Algèbre Linéaire 2

2019-2020

Université de Bordeaux

Table des Matières

CHAPITRE 1

Groupes, anneaux et corps

Ce premier chapitre est consacré à l'introduction de quelques structures algébriques fondamentales de l'algèbre générale, en particulier les groupes, les anneaux et les corps. Nous étudierons particulièrement le groupe symétrique qui sera très utile dans la suite du cours à travers notamment la notion de déterminant.

I. Groupes

I.1. Définition et premières propriétés.

Définition 1. *Un groupe est un ensemble G muni d'une loi de composition interne $*$, c'est-à-dire d'une application :*

$$G \times G \to G$$
$$(x, y) \mapsto x * y$$

vérifiant les propriétés suivantes :

(1) La loi $*$ est associative *: $x * (y * z) = (x * y) * z$ pour tout x, y, z dans G.*

(2) $(G, *)$ possède un élément neutre*, c'est-à-dire un élément $e \in G$ tel que $e * x = x * e = x$ pour tout x dans G.*

(3) Tout élément de G est inversible *: pour tout $x \in G$, il existe $y \in G$ tel que $x * y = y * x = e$.*

Si de plus, $x * y = y * x$ pour tout x, y appartenant à G, on dit que $(G, *)$ est un groupe commutatif (ou abélien).

Remarque 2. *L'axiome d'associativité permet de définir l'opération sur trois éléments et non plus deux, en enlevant les parenthèses. Evidemment par une récurrence triviale le principe se généralise à un nombre quelconque fini d'éléments.*

Remarque 3. *Dans les exemples qui vont suivre, nous verrons que l'opération $*$ peut désigner suivant les cas l'addition $+$, la multiplication $\times$ ou encore la composition $\circ$. Il faut donc en pratique toujours avoir à l'esprit l'opération $*$ associée au groupe.*

Remarque 4. *Notons que l'on a précisé dans* (2) *que le neutre est neutre à gauche et à droite, et dans* (3) *que l'inverse est inverse à gauche et à droite car dans le cas général d'un groupe non commutatif le résultat des opérations $x * y$ et $y * x$ peuvent différer. Cependant les axiomes peuvent être réduits. Par exemple, on obtient une définition équivalente si on remplace les deux derniers axiomes de la définition ci-dessus par les conditions suivantes : il existe un élément e de G qui est neutre à gauche et tout élément admet un inverse à gauche, i.e. pour tout élément a de G, il existe b dans G tel que $b * a = e$ (ce qu'on peut exprimer en disant que tout élément de G admet un symétrique à gauche en association avec e).*

En particulier, si les conditions affaiblies sont vérifiées alors on peut montrer que les deux derniers axiomes de la définition sont vérifiés. En effet, supposons qu'il existe un élément e de G qui est neutre à gauche et tout élément admet un inverse à gauche et montrons successivement que tout élément admet aussi un inverse à droite puis que e est aussi élément neutre à droite.

Soit a un élément de G et b un inverse à gauche. Choisissons de même un inverse c à gauche de b. On a donc : $b * a = e$ et $c * b = e$. Donc $a * b = e * (a * b)$ car e est élément neutre à gauche. Ainsi $a * b = (c * b) * (a * b)$ car $c * b = e$ et donc $a * b = c * (b * a) * b$; par associativité. On obtient $a * b = c * e * b$; puisque $b * a = e$, puis $a * b = c * b$; car $e * b = b$, puisque e est élément neutre à gauche. Enfin $a * b = e$; car c est un inverse à gauche de b. Ainsi b est aussi un inverse à droite de a.

Prouvons maintenant que e est élément neutre à droite. Soit a un élément du groupe. D'après ce qui précède, nous pouvons choisir un élément b qui est inverse à gauche et à droite de a. Alors $a = e * a = (a * b) * a = a * (b * a)$ *par associativité donc* $a = a * e$ *et donc e est bien aussi élément neutre à droite.*

Proposition 5. *Dans un groupe* $(G, *)$*, on a les propriétés suivantes :*

(1) L'élément neutre est unique.

(2) L'inverse d'un élément est unique. On note x^{-1} *l'inverse de* $x \in G$*.*

(3) Le neutre vérifie $e^{-1} = e$*.*

(4) Pour tout x dans G, on a $(x^{-1})^{-1} = x$*.*

(5) Pour tout x et pour tout y dans G, on a $(x * y)^{-1} = y^{-1} * x^{-1}$*.*

(6) Pour tout a, x, y *dans G, si* $a * x = a * y$ *ou si* $x * a = y * a$*, alors* $x = y$*.*

Démonstration. On procède successivement.

(1) Si G possède deux neutres e et e' alors $e * e' = e$ mais aussi $e * e' = e'$ donc $e = e'$.

(2) Supposons que x ait deux inverses y et y'. Alors $x * y = e$ et $x * y' = e$ par définition. On en déduit que $x * y = x * y'$. On multiplie cette identité à gauche par y et on utilise l'associativité : $y * (x * y) = y * (x * y')$ donc $(y * x) * y = (y * x) * y'$. Mais $y * x = e$ donc $e * y = e * y'$ donc $y = y'$.

(3) Puisque e est neutre, en particulier, $e * e = e$. Mais e a un unique inverse donc cet inverse est bien e.

(4) De même, la propriété $x * x^{-1} = x^{-1} * x = e$ montre que x est l'inverse de x^{-1}.

(5) Enfin en utilisant deux fois l'associativité de la loi, que y est l'inverse de y^{-1}, puis que x est l'inverse de x^{-1}, on obtient :

$$\begin{aligned}(x * y) * (y^{-1} * x^{-1}) &= \big((x * y) * y^{-1} * \big)x^{-1} \\ &= \big(x * (y * y^{-1}) * \big)x^{-1} \\ &= \big(x * e\big) * x^{-1} \\ &= x * x^{-1} \\ &= e.\end{aligned}$$

et on en conclut donc que l'inverse de $x * y$ est $y^{-1} * x^{-1}$.

(6) Il suffit d'utiliser l'inverse de a.

□

I.2. Exemples. Dans la suite, K désigne $\mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$.
Débutons avec quelques exemples standards :

— les ensembles $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ tous munis de la somme sont des groupes pour lesquels le neutre est $e = 0$, l'inverse de x est $-x$. Ils sont commutatifs. Observons qu'en revanche $(\mathbb{N}, +)$ n'est pas un groupe.

— les ensembles $\mathbb{Q}^*$, $\mathbb{R}^*$ et $\mathbb{C}^*$ tous munis du produit sont des groupes pour lesquels le neutre est $e = 1$, l'inverse de x est $1/x$. Ils sont commutatifs.

— L'ensemble $\mathrm{M}_{m,n}(K)$ des matrices de taille (m, n) à coefficients dans K muni de la somme est un groupe commutatif, e est la matrice nulle.

— L'ensemble $\mathrm{GL}(n, K)$ des matrices carrées inversibles de taille n à coefficients dans K, muni du produit des matrices. Le neutre est la matrice identité, l'inverse est la matrice inverse. Il est non commutatif si $n > 1$.

— Le groupe symétrique S_n des bijections σ de $\{1, 2, \ldots, n\}$ dans lui-même. Nous allons revenir en détails sur S_n par la suite. Ses éléments σ sont appelés des permutations de $\{1, 2, \ldots, n\}$. L'ensemble S_n, muni de la composition des applications est un groupe dont l'élément neutre est l'identité, noté Id. Il est non commutatif si $n \geq 3$. On note parfois la composition plus simplement $\sigma_1\sigma_2$ au lieu de $\sigma_1 \circ \sigma_2$.

Rappelons maintenant la définition de $\mathbb{Z}/n\mathbb{Z}$, où $n \geq 1$ est un entier. On définit sur $\mathbb{Z}$ la congruence modulo n de la manière suivante : on dit que a est congruent à b si $a-b$ est divisible par n et on note $a = b \bmod n$. C'est une relation d'équivalence. On note $\mathbb{Z}/n\mathbb{Z}$ l'ensemble quotient pour cette relation d'équivalence. On note $a \bmod n$, $\bar{a}$ ou tout simplement a s'il n'y a pas d'ambiguïté, la classe d'équivalence d'un élément $a \in \mathbb{Z}$. Le cardinal de $\mathbb{Z}/n\mathbb{Z}$ est n.

— Venons-en à l'exemple très important du **groupe** commutatif $(\mathbb{Z}/n\mathbb{Z}, +)$. On définit une opération d'addition dans $\mathbb{Z}/n\mathbb{Z}$ en posant :

$$(a \bmod n) + (b \bmod n) = (a+b) \bmod n.$$

Pour avoir un sens, il faut montrer que cette définition ne dépend pas du choix d'un représentant d'une classe de congruence modulo n. Autrement dit, si $a = a' \bmod n$ et $b = b' \bmod n$, il faut montrer que $(a+b) = (a'+b') \bmod n$. Pour cela, on traduit les congruences par des égalités dans $\mathbb{Z}$: il existe u tel que $a = a' + un$ et il existe v tel que $b = b' + vn$. Alors $(a+b) = (a'+b') + (u+v)n$ ce qui conduit à : $(a+b) = (a'+b') \bmod n$. On vérifie aisément que le neutre pour cette opération est $0 \bmod n$ et que tout élément $x \bmod n$ est inversible, d'inverse $-x \bmod n$.

— On peut également utiliser **le groupe multiplicatif** $(\mathbb{Z}/n\mathbb{Z})^*$**.** De la même façon que pour l'addition, on peut définir une multiplication dans $\mathbb{Z}/n\mathbb{Z}$ en posant :

$$(a \bmod n)(b \bmod n) = (ab) \bmod n.$$

Cette loi est associative, commutative, et possède un élément neutre qui est $1 \bmod n$. Par contre, tout élément n'est pas inversible, en particulier $0 \bmod n$ n'est *jamais* inversible. Par exemple, on vérifie facilement que les inversibles modulo 4 sont 1 et 3.

Définition 6. *Pour $n \geq 2$, on note $(\mathbb{Z}/n\mathbb{Z})^*$ l'ensemble des éléments de $\mathbb{Z}/n\mathbb{Z}$ qui sont inversibles pour la multiplication.*

On a donc :

$$(\mathbb{Z}/n\mathbb{Z})^* = \{a \bmod n : \text{ il existe } b \in \mathbb{Z} \text{ tel que } ab = 1 \bmod n\}.$$

Alors $(\mathbb{Z}/n\mathbb{Z})^*$ muni de la multiplication est un groupe commutatif.

Par exemple :

$$(\mathbb{Z}/4\mathbb{Z})^* = \{1,3\}, \quad (\mathbb{Z}/5\mathbb{Z})^* = \{1,2,3,4\}, \quad (\mathbb{Z}/6\mathbb{Z})^* = \{1,5\}.$$

En général, $(\mathbb{Z}/n\mathbb{Z})^*$ n'est donc pas $\mathbb{Z}/n\mathbb{Z}$ privé de 0.

Exercice 7. *Démontrez en détail que la multiplication définie ci-dessus a bien un sens, et que $(\mathbb{Z}/n\mathbb{Z})^*$ muni de cette multiplication est un groupe commutatif.*

Notations. Cette duplicité de structures de groupe lié à $\mathbb{Z}/n\mathbb{Z}$ nous amène aux quelques mots d'avertissement suivants. Pour alléger les notations, on va souvent utiliser la *notation multiplicative* pour un groupe général : $x * y = xy$, et $e = 1$. On dit que xy est le *produit* de x et y. On utilise aussi les raccourcis $x^n = x * \cdots * x$ (n termes), $x^0 = e$, pour $n \in \mathbb{N}$ et $x^{-n} = x^{-1} * \cdots * x^{-1}$.

Lorsque le groupe est commutatif, en particulier lorsque la loi est issue de l'addition usuelle, on emploie souvent la *notation additive* $x * y = x + y$ et $e = 0$. Alors on parle d'opposé plutôt que d'inverse d'un élément, et on note $nx = x + \cdots + x$.

Notons enfin que si l'on dispose de deux groupes $(G, *)$ et $(G', \circ)$, on peut définir le *produit direct* $G \times G'$ de ces deux groupes comme l'ensemble

$$G \times G' = \{(x, y) : x \in G, y \in G'\}.$$

Muni de l'opération : $(x,y) \cdot (x',y') = (x * x', y \circ y')$, c'est un groupe dont le neutre est $(e_G, e_{G'})$ et l'inverse de (x,y) est (x^{-1}, y^{-1}).

I.3. Sous-groupes.

Définition 8. *Soit $(G,*)$ un groupe. Un sous-groupe de G est un sous-ensemble $H \subset G$ tel que la loi $*$ préserve H et $(H,*)$ est un groupe.*

Proposition 9. *Soit $(G,*)$ un groupe et soit $H \subset G$. Alors H est un sous-groupe de G si et seulement si*

$$\begin{array}{ll}(i) & \text{Pour tout } x \in H, y \in H, x * y \in H.\\ (ii) & e \in H\\ (iii) & \text{Pour tout } x \in H, x^{-1} \in H.\end{array}$$

Démonstration. Examinons les propriétés que doit vérifier $H \subset G$ pour être un groupe pour $*$.

Tout d'abord il est nécessaire que la loi $*$ soit interne dans H, c'est-à-dire que $x * y \in H$ pour tout $x \in H$, $y \in H$. Remarquons que la loi $*$ étant associative dans G, elle l'est forcément dans H.

Le neutre de H ne peut être que le neutre de G. En effet, si e' est le neutre de H, on a comme $e' \in G$, $e' * e = e'$. Mais aussi $e' * e' = e'$ en raisonnant dans H. Donc $e' * e = e' * e'$. Comme $e' \in G$, il a un inverse e'^{-1} dans G. On multiplie la précédente égalité à gauche par celui-ci, pour obtenir : $(e'^{-1} * e') * e = (e'^{-1} * e') * e'$ soit $e * e = e * e'$ soit $e = e'$. Donc si $(H,*)$ est un groupe, son neutre est e le neutre de G.

Pour ce qui est de l'inverse, un élément de H a bien toujours un inverse (unique) dans G. Il faut donc que cet inverse appartienne à H.

Réciproquement, si les propriétés (i), (ii), (iii), sont vérifiées, alors $(H,*)$ est bien un groupe. En effet, l'associativité et la propriété $e * x = x$ sont automatiquement vraies dans H puisqu'elles sont vraies dans G. □

La proposition suivante énonce une propriété minimale suffisante à vérifier pour qu'un sous-ensemble de G soit un sous-groupe de G.

Proposition 10. *Soit $(G,*)$ un groupe et soit $H \subset G$. Alors H est un sous-groupe de G si et seulement si, H est non vide, et vérifie :*

$$\text{Pour tout } x \in H, y \in H, x * y^{-1} \in H. \tag{1}$$

Démonstration. Supposons que H soit un sous-groupe de G. Alors on a vu que H vérifie (i), (ii), (iii). En particulier (ii) indique que le neutre e de G appartient à H donc H est non vide. Si x et y sont dans H, d'après (iii), $y^{-1} \in H$ et, d'après (i), $x * y^{-1} \in H$.

Réciproquement, supposons que H est non vide et vérifie (1). Il contient donc un élément x. Donc, par (1), $x * x^{-1} = e \in H$. En appliquant encore (1), on obtient $e * x^{-1} = x^{-1} \in H$. Si x et y sont dans H, on a vu que $y^{-1} \in H$, donc encore avec (1), $x * y = x * (y^{-1})^{-1} \in H$. Donc (i), (ii), (iii) sont vérifiées donc H est bien un sous-groupe de G. □

Exemple 11.

- *$\{e\}$ et G sont des sous-groupes de G.*
- *Pour tout n entier naturel, l'ensemble $n\mathbb{Z} := \{nq,\ q \in \mathbb{Z}\}$ est un sous-groupe de $(\mathbb{Z},+)$. En effet, il est non vide puisque $0 \in n\mathbb{Z}$ et si $x = nk \in n\mathbb{Z}$, $y = n\ell \in \mathbb{Z}$, $x - y = n(k-\ell) \in n\mathbb{Z}$.*

Proposition 12. *Les sous-groupes de $(\mathbb{Z},+)$ sont les $n\mathbb{Z}$, pour n dans $\mathbb{N}$.*

Démonstration. On vient de voir que $n\mathbb{Z}$ est un sous-groupe de $\mathbb{Z}$. Réciproquement, soit H un sous-groupe de $\mathbb{Z}$. Il est, par définition, nécessaire que H contienne le neutre 0. Si H contient un autre élément, disons x, alors il contient nécessairement un élément strictement positif, car si $x < 0$ alors son inverse $-x$ est aussi dans H. Il fait donc sens de considérer n le plus petit élément strictement positif. Prenons maintenant $x \in H$ quelconque. Par division euclidienne il existe q dans $\mathbb{Z}$ et $r \in \{0,1,\dots,n-1\}$ tels que $x = qn + r$. Alors qn, et donc $r = x - qn$ appartient à H. Comme $r \in \{0,1,\dots,n-1\} \cap H$ et que n est le plus petit élément strictement positif de H, il n'y a qu'une possibilité c'est $r = 0$. Donc $x \in n\mathbb{Z}$ et $H = n\mathbb{Z}$. □

Evoquons maintenant la stabilité de la notion de sous-groupes vis-à-vis des opérations ensemblistes usuelles.

La proposition suivante, relative à l'intersection, dont la preuve est laissée au lecteur, sera utile pour la suite.

Proposition 13. *Soit H_1 et H_2 deux sous-groupes de $(G,*)$. L'intersection $H_1 \cap H_2$ de H_1 et H_2 est un sous-groupe de G. De manière générale, une intersection quelconque de sous-groupes est un sous-groupe : si $(H_i)_{i\in I}$ est une collection de sous-groupes de G, l'intersection $\cap_{i\in I} H_i$ est aussi un sous-groupe de G.*

Attention, en revanche la réunion de deux sous-groupes n'est pas un sous-groupe. Par exemple $2\mathbb{Z}\cup 3\mathbb{Z}$ n'est pas un sous-groupe de $(\mathbb{Z},+)$ car $2+3=5$ n'est pas dans cet ensemble.

I.4. Sous-groupes engendrés par une partie.

Définition 14. *Soit $S \subset G$. On appelle* sous-groupe engendré par S *et on note $\langle S\rangle$ l'intersection de tous les sous-groupes de G contenant S.*

Il est facile de vérifier que c'est un sous-groupe de G, et que c'est le plus petit contenant S (au sens où, si H est un sous-groupe de G contenant S, alors $\langle S\rangle \subset H$).

Si $S=\{x\}$ avec $x\in G$, on note $\langle x\rangle = \langle\{x\}\rangle$.

Exemple 15.

(1) Si $S=\{e\}$, $\langle e\rangle = \{e\}$

(2) Si $S=\{2,3\}\subset\mathbb{Z}$, $\langle S\rangle = \mathbb{Z}$. En effet, $1=3-2\in\langle S\rangle$.

Proposition 16. *Soit G un groupe noté multiplicativement.*

(1) $\langle x\rangle = \{x^k : k\in\mathbb{Z}\}$.

(2) Si x et y commutent, i.e. $xy=yx$, alors $\langle x,y\rangle = \{x^k y^\ell : k,\ell\in\mathbb{Z}\}$.

La preuve est laissée au lecteur.

À partir de maintenant, on utilise systématiquement la notation multiplicative $x*y=xy$ et $e=1_G$ pour un groupe général G. Un peu plus tard on simplifiera encore 1_G en 1.

I.5. Ordres.

Définition 17. *L'ordre d'un groupe est le nombre de ses éléments (il peut être infini). On note $|G|$ l'ordre de G.*

Définition 18. *Soit G un groupe et soit $x\in G$. L'ordre de x est le plus petit entier $k\geq 1$, s'il existe, tel que $x^k = 1_G$.*

Si pour tout $k\geq 1$, $x^k \neq 1_G$, on dit que x est d'ordre infini.

Exemple 19.

- *Le neutre 1_G est toujours d'ordre 1.*
- *Dans $(\mathbb{Z},+)$, tout élément non nul est d'ordre infini.*
- *Dans $(\mathbb{Z}/n\mathbb{Z},+)$, tout élément a vérifie $na=0$. Mais a n'est pas forcément d'ordre n. Exemple : dans $\mathbb{Z}/6\mathbb{Z}$, 2 est d'ordre 3.*
- *Dans $(\mathbb{Z}/n\mathbb{Z},+)$, 1 est d'ordre n.*

Proposition 20. *Soit G un groupe et soit $x\in G$, un élément d'ordre k. Si $n\in\mathbb{Z}$ et si $n=kq+r$, avec q,r dans $\mathbb{Z}$ et $0\leq r<k$ est la division euclidienne de n par k, alors*

$$x^n = x^r.$$

De plus on a l'équivalence :

$$x^n = 1_G \iff k \text{ divise } n$$

Démonstration. Si $n = kq + r$ alors $x^n = x^{kq+r} = (x^k)^q.x^r$. Donc, si $x^k = 1$ alors $x^n = x^r$.
En particulier, si k divise n alors $r = 0$ et $x^n = x^r = 1$.
Réciproquement, si $x^n = 1$, alors $x^r = 1$ avec $0 \leq r < k$ mais comme k est le plus petit entier positif avec cette propriété, c'est que $r = 0$ et donc que k divise n. □

Remarque 21.

(1) Si $x^n = 1_G$, il ne faut pas conclure trop rapidement que n est l'ordre de x. Par contre, on sait que l'ordre de x est un diviseur de n, ça limite les possibilités. En fait, on peut alors calculer l'ordre de x en descendant l'arbre des diviseurs de n.

(2) Si G est un groupe fini, alors tout élément est d'ordre fini. En effet, $\{1_G, x, x^2, \ldots, x^n, \ldots\}$ ne peut être infini donc il existe $k < \ell$ tels que $x^k = x^\ell$ d'où on tire $x^{\ell-k} = 1_G$.

On a déjà vu la notion de sous-groupe engendré par un élément $x \in G$. Cette notion va justement nous permettre de définir la notion de groupe cyclique.

Définition 22. *Un groupe G est dit* monogène *s'il existe $x \in G$ tel que $G = \langle x \rangle$. Un groupe monogène fini est dit* cyclique. *Un élément x tel que $G = \langle x \rangle$ est appelé* un générateur *de G.*

Exemple 23. *Le groupe $(\mathbb{Z}/n\mathbb{Z}, +)$ est cyclique d'ordre n. Le groupe $((\mathbb{Z}/5\mathbb{Z})^*, \times)$ est cyclique d'ordre 4. Listez leurs générateurs.*

Proposition 24. *Soit G un groupe et $x \in G$.*

(1) Si x est d'ordre k, $\langle x \rangle = \{1, x, x^2, \ldots, x^{k-1}\}$ et $|\langle x \rangle| = k$.

(2) G est cyclique si et seulement si G contient un élément d'ordre $|G|$.

Démonstration. On a déjà vu que $\langle x \rangle = \{x^n : n \in \mathbb{Z}\}$ donc $\{1, x, x^2, \ldots, x^{k-1}\} \subset \langle x \rangle$. Si x est d'ordre k, $x^n = x^r$ où r est le reste de n dans la division euclidienne par k, $0 \leq r < k$, donc l'inclusion inverse est vérifiée. Il reste à montrer que l'ensemble $\{1, x, x^2, \ldots, x^{k-1}\}$ a exactement k éléments, c'est-à-dire que les x^i, $0 \leq i \leq k-1$ sont distincts. En effet, supposons que, pour $0 \leq i < j \leq k-1$, on ait $x^i = x^j$. Alors, $x^{j-i} = 1$. Mais $1 \leq j - i \leq k-1$, donc c'est en contradiction avec la propriété que k est le plus petit entier positif tel que $x^k = 1$.
Supposons G cyclique. Alors, il existe $x \in G$ tel que $G = \langle x \rangle$, et, d'après la discussion qui précède, $|G|$ est égal à l'ordre de x. Donc G contient bien un élément dont l'ordre vaut $|G|$. Réciproquement, supposons que G contienne un élément x d'ordre $k = |G|$. Alors $\langle x \rangle \subset G$ et $|\langle x \rangle| = k = |G|$ donc on peut conclure que $\langle x \rangle = G$ □

Attention : Un groupe cyclique n'a pas un unique générateur. En fait, il a autant de générateurs qu'il y a d'éléments d'ordre égal à l'ordre de ce groupe.

I.6. Morphisme de groupe.

Définition 25. *Si $(G, \star)$ et $(G', *)$ sont deux groupes, on dit qu'une application $\phi : G \to G'$ est un morphisme de groupes si $\phi(x \star y) = \phi(x) * \phi(y)$, pour tous $x, y \in G$.*

Donnons quelques exemples. L'application constante égale au neutre est un morphisme. L'identité est un morphisme d'un groupe vers lui-même. La fonction exponentielle $\mathbb{C} \to \mathbb{C}^*$, $z \mapsto \mathrm{e}^z$ vérifie $e^{z+z'} = \mathrm{e}^z \times \mathrm{e}^{z'}$. C'est donc un morphisme de groupes de $(\mathbb{C}, +)$ dans $(\mathbb{C}^*, \times)$ et, par restriction, de $(\mathbb{R}, +)$ dans $(\mathbb{R}^*, \times)$.

Proposition 26. *On considère deux groupes G et G' et un morphisme de groupes $\phi : G \to G'$. On note respectivement e et e' les neutres de G et G' Alors $\phi(e) = e'$.*

Démonstration. Il suffit de prendre $x = x' = e$ dans la définition et de simplifier. □

Proposition 27. *On considère deux groupes G et G' et un morphisme de groupes $\phi : G \to G'$. Alors pour tout x dans G , $\phi(x^{-1}) = \phi(x)^{-1}$.*

DÉMONSTRATION. Il suffit de prendre $x' = x^{-1}$ dans la définition et d'utiliser la proposition précédente. □

On laisse le lecteur vérifier l'énoncé suivant.

Proposition 28. *Soit $\phi : G \to G'$ un morphisme de groupes. Alors l'image réciproque $\phi^{-1}(H')$ de tout sous-groupe H' de G' est un sous-groupe de G. L'image directe $\phi(H)$ de tout sous-groupe H de G est un sous-groupe de G'.*

En particulier, pour tout morphisme $\phi : G \to G'$, l'image $\mathrm{Im}(\phi) = \phi(G)$ et le noyau $\ker(\phi) = \phi^{-1}(\{e'\})$ sont des sous-groupes de G'.

On regroupe les définitions suivantes.

Définition 29.

- *Un isomorphisme de groupes est un morphisme de groupes qui est bijectif.*
- *Lorsqu'il existe un isomorphisme du groupe G vers le groupe G', sa bijection réciproque est un isomorphisme du groupe G' vers le groupe G ; on dit alors que les deux groupes sont isomorphes, ce que l'on note $G \simeq G'$.*
- *Un automorphisme du groupe G est un isomorphisme de G vers G. L'ensemble des automorphismes du groupe G est généralement noté $Aut(G)$.*

On vérifie facilement la proposition suivante.

Proposition 30. *$Aut(G)$ est un sous-groupe du groupe des bijections de G dans G, muni de la loi de composition.*

II. Le groupe symétrique S_n

II.1. Définition. On commence par rappeler la notion de permutation, ce seront les éléments du groupe symétrique.

Définition 31. *Une* permutation *σ de $\{1, 2, \ldots, n\}$ est une bijection de $\{1, 2, \ldots, n\}$ dans lui-même.*

Définition 32. *On appelle groupe symétrique de l'ensemble $\{1, 2, \ldots, n\}$ et on note S_n l'ensemble des* permutations *de $\{1, 2, \ldots, n\}$ muni de la loi de composition.*

Une permutation est notée de la façon suivante :

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

Un exemple important de permutation est celui des transpositions.

Définition 33. *Une* transposition *τ est une permutation qui laisse invariant tous les éléments sauf deux qu'elle échange : il existe deux éléments distincts i et j tels que :*

$$\tau(i) = j \text{ et } \tau(j) = i \text{ et } \quad \forall k \neq i \text{ et } k \neq j,\ \tau(k) = k.$$

Voyons maintenant une autre classe de permutations qui généralise la notion de transposition.

