal{F}(X, \mathbb{C}), +)$ est un groupe commutatif. Quel est son élément neutre ?

(e) On note $\mathrm{Bij}(X)$ l'ensemble des bijections de l'ensemble X. Alors $(\mathrm{Bij}(X), \circ)$ est un groupe. Quel est son élément neutre ?

(f) $(\mathrm{GL}_n(\mathbb{C}), \times)$ est un groupe. Quel est son élément neutre ? On rappelle que $\mathrm{GL}_n(\mathbb{C})$ est l'ensemble des matrices $n \times n$ inversibles à coefficients complexes .

Exercice 1. Montrer que $(\mathbb{C}, -)$, $(\mathbb{C}, \times)$, $(M_n(\mathbb{C}), \times)$ et $(\mathcal{F}(\mathbb{R}, \mathbb{R}), \circ)$ ne sont pas des groupes.

0.2 Anneaux

Définition 0.2.1. *Soit A un ensemble muni de deux opérations $+$ et $\times$. On dit que $(A, +, \times)$ est un anneau si les trois conditions suivantes sont satisfaites :*

(a) $(A, +)$ *est un groupe commutatif,*

(b) $(A, \times)$ *est un semi-groupe, c'est-à-dire satisfait les deux conditions suivantes :*

(i) $\forall a, b, c \in A, a \times (b \times c) = (a \times b) \times c$*, associativité,*

(ii) $\exists e \in A, \forall a \in A, a \times e = e \times a = a$*, cet élément* e *est appelé neutre de la multiplication,*

(c) $\forall a, b, c \in A, a \times (b + c) = (a \times b) + (a \times c)$ *et* $(b + c) \times a = b \times a + c \times a$*, appelée distributivité.*

Ici aussi, l'élément neutre multiplicatif est unique (exercice).

Définition 0.2.2. *Un anneau $(A, +, \times)$ est dit commutatif si de plus*

(d) $\forall a, b \in A, a \times b = b \times a$*, c'est-à-dire si l'opération de multiplication est commutative.*

Exemples 0.2.3. Les ensembles de nombres $(\mathbb{Z}, +, \times), (\mathbb{Q}, +, \times), (\mathbb{R}, +, \times), (\mathbb{C}, +, \times)$, les espaces de fonctions $(\mathcal{F}(X, \mathbb{Z}), +, \times), (\mathcal{F}(X, \mathbb{Q}), +, \times), (\mathcal{F}(X, \mathbb{R}), +, \times), (\mathcal{F}(X, \mathbb{C}), +, \times)$, ainsi que les ensembles de congruences $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ pour $n \geq 2$ (voir cours d'arithmétique) sont des anneaux commutatifs. Les ensembles de matrices $(M_n(\mathbb{Z}), +, \times), (M_n(\mathbb{Q}), +, \times), (M_n(\mathbb{R}), +, \times)$, $(M_n(\mathbb{C}), +, \times)$ sont des anneaux non commutatifs. Quels sont leurs neutres additif et multiplicatif ?

0.3 Corps

Définition 0.3.1. *Soit $\mathbb{K}$ un ensemble muni de deux opérations $+$ et $\times$. On dit que $(\mathbb{K}, +, \times)$ est un corps si c'est un anneau commutatif et de plus*

(e) $\forall a \in \mathbb{K} \setminus \{0\}, \exists b \in \mathbb{K} \setminus \{0\}, a \times b = b \times a = e$*, c'est-à-dire si tout élément non-nul admet un inverse pour la multiplication.*

En particulier, un corps est un cas particulier d'anneau. Dans la littérature, il n'est pas toujours requis qu'un corps soit commutatif. Au cours de ce semestre, on supposera toujours que les corps sont commutatifs. Pour désigner un anneau satisfaisant la condition (e) mais pas nécessairement la condition (d) on parlera de *corps gauche*. L'exemple le plus célèbre de corps gauche est celui des quaternions (cf wikipedia).

Exemples 0.3.2. Les ensembles de nombres $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ muni de l'addition et la multiplication sont des corps.

Exercice 2. Soit $\mathbb{D}$ l'ensemble des nombres décimaux. Démontrer que $(\mathbb{D}, +, \times)$ est un anneau commutatif. Est-ce un corps ?

Exercice 3. Montrer que $(\mathbb{Z}, +\times)$ n'est pas un corps. Montrer que $(M_n(\mathbb{C}), +, \times)$ n'est pas un corps gauche.

Exercice 4. (Difficile) Soit $n \geq 2$, montrer que $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ est un corps si et seulement si n est un nombre premier. Notez que dans ce cas c'est un corps avec un nombre fini d'éléments.

Notez qu'il existe en mathématique bien d'autres corps (plus compliqués). Ce semestre, nous pourrions nous restreindre aux cas où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, mais la théorie reste valide pour tout autre corps. On fera parfois l'hypothèse supplémentaire que le corps $\mathbb{K}$ est infini.

0.4 Espaces vectoriels

Définition 0.4.1. *Soit $\mathbb{K}$ un corps et V un ensemble muni d'une opération $+$ et d'une loi externe $\cdot$*

$$\begin{array}{rccc} \cdot : & \mathbb{K} \times V & \to & V \\ & (\lambda, v) & \mapsto & \lambda \cdot v \end{array}$$

On dit que $(V, +, \cdot)$ est un espace vectoriel sur $\mathbb{K}$ (dit aussi $\mathbb{K}$-espace vectoriel) si les conditions suivantes sont satisfaites :

(a) $(V, +)$ *est un groupe commutatif,*

(b) $\forall \lambda, \mu \in \mathbb{K}, \forall v, w \in V$*,*

(i) $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$*,*

(ii) $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$*,*

(iii) $\lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v$*, où* $\lambda\mu$ *désigne la multiplication dans* $\mathbb{K}$*,*

(iv) $1 \cdot v = v$ *où* 1 *désigne le neutre multiplicatif de* $\mathbb{K}$*.*

Les espaces vectoriels sur $\mathbb{R}$ et $\mathbb{C}$ ont éte étudiés en L1. Pour suivre ce cours, il est indispensable d'être à l'aise avec les notions d'espaces vectoriels et d'applications linéaires.

0.5 Algèbres

Définition 0.5.1. *Soit* $\mathbb{K}$ *un corps et* $\mathcal{A}$ *un ensemble muni de deux opérations* $+$ *et* $\times$ *ainsi que d'une loi externe* $\cdot : \mathbb{K} \times \mathcal{A} \to \mathcal{A}$*. On dit que* $(\mathcal{A}, +, \times, \cdot)$ *est une algèbre sur le corps* $\mathbb{K}$ *(en abrégé une* $\mathbb{K}$*-algèbre) si les trois conditions suivantes sont satisfaites :*

(a) $(\mathcal{A}, +, \times)$ *est un anneau,*

(b) $(\mathcal{A}, +, \cdot)$ *est un* $\mathbb{K}$*-espace vectoriel,*

(c) $\forall \lambda \in \mathbb{K}, \forall a, b \in \mathcal{A}, (\lambda \cdot a) \times b = \lambda \cdot (a \times b) = a \times (\lambda \cdot b)$.

Une $\mathbb{K}$-algèbre $\mathcal{A}$ est dite *commutative* si l'anneau $(\mathcal{A}, +, \times)$ est commutatif.

En pratique, on ne note ni les $\times$ ni les $\cdot$ car la relation (c) de compatibilité assure que l'ordre de ces opérations n'a pas d'importance. Par contre, il est indispensable de noter les $+$! ! !

Exemples 0.5.2. Soit $\mathbb{K}$ un corps.

(a) Le corps $\mathbb{K}$ lui même est une $\mathbb{K}$-algèbre commutative.

(b) Soit X un ensemble, l'espace $\mathcal{F}(X, \mathbb{K})$ des fonctions de X dans $\mathbb{K}$ est une $\mathbb{K}$-algèbre commutative.

(c) L'espace des matrices $M_n(\mathbb{K})$ est une $\mathbb{K}$-algèbre, non-commutative dés que $n \geq 2$.

(d) Si V est un $\mathbb{K}$-espace vectoriel, l'espace $\mathcal{L}(V)$ des applications linéaires de V dans lui-même est une $\mathbb{K}$-algèbre, non-commutative dés que $\dim(V) \geq 2$.
Si V est de dimension finie n, cette $\mathbb{K}$-algèbre est isomorphe à (c'est-à-dire la même que) $M_n(\mathbb{K})$ via la correspondance entre matrices et applications linéaires.

(e) L'algèbre des polynômes $\mathbb{K}[X]$ que l'on va étudier au chapitre suivant est une $\mathbb{K}$-algèbre commutative.

(f) Les algèbres de polynômes en plusieurs variables $\mathbb{K}[X, Y]$ ou $\mathbb{K}[X_1, \ldots, X_n]$ sont des $\mathbb{K}$-algèbres commutatives.

Exercice 5. Soit X un ensemble quelconque. On rappelle que $\mathbb{Z}/2\mathbb{Z}$ est le corps (commutatif) à deux éléments avec l'addition $0+0 = 1+1 = 0$, $0+1 = 1$ et la multiplication $0 \times 0 = 0 \times 1 = 0$, $1 \times 1 = 1$. On note $(\mathbb{Z}/2\mathbb{Z})^X$ l'ensemble des fonctions de X dans $\mathbb{Z}/2\mathbb{Z}$.

(a) Montrer que $((\mathbb{Z}/2\mathbb{Z})^X, +, \times, \cdot)$ est une $\mathbb{Z}/2\mathbb{Z}$-algèbre. On identifiera clairement les opérations et la loi externe en question.

(b) Soit $A \in \mathcal{P}(X)$, on appelle fonction indicatrice de A la fonction $\mathbb{1}_A : X \to \mathbb{Z}/2\mathbb{Z}$ donnée par :

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{si } x \notin A. \end{cases}$$

Montrer que l'application

$$\begin{array}{rccc} \Phi : & \mathcal{P}(X) & \to & (\mathbb{Z}/2\mathbb{Z})^X \\ & A & \mapsto & \mathbb{1}_A \end{array}$$

est une bijection.

(c) Montrer que $\forall A, B \in (\mathbb{Z}/2\mathbb{Z})^X, \forall \lambda \in \mathbb{Z}/2\mathbb{Z}$,

$$\Phi(A \Delta B) = \mathbb{1}_A + \mathbb{1}_B, \Phi(A \cap B) = \mathbb{1}_A + \mathbb{1}_B, \text{ et } \Phi(\lambda \cdot A) = \lambda \mathbb{1}_A,$$

où par définition $\lambda \cdot A = \emptyset$ si $\lambda = 0$ et $\lambda \cdot A = A$ si $\lambda = 1$, et $A \Delta B = (A \cup B) \setminus (A \cap B)$ est la différence symétrique.

(d) En déduire que $(\mathcal{P}(X), \Delta, \cap, \cdot)$ est une $\mathbb{Z}/2\mathbb{Z}$-algèbre. On dit que Φ est un isomorphisme de $\mathbb{Z}/2\mathbb{Z}$-algèbres.

L'étude générale de ces structures algèbriques est très complexe. On l'approfondira au cours des semestres et années suivantes. Dans le cadre de ce cours, on étudiera des exemples spécifiques et fondamentaux, comme l'algèbre des polynômes $\mathbb{K}[X]$, le groupe symétrique Σ_n et l'algèbre des matrices complexes $M_n(\mathbb{C})$. Dans le cadre du cours d'arithmétique HLMA304, on étudiera l'anneau des entiers $\mathbb{Z}$ et les anneaux de congruences $\mathbb{Z}/n\mathbb{Z}$.

Chapitre 1

Diagonalisation des endomorphismes

Dans tout le chapitre, E désigne un $\mathbb{K}$-espace vectoriel de dimension n.

1.1 Préliminaires et rappels

1.1.1 Sommes directes

Définition 1.1.1. *Soit E un $\mathbb{K}$-espace vectoriel. Soient F_1, F_2 deux sous-espaces vectoriels. La somme $F_1 + F_2$ des espaces vectoriels est*

$$F_1 + F_2 = \{u_1 + u_2 | u_1 \in F_1, u_2 \in F_2\}.$$

C'est encore un sous-espace vectoriel de E. On dit que la somme $F_1 + F_2$ est directe *si*

$$\forall u_1 \in F_1, \forall u_2 \in F_2, \text{ si } u_1 + u_2 = 0 \text{ alors } u_1 = u_2 = 0.$$

On note alors $F_1 \oplus F_2$.

Proposition 1.1.2. *Soit E un $\mathbb{K}$-espace vectoriel et F_1, F_2 deux sous-espaces vectoriels. Les assertions suivantes sont équivalentes :*

(a) la somme $F_1 \oplus F_2$ est directe, c'est-à-dire $\forall u_1 \in F_1, \forall u_2 \in F_2$, si $u_1 + u_2 = 0$ alors $u_1 = u_2 = 0$.

(b) $\forall u_1, v_1 \in F_1, \forall u_2, v_2 \in F_2$, si $u_1 + u_2 = v_1 + v_2$, alors $u_1 = v_1$ et $u_2 = v_2$ (unicité de la décomposition comme somme),

(c) si L_1 est une famille libre de F_1 et L_2 est une famille libre de F_2, alors $L_1 \cup L_2$ est une famille libre de E,

(d) si $\mathcal{B}_1$ est une base de F_1 et $\mathcal{B}_2$ est une base de F_2, alors $\mathcal{B}_1 \cup \mathcal{B}_2$ est une base de $F_1 + F_2$,

(e) $\dim(F_1 + F_2) = \dim(F_1) + \dim(F_2)$,

(f) $F_1 \cap F_2 = \{0\}$.

Démonstration. Exercice. □

Exercice 6. Montrer que $F_1 \oplus (F_2 \oplus F_3)$ équivaut à $(F_1 \oplus F_2) \oplus F_3$.

Pour la définir la somme directe de k sous-espaces $F_1, \dots, F_k$, on dit que $F_1 \oplus \cdots \oplus F_k$ si $F_1 \oplus \cdots \oplus F_{k-1}$ et $(F_1 \oplus \cdots \oplus F_{k-1}) \oplus F_k$. L'exercice précédent montre que cela ne dépend pas de l'ordre choisi.

L'analogue de la proposition 1.1.2 reste vrai pour les assertions (a) à (e) (exercice).

Exercice 7. Donner un exemple de trois sous-espaces de $\mathbb{R}^2$ tels que $F_1 \cap F_2 \cap F_3 = \{0\}$ mais dont la somme n'est pas directe.

1.1.2 Changements de bases, matrices de passage

Soient $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{B}' = (v_1, \dots, v_n)$ deux bases de E. Soit $x \in E$. Les *coordonnées* de x dans $\mathcal{B}$ forment l'unique n-uplet $(x_1, \dots, x_n) \in \mathbb{K}^n$ tel que

$$x = \sum_{i=1} x_i e_i \quad \text{on note aussi} \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}} \tag{1.1}$$

Notez que cette notation est quelque peu abusive puisque x est un vecteur et les coordonnées forment une colonne. De même, les coorodnnées de x dans $\mathcal{B}'$ forment l'unique n-uplet $(x'_1, \dots, x'_n) \in \mathbb{K}^n$ tel que

$$x = \sum_{i=1} x'_i v_i = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}_{\mathcal{B}'}$$

Ces coordonnées dans des bases différentes sont reliées par la matrice de passage que l'on va définir maintenant. Pour tout vecteur v_j de la base $\mathcal{B}'$, on note $(p_{1j}, \dots, p_{nj})$ ses coordonnées dans $\mathcal{B}$, c'est-à-dire

$$v_j = \sum_{i=1}^{n} p_{ij} e_i = \begin{pmatrix} p_{1j} \\ \vdots \\ p_{nj} \end{pmatrix}_{\mathcal{B}}.$$

La matrice de passage est la matrice

$$P = ((v_1)_{\mathcal{B}} \dots (v_n)_{\mathcal{B}}) = \left(\begin{pmatrix} p_{11} \\ \vdots \\ p_{n1} \end{pmatrix}_{\mathcal{B}} \cdots \begin{pmatrix} p_{1n} \\ \vdots \\ p_{nn} \end{pmatrix}_{\mathcal{B}} \right) = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}$$

ayant comme $j^{\text{ième}}$ colonne les coordonnées de v_j dans $\mathcal{B}$.

Proposition 1.1.3.

$$\forall x \in E, \quad P \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}_{\mathcal{B}'} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}}.$$

Démonstration. Par définition du produit matriciel, il s'agit de montrer que $\forall 1 \leq i \leq n$, on a $x_i = \sum_{j=1}^{n} p_{ij} x'_j$. Or par définition des coordonnées, on a

$$x = \sum_{j=1}^{n} x'_j v_j = \sum_{j=1}^{n} x'_j \left(\sum_{i=1}^{n} p_{ij} e_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} p_{ij} x'_j \right) e_i.$$

L'égalité voulue découle de la comparaison avec (1.1) et de l'unicité des coordonnées dans la base $\mathcal{B}$. □

Soit $\varphi \in \mathcal{L}(E)$ un endomorphisme de E, alors la matrice de φ dans la base $\mathcal{B}$ est

$$A = \mathrm{Mat}_{\mathcal{B}}(\varphi) = ((\varphi(e_1))_{\mathcal{B}} \dots (\varphi(e_n))_{\mathcal{B}})$$

dont les colonnes sont les coordonnées dans $\mathcal{B}$ des images de vecteurs de $\mathcal{B}$ par φ. Avec toujours le même abus de notation, cela signifie que

$$\text{si} \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}}, \quad \text{alors} \quad \varphi(x) = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}}.$$

Pour notre deuxième base $\mathcal{B}'$, on note

$$C = \mathrm{Mat}_{\mathcal{B}'}(\varphi) = ((\varphi(v_1))_{\mathcal{B}'} \dots (\varphi(v_n))_{\mathcal{B}'})$$

Proposition 1.1.4. *Avec les notations ci-dessus :*

$$AP = PC.$$

Comme P est une matrice de passage, elle est inversible. (Pourquoi ?)

Démonstration. On a d'une part

$$PC = (P(\varphi(v_1))_{\mathcal{B}'} \dots P(\varphi(v_n))_{\mathcal{B}'}) = ((\varphi(v_1))_{\mathcal{B}} \dots (\varphi(v_n))_{\mathcal{B}})$$

d'après la Proposition 1.1.3, et d'autre part

$$AP = A\left((v_1)_{\mathcal{B}} \cdots (v_n)_{\mathcal{B}}\right) = (A(v_1)_{\mathcal{B}} \cdots A(v_n)_{\mathcal{B}}) = ((\varphi(v_1))_{\mathcal{B}} \dots (\varphi(v_n))_{\mathcal{B}})$$

par définition de A. □

Définition 1.1.5. *Deux matrices $A, C \in \mathrm{M}_n(\mathbb{K})$ sont dites* conjuguées *si il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $AP = PC$.*

Cela équivaut à $P^{-1}AP = C$ ou encore $A = PCP^{-1}$.

Exercice 8. Soit $P \in \mathrm{GL}_n(\mathbb{K})$, montrer que l'application "conjugaison par P" de $\mathrm{M}_n(\mathbb{K})$ dans lui-même donnée par $A \mapsto P^{-1}AP$ est un isomorphisme de $\mathbb{K}$-algèbre.

On rappelle (et on reformule) le résultat de L1 :

Proposition 1.1.6. *L'application :*

$$\begin{array}{rccc} \mathrm{Mat}_{\mathcal{B}} : & \mathcal{L}(E) & \to & \mathrm{M}_n(\mathbb{K}) \\ & \varphi & \mapsto & \mathrm{Mat}_{\mathcal{B}}(\varphi) \end{array}$$

est un isomorphisme de $\mathbb{K}$-algèbres, ce qui signifie qu'elle est bijective et que

- $\mathrm{Mat}_{\mathcal{B}}(\mathrm{id}_E) = I_n$
- *pour tous $\varphi_1, \varphi_2 \in \mathcal{L}(E)$ et tous $\lambda, \mu \in \mathbb{K}$, on a*

$$\mathrm{Mat}_{\mathcal{B}}(\lambda\varphi_1 + \mu\varphi_2) = \lambda\mathrm{Mat}_{\mathcal{B}}(\varphi_1) + \mu\mathrm{Mat}_{\mathcal{B}}(\varphi_2)$$

- *pour tous $\varphi_1, \varphi_2 \in \mathcal{L}(E)$, on a*

$$\mathrm{Mat}_{\mathcal{B}}(\varphi_1 \circ \varphi_2) = \mathrm{Mat}_{\mathcal{B}}(\varphi_1)\mathrm{Mat}_{\mathcal{B}}(\varphi_2).$$

L'étudiant·e est invité·e à reprouver cette proposition à titre d'exercice.

1.2 Diagonalisation

1.2.1 Définition-vocabulaire

Définition 1.2.1. *Un endomorphisme $\varphi \in \mathcal{L}(E)$ est* diagonalisable *si il existe une base $\mathcal{B}' = (v_1, \ldots, v_n)$ de E telle que* $\mathrm{Mat}_{\mathcal{B}'}(\varphi) = \begin{pmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{pmatrix}$ *soit diagonale.*

Cela signifie que pour tout $1 \leq i \leq n$, il existe $d_i \in \mathbb{K}$ tel que $\varphi(v_i) = d_i v_i$, c'est-à-dire que l'image du vecteur v_i par φ lui est colinéaire.

Définition 1.2.2. *Soit $\varphi \in \mathcal{L}(E)$.*

- *Un scalaire $\lambda \in \mathbb{K}$ est une* valeur propre *de φ si il existe $v \in E \setminus \{0\}$, tel que $\varphi(v) = \lambda v$.*
- *Un tel vecteur v* **non nul** *est appelé* vecteur propre *de φ pour la valeur propre λ.*
- *L'ensemble des valeurs propres de φ est appelé le* spectre *de φ, noté* $\mathrm{Sp}(\varphi)$.
- *Pour $\lambda \in \mathrm{Sp}(\varphi)$, le* sous-espace propre *associé à λ est*

$$E_\lambda = \mathrm{Ker}(\varphi - \lambda\mathrm{id}_E) = \{v \in E | \varphi(v) - \lambda v = 0\}.$$

Avec ce vocabulaire, on peut **reformuler** la notion d'endomorphisme diagonalisable, en disant que $\varphi \in \mathcal{L}(E)$ est diagonalisable si et seulement si l'espace E admet une base $\mathcal{B}'$ formée de vecteurs propres de φ.

Fait 1.2.3. *Soit $\varphi \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$. Les assertions suivantes sont équivalentes :*

(a) $\lambda \in \mathrm{Sp}(\varphi)$.

(b) $\exists v \in E \setminus \{0\}, \varphi(v) = \lambda v$.

(c) $\exists v \in E \setminus \{0\}, (\varphi - \lambda \mathrm{id}_E)(v) = 0$.

(d) $(\varphi - \lambda \mathrm{id}_E)$ *est non-injective.*

(e) $\dim \ker(\varphi - \lambda \mathrm{id}_E) = \dim(E_\lambda) > 0$.

(f) $\dim \mathrm{im}(\varphi - \lambda \mathrm{id}_E) = \mathrm{rg}(\varphi - \lambda \mathrm{id}_E) < n$.

(g) $(\varphi - \lambda \mathrm{id}_E)$ *est non-surjective.*

(h) $(\varphi - \lambda \mathrm{id}_E)$ *est non-bijective.*

Démonstration. Il s'agit de diverses manières équivalentes de formuler que l'endomorphisme $\varphi - \lambda \mathrm{id}_E$ est non- bijectif. On rappelle que par définition :

$$\forall v \in E, (\varphi - \lambda \mathrm{id}_E)(v) = \varphi(v) - \lambda \mathrm{id}_E(v) = \varphi(v) - \lambda v.$$

□

On définit aussi la diagonalisabilité d'une matrice. Pour cela, on rappelle que si $A \in \mathrm{M}_n(\mathbb{K})$, on peut lui associer un endomorphisme $\varphi_A \in \mathcal{L}(\mathbb{K}^n)$ du $\mathbb{K}$-espace vectoriel $\mathbb{K}^n$ comme suit :

$$\begin{array}{rccc} \varphi_A : & \mathbb{K}^n & \to & \mathbb{K}^n \\ & v & \mapsto & Av \end{array}$$

Par construction, on a $\mathrm{Mat}_{\mathcal{B}_{\mathrm{can}}}(\varphi_A) = A$.

Définition 1.2.4. *Une matrice $A \in \mathrm{M}_n(\mathbb{K})$ est* diagonalisable *si l'endomorphisme $\varphi_A \in \mathcal{L}(\mathbb{K}^n)$ est diagonalisable.*

Proposition 1.2.5. *Une matrice $A \in \mathrm{M}_n(\mathbb{K})$ est diagonalisable si et seulement si il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $P^{-1}AP$ soit une matrice diagonale.*

Démonstration. Par définition, la matrice A est diagonalisable si et seulement si il existe une base $\mathcal{B}'$ de $\mathbb{K}^n$ telle que $C := \mathrm{Mat}_{\mathcal{B}'}(\varphi_A)$ soit diagonale. D'après la proposition 1.1.4, toute matrice de φ_A dans une base $\mathcal{B}'$ est de la forme $C = P^{-1}AP$ où P est la matrice de passage, inversible. □

On note que la Définition 1.2.2 et la Fait 1.2.3 sont aussi valables pour les matrices en remplacant φ par A, id_E par I_n et E par $\mathbb{K}^n$.

1.2.2 Caractérisation des endomorphismes diagonalisables

On dispose d'une caractérisation complète :

Théorème 1.2.6. *Soit E un $\mathbb{K}$-espace vectoriel de dimension n et soit $\varphi \in \mathcal{L}(E)$. Alors les assertions suivantes sont équivalentes :*

(a) φ est diagonalisable.

(b) la somme des dimensions des sous-espaces propres $E_\lambda = \mathrm{Ker}(\varphi - \lambda \mathrm{id}_E)$ vaut n, i.e.

$$\sum_{\lambda \in \mathrm{Sp}(\varphi)} \dim(E_\lambda) = n.$$

Démonstration de (a) implique (b) du Théorème 1.2.6. On suppose que φ est diagonalisable, c'est-à-dire qu'on dispose d'une base $\mathcal{B}' = (v_1, \ldots, v_n)$ telle que $\mathrm{Mat}_{\mathcal{B}'}(\varphi) = D = \begin{pmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{pmatrix}$ soit diagonale. Alors on peut calculer le polynôme caractéristique $P_\varphi(X) = \det(D - XI_n) = \prod_{i=1}^n (d_i - X)$. Ses racines, les valeurs propres de φ, sont donc les d_i. (Notez qu'il se peut que $d_i = d_j$ même si $i \neq j$.) On note $\{\lambda_1, \ldots, \lambda_k\} = \mathrm{Sp}(\varphi)$ les valeurs prises par les valeurs propres (racines de P_φ, coefficients diagonaux de D), de sorte que $\lambda_i \neq \lambda_j$ dés que $i \neq j$. Alors

$$\sum_{\lambda \in \mathrm{Sp}(\varphi)} \dim(E_\lambda) = \sum_{j=1}^{k} \dim(\mathrm{Ker}(D - \lambda_j I_n)).$$

On est donc ramené à déterminer la dimension du noyau de la matrice

$$D - \lambda_j I_n = \begin{pmatrix} d_1 - \lambda_j & 0 & \cdots & 0 \\ 0 & d_2 - \lambda_j & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n - \lambda_j \end{pmatrix}$$

Le noyau est l'ensemble de n-uplets $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ satisfaisant le système $\forall 1 \leq i \leq n, (d_i - \lambda_j)x_i = 0$.

Ce sont donc les n-uplets tels que $x_i = 0$ quand $d_i \neq \lambda_j$ et $x_i \in \mathbb{R}$ quelconque quand $d_i = \lambda_j$. En d'autres termes, $\mathrm{Ker}(D - \lambda_j I_n)$ est l'espace vectoriel engendré par les vecteurs v_i pour lesquels $d_i = \lambda_j$, et sa dimension est le nombre de fois que le nombre λ_j apparait sur la diagonale de D. Au total, la somme des dimension est le nombre d'entrées sur la diagonale : n. □

L'étudiant·e est invité·e à relire la preuve ci-dessus dans le cas

$$D = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

Pour prouver la réciproque, on établit d'abord un

Lemme 1.2.7. *Soit $\varphi \in \mathcal{L}(E)$. Soit $\mathrm{Sp}(\varphi) = \{\lambda_1, \dots, \lambda_k\}$ avec λ_i deux à deux distincts, et $E_{\lambda_i} = \mathrm{Ker}(\varphi - \lambda_i \mathrm{id}_E)$. Alors la somme $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$ est directe, c'est-à-dire que si $\forall 1 \leq i \leq k, w_i \in E_{\lambda_i}$ et si $\sum_{i=1}^k w_i = 0$ alors $\forall 1 \leq i \leq n, w_i = 0$. En particulier :*

$$\dim(E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}) = \sum_{i=1}^{k} \dim(E_{\lambda_i}).$$

Démonstration du Lemme 1.2.7. On procède par récurrence sur k. Le résultat est évident si $k = 1$. Supposons le lemme vrai pour $k-1$ sous-espaces propres. On considère $w_i \in E_{\lambda_i}$ tels que

$$\sum_{i=1}^{k} w_i = 0. \tag{1.2}$$

Il s'agit de montrer que $\forall 1 \leq i \leq n, w_i = 0$. Pour cela, on applique φ à (1.2), on déduit

$$0 = \varphi(0) = \varphi\left(\sum_{i=1}^{k} w_i\right) = \sum_{i=1}^{k} \varphi(w_i) = \sum_{i=1}^{k} \lambda_i w_i, \tag{1.3}$$

où la dernière égalité utilise $w_i \in \mathrm{Ker}(\varphi - \lambda_i \mathrm{id}_E)$. En combinant (1.2) et (1.3), on obtient

$$0 = \sum_{i=1}^{k} \lambda_i w_i - \lambda_k \sum_{i=1}^{k} w_i = \sum_{i=1}^{k} (\lambda_i - \lambda_k) w_i = \sum_{i=1}^{k-1} (\lambda_i - \lambda_k) w_i.$$

Comme les vecteurs $(\lambda_i - \lambda_k)w_i$ sont dans E_{λ_i}, il s'agit d'une combinaison linéaire nulle de vecteurs appartenant à $k-1$ sous-espaces propres. Par hypothèse de récurrence, on en déduit que $\forall 1 \leq i \leq k-1, (\lambda_i - \lambda_k)w_i = 0$, donc que $\forall 1 \leq i \leq k-1, w_i = 0$ puisque les λ_i sont deux à deux distincts. En réinjectant dans (1.2), on obtient que $w_k = 0$ aussi. □

Démonstration de (b) implique (a) du Théorème 1.2.6. Si $\sum_{\lambda \in \mathrm{Sp}(\varphi)} \dim(E_\lambda) = n$, alors avec le Lemme 1.2.7, on a $\dim(E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}) = \dim(E)$, donc une décomposition en somme directe :

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}.$$

Dans ce cas, si $\mathcal{B}'_i$ est une base de E_{λ_i} pour chaque $1 \leq i \leq k$, alors $B' = \cup_{i=1}^k \mathcal{B}'_i$ est une base de E, formée de vecteurs propres. Cela signifie que φ est diagonalisable. □

Remarque 1.2.8. La preuve du théorème assure aussi que φ est diagonalisable si et seulement si $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$.

Corollaire 1.2.9. *Soit $\varphi \in \mathcal{L}(E)$. Si φ possède n valeurs propres deux à deux distinctes, alors φ est diagonalisable.*

Démonstration. Notons $\mathrm{Sp}(\varphi) = \{\lambda_1, \ldots, \lambda_n\}$ avec des λ_i deux à deux distincts. Comme $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$ est un sous-espace de E, on a toujours $\dim(E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}) \leq n$, et en particulier ici pour $k = n$.

D'autre part, comme λ_i est une valeur propre, on a nécessairement $\dim(E_{\lambda_i}) \geq 1$ pour tout i. Le Lemme 1.2.7 assure donc que

$$\dim(E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_n}) = \sum_{i=1}^{n} \dim(E_{\lambda_i}) \geq n.$$

Au total, on en déduit $\sum_{i=1}^{n} \dim(E_{\lambda_i}) = n$ donc φ diagonalisable d'après le Théorème 1.2.6. □

1.2.3 Méthode de diagonalisation

On explique ici comment procéder lorsqu'on cherche à diagonaliser une matrice. On traite l'exemple explicite de la matrice

$$A = \begin{pmatrix} 3 & 1 & 1 \\ -8 & -3 & -4 \\ 6 & 3 & 4 \end{pmatrix}$$

On détermine d'abord le spectre de A.

Pour cela, on cherche pour quelles valeurs de λ la matrice $A - \lambda I_3$ est non-inversible. On a

$$A - 3I_3 = \begin{pmatrix} 3-\lambda & 1 & 1 \\ -8 & -3-\lambda & -4 \\ 6 & 3 & 4-\lambda \end{pmatrix}.$$

On utilise le coefficient 1 en haut à droite comme pivot. On effectue $L2 \leftarrow L2 + 4L1$ et $L3 \leftarrow L3 - (4-\lambda)L1$ qui fournit

$$\begin{pmatrix} 3-\lambda & 1 & 1 \\ 4-4\lambda & 1-\lambda & 0 \\ -6+7\lambda-\lambda^2 & -1+\lambda & 0 \end{pmatrix}.$$

On utilise le coefficient central comme pivot pour effectuer $L3 \leftarrow L3 + L2$ qui fournit :

$$\begin{pmatrix} 3-\lambda & 1 & 1 \\ 4-4\lambda & 1-\lambda & 0 \\ -2+3\lambda-\lambda^2 & 0 & 0 \end{pmatrix}.$$

Ainsi, la matrice $A - \lambda I_3$ est inversible si et seulement si $1 - \lambda \neq 0$ et $-2 + 3\lambda - \lambda^2 = -(\lambda - 1)(\lambda - 2) \neq 0$. C'est-à-dire qu'elle est non-inversible ssi $\lambda = 1$ ou $\lambda = 2$. On conclut que le spectre est $\mathrm{Sp}(A) = \{2, 1\}$.

On calcule ensuite les sous espaces propres E_2, E_1 et on en détermine une base.

Pour $E_2 = \mathrm{Ker}(A-2I_3) = \mathrm{Ker}\begin{pmatrix} 1 & 1 & 1 \\ -8 & -5 & -4 \\ 6 & 3 & 2 \end{pmatrix} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : (A-2I_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$,
la troisième ligne est redondante puisque $L_3 = -L_2 - 2L_1$. On obtient le système

$$\left\{ \begin{array}{rrrl} x & +y & +z & =0 \\ -8x & -5y & -4z & =0 \end{array} \right. \Leftrightarrow_{L_2 \to L_2+8L_1} \left\{ \begin{array}{rrrl} x & +y & +z & =0 \\ & 3y & +4z & =0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{rl} x & =\frac{1}{3}z \\ y & =\frac{-4}{3}z \end{array} \right.$$

Ainsi

$$E_2 = \left\{ \begin{pmatrix} \frac{1}{3}z \\ \frac{-4}{3}z \\ z \end{pmatrix} : z \in \mathbb{R} \right\} = \left\{ z \begin{pmatrix} \frac{1}{3} \\ \frac{-4}{3} \\ 1 \end{pmatrix} : z \in \mathbb{R} \right\} = \mathrm{vect} \begin{pmatrix} \frac{1}{3} \\ \frac{-4}{3} \\ 1 \end{pmatrix} = \mathrm{vect} \begin{pmatrix} 1 \\ -4 \\ 3 \end{pmatrix}$$

On pose $v_1 = \begin{pmatrix} 1 \\ -4 \\ 3 \end{pmatrix}$. Ce vecteur forme une base de E_2.

