

Algèbre Linéaire Avancée (1er Semestre)¹

Philippe Michel

¹Monday 16th December, 2024, 17:08

Table des matieres

Introduction	5
Chapitre 1. Le langage des ensembles	7
1.1. La theorie des ensembles	7
1.2. Operations sur les ensembles	12
1.3. Applications entre ensembles	15
1.4. Cardinal d'un ensemble	22
Chapitre 2. Groupes	27
2.1. Groupes abstraits	27
2.2. Le cas du groupe symetrique	30
2.3. Sous-groupes	32
2.4. Morphismes de groupes	36
2.5. Action d'un groupe sur un ensemble	43
Chapitre 3. Anneaux	49
3.1. Anneaux	49
3.2. Elements inversibles	53
3.3. Sous-anneau	55
3.4. Morphismes d'anneaux	56
3.5. Anneau quotient	57
Chapitre 4. Corps	61
4.1. Corps	61
4.2. Construction de corp: corps des fractions	63
4.3. Construction de corps: corps quotient	65
4.4. Caracteristique d'un corps, Sous-corps premier	67
Recapitulatif concernant la caracteristique d'un corps	69
Chapitre 5. Modules et Espaces Vectoriels	71
5.1. Module sur un anneau	71
5.2. Espaces vectoriel	78
5.3. Famille generatrice, libre, base	82
5.4. Espaces vectoriels de dimension infinie	90
Chapitre 6. Applications lineaires	93
6.1. Le Theoreme Noyau-Image	93
6.2. Structure et dimension des espaces d'applications lineaires	95
6.3. Proprietes fonctionelles des coefficients d'une application lineaire	102
Chapitre 7. Matrices	109

7.1. Matrices et applications lineaires	109
7.2. Structure des espaces de matrices	113
7.3. L'algebre des matrices carrees	120
7.4. Changement de base	123
Chapitre 8. Interlude: le corps des nombres complexes	131
8.1. Origine des nombres complexes	131
8.2. Construction matricielle d'extensions quadratiques	132
8.3. Le corps des nombres complexes; proprietes de base	136
8.4. Le plan complexe	142
8.5. Equations polynomiales complexes	144
Chapitre 9. Operations elementaires sur les matrices	151
9.1. Operation elementaires sur les lignes	151
9.2. Echelonnage	154
9.3. Applications	157
9.4. Operation elementaires sur les colonnes	162
Chapitre 10. Determinants	165
10.1. Formes multilineaires	165
10.2. Determinants	179
10.3. Calcul de determinants	185
10.4. Le determinant en caracteristique 2	194
Chapitre 11. Le polynome caracteristique	197
11.1. Le polynome caracteristique d'une matrice	197
11.2. Le polynome caracteristique d'un endomorphisme	201
11.3. Le Theoreme de Cayley-Hamilton	202
Appendice A. L'anneau des polynomes sur un corps	205
A.1. Preliminaire: fonctions polynomiales	205
A.2. Les polynomes sont des suites	206
A.3. Structure d'anneau	209
A.4. Division et factorisation	212
A.5. Application a la construction de corps	220

Introduction

Le terme "Algebre" est derive du mot arabe *al-jabr* qui est tire du titre d'un ouvrage du mathematicien persan *Al-Khwarizmi*, redige vers 825 (source wikipedia) et intitule

Kitab al-mukhtasar fi hisab al-jabr wa-l-muqabala

Abrege du calcul par la restauration et la comparaison.

L'ouvrage fournissait des procedures generales de calcul pour resoudre des problemes pratiques lies aux actes legaux (partage lors d'un heritage, subdivision de terrains et calculs d'aires) qui conduisaient a resoudre des equations lineaires ou quadratiques. Le nom "Al-Khwarizmi" a d'ailleurs donne naissance au mot "Algorithme".

De nos jours le terme "Algebre" designe plutot l'etude et la classification de structures mathematiques formelles liees aux operations. l'*Algebre Lineaire* se concentre plus particulierement sur l'etude des "espaces vectoriels". Cependant avant d'arriver a cette notion, nous auront besoin d'introduire d'autre structures algebrique plus generales,

- Les "groupes",
- les "anneaux"
- et les "corps" (qui sont des anneaux particuliers) ainsi que
- les "modules" sur les anneaux, les espaces vectoriels sont des modules sur des corps.

L'etude des premiers releve de la "theorie des groupes" (qui sera developpee plus en details dans le cours MATH-113) et celle des trois au tres releve de "l'algebre commutative" (qui sera discutee en deuxieme annee) cependant, comme on va le voir, tous ces sujets sont intimement connectes et il est impossible de traiter l'un de ces sujets sans avoir recours aux autres.

Avant cela nous aurons besoin d' introduire le langage des *ensembles*.



CHAPITRE 1

Le langage des ensembles

“Le langage est un ensemble de citations.”

1.1. La theorie des ensembles

La notion d'ensemble (et les operations qui y sont associees comme l'intersection ou la reunion) est tellement naturelle qu'on peut legitimement s'interroger sur le bien-fonde de construire une "theorie des ensembles". Cette necessite, bien reelle, n'est vraiment apparue que dans le cours du 19eme siecle quand certains mathematiens ont obtenus des objets mathematiques (d'origine logique, analytique ou geometrique) semblant posseder des proprietes paradoxales et en tout cas defiant l'intuition primaire. Dans certains cas on a pu montrer qu'une re-interpretation convenable ou le developpement d'une theorie plus rigoureuse permettait de donner un sens a ces objets; dans d'autres, on a realise que de tels objets conduisaient a une contradiction avec les theories existantes ce qui a conduit a une remise en cause des fondements meme sur lesquels le raisonnement mathematique etait basee. La¹ Theorie des Ensembles est l'un des fruits de ces reflexions.

Il est impossible, dans le cadre de ce cours, de presenter une definition rigoureuse de la notion d'ensemble; nous preferons renvoyer le lecteur a un cours plus avance de "logique mathematique" (par exemple MATH-381) et en attendant nous en remettons a l'intuition du lecteur qui est souvent bien suffisante.

Cependant nous voulons insister que le developpement d'une theorie des ensembles ce n'est pas du tout evident. Cela necessite au prealable d'introduire un concept de logique appele *calcul des predicats du premier ordre*: c'est un *langage* forme de *constituants* et muni d'une *syntaxe* permettant creer des phrases (appelees "formules" ou "predicats") qui s'organisent en *proprietes* ou en *relations* et qui permet de modeliser le raisonnement mathematique usuel. Une fois cela defini, on peut construire une *theorie des ensembles* a partir d' *axiomes* convenables de sorte que la theorie soit *consistante* (ie. ne conduise pas a des contradictions comme c'etait le cas avec des constructions moins precises). Il n'y a pas de choix unique pour les axiomes mais la plupart du temps on utilise les axiomes ZF ou ZFC²)

Le calcul des predicats du premier ordre (egalitaire) est un langage dont les phrases sont composees de

- *Divers alphabets*: des ensembles de symboles (usuellement des lettres ou des ensembles de lettres) representant soit des *variables*, $x, y, z \dots$ ou des *constantes* $a, b, c, \dots$ qui permettent d'identifier les divers objets sur lesquels on travaille et egalement les predicats ou des fonctions

$$P(\cdot), Q(\cdot), f(\cdot), \cos(\cdot)$$

¹il y a en fait plusieurs theories possibles

²d'apres Zermelo et Fraenkel

permettant de d'expliciter les relations existant entre les divers ensembles considérés.

– *Quantificateurs logiques*:

– Le quantificateur *universel* $\forall$:

$\forall x P(x)$: "pour tout x , la propriété $P(x)$ est vraie" .

– Le quantificateur *existentiel* $\exists$:

$\exists x P(x)$ ($\exists x |P(x)$) : "il existe x tel que la propriété $P(x)$ est vraie"

ou la variante

$\exists! x P(x)$ (ou $\exists! x |P(x)$) : "il existe un unique x tel que la propriété $P(x)$ est vraie".

– Un symbole pour la relation *d'égalité* = permettant d'exprimer le fait que deux éléments sont les mêmes et peuvent être librement *substitués* dans toute formule impliquant l'un ou l'autre.

– *Connecteurs logiques* reliant les prédicats

$\wedge$: "et", $\vee$: "ou"

$\implies$: "implique"; $\iff$: "équivalent à, si et seulement si"

$\neg$: "négation" "contraposée".

– Des règles syntaxiques de construction des formules (l'orthographe et la grammaire du langage en question).

– D'un *système de deduction* permettant de dériver des propositions (appelées *conclusions*) à partir de propositions existantes (appelées *premices*). Pour initier le processus de deduction, on se donne un ensemble de propositions initiales appelées *axiomes*.

Ce langage est interprété dans le cadre d'un *modèle* (dans notre cas, les ensembles; il peut a priori y avoir plusieurs modèles associés à un langage donné) et il sert à exprimer diverses relations existantes entre les divers objets du *modèle*. En particulier on peut déterminer si certaines de ces formules (celles qui sont "closes": une formule est *close* si toutes les variables qui apparaissent devant ont devant elles l'un des deux quantificateurs logiques $\forall, \exists$) sont "vraies" ou "fausses" quand on leur applique des éléments du modèle et le système de deduction ci-dessus est construit de sorte qu'il preserve ces valeurs de vérité: si des formules "premices" sont "vraies" alors la formule "conclusion" doit être "vraie" (les axiomes initiaux qu'on a pu se donner en départ doivent également être vrais).

1.1.1. Ensembles. La catégorie des *Ensembles* est une collection d'objets (les ensembles) munies d'une relation d'*appartenance* qui lie entre eux certains couples d'ensembles. Soient e, E deux ensembles, si ces ensembles sont liés par cette relation, on le note

$$e \in E.$$

On dit alors que " e est un élément de E " ou que " e appartient à E ".

1.1.2. Sous-ensemble. A partir de cette relation d'appartenance, on forme la relation *d'inclusion*: un ensemble A est contenu (ou inclu) dans un ensemble B

$$A \subset B$$

si tout element de A appartient a B :

$$\forall a, a \in A \implies a \in B.$$

On dit egalement que A est un *sous-ensemble* de B et on le note

$$A \subset B.$$

REMARQUE 1.1.1. les relations d'appartenance $\in$ et d'inclusion $\subset$ sont distinctes. On peut tres bien avoir $A \in B$ (A est un element de B) sans que l'on ait $A \subset B$ et on peut tres bien avoir $A \subset B$ sans que $A \in B$ (A est inclus dans B).

1.1.3. Axiomes de la theorie des ensembles. Les ensembles verifient un certain nombre d'axiomes (une dizaine) qui permettent la construction de nouveaux ensembles a partir d'ensembles primitifs: on va donner quelques uns des ces axiomes:

1.1.3.1. *Existence de l'ensemble vide.* Il existe un ensemble ne contenant aucun autre ensemble comme element et qui est inclut ($\subset$) dans tout ensemble (y compris dans lui-meme): l'*ensemble vide* qu'on note

$$\emptyset.$$

On a donc

$$\forall E, E \not\subset \emptyset \wedge \emptyset \subset E.$$

REMARQUE 1.1.2. Il est important ici de ne pas confondre $\in$ et $\subset$.

1.1.3.2. *Axiome de la double-inclusion.* Deux ensembles sont egaux si ils sont inclus l'un dans l'autre (si ils possedent les meme elements):

$$A \subset B \wedge B \subset A \implies A = B.$$

1.1.3.3. *Ensemble des parties d'un ensemble.* Si A est un ensemble, il existe un ensemble dont les elements sont les sous-ensembles de A ; cet ensemble (unique par l'axiome de la double inclusion) est appelle l'ensemble des parties (ou des sous-ensembles) de A on le note $\mathcal{P}(A)$:

$$\mathcal{P}(A) = \{B, B \subset A\}.$$

En particulier on a toujours

$$\emptyset, A \in \mathcal{P}(A)$$

donc $\mathcal{P}(A)$ contient toujours au moins 1 element (et au moins 2 ssi $A \neq \emptyset$).

1.1.3.4. *Axiome de la reunion.* Soit E un ensemble, il existe un ensemble, la *reunion* de E , qu'on notera

$$\bigcup_E$$

dont les elements sont exactement les elements des elements de E (on rappelle que les element de E sont eux-meme des ensembles.

1.1.3.5. *Axiome de la paire.* Soient A et B deux ensembles, si existe un ensemble (nécessairement unique par l'axiome de la double inclusion) dont les éléments sont exactement A et B , on le note

$$\{A, B\}.$$

En particulier, si $A = B$, on forme l'ensemble (à un élément)

$$\{A, A\} = \{A\}$$

qu'on appelle le *singleton* $\{A\}$.

REMARQUE 1.1.3 (Reunion d'ensembles). Soient A et B deux ensembles, par l'axiome de la paire il existe un ensemble $E = \{A, B\}$ dont les éléments sont les ensembles A et B . Par l'axiome de la reunion, la reunion de $E = \{A, B\}$ est un ensemble composé des éléments de A et des éléments de B : on l'appelle reunion de A et B et on le note

$$\bigcup_{\{A, B\}} = A \cup B = \{e \mid e \in A \wedge a \in B\}.$$

Plus généralement on montre que si I est un ensemble non vide et $(A_i)_{i \in I}$ une famille d'ensembles indexée par I (la donnée pour chaque élément $i \in I$ d'un ensemble A_i) alors il existe un ensemble dont les éléments sont exactement les éléments appartenant à l'un des A_i , on le note

$$\bigcup_{i \in I} A_i.$$

1.1.3.6. ...et 5 autres axiomes supplémentaires dans la théorie ZFC. notamment "l'Axiome de l'infini" et l'Axiome du choix".

EXEMPLE 1.1.1. Quelques ensembles

- On a déjà vu l'ensemble vide qu'on va noter également

$$\emptyset =: 0.$$

- L'ensemble des parties de l'ensemble vide $\mathcal{P}(\emptyset)$ possède l'ensemble vide comme seul élément et on le note

$$\mathcal{P}(\emptyset) = \{\emptyset\} =: 1.$$

- Par l'axiome de la paire l'ensemble suivant existe

$$\{\emptyset, 1\} = \{\emptyset, \{\emptyset\}\} =: 2,$$

puis en itérant (en appliquant la Remarque 1.1.3) on construit

$$3 := \bigcup_{2, \{2\}} = \{\emptyset, \{\emptyset\}, 2\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} = \{0, 1, 2\}, \quad 4 := \{0, 1, 2, 3\}, \dots$$

- On "arrive" alors à construire l'ensemble des entiers naturels:

$$\mathbb{N} = \{0, 1, 2, 3, \dots\}$$

par un processus récursif: si l'entier n a été construit on définit son *successeur* n^+ comme étant l'ensemble obtenu comme reunion

$$n^+ = \bigcup_{\{n, \{n\}\}} = n \cup \{n\}$$

ie. l'ensemble (cet existe par l'axiome de la reunion) dont les éléments sont les éléments de n et le singleton $\{n\}$; on construit alors le successeur de ce n^+ , etc...le

fait de pouvoir repeter cette construction une infinite de fois necessite *l'axiome de l'infini*.

On defini sur $\mathbb{N}$ le relation "inferieur ou egal" $\leq$ en posant pour $m, n \in \mathbb{N}$

$$m \leq n \iff m \subset n$$

et on definit egalement $\geq$, $<$ et $>$.

- Puis on peut a partir de cela construire l'ensemble des entiers relatifs:

$$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$$

(cela necessite la notion de produit cartesien, cf. ci-dessous) et on peut alors etendre la relation $\leq$.

- On construit ensuite l'ensemble des nombres rationnels:

$$\mathbb{Q} = \left\{ \frac{p}{q}, p, q \in \mathbb{Z}, q \neq 0 \right\},$$

auquel on etend la relation $\leq$

- et vous verrez en analyse la construction de l'ensemble des nombres *reels* $\mathbb{R}$,
- et enfin a partir de $\mathbb{R}$, on construira dans ce cours (en admettant l'existence de $\mathbb{R}$) l'ensemble des nombres *complexes* $\mathbb{C}$ et on a donc

$$\emptyset \subset \mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

1.1.4. Notation. Comme on l'a vu dans les exemples, on designera un ensemble et les elements qu'il contient par la notation "crochets":

$$E = \{\dots\}.$$

Entre ces crochets $\{\dots\}$ on mettra soit

- La liste explicite des elements de l'ensemble (si c'est possible) separees par des virgules: on enumere les elements de l'ensemble.
- une formule indiquant qu'on considere les elements d'un autre ensemble (disons F) qui verifient une certaine propriete P codee par une formule logique:
 - $\{0, 1, 2, 3\} = \{m \in \mathbb{N}, m \leq 3\}$.
 - $\mathbb{N} = \mathbb{Z}_{\geq 0} = \{m \in \mathbb{Z}, m \geq 0\}$.
 - $\mathcal{P} = \text{Ensemble des nombres premiers} = \{p \in \mathbb{N}, d|p \implies d = 1 \text{ ou } p\}$.
 - Soit E-EPFL l'ensemble des etudiants de l'EPFL.

$$A := \{e \in \text{E-EPFL}, 3|\text{SCIPER}(e)\},$$

$$B := \{e \in \text{E-EPFL}, 3|\text{SCIPER}(e) - 1\},$$

$$C := \{e \in \text{E-EPFL}, 3|\text{SCIPER}(e) - 2\}.$$

REMARQUE 1.1.4. (Paradoxe de Russell) *L'ensemble ENS de tous les ensembles* n'est PAS un ensemble: en effet si c'etait le cas, on pourrait considerer, suivant Russell, l'ensemble de tous les ensembles *n'appartenant pas a eux-meme*

$$\text{Ncont} = \{E \text{ ensemble}, E \notin E\}$$

et se poser la question de savoir si

$$\text{Ncont} \in \text{Ncont} \text{ ou bien } \text{Ncont} \notin \text{Ncont}.$$

Si on est dans le premier cas, on a $\text{Ncont} \in \text{Ncont}$ ce qui par definition de Ncont implique que $\text{Ncont} \notin \text{Ncont}$. Contradiction.

Si on est dans le second cas, on a $N_{\text{cont}} \notin N_{\text{cont}}$ ce qui par definition de N_{cont} implique que $N_{\text{cont}} \in N_{\text{cont}}$. Contradiction!

Ce probleme qui etait present dans les versions initiales de la theorie des ensembles (theories dites "naives") a ete resolu dans la theorie ZF ou ZFC par l'ajout d'axiomes convenables. Par ailleurs pour donner un sens a la notion "d'ensemble de tous les ensembles" (qui n'est PAS un ensemble), on a introduit des concepts plus "souples" appellees *categories* qui sont exemptes de paradoxe de type Russell; ainsi "l'ensemble" de tous les ensembles ENS forme ce qu'on appelle une categorie.

1.2. Operations sur les ensembles

1.2.1. Union, Intersection. Soient $A, B \subset E$ des sous-ensembles d'un ensemble, on a les operations suivantes

- la reunion de A et B ,

$$A \cup B = \{e \in E | e \in A \text{ ou } e \in B\}.$$

- l'intersection de A et B ,

$$A \cap B = \{e \in E | e \in A \text{ et } e \in B\}.$$

- la difference de A et B ,

$$A - B = A \setminus B = \{a \in A | a \notin B\}.$$

En particulier la difference

$$E - A = \{e \in E, e \notin A\} := A^c$$

s'appelle le complementaire de A dans E .

- la difference symetrique de A et B ,

$$A \Delta B = A \setminus B \cup B \setminus A.$$

- Si $A \cap B = \emptyset$, on dit que A et B sont *disjoints*.

Plus generalement si on dispose de $n \geq 2$ sous-ensembles $E_1, \dots, E_n \subset E$ on note

$$\bigcup_{i=1}^n E_i = E_1 \cup \dots \cup E_n = E_1 \cup (E_2 \cup \dots \cup E_n) = \{e \in E | \text{il existe } i \leq n, e \in E_i\},$$

$$\bigcap_{i=1}^n E_i = E_1 \cap \dots \cap E_n = E_1 \cap (E_2 \cap \dots \cap E_n) = \{e \in E | \text{pour tout } i \leq n, e \in E_i\}.$$

Plus generalement si I est un ensemble et $(E_i)_{i \in I}$ est une famille de sous-ensembles de E indexes par I on definit

$$\bigcup_{i \in I} E_i = \{e \in E | \exists i \in I, e \in E_i\},$$

$$\bigcap_{i \in I} E_i = \{e \in E | \forall i \in I, e \in E_i\}.$$

EXERCICE 1.1. Montrer que

$$A \Delta B = A \cup B - A \cap B.$$

1.2.2. Produit cartésien.

DÉFINITION 1.1. *Etant donne deux ensembles A, B et $a \in A, b \in B$ des elements de A et B respectivement. On definit la paire ordonnee (a, b) comme etant l'ensemble*

$$(a, b) := \{a, \{a, b\}\}$$

obtenu a partir de l'axiome de la paire.

REMARQUE 1.2.1. Notons que si $a \neq b$ alors la paire ordonnee $(a, b) = \{a, \{a, b\}\}$ est distincte de la paire ordonnee $(b, a) = \{b, \{b, a\}\} = \{b, \{a, b\}\}$.

DÉFINITION 1.2. *Le produit cartésien $A \times B$ est l'ensemble des paires ordonnees (a, b) avec a un element de A et b un element de B :*

$$A \times B = \{(a, b), a \in A, b \in B\}.$$

REMARQUE 1.2.2. Si un des facteurs est l'ensemble vide, le produit cartésien est vide:

$$\emptyset \times B = A \times \emptyset = \emptyset.$$

REMARQUE 1.2.3. Les ensembles $A \times B$ et $B \times A$ sont distincts sauf si $A = B$ ou si A ou B est l'ensemble vide.

Si $A = B \neq \emptyset$ on ecrit alors

$$A \times A =: A^2$$

On peut iterer cette construction: si on dispose de $n \geq 1$ ensembles $A_1, \dots, A_n$ le produit

$$A_1 \times \dots \times A_n$$

est l'ensemble des n -uples (ordonnes)

$$(a_1, \dots, a_n), a_1 \in A_1, \dots, a_n \in A_n.$$

Si $A_1 = \dots = A_n = A$ on note ce produit A^n .

1.2.2.1. *L'axiome du choix.* On peut chercher a definir le produit cartésien pour un ensemble arbitraire de facteurs: soit I un ensemble et $(A_i)_{i \in I}$ une famille d'ensembles indexee par I ; on veut construire un ensemble note

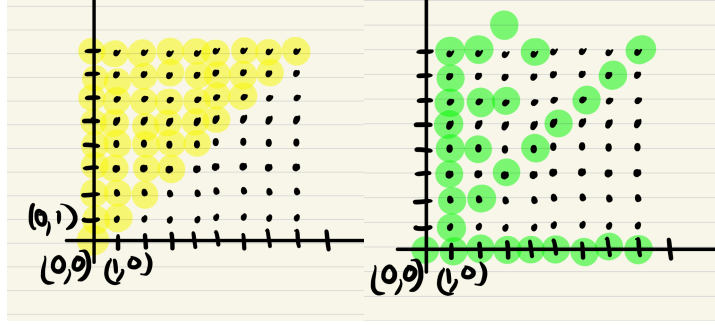
$$\prod_{i \in I} A_i$$

dont les elements sont formes de toutes les familles de la forme

$$(a_i)_{i \in I}, \forall i \in I, a_i \in A_i.$$

Ainsi, exhiber un element de $\prod_{i \in I} A_i$ implique de choisir pour chaque $i \in I$ un element $a_i \in A_i$; cela ne pose pas de probleme si I est fini ou meme si $I = \mathbb{N}$ mais si I est general, des problemes de logique peuvent apparaitre; pouvoir le faire en toute generalite (pour tout ensemble I) implique d'admettre l' *axiome du choix*.

Vous verrez plus tard (notamment en analyse) d'autres formulations et applications de cet axiome.

FIGURE 1. Les relations $\leq$ et $|$ dans $\mathbb{N} \times \mathbb{N}$.

1.2.2.2. *Relation binaire.* Une *relation* (binaire) $\mathcal{R}$ entre (les elements de) deux ensembles A, B est un sous-ensemble

$$\mathcal{R} \subset A \times B.$$

Soient $a \in A$, $b \in B$, on dit que a et b sont *lies par la relation* $\mathcal{R}$ si

$$(a, b) \in \mathcal{R}$$

ce que l'on écrit

$$a \sim_{\mathcal{R}} b \text{ ou bien } a\mathcal{R}b.$$

Si a et b ne sont pas en relation (ie. $(a, b) \notin \mathcal{R}$) on le note

$$a \not\sim_{\mathcal{R}} b \text{ ou bien } a \not\mathcal{R}b.$$

Il se peut que le sous-ensemble $\mathcal{R} \subset A \times B$ ai des proprietes supplementaires qui se traduisent en des proprietes de la relation correspondante.

EXEMPLE 1.2.1. Si $A = B = \mathbb{N}$, on a la relation "inferieur ou egal" $m \leq n$ (par exemple $2 \leq 3$). On a egalement la relation "divise" $m|n$: m divise n si il existe $k \in \mathbb{N}$ tel que $n=m.k$ (ex. $2|8$). Voir la figure 1.2.2.2 pour les representations graphiques de ces relations.

En pratique, le cas le plus important est quand $A = B$. Soit donc une relation $\mathcal{R} \subset A \times A$ de A sur lui-meme. On a les definitions suivantes:

- La relation $\mathcal{R}$ est *reflexive* si

$$\forall a \in A, a\mathcal{R}a$$

(cad $(a, a) \in \mathcal{R}$). En d'autre termes $\Delta A \subset \mathcal{R}$ ou $\Delta A = \{(a, a), a \in A\}$ est appelee la diagonale de $A \times A$. Par exemple pour $\mathbb{N}$, les relations $\leq$ et $|$ sont reflexives.

- La relation $\mathcal{R}$ est *symetrique* si

$$\forall a, a' \in A, a\mathcal{R}a' \iff a'\mathcal{R}a.$$

En d'autre termes la relation $\mathcal{R} \subset A \times A$ est invariante par la symetrie par rapport a la diagonale

$$s_{\Delta} : (a, a') \in A \times A \mapsto (a', a) \in A \times A;$$

c'est a dire

$$s_{\Delta}(\mathcal{R}) = \mathcal{R}.$$

Par exemple sur $\mathbb{N}$, $\leq$ et $|$ ne sont pas symetriques.

- La relation $\mathcal{R}$ est *antisymétrique* si

$$\forall a, a' \in A, a\mathcal{R}a' \text{ et } a'\mathcal{R}a \iff a = a'.$$

Autrement dit la seule possibilité pour que l'on ait à la fois $(a, a') \in \mathcal{R}$ et $(a', a) \in \mathcal{R}$ est que $a = a'$. Par exemple sur $\mathbb{N}$, les relations $\leq$ et $|$ sont antisymétriques.

- La relation $\mathcal{R}$ est *transitive* si

$$\forall a, a', a'' \in A, a\mathcal{R}a' \text{ et } a'\mathcal{R}a'' \implies a\mathcal{R}a''.$$

Par exemple pour $\mathbb{N}$, les relations $\leq$ et $|$ sont transitives.

DÉFINITION 1.3. Une relation $\mathcal{R}$ est dite d'équivalence si elle est réflexive, symétrique et transitive.

Par exemple sur $\mathbb{N}$ la relation "de congruence modulo 3" définie par

$$m \equiv n \pmod{3} \iff 3|m - n$$

est d'équivalence.

Plus généralement pour tout entier $q \neq 0$ la relation "de congruence modulo q " définie par

$$m \equiv n \pmod{q} \iff q|m - n$$

est d'équivalence.

DÉFINITION 1.4. Une relation $\mathcal{R}$ est dite d'ordre si elle est réflexive, antisymétrique et transitive.

Par exemple pour $\mathbb{N}$, les relations $\leq$ et $|$ sont des relations d'ordre.

1.3. Applications entre ensembles

Une autre classe très importante de relation est donnée par les applications entre ensembles.

DÉFINITION 1.5. Soient X et Y des ensembles. Une application (appelée également fonction) f de X (l'espace de départ) vers Y (l'espace d'arrivée) est la donnée pour tout $x \in X$ d'un unique élément $f(x) \in Y$; l'élément $f(x)$ est l'image de x par f . Si $y \in Y$ est de la forme $y = f(x)$ pour un certain $x \in X$ on dit que x est un antécédent de y par f .

Une application est notée

$$f : X \mapsto Y.$$

EXEMPLE 1.3.1. – *Application constante.* Soit $y \in Y$ fixe; l'application qui à tout élément $x \in X$ associe y et l'application constante de valeur y et on la note

$$\underline{y} : x \in X \mapsto y \in Y.$$

– *Application Identité.* Supposons que $Y = X$, l'application identité est celle qui à tout élément $x \in X$ associe x :

$$\text{Id}_X : x \in X \mapsto x \in X.$$

– *Suites:* si $X = \mathbb{N} = \{0, 1, 2, \dots\}$ (ou $\mathbb{N}_{>0} = \{1, 2, \dots\}$) une application de $\mathbb{N}$ vers Y

$$f : n \in \mathbb{N} \mapsto f(n) \in Y$$

s'appelle une *suite* de $\mathbb{N}$ à valeurs dans Y . On note souvent une suite sous la forme

$$(y_n)_{n \geq 0}, \quad y_n = f(n).$$

L'element y_n s'appelle le n -ieme element de la suite.

–*Projection* Soit $A_1, \dots, A_n$ des ensemble et

$$\prod_{i=1}^n A_i$$

leur produit cartésien. Pour $i = 1, \dots, n$ la *projection sur le i -eme facteur* est l'application

$$\pi_i : \begin{array}{ccc} \prod_{i=1}^n A_i & \mapsto & A_i \\ (a_1, \dots, a_n) & \mapsto & a_i \end{array}$$

qui a un n -uple associe la i -eme coordonnee.

1.3.1. Graphe d'une application. On peut donner a la notion d'application une definition purement ensembliste a l'aide du produit cartésien et voir cela en terme de relations. Se donner une application

$$f : X \mapsto Y$$

est equivalent a se donner un sous-ensemble

$$\Gamma \subset X \times Y$$

qu'on appelle un *graphe*:

DÉFINITION 1.6. *Un graphe $\Gamma \subset X \times Y$ est un sous-ensemble de $X \times Y$ tel que pour tout $x \in X$, l'ensemble*

$$\Gamma_x = \{(x, y), y \in Y\} \subset \Gamma$$

(l'ensemble des elements de Γ dont la premiere coordonnee vaut x) possede exactement un element.

REMARQUE 1.3.1. Un graphe Γ definit donc une relation entre X et Y :

$$x \sim_{\Gamma} y \iff (x, y) \in \Gamma.$$

Si $f : X \mapsto Y$ est une application, le graphe associe a f est le sous ensemble

$$\Gamma_f = \{(x, f(x)), x \in X\} \subset X \times Y.$$

Reciproquement si $\Gamma \subset X \times Y$ est un graphe, on lui associe l'application $f_{\Gamma} : X \mapsto Y$ qui a $x \in X$ associe $f(x) := y$ ou y est l'unique element de Y tel que

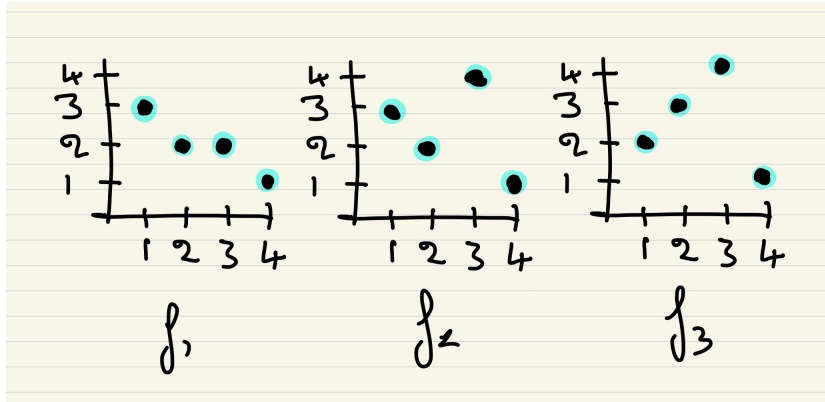
$$(x, y) \in \Gamma.$$

NOTATION 1.1. *On note*

$$\text{Hom}_{ENS}(X, Y) \text{ ou encore } \mathcal{F}(X, Y) \text{ ou encore } Y^X$$

l'ensemble des applications de X vers Y (aussi les fonctions de X a valeurs dans Y).

La realisation ci-dessus des applications entre ensembles en terme de graphes permet de dire que l'ensemble $\text{Hom}_{ENS}(X, Y)$ des applications entre X et Y est un ensemble et plus precisement un sous-ensemble de $\mathcal{P}(X \times Y)$ (on l'identifie avec le sous-ensemble de tous les graphes dans $X \times Y$).

FIGURE 2. Graphes de f_1, f_2, f_3 .

1.3.1.1. *Exemples.* Soit $X = Y = \{1, 2, 3, 4\}$ et posons

$$f_1 : 1 \mapsto 3, 2 \mapsto 2, 3 \mapsto 2, 4 \mapsto 1$$

$$f_2 : 1 \mapsto 3, 2 \mapsto 2, 3 \mapsto 4, 4 \mapsto 1$$

$$f_3 : 1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 4, 4 \mapsto 1.$$

Les graphes de ces applications sont données par les dessins ci-dessus.

– Le graphe de l'application constante $\underline{y} : X \mapsto Y$ est

$$\Gamma(\underline{y}) = \{(x, y), x \in X\} \subset X \times Y.$$

– Quand $X = Y$, le graphe de l'identité Id_X est donné par

$$\Gamma(\text{Id}_X) = \Delta(X) = \{(x, x), x \in X\} \subset X \times X$$

et s'appelle la diagonale de $X \times X$.

1.3.2. Image, preimage.

DÉFINITION 1.7. Soit une application

$$f : X \mapsto Y$$

et $A \subset X$. L'image de A par f est le sous-ensemble de Y

$$f_*(A) = f(A) = \{f(x), x \in A\} \subset Y.$$

On appellera également "image de f ", l'image de l'ensemble de départ X tout entier

$$\text{Im}(f) := f(X).$$

DÉFINITION 1.8. Soit une application

$$f : X \mapsto Y$$

et $B \subset Y$. La preimage de B par f est le sous-ensemble de X

$$f^*(B) = f^{(-1)}(B) := \{x \in X, f(x) \in B\} \subset X.$$

Si $B = \{y\}$ est un singleton

$$f^{(-1)}(\{y\}) = \{x \in X \mid f(x) = y\}$$

est l'ensemble des antécédents de y . On dit quelquefois que la preimage de B est l'ensemble des antécédents des éléments de B par f .

Une application

$$f : X \mapsto Y$$

induit donc naturellement deux applications entre les ensembles des parties de X et Y :

- L'application "image"

$$f(\cdot), f_*, \text{Im}(f) : \mathcal{P}(X) \mapsto \mathcal{P}(Y)$$

qui a un sous-ensemble $A \subset X$ associe son image:

$$f_*(A) = \text{Im}(f)(A) = \{f(x), x \in A\} \subset Y.$$

- L'application "preimage"

$$f^*, f^{(-1)} : \mathcal{P}(Y) \mapsto \mathcal{P}(X)$$

qui a un sous-ensemble $B \subset Y$ associe sa preimage:

$$f^*(B) = f^{(-1)}(B) = \{x \in X, f(x) \in B\} \subset X.$$

REMARQUE 1.3.2. Notons que l'application preimage est toujours defini : si $B \subset Y$ ne possede aucun antecedent dans X alors $f^{(-1)}(B) = \emptyset$.

EXEMPLE 1.3.2. Pour $X = Y = \{1, 2, 3, 4\}$

$$\text{Im}(f_1) = \{1, 2, 3\}, \text{Im}(f_2) = \{1, 2, 3, 4\}, \text{Im}(f_3) = \{1, 2, 3, 4\}$$

$$f_1(\{2, 3\}) = \{2\}, f_2(\{2, 3\}) = \{2, 4\}, f_3(\{2, 3\}) = \{3, 4\}$$

$$f_1^{(-1)}(\{2, 4\}) = \{2, 3\}, f_2^{(-1)}(\{2, 4\}) = \{2, 3\}, f_3^{(-1)}(\{2, 4\}) = \{1, 3\}.$$

EXERCICE 1.2. Montrer que pour $A \subset X$, on a

$$A \subset f^{(-1)}(f(A)).$$

Montrer par un exemple qu'en general on n'a pas l'egalite

$$A = f^{(-1)}(f(A)).$$

Soit $B \subset Y$, existe-t-il des relations d'inclusion entre B et $f(f^{(-1)}(B))$?

1.3.3. Injectivite, surjectivite, application reciproque.

- Une application $f : X \mapsto Y$ est *injective* (f est une injection) si pour tout $y \in Y$, $f^{(-1)}(\{y\})$ (l'ensemble des antecedents de y par f) ne possede pas plus d'un element. On note l'injectivite par

$$f : X \hookrightarrow Y.$$

- Une application $f : X \mapsto Y$ est *surjective* (f est une surjection) si pour tout $y \in Y$, $f^{(-1)}(\{y\})$ (l'ensemble des antecedents de y par f) possede au moins un element. On note la surjectivite par

$$f : X \twoheadrightarrow Y.$$

- Une application $f : X \mapsto Y$ est *bijjective* (f est une bijection) si elle est *injective* et *surjective* : cad si pour tout $y \in Y$, $f^{(-1)}(\{y\})$ (l'ensemble des antecedents de y par f) possede exactement un element. On note la bijectivite par

$$f : X \xrightarrow{\sim} Y \text{ ou } f : X \simeq Y.$$

REMARQUE 1.3.3. Notons qu'une application $f : X \mapsto Y$ est tautologiquement surjective sur son image $\text{Im}(f)$:

$$f : X \twoheadrightarrow \text{Im}(f) \subset Y.$$

En particulier une application injective $f : X \hookrightarrow Y$ définit une bijection

$$f : X \simeq \text{Im}(f).$$

On peut alors identifier les éléments de X à certains éléments de Y via cette dernière bijection (on a "injecté" X dans Y).

NOTATION 1.2. *On note*

$$\text{Inj}(X, Y), \text{Surj}(X, Y), \text{Bij}(X, Y) \subset \text{Hom}_{\text{ENS}}(X, Y)$$

les ensembles d'applications, injectives, surjectives et bijectives de X vers Y .

EXEMPLE 1.3.3. On a:

- (1) f_1 n'est ni injective ($f_1^{-1}(\{2\}) = \{2, 3\}$) ni surjective ($4 \notin \text{Im}(f_1)$). f_2 et f_3 sont bijectives.
- (2) L'application $n \in \mathbb{Z} \mapsto 2n \in \mathbb{Z}$ est injective mais pas surjective.
- (3) L'application $n \in \mathbb{N} \mapsto [n/2] \in \mathbb{N}$ est surjective mais pas injective ($[x]$ désigne la partie entière d'un nombre rationnel x , c'est-à-dire le plus grand entier $\leq x$).
- (4) L'application polynomiale

$$C : (m, n) \mapsto ((m + n)^2 + m + 3n)/2$$

et une bijection entre $\mathbb{N}^2$ et $\mathbb{N}$ (Cantor).

- (5) L'application

$$(m, n) \mapsto m + (n + [(m + 1)/2])^2$$

et une bijection entre $\mathbb{N}^2$ et $\mathbb{N}$.

EXERCICE 1.3. Démontrer (4). Pour cela

- (1) Commencer à vérifier qu'on a bien une application de $\mathbb{N}^2$ vers $\mathbb{N}$.
- (2) Calculer les valeurs $C(m, n)$ pour $(m, n) \leq 5$ et les reporter sur le plan (m, n) .
- (3) Pour montrer l'injectivité et la surjectivité on pourra étudier l'application $(m, n) \mapsto C(m, n)$ quand on la restreint au sous-ensemble

$$D_k = \{(m, n) \in \mathbb{N}^2, m + n = k\}$$

pour $k \geq 0$ un entier et regarder les valeurs que prend cette fonction sur ces ensembles.

Dans le cas des ensembles finis dont on connaît le nombre d'éléments on a les propriétés suivantes liant injectivité, surjectivité, bijectivité au nombre d'éléments, très utiles pour démontrer la bijectivité.

PROPOSITION 1.1. *Soient X et Y des ensembles finis possédant respectivement $|X|$ et $|Y|$ éléments et $f : X \mapsto Y$ une application entre ces ensembles. On a les propriétés suivantes*

- Si $f : X \hookrightarrow Y$ est injective alors $|X| \leq |Y|$.
- Si $f : X \twoheadrightarrow Y$ est surjective alors $|X| \geq |Y|$.
- Si $f : X \hookrightarrow Y$ est injective et $|X| \geq |Y|$ alors $|X| = |Y|$ et f est bijective.
- Si $f : X \twoheadrightarrow Y$ est surjective et $|X| \leq |Y|$ alors $|X| = |Y|$ et f est bijective.

1.3.3.1. *Application reciproque d'une bijection.* Soit $f : X \xrightarrow{\sim} Y$ une bijection, alors pour tout $y \in Y$, $f^{(-1)}(\{y\}) \subset X$ est un ensemble a un seul element

$$f^{(-1)}(\{y\}) = \{x\},$$

a savoir l'unique element x de X tel que $f(x) = y$, ie. l'unique solution de l'equation

$$f(T) = y$$

(dont l'inconnue "T" est a valeur dans X).

On peut donc definir une application (l'application *reciproque* de f)

$$f^{-1} : Y \rightarrow X$$

en posant

$$f^{-1}(y) = x.$$

REMARQUE 1.3.4. On prendra garde que l'application reciproque d'une application bijective $f^{-1} : Y \xrightarrow{\sim} X$ n'existe que si f est bijective alors que l'application *preimage* existe tout le temps.

$$f^{(-1)} : \mathcal{P}(Y) \mapsto \mathcal{P}(X).$$

EXEMPLE 1.3.4. On a

$$\text{Id}_X^{-1} = \text{Id}_X.$$

1.3.3.2. *Involutivite de la reciproque.* On voit que si $f : X \xrightarrow{\sim} Y$ est bijective, sa reciproque $f^{-1} : Y \mapsto X$ est bijective: pour tout $x \in X$, $y \in Y$ on a par definition de la reciproque

$$(1.3.1) \quad f(x) = y \iff x = f^{-1}(y).$$

Ainsi pour tout $x \in X$ il existe bien $y \in Y$ tel que $f^{-1}(y) = x$, c'est $y = f(x)$ et f^{-1} est surjective. Par ailleurs l'ensemble des antecedent de x par f^{-1} est l'ensemble des y tels que $f^{-1}(y) = x$, c'est a dire que $y = f(x)$ et y est unique.

On peut alors se demander quelle est la reciproque de la reciproque: c'est l'application f : on a

$$(f^{-1})^{-1} = f.$$

En effet pour $x \in X$, posons $y := (f^{-1})^{-1}(x)$. On a (appliquant (1.3.1) a f^{-1} au lieu de f puis (1.3.1))

$$(f^{-1})^{-1}(x) = y \iff f^{-1}(y) = x \iff f(x) = y$$

et ainsi pour tout $x \in X$

$$(f^{-1})^{-1}(x) = y = f(x)$$

ce qui est precisement dire que $(f^{-1})^{-1} = f$.

1.3.4. Composition d'applications. Soit X, Y, Z des ensembles et $f : X \mapsto Y$ et $g : Y \mapsto Z$ des applications; a f et g on associe la *composee* de f et g

$$g \circ f : X \mapsto Z$$

est l'application qui va de X a Z en allant, de X a Y via f et de Y a Z via g :

$$\begin{array}{ccc} & Y & \\ f \nearrow & & \searrow g \\ X & \xrightarrow{g \circ f} & Z \end{array}$$

Elle est définie par

$$x \in X \mapsto g \circ f(x) := g(f(x)) \in Z.$$

En d'autres termes on a une application (dite de composition)

$$(1.3.2) \quad \circ : \begin{array}{ccc} \text{Hom}_{ENS}(Y, Z) \times \text{Hom}_{ENS}(X, Y) & \mapsto & \text{Hom}_{ENS}(X, Z) \\ (g, f) & \mapsto & g \circ f \end{array}$$

La composition a les propriétés suivantes:

- Associativité: soient $f : X \mapsto Y$, $g : Y \mapsto Z$, $h : Z \mapsto W$,

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

de sorte que la composée des trois applications s'écrit simplement

$$h \circ g \circ f.$$

- Neutralité de l'identité: soit $f : X \mapsto Y$ alors

$$f \circ \text{Id}_X = f, \text{Id}_Y \circ f = f.$$

- Simplification: soit $f : X \xrightarrow{\sim} Y$ une bijection,

$$f^{-1} \circ f = \text{Id}_X, f \circ f^{-1} = \text{Id}_Y.$$

En particulier

$$\text{Id}_X \circ \text{Id}_X = \text{Id}_X.$$

LEMME 1.1. Soient des applications $f : X \mapsto Y$ et $g : Y \mapsto Z$. Si

- (1) Si f et g sont injectives, $g \circ f$ est injective.
- (2) Si f et g sont surjectives, $g \circ f$ est surjective.
- (3) Si f et g sont bijectives, $g \circ f$ est bijective et

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

Preuve: Pour le (1), il s'agit de montrer que pour tout $z \in Z$, l'image réciproque $(g \circ f)^{-1}(\{z\})$ a au plus un élément. On a

$$(g \circ f)^{-1}(\{z\}) = \{x \in X, g(f(x)) = z\}$$

Si $(g \circ f)^{-1}(\{z\}) = \emptyset$ on a fini. Sinon supposons que $x \in (g \circ f)^{-1}(\{z\})$, on veut montrer que x est unique. Comme g est injective $g^{-1}(\{z\})$ possède au plus un élément et comme

$$z = g \circ f(x) = g(f(x))$$

on voit que $f(x)$ appartient à $g^{-1}(\{z\})$; en particulier $g^{-1}(\{z\})$ est non-vide et s'écrit

$$g^{-1}(\{z\}) = \{y\}$$

pour un certain $y \in Y$ (qui ne dépend que de z); on a donc $f(x) = y$ et donc $x \in f^{-1}(\{y\})$. Comme f est injective, $f^{-1}(\{y\})$ possède au plus un élément et x est celui-ci donc x est l'unique élément de $f^{-1}(\{y\})$ ou y est l'unique élément de $g^{-1}(\{z\})$ et x est donc unique.

Pour (2): comme f est surjective on a $f(X) = Y$ et comme g est surjective on a $g(Y) = Z$ donc

$$g \circ f(X) = g(f(X)) = g(Y) = Z$$

et donc $g \circ f$ est surjective.

Pour (3), $g \circ f$ est injective et surjective par le point (1) et (2) (car f et g le sont) et est donc bijective.

Pour montrer que $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$ (on parle cette fois-ci de reciproques d'applications bijectives) il s'agit de montrer que pour tout $z \in Z$ on a

$$x := (g \circ f)^{-1}(z) = f^{-1} \circ g^{-1}(z) = f^{-1}(g^{-1}(z)) =: x'.$$

Posons $x := (g \circ f)^{-1}(z)$ et $x' := f^{-1}(g^{-1}(z))$. On a

$$g \circ f(x) = z$$

(par definition de la reciproque $(g \circ f)^{-1}$) et on a

$$g \circ f(x') = g(f(f^{-1}(g^{-1}(z))))$$

mais

$$g(f(f^{-1}(g^{-1}(z)))) = g(g^{-1}(z)) = z$$

(car pour tout $u \in X$, $f^{-1}(f(u)) = u$ et $g(g^{-1}(z)) = z$) et donc

$$g \circ f(x') = z = g \circ f(x)$$

et comme $g \circ f$ est injective cela implique que $x' = x$ (car ce sont deux antecedents de z par $g \circ f$). $\square$

En particulier ce lemme dit que l'application de composition 1.3.2 se restreint aux applications bijectives:

$$(1.3.3) \quad \begin{array}{ccc} \text{Bij}(Y, Z) \times \text{Bij}(X, Y) & \mapsto & \text{Bij}(X, Z) \\ (g, f) & \mapsto & g \circ f \end{array}.$$

EXERCICE 1.4. Soient des applications $f : X \mapsto Y$ et $g : Y \mapsto Z$. Montrer que

- (1) Si $g \circ f$ est injective alors f est injective.
- (2) Si $g \circ f$ est surjective alors g est surjective.

Montrer par des exemples que dans le premier cas g n'est pas forcément injective et que dans le second cas f n'est pas forcément surjective.

On suppose que $g \circ f$ est bijective, que peut on dire (ou ne pas dire) de f et de g ?

EXERCICE 1.5. Soit $f : X \mapsto Y$ une application.

- On suppose qu'il existe $g : Y \mapsto X$ telle que $g \circ f = \text{Id}_X$ et $f \circ g = \text{Id}_Y$. Montrer qu'alors f est bijective et que g est sa reciproque.
- Montrer que ce n'est pas forcément vrai si on a seulement que $g \circ f = \text{Id}_X$.

1.4. Cardinal d'un ensemble

DÉFINITION 1.9. Soient X et Y deux ensembles. Si il existe une bijection $f : X \xrightarrow{\sim} Y$, on dit que X et Y ont le meme cardinal et on le note

$$|X| = |Y|.$$

PROPOSITION 1.2. La relation "avoir le meme cardinal" a la proprietes suivantes

- (1) Reflexivite: $|X| = |X|$
- (2) Symetrie: $|X| = |Y| \implies |Y| = |X|$,
- (3) Transitivite: $|X| = |Y|$ et $|Y| = |Z| \implies |X| = |Z|$.

Preuve: Pour la reflexivite, il suffit de prendre Id_X . Pour la Symetrie, si $f : X \simeq Y$ est une bijection, sa reciproque $f^{-1} : Y \simeq X$ est une bijection. Pour la Transitivite, si $f : X \simeq Y$ et $g : Y \simeq Z$ sont des bijections alors $g \circ f : X \mapsto Z$ est encore une bijection. $\square$

DÉFINITION 1.10. *Un ensemble X est fini si il est soit vide, soit en bijection avec un ensemble de la forme $\{1, \dots, n\}$ pour $n \in \mathbb{N}$ un entier ≥ 1 . On écrit alors*

$$|\emptyset| = 0, |X| = n.$$

Un ensemble est infini sinon.

DÉFINITION 1.11. *Un ensemble X est denombrable si il est fini ou a meme cardinal que $\mathbb{N}$. Un ensemble est indenombrable sinon.*

EXEMPLE 1.4.1. (1) Pour tout ensemble X , $|\mathcal{P}(X)| = |\{0, 1\}^X|$: en effet a un sous ensemble $A \subset X$ on associe sa fonction caracteristique

$$1_A : x \in X \rightarrow \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases}$$

et on montre que l'application

$$A \in \mathcal{P}(X) \mapsto 1_A \in \{0, 1\}^X$$

est une bijection.

- (2) Si $|X| = n \in \mathbb{N}$, $|\mathcal{P}(X)| = 2^n$.
- (3) $|\mathbb{Z}|$ est denombrable.
- (4) $\mathbb{Q}$ est denombrable.
- (5) $|X| = |Y| = |\mathbb{N}| \implies |X| \times |Y| = |\mathbb{N}|$.
- (6) (Cantor) Si X est denombrable et infini alors $\mathcal{P}(X)$ n'est pas denombrable.
- (7) $\mathbb{R}$ nest pas denombrable (c'est un corollaire du point precedent).

On va demontrer (6) qui est du a G. Cantor.

Preuve: Si X denombrable infini alors on a une identification $X \xrightarrow{\sim} \mathbb{N}$ et donc

$$\mathcal{P}(X) \xrightarrow{\sim} \mathcal{P}(\mathbb{N}) \xrightarrow{\sim} \{0, 1\}^{\mathbb{N}}.$$

Il suffit donc de montrer que ce dernier ensemble n'est pas denombrable.

URemarquons d'abord qu'une application $f : n \in \mathbb{N} \mapsto f(n) \in \{0, 1\}$ est simplement une *suite* a valeurs dans $\{0, 1\}$.

Supposons qu'il existe une bijection

$$f_{\bullet} : n \in \mathbb{N} \xrightarrow{\sim} f_n(\bullet) \in \{0, 1\}^{\mathbb{N}}.$$

Ainsi, a tout entier n on associe la suite a valeurs dans $\{0, 1\}$,

$$f_n = (f_n(m))_{m \geq 0}$$

et par hypothese, toute suite $f = (f(m))_{m \geq 0} \in \{0, 1\}^{\mathbb{N}}$ est de la forme f_n pour un certain n (unique).

Considerons la suite (dite de Cantor) $f_C \in \{0, 1\}^{\mathbb{N}}$ definie par

$$f_C(n) = \begin{cases} 0 & \text{si } f_n(n) = 1 \\ 1 & \text{si } f_n(n) = 0. \end{cases}$$

Cette suite vaut donc 0 si le n -ieme terme $f_n(n)$ de la n -ieme suite $(f_n(m))_{m \geq 0}$ vaut 1 et 1 si ce terme vaut 0.

Considerons la suite de Cantor f_C : il existe $n_0 \in \mathbb{N}$ telle que

$$f_C = f_{n_0}.$$

Quelle est la valeur de

$$f_C(n_0) = f_{n_0}(n_0)?$$

Il y a deux possibilités 0 ou 1:

- Si $f_C(n_0) = 0$ alors $f_{n_0}(n_0) = 1$ par définition de f_C mais alors $0 = f_C(n_0) = f_{n_0}(n_0) = 1$, une contradiction.
- Si $f_C(n_0) = 1$ alors $f_{n_0}(n_0) = 0$ par définition de f_C mais alors $1 = f_C(n_0) = f_{n_0}(n_0) = 0$, une autre contradiction!

Ainsi la bijection $f_\bullet$ n'existe pas et $\{0, 1\}^{\mathbb{N}}$ n'est pas dénombrable. Cet argument s'appelle l'argument de *la diagonale de Cantor* (il vous rappellera certainement l'argument qui soutient le paradoxe de Russell). $\square$

EXERCICE 1.6. Deducire (7) de (6) (utiliser le développement binaire d'un nombre réel dans $[0, 1[$ mais faire attention que par convention un développement binaire ne se termine pas par une suite constante de 1. Heureusement l'ensemble des suites à valeurs dans $\{0, 1\}$ qui sont ultimement constantes égales à 1 est "petit" mais il faudra dire ce qu'on entend par "petit".

1.4.1. Le Theoreme de Cantor-Bernstein-Schroeder. On peut raffiner la notion d'égalité des cardinaux:

DÉFINITION 1.12. Soient X et Y deux ensembles. Si il existe une application injective entre X et Y , $\phi : X \hookrightarrow Y$, on dit que le cardinal de X est plus petit que celui de Y et on note cette relation $|X| \leq |Y|$. Si de plus $|X| \neq |Y|$, on le note $|X| < |Y|$.

Bien évidemment si les ensembles sont finis cette définition correspond à la notion habituelle de cardinal comme étant le nombre d'éléments.

EXERCICE 1.7. Montrer la transitivité de cette relation:

$$|X| \leq |Y| \text{ et } |Y| \leq |Z| \implies |X| \leq |Z|.$$

En pensant au cas des ensembles finis il est très tentant de penser que cette relation est antisymétrique

$$|X| \leq |Y| \text{ et } |Y| \leq |X| \implies |X| = |Y|.$$

Eh bien c'est vrai et c'est le théorème suivant dont la preuve est donnée en exercice du cours "Structures Algébriques":

THÉORÈME (Cantor-Bernstein-Schroeder). Soit X et Y deux ensembles (pas nécessairement finis). Si il existe une injection $\phi : X \hookrightarrow Y$ et une injection $\psi : Y \hookrightarrow X$ alors il existe une bijection $\varphi : X \simeq Y$. En d'autres termes

$$|X| \leq |Y| \text{ et } |Y| \leq |X| \iff |X| = |Y|.$$

1.4.2. Hypothèse du continu. L'hypothèse du continu est une question posée par G. Cantor:

QUESTION (Cantor). On sait que $|\mathbb{N}| < |\mathbb{R}|$. Existe-il un ensemble $\aleph_1$ tel que

$$|\mathbb{N}| < |\aleph_1| < |\mathbb{R}|.$$

L'hypothèse du continu est que cet ensemble n'existe pas: en d'autres termes le plus "petit" ensemble non-dénombrable est $\mathbb{R}$.

Cette question a motivé en grande partie le développement de la logique mathématique pendant le 20ème siècle.

En 1938, K. Goedel a demontre que l'hypothese du continu ne pouvait etre refutee dans la theorie ZFC: on ne peut pas montrer qu'elle est fausse.

En 1963, P. Cohen a demontre que l'hypothese du continu ne pouvait etre demontree dans la theorie ZFC: on ne peut pas montrer qu'elle est vraie.

En fait l'hypothese du continu est *indecidable* dans la theorie ZFC et une grande partie de la logique mathematique actuelle consiste a trouver un ou des axiomes supplementaires, "naturels" et "minimaux" pour rendre l'hypothese decidable.

CHAPITRE 2

Groupes

"The introduction of the digit 0 or the group concept was general nonsense too, and mathematics was more or less stagnating for thousands of years because nobody was around to take such childish steps..."

2.1. Groupes abstraits

DÉFINITION 2.1. Un groupe $(G, \star, e_G, \cdot^{-1})$ est la donnée d'un quadruple forme de

- d'un ensemble G non-vide,
- d'une application (appelee loi de composition interne)

$$\star : \begin{array}{ccc} G \times G & \mapsto & G \\ (g, g') & \mapsto & \star(g, g') =: g \star g' \end{array}$$

- d'un element $e_G \in G$ (appele element neutre),
- d'une application (appele inversion)

$$\bullet^{-1} : \begin{array}{ccc} G & \mapsto & G \\ g & \mapsto & g^{-1} \end{array}$$

ayant les proprietes suivantes:

- Associativite: $\forall g, g', g'' \in G, (g \star g') \star g'' = g \star (g' \star g'')$.
- Neutralite de e_G : $\forall g \in G, g \star e_G = e_G \star g = g$.
- Inversibilite: $\forall g \in G, g^{-1} \star g = g \star g^{-1} = e_G$.

REMARQUE 2.1.1. Par soucis de concision on omettra l'element neutre et l'inversion (voire de la loi de groupe) dans les donnees: notera souvent un groupe par G ou $(G, \star)$.

REMARQUE 2.1.2. La propriete d'associativite est indispensable et par ailleurs extremement utile: si l'on se donne 3 elements

$$g_1, g_2, g_3 \in G$$

dont on veut former le produit (dans cet ordre): pour cela on calcule $g_{12} = g_1 \star g_2$ puis le produit $g_{12} \star g_3 = (g_1 \star g_2) \star g_3$ et l'associativite nous dit qu'au lieu de cela on aurait pu commencer par calculer $g_{23} = g_2 \star g_3$ et faire le produit

$$g_1 \star g_{23} = g_1 \star (g_2 \star g_3)$$

et l'associativite nous dit que cela ne depend pas de la maniere dont on s'y prend:

$$(g_1 \star g_2) \star g_3 = g_1 \star (g_2 \star g_3)$$

et on peut ecrire sans ambiguite ce produit sans parentheses

$$g_1 \star g_2 \star g_3 = g_1 \star (g_2 \star g_3) = (g_1 \star g_2) \star g_3.$$

De meme si on dispose de n elements $g_1, \dots, g_n \in G$, on defini sans ambiguite leur produit

$$g_1 \star \dots \star g_n = \star_{i=1}^n g_i.$$

PROPOSITION 2.1. (*Proprietes de base de la loi de groupe*) Soit G un groupe. On a

(1) *Involutivite de l'inversion:*

$$\forall g, (g^{-1})^{-1} = g, g^{-1} \star g = e_G.$$

(2) *Unicite de l'element neutre:* soit $e'_G \in G$ tel qu'il existe $g \in G$ verifiant $g \star e'_G = g$ alors $e'_G = e_G$. On a la meme conclusion si il existe g' tel que $e'_G \star g' = e'_G$.

(3) *Unicite de l'inverse:* si $g' \in G$ verifie $g \star g' = e_G$ alors $g' = g^{-1}$ et on a donc egalement $g' \star g = e_G$. De meme si $g' \in G$ verifie $g' \star g = e_G$ alors $g' = g^{-1}$ et on a donc egalement $g \star g' = e_G$.

(4) *Inverse d'un produit:* on a

$$(g \star g')^{-1} = g'^{-1} \star g^{-1}.$$

Preuve: (2) Unicite de l'element neutre: dans l'equation

$$g \star e'_G = g$$

on multiplie a gauche par g^{-1} ce qui donne

$$g^{-1} \star g \star e'_G = e_G \star e'_G = e'_G = g^{-1} \star g = e_G.$$

Pour le deuxieme cas, on multiplie a droite par g'^{-1} .

(3) Unicite de l'inverse: en multipliant l'egalite $g \star g' = e_G$ a gauche par g^{-1} et en utilisant l'associativite on a

$$g \star g' = e_G \implies g^{-1} \star g \star g' = g^{-1} \star e_G$$

et $g^{-1} \star g \star g' = g'$ tandis que $g^{-1} \star e_G = g^{-1}$.

On traite de la meme maniere le cas $g' \star g = e_G$.

(1) Involutivite de l'inversion: en particulier, appliquant ce raisonnement a g^{-1} avec $g' = g$, comme $g \star g^{-1} = e_G$ on obtient que $(g^{-1})^{-1} = g$.

(4) Inverse d'un produit:

$$(g'^{-1} \star g^{-1}) \star (g \star g') = g'^{-1} \star (g^{-1} \star g) \star g' = g'^{-1} \star e_G \star g' = g'^{-1} \star g' = e_G$$

et donc (par unicite de l'inverse)

$$(g \star g')^{-1} = g'^{-1} \star g^{-1}.$$

2.1.1. Exemples de groupes.

- **Le groupe additif des entiers relatifs.** L'ensemble $(\mathbb{Z}, +, 0, -\bullet)$ des entiers relatifs $\mathbb{Z}$ muni de l'addition, du zero 0 et de l'oppose $n \mapsto -n$ forme un groupe d'ordre infini.
- En revanche $(\mathbb{Z} - \{0\}, +, 0, -\bullet)$ forme des entiers non-nuls muni des memes structures ne forme pas un groupe (il manque un element neutre et d'ailleurs il n'est pas stable par addition).
- **Le groupe additif des nombres rationels.** L'ensemble $(\mathbb{Q}, +, 0, -\bullet)$ des nombres rationels $\mathbb{Q}$ muni de l'addition, du zero 0 et de l'oppose $n \mapsto -n$ forme un groupe.
- **Le groupe multiplicatif des nombres rationels.** L'ensemble $(\mathbb{Q}^\times, \times, 1, 1/\bullet)$ avec $\mathbb{Q}^\times = \mathbb{Q} - \{0\}$ est l'ensemble des nombres rationels non-nuls muni de la multiplication, de l'unité 1 et de l'inversion $\lambda \mapsto 1/\lambda$ forme un groupe.
- **Le groupe multiplicatif des entiers relatifs.** De meme le sous-ensemble $\mathbb{Z}^\times := \{\pm 1\}$ muni des memes structures est un groupe.
- **Groupe produit.** soient $(G, \star)$ et $(H, *)$ deux groupes. Le groupe produit $(G \times H, \boxtimes)$ est le groupe associe au produit cartésien

$$G \times H = \{(g, h), g \in G, h \in H\}$$

muni de la loi de composition interne $\boxtimes$ definie par

$$(g, h) \boxtimes (g', h') := (g \star g', h * h').$$

On peut le munir d'un element neutre et d'une inversion pour en faire un groupe (exercice).

- **Groupe trivial.** Soit $G = \{e_G\}$ un ensemble reduit a un seul element. Alors $G \times G$ possede un seul element $((e_G, e_G))$ et la seule application possible de $G \times G$ vers G est donnee par

$$\star : (e_G, e_G) \in G \times G \mapsto e_G \in G;$$

de meme la seule application possible de G vers G est

$$\bullet^{-1} : e_G \in G \mapsto e_G \in G;$$

on verifie facilement que $(G = \{e_G\}, \star, e_G, \bullet^{-1})$ est un groupe appele le groupe trivial.

- Groupe des classes de congruences: Soit $q \in \mathbb{N} - \{0\}$ un entier non-nul. Pour $a \in \mathbb{Z}$, on definit le sous-ensemble de $\mathbb{Z}$

$$a \pmod{q} := \{a + qk, k \in \mathbb{Z}\} \in \mathcal{P}(\mathbb{Z})$$

et qu'on appelle la classe de congruence de a modulo q . L'ensemble de ces sous-ensembles est note

$$\mathbb{Z}/q\mathbb{Z} = \{a \pmod{q}, a \in \mathbb{Z}\} \subset \mathcal{P}(\mathbb{Z});$$

cet ensemble est fini de cardinal q . En effet on montre en utilisant la division euclidienne par q que

$$\mathbb{Z}/q\mathbb{Z} = \{a \pmod{q}, a \in \{0, 1, \dots, q-1\}\}$$

D'autre part, pour $A, B \in \mathcal{P}(\mathbb{Z})$ des sous-ensembles de $\mathbb{Z}$, on a pose

$$A \boxplus B := \{a + b, a \in A, b \in B\} \in \mathcal{P}(\mathbb{Z}),$$

et definit egalement

$$\boxminus A := \{-a, a \in A\} \in \mathcal{P}(\mathbb{Z}).$$

Alors $(\mathbb{Z}/q\mathbb{Z}, \boxplus, 0 \pmod{q}, \boxminus)$ est un groupe commutatif appelle groupe additif des classes des congruences modulo q .

2.1.1.1. *Notation exponentielle.* Soit $g \in G$ un element d'un groupe. Pour tout entier $n \geq 1$, on forme le produit de g avec lui-meme n fois et on le note

$$g \star g \star \dots \star g = g^n.$$

On a donc

$$g^{n+1} = g^n \star g = g \star g^n.$$

On pose ensuite

$$(2.1.1) \quad g^0 = e_G$$

et si $n < 0$ est un entier negatif, on pose

$$g^n = (g^{-1})^{-n} = g^{-1} \star \dots \star g^{-1} (-n = |n| \text{ fois}).$$

cela defini g^n pour $n \in \mathbb{Z}$.

On a alors pour tout $m, n \in \mathbb{Z}$

$$(2.1.2) \quad g^{m+n} = g^m \star g^n.$$

On a alors defini une fonction

$$(2.1.3) \quad \begin{array}{ccc} \mathbb{Z} & \mapsto & G \\ \exp_g : n & \mapsto & \exp_g(n) = g^n =: g^{\mathbb{Z}} \end{array}$$

qu'on appelle *exponentielle* de n dans la base g . On dira alors que l'image

$$\text{Im}(\exp_g) = \exp_g(\mathbb{Z}) = \{g^n, n \in \mathbb{Z}\}$$

est l'ensemble des puissances de g .

2.1.2. Groupes commutatifs. Tous les groupes que nous avons vu possèdent une propriété supplémentaire: la *commutativité*

DÉFINITION 2.2. Soit $(G, \star)$ un groupe. Deux éléments g, g' commutent si

$$g \star g' = g' \star g.$$

Un groupe G est abélien (ou commutatif) si toutes les paires d'éléments de G commutent:

$$\forall g, g' \in G, g \star g' = g' \star g.$$

2.1.2.1. *Notation additive.* Si le groupe G est commutatif, sa loi de groupe sera souvent notée (mais pas toujours) par une addition (par exemple $+_G$), l'élément neutre par le signe "0" (par exemple 0_G) et l'inversion par $-\bullet$: $g \mapsto -g$ (par exemple $-_G$).

L'inverse de g , $-g$ sera alors appelé *l'opposé de g* . De plus, on écrira

$$g +_G g', g +_G 0_G = 0_G +_G g = g, g +_G (-g) = 0_G.$$

Enfin la notation exponentielle pour $g +_G \cdots +_G g$ (n fois) sera remplacée par la notation "multiple": pour $n \geq 1$, on posera

$$n.g = g +_G \cdots +_G g \text{ (} n \text{ fois)}, (-n).g = (-Gg) +_G \cdots +_G (-Gg) \text{ (} n \text{ fois)}, 0.g = 0_G,$$

de sorte que (2.1.2) devient

$$\forall m, n \in \mathbb{Z}, (m+n).g = m.g +_G n.g.$$

On dispose alors d'une application (de multiplication par g) de $\mathbb{Z}$ à valeurs dans G :

$$\cdot.g : \begin{array}{ccc} \mathbb{Z} & \mapsto & G \\ n & \mapsto & n.g \end{array}$$

On dira alors que son image

$$\mathbb{Z}.g = \{n.g, n \in \mathbb{Z}\} \subset G$$

est l'ensemble des multiples de g .

2.1.3. Ordre d'un groupe.

DÉFINITION 2.3. Soit $(G, \star, e_G, \bullet^{-1})$ un groupe, le cardinal $|G|$ de l'ensemble sous-jacent s'appelle également l'ordre du groupe G .

Ainsi $(\mathbb{Z}, +)$ est un groupe d'ordre infini alors que $(\mathbb{Z}^\times, \times)$ est un groupe d'ordre 2 et que $\mathbb{Z}/q\mathbb{Z}$ est d'ordre q .

2.2. Le cas du groupe symétrique

Soit X un ensemble, on note

$$\text{Bij}(X) = \mathfrak{S}(X) = \text{Aut}_{\text{ENS}}(X) = \text{Bij}(X, X) \subset \text{Hom}_{\text{ENS}}(X, X)$$

l'ensemble des bijections de X vers lui-même.

Si X est fini non-vidé (on peut alors supposer que $X = \{1, \dots, n\}$) pour $n \geq 1$ une telle bijection s'appelle alors une *permutation* de X sur lui-même.

Cet ensemble admet des structures supplémentaires

- (1) $\text{Bij}(X)$ est non-vidé: $\text{Id}_X \in \text{Bij}(X)$,
- (2) $\text{Bij}(X)$ est stable par composition des applications (1.3.2): soient $f : X \xrightarrow{\sim} X$, $g : X \xrightarrow{\sim} X$ des bijections alors l'application composée, $f \circ g : X \rightarrow X$ est encore une bijection (la composée d'applications injectives est injective et la composée d'applications surjectives est surjective). On dispose donc d'une application (de composition):

$$\circ : \begin{array}{ccc} \text{Bij}(X) \times \text{Bij}(X) & \mapsto & \text{Bij}(X) \\ (f, g) & \mapsto & f \circ g \end{array}$$

(3) La composition est associative:

$$\forall f, g, h \in \text{Bij}(X), (f \circ g) \circ h = f \circ (g \circ h) =: f \circ g \circ h.$$

(4) L'identite Id_X a la propriete de *neutralite*:

$$\forall f \in \text{Bij}(X), f \circ \text{Id}_X = \text{Id}_X \circ f = f.$$

(5) L'application reciproque $f \mapsto f^{-1}$ envoie $\text{Bij}(X)$ sur $\text{Bij}(X)$

$$\bullet^{-1} : \begin{array}{ccc} \text{Bij}(X) & \mapsto & \text{Bij}(X) \\ f & \mapsto & f^{-1} \end{array}$$

et elle verifie

$$\forall f \in \text{Bij}(X), f \circ f^{-1} = f^{-1} \circ f = \text{Id}_X.$$

Ces proprietes font de l'ensemble $\text{Bij}(X)$ un *groupe* qu'on appelle le *groupe symetrique de X* .

Ce groupe est la plupart du temps hautement non commutatif:

EXERCICE 2.1. Montrer que si X possede 2 elements ou moins alors $\text{Bij}(X)$ est commutatif. Montrer que si X possede au moins 3 elements, il n'est pas commutatif : pour cela choisir trois elements distincts $x_1, x_2, x_3 \in X$ et trouver des bijections σ, τ qui verifient

$$\forall x \in X - \{x_1, x_2, x_3\}, \sigma(x) = x, \tau(x) = x$$

et telles que $\sigma \circ \tau \neq \tau \circ \sigma$.

2.2.1. Exemple: les permutations d'un ensemble fini. Considerons le cas ou X est un ensemble fini, non-vide de cardinal $n \geq 1$; on peut alors supposer que $X = \{1, \dots, n\}$. On note souvent ce groupe Σ_n ou $\mathfrak{S}_n$.

On rappelle qu'alors $\text{Bij}(X)$ est fini de cardinal

$$|\text{Bij}(X)| = n!$$

avec

$$n! = 1.2. \dots .n, \quad n \geq 1, \quad 0! = 1.$$

Preuve: En effet pour definir une bijection $\sigma : \{1, \dots, n\} \xrightarrow{\sim} \{1, \dots, n\}$. On choisit $\sigma(1)$ parmi n elements, puis $\sigma(2)$ parmi les $n - 1$ element restants,... Le mieux est de demontrer cette egalite une recurrence sur n . $\square$

On peut représenter une permutation par un tableau a deux lignes et n colonnes

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

Ainsi l'identite est ainsi codee par

$$\text{Id}_X = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}.$$

Par exemple, pour $n = 4$

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

est la permutation qui envoie

$$1 \mapsto 3, \quad 2 \mapsto 2, \quad 3 \mapsto 4, \quad 4 \mapsto 1$$

et si on compose σ avec elle-meme on obtient

$$\sigma \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix},$$

qui envoie

$$1 \mapsto 3, 2 \mapsto 2, 3 \mapsto 4, 4 \mapsto 1;$$

iterant une fois de plus, on a

$$\sigma \circ \sigma \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \text{Id}_X.$$

2.2.1.1. *Cycles.* Un autre exemple est la permutation cyclique

$$\sigma_{+1} = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ 2 & 3 & \cdots & n & 1 \end{pmatrix}$$

qui envoie

$$1 \mapsto 2, 2 \mapsto 3, \dots, k \mapsto k+1, \dots, n \mapsto 1.$$

Pour les permutations cycliques telle que celle ci-dessus, une autre notation (plus compacte) est tres utile: pour $1 \leq k \leq n$, on se donne

$$\{a_1, \dots, a_k\} \subset \{1, \dots, n\}$$

des elements *distincts* et on pose

$$(a_1 a_2 \cdots a_k)$$

la permutation qui envoie

$$a_1 \mapsto a_2, a_2 \mapsto a_3, \dots, a_k \mapsto a_1$$

et qui envoie chacun des $n - k$ elements de $\{1, \dots, n\} - \{a_1, \dots, a_k\}$ sur lui meme: la permutation $(a_1 a_2 \cdots a_k)$ est appelee *cycle de longueur k*.

Par exemple

$$\sigma_{+1} = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ 2 & 3 & \cdots & n & 1 \end{pmatrix} = (12 \cdots n)$$

est un cycle de longueur n et

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix} = (134)$$

est un cycle de longueur 3.