Définition 34. *Un* cycle *est une permutation particulière qui permute circulairement un sous-ensemble de $\{1, \ldots, n\}$ et laisse les autres éléments inchangés. On note le cycle $\sigma = (a_1, \ldots, a_p)$, $p \geq 2$, si $\sigma(a_1) = a_2$, $\sigma(a_2) = a_3, \ldots, \sigma(a_p) = a_1$. On dit que $p \geq 2$ est* la longueur *du cycle σ et on la note $\ell(\sigma)$.*

Une transposition est donc un cycle de longueur 2.

Exemple 35. *Considérons dans S_5 :*

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 2 & 4 & 5 \end{pmatrix}.$$

On constate que σ est le cycle de longueur 3 qui envoie 1 sur 3, 3 sur 2 et 2 sur 1, que l'on peut aussi noter : $\sigma = (1, 3, 2)$.

Exemple 36. *Les éléments de S_3 sont*

$$\mathrm{Id} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad (1\ 2) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad (1\ 3) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$
$$(2\ 3) = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad (1\ 2\ 3) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \text{ et } (1\ 3\ 2) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

On remarque qu'il y a $6 = 3!$ éléments.

Exercice 37. *Montrez que S_n n'est pas commutatif si $n \geq 3$.*

Exercice 38. *Montrez que S_n a $n!$ éléments c'est-à-dire qu'il est d'ordre $n!$.*

Définition 39. *Le* support *d'une permutation σ est l'ensemble :*

$$\mathrm{Sup}(\sigma) := \{i \in \{1, \dots, n\} : \sigma(i) \neq i\}.$$

En particulier si σ est le cycle $\sigma = (a_1, \dots, a_p)$, alors $\mathrm{Sup}(\sigma) = \{a_1, \dots, a_p\}$.

Proposition 40. *Si S est le support de σ, $\sigma(S) = S$.*

Démonstration. Prenons j dans S, de sorte que $\sigma(j) \neq j$. Comme σ est bijective, elle est injective et par conséquent $\sigma(\sigma(j)) \neq \sigma(j)$ i.e. $\sigma(j)$ est dans S. On a donc $\sigma(S) \subset S$. De plus, σ étant bijectif, S et $\sigma(S)$ ont le même cardinal. Ainsi $\sigma(S) = S$. □

Proposition 41. *Deux permutations de supports disjoints commutent.*

Démonstration. Soit σ_1 une permutation de support S_1 et σ_2 une permutation de support S_2, telles que $S_1 \cap S_2 = \emptyset$. Montrons que $\sigma_1\sigma_2 = \sigma_2\sigma_1$. On peut distinguer trois cas :

— $i \notin S_1 \cup S_2$. Alors, $\sigma_1\sigma_2(i) = \sigma_1(\sigma_2(i)) = \sigma_1(i) = i$, et de même, $\sigma_2\sigma_1(i) = \sigma_2(\sigma_1(i)) = \sigma_2(i) = i$.

— $i \in S_1$. Alors, $i \notin S_2$ donc $\sigma_1\sigma_2(i) = \sigma_1(i)$. De plus, $\sigma_1(i) \in S_1$ d'après la proposition précédente, donc $\sigma_1(i) \notin S_2$ et $\sigma_2(\sigma_1(i)) = \sigma_1(i)$.

— $i \in S_2$. Ce cas est analogue au précédent.

□

Proposition 42. *Un cycle de longueur p est d'ordre p dans S_n.*

Démonstration. Si $c = (a_1, \dots, a_p)$, $c^k(a_1) = a_{k+1}$ donc, si $c^k = 1$, alors $k \geq p$. Il est clair que $c^p = 1$. □

Définition 43. *Si $\sigma \in S_n$ et $i \in \{1, \dots, n\}$, on appelle σ-orbite de i l'ensemble $\mathcal{O}_\sigma(i) = \{\sigma^k(i)\,,\, k \in \mathbb{Z}\}$.*

Proposition 44. *Le cardinal $m = |\mathcal{O}_\sigma(i)|$ de $\mathcal{O}_\sigma(i)$ est inférieur ou égal à n et*

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

Par ailleurs, pour tous $i, j \in \{1, \dots, n\}$, on a soit $\mathcal{O}_\sigma(i) = \mathcal{O}_\sigma(j)$, soit $\mathcal{O}_\sigma(i) \cap \mathcal{O}_\sigma(j) = \emptyset$.

Démonstration. L'inégalité $m \leq n$ est évidente. Soit $d \geq 1$ le plus petit entier positif tel que $i, \sigma(i), \dots, \sigma^d(i)$ ne soient pas deux à deux distincts. On a $d \leq m$ et par définition, il existe $k \in \{0, \dots, d-1\}$ tel que $\sigma^k(i) = \sigma^d(i)$. Par conséquent, $\sigma^{d-k}(i) = i$, ce qui montre que $\{i, \sigma(i), \sigma^2(i), \dots, \sigma^{d-k}(i)\}$ ne sont pas deux à deux distincts, donc $k = 0$.

Montrons maintenant que $d = m$. Pour tout $k \in \mathbb{Z}$, soit $k = qd + r$, $q \in \mathbb{Z}$ et $0 \leq r \leq d-1$ la division euclidienne de k par d. On a $\sigma^k(i) = \sigma^r \circ \sigma^{qd}(i) = \sigma^r(i) \in \{i, \dots, \sigma^{d-1}(i)\}$, car $\sigma^{qd}(i) = \sigma^d \circ \cdots \circ \sigma^d(i)$ (on compose q fois σ^d) et $\sigma^d(i) = i$. Comme par ailleurs $d \leq m$, on a bien le résultat annoncé.

La dernière assertion est évidente car si $\mathcal{O}_\sigma(i) \cap \mathcal{O}_\sigma(j) \neq \emptyset$, il existe $k \in \{0, \dots, m-1\}$ tel que $j = \sigma^k(i)$ et le résultat découle immédiatement par définition des σ-orbites. □

II.2. La décomposition canonique d'une permutation en produit de cycles.

Théorème 45.

Une permutation σ se décompose en un produit de cycles à supports disjoints, de façon unique à l'ordre près.

Démonstration. D'après la proposition 44, les σ-orbites sont deux à deux disjointes. Soient $\mathcal{O}_1, \dots, \mathcal{O}_s$ les σ-orbites de cardinal au moins 2 et $m_1, \dots, m_s \geq 2$ leurs cardinaux respectifs. Pour tout $1 \leq k \leq s$, il existe $i \in \{1, \dots, n\}$ tel que $\mathcal{O}_k = \{i, \sigma(i), \dots, \sigma^{m_k-1}(i)\}$. On définit alors le cycle $c_k := (i, \sigma(i), \dots, \sigma^{m_k-1}(i))$. On remarque immédiatement que la définition de c_k est indépendante de $i \in \mathcal{O}_k$. En outre, les supports des c_k sont les $\mathcal{O}_k$ et sont donc deux à deux disjoints. En particulier, les c_k commutent deux à deux.

Soit $i \in \{1, \dots, n\}$. Si i est l'unique point de sa σ-orbite, alors $c_1 \circ c_2 \circ \cdots \circ c_s(i) = i$ car i n'appartient au support d'aucun cycle c_p. Sinon, il existe un unique k tel que $i \in \mathcal{O}_k$. On a alors $c_1 \circ c_2 \circ \cdots \circ c_s(i) = c_k \circ c_1 \circ \cdots \circ c_{k-1} \circ c_{k+1} \circ \cdots \circ c_s(i) = c_k(i) = \sigma(i)$ où l'on a utilisé dans l'ordre que les cycles c_p commutent deux à deux, que $c_p(i) = i$ si $p \neq k$ car les supports sont deux à deux disjoints, et $c_k(i) = \sigma(i)$ par définition de c_k. Par conséquent, on a montré que $\sigma = c_1 \circ \cdots \circ c_s$, d'où l'existence de la décomposition.

En ce qui concerne l'unicité, il est clair que si $\sigma(i) \neq i$, i doit être dans le support d'un cycle qui envoie i sur $\sigma(i)$, en d'autres termes il n'y a pas de choix pour les cycles c_p. □

Voyons cela sur un exemple :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 4 & 3 & 1 & 7 & 6 & 5 & 2 & 8 & 10 & 9 \end{pmatrix}$$

On part de 1 et on applique successivement σ pour obtenir le premier cycle $c_1 : 1 \to 4 \to 7 \to 2 \to 3 \to 1$, donc $c_1 = (1, 4, 7, 2, 3)$. On recommence en partant d'un élément qui n'a pas été visité : $5 \to 6 \to 5$; puis $8 \to 8$ et $9 \to 10 \to 9$ qui donnent :

$$\sigma = (1, 4, 7, 2, 3)(5, 6)(9, 10).$$

Notons que l'on convient de ne pas écrire le $\circ$ de composition entre les cycles, l'expression ci-dessous désigne $(1, 4, 7, 2, 3) \circ (5, 6) \circ (9, 10)$.

Observons que ceci peut être très utile pour calculer les itérés. Prenons par exemple pour objectif de calculer la permutation σ^{1145}. Parce que les cycles à supports disjoints commutent,

$$\sigma^{1145} = (1, 4, 7, 2, 3)^{1145}(5, 6)^{1145}(9, 10)^{1145}.$$

Il suffit maintenant de réduire les exposants modulo les longueurs respectives des cycles, ce qui donne

$$\sigma^{1145} = (1, 4, 7, 2, 3)^0(5, 6)^1(9, 10)^1 = (5, 6)(9, 10).$$

Corollaire 46. *Soit $\sigma = c_1 \dots c_s$ la décomposition en produits de cycles disjoints de σ, avec $\ell_i := \ell(c_i)$. L'ordre de σ est le ppcm de $\ell_1, \dots, \ell_s$.*

Démonstration. On a déjà vu que $\sigma^k = c_1^k \dots c_s^k$. Comme les supports des c_i sont disjoints, $\sigma^k = 1$ si et seulement si $c_i^k = 1$ pour tout $i = 1, \dots, s$. Comme l'ordre de c_i vaut ℓ_i, $c_i^k = 1$ si et seulement si ℓ_i divise k. Finalement, on a démontré que $\sigma^k = 1$ si et seulement si $\mathrm{ppcm}(\ell_1, \dots, \ell_s)$ divise k donc l'ordre de σ est bien $\mathrm{ppcm}(\ell_1, \dots, \ell_s)$. □

Examinons maintenant d'autres décompositions des permutations.

Théorème 47.

On a :

(1) $c = (a_1, \dots, a_p) = (a_1, a_2)(a_2, a_3) \dots (a_{p-1}, a_p)$.

(2) Toute permutation est un produit de transpositions.

DÉMONSTRATION. Le premier point se vérifie directement, et le deuxième en résulte en combinant avec le théorème 45. □

II.3. La signature.

Définition 48. *Soit $\sigma \in S_n$. Le* nombre d'inversions de σ *est le nombre*

$$\mathcal{I}(\sigma) := |\{(i,j) : 1 \leq i < j \leq n \text{ et } \sigma(i) > \sigma(j)\}|.$$

Nous rappelons que les barres verticales ci-dessus désigne le cardinal de l'ensemble.
Ceci nous permet de définir la notion de signature d'une permutation.

Définition 49. *La* signature de σ *est définie par :*

$$\varepsilon(\sigma) = (-1)^{\mathcal{I}(\sigma)}.$$

Une permutation est dite paire quand elle présente un nombre pair d'inversions, impaire sinon. La signature d'une permutation paire est 1 ; celle d'une permutation impaire est -1.

Exemple 50. *Soit la permutation*

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 5 & 4 & 2 \end{pmatrix}.$$

La liste des paires en inversion est {2, 5}, {3, 4}, {3, 5}, {4, 5}. Il y en a quatre, donc la signature est 1 *et la permutation est paire.*

Proposition 51. *Une transposition est une permutation impaire.*

DÉMONSTRATION. Notons $i < j$ les termes échangés par la transposition, de sorte que

$$\begin{pmatrix} 1 & \dots & i-1 & i & i+1 & \dots & j-1 & j & j+1 & \dots & n \\ 1 & \dots & i-1 & j & i+1 & \dots & j-1 & i & j+1 & \dots & n \end{pmatrix}.$$

Les paires en inversion sont les paires de la forme $\{i,k\}$ avec k compris entre $i+1$ et j et celles de la forme $\{k,j\}$ avec k compris entre $i+1$ et $j-1$. Au total, il y a $(j-i)+(j-i-1)$ soit un nombre impair d'inversions, et l'imparité de la permutation en découle. □

Nous donnons maintenant une manière alternative utile pour calculer la signature d'une permutation.

Proposition 52. *Soit $\sigma \in S_n$. On a*

$$\varepsilon(\sigma) = \prod_{\{i,j\}, i \neq j} \frac{\sigma(j) - \sigma(i)}{j - i},$$

où chaque paire $\{i,j\}$ intervient une fois et une seule (on peut par exemple indexer le produit par l'ensemble des (i,j) tels que $i < j$).

DÉMONSTRATION. Remarquons premièrement que pour tout $i \neq j$, $\frac{\sigma(j)-\sigma(i)}{j-i} = \frac{\sigma(i)-\sigma(j)}{i-j}$, ce qui montre que la formule est bien définie. Par ailleurs, $\frac{\sigma(j)-\sigma(i)}{j-i}$ est positif si l'ordre de i et j est préservé par σ, et négatif si leur ordre est inversé. Ainsi, le signe du terme de droite est bien égal à celui de $\varepsilon(\sigma)$. Enfin, comme σ est une bijection, chaque paire $\{i,j\}$ intervient au numérateur du produit une fois et une seule, ce qui montre que la valeur absolue du membre de droite vaut 1. □

Observant que $\{\pm 1\}$ est un groupe pour la multiplication, on a le résultat suivant.

Théorème 53.

La signature est un morphisme entre $(S_n, \circ)$ et $(\{\pm 1\}, \times)$.

Rappelons que cela signifie que pour toutes permutations $\sigma, \sigma' \in S_n$,

$$\varepsilon(\sigma\sigma') = \varepsilon(\sigma)\varepsilon(\sigma').$$

Démonstration. D'après la proposition 52,

$$\varepsilon(\sigma) = \prod_{i\neq j} \frac{\sigma(j)-\sigma(i)}{j-i} = \prod_{i\neq j} \frac{\sigma(\sigma'(j))-\sigma((\sigma'(i))}{\sigma'(j)-\sigma'(i)}$$

car σ' est une bijection et donc

$$\varepsilon(\sigma)\varepsilon(\sigma') = \prod_{i\neq j} \frac{\sigma(\sigma'(j))-\sigma((\sigma'(i))}{\sigma'(j)-\sigma'(i)} \prod_{i\neq j} \frac{\sigma'(j)-\sigma'(i)}{j-i} = \prod_{i\neq j} \frac{\sigma(\sigma'(j))-\sigma((\sigma'(i))}{j-i} = \varepsilon(\sigma\sigma').$$

□

Corollaire 54. *Si c est un cycle de longueur ℓ, $\varepsilon(c) = (-1)^{\ell-1}$.*

Démonstration. On a vu que $\varepsilon((i,j)) = -1$ et on combine le premier point du théorème 47 avec le résultat précédent. □

On appelle A_n l'ensemble des permutations de signature $+1$. Nous laisserons le soin au lecteur de déduire le résultat suivant du Théorème 53 et de la Proposition 28.

Corollaire 55. *A_n est un sous-groupe de S_n, appelé* le groupe alterné.

III. Anneaux et corps

III.1. Définitions.

Définition 56. *Un* anneau *$(A,+,\cdot)$ est un ensemble muni de deux lois de composition interne $+$ et $\cdot$ telles que :*

(1) $(A,+)$ est un groupe commutatif de neutre noté 0 et appelé zéro. *L'opposé de x pour $+$ est noté $-x$.*

(2) La loi $\cdot$

(a) est associative : $(x\cdot y)\cdot z = x\cdot(y\cdot z)$ pour tout $x,y,z\in A$.

(b) possède un élément neutre noté 1, qui est différent de 0, et appelé un *ou* unité *: $1\cdot x = x\cdot 1 = x$ pour tout $x\in A$.*

(3) La loi $\cdot$ est distributive sur l'addition $+$: pour tout $x,y,z\in A$, $x\cdot(y+z) = x\cdot y + x\cdot z$ et $(x+y)\cdot z = x\cdot z + y\cdot z$.

Si en outre la loi $\cdot$ est commutative, on dit que A est un anneau commutatif.

Remarque 57. *On a : $0\cdot x = x\cdot 0 = 0$ pour tout $x\in A$. En effet,*

$$0\cdot x = (0+0)\cdot x = 0\cdot x + 0\cdot x \Longrightarrow 0\cdot x = 0.$$

Également, on montre que $(-x)\cdot y = -(x\cdot y)$ car :

$$x\cdot y + (-x)\cdot y = (x+(-x))\cdot y = 0\cdot y = 0.$$

Désormais on note $x\cdot y = xy$.

Définition 58. *Soit $(A,+,\cdot)$ un anneau. Un élément $x\in A$ est appelé* un inversible *(ou une unité) de A s'il est inversible pour $\cdot$, c'est-à-dire s'il existe $y\in A$ tel que $xy = yx = 1$. On note $y = x^{-1}$. L'ensemble des inversibles de A est noté A^*.*

Proposition 59. *$(A^*,\cdot)$ est un groupe et $A^*\subset A\setminus\{0\}$.*

Démonstration. La loi $\cdot$ est bien interne dans A^* : si $x,y\in A^*$ alors xy est inversible d'inverse $y^{-1}x^{-1}$. On a $1\in A^*$ car 1 est inversible : $1\cdot 1 = 1$. Par définition, tout élément $x\in A^*$ est inversible, et son inverse x^{-1} est aussi dans A^* puisqu'il est inversible d'inverse x.

On a vu que $0x = 0$ pour tout x donc 0 n'est jamais inversible, d'où l'inclusion $A^*\subset A\setminus\{0\}$. □

Définition 60. $(A^*, \cdot)$ *est appelé groupe des inversibles (ou le groupe des unités) de* A.

Définition 61. *Si* A *est commutatif et si* $A^* = A \setminus \{0\}$, *c'est-à-dire si tout élément non nul de* A *est inversible, on dit que* A est un corps.

Définition 62. *Un* diviseur de zéro *d'un anneau* A *est un élément* $x \in A$ *tel que* $x \neq 0$ *et il existe* $y \neq 0$, $y \in A$, *avec* $xy = 0$ *ou* $yx = 0$. *Un anneau* A *sans diviseur de zéro s'appelle un* anneau intègre.

Remarque 63. *Si* $x \in A^*$ *alors* x *n'est pas un diviseur de zéro de* A. *En effet, si* $xy = 0$, *en multipliant à gauche par* x^{-1}, *on obtient* $y = 0$.

En particulier un corps est un anneau intègre.

III.2. Exemples.

Exemple 64. *Les ensembles* $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ *sont des anneaux commutatifs et intègres, pour les opérations d'addition et de multiplication usuelles. Les anneaux* $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ *sont mêmes des corps.*

Exemple 65. $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ *est un anneau. Il résulte du théorème de Bézout que le groupe* $(\mathbb{Z}/n\mathbb{Z})^*$ *des unités de* $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ *est d'ordre* $\varphi(n)$, *où* φ *est l'indicatrice d'Euler qui à tout entier naturel* n *non nul associe le nombre d'entiers compris entre* 1 *et* n *(inclus) et premiers avec* n. *En particulier,* $\mathbb{Z}/n\mathbb{Z}$ *est un corps si et seulement si* n *est premier. Si* n *n'est pas premier, les éléments non nuls et non inversibles, c'est-à-dire les* $a \bmod n$ *tels que* $\mathrm{pgcd}(a, n) > 1$, *sont tous diviseurs de zéro.*

Exemple 66. *Si* A *est un anneau commutatif, l'ensemble* $M_n(A)$ *des matrices carrées de taille* n *est un anneau de zéro la matrice nulle (i.e. dont tous les coefficients sont* 0*) et de* 1 *la matrice identité* Id_n. *Si* $n \geq 2$ *il n'est pas commutatif et possède des diviseurs de zéros.*

III.3. Sous-anneaux et idéaux. On traite ici du cas d'un anneau commutatif.

Définition 67. *Un sous-anneau* B *d'un anneau commutatif* $(A, +, \cdot)$ *est un sous-ensemble de* A *qui est un anneau pour les mêmes lois (en particulier, il contient* 1_A*).*

Proposition 68. $B \subset A$ *est un sous-anneau de* A *s'il vérifie les conditions suivantes :*

(1) $1_A \in B$,

(2) Pour tout $x, y \in B$, $x - y \in B$,

(3) Pour tout $x, y \in B$, $xy \in B$.

Démonstration. Les conditions (1) et (2) garantissent que $(B, +)$ est un sous-groupe de $(A, +)$ d'après la proposition 10, et la condition (3) que la mutiplication est une loi de composition interne pour B. Comme $1_A \in B$, la multiplication possède bien un élément neutre dans B. Les autres propriétés définissant un anneau sont vraies pour A donc aussi pour B. □

Définition 69. *Un sous-ensemble* $I \subset A$ *est un* idéal *de* A *si :*

(1) $(I, +)$ *est un sous-groupe de* $(A, +)$.

(2) Pour tout $x \in I$, *et tout* $a \in A$, $ax \in I$.

Exemple 70. $\{0\}$ *et* A *sont des idéaux de* A.

Proposition 71. *Soit* I *un idéal de* A. *Alors* I *contient un élément inversible de* A *si et seulement si* $I = A$. *En particulier, si* A *est un corps, ses seuls idéaux sont* $\{0\}$ *et* A.

Démonstration. La condition est clairement nécessaire puisque, si $I = A$, $1 \in I$. Réciproquement, si $x \in I \cap A^*$, alors, en prenant $a = x^{-1}$ dans la définition d'un idéal, I contient $x^{-1}x = 1$ donc $y \cdot 1 = y$ pour tout $y \in A$. □

Exemple 72. *Les idéaux de $\mathbb{Z}$ sont les $n\mathbb{Z}$. En effet, on a déjà vu que les sous-groupes de $\mathbb{Z}$ sont les $n\mathbb{Z}$ donc un idéal de $\mathbb{Z}$ est nécessairement de la forme $n\mathbb{Z}$. De plus, si $x \in n\mathbb{Z}$, c'est-à-dire $x = nq$ pour un $q \in \mathbb{Z}$, et si $a \in \mathbb{Z}$, alors $ax = naq \in n\mathbb{Z}$. Ainsi les $n\mathbb{Z}$ sont bien des idéaux de $\mathbb{Z}$.*

Définition 73. *Si S est une partie de A, on note (S) le plus petit idéal de A contenant S. On l'appelle* l'idéal engendré par S.

On a le cas particulier suivant.

Définition 74. *Un* idéal principal *I d'un anneau A est un idéal engendré par $\{x\}$, où $x \in A$. On dit que x est* un générateur *de I. On note aussi $I = (x)$.*

Notons

$$Ax = \{ax : a \in A\}.$$

Proposition 75. *Pour tout x dans A, $(x) = Ax$.*

Démonstration. Il suffit de prendre $a = 1$ pour montrer que $x \in Ax$. De plus Ax est un idéal de A. En effet :

- $0 \in I$ et, si $y = ax \in I$ et $y' = a'x \in I$, $y - y' = ax - a'x = (a - a')x \in I$. Donc $(I, +)$ est bien un sous-groupe de $(A, +)$.
- Si $y = ax \in I$ et si $b \in A$, $by = bax = (ba)x \in I$.

Enfin, par définition, si I est un idéal de A contenant x alors $Ax \subset I$. Ainsi Ax est bien le plus petit idéal de A contenant x. □

Définition 76. *Un anneau commutatif et intègre dont tous les idéaux sont principaux est appelé un* anneau principal.

Exemple 77. *$\mathbb{Z}$ est un anneau principal.*

Les propriétés arithmétiques de $\mathbb{Z}$ s'étendent à un anneau A principal, et donc en particulier à $K[X]$, où K est un corps, ce que nous verrons en détail dans le chapitre suivant.

Remarque 78. *On a supposé dans tout ce chapitre que les anneaux considérés sont commutatifs. Dans le cas d'un anneau non commutatif, quelques nuances s'imposent : on distingue* idéaux à gauche *et* idéaux à droite*, vérifiant respectivement $a \in A, x \in I \Longrightarrow ax \in I$ et $a \in A, x \in I \Longrightarrow xa \in I$. Un idéal à droite et à gauche est appelé un* idéal bilatère.

CHAPITRE 2

Polynômes à coefficients dans un corps

Ce chapitre est consacré à l'étude des polynômes à coefficients dans un corps K.

I. Généralités

Définition 79. *On a appelle* polynôme à une indéterminée *à coefficients dans K toute suite $(a_k)_{k\in\mathbb{N}}$ d'éléments de K tous nuls à partir d'un certain rang. On note $K[X]$ l'ensemble des polynômes à une indéterminée à coefficients dans K (cette notation sera justifiée ci-dessous).*

Définition 80. *Soit $P = (a_k)_{k\in\mathbb{N}} \in K[X]$. On définit le* degré *de P, noté* $\deg P$ *ou* $d^\circ P$*, comme étant le plus grand entier d tel que $a_d \neq 0$. Par convention, si $a_k = 0$ pour tout $k \in \mathbb{N}$, $\deg P = -\infty$.*

Un polynôme de degré nul est dit constant*, égal à a_0.*

Un polynôme est dit unitaire *si son coefficient de plus haut degré vaut* 1.

On dit d'un polynôme que c'est un monôme *si tous ses coefficients sont nuls, sauf un.*

Définition 81. *Si $P = (a_k)_{k\in\mathbb{N}}$ et $Q = (b_k)_{k\in\mathbb{N}}$ sont deux éléments de $K[X]$, et si $\lambda \in K$, on définit la* somme $P + Q = (a_k + b_k)_{k\in\mathbb{N}} \in K[X]$ *et le* produit $\lambda \cdot P = (\lambda a_k)_{k\in\mathbb{N}}$.

Remarque 82. *Il est évident que* $\deg(P + Q) \leq \max\{\deg P, \deg Q\}$ *et qu'il n'y a pas forcément égalité (l'inégalité est stricte si P et Q ont même degré, et si leurs coefficients de plus haut degré sont opposés).*

Proposition-Définition 83. *Si $P = (a_k)_{k\in\mathbb{N}}$ et $Q = (b_k)_{k\in\mathbb{N}}$ sont deux éléments de $K[X]$, on définit leur* produit *par $P \cdot Q = PQ = (c_k)_{k\in\mathbb{N}}$ où $c_k = \sum_{i=0}^{k} a_i b_{k-i} = \sum_{i=0}^{k} a_{k-i} b_i = \sum_{i+j=k} a_i b_j$. Le produit vérifie* $\deg(PQ) = \deg P + \deg Q$*. En particulier, $PQ = 0$ si et seulement si $P = 0$ ou $Q = 0$.*

Démonstration. On vérifie que PQ est bien un polynôme car, si $k > \deg P + \deg Q$ alors

$$c_k = \sum_{i=0}^{k} a_i b_{k-i} = \sum_{i=0}^{\deg P} a_i b_{k-i} + \sum_{i=\deg P+1}^{k} a_i b_{k-i} = 0$$

car dans la première somme, $k - i > \deg P + \deg Q - \deg P = \deg Q$ donc $b_{k-i} = 0$, alors que dans la deuxième somme, $i > \deg P$ donc $a_i = 0$. Par ailleurs, par le même raisonnement, on montre que le terme de degré $(\deg P + \deg Q)$ de PQ est $a_{\deg P} b_{\deg Q} \neq 0$ et ainsi, $\deg(PQ) = \deg P + \deg Q$. Remarquons que cette formule est également valable si l'un des deux polynômes est nul car dans ce cas $PQ = 0$.

Par conséquent, $PQ = 0$ si et seulement si $P = 0$ ou $Q = 0$. □

La proposition suivante est facile à montrer à partir des définitions.