Pour $E_1 = \mathrm{Ker}(A-I_3) = \mathrm{Ker}\begin{pmatrix} 2 & 1 & 1 \\ -8 & -4 & -4 \\ 6 & 3 & 3 \end{pmatrix} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : (A-I_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$,
les trois lignes sont équivalentes, donc on obtient un système

$$2x + y + z = 0 \Leftrightarrow z = -2x - y.$$

Ainsi

$$E_1 = \left\{ \begin{pmatrix} x \\ y \\ -2x-y \end{pmatrix} : x, y \in \mathbb{R} \right\} = \left\{ x \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} : x, y \in \mathbb{R} \right\} = \mathrm{vect}\left(\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right)$$

On pose $v_2 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$ et $v_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$. La famille (v_2, v_3) est une base de E_1.

Le Lemme 1.2.7 nous assure que $\mathcal{B}' = (v_1, v_2, v_3)$ est une base de $\mathbb{R}^3$. De plus, on a $\varphi_A(v_1) = Av_1 = 2v_1$ puisque $v_1 \in E_2 = \mathrm{Ker}(A - 2I_3)$ (qui signifie $(A - 2I_3)v_1 = 0$) et de même $\varphi_A(v_2) = v_2$ et $\varphi_A(v_3) = v_3$ puisque $v2, v_3 \in E_1$. On a donc

$$\mathrm{Mat}_{(v_1,v_2,v_3)}\varphi_A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = P^{-1}AP$$

où P est la matrice de passage dont la $i^{\text{ième}}$ colonne est formée des coordonnées de v_i dans l'ancienne base canonique $\mathcal{B}_{\text{can}} = (e_1, e_2, e_3)$. En d'autre termes, la $i^{\text{ième}}$ colonne est simplement le vecteur colonnes v_i. On a

$$P = \begin{pmatrix} 1 & 1 & 0 \\ -4 & 0 & 1 \\ 3 & 2 & -1 \end{pmatrix}.$$

On déduit de cela que $A = P \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} P^{-1}$.

Cette formule est très utile si on souhaite calculer les puissances de A. En effet, on a

$$A^n = (PDP^{-1})^n = (PDP^{-1})(PDP^{-1})\dots(PDP^{-1}) = PD^nP^{-1},$$

et $D^n = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^n = \begin{pmatrix} 2^n & 0 & 0 \\ 0 & 1^n & 0 \\ 0 & 0 & 1^n \end{pmatrix} = \begin{pmatrix} 2^n & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Il n'y a plus qu'à calculer P^{-1} et multiplier 3 matrices pour obtenir une expression de A^n en fonction de n.

Exercice 9. Calculer explicitement A^{10}.

1.3 Trigonalisation

1.3.1 Définition

Définition 1.3.1. *Soit $\varphi \in \mathcal{L}(E)$. On dit que φ est* trigonalisable *si il existe une base $\mathcal{B}'$ de E telle que la matrice $\mathrm{Mat}_{\mathcal{B}'}(\varphi)$ soit triangulaire supérieure.*

Comme pour la diagonalisation, on dit qu'une matrice $A \in \mathrm{M}_n(\mathbb{K})$ est trigonalisable si son endomorphisme associé $\varphi_A : \mathbb{K}^n \to \mathbb{K}^n$ donné par $\varphi_A(v) = Av$ est diagonalisable. C'est équivalent à la

Définition 1.3.2. *Soit $A \in \mathrm{M}_n(\mathbb{K})$. On dit que A est* trigonalisable *si il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $P^{-1}AP$ soit triangulaire supérieure.*

Comme d'habitude, P est la matrice de passage dont les colonnes sont les coordonnées des vecteurs de la base $\mathcal{B}'$ de $\mathbb{K}^n$ trigonalisant l'endomorphisme φ_A, coordonnées exprimées dans la base canonique de $\mathbb{K}^n$, et on a $P^{-1}AP = \mathrm{Mat}_{\mathcal{B}'}(\varphi_A)$.

Un résultat essentiel du cours de ce semestre assure que les **tous les endomorphismes d'un $\mathbb{C}$-espace vectoriel de dimension fini sont trigonalisables.** Ce résultat sera prouvé plus tard au cours du semestre.

1.3.2 Un peu de dynamique

On va présenter sur un exemple explicite une méthode générale de trigonalisation. Pour mieux la comprendre, on note d'abord quelques résultats de dynamique.

On rappelle que si $\psi \in \mathcal{L}(E)$, alors la puissance $s^{\text{ième}}$ notée ψ^s désigne la composition s fois de ψ, i.e. $\psi^n = \psi \circ \dots \circ \psi$ avec s facteurs. Si A est la matrice de ψ dans une base donnée, alors A^s (puissance au sens des multiplications de matrices) est la matrice de ψ^s.

Le lemme suivant est tellement élémentaire qu'on le qualifie de simple observation. Il est toutefois très utile.

Observation 1.3.3. *Soit $\psi \in \mathcal{L}(E)$ et $v \in E$. Alors $\forall s \geq 1$,*

$$v \in \ker(\psi^s) \quad \Leftrightarrow \quad \psi(v) \in \ker(\psi^{s-1}).$$

Soit $B \in \mathrm{M}_n(\mathbb{K})$ et $v \in \mathbb{K}^n$. Alors $\forall s \geq 1$,

$$v \in \ker(B^s) \quad \Leftrightarrow \quad Bv \in \ker(B^{s-1}).$$

Démonstration. Si $v \in \ker(\psi^s)$, alors $\psi^s(v) = 0$, donc $\psi^{s-1}(\psi(v)) = 0$, donc $\psi(v) \in \ker(\psi^{s-1})$. Et si $v \notin \ker(\psi^s)$, alors $\psi^s(v) \neq 0$, donc $\psi^{s-1}(\psi(v)) \neq 0$, donc $\psi(v) \notin \ker(\psi^{s-1})$.

L'énoncé matriciel est équivalent. □

On utilisera aussi le

Lemme 1.3.4. *[Lemme dynamique] Soit E un $\mathbb{K}$-espace vectoriel de dimension finie. Soit $\psi \in \mathcal{L}(E)$, alors*

(a) $\forall s \in \mathbb{N}, \mathrm{Ker}(\psi^s) \subset \mathrm{Ker}(\psi^{s+1})$,

(b) il existe $c \in \mathbb{N}$ tel que $\mathrm{Ker}(\psi^c) = \mathrm{Ker}(\psi^{c+\ell})$ pour tout $\ell > 0$.

Soit $B \in \mathrm{M}_n(\mathbb{K})$, alors

(a) $\forall s \in \mathbb{N}, \mathrm{Ker}(B^s) \subset \mathrm{Ker}(B^{s+1})$,

(b) il existe $c \in \mathbb{N}$ tel que $\mathrm{Ker}(B^c) = \mathrm{Ker}(B^{c+\ell})$ pour tout $\ell > 0$.

Démonstration. Soit $x \in \ker(\psi^s)$, alors $\psi^{s+1}(x) = \psi\left(\psi^s(x)\right) = \psi(0) = 0$, donc $x \in \ker(\psi^{s+1})$. Cela montre (a).

Pour montrer (b), on montre d'abord que si $\mathrm{Ker}(\psi^c) = \mathrm{Ker}(\psi^{c+1})$, alors $\mathrm{Ker}(\psi^{c+1}) = \mathrm{Ker}(\psi^{c+2})$. Par récurrence, cela assurera $\mathrm{Ker}(\psi^c) = \mathrm{Ker}(\psi^{c+\ell})$ pour tout $\ell > 0$. Le (a) montre que $\mathrm{Ker}(\psi^{c+1}) \subset \mathrm{Ker}(\psi^{c+2})$. Soit $x \in \mathrm{Ker}(\psi^{c+2})$, alors $\psi^{c+1}(\psi(x)) = 0$, donc $\psi(x) \in \mathrm{Ker}(\psi^{c+1}) = \mathrm{Ker}(\psi^c)$, et donc $0 = \psi^c(\psi(x)) = \psi^{c+1}(x)$, donc $x \in \mathrm{Ker}(\psi^{c+1})$ qui montre l'inclusion réciproque.

Reste à justifier l'existence de c tel que $\mathrm{Ker}(\psi^c) = \mathrm{Ker}(\psi^{c+1})$. Mais sinon, la suite $(\ker(\psi^s))_{s\in\mathbb{N}}$ serait une suite strictement croissante de sous-espaces vectoriels de E, donc on aurait $\dim(\ker(\psi^s)) \geq s$ pour tout $s \in \mathbb{N}$. Cela contredirait la dimension finie de E.

Là encore, l'énoncé matriciel est équivalent. □

1.3.3 Un exemple explicite

On considère la matrice suivante dans $\mathrm{M}_4(\mathbb{R}) \subset \mathrm{M}_4(\mathbb{C})$.

$$A = \begin{pmatrix} -8 & -3 & -3 & 1 \\ 6 & 3 & 2 & -1 \\ 26 & 7 & 10 & -2 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

On se propose de trigonaliser A dans $\mathrm{M}_n(\mathbb{C})$.

On cherche d'abord le spectre, c'est-à-dire les valeurs λ pour lesquelles $A - \lambda I_4$ est non inversible. On a

$$A - \lambda I_4 = \begin{pmatrix} -8-\lambda & -3 & -3 & 1 \\ 6 & 3-\lambda & 2 & -1 \\ 26 & 7 & 10-\lambda & -2 \\ 0 & 0 & 0 & 2-\lambda \end{pmatrix}$$

On effectue les opérations (pivot sur l'entrée m_{13}) suivantes $L2 \leftarrow 3L2 + 2L1$ et $L3 \leftarrow 3L3 + (10-\lambda)L1$. On obtient

$$\begin{pmatrix} -8-\lambda & -3 & -3 & 1 \\ 2(1-\lambda) & 3(3-\lambda) & 0 & * \\ -2-2\lambda+\lambda^2 & -9+3\lambda & 0 & * \\ 0 & 0 & 0 & 2-\lambda \end{pmatrix}$$

où les $*$ désignent des coefficients que l'on n'a pas calculés (et dont nous n'aurons pas besoin). On effectue encore l'opération $l3 \leftarrow 3(1-\lambda)L3 - (-9+3\lambda)L2$ et on obtient :

$$\begin{pmatrix} -8-\lambda & -3 & -3 & 1 \\ 2(1-\lambda) & 3(3-\lambda) & 0 & * \\ 3(1-\lambda)(2-\lambda)^2 & 0 & 0 & * \\ 0 & 0 & 0 & 2-\lambda \end{pmatrix}$$

Cette matrice est de rang 4 si et seulement si $\lambda \notin \{1, 2\}$. Ainsi $\mathrm{Sp}(A) = \{1, 2\}$.

Pour $\lambda = 1$, on détermine (au moyen du pivot de Gauss)

$$E_1^1 = \ker(A - I_4) = \ker \begin{pmatrix} -9 & -3 & -3 & 1 \\ 6 & 2 & 2 & -1 \\ 26 & 7 & 9 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \mathrm{vect} \begin{pmatrix} -2 \\ 1 \\ 5 \\ 0 \end{pmatrix}. \quad \text{On pose } v_1 = \begin{pmatrix} -2 \\ 1 \\ 5 \\ 0 \end{pmatrix}.$$

Pour $\lambda = 2$, on détermine le sous espace propre :

$$E_2^1 = \ker(A - 2I_4) = \ker \begin{pmatrix} -10 & -3 & -3 & 1 \\ 6 & 1 & 2 & -1 \\ 26 & 7 & 8 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \mathrm{vect} \begin{pmatrix} 3 \\ -2 \\ -8 \\ 0 \end{pmatrix}. \quad \text{On pose } v_2 = \begin{pmatrix} 3 \\ -2 \\ -8 \\ 0 \end{pmatrix}.$$

On obtient deux sous-espaces propres de dimension 1. On en déduit que la matrice A n'est pas diagonalisable puisque le somme des dimensions des sous-espaces propres est $1 + 1 = 2 < 4$.

À ce stade, on sait que si on complète la famille (v_1, v_2) en une base $\mathcal{B}' = (v_1, v_2, v_3, v_4)$, on aura :

$$\mathrm{Mat}_{\mathcal{B}'}(\varphi_A) = \begin{pmatrix} 1 & 0 & * & * \\ 0 & 2 & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{pmatrix} \text{ puisque } Av_1 = v_1 \text{ et } Av_2 = 2v_2.$$

Une "astuce" consiste à chercher v_3 dans le sous-espace :

$$E_2^2 = \ker((A-2I_4)^2) = \ker \begin{pmatrix} 4 & 6 & 0 & -1 \\ -2 & -3 & 0 & 1 \\ -10 & -15 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \text{vect} \left(\begin{pmatrix} 3 \\ -2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right)$$

(On laisse en exercice le calcul du carré de la matrice $A_\lambda I_4$ ainsi que la résolution du système.) On choisit alors v_3 de sorte que (v_2, v_3) soit une base de E_2^2 (qui contient E_2^1 d'après le Lemme dynamique 1.3.4 (a)). Ici, on pourrait prendre l'un ou l'autre des deux vecteurs de la base de E_2^2 obtenue puisqu'ils ne sont pas colinéaires à v_2. En général, il faut faire attention.

Arbitrairement, on choisit de poser $v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$. On vérifie que la famille (v_1, v_2, v_3) est libre.

On a $v_3 \in E_2^2 = \text{Ker}((A-2I_4)^2)$, donc $(A-2I_4)v_3 \in \text{Ker}(A-2I_4) = E_2^1$ d'après le Lemme dynamique 1.3.4 (b) appliqué à $B = A - 2I_4$. Ainsi $Av_3 = 2v_3 + w$ avec $w \in E_2^1$. Donc $Av_3 = 2v_3 + tv_2$ pour un certain $t \in \mathbb{R}$ à determiner. En remplacant A, v_3 et v_2 par leurs valeurs, on trouve $t = -1$. On en déduit que quelque soit v_4 (choisi tel que $\mathcal{B}'$ soit une base)

$$\text{Mat}_{\mathcal{B}'}(\varphi_A) = \begin{pmatrix} 1 & 0 & 0 & r_1 \\ 0 & 2 & -1 & r_2 \\ 0 & 0 & 2 & r_3 \\ 0 & 0 & 0 & r_4 \end{pmatrix} \text{ puisque } Av_3 = -v_2 + 2v_3.$$

À ce stade, on a trigonalisé la matrice A. Reste quand même à trouver un v_4 explicite. Par exemple, on pose $v_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$. Et pour trouver r_1, r_2, r_3, r_4, on résoud le système $Av_4 = r_1v_1 + r_2v_2 + r_3v_3 + r_4v_4$. On obtient :

$$\text{Mat}_{\mathcal{B}'}(\varphi_A) = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 2 & -1 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \text{ puisque } Av_4 = v_1 + v_2 + v_3 + 2v_4.$$

En fait, l'astuce peut être améliorée en cherchant w_4 (pour remplacer v_4) dans le sous-espace :

$$E_2^3 = \ker((A-2I_4)^3) = \ker \begin{pmatrix} -4 & -6 & 0 & -2 \\ 2 & 3 & 0 & 1 \\ 10 & 15 & 0 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \text{vect} \left(\begin{pmatrix} 3 \\ -2 \\ -8 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ -2 \end{pmatrix} \right)$$

On choisit alors w_4 de sorte que (v_2, v_3, w_4) soit une base de E_2^3 (qui contient E_2^2). Attention, on ne peut évidemment pas choisir l'un des deux premiers vecteurs de la base puisque ce sont déjà v_2 et v_3. On choisit de poser $w_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -2 \end{pmatrix}$, et on s'assure que w_4 n'est pas combinaison linéaire de v_1, v_2 et v_3.

Alors $\mathcal{B}'' = (v_1, v_2, v_3, w_4)$ est une base de $\mathbb{C}^4$ et si

$$P = (v_1, v_2, v_3, w_4) = \begin{pmatrix} -2 & 3 & 0 & 1 \\ 1 & -2 & 0 & 0 \\ 5 & -8 & 1 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

alors la matrice de l'endomorphisme canoniquement associé à A dans la base $\mathcal{B}''$ est de la forme

$$\mathrm{Mat}_{\mathcal{B}''}(\varphi_A) = T = P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & -1 & * \\ 0 & 0 & 2 & * \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

où les $*$ sont des valeurs inconnues à ce stade.

En effet, on a $v_4 \in E_2^3 = \mathrm{Ker}((A-2I_4)^3)$, donc $(A-2I_4)v_4 \in \mathrm{Ker}((A-2I_4)^2) = E_2^2$. Ainsi $Av_4 = 2v_4 + w$ avec $w \in E_2^2$. Donc $Av_4 = 2v_4 + s_2v_2 + s_{3s}v_3$ pour $s_2, s_3 \in \mathbb{C}$ à determiner. Pour déterminer les $*$ (c'est-à-dire les s_i), on explicite le système

$$Aw_4 = \begin{pmatrix} -10 \\ 8 \\ 30 \\ -4 \end{pmatrix} = s_2 \begin{pmatrix} 3 \\ -2 \\ -8 \\ 0 \end{pmatrix} + s_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ 0 \\ 0 \\ -2 \end{pmatrix}$$

qui donne $s_2 = -4$ et $s_3 = -2$.

On conclut que

$$\mathrm{Mat}_{\mathcal{B}''}(\varphi_A) = T = P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & -1 & -4 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

On se souvient que l'on a fait des choix lorsqu'on a défini v_3 et v_4. D'autres choix auraient donné d'autres valeurs aux entrées $*$.

1.3.4 Forme de Jordan

Définition 1.3.5. *Une matrice $M = (m_{ij}) \in \mathrm{M}_n(\mathbb{K})$ est appelée matrice de Jordan si on a $\forall 1 \leq i, j \leq n$,*

- *les entrées diagonales m_{ii} sont des valeurs propres,*
- $j - i \notin \{0, 1\}$ *implique* $m_{ij} = 0$*,*
- *les entrées surdiagonales $m_{i,i+1}$ prennent valeur 0 ou 1, et de plus si $m_{ii} \neq m_{i+1,i+1}$, alors* $m_{i,i+1} = 0$*.*

Le Théorème de Jordan (à voir plus tard) assure que si φ est un endomorphisme d'un espace vectoriel **complexe**, alors cet espace E admet une base dans laquelle la matrice de φ est de Jordan.

On se propose de construire une telle base pour l'exemple précédent. Notre objectif est de trouver une base $\mathcal{B}'''$ de $\mathbb{C}^4$ encore meilleure que $\mathcal{B}'' = (v_1, v_2, v_3, w_4)$.

On se propose de trouver une base $\mathcal{B}_J = (u_2, u_3, u_4)$ de $E_2^3 = \mathrm{Ker}((A - 2I_4)^3)$ qui satisfasse les relations suivantes :

$Au_4 = u_3 + 2u_4$ et $Au_3 = u_2 + 2u_3$ et bien sûr $Au_2 = 2u_2$ puisque $u_2 \in E_2^1 = \ker(A - 2I_4)$.

Alors par construction, dans la base $\mathcal{B}'' = (v_1, u_2, u_3, u_4)$, la matrice sera "de Jordan", c'est-à-dire que si $P' = (v_1, u_2, u_3, u_4)$ est la matrice de passage (dont les colonnes sont les coordonnées de vecteurs dans la base canonique, c'est-à-dire simplement les vecteurs colonnes v_1, u_2, u_3, u_4), alors

$$P'^{-1} A P' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

Et c'est très facile. Il suffit de prendre $u_4 = v_4 \in E_2^3 \setminus E_2^2$. On a alors

$$Au_4 - 2u_4 = (A - 2I_4)u_4 \in E_2^2 \setminus E_2^1 = \ker\left((A - 2I_4)^2\right) \setminus \ker(A - 2I_4)$$

par application de l'observation 1.3.3 avec $\psi = A - 2I_4$. On pose $u_3 = Au_4 - 2u_4$ qui donne la première relation.

On recommence en posant $u_2 = Au_3 - 2u_3$ qui donne la deuxième relation. On a bien $u_2 \in E_2^1 \setminus \{0\}$ grâce à l'observation qui assure que $u_2 = (A - 2I_4)u_3 \in E_2^1 \setminus E_2^0$ puisque $u_3 \in E_2^2 \setminus E_2^1$.

Si l'on veut expliciter P', il suffit de calculer par opération matricielles avec

$$u_4 = w_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -2 \end{pmatrix} \quad u_3 = (A - 2I_4)u_4 = \begin{pmatrix} -12 \\ 8 \\ 30 \\ 0 \end{pmatrix} \quad u_2 = (A - 2I_4)u_3 = \begin{pmatrix} 6 \\ -4 \\ 64 \\ 0 \end{pmatrix}$$

et donc

$$P' = (v_1, u_2, u_3, u_4) = \begin{pmatrix} -2 & 6 & -12 & 1 \\ 1 & -4 & 8 & 0 \\ 5 & 64 & 30 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

Chapitre 2

Le groupe symétrique $\mathcal{S}_n$

NB : les sections 2.1, 2.4 et 2.5 ne sont pas au programme de l'examen de l'année universitaire 2019-2020.

Définition 2.0.1. *Soit E un ensemble. On appelle groupe symétrique de E l'ensemble des bijections de E dans E muni de la loi de composition $\circ$. On note ce groupe $\mathcal{S}_E$ et on appelle ses éléments des permutations de E.*

Il s'agit bien d'un groupe car

(a) $\forall f, g, h \in \mathcal{S}_E, (f \circ g) \circ h = f \circ (g \circ h)$, associativité

(b) l'application $\mathrm{id}_E : E \to E$ donnée par $\forall x \in E, \mathrm{id}_E(x) = x$ est bien élément neutre : $\forall f \in \mathcal{S}_E, f \circ \mathrm{id}_E = \mathrm{id}_E \circ f = f$,

(c) la bijection réciproque f^{-1} de f fournit bien un inverse $f^{-1} \circ f = f \circ f^{-1} = \mathrm{id}_E$.

2.1 Canonicité et théorème de Cayley

Définition 2.1.1. *Soient $(G_1, \cdot)$ et $(G_2, \cdot)$ deux goupes. Une application $\varphi : G_1 \to G_2$ est un morphisme de groupes si les trois conditions suivantes sont satisfaites*

(a) $\varphi(e_1) = e_2$, où e_i désigne l'élément neutre de G_i,

(b) $\forall x, y \in G_1$, $\varphi(x \cdot y) = \varphi(x) \cdot \varphi(y)$,

(c) $\forall x \in G_1, \varphi(x^{-1}) = \varphi(x)^{-1}$.

L'application φ est appelée isomorphisme de groupes si elle est de plus bijective (et alors la réciproque est aussi un isomorphisme de G_2 vers G_1). On dit que les groupes G_1 et G_2 sont isomorphes s'il existe un isomorphisme de l'un vers l'autre. On note $G_1 \simeq G_2$.

Deux groupes isomorphes sont "identiques" du point de vue de la théorie des groupes. Tout résultat sur le premier se transmet au second immédiatement en appliquant l'isomorphisme φ et réciproquement en appliquant φ^{-1}.

Proposition 2.1.2. *$\varphi : G_1 \to G_2$ est un morphisme de groupe si et seulement si $\forall x, y \in G_1$, $\varphi(x \cdot y) = \varphi(x) \cdot \varphi(y)$.*

Démonstration. La preuve est laissée en exercice. □

Proposition 2.1.3. *Soient E, F deux ensembles. On suppose qu'il existe une bijection $f : E \to F$. Alors il existe un isomorphisme $\varphi : \mathcal{S}_E \to \mathcal{S}_F$.*

Cette proposition assure que le groupe symétrique d'un ensemble E ne dépend (à isomorphisme près) que du cardinal de cet ensemble.

En particulier, si E est de cardinal n (c'est-à-dire si E possède n éléments), alors $\mathcal{S}_E \simeq \mathcal{S}_{\{1,2,\dots,n\}}$, que l'on note aussi plus simplement $\mathcal{S}_n$. C'est un représentant canonique des groupes de permutations d'ensembles à n éléments.

Démonstration. On considère l'application $\varphi : \mathcal{S}_E \to \mathcal{S}_F$ donnée par

$$\forall \sigma \in \mathcal{S}_E, \varphi(\sigma) := f \circ \sigma \circ f^{-1}.$$

On vérifie aisément que cette application est bien à valeurs dans $\mathcal{S}_F$, car l'application $\varphi(\sigma)$ va de F dans F, et est bijective comme composée de trois bijections (faire un diagramme).

De plus, φ est un morphisme. En effet, $\forall \sigma_1, \sigma_2 \in \mathcal{S}_E$,

$$\varphi(\sigma_1 \circ \sigma_2) = f \circ \sigma_1 \circ \sigma_2 \circ f^{-1} = f \circ \sigma_1 \circ f^{-1} \circ f \circ \sigma_2 \circ f^{-1} = \varphi(\sigma_1) \circ \varphi(\sigma_2).$$

Et $\forall \sigma \in \mathcal{S}_E$,

$$\varphi(\sigma)^{-1} = \left(f \circ \sigma \circ f^{-1}\right)^{-1} = (f^{-1})^{-1} \circ \sigma^{-1} \circ f^{-1} = \varphi(\sigma^{-1}).$$

On rappelle que $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$ pour toutes applications f et g bijectives telles que le domaine de définition de f est égal à l'image de g.

Enfin, φ est bijective. En effet, l'application $\psi : \mathcal{S}_F \to \mathcal{S}_F$ donnée par $\forall \tau \in \mathcal{S}_F, \psi(\tau) := f^{-1} \circ \tau \circ f$ est l'inverse de φ, car $\forall \sigma \in \mathcal{S}_E, \psi \circ \varphi(\sigma) = f^{-1} \circ f \circ \sigma \circ f^{-1} \circ f = \sigma$ et de même $\forall \tau \in \mathcal{S}_F, \varphi \circ \psi(\tau) = \tau$. □

Les groupes symétriques fournissent des exemples de groupes très interessants et en un certain sens universels, comme le montre le théorème suivant.

Définition 2.1.4. *Soit G un groupe. Une partie H non-vide de G est appelée sous-groupe de G si les deux conditions suivantes sont satisfaites*

(a) $\forall x, y \in H, xy \in H$,

(b) $\forall x \in H, x^{-1} \in H$.

La conjonction des deux conditions est en fait équivalente à la suivante :
(c) $\forall x, y \in H, xy^{-1} \in H$ (exercice).

Exercice 10. Soit $\varphi : G_1 \to G_2$ un morphisme de groupes. Montrer que $\mathrm{Im}(\varphi)$ est un sous-groupe de G_2.

Théorème 2.1.5 (Théorème de Cayley)**.** *Soit G un groupe, alors G est isomorphe à un sous-groupe de $\mathcal{S}_G$.*

Démonstration. On donne la preuve sous forme d'exercice. Pour tout g dans G, on définit

$$\begin{array}{rccc} \lambda_g : & G & \to & G \\ & x & \mapsto & gx \end{array}$$

Il s'agit de la multiplication à gauche par g.

(a) Montrer que $\lambda_g \in \mathcal{S}_G$.
On définit :

$$\begin{array}{rccc} \lambda : & G & \to & \mathcal{S}_G \\ & g & \mapsto & \lambda_g \end{array}$$

(b) Montrer que λ est un morphisme de groupes.

(c) Montrer que λ est injectif.

(d) Démontrer le théorème.

□

On observe que le théorème de Cayley est vrai même lorsque G est infini.

2.2 Généralités sur $\mathcal{S}_n$

On considère maintenant le groupe $\mathcal{S}_n = \mathcal{S}_{\{1,2,\ldots,n\}}$. On notera souvent un élément $\sigma \in \mathcal{S}_n$ sous la forme

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \ldots & \sigma(n) \end{pmatrix} \tag{2.1}$$

Par exemple, la permutation

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 3 & 5 & 2 \end{pmatrix} \in \mathcal{S}_5$$

désigne la bijection de $\{1,2,3,4,5\}$ donnée par $\sigma_1(1) = 4, \sigma_1(2) = 1, \sigma_1(3) = 3, \sigma_1(4) = 5, \sigma_1(5) = 2$. Cette notation est pratique pour calculer la composition. Par exemple si

$$\sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 2 & 3 & 4 \end{pmatrix} \in \mathcal{S}_5$$

alors

$$\sigma_1 \circ \sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 3 & 5 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 3 & 5 \end{pmatrix}$$

On fera bien **attention** à appliquer d'abord la permutation de droite puis celle de gauche car $\sigma_1 \circ \sigma_2(1) = \sigma_1(\sigma_2(1)) = \sigma_1(5) = 2$. On peut aussi s'en souvenir grâce aux flêches $1 \overset{\sigma_2}{\mapsto} 5 \overset{\sigma_1}{\mapsto} 2$.

Proposition 2.2.1. *Soit $n \geq 1$,*

(a) le cardinal du groupe $\mathcal{S}_n$ est $n!$

(b) si $n \geq 3$, alors $\mathcal{S}_n$ n'est pas commutatif.

Exercice 11. Dresser la liste des éléments de $\mathcal{S}_2$ puis de $\mathcal{S}_3$.

Démonstration. Pour décrire une permutation sous la forme (2.1), comme σ doit être bijective, il y a n possibilités pour le choix de $\sigma(1)$, puis il reste $n-1$ possibilités pour le choix de $\sigma(2)$, puis il reste $n-2$ possibilités pour le choix de $\sigma(3)$, etc. À la fin, il ne reste qu'une possibilité pour $\sigma(n)$. Au total, on a $n \times (n-1) \times \cdots \times 1 = n!$ permutations dans $\mathcal{S}_n$. Cela prouve (a).

Pour le (b), on suppose $n \geq 3$, on pose

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 2 & 1 & 3 & 4 & \dots & n \end{pmatrix} \text{ et } \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 3 & 1 & 2 & 4 & \dots & n \end{pmatrix}$$

alors

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 3 & 2 & 1 & 4 & \dots & n \end{pmatrix} \neq \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 1 & 3 & 2 & 4 & \dots & n \end{pmatrix} = \tau\sigma.$$

□

Exercice 12. Montrer que $\mathcal{S}_2$ est commutatif.

2.3 Cycles et décomposition en produit de cycles

2.3.1 Support d'une permutation

Définition 2.3.1. *Soit $\sigma \in \mathcal{S}_n$ une permutation. L'ensemble*

$$\textit{supp}(\sigma) = \{1 \leq i \leq n | \sigma(i) \neq i\}$$

est appelé support de la permutation σ. Un élément $1 \leq i \leq n$ tel que $\sigma(i) = i$ est appelé point fixe de la permutation σ.

Clairement, l'ensemble des points fixes de σ est le complémentaire de $\mathrm{supp}(\sigma)$ dans l'ensemble $\{1, 2, \dots, n\}$.

Proposition 2.3.2. *Deux permutations à supports disjoints commutent.*

Démonstration. Soient $\sigma, \tau \in \mathcal{S}_n$ telles que $\mathrm{supp}(\sigma) \cap \mathrm{supp}(\tau) = \emptyset$, et soit $1 \leq i \leq n$. Il s'agit de montrer que $\sigma\tau(i) = \tau\sigma(i)$. On distingue trois cas.

Premier cas : si $i \in \mathrm{supp}(\sigma)$. Alors $i \notin \mathrm{supp}(\tau)$, donc $\tau(i) = i$ et a fortiori $\sigma\tau(i) = \sigma(i)$. D'autre part, on observe que $\sigma(i) \in \mathrm{supp}(\sigma)$. En effet, sinon on aurait $\sigma(\sigma(i)) = \sigma(i)$ et en appliquant σ^{-1} il viendrait $\sigma(i) = i$ qui contredirait l'hypothèse. On déduit que $\sigma(i) \notin \mathrm{supp}(\tau)$, donc que $\tau\sigma(i) = \sigma(i)$ aussi.

Deuxième cas : si $i \in \mathrm{supp}(\tau)$, on procède de la même manière.

Troisième cas : si $i \notin \mathrm{supp}(\sigma) \cup \mathrm{supp}(\tau)$. Alors on a $\sigma(i) = i$ et $\tau(i) = i$. Donc $\sigma\tau(i) = \tau\sigma(i) = i$. □

2.3.2 Décomposition en produit de cycles

Définition 2.3.3. *Un cycle de longueur $2 \leq \ell \leq n$ est une permutation $\sigma \in \mathcal{S}_n$ telle que il existe une partie $\{i_1, \ldots, i_\ell\} \subset \{1, \ldots, n\}$ de cardinal ℓ avec*

(a) $\forall 1 \leq s \leq \ell - 1, \sigma(i_s) = i_{s+1}$ *et* $\sigma(i_\ell) = i_1$.

(b) et si $i \notin \{i_1, \ldots, i_\ell\}$, *alors* $\sigma(i) = i$.

On note souvent un tel cycle $\sigma = (i_1 i_2 \ldots i_\ell)$. Un cycle de longueur 2 est appelé transposition.

La première condition peut s'unifier en disant que $\forall 1 \leq s \leq \ell, \sigma(i_s) = i_{s+1}$ où les indices sont considérés modulo ℓ.

On remarque que l'écriture d'un cycle n'est pas unique, par exemple

$$(123) = (231) = (312) \neq (132) = (321) = (213).$$

Théorème 2.3.4. *[Décomposition en produit de cycles à supports disjoints] Soit $\sigma \in \mathcal{S}_n$. Alors il existe $k \geq 0$ et $c_1, \ldots, c_k \in \mathcal{S}_n$ des cycles à supports deux à deux disjoints tels que*

$$\sigma = c_1 \ldots c_k.$$

De plus, cette décomposition est unique à l'ordre des facteurs près.

On illustre d'abord le théorème par un exemple. On considère

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 7 & 3 & 2 & 9 & 4 & 6 & 8 & 1 \end{pmatrix} \in \mathcal{S}_9.$$

On observe que $1 \mapsto 5 \mapsto 9 \mapsto 1$ et que $2 \mapsto 7 \mapsto 6 \mapsto 4 \mapsto 2$. Enfin $3 \mapsto 3$ et $8 \mapsto 8$. On vérifie donc immédiatement que

$$\sigma = (159)(2764)$$

Pour traiter le cas général, on introduit d'abord la notion d'orbite.

Définition 2.3.5. *Soit $\sigma \in \mathcal{S}_n$ et $i \in \{1, \ldots, n\}$. L'orbite du point i sous la permutation σ est*

$$\mathcal{O}_\sigma(i) = \{\sigma^k(i) | k \in \mathbb{Z}\}.$$

Si i est un point fixe de σ, alors $\mathcal{O}_\sigma(i) = \{i\}$ est un singleton.