Transpositions. Une classe particulierement importante de cycles est celle des cycles de longueur 2, $(a_1 a_2)$, $a_1 \neq a_2$. On les appelle *transpositions*: explicitement $(a_1 a_2)$ echange a_1 et a_2 et envoie tous les autres elements sur eux-meme.

Dans le cours MATH-113 vous démontrerez le Theoreme de decomposition suivant

THÉORÈME 2.1. Soit $\mathfrak{S}_n = \text{Bij}(\{1, \dots, n\})$ le groupe de permutations de n elements alors

- (1) Toute permutation s'écrit comme une composee de cycles,
- (2) tout cycle s'écrit comme compose de transpositions,
- (3) et donc toute permutation s'écrit comme compose de transpositions.

Par exemple

$$\sigma = (134) = (34) \circ (14)$$

et (le demontrer)

$$(12 \cdots n) = (2n) \circ (23) \circ \cdots \circ (k-1, k) \circ \cdots \circ (n-2, n-1) \circ (1n)$$

2.3. Sous-groupes

Avec la notion d'ensemble vient la notion de sous-ensemble. De meme avec la notion de *groupe* vient la notion de *sous-groupe* d'un groupe G : un sous-groupe est un sous-ensemble de G qui herite naturellement des structures additionnelles $\star, e_G, \bullet^{-1}$ venant avec la structure de groupe de l'ensemble G .

DÉFINITION 2.4. Soit $(G, \star, e_G, \bullet^{-1})$ un groupe. Un sous-groupe $H \subset G$ est un sous-ensemble de G tel que

- (1) $e_G \in H$.

(2) H est stable pour la loi de composition interne $\star$:

$$\forall h, h' \in H, h \star h' \in H.$$

(3) H est stable par l'inversion:

$$\forall h \in H, h^{-1} \in H.$$

Alors si on note $\star_H$ et $\bullet_H^{-1}$ les restrictions de la loi de composition $\star$ et de l'inversion $\bullet^{-1}$ aux sous-ensembles $H \times H$ et H on a

$$\star_H : \begin{array}{ccc} H \times H & \mapsto & H \\ (h, h') & \mapsto & h \star h' \end{array} \quad \bullet_H^{-1} : \begin{array}{ccc} H & \mapsto & H \\ h & \mapsto & h^{-1} \end{array}$$

et $(H, \star_H, e_G, \bullet_H^{-1})$ forme un groupe.

REMARQUE 2.3.1. Distinguer les restrictions à H de la loi de composition et de l'inversion est formellement correct mais un peu pédant. La convention universelle est d'omettre cette restriction dans les notations et d'écrire $(H, \star, e_H = e_G, \bullet^{-1})$ ou plus simplement $(, \star)$.

En fait il n'est pas nécessaire de vérifier les trois conditions de la définition d'un sous-groupe.

PROPOSITION 2.2 (Critère de sous-groupe). *Pour montrer qu'un sous-ensemble non-vidé*

$$\emptyset \neq H \subset G$$

est un sous-groupe il suffit de vérifier l'un ou l'autre des groupes de propriétés (1) ou (2) ci-dessous:

- (1) (a) $\forall h, h' \in H, h \star h' \in H,$
 (b) $\forall h \in H, h^{-1} \in H.$
 (2) $\forall h, h' \in H, h \star h'^{-1} \in H.$

Preuve: On va montrer que si (2) est vérifiée alors H est un sous-groupe (le cas (1) est encore plus simple):

- En prenant $h' = h$, on a $h \star h^{-1} = e_G \in H$ donc H contient l'élément neutre.
- En appliquant $h \star h'^{-1} \in H$ avec $h = e_G$ on a que si $h' \in H$ alors $h'^{-1} \in H$.
- En appliquant $h \star h'^{-1} \in H$ avec $h \in H$ et $h'' = h'^{-1}$ et en utilisant que $(h'^{-1})^{-1} = h'$, on a que si $h, h' \in H$ alors $h \star h' \in H$.

□

EXEMPLE 2.3.1. Voici quelques exemples de sous-groupes:

- $\{e_G\} \subset G$ est un sous-groupe: le sous-groupe trivial.
- $G \subset G$ est également un sous-groupe.
- l'ensemble vide $\emptyset \subset G$ n'est pas un sous-groupe (il lui manque l'élément neutre).
- $2\mathbb{Z} \subset \mathbb{Z}$ (l'ensemble des entiers pairs) est un sous-groupe.
- $1 + 2\mathbb{Z} \subset \mathbb{Z}$ (l'ensemble des entiers impairs) n'est pas un sous-groupe.
- On peut classer tous les sous-groupes de $\mathbb{Z}$:

THÉORÈME 2.2. *Les sous-groupes de $\mathbb{Z}$ sont exactement les sous-ensembles de la forme*

$$q\mathbb{Z} = \{qk, k \in \mathbb{Z}\} = 0 \pmod{q} \subset \mathbb{Z}$$

pour $q \in \mathbb{Z}$ un entier.

Preuve: Pour tout entier $q \in \mathbb{Z}$, on vérifie par la définition ou le critère de sous-groupe que l'ensemble des multiples de q

$$q.\mathbb{Z} = \{q.n, n \in \mathbb{Z}\} \subset \mathbb{Z}$$

est un sous-groupe.

Montrons que réciproquement, tout sous-groupe de $\mathbb{Z}$ est de la forme $q.\mathbb{Z}$ pour $q \in \mathbb{Z}$. En effet, soit $H \subset \mathbb{Z}$ un sous-groupe. Si $H = \{0\}$ on a terminé car $H = 0.\mathbb{Z}$. Sinon soit $q \in H - \{0\}$; quitte à

remplacer q par $-q$ (qui est encore dans H car H est un sous-groupe) ops $q > 0$. On peut egalement supposer que q est le plus petit entier > 0 contenu dans H . On va montrer qu'alors $H = q\mathbb{Z}$.

Comme $q \in H$ on a $\mathbb{Z}.q \subset H$

Soit $h \in H$ alors par division euclidienne, h peut s'ecrire

$$h = q.k + r$$

avec $k \in \mathbb{Z}$ et $0 \leq r < q$. Mais comme H est un sous-groupe et que h et $q.k = \pm(q + \dots + q)$ ($|k|$ fois) sont dans H ,

$$r = h - q.k \in H.$$

Comme $0 \leq r < q$ on a necessairement $r = 0$ (par definition de q comme plus petit element positif non-nul de H) et donc $h = q.k \in q\mathbb{Z}$. $\square$

- Pour $g \in G$, l'ensemble des puissance de g

$$\exp_g(\mathbb{Z}) = g^{\mathbb{Z}} = \{g^n, n \in \mathbb{Z}\} \subset G$$

est un sous-groupe commutatif de G .

- Si G est commutatif et que la loi de groupe est notee additivement, l'ensemble des multiples de g ,

$$\mathbb{Z}.g = \{n.g, n \in \mathbb{Z}\} \subset G$$

est un sous-groupe commutatif de G .

- Soit X un ensemble $G = \text{Bij}(X)$ et $x \in X$ un element, alors le sous-ensemble

$$\text{Bij}(X)_x = \{\sigma \in \text{Bij}(X), \sigma(x) = x\}$$

est un sous-groupe: on l'appelle *le stabilisateur* de x dans $\text{Bij}(X)$.

Le resultat suivant qu'on demontrera plus tard nous dit que le cas du groupe symetrique est fondamental (voir Exercice 2.6 pour la preuve) :

THÉORÈME 2.3. *Soit G un groupe alors G s'identifie canoniquement a un sous-groupe du groupe symetrique $\mathfrak{S}_G = \text{Bij}(G)$ des permutations de G .*

2.3.1. Le Theoreme de Lagrange.

THÉORÈME 2.4. *Soit G un groupe fini et $H \subset G$ un sous-groupe alors l'ordre de H divise l'ordre de G :*

$$|H| \mid |G|.$$

Preuve: On consider l'ensemble des sous-ensembles de G la forme

$$T_G(H) = \{g.H \subset G, g \in G\} \subset \mathcal{P}(G)$$

avec

$$g.H = \{g.h, h \in H\}.$$

(l'ensemble des translates a gauche de H par les elements de G). On montre que

- les translates recouvrent G :

$$G = \bigcup_{g \in G} g.H$$

- les translates sont disjoints:

$$g.H \cap g'.H \neq \emptyset \iff g.H = g'.H,$$

- les translates ont tous le meme cardinal:

$$\forall g \in G, |g.H| = |H|.$$

En particulier $T_G(H)$ forme une partition de G : il existe un sous-ensemble $G_H \subset G$ tel que

$$G = \bigsqcup_{g \in G_H} g.H$$

et donc

$$|G| = \sum_{g \in G_H} |g.H| \sum_{g \in G_H} |H| = |G_H| \cdot |H|.$$

□

COROLLAIRE 2.1. *Si $|G|$ est un nombre premier, ses seuls sous-groupes sont $\{e_G\}$ et G .*

2.3.1.1. Ordre d'un element. On a vu precedemment que le cardinal d'un groupe etait aussi appelle son *ordre*. L'ordre d'un element $g \in G$ est definit par

DÉFINITION 2.5. *Soit G un groupe et $g \in G$ un element de G . L'ordre de g est l'ordre du sous-groupe $g^{\mathbb{Z}} \subset G$ (ou $\mathbb{Z}.g$ si la notation est additive). On le note*

$$\text{ord}(g) = |g^{\mathbb{Z}}| \text{ (= } |\mathbb{Z}.g| \text{ en notation additive).}$$

COROLLAIRE 2.2. *Soit G une groupe fini. Pour tout $g \in G$, l'ordre de g divise l'ordre de G :*

$$\text{ord}(g) \mid |G|$$

COROLLAIRE 2.3. *Si $|G|$ est un nombre premier, pour tout $g \neq e_G$ on a*

$$g^{\mathbb{Z}} = G.$$

2.3.2. Groupe engendre par un ensemble.

PROPOSITION 2.3. *(Invariance par intersection) Soit G un groupe et $H_1, H_2 \subset G$ deux sous-groupes alors $H_1 \cap H_2$ est un sous-groupe. Plus generalement soit $H_i, i \in I$, $H_i \subset G$ une collection de sous-groupes de G indexes par I alors*

$$\bigcap_{i \in I} H_i \subset G$$

est un sous-groupe de G .

Preuve: On utilise le critere de sous-groupe: d'abord $\bigcap_{i \in I} H_i$ est non-vidé car il contient l'element neutre e_G . Soient $h, h' \in \bigcap_{i \in I} H_i$ montrons que $h \star h'^{-1} \in \bigcap_{i \in I} H_i$. Il s'agit de montrer que pour tout $i \in I$, $h \star h'^{-1} \in H_i$ mais c'est vrai car H_i est un sous-groupe de G . □

DÉFINITION 2.6. *Soit*

$$\mathcal{G}_A = \{H \subset G \text{ sous-groupe} \mid A \subset H\}$$

l'ensemble de tous les sous-groupes de G contenant A (cet ensemble est non-vidé car G est dedans). Alors l'intersection de ses sous-groupes

$$\bigcap_{H \in \mathcal{G}_A} H \subset G$$

est un sous-groupe contenant A et c'est le plus petit (si H est un sous-groupe contenant A alors $\langle A \rangle \subset H$.) Ce sous-groupe

$$\langle A \rangle := \bigcap_{H \in \mathcal{G}_A} H$$

s'appelle le sous-groupe engendre par A .

Si $\langle A \rangle = G$ on dit que G est engendre par A (ou que A est un systeme de generateurs de G).

Voici une caracterisation plus constructive de $\langle A \rangle$ (qui justifie la terminologie):

THÉORÈME 2.5 (Caractérisation linguistique du groupe engendré par un ensemble). Soit $A \subset G$ un ensemble, si $A = \emptyset$ alors $\langle A \rangle = \{e_G\}$, sinon on pose

$$A^{-1} = \{g^{-1}, g \in A\} \subset G$$

l'image de A par l'inversion, alors

$$\langle A \rangle = \{g_1 \star \cdots \star g_n, n \geq 1, g_i \in A \cup A^{-1}\}.$$

En d'autres termes, $\langle A \rangle$ est l'ensemble des éléments de G qu'on peut former en multipliant ensemble des éléments de A et de son inverse A^{-1} de toutes les manières possibles.

Preuve: Si $A = \emptyset$, il est clair que le groupe trivial a les bonnes propriétés. Supposons A non-vide. Il s'agit de montrer que l'ensemble

$$\langle A \rangle' = \{g_1 \star \cdots \star g_n, n \geq 1, g_i \in A \cup A^{-1}\}$$

est un sous-groupe contenant A et qu'il est contenu dans tout sous-groupe $H \supset A$.

Considérant les mots de longueur 1, $g_1, g_1 \in A$ on voit que $A \subset \langle A \rangle'$. Soient

$$g_1 \star \cdots \star g_n, g'_1 \star \cdots \star g'_{n'} \in \langle A \rangle'$$

deux tels mots alors

$$g_1 \star \cdots \star g_n \star (g'_1 \star \cdots \star g'_{n'})^{-1} = g_1 \star \cdots \star g_n \star g'^{-1}_{n'} \star \cdots \star g'^{-1}_1 \in \langle A \rangle'.$$

ainsi $\langle A \rangle'$ est un sous-groupe de G contenant A par conséquent

$$\langle A \rangle \subset \langle A \rangle'.$$

Enfin, si $A \subset H$ est un autre sous-groupe alors $A^{-1} \in H$ (car H est stable par inversion) et pour tout $n \geq 1$ et tout $g_1, \dots, g_n \in A \cup A^{-1} \subset H$ on a $g_1 \star \cdots \star g_n \in H$ car H est stable par $\star$ et donc $\langle A \rangle' \subset H$ et donc

$$\langle A \rangle' \subset \bigcap_{H \in \mathcal{G}_A} H = \langle A \rangle \subset \langle A \rangle'.$$

□

2.3.2.1. *Groupe monogène/cyclique.* Soit $g \in G$ alors le sous-groupe engendré par g , $\langle \{g\} \rangle$ vaut

$$\langle \{g\} \rangle = g^{\mathbb{Z}} = \exp_g(\mathbb{Z}).$$

DÉFINITION 2.7. Un groupe G est dit

- monogène si il est engendré par un seul élément:

$$\exists g \in G, G = \langle \{g\} \rangle = g^{\mathbb{Z}}.$$

On dit que g est un générateur de G .

- cyclique si il est fini et monogène.

EXEMPLE 2.3.2.

- Le groupe $\mathbb{Z}$ est monogène : engendré par 1 ou -1 .
- Le groupe $\mathbb{Z}/q\mathbb{Z}$ est cyclique: il est engendré par $1 \pmod{q}$ et plus généralement par $a \pmod{q}$ pour tout a premier avec q .

2.4. Morphismes de groupes

Les sous-groupes d'un groupe sont les sous-ensembles qui préservent la structure de groupe; les *morphismes* de groupes sont les applications entre deux groupes qui préservent les structures respectives de groupes.

DÉFINITION 2.8. Soient $(G, \star)$ et $(H, *)$ deux groupes, un morphisme de groupes $\varphi : G \mapsto H$ est une application telle que

$$\forall g, g' \in G, \varphi(g \star g') = \varphi(g) * \varphi(g').$$

On notera

$$\text{Hom}_{Gr}(G, H)$$

l'ensemble des morphismes de G vers H .

THÉORÈME 2.6 (Propriété fonctionnelle d'un morphisme). Soit $\varphi : G \mapsto H$ un morphisme de groupes alors

- (1) $\varphi(e_G) = e_H$,
- (2) $\forall g \in G, \varphi(g^{-1}) = \varphi(g)^{-1}$,
- (3) $\forall g, g' \in G, \varphi(g \star g') = \varphi(g) * \varphi(g')$.

Preuve: La troisième identité est juste une répétition de la définition.

Pour la première identité, on a

$$\varphi(g) = \varphi(g \star e_G) = \varphi(g) * \varphi(e_G)$$

et donc $\varphi(e_G) = e_H$ par unicité de l'élément neutre dans H .

Pour la deuxième on a pour tout $g \in G$

$$\varphi(g \star g^{-1}) = \varphi(e_G) = e_H = \varphi(g) * \varphi(g^{-1})$$

et donc $\varphi(g^{-1}) = \varphi(g)^{-1}$ par unicité de l'inverse dans H . □

EXEMPLE 2.4.1. Les applications suivantes sont des morphismes de groupes

- Soit G un groupe (note multiplicativement) et $g \in G$. Montrer que l'application

$$g^\bullet = \exp_g : n \in \mathbb{Z} \mapsto g^n \in G$$

est un morphisme de groupe.

- En particulier pour

$$q \in \mathbb{Z}, [\times q] : \begin{array}{ccc} \mathbb{Z} & \mapsto & \mathbb{Z} \\ n & \mapsto & qn \end{array}$$

est un morphisme de groupes.

- Les fonctions exponentielles et logarithme sont des morphismes de groupes:

$$\exp : \begin{array}{ccc} (\mathbb{R}, +) & \mapsto & (\mathbb{R}_{>0}, \times) \\ x & \mapsto & \exp(x) \end{array}, \log : \begin{array}{ccc} (\mathbb{R}_{>0}, \times) & \mapsto & (\mathbb{R}, +) \\ x & \mapsto & \log(x) \end{array}.$$

- Soit $q \geq 1$ et

$$\bullet(\text{mod } q) : \begin{array}{ccc} \mathbb{Z} & \mapsto & \mathbb{Z}/q\mathbb{Z} \\ a & \mapsto & a(\text{mod } q) \end{array}$$

l'application qui à un entier a associe sa classe de congruence modulo q alors $\bullet(\text{mod } q)$ est un morphisme de $(\mathbb{Z}, +)$ vers $(\mathbb{Z}/q\mathbb{Z}, \boxplus)$.

2.4.1. Noyau, Image. Les morphismes préservent la structure de sous-groupe:

PROPOSITION 2.4. (Invariance des sous-groupes par morphismes) Soit $\varphi \in \text{Hom}_{Gr}(G, H)$ un morphisme de groupes.

- (1) Soit $K \subset G$ un sous-groupe alors $\varphi(K) \subset H$ est un sous-groupe. En particulier l'image de φ ,

$$\text{Im}(\varphi) = \varphi(G) \subset H$$

est un sous-groupe de H .

- (2) Soit $L \subset H$ un sous-groupe de H , alors la préimage

$$\varphi^{(-1)}(L) = \{g \in G, \varphi(g) \in L\} \subset G$$

est un sous-groupe de G . En particulier $\varphi^{(-1)}(\{e_H\})$ est un sous-groupe de G .

Preuve: Soit $h, h' \in \varphi(K)$, on veut montrer que $h * h'^{-1} \in \varphi(K)$. Par definition il existe $k, k' \in K$ tels que $\varphi(k) = h, \varphi(k') = h'$ et

$$h * h'^{-1} = \varphi(k) * \varphi(k')^{-1} = \varphi(k * k'^{-1}) \in \varphi(K)$$

car $k * k'^{-1} \in K$ puisque K est un sous-groupe.

Soit $g, g' \in \varphi^{-1}(L)$ alors montrons que $\varphi(g * g'^{-1}) \in L$. On a

$$\varphi(g * g'^{-1}) = \varphi(g) * \varphi(g')^{-1} \in L$$

car $\varphi(g), \varphi(g') \in L$ par definition et L est un sous-groupe. $\square$

DÉFINITION 2.9. *Le sous-groupe $\varphi^{(-1)}(\{e_H\})$ s'appelle le noyau de φ et est noté*

$$\ker(\varphi) = \varphi^{(-1)}(\{e_H\}) = \{g \in G, \varphi(g) = e_H\}.$$

L'importance du noyau vient du fait qu'il permet de tester facilement si un morphisme est injectif.

THÉORÈME 2.7 (Critere d'injectivite). *Soit $\varphi \in \text{Hom}_{Gr}(G, H)$ un morphisme de groupes alors les proprietes suivantes sont equivalentes*

- (1) φ est injectif,
- (2) $\ker(\varphi) = \{e_G\}$.

Preuve: Supposons φ injectif alors $\ker(\varphi) = \{g \in G, \varphi(g) = e_H\}$ possede au plus un element. Mais comme $\varphi(e_G) = e_H$ on a $\ker(\varphi) = \{e_G\}$.

Supposons que $\ker(\varphi) = \{e_G\}$; on veut montrer que pour tout $h \in H$,

$$\varphi^{(-1)}(\{h\}) = \{g \in G, \varphi(g) = h\}$$

possede au plus un element. Soient $g, g' \in \varphi^{(-1)}(\{h\})$ (si l'ensemble est vide on a fini) alors

$$\varphi(g) = \varphi(g') = h$$

et

$$\varphi(g) * \varphi(g')^{-1} = h * h^{-1} = e_H$$

mais

$$e_H = \varphi(g) * \varphi(g')^{-1} = \varphi(g * g'^{-1})$$

donc $g * g'^{-1} \in \ker(\varphi) = \{e_G\}$ et

$$g * g'^{-1} = e_G \implies g = g'$$

et donc $\varphi^{(-1)}(\{h\})$ possede au plus un element. $\square$

2.4.1.1. *Propriete d'invariance du Noyau.*

THÉORÈME 2.8. *Soit $\varphi : G \mapsto H$ un morphisme de groupes et $\ker(\varphi) \subset G$ son noyau. Alors pour tout $g \in G$ on a l'egalite suivante entre ensembles*

$$g. \ker(\varphi).g^{-1} = \{g.k.g^{-1}, k \in \ker(\varphi)\} = \ker(\varphi).$$

Preuve: Montrons que pour tout g on a

$$g. \ker(\varphi).g^{-1} \subset \ker(\varphi).$$

Il s'agit de montrer que pour $k \in \ker(\varphi)$ on a $g.k.g^{-1} \in \ker(\varphi)$ c'est a dire $\varphi(g.k.g^{-1}) = e_H$

$$\varphi(g.k.g^{-1}) = \varphi(g) * \varphi(k) * \varphi(g^{-1}) = \varphi(g) * e_H * \varphi(g)^{-1} = \varphi(g) * \varphi(g)^{-1} = e_H.$$

Montrons l'inclusion reciproque: comme $g. \ker(\varphi).g^{-1} \subset \ker(\varphi)$, en multipliant cette inclusion a gauche par g^{-1} et a droite par g on a

$$g^{-1}g. \ker(\varphi).g^{-1}.g \subset g^{-1} \ker(\varphi)g$$

et comme

$$g^{-1}g. \ker(\varphi).g^{-1}.g = e_g. \ker(\varphi).e_G = K$$

on a pour tout $g \in G$

$$\ker(\varphi) \subset g^{-1} \ker(\varphi) g.$$

En particulier substituant g par g^{-1} on a

$$\ker(\varphi) \subset g. \ker(\varphi). g^{-1}$$

et on a donc

$$g. \ker(\varphi). g^{-1} = \ker(\varphi).$$

□

DÉFINITION 2.10. *Un sous-groupe $K \subset G$ ayant la propriété que pour tout $g \in G$ on a*

$$g.K.g^{-1} = K$$

est dit normal ou distingue et on le note

$$K \triangleleft G.$$

REMARQUE 2.4.1. Ainsi un noyau est un sous-groupe distingue. Reciproquement on peut montrer que tout sous-groupe distingue est un noyau mais cela necessite la notion de groupe quotient.

EXERCICE 2.2 (Equations dans les groupes). Soit G, H des groupes et $\varphi : G \mapsto H$ un morphisme. Etant donne $h \in H$, on cherche a resoudre l'equation d'inconnue $g \in G$:

$$Eq(\varphi, h) : \quad \varphi(g) = h.$$

L'ensemble des solutions de cette equation n'est autre que la preimage $\varphi^{(-1)}(\{h\})...$

(1) Montrer que

$$\varphi^{(-1)}(\{h\})$$

est soit vide soit qu'il existe $g_0 \in G$ tel que

$$\varphi^{(-1)}(\{h\}) = g_0 \star \ker(\varphi)$$

ou

$$g_0 \star \ker(\varphi) = \{g_0 \star k, k \in \ker(\varphi)\}.$$

(2) Montrer que

$$\varphi^{(-1)}(\{h\}) = \ker(\varphi) \star g_0$$

avec

$$\ker(\varphi) \star g_0 = \{k \star g_0, k \in \ker(\varphi)\}.$$

(3) Quel est l'ensemble de tous les $g_0 \in G$ ayant cette propriété ? Cela vous rappelle t il quelque chose ? (pensez a "equation avec" et "sans second membre", "solution particuliere", "solution generale" ...)

2.4.2. Exemple: ordre d'un element. Soit $g \in G$ un element d'un groupe. On rappelle que l'ordre de g est egal a

$$\text{ord}(g) = |g^{\mathbb{Z}}| = |\exp_g(\mathbb{Z})|,$$

le cardinal de l'image du morphisme "puissances de g "

$$\exp_g : n \in \mathbb{Z} \mapsto g^n \in G.$$

Son noyau, $\ker(\exp_g)$ est un sous-groupe de $\mathbb{Z}$ et donc de la forme

$$\ker(\exp_g) = q.\mathbb{Z}$$

avec $q = q(g) \in \mathbb{N}$ (car tous les sous-groupes de $\mathbb{Z}$ sont de cette forme). On a la caracterisation suivante de l'ordre de g :

THÉORÈME 2.9. *Soit G un groupe, $g \in G$ un element et $q \in \mathbb{N}$ un entier naturel tel que*

$$q\mathbb{Z} = \ker(g^\bullet).$$

- Si $q = 0$ alors $\ker(g^\bullet) = \{0\}$ et $g^\bullet$ est injectif et ainsi on a un isomorphisme de groupes (un morphisme de groupes bijectif)

$$\mathbb{Z} \simeq g^\mathbb{Z};$$

On a alors

$$\text{ord}(g) = |\mathbb{Z}| = \infty.$$

- Si $q > 0$, alors q est le plus petit entier strictement positif verifiant

$$g^q = e_G$$

et on a

$$\text{ord}(g) = |g^\mathbb{Z}| = q.$$

- Si G est fini on a $\text{ord}(g) \mid |G|$ et $g^{|G|} = e_G$

Preuve:

EXERCICE 2.3. Démontrer les affirmations precedentes et en particulier que si $q > 0$ alors

$$g^\mathbb{Z} = \{g^0 = e_G, g, \dots, g^{q-1}\}$$

est fini de cardinal q

□

2.4.3. Groupe quotient. On a vu qu'un noyau d'un morphisme $\varphi : G \rightarrow H$ est un sous-groupe distingue de G . On va voir que reciproquement tout sous-groupe distingue $K \triangleleft G$ est le noyau d'un morphisme de groupe. Pour cela on commencera par definir l'image de ce morphisme: le groupe quotient G/K .

DÉFINITION 2.11. Soit $K \subset G$ un sous-groupe d'un groupe. Une classe a gauche (resp. a droite) de G est un sous-ensemble de G de la forme

$$gK = \{g.k, k \in K\},$$

resp.

$$Kg = \{k.g, k \in K\}$$

pour $g \in G$.

- L'ensemble des classes a gauche de G est note

$$G/K := \{gK, g \in G\} \subset \mathcal{P}(G).$$

On l'appelle egalement le quotient a droite de G par K .

- L'ensemble des classes a droite de G est note

$$K \backslash G := \{Kg, g \in G\} \subset \mathcal{P}(G).$$

On l'appelle egalement le quotient a gauche de G par K .

LEMME 2.1. les classes a gauche (resp. a droite) on les proprietes suivantes

- $e_G K = K$.
- $gK = g'K \iff g' = gk, k \in K$
- $gK \cap g'K \neq \emptyset \iff gK = g'K$.
- $Ke_G = K$.
- $Kg = Kg' \iff g' = kg, k \in K$
- $Kg \cap Kg' \neq \emptyset \iff Kg = Kg'$.

Si G est fini on a

$$|G/K| = |K \backslash G| = |G|/|K|.$$

Preuve: Exercice □

Supposons maintenant que K est distingue dans G . On a alors

LEMME 2.2. *Si K est distingue dans G on a*

$$\forall g \in G, gK = Kg.$$

Ainsi

$$G/K = K \backslash G.$$

Deplus, pour tout $g, g' \in G$ on a

$$gK.g'K := \{gkg'k', k, k' \in K\} = gg'K.$$

Preuve: Exercice □

On notera une classe a gauche (ou a droite,) de la maniere suivante

$$g \pmod{K} := gK = Kg$$

On definit sur l'ensemble G/K la loi de composition interne

$$\cdot_K : G/K \times G/K \rightarrow G/K$$

en posant

$$g \pmod{K} \cdot_K g' \pmod{K} = gK \cdot_K g'K := gK.g'K = gg'K.$$

THÉORÈME 2.10 (Existence du groupe quotient). *Si K est distingue dans G , l'ensemble $(G/K, \cdot_K)$ a une structure de groupe dont l'element neutre est*

$$e_{G/K} = e_G K = K$$

et l'inversion est donnee par

$$(gK)^{-1} = g^{-1}K.$$

Preuve: Exercice. □

DÉFINITION 2.12. *Si K est distingue dans G le groupe $(G/K, \cdot_K)$ est appele groupe quotient de G par K .*

Quotients et morphismes. Le groupe quotient a la propriete suivante par rapport aux morphismes:

THÉORÈME 2.11. *Soit $K \triangleleft G$ un sous-groupe distingue et G/K le groupe quotient. L'application*

$$\bullet \pmod{K} : g \in G \mapsto g \pmod{K} = gK \in G/K$$

est un morphisme de groupes surjectif de noyau K .

Soit $\varphi : G \rightarrow H$ un morphisme de groupe tel que

$$K \subset \ker(\varphi)$$

alors il existe un unique morphisme de groupe

$$\varphi_K : G/K \rightarrow H$$

tel que

$$\forall g \in G, \varphi_K(gK) = \varphi(g).$$

On a alors

$$\ker(\varphi_K) = \ker \varphi \pmod{K} = (\ker \varphi).K = \{k'K, k' \in \ker \varphi\}.$$

Preuve: Exercice. □

THÉORÈME 2.12 (Theoreme Noyau-Image). *Supposons que K est distingue dans G . Avec les notations precedentes, on a*

$$\ker \varphi = K \iff \varphi_K \text{ est injectif}$$

et on a alors un isomorphisme

$$\varphi_K : G/K \simeq \varphi(G) \subset H.$$

En particulier si G est fini on a

$$|G|/|K| = |\varphi(G)|.$$

Ainsi si φ est surjectif et si $\ker \varphi = K$ on a un isomorphisme

$$G/K \simeq H.$$

2.4.4. Operations entre morphismes de groupes.

Notation/Terminologie. On notera

- $\text{Hom}_{Gr}(G, H)$ l'ensemble des morphismes de groupes de G vers H ,
- $\text{Inj}_{Gr}(G, H)$ l'ensemble des morphisme injectifs (qu'on appelle egalement *monomorphismes* de groupes),
- $\text{Surj}_{Gr}(G, H)$ l'ensemble des morphisme surjectifs (qu'on appelle egalement *epimorphismes* de groupes), et
- $\text{Isom}_{Gr}(G, H)$, l'ensemble des morphisme de groupes bijectifs (qu'on appelle lgalement *isomorphismes* de groupes).
- Si $H = G$, on ecrit notera ces ensembles

$$\text{Hom}_{Gr}(G), \text{Inj}_{Gr}(G), \text{Surj}_{Gr}(G), \text{Isom}_{Gr}(G);$$

en particulier l'ensemble des morphismes de G sur lui-meme $\text{Hom}_{Gr}(G)$ est aussi appelle ensemble des *endomorphismes* du groupe G et est egalement note

$$\text{End}_{Gr}(G) := \text{Hom}_{Gr}(G, G).$$

L'ensemble des endomorphismes bijectifs (isomorphismes) de G sur lui-meme est note

$$\text{Aut}_{Gr}(G) := \text{Isom}_{Gr}(G, G)$$

est est appele l'ensemble des automorphismes de G .

Les lois de compositions s'appliquent egalement aux morphismes de groupes:

PROPOSITION 2.5. (*Invariance par composition et par reciproque*) Soient $(G, \star), (H, *), (K, \otimes)$ des groupes et

$$\varphi : G \mapsto H \text{ et } \psi : H \mapsto K$$

des morphismes de groupes alors la composee $\psi \circ \varphi : G \mapsto K$ est un morphisme de groupes.

Supposons que $\varphi : G \mapsto H$ un morphisme de groupes bijectif alors l'application reciproque est un morphisme de groupe bijectif:

$$\varphi^{-1} \in \text{Hom}_{Gr}(H, G).$$

Preuve: Soit $g, g' \in G$ alors

$$\psi \circ \varphi(g \star g') = \psi(\varphi(g \star g')) = \psi(\varphi(g) * \varphi(g')) = \psi(\varphi(g)) \otimes \psi(\varphi(g')) = \psi \circ \varphi(g) \otimes \psi \circ \varphi(g').$$

Supposons que φ soit bijectif. Il faut montrer que pour $h, h' \in H$

$$\varphi^{-1}(h * h') = \varphi^{-1}(h) \star \varphi^{-1}(h').$$

Soit $g = \varphi^{-1}(h)$, $g' = \varphi^{-1}(h')$ alors

$$\varphi(g \star g') = \varphi(g) * \varphi(g') = \varphi(\varphi^{-1}(h)) * \varphi(\varphi^{-1}(h')) = h * h'.$$

Ainsi $g \star g' \in \varphi^{-1}(\{h * h'\})$ mais comme φ est bijective $\varphi^{-1}(\{h * h'\})$ ne possede qu'un seul element et comme $\varphi^{-1}(h * h')$ en fait partie (puisque $\varphi(\varphi^{-1}(h * h')) = h * h'$) on a

$$\varphi^{-1}(h) \star \varphi^{-1}(h') = g \star g' = \varphi^{-1}(h * h')$$

□

On en deduit de la proposition precedente le

COROLLAIRE 2.4. *L'ensemble des automorphismes de G*

$$\text{Aut}_{Gr}(G) \subset \text{Bij}(G)$$

est un sous-groupe pour la composition $\circ$.

Preuve: En effet l'ensemble $\text{Aut}_{Gr}(G) \subset \text{Bij}_{ENS}(G)$ est stable par composition et par reciproque. On applique le critere de sous-groupe. □

2.4.5. Groupes isomorphes. Soient G, H deux groupes tels que $\text{Iso}_{Gr}(G, H) \neq \emptyset$ et il existe donc un isomorphisme de groupes

$$\varphi : G \xrightarrow{\sim} H.$$

On dit alors que G et H sont *isomorphes* et on le note

$$G \simeq_{Gr} H.$$

Si c'est le cas, – pour autant que l'on soit interesse par les structures de groupes – G et H ont exactement les meme proprietes et peuvent etre identifies l'un a l'autre comme groupes via les morphismes φ et φ^{-1} .

EXERCICE 2.4. Montrer que la relation pour deux groupes d'etre isomorphes est une relation d'equivalence dans la categorie des groupes (qui n'est pas un ensemble): elle est reflexive, symetrique et transitive.

EXERCICE 2.5. Soient G et H deux groupes isomorphes (de sorte que $\text{Iso}_{Gr}(G, H) \neq \emptyset$). Montrer que pour tout $\varphi \in \text{Iso}_{Gr}(G, H)$ on a,

(1)

$$\text{Iso}_{Gr}(G, H) = \varphi \circ \text{Aut}_{Gr}(G) = \text{Aut}_{Gr}(H) \circ \varphi$$

avec

$$\varphi \circ \text{Aut}_{Gr}(G) = \{\varphi \circ \psi, \psi \in \text{Aut}_{Gr}(G)\}$$

et

$$\text{Aut}_{Gr}(H) \circ \varphi = \{\psi \circ \varphi, \psi \in \text{Aut}_{Gr}(H)\}.$$

2.5. Action d'un groupe sur un ensemble

l'exemple suivant de morphisme est fondamental en theorie des groupes et en mathematiques en general

DÉFINITION 2.13. Soit $(G, \star)$ un groupe, X un ensemble et $(\text{Bij}(X), \circ)$ le groupe symetrique de X (des bijections de X sur lui-meme). Une action (a gauche) de G sur X est la donnee d'un morphisme

$$\varphi : G \mapsto \text{Bij}(X).$$

On dit alors que G agit sur X (a gauche) a travers le morphisme φ et on le note $G \curvearrowright_{\varphi} X$.

PROPOSITION 2.6. La donnee d'une action $G \curvearrowright_{\varphi} X$ est equivalente a la donnee d'une application (appellee loi de composition externe)

$$\bullet \odot \bullet : \begin{array}{ccc} G \times X & \mapsto & X \\ (g, x) & \mapsto & g \odot x \end{array}$$

verifiant

(1) neutralite de l'element neutre:

$$\forall x \in X, e_G \odot x = x,$$

(2) *associativité*: $\forall x \in X, g, g' \in G,$

$$(g \star g') \odot x = g \odot (g' \odot x).$$

(3) *simplification*: en combinant les deux propriétés précédentes on a $\forall x \in X, g \in G,$

$$g \odot (g^{-1} \odot x) = g^{-1} \odot (g \odot x) = x.$$

Preuve: (à compléter) Dans une direction, on associe à un morphisme $\varphi : G \mapsto X$ l'application

$$\bullet \odot_{\varphi} \bullet : \begin{array}{ccc} G \times X & \mapsto & X \\ (g, x) & \mapsto & g \odot_{\varphi} x \end{array} := \varphi(g)(x).$$

Dans l'autre direction, étant donné une application $\bullet \odot \bullet$, on considère pour tout $g \in G$, l'application

$$\varphi(g) : \begin{array}{ccc} X & \mapsto & X \\ x & \mapsto & \varphi(g)(x) := g \odot x \end{array}.$$

On montre alors que $\varphi(g)$ est une bijection de X sur X , de réciproque

$$\varphi(g)^{-1} = \varphi(g^{-1})$$

et que l'application

$$\varphi : g \mapsto \varphi(g) \in \text{Bij}(X)$$

est un morphisme de groupes. □

EXEMPLE 2.5.1. Soit X un ensemble et $\sigma \in \text{Bij}(X)$ une bijection de X sur X , on a vu que l'application

$$\sigma^{\bullet} : n \in \mathbb{Z} \mapsto \sigma^n \in \text{Bij}(X)$$

est un morphisme de groupes et on obtient donc une action du groupe $(\mathbb{Z}, +)$ sur X qu'on pourrait noter par

$$\mathbb{Z} \curvearrowright_{\sigma} X : n \odot_{\sigma} x := \sigma^n(x).$$

Notons que si on change σ on obtient une autre action $\mathbb{Z} \curvearrowright X$.

2.5.1. Action par translations dans un groupe. Soit $(G, \cdot)$ un groupe et $g \in G$, l'application de translation à gauche par g est l'application

$$t_g : \begin{array}{ccc} G & \mapsto & G \\ g' & \mapsto & g.g' \end{array}.$$

Cette application n'est PAS un morphisme de groupe en général: elle ne l'est que si $g = e_G$. En effet si $g = e_G$, on a $t_g(g') = e_G.g' = g'$ et $t_{e_G} = \text{Id}_G$. Sinon on a

$$t_g(e_G) = g.e_G = g \neq e_G$$

donc t_g , n'est PAS un morphisme de groupes.

En revanche $t_g \in \text{Bij}(G)$. En effet, t_g admet $t_{g^{-1}}$ comme application réciproque:

$$t_{g^{-1}} \circ t_g(g') = g^{-1}.g.g' = g'$$

et donc $t_{g^{-1}} \circ t_g = \text{Id}_G$ et de même $t_g \circ t_{g^{-1}} = \text{Id}_G$.

THÉORÈME 2.13. *L'application translation à gauche*

$$t_{\bullet} : \begin{array}{ccc} G & \mapsto & \text{Bij}(G) \\ g & \mapsto & t_g : g' \mapsto g.g' \end{array}$$

est un morphisme de groupes de $(G, \cdot)$ vers $(\text{Bij}(G), \circ)$. Le morphisme $t_{\bullet}$ définit donc une action à gauche de G sur G qu'on appellera action par translations à gauche et qu'on notera $G \curvearrowright_t G$.

Preuve: Pour tout $g_1, g_2 \in G$ et tout $g' \in G$ on a

$$t_{g_1} \circ t_{g_2}(g') = t_{g_1}(t_{g_2}(g')) = t_{g_1}(g_2.g') = g_1.(g_2.g') = (g_1.g_2).g' = t_{g_1.g_2}(g')$$

et donc

$$t_{g_1} \circ t_{g_2} = t_{g_1.g_2}.$$

On a donc bien un morphisme de groupes. $\square$

REMARQUE 2.5.1. La notation pour la definition equivalente d'une action a gauche dans la Proposition 2.6 est faite pour copier l'action par translation a gauche sur le groupe.

EXERCICE 2.6. Soit G un groupe et

$$t_\bullet : \begin{array}{ccc} G & \mapsto & \text{Bij}(G) \\ g & \mapsto & t_g : G \mapsto G \end{array}$$

l'action par translation a gauche de G vers G .

(1) Montrer que $t_\bullet$ est injective.

REMARQUE 2.5.2. L'image de ce morphisme $t_G \subset \text{Bij}(G)$ est donc un sous-groupe de G : le groupe des translations a gauche sur G . Ainsi on a un isomorphisme de groupes

$$G \xrightarrow{\sim} t_G.$$

Ainsi un groupe quelconque, G , est toujours isomorphe a un sous-groupe d'un groupe de permutation d'un ensemble, $\text{Bij}(G)$.

2.5.2. La conjugaison dans un groupe. Un autre exemple fondamental d'action de groupe est la *conjugaison* d'une groupe sur lui-meme.

Soit $(G, \cdot)$ un groupe et $g \in G$ un element. La conjugaison par g est l'application

$$\text{Ad}_g : \begin{array}{ccc} G & \mapsto & G \\ h & \mapsto & g.h.g^{-1}. \end{array}$$

THÉORÈME 2.14. Pour tout g , l'application $\text{Ad}_g : G \mapsto G$ est un isomorphisme de groupes (ie $\text{Ad}_g \in \text{Aut}_G(G)$) dont l'application reciproque vaut

$$\text{Ad}_g^{-1} = \text{Ad}_{g^{-1}} : G \xrightarrow{\sim} G.$$

De plus l'application

$$\text{Ad}_\bullet : \begin{array}{ccc} G & \mapsto & \text{Bij}(G) \\ g & \mapsto & \text{Ad}_g \end{array}$$

est un morphisme de groupes.

Preuve: Calculons (comme $g.g^{-1} = e_G$)

$$\text{Ad}_g(h.h') = g.h.h'.g^{-1} = g.h.e_G.h'.g^{-1} = g.h.g.g^{-1}.h'.g^{-1} = \text{Ad}_g(h).\text{Ad}_g(h').$$

Verifions que Ad_g est injective en calculant son noyau:

$$\ker(\text{Ad}_g) = \{h \in G, g.h.g^{-1} = e_G\}$$

mais

$$g.h.g^{-1} = e_G \implies g.h = g \implies h = e_G$$

(en multipliant a droite par g et a gauche par g^{-1} . Notons ensuite que pour tout $h' \in G$

$$\text{Ad}_g(g^{-1}.h'.g) = g.g^{-1}.h'.g.g^{-1} = h'$$

donc $h' \in \text{Im}(\text{Ad}_g)$ et l'application est surjective. En fait on a pour tout $h \in G$

$$\text{Ad}_{g^{-1}}(\text{Ad}_g(h)) = h, \text{Ad}_g(\text{Ad}_{g^{-1}}(h)) = h$$

de sorte que $\text{Ad}_{g^{-1}}$ est la reciproque de Ad_g . Ainsi $\text{Ad}_g \in \text{Bij}(G)$.

On a pour tout $g, g' \in G, h \in G$

$$\text{Ad}_g \circ \text{Ad}_{g'}(h) = g.g'.h.g'^{-1}.g^{-1} = \text{Ad}_{g.g'}(h)$$

de sorte que

$$\text{Ad}_g \circ \text{Ad}_{g'} = \text{Ad}_{g.g'}$$

et l'application $\text{Ad} : G \mapsto \text{Bij}(G)$ est bien un morphisme de groupes (dont l'image est contenue dans $\text{Aut}_{Gr}(G)$). $\square$

DÉFINITION 2.14. *L'application de conjugaison*

$$\begin{array}{ccc} \text{Ad} : G & \mapsto & \text{Bij}(G) \\ g & \mapsto & \text{Ad}_g \end{array}$$

etant un morphisme de groupes, elle defini une action a gauche de G sur G (par automorphismes de groupes) qu'on appelle action par conjugaison et qu'on notera $G \curvearrowright_{\text{Ad}} G$.

L'image de ce morphisme

$$\text{Ad}_G = \{\text{Ad}_g, g \in G\} \subset \text{Aut}_{Gr}(G) \subset \text{Bij}(G)$$

(formee d'automorphismes de groupe) et est appelee groupe des automorphismes "interieurs" de G et est notee

$$\text{Ad}_G = \text{Int}_{Gr}(G) = \text{Inn}_{Gr}(G).$$

("Inn" pour "Inner").

REMARQUE 2.5.3. Le noyau de Ad est le sous-groupe

$$\begin{aligned} \ker(\text{Ad}) &= \{g \in G, \text{Ad}_g = \text{Id}_G\} = \{g \in G, \forall h \in G, g.h.g^{-1} = h\} \\ &= \{g \in G, \forall h \in G, g.h = h.g\} \end{aligned}$$

est l'ensemble des elements de G qui commutent avec tous les elements de G , on appelle ce sous-groupe le *centre* de G et on le note

$$Z(G) \subset G.$$

EXERCICE 2.7. (suite de l'exercice 2.5) Soient G et H deux groupes isomorphes (de sorte que $\text{Iso}_{Gr}(G, H) \neq \emptyset$). Montrer que pour tout $\varphi \in \text{Iso}_{Gr}(G, H)$

(1) L'application

$$\text{Ad}_\varphi : \begin{array}{ccc} \text{Aut}_{Gr}(G) & \mapsto & \text{Aut}_{Gr}(H) \\ \phi & \mapsto & \varphi \circ \phi \circ \varphi^{-1} \end{array}$$

est un isomorphisme de groupes entre $\text{Aut}_{Gr}(G)$ et $\text{Aut}_{Gr}(H)$.

REMARQUE. Noter que cette application de conjugaison par φ n'est pas de $\text{Aut}_{Gr}(G)$ vers $\text{Aut}_{Gr}(G)$ (sauf si $G = H$) mais de $\text{Aut}_{Gr}(G)$ vers $\text{Aut}_{Gr}(H)$.

2.5.3. Action a droite d'un groupe sur un ensemble. On peut egalement definir la notion d'action a droite. Pour cela la notion d'antimorphisme est tres utile:

DÉFINITION 2.15. Soient $(G, \star)$ et $(H, *)$ deux groupes, un anti-morphisme de groupes $\varphi : G \mapsto H$ est une application telle que

$$\forall g, g' \in G, \varphi(g \star g') = \varphi(g') * \varphi(g).$$

PROPOSITION 2.7. Une application entre groupes $\varphi : G \rightarrow H$ est un anti-morphisme de groupes ssi

$$\varphi \circ \bullet^{-1} : g \mapsto \varphi(g^{-1})$$

est un morphisme de groupes ou bien ssi

$$\bullet^{-1} \circ \varphi : g \mapsto \varphi(g)^{-1}$$

est un morphisme de groupes.

Preuve: Exercice. □

DÉFINITION 2.16. Soit $(G, \star)$ un groupe, X un ensemble et $(\text{Bij}(X), \circ)$ le groupe symétrique de X (des bijections de X sur lui-même). Une action à droite de G sur X est la donnée d'un antimorphisme de groupes

$$\varphi : G \mapsto \text{Bij}(X).$$

On dit alors que G agit sur X à droite à travers φ et on le note $X \curvearrowright_{\varphi} G$.

PROPOSITION 2.8. La donnée d'une action à droite $X \curvearrowright_{\varphi} G$ est équivalente à la donnée d'une application

$$\bullet \odot \bullet : \begin{array}{ccc} X \times G & \mapsto & X \\ (x, g) & \mapsto & x \odot g \end{array}$$

verifiant

- (1) neutralité de l'élément neutre: $\forall x \in X, x \odot e_G = x$,
- (2) associativité: $\forall x \in X, g, g' \in G, x \odot (g \star g') = (x \odot g) \odot g'$.
- (3) simplification: en combinant les deux propriétés précédentes on a $\forall x \in X, g \in G$,

$$(x \odot g) \odot g^{-1} = (x \odot g^{-1}) \odot g = x.$$

REMARQUE 2.5.4. On voit ainsi que dans une action à droite pour calculer l'action de $g \star g'$ sur x , on fait d'abord "agir" g sur x et ensuite on fait "agir" g' sur le résultat alors que pour une action à gauche c'est g' qui agit en premier et ensuite g agit sur le résultat.

2.5.3.1. Action par translations à droite. Soit $(G, \cdot)$ un groupe et $g \in G$, l'application de translation à droite par g est l'application

$$\text{td}_g : \begin{array}{ccc} G & \mapsto & G \\ g' & \mapsto & g' \cdot g \end{array}$$

Tout comme pour la translation à gauche, cette application n'est PAS un morphisme de groupes en général (sauf si $g = e_G$).

Par ailleurs $\text{td}_g \in \text{Bij}(G)$. En effet, td_g admet $\text{td}_{g^{-1}}$ comme application réciproque: pour tout g' , on a

$$\text{td}_{g^{-1}} \circ \text{td}_g(g') = g' \cdot g \cdot g^{-1} = g'$$

et donc

$$t_{g^{-1}} \circ t_g = \text{Id}_G$$

et de même

$$t_g \circ t_{g^{-1}} = \text{Id}_G.$$

THÉORÈME 2.15. L'application de translation à droite

$$\text{td}_{\bullet} : \begin{array}{ccc} G & \mapsto & \text{Bij}(G) \\ g & \mapsto & \text{td}_g : g' \mapsto g' \cdot g \end{array}$$

est un anti-morphisme de $(G, \cdot)$ vers $(\text{Bij}(G), \circ)$ et définit donc une action à droite de G sur G qu'on appellera action par translations à droite et qu'on notera $G \curvearrowright_{\text{td}} G$ (le premier G est vu comme un ensemble et le second comme le groupe qui agit).

Preuve: Exercice. □

EXERCICE 2.8. Soit X, Y des ensembles, $\mathcal{F}(X, Y)$ l'espace des fonctions (ie. des applications) de X à valeurs dans (ie. vers) Y et $G \curvearrowright X$ un groupe agissant sur X à gauche: $(g, x) \mapsto g \odot x$.

- (1) Montrer que l'application

$$\bullet | \bullet : \begin{array}{ccc} (\mathcal{F}(X, Y), G) & \mapsto & \mathcal{F}(X, Y) \\ (f, g) & \mapsto & f|_g : x \mapsto f|_g(x) := f(g \odot x) \end{array}$$

defini une action à droite de G sur $\mathcal{F}(X, Y)$.

(2) Réciproquement, construire à partir d'une action à droite

$$X \curvearrowright G : (x, g) \mapsto x \odot' g$$

de G sur X , une action à gauche $G \curvearrowright \mathcal{F}(X, Y)$.

CHAPITRE 3

Anneaux

*"Un Anneau pour les gouverner tous,
Un Anneau pour les trouver,
Un Anneau pour les amener tous,
Et dans les ténèbres les lier"*

3.1. Anneaux

DÉFINITION 3.1. Un anneau $(A, +, \cdot, 0_A, 1_A)$ est la donnée, d'un groupe commutatif $(A, +)$ (note additivement) d'élément neutre note 0_A , d'une loi de composition interne (dite de multiplication)

$$\begin{aligned} \bullet \bullet : A \times A &\mapsto A \\ (a, b) &\mapsto a \cdot b \end{aligned}$$

et d'un élément unité $1_A \in A$ ayant les propriétés suivantes

(1) Associativité de la multiplication:

$$\forall a, b, c \in A, (a \cdot b) \cdot c = a \cdot (b \cdot c) = a \cdot b \cdot c.$$

(2) distributivité:

$$\forall a, b, c \in A, (a + b) \cdot c = a \cdot c + b \cdot c, c \cdot (a + b) = c \cdot a + c \cdot b.$$

(3) Neutralité de l'unité:

$$\forall a \in A, a \cdot 1_A = 1_A \cdot a = a.$$

Un anneau est dit commutatif si de plus la multiplication est commutative:

$$\forall a, b \in A, a \cdot b = b \cdot a.$$

LEMME 3.1. Pour tout $a, b \in A$, on a

$$0_A \cdot a = a \cdot 0_A = 0_A,$$

(on dit que l'élément neutre de l'addition 0_A est absorbant). Pour l'opposé, on a

$$(-a) \cdot b = -(a \cdot b) = a \cdot (-b).$$

Preuve: Pour tout a on a

$$a = 1_A \cdot a = (1_A + 0_A) \cdot a = a + 0_A \cdot a$$

et donc $0_A \cdot a = 0_A$. □

EXERCICE 3.1. Montrer que si $1'_A$ a la propriété de neutralité: $\forall a \in A, a \cdot 1'_A = 1'_A \cdot a = a$, alors $1'_A = 1_A$.

EXEMPLE 3.1.1. Quelques exemples importants d'anneaux:

(1) Les ensembles $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ munis de leurs lois usuelles sont des anneaux commutatifs.

- (2) *L'anneau nul*: Soit $\mathbf{Nul} = \{\mathbf{0}\}$ un ensemble non-vidé forme d'un seul element. On muni cet ensemble de l'addition et de la multiplication definies par

$$\mathbf{0} + \mathbf{0} := \mathbf{0}, \mathbf{0} \cdot \mathbf{0} := \mathbf{0}$$

alors

$$(\mathbf{Nul}, +, \cdot, \mathbf{0}, \mathbf{0})$$

est un anneau commutatif qu'on appelle l'anneau nul.

- (3) *Produits d'anneaux*: Soient A et B des anneaux alors le produit $A \times B$ muni de l'addition et de la multiplication "coordonnee par coordonnee"

$$(a, b) + (a', b') = (a +_A a', b +_B b'), (a, b) \cdot (a', b') = (a \cdot_A a', b \cdot_B b')$$

est un anneau avec $(0_A, 0_B)$ comme element neutre et $(1_A, 1_B)$ comme element unite.

Plus generalement si $A_1, \dots, A_n$ sont des anneaux on peut munir le produit

$$A_1 \times \dots \times A_n$$

d'une structure d'anneau par addition et multiplication "coordonnee par coordonnee" dont le neutre et l'unite sont $(0_{A_1}, \dots, 0_{A_n})$ et $(1_{A_1}, \dots, 1_{A_n})$.

- (4) *Anneau de fonctions* Soit X un ensemble et $\mathcal{F}(X; \mathbb{R})$ l'ensemble des fonctions sur X a valeurs dans $\mathbb{R}$: on definit l'addition et la multiplication de deux fonctions $f, g \in \mathcal{F}(X; \mathbb{R})$ par

$$f + g : x \mapsto (f + g)(x) = f(x) + g(x), f \cdot g : x \mapsto (f \cdot g)(x) := f(x) \cdot g(x).$$

Alors si $\underline{0}$ et $\underline{1}$ sont les fonctions constantes egales a 0 et 1, $(\mathcal{F}(X; \mathbb{R}), +, \cdot, \underline{0}, \underline{1})$ est un anneau commutatif.

Plus generalement si $(A, +, \cdot, 0_A, 1_A)$ est un anneau, et que

$$\underline{0}_A, \underline{1}_A : X \mapsto A$$

designent les fonctions de X vers A qui sont constantes egales respectivement a 0_A et 1_A , en posant pour $f, g \in \mathcal{F}(X, A)$

$$f + g : x \mapsto (f + g)(x) = f(x) + g(x) \in A, f \cdot g : x \mapsto (f \cdot g)(x) := f(x) \cdot g(x) \in A,$$

on verifie que

$$(\mathcal{F}(X; A), +, \cdot, \underline{0}_A, \underline{1}_A)$$

est un anneau.

- (5) Soit

$$\mathbb{R}[X] = \{P(X) = a_0 + a_1 \cdot X + a_2 X^2 + \dots + a_d \cdot X^d, d \geq 1, a_0, a_1, \dots, a_d \in \mathbb{R}\}$$

l'ensemble des fonctions polynomiales a coefficients dans $\mathbb{R}$. Alors $\mathbb{R}[X]$ muni de l'addition des polynomes et de la multiplication des polynomes est un anneau dont le neutre est le polynome constant nul 0 et l'element unite est le polynome constant 1.

- (6) Plus generalement on verra plus tard que pour tout anneau commutatif A on peut former l'anneau des polynomes a coefficients dans A , $A[X]$:

$$A[X] = \{P(X) = a_0 + a_1 \cdot X + a_2 X^2 + \dots + a_d \cdot X^d, d \geq 1, a_0, a_1, \dots, a_d \in A\}$$

qui est un anneau commutatif muni des lois d'addition et de multiplication des polynomes usuelles. Formellement, on ne definit PAS $A[X]$ comme l'ensemble des fonctions polynomiales de A a valeurs dans A (ce dernier anneau est en general plus petit) mais comme l'ensemble des symboles $a_0 + a_1 \cdot X + a_2 X^2 + \dots + a_d \cdot X^d$ munis des regles usuelles d'addition et de multiplications des polynomes.

(7) Soit A un anneau commutatif, l'ensemble

$$M_2(A) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix}, a, b, c, d \in A \right\}$$

des matrices 2×2 a coefficients dans A et muni des lois d'addition et de multiplication des matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a+a' & b+b' \\ c+c' & d+d' \end{pmatrix}, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}$$

est un anneau (non-commutatif) d'element nul la matrice nulle

$$0_{M_2(A)} = \begin{pmatrix} 0_A & 0_A \\ 0_A & 0_A \end{pmatrix}$$

et d'unité la matrice identité

$$1_{M_2(A)} = \text{Id}_2 = \begin{pmatrix} 1_A & 0 \\ 0 & 1_A \end{pmatrix}.$$

REMARQUE 3.1.1. On peut definir egalement le produit (externe) d'un scalaire $a' \in A$ et d'une matrice $m = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ en posant

$$a' \cdot m = a' \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} := \begin{pmatrix} a'a & a'b \\ a'c & a'd \end{pmatrix}$$

(on multiplies toutes les coordonnees de la matrice par le scalaire a').

Cette loi de multiplication externe a des proprietes d'associativite et de distributivite relativement a l'addition et au produit dans A et $M_2(A)$: pour $a', a'' \in A$, $m, m' \in M_2(A)$ on a

$$(a' \cdot a'') \cdot m = a' \cdot (a'' \cdot m) = a \cdot a' \cdot m$$

$$(a' + a'') \cdot m = a' \cdot m + a'' \cdot m, \quad a' \cdot (m + m') = a' \cdot m + a' \cdot m'.$$

Exemple: l'anneau des classes de congruences $\mathbb{Z}/q\mathbb{Z}$. Soit $q \geq 1$ un entier et

$$\mathbb{Z}/q\mathbb{Z} = \{a \pmod{q}, a \in \mathbb{Z}\}, \quad a \pmod{q} = a + q\mathbb{Z}$$

l'ensemble des classes de congruence de module q . On rappelle que $(\mathbb{Z}/q\mathbb{Z}, \boxplus, 0 \pmod{q}, \boxminus)$ forme un groupe commutatif qu'on note additivement: pour $a, b \in \mathbb{Z}$ on pose

$$a \pmod{q} \boxplus b \pmod{q} := a + b \pmod{q}.$$

En particulier, on verifie que c'est bien defini: si $a \pmod{q} = a' \pmod{q}$ et $b \pmod{q} = b' \pmod{q}$ alors

$$a + b \pmod{q} = a' + b' \pmod{q}.$$

Pour $a \pmod{q}$, $b \pmod{q}$ des classes de congruences, on pose¹

$$a \pmod{q} \boxtimes b \pmod{q} := a \cdot b \pmod{q}.$$

On verifie a nouveau que c'est bien defini: si $a \pmod{q} = a' \pmod{q}$ et $b \pmod{q} = b' \pmod{q}$ alors

$$a \pmod{q} \boxtimes b \pmod{q} = a \cdot b \pmod{q} = a' \cdot b' \pmod{q} = a' \pmod{q} \boxtimes b' \pmod{q}.$$

L' operation $\boxtimes$ nous fourni une application

$$\begin{aligned} \bullet \boxtimes \bullet : \mathbb{Z}/q\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z} &\mapsto \mathbb{Z}/q\mathbb{Z} \\ (a \pmod{q}, b \pmod{q}) &\mapsto a \cdot b \pmod{q} \end{aligned}$$

¹Remarquer que ce n'est pas exactement la meme operation $\boxtimes$ que dans la serie 1.

qui est bien définie: si $a', b' \in \mathbb{Z}$ sont tels que

$$a' \pmod{q} = a \pmod{q}, \quad b' \pmod{q} = b \pmod{q}$$

alors

$$a'.b' \pmod{q} = a.b \pmod{q}.$$

Ainsi pour tout entier $q \geq 1$, il existe un anneau commutatif fini de cardinal q .

Exemple: l'anneau des endomorphismes d'un groupe commutatif. Soit $(M, +)$ un groupe commutatif noté additivement et $\text{End}(M) := \text{End}_{Gr}(M)$ l'ensemble des endomorphismes de M (les morphismes de groupe de M vers M). Alors, on peut munir $\text{End}(M)$ d'une structure d'anneau (non-commutatif en général):

- (1) L'addition est définie comme suit : soient $\varphi, \psi \in \text{End}(M)$, on pose

$$\varphi + \psi : \begin{array}{ccc} M & \mapsto & M \\ m & \mapsto & (\varphi + \psi)(m) := \varphi(m) + \psi(m) \end{array}$$

alors $\varphi + \psi \in \text{End}(M)$ est bien un morphisme de groupes;

- (2) On définit l'opposé pour l'addition par

$$-\varphi : \begin{array}{ccc} M & \mapsto & M \\ m & \mapsto & (-\varphi)(m) := -\varphi(m) \end{array}$$

et on vérifie que $-\varphi$ est encore un morphisme de groupes: cela utilise le fait que M est commutatif.

- (3) Ainsi on montre que $(\text{End}(M), +)$ forme un groupe commutatif dont l'élément neutre est le morphisme nul:

$$\underline{0}_M : m \in M \mapsto 0_M.$$

- (4) La multiplication des endomorphismes est définie par la composition des applications:

$$\varphi \circ \psi : m \in M \mapsto \varphi \circ \psi(m) = \varphi(\psi(m)).$$

qui a la propriété d'associativité requise (cf. §1.3.4) et pour laquelle l'application identité

$$\text{Id}_M : m \in M \mapsto m$$

(qui est bien un morphisme de groupes) a la propriété de neutralité par rapport à l'addition. On vérifie alors la distributivité de la composition par rapport à l'addition (on utilise à nouveau les propriétés des morphismes de groupes)

$$\forall \varphi, \varphi', \psi \in \text{End}(M), \quad (\varphi + \varphi') \circ \psi = \varphi \circ \psi + \varphi' \circ \psi, \quad \psi \circ (\varphi + \varphi') = \psi \circ \varphi + \psi \circ \varphi'.$$

En effet $\forall m \in M$

$$(\varphi + \varphi') \circ \psi(m) = (\varphi + \varphi')(\psi(m)) = \varphi(\psi(m)) + \varphi'(\psi(m)) = \varphi \circ \psi(m) + \varphi' \circ \psi(m)$$

et

$$\begin{aligned} \psi \circ (\varphi + \varphi')(m) &= \psi((\varphi + \varphi')(m)) = \psi(\varphi(m) + \varphi'(m)) \\ &= \psi(\varphi(m)) + \psi(\varphi'(m)) = \psi \circ \varphi(m) + \psi \circ \varphi'(m) \end{aligned}$$

On obtient ainsi que

$$(\text{End}(M), +, \circ, \underline{0}_M, \text{Id}_M)$$

forme un anneau.

3.2. Elements inversibles

DÉFINITION 3.2. Soit A un anneau. Un element $a \in A$ est inversible si il existe $b \in A$ tel que

$$a.b = b.a = 1_A.$$

On dit alors que b est un inverse (à gauche et à droite) de a (pour la multiplication).

PROPOSITION 3.1. (Unicité de l'inverse) Soit A un anneau et $a \in A$ un element inversible et soit b tel que $a.b = b.a = 1_A$.

Soit b' verifiant

$$a.b' = 1_A$$

alors $b' = b$; de même si b' vérifie

$$b'.a = 1_A$$

alors $b' = b$

Preuve: Supposons que a est inversible avec $a.b = b.a = 1_A$ et soit $b' \in A$ tel que

$$a.b' = 1_A$$

alors

$$a.b' = 1_A \implies b.a.b' = b = 1_A.b' = b'.$$

□

NOTATION 3.1. Par la Proposition précédente si un element $a \in A$ est inversible son inverse est unique. On notera cet inverse

$$a^{-1}.$$

Notons que a^{-1} est également inversible et on a

$$(a^{-1})^{-1} = a.$$

On déduit de cette discussion que

PROPOSITION 3.2. Soit $A^\times$ l'ensemble des elements inversibles d'un anneau A , alors

$$(A^\times, \cdot, 1_A, \bullet^{-1})$$

forme un groupe: le groupe des elements inversibles de A .

REMARQUE 3.2.1. Rappelons que l'on utilise la notation additive pour le groupe commutatif $(A, +)$. En particulier pour tout $a \in A$, l'element $-a$ ("l'inverse" de a pour la loi $+$) sera appelé l'opposé de a :

$$a + (-a) = (-a) + a = 0_A.$$

On réservera le terme "inverse" à la multiplication.

REMARQUE 3.2.2. Par une perversion du vocabulaire, le groupe $A^\times$ est également appelé le groupe des unités de A et ses elements sont des unités de A . Quelque fois quand on voudra parler d'un element a inversible on parlera d'une "unité" de A et on réservera le terme "l'unité de A " à l'element 1_A .

EXEMPLE 3.2.1. (1) On a

$$\mathbb{Z}^\times = \{+1, -1\}, \mathbb{Q}^\times = \mathbb{Q} - \{0\}, \mathbb{R}^\times = \mathbb{R} - \{0\}, \mathbb{C}^\times = \mathbb{C} - \{0\}.$$

par exemple 2 n'est pas inversible dans $\mathbb{Z}$ car son inverse $1/2$ n'est pas entier mais il est inversible dans $\mathbb{Q}$.

(2) On a

$$\text{Nul}(A)^\times = \{0_A\}.$$

- (3) Les matrices inversibles de $\mathbb{R}$ sont celles dont le determinant est inversible:

$$M_2(\mathbb{R})^\times = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix}, a, b, c, d \in \mathbb{R}, \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \in \mathbb{R}^\times = \mathbb{R} - \{0\} \right\}.$$

- (4) Si $(M, +)$ est un groupe commutatif et $\text{End}(M) = \text{End}_{Gr}(M)$ est son anneau d'endomorphismes, le groupe des unites de $\text{End}(M)$ est

$$\text{End}(M)^\times = \text{Aut}_{Gr}(M)$$

le groupe des automorphismes du groupe $(M, +)$.

- (5) Si A et B sont des anneaux, le groupe des elements inversibles du produit $A \times B$ est

$$(A \times B)^\times = A^\times \times B^\times.$$

- (6) Anneau des classes de congruences: les elements inversibles de $\mathbb{Z}/q\mathbb{Z}$ sont les classes de congruences premieres a q :

$$(\mathbb{Z}/q\mathbb{Z})^\times = \{a \pmod{q}, (a, q) = 1\}.$$

En effet si $a \pmod{q} \in (\mathbb{Z}/q\mathbb{Z})^\times$, il existe $d \pmod{q}$ tel que

$$a \pmod{q} \cdot d \pmod{q} = 1 \pmod{q}$$

et donc

$$a \cdot d \pmod{q} = 1 \pmod{q}.$$

Il existe donc $b \in \mathbb{Z}$ tel que

$$a \cdot d = 1 + qb \iff ad - qb = 1.$$

Cela implique que a et d sont premiers entre eux. Cela nous donne l'inclusion $\subset$.

Supposons $(a, q) = 1$ par Bezout il existe $d, b \in \mathbb{Z}$ tel que

$$ad - qb = 1$$

et donc

$$ad \equiv 1 \pmod{q}$$

ce qui nous donne l'inclusion $\supset$.

EXERCICE 3.2. Soit A un anneau commutatif et $M_2(A)$ l'anneau des matrices a coefficients dans

A. Soit $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(A)$, la transposee de la matrice des *cofacteurs* de M est la matrice definie par

$$\text{tcof}(M) := \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

- (1) Montrer que

$$M \cdot \text{tcof}(M) = \text{tcof}(M) \cdot M = \det(M) \cdot \text{Id}_2 = \begin{pmatrix} \det(M) & 0 \\ 0 & \det(M) \end{pmatrix}$$

ou $\det(M)$ (le determinant de M) est defini par

$$\det(M) := ad - bc \in A.$$

- (2) En deduire que

$$M_2(A)^\times =: \text{GL}_2(A) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix}, a, b, c, d \in A, \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \in A^\times \right\}.$$

3.2.0.1. *Divisibilite.*

DÉFINITION 3.3. Soit $(A, +, \cdot)$ un anneau commutatif et $a, c \in A$, on dit que a divise c et on le note

$$a|c$$

si il existe $b \in A$ tel que

$$c = a.b.$$

On dit également que a est un diviseur de b .

EXERCICE 3.3. Soit A un anneau.

- (1) Montrer que la relation de divisibilite est reflexive et transitive.
- (2) Montrer que tout element du groupe des unites $A^\times$ est un diviseur de tout element de A .
- (3) Quels sont les diviseurs de 0_A ? de 1_A ?

3.3. Sous-anneau

DÉFINITION 3.4. Soit $(A, +, \cdot)$ un anneau. Un sous-anneau $B \subset A$ est un sous-groupe de $(A, +)$ qui est

- soit le sous-groupe trivial $\{0_A\}$,
- soit qui contient l'unité 1_A et qui est stable par multiplication:

$$\forall b, b' \in B, \quad b.b' \in B.$$

Ainsi $(B, +, \cdot, 0_A, 1_A)$ est un anneau.

PROPOSITION 3.3. (Critere de sous-anneau) Soit $(A, +, \cdot)$ un anneau et $B \subset A$ un sous-ensemble non-vide; alors B est un sous-anneau ssi $B = \{0_A\}$, ou bien $1_A \in B$ et

$$(3.3.1) \quad \forall b, b', b'' \in B, \quad b.b' - b'' \in B$$

Preuve: Exercice. □

EXEMPLE 3.3.1. (1) La chaine d'inclusions

$$\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

est une chaine de sous-anneaux de $\mathbb{C}$.

- (2) Les seuls sous-anneaux de $\mathbb{Z}$ sont $\{0\}$ et $\mathbb{Z}$.
- (3) Les seuls sous-anneaux de $\mathbb{Z}/q\mathbb{Z}$ sont $\{0 \pmod{q}\}$ et $\mathbb{Z}/q\mathbb{Z}$.
- (4) La chaine d'inclusions

$$M_2(\mathbb{Z}) \subset M_2(\mathbb{Q}) \subset M_2(\mathbb{R}) \subset M_2(\mathbb{C})$$

est une chaine de sous-anneaux.

- (5) Pour tout anneau commutatif, l'ensemble des matrices scalaires

$$A.\text{Id}_2 = \{a.\text{Id}_2 = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}, \quad a \in A\} \subset M_2(A),$$

l'ensemble des matrices diagonales

$$\text{Diag}_2(A) = \left\{ \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}, \quad a, d \in A \right\} \subset M_2(A),$$

et l'ensemble des matrices triangulaires superieures

$$T_{\text{sup},2}(A) = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}, \quad a, b, d \in A \right\} \subset M_2(A)$$

sont des sous-anneaux emboites les uns dans les autres.

l'ensemble des matrices triangulaires inferieures

$$T_{\text{inf},2}(A) = \left\{ \begin{pmatrix} a & 0 \\ c & d \end{pmatrix}, a, c, d \in A \right\} \subset M_2(A)$$

est egalement un sous-anneau.

- (6) Si $B, C \subset A$ sont des sous-anneaux de A alors $B \cap C$ est un sous-anneau de A . Plus generalement pour toute collection $(A_i)_{i \in I}$ de sous-anneaux $A_i \subset A$ de A , l'intersection

$$\bigcap_{i \in I} A_i = \{a \in A, \forall i \in I, a \in A_i\}$$

est un sous-anneau de A . En particulier, pour tout ensemble $X \subset A$ il existe un plus petit sous-anneau de A contenant X (l'intersection de l'ensemble des sous-anneaux de A contenant X): on l'appelle le *sous-anneau engendre par X* et on le note

$$\langle X \rangle \subset A.$$

3.4. Morphismes d'anneaux

DÉFINITION 3.5. Soient $(A, +_A, \cdot_A)$, $(B, +_B, \cdot_B)$ des anneaux. Un morphisme d'anneaux $\varphi : A \mapsto B$ est un morphisme de groupes commutatif $\varphi : (A, +_A) \mapsto (B, +_B)$ tel que

$$\begin{aligned} \varphi(1_A) &= 1_B \text{ ou bien } \varphi(1_A) = 0_B, \\ \forall a, a' \in A, \varphi(a \cdot_A a') &= \varphi(a) \cdot_B \varphi(a'). \end{aligned}$$

REMARQUE 3.4.1. Si $\varphi(1_A) = 0_B$ alors φ est l'application constante nulle $\underline{0}_B$:

$$\forall a \in A, \varphi(a) = \varphi(a) \cdot \varphi(1_A) = 0_B.$$

Le morphisme canonique. Le morphisme canonique associe a un anneau A est l'application

$$\text{Can}_A : \begin{array}{ccc} \mathbb{Z} & \mapsto & A \\ n & \mapsto & n \cdot 1_A \end{array}$$

ou

$$n \cdot 1_A = \begin{cases} 0 & \text{si } n = 0 \\ 1_A + \cdots + 1_A (n \text{ fois}) & \text{si } n > 0 \\ -(1_A + \cdots + 1_A) (|n| \text{ fois}) & \text{si } n < 0. \end{cases}$$

On notera egalement pour $n \in \mathbb{Z}$

$$n_A := \text{Can}_A(n).$$

EXERCICE 3.4. On a deja vu que Can_A est un morphisme de groupes commutatifs (pour l'addition). Verifier que c'est un morphisme d'anneaux.