Proposition 84. *Le produit dans $K[X]$ est associatif, commutatif, distributif par rapport à l'addition. Il existe un élément neutre pour le produit, il s'agit du polynôme constant égal à* 1*. Enfin, il revient au même de multiplier un polynôme par un scalaire que de le multiplier par le polynôme constant égal à ce scalaire, ce qui justifie que l'on note de la même manière la multiplication dans $K[X]$ et la multiplication par les scalaires.*

On note $1 = (1, 0, \dots, 0, \dots)$ le monôme unitaire de degré 0 et $X = (0, 1, 0, \dots, 0, \dots)$ le monôme unitaire de degré 1. On voit alors facilement que pour tout $n \geq 1$, X^n (i.e. le produit de X avec lui-même n fois) est le monôme unitaire de degré n. On convient que $X^0 = 1$. Par conséquent, on écrira dans la suite un polynôme $P = (a_k)_{k\in\mathbb{N}}$ sous la forme $P = \sum_{k\in\mathbb{N}} a_k X^k = \sum_{k=0}^{\deg P} a_k X^k$.

En particulier, si $P = \sum_{k\in\mathbb{N}} a_k X^k$ et $Q = \sum_{k\in\mathbb{N}} b_k X^k$ sont deux éléments de $K[X]$, et si $\lambda \in K$, $P + Q = \sum_{k\in\mathbb{N}}(a_k + b_k)X^k$ et $\lambda \cdot P = \sum_{k\in\mathbb{N}}(\lambda a_k)X^k$. Pour calculer le produit PQ, on utilise la règle habituelle de développement du produit de deux sommes.

Proposition 85. *L'ensemble $(K[X], +, \cdot)$ est un anneau commutatif intègre. Les éléments neutres sont respectivement le polynôme nul (i.e. dont tous les coefficients sont nuls) noté 0, et le polynôme constant égal à 1. L'ensemble $K[X]$ est également un K-espace vectoriel de dimension infinie. Pour tout $N \in \mathbb{N}$, l'ensemble $K_N[X] = \{P \in K[X]\,,\, \deg P \leq N\}$ des polynômes de degré inférieur ou égal à N est un sous-espace vectoriel de $K[X]$ de dimension $N + 1$.*

Démonstration. Le fait que $K[X]$ muni des deux lois ci-dessus est un anneau découle de la proposition précédente, et le fait que c'est un K-espace vectoriel est immédiat. Soit maintenant une famille finie $(P_1, \ldots, P_m)$ d'éléments de $K[X]$ et soit $d = \max\{\deg P_i\,,\, i = 1, \ldots, m\}$. Alors, le polynôme unitaire dont tous les coefficients sont nuls, sauf celui de degré $d+1$, ne peut pas être combinaison linéaire de $P_1, \ldots, P_m$, donc aucune famille finie ne peut être génératrice de $K[X]$.

Le fait que $K_N[X]$ est un sous-espace vectoriel provient du fait que $\deg 0 = -\infty$, $\deg(P + Q) \leq \max\{\deg P, \deg Q\}$ et $\deg(\lambda P) = \deg P$ (sauf si $\lambda = 0$, mais dans ce cas, $\lambda P = 0$). Il est clair que $(X^k)_{0\leq k\leq N}$ est une famille génératrice de $K_N[X]$, et elle est également libre car par définition, un polynôme est nul si et seulement si tous ses coefficients sont nuls. □

Remarque 86. *La remarque 82 montre que l'ensemble des éléments de $K[X]$ de degré exactement N n'est* pas *un sous-espace vectoriel de $K[X]$.*

Définition 87. *Soit $P = \sum_{i\in\mathbb{N}} a_k X^k \in K[X]$. Le* polynôme dérivé *de P, noté P', est le polynôme défini par $P' = \sum_{k\geq 1} k a_k X^{k-1} = \sum_{k\in\mathbb{N}}(k + 1)a_{k+1}X^k$.*

Remarquons que $\deg P' = \deg P - 1$, sauf si P est constant, auquel cas $P' = 0$. En particulier, la dérivation préserve tous les sous-espaces vectoriels $K_N[X]$ de $K[X]$.

Proposition 88. *L'application de $K[X]$ dans $K[X]$ qui à un polynôme associe son polynôme dérivé est linéaire. De plus, pour tous $P, Q \in K[X]$, $(PQ)' = P'Q + PQ'$.*

Démonstration. La linéarité de la dérivation est immédiate. Montrons la formule pour la dérivée du produit. Si $P = \sum_{k\in\mathbb{N}} a_k X^k$ et $Q = \sum_{\ell\in\mathbb{N}} b_\ell X^\ell$, en utilisant en particulier la linéarité de la dérivation on obtient

$$\begin{aligned}
(PQ)' &= \left[(\textstyle\sum_{k\in\mathbb{N}} a_k X^k)(\sum_{k\in\mathbb{N}} b_k X^k)\right]' \\
&= \left[\textstyle\sum_{k\in\mathbb{N}} \sum_{i+j=k} a_i b_j X^k\right]' \\
&= \textstyle\sum_{k\in\mathbb{N}} \sum_{i+j=k} a_i b_j \left(X^k\right)' \\
&= \textstyle\sum_{k\geq 1} \sum_{i+j=k} a_i b_j k X^{k-1} \\
&= \textstyle\sum_{k\geq 1} \sum_{i+j=k} a_i b_j (i+j) X^{k-1} \\
&= \textstyle\sum_{k\geq 1} \sum_{i+j=k} i a_i b_j X^{k-1} + \sum_{k\geq 1} \sum_{i+j=k} a_i j b_j X^{k-1} \\
&= \textstyle\sum_{k\in\mathbb{N}} \sum_{i+j=k} (i+1) a_{i+1} b_j X^k + \sum_{k\in\mathbb{N}} \sum_{i+j=k} a_i (j+1) b_{j+1} X^k \\
&= (\textstyle\sum_{k\in\mathbb{N}}(k+1)a_{k+1}X^k)(\sum_{k\in\mathbb{N}} b_k X^k) + (\sum_{k\in\mathbb{N}} a_k X^k)(\sum_{k\in\mathbb{N}}(k+1)b_{k+1}X^k) \\
&= P'Q + PQ'
\end{aligned}$$

□

Définition 89. *Pour tout $n \geq 1$, on définit par récurrence la* dérivée d'ordre n *d'un polynôme P par $P^{(n)} = (P^{(n-1)})'$. Par convention, $P^{(0)} = P$, $P^{(1)} = P'$.*

Remarque 90.

$$\deg P^{(n)} = \begin{cases} \deg P - n & \text{si } \deg P \geq n \\ -\infty & \text{si } \deg P \leq n-1 \end{cases}$$

Définition 91. *Soient $P = \sum_{k\in\mathbb{N}} a_k X^k$ et $Q = \sum_{k\in\mathbb{N}} b_k X^k$ deux éléments de $K[X]$. Le polynôme* composé *de P par Q est le polynôme noté $P \circ Q$ ou $P(Q)$ défini par*

$$P \circ Q = \sum_{k\in\mathbb{N}} a_k Q^k = \sum_{k\in\mathbb{N}} a_k \left(\sum_{\ell\in\mathbb{N}} b_\ell X^\ell \right)^k$$

i.e. obtenu en "remplaçant" X par Q dans l'expression de P. On a $\deg(P \circ Q) = \deg P \deg Q$.

Le fait que $\deg(P \circ Q) = \deg P \deg Q$ est facile à établir : si $P, Q \neq 0$, et si $\deg P = n$ et $\deg Q = m$, on vérifie facilement que le terme de plus haut degré de $P \circ Q$ est $a_n b_m^n X^{mn}$.

Définition 92. *Soit $P = \sum_{k\in\mathbb{N}} a_k X^k \in \mathbb{C}[X]$. Le polynôme* conjugué *de P est le polynôme $\bar{P}$ dont les coefficients sont les conjugués des coefficients de P i.e.*

$$\bar{P} = \sum_{k\in\mathbb{N}} \bar{a}_k X^k.$$

Remarque 93. *Les propriétés usuelles de la conjugaison sur $\mathbb{C}$ sont vraies dans $\mathbb{C}[X]$:*

$$\overline{P+Q} = \bar{P} + \bar{Q},\ \overline{PQ} = \bar{P}\bar{Q},\ \bar{\bar{P}} = P.$$

Par ailleurs, P est à coefficients réels si et seulement si $\bar{P} = P$. Enfin, $(\bar{P})' = \overline{P'}$.

II. Arithmétique dans $K[X]$

On remarquera que l'arithmétique dans $K[X]$ ressemble à beaucoup d'égards à celle dans $\mathbb{Z}$, en remplaçant généralement dans les énoncés la relation d'ordre sur les nombres entiers par la relation d'ordre sur les degrés des polynômes.

II.1. Division euclidienne et divisibilité.

Théorème 94.

(Division euclidienne). *Soient $A, B \in K[X]$, avec $B \neq 0$. Il existe un unique couple (Q, R) d'éléments de $K[X]$ tel que*

$$A = BQ + R \quad \text{avec} \quad \deg R < \deg B.$$

Le polynôme Q (resp. R) est appelé le quotient (resp. reste) de la division euclidienne de A par B.

Démonstration. Commençons par montrer l'unicité. Si $A = BQ_1 + R_1 = BQ_2 + R_2$ avec $\deg R_i < \deg B$, alors $B(Q_1 - Q_2) = R_2 - R_1$. Or, si $Q_1 \neq Q_2$, $\deg B(Q_1 - Q_2) \geq \deg B$ et par ailleurs, $\deg(R_2 - R_1) \leq \max\{\deg R_1, \deg R_2\} < \deg B$. Donc $Q_1 = Q_2$ et par conséquent, $R_1 = R_2$.

La preuve de l'existence se fait par "récurrence descendante", au moyen de l'algorithme d'Euclide.

Si $\deg A < \deg B$, il n'y a rien à faire : $Q = 0$ et $R = A$. Sinon, soit $a_{k_0} X^{k_0}$ le terme de plus haut degré de A, et $b_\ell X^\ell$ le terme de plus haut degré de B, en particulier a_{k_0} et b_ℓ sont non nuls, et $k_0 \geq \ell$. Considérons la polynôme $A_1 = A - B\frac{a_{k_0}}{b_\ell} X^{k_0-\ell}$. Comme les polynômes A et $B\frac{a_{k_0}}{b_\ell} X^{k_0-\ell}$ ont même degré et même terme de plus haut degré, $\deg A_1 < \deg A$.

Ensuite, on recommence avec le polynôme A_1. Si $\deg A_1 < \deg B$, on pose $Q = \frac{a_{k_0}}{b_\ell} X^{k_0-\ell}$ et $R = A_1$, sinon, en notant $a_{k_1} X^{k_1}$ le terme de plus haut degré de A_1, on définit $A_2 = A_1 - B\frac{a_{k_1}}{b_\ell} X^{k_1-\ell}$ et $\deg A_2 < \deg A_1$.

On construit ainsi des polynômes A_i tels que $A_{i+1} = A_i - B\frac{a_{k_i}}{b_\ell} X^{k_i-\ell}$ et $\deg A_{i+1} < \deg A_i$. Puisque la suite des $\deg A_i$ est strictement décroissante, après un nombre fini p d'étapes, on a forcément $\deg A_p < \deg B$. Il suffit alors de prendre

$$Q = \sum_{i=0}^{p-1} \frac{a_{k_i}}{b_\ell} X^{k_i-\ell} \quad \text{et} \quad R = A_p.$$

□

Définition 95. *Soient $A, B \in K[X]$, avec $B \neq 0$. Si le reste de la division euclidienne de A par B est nul, on dit que B* divise *A et on note $B|A$. On dit aussi que A est un* multiple *de B.*

Remarque 96. *Il est immédiat que :*

(1) tout polynôme B divise 0 puisque $0 = B0$;

(2) si $A|B$ et $A|C$ alors $A|(B + C)$;

(3) si $A|B$ et $B|C$ alors $A|C$.

Nous finissons cette section par un corollaire du théorème de division euclidienne concernant les idéaux de $K[X]$.

Proposition 97. *L'anneau $K[X]$ est principal. De plus, tout idéal $\mathfrak{I} \neq \{0\}$ de $K[X]$ admet un unique générateur unitaire.*

Démonstration. Si $\mathfrak{I} = \{0\}$, il est engendré par 0. Sinon, soit P un polynôme non nul dans $\mathfrak{I}$ et de degré minimal. Comme $\mathfrak{I}$ est un idéal, il contient tous les multiples de P. Inversement si A est un élément quelconque de $\mathfrak{I}$, on effectue la division euclidienne $A = PQ + R$. Comme $\mathfrak{I}$ est un idéal, $R = A - PQ$ appartient à $\mathfrak{I}$, et son degré est strictement inférieur à celui de P. Ainsi R est nul, à cause du choix de P, et A est donc un multiple de P.

Enfin, si P engendre $\mathfrak{I} \neq \{0\}$, et si le coefficient de son terme de plus haut degré est λ, clairement $\lambda^{-1}P$ est unitaire et engendre également $\mathfrak{I}$. Enfin, si deux polynômes unitaires P_1 et P_2 engendrent $\mathfrak{I}$, par définition chacun est divisible par l'autre donc $P_1 = P_2Q_1$ et $P_2 = P_1Q_2$ dont on déduit $P_1 = Q_1Q_2P_1$ ce qui implique que Q_1 et Q_2 sont constants. Mais comme P_1 et P_2 sont unitaires, $Q_1 = Q_2 = 1$. □

II.2. PGCD.

Théorème 98.

Soient $A, B \in K[X]$, l'un des deux au moins n'étant pas nul. Alors, il existe un unique $D \in K[X]$ unitaire tel que

(1) D divise A et B ;

(2) si Q divise A et B, alors Q divise D.

Le polynôme D est appelé le plus grand commun diviseur *de A et B et on le note* $\operatorname{pgcd}(A, B)$.

Démonstration. Montrons d'abord l'unicité du pgcd. Soient D_1 et D_2 qui réalisent les conditions du théorème. D'après 2, $D_1|D_2$ (donc $\deg D_1 \leq \deg D_2$) et $D_2|D_1$ (donc $\deg D_2 \leq D_1$). Par conséquent, $\deg D_1 = \deg D_2$ et comme $D_1|D_2$, $D_2 = \lambda D_1$ avec $\lambda \in K^\star$ (le quotient de la division euclidienne de D_2 par D_1 est de degré 0 car ils sont de même degré). Enfin, D_1 et D_2 sont tous deux unitaires donc $\lambda = 1$.

Montrons maintenant l'existence du pgcd. Nous définissons de proche en proche une suite de polynômes P_i par : $P_0 = A$, $P_1 = B$ puis, pour tout $i \geq 1$, si $P_i \neq 0$, P_{i+1} est le reste de la division euclidienne de P_{i-1} par P_i (en particulier, $\deg P_{i+1} < \deg P_i$). La suite des $\deg P_i$ (pour $i \geq 1$) est donc strictement décroissante et par conséquent, il existe $r \geq 0$ tel que $P_{r+1} = 0$. Le polynôme P_r est non nul et nous allons voir qu'il vérifie les conditions 1 et 2 du théorème.

Si $r = 0$, i.e. $B = 0$, il est clair que $P_0 = A$ divise A et B, et que c'est le pgcd de A et B (quitte à multiplier A par l'inverse du coefficient de son terme de plus haut degré pour le rendre unitaire).

Sinon, $r \geq 1$ et pour tout i compris entre 1 et r, nous avons $P_{i-1} = P_iQ_i + P_{i+1}$. Or, il est clair qu'un polynôme divise P_{i-1} et P_i si et seulement si il divise P_i et P_{i+1} et par conséquent, un polynôme divise A et B si et seulement si il divise P_r et $P_{r+1} = 0$ i.e. si et seulement si il divise P_r. Dès lors, il est clair que P_r vérifie les conditions 1 et 2. Il ne reste plus qu'à le multiplier par l'inverse du coefficient de son terme de plus haut degré pour obtenir le pgcd de A et B. □

Remarque 99. *Comme on le voit dans la preuve, si on ne demande pas au* pgcd *d'être unitaire, il n'est pas unique car si P réalise les conditions 1 et 2 du théorème, pour tout $\lambda \in K^\star$, λP les réalise également. Dans la suite, un tel polynôme non nécessairement unitaire sera néanmoins parfois aussi appelé* un pgcd.

Définition 100. *Soient $A, B \in K[X]$, l'un des deux au moins n'étant pas nul. On dit qu'ils sont* premiers entre eux *si leur* pgcd *est égal à* 1.

Ce qui précède se généralise facilement au cas de plus de deux polynômes.

Théorème 101.

Soient $A_1, \ldots, A_n \in K[X]$ des polynômes non tous nuls. Il existe un unique polynôme unitaire D qui vérifie :

(1) D divise chacun des A_i ;

(2) tout polynôme qui divise chacun des A_i divise D.

Le polynôme D est appelé le plus grand commun diviseur *des A_i et on le note* $\operatorname{pgcd}(A_1, \ldots, A_n)$. *Si* $\operatorname{pgcd}(A_1, \ldots, A_n) = 1$, *on dit que les A_i sont* premiers entre eux dans leur ensemble.

Démonstration. L'unicité se montre comme pour le cas du pgcd de deux polynômes.

L'existence de $\operatorname{pgcd}(A_1, \ldots, A_n)$ se montre par récurrence en vérifiant que si $\operatorname{pgcd}(A_1, \ldots, A_{n-1})$ existe alors $\operatorname{pgcd}(\operatorname{pgcd}(A_1, \ldots, A_{n-1}), A_n)$ a les propriétés voulues. □

II.3. Théorème de Bézout. Algorithme d'Euclide étendu.

Théorème 102.

(Bézout). *Soient $A, B \in K[X]$, l'un des deux au moins n'étant pas nul. Alors, si* $\operatorname{pgcd}(A, B) = D$, *il existe $U, V \in K[X]$ tels que $AU + BV = D$. Si B n'est pas nul, on obtient U et V par le procédé algorithmique suivant : soient P_i, Q_i, U_i et V_i les polynômes calculés itérativement à partir des initialisations :*

$$\begin{array}{lll} P_0 = A & U_0 = 1 & V_0 = 0 \\ P_1 = B & U_1 = 0 & V_1 = 1 \end{array}$$

par division euclidienne :

$$P_{i-1} = P_i Q_i + P_{i+1}$$

où $\deg(P_{i+1}) < \deg(P_i)$, *et les formules*

$$\begin{array}{l} U_{i+1} = U_{i-1} - U_i Q_i \\ V_{i+1} = V_{i-1} - V_i Q_i \end{array},$$

tant que P_i n'est pas nul. Soit r le plus grand entier tel que P_r est non nul. Alors P_r est un pgcd de A et B et $P_r = AU_r + BV_r$.

Démonstration. Nous allons reprendre l'algorithme d'Euclide pour montrer l'existence de U et V (de cette façon nous aurons un procédé pour trouver U et V dans la pratique que l'on appelle algorithme d'Euclide étendu).

Nous avons posé $P_0 = A$, $P_1 = B$, puis avons trouvé des polynômes P_i $(2 \leq i \leq r)$ et Q_i $(1 \leq i \leq r-1)$ tels que

$$\begin{array}{rcl} P_2 & = & P_0 - P_1 Q_1 \\ P_3 & = & P_1 - P_2 Q_2 \\ \vdots & = & \vdots \\ P_r & = & P_{r-2} - P_{r-1} Q_{r-1} \end{array}$$

où $P_r \neq 0$ est, à une constante multiplicative près, le pgcd de A et B (i.e. $P_{r+1} = 0$). L'idée consiste à exprimer progressivement chaque P_i $(i \geq 2)$ en fonction de P_0 et P_1. Pour P_2, c'est déjà fait. Ensuite,

$P_3 = P_1 - P_2Q_2 = P_1 - (P_0 - P_1Q_1)Q_2 = -P_0Q_2 + P_1(1 + Q_1Q_2)$ etc, jusqu'à arriver à exprimer P_r en fonction de P_0 et P_1.

Plus précisément, soit i, $2 \leq i \leq r-1$, et supposons que pour tout k tel que $2 \leq k \leq i$ on ait trouvé U_k, V_k tels que $P_k = AU_k + BV_k$. Alors,

$$P_{i+1} = P_{i-1} - P_iQ_i = (AU_{i-1} + BV_{i-1}) - (AU_i + BV_i)Q_i = A(U_{i-1} - U_iQ_i) + B(V_{i-1} - V_iQ_i)$$

i.e. $U_{i+1} = U_{i-1} - U_iQ_i$ et $V_{i+1} = V_{i-1} - V_iQ_i$. Ceci nous fournit donc un procédé algorithmique pour trouver $U = U_r$ et $V = V_r$. □

Remarque 103.

(1) Il n'y a jamais unicité du couple (U, V) vérifiant $AU + BV = D$. En effet, si (U_0, V_0) est solution alors pour tout $W \in K[X]$, $(U, V) = (U_0 + BW, V_0 - AW)$ est également solution.

(2) En général, s'il existe U, V et P tels que $AU + BV = P$, cela ne signifie pas que $\operatorname{pgcd}(A, B) = P$, *mais seulement que P est un multiple de* $\operatorname{pgcd}(A, B)$ *(en effet* $\operatorname{pgcd}(A, B)$ *divise à la fois A et B, donc P). Cependant, si $P = 1$, son seul diviseur unitaire est lui-même donc on obtient le corollaire suivant.*

Corollaire 104. *Soient $A, B \in K[X]$, l'un des deux au moins n'étant pas nul. Alors, A et B sont premiers entre eux si et seulement si il existe $U, V \in K[X]$ tels que $AU + BV = 1$. En outre, on peut choisir U et V de sorte que* $\deg U < \deg B$ *et* $\deg V < \deg A$.

Démonstration. Il ne reste plus qu'à montrer que l'on peut réaliser les conditions sur les degrés de U et V. Nous savons qu'il existe $U, V \in K[X]$ tels que $AU + BV = 1$. Effectuons la division euclidienne de U par B : il existe $Q, R \in K[X]$ tels que $U = BQ + R$ et $\deg R < \deg B$. On a alors $A(BQ + R) + BV = AR + B(AQ + V) = 1$ i.e. $B(AQ + V) = 1 - AR$. Le membre de droite est de degré inférieur ou égal à $\deg A + \deg R < \deg A + \deg B$. Par conséquent, $\deg(AQ + V) < \deg A$. □

Remarque 105. *Soient $A, B \in K[X]$ non nuls. Soit $D \in K[X]$, unitaire, divisant A et B. Il existe donc $A_1, B_1 \in K[X]$ tels que $A = A_1D$ et $B = B_1D$. Si de plus* $\operatorname{pgcd}(A_1, B_1) = 1$ *alors* $\operatorname{pgcd}(A, B) = D$.

En effet, d'après le théorème de Bézout, il existe $U, V \in K[X]$ tels que $A_1U + B_1V = 1$ et en multipliant cette égalité par D on obtient $AU + BV = D$. D'après le 2 de la remarque 103, $\operatorname{pgcd}(A, B)$ *divise D. Mais comme D divise A et B, ceci entraîne que* $\operatorname{pgcd}(A, B) = D$.

Avant de poursuivre, donnons un exemple de détermination de U et V dans le théorème de Bézout.

Exemple 106. *Un exemple d'exécution de l'algorithme d'Euclide étendu sur les polynômes $A = X^4 - 1$, $B = X^3 - 2X + 1$.*

P_i	U_i	V_i	Q_i	
$X^4 - 1$	1	0		
$X^3 - 2X + 1$	0	1	X	$A = BX + (2X^2 - X - 1)$
$2X^2 - X - 1$	1	$-X$	$(2X + 1)/4$	$B = P_2(X/2 + 1/4) - 5/4(X - 1)$
$-5/4(X - 1)$	$-(2X + 1)/4$	$(2X^2 + X + 4)/4$	$-4/5(2X + 1)$	$P_2 = -4/5P_3(2X + 1)$
0				

On obtient que $(X - 1)$ est le pgcd de A et B, et la relation de Bézout :

$$-5(X - 1) = -(2X + 1)A + (2X^2 + X + 4)B.$$

Nous allons maintenant voir deux conséquences intéressantes du théorème de Bézout.

Corollaire 107. (Lemme de Gauss). *Soient $A, B, C \in K[X]$. Supposons que $C|AB$ et que* $\operatorname{pgcd}(A, C) = 1$. *Alors $C|B$.*

Démonstration. D'après le théorème de Bézout, il existe $U, V \in K[X]$ tels que $AU + CV = 1$. Par conséquent, $ABU + CBV = B$ et comme C divise AB, il divise aussi B. □

Remarque 108. *Grâce au lemme de Gauss, on peut montrer que si* $\text{pgcd}(A,B)=1$*, et si* $U_0,V_0\in K[X]$ *sont tels que* $AU_0+BV_0=1$*, alors l'ensemble des solutions de* $AU+BV=1$ *est* $\mathcal{S}=\{(U_0+BW,V_0-AW)\,,\,W\in K[X]\}$*. On a déjà vu (voir 1 de la remarque 103) que tous les éléments de* $\mathcal{S}$ *sont bien solutions. Montrons la réciproque.*

Si (U,V) *est une solution, on a* $A(U-U_0)=-B(V-V_0)$*. Donc, d'après le lemme de Gauss* A *divise* $(V-V_0)$*, puisque* $\text{pgcd}(A,B)=1$*. Il existe donc* $W\in K[X]$ *tel que* $V=V_0-AW$ *et par conséquent,* $AU=AU_0+BAW$ *i.e.* $U=U_0+BW$*.*

Corollaire 109. *Soient* $A,B,C\in K[X]$*. Supposons que* $A|C$ *et* $B|C$*, et que* $\text{pgcd}(A,B)=1$*. Alors,* $AB|C$*.*

Démonstration. D'après le théorème de Bézout, il existe $U,V\in K[X]$ tels que $AU+BV=1$. Par conséquent, $ACU+BCV=C$. Comme A divise C, AB divise BC. De même, comme B divise C, AB divise AC. D'où, AB divise C. □

II.4. PPCM.

Théorème 110.

Soient $A_1,\dots,A_n\in K[X]$*, aucun n'étant nul. Il existe un unique polynôme unitaire* M *qui vérifie :*

(1) chaque A_i *divise* M *;*

(2) tout polynôme qui est un multiple de chacun des A_i *est divisible par* M*.*

Le polynôme M *est appelé le* plus petit commun multiple *des* A_i*, et on le note* $\text{ppcm}(A_1,\dots,A_n)$*.*

Démonstration. Nous allons montrer que le ppcm s'obtient facilement à partir du pgcd de la façon suivante : pour $n=2$, $\text{ppcm}(A_1,A_2)$ est le quotient de A_1A_2 par $\text{pgcd}(A_1,A_2)$ (à une constante multiplicative près, pour le rendre unitaire) et pour tout $k\geq 2$,

$$\text{ppcm}(A_1,\dots,A_{k+1})=\text{ppcm}(\text{ppcm}(A_1,\dots,A_k),A_{k+1}).$$

Cette deuxième propriété, comme pour le pgcd, est évidente. Montrons la première (l'unicité du ppcm de deux polynômes se montre de manière similaire à celle du pgcd).

Notons $D=\text{pgcd}(A_1,A_2)$. Par hypothèse, il existe $B_1,B_2\in K[X]$ tels que $A_1=DB_1$ et $A_2=DB_2$. Remarquons que $\text{pgcd}(B_1,B_2)=1$ par maximalité du pgcd. Nous voulons montrer que $DB_1B_2=A_1B_2=B_1A_2$ est (toujours à une constante multiplicative près) le ppcm de A_1 et A_2. Il est clair que c'est un multiple de A_1 et de A_2. Prenons un multiple P de A_1 et A_2. Il existe $U,V\in K[X]$ tels que $P=A_1U=A_2V$. En divisant par D, on obtient $B_1U=B_2V$. Comme $\text{pgcd}(B_1,B_2)=1$ et que B_1 divise B_2V, B_1 divise V : il existe $W\in K[X]$ tel que $V=B_1W$. Finalement, nous avons $P=A_2V=DB_1B_2W$, donc P est divisible par DB_1B_2. □

Remarque 111. *Dans la preuve du dernier théorème, il apparaît en particulier qu'il existe* $\lambda\in K^\star$ *tel que* $\text{pgcd}(A,B)\ \text{ppcm}(A,B)=\lambda AB$*.*

II.5. Irréductibilité.