Lemme 2.3.6. *Soit $d := \min\{k \geq 1 | \sigma^k(i) = i\}$, alors*

$$\mathcal{O}_\sigma(i) = \{i, \sigma(i), \sigma^2(i), \ldots, \sigma^{d-1}(i)\}.$$

On observe que l'orbite a cardinal d par minimalité.

Démonstration. On montre d'abord que d est bien défini, en exhibant un $k \geq 1$ tel que $\sigma^k(i) = i$. Pour cela, on observe que les termes de la suite $(\sigma^k(i))_{k\geq 1}$ ne peuvent pas être deux à deux distincts car l'ensemble $\{1, \dots, n\}$ est fini. Donc il existe $k_1, k_2 \geq 1$ tels que $\sigma^{k_1}(i) = \sigma^{k_2}(i)$. (En fait, $k_i \leq n + 1$.) Par symétrie, on peut supposer que $k_1 \geq k_2$. En composant l'égalité par $\sigma^{-k_2} := (\sigma^{-1})^{k_2}$, on obtient $\sigma^{k_1-k_2}(i) = i$.

Reste à démontrer l'égalité par double inclusion. Comme $\supset$ est évidente, il ne reste que l'inclusion réciproque. Soit donc $k \in \mathbb{Z}$. On effectue la division euclidienne de k par d. On obtient $k = md + r$ où $0 \leq r < d$. Mais alors

$$\sigma^k(i) = \sigma^{r+md}(i) = \sigma^r\sigma^d \dots \sigma^d(i) = \sigma^r(i),$$

où il y a m fois σ^d et où $\sigma^d(i) = i$. □

Lemme 2.3.7. *Soit $i, j \in \{1, \dots, n\}$ et $\sigma \in \mathcal{S}_n$. Alors*

- *soit $\mathcal{O}_\sigma(i) = \mathcal{O}_\sigma(j)$,*
- *soit $\mathcal{O}_\sigma(i) \cap \mathcal{O}_\sigma(j) = \emptyset$.*

En d'autres termes, soit les orbites de i et j sont égales, soit elles sont disjointes.

Démonstration. On suppose que $\mathcal{O}_\sigma(i) \cap \mathcal{O}_\sigma(j) \neq \emptyset$. Alors il existe $k_1, k_2 \in \mathbb{Z}$ tels que $\sigma^{k_1}(i) = \sigma^{k_2}(j)$. En appliquant σ^{-k_2}, on obtient $\sigma^{k_1-k_2}(i) = j$. Mais alors $\forall k \in \mathbb{Z}$, on a $\sigma^k(j) = \sigma^{k+k_1-k_2}(i) \in \mathcal{O}_\sigma(i)$, d'où $\mathcal{O}_\sigma(j) \subset \mathcal{O}_\sigma(i)$. L'inclusion réciproque est aussi vraie par symétrie. □

Définition 2.3.8. *Soit X un ensemble. Une famille $(A_i)_{i\in I}$ de parties de X forme une partition de X si les deux conditions suivantes sont satisfaites :*

(a) $X = \cup_{i\in I} A_i$,

(b) $\forall i, j \in I$, *si* $i \neq j$, *alors* $A_i \cap A_j = \emptyset$.

Démonstration du Théorème 2.3.4. On considère toutes les orbites de σ. Dans chaque orbite, on choisit un point privilégié (par exemple, le plus petit pour l'ordre naturel sur $\{1, \dots, n\}$). On peut donc écrire

$$\{1, \dots, n\} = \text{Fix}(\sigma) \sqcup \bigsqcup_{j=1}^{k} \mathcal{O}_\sigma(i_j),$$

où chaque $\mathcal{O}_\sigma(i_j)$ est une orbite de taille $d_j \geq 2$ et $\text{Fix}(\sigma)$ est l'ensemble des points fixes de σ. Le Lemme 2.3.7 assure qu'il s'agit d'une union disjointe, donc d'une partition.

On définit alors pour tout $1 \leq j \leq k$ le cycle $c_j := (i_j \ \sigma(i_j) \ \sigma^2(i_j) \ \dots \ \sigma^{d_j-1}(i_j))$.

Il est alors clair que $\forall s \in \{1, \dots, n\}$, $\sigma(s) = c_1 \dots c_k(s)$. En effet, soit s est un point fixe de σ et de tous les c_j auquel cas le résultat tient, soit s appartient à exactement une orbite $\mathcal{O}_\sigma(i_{j_0})$, et alors $s = \sigma^k(i_{j_0})$ pour un certain k, mais on a $\sigma(s) = \sigma^{k+1}(i_{j_0}) = c_{j_0}(s)$ et tant s que $\sigma(s) = c_{j_0}(s)$ sont fixes pour c_j avec $j \neq j_0$.

L'unicité est laissée à la sagacité des lecteurs.trices. □

Définition 2.3.9. *Soit $\sigma \in \mathcal{S}_n$. Le graphe de Schreier de σ est le graphe orienté dont l'ensemble de sommets est $\{1, \dots, n\}$ et où l'on met une arrête orientée de i vers j si $\sigma(i) = j$.*

Le fait que σ soit une bijection correspond au fait que chaque sommet est source d'une unique arrête et cible d'une unique arrête (et ces deux arrêtes sont la même si et seulement si le sommet correspond à un point fixe de σ).

La décomposition en cycles de σ correspond à la décomposition en cycles du graphe de Schreier de σ.

Exercice 13. Dessiner le graphe de Schreier de

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 7 & 3 & 2 & 9 & 4 & 6 & 8 & 1 \end{pmatrix} \in \mathcal{S}_9.$$

2.4 Ordre d'une permutation

Définition 2.4.1. *Soit G un groupe et g un élément de G. L'ordre de g est :*

$$\textit{ordre}(g) := \min\{k \geq 1 | g^k = \mathrm{id}\}.$$

Dans le cas où l'ensemble est vide, l'ordre de g est infini.

Fait 2.4.2. *Soit G un groupe fini, alors tout élément de G est d'ordre fini inférieur au cardinal de G.*

Démonstration. La suite $(g^k)_{k\geq 1}$ est à valeurs dans l'ensemble G fini. Donc il existe $1 \leq k_1 < k_2 \leq |G|$ tels que $g^{k_1} = g^{k_2}$. Donc $g^{k_2-k_1} = \mathrm{id}$. □

Fait 2.4.3. *Soit $c \in \mathcal{S}_n$ un cycle. Alors l'ordre de c est la longueur de c.*

Démonstration. On note $c = (i_0\, i_1\, i_2 \dots i_{\ell-1})$ un cycle de longueur ℓ. On vérifie immédiatemment que $c^k = (i_0\, i_k\, i_{2k} \dots) \neq \mathrm{id}$ pour tout $1 \leq k \leq \ell - 1$, tandis que $c^\ell = \mathrm{id}$. Notez que les indices sont pris modulo ℓ. □

Exercice 14. Soit G un groupe et $g \in G$. Alors $g^r = \mathrm{id}$ si et seulement si l'ordre de g divise r.

On déduit du Théorème 2.3.4 de décomposition en cycles à supports disjoints.

Corollaire 2.4.4. *Soit $\sigma \in \mathcal{S}_n$. On note $\sigma = c_1 \dots c_k$ sa décomposition en cycles à supports disjoints. Alors*

$$\textit{ordre}(\sigma) = \textit{ppcm}\left(\textit{longueur}(c_1), \dots, \textit{longueur}(c_k)\right).$$

Démonstration. D'après la Proposition 2.3.2, des permutations à supports disjoints commutent, donc $\forall r \geq 1, \sigma^r = c_1^r \dots c_k^r$ qui est encore une décomposition à cycles disjoints, donc $\sigma^r = \mathrm{id}$ si et seulement si $c_j^r = \mathrm{id}$ pour tout $1 \leq j \leq k$. Or d'après l'exercice 14, on a $c_j^r = \mathrm{id}$ si et seulement si la longueur de c_j divise r, c'est-à-dire si r est un multiple de longueur(c_j). □

2.5 Conjugaison

Définition 2.5.1. *Soit G un groupe et soient $\sigma_1, \sigma_2 \in G$. On dit que σ_1 est conjugué à σ_2 si il existe $\sigma \in G$ tel que*

$$\sigma_1 = \sigma\sigma_2\sigma^{-1}.$$

Exercice 15. Montrer que la relation "est conjugé à" est une relation d'équivalence parmi les éléments de G.

On observe par simple calcul dans $\mathcal{S}_n$ (faites-le !) que $\sigma(12\dots\ell)\sigma^{-1} = (\sigma(1)\ \ \sigma(2)\dots\sigma(\ell))$. Plus généralement, on dispose du

Lemme 2.5.2. *Deux cycles de $\mathcal{S}_n$ sont conjugués si et seulement si ils ont même longueur.*

Démonstration. La première implication découle de l'observation précédant le lemme, puisque $\sigma(i_1 i_2 \dots i_\ell)\sigma^{-1} = (\sigma(i_1)\ \ \sigma(i_2)\dots\sigma(i_\ell))$.

Réciproquement, si $(i_1 i_2 \dots i_\ell)$ et $(j_1 j_2 \dots j_\ell)$ sont deux cycles de longueur $\ell \geq 2$, alors n'importe quelle permutation σ telle que $\forall 1 \leq k \leq \ell, \sigma(i_k) = j_k$ satisfait :

$$\sigma(i_1 i_2 \dots i_\ell)\sigma^{-1} = (j_1 j_2 \dots j_\ell).$$

Notez que toute application satisfaisant ces conditions peut se prolonger en une bijection de $\{1, \dots, n\}$ vers lui-même, puisque il reste $n - \ell$ élément à la source dont on n'a pas encore imposé l'image, et $n - \ell$ éléments dans l'ensemble d'arrivée dont on n'a pas encore imposé les antécédants. □

En fait, la preuve montre même plus généralement le

Théorème 2.5.3. *Soient $\sigma_1, \sigma_2 \in \mathcal{S}_n$. Pour $\ell \geq 1$, on note $m_{i,\ell}$ le nombre de cycles de longueur ℓ dans la décomposition en cycles à supports disjoints de σ_i.*

Alors σ_1 et σ_2 sont conjuguées si et seulement si $\forall \ell \geq 1, m_{1,\ell} = m_{2,\ell}$.

Démonstration. Notons $s_1 = c_{11} \dots c_{1k}$ la décomposition en cycles à supports disjoints de σ_1. La première implication découle du calcul

$$\sigma\sigma_1\sigma^{-1} = \sigma c_{11} \dots c_{1k}\sigma^{-1} = \sigma c_{11}\sigma^{-1}\sigma c_{12}\sigma^{-1} \dots \sigma c_{1k}\sigma^{-1}$$

du lemme précédent qui assure que $\sigma c_{1i}\sigma^{-1}$ est un cycle de même longueur que c_{1i}.

Réciproquement, notons

$$\sigma_i = \prod_{\ell=1}^{n}\prod_{j=1}^{m_{1,\ell}} c_{i\ell j} = \prod_{\ell=1}^{n}\prod_{j=1}^{m_{1,\ell}}(t_{i\ell j1}\ \ t_{i\ell j2}\ \ \dots\ \ t_{i\ell j\ell})$$

la décompositon en produit de cycles à supports disjoints de σ_i (on rappelle que $m_{2,\ell} = m_{1,\ell}$ pour tout ℓ), où les cycles $c_{i\ell j} = (t_{i\ell j1} t_{i\ell j2} \dots t_{i\ell j\ell})$ ont longueur ℓ.

Notez que le premier indice indique si l'on considère σ_1 ou σ_2, le second qu'on considère un cycle de longueur ℓ, le troisième qu'on considère le $j^{\text{ième}}$ tel cycle, et le quatrième la position

de l'entier désigné dans ce cycle. Notez aussi qu'on se permet d'utiliser la notation $\prod$ pour le produit car dans cette situation les facteurs commutent (supports disjoints) ; en général, il faudrait préciser l'ordre des facteurs.

Alors comme dans le lemme précédent, n'importe quelle permutation σ telle que $\forall \ell, j, s, \sigma(t_{1\ell js}) = t_{2\ell js}$ satisfait $\sigma\sigma_1\sigma^{-1} = \sigma_2$, et il existe bien de tels permutations. □

Exercice 16. Trouver $\sigma \in \mathcal{S}_8$ explicite telle que

$$\sigma(172)(35)(48)\sigma^{-1} = (12)(34)(567).$$

Combien y a-t-il de tels σ ?

2.6 Transpositions et morphisme de signature

2.6.1 Parties génératrices

Définition 2.6.1. *Soit G un groupe et S une partie de G. On dit que S est une partie génératrice de G, ou encore que G est engendré par S, si pour tout élément g de G, il existe $s_1, \ldots, s_k \in S$ tels que $g = s_1 s_2 \ldots s_k$.*

Exemples 2.6.2.
- Le groupe G est une partie génératrice de G.
- La partie $\{-1, 1\}$ est une partie génératrice du groupe $(\mathbb{Z}, +)$ (exercice).
- Les cycles de $\mathcal{S}_n$ engendrent $\mathcal{S}_n$. Cela découle du Théorème 2.3.4 de décomposition en cycles à supports disjoints.

Il est souvent intéressant de chercher des parties génératrices de petite taille d'un groupe.

Théorème 2.6.3. *Les transpositions engendrent $\mathcal{S}_n$.*

Démonstration. Comme toute permutation de $\mathcal{S}_n$ est un produit de cycles, il suffit de montrer que tout cycle est un produit de transpositions. Soit $c = (12 \ldots \ell)$ un cycle de longueur $\ell \geq 2$. On montre par récurrence sur ℓ que c est un produit de $\ell - 1$ permutations.

C'est trivial si $\ell = 2$. En général, on pose $\tau = (\ell - 1 \ \ \ell) = \tau^{-1}$. On a $c\tau = (12 \ldots \ell - 1)$ qui est produit de $\ell - 2$ transpositions par hypothèse de récurrence, c'est-à-dire $c\tau = \tau_1 \ldots \tau_{\ell-2}$ et donc $c = \tau_1 \ldots \tau_{\ell-2}\tau$. □

La preuve montre en fait que

$$(12 \ldots \ell) = (12)(23)(34) \ldots (\ell - 2 \ \ \ell - 1)(\ell - 1 \ \ \ell).$$

Par ailleurs, elle est valide pour n'importe quel cycle $(i_1 i_2 \ldots i_\ell)$ de longueur ℓ et pas seulement $(12 \ldots \ell)$. Il suffit de "rajouter des i" !

2.6.2 Signature

Définition 2.6.4. *Soit $\sigma \in \mathcal{S}_n$. On note $r = r(\sigma)$ le nombre d'orbites de σ, c'est-à-dire que r est le nombre de cycles dans la décomposition en cycles à supports disjoints plus le nombre de points fixes de σ. On définit la signature de σ comme le nombre*

$$\varepsilon(\sigma) := (-1)^{n-r}.$$

Le nombre $n - r(\sigma)$ est parfois appelé le décrément de la permutation σ.

Par exemple, $\varepsilon(\mathrm{id}) = (-1)^{n-n} = 1$.

Exercice 17. Soit c un ℓ-cycle. Montrer que $\varepsilon(c) = (-1)^{\ell-1}$.

En particulier, les transpositions ont signature (-1). Plus généralement, on dispose du

Lemme 2.6.5. *Soit $\sigma \in \mathcal{S}_n$ et τ une transposition, alors*

$$\varepsilon(\sigma\tau) = -\varepsilon(\sigma).$$

Démonstration. Notons $\tau = (ij)$. Grâce au Lemme 2.3.7, on distingue deux cas :

Premier cas : si $\mathcal{O}_\sigma(i) \cap \mathcal{O}_\sigma(j) = \emptyset$. Quite à conjuguer (c'est-à-dire essentiellement "renuméroter"), on peut supposer qu'on est dans la situation suivante : $\mathcal{O}_\sigma(i) = \{1, 2, \ldots, \ell\}$ et $\sigma|_{\mathcal{O}_\sigma(i)} = (12\ldots\ell)$, $\mathcal{O}_\sigma(j) = \{\ell+1, \ell+2, \ldots, \ell+\ell'\}$ et $\sigma|_{\mathcal{O}_\sigma(j)} = (\ell+1\ \ \ell+2\ldots\ell+\ell')$ avec $1 \leq i \leq \ell$ et $\ell+1 \leq j \leq \ell+\ell'$.

Alors on vérifie par calcul que

$$\sigma\tau|_{\{1,\ldots,\ell+\ell'\}} = (12\ldots i\ \ j+1\ \ j+2\ldots\ell+\ell'\ \ \ell+1\ldots j\ \ i+1\ldots\ell)$$

et que $\sigma\tau|_{\{1,\ldots,n\}\setminus\{1,\ldots,\ell+\ell'\}} = \sigma|_{\{1,\ldots,n\}\setminus\{1,\ldots,\ell+\ell'\}}$. On constate que les deux cycles de σ à supports disjoints contenant i et j fusionnent. Donc $\sigma\tau$ contient un cycle de moins que σ. La formule est valide.

Second cas : si i et j sont dans la même orbite de σ. On peut supposer $1 \leq i < j \leq \ell$ et que $(12\ldots\ell)$ est un cycle de σ. Par le calcul, on obtient qu'en restriction à $\{1, \ldots, \ell\}$, on a

$$(12\ldots\ell)(ij) = (12\ldots i\ \ j+1\ \ j+2\ldots\ell)(i+1\ \ i+2\ldots j),$$

et qu'en restriction au complémentaire, σ et $\sigma\tau$ coïncident. On a donc dédoublé un des cycles de σ. La formule est valide. □

Exercice 18. Illustrer la preuve au moyen des graphes de Schreier de σ, τ et $\sigma\tau$.

Théorème 2.6.6. *Si une permutation σ dans $\mathcal{S}_n$ s'écrit comme produit de k transpositions, alors $\varepsilon(\sigma) = (-1)^k$.*

Ce théorème assure que la signature de σ correspond à la parité du nombre de transpositions dans **toute** écriture de σ comme produit de transpositions. La signature est -1 (resp. 1) si et seulement si cette parité est impair (resp. pair).

Démonstration. Le Théorème 2.6.6 découle du Lemme 2.6.5 par récurrence immédiate sur k. □

Corollaire 2.6.7. *L'application signature $\varepsilon : \mathcal{S}_n \to \{\pm 1\}$ est un morphisme du groupe $(\mathcal{S}_n, \circ)$ vers le groupe $(\{\pm 1\}, \times)$.*

Démonstration. Soient $\sigma, \sigma' \in \mathcal{S}_n$. On écrit $\sigma = \tau_1 \dots \tau_k$ et $\sigma' = \tau'_1 \dots \tau'_{k'}$ comme produits de transpositions. Alors

$$\varepsilon(\sigma\sigma') = \varepsilon(\tau_1 \dots \tau_k \tau'_1 \dots \tau'_{k'}) = (-1)^{k+k'} = (-1)^k (-1)^{k'} = \varepsilon(\tau_1 \dots \tau_k)\varepsilon(\tau'_1 \dots \tau'_{k'}) = \varepsilon(\sigma)\varepsilon(\sigma').$$

D'autre part, $\sigma^{-1} = \tau_k^{-1} \dots \tau_1^{-1}$ est aussi produit de k transpositions, d'où $\varepsilon(\sigma^{-1}) = (-1)^k = \varepsilon(\sigma) = \varepsilon(\sigma)^{-1}$. (On observe que pour tout $a \in \{\pm 1\}$, $a^2 = 1$, donc a est son propre inverse pour la multiplication.) □

On remarquera que le groupe $(\{\pm 1\}, \times)$ est isomorphe au groupe $(\mathbb{Z}/2\mathbb{Z}, +)$ via le morphisme $1 \mapsto 0$ (ce sont les neutres des deux groupes) et $-1 \mapsto 1$.

Exercice 19. Montrer que l'application signature $\varepsilon : \mathcal{S}_2 \to \{\pm 1\}$ induit un isomorphisme de $(\mathcal{S}_2, \circ)$ sur $(\{\pm 1\}, \times)$.

Exercice 20. Soit $(G, \cdot)$ un groupe ayant exactement deux éléments. Montrer qu'il existe un isomorphisme de $(G, \cdot)$ vers $(\mathbb{Z}/2\mathbb{Z}, +)$.

On dit qu'à isomorphisme près, il existe un unique groupe ayant exactement deux éléments. L'étudiant·e intéressé·e pourra montrer qu'il y a, à isomorphisme près, un unique groupe à trois éléments, dont une représentation est $(\mathbb{Z}/3\mathbb{Z}, +)$, mais qu'il y a deux classes d'isomorphisme des groupes à quatre éléments dont des représentations sont $(\mathbb{Z}/4\mathbb{Z}, +)$ et $(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}, +)$.

2.6.3 Le groupe alterné $\mathcal{A}_n$

Définition 2.6.8. *Soit $n \geq 1$. Les éléments de l'ensemble*

$$\mathcal{A}_n := \mathrm{Ker}(\varepsilon) = \{\sigma \in \mathcal{S}_n | \varepsilon(\sigma) = 1\}$$

sont appelés des permutations alternées.

Par exemple, les cycles de longueurs impaires sont alternés d'après l'exercice 17. Le groupe $\mathcal{A}_2$ est réduit à la permutation identité (on dit que c'est le groupe trivial).

Proposition 2.6.9. *(a) $(\mathcal{A}_n, \circ)$ est un sous-groupe de $(\mathcal{S}_n, \circ)$.*

(b) Les 3-cycles engendrent $\mathcal{A}_n$.

Démonstration. Le (a) découle du fait suivant dont la preuve est un exercice :

Fait 2.6.10. *Si $\varphi : G_1 \to G_2$ est un morphisme de groupes, alors l'ensemble*

$$\mathrm{Ker}(\varphi) := \{g \in G_1 | \varphi(g) = e_2\}$$

où e_2 est l'élément neutre de G_2, est un sous-groupe de G_1.

Le (b) découle du fait que tout produit de deux transpositions peut s'écrire comme produit de 3-cycles (exercice). □

Pour $n \geq 5$, une propriété intéressante du groupe alterné est la suivante. Soit G un groupe, on suppose qu'il existe un morphisme $\varphi : \mathcal{A}_n \to G$ surjectif. Alors, soit φ est un isomorphisme, soit G est le groupe trivial ayant un seul élément. On dit que le groupe $\mathcal{A}_n$ est *simple*. On reverra cette notion en L3.

Chapitre 3

Déterminants

Dans ce chapitre, $\mathbb{K}$ désigne un corps, et E désigne un $\mathbb{K}$-espace vectoriel de dimension n.

3.1 Formes n-linéaires alternées et déterminants

3.1.1 Formes n-linéaires alternées

Définition 3.1.1. *Soit E un $\mathbb{K}$-espace vectoriel de dimension n. Une application $\varphi : E^n \to \mathbb{K}$ est n-linéaire si $\forall 1 \leq i \leq n, \forall v_1, \ldots, v_n, v_i' \in E, \forall \lambda, \mu \in \mathbb{K}$*

$$\varphi(v_1, \ldots, v_{i-1}, \lambda v_i + \mu v_i', v_{i+1}, \ldots, v_n) = \lambda\varphi(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n) + \mu\varphi(v_1, \ldots, v_{i-1}, v_i', v_{i+1}, \ldots, v_n)$$

c'est-à-dire si φ est linéaire en chacune de ses n-variables (qui sont chacunes des vecteurs dans E).

Dans ce contexte, il est d'usage courant de désigner φ comme une *forme* n-linéaire. Le mot "forme" est dans ce contexte un simple synonyme du mot "application".

Exercice 21. Soit $n = 3$ et $\varphi : E^3 \to \mathbb{K}$ une application 3-linéaire. Développer

$$\varphi(\lambda_1 v_1 + \mu_1 v_1' + \nu_1 v_1'', \lambda_2 v_2, \lambda_3 v_3 + \mu_3 v_3').$$

Définition 3.1.2. *Une forme $\varphi : E^n \to \mathbb{K}$ est antisymétrique si $\forall 1 \leq i, j \leq n, \forall v_1, \ldots, v_n \in E$, si $i \neq j$, alors*

$$\varphi(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -\varphi(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n).$$

(Les vecteurs en positions différente de i ou j sont les mêmes dans les deux membres de l'égalité.)

Proposition 3.1.3. *Soit $\varphi : E^n \to \mathbb{K}$ une forme, alors les deux assertions suivantes sont équivalentes :*

(a) φ est antisymétrique,

(b) $\forall v_1, \dots, v_n \in E, \forall \sigma \in \mathcal{S}_n, \varphi(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = \varepsilon(\sigma)\varphi(v_1, \dots, v_n)$*, où* $\varepsilon(\sigma)$ *désigne la signature de* σ*.*

Démonstration. (b) implique (a) en prenant $\sigma = (ij)$.

Réciproquement, (a) assure que (b) est vrai pour les transpositions. Soit σ une permutation quelconque, on peut écrire $\sigma = \tau_1 \dots \tau_k$ comme un produit de k transpositions. On a alors pour chaque $1 \leq s \leq k$

$$\varphi(v_{\tau_1 \dots \tau_{s-1}(1)}, \dots, v_{\tau_1 \dots \tau_{s-1}(n)}) = -\varphi(v_{\tau_1 \dots \tau_s(1)}, \dots, v_{\tau_1 \dots \tau_s(n)}).$$

Il suit par récurrence que

$$\varphi(v_1, \dots, v_n) = (-1)^k \varphi(v_{\sigma(1)}, \dots, v_{\sigma(n)}).$$

Le résultat vient car $\varepsilon(\sigma) = (-1)^k$. □

Définition 3.1.4. *Une forme* $\varphi : E^n \to \mathbb{K}$ *est alternée si* $\forall i \neq j, \forall v_1, \dots, v_n \in E$*,*

$$\varphi(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_n) = 0,$$

c'est-à-dire si $v_i = v_j$ *implique* $\varphi(v_1, \dots, v_n) = 0$*.*

Exercice 22. Soit $\varphi : E^n \to \mathbb{K}$ une forme n-linéaire. Montrer que φ est antisymétrique si et seulement si φ est alternée.

Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E, et soit $(v_1, \dots, v_n) \in E^n$. Pour tout $1 \leq j \leq n$, on note $(a_{ij})_{i=1}^n$ les coordonnées de v_j dans la base $\mathcal{B}$, soit

$$v_j = \sum_{i=1}^{n} a_{ij} e_i = a_{1j} e_1 + \dots + a_{nj} e_n = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix},$$

où l'on note les coordonnées en colonne.

Théorème 3.1.5. *Soit* $\varphi : E^n \to \mathbb{K}$ *une forme* n*-linéaire et alternée, alors* $\forall v_1, \dots, v_n \in E$

$$\varphi(v_1, \dots, v_n) = \left(\sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n} \right) \varphi(e_1, \dots, e_n).$$

Démonstration. En utilisant successivement la linéarité en chacune des variables, on calcule

$$\begin{aligned} \varphi(v_1, \dots, v_n) &= \varphi\left(\sum_{i_1=1}^{n} a_{i_1 1} e_{i_1}, \dots, \sum_{i_n=1}^{n} a_{i_n n} e_{i_n} \right) \\ &= \sum_{1 \leq i_1, \dots, i_n \leq n} a_{i_1 1} \dots a_{i_n n} \varphi(e_{i_1}, \dots, e_{i_n}). \end{aligned}$$

La forme φ étant alternée, on a $\varphi(e_{i_1}, \ldots, e_{i_n}) = 0$ dés que les e_{i_s} ne sont pas deux à deux distincts. C'est-à-dire que les seuls termes non-nuls sont ceux pour lesquels il existe σ dans $\mathcal{S}_n$ tels que $\forall 1 \leq s \leq n, i_s = \sigma(s)$. Ainsi

$$\begin{aligned}\varphi(v_1, \ldots, v_n) &= \sum_{\sigma \in \mathcal{S}_n} a_{\sigma(1)1} \ldots a_{\sigma(n)n} \varphi(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) \\ &= \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \ldots a_{\sigma(n)n} \varphi(e_1, \ldots, e_n)\end{aligned}$$

en utilisant le (b) de la Proposition 3.1.3, puisque d'après l'exercice 22, φ est aussi antisymétrique. □

Exercice 23. Faire le calcul complètement détaillé pour $n = 2$ et pour $n = 3$.

3.1.2 Déterminant d'une famille de n vecteurs.

Définition 3.1.6. Déterminant d'une famille de n vecteurs. *Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base du $\mathbb{K}$-espace vectoriel E. Le déterminant du n-uplet de vecteurs $(v_1, \ldots, v_n) \in E^n$ dans la base $\mathcal{B}$ est*

$$det_{\mathcal{B}}(v_1, \ldots, v_n) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(s) a_{\sigma(1)1} \ldots a_{\sigma(n)n}$$

où les coordonnées du vecteurs v_j dans la base $\mathcal{B}$ sont $v_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$.

Le théorème précédent 3.1.5 s'énonce donc : pour toute φ forme n-linéaire alternée,

$$\varphi(v_1, \ldots, v_n) = \det_{\mathcal{B}}(v_1, \ldots, v_n) \varphi(e_1, \ldots, e_n)$$

Exercice 24. Montrer que $\det_{\mathcal{B}}(\mathcal{B}) = \det_{\mathcal{B}}(e_1, \ldots, e_n) = 1$.

Théorème 3.1.7. *L'application $det_{\mathcal{B}} : E^n \to \mathbb{K}$ est une forme n-linéaire alternée.*

Ce théorème assure en particulier que les formes n-linéaires alternées existent bel et bien !

Démonstration. On montre d'abord la linéarité en la $j^{\text{ième}}$ variable v_j.

On note $v_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$ et $v'_j = \begin{pmatrix} a'_{1j} \\ \vdots \\ a'_{nj} \end{pmatrix}$. Soient aussi $\lambda, \mu \in \mathbb{K}$. Alors

$$\begin{aligned}
\det_{\mathcal{B}}\left(v_1, \ldots, \lambda v_j + \mu v'_j, \ldots, v_n\right) &= \det_{\mathcal{B}}\left(\begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, \ldots, \begin{pmatrix} \lambda a_{1j} + \mu a'_{1j} \\ \vdots \\ \lambda a_{nj} + \mu a'_{1j} \end{pmatrix}, \ldots, \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix}\right) \\
&= \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \ldots \left(\lambda a_{\sigma(j)j} + \mu a'_{\sigma(j)j}\right) \ldots a_{\sigma(n)n} \\
&= \lambda \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \ldots a_{\sigma(j)j} \ldots a_{\sigma(n)n} + \mu \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \ldots a'_{\sigma(j)j} \ldots a_{\sigma(n)n} \\
&= \lambda \det_{\mathcal{B}}\left(v_1, \ldots, v_j, \ldots, v_n\right) + \mu \det_{\mathcal{B}}\left(v_1, \ldots, v'_j, \ldots, v_n\right).
\end{aligned}$$

Reste à montrer que le déterminant est alterné. Pour cela, il suffit de montrer qu'il est anti-symétrique. Soit $\theta \in \mathcal{S}_n$, on calcule

$$\begin{aligned}
\det_{\mathcal{B}}\left(v_{\theta(1)}, \ldots, v_{\theta(n)}\right) &= \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{\sigma(1)\theta(1)} \ldots a_{\sigma(n)\theta(n)} \\
&= \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{\sigma(\theta^{-1}(1))1} \ldots a_{\sigma(\theta^{-1}(n))n} \text{ car } \theta(i) = j \text{ équivaut à } i = \theta^{-1}(j), \\
&= \sum_{\tau \in \mathcal{S}_n} \varepsilon(\tau\theta) a_{\tau(1)1} \ldots a_{\tau(n)n} \\
&= \sum_{\tau \in \mathcal{S}_n} \varepsilon(\tau)\varepsilon(\theta) a_{\tau(1)1} \ldots a_{\tau(n)n} \\
&= \varepsilon(\theta) \sum_{\tau \in \mathcal{S}_n} \varepsilon(\tau) a_{\tau(1)1} \ldots a_{\tau(n)n} \\
&= \varepsilon(\theta) \det_{\mathcal{B}}\left(v_1, \ldots, v_n\right)
\end{aligned}$$

À la troisième ligne, on a utilisé le changement de variable $\tau = \sigma\theta^{-1}$ qui est bijectif dans $\mathcal{S}_n$ (cela équivaut à $\sigma = \tau\theta$). □

Corollaire 3.1.8. *Soit $\mathcal{B}$ une base, alors toute forme $\varphi : E^n \to \mathbb{K}$ n-linéaire alternée est colinéaire au déterminant dans la base $\mathcal{B}$, c'est-à-dire qu'il existe $\lambda \in \mathbb{K}$ tel que $\varphi = \lambda det_{\mathcal{B}}$.*

On dit qu'à normalisation près, le déterminant est l'unique forme n-linéaire alternée sur E^n.

Démonstration. Le Théorème 3.1.5 assure que $\forall (v_1, \ldots, v_n) \in E^n$,

$$\varphi(v_1, \ldots, v_n) = \det_{\mathcal{B}}(v_1, \ldots, v_n)\varphi(e_1, \ldots, e_n)$$

donc $\varphi = \lambda \det_{\mathcal{B}}$ avec $\lambda = \varphi(e_1, \ldots, e_n)$. □

Puisque pour toute autre base $\mathcal{B}'$ le déterminant $\det_{\mathcal{B}'}$ dans la base $\mathcal{B}'$ est aussi une forme n-linéaire alternée, on en déduit encore le

Corollaire 3.1.9. *Soit $\mathcal{B}, \mathcal{B}'$ deux bases de E, alors on a*

$$det_{\mathcal{B}'}(\cdot) = det_{\mathcal{B}}(\cdot) det_{\mathcal{B}'}(\mathcal{B}).$$

Comme $\det_{\mathcal{B}'}(\mathcal{B}') = 1$, on déduit que $\det_{\mathcal{B}'}(\mathcal{B}) \neq 0$ pour toutes bases $\mathcal{B}, \mathcal{B}'$.

On déduit du corollaire 3.1.9 le résultat primordial suivant :

Théorème 3.1.10. *[Théorème fondamental du déterminant] Soit E un $\mathbb{K}$-espace vectoriel de dimension n, soit $\mathcal{B}$ une base de E et soit $(v_1, \ldots, v_n) \in E^n$. Les assertions suivantes sont équivalentes :*

(a) $(v_1, \ldots, v_n)$ est une base de E,

(b) $det_{\mathcal{B}}(v_1, \ldots, v_n) \neq 0$.

La seconde assertion s'interprète en disant que le volume du parallélépipède engendré est non nul.

Démonstration. Montrons d'abord que si $(v_1, \ldots, v_n)$ n'est pas une base de E alors $\det_{\mathcal{B}}(v_1, \ldots, v_n) = 0$. En effet si la famille $(v_1, \ldots, v_n)$ est liée, alors il existe j_0 et λ_j tels que $v_{j_0} = \sum_{j \neq j_0} \lambda_j v_j$, mais alors

$$\begin{aligned} \det_{\mathcal{B}}(v_1, \ldots, v_n) &= \det_{\mathcal{B}}(v_1, \ldots, v_{j_0}, \ldots, v_n) \\ &= \det_{\mathcal{B}}(v_1, \ldots, \sum_{j \neq j_0} \lambda_j v_j, \ldots, v_n) \\ &= \sum_{j \neq j_0} \lambda_j \det_{\mathcal{B}}(v_1, \ldots, v_j, \ldots, v_n) = 0 \end{aligned}$$

puisque le déterminant dans $\mathcal{B}$ est alterné, et que le vecteur v_j apparait aux positions distinctes j et j_0 dans le $j^{\text{ième}}$ terme de la somme.