3.4.1. Noyau, Image.

PROPOSITION 3.4. (*Stabilite par morphismes*) Soient $\varphi \in \text{Hom}_{\text{Ann}}(A, B)$ un morphisme alors $\varphi(A) \subset B$ est un sous-anneau. Par ailleurs le sous-groupe $\ker(\varphi)$ est un sous-groupe de $(A, +)$ qui est de plus stable par multiplication (a gauche et a droite) par A :

$$\forall a \in A, k \in \ker(\varphi), a \cdot k, k \cdot a \in \ker(\varphi).$$

Preuve: On sait deja que $\varphi(A)$ est un sous-groupe de $(B, +)$. Si $\varphi(A)$ n'est pas l'anneau nul alors $1_B = \varphi(1_A) \in \varphi(A)$ et pour tout $b, b' \in \varphi(A)$, on a $b = \varphi(a)$, $b' = \varphi(a')$ pour $a, a' \in A$ et

$$b \cdot b' = \varphi(a) \cdot \varphi(a') = \varphi(a \cdot a') \in \varphi(A)$$

ainsi $\varphi(A)$ est stable par produit.

On sait deja que $\ker(\varphi)$ est un sous-groupe de $(A, +)$. De plus $\forall a \in A, k \in \ker(\varphi)$, on a

$$\varphi(a \cdot k) = \varphi(a) \cdot \varphi(k) = \varphi(a) \cdot 0_B = 0_B$$

donc $a \cdot k \in \ker(\varphi)$. □

REMARQUE 3.4.2. Notez que $\ker(\varphi)$ est PAS un sous-anneau en general : il ne contient pas 1_A sauf si $1_B = 0_B$ (c'est à dire sauf si B est l'anneau nul).

EXERCICE 3.5. Soit $\varphi : A \mapsto B$ un morphisme d'anneaux et $\{0_B\} \neq B' \subset B$ un sous-anneau qui n'est pas l'anneau nul. Montrer que l'image reciproque $A' = \varphi^{(-1)}(B')$ est un sous-anneau de A .

Comme φ est un morphisme de groupes additifs on a

PROPOSITION 3.5. *Un morphisme d'anneaux $\varphi \in \text{Hom}_{\text{Ann}}(A, B)$ est injectif ssi $\ker(\varphi) = \{0_A\}$.*

PROPOSITION 3.6. *Soient $\varphi : A \mapsto B$ et $\psi : B \mapsto C$ des morphismes d'anneaux alors*

- $\psi \circ \varphi : A \mapsto C$ est un morphisme d'anneaux.
- Soit $\varphi \in \text{Hom}_{\text{Ann}}(A, B)$ un morphisme d'anneaux bijectif, l'application reciproque $\varphi^{-1} : B \mapsto A$ est un morphisme d'anneaux. On dit que φ est un isomorphisme d'anneaux et on dit que A et B sont des anneaux isomorphes.

Preuve: Exercice. □

NOTATION 3.2. Soient A, B des anneaux. On note

$$\text{Hom}_{\text{Ann}}(A, B), \text{End}_{\text{Ann}}(A) = \text{Hom}_{\text{Ann}}(A, A)$$

$$\text{Isom}_{\text{Ann}}(A, B), \text{Aut}_{\text{Ann}}(A) = \text{Isom}_{\text{Ann}}(A, A)$$

l'ensemble des morphismes d'anneaux entre A et B , des endomorphismes de l'anneau A , des isomorphismes d'anneaux entre A et B et des automorphismes de l'anneau A .

EXERCICE 3.6. L'ensemble des automorphismes $\text{Aut}_{\text{Ann}}(A)$ muni de la composition forme un sous-groupe de $\text{Bij}(A)$.

3.5. Anneau quotient

Dans cette section on va donner une generalisation de la construction de l'anneau $\mathbb{Z}/q\mathbb{Z}$.

3.5.1. Ideal d'un anneau. On a vu que le noyau $\ker(\varphi)$ d'un morphisme d'anneaux $\varphi : A \rightarrow B$ n'est pas un sous-anneau en general. C'est un sous-groupe du groupe additif $(A, +)$ stable par multiplications par les elements de A . On va donner un nom a ces objets.

DÉFINITION 3.6. *Soit A un anneau pas forcément commutatif.*

- *Un ideal (à gauche) de A est un sous-groupe additif $(I, +) \subset (A, +)$ qui est stable par multiplication (à gauche) par les elements de A :*

$$\forall a \in A, b \in I, a.b \in I.$$

- *Un ideal (à droite) de A est un sous-groupe additif $(I, +) \subset (A, +)$ qui est stable par multiplication (à droite) par les elements de A :*

$$\forall a \in A, b \in I, b.a \in I.$$

- *Un ideal bilatere de A est un sous-groupe additif $(I, +) \subset (A, +)$ qui est un ideal à gauche et à droite:*

$$\forall a \in A, b \in I, a.b, b.a \in I.$$

En particulier si A est commutatif les notions d'ideal à gauche, à droite ou bilatere sont toutes les memes.

EXEMPLE 3.5.1. Soit $\varphi : A \mapsto B$ un morphisme d'anneaux alors $\ker(\varphi)$ est un ideal bilatere de A .

EXERCICE 3.7. Soit $I \subset A$ un idéal (à gauche) d'un anneau A . Montrer que si

$$I \cap A^\times \neq \emptyset$$

alors

$$I = A$$

(on commencera par montrer que si $A^\times \cap I \neq \emptyset$ alors $1_A \in I$ et on en déduira que $I = A$).

EXERCICE 3.8. Montrer que les idéaux de l'anneau $\mathbb{Z}$ sont les sous-groupes $q\mathbb{Z}$ pour $q \geq 0$.

3.5.2. Anneau quotient par un idéal. Soit $(A, +, \cdot)$ un anneau et $I \subset A$ un idéal *bilatère* (c'est automatique si A est commutatif). Pour $a \in A$, la classe de congruence de a modulo I est le sous-ensemble

$$a \pmod{I} := a + I = \{a + i, i \in I\} \subset A.$$

Soient $a, a' \in A$; on dit que a est congru à a' modulo I ssi on a

$$a \pmod{I} = a' \pmod{I};$$

on note cette relation

$$a \equiv a' \pmod{I}.$$

EXERCICE 3.9. Montrer que la relation de congruence modulo I , $a \equiv a' \pmod{I}$ est une relation d'équivalence sur A dont les classes d'équivalences sont précisément les classes de congruence $a \pmod{I}$. On pourra commencer par montrer l'équivalence

$$a \equiv a' \pmod{I} \iff a - a' \in I.$$

L'ensemble des classes de congruences modulo I (c'est un sous-ensemble de $\mathcal{P}(A)$) est noté

$$A/I := \{a + I, a \in A\}.$$

On peut munir cet ensemble A/I d'une structure d'anneau qu'on appelle *l'anneau quotient* de A par l'idéal I .

THÉORÈME 3.1. Soit $(A, +, \cdot, 0_A, 1_A)$ un anneau et $I \subset A$ un idéal bilatère et

$$A/I = \{a \pmod{I} = a + I, a \in A\}$$

l'ensemble des classes de congruences modulo I . En particulier on a

$$0_A \pmod{I} = I, 1_A \pmod{I} = 1_A + I.$$

(1) Il existe une (unique) structure d'anneau

$$(A/I, +_I, \cdot_I, 0_{A/I}, 1_{A/I})$$

telle que l'application

$$\pi_I := \bullet \pmod{I} : \begin{array}{ccc} A & \mapsto & A/I \\ a & \mapsto & a \pmod{I} \end{array}$$

soit un morphisme d'anneau surjectif de noyau

$$\ker(\pi_I) = I.$$

On appelle cet anneau l'anneau quotient de A par I et on appelle π_I morphisme canonique de A vers son quotient A/I .

(2) On a en particulier

$$(3.5.1) \quad 0_{A/I} = 0_A \pmod{I} = I, 1_{A/I} = 1_A \pmod{I} = 1_A + I$$

et pour tout $a, b \in A$

$$(3.5.2) \quad a \pmod{I} +_I b \pmod{I} = a + b \pmod{I}, a \pmod{I} \cdot_I b \pmod{I} = a \cdot b \pmod{I}.$$

(3) Soit $\varphi : A \rightarrow B$ un morphisme d'anneaux. On suppose que $I \subset \ker(\varphi)$. Alors il existe un unique morphisme d'anneaux

$$\varphi_I : A/I \rightarrow B$$

tel que

$$(3.5.3) \quad \forall a \in A, \varphi_I(a \pmod I) = \varphi(a).$$

En d'autres termes on a

$$(3.5.4) \quad \varphi = \varphi_I \circ \pi_I;$$

On dit que le morphisme φ se factorise par le morphisme canonique et on le note avec le diagramme suivant

$$\begin{array}{ccc} A & \xrightarrow{\varphi} & B \\ \pi_I \downarrow & \nearrow \varphi_I & \\ A/I & & \end{array}$$

Preuve: Notons que A/I est réduit à un seul élément ssi $I = A$. Alors le résultat est évident.

Si A/I n'est pas réduit à un élément (ie. si $I \subsetneq A$) on a nécessairement

$$\pi_I(0_A) = 0 \pmod I = 0_{A/I}, \pi_I(1_A) = 1 \pmod I = 1_{A/I}$$

ce qui montre qu'on doit avoir (3.5.1). Le fait que π_I doive être un morphisme d'anneaux implique (3.5.2): en effet on doit avoir

$$a + b \pmod I = \pi_I(a + b) = \pi_I(a) +_I \pi_I(b) = a \pmod I +_I b \pmod I$$

et

$$a \cdot b \pmod I = \pi_I(a \cdot b) = \pi_I(a) \cdot_I \pi_I(b) = a \pmod I \cdot_I b \pmod I.$$

Ainsi la structure d'anneau si elle existe est unique (l'application π_I est évidemment surjective: tout élément x de A/I s'écrivant $a + I$ est l'image de a par π_I)

Pour montrer l'existence, on voudrait poser

$$a \pmod I +_I b \pmod I := a + b \pmod I, \quad a \pmod I \cdot_I b \pmod I := a \cdot b \pmod I.$$

Le problème est qu'une classe $a \pmod I$ peut aussi s'écrire $a' \pmod I$ pour tout $a' \in a \pmod I$. On veut que le résultat ne dépende pas du choix de l'élément a' .

Il suffit donc de montrer que si

$$a \pmod I = a' \pmod I \text{ et } b \pmod I = b' \pmod I$$

alors

$$a + b \pmod I = a' + b' \pmod I \text{ et } a \cdot b \pmod I = a' \cdot b' \pmod I.$$

On doit donc montrer que

$$(a + b) - (a' + b') \in I, \quad a \cdot b - a' \cdot b' \in I.$$

On a

$$a - a' \in I, \quad b - b' \in I$$

et donc

$$(a + b) - (a' + b') = c + d \in I + I \subset I$$

car I est un sous groupe de $(A, +)$.

On a

$$\begin{aligned} a \cdot b - a' \cdot b' &= a \cdot b - a \cdot b' + a \cdot b' - a' \cdot b' \\ &= a \cdot (b - b') - (a - a') \cdot b' \in a \cdot I + I \cdot b' \subset I + I \subset I \end{aligned}$$

car I est un idéal (bilatère) de A et donc stable par addition et multiplication à gauche et à droite par des éléments quelconques de A (ici a et b').

Le fait que les lois $+_I$ et $\cdot_I$ soient associatives et distributives et que $0_A \pmod I$ et $1_A \pmod I$ en soit les elements neutre provient des definitions de ces lois et des proprietes correspondantes pour l'anneau $(A, +, \cdot, 0_A, 1_A)$.

Soit $\varphi : A \rightarrow B$ un morphisme tel que $I \subset \ker \varphi$. On veut montrer l'existence de $\varphi_I : A/I \rightarrow B$ verifiant (3.5.3). En particulier, comme π_I est surjectif un tel morphisme si il existe est unique.

Pour montrer l'existence il suffit de montrer que si $a \pmod I = a' \pmod I$ alors

$$\varphi(a) = \varphi(a').$$

Alors on pourra poser sans ambiguïté

$$\varphi_I(a \pmod I) = \varphi(a')$$

pour tout $a' \in a + I$ (c'est a dire pour tout a' tel que $a' + I = a + I$).

On a $a' = a + i$ avec $i \in I$ et donc

$$\varphi(a') = \varphi(a) + \varphi(i) = \varphi(a) + 0_B = \varphi(a)$$

car

$$\varphi(i) = 0_B$$

puisque $I \subset \ker(\varphi)$. □

CHAPITRE 4

Corps

"Le corps conditionne le raisonnement."

4.1. Corps

DÉFINITION 4.1. *Un corps K est un anneau commutatif possédant au moins deux éléments $0_K \neq 1_K$ et tel que tout élément non-nul est inversible:*

$$K^\times = K - \{0_K\}.$$

REMARQUE 4.1.1. Dans cette définition, on demande que K soit commutatif. Il existe des anneaux non-commutatifs dont l'ensemble des éléments inversibles sont exactement les éléments non-nuls. On les appelle *corps gauche* ou *algebres à divisions*.

EXEMPLE 4.1.1. On a $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ sont des corps; $\mathbb{Z}$ n'en est pas un (par exemple 2 n'est pas inversible dans $\mathbb{Z}$).

4.1.1. Exemples de corps finis. Un autre exemple fondamental est celui des corps finis.

THÉORÈME 4.1. *Soit $q \geq 2$ un nombre premier (les seuls diviseurs de q sont 1 et q) alors l'anneau des classes de congruences modulo q $(\mathbb{Z}/q\mathbb{Z}, +, \cdot)$ est un corps (fini de cardinal q).*

Preuve: Comme $q > 1$ (par définition un premier n'est pas égal à 1) on a $0 \pmod{q} \neq 1 \pmod{q}$ (car $q \nmid 1 - 0 = 1$). Ainsi $\mathbb{Z}/q\mathbb{Z}$ a au moins deux éléments.

On a montré au chapitre précédent que pour tout entier $q \geq 1$

$$(\mathbb{Z}/q\mathbb{Z})^\times = \{a \pmod{q}, (a, q) = 1\}.$$

On va voir que si q est premier on a

$$(a, q) \neq 1 \iff q|a \iff a \pmod{q} = 0 \pmod{q}.$$

Cela nous donnera que

$$(\mathbb{Z}/q\mathbb{Z})^\times = \mathbb{Z}/q\mathbb{Z} - 0 \pmod{q}.$$

$\Leftarrow$: si $q|a$ alors q est un diviseur commun de a et q et donc $q|(a, q) > 1$ car $q > 1$ (un premier n'est pas égal à 1).

Montrons $\Rightarrow$: si $(a, q) > 1$ alors a et q admettent un diviseur commun non-trivial (le pgcd (a, q)) et comme q est premier ses seuls diviseurs sont 1 et q et donc $(a, q) = q$ et $q|a$. □

NOTATION 4.1. *Soit $q \geq 2$ un nombre premier, le corps fini à q éléments $(\mathbb{Z}/q\mathbb{Z}, +, \cdot)$ est noté $\mathbb{F}_q$.*

REMARQUE 4.1.2. Réciproquement si $(\mathbb{Z}/q\mathbb{Z}, +, \cdot)$ est un corps alors q est premier: en effet si $q = q_1 \cdot q_2$ est composé (avec $2 \leq q_1, q_2 < q$) alors on a

$$q_1 \pmod{q} \cdot q_2 \pmod{q} = q_1 \cdot q_2 \pmod{q} = q \pmod{q} = 0 \pmod{q}.$$

La classe $q_1 \pmod{q}$ est non-nulle (car q ne divise pas q_1) mais elle n'est pas inversible non-plus: si on avait q'_1 tel que $q'_1 \pmod{q} \cdot q_1 \pmod{q} = 1 \pmod{q}$ on aurait

$$q'_1 \pmod{q} \cdot q_1 \pmod{q} \cdot q_2 \pmod{q} = 0 \pmod{q} = 1 \pmod{q} \cdot q_2 \pmod{q} = q_2 \pmod{q}$$

mais q ne divise pas q_2 (car $1 < q_2 < q$).

PROPOSITION 4.1. *Soit $q \geq 2$ un nombre premier et $\mathbb{F}_q = \mathbb{Z}/q\mathbb{Z}$ le corps à q éléments. Pour tout $x \in \mathbb{F}_q$ on a*

$$x^q = x.$$

REMARQUE 4.1.3. En particulier les fonctions polynomiales sur $\mathbb{F}_q$

$$X : \mathbb{F}_q \mapsto \mathbb{F}_q, X^q : \mathbb{F}_q \mapsto \mathbb{F}_q$$

$$x \mapsto x, x \mapsto x^q$$

sont identiques !

Preuve: Comme $\mathbb{F}_q$ est un corps, son groupe multiplicatif des éléments inversibles vaut

$$\mathbb{F}_q^\times = (\mathbb{Z}/q\mathbb{Z})^\times = \mathbb{F}_q - \{0_{\mathbb{F}_q}\}$$

est d'ordre $q - 1$. Par le théorème de Lagrange,

$$\forall x \in \mathbb{F}_q^\times = \mathbb{F}_q - \{0\}, x^{q-1} = 1_{\mathbb{F}_q}$$

et donc multipliant encore par x

$$\forall x \in \mathbb{F}_q^\times = \mathbb{F}_q - \{0\}, x^q = x$$

et cette dernière égalité est aussi valable pour $x = 0_{\mathbb{F}_q}$. $\square$

4.1.2. Injectivité des morphismes depuis un corps. Comme on va le voir, le fait, dans un corps, de pouvoir inverser tous les éléments non-nuls simplifie considérablement la théorie. Par exemple on a

PROPOSITION 4.2. *Soit K un corps, B un anneau et $\varphi \in \text{Hom}_{\text{Ann}}(K, B)$ un morphisme d'anneaux. Alors si φ n'est pas nul ($\varphi \neq 0_B$) φ est injectif:*

$$\varphi : K \hookrightarrow B.$$

Preuve: Supposons que φ n'est pas nul. Il s'agit de montrer que $\ker \varphi = \{0_K\}$. Soit $x \in K - \{0\}$, alors x est inversible et soit x^{-1} son inverse. On a

$$\varphi(x.x^{-1}) = \varphi(1_K) = \varphi(x).\varphi(x^{-1})$$

et comme $\varphi \neq 0_B$, $\varphi(1_K) = 1_B \neq 0_B$ et $\varphi(x) \neq 0$ et donc $x \notin \ker(\varphi)$. $\square$

REMARQUE 4.1.4. On a même mieux: si $x \in K - \{0\}$ alors $\varphi(x)$ est inversible dans B , d'inverse

$$\varphi(x)^{-1} = \varphi(x^{-1}).$$

Une autre propriété caractéristique des corps est la suivante:

PROPOSITION 4.3. *Soit K un corps alors tout idéal $I \subset K$ est soit $I = \{0_K\}$ ou bien $I = K$.*

Reciproquement, soit A un anneau commutatif possédant au moins deux éléments alors si ses idéaux sont $\{0_K\}$ ou bien K alors K est un corps.

Preuve: Soit $I \subset K$ un idéal non-nul et soit $a \in I - \{0\}$ alors a est inversible et il existe $a^{-1} \in K$ tel que

$$a^{-1}.a = 1.$$

Comme $a \in I$ et que I est un idéal, on a $a^{-1}.a \in I$ et donc $1 \in I$. Pour tout $b \in A$ on a alors

$$b = b.1 \in bI \subset I$$

et donc $A = I$.

Pour la réciproque, prendre $a \in K - \{0\}$ et considérer l'ensemble

$$(a) = a.K = \{ak, k \in K\} \subset K$$

et montrer que c'est un idéal et conclure. $\square$

4.2. Construction de corp: corps des fractions

Etant donne un anneau A , sous certaines hypotheses, on peut construire un corps K (le plus petit possible) dont A est peut etre considere comme un sous-anneau. En particulier si $a \in A - \{0\}$ alors il existe $a^{-1} \in K$ tel que $a.a^{-1} = 1_A = 1_K$. Pour cela il faut que A satisfasse une propriete particuliere: etre *integre*.

LEMME 4.1. *Soit $\{0\} \neq A \subset K$ un sous anneau non-nul d'un corps K alors A est commutatif et*

$$(4.2.1) \quad \forall a, b \in A, \quad a.b = 0 \iff a = 0 \text{ ou } b = 0.$$

Preuve: A est commutatif car K est commutatif. Pour (4.2.1) seule la direction $\implies$ est non evidente: supposons que $a, b \neq 0$ alors il existe $a^{-1} \in K$ tel que $a^{-1}.a = 1_K$ mais alors on a

$$a.b = 0 \implies a^{-1}.a.b = 0_K = b,$$

contradiction. □

DÉFINITION 4.2. *Un anneau A non-nul, commutatif, tel que $\forall a, b \in A$ on ait*

$$a.b = 0 \iff a = 0 \text{ ou } b = 0$$

est dit integre.

REMARQUE 4.2.1. En particulier un corps est integre: appliquer le lemme precedent a $A = K$.

EXERCICE 4.1. Montrer que si $q = q_1.q_2$ avec $q_1, q_2 \neq 1, q$ (des diviseurs non-triviaux de q) alors $(\mathbb{Z}/q\mathbb{Z}, +, \cdot)$ n'est pas integre et donc pas un corps (cf. Remarque 4.1.2)

THÉORÈME 4.2. *Soit A un anneau integre (en particulier commutatif), alors il existe un corps K et un morphisme d'anneau injectif*

$$\iota : A \hookrightarrow K$$

(de sorte qu'on peut considerer A comme un sous-anneau de K en identifiant A a son image $\iota(A) \subset K$) et tel que K a la propriete de minimalite suivante: pour tout corps K' et tout morphisme injectif

$$\iota' : A \hookrightarrow K',$$

il existe un morphisme (necessairement injectif)

$$\iota'_K : K \hookrightarrow K'$$

prolongeant le morphisme ι' (ainsi A et K peuvent etre vus comme des sous-anneaux de K').

REMARQUE 4.2.2. "Prolonge" signifie que pour $a \in A$, on a

$$\iota'_K(\iota(a)) = \iota'(a).$$

DÉFINITION 4.3. *Le corps K s'appelle le corps des fractions K et se note $\text{Frac}(A)$.*

Preuve: Soit A un anneau integre. On considere le produit cartesien

$$A \times (A - \{0\}) = \{(a, b), a, b \in A, b \neq 0\}.$$

On definit sur $A \times (A - \{0\})$ une relation $\sim$ en posant

$$(a, b) \sim (a', b') \iff a.b' = a'.b.$$

Cette relation est une relation d'equivalence (reflexive, symetrique, transitive). En effet

- reflexive: $(a, b) \sim (a, b)$ car $ab = ab$.
- symetrique: $(a, b) \sim (a', b') \iff a'b = ab' \iff (a', b') \sim (a, b)$

– transitive: si $(a, b) \sim (a', b')$ et $(a', b') \sim (a'', b'')$, alors on a

$$a.b' = a'.b, \quad a'.b'' = a''.b'$$

et comme A est commutatif

$$a.b''.b' = a.b'.b'' = a'.b.b'' = a''.b'.b = a''.b.b'.$$

On a donc

$$0_A = a.b''.b' - a''.b.b' = (a.b'' - a''.b).b'$$

et comme A est intègre et $b' \neq 0$ on a

$$a.b'' - a''.b = 0_A \iff a.b'' = a''.b \iff (a, b) \sim (a'', b'').$$

On note

$$K = \text{Frac}(A) = A \times (A - \{0\}) / \sim$$

l'ensemble des classes d'équivalence et on note

$$\frac{a}{b} \in K$$

la classe d'équivalence de la paire (a, b) . On l'appelle la fraction $\frac{a}{b}$ de numérateur a et de dénominateur b .

On munit $\text{Frac}(A)$ d'une structure d'anneau en posant

$$\frac{a}{b} + \frac{c}{d} := \frac{ad + bc}{bd}, \quad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}, \quad -\frac{a}{b} = \frac{-a}{b}$$

$$0_K = \frac{0}{1}, \quad 1_K = \frac{1}{1}.$$

Notons que comme A est intègre, si b et d sont non-nuls et produit $b.d$ est non-nul et

$$(a.d + b.c, b.d), (a.c, b.d) \in A \times (A - \{0\}).$$

On vérifie premièrement que ces définitions ne dépendent pas du choix des représentants de chaque classe d'équivalence: si $\frac{a}{b} = \frac{a'}{b'}$ et $\frac{c}{d} = \frac{c'}{d'}$ cad si

$$(a, b) \sim (a', b'), \quad (c, d) \sim (c', d')$$

alors

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd} = \frac{a'd' + b'c'}{b'd'} = \frac{a'}{b'} + \frac{c'}{d'}$$

et

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{a.c}{b.d} = \frac{a'.c'}{b'.d'} = \frac{a'}{b'} \cdot \frac{c'}{d'}$$

c'est à dire que

$$(ad + bc, bd) \sim (a'd' + b'c', b'd'), \quad (a.c, b.d) \sim (a'.c', b'.d').$$

Par exemple pour la première relation on doit montrer que

$$(ad + bc)b'd' = (a'd' + b'c')bd.$$

On a

$$(ad + bc)b'd' = ab'dd' + bb'cd' = a'bdd' + bb'c'd$$

en utilisant que

$$ab' = a'b, \quad cd' = c'd$$

et donc mettant bd en facteur on obtient

$$(ad + bc)b'd' = (a'd' + b'c')bd.$$

On doit vérifier ensuite que $(K, +, \cdot, 0_K, 1_K)$ forme un anneau (exercice)

Soit $\frac{a}{b} \neq 0_K = \frac{0}{1}$, cela signifie que

$$a.1 \neq b.0 = 0$$

et donc la paire $(b, a) \in A \times (A - \{0\})$ et on a

$$\frac{a}{b} \cdot \frac{b}{a} = \frac{a.b}{a.b} = \frac{1_A}{1_A} = 1_K$$

donc $\frac{a}{b}$ est inversible dans K et K est un corps.

Soit

$$\iota : \begin{array}{ccc} A & \mapsto & K \\ a & \mapsto & \frac{a}{1} \end{array}$$

On verifie que ι est un morphisme d'anneau qui est de plus injectif: en effet

$$\frac{a}{1} = 0_K = \frac{0}{1} \iff a = a.1 = 0.1 = 0.$$

On peut donc identifier a a la fraction $\frac{a}{1}$ et voir A comme un sous-anneau de K .

Soit $\iota' : A \mapsto K'$ un morphisme injectif dans un corps K' . Comme ι' est injectif, pour tout $b \in A - \{0\}$, $\iota'(b) \neq 0_{K'}$ et l'inverse $\iota'(b)^{-1} \in K' - \{0_{K'}\}$ existe.

On defini alors pour toute fraction $\frac{a}{b} \in \text{Frac}(A)$,

$$\iota'_K\left(\frac{a}{b}\right) := \iota'(a).\iota'(b)^{-1}.$$

On verifie alors que l'application

$$\iota'_K : \begin{array}{ccc} \text{Frac}(A) & \mapsto & K' \\ \frac{a}{b} & \mapsto & \iota'(a).\iota'(b)^{-1} \end{array}$$

est bien definie et est un morphisme non-nul de K vers K' et qu'il prolonge $\iota' : A \mapsto K'$. $\square$

NOTATION 4.2. *Dans la suite et pour alléger les notation on identifiera l'anneau A avec son image $\iota(A)$ dans son corps des fraction: ainsi pour $a \in A$ on écrira simplement " a " pour la fraction $\frac{a}{1_A} \in \text{Frac}(A)$.*

REMARQUE 4.2.3. La condition que ι' soit injective est vraiment necessaire (merci a Estelle de l'avoir remarque)

EXERCICE 4.2. Donner un exemple d'un anneau integre A et d'un morphisme d'anneau $\iota : A \mapsto K'$ non-nul et a valeurs dans un corps K' qui n'est pas injectif.

4.3. Construction de corps: corps quotient

Soit A un anneau commutatif. On a vu que etant donne un ideal I on peut fabriquer un autre anneau commutatif, l'anneau *quotient* dont les elements sont les *classes de congruence modulo I*

$$A/I = \{a \pmod I := a + I, a \in A\}$$

et les lois d'addition et de multiplications sont donnees par

$$a \pmod I + a' \pmod I = a + a' \pmod I, a \pmod I . a' \pmod I = a.a' \pmod I$$

et de plus l'application

$$\bullet \pmod I : a \in A \mapsto a \pmod I = a + I \in A/I$$

est un morphisme d'anneaux.

On va donner une condition necessaire et suffisante pour que A/I soit un corps.

DÉFINITION 4.4. *soit A un anneau commutatif. Un ideal $I \subset A$ est maximal si $I \neq A$ et si I est maximal pour l'inclusion parmi tous les ideaux de A distincts de A :*

$$\forall J \subset A, J \neq A \text{ ideal de } A, I \subset J \implies I = J.$$

REMARQUE 4.3.1. L'anneau nul $A = \{0_A\}$ n'admet pas d'ideal $\neq A$ et donc pas d'ideal maximal au sens precedent. Si A n'est pas l'anneau nul alors A admet toujours un ideal maximal (pour des anneaux generaux cela necessite l'axiome du choix).

THÉORÈME 4.3. *L'anneau commutatif A/I est un corps ssi I est un idéal maximal.*

Preuve: On va montrer que

$$I \text{ maximal} \implies A/I \text{ est un corps.}$$

Notons que comme $I \neq A$ on a que A/I n'est pas réduit à la seule classe $I = 0_{A/I}$ (si $a \in A - I$ alors $a \pmod I = a + I \neq I$) donc A/I contient au moins deux éléments distincts:

$$0_A \pmod I = I, 1_A \pmod I = 1_A + I$$

(repretons ce qu'on a dit ci-dessus : si on avait $1_A + I = I$ alors $1_A \in I$ et donc $I \supset \{a.1_A, a \in A\} = A$).

Soit $a \pmod I \in A/I - \{0_{A/I}\}$, on veut montrer que $a \pmod I$ est inversible c'est à dire qu'il existe $b \pmod I$ tel que

$$a \pmod I . b \pmod I = a.b \pmod I = 1_A \pmod I.$$

Cela equivaut à trouver $b \in A$ tel que

$$a.b - 1_A \in I.$$

Comme $a \pmod I \neq 0_A \pmod I = I$ alors $a \notin I$. Considerons l'idéal $J \subset A$ engendré par a et I :

$$J = \langle a, I \rangle_A = A.a + A.I = A.a + I$$

(l'ensemble $A.a + I$ contient a et I ; on vérifie que c'est un idéal de A et tout idéal de A contenant a et I doit contenir cet ensemble).

Comme $a \notin I$ on a $J \neq I$ mais évidemment $I \subset J$. Comme I est maximal et que $J \neq I$ cela implique que

$$J = A.a + I = A.$$

En particulier $1_A \in A.a + I$: il existe $b \in A$ et $i \in I$ tel que

$$1_A = b.a + i$$

et donc

$$a.b - 1_A = -i \in I.$$

La réciproque est laissée en exercice. □

REMARQUE 4.3.2. Voyons directement que $p\mathbb{Z} \subset \mathbb{Z}$ est maximal ssi p est premier. On a d'abord que

$$p\mathbb{Z} \neq \mathbb{Z} \iff p = 0 \text{ ou } p > 1.$$

L'idéal nul (le cas $p = 0$) n'est pas maximal (car contenu dans $2\mathbb{Z} \neq \mathbb{Z}$).

Si $p \geq 2$ est composé, $p = q_1 q_2$ avec $q_1, q_2 > 1$ alors $p\mathbb{Z} \subset q_1\mathbb{Z} \neq \mathbb{Z}$ et n'est donc pas maximal.

Si p est premier et si $p\mathbb{Z} \subset q\mathbb{Z}$ avec $q \geq 2$ alors p est un multiple de q et comme p est premier $p = q$ donc $p\mathbb{Z}$ est maximal.

DÉFINITION 4.5. On dit qu'un idéal $I \subset A$ est premier si $I \neq \{0_A\}$, A et si

$$\forall a, b \in A, a.b \in I \implies a \in I \text{ ou } b \in I.$$

EXERCICE 4.3. Montrer que

$$I \text{ est premier} \iff A/I \text{ est intègre.}$$

Comme un corps est intègre ou à que

$$\{0_A\} \neq I \text{ maximal} \implies I \text{ premier}.$$

4.4. Caractéristique d'un corps, Sous-corps premier

Soit K un corps alors on a vu qu'il existe un morphisme d'anneaux canonique

$$\text{Can}_K : \begin{array}{ccc} \mathbb{Z} & \mapsto & K \\ n & \mapsto & n.1_K = \pm(1_K + \dots + 1_K) \text{ } |n| \text{ fois} \end{array}$$

NOTATION 4.3. Soit K un corps et $n \in \mathbb{Z}$ un entier. On notera

$$n_K = \text{Can}_K(n) = n.1_K$$

l'image de n par le morphisme canonique.

Le noyau de ce morphisme est de la forme

$$\ker(\text{Can}_K) = p.\mathbb{Z}, \quad p \geq 0.$$

DÉFINITION 4.6. L'entier p s'appelle la caractéristique du corps K et se note

$$p =: \text{car}(K).$$

4.4.0.1. *Caractéristique nulle.* Si $\text{car}(K) = p = 0$ alors $\text{Can}_K.\mathbb{Z} \hookrightarrow K$ est injectif et K contient (un anneau isomorphe à) l'anneau $\mathbb{Z}$ et donc contient (un corps isomorphe à) le corps des fractions de $\mathbb{Z}$, le corps des nombres rationnels $\mathbb{Q}$: il existe une injection de corps

$$\iota_K : \mathbb{Q} \hookrightarrow K$$

obtenues en posant pour toute fraction rationnelle $\frac{a}{b} \in \mathbb{Q}$

$$\iota_K\left(\frac{a}{b}\right) = \text{Can}_K(a).\text{Can}_K(b)^{-1} \in K.$$

En effet comme $b \in \mathbb{Z} - \{0\}$ et que l'application Can_K est injective on a $\text{Can}_K(b) \in K - \{0_K\}$ est donc inversible dans K .

NOTATION 4.4. Pour simplifier les notations on identifiera $\mathbb{Q}$ avec son image $\iota_K(\mathbb{Q})$ dans le corps K et on écrira $\frac{a}{b} \in K$ pour l'image de la fraction correspondante $\iota_K(\frac{a}{b})$.

4.4.0.2. *Caractéristique strictement positive.* On a alors

LEMME 4.2. Si $\text{car}(K) > 0$ alors $\text{car}(K) = p$ est un nombre premier.

Preuve: Supposons que p n'est pas premier alors $p > 1$; sinon on aurait $\ker(\text{Can}_K) = 1.\mathbb{Z} = \mathbb{Z}$ et Can_K serait le morphisme nul mais ce n'est pas possible car $\text{Can}_K(1) = 1_K \neq 0_K$.

On a alors $p = q_1.q_2$ avec $2 \leq q_1, q_2 < p$ et on a

$$p_K = 0_K = q_{1K}.q_{2K}$$

et donc ou bien $q_{1K} = 0$ ou bien $q_{2K} = 0$ (car un corps est intègre). Cela signifie que q_1 ou bien q_2 appartient à $\ker(\text{Can}_K) = p.\mathbb{Z}$ mais cela contredit le fait que p est le plus petit entier strictement positif contenu dans $\ker(\text{Can}_K)$. $\square$

Considérons alors l'image $\text{Can}_K(\mathbb{Z}) = \mathbb{Z}.1_K$, c'est un sous-anneau de K .

LEMME 4.3. L'anneau $\text{Can}_K(\mathbb{Z}) = \mathbb{Z}.1_K$ est un corps fini de cardinal p isomorphe au corps $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.

Preuve: Notons que pour tout $n, k \in \mathbb{Z}$ on a

$$\text{Can}_K(n + p.k) = \text{Can}_K(n) + \text{Can}_K(p.k) = \text{Can}_K(n)$$

car $p.k \in \ker(\text{Can}_K)$. Ainsi, la valeur de $\text{Can}_K(n)$ ne dépend que de la classe de congruence $n \pmod{p}$. On peut donc définir une application

$$\iota_K : \begin{array}{ccc} \mathbb{Z}/p\mathbb{Z} & \mapsto & \text{Can}_K(\mathbb{Z}) \\ n \pmod{p} & \mapsto & \text{Can}_K(n) \end{array}$$

Comme l'application

$$n \in \mathbb{Z} \mapsto n \pmod{p} \in \mathbb{Z}/p\mathbb{Z}$$

est un morphisme d'anneaux d'image $\text{Can}_K(\mathbb{Z})$, on en deduit que ι_K est un morphisme d'anneaux non-nul et comme $\mathbb{Z}/p\mathbb{Z}$ est un corps, ce morphisme est injectif: ι_K est un isomorphisme de $\mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$ sur son image $\text{Can}_K(\mathbb{Z})$. □

NOTATION 4.5. *Pour simplifier les notations on identifiera $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ avec l'image $\text{Can}_K(\mathbb{Z}) \subset K$ de $\mathbb{Z}$ dans K par le morphisme canonique. Ainsi on écrira*

$$\text{Can}_K(\mathbb{Z}) = \mathbb{Z}.1_K = \mathbb{F}_p$$

et pour $n \in \mathbb{Z}$ on écrira indifféremment

$$n_K = n.1_K = n \pmod{p}$$

qu'on verra comme un élément de K .

DÉFINITION 4.7. *Le corps $\mathbb{Q} \subset K$ (si $\text{car}(K) = 0$) ou bien $\mathbb{F}_p \subset K$ (si $\text{car}(K) = p > 0$) s'appelle le sous-corps premier de K .*

REMARQUE 4.4.1. On peut montrer (exercice) que si K contient un sous-corps K' isomorphe soit à $\mathbb{Q}$ soit à $\mathbb{F}_p$ pour p premier alors K' est le sous-corps premier de K .

4.4.1. Arithmétique des corps de caractéristique positive: le Frobenius.

PROPOSITION 4.4. *Soit K un corps de caractéristique $p > 0$ alors l'application*

$$\bullet^p : \begin{array}{ccc} K & \mapsto & K \\ x & \mapsto & x^p \end{array}$$

est un morphisme d'anneaux non-nul (donc nécessairement injectif).

Preuve: Comme K est un anneau commutatif, on a pour tout $x, y \in K$

$$(x.y)^p = (x.y) \cdots (x.y) = x^p.y^p.$$

Montrons que

$$(x + y)^p = x^p + y^p.$$

Par la formule du binôme de Newton, on a (à nouveau parce que K est commutatif)

$$(x + y)^p = \sum_{k=0}^p C_p^k x^k . y^{p-k} = x^p + y^p + \sum_{k=1}^{p-1} C_p^k x^k . y^{p-k}$$

avec

$$C_p^k = \frac{p!}{k!(p-k)!} = \frac{p.(p-1).\cdots.(p-k+1)}{k.(k-1).\cdots.2.1} \in \mathbb{N}$$

(on rappelle que C_p^k est le nombre de sous-ensembles de k éléments dans un ensemble de p éléments).

LEMME 4.4. *Soit p un nombre premier et $1 \leq k \leq p-1$ alors C_p^k est divisible par p : il existe $c_{p,k} \in \mathbb{N}$ tel que $C_p^k = p.c_{p,k}$. En particulier $C_p^k_K = 0_K$.*

Preuve: On a

$$C_p^k = p \cdot \frac{(p-1).\cdots.(p-k+1)}{k.(k-1).\cdots.2.1} = p.c_{p,k}$$

avec $c_{p,k}$ a priori un nombre rationnel. On sait que $1.2.\cdots.k$ divise $p.(p-1).\cdots.(p-k+1)$ (car C_p^k est un entier). Comme p est un nombre premier $k! = k.(k-1).\cdots.2.1$ est premier avec p (car tout diviseur premier de $k!$ est $< p$) et comme $k!$ divise $p.(p-1).\cdots.(p-k+1)$, il doit diviser $(p-1).\cdots.(p-k+1)$ et $c_{p,k}$ est premier. □

On a alors

$$(x + y)^p = x^p + y^p + \sum_{k=1}^{p-1} C_p^k \cdot 1_K \cdot x^k \cdot y^{p-k} = x^p + y^p$$

car pour $1 \leq k \leq p-1$,

$$C_p^k \cdot 1_K = c_{p,k} \cdot (p \cdot 1_K) = 0_K.$$

Ainsi $x \mapsto x^p$ est un morphisme d'anneau et comme $1_K^p = 1_K \neq 0_K$ ce morphisme est non-nul. $\square$

DÉFINITION 4.8. Soit K un corps de caractéristique p , le morphisme d'anneau precedent s'appelle le morphisme de Frobenius (ou simplement le Frobenius) de K se note

$$\text{frob}_p : x \in K \mapsto x^p \in K.$$

THÉORÈME 4.4 (Petit Theoreme de Fermat). Soit K un corps de caractéristique positive p et $\text{frob}_p : K \mapsto K$ le Frobenius. Pour tout $x \in \mathbb{F}_p = \mathbb{Z} \cdot 1_K$ on a

$$\text{frob}_p(x) = x^p = x.$$

Recapitulatif concernant la caractéristique d'un corps

Si K est un corps et ≥ 0 sa caractéristique, ie.

$$\ker(\text{Can}_K) = \{n \in \mathbb{Z}, n \cdot 1_K = 0_K\} = p\mathbb{Z}.$$

Si $p = 0$. Alors $\text{Can}_K(\mathbb{Z}) = \{n_K = n \cdot 1_K, n \in \mathbb{Z}\}$ est un sous-anneau isomorphe a $\mathbb{Z}$ et K contient le corps $\mathbb{Q}$ comme sous-corps via le morphisme

$$\bullet_K : \frac{a}{b} \in \mathbb{Q} \mapsto \left(\frac{a}{b}\right)_K := a_K \cdot b_K^{-1} \in K.$$

De plus tout sous-corps $K' \subset K$ isomorphe a $\mathbb{Q}$ est egal a $\mathbb{Q}_K$ et K ne contient aucun sous-corps isomorphe a $\mathbb{F}_p$ pour p premier.

On identifiera $\mathbb{Q}$ avec son image dans K et écrira simplement $\frac{a}{b}$ pour l'image de la fraction $\left(\frac{a}{b}\right)_K = a_K \cdot b_K^{-1}$.

Si $p > 0$. Alors p est premier et

$$\text{Can}_K(\mathbb{Z}) = \{n_K = n \cdot 1_K, n \in \mathbb{Z}\} = \mathbb{Z} \cdot 1_K$$

est (isomorphe au) le corps $\mathbb{F}_p$ a p elements.

De plus si K contient un sous-corps $K' \subset K$ isomorphe a $\mathbb{F}_p$ alors

$$K' = \text{Can}_K(\mathbb{Z}).$$

Enfin K ne contient aucun sous-corps isomorphe a $\mathbb{Q}$ ou a $\mathbb{F}_q$ pour $q \neq p$ premier.

On identifiera $\mathbb{F}_p$ avec le sous-corps de K qui lui est isomorphe $\text{Can}_K(\mathbb{Z}) = \mathbb{Z} \cdot 1_K$ et pour tout $n \in \mathbb{Z}$ on écrira indifferement

$$n_K 0n \cdot 1_K = n \pmod{p}.$$

On a alors

$$n_K = n \cdot 1_K = 0_K \iff n \in p\mathbb{Z}$$

et plus generalement pour tout $x \in K - \{0_K\}$ on a

$$n \cdot x = n \cdot 1_K \cdot x = n_K \cdot x = 0_K \iff n \in p\mathbb{Z}.$$

De plus on a pour tout $x, y \in K$

$$(x + y)^p = x^p + y^p.$$

Enfin (exercice) par le petit Theorem de Fermat pour tout $x \in \mathbb{F}_p \subset K$, on a

$$x^p = x$$

et reciproquement si $x \in K$ verifie $x^p = x$ alors $x \in \mathbb{F}_p$.

CHAPITRE 5

Modules et Espaces Vectoriels

*“An attempt at visualizing the Fourth Dimension:
Take a point, stretch it into a line,
curl it into a circle, twist it into a sphere,
and punch through the sphere.”*

5.1. Module sur un anneau

DÉFINITION 5.1. Soit $(A, +, \cdot)$ un anneau, un A -module (à gauche) est un groupe commutatif $(M, +)$ muni d'une loi de multiplication externe

$$\bullet * \bullet : \begin{array}{ccc} A \times M & \mapsto & M \\ (a, m) & \mapsto & a * m \end{array}$$

(appelée multiplication par les scalaires) ayant les propriétés suivantes:

(1) Associativité: $\forall a, a' \in A, m \in M,$

$$(a \cdot a') * m = a * (a' * m).$$

(2) Distributivité: $\forall a, a' \in A, m, m' \in M,$

$$(a + a') * m = a * m + a' * m, \quad a * (m + m') = a * m + a * m'.$$

(3) Neutralité de 1_A : $\forall m \in M,$

$$1_A * m = m.$$

REMARQUE 5.1.1. On définit de manière analogue la notion de A -module à droite à partir d'une multiplication externe "à droite"

$$\bullet *_d \bullet : \begin{array}{ccc} M \times A & \mapsto & M \\ (m, a) & \mapsto & m *_d a \end{array}$$

vérifiant des propriétés analogues notamment l'associativité

$$\forall a, a' \in A, m \in M, \quad m *_d (a \cdot a') = (m *_d a) *_d a'.$$

EXEMPLE 5.1.1. Quelques exemples de modules sur des anneaux:

- (1) Un anneau A est un A -module sur lui-même pour la multiplication.
- (2) Le singleton élément neutre $\{0_A\}$ est un A -module: le module nul.
- (3) Soit $I \subset A$ un idéal d'un anneau A alors I est un A -module pour la multiplication de A .
- (4) Soit $d \geq 1$, le produit cartésien

$$A^d = A \times \cdots \times A = \{(a_1, \dots, a_d), a_i \in A, i = 1, \dots, d\}$$

est un A -module avec la loi de groupes

$$(a_1, \dots, a_d) + (a'_1, \dots, a'_d) = (a_1 + a'_1, \dots, a_d + a'_d)$$

et la multiplication par les scalaires

$$a \cdot (a_1, \dots, a_d) = (a \cdot a_1, \dots, a \cdot a_d).$$

On dit que A^d est un A -module libre de rang d .

- (5) Soit M un groupe abelien alors M est naturellement un $\mathbb{Z}$ -module pour la loi de multiplication par les scalaires donnee par

$$n.m = \begin{cases} 0_M & \text{si } n = 0 \\ m + m + \cdots + m & (n \text{ fois si } n \geq 1), \\ (-m) + (-m) + \cdots + (-m) & (-n \text{ fois si } n \leq -1) \end{cases}.$$

EXERCICE 5.1. Soit M un A -module, alors M est egalement un $\mathbb{Z}$ -module. Montrer que pour tout $n \in \mathbb{Z}$, on a

$$(n_A) * m = n.m$$

(on rappelle qu'on a note $n_A := \text{Can}_A(n)$) En particulier

$$(-1_A).m = -m.$$

- (6) Soit $\varphi : A \mapsto B$ un morphisme d'anneaux alors $\ker(\varphi) \subset A$ est un A -module pour la multiplication dans A (car $A.\ker \varphi \subset \ker \varphi$). Par ailleurs l'anneau d'arrivee B a une structure de A -module en definissant comme multiplication externe:

$$a.\varphi b := \varphi(a).Bb.$$

- (7) Soit A un anneau, X un ensemble et $\mathcal{F}(X; A)$ l'ensemble des fonction de X a valeurs dans A . On a vu que $\mathcal{F}(X; A)$ a une structure d'anneau; il a egalement une structure de A -module: on definit la multiplication externe d'un element $a \in A$ et d'une fonction $f : X \mapsto A$ par

$$a.f : x \mapsto (a.f)(x) = a.(f(x)).$$

- (8) Soit A un anneau commutatif et $A[X]$ l'anneau des polynomes alors $A[X]$ est naturellement un A -module pour la multiplication d'un polynome par un scalaire: si $P(X) = a_0 + \cdots + a_d.X^d$ alors la multiplication par les scalaires est donnee par

$$a.P(X) = a.a_0 + a.a_1.X + \cdots + a.a_d.X^d.$$

- (9) Soit A un anneau commutatif et

$$A[X]_{\leq d} = \{a_0 + \cdots + a_d.X^d, a_0, \cdots, a_d \in A\}$$

l'anneau des polynomes de degre $\leq d$ alors $A[X]_{\leq d}$ est naturellement un A -module (par contre ce n'est pas un anneau –sauf si $d = 0$: les polynomes constants c'est a dire l'anneau A – car $A[X]_{\leq d}$ n'est pas stable par produit en general).

- (10) Soit A un anneau commutatif et $M_2(A)$ l'anneau des matrice 2×2 a coefficients dans A alors $M_2(A)$ a une structure de A -module en definissant la multiplication par les scalaires par

$$a. \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} := \begin{pmatrix} a.a' & a.b' \\ a.c' & a.d' \end{pmatrix}.$$

Les exemples (7) (si A est commutatif), (8) et (10) sont des cas particuliers de ce qu'on appelle une A -algebre:

DÉFINITION 5.2. Soit A un anneau commutatif. Une A -algebre est un anneau $(B, +_B, \cdot_B)$ possedant une structure de A -module qui verifie la propriete d'associativite suivante pour les deux multiplications:

$$\forall a \in A, b, b' \in B \quad a * (b.Bb') = (a * b).Bb' = b.B(a * b').$$

5.1.1. Sous-module.

DÉFINITION 5.3. Soit M un A -module. Un sous-module $N \subset M$ d'un A -module M est un sous-groupe de $(M, +)$ qui est stable pour la multiplication par les scalaires:

$$\forall a \in A, n \in N, a * n \in N.$$

On a donc $\forall n, n' \in N, a, a' \in A$

$$a * n + a' * n' \in N$$

On a le critere suivant

PROPOSITION 5.1. (Critere de sous-module) Soit $N \subset M$ un sous-ensemble d'un A -module M alors N est un sous-module de M ssi

$$(5.1.1) \quad \forall a \in A, n, n' \in N, a * n + n' \in N.$$

Preuve: Pour tout $n, n' \in N$, et appliquant la condition (5.1.1) a n, n' et $a = -1_A$ on a

$$n + (-1_A) * n' = n - n' \in N$$

donc N verifie le critere de sous-groupe et est donc un sous-groupe de $(M, +)$. Il contient en particulier 0_M et alors pour tout $a \in A$, on a par (5.1.1)

$$a * n + 0_M = a * n \in N.$$

□

EXEMPLE 5.1.2. Exemples de sous-modules

- (1) L'element nul $\{0_M\}$ forme un sous-module de M : le sous-module nul.
- (2) Soit $m \in M$, on note $A.m = \{a.m, a \in A\} \subset M$, alors $A.m$ est un sous-module de A . Soient $m' \in M$, alors

$$A.m + A.m' = \{a.m + a'.m', a, a' \in A\}$$

est un sous-module de M .

- (3) Par exemple, soit A^d le module libre de rank d et

$$\Delta A = \{(a, a \cdots, a) = a.(1, 1, \cdots, 1), a \in A\} \subset A^d$$

est un sous-module de A^d . Plus generalement pour tout $\vec{a} = (a_1, \cdots, a_d) \in A^d$ le sous-ensemble des multiples de $\vec{a}$

$$A.\vec{a} = \{a.\vec{a} = (a.a_1, \cdots, a.a_d), a \in A\}$$

est un sous-module de A^d .

- (4) Soit $1 \leq d \leq d'$ alors

$$A[X]_{\leq d} \subset A[X]_{\leq d'} \subset A[X]$$

est un chaine de sous A -modules.

5.1.2. Module engendre par un ensemble.

PROPOSITION 5.2. Soit $(M, +, *)$ un A -module et M_1, M_2 des sous-modules alors

$$M_1 \cap M_2 \subset M$$

est un sous-module et plus generalement soit $(M_i)_{i \in I}$ une collection de sous-modules alors

$$\bigcap_{i \in I} M_i \subset M$$

est un sous-module.

DÉFINITION 5.4. Soit $X \subset M$ un sous-ensemble d'un A -module, le module engendré par X est le plus petit sous-module de M contenant X (l'intersection de tous les sous-modules contenant X):

$$\langle X \rangle_A := \bigcap_{\substack{X \subset N \subset M \\ N \text{ } A\text{-mod}}} N.$$

REMARQUE 5.1.2. Si $(M, +)$ est un groupe commutatif alors on a vu que c'est naturellement un $\mathbb{Z}$ -module et si $X \subset M$ est un sous-ensemble, le sous-groupe engendré par X $\langle X \rangle \subset M$ est exactement le $\mathbb{Z}$ -module $\langle X \rangle_{\mathbb{Z}}$ engendré par X dans M . Il n'y a donc pas de collision au niveau des notations¹.

PROPOSITION 5.3. Soit $X \subset M$ un ensemble alors $\langle X \rangle_A$ est soit le module nul $\{0_M\}$ si X est vide, soit l'ensemble des combinaisons linéaires d'éléments de X à coefficients dans A :

$$\langle X \rangle_A = \text{CL}_A(X) := \left\{ \sum_{i=1}^n a_i * x_i, n \geq 1, a_1, \dots, a_n \in A, x_1, \dots, x_n \in X \right\}.$$

Preuve: On suppose X non-vide. Soit $X \subset N$ un sous-module contenant X alors pour tout $n \geq 1$, tous $a_1, \dots, a_n \in A$ et tout $x_1, \dots, x_n \in X$ on a

$$a_1 * x_1 + \dots + a_n * x_n \in N$$

par stabilité de N par $+$ et $*$. Donc tout sous-module N contenant X contient $\text{CL}_A(X)$.

Il reste à montrer que $\text{CL}_A(X)$ est un sous-module: soient u et u' des combinaisons linéaires d'éléments de X :

$$u = a_1 * x_1 + \dots + a_n * x_n, u' = a'_1 * x'_1 + \dots + a'_{n'} * x'_{n'}$$

alors

$$u + u' = a_1 * x_1 + \dots + a_n * x_n + a'_1 * x'_1 + \dots + a'_{n'} * x'_{n'}$$

est bien une combinaison linéaire. De plus $\text{CL}_A(X)$ est stable par multiplication par A : pour tout $a \in A$ on a par distributivité et associativité

$$a * u = a * (a_1 * x_1 + \dots + a_n * x_n) = (a.a_1) * x_1 + \dots + (a.a_n) * x_n$$

est bien une combinaison linéaire. □

DÉFINITION 5.5. Si $\langle X \rangle_A = M$, on dit que X est une famille génératrice de M .

DÉFINITION 5.6. Un A -module M est de type fini si il possède une famille génératrice qui est finie.

EXEMPLE 5.1.3. (1) Soit A^d le A -module libre de rang d . La famille suivante est génératrice de A^d (on pose $1 = 1_A, 0 = 0_A$)

$$\mathcal{B}^0 := \{\mathbf{e}_1^0 = (1, 0, \dots, 0), \mathbf{e}_2^0 = (0, 1, 0, \dots, 0), \dots, \mathbf{e}_d^0 = (0, 0, \dots, 1)\}$$

($\mathbf{e}_i^0$ est le d -uplet dont toutes les coordonnées sont nulles sauf la i -ième qui vaut 1). En effet si

$$m = (a_1, \dots, a_d) \in A^d$$

alors

$$m = a_1 \cdot \mathbf{e}_1^0 + \dots + a_d \cdot \mathbf{e}_d^0.$$

On appelle la famille $\mathcal{B}^0$ la base canonique de A^d .

(2) La famille des monômes

$$\{1, X, \dots, X^d, \dots, X^{d+1}, \dots\}$$

est une famille génératrice (infinie) de $A[X]$.

¹Merci à l'étudiante qui a fait cette observation.

(3) La famille des monomes de degre $\leq d$

$$\{1, X, \dots, X^d\}$$

est une famille generatrice de $A[X]_{\leq d}$ (qui est donc un module de type fini)

EXERCICE 5.2. Soient $u_1, \dots, u_d \in A^\times$ des elements inversibles. Montrer que la famille suivante est generatrice de A^d

$$\mathcal{B} := \{\mathbf{e}_1 = (u_1, 0, \dots, 0), \mathbf{e}_2 = (0, u_2, 0, \dots, 0), \dots, \mathbf{e}_d = (0, 0, \dots, u_d)\}.$$

Montrer que l'ecriture d'un element de A^d comme combinaison lineaire des elements de $\mathcal{B}$ est unique.

EXERCICE 5.3. Soient $a, b, c, d \in \mathbb{Z}$ tels que $ad - bc = \pm 1$. Montrer que $\{(a, b), (c, d)\}$ engendre le $\mathbb{Z}$ -module $\mathbb{Z}^2$. Pour cela on montrera que pour tout $(m, n) \in \mathbb{Z}^2$ le systeme lineaire

$$\begin{cases} ax + cy = m \\ bx + dy = n \end{cases}$$

admet une (unique) solution $(x, y) \in \mathbb{Z}^2$ et on montrera que (m, n) s'exprime en fonction de (a, b) et (c, d) .

5.1.3. Morphismes de modules.

DÉFINITION 5.7. Soit A un anneau et M, N des A -modules, un morphisme de A -modules entre M et N est un morphisme de groupes

$$\varphi : M \mapsto N$$

qui est compatible avec les lois de multiplications externes $*_M$ et $*_N$:

$$\forall a \in A, m \in M, \varphi(a *_M m) = a *_N \varphi(m).$$

REMARQUE 5.1.3. Cette definition implique que pour tout $a, a' \in A, m, m' \in M$, on a

$$\varphi(a *_M m + a' *_M m') = a *_N \varphi(m) + a' *_N \varphi(m').$$

Plus generalement pour I un ensemble fini, $(a_i)_{i \in I}$ un I -uplet de scalaires et $(m_i)_{i \in I}$ un I -uplet d'elements de M on a

$$\varphi\left(\sum_{i \in I} a_i *_M m_i\right) = \sum_{i \in I} a_i *_N \varphi(m_i).$$

En d'autres termes, l'image par φ d'une combinaison lineaire est la combinaison lineaire des images.

On dit que φ est une *application A -lineaire*.

LEMME 5.1. (*Critere d'application lineaire*) Soit $\varphi : M \mapsto N$ une application entre deux A -modules alors φ est un morphisme (ie. est A -lineaire) si et seulement si

$$(5.1.2) \quad \forall a \in A, m, m' \in M, \varphi(a *_M m + m') = a *_N \varphi(m) + \varphi(m').$$

Preuve: On applique (5.1.2) avec $a = 1_A$. On a donc

$$\forall m, m' \in M, \varphi(m + m') = \varphi(m) + \varphi(m')$$

donc φ est un morphisme de groupes. On a donc $\varphi(0_M) = 0_N$ et

$$\varphi(a *_M m) = \varphi(a *_M m + 0_M) = a *_N \varphi(m) + 0_N = a *_N \varphi(m).$$

□

5.1.4. Noyau, Image.

PROPOSITION 5.4. Soit $\varphi : M \mapsto N$ un morphisme de A -modules et $M' \subset M$ et $N' \subset N$ des sous-modules alors

$$\varphi(M') \subset N \text{ et } \varphi^{(-1)}(N') \subset M$$

sont des sous-modules de M et N respectivement. En particulier

$$\ker(\varphi) = \varphi^{(-1)}(\{0_N\}) \subset M \text{ et } \operatorname{Im}(\varphi) = \varphi(M) \subset N$$

sont des sous A -modules.

Preuve: Exercice. □

Comme un morphisme de A -module est un morphisme de groupes additifs on a

COROLLAIRE 5.1. L'application A -linéaire $\varphi : M \mapsto M'$ est injective ssi $\ker(\varphi) = \{0_M\}$.

5.1.5. Structure des espaces de morphismes. On a les propriétés de stabilité usuelles pour la composition (similaires à celles pour les morphismes de groupes)

PROPOSITION 5.5. Soient $\varphi : L \mapsto M$ et $\psi : M \mapsto N$ des morphismes de A -modules alors

- $\psi \circ \varphi : L \mapsto N$ est un morphisme de A -modules.
- Si $\varphi : L \mapsto M$ est bijectif alors $\varphi^{-1} : M \mapsto L$ est un morphisme de A -modules.

Preuve: Exercice. □

NOTATION 5.1. On note

$$\operatorname{Hom}_{A\text{-mod}}(M, N), \operatorname{Isom}_{A\text{-mod}}(M, N),$$

$$\operatorname{End}_{A\text{-mod}}(M) = \operatorname{Hom}_{A\text{-mod}}(M, M),$$

$$\operatorname{Aut}_{A\text{-mod}}(M) = \operatorname{Isom}_{A\text{-mod}}(M, M)$$

les ensembles de morphismes, morphismes bijectifs (ou isomorphismes), d'endomorphismes et d'automorphismes des A -modules M et N . On note quelquefois

En particulier on a

COROLLAIRE 5.2. L'ensemble des automorphismes de M , $\operatorname{Aut}_{A\text{-mod}}(M) \subset \operatorname{Bij}(M)$ est un sous-groupe de $\operatorname{Bij}(M)$. Plus précisément $\operatorname{Aut}_{A\text{-mod}}(M)$ est un sous-groupe de $\operatorname{Aut}_{Gr}(M)$. On note également ce groupe (surtout dans le cas où A est un corps)

$$\operatorname{Aut}_{A\text{-mod}}(M) = \operatorname{GL}(M)$$

et on l'appelle le groupe linéaire du A -module M .

On a une propriété supplémentaire de stabilité par somme:

PROPOSITION 5.6. Soient M et N des A -modules alors $\operatorname{Hom}_{A\text{-mod}}(M, N)$ a une structure naturelle de groupe commutatif. Si de plus A est commutatif alors $\operatorname{Hom}_{A\text{-mod}}(M, N)$ a une structure naturelle de A -module.

Preuve: Soient $\varphi, \psi \in \operatorname{Hom}_{A\text{-mod}}(M, N)$, on définit l'addition par

$$\varphi + \psi : m \mapsto (\varphi + \psi)(m) = \varphi(m) + \psi(m) \in N.$$

C'est un morphisme de A -module car N est un A -module:

$$\begin{aligned} (\varphi + \psi)(a * m + m') &= \varphi(a * m + m') + \psi(a * m + m') \\ &= a * \varphi(m) + \varphi(m') + a * \psi(m) + \psi(m') = a * (\varphi + \psi)(m) + (\varphi + \psi)(m'). \end{aligned}$$

et on définit l'opposé $-\varphi$ en posant

$$-\varphi(m) = -(\varphi(m)) \in N$$

et on verifie a nouveau que $-\varphi$ est A -lineaire. L'element neutre est le morphisme nul:

$$\underline{0}_N : m \in M \mapsto 0_N$$

et c'est une application A -lineaire:

$$\forall a \in A, m \in M, \underline{0}_N(a * m) = 0_N = (a * \underline{0}_N)(m).$$

Supposons que A soit commutatif: on definit la multiplication par les scalaires en posant pour $a \in A$

$$a * \varphi : m \mapsto (a * \varphi)(m) := a *_N \varphi(m).$$

L'application $a * \varphi$ est bien un morphisme de A -modules: pour $a' \in A$, on a (par linearite, distributivite et associativite)

$$\begin{aligned} (a * \varphi)(a' *_M m + m') &= a *_N (\varphi(a' *_M m + m')) = a *_N (a' *_N \varphi(m) + \varphi(m')) \\ &= (a.a') *_N \varphi(m) + a *_N \varphi(m') = (a'.a) *_N \varphi(m) + a *_N \varphi(m') = a' *_N a *_N \varphi(m) + (a * \varphi)(m') \\ &= a' *_N (a * \varphi)(m) + (a * \varphi)(m'). \end{aligned}$$

Ici on a utilise de maniere cruciale le fait que A est commutatif et donc $a.a' = a'.a$. $\square$

5.1.6. L'algebre des endomorphismes d'un module. On a vu que l'ensemble des endomorphisme du groupe additif $\text{End}_{Gr}(M)$ muni de la composition et de l'addition est un anneau. Pour les morphismes de A -modules, on a un peu plus. Pour cela nous auront besoin de la definition de A -algebre:

DÉFINITION 5.8. *soit A un anneau commutatif. Une A -algebre associative est un anneau $(B, +, \cdot_B)$ muni d'une structure de A -module, note $*$: $A \times B \mapsto B$ verifiant en plus des axiomes habituels*

– *Distributivite par rapport a la multiplication:*

$$\forall a \in A, b, b' \in B, a * (b \cdot_B b') = (a * b) \cdot_B b' = b \cdot_B (a * b').$$

REMARQUE 5.1.4. Il existe une version plus generale d'algebre qui ne necessite pas que B soit un anneau (en particulier qui ne necessite pas que la multiplication dans B soit associative ni qu'elle possede une unite) mais nous n'en auront pas besoin ici.

EXEMPLE 5.1.4. (1) Les exemples (7) (si A est commutatif), (8) et (10) sont des exemples de A algebres.

(2) Soit B est un anneau et $A \subset B$ est un sous-anneau dont les elements commutent multiplicativement avec tous les elements de B ($\forall a \in A, b \in B, a.b = b.a$) alors B est une A -algebre pour la multiplication dans B .

THÉORÈME 5.1. *Soit M un A -module. L'ensemble $\text{End}_{A-mod}(M)$ des endomorphismes de M comme A -module est un sous-anneau de $(\text{End}_{Gr}(M), +, \circ)$ dont le groupe des unites est le groupe des automorphismes*

$$\text{End}_{A-mod}^\times(M) = \text{Aut}_{A-mod}(M) = \text{GL}_{A-mod}(M).$$

$\text{End}_{A-mod}(M)$ est l'anneau des endomorphismes de (du A -module) M .

De plus, si A est commutatif, $\text{End}_{A-mod}(M)$ possede une structure naturelle de A -module qui en fait une A -algebre et $\text{End}_{A-mod}(M)$ est appelee

Algebre des endomorphismes de (du A -module) M .

Preuve: D'abord Id_M et l'application constante nulle $\underline{0}_M$ qui sont des morphismes de groupes sont egalement des morphismes de A -modules:

$$\forall a \in A, m \in M, \text{Id}_M(a * m) = a * m = a * \text{Id}_M(m), \underline{0}_M(a * m) = 0_M = a * \underline{0}_M(m).$$

On a vu que $\text{End}_{A\text{-mod}}(M)$ est stable par composition et on a vu que la somme de deux endomorphismes est encore un endomorphisme de A -module. Ainsi $\text{End}_{A\text{-mod}}(M)$ est un sous-anneau de $\text{End}_{Gr}(M)$.

Si A est commutatif on a vu que $\text{End}_{A\text{-mod}}(M) = \text{Hom}_{A\text{-mod}}(M, M)$ possède une multiplication par les scalaires qui en fait un A -module ce qui fait de cet anneau une A -algebre: en effet pour tout $\varphi, \psi \in \text{End}_{A\text{-mod}}(M)$ et $a \in A$, on a pour $m \in M$

$$a *_M (\varphi \circ \psi)(m) = a *_M \varphi(\psi(m)) = (a *_M \varphi)(\psi(m)) = ((a *_M \varphi) \circ \psi)(m).$$

De plus on a (par A -linearite de φ)

$$a *_M (\varphi \circ \psi)(m) = a *_M \varphi(\psi(m)) = \varphi(a *_M \psi(m)) = \varphi((a *_M \psi)(m)) = \varphi \circ (a *_M \psi)(m)$$

de sorte que

$$a *_M (\varphi \circ \psi) = (a *_M \varphi) \circ \psi = \varphi \circ (a *_M \psi).$$

□

5.2. Espaces vectoriel

Tout comme les corps sont des cas particuliers d'anneaux, les espaces vectoriels sont des cas particuliers de modules: ce sont les modules dont *l'anneau associe est un corps*. Comme on va le voir les proprietes d'un module sur un corps sont tellement particuliere que cela justifie un changement de terminologie.

DÉFINITION 5.9. *Soit K un corps, un K -espace vectoriel (K -ev) V est simplement un K -module. Les elements de V sont appeles vecteurs de V . Les elements de K sont appeles les scalaires.*

EXEMPLE 5.2.1. Exemples d'espaces vectoriels:

- (1) L'espace vectoriel nul $\{0_K\}$.
- (2) K est un espace vectoriel sur lui-meme.
- (3) Si V et W sont des K -ev leur produit

$$V \times W = \{(v, w), v \in V, w \in W\}$$

muni de l'addition (composante par composante)

$$(v, w) + (v', w') := (v +_V v', w +_W w')$$

et de la mutliplication externe (composante par composante)

$$x.(v, w) := (x.v, x.w)$$

a une structure d'EV dont le vecteur nul est

$$0_{V \times W} = (0_V, 0_W).$$

- (4) En particulier, pour $d \geq 1$, en iterant la construction precedente pour $W = K$ on forme le K -module libre de rank d ,

$$K^d = \{(x_1, \dots, x_d), x_i \in K\}$$

dont l'element neutre est le vecteur nul

$$0_d = (0, \dots, 0).$$

- (5) Si X est un ensemble,

$$\mathcal{F}(X; K) = K^X = \{f : X \mapsto K\}$$

a une structure de K -espace vectoriel.

- (6) Plus generalement si V est un K -espace vectoriel et X est un ensemble,

$$\mathcal{F}(X; V) = V^X = \{f : X \mapsto V\}$$

a une structure de K -espace vectoriel.

NOTATION 5.2. *Pour alléger les notation on notera la multiplication par les scalaires sous la forme d'un point . (le meme point . que pour la multiplication dans le corps K) : pour $\lambda \in K$, $\vec{v} \in V$ on écrira $\lambda.\vec{v}$.*

Les differentes structures associees aux modules sur un anneau ont un nouveau nom quand l'anneau est un corps.

5.2.1. Sous-espace vectoriel.

DÉFINITION 5.10. *Soit V un K -espace vectoriel, un sous-espace vectoriel (SEV) de V est un sous- K module $W \subset V$.*

PROPOSITION 5.7 (Critere de SEV). *Un sous-ensemble $U \subset V$ d'un K -ev est un SEV ssi*

$$\forall \lambda \in K, v, v' \in U, \lambda.v + v' \in U.$$

Preuve: C'est un cas particulier du critere de sous-module. □

EXEMPLE 5.2.2. Exemples de SEV:

- $\{0_V\}, V \subset V$.
- Pour $\mathbf{e} \in V$, $K.\mathbf{e} = \{x.\mathbf{e}, x \in K\}$.
- Si $V' \subset V$ et $W' \subset W$ sont des SEV, $V' \times W'$ en est un.
- $\{(x_1, \dots, x_d) \in K^d, x_1 + \dots + x_d = 0\} \subset K^d$.
- $\{(x_1, \dots, x_d) \in K^d, x_1 + \dots + x_d = 1\} \subset K^d$ n'est pas un SEV.
- Soit $x_0 \in X$, dans $\mathcal{F}(X, V)$ le sous-espaces des fonctions f telles que $f(x_0) = 0_V$.
- Dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$ l'ensemble des fonctions paires (resp. impaires).

$$f : \mathbb{R} \mapsto \mathbb{R}, \forall x \in \mathbb{R}, f(x) = f(-x) \text{ (resp. } f(x) = -f(-x))$$

sont des SEVs.

5.2.2. Applications lineaires.

DÉFINITION 5.11. *Soient V et W deux K -espaces vectoriels; un morphisme $\varphi : V \mapsto W$ de K -modules est appele une application K -lineaire.*

PROPOSITION 5.8 (Critere d'application lineaire). *Une application entre espaces vectoriels $\varphi : V \mapsto W$ est lineaire ssi*

$$\forall \lambda \in K, v, v' \in V, \varphi(\lambda.v + v') = \lambda.\varphi(v) + \varphi(v').$$

Preuve: C'est un cas particulier du critere de morphisme de modules. □

PROPOSITION 5.9. *Si $\varphi : V \mapsto W$ est une application lineaire, le noyau*

$$\ker \varphi = \{v \in V, \varphi(v) = 0_W\} \subset V$$

et l'image

$$\text{Im } \varphi := \{\varphi(v), v \in V\} \subset W$$

sont des sous-espaces vectoriels de V et de W respectivement.

Preuve: C'est un cas particulier du cas des morphismes de modules sur un anneau. □

PROPOSITION 5.10. *Soit $\varphi : V \mapsto W$ est une application lineaire, alors φ est injective ssi*

$$\ker \varphi = \{0_V\}.$$

EXEMPLE 5.2.3. Dans K^d :

$$\mathbf{e}_i^* : \begin{array}{ccc} K^d & \mapsto & K \\ (x_1, \dots, x_d) & \mapsto & x_i \end{array}$$

$$\ker(\mathbf{e}_i^*) = \{(x_1, \dots, 0, \dots, x_d), x_j \in K, j \neq i\}, \text{Im}(\mathbf{e}_i^*) = K.$$

$$\begin{aligned}
S : \begin{array}{ccc} K^d & \mapsto & K \\ (x_1, \dots, x_d) & \mapsto & x_1 + \dots + x_d \end{array} \\
\ker(S) = \{(x_1, \dots, x_d) \in K^d, x_1 + \dots + x_d = 0\}, \operatorname{Im}(S) = K. \\
\varphi : \begin{array}{ccc} K^2 & \mapsto & K^2 \\ (x_1, x_2) & \mapsto & (2x_1 + x_2, x_1 + x_2) \end{array} \\
\ker(\varphi) = \{0_2\}, \operatorname{Im}(\varphi) = K^2.
\end{aligned}$$

NOTATION 5.3. *On notera*

$$\operatorname{Hom}_{K\text{-ev}}(V, W), \operatorname{Isom}_{K\text{-ev}}(V, W),$$

$$\operatorname{End}_{K\text{-ev}}(V) = \operatorname{Hom}_{K\text{-ev}}(V, V), \operatorname{Aut}_{K\text{-ev}}(V) = \operatorname{Isom}_{K\text{-ev}}(V, V)$$

les ensembles des applications lineaires, applications lineaires bijectives (ou isomorphismes), d'endomorphismes et d'automorphismes des K -espaces vectoriels V et W .

Pour simplifier on ecrira souvent

$$\operatorname{Hom}_K(V, W), \operatorname{Isom}_K(V, W), \operatorname{End}_K(V), \operatorname{Aut}_K(V)$$

On rappelle egalement que

PROPOSITION 5.11. *L'ensemble des automorphismes du K -ev V ,*

$$\operatorname{Aut}_{K\text{-ev}}(V) = \operatorname{Isom}_{K\text{-ev}}(V, V)$$

est un groupe pour la composition. On l'appelle egalement le groupe lineaire de V et on le note

$$\operatorname{Aut}_K(V) =: \operatorname{GL}(V).$$

On rappelle que (les applications lineaires etant des applications lineaires entre K -modules) et que K est par definition commutatif on a

PROPOSITION 5.12. *La composee de deux applications K -lineaires est K -lineaire : pour $\varphi \in \operatorname{Hom}_K(U, V)$ et $\psi \in \operatorname{Hom}_K(V, W)$ lineaires, alors $\psi \circ \varphi : U \mapsto W$ est K -lineaire et si φ est bijective alors $\varphi^{-1} : V \mapsto U$ est encore lineaire.*

Une combinaison lineaire de deux applications lineaires est lineaire: $\forall \varphi, \phi : U \mapsto V$ et $\forall \lambda \in K$, l'application

$$\lambda \cdot \varphi + \phi : u \in U \mapsto \lambda \varphi(u) + \phi(u) \in V$$

est K -lineaire.