Proposition-Définition 112. *Soit* $P\in K[X]$*. On dit que* P *est* irréductible *dans* $K[X]$ *si* $\deg P\geq 1$ *et si pour tous* $A,B\in K[X]$*,* $AB=P$ *implique* $\deg A=0$ *ou* $\deg B=0$*. Il est équivalent de dire que pour tout* $Q\in K[X]$*,* $\text{pgcd}(P,Q)=1$ *ou* $\lambda^{-1}P$*, avec* λ *le coefficient du terme de plus haut degré de* P*.*

Démonstration. Supposons d'abord que P est irréductible. Soit $Q\in K[X]$ et $D=\text{pgcd}(P,Q)$: il existe $C\in K[X]$ tel que $P=CD$. Mais par définition, $\deg C=0$ ou $\deg D=0$, d'où le résultat.

Réciproquement, si P n'est pas irréductible, il existe $A,B\in K[X]$ tous deux de degré strictement positif tels que $P=AB$. Quitte à multiplier A par l'inverse du coefficient de son terme de plus haut degré, on peut prendre A unitaire. Mais alors, il est clair que $\text{pgcd}(P,A)=A$ qui est distinct de 1 et P. □

Remarque 113. *Il est immédiat d'après la définition qu'un polynôme de degré 1 est irréductible.*

Proposition 114. *Soient $P, Q_1, Q_2, \dots, Q_n \in K[X]$. On suppose que P est irréductible et divise $\prod_{i=1}^{n} Q_i$. Alors P divise un des Q_i.*

Démonstration. Ceci peut se montrer par récurrence sur n. Le résultat est trivialement vrai pour $n = 1$. Supposons qu'il est vrai pour un certain $n \in \mathbb{N}^\star$, et supposons que P divise $\prod_{i=1}^{n+1} Q_i$. Comme P est irréductible, $\operatorname{pgcd}(P, Q_{n+1})$ vaut soit 1, soit P (à un coefficient multiplicateur près). Dans le deuxième cas, P divise Q_{n+1}. Dans le premier cas, P divise $\prod_{i=1}^{n} Q_i$ d'après le lemme de Gauss, et on peut appliquer l'hypothèse de récurrence : P divise l'un des Q_i, $i \leq n$. □

Remarque 115. *La notion d'irréductibilité de P est liée à K. Par exemple $P = X^2 + 1$ est irréductible dans $\mathbb{R}[X]$ mais pas dans $\mathbb{C}[X]$, dans lequel $P = (X + i)(X - i)$.*

Théorème 116.

Soit $P \in K[X]$, $\deg P \geq 1$. Alors P se décompose de manière unique (à l'ordre des facteurs près) sous la forme

$$P = \lambda P_1^{\alpha_1} P_2^{\alpha_2} \dots P_k^{\alpha_k}$$

avec $\lambda \in K^\star$, $\alpha_i \in \mathbb{N}^\star$ et les $P_i \in K[X]$ unitaires, irréductibles dans $K[X]$, et deux à deux distincts. Les P_i sont appelés les facteurs irréductibles *de P.*

Démonstration. L'existence de la décomposition se montre par récurrence sur le degré de P. Si P est irréductible, il n'y a rien à faire. Sinon, il existe $Q_1, Q_2 \in K[X]$ chacun de degré supérieur ou égal à 1, tels que $P = Q_1 Q_2$. En particulier $\deg Q_i < \deg P$. On recommence ensuite avec Q_1 et Q_2. Ce procédé s'arrête au bout d'un nombre fini d'étapes, puisque à chaque fois le degré des deux nouveaux polynômes qui apparaissent est strictement inférieur à celui du précédent. A la fin, il ne reste plus qu'à factoriser par des constantes les polynômes apparaissant dans la décomposition de manière à les rendre unitaires.

Montrons l'unicité de la décomposition : soient $P = \lambda P_1^{\alpha_1} P_2^{\alpha_2} \dots P_k^{\alpha_k} = \mu Q_1^{\beta_1} Q_2^{\beta_2} \dots Q_\ell^{\beta_\ell}$ deux décompositions de P comme dans le théorème. Pour commencer, il est clair que $\lambda = \mu$: c'est le coefficient du terme de plus haut degré de P. Le polynôme Q_1 est irréductible et divise P. D'après la proposition 114, il divise un des P_i. Quitte à renuméroter, Q_1 divise donc P_1. Comme P_1 et Q_1 sont tous deux irréductibles et unitaires, on a $P_1 = Q_1$. De plus, on a nécessairement $\alpha_1 = \beta_1$ car si ce n'était pas le cas, par le même raisonnement, $P_1 = Q_1$ devrait soit diviser un des P_i, $i \geq 2$, soit un des Q_i, $i \geq 2$. Mais ceci est impossible car $\operatorname{pgcd}(P_1, P_i) = 1 = \operatorname{pgcd}(Q_1, Q_i)$ pour tout $i \geq 2$.

On a donc finalement $P_2^{\alpha_2} \dots P_k^{\alpha_k} = Q_2^{\beta_2} \dots Q_\ell^{\beta_\ell}$ et on recommence le même raisonnement. Au bout de ℓ étapes, on obtient bien que $k = \ell$ et que, quitte à changer l'ordre des Q_i, $Q_i = P_i$, ainsi que $\beta_i = \alpha_i$, pour tout i.

□

Corollaire 117. *Soient $A, B \in K[X]$ non nuls et supposons qu'ils n'ont aucun facteur irréductible en commun. Alors* $\operatorname{pgcd}(A, B) = 1$.

Démonstration. Soit D un facteur irréductible d'un diviseur de A de degré au moins 1. D'après la proposition 114, D divise un facteur irréductible de A, et donc lui est égal. Mais alors, par la même proposition, D ne peut pas diviser B car A et B n'ont pas de facteur irréductible en commun. Par conséquent, $\operatorname{pgcd}(A, B) = 1$. □

Corollaire 118. *Soient $A, B \in K[X]$, non nuls. Si $A = \lambda \prod_{i=1}^{k} P_i^{\alpha_i}$ et $B = \mu \prod_{i=1}^{k} P_i^{\beta_i}$ sont les décompositions en facteurs irréductibles de A et B (ici, on autorise α_i et β_i à être nuls) alors, en posant $\gamma_i = \min\{\alpha_i, \beta_i\}$ et $\delta_i = \max\{\alpha_i, \beta_i\}$, on a*

$$\operatorname{pgcd}(A, B) = \prod_{i=1}^{k} P_i^{\gamma_i} \quad \text{et} \quad \operatorname{ppcm}(A, B) = \prod_{i=1}^{k} P_i^{\delta_i}.$$

Démonstration. Il est clair que le polynôme ci-dessus (appelons-le D) divise A et B et que le deuxième polynôme (appelons-le M) est un multiple de A et B. Le quotient de A (resp. B) par D est $\lambda \prod_{i=1}^{k} P_i^{\alpha_i - \gamma_i}$ (resp. $\mu \prod_{i=1}^{k} P_i^{\beta_i - \gamma_i}$). Or, pour tout i, on a soit $\alpha_i - \gamma_i = 0$, soit $\beta_i - \gamma_i = 0$ par définition des γ_i. Le pgcd de ces deux quotients est donc 1 d'après le corollaire précédent, puisqu'ils n'ont pas de facteur irréductible en commun. D'après la remarque 105, ceci entraîne que $D = \operatorname{pgcd}(A, B)$. Comme $DM = AB$, on a bien $M = \operatorname{ppcm}(A, B)$ d'après la remarque 111. □

III. Racines et factorisation d'un polynôme

III.1. Racines d'un polynôme. Notons $\mathcal{F}(K, K)$ le K-espace vectoriel des fonctions de K dans K. C'est aussi un anneau. À tout polynôme $P = \sum_{k=0}^{n} a_k X^k \in K[X]$, on associe une fonction $\tilde{P} : K \to K$ donnée, pour tout $x \in K$, par $\tilde{P}(x) = \sum_{k=0}^{n} a_k x^k$.

On vérifie immédiatement que l'application $\Phi : K[X] \to \mathcal{F}(K, K)$ qui à $P \in K[X]$ associe $\Phi(P) = \tilde{P}$ est linéaire. C'est également un morphisme d'anneaux. Nous verrons bientôt que Φ est injective si K est de cardinal infini, par exemple si $K = \mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$.

Définition 119. *Pour tout $P \in K[X]$, la fonction $\tilde{P}$ est appelée la* fonction polynôme associée *à P. En général, on la notera également P, ce qui est pleinement justifié lorsque Φ est injective, par exemple pour $K = \mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$. Une fonction $f \in \mathcal{F}(K, K)$ est dite* polynomiale *s'il existe $P \in K[X]$ tel que $\Phi(P) = f$. L'injectivité de Φ revient à dire que deux fonctions polynomiales sont égales si et seulement si elles sont associées au même polynôme.*

Remarque 120. *Etant donnée la façon dont nous avons défini la dérivation dans $K[X]$, si $P \in K[X]$, la dérivée de la fonction associée à P est la fonction associée à la dérivée de P (i.e. $(\Phi(P))' = \Phi(P')$).*

Si $P, Q \in K[X]$, on a aussi $\Phi(P \circ Q) = \Phi(P) \circ \Phi(Q)$. Ceci permet par exemple de déduire, toujours par injectivité de Φ, la formule $(P \circ Q)' = (P' \circ Q)Q'$ dans $\mathbb{R}[X]$ de celle connue pour la dérivée des fonctions composées.

Définition 121. *Soit $P \in K[X]$. On dit que $a \in K$ est une* racine *de P si $P(a) = 0$ (on devrait écrire $\tilde{P}(a) = 0$).*

Proposition 122. *Soient $P \in K[X]$ et $a \in K$. Alors, a est une racine de P si et seulement si $(X - a)$ divise P.*

Démonstration. Supposons que $P(a) = 0$ et effectuons la division euclidienne de P par $(X - a)$. Il existe $Q, R \in K[X]$ tels que $P = (X - a)Q + R$ et $\deg R < \deg(X - a) = 1$, donc R est un polynôme constant. Si on évalue l'égalité précédente en a, on obtient $0 = P(a) = R$, donc $R = 0$, i.e. $(X - a)$ divise P.

Réciproquement, si $(X - a)$ divise P, il existe $Q \in K[X]$ tel que $P = (X - a)Q$, et en évaluant en a, on trouve $P(a) = 0$. □

Définition 123. *Soient $P \in K[X]$ un polynôme non nul et soit $a \in K$ une racine de P. On dit que a est une racine d'*ordre *(ou* multiplicité*) m ($m \in \mathbb{N}^\star$) de P si $(X - a)^m$ divise P et $(X - a)^{m+1}$ ne divise pas P, ce qui est équivalent à : il existe $Q \in K[X]$ tel que $Q(a) \neq 0$ et $P = (X - a)^m Q$.*

L'équivalence des deux propriétés est immédiate d'après la proposition précédente.

Remarque 124.

(1) *L'ordre d'une racine est bien défini : par définition, il vaut au moins 1, et il est majoré par* $\deg P$ *(sauf si* P *est le polynôme nul : dans ce cas, tous les éléments de* K *sont racines du polynôme nul et on peut considérer qu'elles sont d'ordre infini). Si une racine est d'ordre 1, on dit qu'elle est* simple*, si elle est d'ordre 2, on dit qu'elle est* double*.*

(2) *Un polynôme de degré* n *a au plus* n *racines (comptées avec multiplicités). Ainsi, si un polynôme de degré au plus* n *a au moins* $n+1$ *racines, c'est le polynôme nul. En particulier, si la fonction associée à un polynôme est identiquement nulle, ce polynôme est nul lorsque* K *a une infinité d'éléments, ce qui montre l'injectivité de* Φ *dans cette situation.*

(3) *Pour tout nombre premier* p*,* $K = \mathbb{Z}/p\mathbb{Z}$ *est un corps. Soit* $P = X^p - X \in K[X]$*. Comme pour tout* $x \in K$ *on a* $x^p = x$*, la fonction polynomiale associée à* P *est identiquement nulle, et* Φ *n'est donc pas injective dans ce cas.*

Nous supposons maintenant que $K = \mathbb{R}$ ou $K = \mathbb{C}$ (ou éventuellement $K = \mathbb{Q}$).

Proposition 125. *Soient* $P \in K[X]$*,* $a \in K$ *et* $m \geq 1$*. Alors* P *admet* a *comme racine d'ordre* m *si et seulement si* $P(a) = P'(a) = \cdots = P^{(m-1)}(a) = 0$ *et* $P^{(m)}(a) \neq 0$*.*

Démonstration. Commençons par montrer le

Lemme 126. *Soient* $P \in K[X]$ *et* $a \in K$*. Supposons qu'il existe* $k \geq 1$ *et* $Q \in K[X]$ *tels que* $P = (X-a)^k Q$ *et* $Q(a) \neq 0$*. Alors, il existe* $Q_1 \in K[X]$ *tel que* $P' = (X-a)^{k-1} Q_1$ *et* $Q_1(a) \neq 0$*.*

Démonstration. On a $P' = k(X-a)^{k-1}Q + (X-a)^k Q' = (X-a)^{k-1}(kQ + (X-a)Q') = (X-a)^{k-1}Q_1$ avec $Q_1 = kQ + (X-a)Q'$ et donc $Q_1(a) = kQ(a) \neq 0$. □

Revenons à la preuve de la proposition. Si P admet a comme racine d'ordre m, il existe $Q \in K[X]$ tel que $P = (X-a)^m Q$ et $Q(a) \neq 0$. D'après le lemme, on a immédiatement par récurrence que pour tout j compris entre 1 et $m-1$, il existe $Q_j \in K[X]$ tel que $P^{(j)} = (X-a)^{m-j} Q_j$. Par conséquent, $P^{(j)}(a) = 0$ pour tout $0 \leq j \leq m-1$. En outre, les Q_j ne s'annulent pas en a et donc en particulier, $P^{(m-1)} = (X-a)Q_{m-1}$ avec $Q_{m-1}(a) \neq 0$. Mais alors, $P^{(m)} = Q_{m-1} + (X-a)Q'_{m-1}$ et par conséquent, $P^{(m)}(a) = Q_{m-1}(a) \neq 0$.

Réciproquement, cette preuve montre que si l'ordre de la racine a est plus petit (resp. plus grand), il existe $1 \leq k < m-1$ tel que $P^{(k)}(a) \neq 0$ (resp. $P^{(m)}(a) = 0$). □

III.2. Factorisation d'un polynôme dans $\mathbb{C}$ et dans $\mathbb{R}$. Nous ne montrerons pas le théorème suivant (qui est un résultat d'analyse non trivial).

Théorème 127.

(Théorème de d'Alembert). *Tout polynôme de* $\mathbb{C}[X]$ *non constant possède (au moins) une racine.*

Théorème 128.

Soit $P \in \mathbb{C}[X]$ *avec* $\deg P \geq 1$*. Alors,* P *se décompose de manière unique (à l'ordre des facteurs près) sous la forme*

$$P = \lambda(X - z_1)^{\alpha_1}(X - z_2)^{\alpha_2} \dots (X - z_k)^{\alpha_k}.$$

avec $\lambda \in \mathbb{C}^\star$*,* $z_1, \dots, z_k \in \mathbb{C}$ *deux à deux distincts, et* $\alpha_1, \dots, \alpha_k \in \mathbb{N}^\star$ *(on dit que* P *est scindé).*

Démonstration. Il suffit d'appliquer le théorème 116 et de remarquer que d'après le théorème de d'Alembert, les seuls polynômes irréductibles dans $\mathbb{C}[X]$ sont ceux de degré 1. □

Théorème 129.

Soit $P \in \mathbb{R}[X]$ avec $\deg P \geq 1$. Alors, P se décompose de manière unique (à l'ordre des facteurs près) sous la forme

$$P = \lambda P_1^{\alpha_1} P_2^{\alpha_2} \dots P_k^{\alpha_k}$$

avec $\lambda \in \mathbb{R}^\star$, $\alpha_i \in \mathbb{N}^\star$ et les $P_i \in \mathbb{R}[X]$ deux à deux distincts, égaux soit à un polynôme de la forme $X - a$ ($a \in \mathbb{R}$), ou de la forme $X^2 - sX + p$ ($s, p \in \mathbb{R}$, $s^2 - 4p < 0$). On remarquera que quelle que soit leur forme, les P_i sont tous irréductibles dans $\mathbb{R}[X]$.

Démonstration. Commençons par montrer un lemme.

Lemme 130. *Soit $P \in \mathbb{R}[X]$ vu comme un élément de $\mathbb{C}[X]$. Si $\deg P \geq 1$ et si P admet $a \in \mathbb{C}\backslash\mathbb{R}$ pour racine d'ordre m, alors il admet également $\bar{a}$ comme racine d'ordre m.*

Démonstration. Par hypothèse, il existe $Q \in \mathbb{C}[X]$ tel que $P = (X-a)^m Q$ et $Q(a) \neq 0$. Mais alors, $P = \bar{P} = (X - \bar{a})^m \bar{Q}$ et $\bar{Q}(\bar{a}) = \overline{Q(a)} \neq 0$. □

Revenons à la preuve du théorème. On commence par écrire la décomposition de P dans $\mathbb{C}[X]$. Ensuite, on regroupe les termes faisant intervenir des racines complexes conjuguées (non réelles) : elles ont même multiplicité d'après le lemme et donnent donc des termes du type $(X-a)^\alpha (X-\bar{a})^\alpha = (X^2 - (2\mathrm{Re}\, a)X + |a|^2)^\alpha = (X^2 - sX + p)^\alpha$ avec $s^2 - 4p < 0$ puisque le polynôme $X^2 - sX + p$ n'admet pas de racine réelle (donc son discriminant est négatif). L'unicité de l'écriture vient du théorème 116 puisque tous les facteurs sont irréductibles dans $\mathbb{R}[X]$.

□

Exemple 131. *Soit $P = X^4 + 1$. Ses racines complexes sont $e^{-i\frac{3\pi}{4}}, e^{-i\frac{\pi}{4}}, e^{i\frac{\pi}{4}}$ et $e^{i\frac{3\pi}{4}}$ (aucune n'est réelle) donc la décomposition de P en facteurs irréductibles dans $\mathbb{C}[X]$ est*

$$X^4 + 1 = (X - e^{-i\frac{3\pi}{4}})(X - e^{-i\frac{\pi}{4}})(X - e^{i\frac{\pi}{4}})(X - e^{i\frac{3\pi}{4}}).$$

Par ailleurs,

$$P = [(X - e^{i\frac{\pi}{4}})(X - e^{-i\frac{\pi}{4}})][(X - e^{i\frac{3\pi}{4}})(X - e^{-i\frac{3\pi}{4}})] = \Big(X^2 - 2\cos\big(\frac{\pi}{4}\big)X + 1\Big)\Big(X^2 - 2\cos\big(\frac{3\pi}{4}\big)X + 1\Big)$$

et par conséquent, la décomposition de P en facteurs irréductibles dans $\mathbb{R}[X]$ est

$$P = (X^2 - \sqrt{2}X + 1)(X^2 + \sqrt{2}X + 1).$$

CHAPITRE 3

Rappels sur les espaces vectoriels

I. Définitions

Nous commençons par rappeler la notion d'espace vectoriel.

Définition 132. *Soit E un ensemble non vide, muni de deux opérations : une opération (ou loi) interne appelée* addition *et notée $+$:*

$$E \times E \to E \qquad (u, v) \mapsto u + v$$

et une opération (ou loi) externe appelée multiplication *et notée $\cdot$ de $K \times E$ vers E qui à (λ, u) associe $\lambda \cdot u$.*

On dit que $(E, +, \cdot)$ est un espace vectoriel sur K *ou un K-*espace vectoriel *si les propriétés suivantes sont vérifiées :*

- *L'addition sur E est :*
  - *(1) commutative : $x + y = y + x$ pour tout $(x, y) \in E^2$*
  - *(2) associative : $(x + y) + z = x + (y + z)$ pour tout $(x, y, z) \in E^3$*
  - *(3) possède un zéro : il existe $\mathbf{0} \in E$ tel que $x + \mathbf{0} = \mathbf{0} + x = x$, pour tout $x \in E$*
  - *(4) tout élément possède un opposé : pour tout $x \in E$, il existe un élément noté $-x$ tel que $x + (-x) = \mathbf{0}$.*
- *La multiplication externe vérifie :*
  - *(5) $1 \cdot x = x$ pour tout $x \in E$*
  - *(6) $\lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x$ pour tout $(\lambda, \mu) \in K^2$, $x \in E$.*
- *La multiplication est* distributive sur l'addition *:*
  - *(7) $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$ pour tout $\lambda \in K$, pour tout $(x, y) \in E^2$*
  - *(8) $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$ pour tout $(\lambda, \mu) \in K^2$, pour tout $x \in E$.*

On appelle les éléments d'un espace vectoriel des *vecteurs* et les éléments de K des *scalaires*. L'élément $\mathbf{0}$ est appelé le *vecteur nul*.

Les opérations d'espaces vectoriels vérifient quelques propriétés supplémentaires qui se déduisent de celles listées dans la définition :

(1) $\lambda \cdot \mathbf{0} = \mathbf{0}$ pour tout $\lambda \in K$,

(2) $0 \cdot x = \mathbf{0}$ pour tout $x \in E$,

(3) Si $\lambda \cdot x = \mathbf{0}$ alors $\lambda = 0$ ou $x = \mathbf{0}$,

(4) $-x = (-1) \cdot x$ pour tout $x \in E$,

(5) Si $\lambda_1, \ldots, \lambda_k$ appartiennent à K, et si $y_1, y_2, \ldots, y_k$ appartiennent à E, alors

$$x = \lambda_1 y_1 + \cdots + \lambda_k y_k$$

appartient à E. On dit que x est une *combinaison linéaire* des vecteurs $y_1, \ldots, y_k$.

Par exemple K^n est un K-espace vectoriel, ainsi que l'espace vectoriel nul $\{\mathbf{0}\}$ où $\mathbf{0}$ est un "symbole". Une manière de multiplier les exemples est d'appeler la définition suivante.

Définition 133. *Soit $(E, +, \cdot)$ un espace vectoriel sur K. Un sous-espace vectoriel de E est un sous-ensemble F de E vérifiant :*

(1) $F \neq \emptyset$

(2) Pour tout $(x, y) \in F^2$, $x + y \in F$

(3) Pour tout $\lambda \in K$, $x \in F$, $\lambda x \in F$.

En effet il est facile de vérifier que si F est un sous-espace vectoriel de $(E, +, \cdot)$ alors $(F, +, \cdot)$ est lui-même un espace vectoriel sur K. Il est d'ailleurs souvent préférable, pour montrer qu'un certain ensemble est un espace vectoriel, de montrer que c'est un sous-espace vectoriel d'un espace vectoriel déjà connu que de vérifier tous les axiomes.

II. Somme et somme directe de sous-espaces vectoriels

Soit $(E, +, \cdot)$ un espace vectoriel et soient F et G deux sous-espaces vectoriels de E. On rappelle la notion de somme $F + G$: c'est le sous-espace vectoriel de E défini par :

$$F + G = \{y + z \mid y \in F, z \in G\}.$$

L'intersection $F \cap G$ est aussi un sous-espace vectoriel de E et on a, si E est de dimension finie, la formule dite de Grassmann :

$$\dim(F + G) = \dim(F) + \dim(G) - \dim(F \cap G).$$

Définition 134. *On dit que* E *est* la somme directe de F et G, *et on note* $E = F \oplus G$, *si les conditions suivantes sont réalisées :*

(1) $E = F + G$

(2) $F \cap G = \{\mathbf{0}\}$

Théorème 135.

Les conditions suivantes sont équivalentes :

(1) $E = F \oplus G$

(2) Tout vecteur $x \in E$ *s'écrit* de façon unique $x = y + z$ *avec* $y \in F$ *et* $z \in G$.

Si, en outre, E *est de dimension finie, alors on a équivalence de :*

(3) $E = F \oplus G$

(4) $E = F + G$ *et* $\dim(E) = \dim(F) + \dim(G)$

(5) $F \cap G = \{0\}$ *et* $\dim(E) = \dim(F) + \dim(G)$.

Démonstration. Montrons d'abord l'équivalence de (1) et (2). Supposons $E = F \oplus G$, et soit $x \in E$. Comme $E = F + G$, on sait qu'il existe $y \in F$, $z \in G$ tels que $x = y + z$. Si on a deux décompositions de x, on a $x = y + z = y' + Z'$ mais alors $y - y' = z' - z \in F \cap G$. L'hypothèse $F \cap G = \{\mathbf{0}\}$ montre que $y - y' = z' - z = \mathbf{0}$ donc $y = y'$ et $z = z'$.

Réciproquement, si (2) est vrai, alors on a bien sûr $E = F + G$. Il reste à montrer que $F \cap G = \{\mathbf{0}\}$. Si $x \neq \mathbf{0}$, $x \in F \cap G$, on a deux décompositions de $\mathbf{0}$ en la somme d'un vecteur de F et d'un vecteur de G : $\mathbf{0} = \mathbf{0} + \mathbf{0} = x + (-x)$ ce qui contredirait (2).

Les équivalences de (3), (4), (5), se montrent avec la formule $\dim(F+G) = \dim(F)+\dim(G)-\dim(F \cap G)$. □

Exercice : Montrez que, si $E = F \oplus G$, la réunion d'une base de F et d'une base de G est une base de E.

CHAPITRE 4

Déterminant

I. Définition et premières propriétés du déterminant d'une matrice

Définition 136. *Soit $M = (m_{ij})_{1\leq i,j\leq n}$ une matrice carrée d'ordre n à coefficients dans un corps K. On appelle* déterminant de la matrice M *et on note* $\det(M)$ *le scalaire*

$$\det(M) = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} m_{j\sigma(j)}.$$

Remarque 137. *On note en général le déterminant de la matrice M :*

$$\det(M) = \begin{vmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & \ldots & \vdots \\ m_{n1} & \cdots & m_{nn} \end{vmatrix}.$$

Exemple 138. *On a $S_2 = \{\mathrm{Id},\ (1\ 2)\}$ et donc :*

$$\begin{vmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{vmatrix} = m_{11}m_{22} - m_{21}m_{12}.$$

Exemple 139. *On a $S_3 = \{\mathrm{Id},\ (1\ 2\ 3),\ (1\ 3\ 2),\ (1\ 2),\ (1\ 3),\ (2\ 3)\}$ et donc*

$$\begin{vmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{vmatrix} = m_{11}m_{22}m_{33} + m_{12}m_{23}m_{31} + m_{13}m_{21}m_{32} - m_{12}m_{21}m_{33} - m_{13}m_{22}m_{31} - m_{11}m_{23}m_{32}.$$

Proposition 140.
a) *Si $M = (m)$ est une matrice d'ordre 1 alors* $\det(M) = m$.
b) *Si on note I_n la matrice identité d'ordre n alors* $\det(\mathrm{I}_n) = 1$.
c) *Pour tout $\lambda \in K$ et toute matrice carrée M d'ordre n, on a* $\det(\lambda M) = \lambda^n \det(M)$.