Réciproquement, on suppose que $(v_1, \ldots, v_n) = \mathcal{B}'$ est aussi une base de E. Alors l'exercice 24 et le corollaire 3.1.9 assurent que

$$1 = \det_{\mathcal{B}'}(\mathcal{B}') = \det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'}(\mathcal{B})$$

En particulier $\det_{\mathcal{B}}(\mathcal{B}') = \det_{\mathcal{B}}(v_1, \ldots, v_n) \neq 0$. □

3.1.3 Déterminant d'un endomorphisme.

Définition 3.1.11. Déterminant d'un endomorphisme. *Soit $f : E \to E$ un endomorphisme du $\mathbb{K}$-espace vectoriel E de dimension n, et soit $\mathcal{B} = (v_1, \ldots, v_n)$ une base de E. Alors le déterminant de l'endomorphisme f est le déterminant dans la base $\mathcal{B}$ de la famille de vecteurs $f(\mathcal{B}) = (f(v_1), \ldots, f(v_n))$, c'est-à-dire*

$$\det(f) = det_{\mathcal{B}}(f(\mathcal{B})) = det_{\mathcal{B}}\left(f(v_1), \ldots, f(v_n)\right).$$

Cette définition est valide, c'est-à-dire que le déterminant de f est bien défini et indépendant du choix de $\mathcal{B}$ grâce à la :

Proposition 3.1.12. *Soit $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E, alors*

$$det_{\mathcal{B}}(f(\mathcal{B})) = det_{\mathcal{B}'}(f(\mathcal{B}')).$$

Démonstration. Comme f est linéaire, l'application

$$\begin{array}{rccc} \det_{\mathcal{B}} f(\cdot) : & E^n & \to & \mathbb{K} \\ & (u_1, \ldots, u_n) & \mapsto & \det_{\mathcal{B}}(f(u_1), \ldots, f(u_n)) \end{array}$$

est n-linéaire et alternée (vérifiez-le !). Le corollaire 3.1.8 assure alors qu'il existe $\lambda \in \mathbb{K}$ tel que $\det_{\mathcal{B}} f(\cdot) = \lambda \det_{\mathcal{B}}(\cdot)$. L'évaluation sur la base $\mathcal{B} = (v_1, \ldots, v_n)$ et l'exercice 24 donnent $\det_{\mathcal{B}}(f(\mathcal{B})) = \lambda \det_{\mathcal{B}}(\mathcal{B}) = \lambda$. Donc

$$\det_{\mathcal{B}} f(\cdot) = \det_{\mathcal{B}}(f(\mathcal{B})) \det_{\mathcal{B}}(\cdot) \tag{3.1}$$

D'autre part, le corollaire 3.1.9 assure que $\det_{\mathcal{B}} f(\cdot) = \det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'} f(\cdot)$. En évaluant en $\mathcal{B}'$, on obtient

$$\det_{\mathcal{B}} f(\mathcal{B}') = \det_{\mathcal{B}}(f(\mathcal{B})) \det_{\mathcal{B}}(\mathcal{B}') = \det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'} f(\mathcal{B}').$$

L'égalité voulue s'en déduit sachant que $\det_{\mathcal{B}}(\mathcal{B}') \neq 0$. Ceci découle de la remarque après le Corollaire 3.1.9. □

3.1.4 Déterminant d'une matrice carrée.

On rappelle que $\mathbb{K}^n$ est un $\mathbb{K}$-espace vectoriel de dimension n. Sa base canonique $\mathcal{B}_{\text{can}} = (e_1, \ldots, e_n)$ est formée des vecteurs colonnes $e_j = \begin{pmatrix} \delta_{1j} \\ \vdots \\ \delta_{nj} \end{pmatrix}$ avec $\delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$ le symbole de Kronecker. En d'autres termes, e_i a toutes ses coordonnées nulles sauf la $i^{\text{ième}}$ qui vaut 1.

Définition 3.1.13. Déterminant d'une matrice carrée $n \times n$. *Soit $A = (a_{ij})_{1 \leq i,j \leq n}$ une matrice carrée. Alors ses colonnes $c_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$ forment une famille de n vecteurs de $\mathbb{K}^n$. On définit*

$$\det(A) := det_{\mathcal{B}_{\text{can}}}(c_1, \ldots, c_n) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \ldots a_{\sigma(n)n} = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}$$

Exercice 25. Donner des formules explicites pour $n = 1, 2, 3$, ainsi que des moyens mnémotechniques pour s'en souvenir (règle de Sarrus).

Exercice 26. Soit $T = (a_{ij})_{1 \leq i,j \leq n} \in \mathrm{M}_n(\mathbb{K})$ une matrice triangulaire supérieure, c'est-à-dire que $\forall i > j, a_{ij} = 0$. Calculer $\det(T)$. Même question si T est triangulaire inférieure.

On rappelle qu'étant donnés un endomorphisme $f : E \to E$ d'un $\mathbb{K}$-espace vectoriel et une base $\mathcal{B} = (e_1, \ldots, e_n)$ de E, la matrice de f dans $\mathcal{B}$ est

$$\mathrm{Mat}_{\mathcal{B}}(f) = (f(e_1), \ldots, f(e_n)) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = A$$

où la $j^{\text{ième}}$ colonne $f(e_j) = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$ décrit les coordonnées du vecteur $f(e_j)$ dans $\mathcal{B}$, c'est-à-dire $f(e_j) = a_{1j}e_1 + \cdots + a_{nj}e_n$.

Dans ce cas, toutes les définitions du déterminant données ci-dessus sont bien évidemment cohérentes :

$$\det(f) = \mathrm{det}_{\mathcal{B}}\left(f(e_1), \ldots, f(e_n)\right) = \det(\mathrm{Mat}_{\mathcal{B}}(f)) = \det(A)$$

et sa valeur ne dépend pas du choix de la base $\mathcal{B}$. Notez cependant que la matrice $\mathrm{Mat}_{\mathcal{B}}(f) = A$ elle-même dépend de $\mathcal{B}$.

3.1.5 Propriété fondamentale du déterminant.

On rappelle que l'endomorphisme $f : E \to E$ d'un $\mathbb{K}$-espace vectoriel E de dimension finie n est un isomorphisme si et seulement si pour toute base $\mathcal{B} = (e_1, \ldots, e_n)$, la famille $(f(e_1), \ldots, f(e_n))$ est aussi une base, et ceci si et seulement si la matrice $\mathrm{Mat}_{\mathcal{B}}(f)$ est inversible. On déduit donc du Théorème 3.1.10 et des définitions précédentes le

Corollaire 3.1.14. *Soit $f : E \to E$ un endomorphisme du $\mathbb{K}$-espace vectoriel E de dimension n, et soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E, alors les assertions suivantes sont équivalentes :*

(a) f est un isomorphisme (c'est-à-dire f est bijective),

(b) la famille $(f(e_1), \ldots, f(e_n))$ est une base de E,

(c) la matrice $\mathrm{Mat}_{\mathcal{B}}(f)$ est inversible,

(d) $\det(f) = \mathrm{det}_{\mathcal{B}}(f(\mathcal{B})) = \det(\mathrm{Mat}_{\mathcal{B}}(f)) \neq 0$.

3.2 Explication et interprétation géométrique du déterminant

3.2.1 Le cas $n = 1$

Un $\mathbb{K}$-espace vectoriel E de dimension 1 admet une base contenant un seul élément $\mathcal{B} = (e_1)$. Ainsi $E = \{\lambda e_1 | \lambda \in \mathbb{K}\}$, on a donc une bijection entre E et $\mathbb{K}$ qui à tout élément de E fait correspondre sa coordonnée λ dans la base $\mathcal{B} = (e_1)$.

Notez que les formes 1-linéaires $\varphi : E = E^1 \to \mathbb{K}$ sont simplement les formes linéaires, et elles sont toujours alternées puisqu'il n'y a pas de deuxième coordonnée avec laquelle échanger.

Le déterminant du vecteur λe_1 dans la base $\mathcal{B} = (e_1)$ est

$$\det_{\mathcal{B}}(\lambda e_1) = \lambda$$

Ainsi le déterminant d'un vecteur $v = \lambda e_1$ dans la base $\mathcal{B} = (e_1)$ est simplement la longueur "orientée" du vecteur, c'est-à-dire l'abscisse, où le vecteur "unité" est e_1 de longueur $+1$.

Maintenant une application linéaire $f : E \to \mathbb{K}$ est uniquement déterminée par l'image $f(e_1) = a_{11}e_1$ pour un unique $a_{11} \in \mathbb{K}$. En d'autres termes, la matrice de φ dans la base $\mathcal{B} = (e_1)$ est la matrice 1×1 ayant pour seule entrée a_{11}. On a bien $\forall v \in E, f(v) = a_{11}v$.

Le déterminant de l'application linéaire f calculé dans la base $\mathcal{B} = (e_1)$ est donc

$$\det(f) = \det_{\mathcal{B}}(f(\mathcal{B})) = \det_{\mathcal{B}}(f(e_1)) = \det_{\mathcal{B}}(a_{11}e_1) = a_{11}.$$

Le déterminant de f est exactement le "facteur de dilatation" des longueurs par application f, où encore le rapport d'homothétie de f, puisque les applications linéaires en dimension 1 sont des homothéties.

3.2.2 Aire des parallèlogrammes

On voudrait définir une notion d'aire. C'est-à-dire qu'à toute partie P du plan, on veut associer un nombre $\mathrm{Aire}(P) \geq 0$ qui mesure la "superficie" de P. En toute généralité, c'est difficile, et cela fera l'objet du cours de théorie de la mesure en L3. On se restreint ici à la notion d'aire des parallèogrammes. Cette restriction est naturelle, puisque les parallèlogrammes (qui incluent les carrés) serviront de "pièces de base" pour mesurer des aires plus générales.

Étant donnés deux vecteurs $v_1, v_2 \in \mathbb{R}^2$, on définit le parallèlogramme engendré comme la partie

$$P(v_1, v_2) := \{\lambda_1 v_1 + \lambda_2 v_2 | \lambda_1, \lambda_2 \in [0,1]\} = \{v \in \mathbb{R}^2 | \exists \lambda_1, \lambda_2 \in [0,1], v = \lambda_1 v_1 + \lambda_2 v_2\}.$$

Faire un dessin. On voudrait définir une bonne notion d'aire dans le plan, qui respecte notre intuition géométrique.

En observant des dessins correspondants, on réalise que l'on voudrait que l'aire satisfasse les conditions suivantes

(a) pour tout λ, on voudrait $\mathrm{Aire}(P(\lambda v_1, v_2)) = \lambda \mathrm{Aire}(P(v_1, v_2))$

(b) et de même $\mathrm{Aire}(P(v_1, \lambda v_2)) = \lambda \mathrm{Aire}(P(v_1, v_2))$

(c) on voudrait aussi $\mathrm{Aire}(P(v_1 + v_1', v_2)) = \mathrm{Aire}(P(v_1, v_2)) + \mathrm{Aire}(P(v_1', v_2))$

(d) et de même $\mathrm{Aire}(P(v_1, v_2 + v_2')) = \mathrm{Aire}(P(v_1, v_2)) + \mathrm{Aire}(P(v_1, v_2'))$

En d'autres termes, on voudrait que l'application "aire du parallèlogramme" soit linéaire en chacune des variables v_1 et v_2, c'est-à-dire qu'elle soit 2-linéaire.

D'autre part, on souhaiterait que l'aire d'un segment soit nulle, soit :

(e) $\mathrm{Aire}(P(v_1, v_1)) = 0$

qui signifie que l'application "aire du parallèlogramme" est alternée.

Ainsi l'aire du parallèlogramme engendré par v_1 et v_2 se doit d'être une forme 2-linéaire alternée. L'étude de ces formes à la section 3.1 montre qu'à constante multiplicative près, si $v_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ et $v_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$, alors le corollaire 3.1.8 assure que les seules notions d'aires satisfaisant nos conditions sont les

$$\text{Aire}(P(v_1, v_2)) = \lambda(a_{11}a_{22} - a_{12}a_{21}) = \lambda \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Et le λ est une constante de normalisation, qui nous force à choisir arbitrairement une valeur pour au moins un parallèlogramme, par exemple le carré unité. C'est l'unité de mesure d'aire qu'utilisent les physiciens, par exemple le m^2, ou bien le cm^2.

3.2.3 Volume des parallélépipèdes

De la même manière, étant donnés trois vecteurs de l'espace $\mathbb{R}^3$, une bonne notion de volume des parallélépipèdes

$$P(v_1, v_2, v_3) := \{\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 | \lambda_i \in [0, 1]\}$$

doit être une forme 3-linéaire alternée. On déduit de même que

$$\text{Vol}(P(v_1, v_2, v_3)) = \lambda \text{det}_{\mathcal{B}_{\text{can}}}(v_1, v_2, v_3).$$

Le choix d'une base n'est autre que le choix d'un parallélépipède "unité" de volume 1.

Et en dimension n quelconque on peut encore définir une notion de volume par analogie, même si bien sûr l'intuition physique disparait.

Dans le cas d'une application linéaire $f : \mathbb{R}^2 \to \mathbb{R}^2$ ou $f : \mathbb{R}^3 \to \mathbb{R}^3$, le déterminant de f n'est rien d'autre que le facteur par lequel le volume d'une partie est multiplié lorsqu'on lui applique f :

$$\forall P \subset \mathbb{R}^2, \text{Aire}(f(P)) = \det(f)\text{Aire}(P)$$
$$\forall P \subset \mathbb{R}^3, \text{Vol}(f(P)) = \det(f)\text{Vol}(P)$$

Notez que ce facteur multiplicatif est une donnée géométrique. Cette interprétation explique donc bien que le déterminant d'une application linéaire ne dépende pas de la base choisie pour le calculer explicitement.

3.3 Multiplicativité du déterminant

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, on note $\mathcal{L}(E)$ l'ensemble des applications linéaires de E dans E, aussi appelées *endomorphismes* de E.

Théorème 3.3.1. *Soient $f, g \in L(E)$, alors*

$$\det(f \circ g) = \det(f)\det(g)$$

Démonstration. Soit $\mathcal{B}$ une base de E, alors

$$\begin{aligned}\det(f \circ g) &= \det_{\mathcal{B}}(f \circ g(\mathcal{B})) \text{ par définition 3.1.11}\\ &= \det_{\mathcal{B}}(f(g(\mathcal{B})))\\ &= \det_{\mathcal{B}}(f(\mathcal{B}))\det_{\mathcal{B}}(g(\mathcal{B})) \text{ d'après (3.1)}\\ &= \det(f)\det(g).\end{aligned}$$

On rappelle que (3.1) assure $\det_{\mathcal{B}} h(\cdot) = \det_{\mathcal{B}}(h(\mathcal{B}))\det_{\mathcal{B}}(\cdot)$, puisque l'application $\det_{\mathcal{B}}(h(\cdot))$ est n-linéaire alternée, donc proportionnelle au déterminant d'après le corollaire 3.1.8, et la constante de proportionnalité s'obtient en évaluant dans la base $\mathcal{B}$. □

Corollaire 3.3.2. *Soit $f \in \mathcal{L}(E)$, alors f est inversible si et seulement si $\det(f) \neq 0$ et*

$$\det(f^{-1}) = \frac{1}{\det(f)}$$

Démonstration. Le premier point a été vu au corollaire 3.1.14, et alors $f \circ f^{-1} = \mathrm{id}$, d'où

$$1 = \det(\mathrm{id}) = \det(f \circ f^{-1}) = \det(f)\det(f^{-1})$$

□

On note $\mathrm{GL}(E)$ le groupe (pour la composition) des endomorphismes inversibles de E, c'est-à-dire des $f \in \mathcal{L}(E)$ tels que $\det(f) \neq 0$. On a

Corollaire 3.3.3. *L'application $\det : \mathrm{GL}(E) \to \mathbb{K} \setminus \{0\}$ est un morphisme surjectif du groupe $(\mathrm{GL}(E), \circ)$ vers $(\mathbb{K} \setminus \{0\}, \times)$.*

Démonstration. Il ne reste qu'à prouver la surjectivité. On se donne une base $\mathcal{B}$ de E, et pour $\lambda \in \mathbb{K} \setminus \{0\}$, on considère un endomorphisme f_λ dont la matrice dans la base $\mathcal{B}$ est

$$A_\lambda = \mathrm{Mat}_{\mathcal{B}}(f_\lambda) = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

On calcule immédiatement que $\det(A_\lambda) = \lambda$, puisque seule la permutation id contribue un produit non-nul dans la formule de définition. □

Définition 3.3.4. *On appelle groupe spécial linéaire de E et l'on note $\mathrm{SL}(E)$ l'ensemble*

$$\mathrm{SL}(E) := \{g \in \mathrm{GL}(E) | \det(g) = 1\} = \mathrm{Ker}(\det)$$

Il s'agit d'un sous-groupe de $\mathrm{GL}(E)$ puisque c'est le noyau du morphisme déterminant $\det : \mathrm{GL}(E) \to \mathbb{K} \setminus \{0\}$.

Le groupe $\mathrm{SL}(E)$ s'interprète comme le groupe des applications linéaires de E préservant le volume. C'est très explicite dans $\mathbb{R}^2$ et $\mathbb{R}^3$.

3.4 Le polynôme caractéristique

On a vu qu'une valeur propre d'un endomorphisme $\varphi \in \mathcal{L}(E)$ est un scalaire λ tel que $\varphi - \lambda \mathrm{id}_E$ soit non-inversible. C'est le cas si et seulement si $\det(\varphi - \lambda \mathrm{id}_E) = 0$. Il est donc naturel de considérer la fonction $X \mapsto \det(\varphi - X\mathrm{id}_E)$.

Définition 3.4.1. *Soit $\varphi \in \mathcal{L}(E)$, le polynôme caractéristique de φ est*

$$P_\varphi(X) = \det(\varphi - X\mathrm{id}_E).$$

Pour le calculer, on se place dans une base $\mathcal{B}$, et on a

$$P_\varphi(X) = \det(\varphi - X\mathrm{id}_E) = \det(\mathrm{Mat}_\mathcal{B}(\varphi - X\mathrm{id}_E)) = \det(\mathrm{Mat}_\mathcal{B}(\varphi) - XI_n),$$

d'après la Proposition 1.1.6. En particulier, on a $\det(\varphi_A - X\mathrm{id}_E) = \det(A - XI_n)$ pour tout $A \in \mathrm{M}_n(\mathbb{K})$, ce qui justifie la

Définition 3.4.2. *Soit $A \in \mathrm{M}_n(\mathbb{K})$, le polynôme caractéristique de A est*

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

Proposition 3.4.3. *Soit $A = (a_{ij})_{1 \leq i,j \leq n} \in \mathrm{M}_n(\mathbb{K})$, son polynôme caractéristique est un polynôme de la forme*

$$P_A(X) = (-1)^n X^n + (-1)^{n-1}\mathrm{Tr}(A)X^{n-1} + \cdots + \det(A).$$

Ses racines sont exactement les valeurs propres de A.

On rappelle que la trace d'une matrice $A = (a_{ij})$ est la somme de ses entrées diagonales

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

Démonstration. La seconde phrase découle du Fait 1.2.3. Reste à calculer le polynôme caractéristique. Le coefficient constant est obtenu pour $X = 0$, c'est donc bien $P_A(0) = \det(A)$. Pour le reste, on note $m_{ij} = a_{ij} - X\delta_{ij}$ les entrées de la matrice $A - XI_n$, et on utilise la formule

$$\det(A - XI_n) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} m_{\sigma(i)i}.$$

On observe d'abord que si $\sigma \neq \mathrm{id}_{\mathcal{S}_n}$, alors la taille du support de σ est ≥ 2, donc le nombre de points fixes de σ est au plus $n-2$. Il y a dans ce cas au plus $n-2$ facteurs du produit $Q_\sigma(X) = \prod_{i=1}^{n} m_{\sigma(i)i}$ qui contiennent un X, et donc ce produit est un polynôme de degré $\leq n-2$.

Lorsque $\sigma = \mathrm{id}_{\mathcal{S}_n}$, on obtient en développant le polynôme

$$\prod_{i=1}^{n} m_{ii} = \prod_{i=1}^{n}(a_{ii} - X) = (-1)^n X^n + (-1)^{n-1}\mathrm{Tr}(A)X^{n-1} + Q_{\mathrm{id}}(X),$$

où $Q_{\mathrm{id}}(X)$ est un polynôme de degré $\leq n-2$. Au total

$$P_A(X) = (-1)^n X^n + (-1)^{n-1}\mathrm{Tr}(A)X^{n-1} + Q_{\mathrm{id}}(X) + \sum_{\sigma \in \mathcal{S}_n \setminus \{\mathrm{id}\}} \varepsilon(\sigma) Q_\sigma(X)$$

est bien un polynôme de degré n ayant les deux coefficients de plus haut degré voulus. □

Notons que le polynôme caractéristique de $\varphi \in \mathcal{L}(E)$ ne dépend pas du choix de la base, puisqu'un changement de base revient à une conjugaison de matrice : si $A = \mathrm{Mat}_{\mathcal{B}}(\varphi)$, $C = \mathrm{Mat}_{\mathcal{B}'}(\varphi)$ et que P est la matrice de passage, alors $C = P^{-1}AP$ et donc

$$\begin{aligned}\det(C - X_I N) &= \det(P^{-1}AP - P^{-1}XI_nP) = \det(P^{-1}(A - XI_n)P) \\ &= \det(P^{-1})\det(A - XI_n)\det(P) = \frac{1}{\det(P)}\det(A - XI_n)\det(P) \\ &= \det(A - XI_n).\end{aligned}$$

Corollaire 3.4.4. *Soit $\varphi \in \mathcal{L}(E)$. Si son polynôme caractéristique $P_\varphi(X)$ est scindé à racines simples, alors φ est diagonalisable.*

Démonstration. Le polynôme caractéristique $P_\varphi(X)$ est scindé à racines simples si $P_\varphi(X) = \prod_{i=1}^{n}(\lambda_i - X)$ avec des λ_i deux à deux distincts. Comme on obtient n valeurs propres deux à deux distinctes, le corollaire 1.2.9 s'applique. □

3.5 Propriétés calculatoires et techniques de calcul du déterminant

Dans cette partie, on cherche à calculer efficacement le déterminant d'une matrice carrée.

3.5.1 Matrices de tailles 2 et 3

Dans ce cas, on peut utiliser les formules explicites :

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

C'est le produit des entrées de la diagonale descendante (aussi appelée *diagonale*) auquel on soustrait le produit des entrées de la diagonale ascendante (aussi appelée *anti-diagonale*).

Pour les matrices 3×3, on a la règle de Sarrus. (Illustration)

Attention, la règle de Sarrus ne fonctionne pas pour des matrices de taille ≥ 4.

3.5.2 Matrices diagonales et triangulaires

Soit $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \in \mathrm{M}_n(\mathbb{K})$ une matrice carrée $n \times n$.

Définition 3.5.1. *On doit que la matrice A est*

- *diagonale si* $\forall i \neq j, a_{ij} = 0$,
- *triangulaire supérieure si* $\forall i > j, a_{ij} = 0$,
- *strictement triangulaire supérieure si* $\forall i \geq j, a_{ij} = 0$,
- *triangulaire inférieure si* $\forall i < j, a_{ij} = 0$,
- *strictement triangulaire inférieure si* $\forall i \leq j, a_{ij} = 0$.

Il est clair que les matrices diagonales sont les seules matrices à la fois triangulaires supérieures et triangulaires inférieures.

Exercice 27. Soit $T \in \mathrm{M}_n(\mathbb{K})$ une matrice strictement triangulaire (supérieure ou inférieure). Montrer que $T^n = 0$.

Proposition 3.5.2. *Soit T une matrice triangulaire (supérieure ou inférieure), alors*

$$\det(T) = \prod_{i=1}^{n} a_{ii} = a_{11}a_{22}\dots a_{nn}$$

est le produit des termes diagonaux de T.

Démonstration. Il s'agit d'un exercice, on utilise la formule

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}$$

et on vérifie que tous les termes sont nuls hormis celui correspondant à $\sigma = \mathrm{id}$. □

3.5.3 Transposition

Soit $A = (a_{ij})_{1 \leq i,j \leq n} \in \mathrm{M}_n(\mathbb{K})$ une matrice carrée. On rappelle que la transposée de A est la matrice $A^t = (a_{ji})_{1 \leq i,j \leq n}$. Les entrées de A^t sont obtenues à partir de celles de A par symétrie axiale d'axe la diagonale (descendante).

$$A^t = \begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{nn} \end{pmatrix} \qquad \text{quand} \qquad A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

Il est évident que $(A^t)^t = A$ pour tout $A \in \mathrm{M}_n(\mathbb{K})$.

Proposition 3.5.3. *Soit $A \in \mathrm{M}_n(\mathbb{K})$, alors* $\det(A^t) = \det(A)$.

Démonstration. Notons $a'_{ij} = a_{ji}$ les entrées de A^t. On calcule :

$$\begin{aligned}
\det(A^t) &= \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a'_{\sigma(i)i} \\
&= \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} \\
&= \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} a_{\sigma^{-1}(j)j} \text{ par changement d'indice } j = \sigma(i) \\
&= \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma^{-1}) \prod_{j=1}^{n} a_{\sigma^{-1}(j)j} \text{ car } \varepsilon(\sigma^{-1}) = \varepsilon(\sigma)^{-1} \\
&= \sum_{\sigma^{-1}\in\mathcal{S}_n} \varepsilon(\sigma^{-1}) \prod_{j=1}^{n} a_{\sigma^{-1}(j)j} \\
&= \sum_{\theta\in\mathcal{S}_n} \varepsilon(\theta) \prod_{j=1}^{n} a_{\theta(j)j} = \det(A)
\end{aligned}$$

À l'avant dernière ligne, on a utilisé le fait que le passage à l'inverse induit une bijection de $\mathcal{S}_n$ sur lui-même. □

Soit $A = (a_{ij})_{1\leq i,j\lambda n}$ une matrice de taille $n\times n$. Pour $1 \leq i \leq n$, on note $L_i = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix}$ sa $i^{\text{ième}}$ ligne et

$$C_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$$

sa $j^{\text{ième}}$ colonne. De sorte que l'on écrit la matrice A sous les formes suivantes :

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix} = \begin{pmatrix} C_1 & \cdots & C_n \end{pmatrix}$$

Alors la transposée de A s'écrit :

$$A^t = \begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} C_1^t \\ \vdots \\ C_n^t \end{pmatrix} = \begin{pmatrix} L_1^t & \cdots & L_n^t \end{pmatrix}$$

Notez que la transposée d'une matrice rectangle $n \times m$ est une matrice $m \times n$. En particulier, la transposée d'une ligne est une colonne et vice-versa.

On a vu au début du chapitre que le déterminant $\det(A) = \det\begin{pmatrix} C_1 & \cdots & C_n \end{pmatrix}$ de la matrice A est une forme n-linéaire alternée en tant que fonction des colonnes de A.

Corollaire 3.5.4. *Le déterminant de A est aussi une forme n-linéaire alternée en tant que fonction des lignes de A, c'est-à-dire*

- $\forall 1 \leq i \leq n, \forall \lambda, \mu \in \mathbb{K}, \forall L_i' \in \mathrm{M}_{1,n}(\mathbb{K})$

$$\det \begin{pmatrix} L_1 \\ \vdots \\ \lambda L_i + \mu L_i' \\ \vdots \\ L_n \end{pmatrix} = \lambda \det \begin{pmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_n \end{pmatrix} + \mu \det \begin{pmatrix} L_1 \\ \vdots \\ L_i' \\ \vdots \\ L_n \end{pmatrix}$$

- *et si $L_i = L_j$ pour $i \neq j$, alors* $\det(A) = 0$.

Démonstration. Le déterminant $\det(A) = \det(A^t)$ est une forme n-linéaire alternée en les colonnes de A^t, qui sont les lignes de A. □

3.5.4 Matrices élémentaires et opérations sur les lignes et les colonnes

On rappelle la définition des matrices élémentaires de $\mathrm{M}_n(\mathbb{K})$. Pour $1 \leq i, j \leq n$, la matrice élémentaire E_{ij} est la matrice de $\mathrm{M}_n(\mathbb{K})$ dont toutes les entrées sont nulles, sauf l'entrée à l'intersection de la ligne i et de la colonne j. En d'autres termes

$$E_{ij} = (e_{st})_{1 \leq s,t \leq n} \text{ où } e_{st} = \begin{cases} 1 & \text{si } (s,t) = (i,j) \\ 0 & \text{si } (s,t) \neq (i,j) \end{cases}$$

Les multiplications par les matrices élémentaires et leurs multiples s'interprètent en termes de lignes et de colonnes comme suit. Les notations sont définies ci-dessus.

Proposition 3.5.5. *Soit $A \in \mathrm{M}_n(\mathbb{K})$ et $E_{ij} \in \mathrm{M}_n(\mathbb{K})$ une matrice élémentaire. Alors*

(a) $E_{ij}A = \begin{pmatrix} 0 \\ \vdots \\ L_j \\ \vdots \\ 0 \end{pmatrix} \leftarrow i^{\text{ième}}$ *ligne*

Toutes les lignes sont nulles sauf la $i^{\text{ième}}$ où l'on trouve la $j^{\text{ième}}$ ligne L_j de A.

(b) $AE_{ij} = \begin{pmatrix} 0 & \cdots & \underset{j^{\text{ième}}\text{colonne}}{C_i} & \cdots & 0 \end{pmatrix}$

Toutes les colonnes sont nulles sauf la $j^{\text{ième}}$ où l'on trouve la $i^{\text{ième}}$ colonne C_i de A.

Démonstration. Il suffit d'écrire les produits sur une grande feuille de papier. □

On note $I_n \in \mathrm{M}_n(\mathbb{K})$ la matrice identité, dont les entrées sont les symboles de Kronecker $\delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$.

Corollaire 3.5.6. *Soient $\lambda \in \mathbb{K}$, $A \in \mathrm{M}_n(\mathbb{K})$ et E_{ij} une matrice élémentaire, alors*

(a) $(I_n + \lambda E_{ij})\, A = \begin{pmatrix} L_1 \\ \vdots \\ L_i + \lambda L_j \\ \vdots \\ L_n \end{pmatrix} \leftarrow i^{\text{ième}}\textit{ligne}$

(b) $A\,(I_n + \lambda E_{ij}) = \begin{pmatrix} C_1 & \cdots & \underset{j^{\text{ième}}\textit{colonne}}{C_j + \lambda C_i} & \cdots & C_n \end{pmatrix}$

Dans le (a), on dit que l'on a ajouté λ fois la ligne L_j à la $i^{\text{ième}}$ ligne L_i. Dans le (b), on dit que l'on a ajouté λ fois la colonne C_i à la $j^{\text{ième}}$ colonne C_j. Notez que lorsque $i = j$, cela revient à multiplier la ligne ou colonne correspondante par $(1 + \lambda)$.

Démonstration. Il suffit d'ajouter $I_n A = A$ aux résultats de la proposition 3.5.5 □

3.5.5 Matrices de permutations

Soit $\sigma \in \mathcal{S}_n$, on lui associe la matrice $P_\sigma \in \mathrm{M}_n(\mathbb{K})$ dont les entrées sont données par

$$\forall 1 \leq i, j \leq n, p_{ij} = \delta_{i\sigma(j)} = \begin{cases} 1 & \text{si } i = \sigma(j) \\ 0 & \text{si } i \neq \sigma(j) \end{cases}$$

On rappelle que δ_{st} est le symbole de Kronecker, vallant 1 si $s = t$ et 0 sinon.

Les matrices P_σ pour σ dans $\mathcal{S}_n$ sont appelées les matrices de permutations. En d'autres termes, ce sont les matrices ayant toutes leurs entrées nulles sauf n qui vallent 1 de sorte que chaque ligne et chaque colonne contienne exactement un 1.

Exemple 3.5.7. On a $P_{\mathrm{id}} = I_n$. Pour $n = 2$ puis $n = 3$, on a

$$P_{(12)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad P_{(123)} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Dans la proposition suivante, on utilise les notations habituelles pour les colonnes et les lignes de A.

Proposition 3.5.8. *Soient $\sigma, \tau \in \mathcal{S}_n$ et soit $A \in \mathrm{M}_n(\mathbb{K})$. On a*

(a) $P_\sigma^t = P_{\sigma^{-1}} = P_\sigma^{-1}$

(b) $P_\sigma P_\tau = P_{\sigma \circ \tau}$

(c) $\det(P_\sigma) = \varepsilon(\sigma)$

(d) $AP_\sigma = \begin{pmatrix} C_1 & \cdots & C_n \end{pmatrix} P_\sigma = \begin{pmatrix} C_{\sigma(1)} & \cdots & C_{\sigma(n)} \end{pmatrix}$. *On dit que la multiplication à droite par P_σ permute les colonnes de A.*

(e) $P_\sigma A = P_\sigma \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix} = \begin{pmatrix} L_{\sigma^{-1}(1)} \\ \vdots \\ L_{\sigma^{-1}(n)} \end{pmatrix}$. *On dit que la multiplication à gauche par P_σ permute les lignes de A.*

Les points (a) et (b) assurent que l'application $\sigma \mapsto P_\sigma$ est un morphisme du groupe $\mathcal{S}_n$ vers le groupe $\mathrm{GL}_n(\mathbb{K})$. On vérifie aisément que ce morphisme est injectif, de sorte que son image est un sous-groupe de $\mathrm{GL}_n(\mathbb{K})$ isomorphe au groupe $\mathcal{S}_n$.

Démonstration. Elle est laissée en exercice. On pourra se convaincre du résultat en traitant le cas des matrices 3×3. Pour une preuve formelle, on pourra simplement se ramener à la définition du produit matriciel, $AB = C$ signifiant que

$$\forall 1 \leq i, j \leq n, c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj},$$

et au calcul avec le symbole de Kronecker. □

3.5.6 La méthode du pivot de Gauss

On donne ici une méthode numérique efficace, basée sur le célèbre pivot de Gauss. Cette méthode peut s'implémenter pour donner un algorithme explicite. On se contente ici de rappeler la stratégie.

Étant donnée une matrice $A \in \mathrm{M}_n(\mathbb{K})$, la méthode du pivot de Gauss consiste à modifier cette matrice au moyen d'opérations élémentaires sur les lignes afin d'obtenir une matrice triangulaire. Cela permet par exemple de déterminer le rang d'une matrice, et avec plus de soin son inverse si elle est carrée de rang maximal.

Les opérations élémentaires sont les suivantes

- permuter deux lignes i et j, ce qui revient à multiplier à gauche par la matrice de permutation $P_{(ij)}$,
- multiplier une ligne i par un scalaire $\lambda \neq 0$, ce qui revient à multiplier à gauche par la matrice $I_n + (\lambda - 1)E_{ii}$,
- ajouter à une ligne i un multiple λ de la ligne $j \neq i$, ce qui revient à multiplier à gauche par la matrice $I_n + \lambda E_{ij}$.