On en deduit:

THÉORÈME 5.2. *L'ensemble des application lineaires $\operatorname{Hom}_K(V, W)$ a une structure naturelle de K -ev.*

L'ensemble des endomorphismes de V , $\operatorname{End}_K(V)$ muni de l'addition et de la composition a une structure naturelle de K -algebre. Son groupe des unites est le groupe

$$\operatorname{End}_{K\text{-ev}}(V)^\times = \operatorname{Aut}_{K\text{-ev}}(V) = \operatorname{GL}(V)$$

des applications K -lineaires bijectives. C'est un sous-groupe de $\operatorname{Bij}(V)$.

5.2.2.1. *Dual d'un espace vectoriel.* Le cas $W = K$ est important et admet un nom et une notation particuliere:

DÉFINITION 5.12. *Une application lineaire de $\ell : V \mapsto K$ est egalement appelee une forme lineaire. L'espace des formes lineaires $\operatorname{Hom}_K(V, K)$ est egalement note*

$$\operatorname{Hom}_K(V, K) = V^*.$$

On appelle egalement cet espace le dual de V .

5.2.3. Sous-espace engendre par un sous-ensemble. On rappelle egalement que

PROPOSITION 5.13 (Les SEV sont stables par intersection). *Soit W_i , $i \in I$ une famille de SEV de V indexes par un ensemble I alors leur intersection*

$$\bigcap_{i \in I} W_i \subset V$$

est un SEV de V .

DÉFINITION 5.13. *Soit $\mathcal{F} \subset V$ un sous-ensemble, on note*

$$\langle \mathcal{F} \rangle_K = \text{Vect}(\mathcal{F}) \subset V$$

le sous-espace vectoriel (le sous- K module) engendre par $\mathcal{F}$.

On rappelle qu'il s'agit de maniere equivalente

- *de l'intersection de tous les SEV contenant $\mathcal{F}$,*
- *de l'ensemble des combinaisons lineaires d'elements de $\mathcal{F}$ a coefficients dans K*

$$\langle \mathcal{F} \rangle_K = \left\{ \sum_{i=1}^n \lambda_i x_i, n \geq 1, \lambda_1, \dots, \lambda_n \in K, x_1, \dots, x_n \in \mathcal{F} \right\}.$$

Cette notion admet des cas particuliers.

5.2.3.1. *Sommes de SEVs, sommes directes.*

DÉFINITION 5.14. *Soient $X, Y \subset V$ des sous-espaces d'un espace vectoriel.*

Leur somme

$$X + Y = \langle X \cup Y \rangle \subset V$$

est par definition le sous-espace vectoriel engendre par les vecteurs de X et de Y .

LEMME 5.2. *On a*

$$X + Y = \{x + y, x \in X, y \in Y\}.$$

Preuve: Soit $W \subset V$ un SEV contenant X et Y alors W contient $X + Y$ car W est stable par somme. Il reste a montrer que $X + Y$ est un SEV car ce sera necessairement le plus petit contenant X et Y .

Soit $\lambda \in K, x, x' \in X, y, y' \in Y$ alors

$$\lambda(x + y) + (x' + y') = (\lambda.x + x') + (\lambda.y + y') \in X + Y$$

car X et Y sont des SEV. □

NOTATION 5.4. *Si $X \cap Y = \{0_V\}$, on dit que X et Y sont en somme directe et on ecrit*

$$X \oplus Y \subset V$$

pour leur somme.

Si de plus

$$X \oplus Y = V$$

on dit que V est somme directe de X et Y . On dit alors que X et Y sont des espaces supplementaires (dans V).

PROPOSITION 5.14. *Soit $V = X \oplus Y$ la somme directe de deux sous-espaces supplementaires X et Y alors l'écriture de tout vecteur $v \in V \in X \oplus Y$ sous la forme*

$$v = x + y, x \in X, y \in Y$$

est unique.

Preuve: Si $x + y = x' + y'$ alors $x - x' = y' - y$ et donc $x - x' \in X \cap Y = \{0_V\}$ cad que

$$x = x', \text{ et } y = y'.$$

□

EXERCICE 5.4. soit V un K -ev qui est une somme directe de deux SEV $V = X \oplus Y$. Comme on l'a vu tout $v \in V = X \oplus Y$ s'écrit de manière unique

$$v = x + y, \quad x \in X, \quad y \in Y.$$

Montrer que

(1) Les applications

$$\pi_X : \begin{array}{ccc} V & \mapsto & X \\ v & \mapsto & x \end{array}, \quad \pi_Y : \begin{array}{ccc} V & \mapsto & Y \\ v & \mapsto & y \end{array}$$

sont lineaires.

(2) l'EV V est isomorphe à l'espace vectoriel produit $X \times Y$.

5.3. Famille génératrice, libre, base

5.3.1. Famille génératrice. On rappelle la définition qu'on a vu pour les modules:

DÉFINITION 5.15. Soit V un K -e.v. Un sous-ensemble $\mathcal{G} \subset V$ est une famille génératrice si

$$\text{Vect}(\mathcal{G}) = \langle \mathcal{G} \rangle_K = V,$$

ie. tout élément $v \in V$ peut s'écrire sous la forme d'une combinaison linéaire (finie) à coefficients dans K d'éléments de $\mathcal{G}$: pour tout $v \in V$ il existe $n \geq 1$, $x_1, \dots, x_n \in K$, $\mathbf{e}_1, \dots, \mathbf{e}_n \in \mathcal{G}$ tels que

$$(5.3.1) \quad v = \sum_{i=1}^n x_i \mathbf{e}_i.$$

Si V admet une famille génératrice finie, on dit que V est un K -module ou un K -ev de type fini.

DÉFINITION 5.16. Soit V un K -ev de type fini. Si V est non-nul, sa dimension est le cardinal minimum d'une famille génératrice finie de V :

$$\dim(V) := \min_{\mathcal{G} \text{ génératrice}} |\mathcal{G}|.$$

Par convention, la dimension de l'espace vectoriel nul $\{0_V\}$ est

$$\dim(\{0_V\}) = 0$$

(on peut prendre la famille vide comme famille génératrice).

On dira également " K -ev de dimension finie" à la place de " K -ev de type fini".

Un espace vectoriel qui n'est pas de type fini est dit de " ∞ -dimension finie".

On va maintenant se restreindre au cas des espaces vectoriels de dimension finie. A la fin du chapitre, on décrira ce qui se passe pour les espaces vectoriels qui ne sont pas de dimension finie.

Le résultat principal de cette section est le théorème suivant:

THÉORÈME 5.3. Tout K -espace vectoriel de dimension finie $d = \dim V$ est isomorphe (comme K -ev) à l'espace vectoriel K^d (avec la convention que $\{0_K\} = K^0$). En d'autres termes V est isomorphe au K -module libre de rang $d = \dim(V)$, K^d .

Avant de démontrer ce théorème qui nous prendra un peu de temps, examinons sa signification concrète: supposons que $\mathcal{G} = \{\mathbf{e}_1, \dots, \mathbf{e}_d\} \subset V$ soit une famille génératrice finie de V de cardinal $d \geq \dim V$. Tout élément $v \in V$ peut donc se représenter sous la forme d'une combinaison linéaire des $\mathbf{e}_i$

$$v = \sum_{i=1}^d x_i \mathbf{e}_i, \quad x_i \in K.$$

En d'autres termes, on dispose d'une application " ∞ -combinaison linéaire" qui est surjective:

$$CL_{\mathcal{G}} : \begin{array}{ccc} K^d & \mapsto & V \\ (x_1, \dots, x_d) & \mapsto & CL_{\mathcal{G}}(x_1, \dots, x_d) = x_1 \mathbf{e}_1 + \dots + x_d \mathbf{e}_d \end{array}$$

REMARQUE 5.3.1. Cette application *depend* de l'ordre dans lequel on enumere les elements de la famille $\mathcal{G}$: en general

$$x_1.\mathbf{e}_1 + x_2.\mathbf{e}_2 \neq x_1.\mathbf{e}_2 + x_2.\mathbf{e}_1.$$

LEMME 5.3. *L'application $CL_{\mathcal{G}}$ est lineaire.*

Preuve: Soient

$$\mathbf{x} = (x_1, \dots, x_d), \mathbf{y} = (y_1, \dots, y_d) \in K^d$$

et $\lambda \in K$ alors on veut verifier que

$$CL_{\mathcal{G}}(\lambda.\mathbf{x} + \mathbf{y}) = \lambda.CL_{\mathcal{G}}(\mathbf{x}) + CL_{\mathcal{G}}(\mathbf{y}).$$

C'est une consequence de la commutativite et de l'associativite des lois d'addition et de multiplication: on a

$$\begin{aligned} CL_{\mathcal{G}}(\lambda.\mathbf{x} + \mathbf{y}) &= CL_{\mathcal{G}}(\lambda.x_1 + y_1, \dots, \lambda.x_d + y_d) = (\lambda.x_1 + y_1)\mathbf{e}_1 + \dots + (\lambda.x_d + y_d)\mathbf{e}_d \\ &= \lambda.x_1.\mathbf{e}_1 + y_1.\mathbf{e}_1 + \dots + \lambda.x_d.\mathbf{e}_d + y_d.\mathbf{e}_d \\ &= \lambda.(x_1.\mathbf{e}_1 + \dots + x_d.\mathbf{e}_d) + (y_1.\mathbf{e}_1 + \dots + y_d.\mathbf{e}_d) \\ &= \lambda.CL_{\mathcal{G}}(\mathbf{x}) + CL_{\mathcal{G}}(\mathbf{y}). \end{aligned}$$

□

On a donc la definition suivante equivalente d'une famille generatrice:

DÉFINITION. *Soit V un K -e.v. Un sous-ensemble fini*

$$\mathcal{G} = \{\mathbf{e}_1, \dots, \mathbf{e}_d\} \subset V$$

est une famille generatrice (du K -ev V) ssi les conditions equivalentes suivantes sont satisfaites:

(1) *On a*

$$\text{Vect}(\mathcal{G}) = V.$$

(2) *pour tous $v \in V$, il existe $x_1, \dots, x_d \in K$ tels que*

$$v = x_1.\mathbf{e}_1 + \dots + x_d.\mathbf{e}_d.$$

(3) *L'application lineaire*

$$CL_{\mathcal{G}} : \begin{array}{ccc} K^d & \mapsto & V \\ (x_1, \dots, x_d) & \mapsto & x_1.\mathbf{e}_1 + \dots + x_d.\mathbf{e}_d \end{array}$$

est surjective.

Si V admet une famille generatrice finie ou dit que V est un K -ev de type fini ou est de dimension finie. On a alors

$$\dim_K V \leq d.$$

Le Theoreme 5.3 sera alors consequence du

THÉORÈME. *Soit $\mathcal{G} \subset V$ une famille generatrice de V de cardinal $d = \dim V$ alors l'application $CL_{\mathcal{G}}$ est injective et defini donc un isomorphisme*

$$CL_{\mathcal{G}} : K^d \simeq V.$$

Preuve: Soit $\mathcal{G} = \{\mathbf{e}_1, \dots, \mathbf{e}_d\} \subset V$ une famille generatrice de cardinal la dimension $d = \dim V$. Par definition de la dimension, une famille de cardinal $< d$ ne peut etre generatrice. Supposons que $CL_{\mathcal{G}}$ ne soit pas injective: il existe donc $(u_1, \dots, u_d) \neq 0_d$ tel que

$$u_1.\mathbf{e}_1 + \dots + u_d.\mathbf{e}_d = 0_V.$$

comme $(u_1, \dots, u_d)$ est non-nul il existe i tel que $u_i \neq 0_K$. Supposons (quitte a permuter les indices) que $i = d$. On a alors

$$u_d.\mathbf{e}_d = -(u_1.\mathbf{e}_1 + \dots + u_{d-1}.\mathbf{e}_{d-1})$$

et donc comme u_d est inversible (car non-nul)

$$\mathbf{e}_d = y_1 \cdot \mathbf{e}_1 + \cdots + y_{d-1} \cdot \mathbf{e}_{d-1}$$

avec

$$y_i = -u_i \cdot u_d^{-1}.$$

Je dis que la famille $\{\mathbf{e}_1, \dots, \mathbf{e}_{d-1}\}$ engendre V ce qui donnera une contradiction par minimalite de d .

Soit $v \in V$, il existe $x_1, \dots, x_d \in K$ tel que

$$\begin{aligned} v &= x_1 \cdot \mathbf{e}_1 + \cdots + x_{d-1} \cdot \mathbf{e}_{d-1} + x_d \cdot \mathbf{e}_d \\ &= x_1 \cdot \mathbf{e}_1 + \cdots + x_{d-1} \cdot \mathbf{e}_{d-1} + x_d \cdot (y_1 \cdot \mathbf{e}_1 + \cdots + y_{d-1} \cdot \mathbf{e}_{d-1}) \\ &= x'_1 \cdot \mathbf{e}_1 + \cdots + x'_{d-1} \cdot \mathbf{e}_{d-1} \end{aligned}$$

avec

$$x'_i = x_i + x_d y_i = x_i - x_d u_i \cdot u_d^{-1}.$$

Ainsi l'application $CL_{\mathcal{G}}$ est injective et comme elle est surjective (car $\mathcal{G}$ est generatrice) et sa reciproque est egalement lineaire: c'est un isomorphisme de K -espaces vectoriels de K^d vers V $\square$

Le corollaire suivant montre que la dimension determine completement la classe d'isomorphisme des K -ev de dimension finie.

COROLLAIRE 5.3 (Critere dimensionel d'isomorphisme). *Soient V, W des K -ev de dimensions finie d_V et d_W alors V et W sont isomorphes ssi ils ont meme dimension:*

$$V \simeq W \iff d_V = d_W.$$

Preuve: Si $d_V = d_W = d$ alors il existe des isomorphismes

$$\varphi : K^d \simeq V, \psi : K^d \simeq W$$

et alors $\psi \circ \varphi^{-1} : V \rightarrow W$ est un isomorphisme entre V et W .

Reciproquement soit $\varphi : V \simeq W$ un isomorphisme, on veut mq $d_V = d_W$. Soit $\mathcal{G} = \{\mathbf{e}_1, \dots, \mathbf{e}_{d_V}\}$ une famille generatrice de V alors

$$\varphi(\mathcal{G}) = \{\varphi(\mathbf{e}_1), \dots, \varphi(\mathbf{e}_{d_V})\}$$

est generatrice de W : pour tout $w \in W$ il existe $v \in V$ tel que $\varphi(v) = w$. Ecrivons

$$v = x_1 \mathbf{e}_1 + \cdots + x_v \mathbf{e}_v$$

alors

$$w = \varphi(v) = x_1 \varphi(\mathbf{e}_1) + \cdots + x_v \varphi(\mathbf{e}_v)$$

donc w est bien CL des elements de $\{\varphi(\mathbf{e}_1), \dots, \varphi(\mathbf{e}_{d_V})\}$.

Par definition de la dimension on a donc

$$d_W \leq |\varphi(\mathcal{G})| \leq |\mathcal{G}| = d_V.$$

Echangeant V et W (en remplaçant φ par φ^{-1}) on a $d_V \leq d_W$ et donc

$$d_V = d_W.$$

$\square$

5.3.2. Famille libre. La discussion precedente nous conduit naturellement vers le point suivant

Soit $\mathcal{F} = \{\mathbf{e}_1, \dots, \mathbf{e}_f\} \subset V$ une famille de f vecteurs: on dispose alors d'une application lineaire "Combinaison lineaire":

$$CL_{\mathcal{F}} : \begin{array}{ccc} K^f & \mapsto & V \\ (x_1, \dots, x_f) & \mapsto & CL_{\mathcal{F}}(x_1, \dots, x_f) = x_1 \cdot \mathbf{e}_1 + \dots + x_f \cdot \mathbf{e}_f \end{array}$$

dont l'image est

$$CL_{\mathcal{F}}(K^f) = \text{Vect}(\mathcal{F}) := W \subset V$$

est le SEV engendre par $\mathcal{F}$; on s'est pose (dans le cas ou $\mathcal{G}$ etait une famille generatrice de taille minimale) la question de l'injectivite de cette application.

Soit $w \in W$, alors w est combinaison lineaire d'elements de $\mathcal{F}$ et s'ecrit

$$w = x_1 \cdot \mathbf{e}_1 + \dots + x_f \cdot \mathbf{e}_f$$

pour $(x_1, \dots, x_d) \in K^d$ et par definition de l'injectivite, la representation de w sous cette forme est unique:

$$w = x_1 \cdot \mathbf{e}_1 + \dots + x_f \cdot \mathbf{e}_f = x'_1 \cdot \mathbf{e}_1 + \dots + x'_d \cdot \mathbf{e}_f \implies x_1 = x'_1, \dots, x_f = x'_f.$$

D'autre part (par le critere d'injectivite des applications lineaires), l'injectivite est equivalente au fait que

$$\ker(CL_{\mathcal{F}}) = \{\mathbf{x} \in K^f, x_1 \cdot \mathbf{e}_1 + \dots + x_f \cdot \mathbf{e}_f = 0_V\} = \{0_{K^f} = (0, \dots, 0)\}$$

ce qui s'interprete en disant que le vecteur nul 0_V (qui appartient a W) admet une *unique* representation sous forme de combinaison lineaire des $\mathbf{e}_i, i \leq d$: la combinaison *triviale* ou *nulle*:

$$x_1 \cdot \mathbf{e}_1 + \dots + x_f \cdot \mathbf{e}_f = 0_V \iff x_1 = \dots = x_f = 0_K.$$

Cela nous conduit a la definition generale suivante:

DÉFINITION 5.17. *Un sous-ensemble fini $\mathcal{F} = \{\mathbf{e}_1, \dots, \mathbf{e}_f\} \subset V$ d'un espace vectoriel est une famille libre de V si et seulement si l'une des trois conditions equivalentes suivante est satisfaite:*

(1) *L'application lineaire*

$$CL_{\mathcal{F}} : \begin{array}{ccc} K^f & \mapsto & V \\ (x_1, \dots, x_f) & \mapsto & x_1 \cdot \mathbf{e}_1 + \dots + x_f \cdot \mathbf{e}_f \end{array}$$

est injective.

(2) *pour tous $x_1, \dots, x_f, x'_1, \dots, x'_f \in K$*

$$x_1 \cdot \mathbf{e}_1 + \dots + x_f \cdot \mathbf{e}_f = x'_1 \cdot \mathbf{e}_1 + \dots + x'_f \cdot \mathbf{e}_f \implies x_1 - x'_1 = \dots = x_f - x'_f = 0_K.$$

(3) *pour tous $x_1, \dots, x_f \in K$*

$$x_1 \cdot \mathbf{e}_1 + \dots + x_f \cdot \mathbf{e}_f = 0_V \implies x_1 = \dots = x_f = 0_K.$$

Une famille $\mathcal{F}$ qui n'est pas libre est dit liee.

EXEMPLE 5.3.1. Soit $\mathbf{e} \in V - \{0_V\}$ un vecteur non-nul alors $\{\mathbf{e}\}$ est libre: supposons que

$$x \cdot \mathbf{e} = 0_V$$

pour $x \in K$; si $x \neq 0_K$ alors x est inversible et

$$x^{-1} \cdot x \cdot \mathbf{e} = \mathbf{e} = 0_V$$

qui est une contradiction donc $x = 0_K$.

EXEMPLE 5.3.2. Dans K^d , la base canonique

$$\mathcal{B}^0 := \{\mathbf{e}_i^0, i = 1, \dots, d\}$$

qui est generatrice est egalement libre; on rappelle que $\mathbf{e}_i^0$ est le vecteur dont toutes les coordonnees sont nulles sauf la i -eme qui vaut 1,

$$\mathbf{e}_1^0 = (1, 0, \dots, 0), \dots, \mathbf{e}_d^0 = (0, 0, \dots, 1).$$

En effet, pour tout $x_1, \dots, x_d \in K$ on a

$$\sum_{i=1}^d x_i \cdot \mathbf{e}_i^0 = (x_1, x_2, \dots, x_d)$$

et donc si

$$= \sum_{i=1}^d x_i \cdot \mathbf{e}_i^0 = 0_d = (0, \dots, 0)$$

on a

$$x_1 = \dots = x_d = 0.$$

EXEMPLE 5.3.3. Dans $\mathbb{R}^3$, la famille

$$(1, 1, 0), (0, 1, 1), (1, 0, 1)$$

est libre.

En revanche si $\text{car}(K) = 2$ alors la famille est liée:

$$(1, 1, 0) + (0, 1, 1) + (1, 0, 1) = (2, 2, 2) = \underline{0}_3.$$

En fait, cette famille est libre dans K^3 ou K est de caractéristique $\neq 2$.

EXEMPLE 5.3.4. Dans la preuve du Theorem 5.3 on a montré que

PROPOSITION 5.15. Soit V un K -ev de dimension d et $\mathcal{G} = \{\mathbf{e}_1, \dots, \mathbf{e}_d\}$ une famille génératrice de cardinal d alors $\mathcal{G}$ est libre.

On va donner un critère pour qu'une famille soit liée.

PROPOSITION 5.16. Une famille a l éléments $\mathcal{F} = \{\mathbf{e}_1, \dots, \mathbf{e}_l\} \subset V$ est liée ssi il existe $i \in \{1, \dots, l\}$ tel que $\mathbf{e}_i$ peut s'exprimer comme combinaison linéaire des autres éléments de $\mathcal{F}$:

$$\exists i \leq l, \mathbf{e}_i \in \text{Vect}(\mathcal{F} - \{\mathbf{e}_i\}) = \text{Vect}(\{\mathbf{e}_j, j \neq i\}).$$

On a alors

$$W = \text{Vect}(\mathcal{F}) = \text{Vect}(\mathcal{F} - \{\mathbf{e}_i\}).$$

Preuve: Si $\mathcal{F}$ est liée, il existe $x_1, \dots, x_l \in K$ non-tous nuls tels que

$$0_V = x_1 \cdot \mathbf{e}_1 + \dots + x_l \cdot \mathbf{e}_l.$$

Supposons (quitte à renuméroter) que $x_l \neq 0$ alors

$$-x_l \cdot \mathbf{e}_l = x_1 \cdot \mathbf{e}_1 + \dots + x_{l-1} \cdot \mathbf{e}_{l-1}$$

et comme $-x_l$ est inversible

$$\mathbf{e}_l = (x_1 / -x_l) \cdot \mathbf{e}_1 + \dots + (x_{l-1} / -x_l) \cdot \mathbf{e}_{l-1} \in \text{Vect}(\mathcal{F} - \{\mathbf{e}_l\}).$$

Reciproquement si $\mathbf{e}_l \in \text{Vect}(\mathcal{F} - \{\mathbf{e}_l\})$ alors

$$\mathbf{e}_l = y_1 \cdot \mathbf{e}_1 + \dots + y_{l-1} \cdot \mathbf{e}_{l-1}$$

et

$$0_V = y_1 \cdot \mathbf{e}_1 + \dots + y_{l-1} \cdot \mathbf{e}_{l-1} + (-1) \cdot \mathbf{e}_l$$

avec $-1 \neq 0_K$.

On a donc

$$\mathcal{F} = \{\mathbf{e}_1, \dots, \mathbf{e}_i, \dots, \mathbf{e}_l\} \subset \text{Vect}(\mathcal{F} - \{\mathbf{e}_i\})$$

et donc

$$W = \text{Vect}(\mathcal{F}) = \text{Vect}(\mathcal{F} - \{\mathbf{e}_i\}).$$

□

On va maintenant montrer que les familles libres ne peuvent pas être trop grandes.

THÉORÈME 5.4 (Majoration du cardinal d'une famille libre). *Soit V un espace vectoriel non-nul de dimension d et $\mathcal{F} = \{v_1, \dots, v_f\} \subset V$ une famille finie et libre; alors $f \leq d$.*

Preuve: Notons que les vecteurs $v_1, \dots, v_f$ sont tous distincts: si on avait $v_1 = v_2$ alors v_1 serait combinaison lineaire de $v_2, \dots, v_f$.

On procede par recurrence sur d .

Si $d = 1$ alors $V = K \cdot \mathbf{e}$ avec $\mathbf{e} \neq 0_V$; soit $\mathcal{F} = \{v_1, \dots, v_f\}$ une famille libre a f elements. Montrons que $f = 1$.

Notons que $v_1 \neq 0_V$: sinon on aurait

$$0_V = 1.v_1 + 0.v_2 + \dots + 0.v_f$$

et la famille ne serait pas libre. On a pour $i = 1, \dots, f$

$$v_i = x_i \cdot \mathbf{e}$$

avec $x_i \in K$ et $x_1 \neq 0$ (sinon v_1 serait nul). On a alors si $f \geq 2$

$$\mathbf{e} = x_1^{-1}.v_1, \quad v_2 = x_2 \cdot \mathbf{e} = (x_2/x_1).v_1$$

Ainsi v_2 est combinaison lineaire de v_1 contredisant le fait que la famille est libre.

Supposons qu'on a demontre le resultat pour tout espace vectoriel de dimension $\leq d - 1$.

Soit V de dimension $d \geq 1$, $\mathcal{G} = \{\mathbf{e}_1, \dots, \mathbf{e}_d\}$ une famille qui engendre V et

$$\mathcal{F} = \{v_1, \dots, v_f\} \subset V$$

une famille libre a f elements. Montrons que $f \leq d$.

Par definition chaque element de $\mathcal{F}$ est combinaison lineaire des elements de $\mathcal{G}$: pour $i = 1, \dots, f$, il existe $(x_{i,j})_{j \leq d}$ tel que

$$v_i = x_{i,1}\mathbf{e}_1 + \dots + x_{i,d}\mathbf{e}_d, \quad i = 1, \dots, f.$$

Le fait que $\mathcal{F}$ est libre implique que les v_i sont tous non-nuls (cf. ci-dessus). En particulier, il existe un indice $j_0 \in \{1, \dots, d\}$ tel que

$$x_{f,j_0} \neq 0.$$

Supposons (quitte a renumeroter les $\mathbf{e}_j$) que $j_0 = d$; on a donc $x_{f,d} \neq 0$ qui est donc inversible. Posons

$$(5.3.2) \quad v'_i = v_i - (x_{i,d}/x_{f,d}).v_f, \quad i = 1, \dots, f.$$

On a

$$v'_f = v_f - (x_{f,d}/x_{f,d}).v_f = 0_V$$

et en general

$$v'_i = x'_{i,1}\mathbf{e}_1 + \dots + x'_{i,d-1}\mathbf{e}_{d-1} + (x_{i,d} - (x_{i,d}/x_{f,d}).x_{f,d})\mathbf{e}_d = x'_{i,1}\mathbf{e}_1 + \dots + x'_{i,d-1}\mathbf{e}_{d-1}.$$

ainsi la famille

$$\mathcal{F}' = \{v'_i, \quad i \leq f - 1\} \subset V' = \text{Vect}(\{\mathbf{e}_1, \dots, \mathbf{e}_{d-1}\}) \subset V$$

possede $f - 1$ elements et est contenue dans un sous-espace vectoriel V' engendre par $d - 1$ elements donc de dimension $\leq d - 1$. De plus cette famille est libre: supposons que

$$x_1.v'_1 + \dots + x_{f-1}.v'_{f-1} = 0_V;$$

utilisant (5.3.2) on voit que

$$x_1.v_1 + \dots + x_{f-1}.v_{f-1} + y_f.v_f = 0_V$$

pour un certain $y_f \in K$ et comme la famille $\mathcal{F}$ est libre on a

$$x_1 = \dots = x_{f-1} = 0_K.$$

On a alors par recurrence que

$$f - 1 \leq \dim V' \leq d - 1$$

et donc $f \leq d$. □

5.3.3. Base.

DÉFINITION 5.18. Soit V un espace vectoriel de dimension finie. Une famille $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_d\}$ est une base de V si l'une des conditions équivalentes suivantes est vérifiée:

- (1) $\mathcal{B}$ est génératrice et libre,
- (2) L'application combinaison linéaire de $\mathcal{B}$,

$$CL_{\mathcal{B}} : K^d \mapsto V$$

est un isomorphisme,

- (3) Pour tout $v \in V$ il existe un unique uplet $(x_1, \dots, x_d) \in K^d$ tel que v s'écrit sous la forme

$$v = x_1 \cdot \mathbf{e}_1 + \dots + x_d \cdot \mathbf{e}_d.$$

EXEMPLE 5.3.5. Pour $V = K^d$, la base canonique

$$\mathcal{B}^0 = \{\mathbf{e}_1^0, \dots, \mathbf{e}_d^0\}$$

forme (tautologiquement) une base.

On a

THÉORÈME 5.5. Soit V un K -ev de dimension d alors V possède une base et toute base $\mathcal{B}$ de V vérifie

$$(5.3.3) \quad |\mathcal{B}| = \dim(V).$$

REMARQUE 5.3.2. En particulier

$$\dim(K^d) = d.$$

Preuve: On a vu d'une famille génératrice $\mathcal{G}$ de cardinal minimal $\dim V$ est libre et donc forme une base de V .

Si $\mathcal{B}$ est une base de V alors comme elle est génératrice on a

$$|\mathcal{B}| \geq \dim V$$

et comme $\mathcal{B}$ est libre on a par le Théorème 5.4

$$|\mathcal{B}| \leq \dim V.$$

□

Le Théorème d'existence d'une base admet la variante suivante concernant les familles libres et génératrices

THÉORÈME 5.6 (Extraction et Completion). Soit V un K -ev non nul de dimension d . On a

- (1) Une famille génératrice $\mathcal{G}$ de cardinal d est une base.
- (2) Une famille libre $\mathcal{L}$ de cardinal d est une base.
- (3) (Extraction) Soit $\mathcal{G} \subset V$ une famille génératrice alors il existe une base $\mathcal{B}$ de V contenue dans $\mathcal{G}$.
- (4) (Completion) Soit $\mathcal{L} \subset V$ une famille libre alors il existe une base $\mathcal{B}$ de V contenant $\mathcal{L}$.

Preuve: Soit $\mathcal{G}$ une famille génératrice (pas forcément finie); par définition de la dimension $|\mathcal{G}| \geq d$.

Montrons que $\mathcal{G}$ contient une base. L'ensemble $\mathcal{G}$ contient au moins un vecteur non-nul (sinon $V = \text{Vect}(\mathcal{G}) = \{0_V\}$ ce qui est exclu) et la famille réduite à un élément $\{\mathbf{e}\}$ est libre. Soit $\mathcal{B} \subset \mathcal{G}$ une sous-famille libre dont le cardinal $|\mathcal{B}|$ est maximal parmi les sous-familles libres de $\mathcal{G}$. Montrons que $\mathcal{B}$ est génératrice et est donc une base.

On sait déjà que cette famille est finie:

$$\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_{|\mathcal{B}|}\}$$

avec

$$|\mathcal{B}| \leq d.$$

On a les deux cas suivants:

- (1) Si $|\mathcal{B}| = |\mathcal{G}|$ alors $\mathcal{B} = \mathcal{G}$ est generatrice et $\mathcal{B}$ est une base.
- (2) Si $|\mathcal{B}| < |\mathcal{G}|$. Supposons que $\mathcal{B}$ n'est pas generatrice c'est a dire

$$\text{Vect}(\{\mathbf{e}_1, \dots, \mathbf{e}_{|\mathcal{B}|}\}) \neq \text{Vect}(\mathcal{G}) = V,$$

alors il existe $\mathbf{e} \in \mathcal{G}$ tel que

$$\mathbf{e} \notin \text{Vect}(\mathcal{B})$$

c'est a dire que pour tout $x_1, \dots, x_{|\mathcal{B}|} \in K$ on a toujours

$$\mathbf{e} \neq x_1 \mathbf{e}_1 + \dots + x_{|\mathcal{B}|} \mathbf{e}_{|\mathcal{B}|}.$$

Montrons qu'alors la famille $\mathcal{B} \cup \{\mathbf{e}\}$ est encore libre ce qui contredira la maximalite de $|\mathcal{B}|$: supposons que pour $x_1, \dots, x_{|\mathcal{B}|}, x \in K$ on ait

$$x_1 \mathbf{e}_1 + \dots + x_{|\mathcal{B}|} \mathbf{e}_{|\mathcal{B}|} + x \mathbf{e} = 0_V$$

alors

- (a) si $x = 0$ on a

$$x_1 \mathbf{e}_1 + \dots + x_{|\mathcal{B}|} \mathbf{e}_{|\mathcal{B}|} = 0_V$$

et comme $\mathcal{B}$ est libre on a $x_1 = \dots = x_{|\mathcal{B}|} = x = 0$.

- (b) Si $x \neq 0$ alors x est inversible et on a

$$\mathbf{e} = -(x_1/x) \mathbf{e}_1 - \dots - (x_{|\mathcal{B}|}/x) \mathbf{e}_{|\mathcal{B}|}$$

une contradiction: ainsi la famille est libre.

On obtient alors une contradiction avec la maximalite de $|\mathcal{B}|$ ce qui implique que $\mathcal{B}$ est generatrice.

Soit $\mathcal{L} = \{\mathbf{e}_1, \dots, \mathbf{e}_{|\mathcal{L}|}\}$ une famille libre non-void (on sait que $|\mathcal{L}| \leq d$).

Montrons que $\mathcal{L}$ est contenue dans une base. Il existe une famille generatrice finie contenant $\mathcal{L}$: il suffit de prendre la reunion $\mathcal{L} \cup \mathcal{G}$ de $\mathcal{L}$ et d'une famille generatrice finie $\mathcal{G}$ de V (par exemple une base).

Soit $\mathcal{B} \supset \mathcal{L}$ une famille generatrice finie de V contenant $\mathcal{L}$ et dont le cardinal $|\mathcal{B}|$ est minimal parmi toutes les familles generatrices finies de V contenant $\mathcal{L}$. Montrons que $\mathcal{B}$ est libre et est donc une base.

- (1) Si $|\mathcal{B}| = |\mathcal{L}|$ alors $\mathcal{B} = \mathcal{L}$ est generatrice et libre et c'est une base.
- (2) Si $|\mathcal{B}| > |\mathcal{L}|$ ecrivons

$$\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_{|\mathcal{L}|}, \dots, \mathbf{e}_{|\mathcal{B}|}\}$$

et supposons que $\mathcal{B}$ ne soit pas libre: il existe $x_1, \dots, x_{|\mathcal{B}|} \in K$ non tous nuls tels que

$$x_1 \mathbf{e}_1 + \dots + x_{|\mathcal{L}|} \mathbf{e}_{|\mathcal{L}|} + \dots + x_{|\mathcal{B}|} \mathbf{e}_{|\mathcal{B}|} = 0_V.$$

si $x_{|\mathcal{L}|+1} = \dots = x_{|\mathcal{B}|} = 0$ alors on a

$$x_1 \mathbf{e}_1 + \dots + x_{|\mathcal{L}|} \mathbf{e}_{|\mathcal{L}|} = 0_V$$

et comme $\mathcal{L}$ est libre on a

$$x_1 = \dots = x_{|\mathcal{L}|} = x_{|\mathcal{L}|+1} = \dots = x_{|\mathcal{B}|} = 0.$$

Sinon il existe $i > |\mathcal{L}|$ tel que $x_i \neq 0$ disons que c'est $x_{|\mathcal{B}|}$: on a alors

$$\mathbf{e}_{|\mathcal{B}|} = -(x_1/x_{|\mathcal{B}|}) \mathbf{e}_1 - \dots - (x_{|\mathcal{B}|-1}/x_{|\mathcal{B}|}) \mathbf{e}_{|\mathcal{B}|-1}$$

et alors comme $\mathbf{e}_{|\mathcal{B}|}$ est combinaison lineaire des $\mathbf{e}_1, \dots, \mathbf{e}_{|\mathcal{B}|-1}$, la famille $\{\mathbf{e}_1, \dots, \mathbf{e}_{|\mathcal{B}|-1}\}$ contient $\mathcal{L}$ et est generatrice ce qui contredit la minimalite de $|\mathcal{B}|$. Ainsi $\mathcal{B}$ est libre.

□

On a demontre dans la deuxieme partie un resultat un peu plus fort:

THÉORÈME 5.7 (de la base incomplete). *Etant donne $\mathcal{L}$ une famille libre de V et $\mathcal{B} \subset V$ une base, on peut extraire de $\mathcal{B}$ une sous-famille $\mathcal{L}' \subset \mathcal{B}$ de sorte que $\mathcal{L} \sqcup \mathcal{L}'$ forme une base de V .*

EXERCICE 5.5. Montrer que si X et Y sont de dimension finie on a

$$\dim(X \times Y) = \dim(X) + \dim(Y).$$

Montrer que si $V = X \oplus Y$, alors

$$\dim(V) = \dim(X) + \dim(Y).$$

5.3.4. Sous-espaces vectoriels et dimension.

THÉORÈME 5.8 (Bases et SEV). *Soit V un espace vectoriel de dimension finie, et $W \subset V$ un sous-espace vectoriel alors W est de dimension finie et*

- (1) *on a $\dim(W) \leq \dim(V)$.*
- (2) *Si $\dim(W) = \dim(V)$ alors $W = V$.*
- (3) *Si $\mathcal{B}_W$ est une base de W alors il existe une base $\mathcal{B}_V$ de V contenant $\mathcal{B}_W$.*

Preuve: Soit $\mathcal{L} \subset W$ une famille libre et finie de W alors $\mathcal{L}$ est libre dans V et de cardinal $l = |\mathcal{L}| \leq \dim V$. On peut donc supposer que $\mathcal{L} = \{\mathbf{e}_1, \dots, \mathbf{e}_l\}$ est de cardinal maximal (parmi les familles libres et finies de W). On suppose alors qu'il existe $\mathbf{e} \in W$ tel que

$$\mathbf{e} \notin \text{Vect}(\{\mathbf{e}_1, \dots, \mathbf{e}_l\})$$

et on en deduit comme dans le Theoreme d'Extraction/Completion que $\{\mathbf{e}_1, \dots, \mathbf{e}_l, \mathbf{e}\}$ est libre ce qui contredit la maximalite de l . Ainsi

$$\text{Vect}(\{\mathbf{e}_1, \dots, \mathbf{e}_l\}) = W$$

et W est de dimension finie egale a $l \leq \dim V$.

Les deux derniers points resultent du Theoreme d' extraction/completion.

□

- Un sous-espace vectoriel de dimension 1 est appelle *droite vectorielle*.
- Un sous-espace vectoriel de dimension 2 est appelle *plan vectoriel*.
- Un sous-espace vectoriel de dimension $\dim(V) - 1$ est appelle *hyperplan vectoriel*.

5.4. Espaces vectoriels de dimension infinie

DÉFINITION 5.19. *Un K -ev qui ne possede pas de famille generatrice finie est dit de dimension infinie.*

Repetons la definition de famille generatrice:

DÉFINITION 5.20. *Soit V un K -e.v. Un sous-ensemble $\mathcal{G} \subset V$ est une famille generatrice si*

$$\text{Vect}(\mathcal{G}) = V,$$

ie. tout element $v \in V$ peut s'ecrire sous la forme d'une combinaison lineaire (finie) d'elements de $\mathcal{G}$: il existe $d \geq 1$, $\mathbf{e}_1, \dots, \mathbf{e}_d \in \mathcal{G}$, $x_1, \dots, x_d \in K$, tels que

$$(5.4.1) \quad v = x_1 \mathbf{e}_1 + \dots + x_d \mathbf{e}_d.$$

Donnons une definition generale d'une famille libre (pas forcement finie):

DÉFINITION 5.21. *Soit V un K -e.v., un sous-ensemble $\mathcal{L} \subset V$ est une famille libre si tout sous-ensemble fini $\mathcal{L}' \subset \mathcal{L}$ est libre: si $\mathcal{L}' = \{\mathbf{e}_1, \dots, \mathbf{e}_d\}$ (les elements tous distincts), on a*

$$(5.4.2) \quad x_1 \mathbf{e}_1 + \dots + x_d \mathbf{e}_d = 0_V \iff x_1 = \dots = x_d = 0_K.$$

On defini alors ce qu'est une base:

DÉFINITION 5.22. Une base algebrique $\mathcal{B} \subset V$ est une famille libre et generatrice.

PROPOSITION 5.17. Soit $\mathcal{B} \subset V$ une base algebrique. Alors tout element v de V est representable comme combinaison lineaire finie d'elements de $\mathcal{B}$ et une telle representation est unique.

Preuve: L'existence est simplement le fait que $\mathcal{B}$ est generatrice.

Pour l'unicite supposons que

$$v = x_1 \mathbf{e}_1 + \cdots + x_d \mathbf{e}_d = x'_1 \mathbf{e}'_1 + \cdots + x'_{d'} \mathbf{e}'_{d'}$$

pour

$$\mathcal{B}' = \{\mathbf{e}_1, \dots, \mathbf{e}_d\}, \mathcal{B}'' = \{\mathbf{e}'_1, \dots, \mathbf{e}'_{d'}\} \subset \mathcal{B}.$$

Quitte a remplacer $\mathcal{B}$ et $\mathcal{B}'$ par la reunion $\mathcal{B} \cup \mathcal{B}'$ on peut (en ajoutant des coefficients nuls) supposer que $\mathcal{B}' = \mathcal{B}''$: on a

$$v = x_1 \mathbf{e}_1 + \cdots + x_d \mathbf{e}_d = x'_1 \mathbf{e}_1 + \cdots + x'_d \mathbf{e}_d$$

et donc

$$0_V = (x_1 - x'_1) \mathbf{e}_1 + \cdots + (x_d - x'_d) \mathbf{e}_d$$

et comme $\mathcal{B}$ est libre on a

$$x_1 - x'_1 = \cdots = x_d - x'_d = 0_K$$

c'est a dire

$$x_1 = x'_1, \dots, x_d = x'_d.$$

□

EXERCICE 5.6. Soit $\mathcal{F}(\mathbb{N}, \mathbb{R})$ l'espace des fonctions de $\mathbb{N}$ a valeurs reelles (ie. les suites a valeurs reelles). Soit $f : \mathbb{N} \mapsto \mathbb{R}$ une telle fonction; son support est par definition l'ensemble des des point ou f ne s'annule PAS:

$$\text{supp}(f) = f^{(-1)}(\mathbb{R} - \{0\}) = \{n \in \mathbb{N}, f(n) \neq 0\}.$$

Soit $\mathcal{F}_f(\mathbb{N}, \mathbb{R}) \subset \mathcal{F}(\mathbb{N}, \mathbb{R})$ le sous-ensemble des fonctions a support fini.

Pour $m \in \mathbb{N}$ un element, on note $1_{\{m\}}$ la fonction indicatrice de m :

$$1_{\{m\}}(n) = \begin{cases} 1 & \text{si } n = m \\ 0 & \text{si } n \neq m. \end{cases}$$

- (1) Montrer que $\mathcal{F}_f(\mathbb{N}, \mathbb{R})$ est un SEV de $\mathcal{F}(\mathbb{N}, \mathbb{R})$.
- (2) Montrer que la famille

$$\{1_{\{m\}}, m \geq 0\}$$

est une base de $\mathcal{F}_f(\mathbb{N}, \mathbb{R})$.

Il est beaucoup plus difficile d'imaginer une base de l'espace $\mathcal{F}(\mathbb{N}, \mathbb{R})$. Pourtant on a le resultat suivant necessite de travailler dans une theorie des ensembles qui contient l'axiome du choix (par exemple ZFC).

THÉORÈME 5.9 (Existence de bases sous l'axiome du choix). Dans une theorie des ensembles contenant l'axiome du choix, tout espace vectoriel sur un corps K possede une base et toutes les bases de V ont meme cardinal: pour toutes bases $\mathcal{B}, \mathcal{B}'$ il existe une bijection

$$\mathcal{B} \simeq \mathcal{B}'.$$

La dimension de V est de cardinal d'une base:

$$\dim(V) = |\mathcal{B}|.$$

REMARQUE 5.4.1. Le Theoreme de la base incomplete est vrai (sous l'axiome du choix): soit $\mathcal{L} \subset$ une famille libre et $\mathcal{G}$ un famille generatrice. Il existe une famille libre $\mathcal{L}' \subset \mathcal{G}$ telle que $\mathcal{L} \sqcup \mathcal{L}' = \mathcal{B}$ forme une base de V .

Preuve: (idée) Pour démontrer ce théorème, on utilise l'axiome du choix sous la forme équivalente suivante qu'on appelle

LEMME DE ZORN. *Soit E un ensemble ordonné tel que tout sous-ensemble $A \subset E$ totalement ordonné possède une majorant alors E possède un élément maximal.*

On applique le Lemme de Zorn à l'ensemble des familles libres de V ordonné par l'inclusion et on montre qu'une famille libre maximale pour l'inclusion est une base. $\square$

REMARQUE 5.4.2. En fait on peut montrer que le Lemme de Zorn et donc l'axiome du choix sont équivalents à l'existence d'une base pour tout espace vectoriel.

CHAPITRE 6

Applications lineaires

6.1. Le Theoreme Noyau-Image

6.1.1. Rang d'une application lineaire.

PROPOSITION 6.1. Soit $\varphi : V \mapsto W$ une application lineaire avec V de dimension finie. Soit $\mathcal{G} = \{\mathbf{e}_1, \dots, \mathbf{e}_g\} \subset V$ une famille generatrice alors φ est completement determinee par l'ensemble de images des elements de $\mathcal{G}$:

$$\varphi(\mathcal{G}) = \{\varphi(\mathbf{e}_1), \dots, \varphi(\mathbf{e}_g)\} \subset W.$$

En particulier, $\varphi(\mathcal{G})$ est une famille generatrice de $\text{Im}(\varphi) = \varphi(V)$ et on a

$$\dim(\text{Im } \varphi) \leq \dim(V).$$

Preuve: Soit $v \in V$, comme $\mathcal{G}$ est generatrice il existe $x_1, \dots, x_g \in K$ tels que

$$x_1 \cdot \mathbf{e}_1 + \dots + x_g \mathbf{e}_g = v$$

et alors

$$\varphi(v) = x_1 \cdot \varphi(\mathbf{e}_1) + \dots + x_g \varphi(\mathbf{e}_g).$$

Ainsi pour connaitre l'image d'un vecteur v il suffit de connaitre les vecteurs

$$\varphi(\mathbf{e}_1), \dots, \varphi(\mathbf{e}_g)$$

et une decomposition de v en combinaison lineaire d'elements de $\mathcal{G}$.

En particulier pour $w \in \text{Im}(\varphi)$, il existe $v \in V$ tel que $\varphi(v) = w$; ecrivant

$$x_1 \cdot \mathbf{e}_1 + \dots + x_g \mathbf{e}_g = v$$

on a

$$w = \varphi(v) = x_1 \cdot \varphi(\mathbf{e}_1) + \dots + x_g \varphi(\mathbf{e}_g)$$

Ainsi $\varphi(\mathcal{G})$ est generatrice de $\text{Im } \varphi$. En particulier $\text{Im } \varphi$ est de dimension finie et

$$\dim(\text{Im } \varphi) \leq |\varphi(\mathcal{G})|.$$

Ainsi en prenant pour $\mathcal{G}$ une base de V , on aura

$$\dim(\text{Im } \varphi) \leq |\varphi(\mathcal{G})| \leq |\mathcal{G}| = \dim(V).$$

□

DÉFINITION 6.1. Soit $\varphi : V \mapsto W$ une application lineaire. Le rang de φ est la dimension de $\text{Im } \varphi$:

$$\text{rg}(\varphi) = \dim(\text{Im } \varphi).$$

PROPOSITION 6.2 (Inegalite du rang). Soit V de dimension finie. On a

$$\text{rg}(\varphi) \leq \min(\dim V, \dim W).$$

Preuve: On vient de voir que $\text{rg}(\varphi) \leq \dim V$ et que $\text{rg}(\varphi) = \dim \text{Im } \varphi$ comme $\text{Im } \varphi$ est un sev de W on a

$$\text{rg}(\varphi) \leq \dim W.$$

□

REMARQUE 6.1.1. Cette inegalite reste vraie si V ou W sont de dimension infinie.

EXERCICE 6.1. Soient V, W deux espaces vectoriels de dimension finie et $\varphi : V \mapsto W$ une application lineaire. Montrer que

(1) Si φ est injective alors l'image par φ d'une famille libre est libre et

$$\dim(V) \leq \dim(W)$$

(2) Si φ est surjective alors l'image par φ d'une famille generatrice est generatrice et

$$\dim(V) \geq \dim(W).$$

(3) Si φ est bijective, l'image d'une base de V est une base de W et $\dim(V) = \dim(W)$.

EXERCICE 6.2. montrer qu'une application lineaire envoyant une base sur une base est un isomorphisme.

6.1.2. Le Theoreme Noyau-Image.

THEOREME 6.1 (Noyau-Image). Soit $\varphi : V \mapsto W$ une application lineaire avec V de dimension finie. On a

$$\dim V = \dim(\ker \varphi) + \dim(\operatorname{Im} \varphi).$$

Preuve: Notons que si $\mathcal{B}$ est une base alors $\varphi(\mathcal{B})$ est une partie generatrice de $\operatorname{Im} \varphi$ qui est donc de dimension finie de dimension

$$\dim \operatorname{Im} \varphi \leq |\varphi(\mathcal{B})| \leq |\mathcal{B}| = \dim(V).$$

Soit $\{\varphi(\mathbf{e}'_1), \dots, \varphi(\mathbf{e}'_r)\}$ une base de $\operatorname{Im} \varphi$ et $\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ une base de $\ker \varphi$. Montrons que

$$\{\mathbf{e}_1, \dots, \mathbf{e}_k, \mathbf{e}'_1, \dots, \mathbf{e}'_r\}$$

est une base de V . Supposons que

$$x_1 \mathbf{e}_1 + \dots + x_k \mathbf{e}_k + x'_1 \mathbf{e}'_1 + \dots + x'_r \mathbf{e}'_r = 0_V$$

alors

$$0_W = x'_1 \varphi(\mathbf{e}'_1) + \dots + x'_r \varphi(\mathbf{e}'_r)$$

et donc $x'_1 = \dots = x'_r = 0$. On a alors

$$x_1 \mathbf{e}_1 + \dots + x_k \mathbf{e}_k = 0_V$$

et donc $x_1 = \dots = x_k = 0$.

Soit $v \in V$ alors

$$\varphi(v) = x'_1 \varphi(\mathbf{e}'_1) + \dots + x'_r \varphi(\mathbf{e}'_r) = \varphi(x'_1 \mathbf{e}'_1 + \dots + x'_r \mathbf{e}'_r) = \varphi(v').$$

On a

$$\varphi(v - v') = 0_W \implies v - v' \in \ker \varphi$$

et donc

$$v - v' = x_1 \mathbf{e}_1 + \dots + x_k \mathbf{e}_k$$

et

$$v = x_1 \mathbf{e}_1 + \dots + x_k \mathbf{e}_k + x'_1 \mathbf{e}'_1 + \dots + x'_r \mathbf{e}'_r.$$

□

COROLLAIRE 6.1 (Critere de bijectivite). Soit $\varphi : V \mapsto W$ une application lineaire entre espaces de dimension finie. Si

$$\dim(V) = \dim(W)$$

alors est conditions suivantes sont equivalentes

(1) φ est injective.

(2) φ est surjective

(3) φ est bijective.

Preuve: Si φ est injective on a $\dim(\ker \varphi) = 0$ et

$$\dim(W) = \dim(V) = \dim(\operatorname{Im} \varphi) + 0$$

et donc $\dim(\operatorname{Im} \varphi) = \dim(W)$ ce qui implique que $W = \operatorname{Im} \varphi$ et la surjectivite et la bijectivite. Evidemment la bijectivite implique l'injectivite. $\square$

6.1.3. Exemple: les formes lineaires. On rappelle la definition d'un forme lineaire (cf Definition 5.12):

DÉFINITION 6.2. *Une forme lineaire sur V est une application lineaire de V a valeurs dans le corps K (vu comme K -ev sur lui-meme)*

$$\ell : V \mapsto K.$$

On a la proposition suivante:

PROPOSITION 6.3. *Soit ℓ une forme lineaire. Si elle est non-nulle, i.e. $\ell \neq \underline{0}_K$, alors*

$$\operatorname{Im}(\ell) = K, \dim(\ker \ell) = \dim(V) - 1.$$

Preuve: Soit $\ell \neq \underline{0}_K$. Soit $v \in V$ tel que $\ell(v) = \lambda \neq 0$; λ est donc inversible, alors pour tout $x \in K$, on a

$$\ell((x/\lambda).v) = (x/\lambda).\lambda = x$$

donc ℓ est surjective. Ainsi $\operatorname{Im} \ell = K$ est de dimension 1 et $\ker \ell$ est de diemsnion $\dim V - 1$. $\square$

DÉFINITION 6.3. *Soit V de dimension finie. Un sous-espace vectoriel de dimension $\dim V - 1$ est appelle un hyperplan vectoriel.*

PROPOSITION 6.4. *Soit V de dimension finie et $H \subset V$ un hyperplan vectoriel. Il existe une forme lineaire ℓ_H telle que*

$$\ker \ell_H = H.$$

Preuve: Soit $\{\mathbf{e}_1, \dots, \mathbf{e}_{d-1}\}$ une base de H . C'est une famille libre et on peut la completer en une base de V : il existe $\mathbf{e}_d \in V$ tel que

$$\{\mathbf{e}_1, \dots, \mathbf{e}_{d-1}, \mathbf{e}_d\}$$

forme une base de V . Considerons la forme lineaire d -ieme coordonnee:

$$\mathbf{e}_d^* : v = x_1 \mathbf{e}_1 + \dots + x_{d-1} \mathbf{e}_{d-1} + x_d \mathbf{e}_d \in V \mapsto x_d \in K.$$

Alors

$$H = \{v \in V, \mathbf{e}_d^*(v) = 0\}.$$

$\square$

REMARQUE 6.1.2. ℓ_H n'est pas unique: elle depend du choix de $\mathbf{e}_d$.

6.2. Structure et dimension des espaces d'applications lineaires

On rappelle que $(\operatorname{Hom}_{K-ev}(V, W), +, \cdot)$ a une structure naturelle de K -espace vectoriel, ou l'addition est donnee par

$$\varphi + \psi : v \mapsto \varphi(v) + \psi(v)$$

l'element neutre etant l'application identiquement nulle $\underline{0}_W$ et la multiplication externe, est donnee, pour pour $\lambda \in K$ and $\varphi \in \operatorname{Hom}_{K-ev}(V, W)$, par

$$\lambda.\varphi : v \mapsto \lambda.\varphi(v).$$

Rappelons que le fait que $\lambda.\varphi \in \operatorname{Hom}_{K-ev}(V, W)$ provient du fait que K est commutatif: pour $x \in K$

$$\lambda.\varphi(x.v + v') = \lambda(\varphi(x.v + v')) = \lambda(x.\varphi(v) + \varphi(v')) = x.\lambda.\varphi(v) + \lambda.\varphi(v') = x.(\lambda.\varphi)(v) + (\lambda.\varphi)(v').$$

THÉORÈME 6.2 (Dimension de l'espace des applications lineaires). *Si V et W sont de dimensions finies, alors $\text{Hom}_K(V, W)$ est de dimension finie*

$$\dim(\text{Hom}_K(V, W)) = \dim V \cdot \dim W.$$

Preuve: Soit $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_d\}$ une base de V . Soit φ une application lineaire, alors φ est entierement determinee des que l'on connait les valeurs des elements de $\mathcal{B}$

$$\varphi(\mathbf{e}_1), \dots, \varphi(\mathbf{e}_d) \in W.$$

En effet si $v = x_1 \cdot \mathbf{e}_1 + \dots + x_d \cdot \mathbf{e}_d$ alors

$$\varphi(v) = x_1 \cdot \varphi(\mathbf{e}_1) + \dots + x_d \cdot \varphi(\mathbf{e}_d).$$

En d'autres termes on dispose d'une application injective

$$\text{eval}_{\mathcal{B}} : \varphi \in \text{Hom}_K(V, W) \hookrightarrow (\varphi(\mathbf{e}_1), \dots, \varphi(\mathbf{e}_d)) \in W^d.$$

L'application $\text{eval}_{\mathcal{B}}$ est lineaire puisque pour tout $j \leq d$

$$(\lambda\varphi + \psi)(\mathbf{e}_j) = \lambda \cdot \varphi(\mathbf{e}_j) + \psi(\mathbf{e}_j)$$

Par ailleurs, cette application est surjective: soit un uplet

$$(f_1, \dots, f_d) \in W^d$$

alors on associe a $(f_1, \dots, f_d)$ l'application lineaire definie par

$$\varphi(x_1 \cdot \mathbf{e}_1 + \dots + x_d \cdot \mathbf{e}_d) = x_1 \cdot f_1 + \dots + x_d \cdot f_d.$$

Ainsi on a un isomorphisme

$$\text{eval}_{\mathcal{B}} : \text{Hom}_K(V, W) \simeq W^d$$

et (comme la dimension d'un produit d'EVs est la somme des dimensions)

$$\dim(\text{Hom}_{K-\text{ev}}(V, W)) = \dim(W^d) = d \cdot \dim(W).$$

□

On va maintenant decire une base de $\text{Hom}_K(V, W)$.

6.2.1. Formes lineaires, dualite et base duale.

On commence par l'espace des formes lineaires et on rappelle que

DÉFINITION 6.4. *Une application lineaire, $\ell : V \rightarrow K$, de V vers le corps K est appelee "forme lineaire". On note l'espace des formes lineaires par*

$$V^* := \text{Hom}_{K-\text{ev}}(V, K)$$

et on l'appelle le dual de V .

Comme $\dim K = 1$, on a

$$\dim(V^*) = \dim \text{Hom}_K(V, K) = \dim(V) \times 1 = \dim(V).$$

En particulier un espace vectoriel V et son dual V^* sont isomorphes. Pour trouver un tel isomorphisme, on va exhiber une base de V^* .

DÉFINITION 6.5. *Soit $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_d\}$ une base de V , si $v \in V$ s'ecrit*

$$v = x_1 \cdot \mathbf{e}_1 + \dots + x_d \cdot \mathbf{e}_d,$$

pour $i \leq d$, le scalaire x_i est la i -eme coordonnee de v dans la base $\mathcal{B}$. On note ce scalaire

$$x_i = \mathbf{e}_i^*(v).$$

PROPOSITION 6.5. *Pour $i \leq d$, l'application*

$$\mathbf{e}_i^* : v = x_1 \cdot \mathbf{e}_1 + \dots + x_d \cdot \mathbf{e}_d \in V \mapsto \mathbf{e}_i^*(v) = x_i \in K$$

est une forme lineaire. On l'appelle la i -ieme forme lineaire coordonnee relative a la base $\mathcal{B}$ de V .

Preuve: En effet, soit on dit que c'est la composee de deux application lineaires:

$$CL_{\mathcal{B}}^{-1} : \begin{array}{ccc} V & \mapsto & K^d \\ v = x_1.\mathbf{e}_1 + \cdots + x_d.\mathbf{e}_d & \mapsto & (x_1, \cdots, x_d) \end{array}$$

et

$$\bullet_i : \begin{array}{ccc} K^d & \mapsto & K \\ (x_1, \cdots, x_d) & \mapsto & x_i \end{array}$$

Soit on utilise directement le fait que la decomposition en combinaison lineaire est unique:

$$v = x_1.\mathbf{e}_1 + \cdots + x_d.\mathbf{e}_d, \quad v' = x'_1.\mathbf{e}_1 + \cdots + x'_d.\mathbf{e}_d$$

alors

$$\begin{aligned} \lambda.v + v' &= \lambda.x_1.\mathbf{e}_1 + \cdots + \lambda.x_d.\mathbf{e}_d + x'_1.\mathbf{e}_1 + \cdots + x'_d.\mathbf{e}_d \\ &= (\lambda.x_1 + x'_1).\mathbf{e}_1 + \cdots + (\lambda.x_d + x'_d).\mathbf{e}_d \end{aligned}$$

de sorte que par unicite la i -eme coordonnee de $\lambda.v + v'$ est $\lambda.x_i + x'_i$ □

Plus precisement, soit

$$\mathcal{B} = \{\mathbf{e}_1, \cdots, \mathbf{e}_d\}$$

une base de V , on a associe a chaque element $\mathbf{e}_i$ de cette base la forme lineaire " i -ieme coordonnee dans la base $\mathcal{B}$:"

$$\mathbf{e}_i^* : v = x_1.\mathbf{e}_1 + \cdots + x_i.\mathbf{e}_i + \cdots + x_d.\mathbf{e}_d \in V \mapsto x_i \in K.$$

THÉORÈME 6.3. Soit $\mathcal{B}$ une base de V , la famille

$$\mathcal{B}^* := \{\mathbf{e}_1^*, \cdots, \mathbf{e}_d^*\} \subset V^*$$

est une base de V^* . On a

$$\forall i, j \leq d, \quad \mathbf{e}_i^*(\mathbf{e}_j) = \delta_{i=j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}.$$

DÉFINITION 6.6. La base

$$\mathcal{B}^* := \{\mathbf{e}_1^*, \cdots, \mathbf{e}_d^*\} \subset V^*$$

s'appelle la base duale de la base $\mathcal{B}$.

Preuve: Pour $i \leq d$ on a

$$e_i = 1.e_i + \sum_{j \neq i} 0.e_j$$

de sorte que

$$e_i^*(e_i) = 1, \quad e_j^*(e_i) = 0.$$

Montrons que la famille $\mathcal{B}^*$ est libre (comme $\dim(V^*) = \dim(V) = d$ cela montrera qu'elle est generatrice). Supposons que

$$\ell := x_1.\mathbf{e}_1^* + \cdots + x_d.\mathbf{e}_d^* = \sum_{i=1}^d x_i.\mathbf{e}_i^* = \mathbf{0}_K.$$

On a pour $j \leq d$

$$0_K = \ell(\mathbf{e}_j) = \sum_{i=1}^d x_i.\mathbf{e}_i^*(\mathbf{e}_j) = \sum_{i=1}^d x_i.\delta_{i=j} = x_j.$$

□

On a montre que $\mathcal{B}^*$ est une base pour des raisons de cardinal et de dimension. En particulier c'est une famille generatrice et toute forme lineaire est combinaison lineaire des elements de $\mathcal{B}^*$:

COROLLAIRE 6.2. Soit $\ell : V \mapsto K$ une forme lineaire. On a

$$\ell = \sum_{i=1}^d \ell(\mathbf{e}_i) \mathbf{e}_i^*.$$

Autrement dit, les coordonnees de ℓ dans la base $\mathcal{B}^*$ sont donnees par les $(\ell(\mathbf{e}_i))_{i \leq d}$ (ie. les valeurs de ℓ en chacun des $\mathbf{e}_i$, $i \leq d$).

Preuve: On sait qu'il existe $l_i \in K$, $i \leq d$ tel que

$$\ell = \sum_{i \leq d} l_i \mathbf{e}_i^*.$$

Calculant $\ell(\mathbf{e}_i)$ on trouve

$$\ell(\mathbf{e}_i) = \sum_{j \leq d} l_j \mathbf{e}_j^*(\mathbf{e}_i) = \sum_{j \leq d} l_j \delta_{j=i} = l_i.$$

□

REMARQUE 6.2.1. Comment avoir l'idée de cette base duale: on a vu que l'application d' "evaluation le long de la base $\mathcal{B}$ ":

$$\begin{array}{ccc} \text{eval}_{\mathcal{B}} : V^* & \mapsto & K^d \\ \ell & \mapsto & (\ell(\mathbf{e}_1), \dots, \ell(\mathbf{e}_d)) \end{array}$$

est un isomorphisme lineaire.

On rappelle que dans l'espace d'arrivee K^d , on dispose d'une base preferee appelee *la base canonique* de K^d

$$\mathcal{B}_d^0 = \{\mathbf{e}_i^0, i \leq d\} \subset K^d;$$

avec $\mathbf{e}_i^0$ le vecteur dont la i -ieme coordonnee vaut 1 et les autres sont nulles:

$$\mathbf{e}_1^0 = (1, 0, \dots, 0), \dots, \mathbf{e}_d^0 = (0, \dots, 0, 1).$$

La base duale $\mathcal{B}^*$ est alors l'image reciproque par $\text{eval}_{\mathcal{B}}$ de la base canonique $\mathcal{B}_d^0$ de K^d .

Notons egalement que l'isomorphisme "combinaison lineaire dans la base $\mathcal{B}^*$ "

$$CL_{\mathcal{B}^*} : \begin{array}{ccc} K^d & \mapsto & V^* \\ (l_1, \dots, l_d) & \mapsto & l_1 \mathbf{e}_1^* + \dots + l_d \mathbf{e}_d^* \end{array}$$

est l'isomorphisme reciproque de l'isomorphisme $\text{eval}_{\mathcal{B}}$.

REMARQUE 6.2.2. On a deux isomorphismes

$$\text{eval}_{\mathcal{B}} : V^* \simeq K^d, \quad CL_{\mathcal{B}} : K^d \simeq V$$

et donc un isomorphisme "explicite"

$$CL_{\mathcal{B}} \circ \text{eval}_{\mathcal{B}} : V^* \simeq V$$

entre le dual V^* et V . Il faut noter que cet isomorphisme depend du choix de la base $\mathcal{B}$.

EXERCICE 6.3. Soit $V^{**} = (V^*)^*$ le bi-dual de V (le dual du dual V^* de V). On considere l'application:

$$\text{eval}_{\bullet} : \begin{array}{ccc} V & \mapsto & V^{**} = (V^*)^* \\ v & \mapsto & \text{eval}_v \end{array}$$

ou

$$\text{eval}_v : \ell \mapsto \ell(v) \in K$$

est l'application qui a une forme lineaire ℓ associe sa valeur au vecteur v .

- (1) Montrer que eval_v est bien une forme lineaire sur V^* .
- (2) Montrer que $\text{eval}_{\bullet}$ est un isomorphisme.

- (3) Montrer que si on identifie V^{**} à V par l'isomorphisme ci-dessus et que $\mathcal{B} = \{\mathbf{e}_i, i \leq d\}$ est une base de V , la base duale de la base duale

$$\mathcal{B}^{**} = \{(\mathbf{e}_i^*)^*, i = 1, \dots, d\}$$

vaut $\mathcal{B}$.

REMARQUE 6.2.3. A la difference de l'isomorphisme $CL_{\mathcal{B}} \circ \text{eval}_{\mathcal{B}} : V^* \simeq V$ qui depend du choix d'une base. L'isomorphisme $\text{eval}_{\bullet} : V \simeq V^{**}$ n'en depend pas. On dit que le bidual de V est canoniquement isomorphe à V .

REMARQUE 6.2.4. L'application

$$\langle \bullet, \bullet \rangle_{\text{can}, V} : (\ell, v) \in V^* \times V \mapsto \ell(v) \in K$$

est une forme lineaire¹ sur $V^* \times V$ appelee *accouplement canonique* entre V^* et V . Le fait qu'il permette de definir un isomorphisme entre V^* et V fait dire que c'est un *accouplement parfait*.

6.2.2. Representation parametrique et cartesienne d'un SEV. Soit $W \subset V$ un SEV d'un espace vectoriel de dimension finie $d_V = \dim V$ alors W est de dimension finie $d_W = \dim W$.

Soit $\mathcal{G}_W = \{\mathbf{e}_1, \dots, \mathbf{e}_g\}$, $g \geq d_W$ une famille generatrice de W : W est l'ensemble des vecteurs de v de la forme

$$W = \{w \in V, w = x_1 \cdot \mathbf{e}_1 + \dots + x_g \cdot \mathbf{e}_g, x_1, \dots, x_g \in K\}$$

Une telle presentation de W s'appelle une *representation parametrique* de W : chaque vecteur $w \in W$ est obtenu comme somme de vecteurs de la forme

$$x_1 \cdot \mathbf{e}_1 + \dots + x_g \cdot \mathbf{e}_g$$

pour un choix approprie (pas unique en general) de parametres scalaires $x_1, \dots, x_g \in K$. En particulier si $\mathcal{G}_W = \mathcal{B}_W$ est une base de W le nombre de vecteurs $\{\mathbf{e}_i, i \leq g\}$ impliquees dans cette representation est minimal et vaut d_W ; la representation precedente est alors unique.

Par ailleurs un SEV W peut egalement etre represente comme l'ensemble des solutions d'un systeme d'equations lineaires (de second membre nul):

PROPOSITION 6.6 (Representation cartesienne d'un SEV). *Soit $W \subset V$ un SEV (distinct de V). Il existe un entier $d' \geq 1$ et une famille de d' formes lineaires*

$$\mathcal{L} = \{\ell_1, \dots, \ell_{d'}\} \subset V^*$$

telles que

$$W = \{w \in V \text{ tels que } \ell_1(w) = 0, \ell_2(w) = 0, \dots, \ell_{d'}(w) = 0\}.$$

De maniere equivalente, $W = \ker \varphi_{\mathcal{L}}$ avec

$$\varphi_{\mathcal{L}} : w \in V \mapsto (\ell_1(w), \dots, \ell_{d'}(w)) \in K^{d'}.$$

En fait on peut prendre $d' = d_V - d_W$ et la famille

$$\mathcal{L} = \{\ell_1, \dots, \ell_{d_V - d_W}\} \subset V^*$$

forment une famille libre de V^* (ie. les ℓ_i , $i \leq d_V - d_W$ sont lineairement independantes).

Preuve: Soit $\mathcal{B}_W = \{\mathbf{e}_1, \dots, \mathbf{e}_{d_W}\}$ une base de W et

$$\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_{d_W}, \mathbf{e}_{d_W+1}, \dots, \mathbf{e}_{d_V}\}$$

une base de V contenant la base precedente. Soit

$$\mathcal{B}^* = \{\mathbf{e}_1^*, \dots, \mathbf{e}_{d_W}^*, \mathbf{e}_{d_W+1}^*, \dots, \mathbf{e}_{d_V}^*\}$$

la base duale. Alors

$$W = \{v \in V, \mathbf{e}_{d_W+1}^*(v) = \dots = \mathbf{e}_{d_V}^*(v) = 0\}$$

□

¹en fait elle est meme bi-lineaire

La representation

$$W = \{v \in V, \ell_1(v) = \cdots = \ell_{d_V - d_W}(v) = 0\}$$

est appelee representation cartesienne de W d'equations

$$\ell_1(v) = 0, \cdots, \ell_{d_V - d_W}(v) = 0.$$

REMARQUE 6.2.5. Le nombre d' d'equations d'une representation cartesienne est toujours au moins egal a $d_V - d_W$. En effet si $\mathcal{L} = \{\ell_1, \cdots, \ell_{d'}\}$ verifie

$$W = \{v \in V, \ell_1(v) = \cdots = \ell_{d'}(v) = 0\}$$

cela signifie que W est le noyau de l'application lineaire

$$\text{eval}_{\mathcal{L}} : v \in V \mapsto (\ell_1(v), \cdots, \ell_{d'}(v)) \in K^{d'}.$$

On a donc

$$\dim V - \dim W = \dim V - \dim \ker(\text{eval}_{\mathcal{L}}) = \dim(\text{eval}_{\mathcal{L}}(V)) \leq \dim(K^{d'}) = d'$$

EXERCICE 6.4. Dans $\mathbb{Q}^3$, soit $W = \langle (1, 1, 0), (1, 0, 3) \rangle$. Donner une equation cartesienne de W .

EXERCICE 6.5. Dans $\mathbb{Q}^3$, soit $W = \{(x, y, z) \in \mathbb{Q}^3, x + y - z = 0, x - 2y + 3z = 0\}$. Donner une representation parametrique de W .

6.2.3. Une base de $\text{Hom}(V, W)$. Soient V et W des EVs de dimensions finies d et d' .

On a vu que

$$\dim \text{Hom}(V, W) = \dim(W^d) = \dim V \dim W.$$

on va donner une base explicite de cet espace.

Etant donne $\mathcal{B} = \{\mathbf{e}_1, \cdots, \mathbf{e}_d\}$ et $\mathcal{B}' = \{\mathbf{f}_1, \cdots, \mathbf{f}_{d'}\}$ des bases de V et W , on va construire une base de $\text{Hom}(V, W)$: soit

$$\mathcal{B}^* = \{\mathbf{e}_1^*, \cdots, \mathbf{e}_d^*\}$$

la base duale de $\mathcal{B}$, et definissons pour $i \in \{1, \cdots, d'\}$, $j \in \{1, \cdots, d\}$ l'application

$$\mathcal{E}_{ij} : \begin{array}{ccc} V & \mapsto & W \\ v & \mapsto & \mathbf{e}_j^*(v) \cdot \mathbf{f}_i. \end{array}$$

En d'autre termes, si

$$v = x_1 \cdot \mathbf{e}_1 + \cdots + x_d \cdot \mathbf{e}_d,$$

$\mathcal{E}_{ij}(v)$ est egal a $x_j \cdot \mathbf{f}_i$, cad le produit de la j -eme coordonnee de v , x_j dans la base $\mathcal{B}$ et du i -ieme vecteur de la base $\mathcal{B}'$.

En particulier on a pour $k = 1, \cdots, d$

$$\mathcal{E}_{ij}(\mathbf{e}_k) = \begin{cases} \mathbf{f}_i & \text{si } k = j \\ 0_W & \text{si } k \neq j \end{cases}.$$

LEMME 6.1. L'application $\mathcal{E}_{ij} : V \mapsto W$ est lineaire, de rang 1, d'image $K \cdot \mathbf{f}_i$ et de noyau

$$\ker \mathcal{E}_{ij} = \langle \mathcal{B} - \{\mathbf{e}_j\} \rangle = K \cdot \mathbf{e}_1 + \cdots + K \cdot \mathbf{e}_{j-1} + K \cdot \mathbf{e}_{j+1} + \cdots + K \cdot \mathbf{e}_d$$

l'hyperplan vectoriel engendre par les vecteurs de la base $\mathcal{B}$ moins le vecteur $\mathbf{e}_j$.