Démonstration. Le premier point est évident.
Pour le second, on remarque que si on note a_{ij} les coefficients de la matrice identité, alors $a_{j\sigma(j)} = 0$ si $\sigma(j) \neq j$, et donc $\det(\mathrm{I}_n) = \prod_{j=1}^{n} a_{jj} = 1$.
Enfin pour le troisième point si $M = (m_{ij})$ alors $\lambda M = (\lambda m_{ij})$ et donc

$$\det(\lambda M) = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} \lambda m_{j\sigma(j)} = \lambda^n \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} m_{j\sigma(j)} = \lambda^n \det(M).$$

□

Proposition 141. *On a aussi*

$$\det(M) = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} m_{\sigma(j)j}.$$

Démonstration. La preuve repose essentiellement sur deux changements de variable : un premier dans le produit et un second dans la somme. Soit ρ et σ dans S_n, alors comme ρ est une bijection de $\{1,\dots,n\}$, en posant $j=\rho(k)$ on a

$$\prod_{j=1}^{n} m_{j\sigma(j)} = \prod_{k=1}^{n} m_{\rho(k)\sigma(\rho(k))}.$$

En particulier avec $\rho=\sigma^{-1}$, on a $\sigma(\rho(k))=k$ et donc

$$\prod_{j=1}^{n} m_{j\sigma(j)} = \prod_{k=1}^{n} m_{\sigma^{-1}(k)k}.$$

En revenant à la définition du déterminant on obtient, par sommation, que

$$\det(M) = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{k=1}^{n} m_{\sigma^{-1}(k)k}.$$

On se rappelle ensuite que $\varepsilon(\sigma)=\varepsilon(\sigma^{-1})$ de sorte que

$$\det(M) = \sum_{\sigma\in S_n} \varepsilon(\sigma^{-1}) \prod_{k=1}^{n} m_{\sigma^{-1}(k)k}.$$

Comme $\sigma\mapsto\sigma^{-1}$ est une bijection de S_n, en posant $\rho=\sigma^{-1}$ on obtient

$$\det(M) = \sum_{\rho\in S_n} \varepsilon(\rho) \prod_{k=1}^{n} m_{\rho(k)k},$$

et on renomme la variable de sommation σ, au lieu de ρ, pour conclure. □

Corollaire 142. *Une matrice carrée et sa transposée ont des déterminants égaux.*

Démonstration. Soit $M=(m_{ij})$ une matrice carrée. On note ${}^tM=(m^*_{ij})=(m_{ji})$ sa transposée. On a :

$$\det({}^tM) = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} m^*_{j\sigma(j)} = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} m_{\sigma(j)j} = \det(M),$$

d'après la proposition précédente. □

Remarque 143. *Cela signifie que toutes les propriétés du déterminant qui seront établies par rapport aux colonnes d'une matrice seront aussi valables pour les lignes.*

II. Determinant des matrices triangulaires par blocs

Nous allons voir que pour certaines matrices le calcul du déterminant se fait assez facilement. Considérons tout d'abord des matrices dont l'écriture par blocs est

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

où M' est une matrice carrée d'ordre p, M'' est une matrice carrée d'ordre q, N est une matrice à p lignes et q colonnes et le dernier bloc (noté 0 dans la matrice) est la matrice nulle à q lignes et p colonnes, avec $p+q=n$ l'ordre de M.

Proposition 144. *Le déterminant de la matrice* $M = \begin{pmatrix} M' & N \\ 0 & M'' \end{pmatrix}$ *est* $\det(M')\det(M'')$.

Démonstration. Notons $M = (m_{ij})_{1\leq i,j\leq n}$. On a

$$\det(M) = \sum_{\sigma\in S_n} \varepsilon(\sigma) m_{\sigma(1)1}\dots m_{\sigma(n)n}.$$

Pour qu'un terme soit non nul il faut que

$$\sigma(1),\dots,\sigma(p) \in \{1,\dots,p\},$$

et donc

$$\sigma(p+1),\dots,\sigma(n) \in \{p+1,\dots,n\}.$$

Il existe donc des permutations $\sigma' \in S_p$ et $\sigma'' \in S_q$ telles que

$$\begin{aligned}\sigma(j) &= \sigma'(j) \quad \text{si} \quad 1\leq j\leq p,\\ \sigma(p+k) &= p+\sigma''(k) \quad \text{si} \quad 1\leq k\leq q.\end{aligned}$$

On a donc

$$\det(M) = \sum_{\sigma'\in S_p,\sigma''\in S_q} \varepsilon(\sigma) m'_{\sigma'(1)1}\dots m'_{\sigma'(p)p} m''_{\sigma''(1)1}\dots m''_{\sigma''(q)q}.$$

On peut voir une telle permutation σ comme la composée des deux permutations σ' et $p+\sigma''(\cdot - p)$ (à supports disjoints) et donc $\varepsilon(\sigma) = \varepsilon(\sigma')\varepsilon(\sigma'')$. Par conséquent

$$\det(M) = \left(\sum_{\sigma'\in S_p} \varepsilon(\sigma') m'_{\sigma'(1)1}\dots m'_{\sigma'(p)p}\right)\left(\sum_{\sigma''\in S_q} \varepsilon(\sigma'') m''_{\sigma''(1)1}\dots m''_{\sigma''(q)q}\right) = \det(M')\det(M'')$$

ce qui montre la proposition. □

Une simple généralisation de ce résultat donne :

Corollaire 145. *Soit M une matrice de la forme*

$$\begin{pmatrix} M_1 & * & * & * \\ 0 & M_2 & * & * \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & M_r \end{pmatrix}$$

où, pour i compris entre 1 et r, M_i est une matrice carrée d'ordre k_i, avec $k_1+\dots+k_r = n$. Alors on a

$$\det(M) = \det(M_1)\dots\det(M_r)$$

On en déduit alors le déterminant d'une matrice triangulaire supérieure. Rappelons qu'une matrice carrée d'ordre n est dite *triangulaire supérieure (inférieure)* si tous ses coefficients en dessous (respectivement au-dessus) de sa diagonale sont nuls.

Corollaire 146. *Soit M une matrice triangulaire supérieure, c'est-à-dire*

$$M = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ 0 & m_{22} & \cdots & m_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & m_{nn} \end{pmatrix},$$

alors $\det(M) = m_{11}\dots m_{nn}$

Remarque 147. *Le résultat est identique pour une matrice diagonale (matrice triangulaire supérieure dont les coefficients au-dessus de la diagonale sont tous nuls). D'autre part il reste encore vrai pour les matrices triangulaires inférieures (transposées de matrices triangulaires supérieures).*

III. Déterminant d'une famille de vecteurs, multiplicativité et critère d'inversibilité

On définit maintenant le déterminant d'une famille de vecteurs.

Définition 148. *Soient $(x_1,\dots,x_n)$ des vecteurs d'un K-espace vectoriel E de dimension n. On appelle* déterminant *de $x_1,\dots,x_n$ dans une base $\mathcal{B}$ et on note $\det_{\mathcal{B}}(x_1,\dots,x_n)$ (on omet l'indice $\mathcal{B}$ lorsqu'il n'y a pas ambiguïté) le déterminant de la matrice dont les colonnes sont constituées des coordonnées des vecteurs $x_1,\dots,x_n$ dans la base $\mathcal{B}$.*

Proposition 149. *Soient $(x_1,\dots,x_n)$ des vecteurs d'un K-espace vectoriel E de dimension n.*
a) *Soit $\rho\in S_n$, alors*

$$\det(x_{\rho(1)},\dots,x_{\rho(n)})=\varepsilon(\rho)\det(x_1,\dots,x_n).$$

En particulier si on échange deux colonnes d'une matrice, le déterminant est changé en son opposé.
b) *Si $x_i=x_j$ pour $i\neq j$ alors on a $\det(x_1,\dots,x_n)=0$.*
c) *Soient y_k un vecteur de E et λ et μ deux scalaires, alors :*

$$\det(x_1,\dots,\lambda x_k+\mu y_k,\dots,x_n)=\lambda\det(x_1\dots,x_k,\dots,x_n)+\mu\det(x_1\dots,y_k,\dots,x_n).$$

De façon équivalente, le déterminant d'une matrice dépend linéairement de chacun de ses vecteurs colonnes. En particulier il est nul si une colonne est nulle.
d) *Le déterminant d'une famille de vecteurs ne change pas lorsqu'on ajoute à l'un des vecteurs une combinaison linéaire des autres vecteurs ou, de façon équivalente, le déterminant d'une matrice ne change pas lorsqu'on ajoute à l'un de ses vecteurs-colonnes une combinaison linéaire des autres vecteurs-colonnes ; il est nul si l'un des vecteurs-colonnes est combinaison des autres.*

On dit que le déterminant est une forme multilinéaire *(propriété* **c***)* alternée *(propriété* **a***).*

Démonstration. On note $M=(m_{ij})$ la matrice dont les colonnes sont les coordonnées des vecteurs $x_1,\dots,x_n$.
a) On a

$$\begin{aligned}
\det(x_{\rho(1)},\dots,x_{\rho(n)}) &= \sum_{\sigma\in S_n}\varepsilon(\sigma)\prod_{j=1}^{n}m_{j\rho(\sigma(j))}\\
&= \varepsilon(\rho)\sum_{\sigma\in S_n}\varepsilon(\rho)\varepsilon(\sigma)\prod_{j=1}^{n}m_{j\rho\sigma(j)} \ \text{ car } \ \varepsilon(\rho)^2=1\\
&= \varepsilon(\rho)\sum_{\rho\sigma\in S_n}\varepsilon(\rho\sigma)\prod_{j=1}^{n}m_{j\rho\sigma(j)} \ \text{ car } \ \varepsilon(\rho\sigma)=\varepsilon(\sigma)\varepsilon(\rho)\\
&= \varepsilon(\rho)\sum_{\tau\in S_n}\varepsilon(\tau)\prod_{j=1}^{n}m_{j\tau(j)} \ \text{ en posant } \tau=\rho\sigma\\
&= \varepsilon(\rho)\det(x_1,\dots,x_n)
\end{aligned}$$

En particulier si on note τ la transposition $(i\ j)$ alors $\varepsilon(\tau)=-1$ et par conséquent on a

$$\det(x_1,\dots,x_j,\dots,x_i,\dots,x_n)=\det(x_1,\dots,x_{\tau(i)},\dots,x_{\tau(j)},\dots,x_n)=-\det(x_1,\dots,x_i,\dots,x_j,\dots,x_n).$$

b) On a d'après la propriété précédente

$$\begin{aligned}
\det(x_1,\dots,x_i,\dots,x_j,\dots,x_n) &= -\det(x_1,\dots,x_j,\dots,x_i,\dots,x_n)\\
&= -\det(x_1,\dots,x_i,\dots,x_j,\dots,x_n) \ \text{ car } x_i=x_j
\end{aligned}$$

et donc ce déterminant est nul.
c) Soit

$$x_k=\begin{pmatrix}a_{1k}\\ \vdots\\ a_{nk}\end{pmatrix} \text{ et } y_k=\begin{pmatrix}b_{1k}\\ \vdots\\ b_{nk}\end{pmatrix},$$

alors en utilisant la deuxième formule du déterminant

$$\begin{aligned}\det(x_1,\dots,\lambda x_k+\mu y_k,\dots,x_n) &= \sum_{\sigma\in S_n}\varepsilon(\sigma)(\lambda a_{\sigma(k)k}+\mu b_{\sigma(k)k})\prod_{j=1,j\neq k}^{n} m_{\sigma(j)j}\\ &= \lambda\sum_{\sigma\in S_n}\varepsilon(\sigma)a_{\sigma(k)k}\prod_{j=1,j\neq k}^{n} m_{\sigma(j)j}+\mu\sum_{\sigma\in S_n}\varepsilon(\sigma)b_{\sigma(k)k}\prod_{j=1,j\neq k}^{n} m_{\sigma(j)j}\\ &= \lambda\det(x_1\dots,x_k,\dots,x_n)+\mu\det(x_1\dots,y_k,\dots,x_n)\end{aligned}$$

d) Par la linéarité du déterminant par rapport à chaque variable (propriété précédente) on a

$$\det(x_1,\dots,x_k+\sum_{l=1,l\neq k}^{n}\lambda_l x_l,\dots,x_n)=\det(x_1,\dots,x_k,\dots,x_n)+\sum_{l=1,l\neq k}^{n}\lambda_l\det(x_1,\dots,x_l,\dots,x_n).$$

Or tous les déterminants de la somme sont nuls d'après la propriété **b)**, car ils ont deux colonnes identiques. Ceci montre le résultat voulu. En particulier, si un vecteur est égal à une combinaison linéaire des autres vecteurs, on ne change pas le déterminant si on ajoute au vecteur l'opposé de la combinaison linéaire. Donc on est ramené à calculer le déterminant d'une matrice dont une colonne est nulle et par conséquent ce déterminant est nul. □

Le prochain résultat établit le caractère multiplicatif du déterminant.

Théorème 150.

Soient M et N deux matrices carrées d'ordre n, alors

$$\det(MN)=\det(M)\det(N).$$

Démonstration. Notons par $M_1,\dots,M_n$ les colonnes de M de sorte que celles du produit MN sont $\sum_{i=1}^{n} n_{i,1}M_i,\dots,\sum_{i=1}^{n} n_{i,n}M_i$. En utilisant la propriété **c)** de la proposition 149 on obtient

$$\begin{aligned}\det(MN) &= \det(\sum_{i_1=1}^{n} n_{i_1,1}M_{i_1},\dots,\sum_{i_n=1}^{n} n_{i_n,n}M_{i_n})\\ &= \sum_{1\leqslant i_1,\dots,i_n\leqslant n} n_{i_1,1}\dots n_{i_n,n}\det(M_{i_1},\dots,M_{i_n}),\end{aligned}$$

et grâce à la propriété **b)** de Proposition 149 on peut restreindre la somme précédente aux indices $i_1,\dots,i_n$ distincts deux à deux. Donc

$$\begin{aligned}\det(MN) &= \sum_{\sigma\in S_n} n_{\sigma(1),1}\dots n_{\sigma(n),n}\det(M_{\sigma(1)},\dots,M_{\sigma(n)})\\ &= \sum_{\sigma\in S_n} n_{\sigma(1),1}\dots n_{\sigma(n),n}\,\varepsilon(\sigma)\det(M_1,\dots,M_n)\\ &= \Big(\sum_{\sigma\in S_n}\varepsilon(\sigma)\,n_{\sigma(1),1}\dots n_{\sigma(n),n}\Big)\det M\\ &= \det(N)\det(M),\end{aligned}$$

grâce à la propriété **a)** de la proposition 149 et à la proposition 141. □

Remarque 151. *Le déterminant n'est pas additif i.e. on a en général*

$$\det(M+N)\neq\det(M)+\det(N).$$

On peut prendre par exemple le cas des matrices carrées :

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } N = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Corollaire 152. *Soient M et N deux matrices carrées d'ordre n, alors*

$$\det(MN) = \det(NM).$$

Démonstration. D'après le théorème précédent on a :

$$\det(MN) = \det(M)\det(N) = \det(N)\det(M) = \det(NM).$$

□

Théorème 153.

Soit M une matrice carrée d'ordre n, alors

$$M \text{ est inversible} \iff \det(M) \neq 0,$$

et dans ce cas

$$\det(M^{-1}) = (\det(M))^{-1}.$$

Démonstration. On a

$$M \text{ est inversible} \iff MM^{-1} = \mathrm{I}_n \Rightarrow \det(M)\det(M^{-1}) = \det(MM^{-1}) = \det(\mathrm{I}_n) = 1,$$

et donc on en déduit que $\det(M) \neq 0$.

Pour la réciproque on va démontrer la contraposée. Si la matrice M n'est pas inversible, en particulier, comme c'est une matrice carrée, cela implique que son image n'est pas $\mathbb{R}^n$ tout entier, or cette image est l'espace vectoriel engendré par les vecteurs $Me_1, \ldots, Me_n$, où $e_1, \ldots, e_n$ sont les vecteurs de la base canonique. Pour tout i, Me_i est la i ème colonne de M. Ainsi les vecteurs-colonnes de M sont liés, donc l'un des vecteurs est combinaison linéaire des autres et donc $\det(M) = 0$. On en déduit $\det(M) \neq 0 \Rightarrow$ M est inversible. □

On déduit immédiatement de ce qui précède le

Corollaire 154. *Pour tout $n \geq 1$, l'application*

$$\det : \mathrm{GL}_n(K) \to K^*$$

est un morphisme de groupe.

IV. Déterminant de matrices semblables et determinant d'un endomorphisme

Rappelons que deux matrices carrées d'ordre n, M et N, sont *semblables* s'il existe une matrice inversible P telle que $N = P^{-1}MP$. Pour que deux matrices soient semblables, il faut et il suffit qu'elles soient les matrices du même endomorphisme d'un espace vectoriel E dans deux bases.

Proposition 155. *Des matrices semblables ont même déterminant.*

Démonstration. On a

$$\det(N) = \det(P^{-1}MP) = \det(P^{-1})\det(M)\det(P) = \frac{1}{\det(P)}\det(M)\det(P) = \det(M).$$

□

Nous pouvons ainsi définir le déterminant d'un endomorphisme.

Théorème 156.

Soit E un K-espace vectoriel de dimension n et f un endomorphisme de E. Alors le scalaire $\det_{\mathcal{B}}(f(e_1),\dots,f(e_n))$ ne dépend pas de la base $\mathcal{B}=(e_1,\dots,e_n)$ choisie. On l'appelle déterminant de l'endomorphisme f *et on le note* $\det(f)$.

Démonstration. Soient $\mathcal{B}=(e_1,\dots,e_n)$ et $\mathcal{B}'=(e'_1,\dots,e'_n)$ deux bases de E. On note M et N les matrices de l'endomorphisme f de E dans ces deux bases et P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Alors on a $N=P^{-1}MP$ et d'après la proposition précédente $\det(M)=\det(N)$, d'où le résultat. □

Théorème 157.

Soient f et g deux endomorphismes d'un K-espace vectoriel E. Alors les propriétés suivantes sont vérifiées :
a) $\det(f\circ g)=\det(f)\det(g)$;
b) f est bijectif si et seulement si $\det(f)\neq 0$ *et alors on a* $\det(f^{-1})=\dfrac{1}{\det(f)}$.

Démonstration. Ces résultats s'obtiennent facilement lorsqu'on passe à l'écriture matricielle. □

V. Développement par rapport à une ligne ou une colonne

Dans la pratique, il est extrêmement rare que l'on calcule un déterminant par la formule directe, qui est trop compliquée. On a recours à des astuces. Nous allons donner une formule qui permet effectivement de calculer le déterminant d'une matrice carrée d'ordre n.

Définition 158. *Soit $M=(m_{ij})_{1\leq i,j\leq n}$ une matrice carrée d'ordre n. On note M_{ij} la matrice d'ordre $n-1$, obtenue à partir de M en supprimant la i-ème ligne et la j-ème colonne (tous les autres coefficients restent dans le même ordre). On appelle* mineur *de m_{ij} le scalaire* $\det(M_{ij})$ *et* cofacteur *de m_{ij} le scalaire* $\Delta_{ij}=(-1)^{i+j}\det(M_{ij})$.

Théorème 159.

Avec les notations de la définition précédente on a :
a) développement par rapport à la colonne j, $j=1,\dots,n$,

$$\det(M)=(-1)^{1+j}\,m_{1j}\det(M_{1j})+\cdots+(-1)^{n+j}\,m_{nj}\det(M_{nj})=m_{1j}\,\Delta_{1j}+\cdots+m_{nj}\,\Delta_{nj};$$

b) développement par rapport à la ligne i, $i=1,\dots,n$,

$$\det(M)=(-1)^{i+1}\,m_{i1}\det(M_{i1})+\cdots+(-1)^{i+n}\,m_{in}\det(M_{in})=m_{i1}\,\Delta_{i1}+\cdots+m_{in}\,\Delta_{in}.$$

Démonstration. Démontrons la première formule (la seconde se démontre de la même façon ou s'obtient en considérant la transposée de M). Pour une colonne de M, on peut écrire

$$\begin{pmatrix} m_{1j}\\ m_{2j}\\ \vdots\\ m_{nj}\end{pmatrix}=\begin{pmatrix} m_{1j}\\ 0\\ \vdots\\ 0\end{pmatrix}+\begin{pmatrix} 0\\ m_{2j}\\ \vdots\\ 0\end{pmatrix}+\cdots+\begin{pmatrix} 0\\ \vdots\\ 0\\ m_{nj}\end{pmatrix}.$$

Comme le déterminant est multi-linéaire, il est linéaire en la j-ème colonne. Si on note N_{ij} la matrice obtenue à partir de M en remplaçant tous les éléments de la j-ème colonne, sauf celui dans la i-ème ligne, par 0, on obtient

$$\det(M) = \det(N_{1j}) + \cdots + \det(N_{nj}).$$

Si l'on fait passer la j-ème colonne de N_{ij} à la première place, sans changer l'ordre des autres, on effectue $j-1$ transpositions de colonnes, donc le déterminant est multiplié par $(-1)^{j-1}$; de même, si l'on fait passer la i-ème ligne de N_{ij} à la première place (sans changer l'ordre des autres) le déterminant est multiplié par $(-1)^{i-1}$. On a donc

$$\det(N_{ij}) = (-1)^{(i-1)+(j-1)} \begin{vmatrix} m_{ij} & m_{i1} & \cdots & m_{i\,j-1} & m_{i\,j+1} & \cdots & m_{in} \\ 0 & & & & & & \\ \vdots & & & M_{ij} & & & \\ 0 & & & & & & \end{vmatrix} = (-1)^{i+j}\, m_{ij} \det(M_{ij})$$

d'après la Proposition 144. On obtient donc le a). □

Exemple 160. *Il est avantageux de développer selon une ligne ou une colonne où il y a beaucoup de zéros :*

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 0 & 5 & 6 \\ 0 & 0 & 3 & 0 \\ 1 & 6 & 4 & 2 \end{vmatrix} = (-1)^{3+1}.0.\begin{vmatrix} 2 & 3 & 4 \\ 0 & 5 & 6 \\ 6 & 4 & 2 \end{vmatrix} + (-1)^{3+2}.0.\begin{vmatrix} 1 & 3 & 4 \\ 4 & 5 & 6 \\ 1 & 4 & 2 \end{vmatrix} + (-1)^{3+3}.3.\begin{vmatrix} 1 & 2 & 4 \\ 4 & 0 & 6 \\ 1 & 6 & 2 \end{vmatrix} + (-1)^{3+4}.0.\begin{vmatrix} 1 & 2 & 3 \\ 4 & 0 & 5 \\ 1 & 6 & 4 \end{vmatrix}$$

$$= 3.\begin{vmatrix} 1 & 2 & 4 \\ 4 & 0 & 6 \\ 1 & 6 & 2 \end{vmatrix} = 3.\begin{vmatrix} 1 & 2 & 4 \\ 4 & 0 & 6 \\ -2 & 0 & -10 \end{vmatrix} \quad L_3 \leftarrow L_3 - 3L_1$$

$$= 3.(-1)^{1+2}.2.\begin{vmatrix} 4 & 6 \\ -2 & -10 \end{vmatrix} = -6(4.(-10) - 6.(-2)) = 168.$$

VI. Calcul de l'inverse d'une matrice

Soit $M = (m_{ij})_{1\le i,j\le n}$ une matrice carrée. On peut lui associer n^2 cofacteurs Δ_{ij} pour $1 \le i, j \le n$. Il est donc possible de construire une matrice carrée d'ordre n dont les coefficients sont les cofacteurs de M, c'est la matrice des cofacteurs dont la définition est donnée ci-dessous.

Définition 161. *Soit $M = (m_{ij})_{1\le i,j\le n}$ une matrice carrée. On appelle* comatrice *de M la matrice notée* $\mathrm{Com}(M)$ *telle que le coefficient de la i-ème ligne et de la j-ème colonne soit exactement le cofacteur Δ_{ij} de M.*

Nous donnons maintenant une relation entre une matrice et sa comatrice, ainsi qu'une expression de l'inverse d'une matrice.

Théorème 162.

Soit $M = (m_{ij})_{1\le i,j\le n}$ une matrice carrée. On a

$$M\,{}^t\mathrm{Com}(M) = \det(M) I_n \text{ et } {}^t\mathrm{Com}(M)\, M = \det(M) I_n.$$

De plus si M est inversible ($\det(M) \neq 0$) *alors on a*

$$M^{-1} = \frac{1}{\det(M)}\, {}^t\mathrm{Com}(M).$$

Démonstration. Le terme d'indice (i, j) du produit $M\,{}^t\mathrm{Com}(M)$ est $\sum_{k=1}^{n} m_{ik}\Delta_{jk}$ (on multiplie par la transposée). Distinguons deux cas :

— Si $j = i$ on reconnaît la deuxième formule du théorème 159 et donc on obtient $\det(M)$.

— Sinon c'est le développement suivant la j-ème ligne, du déterminant de la matrice obtenue à partir de M en ayant remplacé la j-ème ligne par la i-ème. Cette matrice a deux lignes égales, donc son déterminant est nul. On obtient donc $M\,{}^t\mathrm{Com}(M) = \det(M) I_n$.

L'autre égalité s'obtient de la même manière en utilisant le développement suivant les colonnes.

Pour le dernier résultat, lorsque M est inversible il suffit de diviser tous les termes dans la première formule par $\det(M)$. □

Exemple 163. *Soit* $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *une matrice inversible c'est-à-dire telle que* $\det(M) = ad - bc \neq 0$. *Alors l'inverse de* M *est donné par*

$$M^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

VII. Système de Cramer

On considère ici des systèmes linéaires qui ont un nombre d'équations égal au nombre d'inconnues.

Définition 164. *Soit* n *un entier naturel non nul. On appelle* système d'équations linéaires *de* n *équations à* n *inconnues sur le corps* K*, le système :*

$$(S) \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

où les a_{ij} *et* b_i *sont des éléments de* K *pour* $i = 1, \ldots, n$ *et* $j = 1, \ldots, n$.

On peut écrire matriciellement ce système de la manière suivante. Soit

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ldots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

la matrice du système. On pose encore

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ et } B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Alors le système est équivalent à l'équation matricielle $AX = B$.

Théorème 165.

On dit qu'un système linéaire de n *équations à* n *inconnues est de* Cramer*, si sa matrice est inversible. Alors le système a une unique solution qui est*

$$X = A^{-1}B.$$

Démonstration. Le résultat est immédiat. □

Proposition 166. *Soit* (S) *un système de Cramer, d'équation matricielle* $AX = B$*. Pour* i*,* $1 \leq i \leq n$*, on définit la matrice* A_i *obtenue en remplaçant dans* A *la* i*-ème colonne par le second membre* B*. Alors la solution du système est donnée par : pour tout* i *compris entre* 1 *et* n

$$x_i = \frac{\det(A_i)}{\det(A)}.$$

DÉMONSTRATION. Comme le système est de Cramer, il admet une unique solution. Si on note $C_1, \ldots, C_n$ les vecteurs colonnes de la matrice A du système, on peut écrire

$$x_1C_1 + x_2C_2 + \cdots + x_nC_n = B.$$

Maintenant, si A_i est la matrice obtenue en remplaçant dans A la i-ème colonne par le second membre B, on a

$$\begin{aligned}
\det(A_i) &= \det(C_1, \ldots, C_{i-1}, B, C_{i+1}, \ldots, C_n) \\
&= \det(C_1, \ldots, C_{i-1}, x_1C_1 + x_2C_2 + \cdots + x_nC_n, C_{i+1}, \ldots, C_n) \\
&= \sum_{k=1}^{n} x_k \det(C_1, \ldots, C_{i-1}, C_k, C_{i+1}, \ldots, C_n) \\
&= x_i \det(C_1, \ldots, C_{i-1}, C_i, C_{i+1}, \ldots, C_n) = x_i \det(A)
\end{aligned}$$

car l'application déterminant est une forme multi-linéaire alternée. Comme $\det(A) \neq 0$, ceci montre le résultat. □

Exemple 167. *On souhaite résoudre le système :*

$$(S) \begin{cases} 2x_1 + 3x_2 + x_3 = 9 \\ x_1 + 2x_2 + 3x_3 = 6 \\ 3x_1 + x_2 + 2x_3 = 8 \end{cases}.$$

La matrice du système est $A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$. *Son déterminant est égal à* 18*, c'est donc un système de Cramer. La solution est donc*

$$\begin{aligned}
x_1 &= \frac{1}{18} \begin{vmatrix} 9 & 3 & 1 \\ 6 & 2 & 3 \\ 8 & 1 & 2 \end{vmatrix} = \frac{35}{18}, \\
x_2 &= \frac{1}{18} \begin{vmatrix} 2 & 9 & 1 \\ 1 & 6 & 3 \\ 3 & 8 & 2 \end{vmatrix} = \frac{29}{18}, \\
x_3 &= \frac{1}{18} \begin{vmatrix} 2 & 3 & 9 \\ 1 & 2 & 6 \\ 3 & 1 & 8 \end{vmatrix} = \frac{5}{18}.
\end{aligned}$$

CHAPITRE 5

Réduction des endomorphismes

Pour décrire un endomorphisme f d'un espace vectoriel E, on cherche une base de E dans laquelle la matrice de f soit la plus simple possible. Pour diverses raisons, on voudrait que cette matrice soit *diagonale*, c'est-à-dire que les coefficients en dehors de la diagonale soient nuls. En d'autre termes, les vecteurs de cette base sont tels que leur image par f leur est colinéaire. On les appelle des *vecteurs propres* de f. L'avantage d'avoir une matrice diagonale est qu'il est alors plus facile de faire des calculs faisant intervenir f, par exemple le calcul des itérés de f.