(Notez que les mêmes opérations sur les colonnes sont aussi faisables, et qu'elles correspondent à des multiplications par les mêmes matrices, mais à droite.)

En procédant systématiquement, on peut toujours obtenir (après un certain nombre de telles opérations) une matrice triangulaire supérieure T, de sorte que $T = A\Pi$ où Π est le produit des matrices élémentaires correspondantes aux opérations (la matrice de la première opération tout à gauche et celle de la dernière tout à droite).

Le déterminant du produit Π est facile à calculer, puisque c'est le produit des déterminants des matrices élémentaires, et que

$$\det(P_\sigma) = \varepsilon(\sigma), \quad \det(I_n + (\lambda - 1)E_{ii}) = \lambda \neq 0, \quad \det(I_n + \lambda E_{ij}) = 1 \text{ si } i \neq j.$$

De plus, cela implique $\det(\Pi) \neq 0$. On a donc

$$\det(A) = \frac{\det(T)}{\det(\Pi)},$$

et le déterminant d'une matrice triangulaire est le produit de ses entrées diagonales.

Numériquement, cette méthode est la plus efficace, sauf dans des cas de matrices "clair-semées" ("sparse" en anglais) qui ont beaucoup d'entrées nulles. En effet, on vérifie que le nombre d'opérations est $O(n^2)$ où n est la taille de la matrice, alors qu'un calcul naïf utilisant la formule de la définition se fait en $O(n!) = O(e^{n\log(n)})$ opérations.

3.5.7 Matrices triangulaires par blocs

Définition 3.5.9. *Une matrice $A \in \mathrm{M}_n(\mathbb{K})$ est triangulaire supérieure par blocs si il existe $s \in \mathbb{N}^*$ et $k_1, \ldots, k_s \in \mathbb{N}^*$ tels que $k_1 + \cdots + k_s = n$ et que A soit de la forme :*

$$A = \left(\begin{array}{c|c|c|c|c} A_{11} & A_{12} & A_{13} & \cdots & A_{1s} \\ \hline 0 & A_{22} & A_{23} & \cdots & A_{2s} \\ \hline 0 & 0 & A_{33} & \cdots & A_{3s} \\ \hline \vdots & \vdots & \ddots & \ddots & \vdots \\ \hline 0 & 0 & 0 & \cdots & A_{ss} \end{array}\right), \tag{3.2}$$

où la matrice en position i, j est une matrice de taille $k_i \times k_j$.

On observe que les matrices $A_{ii} \in \mathrm{M}_{k_i}(\mathbb{K})$ sont carrées.

Proposition 3.5.10. *Soit $A \in \mathrm{M}_n(\mathbb{K})$ une matrice triangulaire par bloc de la forme (3.2), alors*

$$\det(A) = \det(A_{11})\det(A_{22})\ldots\det(A_{ss}) = \prod_{i=1}^{s} \det(A_{ii}).$$

Quand les blocs sont de taille 1, on retrouve la formule bien connue pour les matrices triangulaires supérieures. Bien sûr, la proposition est aussi valide pour les matrices triangulaire inférieures par blocs.

Lemme 3.5.11. *Soit $B \in \mathrm{M}_{k_1}(\mathbb{K})$ et $k_1 + k_2 = n$, alors*

$$\det\left(\begin{array}{c|c} B & 0 \\ \hline 0 & I_{k_2} \end{array}\right) = \det(B) \quad \textit{et} \quad \det\left(\begin{array}{c|c} I_{k_2} & 0 \\ \hline 0 & B \end{array}\right) = \det(B).$$

Démonstration du Lemme 3.5.11. On montre la première égalité, la seconde est similaire. On utilise la formule de la définition 3.1.13. On note a_{ij} les entrées de la matrice, avec $a_{ij} = b_{ij}$ pour $i, j \leq k_1$, $a_{ii} = 1$ pour $k_1 + 1 \leq i \leq n$ et $a_{ij} = 0$ sinon.

On observe que si une permutation $\sigma \in \mathcal{S}_n$ satisfait $\sigma(i) = i$ pour tout $k_1 + 1 \leq i \leq n$, alors elle correspond bijectivement à une permutation de $\mathcal{S}_{k_1}$ et $a_{\sigma(i)i} = 1$ pour tout $k_1 + 1 \leq i \leq n$, donc $\varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i} = \varepsilon(\sigma) \prod_{i=1}^{k_1} a_{\sigma(i)i}$.

D'autre part si une telle permutation ne satisfait pas cette condition, alors il existe $k_1 + 1 \leq i \leq n$ tel que $\sigma(i) \neq i$, donc $a_{\sigma(i)i} = 0$. Ainsi

$$\det\left(\begin{array}{c|c} B & 0 \\ \hline 0 & I_{k_2} \end{array}\right) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i} = \sum_{\sigma \in \mathcal{S}_{k_1}} \varepsilon(\sigma) \prod_{i=1}^{k_1} b_{\sigma(i)i} = \det(B).$$

□

Démonstration de la Proposition 3.5.10. Par une récurrence immédiate sur s, il suffit de traiter le cas $s = 2$. Pour cela, on utilise la méthode du pivot de Gauss. On sait qu'alors $\det(A) = \frac{\det(T)}{\det(\Pi)}$ où Π est entièrement déterminée par les opérations sur les lignes, et T est la matrice triangulaire supérieure obtenue à l'issue de l'algorithme.

La forme triangulaire par blocs nous assure que si $1 \leq i \leq k_1$ et $k_1 + 1 \leq j \leq n$, alors aucune opération (de permutations ou de mulitplications par des matrices de transvections) n'est effectuée entre les lignes L_i et L_j. Il s'ensuit que Π est diagonale par bloc de la forme

$$\Pi = \left(\begin{array}{c|c} \Pi_1 & 0 \\ \hline 0 & \Pi_2 \end{array}\right)$$

avec $\Pi_1 \in \mathrm{M}_{k_1}(\mathbb{K})$ qui encode les opérations sur les lignes 1 à k_1, qui correspondent aux opérations du pivot de Gauss sur A_{11} et $\Pi_2 \in \mathrm{M}_{k_2}(\mathbb{K})$ qui encode les opérations sur les lignes $k_1 + 1$ à n, correspondants au pivot de Gauss sur A_{22}. On a donc

$$\det(\Pi) = \det\left(\begin{array}{c|c} \Pi_1 & 0 \\ \hline 0 & I_{k_2} \end{array}\right) \det\left(\begin{array}{c|c} I_{k_1} & 0 \\ \hline 0 & \Pi_2 \end{array}\right) = \det(\Pi_1)\det(\Pi_2) \quad \text{d'après le Lemme 3.5.11.}$$

D'autre part, T est triangulaire par blocs

$$T = \left(\begin{array}{c|c} T_1 & * \\ \hline 0 & T_2 \end{array}\right)$$

où T_i résulte du pivot de Gauss appliqué à A_{ii}. De plus $\det(T) = \det(T_1)\det(T_2)$ puisque ces trois matrices sont triangulaires (pas seulement par blocs) donc leurs déterminants sont les produits des entrées diagonales. On conclut que

$$\det(A) = \frac{\det(T)}{\det(\Pi)} = \frac{\det(T_1)\det(T_2)}{\det(\Pi_1)\det(\Pi_2)} = \det(A_{11})\det(A_{22}).$$

□

3.5.8 Développement par rapport à une ligne ou une colonne

Définition 3.5.12. *Soit $A = (a_{ij})_{1\leq i,j\leq n} \in \mathrm{M}_n(\mathbb{K})$ une matrice de taille $n \times n$. Pour tout $1 \leq i, j \leq n$, on note*

$$A^{ij} = (a_{st})_{s\neq i, t\neq j} \in \mathrm{M}_{n-1}(\mathbb{K})$$

la matrice de taille $(n-1)\times(n-1)$ obtenue en effacant la ligne i et la colonne j de A. Le $ij^{ième}$ cofacteur *de A est le coefficient*

$$\tilde{a}_{i,j} := (-1)^{i+j} \det(A^{ij}).$$

Exemple 3.5.13. Soit $A = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 1 & -3 \\ 0 & 1 & 4 \end{pmatrix}$, alors $A^{23} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ et $\tilde{a}_{23} = -1$. Plus généralement, la comatrice de A (voir Définition 3.5.15) est

$$\tilde{A} = \begin{pmatrix} 7 & -8 & 2 \\ 4 & 4 & -1 \\ 3 & 3 & 3 \end{pmatrix}$$

Proposition 3.5.14. *Soit $A \in \mathrm{M}_n(\mathbb{K})$*

(a) **Développement par rapport à une ligne :** $\forall 1 \leq i \leq n$

$$\det(A) = \sum_{j=1}^{n} a_{ij}(-1)^{i+j} \det(A^{ij}) = \sum_{j=1}^{n} a_{ij}\tilde{a}_{ij}.$$

(b) Développement par rapport à une colonne : $\forall 1 \leq j \leq n$

$$\det(A) = \sum_{i=1}^{n} a_{ij}(-1)^{i+j} \det(A^{ij}) = \sum_{i=1}^{n} a_{ij}\tilde{a}_{ij}.$$

Cette proposition est très utile pour calculer le déterminant d'une matrice ayant une ligne ou une colonne avec beaucoup de zéros.

Démonstration. On prouve la formule de développement par rapport à une colonne (b). Le développement par rapport à une ligne est similaire.

$$\begin{aligned}
\det(A) &= \det(C_1, \ldots, C_j, \ldots, C_n) \\
&= (-1)^{j-1} \det(C_j, C_1, \ldots, C_{j-1}, C_{j+1}, \ldots, C_n) \\
&\qquad \text{obtenu en permutant les colonnes par } (1\ 2 \ldots j) \\
&= (-1)^{j-1} \det\left(\sum_{i=1}^{n} a_{ij} e_i, C_1, \ldots, \hat{C}_j, \ldots, C_n\right) \\
&\qquad \text{où la notation } \hat{C}_j \text{ signifie qu'on a retiré la colonne } C_j \\
&= (-1)^{j-1} \sum_{i=1}^{n} a_{ij} \det(e_i, C_1, \ldots, \hat{C}_j, \ldots, C_n) \quad \text{par linéarité.}
\end{aligned}$$

Or on peut écrire la matrice :

$$(e_i, C_1, \dots, \hat{C}_j, \dots, C_n) = \begin{pmatrix} 0 & L_1^j \\ \vdots & \vdots \\ 1 & L_i^j \\ \vdots & \vdots \\ 0 & L_n^j \end{pmatrix}$$

où L_i^j est la ligne i de la matrice A à laquelle on a retiré la $j^{\text{ième}}$ entrée (situ'ee sur la colonne C_j). En appliquant la permutation $(1\,2\dots i)$ de signature $(-1)^{i-1}$ aux lignes de cette matrice, on obtient :

$$\det \begin{pmatrix} 0 & L_1^j \\ \vdots & \vdots \\ 1 & L_i^j \\ \vdots & \vdots \\ 0 & L_n^j \end{pmatrix} = (-1)^{i-1} \det \begin{pmatrix} 1 & L_i^j \\ 0 & L_1^j \\ \vdots & \vdots \\ \hat{0} & \hat{L_i^j} \\ \vdots & \vdots \\ 0 & L_n^j \end{pmatrix} = (-1)^{i-1} \det \begin{pmatrix} 1 & L_i^j \\ 0 & A^{ij} \end{pmatrix}$$ où la ligne avec les chapeaux a été retirée.

La formule de calcul du déterminant des matrices triangulaires par blocs de la Proposition 3.5.10 assure que

$$\det \begin{pmatrix} 1 & L_i^j \\ 0 & A^{ij} \end{pmatrix} = \det(1) \det(A^{ij}) = \det(A^{ij}).$$

On obtient bien au total :

$$\det(A) = (-1)^{j-1} \sum_{i=1}^{n} a_{ij} (-1)^{i-1} \det(A^{ij}) = \sum_{i=1}^{n} a_{ij} (-1)^{i+j} \det(A^{ij}).$$

□

3.5.9 Comatrice et formule de Cramer

Définition 3.5.15. *Soit $A \in \mathrm{M}_n(\mathbb{K})$. La* comatrice *de A est la matrice de taille $n \times n$*

$$\tilde{A} = (\tilde{a}_{ij})_{1 \le i,j \le n} \in \mathrm{M}_n(\mathbb{K}),$$

où $\tilde{a}_{i,j} := (-1)^{i+j} \det(A^{ij})$ sont les cofacteurs *de A.*

Proposition 3.5.16. *Soit $A \in \mathrm{M}_n(\mathbb{K})$, et $\tilde{A} = ((-1)^{i+j} \det(A^{ij}))_{1 \le i,j \le n}$ sa comatrice. On a*

$$A\tilde{A}^t = \tilde{A}^t A = \det(A) I_n.$$

On déduit immédiatement le

Corollaire 3.5.17. *Si $A \in \mathrm{GL}_n(\mathbb{K})$, alors*

$$A^{-1} = \frac{1}{\det(A)} \tilde{A}^t.$$

Cette formule élégante peut être utilisée explicitement pour les matrices 2×2 ou 3×3, mais pour des matrices plus grandes, elle n'est pas pratique calculatoirement. Par contre, elle est très utile d'un point de vue théorique, par exemple, comme le déterminant est une fonction polynomiale des entrées de la matrice (polynôme en n^2 variables), le corollaire 3.5.17 assure que l'application $A \mapsto A^{-1}$ qui à une matrice associe son inverse est une application continue de $\mathrm{GL}_n(\mathbb{C})$ dans lui-même.

Démonstration de la Proposition 3.5.16. Soit $M := A\tilde{A}^t = (m_{ij})_{1 \leq i,j \leq n}$. Pour obtenir $A\tilde{A}^t = \det(A) I_n$, il s'agit de montrer que $m_{ij} = \det(A)\delta_{ij}$. (L'autre produit s'obtient par un calcul similaire.) On note a_{ij} les entrées de A et $\tilde{a}'_{ij} = \tilde{a_{ji}}$ les entrées de $\tilde{A}^t$. Par définition du produit matriciel, on a

$$m_{ij} = \sum_{k=1}^{n} a_{ik} \tilde{a}'_{kj} = \sum_{k=1}^{n} a_{ik} \tilde{a}_{jk}.$$

Lorsque $i = j$, la somme précédente vaut $\det(A)$ d'après la Proposition 3.5.14. Reste à montrer

qu'elle vaut zéro lorsque $i \neq j$. Pour cela, on adapte la preuve de la Proposition 3.5.14.

$$\sum_{k=1}^{n} a_{ik}\tilde{a}_{jk} = \sum_{k=1}^{n} a_{ik}(-1)^{j+k}\det(A^{jk})$$

$$= \sum_{k=1}^{n}(-1)^{j+k}\det\left(\begin{array}{c|c} a_{ik} & 0 \\ \hline a_{1k} & \\ \dots & \\ \hat{a}_{jk} & A^{jk} \\ \vdots & \\ a_{nk} & \end{array}\right) \quad \text{comme déterminant d'une matrice triangulaire par blocs,}$$

$$= \sum_{k=1}^{n}(-1)^{j-1}\det\begin{pmatrix} 0 & \dots 0 & a_{ik} & 0 \dots & 0 \\ a_{11} & \dots & a_{1k} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ \hat{a}_{j1} & \dots & \hat{a}_{jk} & \dots & \hat{a}_{jn} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nk} & \dots & a_{nn} \end{pmatrix} \quad \text{en permutant les colonnes selon } (1\,2\dots k),$$

$$= (-1)^{j-1}\det\begin{pmatrix} a_{i1} & \dots & a_{ik} & \dots & a_{in} \\ a_{11} & \dots & a_{1k} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ \hat{a}_{j1} & \dots & \hat{a}_{jk} & \dots & \hat{a}_{jn} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nk} & \dots & a_{nn} \end{pmatrix} \quad \text{par linéarité en la première ligne,}$$

$$= (-1)^{j-1}\det\begin{pmatrix} L_i \\ L_1 \\ \dots \\ \hat{L}_j \\ \dots \\ L_n \end{pmatrix} = 0 \quad \text{car la matrice possède deux lignes identiques.}$$

□

Corollaire 3.5.18. *Soit $A \in \mathrm{GL}_n(\mathbb{K})$ et $b \in \mathbb{K}^n$. Alors il existe un unique vecteur $x \in \mathbb{K}^n$ tel que $Ax = b$. De plus, si on note $A = (C_1 \ \dots \ C_n)$ donnée par ses colonnes, $A_{b,i} = (C_1 \dots C_{i-1}\ b\ C_{i+1} \dots C_n)$ obtenue en remplacant la $i^{\text{ième}}$ colonne de A par b et*

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ et } b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

alors

$$\forall 1 \leq i \leq n, \quad x_i = \frac{\det(A_{b,i})}{\det(A)},$$

Démonstration. Comme A est inversible, un tel x existe bien, est unique et vaut $x = A^{-1}b = \frac{1}{\det(A)}\tilde{A}^t b$ d'après la Proposition 3.5.17. Par produit entre une matrice et un vecteur colonne, on déduit l'expression suivante

$$\begin{aligned} x_i &= \frac{1}{\det(A)} \sum_{j=1}^{n} \tilde{a}'_{ij} b_j = \frac{1}{\det(A)} \sum_{j=1}^{n} \tilde{a}_{ji} b_j = \frac{1}{\det(A)} \sum_{j=1}^{n} (-1)^{i+j} \det(A^{ji}) b_j \\ &= \frac{1}{\det(A)} \sum_{j=1}^{n} \det(C1 \dots C_{i-1}\ e_j\ C_{i+1} \dots C_n) b_j \quad \text{par développement par rapport à la colonne } i \\ &= \frac{1}{\det(A)} \det\left(C_1 \dots C_{i-1} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} C_{i+1} \dots C_n \right) = \frac{\det(A_{b,i})}{\det(A)}. \end{aligned}$$

On fera bien attention que i désigne ici des colonnes et j des lignes. □

Chapitre 4

L'algèbre des polynômes $\mathbb{K}[X]$

Chacun sait qu'un polynôme est une expression de la forme

$$\sum_{k=0}^{n} a_k X^k = a_0 + a_1 X + a_2 X^2 + \cdots + a_n X^n, \tag{4.1}$$

où $n \in \mathbb{N}$ et les a_i sont des "nombres". De plus, des polynômes peuvent s'additionner :

$$\sum_{k=0}^{n} a_k X^k + \sum_{k=0}^{m} b_k X^k = \sum_{k=0}^{\max(n,m)} (a_k + b_k) X^k,$$

où l'on convient que $a_k = 0$ si $k > n$ et $b_k = 0$ si $k > m$. Des polynômes peuvent aussi se multiplier par un "nombre" λ comme suit :

$$\lambda \left(\sum_{k=0}^{n} a_k X^k \right) = \sum_{k=0}^{n} (\lambda a_k) X^k$$

et se multiplier entre eux comme suit :

$$\left(\sum_{=0}^{n} a_k X^k \right) \times \left(\sum_{k=0}^{m} b_k X^k \right) = \sum_{k=0}^{n+m} c_k X^k,$$

où $c_k = \sum_{i=0}^{k} a_i b_{k-i} = a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0 = \sum_{i+j=k} a_i b_j$.

Exercice 28. Développer complètement le produit $(a_0 + a_1 X + a_2 X^2 + a_3 X^3)(b_0 + b_1 X + b_2 X^2 + b_3 X^3 + b_4 X^4)$ pour retrouver les formules des coefficients c_k.

On vérifie sans difficulté majeure (c'est juste un exercice quelque peu fastidieux) que ces trois opérations donnent à l'ensemble des polynômes une structure de $\mathbb{K}$-algèbre dés que l'on s'assure que les "nombres" a_i, b_j appartiennent à un corps $\mathbb{K}$ fixé.

On a donc envie de définir la $\mathbb{K}$-algèbre des polynômes sur $\mathbb{K}$, que l'on notera $\mathbb{K}[X]$, comme l'ensemble des expressions de la forme (4.1) muni de ces 3 opérations. Si c'est certes comme

cela qu'il faut y penser, il ne s'agit pas d'une définition rigoureuse, puisque nous n'avons pas expliqué ce qu'est "l'indéterminée" X.

On peut penser que X est une variable et que les polynômes sont des fonctions de cette variable. Se pose alors la question de savoir à quel espace X cette variable appartient-elle ? On peut proposer le choix naturel $X = \mathbb{K}$, cela fonctionnerait avec $\mathbb{R}$ ou $\mathbb{C}$ mais pas avec n'importe quel corps. En fait, il s'avère préferable de distinguer un polynôme de sa fonction polynômiale naturellement associée.

4.1 Définition de l'algèbre $\mathbb{K}[X]$

Soit $\mathbb{K}$ un corps (dans les exercices, on supposera $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, rarement $\mathbb{Q}$). On note $\mathcal{P}$ l'ensemble des suites $a = (a_k)_{k\in\mathbb{N}}$ telles que

(a) $\forall k \in \mathbb{N}, a_k \in \mathbb{K}$, c'est-à-dire que la suite a est à valeurs dans $\mathbb{K}$,

(b) $\exists n \in \mathbb{N}, \forall k > n, a_k = 0$, c'est-à-dire que la suite a est stationnaire de limite nulle (ici n dépend de a).

On munit l'espace $\mathcal{P}$ d'une opération d'addition terme à terme

$$\begin{array}{rccc} + : & \mathcal{P}\times\mathcal{P} & \to & \mathcal{P} \\ & ((a_k)_{k\in\mathbb{N}}, (b_k)_{k\in\mathbb{N}}) & \mapsto & (a_k + b_k)_{k\in\mathbb{N}} \end{array}$$

d'une loi externe de produit par des scalaires (éléments de $\mathbb{K}$) terme à terme

$$\begin{array}{rccc} \cdot : & \mathbb{K}\times\mathcal{P} & \to & \mathcal{P} \\ & (\lambda, (a_k)_{k\in\mathbb{N}}) & \mapsto & (\lambda a_k)_{k\in\mathbb{N}} \end{array}$$

et d'une opération de multiplication **pas du tout** terme à terme !

$$\begin{array}{rccc} \times : & \mathcal{P}\times\mathcal{P} & \to & \mathcal{P} \\ & ((a_k)_{k\in\mathbb{N}}, (b_k)_{k\in\mathbb{N}}) & \mapsto & (c_k)_{k\in\mathbb{N}} \end{array}$$

où

$$\forall k \in \mathbb{N}, c_k = \sum_{i=0}^{k} a_i b_{k-i} = a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0 = \sum_{i+j=k} a_i b_j.$$

Proposition 4.1.1. *$(\mathcal{P}, +, \times, \cdot)$ est une $\mathbb{K}$-algèbre commutative.*

Démonstration. La preuve est laissée en exercice. On observera que le neutre additif est la suite nulle $0 = (0, 0, 0, \dots)$ tandis que le neutre multiplicatif est la suite $1 = (1, 0, 0, 0, \dots)$. □

À ce stade, on "sent" bien la proximité entre les opérations des expressions polynômiales et la $\mathbb{K}$-algèbre $\mathcal{P}$. La proposition suivante et sa preuve font le lien entre les deux.

On choisit de noter X la suite $X = (0, 1, 0, 0, \dots)$. On calcule alors par récurrence

$$X^2 = X \times X = (0, 0, 1, 0, 0, \dots) \text{ et } \forall n \in \mathbb{N}, X^n = \underbrace{X \times \cdots \times X}_{n} = (\underbrace{0, \dots, 0}_{n}, 1, 0, 0, \dots)$$

où par convention $X^0 = 1 = (1, 0, 0, \dots)$.

Proposition 4.1.2. *La famille $(X^n)_{n\in\mathbb{N}}$ est une base de la $\mathbb{K}$-algèbre $\mathcal{P}$ en tant qu'espace vectoriel.*

Démonstration. Pour toute suite $a = (a_k)_{k\in\mathbb{N}}$ dont tous les termes sont nuls au-delà du rang n :

$$\begin{aligned}(a_k)_{k\in\mathbb{N}} &= (a_0, a_1, a_2, a_3, \dots)\\ &= a_0(1,0,0,\dots) + a_1(0,1,0,0,\dots) + a_2(0,0,1,0,0,\dots) + \cdots + a_n(\underbrace{0,\dots,0}_{n},1,0,0,\dots)\\ &= a_0 + a_1X + a_2X^2 + \cdots + a_nX^n = \sum_{k=0}^{n} a_kX^k.\end{aligned}$$

Donc la famille est génératrice. Elle est également libre par unicité des coefficients, qui proviennent directement des valeurs de la suite $a = (a_k)_{k\in\mathbb{N}}$. Ainsi $(X^n)_{n\in\mathbb{N}}$ est bien une base de l'espace vectoriel $\mathcal{P}$. □

La preuve précédente a fait réapparaitre les expression polynômiales, donc on pourrait craindre de "tourner en rond". Ce n'est pas le cas, car ici X est bien défini : c'est la suite $(0, 1, 0, 0\dots)$.

Définition 4.1.3. *Une $\mathbb{K}$-algèbre $\mathcal{A}$ telle qu'il existe X dans $\mathcal{A}$ de sorte que la famille $(X^n)_{n\in\mathbb{N}}$ soit une base de $\mathcal{A}$ est appelée algèbre des polynômes sur $\mathbb{K}$.*

Ici, la puissance X^n est la multiplication dans $\mathcal{A}$ de n facteurs X.

Proposition 4.1.4. *Deux algèbres des polynômes sur $\mathbb{K}$ sont isomorphes en tant que $\mathbb{K}$-algèbre.*

On renvoie aux semestres ultérieurs du cursus pour une définition précise de la notion d'isomorphisme d'algèbres. La proposition assure essentiellement que le choix de noter "l'indeterminée" X ou X' ou Y ou Z ou T ou même $(0, 1, 0, 0, \dots)$ n'a pas d'importance. L'algèbre obtenue est "la même" pourvu que la suite de ses puissances soit une base. On peut donc considérer qu'il y a une seule algèbre des polynômes $\mathbb{K}[X]$.

La base $(X^n)_{n\in\mathbb{N}}$ est dite *base canonique* de la $\mathbb{K}$-algèbre $\mathbb{K}[X]$.

La construction de $\mathcal{P}$ a eu pour seule utilité de nous garantir l'**existence** de cette algèbre. Cette preuve d'existence peut sembler incongrue. Elle l'est moins si on pense à la question de l'existence du nombre imaginaire i et plus généralement de l'ensemble $\mathbb{C}$ des nombres complexes, qui a dérouté les mathématiciens pendant plusieurs siècles. Aujourd'hui, les mathématiciens définissent le corps $\mathbb{C}$ comme la $\mathbb{R}$-algèbre quotient $\mathbb{R}[X]/(X^2 + 1)$ (cf cours de L3 et M1).

Noter aussi que si l'on supprime la condition (b) de stationnarité dans la définition de la $\mathbb{K}$-algèbre $\mathcal{P}$, on obtient alors la $\mathbb{K}$-algèbre des séries formelles que l'on retrouvera lors de l'étude des séries entières au S4.

4.2 Degré d'un polynôme

Définition 4.2.1. *Soit $P = \sum_{k=0}^{n} a_kX^k \in \mathbb{K}[X]$ un polynôme.*

(a) Si $\forall k \in \mathbb{N}, a_k = 0$, alors $P = 0$ est appelé polynôme nul. Par convention, $\deg(0) = -\infty$.

(b) Si $P \neq 0$, on définit son degré comme

$$\deg(P) = \max\{k \in \mathbb{N} | a_k \neq 0\}.$$

Notons que le maximum existe bien vu que l'ensemble est fini et non vide. Le coefficient a_d où $d = \deg(P)$ est appelé coefficient dominant de P. Si $a_d = 1$, on dit que P est un polynôme unitaire.

Proposition 4.2.2. *On a $\forall P, Q \in \mathbb{K}[X]$,*

(a) $\deg(PQ) = \deg(P) + \deg(Q)$,

(b) $\deg(P+Q) \leq \max(\deg(P), \deg(Q))$,

(c) si $\deg(P) \neq \deg(Q)$, *alors* $\deg(P+Q) = \max(\deg(P), \deg(Q))$.

Démonstration. Les points (b) et (c) sont évidents. Pour prouver (a), on traite d'abord le cas où P ou Q est nul, alors $PQ = 0$ et la formule est valide. Supposons maintenant $P \neq 0$ et $Q \neq 0$. On montre l'égalité par récurrence sur $n = \deg(P) + \deg(Q)$.

On énonce précisement l'hypothèse de récurrence au rang n.

(HR$_n$) $\forall P, Q \in \mathbb{K}[X]$, si $\deg(P) + \deg(Q) = n$, alors $\deg(PQ) = \deg(P) + \deg(Q)$.

(HR$_0$) est vraie car si $n = \deg(P) + \deg(Q) = 0$ alors P et Q sont tous deux des constantes non-nulles. Leur produit aussi, donc $\deg(PQ) = 0 = 0 + 0 = \deg(P) + \deg(Q)$.

Supposons avoir démontré (HR$_k$) pour tout $k \leq n$. Déduisons-en (HR$_{n+1}$). On note $P = a_{d_1}X^{d_1} + P_1$ et $P = b_{d_2}X^{d_2} + P_2$ où $\deg(P_1) < d_1 = \deg(P)$ et $\deg(P_2) < d_2 = \deg(Q)$. On développe le produit comme suit :

$$PQ = (a_{d_1}X^{d_1} + P_1)(b_{d_2}X^{d_2} + P_2) = a_{d_1}b_{d_2}X^{d_1+d_2} + \underbrace{b_{d_2}X^{d_2}P_1 + a_{d_1}X^{d_1}P_2 + P_1P_2}_{\deg \leq n}.$$

Le terme sur l'accolade a degré $\leq n$ d'après l'hypothèse de récurrence utilisée pour les couples de polynômes $(b_{d_2}X^{d_2,}P_1), (a_{d_1}X^{d_1}, P_2), (P_1, P_2)$ dont les sommes de degrés sont toutes $\leq n$ et la proposition (b), et donc $\deg(PQ) = n = d + p = \deg(P) + \deg(Q)$ par (c). □

On dit qu'on fait une récurrence *forte* si on suppose connues toutes les (HR$_k$) pour $k \leq n$ afin de montrer (HR$_{n+1}$). Toutefois la différence entre récurrence forte et récurrence usuelle est minime, car une récurrence forte pour l'énoncé (HR_n) ne diffère pas d'une récurrence usuelle pour l'énoncé

$$(\text{HR}'_n) : \forall k \leq n, (\text{HR}_k) estvraie.$$

Corollaire 4.2.3. *Soit $P, Q \in \mathbb{K}[X]$. Si $PQ = 0$, alors $P = 0$ ou $Q = 0$.*

On dit que l'anneau $\mathbb{K}[X]$ est *intègre*.

Exercice 29. Montrer que si $n \geq 2$, l'anneau $M_n(\mathbb{C})$ n'est pas intègre, c'est-à-dire qu'il existe $P, Q \in M_n(\mathbb{C})$ tous deux non-nuls tels que $PQ = 0$.

Exercice 30. Montrer que les seuls éléments inversibles pour le produit de l'anneau $\mathbb{K}[X]$ sont les polynômes de degré 0, c'est-à-dire les constantes non-nulles.

Définition 4.2.4. *On note $\mathbb{K}_n[X]$ l'ensemble des polynômes de degré $\leq n$.*

D'après la Proposition 4.2.2 (b), $\mathbb{K}_n[X]$ est un sous-espace vectoriel de $\mathbb{K}[X]$. De plus, la famille $(1, X, X^2, \ldots, X^n)$ en est une base (exercice), donc $\dim(\mathbb{K}_n[X]) = n + 1$.

Pour $n \geq 1$, on observe que $\mathbb{K}_n[X]$ n'est pas une sous-algèbre de $\mathbb{K}[X]$ car cet ensemble n'est pas stable par produit (en effet, $X^n X^n = X^{2n} \notin \mathbb{K}_n[X]$).

4.3 Arithmétique de $\mathbb{K}[X]$

4.3.1 Division euclidienne

L'outil principal de l'arithmétique de $\mathbb{K}[X]$ est la division euclidienne :

Théorème 4.3.1. *Soit $A, B \in \mathbb{K}[X]$ avec $B \neq 0$. Alors il existe un unique couple (Q, R) dans $\mathbb{K}[X]^2$ tel que*

$$A = BQ + R \text{ et } \deg(R) < \deg(B).$$

On appelle Q le *quotient* et R le *reste* de la division euclidienne de A par B.

Démonstration. On montre d'abord l'**unicité**. Supposons que $A = BQ_1 + R_1$ et $A = BQ_2 + R_2$ avec $\deg(R_i) < \deg(B)$. On aurait alors $BQ_1 + R_1 = BQ_2 + R_2$, d'où

$$B(Q_1 - Q_2) = R_2 - R_1.$$

On en déduit que $R_2 - R_1$ est un multiple de B de degré $< \deg(B)$. D'après la Proposition 4.2.2, cela force $R_2 - R_1 = 0$, donc $R_2 = R_1$.

Il vient alors $B(Q_1 - Q_2) = 0$, et comme l'anneau $\mathbb{K}[X]$ est intègre, on déduit que $B = 0$, cas exclu par hypothèse, ou que $Q_1 - Q_2 = 0$, donc $Q_1 = Q_2$.

Reste à voir l'**existence**. Si $\deg(B) = 0$, alors $B = \lambda$ est un polynôme constant et $(Q, R) = (\frac{A}{\lambda}, 0)$ convient. Supposons donc $\deg(B) \geq 1$.

On procède par récurrence sur $n = \deg(A)$. On énonce l'hypothèse de récurrence :

$$(\text{HR}_n) \forall A \in \mathbb{K}[X], \text{ si } \deg(A) \leq n, \text{ alors } \exists Q, R \in \mathbb{K}[X], A = BQ + R \text{ et } \deg(R) < \deg(B).$$

Pour initialiser, on observe que si $n < \deg(B)$, alors $(Q, R) = (0, A)$ convient. D'où (HR_0).

Supposons maintenant $n \geq \deg(B) = m$, et supposons (HR_{n-1}). Déduisons-en (HR_n). Il suffit clairement de traiter le cas où $\deg(A) = n$. On note a_n le coefficient dominant de A et b_m celui de B. On pose

$$C = A - a_n X^{n-m} \frac{B}{b_m}. \tag{4.2}$$

Comme le polynôme $X^{n-m}\frac{B}{b_m}$ est unitaire de degré $n = \deg(A)$, on déduit que $\deg(C) \leq n - 1$. On peut donc appliquer l'hypothèse de récurrence à la division euclidienne de C par B. On obtient $C = BQ_1 + R_1$ avec $\deg(R_1) < \deg(B)$. L'équation (4.2) donne

$$A = B(Q_1 + \frac{a_n}{b_m} X^{n-m}) + R_1,$$

qui fournit le résultat avec $Q = Q_1 + \frac{a_n}{b_m} X^{n-m}$ et $R = R_1$. □

La preuve fournit un algorithme effectif permettant d'effectuer les divisions euclidiennes.

Exercice 31. Effectuer la division euclidienne de $A = 3X^5+2X^4-X^2+1$ par $B = X^3+X^2+2$. (On trouve $Q = 3X^2 - X + 1$ et $R = -8X^2 + 2X - 1$.)