Preuve: Comme $\mathbf{e}_j^*$ est lineaire on a

$$\mathcal{E}_{ij}(\lambda \cdot v + v') = \mathbf{e}_j^*(\lambda \cdot v + v') \cdot \mathbf{f}_i = (\lambda \cdot x_j + x'_j) \cdot \mathbf{f}_i = \lambda \cdot x_j \cdot \mathbf{f}_i + x'_j \cdot \mathbf{f}_i = \lambda \mathcal{E}_{ij}(v) + \mathcal{E}_{ij}(v').$$

Il est clair que $\text{Im } \mathcal{E}_{ij} \subset K \cdot \mathbf{f}_i$ et comme $\mathcal{E}_{ij}(\mathbf{e}_j) = \mathbf{f}_i$ on a egalite. Ainsi $\text{rg}(\mathcal{E}_{ij}) = 1$ ($\mathbf{f}_i \neq 0_W$, ce vecteur etant dans une base).

Par ailleurs ($\mathbf{f}_i \neq 0_W$) il est clair que $\mathcal{E}_{ij}(v) = x_j \cdot \mathbf{f}_i = 0_W$ si et seulement si la j -eme coordonnee x_j de v dans la base $\mathcal{B}$ est nulle. $\square$

DÉFINITION 6.7. Soit V, W des K -EV de dimensions finies d, d' et

$$\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_d\} \text{ et } \mathcal{B}' = \{\mathbf{f}_1, \dots, \mathbf{f}_{d'}\}$$

des bases de V et W et $\mathcal{B}^* = \{\mathbf{e}_1^*, \dots, \mathbf{e}_d^*\} \subset V^*$ la base duale de $\mathcal{B}$.

Pour $i \leq d', j \leq d$ les applications lineaires definies par

$$\mathcal{E}_{i,j} : v \in V \mapsto \mathbf{e}_j^*(v) \cdot \mathbf{f}_i \in W$$

sont appelees applications lineaires elementaires associees aux bases $\mathcal{B}$ et $\mathcal{B}'$.

THÉORÈME 6.4 (Une base de l'espace des applications lineaires). La famille des applications lineaires elementaires

$$\mathcal{B}_{\mathcal{B}', \mathcal{B}} := \{\mathcal{E}_{i,j}, i \leq d', j \leq d\} \subset \text{Hom}_{K\text{-ev}}(V, W)$$

forme une base de $\text{Hom}_{K\text{-ev}}(V, W)$.

Preuve: Comme le cardinal de cette famille vaut $\dim(V) \dim(W) = \dim \text{Hom}_{K\text{-ev}}(V, W)$ il suffit de montrer qu'elle est libre: soit $m_{ij} \in K, i \leq d', j \leq d$ des scalaires tels que

$$\sum_{i,j} m_{ij} \mathcal{E}_{i,j} = \mathbf{0}_W.$$

On a donc pour chaque $k \leq d$

$$\left(\sum_{i,j} m_{ij} \mathcal{E}_{i,j} \right) (\mathbf{e}_k) = \sum_i m_{ik} \mathbf{f}_i = \mathbf{0}_W.$$

Comme $\mathcal{B}'$ est une base de W on a pour tout $i \leq d'$,

$$m_{ik} = 0$$

et donc pour tout i, j on a $m_{ij} = 0$. □

6.2.3.1. *Preuve directe que $(\mathcal{E}_{i,j})_{i,j}$ est generatrice.* On peut en fait montrer directement (sans utiliser la dimension) que $\mathcal{B}_{\mathcal{B}', \mathcal{B}}$ est generatrice: soit $\varphi : V \mapsto W$ une application lineaire, on cherche a trouver d, d' scalaires $(m_{i,j})_{i \leq d', j \leq d}$ tels que

$$\varphi = \sum_{i,j} m_{i,j} \mathcal{E}_{i,j} = \sum_{i,j} m_{i,j} \mathbf{e}_j^* \cdot \mathbf{f}_i.$$

Supposons qu'on dispose d'une telle de composition et calculons pour $k \leq d$

$$\varphi(\mathbf{e}_k) = \sum_{i,j} m_{i,j} \mathbf{e}_j^*(\mathbf{e}_k) \cdot \mathbf{f}_i = \sum_i m_{i,k} \mathbf{f}_i$$

et donc pour $i \leq d'$, $m_{i,k}$ est la i -ieme coordonnee de $\varphi(\mathbf{e}_k)$ dans la base $\mathcal{B}'$:

$$m_{i,k} = \mathbf{f}_i^*(\varphi(\mathbf{e}_k)).$$

Considerons alors la combinaison lineaire d'applications elementaires

$$\varphi' = \sum_{i,j} \mathbf{f}_i^*(\varphi(\mathbf{e}_j)) \mathcal{E}_{i,j}.$$

La raisonnement precedent montre que pour tout $\mathbf{e}_k \in \mathcal{B}$ on a

$$\varphi(\mathbf{e}_k) = \varphi'(\mathbf{e}_k).$$

Comme les deux applications lineaires prennent les memes valeurs sur une famille generatrice, elles sont egales: on a donc

$$(6.2.1) \quad \varphi = \sum_{i,j} \mathbf{f}_i^*(\varphi(\mathbf{e}_j)) \mathcal{E}_{i,j} = \sum_{i,j} m_{i,j} \mathcal{E}_{i,j}$$

avec

$$m_{i,j} = \mathbf{f}_i^*(\varphi(\mathbf{e}_j)).$$

REMARQUE 6.2.6. Comme la notation l'indique $\mathcal{B}_{\mathcal{B}', \mathcal{B}}$ depend du choix d'une base de $\mathcal{B}$ et d'une base de $\mathcal{B}'$. Les applications $\mathcal{E}_{ij}$ sont appelees *applications elementaires* associees aux bases $\mathcal{B}$ et $\mathcal{B}'$.

EXEMPLE 6.2.1. Soit $V = \mathbb{R}^3$, $W = \mathbb{R}^2$ et prenons les bases canoniques

$$\mathcal{B} = \mathcal{B}_3^0 = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}, \quad \mathcal{B}' = \mathcal{B}_2^0 = \{(1, 0), (0, 1)\}.$$

On dispose de 6 applications lineaires elementaires

$$\mathcal{E}_{11}, \mathcal{E}_{12}, \mathcal{E}_{13}, \mathcal{E}_{21}, \mathcal{E}_{22}, \mathcal{E}_{23}$$

et par exemple

$$\mathcal{E}_{12}(x, y, z) = y(1, 0) = (y, 0), \quad \mathcal{E}_{23}(x, y, z) = z(0, 1) = (0, z).$$

Soit l'application lineaire de $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ donnee par

$$\varphi(x, y, z) = (2x + 4y, y + 3z)$$

alors

$$\varphi = 2\mathcal{E}_{11} + 4\mathcal{E}_{12} + \mathcal{E}_{22} + 3\mathcal{E}_{23}.$$

DÉFINITION 6.8. L'ensemble des d, d' scalaires $(m_{i,j})_{i \leq d', j \leq d}$ donnees par

$$(6.2.2) \quad m_{i,j} = \mathbf{f}_i^*(\varphi(\mathbf{e}_j)).$$

sont les coefficients de φ dans la base $\mathcal{B}_{\mathcal{B}', \mathcal{B}}$ ou encore la matrice de φ relative aux bases $\mathcal{B}$, $\mathcal{B}'$.

6.3. Proprietes fonctionnelles des coefficients d'une application lineaire

Dans cette section on va voir comment la donnee des coefficients (relative a des bases choisies) d'une application lineaire permet de faire des calculs effectifs.

6.3.1. Image d'un vecteur. Soient V, W de dimensions d, d' finies et de bases

$$\mathcal{B} = \{\mathbf{e}_j, j \leq d\}, \mathcal{B}' = \{\mathbf{f}_i, i \leq d'\}.$$

Soit

$$\mathcal{B}_{\mathcal{B}', \mathcal{B}} = \{\mathcal{E}_{ij} = \mathbf{e}_j^* \cdot \mathbf{f}_i, i \leq d', j \leq d\} \subset \text{Hom}_{K-év}(V, W)$$

la base de l'espace des application lineaires formee des applications elementaires.

PROPOSITION 6.7. Soit $\varphi : V \rightarrow W$ une application lineaire et $(m_{ij})_{i \leq d', j \leq d}$ les coordonnees de φ dans la base $\mathcal{B}_{\mathcal{B}', \mathcal{B}}$. Alors pour $k = 1, \dots, d$ le d' -uplet

$$(m_{i,k})_{i \leq d'}$$

sont les coordonnees de $\varphi(\mathbf{e}_k)$ dans la base $\mathcal{B}'$:

$$(6.3.1) \quad \varphi(\mathbf{e}_k) = \sum_{i \leq d'} m_{ik} \mathbf{f}_i.$$

Preuve: On a

$$\varphi(\mathbf{e}_k) = \left(\sum_{i,j} m_{ij} \mathcal{E}_{ij} \right) (\mathbf{e}_k) = \sum_{i,j} m_{ij} \mathcal{E}_{ij}(\mathbf{e}_k) = \sum_{i \leq d'} m_{ik} \mathbf{f}_i.$$

□

Soit $v \in V$ un vecteur de coordonnees $(x_j)_{j \leq d}$ dans la base $\mathcal{B}$. Calculons les coordonnees $(y_i)_{i \leq d'}$ de $\varphi(v) \in W$ dans la base $\mathcal{B}'$:

PROPOSITION 6.8. Avec les notations precedentes, si $v = \sum_{j=1}^d x_j \mathbf{e}_j$, on a

$$\varphi(v) = \sum_{i=1}^{d'} y_i \mathbf{f}_i \text{ avec } y_i = \sum_{j=1}^d m_{ij} x_j.$$

Preuve: on a

$$v = \sum_{j \leq d} x_j \mathbf{e}_j, \quad \varphi(v) = \sum_{i \leq d'} y_i \mathbf{f}_i$$

et

$$\varphi(\mathbf{e}_j) = \sum_{i \leq d'} m_{ij} \mathbf{f}_i.$$

Ainsi on a

$$\varphi(v) = \sum_{j \leq d} x_j \varphi(\mathbf{e}_j) = \sum_{j \leq d} x_j \left(\sum_{i \leq d'} m_{ij} \mathbf{f}_i \right) = \sum_{i \leq d'} \left(\sum_{j \leq d} m_{ij} x_j \right) \mathbf{f}_i$$

On a donc

$$y_i = \sum_{j \leq d} m_{ij} x_j.$$

□

6.3.2. Combinaison lineaire d'applications lineaires.

PROPOSITION 6.9. *Soit*

$$\varphi, \psi : V \mapsto W$$

deux applications lineaires et $(m_{ij})_{i \leq d', j \leq d}$, $(n_{ij})_{i \leq d', j \leq d}$ leurs coordonnees dans la base $\mathcal{B}_{\mathcal{B}', \mathcal{B}}$. Pour tout $\lambda \in K$, $\lambda \varphi + \psi$ est lineaire et ses coordonnees dans la base $\mathcal{B}_{\mathcal{B}', \mathcal{B}}$ sont donnees par

$$(\lambda m_{ij} + n_{ij})_{i \leq d', j \leq d}.$$

Preuve: En effet pour tout EV E et toute base $\mathcal{B}_E$ de E et tout vecteur $\mathbf{g} \in \mathcal{B}_E$ de cette base, la fonction coordonnee $\mathbf{g}^* : E \mapsto K$ qui a un element associe sa coordonne suivant le vecteur $\mathbf{g}$ est une forme lineaire. On applique cela a $\text{Hom}(V, W)$ et aux vecteurs de la base $\mathcal{B}_{\mathcal{B}', \mathcal{B}}$.

Alternativement on a la formule

$$m_{ij}(\varphi) = \mathbf{f}_i^*(\varphi(\mathbf{e}_j))$$

et l'application

$$\varphi \mapsto \mathbf{f}_i^*(\varphi(\mathbf{e}_j)) \in K$$

est lineaire:

$$\begin{aligned} m_{ij}(\lambda \varphi + \psi) &= \mathbf{f}_i^*((\lambda \varphi + \psi)(\mathbf{e}_j)) = \mathbf{f}_i^*(\lambda \varphi(\mathbf{e}_j) + \psi(\mathbf{e}_j)) = \\ &= \lambda \mathbf{f}_i^*(\varphi(\mathbf{e}_j)) + \mathbf{f}_i^*(\psi(\mathbf{e}_j)) = \lambda m_{ij}(\varphi) + m_{ij}(\psi). \end{aligned}$$

□

6.3.3. Composition d'applications lineaires. Soient U, V, W trois espaces vectoriels. Soient deux applications lineaires

$$\varphi : U \mapsto V, \quad \psi : V \mapsto W \quad \text{et} \quad \psi \circ \varphi : U \mapsto W$$

leur composee. Soient

$$\mathcal{B} = \{\mathbf{e}_k, k \leq d\}, \mathcal{B}' = \{\mathbf{f}_j, j \leq d'\}, \mathcal{B}'' = \{\mathbf{g}_i, i \leq d''\}$$

des bases de U, V et W , on dispose alors de bases

$$\mathcal{B}_{\mathcal{B}', \mathcal{B}} = \{\mathbf{e}_k^* \cdot \mathbf{f}_j\}, \mathcal{B}_{\mathcal{B}'', \mathcal{B}'} = \{\mathbf{f}_j^* \cdot \mathbf{g}_i\}, \mathcal{B}_{\mathcal{B}'', \mathcal{B}} = \{\mathbf{e}_k^* \cdot \mathbf{g}_i\}$$

pour

$$\text{Hom}(U, V), \text{Hom}(V, W), \text{Hom}(U, W),$$

THÉORÈME 6.5. Soient $(n_{jk})_{j \leq d', k \leq d}$ les coordonnées de φ dans la base $\mathcal{B}_{\mathcal{B}', \mathcal{B}}$ et $(m_{ij})_{i \leq d'', j \leq d'}$ les coordonnées de ψ dans la base $\mathcal{B}_{\mathcal{B}'', \mathcal{B}'}$. Alors les coordonnées $(l_{ik})_{i \leq d'', k \leq d}$ de $\psi \circ \varphi$ dans la base $\mathcal{B}_{\mathcal{B}'', \mathcal{B}}$ sont données par

$$l_{ik} = \sum_{j=1}^{d'} m_{ij} \cdot n_{jk}.$$

Preuve: Ecrivons

$$\varphi = \sum_{j \leq d'} \sum_{k \leq d} n_{jk} \mathbf{e}_k^* \cdot \mathbf{f}_j, \quad \psi = \sum_{j \leq d'} \sum_{i \leq d''} m_{ij} \mathbf{f}_j^* \cdot \mathbf{g}_i.$$

On a pour tout $k \leq d$ et $j \leq d'$

$$\varphi(\mathbf{e}_k) = \sum_{j \leq d'} n_{jk} \mathbf{f}_j, \quad \psi(\mathbf{f}_j) = \sum_{i \leq d''} m_{ij} \mathbf{g}_i$$

et

$$\psi(\varphi(\mathbf{e}_k)) = \sum_{j \leq d'} n_{jk} \psi(\mathbf{f}_j) = \sum_{j \leq d'} n_{jk} \sum_{i \leq d''} m_{ij} \mathbf{g}_i = \sum_{i \leq d''} \left(\sum_{j \leq d'} m_{ij} n_{jk} \right) \cdot \mathbf{g}_i = \sum_{i \leq d''} l_{ik} \cdot \mathbf{g}_i$$

Ainsi

$$l_{ik} = \sum_{j \leq d'} m_{ij} n_{jk}.$$

□

6.3.4. Application lineaire duale. Soit $\varphi : V \mapsto W$ une application lineaire et $\ell' : W \rightarrow K$ une forme lineaire. Alors la composee

$$\ell' \circ \varphi : v \in V \rightarrow \ell'(\varphi(v)) \in K$$

est une forme lineaire sur V . On la note

$$\varphi^*(\ell') := \ell' \circ \varphi.$$

En effet $\varphi^*(\ell')$ est a valeurs dans K et est lineaire comme composee de deux applications lineaires.

REMARQUE 6.3.1. Avec nos notations, on a la formule dite d' *adjonction* : pour tout $v \in V$, $\ell' \in W^*$ on a

$$(6.3.2) \quad \ell'(\varphi(v)) = \varphi^*(\ell')(v).$$

Ainsi a toute forme lineaire $\ell' \in W^*$ on a associe une forme lineaire $\varphi^*(\ell') \in V^*$ a l'aide de φ .

DÉFINITION 6.9. Soit $\varphi : V \mapsto W$ une application lineaire. L'application duale φ^* de φ est l'application

$$\varphi^* : W^* \mapsto V^*$$

qui associe a une forme lineaire $\ell' : w \in W \mapsto \ell'(w) \in K$, la forme lineaire sur V obtenue par pre-composition par φ :

$$\varphi^*(\ell') := \ell' \circ \varphi : \begin{array}{ccc} V & \mapsto & K \\ v & \mapsto & \ell'(\varphi(v)) \end{array}.$$

EXEMPLE 6.3.1. Soit $U \subset V$ un SEV d'un EV V alors l'injection

$$\iota_U : u \in U \hookrightarrow u \in V$$

est lineaire et son application lineaire duale

$$\iota_U^* = \ell|_U : \ell \in V^* \mapsto \ell|_U \in U^*$$

est simplement la restriction de ℓ a U :

$$\iota_U^*(\ell)(u) = \ell(\iota_U(u)) = \ell(u).$$

PROPOSITION 6.10. *L'application duale*

$$\varphi^* : \ell' \in W^* \mapsto \ell \circ \varphi \in V^*$$

est lineaire:

$$\varphi^* \in \text{Hom}_K(W^*, V^*).$$

Preuve: Soit $\ell'_1, \ell'_2 \in W^*$ et $\lambda \in K$, on veut montrer que

$$\varphi^*(\lambda \ell'_1 + \ell'_2) = \lambda \varphi^*(\ell'_1) + \varphi^*(\ell'_2).$$

Pour tout $v \in V$ on a

$$\varphi^*(\lambda \ell'_1 + \ell'_2)(v) = (\lambda \ell'_1 + \ell'_2)(\varphi(v)) = \lambda \ell'_1(\varphi(v)) + \ell'_2(\varphi(v)) = \lambda \varphi^*(\ell'_1)(v) + \varphi^*(\ell'_2)(v).$$

□

On laisse en exercice la preuve des proprietes fonctionnelles de l'application duale.

EXERCICE 6.6. Soit $\varphi : V \mapsto W$ une application lineaire entre deux espaces de dimensions finies.

(1) (Linearite) Montrer que l'application

$$\bullet^* : \varphi \in \text{Hom}(V, W) \mapsto \varphi^* \in \text{Hom}(W^*, V^*)$$

qui a une application lineaire associe l'application lineaire duale est elle meme lineaire:

$$(\lambda \varphi + \varphi')^* = \lambda \varphi^* + \varphi'^*$$

En d'autres termes

$$\bullet^* \in \text{Hom}(\text{Hom}(V, W), \text{Hom}(W^*, V^*)).$$

(2) (Anti-morphisme) Soit $\psi : W \mapsto Z$. Montrer que

$$(\psi \circ \varphi)^* = \varphi^* \circ \psi^*.$$

(3) (Involutive) Montrer que si le bi-dual V^{**} est identifie (canoniquement) a V via l'isomorphisme

$$\text{eval}_\bullet : v \in V \mapsto (\ell \mapsto \ell(v)) \in V^{**}$$

alors la duale de la duale qu'une application est l'application elle-meme:

$$(\varphi^*)^* = \varphi.$$

REMARQUE 6.3.2. La propriete d'adjonction s'ecrit de la maniere suivante en termes des accouplements canoniques de V et W : rappelons les notations

$$\langle \bullet, \bullet \rangle_{\text{can}, V} : (\ell, v) \mapsto \langle \ell, v \rangle_{\text{can}, V} = \ell(v) \in K.$$

$$\langle \bullet, \bullet \rangle_{\text{can}, W} : (\ell', w) \mapsto \langle \ell', w \rangle_{\text{can}, W} = \ell'(w) \in K.$$

On a pour $v \in V, \ell' \in W^*$

$$\langle \ell', \varphi(v) \rangle_{\text{can}, W} = \langle \varphi^*(\ell'), v \rangle_{\text{can}, V}$$

Le resultat suivant calcule les coefficient de l'application duale.

THEOREME 6.6. Soit $\varphi : V \mapsto W$ une application lineaire et $\varphi^* : W^* \rightarrow V^*$ l'application lineaire duale; soient $\mathcal{B}$ et $\mathcal{B}'$ des bases de V et V' et $(m_{ij})_{i \leq d', j \leq d}$ les coefficients de φ dans la base $\mathcal{B}_{\mathcal{B}', \mathcal{B}}$; soient $(m_{ji}^*)_{j \leq d, i \leq d'}$ les coefficients de φ^* dans la base

$$\mathcal{B}_{\mathcal{B}^*, \mathcal{B}'^*} \subset \text{Hom}(W^*, V^*)$$

associee aux bases duales $\mathcal{B}^* \subset V^*$ et $\mathcal{B}'^* \subset W^*$. On a

$$m_{ji}^* = m_{ij}, \quad i \leq d', \quad j \leq d.$$

Preuve: Soient $(m_{ji}^*)_{j \leq d, i \leq d'}$ les coefficients de φ^* relatifs aux bases $\mathcal{B}^*, \mathcal{B}'^*$. Par la formule generale (6.3.1) appliquees a φ^* , on a pour $i = 1, \dots, d'$

$$\varphi^*(\mathbf{f}_i^*) = \sum_{j=1}^d m_{ji}^* \mathbf{e}_j^*.$$

On va calculer les m_{ji}^* en evaluant cette forme lineaire $\varphi^*(\mathbf{f}_i^*)$ sur les vecteurs $\mathbf{e}_{j'}$, $j' \leq d$: on a d'une part (par definition de l'application duale)

$$\varphi^*(\mathbf{f}_i^*)(\mathbf{e}_{j'}) = \mathbf{f}_i^*(\varphi(\mathbf{e}_{j'})) = \mathbf{f}_i^*\left(\sum_{i'=1}^{d'} m_{i'j'} \mathbf{f}_{i'}\right) = \sum_{i'=1}^{d'} m_{i'j'} \mathbf{f}_i^*(\mathbf{f}_{i'}) = m_{ij'}$$

car $\mathbf{f}_i^*(\mathbf{f}_{i'}) = \delta_{i=i'}$ et donc un seul terme survit dans la somme precedente. D'autre part, on a

$$\varphi^*(\mathbf{f}_i^*)(\mathbf{e}_{j'}) = \sum_{j=1}^d m_{ji}^* \mathbf{e}_j^*(\mathbf{e}_{j'}) = m_{j'i}^*$$

car $\mathbf{e}_j^*(\mathbf{e}_{j'}) = \delta_{j=j'}$ et donc un seul terme survit dans la somme precedente. Ainsi pour tout $i \leq d', j' \leq d$ on a

$$m_{j'i}^* = m_{ij'}.$$

□

REMARQUE 6.3.3. Voici une autre presentation de la meme preuve si on est a l'aise avec le bidual. On a vu que si on identifie V^{**} a V via l'isomorphisme

$$\text{eval}_\bullet : v \mapsto \text{eval}_v : \ell \mapsto \ell(v),$$

alors la base duale de la base duale est la base elle-meme:

$$\mathcal{B}^{**} = \mathcal{B}, \mathcal{B}'^{**} = \mathcal{B}'.$$

On a vu egalement que

$$m_{j,i}^* = \mathbf{e}_j^{**}(\varphi^*(\mathbf{f}_i^*)).$$

Par definition de $\mathbf{e}_j^{**}$, puis de φ^* on a

$$\mathbf{e}_j^{**}(\varphi^*(\mathbf{f}_i^*)) = \varphi^*(\mathbf{f}_i^*)(\mathbf{e}_j) = \mathbf{f}_i^*(\varphi(\mathbf{e}_j)) = m_{i,j}.$$

THÉORÈME 6.7 (Rang de l'application duale). Soit $\varphi : V \mapsto W$ une application lineaire et $\varphi^* : W^* \mapsto V^*$ sa duale, alors on a

$$\text{rg}(\varphi) = \dim(\text{Im } \varphi) = \dim(\text{Im } \varphi^*) = \text{rg}(\varphi^*).$$

Preuve: Soit $r = \dim(\text{Im } \varphi)$ et

$$\{\mathbf{f}_1 = \varphi(\mathbf{e}_1), \dots, \mathbf{f}_r = \varphi(\mathbf{e}_r)\} \subset W$$

une base de $\text{Im } \varphi$. On complete cette base en une base de W

$$\mathcal{B}' = \{\mathbf{f}_i, i \leq d'\} \subset W.$$

D'autre part on a vu dans la preuve du Thm Noyau-Image que si

$$\{\mathbf{e}_{r+1}, \dots, \mathbf{e}_{d-r}\} \subset \ker(\varphi)$$

est une base du noyau de φ alors

$$\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_r, \mathbf{e}_{r+1}, \dots, \mathbf{e}_d\}$$

est une base de V .

On a

$$\text{rg}(\varphi^*) = \dim\{\text{Vect}\{\varphi^*(\mathbf{f}_1^*), \dots, \varphi^*(\mathbf{f}_{d'}^*)\}\}.$$

Ecrivons pour $i = 1, \dots, d'$

$$\varphi^*(\mathbf{f}_i^*) = \sum_{j=1}^d m_{ji}^* \mathbf{e}_j^*.$$

Par le Theoreme 6.6, on a

$$\varphi^*(\mathbf{f}_i^*) = \sum_{j=1}^d m_{ij} \mathbf{e}_j^*$$

avec m_{ij} defini par (cf. (6.3.1))

$$\varphi(\mathbf{e}_j) = \sum_{i=1}^{d'} m_{ij} \mathbf{f}_i.$$

Si $j > r$ alors $\mathbf{e}_j \in \ker(\varphi)$ et $\varphi(\mathbf{e}_j) = 0_W$: $\forall i \leq d'$, $m_{ij} = 0$ et donc

$$\forall i \leq d', \varphi^*(\mathbf{f}_i^*) = \sum_{j=1}^r m_{ij} \mathbf{e}_j^*.$$

Ainsi

$$\text{Vect}\{\varphi^*(\mathbf{f}_1^*), \dots, \varphi^*(\mathbf{f}_{d'}^*)\} \subset \text{Vect}\{\mathbf{e}_1^*, \dots, \mathbf{e}_r^*\}.$$

De plus, on a pour $i \leq r$

$$\varphi^*(\mathbf{f}_i^*) = \sum_{j=1}^r m_{ij} \mathbf{e}_j^*$$

avec

$$m_{ij} = \mathbf{f}_i^*(\varphi(\mathbf{e}_j)) = \mathbf{f}_i^*(\mathbf{f}_j) = \delta_{i=j}.$$

Ainsi si $i \leq r$, on a

$$\varphi^*(\mathbf{f}_i^*) = \mathbf{e}_i^*$$

et

$$\text{Im}(\varphi^*) = \text{Vect}\{\mathbf{e}_i^*, i \leq r\}.$$

Comme la famille $\{\mathbf{e}_i^*, i \leq r\}$ est libre l'espace engendre est de dimension r . □

CHAPITRE 7

Matrices

- *M: Do you know what I'm talking about ?*
- *N: The Matrix ?*
- *M: Do you want to know what IT is ?*
The Matrix is everywhere. It is all around us.
Even now, in this very room.

7.1. Matrices et applications lineaires

Soient V, W des EVs de dimensions finies munis de bases

$$\mathcal{B} = \{\mathbf{e}_j, j \leq d\}, \mathcal{B}' = \{\mathbf{f}_i, i \leq d'\}.$$

Alors on a des isomorphismes d'espaces vectoriels

$$CL_{\mathcal{B}} : K^d \simeq V, CL_{\mathcal{B}'} : K^{d'} \simeq W$$

qui permettent d'identifier V et W aux espaces produits K^d et $K^{d'}$ et d'identifier des vecteurs $v \in V$ et $w \in W$ avec des uplets

$$(x_j)_{j \leq d} = (x_1, \dots, x_d) \in K^d, (y_i)_{i \leq d'} = (y_1, \dots, y_{d'}) \in K^{d'}.$$

On dispose egalement d'une base

$$\mathcal{B}_{\mathcal{B}', \mathcal{B}} = \{\mathcal{E}_{ij} = \mathbf{e}_j^* \cdot \mathbf{f}_i, i \leq d', j \leq d\}$$

de $\text{Hom}_K(V, W)$ de sorte que l'application

$$(7.1.1) \quad CL_{\mathcal{B}', \mathcal{B}} : (m_{ij})_{i \leq d', j \leq d} \in (K^{d'})^d \mapsto \varphi = \sum_{i \leq d'} \sum_{j \leq d} m_{ij} \mathcal{E}_{ij} \in \text{Hom}_K(V, W)$$

est un isomorphisme d'espaces vectoriels entre $(K^{d'})^d$ et $\text{Hom}_K(V, W)$; cet isomorphisme permet d'identifier toute application lineaire φ avec un $d' \times d$ uplet $(m_{ij})_{i \leq d', j \leq d}$.

DÉFINITION 7.1. *L'espace vectoriel $(K^{d'})^d$ s'appelle l'espace des matrices de dimension $d' \times d$ a coefficients dans K et est note*

$$M_{d' \times d}(K) = \{(m_{ij})_{i \leq d', j \leq d}, m_{ij} \in K\}.$$

Un element de $M_{d' \times d}(K)$ est appelle matrice de dimensions $d' \times d$ ou juste une matrice $d' \times d$.

On a coutume de représenter une matrice $(m_{ij})_{i \leq d', j \leq d}$ comme un "tableau" de dimension 2 possédant d' lignes et d colonnes: ainsi m_{ij} est le coefficient de ce tableau qui se trouve a l'intersection de la i -ieme ligne et de la j -ieme colonne compte a partir du coin superieur gauche.

$$M = (m_{ij})_{i \leq d', j \leq d} = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1d} \\ m_{21} & m_{22} & \cdots & m_{2d} \\ \vdots & \vdots & \cdots & \vdots \\ m_{d'1} & m_{d'2} & \cdots & m_{d'd} \end{pmatrix}.$$

REMARQUE 7.1.1. Habituellement quand on repere un point dans le plan, la premiere coordonnee i donne la "position horizontale" et la seconde j la "position verticale". On prend ici la convention symetrique et il y a de bonnes raisons pour cela notamment lies au sens de l'ecriture gauche-droite.

DÉFINITION 7.2. Soient $\mathcal{B} \subset V$, $\mathcal{B}' \subset W$ des bases comme ci-dessous et $\mathcal{B}_{\mathcal{B}', \mathcal{B}} \subset \text{Hom}(V, W)$ la base de $\text{Hom}(V, W)$ associee. L'application reciproque $CL_{\mathcal{B}_{\mathcal{B}', \mathcal{B}}}^{-1}$ sera egalement notee

$$\text{mat}_{\mathcal{B}', \mathcal{B}} : \text{Hom}(V, W) \mapsto M_{d' \times d}(K).$$

Explicitement, si on la la decomposition $\varphi = \sum_{i \leq d'} \sum_{j \leq d} m_{ij}(\varphi) \mathcal{E}_{ij}$ alors on a

$$\text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi) = (m_{ij}(\varphi))_{i \leq d', j \leq d} = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1d} \\ m_{21} & m_{22} & \cdots & m_{2d} \\ \vdots & & \cdots & \vdots \\ m_{d'1} & m_{d'2} & \cdots & m_{d'd} \end{pmatrix}.$$

La matrice $\text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi)$ est appelee matrice associee a φ dans les bases $\mathcal{B}, \mathcal{B}'$. Rappelons que pour tout $1 \leq j \leq d$, $(m_{i,j}(\varphi))_{i \leq d'}$ est l'ensemble des coordonnees de l'image $\varphi(\mathbf{e}_j)$ de $\mathbf{e}_j \in \mathcal{B}$ dans la base $\mathcal{B}'$: ie.

$$\varphi(\mathbf{e}_j) = \sum_{1 \leq i \leq d'} m_{ij}(\varphi) \mathbf{f}_i.$$

EXEMPLE 7.1.1. Soit $V = \mathbb{R}^3$, $W = \mathbb{R}^2$ et prenons les bases canoniques

$$\mathcal{B} = \mathcal{B}_3^0 = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}, \quad \mathcal{B}' = \mathcal{B}_2^0 = \{(1, 0), (0, 1)\}.$$

On dispose de 6 applications lineaires elementaires

$$\mathcal{E}_{11}, \mathcal{E}_{12}, \mathcal{E}_{13}, \mathcal{E}_{21}, \mathcal{E}_{22}, \mathcal{E}_{23}$$

et par exemple

$$\mathcal{E}_{12}(x, y, z) = y(1, 0) = (y, 0), \quad \mathcal{E}_{23}(x, y, z) = z(0, 1) = (0, z).$$

Soit l'application lineaire de $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ donnee par

$$\varphi(x, y, z) = (2x + 4y, y + 3z)$$

alors

$$\varphi = 2\mathcal{E}_{11} + 0\mathcal{E}_{21} + 4\mathcal{E}_{12} + \mathcal{E}_{22} + 0\mathcal{E}_{13} + 3\mathcal{E}_{23}$$

et la matrice associee a φ vaut

$$\text{mat}_{\mathcal{B}_2^0, \mathcal{B}_3^0}(\varphi) = \begin{pmatrix} 2 & 4 & 0 \\ 0 & 1 & 3 \end{pmatrix}.$$

7.1.0.1. *Matrice nulle.* Si $\varphi = \underline{0}_W$ alors

$$\text{mat}_{\mathcal{B}', \mathcal{B}}(\underline{0}_W) = (0_K)_{i,j} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} = \underline{0}_{d' \times d}$$

est la matrice nulle.

7.1.0.2. *Matrices elementaires.* Une base de $M_{d' \times d}(K)$ est obtenue en transportant une base de $\text{Hom}_K(V, W)$ via cet isomorphisme, en particulier la base des applications elementaires

$$\mathcal{E}_{ij} = \mathbf{e}_j^* \cdot \mathbf{f}_i.$$

On note $E_{ij} = \text{mat}_{\mathcal{B}, \mathcal{B}'}(\mathcal{E}_{ij})$ la matrice correspondante qu'on appelle *matrice elementaire*. Ainsi, E_{ij} est la matrice dont le coefficient a l'intersection de la i -ieme ligne et de la j -ieme colonne vaut 1 et tous les autres coefficients sont nuls: pour $k \leq d', l \leq d$, on a

$$E_{ij,kl} = \delta_{k=i} \cdot \delta_{l=j}.$$

L'ensemble des matrices elementaires

$$\mathcal{B}_{d' \times d}^0 := \{E_{ij}, i \leq d', j \leq d\}$$

est forme une base de $M_{d' \times d}(K)$ qu'on appelle la *base canonique* de $M_{d' \times d}(K)$.

La base *duale* de la base canonique dans l'espace des formes lineaires

$$M_{d' \times d}(K)^* = \text{Hom}(M_{d' \times d}(K), K)$$

est notees

$$\mathcal{B}_{d' \times d}^{0,*} := \{E_{ij}^*, i \leq d', j \leq d\}.$$

Pour $i \leq d', j \leq d$ et $m \in M_{d' \times d}(K)$ une matrice,

$$E_{ij}^*(m) = m_{ij},$$

est le (i, j) -ieme coefficient de m .

7.1.0.3. *Matrices carrees.* Si $d' = d$ on dit que la matrice est carree et notera l'espaces des matrices carrees de taille d par

$$M_d(K) = M_{d \times d}(K).$$

Ces matrices codent les applications lineaires de $\text{Hom}(V, W)$ si $\dim V = \dim W$. En particulier si $V = W$ les elements de l'algebre des endomorphismes $\text{End}(V)$ sont codes par des matrices carrees.

7.1.0.4. *Matrice Identite.* Si $V = W$, $\mathcal{B} = \mathcal{B}'$ et $\varphi = \text{Id}_V$ est l'identite alors

$$(7.1.2) \quad \text{mat}_{\mathcal{B}, \mathcal{B}}(\text{Id}_V) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = (\delta_{i=j})_{i,j} =: \text{Id}_d \in M_{d \times d}(K).$$

est appelee matrice identite de rang d et est notee Id_d .

REMARQUE 7.1.2. En revanche si $\mathcal{B}' \neq \mathcal{B}$ la matrice $\text{mat}_{\mathcal{B}', \mathcal{B}}(\text{Id}_V)$ n'est pas egale a la matrice identite Id_d .

7.1.0.5. *Matrices scalaires.* Plus generalement notons pour $\lambda \in K$

$$[\times \lambda]: \begin{matrix} V & \mapsto & V \\ v & \mapsto & \lambda.v \end{matrix}$$

l'application lineaire de multiplication par le scalaire λ .

Sa matrice associee $\text{mat}_{\mathcal{B}, \mathcal{B}}([\times \lambda])$ vaut

$$\lambda \cdot \text{Id}_d = \lambda \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix}.$$

Elle est appelee matrice scalaire associee a λ . On note

$$K \cdot \text{Id}_d = \{\lambda \cdot \text{Id}, \lambda \in K\} \subset M_d(K)$$

l'ensemble des matrices scalaires. C'est un SEV de dimension 1 isomorphe a K et de base la matrice identite $\{\text{Id}_d\}$.

7.1.0.6. *Matrices colonnes.*

$$M_{d' \times 1}(K) =: \text{Col}_{d'}(K)$$

sont des matrices "colonnes" de hauteur d' . on posera

$$\text{Col}((x_i)_{i \leq d'}) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{d'} \end{pmatrix}.$$

7.1.0.7. *Matrices lignes.* Les element de

$$M_{1 \times d}(K) =: \text{Lig}_d(K)$$

sont des matrices "lignes" de longueur d : on posera

$$\text{Lig}((x_j)_{j \leq d}) = (x_1, \quad \dots, \quad x_d)$$

qui n'est autre que l'application identite de l'espace des matrices lignes.

DÉFINITION 7.3. Soient $\mathcal{B} \subset V$ une base. Soit

$$v = x_1 \cdot \mathbf{e}_1 + \dots + x_d \cdot \mathbf{e}_d \in V$$

un vecteur decompose dans la base $\mathcal{B}$. Alors la matrices

$$\text{Col}_{\mathcal{B}}(v) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix}, \quad \text{Lig}_{\mathcal{B}}(v) = (x_1 \quad \dots \quad x_d)$$

sont appelees respectivement

- la matrice colonne associee a v dans la base $\mathcal{B}$,
- La matrice ligne associee a v dans la base $\mathcal{B}$,

Ces applications sont des isomorphisme entre V et $\text{Col}_d(K)$ et $\text{Lig}_d(K)$.

7.1.0.8. *Colonnes et lignes extraites d'une matrice.*

DÉFINITION 7.4. Soit une matrice

$$M = \begin{pmatrix} m_{11} & m_{12} & \dots & m_{1d} \\ m_{21} & m_{22} & \dots & m_{2d} \\ \vdots & & \dots & \vdots \\ m_{d'1} & m_{d'2} & \dots & m_{d'd} \end{pmatrix} \in M_{d' \times d}(K).$$

Pour $j \leq d$ (resp. $i \leq d'$), la j -ieme colonne de M (resp. la i -ieme ligne de M) est la matrice colonne (resp. ligne)

$$\text{Col}_j(M) = \begin{pmatrix} m_{1j} \\ m_{2j} \\ \vdots \\ m_{d'j} \end{pmatrix} \in \text{Col}_{d'}(K), \quad \text{resp.} \quad \text{Lig}_i(M) = (m_{i1} \ m_{i2} \ \dots \ m_{id}) \in \text{Lig}_d(K)$$

EXEMPLE 7.1.2. Si

$$M = (m_{ij})_{i \leq d', j \leq d} = \text{mat}_{\mathcal{B}'\mathcal{B}}(\varphi)$$

alors on a vu que pour $j \leq d$ les coordonnees de $\varphi(\mathbf{e}_j)$ dans la base $\mathcal{B}'$ sont donnees par le vecteur ligne $(m_{ij})_{i \leq d'}$ dont le vecteur colonne associe est la j -ieme colonne de la matrice M :

$$\text{Col}_j(M) = \begin{pmatrix} m_{1j} \\ m_{2j} \\ \vdots \\ m_{d'j} \end{pmatrix}.$$

7.2. Structure des espaces de matrices

7.2.1. Addition et multiplication par les scalaires. Les espaces de matrices $M_{d',d}(K)$ sont naturellement des K -ev pour les lois d'addition et de multiplication par les scalaires evidentes: si

$$M = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1d} \\ m_{21} & m_{22} & \cdots & m_{2d} \\ \vdots & \vdots & \cdots & \vdots \\ m_{d'1} & m_{d'2} & \cdots & m_{d'd} \end{pmatrix}, \quad M' = \begin{pmatrix} m'_{11} & m'_{12} & \cdots & m'_{1d} \\ m'_{21} & m'_{22} & \cdots & m'_{2d} \\ \vdots & \vdots & \cdots & \vdots \\ m'_{d'1} & m'_{d'2} & \cdots & m'_{d'd} \end{pmatrix} \in M_{d' \times d}(K)$$

$$\lambda.M + M' = (\lambda.m_{ij} + m'_{ij})_{ij} = \begin{pmatrix} \lambda.m_{11} + m'_{11} & \lambda.m_{12} + m'_{12} & \cdots & \lambda.m_{1d} + m'_{1d} \\ \lambda.m_{21} + m'_{21} & \lambda.m_{22} + m'_{22} & \cdots & \lambda.m_{2d} + m'_{2d} \\ \vdots & \vdots & \cdots & \vdots \\ \lambda.m_{d'1} + m'_{d'1} & \lambda.m_{d'2} + m'_{d'2} & \cdots & \lambda.m_{d'd} + m'_{d'd} \end{pmatrix}$$

de sorte que l'application

$$\text{mat}_{\mathcal{B}', \mathcal{B}} : \varphi \in \text{Hom}(V, W) \mapsto \text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi) \in M_{d' \times d}(K)$$

est un isomorphisme de K -ev.

Il est facile de verifier que les applications lignes et colonnes

$$\text{Col}_i : M_{d' \times d}(K) \mapsto \text{Col}_{d'}(K), \quad \text{Lig}_j : M_{d' \times d}(K) \mapsto \text{Lig}_d(K)$$

sont lineaires.

7.2.2. Multiplication de matrices. Soient U, V, W trois espaces vectoriels munis de bases

$$\mathcal{B} = \{\mathbf{e}_k, k \leq d\}, \mathcal{B}' = \{\mathbf{f}_j, j \leq d'\}, \mathcal{B}'' = \{\mathbf{g}_i, i \leq d''\}.$$

On dispose alors de bases

$$\mathcal{B}_{\mathcal{B}', \mathcal{B}} = \{\mathbf{e}_k^* \cdot \mathbf{f}_j\}, \quad \mathcal{B}_{\mathcal{B}'', \mathcal{B}'} = \{\mathbf{f}_j^* \cdot \mathbf{g}_i\}, \quad \mathcal{B}_{\mathcal{B}'', \mathcal{B}} = \{\mathbf{e}_k^* \cdot \mathbf{g}_i\}$$

pour

$$\text{Hom}_{K\text{-ev}}(U, V), \quad \text{Hom}_{K\text{-ev}}(V, W), \quad \text{Hom}_{K\text{-ev}}(U, W).$$

Soient

$$\varphi : U \mapsto V, \quad \psi : V \mapsto W$$

deux applications lineaires et

$$\psi \circ \varphi : U \mapsto W$$

leur composee.

Soient alors

$$N := \text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi) = (n_{jk})_{j \leq d', k \leq d} \in M_{d' \times d}(K)$$

et

$$M := \text{mat}_{\mathcal{B}'', \mathcal{B}'}(\psi) = (m_{ij})_{i \leq d'', j \leq d'} \in M_{d'' \times d'}(K)$$

et

$$L := \text{mat}_{\mathcal{B}'', \mathcal{B}}(\psi \circ \varphi) = (l_{ik})_{i \leq d'', k \leq d} \in M_{d'' \times d}(K)$$

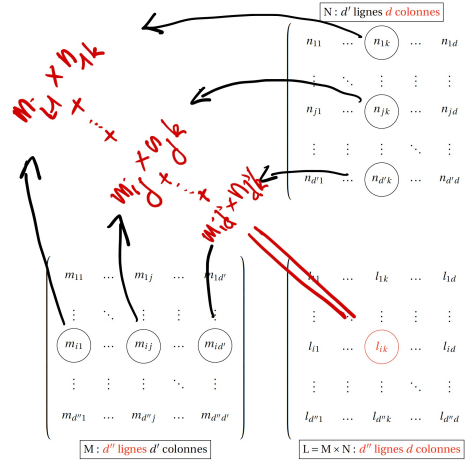


FIGURE 1. Calcul des coordonnees du produit de deux matrices

On a vu (Thm 6.5) que les $(l_{ik})_{i \leq d'', k \leq d}$ pouvaient s'exprimer en fonction des $(m_{ij})_{i \leq d'', j \leq d'}$ et des $(n_{jk})_{j \leq d', k \leq d}$. Plus precisement, on a

$$l_{ik} = \sum_{j=1}^{d'} m_{ij} \cdot n_{jk}.$$

On definit ainsi une loi de multiplication (externe) sur les espaces de matrices en posant:

DÉFINITION 7.5. Soient $d, d', d'' \geq 1$ et $M \in M_{d'' \times d'}(K)$, $N \in M_{d' \times d}(K)$, on defini le produit des matrices M et N comme etant la matrice

$$L := M.N \in M_{d'' \times d}(K)$$

avec

$$L = (l_{ik})_{i \leq d'', k \leq d} \in M_{d'' \times d}(K) \text{ et } l_{ik} := \sum_{j=1}^{d'} m_{ij} \cdot n_{jk}.$$

Soient $d, d', d'' \geq 1$, on a donc defini une application "produit de matrices"

$$(7.2.1) \quad \bullet \bullet : \begin{array}{ccc} M_{d'' \times d'}(K) \times M_{d' \times d}(K) & \mapsto & M_{d'' \times d}(K) \\ (M, N) & \mapsto & L = M.N \end{array}$$

REMARQUE 7.2.1. Notons que ce produit est entre deux espaces de matrices de tailles qui peuvent etre differentes $d'' \times d'$ et $d' \times d$ et a valeurs dans un troisieme espace de matrices dont les tailles peuvent encore etre differente (ie $d'' \times d$). La contrainte la plus importantw est que la deuxieme dimension (d') du premier espace de matrices soit egale a la premiere dimension du premier espace de matrices. Le resultat est a valeurs dans l'espace des matrices de tailles les deux dimensions "extremes" (ie $d'' \times d$).

EXEMPLE 7.2.1. Quelques cas particuliers importants:

- Si $d = 1$: on dispose d'une multiplication "externe" (a gauche) a valeurs dans les matrices colonnes: on a $M_{d' \times 1}(K) = \text{Col}_{d'}(K)$ et donc

$$\bullet \bullet : M_{d'' \times d'}(K) \times \text{Col}_{d'}(K) \mapsto \text{Col}_{d''}(K).$$

- Si $d'' = d' = d$: les matrices sont toutes carrees et on dispose d'une multiplication "interne" sur l'espace des matrices carrees de taille d :

$$\bullet \times \bullet : M_d(K) \times M_d(K) \mapsto M_d(K).$$

THÉORÈME 7.1 (Propriétés fonctionnelles du produit de matrices). Soient $d, d', d'' \geq 1$ et $M_{d'' \times d'}(K)$, $M_{d' \times d}(K)$, $M_{d'' \times d}(K)$ les espaces de matrices correspondants.

L'application "produit de matrices"

$$\begin{aligned} M_{d'' \times d'}(K) \times M_{d' \times d}(K) &\mapsto M_{d'' \times d}(K) \\ (M, N) &\mapsto M.N \end{aligned}$$

a les propriétés suivantes

(1) Distributive à gauche: pour $\lambda \in K$, $M, M' \in M_{d'' \times d'}(K)$, $N \in M_{d' \times d}(K)$,

$$(\lambda.M + M').N = \lambda.M.N + M'.N.$$

(2) Distributive à droite: pour $\lambda \in K$, $M \in M_{d'' \times d'}(K)$, $N, N' \in M_{d' \times d}(K)$,

$$M.(\lambda.N + N') = \lambda.M.N + M.N'.$$

(3) Neutralité de l'identité: pour $M \in M_{d'' \times d'}(K)$,

$$\text{Id}_{d''}.M = M, M.\text{Id}_{d'} = M$$

(4) La matrice nulle est absorbante: pour $M \in M_{d'' \times d'}(K)$,

$$0_{d'' \times d''}.M = 0_{d'' \times d'}, M.0_{d' \times d} = 0_{d'' \times d}.$$

(5) Associativité: Soit $d''' \geq 1$ et $L \in M_{d''' \times d''}(K)$, $M \in M_{d'' \times d'}(K)$, $N \in M_{d' \times d}(K)$ alors

$$(L.M).N = L.(M.N) \in M_{d''' \times d}(K)$$

Preuve: On démontre ces énoncés soit par un calcul direct, soit sans faire de calcul mais en interprétant le produit de matrices en terme de composition d'applications linéaires. On utilise le Théorème 7.2 ci-dessous et les propriétés d'associativité et de distributivité des applications linéaires par rapport à la composition et l'addition (qu'on a plus ou moins vu précédemment) et qu'on liste dans le Théorème 7.2.2. $\square$

Le Théorème ci-dessous est une tautologie puisqu'on a défini le produit des deux matrices précisément pour être compatible avec la composition d'applications linéaires.

THÉORÈME 7.2. Soit U, V, W des espaces vectoriels de dimensions d, d', d'' et $\mathcal{B}, \mathcal{B}', \mathcal{B}''$ des bases. Soient des applications linéaires

$$\varphi : U \mapsto V, \psi : V \mapsto W.$$

On note les coefficients des matrices de φ, ψ et $\psi \circ \varphi$ dans les bases adéquates par

$$\begin{aligned} \text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi) &= (n_{jk})_{jk}, \quad \text{mat}_{\mathcal{B}'', \mathcal{B}'}(\psi) = (m_{ij})_{ij} \\ \text{mat}_{\mathcal{B}'', \mathcal{B}}(\psi \circ \varphi) &= (l_{ik})_{ik} \end{aligned}$$

alors on a

$$(7.2.2) \quad \text{mat}_{\mathcal{B}'', \mathcal{B}}(\psi \circ \varphi) = \text{mat}_{\mathcal{B}'', \mathcal{B}'}(\psi) \cdot \text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi)$$

Autrement dit on a

$$\begin{pmatrix} l_{11} & \cdots & l_{1d} \\ l_{21} & \cdots & l_{2d} \\ \vdots & \cdots & \vdots \\ \vdots & \cdots & \vdots \\ l_{d''1} & \cdots & l_{d''d} \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1d'} \\ m_{21} & m_{22} & \cdots & m_{2d'} \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ m_{d''1} & m_{d''2} & \cdots & m_{d''d'} \end{pmatrix} \cdot \begin{pmatrix} n_{11} & \cdots & n_{1d} \\ n_{21} & \cdots & n_{2d} \\ \vdots & \cdots & \vdots \\ n_{d'1} & \cdots & n_{d'd} \end{pmatrix}$$

Le résultat suivant est obtenu en démontrant l'égalité de diverses applications linéaires en vérifiant que deux applications prennent la même valeur pour tout vecteur de l'espace de départ.

THÉORÈME (Propriétés fonctionnelles de la composition des applications linéaires). Soient U, V, W, Z des espaces vectoriels de dimensions finies.

L'application "composition"

$$\begin{aligned} \bullet \circ \bullet : \text{Hom}_K(V, W) \times \text{Hom}_K(U, V) &\mapsto \text{Hom}_K(U, W) \\ (\psi, \varphi) &\mapsto \psi \circ \varphi \end{aligned}$$

a les propriétés suivantes

(1) Distributive à gauche: pour $\lambda \in K$, $\psi, \psi' \in \text{Hom}_K(V, W)$, $\varphi \in \text{Hom}_K(U, V)$,

$$(\lambda \cdot \psi + \psi') \circ \varphi = \lambda \cdot \psi \circ \varphi + \psi' \circ \varphi.$$

(2) Distributive à droite: pour $\lambda \in K$, $\psi \in \text{Hom}_K(V, W)$, $\varphi, \varphi' \in \text{Hom}_K(U, V)$,

$$\psi \circ (\lambda \cdot \varphi + \varphi') = \lambda \cdot \psi \circ \varphi + \psi \circ \varphi'.$$

(3) Neutralité de l'identité: pour $\psi \in \text{Hom}_K(V, W)$,

$$\text{Id}_W \circ \psi = \psi, \quad \psi \circ \text{Id}_V = \psi.$$

(4) L'application linéaire nulle est absorbante: soit Z un K -ev et

$$\underline{0}_Z : W \mapsto Z, \quad \underline{0}'_Z : V \mapsto Z, \quad \underline{0}_W : V \mapsto W, \quad \underline{0}'_W : U \mapsto W, \quad \underline{0}_V : U \mapsto V$$

les applications constantes nulles; on a pour $\psi \in \text{Hom}_K(V, W)$,

$$\underline{0}_Z \circ \psi = \underline{0}'_Z, \quad \psi \circ \underline{0}_V = \underline{0}_W.$$

(5) Associativité: Soit $\theta \in \text{Hom}_K(W, Z)$, $\psi \in \text{Hom}_K(V, W)$, $\varphi \in \text{Hom}_K(U, V)$ alors

$$(\theta \circ \psi) \circ \varphi = \theta \circ (\psi \circ \varphi) \in \text{Hom}_K(U, Z)$$

7.2.2.1. Image de vecteurs. La multiplication matricielle permet également de calculer l'image d'un vecteur par une application linéaire:

PROPOSITION 7.1. Soit $\mathcal{B} \subset V$, $\mathcal{B}' \subset W$ des bases, $v \in V$ un vecteur de coordonnées $(x_j)_{j \leq d}$ dans la base $\mathcal{B}$ (ie. $v = x_1 \cdot \mathbf{e}_1 + \dots + x_d \cdot \mathbf{e}_d$) et $(y_i)_{i \leq d'}$ les coordonnées de $\varphi(v)$ dans la base $\mathcal{B}'$ (ie. $\varphi(v) = y_1 \cdot \mathbf{f}_1 + \dots + y_{d'} \cdot \mathbf{f}_{d'}$). On associe à v et $\varphi(v)$ leurs matrices colonnes (de hauteurs d et $d' =$

$$\text{Col}_{\mathcal{B}}(v) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix}, \quad \text{Col}_{\mathcal{B}'}(\varphi(v)) = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{d'} \end{pmatrix}$$

alors on a la relation

$$\text{Col}_{\mathcal{B}'}(\varphi(v)) = \text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi) \cdot \text{Col}_{\mathcal{B}}(v).$$

Autrement dit si $\text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi) = (m_{ij})_{i \leq d', j \leq d}$, on a

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{d'} \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & \dots & m_{1d} \\ m_{21} & m_{22} & \dots & m_{2d} \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ m_{d'1} & m_{d'2} & \dots & m_{d'd} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix}$$

7.2.2.2. *Produit de matrices elementaires.*

PROPOSITION 7.2. Soit $E_{ij} \in M_{d'' \times d'}$ et $E_{j'k} \in M_{d' \times d}$ alors

$$E_{ij} \cdot E_{j'k} = \delta_{j=j'} E_{ik}.$$

Preuve: On raisonne en terme d'applications lineaires elementaires $\mathcal{E}_{ij}$, $\mathcal{E}_{j'k}$: on a

$$\mathcal{E}_{ij} \circ \mathcal{E}_{j'k}(\mathbf{e}_{k'}) = \mathcal{E}_{ij}(\delta_{k'=k} \mathbf{f}_{j'}) = \delta_{k'=k} \delta_{j=j'} \mathbf{g}_i = \delta_{j=j'} \mathcal{E}_{ik}(\mathbf{e}_{k'}).$$

□

7.2.2.3. *Le cas des isomorphis mes.* On considere le cas ou $\varphi : U \mapsto V$ est un isomorphisme et $\psi = \varphi^{-1} : V \mapsto W = U$ est l'application reciproque. En particulier U et V sont de meme dimension: $d = d' = d''$.

PROPOSITION 7.3. soit $\varphi : V \simeq W$ un isomorphisme lineaire et $\varphi^{-1} : W \mapsto V$ la reciproque. On a les relations

$$\text{mat}_{\mathcal{B}, \mathcal{B}'}(\varphi^{-1}) \cdot \text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi) = \text{Id}_d,$$

$$\text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi) \cdot \text{mat}_{\mathcal{B}, \mathcal{B}'}(\varphi^{-1}) = \text{Id}_d.$$

En particulier si $V = W$ et $\varphi = \text{Id}_V$ est l'identite on a

$$(7.2.3) \quad \text{mat}_{\mathcal{B}', \mathcal{B}}(\text{Id}_V) \cdot \text{mat}_{\mathcal{B}, \mathcal{B}'}(\text{Id}_V) = \text{Id}_d.$$

Preuve: On applique la relation (7.2.2) a la suite de K -EVs $V, W, V, \mathcal{B}, \mathcal{B}', \mathcal{B}'' = \mathcal{B}$ et $\psi = \varphi^{-1}$. On a donc

$$\psi \circ \varphi = \text{Id}_V, \quad \varphi \circ \psi = \text{Id}_W.$$

On a donc par (7.2.2)

$$\text{mat}_{\mathcal{B}, \mathcal{B}}(\text{Id}_V) = \text{mat}_{\mathcal{B}, \mathcal{B}'}(\varphi^{-1}) \cdot \text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi)$$

Comme

$$\text{mat}_{\mathcal{B}, \mathcal{B}}(\text{Id}_V) = \text{Id}_d$$

on obtient

$$\text{mat}_{\mathcal{B}, \mathcal{B}'}(\varphi^{-1}) \cdot \text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi) = \text{Id}_d.$$

L'autre relation se demontre de la meme maniere. □

7.2.3. Rang d'une matrice. On a defini le rang d'une application lineaire $\varphi : V \mapsto W$ comme etant la dimension de l'image

$$\text{rg}(\varphi) = \dim \varphi(V).$$

Soit $M = \text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi)$ la matrice associee. Comme l'image $\varphi(V)$ est le SEV engendre par

$$\{\varphi(\mathbf{e}_j), j \leq d\} \subset W,$$

l'image $\varphi(V)$ s'identifie avec le SEV de l'espace vectoriel des matrices colonnes $\text{Col}_{d'}(K)$ engendre par les j -colonnes de M ,

$$\{\text{Col}_j(M) = \text{Col}_{\mathcal{B}'}(\varphi(\mathbf{e}_j)), j \leq d\}.$$

La dimension de l'espace engendre par ces matrices colonnes est donc de dimension $r = \text{rg}(\varphi)$:

DÉFINITION 7.6. Soit $M \in M_{d' \times d}(K)$, le rang d'une matrice M est la dimension de l'espace engendre par les d colonnes de M dans $\text{Col}_{d'}(K)$:

$$\text{rg}(M) = \dim \text{Vect}(\{\text{Col}_j(M), j \leq d\}).$$

Autrement dit $\text{rg}(M)$ est la taille maximale d'une sous-famille libre de la famille $\{\text{Col}_j(M), j \leq d\}$ des colonnes de M .

Compte-tenu de la discussion precedente on a

$$(7.2.4) \quad \text{rg}(\text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi)) = \text{rg}(\varphi).$$

REMARQUE 7.2.2. On a $\text{rg}(M) \leq d$ (puisque d vecteurs engendrent un espace de dimension au plus d) et

$$\text{rg}(M) \leq d' = \dim \text{Col}_{d'}(K).$$

Ainsi

$$\text{rg}(M) \leq \min(d, d').$$

7.2.3.1. *Exemple d'une matrice de rang donne.* Soit $\varphi : V \mapsto W$ telle que $\text{rg}(\varphi) = r$. Soit

$$\mathcal{I} := \{\mathbf{f}_i = \varphi(\mathbf{e}_i), i = 1, \dots, r\}$$

une base de $\text{Im}(\varphi)$; complétons $\mathcal{I}$ en une base de W

$$\mathcal{B}' = \mathcal{I} \sqcup \{\mathbf{f}_{r+1}, \dots, \mathbf{f}_{d'}\} = \{\mathbf{f}_1, \dots, \mathbf{f}_{d'}\}$$

et soit

$$\mathcal{K} = \{\mathbf{e}_{r+1}, \dots, \mathbf{e}_d\} \subset \ker(\varphi)$$

une base de $\ker(\varphi)$, on a vu que

$$\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_r\} \sqcup \mathcal{K} \subset V$$

est une base de V . On a alors

$$\varphi(\mathbf{e}_i) = \begin{cases} \mathbf{f}_i & i = 1, \dots, r \\ 0_W & i \geq r+1 \end{cases}$$

et donc

$$(7.2.5) \quad \text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & \dots & \dots & 1 & \vdots \\ \vdots & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix} = \begin{pmatrix} & & & & 0 \\ & & & & \vdots \\ & & & & \vdots \\ & & \text{Id}_r & & \vdots \\ & & & & \vdots \\ 0 & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix} =: I_{d' \times d}(r)$$

Il est clair que les r premières colonnes de la matrice $I_{d' \times d}(r)$ forment une famille libre et la matrice est bien de rang r .

EXERCICE 7.1. Déterminer le rang de la matrice

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

en fonction de la caractéristique du corps K .

7.2.4. Transposition. La transposition est l'application qui transforme une matrice par symétrie par rapport à la première diagonale $i = j$:

DÉFINITION 7.7. La transposition est l'application des matrices $d' \times d$ vers les matrices $d \times d'$ définie par

$${}^t \bullet : \begin{matrix} M_{d' \times d}(K) & \mapsto & M_{d \times d'}(K) \\ M = (m_{ij})_{i \leq d', j \leq d} & \mapsto & {}^t M = (m_{ji}^*)_{j \leq d, i \leq d'} \end{matrix}$$

avec

$$m_{ji}^* = m_{ij}, \quad j \leq d, i \leq d'.$$

Autrement dit si

$$M = (m_{ij})_{i \leq d', j \leq d}, \quad {}^t M = (m_{ji}^*)_{j \leq d, i \leq d'} = (m_{ij})_{j \leq d, i \leq d'}$$

$$M = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1d} \\ m_{21} & m_{22} & \cdots & m_{2d} \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ m_{d'1} & m_{d'2} & \cdots & m_{d'd} \end{pmatrix}, \quad {}^tM = \begin{pmatrix} m_{11} & m_{21} & \cdots & \cdots & m_{d'1} \\ m_{12} & m_{22} & \cdots & \cdots & m_{d'2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ m_{1d} & m_{2d} & \cdots & \cdots & m_{d'd} \end{pmatrix}$$

La transposition est l'opération matricielle qui correspond à prendre la duale d'une application lineaire.

Rappelons que si V et W sont des K -EV de dimensions finies, à toute application lineaire $\varphi \in \text{Hom}(V, W)$ on associe une application lineaire duale $\varphi^* \in \text{Hom}(W^*, V^*)$ donnees par

$$\ell' \in W^* \mapsto \varphi^*(\ell') = \ell' \circ \varphi : v \mapsto \ell'(\varphi(v)).$$

Munissons V et W de bases $\mathcal{B} = \{\mathbf{e}_j, j \leq d\}$ et $\mathcal{B}' = \{\mathbf{f}_i, i \leq d'\}$; les espaces duaux V^* et W^* sont munis des bases duales $\mathcal{B}^* = \{\mathbf{e}_j^*, j \leq d\}$ et $\mathcal{B}'^* = \{\mathbf{f}_i^*, i \leq d'\}$. On rappelle qu'on a demontre le

THÉORÈME (Matrice de l'application duale). *Soit $\varphi : V \mapsto W$ une application lineaire et $\varphi^* : W^* \rightarrow V^*$ sa duale; $\mathcal{B}$ et $\mathcal{B}'$ des bases de V et V' et*

$$\text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi) = (m_{ij})_{i \leq d', j \leq d}$$

la matrice de φ dans les bases $\mathcal{B}$ et $\mathcal{B}'$ et soit

$$\text{mat}_{\mathcal{B}^*, \mathcal{B}'^*}(\varphi^*) = (m_{ji}^*)_{j \leq d, i \leq d'}$$

la matrice de φ^* dans les bases duales $\mathcal{B}'^* \subset W^*$ et $\mathcal{B}^* \subset V^*$ alors on a

$$m_{ji}^* = m_{ij}, \quad i \leq d', \quad j \leq d$$

En d'autres termes

$$\text{mat}_{\mathcal{B}^*, \mathcal{B}'^*}(\varphi^*) = {}^t \text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi).$$

THÉORÈME 7.3. (Proprietes fonctionelles de la transposition) *La transposition a les proprietes suivantes:*

- (1) *Linearite:* ${}^t(\lambda.M + M') = \lambda {}^tM + {}^tM'$.
- (2) *Involutive:* ${}^t({}^tM) = M$.
- (3) *Anti-multiplicativite:* pour $M \in M_{d'', d'}(K)$, $N \in M_{d', d}(K)$, $M.N \in M_{d'', d}(K)$ et

$${}^t(M.N) = {}^tN.{}^tM.$$

Preuve: Seul le dernier point est un peu plus difficile: on peut le verifier par un calcul explicite sur les produits de matrices ou l'obtenir de maniere abstraite. Pour cela on note que si on a

$$\varphi : U \mapsto V, \quad \psi : V \mapsto W, \quad \psi \circ \varphi : U \mapsto W$$

alors on a les applications duales

$$\varphi^* : V^* \mapsto U^*, \quad \psi^* : W^* \mapsto V^*, \quad (\psi \circ \varphi)^* : W^* \mapsto U^*$$

On a d'autre part la composee

$$\varphi^* \circ \psi^* : W^* \mapsto U^*$$

et il suffira de montrer que

$$(\psi \circ \varphi)^* = \varphi^* \circ \psi^*$$

(et de passer aux matrices). On a par definition, pour $\ell'' \in W^*$ et par associativite

$$(\psi \circ \varphi)^*(\ell'') = \ell'' \circ (\psi \circ \varphi) = (\ell'' \circ \psi) \circ \varphi = \varphi^*(\ell'' \circ \psi) = \varphi^*(\psi^*(\ell'')) = \varphi^* \circ \psi^*(\ell'')$$

□

Compte tenu de l'interpretation du rang d'une matrice comme rang d'une application lineaire (cf. (7.2.4)), on deduit du Theoreme 6.7 qui dit que

$$\text{rg}(\varphi) = \text{rg}(\varphi^*),$$

le

THÉORÈME 7.4 (Invariance du rang par transposition). *Soit $M \in M_{d' \times d}(K)$ on a*

$$\text{rg}(M) = \text{rg}({}^t M).$$

Comme la transposee d'une matrice transforme les colonnes en lignes on obtient:

COROLLAIRE 7.1. *La rang d'une matrice est egal a la dimension de l'espace K^d engendre par les vecteurs lignes de M*

$$\text{rg}(M) = \dim_K \text{Vect}(\text{Lig}_j(M), j = 1, \dots, d').$$

7.3. L'algebre des matrices carrees

Si $d' = d$, on obtient l'espace vectoriel des matrices carres

$$M_{d \times d}(K) = M_d(K)$$

qui est de dimension $\dim M_d(K) = d^2$.

7.3.1. Structure d'anneau. Comme on l'a vu, la multiplication des matrices

$$(M, M') \in M_d(K) \times M_d(K) \mapsto M.M' \in M_d(K)$$

est alors une loi de composition interne et par le Theoreme 7.1, on a

THÉORÈME 7.5. *L'espace $M_d(K)$ muni de l'addition des matrices et de la multiplication est un anneau (non-commutatif en general) dont l'element neutre est la matrice carree nulle $\underline{0}_d = \underline{0}_{d \times d}$ et dont l'unité est la matrice identite Id_d . De plus la structure de K -EV de $M_d(K)$ fait de l'anneau $(M_d(K), +, \cdot)$ une K -algebre (de dimension d^2).*

On l'appelle l'algebre des matrices carres de dimension d (ou de rang d) sur le corps K (ou a coefficient dans K).

REMARQUE 7.3.1. Ici "dimension d " designe a la taille des matrice, pas a la dimension de l'espace des matrices $M_d(K)$ (qui est d^2).

7.3.2. Lien avec l'algebre des endomorphismes. Soit V de dimension d . On rappelle que l'ensemble des endomorphismes de V , $\text{End}(V) = \text{Hom}(V, V)$ est non seulement un K -espace vectoriel (pour l'addition des applications lineaires) mais egalement possede une structure d'anneau (et donc de K -algebre) ou la "multiplication" est donnee par la composition des endomorphismes: pour $\varphi, \psi \in \text{End}(V)$

$$\varphi \circ \psi : V \xrightarrow{\psi} V \xrightarrow{\varphi} V.$$

L'element neutre est l'endomorphisme nul $\underline{0}_V$ et l'element unite est l'application identite Id_V .

Soit $\mathcal{B}$ une base de V , on dispose alors d'un isomorphisme d'espaces vectoriels

$$\text{mat}_{\mathcal{B}, \mathcal{B}} : \varphi \in \text{End}(V) \mapsto \text{mat}_{\mathcal{B}, \mathcal{B}}(\varphi) \in M_d(K).$$

Pour simplifier les notations on ecrira cet isomorphisme $\text{mat}_{\mathcal{B}}$ (ou juste mat si la base $\mathcal{B}$ est implicite) et la matrice associee a un endomorphisme φ sera notee

$$\text{mat}_{\mathcal{B}}(\varphi) := \text{mat}_{\mathcal{B}, \mathcal{B}}(\varphi).$$

THÉORÈME 7.6. *Soit V de dimension finie d et $\mathcal{B}$ une base de V , l'application*

$$\text{mat}_{\mathcal{B}} : \text{End}(V) \mapsto M_d(K)$$

est un isomorphisme d'anneaux (et donc de K -algebres) pour les lois d'addition et de multiplication decrites precedemment.

Preuve: On sait déjà que $\text{mat}_{\mathcal{B}}$ est un isomorphisme d'espace vectoriel (et est donc bijectif). Pour montrer qu'on a un isomorphisme d'anneaux, il suffit de vérifier que c'est morphisme d'anneaux non-nul: on doit vérifier que

$$\text{mat}_{\mathcal{B}}(\text{Id}_V) = \text{Id}_d$$

ce qu'on a déjà vu et que pour $\varphi, \psi \in \text{End}(V)$

$$\text{mat}_{\mathcal{B}}(\varphi \circ \psi) = \text{mat}_{\mathcal{B}}(\varphi) \cdot \text{mat}_{\mathcal{B}}(\psi).$$

Mais c'est –aux notations pres– un cas particulier pour $U = V = W$ du Theorem 7.2: si $\text{mat}_{\mathcal{B}}(\varphi) = M = (m_{ij})_{i,j \leq d}$ et $\text{mat}_{\mathcal{B}}(\psi) = N = (n_{ij})_{i,j \leq d}$ alors

$$M \cdot N = L = (l_{ik})_{i,k \leq d}$$

avec

$$l_{ik} = \sum_{j=1 \dots d} m_{ij} \cdot n_{jk}$$

et

$$L = (l_{ik})_{i,k \leq d} = \text{mat}_{\mathcal{B}}(\varphi \circ \psi)$$

par le Thm 6.5. □

REMARQUE 7.3.2. Comme on a vu, étant donné un endomorphisme $\varphi : V \mapsto V$, on aurait pu prendre deux bases $\mathcal{B}, \mathcal{B}' \subset V$ et associer la matrice $\text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi)$ à φ . Un des avantages de choisir $\mathcal{B}' = \mathcal{B}$ est que l'identité Id_V est alors représentée par la matrice identité Id_d , mais l'avantage principal de choisir $\mathcal{B}' = \mathcal{B}$ est le Theorem 7.6.

7.3.2.1. La transposition est un antimorphisme. Si une matrice M est carrée $d \times d$ sa transposée ${}^t M$ est encore carrée $d \times d$. Compte tenu des propriétés générales de la transposition (cf. Prop 7.3), on a

PROPOSITION 7.4. La transposition

$${}^t \bullet : M_d(K) \mapsto M_d(K)$$

est un endomorphisme de $M_d(K)$ qui est

(1) *Involutif:*

$${}^t({}^t M) = M.$$

(2) *En particulier ${}^t \bullet$ est inversible et son inverse est lui-même:*

$${}^t({}^t \bullet) = \text{Id}_{M_d(K)}, \quad ({}^t \bullet)^{-1} = {}^t \bullet.$$

(3) *Anti-multiplicatif: ${}^t(M \cdot N) = {}^t N \cdot {}^t M$.*

REMARQUE 7.3.3. On dit que la transposition est un anti-automorphisme d'algèbres.

7.3.3. Le groupe linéaire.

DÉFINITION 7.8. Soit V un K -EV de dimension finie. Le groupe linéaire de V est le groupe (pour la composition dans $\text{End}(V)$) des éléments inversibles de l'algèbre $\text{End}_K(V)$; son élément neutre est l'identité Id_V et on note ce groupe

$$\text{GL}(V) = \text{End}_K(V)^\times = \{\varphi : V \mapsto V, \varphi \text{ est bijectif}\}.$$

Soit $d \geq 1$. Le groupe linéaire de rang d sur K est le groupe des matrices carrées inversibles dans l'algèbre $M_d(K)$ pour la multiplication des matrices; son élément neutre est la matrice identité Id_d et on note ce groupe

$$\text{GL}_d(K) = M_d(K)^\times = \{M \in M_d(K), \exists M' \in M_d(K), M \cdot M' = M' \cdot M = \text{Id}_d\}.$$

On a alors

PROPOSITION 7.5. *L'application $\text{mat}_{\mathcal{B}} : \text{End}(V) \mapsto M_d(K)$ induit un isomorphisme de groupes*

$$\text{mat}_{\mathcal{B}} : \text{GL}(V) \mapsto \text{GL}_d(K)$$

et en particulier

$$\text{mat}_{\mathcal{B}}(\varphi^{-1}) = \text{mat}_{\mathcal{B}}(\varphi)^{-1}.$$

7.3.3.1. *Critere d'inversibilite.* Dans $\text{End}_K(V)$, on a le critere d'inversibilite suivant

THÉORÈME 7.7 (Critere d'inversibilite des endomorphismes). *Soit $\varphi : V \mapsto V$ alors les conditions suivantes son equivalentes:*

- (1) φ est inversible (ie. bijective),
- (2) φ est injective,
- (3) φ est surjective,
- (4) $\text{rg}(\varphi) = d$,
- (5) φ transforme une base de V en une famille libre,
- (6) φ transforme une base de V en une famille generatrice

On en deduit de ce critere et de l'isomorphisme $\text{mat}_{\mathcal{B}} : \text{End}(V) \simeq M_d(K)$ le critere d'inversibilite suivant

THÉORÈME 7.8 (Critere d'inversibilite pour les matrices (via les colonnes)). *Soit une matrice carree $M = (m_{ij})_{i,j \leq d} \in M_d(K)$, les conditions suivantes sont equivalentes*

- (1) M est inversible, ie. $M \in \text{GL}_d(K)$,
- (2) $\text{rg}(M) = d$,
- (3) $\{\text{Col}_i(M), i = 1, \dots, d\}$ forme une famille libre de $\text{Col}_d(K)$,
- (4) $\{\text{Col}_i(M), i = 1, \dots, d\}$ forme une famille generatrice de $\text{Col}_d(K)$.