Il n'existe pas toujours de base de E formée de vecteurs propres de f. Par exemple la rotation R_α dans le plan réel orienté, de matrice

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

dans la base canonique, n'a pas de vecteur propre si $\alpha \not\equiv 0 \mod \pi$, puisqu'aucun vecteur non nul x n'est colinéaire à $R_\alpha(x)$. Le problème dans ce cas se résout en passant du corps $\mathbb{R}$ au corps $\mathbb{C}$: nous verrons qu'un endomorphisme d'un espace vectoriel complexe a toujours un vecteur propre. Néanmoins, même sur $\mathbb{C}$, il peut ne pas y avoir "assez" de vecteurs propres pour former une base ; c'est par exemple le cas pour l'endomorphisme de $\mathbb{C}^2$ de matrice

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

dont tous les vecteurs propres sont colinéaires à e_1 et donc ne peuvent pas former une base. On cherchera pour ces endomorphismes une base dans laquelle leur matrice est aussi simple que possible, c'est-à-dire aussi proche que possible d'une forme diagonale.

A partir de maintenant, les espaces vectoriels considérés sont des espaces vectoriels sur K de dimension finie, où K est le corps des nombres réels $\mathbb{R}$ ou le corps des nombres complexes $\mathbb{C}$. En fait les résultats sont encore vrais si K est un corps commutatif contenant $\mathbb{Q}$.

I. Diagonalisation

I.1. Valeur propre et vecteur propre.

Définition 1. *Soit $f \in \mathcal{L}(E)$. On dit que $\lambda \in K$ est* valeur propre *de f s'il existe un vecteur non nul x de E tel que $f(x) = \lambda x$; x est alors appelé* vecteur propre *de f associé à la valeur propre λ.*

Remarque 2. *Tous les multiples non nuls d'un vecteur propre de f sont encore des vecteurs propres de f pour la même valeur propre. L'ensemble des valeurs propres d'un endomorphisme f s'appelle le spectre* de f et est noté $\mathrm{Sp}(f)$.

Exemple 3. *Si f est une homothétie d'un espace vectoriel E, $f = \lambda \mathrm{Id}_E$, alors tout vecteur non nul est un vecteur propre associé à la valeur propre λ.*

Exemple 4. *Comme on l'a vu dans l'introduction il existe des endomorphismes qui n'ont pas de valeur propre ni de vecteur propre ; par exemple les rotations d'angle différent de $0 \mod \pi$ dans le plan réel.*

Dans le théorème suivant nous caractérisons de plusieurs façons les valeurs propres d'un endomorphisme.

Théorème 5.

Soient $f \in \mathcal{L}(E)$ et λ un scalaire. Les assertions suivantes sont équivalentes :
(i) *λ est valeur propre de f ;*
(ii) *l'endomorphisme* $f - \lambda \mathrm{Id}_E$ n'est pas injectif, *i.e. son* noyau *vérifie*

$$\mathrm{Ker}(f - \lambda \mathrm{Id}_E) = \{x \in E,\ (f - \lambda \mathrm{Id}_E)(x) = 0\} \neq \{0\};$$

(iii) $\det(f - \lambda \mathrm{Id}_E) = 0$ *;*
(iv) $\det(M - \lambda \mathrm{I}_n) = 0$ *où M est la matrice de f dans n'importe quelle base de E.*

Démonstration. Pour que λ soit valeur propre de f il faut et il suffit qu'il existe un vecteur non nul x de E tel que $f(x) = \lambda x$, c'est-à-dire que l'on ait $(f - \lambda \mathrm{Id}_E)(x) = 0$, ou encore que le noyau $\mathrm{Ker}(f - \lambda \mathrm{Id}_E) \neq \{0\}$. Ceci entraîne l'équivalence de (i) et (ii). Pour que l'endomorphisme $f - \lambda \mathrm{Id}_E$ de l'espace vectoriel *de dimension finie* E ne soit pas injectif il faut et il suffit qu'il ne soit pas bijectif, c'est-à-dire que son déterminant soit nul, d'où l'équivalence de (ii) et (iii). Enfin par définition du déterminant d'un endomorphisme, (iii) et (iv) sont équivalentes. □

Ce qui précède montre que l'ensemble des vecteurs propres associés à une valeur propre λ auquel on ajoute le vecteur nul est exactement $\mathrm{Ker}(f - \lambda \mathrm{Id}_E)$.

Définition 6. *Soit λ une valeur propre d'un endomorphisme f. On appelle* sous-espace propre *associé à λ, le sous-espace vectoriel de E défini par $E_\lambda = \mathrm{Ker}(f - \lambda \mathrm{Id}_E)$.*

Remarque 7. *C'est en cherchant le noyau de l'application $f - \lambda \mathrm{Id}_E$ que l'on détermine les vecteurs propres associés à la valeur propre λ.*

I.2. Polynôme caractéristique.

Définition 8. *Le* polynôme caractéristique *de $f \in \mathcal{L}(E)$ est défini par $\chi_f(X) = \det(f - X\mathrm{Id}_E)$.*

Si E est un K-espace vectoriel alors $\chi_f(X) \in K[X]$. Par définition, si M est la matrice de f dans une base quelconque $\mathcal{B}$ de E, alors

$$\chi_f(X) = \det(M - X\mathrm{I}_n).$$

Le lecteur attentif remarquera que dans le chapitre précédent, on a uniquement défini le déterminant d'une matrice à coefficients dans un corps, or les coefficients de $M - X\mathrm{I}_n$ sont à valeurs dans $K[X]$ qui est un anneau. Cependant, comme il est commutatif cela ne pose pas de problème au niveau de la définition et de plus, on considérera principalement la fonction polynomiale associée.

Par extension, on définit le polynôme caractéristique d'une matrice carrée.

Théorème 9.

Les valeurs propres d'un endomorphisme f sur un K-espace vectoriel E sont exactement les racines de son polynôme caractéristique qui sont dans K.

Démonstration. On a les équivalences suivantes :

$$\begin{aligned} \lambda \in K \text{ est valeur propre de } f &\iff (f - \lambda \mathrm{Id}_E) \text{ est non injectif} \iff \det(f - \lambda \mathrm{Id}_E) = 0 \\ &\iff \chi_f(\lambda) = 0 \iff \lambda \text{ est une racine de } \chi_f \text{ dans } K \end{aligned}$$

□

Soit M une matrice d'ordre n à coefficients dans K. Soit X une indéterminée, alors on peut écrire :

$$M - X\mathrm{I}_n = \begin{pmatrix} m_{11} - X & m_{12} & \cdots & m_{1n} \\ m_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & m_{(n-1)n} \\ m_{n1} & \cdots & m_{n(n-1)} & m_{nn} - X \end{pmatrix}$$

Le déterminant de cette matrice est une somme de produits de coefficients de la matrice, c'est donc bien un polynôme en X à coefficients dans K. Chacun des produits comporte n termes, qui sont des coefficients de chacune des colonnes ; ce sont donc des polynômes de degré au plus n. De plus un seul produit est de degré n, c'est le produit de tous les coefficients diagonaux (il correspond à la permutation $\sigma = \mathrm{Id}$). Comme la signature de l'identité est égale à $+1$, le terme de plus haut degré est $(-1)^n X^n$. On a donc montré le résultat suivant.

Proposition 10. *Le polynôme caractéristique d'une matrice d'ordre n (ou d'un endomorphisme d'un espace vectoriel de dimension n) est un polynome de degré n dont le coefficient dominant est $(-1)^n$.*

Comme un polynôme de degré n a au plus n racines on obtient :

Corollaire 11. *En dimension n un endomorphisme (ou une matrice d'ordre n) a au plus n valeurs propres distinctes.*

Exemple 12. *Le polynôme caractéristique de la matrice*

$$M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

est

$$\chi_M(X) = \begin{vmatrix} -X & 1 \\ 1 & -X \end{vmatrix} = X^2 - 1.$$

Les valeurs propres sont ± 1. Les sous-espaces propres associés à ces valeurs propres se déterminent alors de la manière suivante :

$$\begin{pmatrix} x \\ y \end{pmatrix} \in E_1 \iff \begin{cases} -x + y = 0 \\ x - y = 0 \end{cases} \iff \{x = y \iff E_1 = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle.$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \in E_{-1} \iff \begin{cases} x + y = 0 \\ x + y = 0 \end{cases} \iff \{y = -x \iff E_{-1} = \langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rangle.$$

Exemple 13. *La rotation plane R_θ a pour matrice dans la base canonique*

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Son polynôme caractéristique est

$$\chi_{R_\theta}(X) = \begin{vmatrix} \cos\theta - X & -\sin\theta \\ \sin\theta & \cos\theta - X \end{vmatrix} = (\cos\theta - X)^2 + (\sin\theta)^2 = X^2 - 2X\cos\theta + 1.$$

Ce polynôme n'a pas de racines réelles si $\sin\theta$ est non nul. Par contre dans $\mathbb{C}$ il a deux racines $e^{\pm i\theta}$, qui sont donc les valeurs propres de R_θ. On peut vérifier que les vecteurs propres associés sont colinéaires à

$$\begin{pmatrix} 1 \\ \mp i \end{pmatrix}.$$

Exemple 14. *Soit $M = (m_{ij})_{1 \le i,j \le n}$ une matrice triangulaire. Alors $M - X\mathrm{I}_n$ est aussi une matrice triangulaire et le polynôme caractéristique (déterminant d'une matrice triangulaire) est donc le produit des coefficients diagonaux i.e. $\chi_M(X) = (m_{11} - X) \cdots (m_{nn} - X)$. Les valeurs propres de M sont donc les coefficients diagonaux de M. En particulier ce résultat est vrai pour une matrice diagonale.*

Définition 15. *Soit f un endomorphisme d'un K-espace vectoriel E.*

(1) L'ordre de multiplicité d'une valeur propre λ *de* f *est l'ordre de multiplicité de la racine* λ *dans le polynôme caractéristique de* f*.*

(2) Un polynôme de $K[X]$ *est dit* scindé *sur* K *s'il se décompose dans* $K[X]$ *comme produit de polynômes de degré 1.*

(3) Par extension on dira qu'un endomorphisme est scindé si son polynôme caractéristique l'est.

Si un endomorphisme f d'un espace vectoriel E de dimension n, est scindé, son polynôme caractéristique a donc la forme suivante :

$$\chi_f(X) = (-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i},$$

où r, $1 \leq r \leq n$, est le nombre de valeurs propres distinctes, les $(\lambda_i)_{1\leq i\leq r}$ sont les différentes valeurs propres et les $(\alpha_i)_{1\leq i\leq r}$ sont leurs ordres de multiplicité respectifs. On a de plus $\sum_{i=1}^{r} \alpha_i = n$.

I.3. Étude des sous-espaces propres.

Théorème 16.

Soit $f \in \mathcal{L}(E)$*. On suppose que* f *possède au moins une valeur propre et on considère* $\{\lambda_1, \ldots, \lambda_r\}$ *l'ensemble des valeurs propres deux à deux distinctes de* f*. Alors les sous-espaces propres* $(E_{\lambda_i})_{1\leq i\leq r}$ *sont en* somme directe.

Rappelons que r sous-espaces vectoriels $(E_{\lambda_i})_{1\leq i\leq r}$ sont dits en somme directe si et seulement si l'une deux propriétés équivalentes suivantes est vérifiée :

(**a**) $x_1 + x_2 + \cdots + x_r = 0$, avec pour $i = 1, \ldots, r$, $x_i \in E_{\lambda_i} \Rightarrow \forall\ 1 \leq i \leq r,\ x_i = 0$;

(**b**) pour tout $2 \leq k \leq r$, $E_{\lambda_k} \cap (E_{\lambda_1} + \cdots + E_{\lambda_{k-1}}) = \{0\}$.

Démonstration. On va établir ce résultat par récurrence i.e. on va montrer que si pour un certain $1 \leq l \leq r-1$, les E_{λ_i} pour $1 \leq i \leq l$ sont en somme directe, alors il en est de même des E_{λ_i} pour $1 \leq i \leq l+1$.

Soit donc

$$x \in E_{\lambda_{l+1}} \cap (E_{\lambda_1} + \cdots + E_{\lambda_l}).$$

On a $x = x_1 + \cdots + x_l$ avec pour $1 \leq i \leq l$, $x_i \in E_{\lambda_i}$. En prenant l'image par f de cet élément, on obtient :

$$f(x) = \lambda_{l+1} x = \lambda_{l+1}(x_1 + \cdots + x_{k-1}),$$

et d'autre part

$$f(x) = f(x_1) + \cdots + f(x_l) = \lambda_1 x_1 + \cdots + \lambda_l x_l.$$

Ceci entraîne que

$$(\lambda_{l+1} - \lambda_1)x_1 + \cdots + (\lambda_{l+1} - \lambda_l)x_l = 0.$$

Or par hypothèse de récurrence $E_{\lambda_1}, \ldots, E_{\lambda_l}$ sont en somme directe, et comme les valeurs propres sont deux à deux distinctes, on obtient $x_i = 0$ pour $1 \leq i \leq l$, c'est-à-dire $E_{\lambda_{l+1}} \cap (E_{\lambda_1} + \cdots + E_{\lambda_l}) = \{0\}$.

□

Théorème 17.

Soient $f \in \mathcal{L}(E)$ *et* λ *une valeur propre de* f*. Alors la dimension du sous-espace propre associé à la valeur propre* λ *est inférieure ou égale à la multiplicité de la valeur propre* λ*, c'est-à-dire :* $\dim(E_\lambda) \leq \alpha_\lambda$.
En particulier, si λ *est une valeur propre simple (multiplicité égale à* 1*) alors* $\dim(E_\lambda) = 1$

Démonstration. On considère $(e_1, \dots, e_k, e_{k+1}, \dots, e_n)$ une base de E telle que $(e_1, \dots, e_k)$ soit une base de E_λ. Dans cette base la matrice de f s'écrit

$$M = \left(\begin{array}{ccc|c} \lambda & & & \\ & \ddots & & M_1 \\ & & \lambda & \\ \hline 0 & \cdots & 0 & \\ \vdots & \ddots & \vdots & M_2 \\ 0 & \cdots & 0 & \end{array}\right)$$

Le polynôme caractéristique de f est donc

$$\chi_f(X) = \det(M - X\mathrm{I}_n) = (\lambda - X)^k \det(M_2 - X\mathrm{I}_{n-k}).$$

Ainsi λ est racine de χ_f d'ordre α_λ supérieur ou égal à $k = \dim(E_\lambda)$. Pour le second point un sous-espace propre contient au-moins un vecteur propre (vecteur non nul) donc on a $1 \leq \dim(E_\lambda)$. Si on applique alors le premier point avec $\alpha_\lambda = 1$, on obtient le résultat. □

I.4. Endomorphismes diagonalisables.

Définition 18. *On dit que $f \in \mathcal{L}(E)$ est* diagonalisable *s'il existe une base de E constituée de vecteurs propres.*

Remarque 19. *Dans une telle base la matrice de f s'écrit*

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

Théorème 20.

Soit $f \in \mathcal{L}(E)$, dont $\{\lambda_1, \dots, \lambda_r\}$ est l'ensemble des valeurs propres deux à deux distinctes. L'endomorphisme f est diagonalisable si et seulement si ses sous-espaces propres vérifient : $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$.

Démonstration. Supposons que f est diagonalisable. Quitte à réordonner les vecteurs, la base de vecteurs propres de f est de la forme

$$(e_{1,\lambda_1}, \dots, e_{\beta_1,\lambda_1}, \dots, e_{1,\lambda_r}, \dots, e_{\beta_r,\lambda_r})$$

où pour $1 \leq i \leq r$ et pour $1 \leq j \leq \beta_i$, e_{j,λ_i} est un vecteur propre associé à la valeur propre λ_i. Clairement, pour tout i, $\beta_i \leq \dim(E_{\lambda_i})$ puisque pour tout i, les e_{j,λ_i}, $1 \leq j \leq \beta_i$, forment une famille libre d'éléments de E_{λ_i}. Donc

$$n = \beta_1 + \cdots + \beta_r \leq \dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_r}) \leq n$$

la dernière inégalité venant du fait que les E_{λ_i} sont en somme directe. Mais alors cette inégalité est en fait une égalité d'où $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$ (et également $\beta_i = \dim(E_{\lambda_i})$ pour tout i).

Réciproquement, si les sous-espaces propres de f vérifient

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

Alors en notant une base de E_{λ_i}, $\mathcal{B}_i = (e_{1,\lambda_i}, \dots, e_{\beta_i,\lambda_i})$, on a $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ est une base de E formée de vecteurs propres. □

Théorème 21.

Soit $f \in \mathcal{L}(E)$, avec E un K-espace vectoriel. f est diagonalisable si et seulement si les deux conditions suivantes sont vérifiées :

(a) *f est scindé ;*

(b) *Pour toute valeur propre λ de f on a $\dim(E_\lambda) = \alpha_\lambda$ qui est l'ordre de multiplicité de λ.*

DÉMONSTRATION. Si f est diagonalisable alors d'après le théorème 20 il existe une base de vecteurs propres, $(e_{1,\lambda_1}, \ldots, e_{\beta_1,\lambda_1}, \ldots, e_{1,\lambda_r}, \ldots, e_{\beta_r,\lambda_r})$, où $\beta_i = \dim(E_{\lambda_i})$, dans laquelle sa matrice s'écrit

$$M = \begin{pmatrix} \lambda_1 & 0 & \ldots & \ldots & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & \lambda_1 & \ddots & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \lambda_r & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & \ldots & \ldots & 0 & \lambda_r \end{pmatrix}$$

Son polynôme caractéristique est égal à $(-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{\beta_i}$. Il est donc scindé, et pour $1 \leq i \leq r$, $\dim(E_{\lambda_i}) = \beta_i =$ l'ordre de multiplicité de λ_i.
Réciproquement, pour tout sous-espace propre E_{λ_i} on a d'après la propriété (b), $\dim(E_{\lambda_i}) = \alpha_i := \alpha_{\lambda_i}$, et d'après (a),

$$\sum_{i=1}^{r} \alpha_i = n = \dim(E).$$

Donc comme les sous-espaces propres sont en somme directe, on obtient que $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$, c'est-à-dire que f est diagonalisable. □

Corollaire 22. *Soit $f \in \mathcal{L}(E)$. Une condition suffisante pour que f soit diagonalisable est que f ait exactement n valeurs propres deux à deux distinctes.*

DÉMONSTRATION. Si f possède n valeurs propres deux à deux distinctes, f est scindé, et de plus tout sous-espace propre est de dimension 1. L'endomorphisme f est donc diagonalisable d'après le théorème précédent. □

Attention ce n'est pas une condition nécessaire. En effet, les homothéties $\lambda \mathrm{Id}_E$, par exemple, sont diagonalisables mais n'ont qu'une seule valeur propre λ. Ce sont d'ailleurs les seuls endomorphismes diagonalisables qui ont une seule valeur propre.

I.5. Exemple de diagonalisation.

Soit f l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base canonique est

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

Le polynôme caractéristique de f est alors $\chi_f(X) = (1 - X)^2(2 - X)$. f a donc une valeur propre double $\lambda = 1$, et une valeur propre simple $\lambda = 2$. Recherchons maintenant les sous-espaces propres. E_1 le sous-espace propre associé à la valeur propre $\lambda = 1$ est déterminé par :

$$(M - 1\mathrm{I}_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{cases} 4x - 2y + 3z = 0 \\ 4x - 2y + 3z = 0 \end{cases} \iff \{\ 4x - 2y + 3z = 0$$

C'est donc un sous-espace vectoriel de $\mathbb{R}^3$ de dimension 2, engendré par deux vecteurs u_1 et u_2 de coordonnées dans la base canonique

$$\begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix} \text{ et } \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

Le sous-espace propre E_2 associé à la valeur propre $\lambda = 2$ est déterminé par :

$$(M - 2\mathrm{I}_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{cases} -x & = & 0 \\ 4x - 3y + 3z & = & 0 \\ 4x - 2y + 2z & = & 0 \end{cases} \iff \begin{cases} x & = & 0 \\ y & = & z \end{cases}$$

C'est donc un sous-espace vectoriel de $\mathbb{R}^3$ de dimension 1, engendré par un vecteur u_3 de coordonnées dans la base canonique

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

La somme des dimensions des sous-espaces propres est égale à 3, c'est-à-dire à celle de l'espace vectoriel E (ici $\mathbb{R}^3$), f est donc diagonalisable et sa matrice dans la base (u_1, u_2, u_3) est

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

II. Trigonalisation

II.1. Endomorphismes trigonalisables.

Définition 23. *Soit $f \in \mathcal{L}(E)$. f est dit* trigonalisable *s'il existe une base de E dans laquelle sa matrice est triangulaire supérieure ou inférieure.*

Définition 24. *Une matrice carrée M d'ordre n est dite* trigonalisable *si elle est semblable à une matrice triangulaire T, supérieure ou inférieure, c'est-à-dire s'il existe une matrice inversible P telle que $M = P^{-1}TP$.*

On peut remarquer que tout endomorphisme (ou matrice) diagonalisable est trigonalisable. D'autre part, si $f \in \mathcal{L}(E)$ est trigonalisable cela signifie qu'il existe une base B de E dans laquelle la matrice de f s'écrit

$$M = \begin{pmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}.$$

Son polynôme caractéristique est donc égal à $\chi_f(X) = (-1)^n \prod_{i=1}^{n} (X - a_{ii})$. Il est donc scindé et ses valeurs propres sont exactement les $(a_{ii})_{1 \leq i \leq n}$.

Théorème 25.

$f \in \mathcal{L}(E)$ est trigonalisable si et seulement si f est scindé.

Démonstration. Si f est trigonalisable, d'après la remarque ci-dessus, f est scindé. Réciproquement supposons que f soit scindé et montrons par récurrence sur la dimension n de E que f est trigonalisable. Si $n = 1$, la propriété est vraie. Supposons qu'elle soit vraie jusqu'au rang $n - 1$. Comme χ_f est scindé, il existe au moins une valeur propre λ. Soit alors u_1 un vecteur propre associé à cette valeur propre. On

considère des vecteurs $(u_2, \dots, u_n)$ de façon à ce que $\mathcal{B} = (u_1, u_2, \dots, u_n)$ soit une base de E. Dans cette base la matrice de f s'écrit

$$M = \begin{pmatrix} \lambda & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & M_1 & \\ 0 & & & \end{pmatrix}.$$

Soit E' le sous espace vectoriel de E engendré par $(u_2, \dots, u_n)$. L'endomorphisme $g = \pi \circ f_{|E'}$ de E', où π est la projection de E sur E' parallèlement à u_1 a pour matrice M_1 dans la base $(u_2, \dots, u_n)$. On a

$$\chi_f(X) = (\lambda - X)\det(M_1 - X\mathrm{I}_{n-1}) = (\lambda - X)\chi_g(X),$$

et comme χ_f est scindé, χ_g l'est aussi. Donc par hypothèse de récurrence, il existe une base $(u'_2, \dots, u'_n)$ de E' dans laquelle la matrice de g est triangulaire supérieure. On vérifie alors immédiatement que dans la base $\mathcal{B}' = (u_1, u'_2, \dots, u'_n)$ de E la matrice de f est triangulaire supérieure. □

Corollaire 26. *Tout endomorphisme sur un espace vectoriel complexe est trigonalisable.*

Démonstration. Cela provient du fait que tout polynôme dans $\mathbb{C}[X]$ est scindé. □

II.2. Exemple de trigonalisation.

On considère f l'endomorphisme de $E = \mathbb{R}^4$ dont la matrice dans la base canonique est

$$M = \begin{pmatrix} 1 & -3 & 0 & 3 \\ -2 & -6 & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -4 & 0 & 8 \end{pmatrix}.$$

Le calcul du polynôme caractéristique donne $\chi_f(X) = \chi_M(X) = (1 - X)^4$. Il y a donc une seule valeur propre $\lambda = 1$ d'ordre 4. On détermine maintenant le sous-espace propre E_1 associé à cette valeur propre.

$$(M - 1\mathrm{I}_4)\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{cases} -3y + 3t &= 0 \\ -2x - 7y + 13t &= 0 \\ -3y + 3t &= 0 \\ -x - 4y + 7t &= 0 \end{cases} \iff \begin{cases} x &= 3t \\ y &= t \end{cases}.$$

L'unique sous-espace propre E_1 est donc de dimension 2 ($< 4 = \dim(\mathbb{R}^4)$ donc f n'est pas diagonalisable), il est engendré par les vecteurs u_1 et u_2 de coordonnées dans la base canonique

$$\begin{pmatrix} 3 \\ 1 \\ 0 \\ 1 \end{pmatrix} \text{ et } \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

On complète (u_1, u_2) par (u_3, u_4) pour obtenir une base de $\mathbb{R}^4$. Ici on peut par exemple choisir $u_3 = e_1$ et $u_4 = e_2$. On a donc les relations suivantes :

$$e_1 = u_3,\ e_2 = u_4,\ e_3 = u_2 \text{ et } e_4 = u_1 - 3u_3 - u_4.$$

Ainsi :

$$\begin{cases} f(u_3) &= f(e_1) = e_1 - 2e_2 - e_4 = -u_1 + 4u_3 - u_4 \\ f(u_4) &= f(e_2) = -3e_1 - 6e_2 - 3e_4 - 4e_4 = -4u_1 - 3u_2 + 9u_3 - 2u_4 \end{cases}$$

La matrice de f dans la base (u_1, u_2, u_3, u_4) s'écrit alors

$$\widetilde{M} = \begin{pmatrix} 1 & 0 & -1 & -4 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 4 & 9 \\ 0 & 0 & -1 & -2 \end{pmatrix}.$$

On considère la sous-matrice

$$M_1 = \begin{pmatrix} 4 & 9 \\ -1 & -2 \end{pmatrix}$$

qui est la matrice de l'endomorphisme $\pi \circ f_{|E'}$ de E', où π est la projection de E sur $E' = \langle u_3, u_4\rangle$ parallèlement à $\langle u_1, u_2\rangle$. On va maintenant trigonaliser cette matrice. Son polynôme caractéristique est

$$\chi_{M_1}(X) = (1-X)^2.$$

Cette matrice n'a qu'une seule valeur propre double qui est 1. Le sous-espace propre associé a cette valeur propre est déterminé par

$$\left\{ \begin{array}{rcl} 3x+9y & = & 0 \\ -x-3y & = & 0 \end{array} \right. \iff \{x = -3y\ .$$

Sa dimension est donc 1, et il est engendré par v_1 de coordonnées

$$\begin{pmatrix} -3 \\ 1 \end{pmatrix}$$

dans la base (u_3, u_4). On le complète en une base du sous-espace avec le vecteur v_2 de coordonnées

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

Les vecteurs correspondants dans l'espace $E = \mathbb{R}^4$ sont donc $v_1' = -3u_3 + u_4$ et $v_2' = u_4$. On vérifie que

$$\left\{ \begin{array}{rcl} f(v_1') & = & -u_1 - 3u_2 + v_1' \\ f(v_2') & = & -4u_1 - 3u_2 - 3v_1' + v_2' \end{array} \right. .$$

La matrice de f dans la base (u_1, u_2, v_1', v_2') s'écrit alors

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

III. Polynômes d'endomorphismes - Polynôme minimal

III.1. Polynômes d'endomorphismes.