4.3.2 Divisibilité et PGCD

Définition 4.3.2. *Soit $A, B \in \mathbb{K}[X]$, on dit que B divise A si il existe Q dans $\mathbb{K}[X]$ tel que $A = BQ$. On dit aussi que A est un multiple de B.*

Il est évident que B divise A si et seulement si le reste de la division euclidienne de A par B est nul.

Exemples 4.3.3. (a) $X - 1$ divise $X^2 - 3X + 2 = (X - 1)(X - 2)$.

(b) Si $\lambda \in \mathbb{K} \setminus \{0\}$, alors λA divise A. En effet, on pose $Q = \frac{1}{\lambda} \in \mathbb{K}[X]$.

(c) Si $\lambda \in \mathbb{K} \setminus \{0\}$, alors λ divise A. En effet, on pose $Q = \frac{1}{\lambda}A \in \mathbb{K}[X]$.

Les diviseurs de type (b) et (c) ci-dessus sont dits *triviaux*.

Exercice 32. Soient $P, Q, R \in \mathbb{K}[X]$. Montrer que si P divise Q et Q divise R, alors P divise R.

Théorème 4.3.4. *Soit $A, B \in \mathbb{K}[X]$. Il existe un unique polynôme $D \in \mathbb{K}[X]$ unitaire tel que pour tout $P \in \mathbb{K}[X]$, P divise A et B si et seulement si P divise D.*

De plus, il existe $S, T \in \mathbb{K}[X]$ tels que

$$D = SA + TB. \tag{4.3}$$

Le polynôme D est appelé *plus grand diviseur commun* de A et B, noté $D = \text{pgcd}(A, B)$. L'égalité (4.3) est appelée identité de Bézout. Noter que S et T ne sont pas uniques.

On rappelle quelques propriétés du PGCD qui seront utiles dans la démonstration du théorème.

Propriétés 4.3.5. *$\forall A, B \in \mathbb{K}[X]$, on a*

(a) $\forall P, Q_1, Q_2 \in \mathbb{K}[X]$, si P divise A et B, alors P divise $AQ_1 + BQ_2$,

(b) $pgcd(A, 0) = \frac{A}{coeff.dom.(A)}$,

(c) $\forall \lambda, \mu \in \mathbb{K} \setminus \{0\}, pgcd(\lambda A, \mu B) = pgcd(A, B)$.

(d) $\forall Q \in \mathbb{K}[X], pgcd(AQ, BQ) = pgcd(A, B)\frac{Q}{coeff.dom.(Q)}$.

Démonstration. Il s'agit d'un exercice. On pourra s'inspirer de l'arithmétique de $\mathbb{Z}$. □

Démonstration du Théorème 4.3.4. On montre d'abord l'**unicité**. Si D_1 et D_2 ont tous deux cette propriété, alors $\forall P \in \mathbb{K}[X]$, P divise D_1 si et seulement si P divise D_2. On en déduit que D_1 divise D_2, donc que $\deg(D_1) \leq \deg(D_2)$, et symétriquement que D_2 divise D_1, donc que $\deg(D_2) \leq \deg(D_1)$. Ainsi $\deg(D_1) = \deg(D_2)$. Or les seuls diviseurs de D_1 de même degré sont les multiples de D_1 par un facteur scalaire. L'égalité $D_1 = D_2$ découle alors du caractère unitaire.

Reste à prouver l'**existence**. On procède par récurrence sur $m = \deg(B)$.

On énonce précisement l'hypothèse de récurrence au rang m : (HR$_m$) $\forall A \in \mathbb{K}[X], \forall B \in \mathbb{K}_m[X]$, il existe $D \in \mathbb{K}[X]$ tel que tout polynôme P divise A et B si et seulement si P divise D. De plus, il existe $S, T \in \mathbb{K}[X]$ tels que $D = SA + TB$.

Initialisation : pour $m = -\infty$, alors $B = 0$. Comme tout polynôme P divise le polynôme nul (car $P \times 0 = 0$), on déduit qu'un polynôme P divise A et 0 si et seulement si P divise A. Le théorème est vrai avec D égal à l'unitarisé de A (c'est-à-dire A divisé par son coefficient dominant), S le polynôme constant égal à l'inverse du coefficient dominant de A et T n'importe quel élément de $\mathbb{K}[X]$. On a montré (HR$_{-\infty}$)

Hérédité : supposons (HR$_{m-1}$) et déduisons-en (HR$_m$). On peut supposer $\deg(B) = m$.

On effectue la division euclidienne de A par B. On obtient Q, R tels que $A = BQ + R$ et $\deg(R) < \deg(B)$. On vérifie qu'un polynôme P divise A et B si et seulement si il divise B et R.

En effet, si P divise A et B, alors P divise B et aussi $A - BQ = R$. Réciproquement, si P divise B et R, alors P divise B et aussi $A = BQ + R$.

On utilise alors l'hypothèse de récurrence (HR$_{m-1}$) pour B et R (c'est possible car $\deg(R) \leq m - 1$). On déduit qu'il existe un unique polynôme unitaire D tel que P divise B et R si et seulement si P divise D. Cela prouve la première partie.

De plus notre hypothèse de récurrence fournit une identité de Bézout, soit $T_1, S_1 \in \mathbb{K}[X]$ tels que

$$\begin{aligned} D &= S_1 B + T_1 R \\ &= S_1 B + T_1(A - BQ) \\ &= (S_1 - T_1 Q)B + T_1 A. \end{aligned}$$

Cela prouve (4.3) avec $S = T_1$ et $T = S_1 - T_1 Q$. □

La preuve fournit un algorithme effectif (appelé algorithme d'Euclide) permettant de calculer D et d'obtenir une identité de Bézout.

Exercice 33. Soit $A = X^4 - 2X^3 + 3X^2 - 5X - 2$ et $B = X^3 - X^2 - X - 2$. Déterminer $D = PGCD(A, B)$ ainsi qu'une identité de Bézout.

On suit l'algorithme d'Euclide de la preuve du Théorème 4.3.4. Pour cela, on effectue les divisions euclidiennes successives

$$\begin{aligned} A &= BQ_1 + R_1 \\ B &= R_1 Q_2 + R_2 \\ R_1 &= R_2 Q_2 + R_3 \text{ etc.} \end{aligned}$$

L'unitarisé du premier reste non-nul fournit le PGCD, et on "remonte" les équations pour obtenir une égalité de Bézout. On obtient ici

$$\begin{aligned} \text{pgcd}(A, B) &= \frac{R_3}{\text{coeff.dom.}(R_3)} = X - 2 \\ S &= -\frac{3}{7}X - \frac{1}{7} \\ T &= \frac{3}{7}X^2 - \frac{2}{7}X + \frac{8}{7}. \end{aligned}$$

4.3.3 Le point de vue "idéal"

Définition 4.3.6. *Soit $(A, +, \times)$ un anneau commutatif. On dit qu'une partie $I \subset A$ non-vide est un idéal de A si les deux conditions suivantes sont satisfaites :*

(a) $\forall a, b \in I, a + b \in I$ *et* $-a \in I$ *(c'est-à-dire que I est un sous-groupe du groupe additif $(A, +)$,*

(b) $\forall a \in I, \forall b \in A, ab \in I$ *(on dit parfois que les idéaux sont "fortement stables" par produits).*

Exemples 4.3.7. (a) Si $a_0 \in A$, on vérifie que la partie $a_0A = \{a_0b | b \in A\}$ est un idéal de A (exercice). Un tel idéal est appelé *idéal principal*. On dit que a_0 est un *générateur* de cet idéal.

(b) Si $a_0, a_1 \in A$, la partie $a_0A + a_1A = \{a_0b_0 + a_1b_1 | b_i \in A\}$ est un idéal (exercice).

Définition 4.3.8. *Un anneau principal est un anneau pour lequel tous les idéaux sont des idéaux principaux.*

On verra dans le cours d'arithmétique que l'anneau $\mathbb{Z}$ est principal.

Proposition 4.3.9. *L'anneau $\mathbb{K}[X]$ est un anneau principal, c'est-à-dire que si I est un idéal de $\mathbb{K}[X]$, alors il existe $P_0 \in \mathbb{K}[X]$ tel que $I = P_0\mathbb{K}[X] = \{P_0Q | Q \in \mathbb{K}[X]\}$.*

On verra en L3 qu'il existe des anneaux non-principaux. C'est le cas par exemple de l'anneau $\mathbb{Z}[X]$ des polynômes à coefficients entiers, où l'idéal $2\mathbb{Z}[X] + X\mathbb{Z}[X]$ n'est pas principal.

Démonstration. Laissée en exercice (cf TD). On pourra considérer P_0 un élément de I non-nul de degré minimal. □

La proposition 4.3.9 assure que pour tout A, B dans $\mathbb{K}[X]$, l'idéal $A\mathbb{K}[X] + B\mathbb{K}[X]$ est un idéal principal, donc de la forme $P_0\mathbb{K}[X]$.

Proposition 4.3.10. *L'unique polynôme unitaire $P_0 \in \mathbb{K}[X]$ tel que*

$$A\mathbb{K}[X] + B\mathbb{K}[X] = P_0\mathbb{K}[X]$$

est le PGCD de A et B, soit $P_0 = pgcd(A, B)$.

On peut donc définir le PGCD de A et B comme le générateur unitaire de l'idéal $A\mathbb{K}[X] + B\mathbb{K}[X]$. Il s'agit d'une reformulation de la notion de PGCD. Avant de la vérifier, on reformule la divisibilité en termes d'inclusion d'idéaux.

Lemme 4.3.11. *Soit $P, Q \in \mathbb{K}[X]$, alors P divise Q si et seulement si $Q\mathbb{K}[X] \subset P\mathbb{K}[X]$.*

Démonstration. Si $Q\mathbb{K}[X] \subset P\mathbb{K}[X]$, alors en particulier $Q \in P\mathbb{K}[X]$, donc il existe T tel que $Q = PT$, donc P divise Q.

Réciproquement, si P divise Q, il existe T tel que $Q = PT$ donc

$$\begin{aligned} Q\mathbb{K}[X] &= \{QS | S \in \mathbb{K}[X]\} \\ &= \{PTS | S \in \mathbb{K}[X]\} \\ &\subset \{PS' | S' \in \mathbb{K}[X]\} = P\mathbb{K}[X]. \end{aligned}$$

□

Démonstration de la Proposition 4.3.10. D'après le Théorème 4.3.4, il s'agit de montrer qu'un polynôme P divise A et B si et seulement si P divise P_0.

Si P divise A et B, alors le lemme précédent assure que $A\mathbb{K}[X] \subset P\mathbb{K}[X]$ et $B\mathbb{K}[X] \subset P\mathbb{K}[X]$. Il s'ensuit que $A\mathbb{K}[X] + B\mathbb{K}[X] = P_0\mathbb{K}[X] \subset P\mathbb{K}[X]$, donc que P divise P_0.

Réciproquement si P divise P_0, alors $P_0\mathbb{K}[X] = A\mathbb{K}[X] + B\mathbb{K}[X] \subset P\mathbb{K}[X]$. A fortiori $A\mathbb{K}[X] \subset P\mathbb{K}[X]$ et $B\mathbb{K}[X] \subset P\mathbb{K}[X]$ donc P divise A et B. □

Exercice 34. Soit $A, B \in \mathbb{K}[X]$.

(a) Montrer qu'il existe un unique polynôme unitaire M tel que pour tout polynôme P on ait A et B divisent P si et seulement si M divise P.

Indication : on pourra considérer les idéaux $A\mathbb{K}[X]$ et $B\mathbb{K}[X]$.

Ce polynôme M est appelé plus petit multiple commun de A et B, noté $\text{ppcm}(A, B)$.

(b) On suppose A et B unitaires, montrer que

$$\text{pgcd}(A, B)\text{ppcm}(A, B) = AB.$$

4.3.4 Factorisation

Définition 4.3.12. *Un polynôme $P \in \mathbb{K}[X]$ est irréductible si il admet exactement deux diviseurs unitaires.*

Le polynôme nul admet une infinité de diviseurs unitaires, donc n'est pas irréductible. Les polynômes de degré 0 n'admettent qu'un diviseur unitaire, donc ne sont pas irréductibles. Tout autre polynôme admet **au moins** deux diviseurs unitaires : 1 et son unitarisé.

Exemples 4.3.13. (a) Les polynômes du premier degré sont irréductibles. Les deux seuls diviseurs unitaires de $a_0 + a_1X$ sont 1 et $\frac{a_0}{a_1} + X$.

(b) Le polynôme $X^2 + 1$ est irréductible dans $\mathbb{R}[X]$. En effet, s'il avait un autre diviseur unitaire, celui-ce serait de degré 1, et on aurait nécessairement $X^2 + 1 = (X + a)(X + b)$ (vérifiez-le !), mais alors en remplacant X par $-a$ on aurait $0 < (-a)^2 + 1 = (-a + a)(-a + b) = 0$ ce qui est absurde.

(c) Le polynôme $X^2 + 1$ n'est pas irréductible dans $\mathbb{C}[X]$, en effet $X^2 + 1 = (X - i)(X + i)$. Il a 4 diviseurs unitaires : $1, X^2 + 1, X - i, X + i$.

(d) Le polynôme $X^2 - 2$ est irréductible dans $\mathbb{Q}[X]$, mais pas dans $\mathbb{R}[X]$ car $X^2 - 2 = (X - \sqrt{2})(X + \sqrt{2})$.

Exercice 35. Soient $P, Q \in \mathbb{K}[X]$. On suppose que P divise Q et $\deg(P) = \deg(Q)$. Montrer qu'il existe $\lambda \in \mathbb{K} \setminus \{0\}$ tel que $P = \lambda Q$.

On note que l'irréductibilité dépend de l'algèbre $\mathbb{K}[X]$, pas seulement du polynôme P. Tout polynôme se décompose en produit de facteurs irréductibles :

Théorème 4.3.14. *Soit $A \in \mathbb{K}[X]$. On suppose $A \neq 0$. Alors*

(a) il existe des polynômes irréductibles unitaires $P_i \in \mathbb{K}[X]$ de sorte que

$$A = \lambda \Pi_{i=1}^{r} P_i,$$

où λ est le coefficient dominant de A.

(b) De plus cette décomposition est unique à l'ordre des facteurs près. C'est-à-dire que si $A = \mu \Pi_{j=1}^{s} Q_j$, alors $\lambda = \mu$, $r = s$ et il existe une application $\varphi : \{1, \ldots, r\} \to \{1, \ldots, s\}$ bijective telle que $\forall i \in \{1, \ldots, r\}$, $P_i = Q_{\varphi(i)}$.

Les polynômes irréductibles jouent dans $\mathbb{K}[X]$ le rôle des nombres premiers dans l'anneau $\mathbb{Z}$.

On écrira parfois cette décomposition sous la forme $A = \lambda \Pi_{i=0}^{n} P_i^{\alpha_i}$ où $\alpha_i \geq 1$ et les P_i sont deux à deux distincts. Il suffit de regrouper les facteurs identiques.

Un anneau intègre possedant une telle décomposition est dit *factoriel*. En fait, tout anneau principal est factoriel (l'étudiant.e interessé.e par cette assertion vérifiera que nous n'utilisons aucune spécificité de $\mathbb{K}[X]$ pour établir le Théorème 4.3.14 de décomposition en irréductibles). L'anneau $\mathbb{Z}[X]$ aussi est factoriel, bien qu'il ne soit pas principal.

Lemme 4.3.15. *Soit $A \in \mathbb{K}[X]$ de degré ≥ 1. Alors il existe P irréductible unitaire non-constant divisant A.*

Démonstration. Soit $E = \{\deg(Q) | Q \text{ unitaire divise } A \text{ et } \deg(Q) \geq 1\}$. Cette partie de $\mathbb{N}$ est non-vide car elle contient $\deg(A)$. Soit m son minimum et P un polynome de E de degré $m \geq 1$. Si P n'était pas irréductible, on aurait $P = UV$ avec $m > \deg(U) \geq 1$ ou $m > \deg(V) \geq 1$, ce qui contredirait la minimalité de m. □

*Démonstration de l'***existence** *(a) dans le Théorème 4.3.14.* On procède par récurrence sur le degré de A. C'est trivial si $\deg(A) = 0$. Supposons que la factorisation existe pour des polynômes de degré $\leq n - 1$ et supposons que $\deg(A) = n$. Le Lemme 4.3.15 assure l'existence de P_1 irréductible unitaire non-constant divisant A, d'où $A = P_1 A_1$. L'hypothèse de récurrence s'applique à A_1 pour fournir $A_1 = \lambda \Pi_{i=2}^{r} P_i$. □

Lemme 4.3.16 (Lemme de Gauss)**.** *Soit $A, B, C \in \mathbb{K}[X]$. Si A divise BC et pgcd$(A, C) = 1$, alors A divise B.*

Démonstration. D'après Bézout il existe S, T dans $\mathbb{K}[X]$ tels que $1 = SA + TC$. Alors $B = SAB + TCB$. On observe que A divise SAB et que A divise TCB (car divise BC), donc A divise leur somme $SAB + TCB = B$. □

Exercice 36. Soit $P \in \mathbb{K}[X]$ un polynôme irréductible unitaire. Soient $A_1, \ldots, A_r \in \mathbb{K}[X]$.

(i) Montrer que $\text{pgcd}(A_1, P) = \begin{cases} 1 & \text{si } P \text{ ne divise pas } A_1, \\ P & \text{si } P \text{ divise } A_1. \end{cases}$

(ii) En déduire que si P divise $A_1 \ldots A_r$, alors il existe i tel que P divise A_i.

*Démonstration de l'***unicité** *(b) dans le Théorème 4.3.14.* On se contente d'esquisser la preuve. L'étudiant.e est invité.e à écrire une démonstration rigoureuse par récurrence sur le degré de A.

L'exercice précédent montre que P_1 divise l'un des facteurs Q_j. Les deux étant irréductibles, on obtient $P_1 = Q_j$. On simplifie l'égalité par ces deux facteurs, et on a ramené le problème à des polynômes de degré moindre. □

Corollaire 4.3.17. *Soit $A, B \in \mathbb{K}[X]$ tous deux non-nuls. Soient $P_1, \ldots, P_n$ les polynômes irréductibles unitaires divisant A ou B. On peut alors écrire $A = \lambda \Pi_{i=1}^n P_i^{\alpha_i}$ et $B = \mu \Pi_{i=1}^n P_i^{\beta_i}$, où $\lambda, \mu \in \mathbb{K}$ et $\alpha_i, \beta_i \in \mathbb{N}$. Alors*

$$pgcd(A, B) = \Pi_{i=1}^n P_i^{\min(\alpha_i, \beta_i)}.$$

Démonstration. Soit $D = \Pi_{i=1}^n P_i^{\min(\alpha_i, \beta_i)}$. On a D divise A et B (donc si P divise D, il divise aussi A et B).

D'autre part, si P divise A et B, tout facteur irréductible de P est nécessairement parmi les P_i pour $1 \leq i \leq n$. On peut donc écrire $P = \nu \Pi_{i=1}^n P_i^{\gamma_i}$. Comme P divise A, on a pour tout i, $\gamma_i \leq \alpha_i$. De même, comme P divise B, on a pour tout i, $\gamma_i \leq \beta_i$. On déduit que P divise D. □

Exercice 37. Donner une expression similaire pour le PPCM de A et B. En déduire que si A et B sont unitaires, alors $\text{pgcd}(A, B)\text{ppcm}(A, B) = AB$.

4.4 Racines des polynômes

4.4.1 Application polynomiale associée et racines

Définition 4.4.1. *Pour $P = \sum_{k=0}^n a_k X^k$ un polynôme de $\mathbb{K}[X]$, on définit son application polynomiale associée comme l'application :*

$$\begin{array}{rccc} \widetilde{P}: & \mathbb{K} & \to & \mathbb{K} \\ & x & \mapsto & \widetilde{P}(x) := \sum_{k=0}^n a_k x^k. \end{array}$$

Proposition 4.4.2. *L'application*

$$\begin{array}{rccc} \widetilde{\cdot}: & \mathbb{K}[X] & \to & \mathcal{F}(\mathbb{K}, \mathbb{K}) \\ & P & \mapsto & \widetilde{P} \end{array}$$

est un morphisme de $\mathbb{K}$-algèbres. C'est-à-dire que c'est une application linéaire du $\mathbb{K}$-espace vectoriel $\mathbb{K}[X]$ vers le $\mathbb{K}$-espace vectoriel $\mathcal{F}(\mathbb{K}, \mathbb{K})$ telle que $\forall P, Q \in \mathbb{K}[X], \widetilde{PQ} = \widetilde{P}\widetilde{Q}$ et $\widetilde{1} = 1$.

Noter que dans la dernière égalité, le 1 du membre de gauche est le polynôme constant unitaire de $\mathbb{K}[X]$ alors que le 1 du membre de droite est la fonction constante $\forall x \in \mathbb{K}, 1(x) = 1$.

Pour alléger les notations, on notera souvent aux sections suivantes P à la place de $\widetilde{P}$.

Démonstration. La preuve est immédiate quand on a bien compris les définitions. □

Définition 4.4.3. *Soit $P \in \mathbb{K}[X]$ et $a \in \mathbb{K}$. On dit que a est racine de P si $\tilde{P}(a) = 0$.*

La proposition suivante est fondamentale. Sa preuve est très simple.

Proposition 4.4.4. *Soit $P \in \mathbb{K}[X]$ et $a \in \mathbb{K}$. Alors a est racine de P si et seulement si le polynôme $X - a$ divise P.*

On a déjà utilisé cette proposition à l'exemple 4.3.13 (b).

Démonstration. Si $X - a$ divise P, il existe $Q \in \mathbb{K}[X]$ tel que $P = (X - a)Q$. On a alors $\widetilde{P}(a) = (a - a)\widetilde{Q}(a) = 0$.

Réciproquement, on effectue la division euclidienne de P par $X - a$, on obtient Q, R tels que $P = (X - a)Q + R$ avec $\deg(R) < \deg(X - a) = 1$, donc R est constant. De plus $\widetilde{P}(a) = 0 = (a - a)\widetilde{Q}(a) + \widetilde{R}(a) = \widetilde{R}(a)$. Donc $R = 0$. □

Corollaire 4.4.5. *Soit $0 \neq P \in \mathbb{K}[X]$. Si $\deg(P) = n$, alors P a au plus n racines.*

Démonstration. Immédiate par récurrence sur le degré. Évident si $\deg(P) = 0$. Si $\deg(P) = n + 1$, soit P n'a pas de racine et le résultat tient, soit P a une racine a, et alors $P = (X - a)Q$ où $\deg(Q) = n$ a au plus n racines par hypothèse de récurrence. □

Définition 4.4.6. *Soit $P \in \mathbb{K}[X]$, $a \in \mathbb{K}$ et $k \in \mathbb{N} \setminus \{0\}$. On dit que a est racine d'ordre au moins k de P si le polynôme $(X - a)^k$ divise P.*

Il s'ensuit qu'on dit que a est racine *d'ordre exactement* k si $(X - a)^k$ divise P, mais $(X - a)^{k+1}$ ne divise pas P.

On parle aussi de la *multiplicité* d'une racine pour désigner son ordre.

Proposition 4.4.7. *Soit $P \in \mathbb{K}[X]$. On note $a_1, \ldots, a_r \in \mathbb{K}$ ses racines et $m_1, \ldots, m_r$ leurs multiplicités. Alors $m_1 + \cdots + m_r \leq \deg(P)$.*

Démonstration. Il suffit de remarquer que les $(X - a_i)$ sont des facteurs irréductibles de P. □

Proposition 4.4.8. *Soit $\mathbb{K}$ un corps* **infini**. *L'application $\widetilde{\cdot} : \mathbb{K}[X] \to \mathcal{F}(\mathbb{K}, \mathbb{K})$ qui à tout polynôme P associe son application polynomiale $\widetilde{P}$ est injective.*

Il s'ensuit que cette application réalise un isomorphisme d'algèbres entre $\mathbb{K}[X]$ et son image qui est l'ensemble des applications polynomiales de $\mathbb{K}$ dans $\mathbb{K}$.

Cette proposition assure que nous aurions pu définir les polynômes comme des applications polynomiales, à la condition que le corps $\mathbb{K}$ soit infini, ce qui n'est pas toujours le cas.

Démonstration. Comme cette application est linéaire d'après la Proposition 4.4.2, il suffit de montrer que $\mathrm{Ker}(\widetilde{\cdot}) = \{0\}$. Soit donc P un polynôme tel que $\widetilde{P} = 0$ est la fonction nulle. Alors tout élément a de $\mathbb{K}$ est une racine de P. Le corps $\mathbb{K}$ étant infini, P a une infinité de racines. Le Corollaire 4.4.5 assure que $P = 0$. □

Exercice 38. Soit $\mathbb{K}$ un corps fini. Trouver un polynôme non-nul dont l'application polynomiale associée est nulle.

4.4.2 Polynôme dérivé et formule de Taylor

Définition 4.4.9. *Soit $P = \sum_{k=0}^{n} a_k X^k \in \mathbb{K}[X]$ un polynôme. On définit son polynôme dérivé comme*

$$P' = \sum_{k=1}^{n} k a_k X^{k-1}.$$

On définit par induction le $(k+1)^{ième}$ polynôme dérivé $P^{(k+1)} = (P^{(k)})'$ comme le polynôme dérivé du $k^{ième}$ polynôme dérivé.

On note que $\deg(P') = \deg(P) - 1$ dés que $\deg(P) \geq 1$. Bien sûr, dans le cas d'un polynôme réel $P \in \mathbb{R}[X]$, on retrouve la dérivation usuelle, c'est-à-dire que $(\widetilde{P})' = \widetilde{(P')}$.

Théorème 4.4.10 (Formule de Taylor)**.** *Soit $P \in \mathbb{C}_n[X]$ un polynôme complexe de degré $\leq n$. Soit $a \in \mathbb{C}$, on a*

$$P = \sum_{i=0}^{n} \frac{P^{(i)}(a)}{i!}(X-a)^i = P(a) + P'(a)(X-a) + \frac{P''(a)}{2}(X-a)^2 + \cdots + \frac{P^{(n)}(a)}{n!}(X-a)^n.$$

Ce théorème est énoncé dans $\mathbb{C}$. Il s'ensuit qu'il est toujours valide dans les sous-corps de $\mathbb{C}$ comme $\mathbb{R}$ et $\mathbb{Q}$ qui nous interessent ce semestre. Cette formule n'est pas valide dans des corps finis. Par exemple dans $\mathbb{Z}/p\mathbb{Z}$ on a $p = 0$, donc il n'est pas possible de diviser par $p! = 0$. En fait, la bonne hypothèse sur le corps $\mathbb{K}$ n'est pas qu'il soit infini, mais qu'il soit de *caractéristique* 0, c'est-à-dire qu'il contienne $\mathbb{Q}$ (on renvoie aux années ultérieures pour la définition précise de la caractéristique d'un corps).

Démonstration. On sait que la famille $\{(X-a)^k\}_{k=0}^{n}$ est une base de $\mathbb{C}_n[X]$. Donc $\exists c_0, \ldots, c_n \in \mathbb{K}$ tels que $P = \sum_{i=0}^{n} c_i (X-a)^i$. Reste à identifier les coefficients c_i.

Pour $k = 0$, on a $P(a) = c_0$.

Pour $k = 1$, on a $P' = \sum_{k=1}^{n} k c_k (X-a)^{k-1}$ dont le coefficient constant est $P'(a) = c_1$.

Pour $k = 2$, on a $P'' = \sum_{k=2}^{n} k(k-1) c_k (X-a)^{k-2}$ dont le coefficient constant est $P''(a) = 2c_2$.

Pour $k = 3$, on a $P^{(3)} = \sum_{k=3}^{n} k(k-1)(k-2) c_k (X-a)^{k-3}$ dont le coefficient constant est $P^{(3)}(a) = 3 \times 2 c_3 = 3! c_3$.

Par récurrence immédiate, on a $P^{(i)} = \sum_{k=i}^{n} \frac{k!}{(k-i)!} c_k (X-a)^{k-i}$ dont le coefficient constant est $P^{(i)}(a) = i! c_i$. □

Corollaire 4.4.11. *Soit $P \in \mathbb{C}[X]$ et a un nombre complexe. Alors a est racine au moins $k^{ième}$ de P si et seulement si $P(a) = P'(a) = \cdots = P^{(k-1)}(a) = 0$.*

En particulier, a est racine au moins double de P si et seulement si $P(a) = P'(a) = 0$.

Démonstration. Soit $n = \deg(P)$. D'après la formule de Taylor, on a si $P(a) = P'(a) = \cdots = P^{(k-1)}(a) = 0$ alors $P = \sum_{i=k}^{n} \frac{P^{(i)}(a)}{i!}(X-a)^i = (X-a)^k Q$ pour un certain polynôme Q. Donc a est racine au moins $k^{\text{ième}}$ de P.

Réciproquement, si $P = (X-a)^k Q$, l'unicité des coefficients dans la base $\{(X-a)^k\}_{k=0}^{n}$ assure que $P = \sum_{i=k}^{n} \frac{P^{(i)}(a)}{i!}(X-a)^i$, donc que $P(a) = P'(a) = \cdots = P^{(k-1)}(a) = 0$. □

4.5 Décomposition dans $\mathbb{C}[X]$ et $\mathbb{R}[X]$

4.5.1 Décomposition dans $\mathbb{C}[X]$

La propriété principale du corps $\mathbb{C}$ est le résultat suivant :

Théorème 4.5.1. *[Théorème de d'Alembert-Gauss] Tout polynôme de $\mathbb{C}[X]$ admet au moins une racine dans $\mathbb{C}$.*

On dit d'un corps $\mathbb{K}$ qu'il est *algébriquement clos* si tout polynôme de $\mathbb{K}[X]$ admet une racine dans $\mathbb{K}$. Le téorème s'énonce donc aussi : $\mathbb{C}$ est algébriquement clos.

La connaissance de la démonstration, difficile, ne figure pas au programme de ce semestre. Elle est donnée en section 4.7 par souci de complétude.

Corollaire 4.5.2. *Les polynômes irréductibles de $\mathbb{C}[X]$ sont exactement les polynômes de degré 1.*

Démonstration. On a vu à l'exemple 4.3.13 que les polynômes de degré 1 sont toujours irréductibles. Reste à voir que ce sont les seuls. Soit P un polynôme de degré ≥ 2. D'après le Théorème 4.5.1 de d'Alembert-Gauss, P admet une racine a, donc la Proposition 4.4.4 assure que P possède au moins trois diviseurs unitaires : $1, X-a$ et l'unitarisé de P. Il n'est donc pas irréductible. □

Définition 4.5.3. *Un polynôme est dit scindé s'il est égal à un produit de facteurs du premier degré.*

Corollaire 4.5.4. *Tout polynôme de $\mathbb{C}[X]$ est scindé.*

Démonstration. Cela découle du Théorème 4.3.14 de décomposition en facteurs irréductibles et du Corollaire 4.5.2. □

La Proposition 4.4.7 peut être améliorée dans $\mathbb{C}[X]$.

Corollaire 4.5.5. *Soit $P \in \mathbb{C}[X]$. On note $a_1, \ldots, a_r \in \mathbb{C}$ ses racines et $m_1, \ldots, m_r$ leurs multiplicités. Alors $m_1 + \cdots + m_r = \deg(P)$.*

4.5.2 Décomposition dans $\mathbb{R}[X]$

La situation est un peu plus compliquée dans $\mathbb{R}[X]$.

Théorème 4.5.6. *Les polynômes irréductibles de $\mathbb{R}[X]$ sont exactement les polynômes de degré 1 et les polynômes de degré 2 à discriminant strictement négatif.*

On rappelle que le discriminant du polynôme $aX^2 + bX + c$ est $\Delta = b^2 - 4ac$. On rappelle que $P = aX^2 + bX + c \in \mathbb{R}[X]$ a une racine réelle si et seulement si $\Delta \geq 0$ (et cette racine est unique si et seulement si $\Delta = 0$).

Démonstration. On sait que les polynômes de degré 1 sont irréductibles. Un polynôme du degré 2 est réductible si et seulement si il admet un diviseur de degré 1, donc une racine, c'est-à-dire si et seulement si son discriminant est ≥ 0.

Reste à voir que les polynômes de degré ≥ 3 ne sont pas irréductibles. Soit P un tel polynôme. S'il admet un racine réelle, alors il n'est pas irréductible. Sinon, on utilise l'inclusion $\mathbb{R}[X] \subset \mathbb{C}[X]$. On peut donc voir P comme un polynôme complexe, qui admet au moins une racine $a \in \mathbb{C}$ d'après le théorème de d'Alembert-Gauss. Par hypothèse, $a \notin \mathbb{R}$. Comme les coefficients c_k de P sont réels, on a

$$0 = \overline{0} = \overline{P(a)} = \overline{\sum_{k=0}^{n} c_k a^k} = \sum_{k=0}^{n} \bar{c_k}\bar{a}^k = \sum_{k=0}^{n} c_k \overline{a}^k = P(\overline{a}),$$

où le surlignage désigne la conjugaison des nombres complexes. Donc $\bar{a}$ est aussi une racine. Les polynômes $X - a$ et $X - \bar{a}$ sont irréductibles et distincts dans $\mathbb{C}[X]$ et ils divisent tous deux P dans $\mathbb{C}[X]$. Il s'ensuit que leur produit $B = (X - a)(X - \bar{a}) = X^2 - 2\text{Re}(a)X + |a|^2$, qui a coefficients réels, divise P dans $\mathbb{C}[X]$.

On affirme que B divise aussi P dans $\mathbb{R}[X]$. Cette affirmation fournit un diviseur strict de P, montrant que ce dernier est réductible.

Mais il faut encore prouver l'affirmation. Pour cela, on effectue la division euclidienne de P par B dans $\mathbb{R}[X]$, on obtient Q, R dans $\mathbb{R}[X]$ tels que $P = BQ + R$ et $\deg(R) < 2$. Comme $\mathbb{R}[X] \subset \mathbb{C}[X]$, l'égalité $P = BQ + R$ est aussi la division euclidienne de P par B dans $\mathbb{C}[X]$. Or par unicité du quotient et du reste dans la division euclidienne, on déduit du fait que B divise P dans $\mathbb{C}[X]$ que $R = 0$, et donc que $P = BQ$ avec $Q \in \mathbb{R}[X]$. □

La preuve fournit une méthode pour factoriser un polynôme dans $\mathbb{R}[X]$ dés lors qu'on connait sa factorisation dans $\mathbb{C}[X]$. Il suffit de regrouper les racines complexes par paires conjuguées, dont les produits sont les diviseurs irréductibles de degré 2 de notre polynôme, les racines réelles fournissant les diviseurs irréductibles de degré 1.