Preuve: On prend $V = K^d$. La matrice M est la matrice $\text{mat}_{\mathcal{B}_d^0}(\varphi)$ de l'endomorphisme $\varphi = \varphi_M$ de K^d qui a un vecteur $\mathbf{e}_j^0$, $j \leq d$ de la base canonique, associe le vecteur $\varphi_M(\mathbf{e}_j)$, $j \leq d$ dont les coordonnees dans $\mathcal{B}_d^0$ sont les $(m_{ij})_{i \leq d}$.

La matrice M est inversible si et seulement si φ est inversible et on applique le critere precedent. $\square$

REMARQUE 7.3.4. Notons qu'alors l'inverse de M est la matrice

$$M^{-1} = M' = \text{mat}_{\mathcal{B}_d^0}(\varphi^{-1}) :$$

en effet

$$M.M' = \text{mat}_{\mathcal{B}_d^0}(\varphi).\text{mat}_{\mathcal{B}_d^0}(\varphi^{-1}) = \text{mat}_{\mathcal{B}_d^0}(\varphi.\varphi^{-1}) = \text{mat}_{\mathcal{B}_d^0}(\text{Id}_{K^d}) = \text{Id}_d$$

et de meme $M'.M = \text{Id}_d$. Ainsi M' est l'inverse de M .

7.3.3.2. *Transposition.* soit $\varphi \in \text{End}(V)$ et $\varphi^* \in \text{End}(V^*)$ sa duale alors

$$\text{rg}(\varphi) = \text{rg}(\varphi^*)$$

et

$$\varphi \in \text{GL}(V) \iff \varphi^* \in \text{GL}(V^*).$$

Cela se traduit en terme de matrices.

Soit $M \in M_d(K)$ on a vu que

$$\text{rg}(M) = \text{rg}({}^t M)$$

et donc M est inversible (de rang d) ssi ${}^t M$ est inversible.

Comme la transposition echange lignes et colonnes on obtient

THÉORÈME 7.9 (Critere d'inversibilite pour les matrices (via les lignes)). *Soit une matrice carree $M = (m_{ij})_{i,j \leq d} \in M_d(K)$, les conditions suivantes sont equivalentes*

- (1) M est inversible, ie. $M \in \text{GL}_d(K)$,
- (2) ${}^t M$ est inversible, ie. ${}^t M \in \text{GL}_d(K)$,

- (3) $\text{rg}({}^t M) = d$,
- (4) $\{\text{Lig}_i(M), i = 1, \dots, d\}$ forme une famille libre de $\text{Lig}_d(K)$,
- (5) $\{\text{Lig}_i(M), i = 1, \dots, d\}$ forme une famille generatrice de $\text{Lig}_d(K)$.

La transposition appliquee au groupe lineaire a les proprietes suivantes:

PROPOSITION 7.6. *La transposition est une bijection de $\text{GL}_d(K)$ sur lui-meme qui verifie:*

$$\forall M, N \in \text{GL}_d(K), ({}^t M)^{-1} = {}^t(M^{-1}), {}^t(M.N) = {}^t N.{}^t M.$$

Preuve: Si M est inversible on a

$$M.M^{-1} = M^{-1}.M = \text{Id}_d$$

et donc

$${}^t(M.M^{-1}) = {}^t(M^{-1}).{}^t M = {}^t(M^{-1}.M) = {}^t(M^{-1}).{}^t(M) = {}^t(\text{Id}_d) = \text{Id}_d.$$

Ainsi ${}^t M$ est inversible d'inverse ${}^t(M^{-1})$. □

EXERCICE 7.2. Soit

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

une matrice carree de taille 2.

- (1) Calculer M^2 et montrer qu'il existe $t, \Delta \in K$ (qui dependent de M et qu'on calculera) tels que

$$M^2 - t.M + \Delta.\text{Id}_2 = 0_2.$$

- (2) Montrer que $M \mapsto t(M)$ est lineaire: pour $\lambda \in K, M, N \in M_2(K)$

$$t(\lambda.M + N) = \lambda.t(M) + t(N).$$

- (3) Montrer que $M \mapsto \Delta(M)$ est multiplicative:

$$\Delta(M.N) = \Delta(M).\Delta(N).$$

- (4) Montrer que M est inversible ssi $\Delta(M) \neq 0_K$ et qu'alors

$$M^{-1} = \frac{1}{\Delta(M)}(t(M)\text{Id}_2 - M).$$

7.4. Changement de base

La question est la suivante: soit $\text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi)$ la matrice associee a $\varphi : V \mapsto W$ dans des bases $\mathcal{B} \subset V$ et $\mathcal{B}' \subset W$; soit

$$\mathcal{B}_n = \{\mathbf{e}_{nj}, j \leq d\} \subset V, \mathcal{B}'_n = \{\mathbf{f}_{ni}, i \leq d\} \subset W$$

de nouvelles bases, quelle est la relation entre la matrice de φ dans les bases $\mathcal{B}, \mathcal{B}'$, $\text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi)$ et la matrice de φ dans les bases $\mathcal{B}_n, \mathcal{B}'_n$, $\text{mat}_{\mathcal{B}'_n, \mathcal{B}_n}(\varphi)$? La proposition suivante repond a cette question.

THÉORÈME 7.10 (Formule de changement de base). *Soient $\mathcal{B}, \mathcal{B}_n \subset V$ et $\mathcal{B}', \mathcal{B}'_n \subset W$ des bases de V et W . On a la relation*

$$\text{mat}_{\mathcal{B}'_n, \mathcal{B}_n}(\varphi) = \text{mat}_{\mathcal{B}'_n, \mathcal{B}'}(\text{Id}_W).\text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi).\text{mat}_{\mathcal{B}, \mathcal{B}_n}(\text{Id}_V).$$

Preuve: On a evidemment

$$\varphi = \text{Id}_W \circ \varphi \circ \text{Id}_V.$$

Il suffit alors d'appliquer deux fois la relation (7.2.2) avec des bases convenables: une fois pour $\varphi \circ \text{Id}_V = \varphi$ et l'autre pour $\text{Id}_W \circ \varphi = \varphi$. □

DÉFINITION 7.9. La matrice carree de taille $d = \dim V$,

$$\text{mat}_{\mathcal{B}, \mathcal{B}_n} := \text{mat}_{\mathcal{B}, \mathcal{B}_n}(\text{Id}_V)$$

est appelle matrice de changement de base, de la base $\mathcal{B}$ a la base $\mathcal{B}_n$ ou encore la matrice de passage de $\mathcal{B}$ a $\mathcal{B}_n$.

Sa j -ieme colonne est formee par les coordonnees du j -ieme vecteur $\mathbf{e}_{nj}$ exprime comme combinaison lineaire dans la base $\mathcal{B}$.

La formule de changement de base se reecrit alors

$$\text{mat}_{\mathcal{B}'_n, \mathcal{B}_n}(\varphi) = \text{mat}_{\mathcal{B}'_n, \mathcal{B}'} \cdot \text{mat}_{\mathcal{B}', \mathcal{B}}(\varphi) \cdot \text{mat}_{\mathcal{B}, \mathcal{B}_n}.$$

REMARQUE 7.4.1. On utilise la terminologie (par forcement standard) "matrice de passage de $\mathcal{B}$ a $\mathcal{B}_n$ " car cette matrice permet de calculer la matrice d'une application lineaire φ quand la base de depart est la base $\mathcal{B}_n$ a partir d'une matrice de la meme application quand la base de depart est la base $\mathcal{B}$ et elle permet donc de "passer" d'une matrice d'une application exprimee dans la base $\mathcal{B}$ a sa matrice exprimee dans la base $\mathcal{B}_n$.

Notons que la matrice de passage $\text{mat}_{\mathcal{B}, \mathcal{B}_n}$ est inversible par le critere d'inversibilite. On va calculer son inverse:

PROPOSITION 7.7. Soit trois bases $\mathcal{B}, \mathcal{B}_1, \mathcal{B}_2 \subset V$ on a

(1) Formule d'inversion:

$$\text{mat}_{\mathcal{B}, \mathcal{B}_1} \cdot \text{mat}_{\mathcal{B}_1, \mathcal{B}} = \text{Id}_d.$$

En particulier une matrice de passage est inversible (dans $M_d(K)$) et son inverse est la matrice de passage de la base initiale a la nouvelle base:

$$\text{mat}_{\mathcal{B}, \mathcal{B}_1}^{-1} = \text{mat}_{\mathcal{B}_1, \mathcal{B}}.$$

(2) Formule de transitivite:

$$\text{mat}_{\mathcal{B}, \mathcal{B}_2} = \text{mat}_{\mathcal{B}, \mathcal{B}_1} \cdot \text{mat}_{\mathcal{B}_1, \mathcal{B}_2}.$$

Preuve: Cela resulte de (7.2.3) et de (7.2.2) appliques a $\varphi = \psi = \text{Id}_V$ et a des bases convenables. $\square$

7.4.0.1. *Cas des endomorphismes.* Si $V = W$ et qu'on prend $\mathcal{B}' = \mathcal{B}$ et qu'on se donne une nouvelle base $\mathcal{B}_n = \mathcal{B}'_n$, la formule de changement de base devient alors

$$\text{mat}_{\mathcal{B}_n}(\varphi) = \text{mat}_{\mathcal{B}_n, \mathcal{B}} \cdot \text{mat}_{\mathcal{B}}(\varphi) \cdot \text{mat}_{\mathcal{B}, \mathcal{B}_n} = \text{mat}_{\mathcal{B}, \mathcal{B}_n}^{-1} \cdot \text{mat}_{\mathcal{B}}(\varphi) \cdot \text{mat}_{\mathcal{B}, \mathcal{B}_n}.$$

EXEMPLE 7.4.1. Prenons $V = K^2$ et $\mathcal{B} = \{(1,0), (0,1)\}$ la base canonique. Soit $\mathcal{B}_n = \{(1,3), (1,2)\}$, c'est une base de K^2 (quelque soit la carateristique) et la matrice de passage de $\mathcal{B}$ a $\mathcal{B}_n$ vaut

$$\text{mat}_{\mathcal{B}, \mathcal{B}_n} = \begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix}$$

et la matrice de passage de $\mathcal{B}_n$ a $\mathcal{B}$ est l'inverse

$$\text{mat}_{\mathcal{B}_n, \mathcal{B}} = - \begin{pmatrix} 2 & -1 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 3 & -1 \end{pmatrix}$$

7.4.1. Matrices équivalentes. Soit $\varphi : V \mapsto W$ et $\mathcal{B}, \mathcal{B}_n, \mathcal{B}', \mathcal{B}'_n$ des paires de bases de V et W alors les matrices représentant φ dans ces bases

$$M = \text{mat}_{\mathcal{B}'\mathcal{B}}(\varphi), \quad N = \text{mat}_{\mathcal{B}'_n\mathcal{B}_n}(\varphi)$$

sont liées par la relation

$$N = A.M.B$$

avec

$$A = \text{mat}_{\mathcal{B}'_n\mathcal{B}'}, \quad B = \text{mat}_{\mathcal{B}\mathcal{B}_n}$$

les matrices de changement de bases qui sont inversibles. Comme M et N représentent la même application linéaire on peut dire qu'elles sont d'une certaine manière équivalentes. Cela induit la définition purement matricielle suivante:

DÉFINITION 7.10. Deux matrices $M, N \in M_{d' \times d}(K)$ sont dites équivalentes si il existe des matrices inversibles $A \in \text{GL}_{d'}(K)$, $B \in \text{GL}_d(K)$ telles que

$$N = A.M.B.$$

Par la formule de changement de bases on a:

PROPOSITION 7.8. Deux matrices $M, N \in M_{d' \times d}(K)$ sont équivalentes ssi il existe V de dimension d et W de dimension d' , des bases $\mathcal{B}, \mathcal{B}_n \subset V$ et $\mathcal{B}', \mathcal{B}'_n \subset W$ et une application linéaire $\varphi : V \mapsto W$ telle que

$$M = \text{mat}_{\mathcal{B}'\mathcal{B}}(\varphi), \quad N = \text{mat}_{\mathcal{B}'_n\mathcal{B}_n}(\varphi)$$

Preuve: Le fait que des matrices M et N qui sont les matrices d'un même endomorphisme φ dans différentes bases, vérifient la relation

$$N = A.M.B$$

avec A et B inversibles résulte de la formule de changement de base en prenant A et B des matrices de passage convenable.

Reciproquement, supposons que l'on ait la relation

$$N = A.M.B$$

avec A et B inversibles. Soit $V = K^d$, $W = K^{d'}$ et $\mathcal{B} \subset V, \mathcal{B}' \subset W$ les bases canoniques et $\varphi_K^d \mapsto K^{d'}$ l'unique application linéaire qui envoie le j -ième vecteur de la base canonique $\mathcal{B}$ vers le vecteur de W dont les coordonnées dans la base canonique $\mathcal{B}'$ soient données par la j -ième colonne de M : on a donc

$$M = \text{mat}_{\mathcal{B}'\mathcal{B}}(\varphi).$$

Soit $\mathcal{B}_n$ la base formée des vecteurs de K^d dont le j -ième vecteur a pour coordonnées (dans la base canonique $\mathcal{B}$) la j -ième colonne de B ; en effet ces vecteurs forment une base car comme B est inversible, donc de rang d , les vecteurs colonnes de B forment une famille génératrice de l'espace des vecteurs colonnes de taille d qui est donc libre. On a donc

$$B = \text{mat}_{\mathcal{B}\mathcal{B}_n}.$$

Soit $\mathcal{B}'_n$ la base formée des vecteurs de $K^{d'}$ dont le i -ième vecteur a pour coordonnées (dans la base canonique $\mathcal{B}'$) la j -ième colonne de A^{-1} : on a donc

$$A^{-1} = \text{mat}_{\mathcal{B}'\mathcal{B}'_n} \text{ et donc } A = \text{mat}_{\mathcal{B}'_n\mathcal{B}'}$$

Alors la formule de changement de base nous dit que

$$N = A.M.B = \text{mat}_{\mathcal{B}'_n\mathcal{B}'}.\text{mat}_{\mathcal{B}'\mathcal{B}}(\varphi).\text{mat}_{\mathcal{B}\mathcal{B}_n} = \text{mat}_{\mathcal{B}'_n\mathcal{B}_n}(\varphi)$$

C'est à dire

$$N = \text{mat}_{\mathcal{B}'_n\mathcal{B}_n}(\varphi).$$

□

PROPOSITION. La relation "etre equivalente" est une relation d'equivalence (reflexive, symetrique, transitive) sur $M_{d' \times d}(K)$.

Preuve: Ecrivons la relation $M \sim N$. Reflexive: on a $M = \text{Id}_{d'} M \text{Id}_d$ donc $M \simeq M$.

Symetrique: si $M \simeq N$ on a $N = AMB$, $A \in \text{GL}_{d'}(K)$, $B \in \text{GL}_d(K)$ et

$$A^{-1}NB^{-1} = A^{-1}AMB B^{-1} = M$$

et $N \sim M$.

Transitive: si $M \sim N$ et $N \sim P$ alors

$$P = ANB, N = A'MB' \implies P = AA'MB'B$$

et $AA' \in \text{GL}_{d'}(K)$, $B'B \in \text{GL}_d(K)$ ainsi $M \sim P$. □

On en deduit le resultat suivant

THÉORÈME 7.11. Soient $M, N \in M_{d' \times d}(K)$. Les conditions suivantes sont equivalentes

- (1) M et N sont equivalentes,
- (2) $\text{rg}(M) = \text{rg}(N)$,
- (3) M et N sont equivalentes a $I_{d' \times d}(r)$.

Preuve: Par la proposition precedente, deux matrices sont equivalentes ssi elle representent la meme application lineaire φ dans des bases differentes. En particulier, elles ont le meme rang (celui de φ).

Si M et N ont meme rang elles sont les matrices d'applications lineaires φ, φ' de meme rang. On a vu qu'une application lineaire φ de rang r admettait pour matrice

$$I_{d' \times d}(r) = \begin{pmatrix} & & 0 & 0 \\ & \text{Id}_r & \vdots & \vdots \\ & & \vdots & \vdots \\ 0 & \dots & \dots & 0 & 0 \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix}$$

dans des bases convenables (cf. §7.2.3.1) et donc, par la proposition precedente, toute matrice equivalente a $I_{d' \times d}(r)$ est la matrice de φ dans des bases convenables. Ainsi les matrices de M et N sont equivalentes a $I_{d' \times d}(r)$.

Finalement si les matrices de M et N sont equivalentes a $I_{d' \times d}(r)$ alors elles sont equivalentes (par transitivite de la relation d'equivalence). □

REMARQUE 7.4.2. La proposition precedente nous dit que toute matrice $d' \times d$ est equivalente a une des matrices de la forme

$$\{I_{d' \times d}(r), 0 \leq r \leq \min(d, d')\}$$

et comme ces matrices sont de rang distincts elle ne sont pas equivalentes: ces matrices forment un ensemble de representants des differentes classes d'equivalence de la relation equivalence de matrices sur $M_{d' \times d}(K)$. Ainsi l'ensemble des classes d'equivalences

$$M_{d' \times d}(K) / \sim \simeq \{I_{d' \times d}(r), 0 \leq r \leq \min(d, d')\}$$

est un ensemble fini de $\min(d, d') + 1$ elements.

7.4.2. Matrices semblables/conjuguees. Supposons maintenant que

$$\varphi : V \mapsto V$$

soit un endomorphisme et soit $\mathcal{B}, \mathcal{B}_n$ des bases de V . Posons encore

$$M = \text{mat}_{\mathcal{B}\mathcal{B}}, N = \text{mat}_{\mathcal{B}_n\mathcal{B}_n} \in M_d(K).$$

On a alors par changement de base

$$N = C.M.D$$

avec

$$C = \text{mat}_{\mathcal{B}_n \mathcal{B}}, D = \text{mat}_{\mathcal{B} \mathcal{B}_n} = (\text{mat}_{\mathcal{B}_n \mathcal{B}})^{-1} = C^{-1}$$

ou encore

$$N = C.M.C^{-1}.$$

Ainsi, la formule de changement de base met en evidence une autre relation sur $M_d(K)$:

DÉFINITION 7.11. *On dit que deux matrices M, N sont semblables ou conjuguées si il existe $C \in GL_d(K)$ tel que*

$$N = C.M.C^{-1}.$$

*La relation "etre semblables" ou "etre conjuguées" est une relation d'équivalence.
Une classe d'équivalence pour cette relation, l'ensemble des matrices de la forme*

$$M^{\natural} := \text{Ad}(GL_d(K))(M) = \{C.M.C^{-1}, C \in GL_d(K)\}$$

est appelée classe de conjugaison (de M) et on note

$$M_d(K)^{\natural} = \{M^{\natural}\} = M_d(K) / \sim$$

l'ensemble des classes de conjugaison.

EXERCICE 7.3. Verifier directement a partir de la definition que l'on a bien une relation d'équivalence (reflexive, symetrique, transitive).

REMARQUE 7.4.3. On a vu que deux matrices representant le meme endomorphisme sont conjuguées. La reciproque est vraie:

PROPOSITION 7.9. *Deux matrices $M, N \in GL_d(K)$ sont semblables ssi M et N sont les matrices d'un meme endomorphisme dans des bases convenables: il existe un espace vectoriel de dimension d , V , deux bases $\mathcal{B}, \mathcal{B}_n \subset V$ et une application lineaire $\varphi : V \mapsto V$ telle que*

$$M = \text{mat}_{\mathcal{B}}(\varphi), N = \text{mat}_{\mathcal{B}_n}(\varphi).$$

EXERCICE 7.4. Completer la preuve et montrer que si $M = \text{mat}_{\mathcal{B}}(\varphi)$ est la matrice representant un endomorphisme $\varphi \in \text{End}(V)$ dans une base $\mathcal{B} \subset V$ alors $M^{\natural}$ est l'ensemble des matrices $\text{mat}_{\mathcal{B}'}(\varphi)$ quand $\mathcal{B}'$ parcourt toutes les bases de V .

REMARQUE 7.4.4. Deux matrices $M, N \in M_d(K)$ carrees de meme taille qui sont semblables sont équivalentes (prendre $A = C, B = C^{-1}$) et en particulier ont meme rang. La reciproque n'est pas vraie.

REMARQUE 7.4.5. On a vu que pour la relation "équivalence de matrices" dans $M_{d' \times d}(K)$ l'espace quotient des classes d'équivalences etait tres simple: c'est un ensemble fini de $\min(d, d') + 1$ elements representes par les matrices standard de rang $0 \leq r \leq \min(d, d')$

$$I_{d' \times d}(r), r = 0, \dots, \min(d, d').$$

Il est beaucoup plus difficile de decrire $M_d(K)^{\natural}$, l'ensemble des differentes classes de conjugaisons de matrices dans $M_d(K)$. Si le corps K est *algebriquement clos* (par exemple $K = \mathbb{C}$) cette classification est donnee par la *decomposition de Jordan* qui releve du semestre prochain. Et avant cela vous aurez besoin de la notion de polynome caracteristique et du Theoreme de Cayley-Hamilton.

7.4.3. Action par conjugaison.

DÉFINITION 7.12. Soit $C \in GL_d(K)$ une matrice inversible. Note note $\text{Ad}(C)$ l'application dite de conjugaison par C :

$$\text{Ad}(C) : \begin{array}{ccc} M_d(K) & \mapsto & M_d(K) \\ M & \mapsto & C.M.C^{-1}. \end{array}$$

Ainsi deux matrices sont semblables si et seulement si elles sont image l'une de l'autre par conjugaison par une matrice inversible.

EXEMPLE 7.4.2. Si $C = \text{mat}_{\mathcal{B}_1, \mathcal{B}}$ est une matrice de changement de base (de la base $\mathcal{B}$ à la base $\mathcal{B}_1$) alors la formule de changement de base pour les matrices carrées s'écrit

$$\text{mat}_{\mathcal{B}_1}(\varphi) = \text{Ad}(\text{mat}_{\mathcal{B}_1, \mathcal{B}})(\text{mat}_{\mathcal{B}}(\varphi)).$$

Propriétés fonctionnelles de la conjugaison.

PROPOSITION 7.10. La conjugaison $\text{Ad}(C)$ est un automorphisme de l'algèbre $M_d(K)$:

- (1) *Linearité*: On a $\text{Ad}(C)(\lambda.M + N) = \lambda\text{Ad}(C)(M) + \text{Ad}(C)(N)$.
- (2) *Multiplicativité*: $\text{Ad}(C)(M.N) = \text{Ad}(C)(M).\text{Ad}(C)(N)$.
- (3) *Inversibilité*: $\text{Ad}(C)$ est bijective et $\text{Ad}(C)^{-1} = \text{Ad}(C^{-1})$.

Preuve: On a

$$\begin{aligned} \text{Ad}(C)(\lambda.M + N) &= C.(\lambda.M + N).C^{-1} = (\lambda.C.M + C.N).C^{-1} \\ &= \lambda.C.M.C^{-1} + C.N.C^{-1} = \lambda\text{Ad}(C)(M) + \text{Ad}(C)(N). \end{aligned}$$

On a

$$\text{Ad}(C)(M.N) = C.M.N.C^{-1} = C.M.\text{Id}_d.N.C^{-1} = C.M.C^{-1}.C.N.C^{-1} = \text{Ad}(C)(M).\text{Ad}(C)(N).$$

Par ailleurs

$$\text{Ad}(C^{-1})(\text{Ad}(C)(M)) = C^{-1}.C.M.C^{-1}.C = M$$

et donc

$$\text{Ad}(C^{-1}) \circ \text{Ad}(C) = \text{Id}_{M_d(K)}$$

□

PROPOSITION 7.11. On dispose donc d'une application

$$\text{Ad}(\bullet) : C \in GL_d(K) \mapsto \text{Ad}(C) \in \text{Aut}(M_d(K)) \simeq GL_{d^2}(K)$$

appelée application adjointe.

L'application adjointe $\text{Ad}(\bullet)$ est un morphisme de groupes et définit donc une action à gauche $GL_d(K) \curvearrowright M_d(K)$. Son noyau est formé par les matrices scalaires:

$$\ker \text{Ad} = K^\times \text{Id}.$$

Preuve: On a déjà vu que $\text{Ad}(C)^{-1} = \text{Ad}(C^{-1})$. Reste à voir que

$$\text{Ad}(B.C) = \text{Ad}(B) \circ \text{Ad}(C).$$

On a

$$\text{Ad}(B.C)(M) = B.C.M.(B.C)^{-1} = B.C.M.C^{-1}.B^{-1} = \text{Ad}(B)(\text{Ad}(C)(M)).$$

Soit $C = (c_{kl})_{k,l \leq d}$ une matrice inversible telle que pour tout M on ait

$$C.M.C^{-1} = M.$$

On a donc pour tout M

$$C.M = M.C.$$

En particulier $\forall i, j \leq d$

$$C.E_{ij} = E_{ij}.C.$$

On a par la proposition 7.2

$$(\sum_{k,l} c_{kl} E_{kl}).E_{ij} = \sum_{k,l} c_{kl} E_{kl}.E_{ij} = \sum_{k,l} c_{kl} \delta_{l=i} E_{kj} = \sum_k c_{ki} E_{kj}$$

et

$$E_{ij}.(\sum_{k,l} c_{kl} E_{kl}) = \sum_{k,l} c_{kl} E_{ij}.E_{kl} = \sum_{k,l} c_{kl} \delta_{k=j} E_{il} = \sum_l c_{jl} E_{il}$$

On a donc necessairement dans les sommes ci-dessus $c_{ki} = 0$ si $k \neq j$ et comme c'est valable pour tout j on voit que $c_{ij} = 0$ sauf si $i = j$. on a donc

$$C.E_{ij} = c_{ii} E_{ij} = E_{ij}.C = c_{jj} E_{ij}$$

ce qui force les c_{ii} a etre tous egaux et donc $C = c_{11}.Id_d$ est une matrice scalaire. $\square$

DÉFINITION 7.13. *L' image $\text{Ad}(\text{GL}_d(K)) \subset \text{Aut}(M_d(K))$ est appelee groupe des automorphismes interieurs de $M_d(K)$ et est notee*

$$\text{Int}(M_d(K)) \subset \text{Aut}_K(M_d(K)).$$

La relation "etre semblable" est une relation d'equivalence. On peut soit le verifier directement a l'aide des proprietes fonctionelles de la conjugaison soit en notant que celle relation est definie via l'action par conjugaison $\text{GL}_d(K) \curvearrowright M_d(K)$: on a vu en exercice que etant donne une action d'un groupe sur un ensemble

$$G \curvearrowright X$$

la relation sur X donnee par

$$x \sim_G x' \iff \exists g \in G, x' = g \star x$$

est une relation d'equivalence (la relation d'appartenance a la meme G -orbite: $x' \in G \star x$).

En effet une telle relation est

- Symetrique: $x = e_G \star x$
- Reflexive:

$$x' = g \star x \implies x = g^{-1} \star x'.$$

- Transitive:

$$x'' = g' \star x', x' = g \star x \implies x'' = g' \star (g \star x) = (g'.g) \star x$$

Ici l'action est

$$C \star M = C.M.C^{-1}.$$

7.4.4. Conjugaison des endomorphismes. On peut egalement definir une notion de conjugaison pour l'algebre (abstraite) $\text{End}(V)$ des endomorphismes d'un espace V en disant que $\varphi, \phi \in \text{End}(V)$ sont conjugues si il existe $\psi \in \text{Aut}(V)$ tel que

$$\phi = \psi \circ \varphi \circ \psi^{-1}.$$

Si on choisit une base $\mathcal{B}$ de V et qu'on l'utilise pour identifier $\text{End}(V)$ avec $M_d(K)$ on obtient exactement la meme notion ($C = \text{mat}_{\mathcal{B}}(\psi)$).

EXERCICE 7.5. Soit V et W des espaces vectoriels de dimension finie de meme dimension alors $\text{End}(V)$ et $\text{End}(W)$ sont des K -EV isomorphes car de meme dimension d^2). Montrer qu'ils sont isomorphes en tant que K -algebres; pour cela construire un isomorphisme de K -algebres

$$\text{End}(W) \simeq \text{End}(V)$$

a partir d'un isomorphisme $\psi : V \simeq W$.

CHAPITRE 8

Interlude: le corps des nombres complexes

*"... eine feine und wunderbare Zuflucht des menschlichen Geistes,
beinahe ein Zwitterwesen zwischen Sein und Nichtsein."*

"Even better than the real thing."

8.1. Origine des nombres complexes

Les nombres complexes sont nés pendant la renaissance italienne dans le but de résoudre des équations polynomiales: étant donné $a_0, \dots, a_{d-1}, a_d \in \mathbb{Z}$, on cherchait à trouver les nombres z vérifiant

$$a_d z^d + a_{d-1} z^{d-1} + \dots + a_1 z + a_0 = 0.$$

En particulier pour $d = 2$, on savait que les solutions d'une équation quadratique

$$az^2 + bz + c = 0$$

étaient de la forme

$$z_{\pm} = \frac{-b \pm \sqrt{\Delta}}{2a}$$

avec

$$\Delta = b^2 - 4ac$$

pour peu que Δ soit positif ou nul. On n'avait pas de problème à travailler avec les nombres tels que $\sqrt{\Delta}$, même si Δ n'est pas le carré d'un entier car on définissait ce nombre comme le côté d'un carré d'aire Δ . En revanche on évitait soigneusement les cas où $\Delta < 0$.

Les mathématiciens se sont également intéressés aux équations cubiques et quartiques (de degré 3 ou 4), notamment les mathématiciens de la renaissance italienne (Del Ferro, Tartaglia, Cardano, Ferrari, Bombelli)

$$az^3 + bz^2 + cz + d = 0, \quad az^4 + bz^3 + cz^2 + dz + e = 0, \quad a, b, c, d, e \in \mathbb{Z}.$$

Dans son ouvrage *Ars Magna* (1545), Cardano (suivant del Ferro) a donné une méthode algorithmique pour trouver les solutions de nombreuses familles d'équations cubiques.

L'une d'elles était soigneusement évitée

$$(8.1.1) \quad z^3 = 15z + 4.$$

Bien qu'elle admette, 4 comme solution (tout a fait naturelle), la méthode suivie par Cardano le conduisait à résoudre l'équation

$$x^2 + 121 = 0.$$

Cardano s'est refusé à introduire la solution formelle

$$\sqrt{-121} = 11\sqrt{-1}$$

dans ses formules generales. C'est Bombelli¹ qui, 30 ans plus tard, sautant le pas introduisit les regles de calcul impliquant des nombres imaginaires tels que $\sqrt{-121}$ et il retrouvera ainsi la solution 4 de (8.1.1) a partir des formules generales de del Ferro et Cardano².

Dans ce chapitre, on va construire concretement le corps des nombres complexes comme une sous-algebre de l'algebre des matrices reelles 2×2 , $M_2(\mathbb{R})$. C'est en fait un cas particulier d'une construction generale basee sur l'anneau des polynomes a coefficients dans un corps K ,

$$K[X] = \{a_0 + a_1.X + \dots + a_d.X^d, d \geq 0, a_0, \dots, a_d \in K\}$$

qu'on verra au chapitre sur les anneaux de polynomes.

8.2. Construction matricielle d'extensions quadratiques

On commence par une construction generale (la solution d'un exercices d'une des series precedentes).

On rappelle que pour toute matrice

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(K)$$

son determinant est le scalaire

$$\det(M) = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Ce dernier verifie (par calcul direct)

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

et on a

$$M \in \text{GL}_2(K) \text{ (} M \text{ est inversible) ssi } \det(M) \neq 0$$

et on a alors

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

THÉORÈME 8.1. *Soit K un corps et $M_2(K)$ l'algebre des matrices 2×2 a coefficients dans K . Soit $d \in K - K^2$ un element de K qui n'est pas un carre: $\forall x \in K, x^2 - d \neq 0$ et*

$$I_d := \begin{pmatrix} 0 & d \\ 1 & 0 \end{pmatrix}.$$

Alors la matrice I_d verifie

$$I_d^2 = d.\text{Id}_2.$$

Soit

$$K[I_d] = K.\text{Id}_2 + K.I_d = \left\{ Z = x.\text{Id}_2 + y.I_d = \begin{pmatrix} x & dy \\ y & x \end{pmatrix}, x, y \in K \right\} \subset M_2(K)$$

le SEV de $M_2(K)$ engendre par Id_2 et I_d . Alors $K[I_d]$ a les proprietes suivantes:

- (1) $\{\text{Id}_2, I_d\}$ est une base de $K[I_d]$ et donc $\dim_K(K[I_d]) = 2$.
- (2) $K[I_d]$ muni du produit de matrices est un sous-anneau commutatif de $M_2(K)$ et c'est meme un corps : toute matrice non-nulle de $K[I_d]$ est inversible dans $K[I_d]$.

¹un cratere de la lune porte son nom.

²on renvoie a <https://www.youtube.com/watch?v=cUzklzVXJwo&t=1072s> pour une video passionnante expliquant cette histoire

(3) Plus précisément soit

$$Z = x\text{Id}_2 + y.I_d = \begin{pmatrix} x & dy \\ y & x \end{pmatrix}$$

alors

$$\det(Z) = x^2 - dy^2$$

et si $\det(Z) \neq 0$ (alors Z est inversible) on a

$$Z^{-1} = \frac{1}{x^2 - dy^2} (x.\text{Id}_2 - y.I_d) = \begin{pmatrix} \frac{x}{x^2 - dy^2} & d\frac{-y}{x^2 - dy^2} \\ \frac{-y}{x^2 - dy^2} & \frac{x}{x^2 - dy^2} \end{pmatrix} \in K[I_d].$$

Preuve: On a

$$Z = x\text{Id}_2 + y.I_d = \begin{pmatrix} x & dy \\ y & x \end{pmatrix} = \mathbf{0}_2 \iff x = y = 0$$

donc $\{\text{Id}_2, I_d\}$ est libre et elle est génératrice de $K[I_d]$ par définition.

Montrons que c'est un sous-anneau de $M_2K(K)$: on a évidemment $\text{Id}_2 \in K[I_d]$ et il reste à montrer que $K[I_d]$ est stable par produit: soient

$$Z = x\text{Id}_2 + y.I_d = \begin{pmatrix} x & dy \\ y & x \end{pmatrix}, \quad Z' = x'\text{Id}_2 + y'.I_d = \begin{pmatrix} x' & dy' \\ y' & x' \end{pmatrix} \in K[I_d]$$

on veut montrer que

$$Z.Z' \in K[I_d].$$

On peut prendre brutalement le produit de matrices et on trouve

$$Z.Z' = \begin{pmatrix} xx' + dyy' & (xy' + yx')d \\ xy' + yx' & xx' + dyy' \end{pmatrix} = (xx' + dyy')\text{Id}_2 + (xy' + yx')I_d \in K[I_d].$$

On peut également faire le calcul de manière plus conceptuelle à partir de l'équation

$$I_d^2 = I_d.I_d = \begin{pmatrix} 0 & d \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & d \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0.0 + d.1 & 0.d + d.0 \\ 1.0 + 0.1 & 01.d + 0.0 \end{pmatrix} = \begin{pmatrix} d & 0 \\ 0 & d \end{pmatrix} = d.\text{Id}_2;$$

comme $\text{Id}_2^2 = \text{Id}_2$ et $I_d^2 = d.\text{Id}_2$, on a par distributivité et associativité

$$\begin{aligned} Z.Z' &= (x\text{Id}_2 + y.I_d).(x'\text{Id}_2 + y'.I_d) = xx'.\text{Id}_2 + (xy' + yx')I_d + yy'.d\text{Id}_2 \\ &= (xx' + dyy')\text{Id}_2 + (xy' + yx')I_d \in K[I_d]. \end{aligned}$$

Comme (K est commutatif)

$$xx' + dyy' = x'x + dy'y, \quad xy' + yx' = x'y + y'x$$

on a donc

$$Z.Z' = Z'.Z$$

et donc l'anneau $K[I_d]$ est commutatif.

Montrons que tout élément non-nul est inversible (et que son inverse est contenu dans $K[I_d]$): soit

$$Z = Z = x\text{Id}_2 + y.I_d = \begin{pmatrix} x & dy \\ y & x \end{pmatrix}$$

alors

$$\det Z = x^2 - dy^2.$$

Supposons que $\det Z = 0$ alors

$$x^2 = dy^2;$$

si $y = 0$ alors $x = 0$ et $Z = \mathbf{0}_2$. Si $y \neq 0$ alors

$$d = (x/y)^2 \in K^2$$

ce qui contredit l'hypothese que d n'est pas un carre. Ainsi

$$Z \neq \mathbf{0}_2 \iff \det Z = x^2 - dy^2 \neq 0 \iff Z \in GL_2(K).$$

Ainsi

$$Z^{-1} = \frac{1}{\det Z} \begin{pmatrix} x & -dy \\ -y & x \end{pmatrix} = \frac{1}{x^2 - dy^2} (x \cdot \text{Id}_2 - y \cdot I_d) \in K[I_d]$$

□

8.2.0.1. *Conjugaison algebrique.* Etant donne $Z = x\text{Id}_2 + yI_d \in K[I_d]$, on pose

$$\overline{Z} = x\text{Id}_2 - yI_d \in K[I_d]$$

qu'on appelle le *conjugue algebrique* de Z . La conjugaison algebrique $Z \mapsto \overline{Z}$ a les proprietes suivantes:

PROPOSITION 8.1. *L'application*

$$\overline{\bullet} : \begin{array}{ccc} K[I_d] & \mapsto & K[I_d] \\ Z & \mapsto & \overline{Z} \end{array}$$

verifie

(1) *Est lineaire:* $\forall \lambda \in K, Z, Z' \in K[I_d]$,

$$\overline{\lambda \cdot Z + Z'} = \lambda \overline{Z} + \overline{Z'}.$$

(2) *Est involutive (en particulier bijective)*

$$\overline{\overline{Z}} = Z.$$

(3) *Est un morphisme de corps: en particulier en on a*

$$\overline{Z \cdot Z'} = \overline{Z} \cdot \overline{Z'}.$$

(4) *On a*

$$Z \cdot \overline{Z} = (x^2 - dy^2) \text{Id}_2.$$

En particulier si $Z \neq \mathbf{0}_2$, on a

$$Z^{-1} = \frac{1}{x^2 - dy^2} \overline{Z}.$$

Preuve: On peut demontrer cela par un calcul direct. □

REMARQUE 8.2.1. Notons que dans $M_2(K)$, on peut trouver un grand nombre de matrices I'_d verifiant

$$I_d'^2 = d \cdot \text{Id}_2,$$

en effet pour tout $C \in GL_2(K)$ la matrice conjuguee

$$\text{Ad}(C)(I_d) = C \cdot I_d \cdot C^{-1}$$

a cette propriete.

8.2.1. Notation algebrique. L'application

$$\lambda \in K \mapsto \lambda \cdot \text{Id}_2 = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \in K \cdot \text{Id}_2 \subset M_2(K)$$

identifie K avec l'ensemble des matrices scalaires qui forme un sous-corps de $M_2(K)$. Comme $K[I_d]$ contient $K \cdot \text{Id}_2$, on peut de cette maniere voir K comme un sous-corps de $K[I_d]$. Comme I_d verifie

$$I_d^2 = d \cdot \text{Id}_2.$$

Si on identifie K au corps des matrices scalaires, d est identifie a $d \cdot \text{Id}_2$ et la matrice I_d est une "racine carree" de d , une autre racine carree etant $-I_d$.

Si on a juste besoin de travailler avec le corps $K[I_d]$, plutot que d'ecrire ses elements sous forme de matrices, on ecira

- 1 pour Id_2 , x pour la matrice scalaire $x.\text{Id}_2$,
- $\sqrt{d}$ pour la matrice I_d , et $y\sqrt{d}$ pour la matrice $y.I_d$
- et a la place de

$$Z = x.\text{Id}_2 + yI_d = \begin{pmatrix} x & dy \\ y & x \end{pmatrix} \text{ on ecrira } z = x + y\sqrt{d}.$$

- On ecrira egalement $K[\sqrt{d}]$ pour $K[I_d]$. Cette ecriture permet de représenter naturellement K comme sous-corps de $K[\sqrt{d}]$:

$$K = \{x + 0.\sqrt{d}, x \in K\} \subset K[\sqrt{d}].$$

Ainsi les sommes, produits et conjugue algebrique s'ecrivent $Z + Z'$ et $Z.Z'$, $\bar{Z}$ s'ecrivent sous la forme

$$z + z' = x + x' + (y + y')\sqrt{d}, \quad z.z' = xx' + dy y' + (xy' + yx')\sqrt{d}, \quad \bar{z} = x - y\sqrt{d}.$$

REMARQUE 8.2.2. Notons egalement qu'on peut ecrire

$$y\sqrt{d} = \sqrt{d}y$$

(car $y.I_d = y.\text{Id}_2.I_d = I_d.y.\text{Id}_2$).

Avec cette ecriture la relation (4) devient

$$(8.2.1) \quad z.\bar{z} = x^2 - dy^2,$$

et si $z \neq 0$ on a

$$(8.2.2) \quad z^{-1} = \frac{1}{x^2 - dy^2} \bar{z} = \frac{x}{x^2 - dy^2} - \frac{y}{x^2 - dy^2} \sqrt{d}.$$

DÉFINITION 8.1. Le scalaire $x^2 - dy^2 \in K$ (le determinant de la matrice Z) est appelle norme algebrique de z et est note

$$\text{Nr}_K(z) = \text{Nr}_K(x + y\sqrt{d}) = z\bar{z} = x^2 - dy^2.$$

Comme le determinant est multiplicatif ($\det(Z.Z') = \det(Z).\det(Z')$), la norme algebrique est multiplicative

$$(8.2.3) \quad \text{Nr}_K(z.z') = \text{Nr}_K(z) \text{Nr}_K(z'),$$

et on rappelle que

$$\text{Nr}_K(z) = 0 \iff z = 0.$$

Comme $K[\sqrt{d}]$ est un K -ev de dimension 2, on dit que le corps $K[\sqrt{d}]$ est une extension quadratique du corps K .

REMARQUE 8.2.3. LK'algebre $M_2(K)$ contient beaucoup de "racines carrees" de d : pour tout $C \in \text{GL}_2(K)$

$$I'_d = \text{Ad}(C)(I_d) = C.I_d.C^{-1}$$

verifie

$$I_d'^2 = \text{Id}_2.$$

8.3. Le corps des nombres complexes; proprietes de base

Prenons $K = \mathbb{R}$ alors $d = -1$ n'est pas un carre car -1 est negatif. La matrice I_{-1} vaut alors

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

DÉFINITION 8.2. *Le sous-corps de $M_2(\mathbb{R})$*

$$\mathbb{R}[I] = \mathbb{R}.\text{Id}_2 + \mathbb{R}.I = \left\{ Z = x. \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + y \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} x & -y \\ y & x \end{pmatrix}, x, y \in \mathbb{R} \right\}$$

est appele corps des nombres complexes et est note $\mathbb{C}$. La conjugaison algebrique

$$Z = x\text{Id}_2 + yI \mapsto x\text{Id}_2 - yI$$

s'appelle conjugaison complexe.

Comme precedement, on note les nombres complexes de maniere condensee en ecrivant

$$i = \sqrt{-1}$$

a la place de I et

$$z = x + iy = x + yi \text{ a la place de } Z = x\text{Id}_2 + yI = \begin{pmatrix} x & -y \\ y & x \end{pmatrix}.$$

On a alors

$$z + z' = x + x' + (y + y')i, \quad z.z' = xx' - yy' + (xy' + yx')i, \quad \bar{z} = x - yi$$

et

$$\text{Nr}_{\mathbb{R}}(z) = z.\bar{z} = x^2 + y^2$$

et (8.2.3) devient

$$\text{Nr}_{\mathbb{R}}(z) \text{Nr}_{\mathbb{R}}(z') = (x^2 + y^2)(x'^2 + y'^2) = \text{Nr}_{\mathbb{R}}(z.z') = (xx' - yy')^2 + (xy' + yx')^2.$$

REMARQUE 8.3.1. On a

$$i^3 = -i, i^4 = 1, i^5 = i, \dots$$

et donc

$$i^n = \pm 1 \text{ ou bien } \pm i$$

suisant la classe de congruence $n \pmod{4}$.

DÉFINITION 8.3. *Le reel x est appele "partie reelle" de z et le reel y est la "partie imaginaire" de z*

$$x = \text{Re} z, \quad y = \text{Im} z.$$

Dans la notation matricielle, la conjugaison algebrique est donnee par la transposition:

$$Z = x.\text{Id}_2 + y.I \mapsto {}^t Z = x.\text{Id}_2 - y.I.$$

Avec la notation simplifiee la conjugaison algebrique

$$z = x + iy \mapsto \bar{z} = x - yi$$

s'appelle la conjugaison complexe. On a alors

$$z.\bar{z} = \text{Nr}_{\mathbb{R}}(z) = x^2 + y^2 \geq 0.$$

Comme ce reel est positif ou nul, il admet deux racine carrees dans $\mathbb{R}$, on note $|z|$ celle qui est positive ou nulle:

$$|z| = (z.\bar{z})^{1/2} = (x^2 + y^2)^{1/2} \geq 0;$$

on l'appelle le module de z .

PROPOSITION 8.2. *On a la proprietes suivantes:*

(1) *Les applications "partie reel" et "imaginaire"*

$$\text{Re}, \text{Im} : \mathbb{C} \mapsto \mathbb{R}$$

sont lineaires:

$$\lambda \in \mathbb{R}, \text{Re}(\lambda.z + z') = \lambda.\text{Re}z + \text{Re}z', \text{Im}(\lambda.z + z') = \lambda.\text{Im}z + \text{Im}z'.$$

Les noyaux valent $\ker(\text{Im}) = \mathbb{R}$ et $\ker(\text{Re}) = \mathbb{R}.i$ est l'ensemble des nombres complexes imaginaires purs.

(2) *La conjugaison complexe*

$$\bar{\bullet} : z \in \mathbb{C} \mapsto \bar{z} \in \mathbb{C}$$

est un automorphisme du corps $\mathbb{C}$: in particulier

$$\lambda \in \mathbb{R}, \overline{\lambda.z + z'} = \lambda.\bar{z} + \bar{z}', \overline{z.z'} = \bar{z}.\bar{z}'.$$

De plus $\bar{\bullet}$ est involutif

$$\overline{\bar{z}} = z$$

et on a

$$\bar{z} = z \iff z = x \in \mathbb{R}.$$

(3) *L'application module*

$$z \mapsto |z| = (z.\bar{z})^{1/2}$$

est multiplicative:

$$|z.z'| = |z|.|z'|$$

et on a

$$z = 0 \iff |z| = 0$$

et pour tout $x \in \mathbb{R} \subset \mathbb{C}$ on a

$$(8.3.1) \quad |x| = |x|_{\mathbb{R}} = \max(x, -x)$$

Autrement dit, le module d'un nombre reel est egal a la "valeur absolue" usuelle de ce nombre reel.

Preuve: (1) Les applications $\text{Re} : \mathbb{C} \mapsto \mathbb{R}$ et $\text{Im} : \mathbb{C} \mapsto \mathbb{R}$ sont lineaires car ce sont les formes lineaires " premiere et seconde coordonnee" de la base $\{\text{Id}_2, I\}$ et on peut egalement le verifier directement.

Ces formes lineaires sont non-nulles donc surjectives sur $\mathbb{R}$. On a

$$\ker(\text{Re}) = \{0 + iy, y \in \mathbb{R}\} = \mathbb{R}.i, \ker(\text{Im}) = \{x + 0i, x \in \mathbb{R}\} = \mathbb{R}.$$

(2) La conjugaison algebrique est un cas particulier de conjugaison algebrique et a les meme proprietes de lineairite, multiplicativite et involutivite.

– On a

$$\bar{z} = z \iff \bar{z} = x - iy = x + iy = z \iff 2iy = 0 \iff y = 0 \iff z = x \in \mathbb{R}.$$

(en effet $2.i$ est non nul donc inversible dans $\mathbb{C}$).

(3) La multiplicativite du module provient de la multiplicativite de la conjugaison complexe (et le fait que $\mathbb{C}$ est commutatif.)

– On a de plus

$$z = 0 \iff x + iy = 0 \iff (x, y) = (0, 0) \iff x^2 + y^2 = 0 \iff |z| = 0.$$

(en effet comme $x^2, y^2 \geq 0$ on ne peut avoir $x^2 + y^2 = 0$ que si $x = y = 0$).

– Soit $z = x \in \mathbb{R}$ alors

$$|z| = |x + i.0| = (x^2 + 0^2)^{1/2} = (x^2)^{1/2} = \max(x, -x) = |x|_{\mathbb{R}}.$$

□

REMARQUE 8.3.2. On notera également la formule d'inversion suivante qui est un cas particulier de la formule d'inversion dans $K[\sqrt{d}]$ (8.2.2):

$$(8.3.2) \quad \forall z \in \mathbb{C}^\times, \quad z^{-1} = \frac{\bar{z}}{|z|^2} = \frac{x - iy}{x^2 + y^2} = \frac{x}{x^2 + y^2} - i \frac{y}{x^2 + y^2}.$$

Pour retrouver cette formule il suffit de se souvenir que

$$z \cdot \bar{z} = |z|^2 = (x^2 + y^2)$$

et si $|z|^2 = x^2 + y^2 \neq 0$ on a

$$z \cdot \frac{\bar{z}}{|z|^2} = 1.$$

8.3.1. Nombres complexes de module 1; décomposition polaire. Considérons le module mais restreint au groupe multiplicatif $\mathbb{C}^\times = \mathbb{C} - \{0\}$:

$$|\bullet| : \begin{array}{ccc} \mathbb{C}^\times & \mapsto & \mathbb{R}_{>0} \\ z & \mapsto & |z| = (x^2 + y^2)^{1/2}. \end{array}$$

Comme le module $|\bullet|$ est multiplicatif, sa restriction à $\mathbb{C}^\times$ est un morphisme de groupe (multiplicatif) à valeurs dans $\mathbb{R}_{>0}$; ce morphisme est surjectif (car pour $x \in \mathbb{R}_{>0}$, $|x| = x$) et son noyau est

$$\ker |\bullet| = \mathbb{C}^{(1)} = \{z \in \mathbb{C}, |z| = 1\},$$

l'ensemble des nombres complexes de module 1.

En particulier $\mathbb{C}^{(1)}$ est un sous-groupe de $\mathbb{C}^\times$ (pour la multiplication).

PROPOSITION 8.3. On a un isomorphisme de groupes

$$\text{pol} : \mathbb{C}^\times \simeq \mathbb{R}_{>0} \times \mathbb{C}^{(1)}$$

donnée par

$$z \in \mathbb{C}^\times \mapsto \text{pol}(z) = (|z|, z/|z|)$$

Preuve: Soit $z \in \mathbb{C}^\times$. On a $|z| > 0$ et comme $||z|| = |z|$ ($|z|$ est un nombre réel positif de sorte que son module est égal à sa valeur absolue et donc à $|z|$), on a

$$|z/|z|| = |z|/|z| = |z|/|z| = 1.$$

Ainsi

$$\text{pol}(z) \in \mathbb{R}_{>0} \times \mathbb{C}^{(1)}.$$

De plus on a

$$|z \cdot z'| = |z| \cdot |z'| \text{ et } z \cdot z' / |z \cdot z'| = (z/|z|) \cdot (z'/|z'|).$$

Ce morphisme de groupe pol est injectif:

$$(|z|, z/|z|) = (1, 1) \implies |z| = 1 = z/|z| \implies z = 1.$$

Il est également surjectif : pour tout $\rho > 0$ et $z^{(1)} \in \mathbb{C}^{(1)}$, on a

$$\text{pol}(\rho \cdot z^{(1)}) = (|\rho \cdot z^{(1)}|, \rho \cdot z^{(1)} / |\rho \cdot z^{(1)}|) = (\rho, z^{(1)});$$

en effet

$$|\rho \cdot z^{(1)}| = |\rho| \cdot |z^{(1)}| = \rho \cdot 1 = \rho$$

car $\rho \in \mathbb{R}_{>0}$. □

DÉFINITION 8.4. Soit $z \in \mathbb{C}^\times$, $\text{pol}(z) = (|z|, z/|z|)$ s'appelle la décomposition polaire de z .

(1) Le premier terme $|z|$ est le module et se note aussi $\rho(z) = r(z) > 0$,

(2) le second terme $z/|z| \in \mathbb{C}^{(1)}$ est appelé argument complexe de z et on le note

$$z/|z| = e^{i\theta(z)}.$$

(3) Si on decompose l'argument complexe en partie reelle et imaginaire,

$$z/|z| = e^{i\theta(z)} = \operatorname{Re}(z/|z|) + i \operatorname{Im}(z/|z|) = c(z) + s(z).i$$

on a donc

$$c(z)^2 + s(z)^2 = 1$$

- le reel $c(z) \in [-1, 1]$ s'appelle le cosinus de z ,
- le nombre $s(z) \in [-1, 1]$ s'appelle le sinus de z .

On a donc

$$z = x + iy = \rho(z).e^{i\theta(z)} = \rho(z)(c(z) + is(z)), \quad x = \rho(z)c(z), \quad y = \rho(z)s(z).$$

REMARQUE 8.3.3. Compte tenu des definitions, on a

$$\begin{aligned} \rho(z) &= |z| = (x^2 + y^2)^{1/2}, \\ c(z) &= \frac{x}{(x^2 + y^2)^{1/2}}, \quad s(z) = \frac{y}{(x^2 + y^2)^{1/2}} \end{aligned}$$

8.3.2. Formules de trigonometrie. On retrouve les formules habituelles de trigonometrie:

8.3.2.1. *Formules de produit.* Pour $z, z' \in \mathbb{C}^\times$

$$\begin{aligned} (8.3.3) \quad \rho(z.z') &= |z.z'| = |z|.|z'| = \rho(z).\rho(z'), \quad e^{i\theta(z.z')} = e^{i\theta(z)}.e^{i\theta(z')} \\ c(z.z') &= c(z).c(z') - s(z).s(z'), \quad s(z.z') = s(z).c(z') + s(z').c(z). \end{aligned}$$

Preuve: Les premieres identites resultent du fait que $\operatorname{pol}(\bullet)$ est un morphisme de groupes. Ecrivant

$$\begin{aligned} e^{i\theta(z.z')} &= c(z.z') + is(z.z') = \\ e^{i\theta(z)}.e^{i\theta(z')} &= (c(z) + is(z)).(c(z') + is(z')) \end{aligned}$$

on obtient en developpant (suivant la regle de produit des complexes)

$$\begin{aligned} c(z.z') + is(z.z') &= c(z)c(z') + is(z)c(z') + ic(z)s(z') + i^2s(z)s(z') \\ &= c(z)c(z') - s(z)s(z') + i(s(z)c(z') + c(z)s(z')). \end{aligned}$$

□

8.3.2.2. *Formule d'inversion.* Pour $z \in \mathbb{C}^\times$, on a

$$\begin{aligned} \rho(z^{-1}) &= |z^{-1}| = \rho(z)^{-1} = |z|^{-1} \\ e^{i\theta(z^{-1})} &= c(z^{-1}) + is(z^{-1}) = (e^{i\theta(z)})^{-1} = \overline{e^{i\theta(z)}} = c(z) - is(z). \end{aligned}$$

En particulier on a

$$c(z) = c(z^{-1}), \quad s(z) = -s(z^{-1}).$$

Preuve: Cela resulte a nouveau du fait que $\operatorname{pol}(\bullet)$ est un morphisme de groupes. De plus, on a vu que (8.3.2)

$$(e^{i\theta(z)})^{-1} = \frac{\overline{e^{i\theta(z)}}}{|e^{i\theta(z)}|^2} = \overline{e^{i\theta(z)}} = c(z) - is(z)$$

car $|e^{i\theta(z)}| = 1$.

□

8.3.2.3. *Formule de l'angle double.* On a

$$|z^2| = |z|^2, \quad c(z^2) = c(z)^2 - s(z)^2, \quad s(z^2) = 2s(z)c(z).$$

Preuve: Appliquer la formule du produit a $z' = z$.

□

Plus generalement on a les

8.3.2.4. *Formules de de Moivre.* Pour tout entier $n \geq 0$, on a³

$$(8.3.4) \quad \begin{aligned} |z^n| &= |z|^n, \quad e^{i\theta(z^n)} = (e^{i\theta(z)})^n \\ c(z^n) &= \sum_{0 \leq k \leq n/2} C_n^{2k} (-1)^k c(z)^{n-2k} s(z)^{2k}, \\ s(z^n) &= \sum_{0 \leq k \leq \frac{n-1}{2}} C_n^{2k+1} (-1)^k c(z)^{n-2k-1} s(z)^{2k+1}. \end{aligned}$$

Preuve: Les premieres identites resultent a nouveau du fait que $\text{pol}(\bullet)$ est un morphisme de groupes.

Pour les deux autres on ecrit

$$e^{i\theta(z^n)} = c(z^n) + is(z^n) = (e^{i\theta(z)})^n = (c(z) + is(z))^n.$$

Par la formule du binome de Newton cela vaut

$$\sum_{0 \leq k \leq n} C_n^k c(z)^{n-k} i^k s(z)^k.$$

On a

$$i^k = \begin{cases} (-1)^{k/2} & k \text{ pair} \\ (-1)^{(k-1)/2} i & k \text{ impair} \end{cases}$$

et on decompose la somme precedente suivant ces deux possibilites: la somme precedente s'ecrit

$$c(z^n) + is(z^n) = \sum_{\substack{0 \leq k \leq n \\ n \equiv 0 \pmod{2}}} C_n^k c(z)^{n-k} (-1)^{k/2} s(z)^k + \sum_{\substack{0 \leq k \leq n \\ n \equiv 1 \pmod{2}}} C_n^k c(z)^{n-k} i \cdot (-1)^{\frac{k-1}{2}} s(z)^k.$$

On met i en facteur dans le second terme et on identifie les parties reelles et imaginaires des complexes de part et d'autre de cette identite: remplaçant k par $2k \leq n$ dans la premiere somme et k par $2k+1 \leq n$ dans la seconde, on obtient les identites annoncees. $\square$

EXEMPLE 8.3.1. Par exemple pour $n = 2$, on obtient

$$c(z^2) = c(z)^2 - s(z)^2, \quad s(z^2) = 2c(z)s(z).$$

Pour $k = 3$, on obtient

$$c(z^3) = c(z)^3 - 3c(z)s(z)^2, \quad s(z^3) = 3c(z)^2s(z) - s(z)^3.$$

Pour $n = 4$, on obtient

$$c(z^4) = c(z)^4 - 6c(z)^2s(z)^2 + s(z)^4, \quad s(z^4) = 4c(z)^3s(z) - 4c(z)s(z)^3.$$

8.3.3. Argument (reel) d'un nombre complexe. Dans ce cours qui est de nature algebrigue, on a resiste jusqu'a present a parler d'*argument d'un nombre complexe*. La raison est la definition precise necessite des notions elaborees d'analyse (notamment la definition de l'exponentielle sur les complexes). On peut parler d'*argument reel* d'un nombre complexe une fois qu'on a demontrer (ou admis) le resultat suivant:

THÉORÈME 8.2 (Existence de l'exponentielle complexe). *Il existe un unique morphisme de groupe*

$$e^{i\bullet} : (\mathbb{R}, +) \mapsto (\mathbb{C}^{\times}, \times) \\ \theta \mapsto \exp(i\theta)$$

qui est derivable (comme fonction de $\mathbb{R}$ a valeurs dans $\mathbb{C} \simeq \mathbb{R}^2$) et qui verifie

$$e^{i\bullet}'(0) = i.$$

Ce morphisme est surjectif et son noyau est de la forme

$$\ker e^{i\bullet} = 2\pi\mathbb{Z}$$

³d'apres Abraham de Moivre (1667-1754)

ou π est un nombre reel dont le developpement decimal commence par $\pi = 3.14159 \dots$.

REMARQUE 8.3.4. On dit qu'une fonction a valeurs complexes

$$f : \theta \in \mathbb{R} \mapsto f(\theta) \in \mathbb{C}$$

est derivable sur $\mathbb{R}$ si les fonctions associees "partie reel" et "partie imaginaire" sont derivables: on ecrit

$$f(\theta) = \operatorname{Re} f(\theta) + i \operatorname{Im} f(\theta)$$

et on demande que les deux fonctions

$$\operatorname{Re} f, \operatorname{Im} f : \theta \in \mathbb{R} \mapsto \operatorname{Re} f(\theta), \operatorname{Im} f(\theta) \in \mathbb{R}$$

soient derivables sur $\mathbb{R}$.

REMARQUE 8.3.5. On peut montrer que si un morphisme de groupes

$$\varphi : \mathbb{R} \mapsto \mathbb{C}^\times$$

est continu (ie. ses parties reelles et imaginaires sont continuees) alors il est automatiquement derivable et meme infiniment derivable.

Admettant ce Theoreme, on obtient par surjectivite que pour tout $z \in \mathbb{C}^{(1)}$ il existe $\theta \in \mathbb{R}$ tel que

$$z = e^{i\theta}.$$

D'autre part, comme $e^{i\bullet}$ est un morphisme de groupes, l'ensemble des θ' verifiant $z = e^{i\theta'}$ (l'ensemble des antecedents de z , $(e^{i\bullet})^{-1}(\{z\})$) est egale a la classe de θ modulo 2π (cf. Exercice 2.2)

$$(e^{i\bullet})^{-1}(\{z\}) = \theta + \ker(e^{i\bullet}) = \theta + 2\pi\mathbb{Z} = \{\theta + 2\pi.k, k \in \mathbb{Z}\}.$$

On obtient alors un isomorphisme de groupe (qu'on notera encore $e^{i\bullet}$)

$$e^{i\bullet} : \begin{array}{ccc} \mathbb{R}/2\pi\mathbb{Z} & \simeq & \mathbb{C}^{(1)} \\ \theta + 2\pi\mathbb{Z} & \mapsto & z = e^{i\theta}. \end{array}$$

La reciproque de cette bijection s'appelle *l'argument (reel)*:

DÉFINITION 8.5. Soit z un nombre complexe de module 1 L'argument reel (encore appelle "angle") de z ,

$$\arg(z) := \theta \pmod{2\pi} = \theta + 2\pi\mathbb{Z} \in \mathbb{R}/2\pi\mathbb{Z}$$

est l'unique classe $\theta \pmod{2\pi} \in \mathbb{R}/2\pi\mathbb{Z}$ telle que $e^{i\theta} = z$.

Plus generalement, pour $z \in \mathbb{C}^\times$, on defini son argument par

$$\arg(z) := \arg(z/|z|) \in \mathbb{R}/2\pi\mathbb{Z}.$$

Notons que l'application

$$\arg : \mathbb{C}^\times \mapsto \mathbb{R}/2\pi\mathbb{Z}$$

est un morphisme de groupes: $\forall z, z' \in \mathbb{C}^\times$ on a

$$\arg(1) = 0, \arg(z.z') = \arg(z) + \arg(z'), \arg(1/z) = -\arg(z).$$

et la decomposition polaire se reecrit sous la forme de l'isomorphisme

$$\operatorname{pol} : \begin{array}{ccc} \mathbb{C}^\times & \simeq & \mathbb{R}_{>0} \times \mathbb{R}/2\pi\mathbb{Z} \\ z & \mapsto & (|z|, \arg(z)) \end{array}.$$

DÉFINITION 8.6. Soit $\theta \in \mathbb{R}$, le cosinus et le sinus de θ sont defini par

$$\cos(\theta) = \operatorname{Re}(e^{i\theta}), \sin(\theta) = \operatorname{Im}(e^{i\theta}).$$

On a donc

$$e^{i\theta} = \cos(\theta) + i \sin(\theta).$$

En particulier on a

$$1 = e^{i0} = \cos(0) + i \sin(0)$$

et donc

$$\cos(0) = 1, \sin(0) = 0.$$

8.3.4. Formules de trigonometrie classiques. On "retrouve" les formules de trigonometrie sous leur forme usuelle:

8.3.4.1. *Formule des sommes.* On a

$$\cos(\theta + \theta') = \operatorname{Re}(e^{i(\theta+\theta')}) = \operatorname{Re}(e^{i\theta} \cdot e^{i\theta'}) = \cos(\theta) \cos(\theta') - \sin(\theta) \sin(\theta')$$

et

$$\sin(\theta + \theta') = \operatorname{Im}(e^{i(\theta+\theta')}) = \operatorname{Im}(e^{i\theta} \cdot e^{i\theta'}) = \sin(\theta) \cos(\theta') + \cos(\theta) \sin(\theta').$$

Preuve: On a

$$e^{i\theta+\theta'} = \cos(\theta + \theta') + i \sin(\theta + \theta') = e^{i\theta} \cdot e^{i\theta'} = (\cos(\theta) + i \sin(\theta)) \cdot (\cos(\theta') + i \sin(\theta'))$$

et on obtient le result en developpant et en isolant les parties reeles et imaginaires. $\square$

8.3.4.2. *Formule de l'angle oppose.* On a

$$\cos(-\theta) = \cos(\theta), \sin(-\theta) = -\sin(\theta).$$

Preuve: En effet comme on a un morphisme de groupes

$$e^{-i\theta} = \cos(-\theta) + i \sin(-\theta) = 1/e^{i\theta} = \overline{e^{i\theta}} = \cos(\theta) - i \sin(\theta).$$

$\square$

8.3.4.3. *Formule de l'angle double.* En prenant $\theta' = \theta$ on obtient

$$\cos(2\theta) = \cos(\theta)^2 - \sin(\theta)^2, \sin(2\theta) = 2 \sin(\theta) \cos(\theta)$$

et plus generalement

8.3.4.4. *Formules de de Moivre.*

$$e^{in\theta} = \cos(n\theta) + i \sin(n\theta) = (e^{i\theta})^n = (\cos(\theta) + i \sin(\theta))^n$$

et en developpant par le binome de Newton et identifiant parties reelles et imaginaires, on obtient

$$\cos(n\theta) = \sum_{0 \leq k \leq n/2} C_n^{2k} (-1)^k \cos(\theta)^{n-2k} \sin(\theta)^{2k}.$$

$$\sin(n\theta) = \sum_{0 \leq k \leq (n-1)/2} C_n^{2k+1} (-1)^k \cos(\theta)^{n-2k-1} \sin(\theta)^{2k+1}.$$

8.4. Le plan complexe

Comme $\mathbb{C}$ est un $\mathbb{R}$ -ev de dimension 2, on peut identifier $\mathbb{C}$ a $\mathbb{R}^2$ en choisissant une base. Ainsi si on prend pour base $\{\operatorname{Id}, I\}$ l'isomorphisme est donne par les parties reele et imaginaire:

$$\begin{aligned} (\operatorname{Re}, \operatorname{Im}) : \quad \mathbb{C} &\mapsto \mathbb{R}^2 \\ z = x \cdot \operatorname{Id} + y \cdot I &\mapsto (x, y). \end{aligned}$$

On parle alors du plan complexe et on represente un nombre complexe par un point dans le plan reel $\mathbb{R}^2$. Le groupe des nombres complexes de module 1 est alors identifie avec le cercle unite

$$S^1 = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 1\}.$$

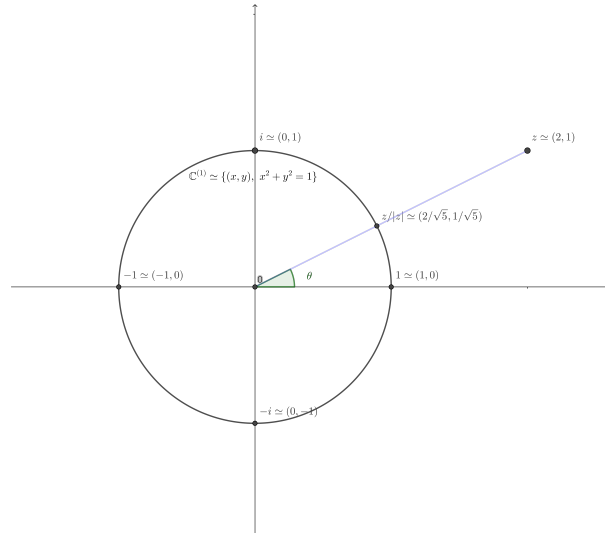


FIGURE 1. Le plan complexe et le cercle unite.

8.4.1. Le plan euclidien. L'espace $\mathbb{R}^2$ est muni d'une distance appelée *distance euclidienne*:

$$d_2((x, y), (x', y')) = \|(x - x', y - y')\|_2 := ((x - x')^2 + (y - y')^2)^{1/2}.$$

Rappelons qu'une distance sur un ensemble X est une application

$$d : \begin{array}{ccc} X \times X & \mapsto & \mathbb{R}_{\geq 0} \\ (v, w) & \mapsto & d(v, w) \end{array}$$

verifiant

- (1) Separation: $d(v, w) = 0 \iff v = w$.
- (2) Symetrie: $d(v, w) = d(w, v)$.
- (3) Inegalite du triangle: $d(u, w) \leq d(u, v) + d(v, w)$.

DÉFINITION 8.7. Une *isometrie (euclidienne)* de $\mathbb{R}^2$ est une application $\varphi : \mathbb{R}^2 \mapsto \mathbb{R}^2$ preservant la distance euclidienne:

$$d_2(\varphi(v), \varphi(w)) = d_2(v, w).$$

EXEMPLE 8.4.1. La translation de vecteur $v_0 \in \mathbb{R}^2$:

$$t_{v_0} : v \in \mathbb{R}^2 \mapsto v + v_0.$$

THÉORÈME 8.3. Une isometrie est bijective et sa reciproque est encore une isometrie. L'ensemble des isometrie $\text{Isom}(\mathbb{R}^2) \subset \text{Bij}(\mathbb{R}^2)$ est un sous-groupe du groupe des bijections de $\mathbb{R}^2$.

Grace a l'isomorphisme de $\mathbb{R}$ -ev $\mathbb{C} \simeq \mathbb{R}^2$ ci-dessus on peut realiser les isometries en terme de transformations simples sur le corps des nombres complexes (on admettra le resultat suivante)

THÉORÈME 8.4. Quand on identifie $(x, y) \in \mathbb{R}^2$ avec le nombre complexe $z = x + iy$ toute isometrie de $\mathbb{R}^2$ est de la forme suivante

- Rotation: il existe $\alpha \in \mathbb{C}^{(1)}$ et $z_0 \in \mathbb{C}$ tels que

$$r_{\alpha, z_0} : z \mapsto \alpha \cdot z + z_0.$$

- Symetrie: il existe $\alpha \in \mathbb{C}^{(1)}$ et $z_0 \in \mathbb{C}$ tels que

$$s_{\alpha, z_0} : z \mapsto \alpha \cdot \bar{z} + z_0.$$

On a la classification suivante plus fine des rotations et des translations. Rappelons que si $\varphi : X \mapsto X$ est une application, un point fixe de φ est un élément $x \in X$ tel que

$$\varphi(x) = x.$$

THÉORÈME 8.5. *La rotation r_{α, z_0} peut être de deux types*

- Si $\alpha = 1$, alors $r_{1, z_0} : z \mapsto z + z_0$ est une translation (par z_0). On dit également que c'est une rotation triviale ou d'angle nul. Si $z_0 = 0$ alors c'est l'identité et tous les points de $\mathbb{C}$ sont fixes. Si $z_0 \neq 0$ alors la translation n'a aucun point fixe.
- Si $\alpha \neq 1$, alors r_{α, z_0} possède un unique point fixe: un point z_f vérifiant

$$r_{\alpha, z_0}(z_f) = z_f$$

donne par

$$z_f = \frac{z_0}{(1 - \alpha)}.$$

Si $\theta \pmod{2\pi} = \arg(\alpha)$ est l'argument de α on dit que r_{α, z_0} est une rotation d'angle θ .

La symétrie s_{α, z_0} peut être de deux types

- L'ensemble des points fixes de s_{α, z_0} est une droite et la symétrie est appelée symétrie orthogonale par rapport à cette droite de points fixes.
- L'ensemble des points fixes de s_{α, z_0} est vide; il existe alors une unique droite de $\mathbb{C}$ telle que s_{α, z_0} est la composée d'une symétrie orthogonale par rapport à cette droite et d'une translation par un complexe parallèle à cette droite. On dit alors que s_{α, z_0} est une symétrie glissée (par rapport à cette droite).

EXEMPLE 8.4.2. Par exemple

$$z \mapsto i.z$$

est la rotation d'angle $\pi/2$ (dans le sens inverse des aiguilles d'une montre) et de centre l'origine et

$$z \mapsto \bar{z}$$

est la symétrie orthogonale par rapport à l'axe des x . Par contre

$$z \mapsto \bar{z} + 1$$

est une symétrie glissée par rapport à l'axe des x .

L'intérêt de représenter les isométries sous forme de transformations sur les nombres complexes c'est qu'il est plus facile de calculer leur composée ou leurs espaces de points fixes: par exemple s_{α, z_0} est la composée de la symétrie $z \mapsto \bar{z}$, de la rotation $z' \mapsto \alpha z'$ et de la translation $z'' \mapsto z'' + z_0$.

8.5. Equations polynomiales complexes

Comme on l'a expliqué, le corps des nombres complexes $\mathbb{C}$ a été introduit (pas sous forme de matrices) dans la renaissance italienne dans l'étude des équations polynomiales: l'étude des solutions z des équations de la forme

$$(8.5.1) \quad P(z) = a_d.z^d + a_{d-1}.z^{d-1} + \cdots + a_1.z + a_0 = 0,$$

avec $a_0, \dots, a_d \in \mathbb{R}$ des nombres réels⁴.