Soit E un K-espace vectoriel de dimension n et f un endomorphisme de E. On introduit les notations suivantes :

$$f^0 = \mathrm{Id}_E, \quad f^1 = f, \quad f^2 = f \circ f, \quad f^k = \overbrace{f \circ f \circ \cdots \circ f}^{k \text{ fois}}.$$

Plus généralement, si $Q(X) = a_0 + a_1 X + \cdots + a_k X^k$ est un polynôme de $K[X]$, alors on définit le *polynôme d'endomorphismes* $Q(f)$ par :

$$Q(f) = a_0 \mathrm{Id}_E + a_1 f + \cdots + a_k f^k.$$

Si M est la matrice de f dans une base $\mathcal{B}$ de E alors le *polynôme de matrices* $Q(M)$ défini par :

$$Q(M) = a_0 \mathrm{I}_n + a_1 M + \cdots + a_k M^k,$$

est la matrice de $Q(f)$ dans la base $\mathcal{B}$.

Proposition 27. *Pour tout endomorphisme f de E, il existe un polynôme non nul Q de $K[X]$ tel que $Q(f) = 0$.*

Démonstration. E est un K-espace vectoriel de dimension n, donc $\mathcal{L}(E)$ est un K-espace vectoriel de dimension n^2. Par conséquent, les n^2+1 endomorphismes $\mathrm{Id}_E,\ f,\ f^2, \ldots,\ f^{n^2}$ sont liés. Il existe donc des coefficients $a_0,\ a_1, \ldots,\ a_{n^2}$ de K non tous nuls, tels que $a_0 \mathrm{Id}_E + a_1 f + \cdots + a_{n^2} f^{n^2} = 0$. C'est-à-dire que le polynôme non nul

$$Q(X) = a_o + a_1 X + \cdots + a_{n^2} X^{n^2}$$

de $K[X]$ vérifie $Q(f) = 0$. □

III.2. Polynôme minimal.

Théorème 28.

Soient E un K-espace vectoriel et f un endomorphisme de E. L'ensemble $\mathfrak{I} = \{P \in K[X],\ P(f) = 0\}$ des polynômes annulateurs de f est un idéal différent de $\{0\}$. Cet idéal est appelé idéal annulateur *de f. Le générateur unitaire de cet idéal s'appelle le* polynôme minimal *de f et est noté $\mu_f(X)$.*

Démonstration. D'après la proposition précédente cet ensemble contient un élément non nul. Donc $\mathfrak{I}$ est différent de $\{0\}$. Soient P_1 et P_2 deux éléments de $\mathfrak{I}$. Alors $(P_1 - P_2)(f) = P_1(f) - P_2(f) = 0$, c'est-à-dire que $P_1 - P_2$ appartient à $\mathfrak{I}$. De plus si $P \in \mathfrak{I}$ et si $Q \in K[X]$ alors on a $(PQ)(f) = P(f)Q(f) = 0$. Ainsi $PQ \in \mathfrak{I}$. Ainsi $\mathfrak{I}$ est un idéal de $K[X]$. □

Remarque 29. *Le polynôme minimal d'un endomorphisme f est donc l'unique polynôme de plus bas degré et unitaire (de coefficient dominant égal à 1), vérifiant $\mu_f(f) = 0$. Il est de degré au moins 1 puisque les polynômes de degré 0 n'annulent aucun endomorphisme.*

On peut de manière similaire définir le polynôme minimal μ_M d'une matrice carrée M. Si M est la matrice de f dans n'importe quelle base $\mathcal{B}$ de E, $\mu_M(X) = \mu_f(X)$ et donc $\mu_f(M) = 0$.

Exemple 30. *Soit f une homothétie d'un K-espace vectoriel E. On a $f = \lambda \mathrm{Id}_E$, avec $\lambda \in K$. Donc le polynôme $X - \lambda$ est un polynôme annulateur de f. Comme il est de degré 1 et unitaire, c'est le générateur unitaire de l'idéal annulateur de f, c'est-à-dire que c'est le polynôme minimal de f. Réciproquement, si le polynôme minimal de f est $X - \lambda$, alors on a $f - \lambda \mathrm{Id}_E = 0$, ou encore $f = \lambda \mathrm{Id}_E$. Par conséquent, f est une homothétie. Nous venons donc de montrer que le polynôme minimal d'un endomorphisme est de degré 1 si et seulement si cet endomorphisme est une homothétie.*

III.3. Théorème de Cayley-Hamilton.

Théorème 31.

Soient E un K-espace vectoriel et f un endomorphisme de E. Le polynôme minimal de f divise le polynôme caractéristique de f.

Démonstration. D'après le théorème précédent, il suffit de montrer que $\chi_f(f) = 0$. Soit M la matrice de f dans une base $\mathcal{B}$ de E. C'est une matrice carrée d'ordre n. On pose $N = M - X\mathrm{I}_n$ et on note $\mathrm{Com}(N)$ sa comatrice. Les coefficients de $\mathrm{Com}(N)$ et donc de ${}^t\mathrm{Com}(N)$ sont des polynômes de degré inférieur ou égal à $n - 1$. On peut donc écrire

$$ {}^t\mathrm{Com}(N) = \sum_{j=0}^{n-1} A_j X^j $$

où les A_j sont des matrices carrées à coefficients dans K. D'après le théorème 162 du chapitre précédent on a

$$ N\,{}^t\mathrm{Com}(N) = \det(N)\mathrm{I}_n = \chi_f(X)\mathrm{I}_n, $$

c'est-à-dire

$$ (M - X\mathrm{I}_n)\sum_{j=0}^{n-1} A_j X^j = \chi_f(X)\mathrm{I}_n. $$

En notant maintenant

$$ \chi_f(X) = \sum_{j=0}^{n} a_j X^j $$

(avec $a_n = (-1)^n$) on obtient en identifiant les coefficients de même degré :

$$MA_0 = a_0\mathrm{I}_n,\ MA_1 - A_0 = a_1\mathrm{I}_n,\ldots,\ MA_{n-1} - A_{n-2} = a_{n-1}\mathrm{I}_n,\ -A_{n-1} = a_n\mathrm{I}_n.$$

On en déduit

$$\chi_f(M) = \sum_{j=0}^{n} a_j M^j = \sum_{j=0}^{n} M^j(a_j\mathrm{I}_n) = MA_0 + M(MA_1 - A_0) + \cdots + M^{n-1}(MA_{n-1} - A_{n-2}) + M^n(-A_{n-1}) = 0,$$

car les termes se regroupent et s'annulent. □

Remarque 32. *Un énoncé équivalent de ce résultat est que χ_f appartient à l'idéal annulateur*

$$\mathfrak{I} = \{P \in K[X],\ P(f) = 0\},$$

ou encore que $\chi_f(f) = 0$. Comme le polynôme caractéristique est de degré n, on en déduit que le polynôme minimal est au plus de degré n.

Exemple 33. *Soit la matrice*

$$M = \begin{pmatrix} 0 & 3 \\ -2 & -5 \end{pmatrix}.$$

Le polynôme caractéristique de cette matrice est

$$\begin{aligned} \chi_M(X) &= \begin{vmatrix} -X & 3 \\ -2 & -X-5 \end{vmatrix} = \begin{vmatrix} -X-2 & -X-2 \\ -2 & -X-5 \end{vmatrix} = (-X-2)\begin{vmatrix} 1 & 1 \\ -2 & -X-5 \end{vmatrix} \\ &= -(X+2)\begin{vmatrix} 1 & 0 \\ -2 & -X-3 \end{vmatrix} = -(X+2)(-X-3) = (X+2)(X+3). \end{aligned}$$

Comme le polynôme minimal, divise le polynôme caractéristique, et qu'il est unitaire de degré au moins 1, on en déduit qu'il est égal à $X+2$, $X+3$ ou $(X+2)(X+3)$. Comme $M + 2\mathrm{I}_2 \neq 0$ et que $M + 3\mathrm{I}_2 \neq 0$, on obtient que $\mu_M(X) = (X+2)(X+3)$.

Proposition 34. *λ est valeur propre de f si et seulement si λ est une racine de son polynôme minimal.*

Démonstration. Si λ est racine de μ_f alors comme μ_f divise χ_f, λ est racine de χ_f, et donc d'après le théorème 9, λ est une valeur propre de f.
Réciproquement, soit λ une valeur propre de f dont on note x un vecteur propre associé. On effectue la division euclidienne de μ_f par $X - \lambda$, et on obtient $\mu_f(X) = Q(X)(X - \lambda) + c$, où $\deg(c) \leq 0$, i.e. $c \in K$. On en déduit que

$$0 = \mu_f(f) = Q(f) \circ (f - \lambda\mathrm{Id}_E) + c\mathrm{Id}_E.$$

Si on applique cette relation au vecteur x, on trouve

$$0 = Q(f) \circ (f - \lambda\mathrm{Id}_E)(x) + c\mathrm{Id}_E(x) = Q(f)(f(x) - \lambda x) + cx = Q(f)(0) + cx = cx.$$

Or comme x n'est pas nul, ceci entraîne que $c = 0$, de sorte que le polynôme minimal s'écrit

$$\mu_f(X) = Q(X)(X - \lambda).$$

Cela signifie que λ est racine de μ_f. □

Remarque 35. *On déduit immédiatement de la proposition précédente que si χ_f est scindé à racines simples alors $\mu_f = \chi_f$.*

La proposition précédente implique que sur $\mathbb{C}$, puisque χ_f est scindé, alors tout facteur irréductible de χ_f est aussi facteur irréductible de μ_f (avec évidemment la possibilité d'une multiplicité inférieure, mais non nulle). En fait, ce résultat est également vrai sur $\mathbb{R}$ (ou $\mathbb{Q}$ mais nous ne traiterons pas ce cas) :

Corollaire 36. *Soit f un endomorphisme d'un K-espace vectoriel E avec $K = \mathbb{R}$ ou $\mathbb{C}$. Tout facteur irréductible de χ_f est également facteur irréductible de μ_f.*

Démonstration. Il suffit de vérifier le résultat sur $\mathbb{R}$. D'après le théorème 129 du chapitre 2, les facteurs irréductibles de χ_f sont soit du type $X-\lambda$ avec $\lambda\in\mathbb{R}$, soit du type X^2-sX+p avec $s^2-4p<0$. Dans le premier cas, la proposition 34 s'applique directement, dans le deuxième cas le polynôme X^2-sX+p admet deux racines complexes (non réelles) conjuguées λ et $\bar\lambda$. Ces deux racines sont donc racines du polynôme caractéristique de $\tilde f$, l'endomorphisme du *complexifié* $E_{\mathbb{C}}$ de l'espace vectoriel E qui coïncide avec f sur E (cette notion n'est pas au programme. En termes plus concrets, une base du $\mathbb{R}$-espace vectoriel E définit une base du $\mathbb{C}$-espace vectoriel $E_{\mathbb{C}}$. Dans cette base, l'endomorphisme f a une matrice à coefficients réels et cette même matrice définit l'application $\mathbb{C}$-linéaire $\tilde f$ dans la base correspondante. Ceci entraîne que $\chi_{\tilde f}=\chi_f$ et également que $\mu_{\tilde f}=\mu_f$ car la matrice de $\tilde f$ dans la base correspondante est à coefficients réels, donc invariante par conjugaison, et par conséquent $\mu_{\tilde f}$ est à coefficients réels). Par la proposition 34, λ et $\bar\lambda$ sont donc racines de $\mu_{\tilde f}=\mu_f$ ce qui entraîne que X^2-sX+p divise μ_f. □

Exemple 37. *Si on reprend l'exemple précédent, on a trouvé que le polynôme caractéristique de la matrice M est $\chi_M(X)=(X+2)(X+3)$. C'est-à-dire que les valeurs propres de M sont -2 et -3. Donc d'après ce qui précède, ce sont des racines du polynôme minimal. D'autre part comme il est unitaire et qu'il divise le polynôme caractéristique, on obtient directement que $\mu_M(X)=(X+2)(X+3)$.*

Exemple 38. *Soit la matrice*

$$M=\begin{pmatrix}1&2&3\\0&1&-5\\0&0&2\end{pmatrix}.$$

Comme la matrice est triangulaire, le polynôme caractéristique de cette matrice est $\chi_M(X)=(1-X)(1-X)(2-X)$. Les valeurs propres de M sont donc 1 et 2. Par définition du polynôme minimal et d'après le théorème de Cayley-Hamilton, on a $\mu_M(X)=(X-1)(X-2)$ ou $(X-1)^2(X-2)$, car toute valeur propre est racine de ce polynôme. Or lorsqu'on effectue le calcul suivant :

$$(M-\mathrm{I}_3)(M-2\mathrm{I}_3)=\begin{pmatrix}0&2&3\\0&0&-5\\0&0&1\end{pmatrix}\begin{pmatrix}-1&2&3\\0&-1&-5\\0&0&0\end{pmatrix}=\begin{pmatrix}0&-2&-10\\0&0&0\\0&0&0\end{pmatrix}\neq 0$$

Donc le polynôme minimal est $\mu_M(X)=(X-1)^2(X-2)$.

III.4. Lemme de décomposition des noyaux.

Les résultats que l'on démontre ici reposent sur l'énoncé suivant.

Lemme 39. *Soit f un endomorphisme d'un K-espace vectoriel E et P un polynôme de $K[X]$.*

(a) *Le sous-espace vectoriel* $\mathrm{Ker}P(f)$ *est stable par f, c'est-à-dire*

$$f\left(\mathrm{Ker}P(f)\right)\subset\mathrm{Ker}P(f).$$

(b) *Si $P=P_1P_2$ avec P_1 et P_2 deux polynômes premiers entre eux, alors*

$$\mathrm{Ker}P(f)=\mathrm{Ker}P_1(f)\oplus\mathrm{Ker}P_2(f).$$

(c) *Si de plus $P(f)=0$ alors pour $i\neq j$, la projection sur* $\mathrm{Ker}P_i(f)$ *parallèlement à* $\mathrm{Ker}P_j(f)$ *est un polynôme en f.*

Démonstration. Remarquons d'abord que f et $P(f)$ commutent. Si x appartient au noyau de $P(f)$ on a $P(f)\big(f(x)\big)=f\big(P(f)(x)\big)=0$, de sorte que $f(x)$ est encore dans le noyau de $P(f)$. Cela montre (a).

Comme P_1 et P_2 sont premiers entre eux, d'après le théorème de Bézout il existe Q_1 et Q_2 deux polynômes de $K[X]$ tels que

$$Q_1P_1+Q_2P_2=1.$$

En appliquant cette relation à f on obtient

$$Q_1(f)P_1(f)+Q_2(f)P_2(f)=\mathrm{Id}_E.$$

Soit $x\in\mathrm{Ker}P_1(f)\cap\mathrm{Ker}P_2(f)$, on a $P_1(f)(x)=P_2(f)(x)=0$, et donc d'après l'égalité ci-dessus

$$x=\mathrm{Id}_E(x)=Q_1(f)P_1(f)(x)+Q_2(f)P_2(f)(x)=0.$$

$\mathrm{Ker}P_1(f)$ et $\mathrm{Ker}P_2(f)$ sont donc en somme directe. D'autre part, pour tout x dans $\mathrm{Ker}P(f)$ on a :

$$x = \mathrm{Id}_E(x) = Q_1(f)P_1(f)(x) + Q_2(f)P_2(f)(x).$$

Or, comme

$$P_2(f)\big(Q_1(f)P_1(f)(x)\big) = Q_1(f)\big(P(f)(x)\big) = 0,$$

$Q_1(f)P_1(f)(x)$ est dans le noyau de $P_2(f)$; de même le vecteur $Q_2(f)P_2(f)(x)$ est dans le noyau de $P_1(f)$. Cela montre (b).

Supposons maintenant que $P(f) = 0$ i.e. $E = \mathrm{Ker}P(f)$ et montrons que la projection sur $\mathrm{Ker}P_2(f)$ parallèlement à $\mathrm{Ker}P_1(f)$ est un polynôme en f. Notons $R_i = P_iQ_i$ et $p_i = R_i(f)$. On a $\mathrm{Id}_E = p_1 + p_2$ et par ailleurs, $p_1 \circ p_2 = p_2 \circ p_1 = (R_1R_2)(f) = (Q_1Q_2)(f) \circ P(f) = 0$. Ainsi,

$$p_1 = p_1 \circ \mathrm{Id}_E = p_1 \circ p_1 + p_1 \circ p_2 = p_1^2$$

L'application linéaire p_1 est diagonalisable et ses valeurs propres sont 0 et/ou 1. En effet, pour tout x dans E, on peut écrire $x = (x - p_1(x)) + p_1(x)$ avec $(x - p_1(x))$ (resp. $p_1(x)$) appartenant à l'espace propre associé à la valeur propre 0 (resp. 1) et on sait que ces espaces propres sont en somme directe. C'est le projecteur sur $\mathrm{Im}\, p_1$, parallèlement à $\mathrm{Ker}\, p_1$. Il nous suffit donc de montrer que $\mathrm{Im}\, p_1 = \mathrm{Ker}P_2(f)$ et $\mathrm{Ker}\, p_1 = \mathrm{Ker}P_1(f)$.

– Soit $y = p_1(x) \in \mathrm{Im}\, p_1$. On a $P_2(f)(y) = P_2(f)R_1(f)(x) = Q_1(f)P(f)(x) = 0$ donc $\mathrm{Im}\, p_1 \subset \mathrm{Ker}P_2(f)$. Réciproquement, si $x \in \mathrm{Ker}P_2(f)$ alors $p_2(x) = Q_2(f)P_2(f)(x) = 0$. Comme $x = p_1(x) + p_2(x)$, on a finalement $x = p_1(x) \in \mathrm{Im}\, p_1$. Donc $\mathrm{Im}\, p_1 = \mathrm{Ker}P_2(f)$.

– si $x \in \mathrm{Ker}P_1(f)$, alors $p_1(x) = Q_1(f)P_1(f)(x) = 0$ donc $\mathrm{Ker}P_1(f) \subset \mathrm{Ker}\, p_1$. Réciproquement, soit $x \in \mathrm{Ker}\, p_1$. On a $x = p_1(x) + p_2(x) = p_2(x)$ donc $P_1(f)(x) = P_1(f)R_2(f)(x) = Q_2(f)P(f)(x) = 0$. Finalement, $\mathrm{Ker}\, p_1 = \mathrm{Ker}P_1(f)$.

□

Corollaire 40. *Soit f un endomorphisme d'un K-espace vectoriel E et l polynômes $P_1, \ldots, P_l$ de $K[X]$ premiers entre eux deux à deux. On pose $P = P_1P_2 \ldots P_l$. Alors*

$$\mathrm{Ker}P(f) = \mathrm{Ker}P_1(f) \oplus \cdots \oplus \mathrm{Ker}P_l(f)$$

et si de plus $P(f) = 0$ alors pour tout i, la projection sur $\mathrm{Ker}P_i(f)$ parallèlement à $\oplus_{j\neq i}\mathrm{Ker}P_j(f)$ est un polynôme en f.

Démonstration. On procède par récurrence sur l. Le cas $l = 1$ est trivial. Le cas $l = 2$ est exactement celui du Lemme 39. On suppose la propriété vraie au rang $l - 1$ et on montre qu'elle l'est encore au rang l. Comme les polynômes sont premiers entre eux deux à deux P_l est premier avec $P_1P_2 \ldots P_{l-1}$ (en effet un facteur irréductible d'un diviseur commun à $P_1P_2 \ldots P_{l-1}$ et P_l divise un des P_i, pour $1 \leq i \leq l - 1$, d'après la proposition 114 du chapitre 2. Or, P_i et P_l sont premiers entre eux). Donc par l'hypothèse de récurrence on a :

$$\begin{aligned}\mathrm{Ker}P(f) &= \mathrm{Ker}(P_1P_2 \ldots P_{l-1})(f) \oplus \mathrm{Ker}P_l(f) = \big(\mathrm{Ker}P_1(f) \oplus \cdots \oplus \mathrm{Ker}P_{l-1}(f)\big) \oplus \mathrm{Ker}P_l(f)\\ &= \mathrm{Ker}P_1(f) \oplus \cdots \oplus \mathrm{Ker}P_{l-1}(f) \oplus \mathrm{Ker}P_l(f).\end{aligned}$$

Finalement, si $P(f) = 0$, en remplaçant P_1 par P_i et P_2 par $\prod_{j\neq i} P_j$, on peut directement appliquer le lemme 39(c) pour montrer que chaque projection est un polynôme en f. □

III.5. Diagonalisation à l'aide du polynôme minimal.

Théorème 41.

Soient E un K-espace vectoriel et f un endomorphisme de E de valeurs propres distinctes $\{\lambda_1, \dots \lambda_r\}$. Les assertions suivantes sont équivalentes :
(i) *f est diagonalisable ;*
(ii) $\mu_f(X) = \prod_{i=1}^{r}(X - \lambda_i)$;
(iii) *μ_f est scindé et à racines simples ;*
(iv) *f admet un polynôme annulateur scindé à racines simples.*

DÉMONSTRATION. $(i) \Rightarrow (ii)$: d'après la proposition 34 les valeurs propres sont les racines de μ_f. Donc le polynôme

$$M(X) = \prod_{i=1}^{r}(X - \lambda_i)$$

divise μ_f.
Si x_i est un vecteur propre associé à λ_i, alors il est clair que $(f - \lambda_i \mathrm{Id}_E)(x_i) = 0$ et donc

$$M(f)(x_i) = \left(\prod_{j \neq i}(f - \lambda_j \mathrm{Id}_E)\right)(f - \lambda_i \mathrm{Id}_E)(x_i) = 0.$$

Si f est diagonalisable, alors en vertu du théorème 20, $E = \bigoplus_{i=1}^{r} E_{\lambda_i}$. Donc tout $x \in E$ se décompose en $x = x_1 + \cdots + x_r$ avec $x_i \in E_{\lambda_i}$, puis d'après le résultat précédent

$$M(f)(x) = \sum_{i=1}^{r} M(f)(x_i) = 0.$$

On obtient ainsi (ii).
$(ii) \Rightarrow (iii) \Rightarrow (iv)$: trivial.
$(iv) \Rightarrow (i)$: soit P un polynôme annulateur de f scindé à racines simples, que l'on peut toujours supposer unitaire (quitte à le multiplier par une constante) :

$$P(X) = \prod_{i=1}^{s}(X - \mu_i) \quad \text{avec les } \mu_i \text{ deux à deux distincts.}$$

Comme les polynômes $X - \mu_i$ sont premiers entre eux on peut appliquer le corollaire 40, ce qui donne

$$E = \mathrm{Ker}[P(f)] = \bigoplus_{i=1}^{s} \mathrm{Ker}(f - \mu_i \mathrm{Id}_E).$$

Si dans cette décomposition, on ne garde que les μ_i tels que $\mathrm{Ker}(f - \mu_i \mathrm{Id}_E)$ ne soit pas réduit à $\{0\}$, on a obtenu une décomposition de E en somme de sous-espaces propres de f et f est donc diagonalisable. □

Corollaire 42. *Soit $f \in \mathcal{L}(E)$ diagonalisable. Soit F un sous-espace de E, stable par f. Alors* l'endomorphisme induit $f_{|F}$ *est diagonalisable.*

DÉMONSTRATION. Si f est diagonalisable, alors d'après le résultat précédent, il existe $P \in K[X]$ scindé à racines simples tel que $P(f) = 0$. On a bien sur aussi $P(f_F) = 0$. Donc à nouveau d'après le théorème 27, f_F est diagonalisable. □

Corollaire 43. *Soient f et g deux endomorphismes diagonalisables de E qui commutent. Alors il existe une base de diagonalisation commune pour f et g.*

De même, si A et B sont deux matrices diagonalisables qui commutent, alors il existe une matrice inversible P et deux matrices diagonales D et D' telles que

$$A = PDP^{-1} \quad B = PD'P^{-1}$$

Démonstration. Si f est diagonalisable, alors on a $E = \oplus_{i=1}^{r} E_{\lambda_i}$. Comme f et g commutent il est facile de voir que chaque E_{λ_i} est stable par g. En effet, si $x \in E_{\lambda_i}$ alors $f(g(x)) = g(f(x)) = g(\lambda_i x) = \lambda_i g(x)$ i.e. $g(x) \in E_{\lambda_i}$.

L'endomorphisme g_i, induit par g sur E_{λ_i} est d'après le résultat précédent diagonalisable. On peut donc trouver une base $(e_1^i, \ldots, e_{\gamma_i}^i)$ de E_{λ_i} qui est une base de vecteurs propres pour g_i. Bien sûr, chaque e_k^i est un vecteur propre de f associé à la valeur propre λ_i.
En rassemblant les bases obtenues pour chaque E_{λ_i}, on obtient une base de E, qui est une base commune de vecteurs propres pour f et g. □

Ce résultat peut s'étendre sans trop de difficultés au cas de p endomorphismes diagonalisables qui commutent.

IV. Sous-espaces caractéristiques

Dans cette partie on suppose que f est un endomorphisme scindé d'un espace vectoriel E de dimension n. On notera $\{\lambda_1, \ldots \lambda_r\}$ l'ensemble de ses valeurs propres distinctes. Son polynôme caractéristique s'écrit donc

$$\chi_f(X) = (-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i},$$

où α_i est l'ordre de multiplicité de la valeur propre λ_i. De plus, par le Théorème de Cayley-Hamilton son polynôme minimal s'écrit

$$\mu_f(X) = \prod_{i=1}^{r} (X - \lambda_i)^{\beta_i},$$

avec pour $1 \leq i \leq r$, $1 \leq \beta_i \leq \alpha_i$.

IV.1. Définition et premières propriétés.

Définition 44. *On reprend les notations ci-dessus. On appelle* sous-espace caractéristique *de f relatif à la valeur propre λ_i le sous-espace vectoriel de E défini par $N_{\lambda_i} = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\alpha_i}\big]$, où α_i est l'ordre de multiplicité de la valeur propre λ_i.*

Remarque 45. *Si $\alpha_i = 1$ alors $N_{\lambda_i} = E_{\lambda_i}$, autrement dit pour une valeur propre simple le sous-espace caractéristique est égal au sous-espace propre.*

Exemple 46. *Soient E un espace vectoriel de dimension 3 et f l'endomorphisme de E de matrice*

$$M = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

dans une base (e_1, e_2, e_3) de E. Son polynôme caractéristique est $\chi_f(X) = (2 - X)^2(3 - X)$, de sorte que son polynôme minimal est $\mu_f(X) = (X - 2)^k(X - 3)$ avec $k = 1$ ou $k = 2$. Par le calcul on vérifie que la matrice $(M - 2\mathrm{I}_3)(M - 3I_3)$ n'est pas nulle, donc $\mu_f(X) = (X - 2)^2(X - 3)$. Comme les racines de μ_f ne sont pas toutes simples, f n'est pas diagonalisable. Comme 3 est une valeur propre simple, E_3 est un sous-espace de dimension 1, et d'après la forme de la matrice, il est engendré par e_3. L'endomorphisme f n'étant pas diagonalisable, E_2 est de dimension 1, et au vu de M il est engendré par e_1. Pour déterminer N_2 il faut résoudre $(M - 2\mathrm{I}_3)^2 v = 0$ et on trouve que c'est le plan engendré par e_1 et e_2. D'autre part, on a $N_3 = E_3$, car 3 est une valeur propre simple. On peut remarquer que (e_1, e_2, e_3) est une base de E, et par conséquent que $E = N_2 \oplus N_3$. Ce résultat va être montré de manière générale dans la suite.