4.5.3 Factorisations explicites

Racines de l'unité

Proposition 4.5.7. *Soit $n \geq 1$ entier. Dans $\mathbb{C}[X]$, on a*

$$X^n - 1 \;=\; \prod_{k=0}^{n-1}(X - e^{\frac{i2k\pi}{n}}) = \begin{cases} \prod_{k=-\frac{n}{2}+1}^{\frac{n}{2}}(X - e^{\frac{i2k\pi}{n}}) & \textit{si } n \textit{ pair,} \\ \prod_{k=-\frac{n-1}{2}}^{\frac{n-1}{2}}(X - e^{\frac{i2k\pi}{n}}) & \textit{si } n \textit{ impair.} \end{cases}$$

Démonstration. On a pour tout $k \in \mathbb{Z}$,

$$\left(e^{\frac{i2k\pi}{n}}\right)^n = e^{i2k\pi} = 1.$$

Donc les $e^{\frac{i2k\pi}{n}}$ sont des racines de $X^n - 1$. De plus, si $k \neq k' \text{mod} n$, alors $e^{\frac{i2k\pi}{n}} \neq e^{\frac{i2k'\pi}{n}}$, car $e^{\frac{i2(k-k')\pi}{n}} \neq 1$ puisque $0 < \left|\frac{k-k'}{n}\right| < 1$ donc $\frac{2(k-k')\pi}{n} \notin 2\pi\mathbb{Z}$.

Dans chacun des produits décrits, on fait apparaitre n racines de $X^n - 1$ deux à deux distinctes puisque les congruences modulo n de k parcourent exactement $\{0, \ldots, n-1\}$. Ce sont donc n racines distinctes, donc les polynômes $X - e^{\frac{i2k\pi}{n}}$ sont des facteurs irréductibles de $X^n - 1$ deux à deux premiers entre eux. Il n'y en a pas d'autres car le degré de $X^n - 1$ est n. □

Les nombres $e^{\frac{i2k\pi}{n}}$ pour $k \in \mathbb{Z}$ sont appelés racines $n^{\text{ième}}$ de l'unité. Ils dépendent de k seulement modulo n. Il est intéressant d'observer que si l'on trace les racines de l'unité dans le plan complexe, elles forment les sommets d'un polygone régulier à n sommet. De plus, pour $k = 0$, on obtient que 1 est une racine de l'unité (et pour $k = \frac{n}{2}$ on obtient que -1 est racine de l'unité si et seulement si n est pair).

Corollaire 4.5.8. *Soit $n \geq 1$ et $a \in \mathbb{C}$. Si $z_0 \in \mathbb{C}$ est tel que $z_0^n = 1$, alors*

$$X^n - a \;=\; \prod_{k=0}^{n-1}(X - z_0 e^{\frac{i2k\pi}{n}})$$

Pour trouver une telle solution particulière z_0, il suffit de la chercher sous forme polaire $z_0 = \rho e^{i\theta}$. En effet, si $a = \rho_0 e^{i\theta_0}$, alors $z_0^n = (\rho e^{i\theta})^n = \rho^n e^{in\theta}$, qui est égal à a si et seulement si $\rho^n = \rho_0$ et $n\theta = \theta_0 \text{mod} 2\pi$. Il suffit de prendre $\rho = \rho_0^{\frac{1}{n}}$ et $\theta = \frac{\theta_0}{n}$.

Démonstration. Les $z_0 e^{\frac{i2k\pi}{n}}$ forment bien n racines deux à deux distinctes du polynôme $X^n - a$ de degré n. □

Corollaire 4.5.9. *Soit $n \geq 1$, la décomposition de $X^n - 1$ en facteurs irréductibles de $\mathbb{R}[X]$ est la suivante :*

$$X^n - 1 \;=\; \begin{cases} (X-1)(X+1)\prod_{k=1}^{\frac{n}{2}-1}(X^2 - 2\cos(\frac{2k\pi}{n})X + 1) & \textit{si } n \textit{ pair,} \\ (X-1)\prod_{k=1}^{\frac{n-1}{2}}(X^2 - 2\cos(\frac{2k\pi}{n})X + 1) & \textit{si } n \textit{ impair.} \end{cases}$$

Démonstration. Il suffit de regrouper les racines complexes conjuguées dans la proposition 4.5.7 □

Exercice 39. Factoriser dans $\mathbb{R}[X]$ les polynômes $X^6 + 8$ et $X^7 - 2$.

Supposons que nous sommes capables de factoriser dans $\mathbb{C}$ le polynôme $P \in \mathbb{C}[X]$. Alors le corollaire 4.5.8 nous permet de factoriser le polynôme $P(X^n)$ pour tout $n \geq 1$. En effet, il suffit de factoriser $X^n - a$ pour chacune des racines a de P et de regrouper les termes.

Pour les polynômes de degré 2, 3 et 4, on dispose de formules explicites (bien connues pour 2, beaucoup plus compliquées et moins utiles sans ordinateur pour 3 et 4). On peut donc factoriser dans $\mathbb{C}[X]$ tout polynôme de la forme $P(X^n)$ où $\deg(P) \leq 4$.

Par contre, il existe des polynômes de degré 5 pour lesquels aucune formule algèbrique (utilisant les opérations usuelles et les racines $n^{\text{ième}}$) n'existe. Ce résultat spectaculaire est dû au mathématicien Évariste Galois (1811-1832). Il fait l'objet du cours de master de théorie de Galois. L'idée principale est d'établir une correspondance entre les formules racines $n^{\text{ième}}$, les extensions de corps, et des structures de groupes finis. C'est d'ailleurs dans les travaux de Galois qu'apparait pour la première fois la notion de groupe vue au chapitre 0.

Polynômes de degré 2

On rappelle ici l'origine des formules pour les racines des polynômes de degré 2.

Soient $a, b, c \in \mathbb{C}$. On suppose $a \neq 0$. On observe que

$$aX^2 + bX + c = a\left(\left(X + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a^2}\right) = a\left(\left(X + \frac{b}{2a}\right)^2 - \frac{\Delta}{4a^2}\right)$$

où l'on pose $\Delta = b^2 - 4ac$. Notez que Δ, ou plutôt $\frac{\Delta}{4a^2}$ mesure "à quel point $aX^2 + bX + c$ n'est pas un carré de polynôme de degré 1".

On pose alors $Y = X + \frac{b}{2a}$, et résoudre $aX^2 + bX + c = 0$ équivaut à résoudre $Y^2 - \frac{\Delta}{4a^2} = 0.$, ce que l'on peut faire grâce au corollaire 4.5.8. En effet, si on note δ_1 et $\delta_2 = -\delta_1$ les deux racines de Δ, alors les racines de $\frac{\Delta}{4a^2}$ sont $\frac{\delta_1}{2a}$ et $\frac{\delta_2}{2a} = \frac{-\delta_1}{2a}$. On obtient

$$Y^2 - \frac{\Delta}{4a^2} = \left(Y - \frac{\delta_1}{2a}\right)\left(Y - \frac{\delta_2}{2a}\right) = \left(Y - \frac{\delta_1}{2a}\right)\left(Y + \frac{\delta_1}{2a}\right).$$

Il s'ensuit en revenant à la variable X que

$$aX^2 + bX + c = a\left(X + \frac{b}{2a} - \frac{\delta_1}{2a}\right)\left(X + \frac{b}{2a} + \frac{\delta_1}{2a}\right) = a\left(X - \frac{-b - \delta_1}{2a}\right)\left(X - \frac{-b + \delta_1}{2a}\right).$$

On retrouve la formule bien connue depuis la classe de seconde.

Notez que si a, b, c sont réels, alors l'équation $Y^2 - \frac{\Delta}{4a^2} = 0$ admet des solutions réelles si et seulement si $\Delta \geq 0$.

4.6 Relations racines-coefficients d'un polynôme

NB : Cette partie ne figure pas au programme de l'examen lors de l'année universitaire 2019-2020.

On considère un polynôme $P = \sum_{j=0}^{n} c_j X^j \in \mathbb{K}[X]$ de degré n scindé, c'est-à-dire tel qu'il existe $x_1, \ldots, x_n \in \mathbb{K}$ avec

$$P = \sum_{j=0}^{n} c_j X^j = c_n \prod_{i=1}^{n} (X - x_i).$$

En développant le produit, on obtient les *relations racines-coefficients* suivantes, dites aussi relations de Viète, pour $1 \leq k \leq n$:

$$(-1)^k \frac{c_{n-k}}{c_n} = \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} x_{i_1} x_{i_2} \ldots x_{i_k}. \tag{4.4}$$

Les deux cas particulier suivants sont à connaitre. Pour $k = 1$,

$$-\frac{c_{n-1}}{c_n} = \sum_{i=1}^{n} x_i,$$

cela assure que si P unitaire ($c_n = 1$), alors la somme des racines est l'opposé du coefficient de $X^{\deg(P)-1}$.

Pour $k = n$, on obtient que le coefficient constant est au signe près le produit des racines :

$$(-1)^n \frac{c_0}{c_n} = x_1 x_2 \ldots x_n.$$

Il faut aussi être capable de retrouver ces formules pour de petites valeurs de n.
Quand $n = 2$, on note $P = aX^2 + bX + c = a(X - x_1)(X - x_2)$. On obtient

$$\begin{aligned} -\frac{b}{a} &= x_1 + x_2 \\ \frac{c}{a} &= x_1 x_2 \end{aligned}$$

Quand $n = 3$, on note $P = aX^3 + bX^2 + cX + d = a(X - x_1)(X - x_2)(X - x_3)$. On obtient

$$\begin{aligned} -\frac{b}{a} &= x_1 + x_2 + x_3 \\ \frac{c}{a} &= x_1 x_2 + x_1 x_3 + x_2 x_3 \\ -\frac{d}{a} &= x_1 x_2 x_3. \end{aligned}$$

Les lecteurs.trices sont invité.e.s à écrire les relations pour $n = 4$.

Exercice 40. Trouver $\lambda \in \mathbb{R}$ tel que le polynôme $X^3 - 7X + \lambda$ possède deux racines dont l'une est le double de l'autre.

Définition 4.6.1. *Soit $Q = Q(X_1, \ldots, X_n) \in \mathbb{K}[X_1, X_2, \ldots, X_n]$ un polynôme en plusieurs indéterminées. On dit que Q est un polynôme symétrique si on a*

$$\forall \sigma \in \mathcal{S}_n, Q(X_{\sigma(1)}, \ldots, X_{\sigma(n)}) = Q(X_1, \ldots, X_n).$$

Exercice 41. Montrer que l'ensemble des polynômes symétriques est une sous-$\mathbb{K}$-algèbre de la $\mathbb{K}$-algèbre $\mathbb{K}[X_1, X_2, \ldots, X_n]$ des polynômes en n indéterminées.

À n fixé, les polynômes du membre de droite de (4.4)

$$E_k(X_1, \ldots, X_n) := \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} X_{i_1} X_{i_2} \ldots X_{i_k}$$

sont des polynômes symétriques puisque $P = c_n \prod_{i=1}^n (X - x_i) = c_n \prod_{i=1}^n (X - x_{\sigma(i)})$ pour tout $\sigma \in \mathcal{S}_n$. En fait, ces polynômes symétriques "élémentaires" permettent d'engendrer tous les polynômes symétriques, comme l'assure le théorème suivant.

Théorème 4.6.2. *Soit Q un polynôme symétrique dans $\mathbb{K}[X_1, X_2, \ldots, X_n]$, alors il existe un polynôme R en n indéterminées tel que*

$$Q(X_1, \ldots, X_n) = R(E_1, \ldots, E_n) = R\left(E_1(X_1, \ldots, X_n), \ldots, E_n(X_1, \ldots, X_n)\right).$$

La preuve de cet énoncé dépasse le cadre de ce cours. Son utilité est la suivante : si on dispose d'un polynôme symétrique Q en n variables, et qu'on souhaite l'évaluer sur les racines, disons complexes, d'un polynôme P de degré n. Alors il suffit d'évaluer le polynôme R correspondant, aussi en n variables, sur les **coefficients** normalisés de P. En particulier, il n'y a pas besoin de factoriser P.

Exercice 42. Calculer $P(x_1, x_2) = (x_1 x_2)^3 + 5(x_1 + x_2)^2$ pour x_1, x_2 les racines du polynôme $X^2 - 3X + 1$.

4.7 Preuve du Théorème 4.5.1 de d'Alembert-Gauss (Hors-programme)

Dans cette partie, on démontre le Théorème 4.5.1 de d'Alembert-Gauss. On considère un polynôme complexe $P \in \mathbb{C}[X]$ de degré ≥ 1, et on doit montrer qu'il admet une racine. On doit donc trouver $z_0 \in \mathbb{C}$ tel que $P(z_0) = 0$.

Il est clair que multiplier P par une constante non-nulle ne change pas cette propriété, donc on peut supposer que P est unitaire et on note $P = a_0 + a_1 X + \cdots + a_{n-1} X^{n-1} + X^n$.

La preuve est largement basée sur des outils analytiques, et donc hors du cadre de ce cours. On la donne ici par souci de complétude pour les étudiant.e.s interessé.e.s.

On observe tout d'abord que la fonction $\mathbb{C} \to \mathbb{R}$ donnée par $z \mapsto |P(z)|$ est continue sur $\mathbb{C}$, comme composée d'une application polynomiale par la norme complexe qui est continue. On

rappelle qu'une fonction $f : \mathbb{C} \to \mathbb{C}$ est continue en z_0 si $\forall \varepsilon > 0, \exists \delta > 0$, $|z - z_0| < \delta$ implique $|f(z) - f(z_0)| < \varepsilon$. C'est la même définition que pour des fonctions réelles, où la norme remplace la valeur absolue.

On observe ensuite que cette fonction $|P|$ admet un minimum dans $\mathbb{C}$. Comme la partie $A_1 = \{|P(z)| : z \in \mathbb{C}\} \subset \mathbb{R}$ est non-vide et minorée, il est clair qu'elle admet une borne inférieure. Reste à voir qu'elle est atteinte. On note alors que

$$|P(z)| = |z|^n \left(1 + \frac{|a_{n-1}|}{|z|} + \cdots + \frac{|a_0|}{|z|^n}\right) \underset{|z|\to+\infty}{\longrightarrow} +\infty$$

car chacune des fractions tend vers 0. Cela signifie que $\forall M > 0, \exists R > 0$, si $|z| > R$, alors $|P(z)| > M$. En particulier, on l'utilise pour $M_0 = \inf(A_1)+1$, et on obtient R_0 tel que $|z| > R_0$ implique $|P(z)| > M_0 = \inf(A_1) + 1$. Il s'ensuit que $\inf(A_1) = \inf(A_2)$ où $A_2 = \{|P(z)| : |z| \leq R_0\}$.

On considère maintenant une suite $(z_n)_{n\in\mathbb{N}}$ de nombres complexes tels que $|z_n| \leq R_0$ et $|P(z_n)| \to \inf(A_2) = \inf(A_1)$. Cette suite existe par définition de la borne inférieure. Chaque nombre complexe z_n s'écrit de manière unique $z_n = x_n + iy_n$ avec x_n et y_n réels, tous deux dans l'intervalle $[-R_0, R_0]$.

Le fameux théorème de Bolzano-Weierstrass assure alors qu'il existe une extraction φ_1 telle que $(x_{\varphi_1(n)})$ converge vers un réel x_∞ dans $[-R_0, R_0]$. Une seconde application de ce théorème, à la suite $(y_{\varphi_1(n)})$ assure maintenant l'existence d'une extraction φ_2 telle que la sous-suite $(y_{\varphi_1(\varphi_2(n))})$ converge vers un réel $y_\infty \in [-R_0, R_0]$. On pose alors $\varphi = \varphi_1 \circ \varphi_2$, et on a bien $x_{\varphi(n)} \to x_\infty$ et $y_{\varphi(n)} \to y_\infty$. Il s'ensuit que $z_{\varphi(n)} \to z_0 := x_\infty + iy_\infty$. Par continuité de $|P|$, on conclut que

$$|P(z_{\varphi(n)}| \to |P(z_0)| = \inf(A_2) = \inf(A_1) = \inf\{|P(z)| : z \in \mathbb{C}\}.$$

Ainsi, la fonction $|P|$ admet un minimum en z_0. Jusque là, la preuve utilise des arguments vus lors du cours d'analyse réelle de première année. La suite est plus originale. Elle relève de l'analyse complexe qui sera étudiée lors des semestres ultérieurs.

On veut montrer que $P(z_0) = 0$. On procède par l'absurde, et on suppose que $P(z_0) \neq 0$. On pose alors

$$Q(X) = \frac{P(z_0 + X)}{P(z_0)} = \sum_{i=0}^{n} b_i X^i.$$

On observe que $b_0 = 0$. On va montrer qu'il existe $z \in \mathbb{C}$ tel que $|Q(z)| < 1$. Cela terminera la preuve en fournissant une contradiction puisque $|P(z_0)| = \inf\{|P(z)| : z \in \mathbb{C}\}$ assure que $\forall z \in \mathbb{C}, |Q(z)| \geq 1$.

On considère $k = \min\{1 \leq i \leq n | b_i \neq 0\}$, qui existe puisque $\deg(P) \geq 1$. On a alors

$$Q(z) = 1 + b_k z^k (1 + \phi(z)), \text{ où } \phi(z) = \sum_{i=1}^{n-k} \frac{b_{i+k}}{b_k} z^i \underset{|z|\to 0}{\longrightarrow} 0.$$

Soit $r > 0$ tel que $|\phi(z)| < \frac{1}{2}$ dés que $|z| < r$.

On note sous forme polaire $b_k = |b_k|e^{i\theta}$. On considère alors $z := \rho e^{i\frac{\theta+\pi}{k}}$ avec $0 < \rho < \min\left(r, \frac{2}{3}\left(\frac{1}{|b_k|}\right)^{\frac{1}{k}}\right)$, et on calcule :

$$\begin{aligned} Q(z) &= 1 + b_k\left(\rho e^{i\frac{\theta+\pi}{k}}\right)^k (1+\phi(z)) \\ &= 1 - |b_k|\,\rho^k\,(1+\phi(z))\,. \end{aligned}$$

On rappelle l'inégalité triangulaire $|a-b| \leq ||a|-|b||$ pour tous $a, b \in \mathbb{C}$ et que $|b_k\rho^k| < \frac{2}{3}$ par choix de ρ et que $|1+\phi(z)| < \frac{3}{2}$ par choix de r. On déduit donc que $0 < \left||b_k|\rho^k\,(1+\phi(z))\right| < 1$ et

$$|Q(z)| \leq 1 - \left||b_k|\rho^k\,(1+\phi(z))\right| < 1.$$

En fait, on a fait un développement limité complexe $Q(z) = 1 + b_k z^k + o_{z\to 0}(z^k)$ et utilisé le fait que la fonction z^k prend des valeurs négative dans tout voisinage complexe de 0. L'argument montre plus généralement que les fonctions holomorphes n'admettent pas d'extremum local (à revoir plus tard).

Chapitre 5

Polynômes d'endomorphismes

5.1 Avant-propos : caractérisation de la trigonalisabilité au moyen du polynôme caractéristique

Théorème 5.1.1. *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et soit $\varphi \in \mathcal{L}(E)$. Les assertions suivantes sont équivalentes :*

(a) φ est trigonalisable

(b) le polynôme caractéristique $P_\varphi(X)$ est scindé dans $\mathbb{K}[X]$.

On rappelle que $P_\varphi(X)$ est scindé si et seulement si sa décomposition en facteurs irréductibles ne comporte que des facteurs de degré 1, c'est-à-dire si il existe $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ et $m_1, \ldots, m_k \in \mathbb{N}$ tels que

$$P_\varphi(X) = \prod_{i=1}^{k} (\lambda_i - X)^{m_i} \tag{5.1}$$

Par la suite on utilisera la notation (5.1) en supposant que les λ_i sont **deux-à-deux distincts**, lorsque le polynôme caractéristique sera scindé.

Corollaire 5.1.2. *Soit E un $\mathbb{C}$-espace vectoriel et $\varphi \in \mathcal{L}(E)$, alors φ est trigonalisable.*

Démonstration. Le théorème 4.5.1 de d'Alembert-Gauss assure que tout polynôme de $\mathbb{C}[X]$ est scindé. □

Notez bien que ce corollaire n'est valide que dans $\mathbb{C}$, et pas dans $\mathbb{R}$. Toutefois, si l'on dispose d'un espace vectoriel réel, on peut toujours le plonger dans un espace vectoriel complexe de même dimension, et donc trigonaliser dans l'espace complexifier. En ce qui concerne les matrices, c'est encore plus clair, puisque $\mathrm{M}_n(\mathbb{R}) \subset M_n(\mathbb{C})$.

Le corollaire 5.1.2 est valide dans tout corps où l'on dispose d'un théorème de d'Alembert-Gauss. De tels corps sont dits *algébriquement clos*. On verra au cours de M1 de théorie des corps que tout corps $\mathbb{K}$ peut s'étendre en un corps algébriquement clos, mais que celui-ci peut être beaucoup plus grand.

Démonstration du Théorèm 5.1.1. On montre d'abord que si φ est trigonalisable, alors son polynôme caractéristique est scindé.

Supposons donc que la matrice de φ dans une base $\mathcal{B}'$ soit triangulaire supérieure, et notons-la

$$\mathrm{Mat}_{\mathcal{B}'}(\varphi) = \begin{pmatrix} d_1 & * & \cdots & * & * \\ 0 & d_2 & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & d_{n-1} & * \\ 0 & 0 & \cdots & 0 & d_n \end{pmatrix}$$

Alors le polynôme caractéristique est $P_\varphi(X) = \det(\varphi - X\mathrm{id}_E)$. Pour le calculer, on exprime la matrice de l'endomorphisme $\varphi - X\mathrm{id}_E$ dans la base de notre choix, puisque le résultat est indépendant de la base choisie d'après la proposition 3.1.12. On choisit la base $\mathcal{B}'$, et on a

$$P_\varphi(X) = \det\left(\mathrm{Mat}_{\mathcal{B}'}(\varphi) - XI_n\right) = \det \begin{pmatrix} d_1 - X & * & \cdots & * & * \\ 0 & d_2 - X & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & d_{n-1} - X & * \\ 0 & 0 & \cdots & 0 & d_n - X \end{pmatrix} = \prod_{j=1}^{n} (d_j - X)$$

calculé facilement car la matrice est triangulaire. On conclut que $P_\varphi(X)$ est scindé.

On démontre maintenant la réciproque par récurrence sur $n = \dim(E)$. Pour $n = 1$, il n'y a rien à démontrer. Supposons le résultat vrai jusquà $n-1$ et que $P_\varphi(X) = \prod_{i=1}^{n}(d_i - X)$ est scindé. Soit v_1 un vecteur propre pour la valeur propre d_1, alors $\varphi(v_1) = d_1 v_1$. On complète pour obtenir une base $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ de E. Alors

$$\mathrm{Mat}_{\mathcal{B}}(\varphi) = \left(\begin{array}{c|c} d_1 & L \\ \hline 0 & S \end{array}\right)$$

est une matrice par blocs où L est une ligne de longueur $n-1$ et 0 une colonne de $n-1$ zéros et $S \in \mathrm{M}_{n-1}(\mathbb{K})$.

L'hypothèse de récurrence appliqué à l'endomorphisme φ_S de $\mathbb{K}^{n-1}$ canoniquement associé à S assure l'existence d'une matrice $Q \in \mathrm{GL}_{n-1}(\mathbb{K})$ telle que $Q^{-1}SQ$ soit triangulaire supérieure. On pose alors

$$P = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q \end{array}\right) \in \mathrm{GL}_n(\mathbb{K})$$

On a

$$P^{-1}\mathrm{Mat}_{\mathcal{B}}(\varphi)P = \left(\begin{array}{c|c} d_1 & LQ \\ \hline 0 & Q^{-1}SQ \end{array}\right) \quad \text{triangulaire supérieure.}$$

De plus, $P^{-1}\mathrm{Mat}_{\mathcal{B}}(\varphi)P$ est la matrice de φ dans la base $\mathcal{B}' = (u_1, u_2, \ldots, u_n)$ où les vecteurs u_j sont donnés par les colonnes de P comme suit : la $j^{\text{ième}}$ colonne est constituée des coordonnées de u_j dans la base $\mathcal{B} = (v_1, \ldots, v_n)$. En particulier, $u_1 = v_1$. □

Notez la différence de notation entre la preuve ci-dessus et (5.1). Dans le premier cas, on regroupe les racines multiples, dans le deuxième cas on ne le fait pas. Les λ_i sont deux-à-deux distincts, pas les d_j.

Proposition 5.1.3. *Soit $\varphi \in \mathcal{L}(E)$, soit $m(\lambda)$ la multiplicité de λ comme racine du polynôme caractéristique $P_\varphi(X)$. On a*

$$\dim(E_\lambda) \leq m(\lambda).$$

Démonstration. Soit $(v_1, \ldots, v_k)$ une base de E_λ. En particulier $\dim(E_\lambda) = k$. On la complète pour obtenir une base $\mathcal{B}' = (v_1, \ldots, v_n)$ de E. Alors puisque $\forall 1 \leq i \leq k, \varphi(v_i) = \lambda v_i$, on a

$$\mathrm{Mat}_{\mathcal{B}'}(\varphi) = \left(\begin{array}{ccc|ccc} \lambda & \cdots & 0 & * & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda & * & \cdots & * \\ \hline 0 & \cdots & 0 & * & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & * & \cdots & * \end{array}\right) = \left(\begin{array}{c|c} \lambda I_k & B \\ \hline 0 & C \end{array}\right).$$

Donc $P_\varphi(X) = (\lambda - X)^k P_C(X)$. Comme $(\lambda - X)^k$ divise $P_\varphi(X)$, la multiplicité de λ est supérieure ou égale à k. □

5.2 Algèbre engendrée par un endomorphisme

Définition 5.2.1. *Soit E un $\mathbb{K}$-espace vectoriel et $\varphi \in \mathcal{L}(E)$ un endomorphisme. On définit l'application* évaluation en φ

$$\begin{array}{rccc} \mathrm{ev}_\varphi : & \mathbb{K}[X] & \to & \mathcal{L}(E) \\ & Q & \mapsto & Q(\varphi) \end{array}$$

donnée par $Q(\varphi) = \sum_{\ell=0}^d c_\ell \varphi^\ell$ si $Q = \sum_{\ell=0}^d c_\ell X^\ell$. On rappelle $\varphi^0 = \mathrm{id}$.

Cette application est bien définie puisque les composées φ^ℓ sont encore des endomorphismes, dont on peut faire des combinaisons linéaires. Comme toujours, il y a un pendant matriciel, peut-être plus naturel :

Définition 5.2.2. *Soit $A \in \mathrm{M}_n(\mathbb{K})$, on définit l'application* évaluation en A

$$\begin{array}{rccc} \mathrm{ev}_A : & \mathbb{K}[X] & \to & \mathrm{M}_n(\mathbb{K}) \\ & Q & \mapsto & Q(A) \end{array}$$

donnée par $Q(A) = \sum_{\ell=0}^d c_\ell A^\ell$ si $Q = \sum_{\ell=0}^d c_\ell X^\ell$. On rappelle $A^0 = I_n$.

En d'autres termes, on applique le polynôme Q à φ ou à A. Ces applications respectent les structures algébriques, c'est-à-dire les opérations, au sens suivant :

Proposition 5.2.3. *L'application* $\mathrm{ev}_\varphi : \mathbb{K}[X] \to \mathcal{L}(E)$ *est un morphisme de* $\mathbb{K}$*-algèbres entre* $(\mathbb{K}[X], +, \times, \cdot)$ *et* $(\mathcal{L}(E), +, \circ, \cdot)$.

L'application $\mathrm{ev}_A : \mathbb{K}[X] \to \mathrm{M}_n(\mathbb{K})$ *est un morphisme de* $\mathbb{K}$*-algèbres entre* $(\mathbb{K}[X], +, \times, \cdot)$ *et* $(\mathrm{M}_n(\mathbb{K}), +, \times, \cdot)$.

Démonstration. Il s'agit de vérifier que $\forall \lambda_1, \lambda_2 \in \mathbb{K}, \forall Q_1, Q_2 \in \mathbb{K}[X]$,

$$\mathrm{ev}_\varphi(\lambda_1 Q_1 + \lambda_2 Q_2) = \lambda_1 \mathrm{ev}_\varphi(Q_1) + \lambda_2 \mathrm{ev}_\varphi(Q_2) \quad \text{et} \quad \mathrm{ev}_\varphi(Q_1 Q_2) = \mathrm{ev}_\varphi(Q_1) \circ \mathrm{ev}_\varphi(Q_2)$$

et similairement

$$\mathrm{ev}_A(\lambda_1 Q_1 + \lambda_2 Q_2) = \lambda_1 \mathrm{ev}_A(Q_1) + \lambda_2 \mathrm{ev}_A(Q_2) \quad \text{et} \quad \mathrm{ev}_A(Q_1 Q_2) = \mathrm{ev}_A(Q_1) \times \mathrm{ev}_A(Q_2).$$

□

Exemple 5.2.4. Prenons $P = X^2 + 2$ et $Q = X^3 + X - 1$, alors $P(\varphi) = \varphi^2 + 2\mathrm{id}$ et $Q(\varphi) = \varphi^3 + \varphi - \mathrm{id}$ et $PQ = X^5 + 3X^3 - X^2 + 2X - 2$ (vérifiez-le !). D'autre part :

$$\begin{aligned}
\forall v \in E, P(\varphi) \circ Q(\varphi)(v) &= (\varphi^2 + 2\mathrm{id}) \circ (\varphi^3 + \varphi - \mathrm{id})(v) \\
&= (\varphi^2 + 2\mathrm{id})(\varphi^3(v) + \varphi(v) - v) \\
&= \varphi^2(\varphi^3(v) + \varphi(v) - v) + 2(\varphi^3(v) + \varphi(v) - v) \\
&= \varphi^5(v) + \varphi^3(v) - \varphi^2(v) + 2\varphi^3(v) + 2\varphi(v) - 2v \\
&= \varphi^5(v) + 3\varphi^3(v) - \varphi^2(v) + 2\varphi(v) - 2v \\
&= (\varphi^5 + 3\varphi^3 - \varphi^2 + 2\varphi - 2\mathrm{id})(v) = (PQ)(\varphi)(v).
\end{aligned}$$

5.3 Le polynôme minimal

On note $\mathbb{K}[\varphi] \subset \mathcal{L}(E)$ l'image du morphisme ev_φ et $\mathbb{K}[A] \subset \mathrm{M}_n(\mathbb{K})$ l'image de ev_A. On va utiliser l'exercice suivant :

Exercice 43. Soit $f : \mathcal{A}_1 \to \mathcal{A}_2$ un morphisme de $\mathbb{K}$-algèbres, alors le noyau $\ker(f) := \{x \in \mathcal{A}_1 : f(x) = 0\}$ est un idéal de $\mathcal{A}_1$.

Ainsi $\ker(\mathrm{ev}_\varphi)$ est un idéal de $\mathbb{K}[X]$. Cet anneau étant principal, cf proposition 4.3.9, l'idéal $\ker(\mathrm{ev}_\varphi)$ est engendré par un unique polynôme unitaire noté Π_φ, appelé polynôme minimal de φ. On note $\ker(\mathrm{ev}_\varphi) = \Pi_\varphi \mathbb{K}[X]$.

Les éléments du noyau $\ker(\mathrm{ev}_\varphi)$ sont les polynômes Q tels que $Q(\varphi) = 0$. On les appelle *polynômes annulateurs* de φ.

Définition 5.3.1. *Le* polynôme minimal Π_φ *de l'endomorphisme* $\varphi \in \mathcal{L}(E)$ *est l'unique générateur unitaire de l'idéal annulateur* $\ker(\mathrm{ev}_\varphi) = \{Q \in \mathbb{K}[X] : Q(\varphi) = 0\}$. *C'est aussi le polynôme annulateur non-nul de* φ *de degré minimal.*

Pour $A \in \mathrm{M}_n(\mathbb{K})$, on définit aussi Π_A comme le générateur unitaire de l'idéal annulateur $\ker(\mathrm{ev}_A) = \{Q \in \mathbb{K}[X] : Q(A) = 0\}$. C'est aussi le polynôme annulateur non-nul de A de degré minimal.

Fait 5.3.2. *Si $A = \mathrm{Mat}_{\mathcal{B}}(\varphi)$ dans une base $\mathcal{B}$ de E, alors $\Pi_\varphi = \Pi_A$. En particulier, si $A \in \mathrm{M}_n(\mathbb{K})$ et $P \in \mathrm{GL}_n(\mathbb{K})$, alors $\Pi_{P^{-1}AP} = \Pi_A$.*

Démonstration. Le choix d'une base $\mathcal{B}$ de E induit un isomorphisme de $\mathbb{K}$-algèbres

$$\begin{array}{rccc} \mathrm{Mat}_{\mathcal{B}} : & \mathcal{L}(E) & \to & \mathrm{M}_n(\mathbb{K}) \\ & \varphi & \mapsto & \mathrm{Mat}_{\mathcal{B}}(\varphi) \end{array}$$

entre $(\mathcal{L}(E), +, \circ, \cdot)$ et $(\mathrm{M}_n(\mathbb{K}), +, \times, \cdot)$. En particulier pour tout polynôme $Q \in \mathbb{K}[X]$, on a $Q(\varphi) = 0$ si et seulement si $Q(\mathrm{Mat}_{\mathcal{B}}(\varphi)) = 0$, donc les idéaux $\ker(\mathrm{ev}_\varphi)$ et $\ker(\mathrm{ev}_{\mathrm{Mat}_{\mathcal{B}}(\varphi)})$ coïncident. En particulier, le polynôme minimal ne dépend pas de la base choisie pour calculer la matrice.

La deuxième partie en découle. En effet, avec les notations usuelles, on a $\mathrm{Mat}_{\mathcal{B}_{\mathrm{can}}}(\varphi_A) = A$ et $P^{-1}AP = \mathrm{Mat}_{\mathcal{B}'}(\varphi_A)$ où $\mathcal{B}'$ est la base de $\mathbb{K}^n$ formées par les vecteurs colonnes de P. □

Proposition 5.3.3. *L'application quotient $\widetilde{\mathrm{ev}}_\varphi : \mathbb{K}[X]/\Pi_\varphi\mathbb{K}[X] \to \mathbb{K}[\varphi]$ induit un isomorphisme de $\mathbb{K}$-algèbres.*

Ceci implique en particulier que la sous-algèbre $\mathbb{K}[\varphi]$ de $\mathcal{L}(E)$ est commutative (puisque isomorphe au quotient encore commutatif d'une algèbre commutative), alors que $\mathcal{L}(E)$ ne l'est pas. De même la sous-algèbre $\mathbb{K}[A]$ est commutative, mais pas $\mathrm{M}_n(\mathbb{K})$.

Démonstration. On rappelle que si I est un idéal de la $\mathbb{K}$-algèbre $\mathcal{A}$, alors l'algèbre quotient $\mathcal{A}/I$ est l'ensemble des parties $a + I$ où $a \in \mathcal{A}$ munie des opérations $(a + I) + (b + I) = (a + b) + I$, $(a + I)(b + I) = ab + I$ et $\lambda(a + I) = \lambda a + I$ pour tous $a, b \in \mathcal{A}$ et $\lambda \in \mathbb{K}$. Ces opérations ne dépendent pas du choix de a ou a' lorsque $a + I = a' + I$ (exercice).

Tout morphisme de $\mathbb{K}$-algèbre devient injectif par passage au quotient par son noyau, c'est-à-dire que si $f : \mathcal{A}_1 \to \mathcal{A}_2$ est un morphisme de $\mathbb{K}$-algèbres, alors $\widetilde{f} : \mathcal{A}_1/\ker(f) \to \mathcal{A}_2$ est un morphisme injectif (exercice).