DÉFINITION 8.8. Soit

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

un polynôme à coefficient dans $\mathbb{C}$. L'ensemble des racines de P dans $\mathbb{C}$, $\text{Rac}_P(\mathbb{C})$ est l'ensemble des solutions dans $\mathbb{C}$ de l'équation $P(z) = 0$:

$$\text{Rac}_P(\mathbb{C}) = \{z \in \mathbb{C}, P(z) = 0\}.$$

⁴en fait c'était plutôt les nombres rationnels car le corps des réels n'existait pas encore mais on s'autorisait à extraire des racines n -ièmes de nombres rationnels positifs ou nuls

On rappelle (cf. Thm A.6 dans le chapitre sur les polynomes) que

$$|\text{Rac}_P(\mathbb{C})| \leq \deg P \leq d.$$

En particulier pour $d = 2$ (les equations quadratiques) on obtient

$$(8.5.2) \quad az^2 + bz + c = 0, \quad a, b, c \in \mathbb{R}, \quad a \neq 0$$

Rappelons d'abord la methode permettant de trouver la forme generale des solutions qui consiste a "completer le carre": on a

$$az^2 + bz + c = a(z^2 + \frac{b}{a}z + \frac{c}{a}) = a(z^2 + 2\frac{b}{2a}z + \frac{c}{a})$$

on reconnait dans $z^2 + 2\frac{b}{2a}z$ le debut d'un carre:

$$z^2 + 2\frac{b}{2a}z = z^2 + 2\frac{b}{2a}z + (\frac{b}{2a})^2 - (\frac{b}{2a})^2 = (z + \frac{b}{2a})^2 - (\frac{b}{2a})^2$$

et l'equation devient

$$a((z + \frac{b}{2a})^2 - (\frac{b}{2a})^2 + \frac{c}{a}) = 0 \iff Z^2 - (\frac{b}{2a})^2 + \frac{c}{a} \iff Z^2 = \frac{\Delta}{4a^2}$$

en posant $Z = z + \frac{b}{2a}$. Si $\Delta \geq 0$ on obtient comme solutions de cette equation

$$Z_{\pm} = \pm \frac{\sqrt{\Delta}}{2a}$$

dont on deduit les formules bien connues

$$z_{\pm} = \frac{-b \pm \sqrt{\Delta}}{2a}.$$

Si $\Delta < 0$ les equations precedentes n'ont pas de solutions dans $\mathbb{R}$; en particulier c'est le cas de l'equation

$$z^2 + 1 = 0$$

dont le discriminant vaut $-4 < 0$. On⁵ a alors introduit "formellement" une solution i verifiant

$$i^2 = -1$$

qu'on a appelle nombre "imaginaire" et on a ainsi obtenu le corps abstrait des nombres complexes $\mathbb{C}$. On a alors trouve dans $\mathbb{C}$ des solutions de toutes les equations quadratiques a coefficients reels : elles sont donnees par la formule usuelle

$$z_{\pm} = \frac{-b \pm \sqrt{\Delta}}{2a}$$

ou $\sqrt{\Delta}$ est l'une des racines carrees de Δ si $\Delta \geq 0$ et si $\Delta < 0$ on prend

$$\sqrt{\Delta} := \sqrt{|\Delta|}.i$$

8.5.1. Equations quadratiques a coefficients complexes. Considerons maintenant la meme equation

$$(8.5.3) \quad az^2 + bz + c = 0$$

mais avec $a, b, c \in \mathbb{C}$. Les meme manipulations algebriques nous disent que les solutions de cette equation devraient etre de la forme

$$z_{\pm} = \frac{-b \pm \sqrt{\Delta}}{2a}, \quad \Delta = b^2 - 4ac \in \mathbb{C}.$$

Ce qui nous reduit a trouver les solutions de l'equation quadratique "monomiale"

$$Z^2 = \Delta$$

⁵Bombelli le premier

pour $\Delta \in \mathbb{C}$. Pour cela on écrit $\Delta = A + IB$ et $Z = X + iY$ et on a donc

$$Z^2 = X^2 - Y^2 + 2XY.i = A + iB$$

ce qui nous amène à un système de deux équations polynomiales à coefficients dans $\mathbb{R}$ en deux inconnues X, Y dans $\mathbb{R}$:

$$X^2 - Y^2 = A, \quad 2XY = B.$$

On peut supposer que $B \neq 0$ car sinon on a $\Delta = A \in \mathbb{R}$ et on sait résoudre l'équation (même si $A < 0$). On a donc $X, Y \neq 0$ et on peut écrire $Y = B/2X$ et substituer:

$$X^2 - B^2/(4X^2) = A \iff 4X^4 - 4AX^2 - B^2 = 0, \quad X \neq 0$$

Posant $U = 2X^2$ on doit résoudre l'équation quadratique

$$U^2 - 2AU - B^2 = 0$$

dont le discriminant vaut

$$\Delta' = 4(A^2 + B^2) > 0.$$

On trouve donc deux racines réelles

$$U_{\pm} = A \pm \sqrt{A^2 + B^2}.$$

Comme $\sqrt{A^2 + B^2} > A$, l'une de ses solutions est positive et l'autre négative mais comme $U = X^2$ et que $X \in \mathbb{R}$ on doit avoir $U \geq 0$ et on prend

$$U_+ = A + \sqrt{A^2 + B^2}$$

et on prend

$$X_{\pm} = \pm \sqrt{U_+}.$$

On trouve alors $Y_{\pm} = \pm B/(2\sqrt{U_+})$ et on obtient deux solutions

$$Z_{\pm} = \pm(\sqrt{U_+} + iB/(2\sqrt{U_+})).$$

8.5.2. Equations monomiales. Les équations monomiales sont celles de la forme

$$X^d - w = 0$$

pour $d \geq 1$ et $w \in \mathbb{C}$. Si $w = 0$ alors $z = 0$ est la seule racine.

Si $w \neq 0$ alors l'existence de l'exponentielle complexe garantit l'existence de n solutions distinctes: soit $z \in \text{Rac}_{X^d - w}(\mathbb{C})$ alors on a

$$|z|^d = |w|$$

et donc

$$|z| = |w|^{1/d}.$$

Pour l'argument on a

$$d \arg(z) = \arg(w) \pmod{2\pi}.$$

On réécrit cela sous la forme

$$d \arg(z) = \arg(w) + 2\pi\mathbb{Z} \iff \arg(z) = \frac{\arg(w)}{d} + 2\pi\frac{1}{d}\mathbb{Z}$$

Ainsi $\arg(z)$ prend d valeurs distinctes modulo 2π :

$$\arg(z) = \frac{\arg(w)}{d} + 2\pi\frac{k}{d}, \quad 0 \leq k \leq d-1$$

et

$$\text{Rac}_{X^d - w}(\mathbb{C}) = \{|w|^{1/d} e^{i\frac{\arg(w)}{d} + i2\pi\frac{k}{d}}, \quad 0 \leq k \leq d-1\}$$

notons que

$$e^{i\frac{\arg(w)}{d} + i2\pi\frac{k}{d}} = e^{i\frac{\arg(w)}{d}} \omega_d^k, \quad \text{avec } \omega_d := e^{i\frac{2\pi}{d}}.$$

Ainsi on a

$$(8.5.4) \quad \text{Rac}_{X^d - w}(\mathbb{C}) = \{|w|^{1/d} e^{i\frac{\arg(w)}{d}} \omega_d^k, \quad 0 \leq k \leq d-1\}$$

8.5.3. Racines de l'unité. En particulier si $w = 1$ on obtient

DÉFINITION 8.9. Pour $d \geq 1$ l'ensemble des racines de l'équation

$$z^d = 1,$$

$$\mu_d := \text{Rac}_{X^d-1}(\mathbb{C}) = \{\omega_d^k, 0 \leq k \leq d-1\}$$

est appelée ensemble des racines d -ièmes de l'unité

On a donc

$$\text{Rac}_{X^d-w}(\mathbb{C}) = |w|^{1/d} e^{i \frac{\arg(w)}{d}} \cdot \mu_d$$

Notons que μ_d est un sous-groupe du groupe multiplicatif $\mathbb{C}^\times$: en effet c'est un noyau

$$\mu_d = \ker(\bullet^d : \begin{matrix} \mathbb{C}^\times & \mapsto & \mathbb{C}^\times \\ z & \mapsto & z^d \end{matrix}).$$

REMARQUE 8.5.1. Pour une équation monomiale générale, l'ensemble des solutions (8.5.4) s'écrit donc

$$\text{Rac}_{X^d-w}(\mathbb{C}) = z_0 \cdot \mu_d, \quad z_0 = e^{i \frac{\arg(w)}{d}}.$$

C'est un cas particulier de résolution d'équations dans les groupes, cf. Exo 2.2 (pour le groupe $(\mathbb{C}^\times, \times)$).

Notons également que

$$\mu_d = \omega_d^{\mathbb{Z}};$$

ce groupe est donc cyclique de générateur $\omega_d = e^{i \frac{2\pi}{d}}$. En fait c'est un cas particulier d'un résultat général purement algébrique:

THÉORÈME 8.6. Soit K un corps et $\mu \subset K^\times$ un sous-groupe fini du groupe multiplicatif $(K^\times, \times)$. Alors μ est cyclique et si on note $d = |\mu|$ son cardinal alors

$$\mu = \mu_d(K) = \text{Rac}_{X^d-1}(K) = \{\omega \in K, \omega^d = 1\}$$

est le groupe des racines d -ièmes de l'unité de K .

On rappelle que de part la théorie des groupes cycliques le groupe $\mu_d(K)$ possède

$$\varphi(d) = |\{0 \leq k \leq d-1, (k, d) = 1\}|$$

générateurs données pour tout générateur ω_0 de μ_d par

$$\mu_d^* = \{\omega_0^k, 0 \leq k \leq d-1, (k, d) = 1\}.$$

Ce sont également les éléments du groupe $\mu_d(K)$ d'ordre d exactement:

$$\mu_d^* = \{\omega \in K, \omega^d = 1, \forall d' | d, \omega^{d'} \neq 1\}.$$

On appelle μ_d^* des racines primitives d -ièmes de l'unité de K .

8.5.4. Racines complexes de l'unité ayant des arguments particuliers. Il y a extrêmement peu de nombres complexes de module 1 pour lesquels on dispose d'une formule simple pour leur argument réel et il y a de bonnes raisons à cela. Pour $d \geq 1$ un entier on pose

$$\omega_d = e^{i2\pi/d}.$$

On va calculer quelques ω_d .

Pour cela on remarque que comme $\ker(e^{i\bullet}) = 2\pi\mathbb{Z}$ et que $e^{i\bullet}$ est surjective sur $\mathbb{C}^{(1)}$, $e^{i\bullet}$ induit une bijection

$$e^{i\bullet} : [0, 2\pi[\simeq \mathbb{C}^{(1)}.$$

On peut commencer:

8.5.4.1. $d = 1$. On a

$$\omega_1 = e^{i0} = 1$$

car un morphisme de groupe envoie l'élément neutre sur l'élément neutre.

8.5.4.2. $d = 2$. On a (formule d'Euler)

$$\omega_2 = e^{i\pi} = -1.$$

En effet on a

$$(\omega_2)^2 = e^{i2\pi} = 1$$

donc ω_2 est une racine carree de 1 et donc vaut ± 1 . Comme on sait que $e^{i0} = 1$ et que $e^{i\pi} \neq e^{i0}$ c'est que $\omega_2 = -1$.

8.5.4.3. $d = 4$. On a

$$\omega_4 = e^{i\pi/2} = i.$$

Preuve: Exercice. □

8.5.4.4. $d = 8$. On a

$$\omega_8 = \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}.$$

Preuve: Exercice. □

8.5.4.5. $d = 3$. On a

$$\omega_3 = \frac{-1 + i\sqrt{3}}{2}.$$

Preuve: Exercice. □

8.5.4.6. $d = 5$. On a

$$\omega_5 = \cos(2\pi/5) + i\sin(2\pi/5)$$

avec

$$\cos(2\pi/5) = -\frac{1 + \sqrt{5}}{4}, \quad \sin(2\pi/5) = \sqrt{1 - \left(\frac{1 + \sqrt{5}}{4}\right)^2}.$$

Preuve: Exercice. □

8.5.4.7. *Formule de l'angle moitie.* Le calcul de $\omega_2, \omega_4, \omega_8$ proviennent d'un principe general: si on connait $\omega_d = e^{i2\pi/d}$ alors on saura exprimer simplement $\omega_{2d} = e^{i2\pi/2d}$ des parties reelles et imaginaires de ω_d . En effet

$$\omega_{2d}^2 = \omega_d$$

et ω_{2d} est solution de l'equation

$$X^2 = \omega_d$$

que l'on sait resoudre sur les complexes. On obtient ainsi

$$\omega_6 = \frac{\sqrt{3} + i}{2}.$$

On voit que les parties reelles et imaginaires de tous ces nombres complexes s'expriment par extractions successives de racines carrees. Une condition geometrique equivalente de cette propriete est la suivante:

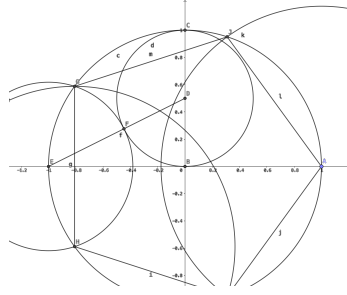
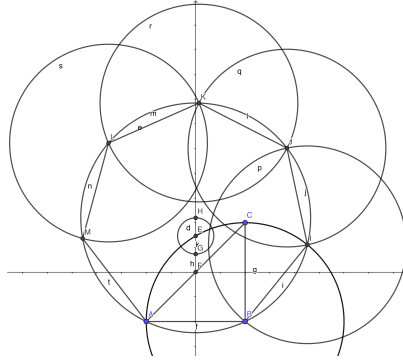
DÉFINITION 8.10 (Constructibilite a la regle et au compas). *Soit $P_0 = (0, 0)$ et $P_1 = (1, 0)$. Un point P du plan est constructible a la regle et au compas a partir d'un ensemble fini de points $\mathcal{P}_n = \{P_0, P_1, \dots, P_n\}$ contenant P_0 et P_1 si P est obtenu soit*

- *comme l'intersection de deux droites passant par des points distincts de $\{P_0, P_1, \dots, P_n\}$*
- *de l'intersection d'une droite passant par deux points distincts de $\{P_0, \dots, P_n\}$ et d'un cercle dont le centre est contenu dans $\{P_0, P_1, \dots, P_n\}$ et le rayon est egal a la distance $|P_i P_j|$ pour $0 \leq i, j \leq n$.*
- *de l'intersection de deux cercles centres en des elements de $\mathcal{P}_n$ et de rayons $|P_i P_j|$ et $|P_k P_l|$.*

Un point P est constructible a la regle et au compas si il existe un ensemble de points

$$\{P_0, P_1, \dots, P_n, P_{n+1}\}$$

avec $P_{n+1} = P$ tel que pour tout $i \geq 2$, P_i soit constructible a la regle et au compas a partir de $\{P_0, P_1, \dots, P_{i-1}\}$.

FIGURE 2. Construction a la regle et au compas d'un pentagone régulier (ω_5).FIGURE 3. Construction (fausse !) a la regle et au compas d'un heptagone régulier (ω_7).

En fait il n'y a pas beaucoup d'autre cas de racine de l'unité constructibles:

THÉORÈME 8.7 (Gauss-Wantzel). *On peut exprimer les parties réelles et imaginaires du nombre complexe $\omega_d = e^{i2\pi/d}$ par extraction successive de racines carrées (ou de manière équivalente, est constructible a la regle et au compas) si et seulement si*

$$d = 2^k \text{ ou bien } d = 2^k \prod_i p_i$$

ou $\prod_i p_i$ est un produit (non-vide) de nombres premiers tous distincts et "de Fermat": on dit qu'un nombre premier p_i est de Fermat si $p_i = F_{f_i} := 2^{2^{f_i}} + 1$ avec $f_i \geq 0$ un entier.

REMARQUE 8.5.2. Les nombres premiers $F_0 = 3, F_1 = 5, F_2 = 17$ sont de Fermat et Gauss est devenu célèbre quand a 19 ans il a montré que la condition était suffisante et a exprimé ω_{17} sous cette forme; un peu plus tard Wantzel a montré qu'elle était nécessaire. Les autres premiers de Fermat connus sont $F_3 = 257$ et $F_4 = 65537$; les entiers $F_5, \dots, F_{32}$ ne sont pas premiers et on ne sait pas si F_{33} ou les entiers de Fermat suivants sont premiers ou pas.

8.5.5. Equations de degré supérieur. On a également pu résoudre dans $\mathbb{C}$ de nombreuses autres équations polynomiales à coefficient réels. En particulier pour les équations de degré 2, 3 ou 4, on (les italiens) a pu obtenir des expressions algébriques explicites pour les solutions des équations polynomiales en fonction des coefficients du polynôme (formules de Cardan) ainsi que pour des polynômes de degré supérieur mais spéciaux cela en extrayant des racines carrées, cubiques ou quartiques ou d'ordre supérieur: on parle d'équation résolubles par radicaux.

Le résultat le plus général est dû à Gauss qui a démontré le

THÉORÈME (fondamental de l'algèbre). Soit $P(X) \in \mathbb{R}[X] = a_d \cdot z^d + a_{d-1} \cdot z^{d-1} + \cdots + a_1 \cdot z + a_0$ un polynôme réel non-constant alors l'équation (8.5.1) admet au moins une solution dans $\mathbb{C}$: il existe $z \in \mathbb{C}$ tel que $P(z) = 0$. En fait c'est également vrai si $P(X) \in \mathbb{C}[X]$ c'est à dire si l'équation polynomiale est à coefficient dans $\mathbb{C}$. On dit que $\mathbb{C}$ est algébriquement clos.

REMARQUE 8.5.3. Ce theoreme n'est pas constructif : il demontre l'existence de solutions mais ne donne pas d'expression des solutions en fonctions des coefficients de P (comme c'est le cas pour les equations quadratiques ou cubiques ou quartiques). Ce probleme a ete analyse en details par Abel et Galois. En particulier Abel a donne un polynome explicite

$$X^5 - X - 1$$

dont les racines ne peuvent s'exprimer par l'extractino de racines carrees, cubiques, quartique, quintiques (ou de tout ordre) de nombres rationnels (cette equation n'est pas resoluble par radicaux). Galois a ensuite donne une condition necessaire et suffisante (en terme d'un certain groupe associe au polynome) pour decider si l'equation est resoluble par radicaux ou pas. C'est l'objet de ce qu'on appelle la *Theorie de Galois*.

EXERCICE 8.1. Demonstrer la partie facile du Theoreme de Gauss: si tout polynome a coefficient reel admet une racine alors tout polynome a coefficient complexes admet une racine.

Pour cela considerer

$$P(X) = a_d \cdot z^d + a_{d-1} \cdot z^{d-1} + \cdots + a_1 \cdot z + a_0 \in \mathbb{C}[X]$$

et

$$\overline{P}(X) = \overline{a_d} \cdot z^d + \overline{a_{d-1}} \cdot z^{d-1} + \cdots + \overline{a_1} \cdot z + \overline{a_0}$$

et montrer que $Q(X) = P(X) \cdot \overline{P}(X) \in \mathbb{R}[X]$ et conclure.

On n'a pas encore les moyens de demontre ce resultat fondamental. On peut le faire soit

- (1) Avec de l'analyse reele classique (theoreme des valeurs intermediaires) et de la *Theorie de Galois*.
- (2) Ou bien avec de l'analyse complexe: soit

$$z \in \mathbb{C} \mapsto P(z) \in \mathbb{C}$$

un polynome non-constant qui ne s'annule pas sur $\mathbb{C}$ alors la fonction

$$z \mapsto 1/P(z)$$

est holomorphe sur $\mathbb{C}$ et bornee; cela implique necessairement qu'elle est constante et donc que $P(z)$ est constant.

CHAPITRE 9

Operations elementaires sur les matrices

*The first matrix I designed was quite naturally perfect.
It was a work of art. Flawless. Sublime.
A triumph only equaled by its monumental failure.*

9.1. Operation elementaires sur les lignes

Soit $M = (m_{ij}) \in M_{d' \times d}(K)$ une matrice. Pour simplifier les notations on écrira sa i -ieme ligne ($i \leq d'$)

$$L_i = L_i(M) = \text{Lig}_i(M) = (m_{ij})_{j \leq d}$$

DÉFINITION 9.1. *Les operations elementaires sur les lignes d'une matrice sont les applications suivantes de $M_{d' \times d}(K)$ vers $M_{d' \times d}(K)$: pour $i, j \in \{1, \dots, d'\}$ et $\lambda \in K^\times$ et $\mu \in K$*

(I) *Transposition: Echanger deux lignes $i \neq j \leq d'$ de M :*

$$L_i \longleftrightarrow L_j$$

(II) *Dilatation: Multiplier la i -eme ligne par un scalaire $\lambda \neq 0$:*

$$L_i \rightarrow \lambda.L_i.$$

(III) *Combinaison Lineaire: Additionner a la ligne i un multiple scalaire de la j -ieme ligne pour $i \neq j$: $\mu \in K$*

$$L_i \rightarrow L_i + \mu L_j$$

Ces transformations sont appelees transformations *elementaires*.

EXEMPLE 9.1.1. Considerons la matrice

$$(9.1.1) \quad M = \begin{pmatrix} 0 & 1 & 1 \\ 2 & 2 & 2 \\ 2 & 1 & 2 \end{pmatrix}.$$

On lui applique la transposition $L_1 \leftrightarrow L_2$ et on obtient

$$M_1 = \begin{pmatrix} 2 & 2 & 2 \\ 0 & 1 & 1 \\ 2 & 1 & 2 \end{pmatrix}.$$

On applique $L_1 \rightarrow (1/2).L_1$ et on obtient

$$M_2 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 2 & 1 & 2 \end{pmatrix}.$$

On applique $L_3 \rightarrow L_3 - 2.L_1$ et on obtient

$$M_3 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

On applique $L_3 \rightarrow L_3 + L_2$ et on obtient

$$M_4 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

On applique $L_1 \rightarrow L_1 - L_2$ et on obtient

$$M_5 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

On applique $L_2 \rightarrow L_2 - L_3$ et on obtient

$$M_6 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \text{Id}_3.$$

PROPOSITION 9.1. *Ces trois operations sont des applications lineaires bijectives*

$$(I), (II), (III) : M_{d' \times d}(K) \mapsto M_{d' \times d}(K).$$

Preuve: La linearite vient du fait que les applications

$$\text{Lig}_i(\bullet), \text{Lig}_j(\bullet) : M \in M_{d' \times d}(K) \mapsto M_i \in \text{Lig}_d(K)$$

sont lineaires et que l'application

$$(\text{Lig}_i + \mu \text{Lig}_j)(\bullet) : M \in M_{d' \times d}(K) \mapsto L_i + \mu.L_j \in \text{Lig}_d(K)$$

est lineaire. Elle sont bijectives car elle admettent des applications reciproques:

(I) Echanger les deux memes lignes $i, j \leq d'$ de M :

$$L_i \longleftrightarrow L_j$$

(II) Multiplier la i -eme ligne par le scalaire λ^{-1} :

$$L_i \rightarrow \lambda^{-1}.L_i.$$

(III) Soustraire a la ligne i un multiple scalaire de la j -ieme ligne: $\mu \in K$

$$L_i \rightarrow L_i - \mu L_j$$

□

REMARQUE 9.1.1. On peut etendre les transformations (I) et (II) au cas $i = j$:

- On a $T_{ii} = \text{Id}_{M_{d' \times d}(K)}$.
- On $Cl_{ii, \mu} = D_{i, 1+\mu}$ et pour que ces transformation soit inversible il faut que $\mu \neq -1$

PROPOSITION 9.2. *Les trois operations elementaires sont obtenues par multiplication a gauche de M par des matrices convenables: pour $1 \leq i \neq j \leq d'$*

- (I) $T_{ij} \cdot \bullet : M \mapsto T_{ij} \cdot M$
- (II) $D_{i, \lambda} \cdot \bullet : M \mapsto D_{i, \lambda} \cdot M$
- (III) $Cl_{ij, \mu} \cdot \bullet : M \mapsto Cl_{ij, \mu} \cdot M$.

ou les matrices carrees $T_{ij}, D_{i, \lambda}, Cl_{ij, \mu} \in M_{d'}(K)$ sont definies par:

$$T_{ij} = \text{Id}_{d'} - E_{ii} - E_{jj} + E_{ij} + E_{ji}.$$

$$D_{i, \lambda} = \text{Id}_{d'} + (\lambda - 1).E_{ii}, \lambda \neq 0$$

$$Cl_{ij, \mu} = \text{Id}_{d'} + \mu.E_{ij}, i \neq j \text{ ou } \mu \neq -1 \text{ si } i = j.$$

Preuve: Notons $E_{ij} = (e_{ij,kl})_{k,j \leq d'}$ la matrice elementaire sous forme de coefficients: on a

$$e_{ij,kl} = \delta_{k=i} \cdot \delta_{l=j}$$

On a donc pour $1 \leq k, l \leq d'$

$$(E_{ij}.M)_{kl} = \sum_{u \leq d'} e_{ij,ku} \cdot m_{ul} = \sum_{u \leq d'} \delta_{k=i} \delta_{u=j} \cdot m_{ul} = \delta_{k=i} m_{jl}.$$

Ainsi le produit $E_{ij}.M$ est la matrice dont la i -ieme ligne est la j -ieme ligne $L_j = (m_{jl})_{l \leq d'}$ et dont toutes les autres coordonnees sont nulles.

– Ainsi $(\text{Id}_{d'} + \mu.E_{ij}).M$ est la matrice formee a partir de M et ou la i -ligne L_i est remplacee par $L_i + \mu.L_j$.

– En particulier, si $i = j$, $(\text{Id}_{d'} + \mu.E_{ii}).M$ est la matrice forme a partir de M et ou la i -ligne L_i est remplacee par $L_i + \mu.L_i = (1 + \mu).L_i$. Ainsi en prenant $\lambda = 1 + \mu$, on multiplie la i -ieme ligne de M par λ .

– De meme $(\text{Id}_{d'} - E_{ii} - E_{jj}).M$ est la matrice M ou les lignes i et j sont remplacees par la ligne nulle $(0)_{l \leq d'}$ et

$$(\text{Id}_{d'} - E_{ii} - E_{jj}).M + (E_{ij} + E_{ji}).M$$

est la matrice precedente ou la ligne L_j est ajoutee a la i -ieme ligne et ou la ligne L_j est ajoutee a la j -ieme ligne de M et c'est donc la matrice M ou les ligne i et j ont ete echangees. $\square$

REMARQUE 9.1.2. En particulier, le fait que ces applications sont lineaires provient du fait que pour toute matrice $D \in M_{d'}(K)$ la multiplication a gauche par D

$$D.\bullet : M \in M_{d' \times d}(K) \mapsto D.M \in M_{d' \times d}(K)$$

est lineaire (par distributivite de la multiplication a gauche, Thm. 7.1).

De plus si D est inversible: $D \in \text{GL}_{d'}(K)$ alors $D.\bullet$ est inversible d'inverse $D^{-1}.\bullet$: en effet

$$D^{-1}.(D.M) = (D^{-1}.D).M = \text{Id}_{d'}.M = M, \quad D.(D^{-1}.M) = (D.D^{-1}).M = \text{Id}_{d'}.M = M.$$

Notons que les matrices T_{ij} , $D_{i,\lambda}$, $Cl_{ij,\mu}$ sont inversibles (si $\lambda \neq 0$ ou $i \neq j$ pour $Cl_{ij,\mu}$) et on a

$$T_{ij}^{-1} = T_{ij}, \quad D_{i,\lambda}^{-1} = D_{i,\lambda^{-1}}, \quad Cl_{ij,\mu}^{-1} = Cl_{ij,-\mu}.$$

REMARQUE 9.1.3. On peut verifier directement que

$$T_{ij}.T_{ij} = \text{Id}_{d'}, \quad D_{i,\lambda}.D_{i,\lambda^{-1}} = \text{Id}_{d'}, \quad Cl_{ij,\mu}.Cl_{ij,-\mu} = \text{Id}_{d'}$$

en utilisant que

$$E_{ij}.E_{kl} = \delta_{j=k} E_{il}$$

DÉFINITION 9.2. Les matrices

$$T_{ij}, \quad D_{i,\lambda}, \quad \lambda \neq 0, \quad Cl_{ij,\mu}$$

pour $i, j \leq d'$, $\lambda \neq 0$, et si $i = j$, $\mu \neq -1$ sont appelees matrices de transformations elementaires.

REMARQUE 9.1.4. On ne confondra pas les matrices de transformations elementaires avec les matrices elementaires qui sont les matrices E_{ij} .

DÉFINITION 9.3. On dit que N est ligne-equivalente a M ssi il existe une suite de transformations elementaires qui transforme M en N .

– De maniere equivalente, N est ligne-equivalente a M ssi il existe une suite finie de matrices des transformations elementaires telle que N est obtenue a partir de M par multiplications a gauche par cette suite de matrices.

EXEMPLE 9.1.2. La matrice M de (9.1.1) est ligne equivalente a la matrice identite Id_3 : on a

$$\text{Id}_3 = Cl_{23,-1} Cl_{12,-1} Cl_{32,1} Cl_{31,-2} D_{1,1/2} T_{12} M$$

PROPOSITION 9.3. La relation etre "ligne-equivalente" est une relation d'equivalence sur $M_{d' \times d}(K)$.

– De plus deux matrices M, N ligne-equivalentes sont equivalentes au sens de la notion d'equivalence de deux matrices de la Definition 7.10.

Preuve: Comme toutes les transformations elementaires sont inversibles et que leur inverse sont des transformations elementaires, cette relation est reflexive, symetrique et transitive.

Si M et N sont lignes-equivalentes, alors

$$N = A.M = A.M.Id_d$$

ou ou A le produit des matrices de transformations elementaires qui permettront de passer de M a N et M et N sont donc equivalentes. $\square$

COROLLAIRE. Si M et N sont lignes equivalentes alors

$$\text{rg}(M) = \text{rg}(N).$$

Preuve: En effet si elles sont lignes-equivalentes elles sont equivalentes et donc ont meme rang. $\square$

PROPOSITION 9.4. Si $N \in M_{d' \times d}(K)$ est ligne-equivalente a M alors toute ligne de N est combinaison lineaire des lignes de M :

$$\forall i \leq d', \text{Lig}_i(N) \in \langle \text{Lig}_1(M), \dots, \text{Lig}_{d'}(M) \rangle \subset K^d$$

et inversement les lignes de M sont combinaisons lineaires des lignes de N . En particulier les SEV engendres par les lignes de M et de N sont les memes

$$\langle \text{Lig}_1(M), \dots, \text{Lig}_{d'}(M) \rangle = \langle \text{Lig}_1(N), \dots, \text{Lig}_{d'}(N) \rangle \subset K^d$$

Preuve: Par definition des transformations elementaires, les lignes de N sont des combinaisons lineaires des lignes de M . Mais comme la relation "ligne-equivalente" est une relation d'equivalence les lignes de M sont CL des lignes de N . $\square$

9.2. Echelonnage

DÉFINITION 9.4. Une matrice $M = (m_{ij}) \in M_{d' \times d}(K)$ est echelonnee si elle est nulle ou bien si

(1) Il existe $1 \leq r \leq d$ et $1 \leq j_1 < \dots < j_r \leq d$ tels que

- Pour la ligne L_1 , le premier terme non-nul est le j_1 -ieme: on a $m_{1j} = 0$ pour tout $j < j_1$ et $m_{1j_1} \neq 0$,
- Pour la ligne L_2 , le premier terme non-nul est le j_2 -ieme: on a $m_{2j} = 0$ pour tout $j < j_2$ et $m_{2j_2} \neq 0$,
- $\vdots$
- Pour la ligne L_r , le premier terme non-nul est le j_r -ieme: on a $m_{rj} = 0$ pour tout $j < j_r$ et $m_{rj_r} \neq 0$

(2) Si $r < d$ les lignes $L_{r+1}, \dots, L_{d'}$ sont toutes nulles.

Si M est non-nulle les $j_1 < \dots < j_r$ sont appeles les echelons de M et les m_{ij_i} , $1 \leq i \leq r$ sont les pivots de M .

La matrice ci-dessous a $r = 3$ echelons: $j_1 = 2, j_2 = 4, j_3 = 5$

$$\begin{pmatrix} 0 & m_{12} & m_{13} & m_{14} & \cdots & \cdots & m_{1d} \\ 0 & 0 & 0 & m_{24} & \cdots & \cdots & m_{2d} \\ 0 & 0 & 0 & 0 & m_{35} & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

DÉFINITION 9.5. Une matrice est echelonnee reduite si le seul coefficient non-nul d'une colonne contenant un pivot est le pivot lui-meme et il vaut 1:

– pour tout $i = 1, \dots, r$

$$m_{ij_i} = 1.$$

– Pour tout $i = 1, \dots, r$ et tout $1 \leq i' \neq i \leq d'$, on a

$$m_{i'j_i} = 0.$$

La matrice ci-dessous a $r = 3$ echelons: $j_1 = 2, j_2 = 4, j_3 = 5$ et est echelonnee reduite.

$$\begin{pmatrix} 0 & 1 & m_{13} & 0 & 0 & \cdots & m_{1d} \\ 0 & 0 & 0 & 1 & 0 & \cdots & m_{2d} \\ 0 & 0 & 0 & 0 & 1 & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

THÉORÈME 9.1 (Gauss). *Toute matrice est ligne-equivalente a une matrice echelonnee reduite.*

Preuve: Si $M = 0_{d' \times d}$ on a termine. Si $M \neq 0_{d' \times d}$, soit j_1 le plus petit indice d'une colonne non-nulle. Soit $m_{ij_1} \neq 0$. Quitte a remplacer M par $T_{1i}.M$ ops $i = 1$.

On peut remplacer la premiere ligne L_1 par $m_{ij_1}^{-1}.L_1$ et supposons que $m_{1,j_1} = 1$. En remplaçant les $L_i, i > 1$ par $L_i - m_{ij_1}.L_1$ annule les autres coefficients de la colonne j_1 et on obtient une matrice ligne-equivalente de la forme (ici $j_1 = 3$)

$$M' = \begin{pmatrix} 0 & 0 & 1 & * & * & \cdots & * \\ 0 & 0 & 0 & m'_{2,j_1+1} & * & \cdots & * \\ 0 & 0 & 0 & * & * & \cdots & \cdots \\ 0 & 0 & 0 & * & * & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & m'_{d',j_1+1} & * & * & * \end{pmatrix}$$

On repete la procedure avec la matrice extraite de M' a partir de la deuxieme ligne et de la $j_1 + 1$ -ieme colonne. On effectue des operations sur les lignes a partir de la deuxieme et donc sans changer la premiere. La matrice M est remplacée par une matrice de la forme

$$M'' = \begin{pmatrix} 0 & 0 & 1 & * & m''_{1j_2} & * & * & \cdots & * \\ 0 & 0 & 0 & 0 & 1 & * & * & \cdots & * \\ 0 & 0 & 0 & 0 & 0 & * & * & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & * & * & * & * \\ \vdots & \vdots & \vdots & 0 & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & * & * & * & * \end{pmatrix}$$

et on peut alors remplacer la premiere ligne L_1'' par $L_1'' - m''_{1j_2}.L_2''$ pour forcer le coefficient au dessus du deuxieme pivot a etre egal a 0. Notons que cette transformation ne modifie pas les coefficients de la ligne L_1 qui sont en position $< j_2$ car les coefficients de L_2'' dans ces positions sont nuls.

On repete l'operation *ad nauseam*.

□

EXEMPLE 9.2.1. L'exemple 9.1.1 est l'echelonnage de la matrice

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

en la matrice echelonnee reduite Id_3 .

THÉORÈME 9.2 (Gauss). *Deux matrices ligne-equivalentes et echelonnees reduites sont egales.*

PREUVE. (due a Yinghan).

EXERCICE 9.1. (★★) Soient $R, R' \in M_{d' \times d}(K)$ deux matrices echelonnees reduites et qui sont lignes equivalentes. On veut montrer que

$$R = R'.$$

Pour $L = (l_1, l_2, \dots, l_d) \in K^d$ un vecteur ligne et $1 \leq j \leq d$, on note

$$e_j^*(L) = l_j$$

la j -ieme coordonnee (dans la base canonique) de L .

Soient $L_1, \dots, L_r, L'_1, \dots, L'_r \subset K^d$ les lignes non-nulles de R et R' (comme R et R' sont lignes equivalentes elles ont meme rang donc $r = r'$), et soit

$$1 \leq j_1 < \dots < j_r \leq d, 1 \leq j'_1 < \dots < j'_r \leq d$$

les positions des pivots de R et R' et

$$W(R) = \text{Vect}(\{L_1, \dots, L_r\}), W(R') = \text{Vect}(\{L'_1, \dots, L'_r\}) \subset K^d$$

les espaces vectoriels engendres par les lignes (non-nulles) de R et R' . On notera egalement pour $1 \leq i \leq r$

$$W_i(R) = \text{Vect}(\{L_i, L_{i+1}, \dots, L_r\}), W_i(R') = \text{Vect}(\{L'_i, L'_{i+1}, \dots, L'_r\})$$

les SEV engendres par les lignes $L_j, j \geq i$ et $L'_j, j \geq i$. En particulier $W_1(R) = W(R)$, $W_r(R) = K.L_r$ et $W_{i+1}(R) \subset W_i(R)$.

- (1) Pourquoi a t'on $W(R) = W(R')$?
- (2) Montrer que pour $1 \leq i, k \leq r$, on a

$$e_{j_i}^*(L_k) = \delta_{k=i}$$

et en deduire que pour tout $L \in W(R)$ on a

$$L = \sum_{i=1}^r e_{j_i}^*(L) L_i$$

(pour la deuxiemem partie, on ecrira L comme CL des $L_i, i \leq r$) et on identifiera les coefficients en appliquant les formes lineaires $e_{j_i}^*$.

- (3) Montrer que pour $L \in K^d$, on a

$$L \in W(R) \implies \forall j < j_1, e_j^*(L) = 0.$$

- (4) En deduire que $j'_1 \geq j_1$ puis que $j'_1 = j_1$ (en observant que R et R' ont des roles symetriques).
- (5) Montrer que pour $L \in W(R)$, on a

$$L \in W_i(R) \iff \forall j < j_i, e_j^*(L) = 0.$$

- (6) Montrer que pour tout $1 \leq i \leq r$ et tout $j < j'_i$ on a $e_j^*(L'_i) = 0$.
- (7) Montrer que $L'_2 \in W_2(R)$ (utiliser que que $j'_2 > j'_1 = j_1$), puis que $j'_2 \geq j_2$ et enfin que $j'_2 = j_2$.
- (8) Montrer (par recurrence) que pour $i = 1, \dots, r, j_i = j'_i$.
- (9) En deduire que pour $i = 1, \dots, r, L'_i = L_i$ puis que $R = R'$ (on appliquera la premiere partie de la Question 2 aux L'_k en utilisant que $j'_i = j_i$).

□

REMARQUE 9.2.1. Les matrices suivantes ne sont pas lignes equivalentes (quelque soit la caracteristique): elles sont echelonnees reduites et distinctes;

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

COROLLAIRE 9.1. (*Unicité de la forme échelonnée réduite*) Soit $M \in M_{d' \times d}(K)$ une matrice alors M est ligne-équivalente à une unique matrice échelonnée réduite (qu'on appelle la forme échelonnée réduite de M).

Preuve: Si M est ligne-équivalente à deux matrices échelonnées réduites R, R' alors R et R' sont ligne-équivalentes (car c'est une relation d'équivalence) et donc $R = R'$. $\square$

9.3. Applications

9.3.1. Calcul du rang. Comme on a observé si M et N sont lignes-équivalentes elles sont équivalentes; on a donc

PROPOSITION 9.5. Si M et N sont lignes équivalentes

$$\text{rg}(M) = \text{rg}(N).$$

Ensuite on a

PROPOSITION 9.6. Si R est échelonnée avec r échelons alors

$$\text{rg}(R) = r.$$

Preuve: Il s'agit de voir que R possède exactement r lignes linéairement indépendantes (cf. Corollaire 7.1). Comme R est échelonnée, elle possède $d' - r$ ligne nulles et r lignes de la forme

$$L_i = (0, \dots, m_{ij_i}, *, \dots, *), \quad i \leq r$$

ou $m_{ij_i} \neq 0$ est en position j_i , $i \leq r$ sur la ligne L_i . Si

$$x_1.L_1 + \dots + x_r.L_r = \mathbf{0}_d$$

la coordonnée j_1 de cette expression donne

$$x_1 m_{1j_1} = 0$$

et donc $x_1 = 0$ (car $m_{1j_1} \neq 0$), ensuite (sachant que $x_1 = 0$) la coordonnée j_2 devient $x_2 m_{2j_2} = 0 \implies x_2 = 0, \dots$, et enfin $x_r m_{rj_r} = 0 \implies x_r = 0$. $\square$

9.3.2. Application aux matrices inversibles.

PROPOSITION 9.7 (Critère d'inversibilité par opérations élémentaires). Soit $M \in M_d(K)$ une matrice carrée alors M est inversible ssi M est ligne équivalente à la matrice identité Id_d .

Preuve: La matrice M est inversible ssi elle est de rang d . Une matrice échelonnée réduite carrée de taille d et de rang d possède d échelons et est donc triangulaire supérieure avec des 1 sur la diagonale; comme elle est réduite, on dessus de chaque 1 on n'a que des 0 et la matrice ne peut être que l'identité. $\square$

9.3.2.1. Engendrement du groupe linéaire par les matrices de transformations élémentaires.

THÉORÈME 9.3. Le groupe linéaire $\text{GL}_d(K)$ est engendré par les matrices des transformations élémentaires

$$T_{ij}, D_{i,\lambda}, Cl_{ij,\mu}, \quad i, j \leq d, \quad \lambda, \mu \in K, \quad \lambda \neq 0, \quad \text{et si } i = j, \quad \mu \neq -1.$$

En d'autres termes (puisque l'ensemble des matrices de transformations élémentaires est stable par inverse) tout matrice $M \in \text{GL}_d(K)$ s'écrit comme un produit fini de ces matrices.

Preuve: Si M est inversible elle est ligne équivalente à l'identité ce qui signifie qu'on peut multiplier à gauche M par un produit Π de $n \geq 1$ matrices de transformations élémentaires et obtenir Id_d :

$$\Pi.M = \text{Id}_d.$$

On a donc

$$M = \Pi^{-1}$$

est un produit d'inverses de matrices de transformations élémentaires et donc un produit de matrices de transformations élémentaires. $\square$

9.3.2.2. *Inversion de matrices par la methode de Gauss.* Cette preuve donne une methode systematique pour inverser une matrice: supposons qu'apres une suite de transformations elementaires on passe de la matrice inversible M a la matrice identite: il existe des matrices de transformations elementaires

$$T_1, T_2, \dots, T_n$$

telles que

$$T_n \cdots T_2 T_1 M = \text{Id}$$

alors

$$M^{-1} = T_n \cdots T_2 T_1.$$

En pratique, on utilise la methode des *vases communicants*: on ecrit l'une a cote de l'autre

$$M \text{ et } \text{Id}_d.$$

Ensuite

- 1. On effectue la premiere transformation elementaire permettant d'echelonner M et on fait la meme transformation sur la matrice Id_d , ce qui revient a multiplier M et Id_d a gauche par T_1 , ce qui donne

$$T_1.M \text{ et } T_1.\text{Id}_d.$$

- 2. On effectue la deuxieme transformation elementaire sur $T_1.M$ et on fait la meme transformation sur la matrice $T_1.\text{Id}_d$, ce qui revient a multiplier les deux matrices a gauche par T_2 , ce qui donne

$$T_2.T_1.M \text{ et } T_2.T_1.\text{Id}_d.$$

- $\vdots$

- n . On effectue la n -ieme transformation elementaire sur $T_{n-1} \cdots T_1.M$ et on fait la meme transformation sur la matrice $T_{n-1} \cdots T_1.\text{Id}_d$, ce qui revient a multiplier les deux matrices a gauche par T_n ce qui donne

$$T_n \cdots T_2 T_1.M = \text{Id}_d \text{ et } T_n \cdots T_2 T_1 = M^{-1}.$$

9.3.3. Extraction d'une base d'une famille generatrice. Soit V un K -EV de diemsnino $d \geq 1$ et

$$\mathcal{G} = \{w_1, \dots, w_l\} \subset V$$

une famille de vecteurs (lignes) et

$$W = \langle \mathcal{G} \rangle$$

l'espace vectoriel qu'ils engendrent. On cherche une base de W .

On choisit $\mathcal{B} = \{e_i, i \leq d\} \subset V$ un base et on identifie alors V a K^d de cette maniere; on associe a chaque w_i son vecteur ligne

$$L_i = \text{Lig}_{\mathcal{B}}(w_i) \in K^d, i \leq l$$

dans cette base. On a donc

$$\langle L_i, i \leq l \rangle = \text{Lig}_{\mathcal{B}}(\langle \mathcal{G} \rangle) = \text{Lig}_{\mathcal{B}}(W).$$

PROPOSITION 9.8 (Description matricielle d'une base d'un SEV). *Soit $M \in M_{l \times d}(K)$ la matrice dont les l lignes sont formees des vecteurs lignes $L_i, i \leq l$. Soit R la matrice echelonnee reduite associee a M et*

$$L'_i = \text{Lig}_i(R), i \leq l$$

l'ensemble des lignes de R alors si R possede r echelons on a

$$\dim W = r$$

et les vecteurs de V correspondants aux r premieres lignes

$$\mathcal{B}_W = \{w'_i = \text{Lig}_{\mathcal{B}}^{-1}(L'_i), i \leq r\}$$

forment une base de W (et les $l - r$ autres vecteurs sont nuls).

On peut alors compléter $\mathcal{B}_W$ en une base $\mathcal{B}$ de V en prenant

$$\mathcal{B} = \mathcal{B}_W \sqcup \{\mathbf{e}_j, j \text{ n'est pas un echelon de } R\}.$$

Preuve: Les $\{L'_i, i \leq r\}$ forment une famille libre et par la proposition 9.4

$$\langle \{L'_i, i \leq r\} \rangle = \langle \{L_i, i \leq l\} \rangle = \text{Lig}_{\mathcal{B}}(W)$$

et comme les L'_i sont nuls pour $i > r$, on a

$$W = \langle \{w_i, i \leq l\} \rangle = \langle \{w'_i, i \leq l\} \rangle = \langle \{w'_i, i \leq r\} \rangle.$$

Pour la completion on note que la famille possède $\dim V$ elements et est libre. $\square$

9.3.4. Resolution de systemes lineaires. Soit $\varphi : V \mapsto W$ une application lineaire entre espaces vectoriels de dimension finies ($d = \dim V$ et $d' = \dim W$). Le probleme qu'on se pose est le suivant:

Etant donne $w \in W$, trouver les $v \in V$ tels que

$$(9.3.1) \quad \varphi(v) = w.$$

Autrement dit, il s'agit de determiner si w appartient a $\varphi(V)$, l'image de V par φ et de calculer l'ensemble des antecedents de w

$$\text{Sol}_{\varphi}(w) = \varphi^{-1}(\{w\}) = \{v \in V, \varphi(v) = w\}.$$

L'equation (9.3.1) s'appelle un *systeme lineaire*.

Rappelons (dans le cadre plus general des groupes quelconques) la structure generale de l'ensemble des solutions de cette equation.

THÉORÈME 9.4 (Resolution d'equations dans les groupes). *Soit $\varphi : G \mapsto H$ un morphisme de groupes alors pour tout $h \in H$, on pose*

$$\text{Sol}_{\varphi}(h) = \varphi^{-1}(\{h\}) = \{g \in G, \varphi(g) = h\} \subset G$$

la preimage de h par φ . En particulier $\text{Sol}_{\varphi}(e_H) = \ker \varphi$. Alors $\text{Sol}_{\varphi}(h)$ est

- soit l'ensemble vide (ssi $h \notin \varphi(G)$),
- soit il existe $g_0 \in \text{Sol}_{\varphi}(h)$ (ce qui equivaut a dire que $h \in \varphi(G)$) et

$$\text{Sol}_{\varphi}(h) = g_0 \cdot \text{Sol}_{\varphi}(e_H) = g_0 \cdot \ker \varphi = \{g_0 \cdot k, \varphi(k) = e_H\}.$$

Preuve: Si $\varphi^{-1}(\{h\}) \neq \emptyset$, soit $g_0 \in G$ tel que $\varphi(g_0) = h$. Alors pour tout g tel que $\varphi(g) = h$ on a

$$\varphi(g_0^{-1} \cdot g) = \varphi(g_0)^{-1} \cdot \varphi(g) = h^{-1} \cdot h = e_H$$

et donc $g = g_0 \cdot k$ avec $k = g_0^{-1} \cdot g \in \ker \varphi$ ce qui montre que

$$\text{Sol}_{\varphi}(h) \subset g_0 \cdot \text{Sol}_{\varphi}(e_H).$$

Reciproquement pour $k \in \ker \varphi$

$$\varphi(g_0 \cdot k) = \varphi(g_0) \cdot \varphi(k) = \varphi(g_0) = h$$

ce qui montre

$$\text{Sol}_{\varphi}(h) \supset g_0 \cdot \text{Sol}_{\varphi}(e_H).$$

$\square$

Appliquant ce resultat general au cas des especes vectoriels (vus comme groupes additifs) $G = V, H = W$ et une application lineaire $\varphi : V \mapsto W$ on obtient

THÉORÈME 9.5 (Resolution d'équations dans les espaces vectoriels). *Soit $\varphi : V \mapsto W$ une application lineaire entre deux espaces vectoriels de dimension finie. Pour tout $w \in W$, on pose*

$$\text{Sol}_\varphi(w) = \varphi^{-1}(\{w\}) = \{v \in V, \varphi(v) = w\} \subset V$$

la preimage de w par φ . En particulier $\text{Sol}_\varphi(\mathbf{0}_W) = \ker \varphi$. Alors $\text{Sol}_\varphi(w)$ est

- *soit $w \notin \varphi(V)$ et $\text{Sol}_\varphi(w)$ est l'ensemble vide,*
- *soit $w \in \varphi(V)$ et il existe $v^0 \in V$ tel que $\varphi(v^0) = w$ et alors*

$$\text{Sol}_\varphi(w) = v^0 + \text{Sol}_\varphi(\mathbf{0}_d) = v^0 + \ker \varphi = \{v^0 + k, k \in \ker \varphi\}.$$

Le corollaire immediat suivant peut alors etre couple avec le Theoreme Noyau-Image:

COROLLAIRE 9.2. *Avec les notations precedente, on a en particulier*

- *si $\dim \ker \varphi = 0$ (cad. $\ker \varphi = \{\mathbf{0}_V\}$ et φ est injective), $\text{Sol}_\varphi(w)$ possede 0 ou 1 element pour tout w .*
- *si $\text{rg} \varphi = \dim \varphi(V) = \dim(W)$ (cad. $\varphi(V) = W$ et φ est surjective) $\text{Sol}_\varphi(w)$ possede au moins un element pour tout w .*
- *Si $\dim V = \dim W$ et que φ est ou bien injective ou bien surjective, φ est bijective et pour tout w , $\text{Sol}_\varphi(w)$ possede exactement un element.*

On va maintenant resoudre ce systeme "abstrait" en le transformant en un probleme concret. Pour cela on se donne des bases

$$\mathcal{B} \subset V, \mathcal{B}' \subset W$$

et

$$M = (m_{ij})_{ij} = \text{mat}_{\mathcal{B}'\mathcal{B}}(\varphi)$$

la matrice de φ dans ces bases. Soient $(v_j)_{j \leq d}$ les coordonnees d'un vecteur $v \in V$ et $(w_i)_{i \leq d'}$ celles de $w \in W$. L'equation (9.3.1) est equivalente au systeme lineaire a d' equations et d inconnues dans K , v_j , $j \leq d$

$$\begin{aligned} m_{11}.v_1 + \cdots + m_{1d}.v_d &= w_1 \\ m_{21}.v_1 + \cdots + m_{2d}.v_d &= w_2 \\ &\vdots \\ m_{d'1}.v_1 + \cdots + m_{d'd}.v_d &= w_{d'} \end{aligned}$$

ou a l'equation matricielle

$$(9.3.2) \quad M \cdot \text{Col}(v) = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1d} \\ m_{21} & m_{22} & \cdots & m_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ m_{d'1} & m_{d'2} & \cdots & m_{d'd} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_{d'} \end{pmatrix} = \text{Col}(w)$$

On cherche alors une condition necessaire et suffisante sur les $(w_i)_{i \leq d'}$ pour que ces equations admettent des solutions $(v_j)_{j \leq d}$.

REMARQUE 9.3.1. En particulier si $w = \mathbf{0}_{d'}$ est le vecteur nul, les solutions nous donnerons les coordonnees des elements du noyau $\ker \varphi$.

DÉFINITION 9.6. *L'equation lineaire (9.3.2) pour un vecteur general w s'appelle equation (ou systeme) lineaire avec second membre (ou non-homogene).*

L'equation lineaire (9.3.2) pour le vecteur nul $\mathbf{0}_W$ s'appelle equation (ou systeme) lineaire sans second membre ou homogene.

Le Theoreme 9.5 et son corollaire 9.2 se reecrivent alors

THÉORÈME 9.6 (Resolution d'équations lineaires). Soit $M = (m_{ij})_{i \leq d', j \leq d}$ une matrice. Pour

toute matrice colonne $w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_{d'} \end{pmatrix} \in \text{Col}_{d'}(K)$, on pose

$$\text{Sol}_M(w) = \{v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{pmatrix} \in \text{Col}_d(K), M.v = w\} \subset \text{Col}_d(K)$$

l'ensemble des solution de l'équation matricielle

$$\begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1d} \\ m_{21} & m_{22} & \cdots & m_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ m_{d1} & m_{12} & \cdots & m_{d'd} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_{d'} \end{pmatrix}$$

Alors $\text{Sol}_M(w)$ est

- soit l'ensemble vide si w n'est pas de la forme $w = M.v_0$ pour $v_0 \in \text{Col}_d(K)$,
- soit de la forme

$$\text{Sol}_M(w) = v_0 + \text{Sol}_M(\mathbf{0}_{d'}) = \{v_0 + k, k \text{Sol}_M(\mathbf{0}_{d'})\}$$

pour tout $v_0 \in \text{Col}_d(K)$ tel que $w = M.v_0$

9.3.4.1. *Systemes lineaires et reduction de matrices.* Pour trouver ces conditions, on applique une suite de transformations elementaires de part et d'autre de l'egalite (9.3.2) de maniere a echelonner-reduire la matrice de gauche. On multiplie les deux termes par un produit $\Pi_n = E_n \cdots E_1$ de matrices de transformations elementaires. Ici, on ne fixe pas la valeurs de w mais on considere ses coordonnees comme des *variables*:

$$\Pi_n \cdot \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1d} \\ m_{21} & m_{22} & \cdots & m_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ m_{d'1} & m_{d'2} & \cdots & m_{d'd} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{pmatrix} = \Pi_n \cdot \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_{d'} \end{pmatrix}$$

On obtient alors un produit dont la premiere matrice est reduite (supposons que le premier pivot soit $j_1 = 1$)

$$\begin{pmatrix} 1 & * & 0 & 0 & * & * \\ 0 & 0 & 1 & 0 & * & * \\ 0 & 0 & 0 & 1 & * & * \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{pmatrix} = \begin{pmatrix} \vdots \\ w'_r \\ 0 \\ 0 \end{pmatrix} = \Pi_n \cdot \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_{d'} \end{pmatrix} = \begin{pmatrix} \vdots \\ w'_r \\ w'_{r+1} \\ \vdots \end{pmatrix}.$$

ou les

$$w'_i = w'_i(w_1, \dots, w_{d'}), \quad i \leq d'$$

sont des combinaisons lineaires des w_i , $i \leq d'$. Notons egalement que comme les lignes d'indice $\geq r + 1$ sont nulles le premier produit fournit un vecteur colonne dont les coordonnees d'indice $\geq r + 1$ sont nulles.

DÉFINITION 9.7. Les inconnues v_{j_i} pour j_i , $1 \leq i \leq r$ etant un echelon sont appellees inconnues principales du systeme. Les inconnues v_j pour $j \leq d$ qui n'est pas un echelon sont appellees inconnues libres du systeme.

On en retire plusieurs informations:

- (1) Le nombre d'échelons est egal au rang de M qui est le rang de φ .
 (2) Les egalites obtenues

$$w'_{r+1} = \dots = w'_{d'} = 0$$

forment un systeme de $d' - r$ equations qui sont les equations cartesiennes l'image $\varphi(V)$:

$$\varphi(V) = \{(w_i)_{i \leq d}, w'_k(w_1, \dots, w_d) = 0, k \geq r+1\} \subset W.$$

- (3) Si $w \in W$ ne satisfait pas les equations ci-dessus alors $w \notin \varphi(V)$ et l'ensemble des solutions est vide.
 (4) Si $w \in W$ satisfait les equations ci-dessus alors $w \in \varphi(V)$ et l'ensemble des solutions est non-vide. On obtient toutes les solutions
- en fixant de maniere arbitraire les inconnues libres v_j (j pas un echelon),
 - puis en resolvant le systeme echelone (dont les inconnues sont les variables principales v_{j_i} , $i \leq r$) en fonctions des inconnues libres prealablement fixees et des $w'_i(w)$, $i \leq r$: on resoud chacune des equations

$$v_{j_i} + \dots = w'_i(w), i \leq r$$

independamment l'une de l'autre; elles ont chacune une solution unique.

Par exemple on peut fixer $v_j^0 = 0$ si j n'est pas un echelon et on ptouve alors $v_{j_i}^0 = w'_i$ pour $i \leq r$.

- (5) Alternativement on obtient toutes les solutions en calculant resolvant le systeme en prenant $w = \mathbf{0}$ le vecteur nul, et en obtenant une relation lineaire entre chaque v_{j_i} , $i \leq r$ et les inconnues libres. Cela nous donne les vecteur du noyau $\ker \varphi$: une base du noyau (qui est de dimension $d - r$) est obtenue en fixant une des inconnue libre egale a 1, et toutes les autres inconnues libres egales a 0 et en fixant (de maniere unique) les inconnues principales de sorte que le systeme d'equations

$$\begin{pmatrix} 1 & * & 0 & 0 & * & * \\ 0 & 0 & 1 & 0 & * & * \\ 0 & 0 & 0 & 1 & * & * \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

soit satisfait.

Ensuite etant donne $w \in \varphi(V)$, on calcule alors une solution particuliere v^0 comme ci-dessus et on lui ajoute un vecteur arbitraire du noyau $\ker \varphi$.

9.4. Operation elementaires sur les colonnes

Soit $M = (m_{ij}) \in M_{d' \times d}(K)$ une matrice. Pour simplifier les notations on ecrira sa i -ieme ligne ($i \leq d'$)

$$C_i = C_i(M) = \text{Col}_i(M) = (m_{ij})_{i \leq d'}$$

DÉFINITION 9.8. Les operations elementaires sur les colonnes d'une matrice sont les applications suivantes de $M_{d' \times d}(K)$ vers $M_{d' \times d}(K)$: pour $i, j \in \{1, \dots, d\}$ et $\lambda \in K^\times$ et $\mu \in K$

- (I) Transposition: Echanger deux colonnes $i \neq j \leq d'$ de M :

$$C_i \longleftrightarrow C_j$$

- (II) Dilatation: Multiplier la i -eme colonne par un scalaire $\lambda \neq 0$:

$$C_i \rightarrow \lambda.C_i.$$

- (III) Combinaison Lineaire: Additionner a la colonne i un multiple scalaire de la j -ieme colonne pour $i \neq j$: $\mu \in K$

$$C_i \rightarrow C_i + \mu C_j$$

Ces transformations sont appelees transformations elementaires sur les colonnes d'une matrice.

On rappelle que les transformations sur les lignes sont donnees par des multiplications a gauche par des matrices inversibles de transformations elementaires (sur les lignes):

$$M' \mapsto T_l.M'$$

Comme la transposition d'une matrice

$$M \leftrightarrow M' = {}^tM$$

transforme la i -ieme colonne de M en la i -ieme ligne de M' et que

$${}^tT_l.M' = {}^tM'.{}^tT_l = M.{}^tT_l,$$

on obtient immediatement

PROPOSITION 9.9. *Une operation elementaire sur les colonnes d'une matrice M equivaut a une operation elementaire sur les lignes de $M' = {}^tM$.*

Une telle transformation est donnee par multiplication par la droite

$$M \mapsto M.{}^tT_l$$

par la transposee d'une matrice de transformation elementaire sur les lignes T_l en composant les operations suivantes

$$M \mapsto {}^tM \mapsto T_l.{}^tM \mapsto {}^tT_l.{}^tM = M.{}^tT_l = M.T_c.$$

Il en resulte que des transformations sont bijectives et lineaires.

DÉFINITION 9.9. *On dit que N est colonne-equivalente a M ssi il existe une suite de transformations elementaires qui transforme M en N .*

– *De maniere equivalente, N est colonne-equivalente a M ssi il existe une suite finie de matrices de transformations elementaires (sur les colonnes) telle que N est obtenue a partir de M par multiplications a droite par cette suite de matrices.*

PROPOSITION 9.10. *La relation etre "colonne-equivalente" est une relation d'equivalence sur $M_{d' \times d}(K)$.*

– *De plus deux matrices M, N colonnes-equivalentes sont equivalentes au sens de la notion d'equivalence de deux matrices de la Definition 7.10. En particulier elles ont meme rang.*

CHAPITRE 10

Determinants

That object was to present the subject as a continuous chain of arguments, separated from all accessories of explanation or illustration, a form which I venture to think better suited for a treatise on exact science than the semi-colloquial semi-logical form often adopted by Mathematical writers.

Lewis Carroll (1867)

10.1. Formes multilinéaires

DÉFINITION 10.1. Soit V un K -espace vectoriel et $n \geq 1$ un entier. Une forme multilinéaire en n variables sur V est une application

$$\Lambda : \begin{matrix} V^n & \mapsto & K \\ (v_1, \dots, v_n) & \mapsto & \Lambda(v_1, \dots, v_n) \end{matrix}$$

telle que pour tout $i = 1, \dots, n$ et tout choix de $n-1$ vecteurs $v_j \in V$, $j \neq i$, l'application Λ "restreinte à la i -ième composante"

$$v_i \in V \mapsto \Lambda(v_1, \dots, v_i, \dots, v_n) \in K$$

est linéaire:

$$\Lambda(v_1, \dots, \lambda v_i + v'_i, \dots, v_n) = \lambda \Lambda(v_1, \dots, v_i, \dots, v_n) + \Lambda(v_1, \dots, v'_i, \dots, v_n).$$

L'ensemble des formes multilinéaires en n variables sur V est noté

$$\text{Mult}^{(n)}(V, K) \text{ ou bien } (V^*)^{\otimes n} (\text{notation "produit tensoriel"}).$$

REMARQUE 10.1.1. Si $n = 1$ c'est la définition usuelle d'une forme linéaire. Si $n = 2$ on parle de forme bi-linéaire, $n = 3$ tri-linéaire, etc...

REMARQUE 10.1.2. Quelques exemples en basse dimension:

- Si $V = K$, $n = 2$ l'application

$$\Pi_2 : \begin{matrix} K^2 & \mapsto & K \\ (x_1, x_2) & \mapsto & \Pi_2(x_1, x_2) = x_1 \cdot x_2 \end{matrix}$$

est multilinéaire. Plus généralement

$$\Pi_n : \begin{matrix} K^n & \mapsto & K \\ (x_1, \dots, x_n) & \mapsto & \Pi_n(x_1, \dots, x_n) = x_1 \times \dots \times x_n \end{matrix}$$

est multilinéaire.

- Soit $V = K^2$ et $n = 2$, on a l'application "produit scalaire"

$$\bullet \bullet : \begin{matrix} K^2 \times K^2 & \mapsto & K \\ ((x_1, y_1), (x_2, y_2)) & \mapsto & (x_1, y_1) \cdot (x_2, y_2) = x_1 \cdot x_2 + y_1 \cdot y_2 \end{matrix}$$

qui est bilinéaire.

– Soit $V = K^2$ et $n = 2$, on a l'application "produit alterne"

$$\bullet \wedge \bullet : \begin{array}{ccc} K^2 \times K^2 & \mapsto & K \\ ((x_1, y_1), (x_2, y_2)) & \mapsto & (x_1, y_1) \wedge (x_2, y_2) = x_1 \cdot y_2 - y_1 \cdot x_2 \end{array}$$

qui est bilinéaire.

EXEMPLE 10.1.1. Soient $\ell_1, \dots, \ell_n : V \mapsto K$ des formes lineaires, alors l'application

$$\ell_1 \otimes \dots \otimes \ell_n : V^n \mapsto K$$

definie par

$$\ell_1 \otimes \dots \otimes \ell_n(v_1, \dots, v_n) = \prod_{i=1}^n \ell_i(v_i) = \ell_1(v_1) \cdot \dots \cdot \ell_n(v_n)$$

est une forme multilinéaire en n variables. C'est en fait l'exemple principal. En effet soit $i \in [1, d]$ fixons des vecteurs v_j pour chaque $j \in [1, d]$ different de i ; l'application

$$v \mapsto \ell_1(v_1) \cdot \dots \cdot \ell_i(v) \cdot \dots \cdot \ell_n(v_n) = \left(\prod_{j \neq i} \ell_j(v_j) \right) \ell_i(v)$$

est un multiple scalaire (de facteur $(\prod_{j \neq i} \ell_j(v_j))$) de la forme lineaire $v \mapsto \ell_i(v)$ et est donc une forme lineaire en v .

REMARQUE 10.1.3. On prendra garde de distinguer la fonction $\ell_1 \otimes \dots \otimes \ell_n$ du produit $\ell_1 \cdot \dots \cdot \ell_n$: le produit $\ell_1 \cdot \dots \cdot \ell_n$ est la fonction d'UNE variable

$$\ell_1 \cdot \dots \cdot \ell_n : v \in V \mapsto \ell_1(v) \cdot \dots \cdot \ell_n(v)$$

alors que la fonction $\ell_1 \otimes \dots \otimes \ell_n$ est une fonction de n variables

$$\ell_1 \otimes \dots \otimes \ell_n : (v_1, \dots, v_n) \in V^n \mapsto \ell_1(v_1) \cdot \dots \cdot \ell_n(v_n) \in K.$$

On a en fait

$$\ell_1 \cdot \dots \cdot \ell_n(v) = \ell_1 \otimes \dots \otimes \ell_n(v, \dots, v).$$

REMARQUE 10.1.4. Notons egalement que l'ordre importe : si $\ell_1 \neq \ell_2$ alors

$$\ell_1 \otimes \ell_2 \otimes \dots \otimes \ell_n \neq \ell_2 \otimes \ell_1 \otimes \dots \otimes \ell_n$$

alors que pour le produit usuel

$$\ell_1 \cdot \ell_2 \cdot \dots \cdot \ell_n \neq \ell_2 \cdot \ell_1 \cdot \dots \cdot \ell_n.$$

Par exemple is $\ell_1 = \mathbf{e}_1^*$, $\ell_2 = \mathbf{e}_2^*$ on a pour $(v_1, v_2) = (\mathbf{e}_1, \mathbf{e}_2)$

$$\mathbf{e}_1^* \otimes \mathbf{e}_2^*(\mathbf{e}_1, \mathbf{e}_2) = \mathbf{e}_1^* \otimes (\mathbf{e}_1) \cdot \mathbf{e}_2^*(\mathbf{e}_2) = 1 \cdot 1 = 1$$

alors que

$$\mathbf{e}_1^* \otimes \mathbf{e}_2^*(\mathbf{e}_2, \mathbf{e}_1) = \mathbf{e}_1^* \otimes (\mathbf{e}_2) \cdot \mathbf{e}_2^*(\mathbf{e}_1) = 0 \cdot 0 = 0.$$

REMARQUE 10.1.5. Attention, V^n est muni d'une structure naturelle de K -ev en posant

$$\lambda \cdot (v_1, \dots, v_n) + (v'_1, \dots, v'_n) = (\lambda \cdot v_1 + v'_1, \dots, \lambda \cdot v_n + v'_n)$$

mais une application $\Lambda : V^n \mapsto K$ qui est lineaire pour cette structure (une forme lineaire sur V^n) n'est pas multilinéaire en general.

Par exemple prenons $V = K$, $n = 2$ et considerons la forme lineaire

$$\Sigma : (x_1, x_2) \in K^2 \mapsto x_1 + x_2 \in K.$$

Fixons x_2 et calculons

$$\Sigma(\lambda x_1 + x'_1, x_2) = \lambda x_1 + x'_1 + x_2$$

et si la forme etait lineaire en la variable x_1 on aurait

$$\Sigma(\lambda x_1 + x'_1, x_2) = \lambda \Sigma(x_1, x_2) + \Sigma(x'_1, x_2) = \lambda x_1 + x_2 + x'_1 + x_2$$

qui ne vaut pas $\lambda x_1 + x'_1 + x_2$ (sauf si $x_2 = 0_K$).

Notons également que si Λ est multilinéaire alors pour tout $i \leq n$ pour tout choix de $n-1$ vecteurs $v_j \in V$ $j \neq i$, l'application

$$v_i \mapsto \Lambda(v_1, \dots, v_i, \dots, v_d)$$

est une forme linéaire et sa valeur en 0_V est nulle

$$\Lambda(v_1, \dots, 0_V, \dots, v_d) = 0_K$$

(le 0_V est placé "en position i "). C'est n'est pas forcément le cas d'une forme linéaire sur l'espace vectoriel V^n (sauf si $(v_1, \dots, 0_V, \dots, v_d)$ est dans le noyau).

REMARQUE 10.1.6. Soient $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n \in K$, $v_1, \dots, v_n \in V$ et Λ une forme multilinéaire alors

$$\Lambda(\lambda_1.v_1 + \mu_1.v'_1, \dots, \lambda_n.v_n + \mu_n.v'_n)$$

est la somme de 2^n termes (2^n est le nombre de décompositions de l'ensemble $\{1, \dots, n\}$ en deux sous-ensembles disjoints):

$$\sum_{I \sqcup J = \{1, \dots, n\}} \left(\prod_{i \in I} \lambda_i \right) \cdot \left(\prod_{j \in J} \mu_j \right) \Lambda(w_{IJ,1}, \dots, w_{IJ,n})$$

avec

$$w_{IJ,i} = \begin{cases} v_i & \text{si } i \in I \\ v'_i & \text{si } i \in J \end{cases}.$$

En particulier

$$\Lambda(\lambda_1.v_1, \dots, \lambda_n.v_n) = \lambda_1 \cdot \dots \cdot \lambda_n \cdot \Lambda(v_1, \dots, v_n)$$

et

$$\Lambda(\lambda.v_1, \dots, \lambda.v_n) = \lambda^n \cdot \Lambda(v_1, \dots, v_n).$$

PROPOSITION 10.1. L'ensemble $\text{Mult}^{(n)}(V, K) = (V^*)^{\otimes n}$ des formes multilinéaires en n variables est un K -espace vectoriel quand on le muni de l'addition et de la multiplication par les scalaires usuelle pour les fonctions de V^n à valeurs dans K : $\forall \Lambda, \Xi \in (V^*)^{\otimes n}$ et pour $\lambda \in K$, la fonction

$$(\lambda\Lambda + \Xi)(v_1, \dots, v_n) = \lambda\Lambda(v_1, \dots, v_n) + \Xi(v_1, \dots, v_n)$$

est encore une forme multilinéaire.

PREUVE. Exercice. □

THÉORÈME 10.1 (Dimension et base de l'espace des formes multilinéaires). Soit $d = \dim V$, $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_d\} \subset V$ une base et $\mathcal{B}^* = \{\mathbf{e}_1^*, \dots, \mathbf{e}_d^*\} \subset V^*$ la base duale. Alors $V^{*\otimes n}$ est de dimension finie égale à d^n ; une base de $V^{*\otimes n}$ est donnée par l'ensemble des formes multilinéaires de la forme

$$\mathbf{e}_{j_1}^* \otimes \dots \otimes \mathbf{e}_{j_n}^*, \text{ quand } j_1, \dots, j_n \text{ parcourent } \{1, \dots, d\}.$$

On note cette base

$$(\mathcal{B}^*)^{\otimes n} = \{\mathbf{e}_{j_1}^* \otimes \dots \otimes \mathbf{e}_{j_n}^*, (j_1, \dots, j_n) \in [1, d]^n\}.$$

Pour tout $\Lambda \in (V^*)^{\otimes n}$ on a la décomposition

$$(10.1.1) \quad \Lambda = \sum_{j_1, \dots, j_n \leq d} \dots \sum \Lambda(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_n}) \mathbf{e}_{j_1}^* \otimes \dots \otimes \mathbf{e}_{j_n}^*$$

Preuve: On va montrer que la famille $(\mathcal{B}^*)^{\otimes n}$ est libre: soit une familles de d^n scalaires

$$\lambda_{j_1, \dots, j_n}, j_1, \dots, j_n \leq d$$

tels que

$$\sum_{j_1, \dots, j_n \leq d} \dots \sum \lambda_{j_1, \dots, j_n} \mathbf{e}_{j_1}^* \otimes \dots \otimes \mathbf{e}_{j_n}^* = \underline{0},$$

on veut montrer que

$$\forall j_1, \dots, j_n \leq d, \lambda_{j_1, \dots, j_n} = 0.$$

Soient $i_1, \dots, i_n \leq d$, on evalue de deux manieres $\Lambda(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_n})$. On a d'abord

$$\Lambda(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_n}) = 0.$$

D'autre part

$$\begin{aligned} \Lambda(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_n}) &= \sum_{j_1, \dots, j_n \leq d} \lambda_{j_1, \dots, j_n} \mathbf{e}_{j_1}^* \otimes \dots \otimes \mathbf{e}_{j_n}^* (\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_n}) \\ &= \sum_{j_1, \dots, j_n \leq d} \lambda_{j_1, \dots, j_n} \mathbf{e}_{j_1}^* (\mathbf{e}_{i_1}) \dots \mathbf{e}_{j_n}^* (\mathbf{e}_{i_n}) = \sum_{j_1, \dots, j_n \leq d} \lambda_{j_1, \dots, j_n} \delta_{j_1=i_1} \dots \delta_{j_n=i_n} \\ &= \lambda_{i_1, \dots, i_n} \end{aligned}$$

Ainsi

$$\forall i_1, \dots, i_n \leq d, \lambda_{i_1, \dots, i_n} = 0$$

et c'est ce que l'on voulait.

On va montrer que la famille est generatrice. On sait deja que c'est le cas pour $n = 1$ (le cas des formes lineaires).