Proposition 47. *Le sous-espace propre E_{λ_i} associé à la valeur propre λ_i est inclus dans le sous-espace caractéristique N_{λ_i}.*

Démonstration. Soit $x \in E_{\lambda_i}$, c'est-à-dire tel que $(f - \lambda_i \mathrm{Id}_E)(x) = 0$. On a

$$(f - \lambda_i \mathrm{Id}_E)^{\alpha_i}(x) = (f - \lambda_i \mathrm{Id}_E)^{\alpha_i - 1}\big((f - \lambda_i \mathrm{Id}_E)(x)\big) = 0.$$

Ainsi x appartient à N_{λ_i}. □

En particulier ceci entraîne que la dimension de N_{λ_i} est toujours supérieure ou égale à 1, car celle de E_{λ_i} l'est.

Proposition 48. *Les sous-espaces caractéristiques sont stables par f.*

Démonstration. Soit $x \in N_{\lambda_i} = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\alpha_i}\big]$, c'est-à-dire tel que $\big[(f - \lambda_i \mathrm{Id}_E)^{\alpha_i}\big](x) = 0$. On a alors

$$\big[(f - \lambda_i \mathrm{Id}_E)_i^{\alpha}\big] f(x) = f\big(\big[(f - \lambda_i \mathrm{Id}_E)^{\alpha_i}\big](x)\big) = 0.$$

Donc $f(x) \in N_{\lambda_i}$. □

Théorème 49.

On a

$$E = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_r}.$$

Démonstration. On a

$$\chi_f(X) = (-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i}.$$

D'après le Théorème de Cayley-Hamilton, $\chi_f(f) = 0$, c'est-à-dire que $E = \mathrm{Ker}\big[\chi_f(f)\big]$. De plus si $i \neq j$, alors $(X - \lambda_i)^{\alpha_i}$ et $(X - \lambda_j)^{\alpha_j}$ sont des polynômes premiers entre eux. Le corollaire 40 entraîne que

$$\mathrm{Ker}\big[\chi_f(f)\big] = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_r},$$

et donc le résultat. □

Proposition 50. *Pour tout $1 \leq i \leq r$, on a $N_{\lambda_i} = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\beta_i}\big]$, où β_i est l'ordre de multiplicité de la racine λ_i dans le polynôme minimal.*

Démonstration. On observe les faits suivants :

— Comme $\beta_i \leq \alpha_i$ la même démonstration que pour la proposition 47 montre que :

$$\mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\beta_i}\big] \subset \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\alpha_i}\big] = N_{\lambda_i}.$$

— On a $\mu_f(f) = 0$, c'est-à-dire $E = \mathrm{Ker}\big[\mu_f(f)\big]$. De plus, si $i \neq j$, alors $(X - \lambda_i)^{\beta_i}$ et $(X - \lambda_j)^{\beta_j}$ sont des polynômes premiers entre eux. Donc par le corollaire 40 on obtient

$$E = \mathrm{Ker}\big[\mu_f(f)\big] = \mathrm{Ker}\big[(f - \lambda_1 \mathrm{Id}_E)^{\beta_1}\big] \oplus \cdots \oplus \mathrm{Ker}\big[(f - \lambda_r \mathrm{Id}_E)^{\beta_r}\big].$$

En utilisant l'inclusion ci-dessus et le fait que

$$E = \mathrm{Ker}\big[\chi_f(f)\big] = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_r},$$

on en déduit que $N_{\lambda_i} = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\beta_i}\big]$ pour tout $1 \leq i \leq r$. □

Théorème 51.

f est diagonalisable si et seulement si les deux conditions suivantes sont vérifiées :
(i) f est scindé.
(ii) Pour toute valeur propre λ_i de f on a $N_{\lambda_i} = E_{\lambda_i}$.

DÉMONSTRATION. On a déjà vu que si f est diagonalisable alors f est scindé et de plus on a

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

D'autre part, f étant scindé, d'après le théorème 49 on a aussi $E = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_r}$ et d'après la proposition 47, $E_{\lambda_i} \subset N_{\lambda_i}$. Tout ceci entraîne que $E_{\lambda_i} = N_{\lambda_i}$.

Réciproquement, si f est scindé d'après le théorème 49, $E = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_r}$. Ainsi comme $E_{\lambda_i} = N_{\lambda_i}$, $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$, c'est-à-dire que f est diagonalisable. □

IV.2. Applications linéaires restreintes.

Proposition 52. *La restriction $f_i = f|_{N_{\lambda_i}}$ de f au sous-espace caractéristique N_{λ_i} est un endomorphisme de N_{λ_i}. Les polynômes minimal et caractéristique de f_i sont*

$$\mu_{f_i}(X) = (X - \lambda_i)^{\beta_i} \text{ et } \chi_{f_i}(X) = (-1)^{\alpha_i}(X - \lambda_i)^{\alpha_i}.$$

De plus on a $\dim(N_{\lambda_i}) = \alpha_i$.

DÉMONSTRATION. La restriction de f à N_{λ_i} est une application linéaire de N_{λ_i} dans $f\big(N_{\lambda_i}\big)$. Or d'après la proposition 48, N_{λ_i} est stable par f et donc par sa restriction. On en déduit que $f_i \in \mathcal{L}\big(N_{\lambda_i}\big)$.

On sait que

$$N_{\lambda_i} = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\alpha_i}\big] = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\beta_i}\big].$$

On en déduit donc $(f_i - \lambda_i \mathrm{Id}_{N_{\lambda_i}})^{\beta_i} = 0$, c'est-à-dire que le polynôme minimal de f_i divise le polynôme $(X - \lambda_i)^{\beta_i}$. Supposons $\deg(\mu_{f_i}) = \gamma_i < \beta_i$, alors le polynôme

$$Q(X) = (X - \lambda_i)^{\gamma_i} \prod_{j \neq i} (X - \lambda_j)^{\beta_j}$$

annule f, car annule la restriction de f à N_{λ_j} pour tout j, et

$$\gamma_i + \sum_{j \neq i} \beta_j < \sum_{j=1}^{r} \beta_j = \deg(\mu_f).$$

Ceci est impossible car le polynôme minimal de f est le polynôme unitaire de plus bas degré vérifiant cette propriété. Ainsi on a $\gamma_i = \beta_i$ et $\mu_{f_i}(X) = (X - \lambda_i)^{\beta_i}$.

Notons $\dim(N_{\lambda_i}) = \gamma_i$. Comme E est la somme directe des N_{λ_i} il existe une base $\mathcal{B}$ de E telle que $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$, où $\mathcal{B}_i = (e_{1,i}, \ldots, e_{\gamma_i,i})$ est une base de N_{λ_i}. Dans cette base la matrice de f s'écrit

$$M = \begin{pmatrix} M_1 & & & & \\ & M_2 & & & \\ & & \ddots & & \\ & & & M_{r-1} & \\ & & & & M_r \end{pmatrix}$$

où les M_i sont les matrices des f_i dans les bases $\mathcal{B}_i$. Or le polynôme minimal de f_i est $\mu_{f_i}(X) = (X-\lambda_i)^{\beta_i}$. Donc la seule valeur propre de M_i (c'est-à-dire de f_i) est λ_i et son polynôme caractéristique est égal à $(-1)^{\gamma_i}(X-\lambda_i)^{\gamma_i}$. On obtient donc l'égalité suivante :

$$\begin{aligned}(-1)^n\prod_{i=1}^r(X-\lambda_i)^{\alpha_i} &= \chi_f(X) = \prod_{i=1}^r \det(M_i - X\mathrm{I}_{\gamma_i}) = \prod_{i=1}^r \chi_{f_i}(X) = \prod_{i=1}^r (-1)^{\gamma_i}(X-\lambda_i)^{\gamma_i} \\ &= (-1)^n\prod_{i=1}^r(X-\lambda_i)^{\gamma_i}\end{aligned}$$

car

$$\sum_{i=1}^r \gamma_i = \sum_{i=1}^r \dim(N_{\lambda_i}) = \dim(E) = n.$$

On en déduit que $\gamma_i = \alpha_i$, c'est-à-dire que $\dim(N_{\lambda_i}) = \alpha_i$. De plus on obtient que $\chi_{f_i}(X) = (-1)^{\alpha_i}(X-\lambda_i)^{\alpha_i}$. □

IV.3. Trigonalisation des matrices en blocs relatifs aux sous-espaces caractéristiques.

On a vu que si on prend $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ avec $\mathcal{B}_i = (e_{1,i}, \ldots, e_{\alpha_i,i})$ une base de N_{λ_i}, alors la matrice de f dans $\mathcal{B}$ s'écrit

$$M = \begin{pmatrix} M_1 & & & & \\ & M_2 & & & \\ & & \ddots & & \\ & & & M_{r-1} & \\ & & & & M_r \end{pmatrix}.$$

Pour trigonaliser M il suffit alors de trigonaliser chaque M_i.

Exemple : On veut trigonaliser l'endomorphisme f dont la matrice dans la base canonique est

$$\begin{pmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{pmatrix}.$$

On commence par chercher le polynôme caractéristique de f. On obtient

$$\chi_f(X) = (X-1)^2(X+1)^2.$$

f a donc deux valeurs propres doubles, 1 et -1. On détermine alors les sous-espaces caractéristiques. Pour $\lambda = -1$, on obtient :

$$N_{-1} = \mathrm{Ker}[(f + 1\mathrm{Id}_E)^2] = \mathrm{Ker}[(M + \mathrm{I}_4)^2].$$

On a donc

$$(M+\mathrm{I}_4)^2\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{cases} 12z - 8t &= 0 \\ 8z - 4t &= 0 \\ 12z - 8t &= 0 \\ 8z - 4t &= 0 \end{cases} \iff \begin{cases} 3z &= 4t \\ 2z &= t \end{cases}$$

Une base de N_{-1} est donc donnée par u_1 et u_2 de coordonnées dans la base canonique

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \text{ et } \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Pour $\lambda = 1$, $N_1 = \mathrm{Ker}[(f - 1\mathrm{Id}_E)^2] = \mathrm{Ker}[(M - \mathrm{I}_4)^2]$. On a donc

$$(M-\mathrm{I}_4)^2\begin{pmatrix}x\\y\\z\\t\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}\iff\begin{cases}-12x+16y+12z-16t=0\\-16x+20y+16z-20t=0\end{cases}\iff\begin{cases}x=z\\y=t\end{cases}.$$

Une base de N_1 est donc donnée par u_3 et u_4 de coordonnées dans la base canonique $\begin{pmatrix}1\\0\\1\\0\end{pmatrix}$ et $\begin{pmatrix}0\\1\\0\\1\end{pmatrix}$.

La matrice de passage de la base canonique à la nouvelle base (u_1,u_2,u_3,u_4), ainsi que son inverse, sont données par

$$P=\begin{pmatrix}1&0&1&0\\0&1&0&1\\0&0&1&0\\0&0&0&1\end{pmatrix}\quad\text{et}\quad P^{-1}=\begin{pmatrix}1&0&-1&0\\0&1&0&-1\\0&0&1&0\\0&0&0&1\end{pmatrix}.$$

Donc la matrice de f dans la nouvelle base s'écrit :

$$\widetilde{M}=P^{-1}MP=\begin{pmatrix}3&-4&0&0\\4&-5&0&0\\0&0&3&-2\\0&0&2&1\end{pmatrix}.$$

On trigonalise alors chacune des sous matrices

$$M_1=\begin{pmatrix}3&-4\\4&-5\end{pmatrix}\quad\text{et}\quad M_2=\begin{pmatrix}3&-2\\2&1\end{pmatrix}.$$

Le polynôme caractéristique de M_1 est égal à $(X+1)^2$. Le sous-espace propre de M_1 associé à la valeur propre -1 est déterminé par

$$(M_1+\mathrm{I}_2)\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}\iff\begin{cases}4x_1-4x_2&=&0\\4x_1-4x_2&=&0\end{cases}\iff x_1=x_2.$$

Une base de cet espace est donc $\begin{pmatrix}1\\1\end{pmatrix}$ dans (u_1,u_2), i.e.

$$v_1=u_1+u_2=\begin{pmatrix}1\\1\\0\\0\end{pmatrix}.$$

On complète par un second vecteur de N_{-1} non colinéaire à v_1, par exemple $v_2=u_1$.
Le polynôme caractéristique de M_2 est égal à $(X-1)^2$. Le sous-espace propre de M_2 associé à la valeur propre 1 est déterminé par

$$(M_2-\mathrm{I}_2)\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}\iff\begin{cases}2x_3-2x_4&=&0\\2x_3-2x_4&=&0\end{cases}\iff\{\ x_3\ =\ x_4$$

Une base de cet espace est donc $\begin{pmatrix}1\\1\end{pmatrix}$ dans (u_3,u_4), i.e.

$$v_3=u_3+u_4=\begin{pmatrix}0\\0\\1\\1\end{pmatrix}.$$

On complète par un second vecteur de N_1 non colinéaire à v_3, par exemple $v_4 = u_3$.
La matrice de passage de la base (u_1, u_2, u_3, u_4) à la base (v_1, v_2, v_3, v_4) et son inverse, sont égales à

$$P' = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad P'^{-1} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

Donc la matrice de f dans la base (v_1, v_2, v_3, v_4) s'écrit :

$$M' = (P')^{-1}\widetilde{M}(P') = \begin{pmatrix} -1 & 4 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

V. Endomorphismes nilpotents

V.1. Caractérisation des endomorphismes nilpotents.

Définition 53. *On dit qu'un endomorphisme f (respectivement une matrice M) est* nilpotent *(respectivement* nilpotente*) si et seulement si il existe un entier $k \geq 1$ tel que $f^k = 0$, l'endomorphisme identiquement nul, (respectivement $M^k = 0$, la matrice nulle). Le plus petit entier $k \geq 1$ vérifiant ceci, est alors appelé l'*indice de nilpotence *de f (respectivement de M).*

Proposition 54. *Un endomorphisme d'un espace vectoriel de dimension n est nilpotent si et seulement si son polynôme caractéristique est $(-1)^n X^n$.*

Démonstration. Si f est un endomorphisme nilpotent, alors il existe un entier $k \geq 1$ tel que $f^k = 0$, c'est-à-dire que le polynôme X^k appartient à l'idéal annulateur de f. Donc par définition, le polynôme minimal de f divise X^k ; d'après le corollaire 36, il est de la forme X^j avec j un entier tel que $1 \leq j \leq k$. L'unique racine de ce polynôme est donc 0, et donc l'unique valeur propre de f est 0. Comme l'espace vectoriel est de dimension n, le polynôme caractéristique de f est donc $(-1)^n X^n$. Réciproquement, si le polynôme caractéristique de f est $(-1)^n X^n$, d'après le Théorème de Cayley-Hamilton $\chi_f(f) = f^n = 0$, c'est-à-dire que f est nilpotent. □

Une conséquence de ce résultat est que le polynôme minimal d'un endomorphisme nilpotent est égal à X^p avec $p \leq n$ (ceci montre au passage que l'indice de nilpotence est inférieur ou égal à la dimension de E). D'autre part, comme il n'a qu'une seule valeur propre qui est zéro, il n'est pas diagonalisable sauf s'il est nul. Mais il est trigonalisable.

V.2. Décomposition de Dunford.

Théorème 55.

Soit f un endomorphisme d'un espace vectoriel E. Si f est scindé alors il existe un unique couple (u, v) d'endomorphismes tel que
- *$f = u + v$.*
- *u est diagonalisable et v est nilpotent.*
- *u et v commutent.*

De plus, u et v sont des polynômes en f.

Démonstration. Montrons que le couple (u, v) existe. Comme f est scindé d'après le théorème 34 on a $E = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_r}$, où les $N_{\lambda_i} = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\alpha_i}\big]$ sont les sous-espaces caractéristiques de f correspondant aux différentes valeurs propres $(\lambda_i)_{1 \leq i \leq r}$ de f. On définit alors u et v par leurs restrictions aux sous-espaces caractéristiques de la manière suivante :

$$u_{|N_{\lambda_i}} = \lambda_i \mathrm{Id}_{N_{\lambda_i}} \text{ et } v_{|N_{\lambda_i}} = f_{|N_{\lambda_i}} - \lambda_i \mathrm{Id}_{N_{\lambda_i}}.$$

Comme les sous-espaces caractéristiques sont supplémentaires et que la restriction de u à chaque N_{λ_i} est diagonale, u est diagonalisable. D'autre part, par définition chaque restriction de v est nilpotente d'ordre $n_i \leq \beta_i \leq \alpha_i$ et donc v sera aussi nilpotente d'ordre $\max_{1\leq i\leq r}(n_i)$. Enfin il est clair que pour tout $1 \leq i \leq r$, $u_{|N_{\lambda_i}}$ commute avec $v_{|N_{\lambda_i}}$ car l'identité commute avec tout endomorphisme, donc u et v commutent.

De façon alternative, on peut définir u et v de la manière suivante : si on note p_i la projection sur N_{λ_i} parallèlement à $\sum_{j\neq i} N_{\lambda_j}$, alors $u = \sum_i \lambda_i p_i$ et $v = f - u$. D'après le corollaire 40, u et v sont des polynômes en f et commutent donc.

Pour l'unicité de la décomposition, considérons un autre couple (u', v') vérifiant les trois propriétés du théorème. Comme u' et v' commutent, ils commutent aussi avec $f = u' + v'$ et par conséquent aussi avec u et v qui sont des polynômes en f. D'après le corollaire 43, u et u' sont diagonalisables dans une même base, ce qui entraîne que $u - u'$ est diagonalisable également. Or, $u - u' = v' - v$ est nilpotente car

$$(v - v')^{p+q} = \sum_{i+j=p+q} \binom{p+q}{i} (-1)^j v^i v'^j = 0$$

si p (resp. q) est l'ordre de nilpotence de v (resp. v'). Ceci entraîne que $u - u' = 0$, et $v = v'$. □

Exemple 56. *Soit la matrice*

$$M = \begin{pmatrix} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \end{pmatrix}.$$

On vérifie par le calcul que son polynôme caractéristique est $\chi_M(X) = -X^2(X+1)$. *Les valeurs propres sont donc :* -1 *valeur propre simple et* 0 *valeur propre double. Le sous-espace caractéristique associé à la valeur propre* -1*, est égal au sous-espace propre associé à cette même valeur propre. Pour le déterminer on résout* $(M+\mathrm{I}_3)v = 0$. *On trouve qu'il est engendré par* $e_1 = (1, -1, -2)$. *Pour le sous-espace caractéristique associé à la valeur propre* 0 *on résout* $(M - 0.\mathrm{I}_3)^2 v = M^2 v = 0$ *et on constate qu'il est engendré par* $e_2 = (1, 2, 0)$ *et* $e_2 = (0, 1, 1)$. *Nous allons maintenant écrire la matrice dans la nouvelle base. La matrice de passage de la base canonique de* $\mathbb{R}^3$ *à la base* (e_1, e_2, e_3) *est*

$$P = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ -2 & 0 & 1 \end{pmatrix} \text{ et son inverse } P^{-1} = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 4 & -2 & 3 \end{pmatrix}.$$

On en déduit que

$$\widetilde{M} = P^{-1}MP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \widetilde{D} + \widetilde{N}.$$

La décomposition de Dunford de M *est donc* $M = D + N$ *avec*

$$D = P\widetilde{D}P^{-1} = \begin{pmatrix} -2 & 1 & -1 \\ 2 & -1 & 1 \\ 4 & -2 & 2 \end{pmatrix} \quad \text{et} \quad N = P\widetilde{N}P^{-1} = M - D = \begin{pmatrix} 4 & -2 & 3 \\ 8 & -4 & 6 \\ 0 & 0 & 0 \end{pmatrix}.$$

On peut vérifier que $DN = ND$.

V.3. Décomposition de Jordan.

Définition 57. *On appelle* bloc de Jordan *de taille* p *et de valeur propre* λ *une matrice carrée d'ordre* p *de la forme*

$$J_{\lambda,p} = \begin{pmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & \dots & 0 & \lambda \end{pmatrix}.$$

Théorème 58.

Soit f un endomorphisme scindé de E. On suppose que le polynôme caractéristique de f est

$$(-1)^n \prod_{i=1}^{r} (X-\lambda_i)^{\alpha_i}, \qquad \text{avec } \sum_{i=1}^{r} \alpha_i = n,$$

où les λ_i, $1 \leq i \leq r$ sont deux à deux distincts. Il existe alors une base de E dans laquelle la matrice de f est diagonale par blocs de la forme :

$$\begin{pmatrix} J_{\lambda_1,n_{1,1}} & 0 & \cdots & 0 \\ 0 & J_{\lambda_1,n_{1,2}} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & J_{\lambda_r,n_{r,s_r}} \end{pmatrix},$$

où pour $1 \leq i \leq r$ et $1 \leq j \leq s_i$, le bloc diagonal $J_{\lambda_i,n_{i,j}}$ est un bloc de Jordan de taille $n_{i,j}$ de valeur propre λ_i. Pour $1 \leq i \leq r$, on a $\sum_{j=1}^{s_i} n_{i,j} = \alpha_i$. L'entier $n_{i,1}$ est égal à la multiplicité de λ_i comme racine du polynôme minimal.

La démonstration que nous donnerons est empruntée au livre *les maths en tête, Mathématiques pour M', Algèbre* par Xavier Gourdon (éditions Ellipses).

On peut remarquer que la matrice de Jordan donnée par le théorème est triangulaire supérieure puisque les blocs de Jordan le sont. Sur la diagonale figurent les valeurs propres, avec la multiplicité qu'elles ont dans le polynôme caractéristique, et les coefficients de la diagonale secondaire $\{(i,j), j = i+1\}$ sont égaux à 1 ou 0. Tous les autres coefficients sont nuls.

La preuve du théorème 58 découle immédiatement du résultat dans le cas particulier des endomorphismes nilpotents. En effet, il suffit de l'appliquer à chaque endomorphisme $f_{|N_{\lambda_i}} - \lambda_i \mathrm{Id}_{N_{\lambda_i}}$ où les N_{λ_i} sont les sous-espaces caractéristiques de f. Nous allons donc montrer le théorème 58 en supposant f nilpotent.

Soit $p \in \mathbb{N}^*$ l'indice de nilpotence de f i.e. $f^p = 0$ et $f^{p-1} \neq 0$. Pour tout i, notons $F_i = \operatorname{Ker} f^i$.

Commençons par montrer que

$$\{0\} = F_0 \subsetneq F_1 \subsetneq \cdots \subsetneq F_{p-1} \subsetneq F_p = E$$

Les inclusions sont évidentes car pour tout $i \geq 1$, $f^i(x) = f(f^{i-1}(x)) = 0$ si $x \in \operatorname{Ker} f^{i-1}$. Il s'agit de voir que ces inclusions sont toutes strictes. En fait, on peut définir p de la façon suivante : la suite des $n_i = \dim F_i$ est croissante par la remarque que l'on vient de faire et stationnaire puisque c'est une suite croissante d'entiers, majorée par $n = \dim E$. On peut alors définir p par $p = \min\{i \in \mathbb{N} \,|\, n_{i+1} = n_i\}$. Puisque $F_n = E$, il suffit de vérifier que pour tout $i \geq p$, $F_{i+1} = F_i$. Or, si $i \geq p$ (avec p comme on vient de le définir) alors

$$F_{i+1} = \{x \in E \,|\, f^{p+1}(f^{i-p}(x)) = 0\} = \{x \in E \,|\, f^{i-p}(x) \in F_{p+1}\} = \{x \in E \,|\, f^{i-p}(x) \in F_p\} = F_i.$$

Remarquons aussi que par construction, pour tout $i \geq 1$, $f(F_i) \subset F_{i-1}$.

Nous allons maintenant montrer l'existence de sous-espaces vectoriels $G_1, \ldots, G_p$ et $H_1, \ldots H_{p-1}$ de E tels que

(1) $\forall i$, $1 \leq i \leq p$, $F_i = G_i \oplus F_{i-1}$,

(2) $\forall i$, $1 \leq i \leq p-1$, f envoie injectivement G_{i+1} dans G_i,

(3) $\forall i$, $1 \leq i \leq p-1$, $G_i = f(G_{i+1}) \oplus H_i$.

Pour commencer, soit G_p un supplémentaire de F_{p-1} dans $F_p = E$, de sorte que $F_p = G_p \oplus F_{p-1}$. On a $\operatorname{Ker} f \cap G_p = F_1 \cap G_p \subset F_{p-1} \cap G_p = \{0\}$ et $f(G_p) \subset f(F_p) \subset F_{p-1}$. Par conséquent, f envoie injectivement G_p dans F_{p-1}.

Ensuite, $f(G_p) \cap F_{p-2} = \{0\}$. En effet, soit $x \in f(G_p) \cap F_{p-2}$ Il existe $y \in G_p$ tel que $f(y) = x$ et on a $0 = f^{p-2}(x) = f^{p-1}(y)$ donc $y \in F_{p-1} \cap G_p = \{0\}$ donc $y = 0$ et $x = f(y) = 0$.

On a donc $f(G_p) \oplus F_{p-2} \subset F_{p-1}$ et il existe un sous-espace vectoriel H_{p-1} tel que $f(G_p) \oplus F_{p-2} \oplus H_{p-1} = F_{p-1}$. Si on pose $G_{p-1} = f(G_p) \oplus H_{p-1}$, on a donc $F_{p-1} = G_{p-1} \oplus F_{p-2}$ et f envoie injectivement G_p dans G_{p-1}.

Les propriétés (1), (2) et (3) sont donc montrées pour $i = p-1$. Pour les montrer pour $i \in \{1, \dots, p-2\}$, nous allons effectuer une « récurrence descendante ». Supposons le résultat prouvé pour $i+1 \in \{2, \dots, p-1\}$ et montrons-le pour i.

Regardons la restriction de f à G_{i+1}. On a $\operatorname{Ker} f \cap G_{i+1} = F_1 \cap G_{i+1} \subset F_i \cap G_{i+1} = \{0\}$ et $f(G_{i+1}) \subset f(F_{i+1}) \subset F_i$. Par conséquent, f envoie injectivement G_{i+1} dans F_i.

Par ailleurs, $f(G_{i+1}) \cap F_{i-1} = \{0\}$. En effet, soit $x \in f(G_{i+1}) \cap F_{i-1}$. Il existe $y \in G_{i+1}$ tel que $x = f(y)$. Or, $x \in F_{i-1}$ donc $0 = f^{i-1}(x) = f^i(y)$ d'où $y \in F_i \cap G_{i+1} = \{0\}$, i.e. $y = 0$ donc $x = f(y) = 0$.

On a donc $f(G_{i+1}) \oplus F_{i-1} \subset F_i$ et il existe un sous-espace vectoriel H_i tel que $f(G_{i+1}) \oplus F_{i-1} \oplus H_i = F_i$. On pose alors $G_i = f(G_{i+1}) \oplus H_i$, de sorte que $F_i = G_i \oplus F_{i-1}$ et f envoie injectivement G_{i+1} dans G_i.

Les sous-espaces vectoriels $G_p, \dots, G_1$ et $H_{p-1}, \dots, H_1$ sont ainsi construits de proche en proche. Les propriétés (1), (2) et (3) pour $i = 1$ sont $\operatorname{Ker} f = F_1 = G_1 \oplus F_0 = G_1 \oplus \{0\} = G_1$ et $G_1 = f(G_2) \oplus H_1$.

En résumé, la suite $G_1, \dots, G_p$ vérifie $E = G_1 \oplus \cdots \oplus G_p$, $G_1 = F_1 = \operatorname{Ker} f$ et pour tout $2 \leq i \leq p$, f envoie injectivement G_i dans G_{i-1}.

Maintenant, si $(e_1^i, \dots, e_{q_i}^i)$ est une base de G_i, $(f(e_1^i), \dots, f(e_{q_i}^i))$ est une famille libre de G_{i-1} que l'on peut compléter en une base de G_{i-1}. En partant d'une base de G_p et en construisant les autres bases de la sorte, de proche en proche, puis en réunissant toutes ces bases, on obtient une base de E. Si on la réordonne correctement, la matrice de f dans cette base a la forme voulue.