Reste à voir la surjectivité de ev_φ qui provient de la définition de $\mathbb{K}[\varphi]$. □

Corollaire 5.3.4. *Soit $\varphi \in \mathcal{L}(E)$ et soit $Q \in \mathbb{K}[X]$. Si P est annulateur de φ, i.e. si $Q(\varphi) = 0$, alors le polynôme minimal Π_φ divise Q.*

Démonstration. On a $Q \in \ker(\mathrm{ev}_\varphi) = \Pi_\varphi\mathbb{K}[X]$. □

Exercice 44. Démontrer directement le corollaire 5.3.4 en utilisant que Π_φ est un polynôme annulateur de degré minimal et la division euclidienne de Q par Π_φ.

5.3.1 Une application

Une utilisation du polynôme minimal est la suivante : étant donné un polynôme $Q \in \mathbb{K}[X]$, on souhaite calculer l'endomorphisme (ou sa matrice) $Q(\varphi)$. On effectue la division euclidienne de Q par Π_φ, notée $Q = Q_1\Pi_\varphi + R$ avec $\deg(R) < \deg(\Pi_\varphi)$, alors

$$Q(\varphi) = Q_1(\varphi) \circ \Pi_\varphi(\varphi) + R(\varphi) = Q_1(0) + R(\varphi) = R(\varphi).$$

Il suffit donc d'évaluer φ sur le reste, dont le degré est toujours $< \deg(\Pi_\varphi)$. C'est très utile lorsque le degré de Q est grand.

Il s'avère donc très utile de connaitre le polynôme minimal d'un endomorphisme ou d'une matrice.

5.4 Théorème de Cayley-Hamilton

Théorème 5.4.1. *[Théorème de Cayley-Hamilton] Soit $\varphi \in \mathcal{L}(E)$. Le polynôme caractéristique $P_\varphi(X)$ est annulateur de φ, c'est-à-dire $P_\varphi(\varphi) = 0$.*

Ce théorème est équivalent au fait que le polynôme minimal Π_φ divise le polynôme caractéristique P_φ.La preuve utilise le :

Lemme 5.4.2. *[Lemme des matrices compagnons] Soit $n \geq 2$, et $a_0, a_1, \ldots, a_{n-1} \in \mathbb{K}$. On considère la matrice compagnon*

$$C = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & 0 & -a_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 & 0 & -a_{n-3} \\ 0 & 0 & 0 & \cdots & 1 & 0 & -a_{n-2} \\ 0 & 0 & 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix} \in \mathrm{M}_n(\mathbb{K}).$$

Alors son polynôme caractéristique est

$$P_C(X) = (-1)^n(X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0).$$

En particulier, si on se donne un polynôme unitaire de degré n dans $\mathbb{K}[X]$, il existe toujours une matrice $C \in \mathrm{M}_n(\mathbb{K})$ dont il est le polynôme caractéristique (on peut changer le signe de la dernière colonne quand n est impair). On dit alors que c'est la matrice compagnon du polynôme.

Démonstration. On montre d'abord par récurrence sur n que

$$\det \begin{pmatrix} -X & 0 & 0 & \cdots & 0 & 0 & -a_0 \\ 1 & -X & 0 & \cdots & 0 & 0 & -a_1 \\ 0 & 1 & -X & \cdots & 0 & 0 & -a_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & -X & 0 & -a_{n-3} \\ 0 & 0 & 0 & \cdots & 1 & -X & -a_{n-2} \\ 0 & 0 & 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix} = (-1)^n(a_{n-1}X^{n-1}+a_{n-2}X^{n-2}+\cdots+a_1X+a_0).$$

C'est immédiat si $n = 2$. Supposons avoir montré le résultat pour $n-1$ et considérons C de taille n. On développe alors par rapport à la dernière ligne pour obtenir :

$$-a_{n-1}\det\begin{pmatrix} -X & 0 & 0 & \cdots & 0 \\ 1 & -X & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \ddots & -X & 0 \\ 0 & 0 & \cdots & 1 & -X \end{pmatrix} - \det\begin{pmatrix} -X & 0 & 0 & \cdots & -a_0 \\ 1 & -X & 0 & \cdots & -a_1 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \ddots & -X & -a_{n-3} \\ 0 & 0 & \cdots & 1 & -a_{n-2} \end{pmatrix}$$

$$= -a_{n-1}(-X)^{n-1} - (-1)^{n-1}(a_{n-2}X^{n-2} + \cdots + a_1X + a_0),$$

en utilisant l'hypothèse de récurrence. Cela donne le résultat voulu.

On revient alors à $P_C(X)$ calculé en développant $C - XI_n$ par rapport à la dernière ligne. On obtient similairement

$$P_C(X) = -\det\begin{pmatrix} -X & 0 & 0 & \cdots & 0 & 0 & -a_0 \\ 1 & -X & 0 & \cdots & 0 & 0 & -a_1 \\ 0 & 1 & -X & \cdots & 0 & 0 & -a_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & -X & 0 & -a_{n-4} \\ 0 & 0 & 0 & \cdots & 1 & -X & -a_{n-3} \\ 0 & 0 & 0 & \cdots & 0 & 1 & -a_{n-2} \end{pmatrix} - (X + a_{n-1})(-X)^{n-1}$$

qui donne le résultat. □

Démonstration du théorème 5.4.1 de Cayley-Hamilton. Il s'agit de montrer que le polynôme d'endomorphisme $P_\varphi(\varphi)$ est nul dans $\mathcal{L}(E)$. Pour cela, on montre que $\forall x \in E, P_\varphi(\varphi)(x) = 0$.

Soit donc $x \in E$. On considère

$$k := \min\{s \geq 0 : \varphi^s(x) \in \mathrm{vect}(x, \varphi(x), \ldots, \varphi^{s-1}(x))\}.$$

Le minimum existe puisque E est de dimension finie, donc la famille $(x, \varphi(x), \ldots, \varphi^n(x))$ est liée, ayant $\dim(E) + 1$ éléments. Pour ce k minimal, la famille $(x, \varphi(x), \ldots, \varphi^{k-1}(x))$ est libre et il existe $a_0, \ldots, a_{k-1} \in \mathbb{K}$ tels que

$$\varphi^k(x) = a_0x + a_1\varphi(x) + \cdots + a_{k-1}\varphi^{k-1}(x). \tag{5.2}$$

On complète cette famille pour obtenir une base $\mathcal{B} = (x, \varphi(x), \ldots, \varphi^{k-1}(x), v_{k+1} \ldots, v_n)$ de E. Par construction, on a

$$\mathrm{Mat}_{\mathcal{B}}(\varphi) = \left(\begin{array}{|c|c|} \hline C & A \\ \hline 0 & B \\ \hline \end{array}\right)$$

matrice par bloc où C est une matrice compagnon

$$C = \left(\begin{array}{cccccc|c} 0 & 0 & 0 & \cdots & 0 & 0 & a_0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & a_1 \\ 0 & 1 & 0 & \cdots & 0 & 0 & a_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \hline 0 & 0 & 0 & \ddots & 0 & 0 & a_{k-3} \\ 0 & 0 & 0 & \cdots & 1 & 0 & a_{k-2} \\ 0 & 0 & 0 & \cdots & 0 & 1 & a_{k-1} \end{array}\right) \in \mathrm{M}_k(\mathbb{K}).$$

En utilisant le calcul des déterminants de matrices triangulaires par blocs, on a $P_\varphi(X) = P_B(X)P_C(X)$, et donc $P_\varphi(\varphi) = P_B(\varphi) \circ P_C(\varphi)$ qui assure

$$\begin{aligned} P_\varphi(\varphi)(x) &= P_B(\varphi)(P_C(\varphi)(x)) = P_B(\varphi)\left((-1)^k(\varphi^k - a_{k-1}\varphi^{k-1} - \cdots - a_1\varphi - a_0\mathrm{id})(x)\right) \\ &= P_B(\varphi)\left((-1)^k(\varphi^k(x) - a_{k-1}\varphi^{k-1}(x) - \cdots - a_1\varphi(x) - a_0x)\right) = P_B(\varphi)(0) = 0 \end{aligned}$$

grâce au lemme 5.4.2 des matrices compagnons et à l'expression (5.2). □

5.5 Lemme des noyaux

Fait 5.5.1. *Soit $\varphi \in \mathcal{L}(E)$ et $Q \in \mathbb{K}[X]$, alors $\ker(Q(\varphi))$ est stable par φ.*

Démonstration. Soit $x \in \ker(Q(\varphi))$. Alors

$$Q(\varphi)(\varphi(x)) = \varphi \circ Q(\varphi)(x) = \varphi(0) = 0,$$

donc $\varphi(x) \in \ker(Q(\varphi))$.

On rappelle que

$$Q(\varphi) \circ \varphi = \left(\sum_{t=0}^{d} c_t\varphi^t\right) \circ \varphi = \sum_{t=0}^{d} c_t\varphi^{t+1} = \varphi \circ \left(\sum_{t=0}^{d} c_t\varphi^t\right) = \varphi \circ Q(\varphi)$$

□

Lemme 5.5.2. *[Lemme des noyaux] Soit $\varphi \in \mathcal{L}(E)$, et soient $Q_1, Q_2 \in \mathbb{K}[X]$ deux polynômes tels que $\mathrm{pgcd}(Q_1, Q_2) = 1$. On suppose que le polynôme produit Q_1Q_2 est annulateur de φ, c'est-à-dire que $(Q_1Q_2)(\varphi) = 0$. Alors*

(a) $E = \ker(Q_1(\varphi)) \oplus \ker(Q_2(\varphi))$,

(b) $\ker(Q_1(\varphi))$ *et* $\ker(Q_2(\varphi))$ *sont stables par φ,*

(c) *et* $\ker(Q_1(\varphi)) = \mathrm{im}(Q_2(\varphi))$ *et* $\ker(Q_2(\varphi)) = \mathrm{im}(Q_1(\varphi))$.

Démonstration. On utilise l'égalité de Bézout, qui assure l'existence de $U, V \in \mathbb{K}[X]$ tels que $Q_1U+Q_2V = \mathrm{pgcd}(Q_1, Q_2) = 1$. En termes de polynômes de l'endomorphisme φ, cela fournit :

$$\mathrm{id}_E = Q_1(\varphi) \circ U(\varphi) + Q_2(\varphi) \circ V(\varphi). \tag{5.3}$$

On montre d'abord que $\mathrm{im}(Q_1(\varphi)) + \mathrm{im}(Q_2(\varphi)) = E$. On utilise (5.3) qui assure que pour tout $x \in E$, on a

$$x = (Q_1(\varphi))\,(U(\varphi)(x)) + (Q_2(\varphi))\,(V(\varphi)(x)) \in \mathrm{im}(Q_1(\varphi)) + \mathrm{im}(Q_2(\varphi)).$$

D'autre part, on a $(Q_1Q_2)(\varphi) = 0 = Q_1(\varphi) \circ Q_2(\varphi)$. Donc $\forall x \in E$, $Q_1(\varphi) \circ Q_2(\varphi)(x) = 0 = (Q_1(\varphi))(Q_2(\varphi)(x))$, donc $\mathrm{im}(Q_2(\varphi)) \subset \ker(Q_1(\varphi))$. Par symétrie, $\mathrm{im}(Q_1(\varphi)) \subset \ker(Q_2(\varphi))$.

Puisque la somme des images forme déjà l'espace entier, on déduit a fortiori

$$E = \ker(Q_1(\varphi)) + \ker(Q_2(\varphi)).$$

Vérifions que la somme est directe. On peut réécrire l'égalité de Bézout sous la forme

$$\mathrm{id}_E = U(\varphi) \circ Q_1(\varphi) + V(\varphi) \circ Q_2(\varphi).$$

Donc si $x \in \ker(Q_1(\varphi)) \cap \ker(Q_2(\varphi))$, alors

$$x = (U(\varphi))\,(Q_1(\varphi)(x)) + (V(\varphi))\,(Q_2(\varphi)(x)) = (U(\varphi))(0) + (V(\varphi))(0) = 0 + 0 = 0$$

qui prouve $\ker(Q_1(\varphi)) \cap \ker(Q_2(\varphi)) = \{0\}$. A fortiori :

$$E = \ker(Q_1(\varphi)) \oplus \ker(Q_2(\varphi)).$$

Dés lors, si l'inclusion de $\mathrm{im}(Q_2(\varphi))$ dans $\ker(Q_1(\varphi))$ était stricte, alors l'inclusion de $\mathrm{im}(Q_2(\varphi)) + \mathrm{im}(Q_1(\varphi))$ dans E serait stricte aussi. C'est absurde puisque la somme des images est E. On a donc $\mathrm{im}(Q_2(\varphi)) = \ker(Q_1(\varphi))$. Et aussi $\mathrm{im}(Q_1(\varphi)) = \ker(Q_2(\varphi))$ par symétrie.

La stabilité par φ découle du fait 5.5.1. □

5.6 Caractérisation de la diagonalisabilité par le polynôme minimal

Le lemme des noyaux 5.5.2 permet de montrer le :

Théorème 5.6.1. *Un endomorphisme $\varphi \in \mathcal{L}(E)$ est diagonalisable si et seulement si son polynôme minimal est scindé à racines simples.*

La preuve utilise le

Lemme 5.6.2. *Soit $\varphi \in \mathcal{L}(E)$, $Q \in \mathbb{K}[X]$ et $v \in E_\lambda$ un vecteur propre. Alors*

$$Q(\varphi)(v) = Q(\lambda)v.$$

Démonstration du Lemme 5.6.2. On a par récurrence immédiate $\forall k \in \mathbb{N}, \varphi^k(v) = \lambda^k v$. Par addition, on a $\sum_{k=0}^{d} a_k \varphi^k(v) = \sum_{k=0}^{d} a_k \lambda^k v$. □

Démonstration du théorème 5.6.1. Supposons φ diagonalisable et notons $\mathrm{Sp}(\varphi) = \{\lambda_1, \ldots, \lambda_k\}$, alors

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$$

est une décomposition en somme directe. On vérifie que le polynôme $Q = \prod_{i=1}^{k}(\lambda_i - X)$ est annulateur. En effet, tout élément $x \in E$ s'écrit $x = x_1 + \cdots + x_k$ avec $x_i \in E_{\lambda_i}$. En utilisant le Lemme 5.6.2, on calcule

$$Q(\varphi)(x) = \sum_{i=1}^{k} Q(\varphi)(x_i) = \sum_{i=1}^{k} Q(\lambda_i) x_i = 0.$$

Réciproquement, si un polynôme $S(X) = \prod_{i=1}^{m}(\mu_i - X)$ scindé à racines simples est annulateur de φ, alors le lemme des noyaux 5.5.2 garantit que

$$E = \mathrm{Ker}(\varphi - \mu_1 \mathrm{id}) \oplus \cdots \oplus \mathrm{Ker}(\varphi - \mu_m \mathrm{id})$$

est une somme directe. Une base adaptée est une base formée de vecteurs propres, donc φ est diagonalisable. □

En fait, le polynôme $Q = \Pi_\varphi$ est bien le polynôme minimal. Il suffit de voir que si Q_1 est un diviseur strict, alors c'est un produit de facteurs $\lambda_i - X$ mais il manque au moins une valeur propre i_0. Et alors $Q_1(\varphi)$ n'annule pas les vecteurs de $E_{\lambda_{i_0}}$.

5.7 Méthode générale de trigonalisation

Définition 5.7.1. *Pour chacune des valeurs propres $\lambda \in \mathrm{Sp}(\varphi)$, on définit la suite des sous-espaces caractéristiques par*

$$E_\lambda^s = \mathrm{Ker}\left((\varphi - \lambda \mathrm{id})^s\right) \quad \textit{pour } s \in \mathbb{N} \setminus \{0\}. \tag{5.4}$$

Ces sous-espaces caractéristiques sont alors emboités par ordre croissant et tous égaux à partir d'un certain rang. Plus précisément, il existe un entier $c_\lambda \in \mathbb{N} \setminus \{0\}$ tel que

$$E_\lambda^1 \subset E_\lambda^2 \subset \cdots \subset E_\lambda^{c_\lambda - 1} \subset E_\lambda^{c_\lambda} \quad \text{et} \quad E_\lambda^{c_\lambda} = E_\lambda^{c_\lambda + \ell}, \forall \ell > 0. \tag{5.5}$$

Cela découle du lemme dynamique 1.3.4 appliqué à $\psi = \varphi - \lambda \mathrm{id}$.

Théorème 5.7.2. *Soit $\varphi \in \mathcal{L}(E)$. On suppose que φ admet un polynôme annulateur scindé*

$$Q(X) = \prod_{i=1}^{k} (\lambda_i - X)^{c_i}.$$

Alors il existe une base $\mathcal{B}'$ de E pour laquelle la matrice de φ est diagonale par blocs triangulaires comme suit :

$$\mathrm{Mat}_{\mathcal{B}'}(\varphi) = \left(\begin{array}{|c|c|c|c|c|} \hline T(\lambda_1) & 0 & \cdots & 0 & 0 \\ \hline 0 & T(\lambda_2) & \cdots & 0 & 0 \\ \hline \vdots & \vdots & \ddots & \vdots & \vdots \\ \hline 0 & 0 & \cdots & T(\lambda_{k-1}) & 0 \\ \hline 0 & 0 & \cdots & 0 & T(\lambda_k) \\ \hline \end{array}\right)$$

avec $T(\lambda_i) \in \mathrm{M}_{m_i}(\mathbb{K})$ de taille $m_i \times m_i$ de la forme suivante par blocs

$$T(\lambda) = \left(\begin{array}{|ccc|ccc|c|ccc|} \hline \lambda & \cdots & 0 & * & \cdots & * & & * & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda & * & \cdots & * & & * & \cdots & * \\ \hline 0 & \cdots & 0 & \lambda & \cdots & 0 & & * & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \lambda & & * & \cdots & * \\ \hline \multicolumn{3}{|c|}{\vdots} & \multicolumn{3}{c|}{\vdots} & \ddots & \multicolumn{3}{c|}{\vdots} \\ \hline 0 & \cdots & 0 & 0 & \cdots & 0 & & \lambda & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & & 0 & \cdots & \lambda \\ \hline \end{array}\right).$$

Corollaire 5.7.3. *Un endomorphisme $\varphi \in \mathcal{L}(E)$ est trigonalisable si et seulement si son polynôme minimal est scindé.*

Démonstration. Si φ est trigonalisable, alors le polynôme caractéristique est scindé, et a fortiori le polynôme minimal aussi d'après le théorème de Cayley-Hamilton 5.4.1. L'implication réciproque découle du Théorème 5.7.2. □

Démonstration du Théorème 5.7.2. Le lemme des noyaux assure que

$$E = \bigoplus_{i=1}^{m} \mathrm{Ker}((\varphi - \lambda_i \mathrm{id})^{c_i}) = \bigoplus_{i=1}^{m} E_{\lambda_i}^{c_i}.$$

D'autre part, le Fait 5.5.1 assure que pour tout i le sous-espace $E_{\lambda_i}^{c_i} = \ker\left((\varphi - \lambda_i \mathrm{id})^{c_i}\right)$ est stable par φ. Donc la forme de la matrice découle du Lemme **??**. On étudie maintenant le bloc $T(\lambda_i)$.

On choisit alors une base de $E_{\lambda_i}^{c_i}$ adaptée à la suite d'inclusions (5.5), c'est-à-dire que $\mathcal{B}_{i,1}$ est une base de $E_{\lambda_i}^1$, que l'on complète de sorte que $\mathcal{B}_{i,1} \cup \mathcal{B}_{i,2}$ soit une base de $E_{\lambda_i}^1$, etc., que l'on complète de sorte que $\mathcal{B}_{i,1} \cup \cdots \cup \mathcal{B}_{i,c_i}$ soit une base de $E_{\lambda_i}^{c_i}$.

Soit $v \in E_\lambda^s = \ker((\varphi - \lambda \mathrm{id})^s)$, l'observation 1.3.3 avec $\psi = \varphi - \lambda \mathrm{id}$ assure que

$$w := (\varphi - \lambda \mathrm{id})(v) = \varphi(v) - \lambda v \in \ker((\varphi - \lambda \mathrm{id})^{s-1}) = E_\lambda^{s-1}$$

qu'on peut réécrire $\varphi(v) = \lambda v + w$.

La matrice de φ dans la base $\cup_{i=1}^{k} \mathcal{B}_{i,1} \cup \cdots \cup \mathcal{B}_{i,c_i}$ a bien la forme voulue. Le premièr bloc de colonnes corresponde aux vecteurs de $\mathcal{B}_{i,1}$, le deuxième à ceux de $\mathcal{B}_{i,2}$, le dernier à ceux de $\mathcal{B}_{i,c_i}$. □

Corollaire 5.7.4. *Soit $\varphi \in \mathcal{L}(E)$, pour chaque $\lambda_i \in \mathrm{Sp}(\varphi)$, on a*

$$\dim(E^1_{\lambda_i}) \leq \dim(E^{c_i}_{\lambda_i}) = m(\lambda_i),$$

où $m(\lambda)$ est la multiplicité de λ comme racine du polynôme caractéristique et $E^1_\lambda = \ker(\varphi - \lambda \mathrm{id})$ est le sous-espace propre associé à λ.

Démonstration. On sait que le polynôme caractéristique ne dépend pas de la base choisie. Si on choisit la base $\mathcal{B}'$, alors la matrice étant triangulaire, le déterminant est le produit des termes diagonaux, on obtient :

$$P_\varphi(X) = \det(\varphi - X\mathrm{id}) = \det(\mathrm{Mat}_{\mathcal{B}'}(\varphi) - XI_{\dim(E)}) = \prod_{i=1}^{k} (\lambda_i - X)^{\dim(E_i)}.$$

Il n'y a plus qu'à comparer avec (5.1), et utiliser l'unicité de la décomposition en facteurs irréductibles dans $\mathbb{K}[X]$. □

Corollaire 5.7.5. *Soit $\varphi \in \mathcal{L}(E)$ tel que $P_\varphi(X)$ est scindé. Alors φ est diagonalisable si et seulement si $\forall \lambda \in \mathrm{Sp}(\varphi), \dim(E^1_\lambda) = m(\lambda)$.*

5.8 Forme de Jordan

Théorème 5.8.1. *Soit $\varphi \in \mathcal{L}(E)$. On suppose que le polynôme caractéristique $P_\varphi(X)$ est scindé et on utilise les notations (5.1), (5.4), (5.5). Alors il existe une base $\mathcal{B}''$ telle que*

$$\mathrm{Mat}_{\mathcal{B}''}(\varphi) = \left(\begin{array}{|c|c|c|c|c|} \hline T'(\lambda_1) & 0 & \cdots & 0 & 0 \\ \hline 0 & T'(\lambda_2) & \cdots & 0 & 0 \\ \hline \vdots & \vdots & \ddots & \vdots & \vdots \\ \hline 0 & 0 & \cdots & T'(\lambda_{k-1}) & 0 \\ \hline 0 & 0 & \cdots & 0 & T'(\lambda_k) \\ \hline \end{array}\right)$$

avec $T'(\lambda_i) \in \mathrm{M}_{m_i}(\mathbb{K})$ de taille $m_i \times m_i$ de la forme suivante par blocs

$$T'(\lambda) = \left(\begin{array}{|c|c|c|c|} \hline J_1(\lambda) & 0 & \cdots & 0 \\ \hline 0 & J_2(\lambda) & \cdots & 0 \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & \cdots & J_r(\lambda) \\ \hline \end{array}\right)$$

où les blocs $J_i(\lambda)$ sont des blocs de Jordan élémentaires

$$J_i(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}$$

ayant toujours λ sur la diagonale, 1 sur la sur-diagonale et 0 ailleurs.

De plus, le plus gros bloc de Jordan élémentaire de la matrice $T'(\lambda_i)$ est de taille $c_{\lambda_i} \times c_{\lambda_i}$.

Exemple 5.8.2. La matrice suivante de type $T'(\lambda)$ possède 2 blocs de Jordan élémetaire de taille 1 (ce sont juste des entrées diagonales), un bloc de taille 2 et un bloc de taille 3 :

$$T'(\lambda) = \left(\begin{array}{c|c|cc|ccc} \lambda & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & \lambda & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \lambda & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \lambda & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda \end{array}\right)$$

Démonstration du Théorème 5.8.1. On rappelle (5.5), et l'on note $d_s = \dim(E_\lambda^s) - \dim(E_\lambda^{s-1})$ la dimension d'un sous-espace supplémentaire de E_λ^{s-1} dans E_λ^s.

On rappelle que si $H \subset F$ sont deux sous-espaces de E, alors il existe un sous-espace H' tel que $H' \oplus H = F$. Un tel sous-espace est appelé sous-espace supplémentaire de H. On a toujours $\dim(H') = \dim(F) - \dim(H)$ et si on prend une base $(w_1, \ldots, w_h)$ de H et qu'on la complète en une base $(w_1, \ldots, w_f)$ de F, alors $(w_{h+1}, \ldots, w_f)$ engendre (et est une base de) un sous-espace supplémentaire H' de H dans F. Il n'y a bien sûr pas unicité de sous-espace supplémentaire.

On note $\psi = \varphi - \lambda \mathrm{id}$.

On se place dans $E_\lambda^{c_\lambda}$ et on prend $v_1, \ldots, v_{d_{c_\lambda}}$ une base d'un supplémentaire de $E_\lambda^{c_\lambda - 1}$ dans $E_\lambda^{c_\lambda}$. Alors pour chaque $r \leq c_\lambda - 1$, la famille $\psi^r(v_1), \ldots, \psi^r(v_{d_{c\lambda}})$ est une famille libre engendrant un sous-espace de $E_\lambda^{c_\lambda - r + 1}$ en somme directe avec $E_\lambda^{c_\lambda - r}$. Cela découle de l'observation 1.3.3.

Il reste juste à vérifier que cette famille est libre. On le fait pour $r = 1$ pour aléger les notations, le cas général en découle par récurrence. Si $\sum_i \mu_i \psi(v_i) = 0$, alors $\psi(\sum_i \mu_i v_i) = 0$. Supposons par l'absurde que $0 \neq \sum_i \mu_i v_i \in E_\lambda^{c_\lambda} \setminus E_\lambda^{c_\lambda - 1}$, alors $\psi(\sum_i \mu_i v_i) \in E_\lambda c_\lambda - 1 \setminus E_\lambda^{c_\lambda - 2}$ par l'observation, donc $\psi(\sum_i \mu_i v_i) \neq 0$ qui serait une contradiction. C'est donc que $\sum_i \mu_i v_i = 0$. La famille des v_i étant libre, tous les μ_i sont nuls, et la famille $\psi(v_1), \ldots, \psi(v_{d_{c\lambda}})$ est bien libre.

Il s'ensuit que la famille $(\psi^r(v_i))$ pour $0 \leq r \leq c_\lambda - 1$ et $1 \leq i \leq d_{c_\lambda}$ est libre. Mais ce n'est pas forcément une base du sous-espace caractéristique $E_\lambda^{c_\lambda}$. Il faut donc la compléter.

Pour cela, on recommence avec $E_\lambda^{c_\lambda-1}$. On complète la famille $(\psi(v_i))_{1\leq i\leq d_{c_\lambda}}$ avec des vecteurs $v_1^1,\dots,v_{d_1'}^1$ où $d_1'=d_{c_\lambda-1}-d_{c_\lambda}$ afin d'obtenir une base d'un supplémentaire de $E_\lambda^{c_\lambda-2}$ dans $E_\lambda^{c_\lambda-1}$. On rajoute aussi leurs images successives $(\psi^r(v_i^1))$ pour $0\leq r\leq c_\lambda-2$.

Etc. Pour E_λ^s, on complete la famille des $(\psi^{s-\alpha}(v_i^\alpha))$ pour $0\leq\alpha\leq s-1$ avec des vecteurs v_i^s afin d'obtenir une base d'un supplémentaire de E_λ^{s-1} dans E_λ^s et on rajoute leurs images par ψ^r pour $0\leq r\leq s-1$.

À la fin, on obtient une base $\mathcal{B}_\lambda''$ de $E_\lambda^{c_\lambda}$ tout entier. On pose $\mathcal{B}''=\mathcal{B}_{\lambda_1}''\cup\cdots\cup\mathcal{B}_{\lambda_k}''$.

Prenons maintenant les vecteurs $\psi^{c_\lambda-1}(v_i)\dots,\psi(v_i),v_i$ et exprimons leurs images par φ dans $\mathcal{B}''$. On a d'abord

$$\varphi(\psi^{c_\lambda-1}(v_i))=\psi^{c_\lambda-1}(v_i)\quad\text{car }\psi^{c_\lambda-1}(v_i)\in E_\lambda^1$$

Ensuite pour chaque s on a par définition de $\psi=\varphi-\lambda\mathrm{id}$ que

$$\varphi(\psi^s(v_i))=\psi^{s+1}(v_i)+\lambda\psi^s(v_i).$$

Il s'ensuit que les coordonnées $\psi^{c_\lambda-1}(v_i)\dots,\psi(v_i),v_i$ correspondent aux blocs de Jordan élémentaires de taille $c_\lambda\times c_\lambda$. En particulier, il y a d_{c_λ} tels blocs.

On procède de même avec les sous-familles $\psi^r(v_i^s)\dots,\psi(v_i^s),v_i^s$ qui vont donner des blocs de Jordan de taille $s\times s$. Il y aura d_s-d_{s+1} tels blocs. □

Le polynôme minimal peut se déterminer facilement quand on dispose d'une forme de Jordan. Plus précisement :

Corollaire 5.8.3. *[Corollaire du théorème 5.8.1] Soit $\varphi\in\mathcal{L}(E)$. On suppose que le polynôme caractéristique $P_\varphi(X)$ est scindé. On adopte les notations (5.1), (5.4), (5.5). Alors*

$$\Pi_\varphi=\prod_{i=1}^k(\lambda_i-X)^{c_{\lambda_i}}.$$

Ce corollaire découle du théorème 5.8.1 et du fait suivant :

Fait 5.8.4. *Soit $N\in\mathrm{M}_s(\mathbb{K})$ le bloc de Jordan de taille $s\times s$ pour $\lambda=0$:*

$$N=\begin{pmatrix}0&1&0&\cdots&0&0\\0&0&1&\cdots&0&0\\0&0&0&\ddots&0&0\\\vdots&\vdots&\vdots&\ddots&\vdots&\vdots\\0&0&0&\cdots&0&1\\0&0&0&\cdots&0&0\end{pmatrix}=(n_{ij})_{1\leq i,j\leq s}$$

avec $n_{ij}=\begin{cases}1 & \textit{si } i=j+1,\\ 0 & \textit{sinon.}\end{cases}$ *Alors $N^s=0$ et $\forall s'<s, N^{s'}\neq 0$.*

Démonstration. On vérifie par un calcul immédiat (faites-le !) que $\forall k\in\mathbb{N}, N^k=\begin{cases}1 & \text{si } i=j+k,\\ 0 & \text{sinon.}\end{cases}$ □

Démonstration du Corollaire 5.8.3. Il découle du fait 5.8.4 que l'endomorphisme $(\varphi - \lambda_i \mathrm{id})^{c_{\lambda_i}}$ annule la restriction à $E_{\lambda_i}^{c_{\lambda_i}}$ de φ (car le bloc correspondant à la famille $\mathcal{B}''_{\lambda_i}$ est obtenu à partir de $T'(\lambda_i)$ en annulant toutes les entrées diagonales, c'est-à-dire qu'il est diagonal par blocs, d'entrées des matrices de Jordan pour $\lambda_i - \lambda_i = 0$ dont la plus grande a taille $c_{\lambda_i} \times c_{\lambda_i}$), alors que $(\varphi - \lambda_i \mathrm{id})^{c_{\lambda_i}-1}$ ne l'annule pas.

On note $Q(\varphi) := \prod_{i=1}^{k} (\varphi - \lambda_i \mathrm{id})^{c_{\lambda_i}}$. Sa matrice dans la base $\mathcal{B}''$ est obtenue à partir de celle de φ en annulant tous les termes diagonaux (vérifiez-le !).

Donc le polynôme d'endomorphisme $Q(\varphi)$ s'annule sur tout l'espace $E = \bigoplus_{i=1}^{k} E_{\lambda_i}^{c_{\lambda_i}}$ (car chaque facteur s'annule sur un des sous-espaces de la somme directe) alors que ses diviseurs stricts ne s'annulent pas. Ainsi Q est un polynôme annulateur, donc un multiple de Π_φ, mais ses diviseurs ne sont pas annulateurs, donc $Q = \Pi_\varphi$. □

Exemple 5.8.5. Soit

$$B = \left(\begin{array}{c|ccc|cc|cc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{array}\right)$$

alors son polynôme caractéristique est $P_B(X) = (1-X)(2-X)^3(5-X)^4$, tandis que son polynôme minimal est $\Pi_B(X) = (1-X)(2-X)^3(5-X)^2$.

Exercice 45. Évaluer $Q(A)$ où $Q = (X^{32} + X^8 + 14X^3 - 5)(1-X)(2-X)^3 + X$ et A_{11} est la matrice 4×4 étudiée la semaine précédente.

5.9 Application aux suites récurrentes linéaires

Définition 5.9.1. *Une* suite récurrente linéaire *d'ordre p est une suite donnée par*

$$u_0, \ldots, u_p \in \mathbb{K} \quad \textit{et} \quad \forall n \geq p, \quad u_{n+1} = a_{p-1}u_n + a_{p-2}u_{n-1} + \cdots + a_0 u_{n-p+1}$$

où les $a_i \in \mathbb{K}$ sont des coefficients fixés.

La réduction des matrices est très utile pour le calcul des suites récurrentes linéaires.
En effet, à une telle suite $(u_n)_{n\in\mathbb{N}}$, on peut associer la suite vectorielle de terme général

$$v_n = \begin{pmatrix} u_n \\ \vdots \\ u_{n+p} \end{pmatrix}$$

Cette suite vectorielle est récurrente d'ordre 1, donnée par

$$v_0 = \begin{pmatrix} u_0 \\ \vdots \\ u_p \end{pmatrix} \text{ et } \forall n \in \mathbb{N}, \quad v_{n+1} = \begin{pmatrix} u_{n+1} \\ u_{n+2} \\ \vdots \\ u_{n+p} \\ u_{n+p+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ a_0 & a_1 & a_2 & \cdots & a_{p-1} \end{pmatrix} \begin{pmatrix} u_n \\ u_{n+1} \\ \vdots \\ u_{n+p-1} \\ u_{n+p} \end{pmatrix} = Cv_n,$$

où C est la transposée d'une matrice compagnon (on peut encore dire que c'est une matrice compagnon).

On en déduit que $v_n = C^n v_0$, et la réduction (par exemple une diagonalisation quand c'est possible) permet de calculer C^n.

Exercice 46. On considère la suite de Fibonnacci, donnée par $u_0 = 0, u_1 = 1$ et $\forall n \geq 1, u_{n+1} = u_n - u_{n-1}$.

Déterminer la matrice compagnon. La diagonaliser explicitement. En déduire que

$$\forall n \in \mathbb{N}, \quad u_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right).$$

Table des matières