Pour se donner une idee, on traite le cas $n = 2$ (les formes bilineaires). Soit $\Lambda : V \times V \mapsto K$ une forme bilineaire et $v_1, v_2 \in V$. On ecrit pour $i = 1, 2$

$$v_i = \sum_{j=1}^d x_{ij} \mathbf{e}_j = \sum_{j=1}^d \mathbf{e}_j^*(v_i) \mathbf{e}_j$$

et alors on a

$$\Lambda(v_1, v_2) = \Lambda\left(\sum_{j_1=1}^d x_{1j_1} \mathbf{e}_{j_1}, v_2\right).$$

On a par linearite en la premiere variable

$$\Lambda(v_1, v_2) = \sum_{j_1 \leq d} x_{1j_1} \Lambda(\mathbf{e}_{j_1}, v_2) = \sum_{j_1 \leq d} x_{1j_1} \Lambda(\mathbf{e}_{j_1}, \sum_{j_2=1}^d x_{2j_2} \mathbf{e}_{j_2})$$

et par linearite en la deuxieme variable on a

$$\Lambda(\mathbf{e}_{j_1}, \sum_{j_2=1}^d x_{2j_2} \mathbf{e}_{j_2}) = \sum_{j_2=1}^d x_{2j_2} \Lambda(\mathbf{e}_{j_1}, \mathbf{e}_{j_2})$$

et donc

$$\begin{aligned} \Lambda(v_1, v_2) &= \sum_{j_1, j_2 \leq d} \Lambda(\mathbf{e}_{j_1}, \mathbf{e}_{j_2}) x_{1j_1} x_{2j_2} = \sum_{j_1, j_2 \leq d} \Lambda(\mathbf{e}_{j_1}, \mathbf{e}_{j_2}) \mathbf{e}_{j_1}^*(v_1) \mathbf{e}_{j_2}^*(v_2) \\ &= \sum_{j_1, j_2 \leq d} \Lambda(\mathbf{e}_{j_1}, \mathbf{e}_{j_2}) \mathbf{e}_{j_1}^* \otimes \mathbf{e}_{j_2}^*(v_1, v_2). \end{aligned}$$

Ainsi

$$\Lambda = \sum_{j_1, j_2 \leq d} \Lambda(\mathbf{e}_{j_1}, \mathbf{e}_{j_2}) \mathbf{e}_{j_1}^* \otimes \mathbf{e}_{j_2}^*$$

ce qui est la formule (10.1.1) pour $n = 2$.

On traite maintenant le cas general. On montre par recurrence sur n que la famille $(\mathcal{B}^*)^{\otimes n}$ est generatrice et que l'on a l'egalite (10.1.1):

$$\Lambda = \sum_{j_1, \dots, j_n \leq d} \Lambda(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_n}) \mathbf{e}_{j_1}^* \otimes \dots \otimes \mathbf{e}_{j_n}^*.$$

On a déjà montré cela pour $n = 1$ et $n = 2$. Supposons que (10.1.1) est vraie pour tout $\Lambda \in \text{Mult}^{(n)}(V, K)$ et montrons la formule pour $\Lambda \in \text{Mult}^{(n+1)}(V, K)$. Soit

$$\Lambda : (v_1, \dots, v_{n+1}) \rightarrow \Lambda(v_1, \dots, v_{n+1}) \in K$$

une forme multilinéaire en $n + 1$ variables. Pour tout $v_{n+1} \in V$ la fonction

$$\Lambda_{v_{n+1}} : (v_1, \dots, v_n) \rightarrow \Lambda(v_1, \dots, v_n, v_{n+1}) \in K$$

est une forme multilinéaire en n variables. Elle s'écrit donc

$$\Lambda_{v_{n+1}} = \sum_{j_1, \dots, j_n \leq d} \Lambda_{v_{n+1}}(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_n}) \mathbf{e}_{j_1}^* \otimes \dots \otimes \mathbf{e}_{j_n}^*$$

et on rappelle que

$$\Lambda_{v_{n+1}}(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_n}) = \Lambda(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_n}, v_{n+1}).$$

Pour chaque $(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_n}) \in \mathcal{B}^n$, les fonctions

$$v_{n+1} \in V \rightarrow \Lambda(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_n}, v_{n+1}) \in K$$

sont des formes linéaires en v_{n+1} et s'écrivent donc comme CL de la base $\mathcal{B}^* = \{\mathbf{e}_1^*, \dots, \mathbf{e}_d^*\}$

$$\Lambda(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_n}, \bullet) = \sum_{j_{n+1} \leq d} \Lambda(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_n}, \mathbf{e}_{j_{n+1}}) \mathbf{e}_{j_{n+1}}^*(\bullet).$$

On a donc

$$\begin{aligned} \Lambda(v_1, \dots, v_n, v_{n+1}) &= \sum_{j_1, \dots, j_n \leq d} \Lambda_{v_{n+1}}(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_n}) \mathbf{e}_{j_1}^* \otimes \dots \otimes \mathbf{e}_{j_n}^*(v_1, \dots, v_n) \\ &= \sum_{j_1, \dots, j_n \leq d} \mathbf{e}_{j_1}^*(v_1) \dots \mathbf{e}_{j_n}^*(v_n) \sum_{j_{n+1} \leq d} \Lambda(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_n}, \mathbf{e}_{j_{n+1}}) \mathbf{e}_{j_{n+1}}^*(v_{n+1}) \\ &= \sum_{j_1, \dots, j_n, j_{n+1} \leq d} \Lambda(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_n}, \mathbf{e}_{j_{n+1}}) \mathbf{e}_{j_1}^*(v_1) \dots \mathbf{e}_{j_n}^*(v_n) \mathbf{e}_{j_{n+1}}^*(v_{n+1}) \\ &= \sum_{j_1, \dots, j_n, j_{n+1} \leq d} \Lambda(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_n}, \mathbf{e}_{j_{n+1}}) \mathbf{e}_{j_1}^* \otimes \dots \otimes \mathbf{e}_{j_{n+1}}^*(v_1, \dots, v_n, v_{n+1}). \end{aligned}$$

Ainsi on a montré que

$$\Lambda = \sum_{j_1, \dots, j_n, j_{n+1} \leq d} \Lambda(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_n}, \mathbf{e}_{j_{n+1}}) \mathbf{e}_{j_1}^* \otimes \dots \otimes \mathbf{e}_{j_{n+1}}^*.$$

La famille $(\mathcal{B}^*)^{\otimes n+1}$ est donc génératrice. □

EXEMPLE 10.1.2. Pour $V = K^2$ l'espace $\text{Mult}^2(K^2, K)$ est de dimension $2^2 = 4$ et une base est donnée en terme de la base canonique $\mathcal{B}^0\{\mathbf{e}_1^0, \mathbf{e}_2^0\}$:

$$(\mathcal{B}^0)^{\otimes 2} = \{\mathbf{e}_1^0 \otimes \mathbf{e}_1^0, \mathbf{e}_1^0 \otimes \mathbf{e}_2^0, \mathbf{e}_2^0 \otimes \mathbf{e}_1^0, \mathbf{e}_2^0 \otimes \mathbf{e}_2^0\}$$

et le produit scalaire s'écrit

$$\bullet \cdot \bullet = \mathbf{e}_1^0 \otimes \mathbf{e}_1^0 + \mathbf{e}_2^0 \otimes \mathbf{e}_2^0$$

et le produit alterne

$$\bullet \wedge \bullet = \mathbf{e}_1^0 \otimes \mathbf{e}_2^0 - \mathbf{e}_2^0 \otimes \mathbf{e}_1^0$$

10.1.1. Formes symetriques/alternees. A partir d'une forme multilinéaire en n variables on peut en obtenir des nouvelles par "permutation" des variables: par exemple soit $n \geq 2$ et $\Lambda \in \text{Mult}^{(n)}(V, K)$, une forme multilinéaire; on définit alors la forme multilinéaire

$$(12).\Lambda : (v_1, v_2, v_3, \dots, v_n) \mapsto \Lambda(v_2, v_1, v_3, \dots, v_n)$$

en échangeant v_1 et v_2 . Cette forme est à nouveau multilinéaire (le vérifier).

Plus généralement pour $1 \leq i < j \leq n$, on pose

$$(ij).\Lambda : (v_1, \dots, v_i, \dots, v_j, \dots, v_n) \mapsto \Lambda(v_1, \dots, v_j, \dots, v_i, \dots, v_n).$$

DÉFINITION 10.2. Une forme multilinéaire

$$\Lambda : V^n \mapsto K$$

est dite

- Symétrique si $\forall i \neq j \leq n$

$$(ij).\Lambda = \Lambda$$

c'est à dire $\forall (v_1, \dots, v_n) \in V^n$, on a

$$\Lambda(v_1, \dots, v_j, \dots, v_i, \dots, v_n) = \Lambda(v_1, \dots, v_i, \dots, v_j, \dots, v_n).$$

Autrement dit si la valeur de Λ ne change pas quand on échange deux composantes.

- Alternée si $\forall i \neq j \leq n$

$$(ij).\Lambda = -\Lambda$$

c'est à dire $\forall (v_1, \dots, v_n) \in V^n$, on a

$$\Lambda(v_1, \dots, v_j, \dots, v_i, \dots, v_n) = -\Lambda(v_1, \dots, v_i, \dots, v_j, \dots, v_n).$$

Autrement dit si sa valeur est changée en son opposée si on échange deux composantes distinctes.

L'ensemble des formes multilinéaires symétriques en n variables sur V est noté

$$\text{Sym}^{(n)}(V; K).$$

L'ensemble des formes multilinéaires alternées en n variables sur V est noté

$$\text{Alt}^{(n)}(V; K).$$

PROPOSITION 10.2. Les ensembles $\text{Sym}^{(n)}(V; K)$ et $\text{Alt}^{(n)}(V; K)$ sont des SEV de l'espace vectoriel $\text{Mult}^{(n)}(V; K)$.

Preuve: Exercice. Pour cela on utilisera (après l'avoir démontré) le fait que l'application

$$(ij) : \Lambda \in \text{Mult}^{(n)}(V, K) \rightarrow (ij)\Lambda \in \text{Mult}^{(n)}(V, K)$$

est linéaire. □

EXEMPLE 10.1.3. On reprend certains exemples vus précédemment pour $V = K^2$:

- L'application "produit scalaire"

$$\begin{aligned} \bullet \bullet : K^2 \times K^2 &\mapsto K \\ ((x_1, y_1), (x_2, y_2)) &\mapsto (x_1, y_1) \cdot (x_2, y_2) = x_1 \cdot x_2 + y_1 \cdot y_2 \end{aligned}$$

qui est bilinéaire symétrique:

$$(x_1, y_1) \cdot (x_2, y_2) = x_1 \cdot x_2 + y_1 \cdot y_2 = (x_2, y_2) \cdot (x_1, y_1)$$

- L'application "produit alterné"

$$\begin{aligned} \bullet \wedge \bullet : K^2 \times K^2 &\mapsto K \\ ((x_1, y_1), (x_2, y_2)) &\mapsto (x_1, y_1) \wedge (x_2, y_2) = x_1 \cdot y_2 - y_1 \cdot x_2 \end{aligned}$$

qui est bilinéaire alternée:

$$(x_1, y_1) \wedge (x_2, y_2) = x_1 y_2 - y_1 x_2 = -(x_2 y_1 - y_2 x_1) = -(x_2, y_2) \wedge (x_1, y_1).$$

REMARQUE 10.1.7. Pourquoi demande t on que

$$(ij)\Lambda = \pm\Lambda$$

et pas

$$(ij)\Lambda = 3\Lambda ?$$

Soit λ un scalaire tel que

$$(ij)\Lambda = \lambda\Lambda.$$

En reappliquant (ij) a l'egalite precedente on a

$$(ij)(ij)\Lambda = \lambda(ij)\Lambda = \lambda^2\Lambda.$$

Mais

$$(ij)(ij)\Lambda = \Lambda$$

de sorte que si $\Lambda \neq 0$ on obtient

$$\lambda^2 = 1 \iff \lambda = \pm 1.$$

10.1.2. Permutation et signature. La transformation

$$(ij).\Lambda \mapsto \Lambda$$

est un cas particulier d'une transformation plus generale: soit $n \geq 1$ et

$$\sigma : i \in \{1, \dots, n\} \mapsto \sigma(i) \in \{1, \dots, n\}$$

une permutation de $\{1, \dots, n\}$, on dicit alors pour tout n -uplet

$$(v_1, \dots, v_n) \in V^n$$

un nouvel uplet obtenu par permutation des indices en posant

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

On defini alors pour toute forme multilinaire $\Lambda \in \text{Mult}^{(n)}(V, K)$ une nouvelle fonction obtenue par precomposition par $\bullet^\sigma$:

$$\sigma.\Lambda : (v_1, \dots, v_i, \dots, v_n) \mapsto \Lambda((v_1, \dots, v_i, \dots, v_n)^\sigma) = \Lambda(v_{\sigma(1)}, \dots, v_{\sigma(i)}, \dots, v_{\sigma(n)}).$$

On verifie facilement que si Λ est multilinaire alors $\sigma.\Lambda$ est encore multilinaire.

THÉORÈME 10.2 (Action par permutation sur les formes multilinaires). *Pour tout $\sigma \in \mathfrak{S}_n$, l'application*

$$\sigma.\bullet : \Lambda \in \text{Mult}^{(n)}(V, K) \mapsto \sigma.\Lambda \in \text{Mult}^{(n)}(V, K)$$

definit un automorphisme du K -ev $\text{Mult}^{(n)}(V, K)$.

Plus precisement, l'application

$$\sigma \in \mathfrak{S}_n \mapsto \sigma.\bullet \in \text{Aut}(\text{Mult}^{(n)}(V, K))$$

verifie

– Soit Id_n la permutation triviale. On a $\forall \Lambda, \text{Id}_n.\Lambda = \Lambda$ autrement dit

$$\text{Id}_n.\bullet = \text{Id}_{\text{Mult}^{(n)}(V, K)}.$$

– $\forall \Lambda, \forall \sigma, \tau \in \mathfrak{S}_n$, on a

$$(\sigma \circ \tau).\Lambda = \sigma.(\tau.\Lambda)$$

autrement dit

$$(\sigma \circ \tau).\bullet = (\sigma.\bullet) \circ (\tau.\bullet) = \sigma.(\tau.\bullet).$$

En particulier, pour tout σ

$$(\sigma.\bullet) \circ (\sigma^{-1}.\bullet) = \text{Id}_n.\bullet = \text{Id}_{\text{Mult}^{(n)}(V, K)}$$

et donc $\sigma.\bullet$ est un automorphisme lineaire de $\text{Mult}^{(n)}(V, K)$ de reciproque $\sigma^{-1}.\bullet$.

Ainsi

$$\sigma \mapsto \sigma \bullet$$

definit une action à gauche $\mathfrak{S}_n \curvearrowright \text{Mult}^{(n)}(V, K)$ par automorphismes linéaires.

Preuve: On va montrer que

$$(\sigma \circ \tau) \bullet = (\sigma \bullet) \circ (\tau \bullet) = \sigma \bullet (\tau \bullet).$$

et le reste s'en déduit. On a, pour toute forme multilinéaire Λ et tout uplet $(v_1, \dots, v_n) \in V^n$

$$(\sigma \circ \tau) \bullet \Lambda(v_1, \dots, v_n) = \Lambda(v_{\sigma(\tau(1))}, \dots, v_{\sigma(\tau(n))}).$$

Par ailleurs

$$\sigma \bullet (\tau \bullet \Lambda)(v_1, \dots, v_n) = (\tau \bullet \Lambda)((v_1, \dots, v_n)^\sigma) = \tau \bullet \Lambda(v_{\sigma(1)}, \dots, v_{\sigma(n)}).$$

Pour calculer cette dernière expression, faisons le changement de variable

$$w_1 = v_{\sigma(1)}, \dots, w_n = v_{\sigma(n)}.$$

On a alors

$$\tau \bullet \Lambda(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = \tau \bullet \Lambda(w_1, \dots, w_n) = \Lambda((w_1, \dots, w_n)^\tau) = \Lambda(w_{\tau(1)}, \dots, w_{\tau(n)})$$

et

$$w_{\tau(i)} = v_{\sigma(\tau(i))} = v_{\sigma \circ \tau(i)}$$

et ainsi

$$\sigma \bullet (\tau \bullet \Lambda)(v_1, \dots, v_n) = \Lambda(v_{\sigma \circ \tau(1)}, \dots, v_{\sigma \circ \tau(n)}) = ((\sigma \circ \tau) \bullet \Lambda)(v_1, \dots, v_n)$$

□

Pour décrire cette action il est utile de savoir comment $\mathfrak{S}_n$ agit sur une base de $\text{Mult}^{(n)}(V; K)$, notamment la base

$$(\mathcal{B}^*)^{\otimes n} = \{\mathbf{e}_{j_1}^* \otimes \dots \otimes \mathbf{e}_{j_n}^*, (j_1, \dots, j_n) \in \{1, \dots, d\}^n\}.$$

On a le lemme suivant:

LEMME 10.1. Soit V un K -EV de dimension finie, $n \geq 1$ un entier, $\ell_1, \dots, \ell_n \in V^*$, n forme linéaires et $\sigma \in \mathfrak{S}_n$ une permutation, on a

$$\sigma \bullet \ell_1 \otimes \dots \otimes \ell_n = \ell_{\sigma^{-1}(1)} \otimes \dots \otimes \ell_{\sigma^{-1}(n)}$$

ou σ^{-1} est la permutation inverse.

Preuve: Pour $(v_1, \dots, v_n) \in V^n$ on a

$$\sigma \bullet (\ell_1 \otimes \dots \otimes \ell_n)(v_1, \dots, v_n) = \ell_1 \otimes \dots \otimes \ell_n(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = \prod_{i=1}^n \ell_i(v_{\sigma(i)}).$$

Faisons le changement de variable $j = \sigma(i)$, $i = 1, \dots, n$, on a alors $i = \sigma^{-1}(j)$ et

$$\begin{aligned} \prod_{i=1}^n \ell_i(v_{\sigma(i)}) &= \prod_{j=1}^n \ell_{\sigma^{-1}(j)}(v_j) = \ell_{\sigma^{-1}(1)}(v_1) \cdot \dots \cdot \ell_{\sigma^{-1}(n)}(v_n) \\ &= \ell_{\sigma^{-1}(1)} \otimes \dots \otimes \ell_{\sigma^{-1}(n)}(v_1, \dots, v_n) \end{aligned}$$

□

Rappels sur la signature. On rappelle que si K est un corps de caracteristique $\neq 2$, le groupe symetrique $\mathfrak{S}_n$ possede un (unique) morphisme non-trivial de $\mathfrak{S}_n$ vers le groupe multiplicatif $(K^\times, \times)$ appele *signature* et que ce morphisme est a valeurs dans le sous-groupe $\{\pm 1\} \subset K^\times$

$$\begin{aligned} \text{sign} : \mathfrak{S}_n &\mapsto \{\pm 1\} \\ \sigma &\mapsto \text{sign}(\sigma) \end{aligned}$$

defini de la maniere suivante: si σ est la composee de $t \geq 0$ transpositions

$$\sigma = \tau_1 \circ \cdots \circ \tau_t$$

alors

$$\text{sign}(\sigma) = (-1)^t.$$

REMARQUE 10.1.8. On rappelle que toute permutation est la composee de transpositions (ie. l'ensemble des transpositions $\{(ij), 1 \leq i < j \leq n\}$ engendre $\mathfrak{S}_n$). En particulier il existe au plus un morphisme de groupes $\mathfrak{S}_n \rightarrow K^\times$ prenant la valeurs -1 sur toute transposition. Rappelons que cette decomposition en transpositions n'est pas unique. En revanche la parite $t \pmod{2}$ du nombre de ces transpositions est unique et ainsi

$$\text{sign}(\sigma) = (-1)^t = \begin{cases} 1 & \text{si } t \equiv 0 \pmod{2} \\ -1 & \text{si } t \equiv 1 \pmod{2} \end{cases}$$

est bien definie.

THÉORÈME 10.3. *Les formes multilineaires alternees $\text{Alt}^{(n)}(V; K)$ (resp. symetriques $\text{Sym}^{(n)}(V; K)$) sont exactement les formes multilineaires verifiant*

$$(10.1.2) \quad \forall \sigma \in \mathfrak{S}_n, \sigma.\Lambda = \text{sign}(\sigma)\Lambda \text{ (resp. } \sigma.\Lambda = \Lambda).$$

Preuve: Il est clair qu'une forme verifiant (10.1.2) est alternee (resp. symetrique) puisque la signature de la transposition τ_{ij} echangeant $i \neq j$ vaut -1 . Inversement soit Λ une forme alternee; pour tout $\sigma \in \mathfrak{S}_n$, si on ecrit $\sigma = \tau_1 \circ \cdots \circ \tau_t$ alors

$$\sigma.\Lambda = (\tau_1 \circ \cdots \circ \tau_t).\Lambda = (-1)(\tau_1 \circ \cdots \circ \tau_{t-1}).\Lambda = \cdots = (-1) \cdots (-1)\Lambda = (-1)^t \Lambda = \text{sign}(\sigma)\Lambda$$

puisque sign est un morphisme de groupes. $\square$

10.1.3. Dimension des espaces de formes symetriques ou alternees. On va s'interesse particulierement a l'espace des formes alternees.

THÉORÈME 10.4 (Dimension des espaces de formes alternees). *On suppose que $\text{car}(K) \neq 2$. Soit $d = \dim V$. On a*

$$\dim \text{Alt}^{(n)}(V; K) = \begin{cases} 0 & \text{si } n > d \\ 1 & \text{si } n = d \\ C_d^n & \text{si } n \leq d \end{cases}$$

REMARQUE 10.1.9. Si $\text{car}(K) = 2$ alors $-1_K = 1_K$ et

$$\text{Sym}^{(n)}(V; K) = \text{Alt}^{(n)}(V; K).$$

Le theoreme est faux: pour $d = n = 2$ les produit scalaire $\bullet \bullet$ et le produit alterne $\bullet \wedge \bullet$ sont alternees et lineairement independants.

Preuve: (debut) On va seulement demontrer les cas $n > d$ et $n = d$ (qui est celui qui nous interesse le plus).

Notons que si Λ est alternee alors on a

$$\Lambda(v_1, \dots, v, \dots, v, \dots, v_n) = -\Lambda(v_1, \dots, v, \dots, v, \dots, v_n)$$

et donc

$$2\Lambda(v_1, \dots, v, \dots, v, \dots, v_n) = 0_K$$

et donc (car $2_K \neq 0_K$)

$$\Lambda(v_1, \dots, v, \dots, v, \dots, v_n) = 0_K.$$

Plus generalement si la famille

$$\{v_1, \dots, v_n\} \subset V$$

est liee alors

$$\Lambda(v_1, \dots, \dots, v_n) = 0.$$

En effet si la famille est liee, il existe i tel que v_i est combinaison lineaire des autres vecteurs: supposons par exemple que ce soit v_n :

$$v_n = x_1.v_1 + \dots + x_{n-1}.v_{n-1}$$

alors

$$\begin{aligned} \Lambda(v_1, \dots, \dots, v_n) &= \Lambda(v_1, \dots, \dots, v_{n-1}, x_1.v_1 + \dots + x_{n-1}.v_{n-1}) \\ &= x_1\Lambda(v_1, \dots, v_{n-1}, v_1) + \dots + x_{n-1}\Lambda(v_1, \dots, v_{n-1}, v_{n-1}) = 0. \end{aligned}$$

car on a toujours deux vecteurs egaux dans chacun des $n-1$ termes de la somme.

En particulier si $n > d$ une famille $\{v_1, \dots, v_n\}$ de n vecteurs est toujours liee et donc

$$\Lambda(v_1, \dots, v_n) = 0.$$

Cela montre que pour $n > d$

$$\text{Alt}^{(n)}(V; K) = \{0\}.$$

Le cas $n = d$. Supposons que $n = d$. Soit $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_d\}$ une base de V et $\mathcal{B}^* = \{\mathbf{e}_1^*, \dots, \mathbf{e}_d^*\}$ la base duale.

Comme Λ est multilinaire, elle se decompose dans la base $(\mathcal{B}^*)^{\otimes d}$:

$$(10.1.3) \quad \Lambda = \sum_{j_1, \dots, j_d \leq d} \Lambda(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_d}) \mathbf{e}_{j_1}^* \otimes \dots \otimes \mathbf{e}_{j_d}^*$$

et Λ est completement determinee si on connait les valeurs $\Lambda(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_d})$ pour tout les choix possibles de $j_1, \dots, j_d \in \{1, \dots, d\}$.

Notons que si pour $i \neq i'$ on a $j_i = j_{i'}$ alors

$$\Lambda(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_i}, \dots, \mathbf{e}_{j_{i'}}, \mathbf{e}_{j_d}) = \Lambda(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_i}, \dots, \mathbf{e}_{j_i}, \mathbf{e}_{j_d}) = 0.$$

Ainsi la somme (10.1.3) est restreinte aux $j_1, \dots, j_d \in \{1, \dots, d\}$ qui sont *distincts*. Mais cela signifie que l'application

$$i \in \{1, \dots, d\} \mapsto j_i \in \{1, \dots, d\}$$

est une permutation σ de $\{1, \dots, d\}$.

Etant donne une telle permutation σ ; comme Λ est alternee on a

$$\Lambda(\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_d}) = \Lambda(\mathbf{e}_{\sigma(1)}, \dots, \mathbf{e}_{\sigma(d)}) = \text{sign}(\sigma) \Lambda(\mathbf{e}_1, \dots, \mathbf{e}_d).$$

On a donc

$$\begin{aligned} \Lambda &= \sum_{\sigma \in \mathfrak{S}_d} \Lambda(\mathbf{e}_{\sigma(1)}, \dots, \mathbf{e}_{\sigma(d)}) \mathbf{e}_{\sigma(1)}^* \otimes \dots \otimes \mathbf{e}_{\sigma(d)}^* \\ (10.1.4) \quad &= \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) \Lambda(\mathbf{e}_1, \dots, \mathbf{e}_d) \mathbf{e}_{\sigma(1)}^* \otimes \dots \otimes \mathbf{e}_{\sigma(d)}^* \\ &= \Lambda(\mathbf{e}_1, \dots, \mathbf{e}_d) \Delta_{\mathcal{B}} \end{aligned}$$

avec

$$(10.1.5) \quad \Delta_{\mathcal{B}} := \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) \mathbf{e}_{\sigma(1)}^* \otimes \dots \otimes \mathbf{e}_{\sigma(d)}^*.$$

Pour conclure il suffira alors de montrer que

THÉORÈME 10.5. *La forme multilinéaire $\Delta_{\mathcal{B}} : V^d \mapsto K$*

$$\Delta_{\mathcal{B}} := \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) \mathbf{e}_{\sigma(1)}^* \otimes \cdots \otimes \mathbf{e}_{\sigma(d)}^*$$

est alternee et non-nulle.

En effet, si on admet que $\Delta_{\mathcal{B}}$ est une forme alternee non-nulle la formule (10.1.4) nous dit que tout forme alternee est proportionnelle a $\Delta_{\mathcal{B}}$; cela montre que $\{\Delta_{\mathcal{B}}\}$ est une base de $\text{Alt}^{(d)}(V; K)$ et que

$$\dim \text{Alt}^{(d)}(V; K) = 1.$$

Preuve: (du Theoreme 10.5) Evaluons cette forme en $(\mathbf{e}_1, \dots, \mathbf{e}_d)$:

$$\begin{aligned} \Delta_{\mathcal{B}}(\mathbf{e}_1, \dots, \mathbf{e}_d) &= \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) (\mathbf{e}_{\sigma(1)}^* \otimes \cdots \otimes \mathbf{e}_{\sigma(d)}^*)(\mathbf{e}_1, \dots, \mathbf{e}_d) \\ &= \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \mathbf{e}_{\sigma(1)}^*(\mathbf{e}_1) \cdots \mathbf{e}_{\sigma(d)}^*(\mathbf{e}_d) \\ &= \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \delta_{\sigma(1)=1} \cdots \delta_{\sigma(d)=d}. \end{aligned}$$

Ainsi, le seul terme non-nul de cette somme est celui ou

$$\sigma(1) = 1, \sigma(2) = 2, \dots, \sigma(d) = d$$

c'est a dire la permutation triviale: on a donc

$$(10.1.6) \quad \Delta_{\mathcal{B}}(\mathbf{e}_1, \dots, \mathbf{e}_d) = \sigma(\text{Id}_d) = 1.$$

La forme est non-nulle; montrons qu'elle est alternee.

Soit τ une permutation; calculons

$$\begin{aligned} \tau.\Delta_{\mathcal{B}} &= \tau \left(\sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \mathbf{e}_{\sigma(1)}^* \otimes \cdots \otimes \mathbf{e}_{\sigma(d)}^* \right) \\ &= \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \tau.\mathbf{e}_{\sigma(1)}^* \otimes \cdots \otimes \mathbf{e}_{\sigma(d)}^* \end{aligned}$$

car l'action est lineaire. Par le Lemme 10.1 on a donc

$$\begin{aligned} \tau.\Delta_{\mathcal{B}} &= \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \mathbf{e}_{\sigma(\tau^{-1}(1))}^* \otimes \cdots \otimes \mathbf{e}_{\sigma(\tau^{-1}(d))}^* \\ &= \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \mathbf{e}_{\sigma \circ \tau^{-1}(1)}^* \otimes \cdots \otimes \mathbf{e}_{\sigma \circ \tau^{-1}(d)}^* \end{aligned}$$

Faisons le changement de variable

$$\sigma' = \sigma \circ \tau^{-1}, \text{ i.e. } \sigma = \sigma' \circ \tau.$$

On a alors

$$\tau.\Delta_{\mathcal{B}} = \sum_{\sigma' \in \mathfrak{S}_n} \text{sign}(\sigma' \circ \tau) \mathbf{e}_{\sigma'(1)}^* \otimes \cdots \otimes \mathbf{e}_{\sigma'(d)}^*.$$

On a

$$\text{sign}(\sigma' \circ \tau) = \text{sign}(\sigma') \text{sign}(\tau)$$

et donc

$$\tau.\Delta_{\mathcal{B}} = \text{sign}(\tau) \sum_{\sigma' \in \mathfrak{S}_n} \text{sign}(\sigma') \mathbf{e}_{\sigma'(1)}^* \otimes \cdots \otimes \mathbf{e}_{\sigma'(d)}^* = \text{sign}(\tau) \Delta_{\mathcal{B}}.$$

et $\Delta_{\mathcal{B}}$ est bien une forme alternee. □

Explicitement, si les vecteurs v_i , $i = 1, \dots, d$ écrivent dans la base $\mathcal{B}$

$$v_i = \sum_{j=1}^d x_{ij} \mathbf{e}_j, \quad i \leq d$$

avec $x_{ij} = \mathbf{e}_j^*(v_i)$, on obtient

$$\Delta_{\mathcal{B}}(v_1, \dots, v_d) = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) x_{1\sigma(1)} \cdots x_{d\sigma(d)}.$$

DÉFINITION. Soit V un K -EV de dimension d et $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_d\}$ une base de V .

La forme alternée $\Delta_{\mathcal{B}}$ est appelée déterminant de V dans la base $\mathcal{B}$. On la note également

$$\det_{\mathcal{B}} = \Delta_{\mathcal{B}} = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) \mathbf{e}_{\sigma(1)}^* \otimes \cdots \otimes \mathbf{e}_{\sigma(d)}^*.$$

Si on écrit $v_i = \sum_{j=1}^d x_{ij} \mathbf{e}_j$, $i \leq d$, on a

$$\det_{\mathcal{B}}(v_1, \dots, v_d) = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) x_{1\sigma(1)} \cdots x_{d\sigma(d)}.$$

La forme $\det_{\mathcal{B}}$ est l'unique forme multilinéaire alternée Λ vérifiant

$$(10.1.7) \quad \Lambda(\mathbf{e}_1, \dots, \mathbf{e}_d) = 1.$$

C'est une base de $\text{Alt}^{(d)}(V, K)$ et pour $\Lambda \in \text{Alt}^{(d)}(V, K)$ on a

$$(10.1.8) \quad \Lambda = \Lambda(\mathbf{e}_1, \dots, \mathbf{e}_d) \det_{\mathcal{B}}.$$

En particulier, pour $V = K^d$ et $\mathcal{B} = \mathcal{B}_d^0$ la base canonique on note simplement $\Delta_d = \det_d$. Ainsi si pour $i = 1, \dots, d$ on a $v_i = (x_{ij})_{j \leq d}$

$$\det_d(v_1, \dots, v_d) = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) x_{1\sigma(1)} \cdots x_{d\sigma(d)}.$$

REMARQUE 10.1.10. La forme $\det_{\mathcal{B}}$ dépend de la base $\mathcal{B}$ puisqu'elle s'exprime comme un polynôme en les coefficients des vecteurs v_i exprimés dans la base $\mathcal{B}$. Si on choisit une autre base $\mathcal{B}' = \{\mathbf{e}'_1, \dots, \mathbf{e}'_d\}$ alors $\det_{\mathcal{B}'}$ est une autre forme alternée, non-nulle et on a une relation de proportionnalité

$$\det_{\mathcal{B}} = \det_{\mathcal{B}'}(\mathbf{e}'_1, \dots, \mathbf{e}'_d) \det_{\mathcal{B}'}$$

avec $\det_{\mathcal{B}'}(\mathbf{e}'_1, \dots, \mathbf{e}'_d) \in K^\times$ car sinon on obtiendrait que $\det_{\mathcal{B}} \equiv 0$. Échangeant les rôles de $\mathcal{B}$ et $\mathcal{B}'$ on obtient

$$\det_{\mathcal{B}'}(\mathbf{e}'_1, \dots, \mathbf{e}'_d) = \det_{\mathcal{B}}(\mathbf{e}_1, \dots, \mathbf{e}_d)^{-1}.$$

10.1.4. Interlude: Principe général de symétrisation. On va ici donner un principe général de construction des formes alternées.

Pour illustrer dans un cas simple ce processus on rappelle comment on construit une fonction paire ou impaire à partir d'une fonction générale $f : \mathbb{R} \mapsto \mathbb{R}$: on pose

$$f_+(x) := f(x) + f(-x), \quad f_-(x) := f(x) - f(-x);$$

alors f_+ est une fonction paire

$$f_+(-x) = f(-x) + f(-(-x)) = f(-x) + f(x) = f_+(x)$$

et f_- est impaire

$$f_-(-x) = f(-x) - f(-(-x)) = f(-x) - f(x) = -f_-(x).$$

REMARQUE 10.1.11. De plus on a

$$f(x) = \frac{1}{2} f_+(x) + \frac{1}{2} f_-(x).$$

Le cas precedent et le cas de l'action du groupe symetrique $\mathfrak{S}_n \curvearrowright \text{Mult}^{(n)}(V, K)$ sont des cas particulier du contexte suivant: soit W un K -EV et G un groupe fini agissant a gauche sur W lineairement: l'action de G est donnee par un morphisme de G vers le groupe des automorphismes lineaires de W

$$\iota : G \rightarrow \text{GL}(W).$$

On notera cette action

$$g \cdot w = \iota(g)(w).$$

EXEMPLE 10.1.4. Le groupe $\{\pm 1\}$ agit sur les fonctions $f : \mathbb{R} \rightarrow \mathbb{R}$ par ($\varepsilon = \pm 1$)

$$\varepsilon \cdot f(x) = f(\varepsilon x).$$

Le groupe $\mathfrak{S}_n$ agit sur $\text{Mult}^{(n)}(V; K)$ par

$$\sigma \cdot \Lambda : (v_1, \dots, v_n) \in V^n \mapsto \Lambda(v_{\sigma(1)}, \dots, v_{\sigma(n)}) \in K.$$

Supposons qu'on veuille trouver un vecteur w_1 *invariant* sous l'action de G : tel que

$$\forall g \in G, g \cdot w_1 = w_1.$$

EXEMPLE 10.1.5. Par exemple pour $G = \{\pm 1\}$ et $W = \mathcal{F}(\mathbb{R}; \mathbb{R})$ on veut que

$$f(\pm x) = f(x)$$

c'est a dire que f est paire.

Alors on considere pour tout $w \in W$ la somme des transformes de w par tous les elements de G

$$w_1 := \sum_{h \in G} h \cdot w$$

Alors w_1 est invariant:

$$\forall g \in G, g \cdot w_1 = w_1.$$

En effet comme l'action est lineaire

$$\begin{aligned} g \cdot w_1 &= g \cdot \left(\sum_{h \in G} h \cdot w \right) = \sum_{h \in G} g \cdot h \cdot w \\ &= \sum_{h \in G} (g \cdot h) \cdot w \\ &= \sum_{h' \in G} h' \cdot w = w_1 \end{aligned}$$

en faisant le changement de variable $h' = g \cdot h$ car la translation

$$h \in G \mapsto g \cdot h \in G$$

est une bijection de G sur G .

Cela permet d'obtenir les fonctions paires. Pour les fonctions impaire on a la variante suivante:

THÉORÈME 10.6 (Processus de symetrisation pour l'action d'un groupe fini). *Soit K un corps, $(G, \cdot)$ un groupe fini, W un K -ev de dimension finie et*

$$\iota : G \mapsto \text{GL}(W)$$

une action a gauche de G sur W qui est lineaire: ι est morphisme de groupe de G vers le groupe des automorphismes de W . On notera cette action

$$g \cdot w = \iota(g)(w).$$

Soit

$$\chi : G \mapsto (K^\times, \times)$$

un morphisme de G vers le groupe multiplicatif de K (on dit que χ est un caractere de G a valeurs dans $K^\times$). Soit $w \in W$ un vecteur, alors le vecteur

$$w_\chi := \sum_{h \in G} \chi(h)^{-1} \cdot h \cdot v$$

verifie pour tout $g \in G$

$$g \cdot w_\chi = \chi(g) \cdot w_\chi.$$

REMARQUE 10.1.12. Au semestre prochain vous verrez la notion de vecteur propre et de valeur propre pour un endomorphisme: le vecteur w_χ est un vecteur propre pour chaque endomorphisme $\iota(g)$ de valeur propre $\chi(g)$.

Preuve: Comme $g \cdot \bullet$ est lineaire, on a

$$g \cdot w_\chi = g \cdot \left(\sum_{h \in G} \chi(h)^{-1} \cdot h \cdot w \right) = \sum_{h \in G} \chi(h)^{-1} \cdot g \cdot h \cdot w = \sum_{h \in G} \chi(h)^{-1} \cdot (g \cdot h) \cdot w.$$

Posons $h' = g \cdot h$ alors quand h parcourt G , h' parcourt G , on a donc (changement de variable $h = g^{-1} \cdot h'$)

$$\sum_{h \in G} \chi(h)^{-1} \cdot (g \cdot h) \cdot w = \sum_{h' \in G} \chi(g^{-1} \cdot h')^{-1} \cdot h' \cdot w = \chi(g) \sum_{h' \in G} \chi(h')^{-1} \cdot h' \cdot w = \chi(g) \cdot w_\chi;$$

en effet comme χ est un morphisme

$$\chi(g^{-1} \cdot h')^{-1} = \chi(g^{-1})^{-1} \cdot \chi(h')^{-1} = \chi(g) \cdot \chi(h')^{-1}.$$

□

10.1.5. Application a la construction de formes alternees. Prenant $W = \text{Mult}^{(n)}(V; K)$, $G = \mathfrak{S}_n$ agissant par

$$\sigma \cdot \Lambda(v_1, \dots, v_n) = \Lambda(v_{\sigma(1)}, \dots, v_{\sigma(n)})$$

et $\chi = \text{sign} : \mathfrak{S}_n \rightarrow \{\pm 1\}$, et appliquant le Theoreme 10.6 et utilisant que fait que comme $\text{sign}(\sigma) \in \{\pm 1\}$ on a $\text{sign}(\sigma) = \text{sign}(\sigma)^{-1}$, on obtient

COROLLAIRE 10.1. Soit Λ une forme multilineaire en n variables sur V alors

$$\Lambda_{\text{sign}} = \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \sigma \cdot \Lambda$$

est alternee.

REMARQUE 10.1.13. On a demontre que si $n > d$, $\text{Alt}^{(n)}(V; K) = \{\underline{0}_K\}$ donc pour toute forme multilineaire Λ en $n > d$ variables

$$\Lambda_{\text{sign}} = \underline{0}_K.$$

Par contre pour $n \leq d$ cette construction produit souvent une forme alternee non-nulle et cela permet des calculer les dimensions des $\text{Alt}^{(n)}(V; K)$ si $n \leq d$.

Pour $n = d$ et

$$\Lambda = \mathbf{e}_1^* \otimes \dots \otimes \mathbf{e}_d^*$$

on a une forme alternee

$$(\mathbf{e}_1^* \otimes \dots \otimes \mathbf{e}_d^*)_{\text{sign}} = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) \sigma \cdot (\mathbf{e}_1^* \otimes \dots \otimes \mathbf{e}_d^*)$$

qui est donc un multiple de $\det_{\mathcal{B}}$. On a

$$\begin{aligned} (\mathbf{e}_1^* \otimes \cdots \otimes \mathbf{e}_d^*)_{\text{sign}}(\mathbf{e}_1, \dots, \mathbf{e}_d) &= \sum_{\sigma} \text{sign}(\sigma) \sigma.(\mathbf{e}_1^* \otimes \cdots \otimes \mathbf{e}_d^*)(\mathbf{e}_1, \dots, \mathbf{e}_d) \\ &= \sum_{\sigma} \text{sign}(\sigma) \mathbf{e}_1^* \otimes \cdots \otimes \mathbf{e}_d^*(\mathbf{e}_{\sigma(1)}, \dots, \mathbf{e}_{\sigma(d)}) \\ &= \sum_{\sigma} \text{sign}(\sigma) \prod_{i=1}^d \delta_{i=\sigma(i)} = \text{sign}(\text{Id}_d) 1 = 1. \end{aligned}$$

Ainsi on a

$$(\mathbf{e}_1^* \otimes \cdots \otimes \mathbf{e}_d^*)_{\text{sign}} = \det_{\mathcal{B}}.$$

10.2. Determinants

10.2.1. Determinant relatif à une base. Rappelons la

DÉFINITION 10.3. Soit V un K -EV de dimension d et $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_d\}$ une base de V . La forme alternée définie par

$$\det_{\mathcal{B}} = \Delta_{\mathcal{B}} = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) \mathbf{e}_{\sigma(1)}^* \otimes \cdots \otimes \mathbf{e}_{\sigma(d)}^*$$

est appelée *determinant de V dans la base $\mathcal{B}$* .

La forme $\det_{\mathcal{B}}$ est l'unique forme multilinéaire alternée Λ vérifiant

$$(10.2.1) \quad \Lambda(\mathbf{e}_1, \dots, \mathbf{e}_d) = 1.$$

C'est une base de $\text{Alt}^{(d)}(V, K)$ et pour $\Lambda \in \text{Alt}^{(d)}(V, K)$ on a

$$(10.2.2) \quad \Lambda = \Lambda(\mathbf{e}_1, \dots, \mathbf{e}_d) \det_{\mathcal{B}}.$$

Si on prend $V = K^d$ et $\mathcal{B} = \mathcal{B}_d^0$ la base canonique on note simplement $\Delta_d = \det_d$. Ainsi si $v_i = (x_{ij})_{j \leq d}$

$$\det_d(v_1, \dots, v_d) = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) x_{1\sigma(1)} \cdots x_{d\sigma(d)}.$$

10.2.1.1. Expression explicite de $\det_{\mathcal{B}}$.

THÉORÈME 10.7 (Formules combinatoire pour le déterminant). Soient $v_1, \dots, v_d$ des vecteurs et $(x_{ij})_{j \leq d}$ leurs coordonnées dans la base $\mathcal{B}$:

$$v_i = \sum_{j=1}^d x_{ij} \mathbf{e}_j.$$

On a les formules suivantes

$$(10.2.3) \quad \det_{\mathcal{B}}(v_1, \dots, v_d) = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) \prod_{i=1}^d x_{i\sigma(i)} = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) x_{1\sigma(1)} \cdots x_{d\sigma(d)}.$$

$$(10.2.4) \quad \det_{\mathcal{B}}(v_1, \dots, v_d) = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) \prod_{j=1}^d x_{\sigma(j)j} = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) x_{\sigma(1)1} \cdots x_{\sigma(d)d}.$$

Preuve: On a déjà vu la première formule. Pour la deuxième, écrivons $j = \sigma(i)$, on a alors $i = \sigma^{-1}(j)$ et quand i parcourt $\{1, \dots, d\}$, j parcourt également $\{1, \dots, d\}$. On a donc

$$\det_{\mathcal{B}}(v_1, \dots, v_d) = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) \prod_{i=1}^d x_{i\sigma(i)} = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) \prod_{j=1}^d x_{\sigma^{-1}(j)j}.$$

On fait le changement de variable $\sigma \mapsto \sigma^{-1}$ et la somme s'écrit

$$\det_{\mathcal{B}}(v_1, \dots, v_d) = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma^{-1}) \prod_{j=1}^d x_{\sigma(j)j}$$

et comme

$$\text{sign}(\sigma^{-1}) = \text{sign}(\sigma)^{-1} = \text{sign}(\sigma)$$

car $\text{sign}(\sigma) = \pm 1$ on obtient

$$\det_{\mathcal{B}}(v_1, \dots, v_d) = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) \prod_{j=1}^d x_{\sigma(j)j}.$$

□

REMARQUE 10.2.1. Les formules (10.2.3) ou (10.2.4) auraient pu être prise comme définissant le déterminant de d vecteurs exprimés dans une base $\mathcal{B}$ dans un espace de dimension d sans jamais parler de formes multilinéaires alternées et c'est ce qu'on trouve dans de nombreux cours d'algèbre linéaires.

10.2.2. Déterminant d'un endomorphisme. Soit $\varphi : V \mapsto V$ un endomorphisme. À toute forme multilinéaire Λ (en n variables) on associe une nouvelle forme (inspirée de la construction de l'application adjointe pour les formes linéaires) en posant

$$\varphi^*(\Lambda)(v_1, \dots, v_n) := \Lambda(\varphi(v_1), \dots, \varphi(v_n)).$$

On vérifie que $\varphi^*(\Lambda)$ est multilinéaire et que si Λ est alternée ou symétrique $\varphi^*(\Lambda)$ est alternée ou symétrique: si $\sigma \in \mathfrak{S}_n$ est une permutation, on a

$$\begin{aligned} \sigma.(\varphi^*\Lambda)(v_1, \dots, v_n) &= \varphi^*\Lambda(v_{\sigma(1)}, \dots, v_{\sigma(n)}) \\ &= \Lambda(\varphi(v_{\sigma(1)}), \dots, \varphi(v_{\sigma(n)})) = \varphi^*(\sigma.\Lambda)(v_1, \dots, v_n) \end{aligned}$$

et

$$\sigma.\Lambda = \pm \Lambda \implies \sigma.(\varphi^*\Lambda) = \pm \varphi^*\Lambda.$$

REMARQUE 10.2.2. Cette notation $\varphi^*(\Lambda)$ est analogue avec la notation pour l'application linéaire duale dans le cas des formes linéaires (ie. les formes multilinéaires en une variable). Il faut cependant remarquer que $\varphi^*(\Lambda)$ est la composée $\Lambda \circ \varphi^{\otimes n}$ ou $\varphi^{\otimes n} : V^n \mapsto V^n$ est l'application

$$\varphi^{\otimes n} : (v_1, \dots, v_n) \mapsto (\varphi(v_1), \dots, \varphi(v_n)).$$

Ainsi on aurait pu/du poser $(\varphi^{\otimes n})^*(\Lambda)$ au lieu de $\varphi^*(\Lambda)$.

En particulier si $n = d$, $\varphi^*(\det_{\mathcal{B}})$ est proportionnel à $\det_{\mathcal{B}}$:

$$\varphi^*(\det_{\mathcal{B}}) = \lambda_{\varphi} \det_{\mathcal{B}}.$$

En fait si Λ est n'importe quelle autre forme alternée, on a $\Lambda = \lambda \det_{\mathcal{B}}$, $\lambda \in K$ (car $\text{Alt}^{(d)}(V; K)$ est de dimension 1) et

$$\varphi^*\Lambda = \varphi^*(\lambda \det_{\mathcal{B}}) = \lambda \varphi^*(\det_{\mathcal{B}}) = \lambda \lambda_{\varphi} \det_{\mathcal{B}} = \lambda_{\varphi} \lambda \det_{\mathcal{B}} = \lambda_{\varphi} \Lambda.$$

Le facteur de proportionnalité $\lambda_{\varphi} \in K$ s'appelle le déterminant de φ et est noté $\det \varphi$.

DÉFINITION 10.4. Le déterminant de φ , $\det \varphi \in K$ est le scalaire vérifiant pour tout $\Lambda \in \text{Alt}^{(d)}(V; K)$

$$(10.2.5) \quad \varphi^*(\Lambda) = \det(\varphi) \Lambda.$$

En particulier $\det(\varphi)$ ne dépend pas du choix d'une base de V et pour toute base $\mathcal{B} \subset V$ on a

$$\varphi^*(\det_{\mathcal{B}}) = \det(\varphi) \det_{\mathcal{B}}.$$

THÉORÈME 10.8 (Propriétés fonctionnelles du déterminant). Soit $\varphi : V \mapsto V$ un endomorphisme. L'application $\det : \text{End}(V) \mapsto K$ a les propriétés suivantes

(1) Homogénéité: soit $\lambda \in K$ alors

$$\det(\lambda \cdot \varphi) = \lambda^d \cdot \det(\varphi).$$

(2) Multiplicativité: on a

$$\det(\psi \circ \varphi) = \det(\psi) \det(\varphi) = \det(\varphi) \det(\psi) = \det(\varphi \circ \psi).$$

(3) Critère d'inversibilité: on a

$$\det(\varphi) \neq 0 \iff \varphi \in \text{GL}(V).$$

(4) Invariance par conjugaison: pour tout $\varphi \in \text{End}(V)$ et $\psi \in \text{GL}(V)$ on a

$$\det(\text{Ad}(\psi)(\varphi)) = \det(\psi \varphi \psi^{-1}) = \det(\varphi).$$

(5) Morphisme: L'application

$$\det : \text{GL}(V) \mapsto K^\times$$

est un morphisme de groupes. En particulier $\det(\text{Id}_V) = 1$.

Preuve: Soit $\det(\varphi)$ tel que

$$\varphi^*(\det_{\mathcal{B}}) = \det(\varphi) \det_{\mathcal{B}}.$$

Soit Λ une forme alternée quelconque, alors

$$\Lambda = \lambda \cdot \det_{\mathcal{B}}, \quad \lambda \in K$$

et

$$\varphi^*(\Lambda) = \varphi^*(\lambda \cdot \det_{\mathcal{B}}) = (\lambda \cdot \det_{\mathcal{B}}) \circ \varphi^{\otimes d} = \lambda \cdot (\det_{\mathcal{B}} \circ \varphi^{\otimes d}) = \lambda \cdot \varphi^*(\det_{\mathcal{B}}) = \lambda \cdot \det(\varphi) \det_{\mathcal{B}} = \det(\varphi) \Lambda.$$

– Homogénéité: on calcule pour Λ une forme alternée quelconque

$$(\lambda \cdot \varphi)^*(\Lambda)(v_1, \dots, v_d) = \Lambda(\lambda \cdot \varphi(v_1), \dots, \lambda \cdot \varphi(v_d)) = \lambda^d \Lambda(\varphi(v_1), \dots, \varphi(v_d)) = \lambda^d \varphi^*(\Lambda)(v_1, \dots, v_d)$$

car Λ est multilinéaire en d variables. Ainsi par (10.2.5)

$$(\lambda \cdot \varphi)^*(\Lambda) = \det(\lambda \varphi) \Lambda = \lambda^d \det(\varphi) \Lambda.$$

– Multiplicativité: Soient $\varphi, \psi \in \text{End}(V)$, on a

$$(\psi \circ \varphi)^* \Lambda = \Lambda \circ \psi^{\otimes d} \circ \varphi^{\otimes d} = \varphi^*(\psi^* \Lambda) = \varphi^* \circ \psi^*(\Lambda).$$

En effet

$$\begin{aligned} (\psi \circ \varphi)^*(\Lambda)(v_1, \dots, v_n) &= \Lambda(\psi(\varphi(v_1)), \dots, \psi(\varphi(v_n))) \\ &= (\psi^* \Lambda)(\varphi(v_1), \dots, \varphi(v_n)) \\ &= \varphi^*(\psi^* \Lambda)(v_1, \dots, v_n) \end{aligned}$$

Par (10.2.5) on a donc

$$(\psi \circ \varphi)^* \Lambda = \det(\psi \circ \varphi) \Lambda$$

et

$$\varphi^* \circ \psi^*(\Lambda) = \det(\varphi) \psi^*(\Lambda) = \det(\varphi) \det(\psi) \Lambda$$

Ainsi

$$\det(\psi \circ \varphi) = \det(\psi) \det(\varphi);$$

de plus comme K est commutatif

$$\det(\psi \circ \varphi) = \det(\psi) \det(\varphi) = \det(\varphi) \det(\psi) = \det(\varphi \circ \psi).$$

Si $\psi = \text{Id}_V$, on a bien sur

$$\det(\text{Id}_V) = 1$$

car

$$\text{Id}_V^* \Lambda = \Lambda.$$

– *Critere d'inversibilite (condition necessaire)* Si φ est inversible, on a

$$\det(\text{Id}_V) = 1 = \det(\varphi^{-1} \circ \varphi) = \det(\varphi^{-1}) \det(\varphi)$$

ce qui implique que $\det(\varphi^{-1}), \det(\varphi)$ sont non-nuls et inverse l'un de l'autre:

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

– *Morphisme*: On a donc montre que

$$\det : \text{GL}(V) \mapsto K^\times$$

est un morphisme de groupes.

– *Critere d'inversibilite (condition suffisante)* Soit $\varphi \in \text{End}(V) - \text{GL}(V)$ (qui n'est pas inversible) alors

$$\{\varphi(\mathbf{e}_1), \dots, \varphi(\mathbf{e}_d)\}$$

n'est pas une base et est donc liee. En particulier

$$\det(\varphi) = \det(\varphi) \det_{\mathcal{B}}(\mathbf{e}_1, \dots, \mathbf{e}_d) = \det_{\mathcal{B}}(\varphi(\mathbf{e}_1), \dots, \varphi(\mathbf{e}_d)) = 0.$$

□

DÉFINITION 10.5. *Le noyau du morphisme $\det : \text{GL}(V) \mapsto K^\times$ est appelle "groupe special lineaire de V " et on le note*

$$\text{SL}(V) = \ker \det = \{\varphi \in \text{GL}(V), \det \varphi = 1\}.$$

C'est un sous-groupe distingue de $\text{GL}(V)$ (car c'est un noyau).

10.2.3. Determinant d'une matrice.

DÉFINITION 10.6. *Soit $M \in M_d(K)$ une matrice carree de coefficients $M = (m_{ij})_{ij \leq d}$. Le determinant $\det(M)$ de M est (de maniere equivalente):*

(1) *Le scalaire*

$$\det M = \det(\varphi_M)$$

ou $\varphi_M : K^d \mapsto K^d$ est l'application lineaire sur K^d dont la matrice dans la base canonique

$$\text{mat}_{\mathcal{B}^0}(\varphi_M) = M.$$

(2) *Le determinant –relatif a la base canonique $\mathcal{B}_{\text{Col}_d}^0$ de l'espace vectoriel $\text{Col}_d(K)$ des vecteurs colonnes de hauteur d – de l'ensemble des vecteurs colonnes de la matrice M :*

$$\det(M) = \det_{\mathcal{B}_{\text{Col}_d}^0}(\text{Col}_1(M), \dots, \text{Col}_d(M))$$

(3) *Le determinant – relatif a la base canonique $\mathcal{B}_{\text{Lig}_d}^0$ de l'espace vectoriel $\text{Lig}_d(K)$ des vecteurs lignes de longueur d – des vecteurs lignes de la matrice M dans l'espace des vecteurs lignes $\text{Lig}_d(K)$:*

$$\det(M) = \det_{\mathcal{B}_{\text{Lig}_d}^0}(\text{Lig}_1(M), \dots, \text{Lig}_d(M))$$

(4) *La somme*

$$(10.2.6) \quad \det(M) = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) m_{\sigma(1)1} \cdots m_{\sigma(d)d}.$$

(5) *La somme*

$$(10.2.7) \quad \det(M) = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) m_{1\sigma(1)} \cdots m_{d\sigma(d)}.$$

REMARQUE 10.2.3. En sens inverse, pour tout K -EV V , tout endomorphisme $\varphi : V \rightarrow V$ et toute base $\mathcal{B} \subset V$, on a

$$\det(\text{mat}_{\mathcal{B}}(\varphi)) = \det(\varphi).$$

Preuve: (de l'équivalence de la première définition avec les autres) Soit $\varphi_M : K^d \mapsto K^d$ telle que $\text{mat}_{\mathcal{B}^0}(\varphi_M) = M$. C'est à dire que la j -ième colonne de M est formée par les coordonnées de $\varphi_M(\mathbf{e}_j)$ dans la base canonique:

$$\varphi_M(\mathbf{e}_j) = \sum_{i=1}^d m_{ij} \mathbf{e}_i.$$

Par définition

$$\det(M) := \det(\varphi_M)$$

ou $\det(\varphi_M)$ vérifie

$$\varphi^*(\det_{\mathcal{B}^0}) = \det(\varphi_M) \det_{\mathcal{B}^0}.$$

Évaluons cette égalité à $(\mathbf{e}_1, \dots, \mathbf{e}_d)$. On obtient

$$\det_{\mathcal{B}^0}(\varphi(\mathbf{e}_1), \dots, \varphi(\mathbf{e}_d)) = \det(\varphi_M) \det_{\mathcal{B}^0}(\mathbf{e}_1, \dots, \mathbf{e}_d) = \det(\varphi_M) = \det(M).$$

– Cela montre l'équivalence de la première et de la deuxième définition.

– La quatrième égalité (10.2.6) provient du fait que les coordonnées du vecteur colonne $\text{Col}_j(M)$ sont données par les $(m_{ij})_{i \leq d}$ et du Théorème 10.7 : on a

$$\begin{aligned} \det_{\mathcal{B}^0}(\varphi(\mathbf{e}_1), \dots, \varphi(\mathbf{e}_d)) &= \det_{\mathcal{B}^0}((m_{11}, \dots, m_{d1}), \dots, (m_{1d}, \dots, m_{dd})) \\ &= \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) \prod_{i=1}^d m_{\sigma(1)1} \cdots m_{\sigma(d)d} \end{aligned}$$

(on a posé $v_i = (m_{1i}, \dots, m_{di})$ le vecteur correspondant à la i -ème colonne).

– La troisième et la cinquième égalité (10.2.7) proviennent à nouveau du Théorème 10.7: on a

$$\begin{aligned} \det_{\mathcal{B}^0_{\text{Lig}_d}}(\text{Lig}_1(M), \dots, \text{Lig}_d(M)) &= \det_{msB_d^0}((m_{11}, \dots, m_{1d}), \dots, (m_{d1}, \dots, m_{dd})) \\ &= \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) \prod_{i=1}^d m_{1\sigma(1)} \cdots m_{1\sigma(d)} \end{aligned}$$

□

REMARQUE 10.2.4. Pour tout endomorphisme $\varphi : V \rightarrow V$ et toute base $\mathcal{B}$ on a

$$\det(\varphi) = \det(\text{mat}_{\mathcal{B}}(\varphi)).$$

EXEMPLE 10.2.1. Si $d = 2$ et

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$$

et $\mathfrak{S}_2 = \{\text{Id}_2, (12)\}$ On trouve

$$\det(M) = m_{11}m_{22} - m_{12}m_{21}.$$

Autrement dit si

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$\det(M) = ad - bc.$$

Si $d = 3$,

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix}$$

$$\mathfrak{S}_3 = \{\text{Id}_3, (12), (13), (23), (123), (132)\}$$

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix}$$

FIGURE 1. Règle de Sarrus

$$\det(M) = m_{11}m_{22}m_{33} - m_{12}m_{21}m_{33} - m_{13}m_{22}m_{31} - m_{11}m_{23}m_{31} + m_{12}m_{23}m_{31} + m_{13}m_{21}m_{32}.$$

On réécrit quelquefois ce déterminant en groupant ensemble les termes avec un + et ceux avec – pour calculer selon la règle de Sarrus.

$$\det(M) = m_{11}m_{22}m_{33} + m_{12}m_{23}m_{31} + m_{13}m_{21}m_{32} - m_{12}m_{21}m_{33} - m_{13}m_{22}m_{31} - m_{11}m_{23}m_{31}.$$

Il résulte de cette définition et des propriétés du déterminant d'une application linéaire et de (10.2.6) et (10.2.7) que:

THÉORÈME 10.9 (Propriétés fonctionnelles du déterminant des matrices). *Le déterminant d'une matrice a les propriétés suivantes*

(1) *Homogénéité: soit $\lambda \in K$ alors*

$$\det(\lambda.M) = \lambda^d \cdot \det(M).$$

(2) *Invariance par transposition:*

$$\det(M) = \det({}^t M).$$

(3) *Multiplicativité: on a*

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

(4) *Critère d'inversibilité: on a*

$$\det(M) \neq 0 \iff M \in \text{GL}_d(K).$$

(5) *Invariance par conjugaison: pour $C \in \text{GL}_d(K)$*

$$\det(\text{Ad}(C)M) = \det(CMC^{-1}) = \det(M).$$

(6) *Morphisme: L'application*

$$\det : \text{GL}_d(K) \mapsto K^\times$$

est un morphisme de groupes. En particulier $\det(\text{Id}_d) = 1$.

Preuve: Rappelons que si $M = \text{mat}_{\mathcal{B}_0}(\varphi)$, $N = \text{mat}_{\mathcal{B}_0}(\psi)$ alors $M.N = \text{mat}_{\mathcal{B}_0}(\varphi \circ \psi)$ et

$$\det(M.N) = \det(\varphi \circ \psi) = \det(\varphi) \det(\psi) = \det(M) \det(N).$$

Cela montre la multiplicativité qui permet de montrer le critère d'inversibilité ou le fait qu'on a un morphisme.

Pour montrer que (on pose ${}^t M = (m_{ij}^*)_{i,j} = (m_{ji})_{i,j}$)

$$\det(M) = \det({}^t M)$$

on remarque que

$$\begin{aligned}\det M &= \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) m_{\sigma(1)1} \cdots m_{\sigma(d)d} = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) m_{1\sigma(1)} \cdots m_{d\sigma(d)} \\ &= \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) m_{\sigma(1)1}^* \cdots m_{\sigma(d)d}^* = \det({}^t M)\end{aligned}$$

□

COROLLAIRE 10.2. (*Invariance du determinant par dualite*) Soit $\varphi \in \text{End}(V)$ et $\varphi^* \in \text{End}(V^*)$ l'application lineaire duale. On

$$\det \varphi^* = \det \varphi.$$

Preuve: C'est un corollaire de (2) du Theorem 10.9. □

COROLLAIRE 10.3. Soient M et N deux matrices semblables (ie. conjuguées): il existe $P \in \text{GL}_d(K)$ tel que

$$N = P.M.P^{-1}.$$

Alors

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

Le determinant ne depend que de la classe de conjugaison (d'une matrice ou d'un endomorphisme).

Preuve: On a

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

car la corps K est commutatif. □

REMARQUE 10.2.5. Ce resultat s'interprete en terme de changement de base: si $M = \text{mat}_{\mathcal{B}}(\varphi)$ est la matrice dans une certaine base d'une application lineaire φ et $N = \text{mat}_{\mathcal{B}' }(\varphi)$ est la matrice de la meme application calculee dans une autre base. On a par la formule de changement de base

$$N = P.M.P^{-1}$$

ou $P = \text{mat}_{\mathcal{B}' \mathcal{B}}$ est une matrice de changement de base et on obtient que

$$\det N = \det M = \det \varphi.$$

DÉFINITION 10.7. Le noyau du morphisme $\det : \text{GL}_d(K) \mapsto K^\times$ est appelle "groupe special lineaire des matrices de taille d " et on le note

$$\text{SL}_d(K) = \ker \det = \{M \in \text{GL}_d(K), \det M = 1\}.$$

C'est un sous-groupe distingue de $\text{GL}_d(K)$ (car c'est un noyau).

10.3. Calcul de determinants

Pour calculer explicitement des determinants il est pratique de les noter

$$\det(M) = |M| = \begin{vmatrix} m_{11} & \cdots & m_{1d} \\ \vdots & \ddots & \vdots \\ m_{d1} & \cdots & m_{dd} \end{vmatrix}$$

10.3.1. Matrices blocs.

THÉORÈME 10.10 (Determinant des matrices par blocs). *Supposons que la matrice $M \in M_d(K)$ s'écrive sous forme triangulaire supérieure par blocs :*

$$M = \begin{pmatrix} M_1 & * \\ \mathbf{0} & M_2 \end{pmatrix}, \quad M_1 \in M_{d_1}(K), \quad M_2 \in M_{d_2}(K), \quad d_1 + d_2 = d$$

alors

$$\det(M) = \det(M_1) \det(M_2)$$

On va donner deux preuves.

Preuve: (Méthode purement combinatoire) Notons que pour $j \leq d_1$ et $i > d_1$ on a $m_{ij} = 0$. On considère l'expression du déterminant sous la forme

$$\det(M) = \det({}^t M) = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) m_{\sigma(1)1} \cdots m_{\sigma(d)d}.$$

Dans cette somme, on voit donc que les σ tels qu'il existe $1 \leq j \leq d_1$ vérifiant $\sigma(j) > d_1$ ont une contribution nulle car $m_{\sigma(j)j} = 0$. Ainsi la somme définissant le déterminant est le long de l'ensemble $\mathfrak{S}_{d,d_1}$ des permutations σ vérifiant

$$\sigma(\{1, \dots, d_1\}) \subset \{1, \dots, d_1\}$$

et donc

$$\sigma(\{d_1 + 1, \dots, d_1 + d_2\}) \subset \{d_1 + 1, \dots, d_1 + d_2\}.$$

Notons qu'une telle permutation σ induit alors (par restriction) deux permutations

$$\sigma_1 = \sigma|_{\{1, \dots, d_1\}} \in \mathfrak{S}_{d_1}$$

$$\sigma_2 = \sigma|_{\{d_1+1, \dots, d_1+d_2\}} \in \mathfrak{S}_{\{d_1+1, \dots, d_1+d_2\}} \simeq \mathfrak{S}_{d_2}$$

et on a

$$\sigma = \sigma_1 \circ \sigma_2$$

en considérant σ_1 comme la permutation de $\{1, \dots, d\}$ qui permute le sous-ensemble $\{1, \dots, d_1\}$ par σ_1 et qui est l'identité sur $\{d_1 + 1, \dots, d_1 + d_2\}$ (et similairement pour σ_2). En particulier on a

$$\text{sign}(\sigma) = \text{sign}(\sigma_1) \text{sign}(\sigma_2).$$

On laisse le lemme suivant au lecteur:

LEMME 10.2. *L'ensemble $\mathfrak{S}_{d,d_1}$ est un sous groupe de $\mathfrak{S}_d$ et l'application*

$$\sigma \mapsto (\sigma_1, \sigma_2)$$

est un isomorphisme de groupes

$$\mathfrak{S}_{d,d_1} \simeq \mathfrak{S}_{d_1} \times \mathfrak{S}_{\{d_1+1, \dots, d_1+d_2\}} \simeq \mathfrak{S}_{d_1} \times \mathfrak{S}_{d_2}.$$

On peut donc réécrire

$$\begin{aligned} \det(M) &= \sum_{\sigma_1 \in \mathfrak{S}_{d_1}} \sum_{\sigma_2 \in \mathfrak{S}_{d_2}} \text{sign}(\sigma_1) \text{sign}(\sigma_2) \prod_{i=1}^{d_1} m_{\sigma_1(i)i} \times \prod_{i=1}^{d_2} m_{d_1+\sigma_2(i), d_1+i} \\ &= \left(\sum_{\sigma_1 \in \mathfrak{S}_{d_1}} \text{sign}(\sigma_1) \prod_{i=1}^{d_1} m_{\sigma_1(i)i} \right) \times \left(\sum_{\sigma_2 \in \mathfrak{S}_{d_2}} \text{sign}(\sigma_2) \prod_{i=1}^{d_2} m_{d_1+\sigma_2(i), d_1+i} \right) = \det(M_1) \det(M_2). \end{aligned}$$

□

COROLLAIRE 10.4. Soit $k \geq 2$ un entier, si M est une matrice triangulaire superieure a k blocs

$$M = \begin{pmatrix} M_1 & * & * \\ \mathbf{0} & \ddots & * \\ \mathbf{0} & \mathbf{0} & M_k \end{pmatrix}, \quad M_i \in M_{d_i}(K), \quad i \leq k, \quad d_1 + \dots + d_k = d$$

on a

$$\det M = \det(M_1) \cdot \dots \cdot \det(M_k).$$

En particulier, si M est triangulaire superieure ($k = d$) –par exemple diagonale–

$$M = \begin{pmatrix} \lambda_1 & * & \dots & \dots \\ 0 & \lambda_2 & * & * \\ \vdots & 0 & \ddots & * \\ 0 & \dots & \dots & \lambda_d \end{pmatrix},$$

on a

$$\det M = \lambda_1 \cdot \dots \cdot \lambda_d.$$

10.3.1.1. *Matrices triangulaires inferieures par blocs.* Une matrice M est triangulaire inferieure par blocs si elle est de la forme

$$M = \begin{pmatrix} M_1 & \mathbf{0} \\ * & M_2 \end{pmatrix}, \quad M_1 \in M_{d_1}(K), \quad M_2 \in M_{d_2}(K), \quad d_1 + d_2 = d.$$

THÉORÈME 10.11. Supposons que la matrice $M \in M_d(K)$ s'ecrive sous forme triangulaire inferieure par blocs:

$$M = \begin{pmatrix} M_1 & \mathbf{0} \\ * & M_2 \end{pmatrix}, \quad M_1 \in M_{d_1}(K), \quad M_2 \in M_{d_2}(K), \quad d_1 + d_2 = d.$$

alors

$$\det(M) = \det(M_1) \det(M_2).$$

Soit $k \geq 2$ un entier, si M est une matrice triangulaire inferieure a k blocs

$$M = \begin{pmatrix} M_1 & \mathbf{0} & \mathbf{0} \\ * & \ddots & \mathbf{0} \\ * & * & M_k \end{pmatrix}, \quad M_i \in M_{d_i}(K), \quad i \leq k, \quad d_1 + \dots + d_k = d$$

on a

$$\det M = \det(M_1) \cdot \dots \cdot \det(M_k).$$

Preuve: Sa transposee tM est alors triangulaire superieure par blocs de la forme

$${}^tM = \begin{pmatrix} {}^tM_1 & * \\ \mathbf{0} & {}^tM_2 \end{pmatrix}.$$

alors on a par invariance du determinant par transposition

$$\det(M) = \det({}^tM) = \det({}^tM_1) \det({}^tM_2) = \det(M_1) \det(M_2).$$

□

Preuve: (Par factorisation) Ecrivons

$$M = \begin{pmatrix} M_1 & M_3 \\ \mathbf{0} & M_2 \end{pmatrix}, \quad M_3 \in \text{mat}_{d_1 \times d_2}(K).$$

Notons que si M_1 ou M_2 n'est pas inversible la matrice M n'est pas inversible: c'est clair si M_1 n'est pas inversible car la famille des d_1 premieres colonnes sera liee et si M_2 n'est pas inversible la famille des d_2 dernieres colonnes sera liee: dans ces deux cas $\det M = 0 = \det(M_1) \det(M_2)$.

Si M_1 et M_2 sont inversibles, on a la factorisation

$$\begin{aligned} M &= \begin{pmatrix} M_1 & M_3 \\ \mathbf{0} & M_2 \end{pmatrix} = \begin{pmatrix} M_1 & \mathbf{0} \\ \mathbf{0} & \text{Id}_{d_2} \end{pmatrix} \begin{pmatrix} \text{Id}_{d_1} & M_1^{-1}M_3 \\ \mathbf{0} & M_2 \end{pmatrix} \\ &= \begin{pmatrix} M_1 & \mathbf{0} \\ \mathbf{0} & \text{Id}_{d_2} \end{pmatrix} \begin{pmatrix} \text{Id}_{d_1} & M_1^{-1}M_3 \\ \mathbf{0} & M_2 \end{pmatrix} \begin{pmatrix} \text{Id}_{d_1} & M_1^{-1}M_3 \\ \mathbf{0} & M_2 \end{pmatrix} \\ &= \begin{pmatrix} M_1 & \mathbf{0} \\ \mathbf{0} & \text{Id}_{d_2} \end{pmatrix} \begin{pmatrix} \text{Id}_{d_1} & M_1^{-1}M_3M_2^{-1} \\ \mathbf{0} & \text{Id}_{d_2} \end{pmatrix} \begin{pmatrix} \text{Id}_{d_1} & \mathbf{0} \\ \mathbf{0} & M_2 \end{pmatrix} \\ &= M'_1 M'_3 M'_2 \end{aligned}$$

On a donc

$$\det(M) = \det \begin{pmatrix} M_1 & \mathbf{0} \\ \mathbf{0} & \text{Id}_{d_2} \end{pmatrix} \det \begin{pmatrix} \text{Id}_{d_1} & M_1^{-1}M_3M_2^{-1} \\ \mathbf{0} & \text{Id}_{d_2} \end{pmatrix} \det \begin{pmatrix} \text{Id}_{d_1} & \mathbf{0} \\ \mathbf{0} & M_2 \end{pmatrix}$$

et il suffit de montrer que ces determinants valent

$$\det(M'_1) = \det(M_1), \det(M'_3) = 1, \det(M'_2) = \det(M_2)$$

respectivement.

On a

$$\det \begin{pmatrix} M_1 & \mathbf{0} \\ \mathbf{0} & \text{Id}_{d_2} \end{pmatrix} = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) m'_{1,\sigma(1)1} \cdots m'_{1,\sigma(d)d}.$$

Notons que pour $j \geq d_1 + 1$, la j -ieme colonne n'a qu'un seul terme non-nul, le j -ieme; on a donc $m'_{1,\sigma(j)j} = 0$ sauf si $\sigma(j) = j$ auquel cas $m'_{jj} = 1$. Ainsi la somme porte sur les permutations σ telles que $\sigma(j) = j$ pour tout $j \geq d_1 + 1$ c'est a dire les permutations qui fixent tous les elements entre $d_1 + 1$ et d . L'ensemble de ces permutations $\mathfrak{S}_{d, \geq d_1+1}$ forme un sous-groupe isomorphe a $\mathfrak{S}_{d_1}$ (en envoyant une permutation de $\mathfrak{S}_{d_1}$ sur la permutation de $\{1, \dots, d\}$ qui permute les elements de 1 a d_1 et fixe les autres)

$$\sigma_1 \in \mathfrak{S}_{d_1} \mapsto \sigma \in \mathfrak{S}_{d, \geq d_1+1} : \begin{cases} j \mapsto \sigma_1(j), & j \leq d_1 \\ j \mapsto j, & j \geq d_1 + 1. \end{cases}$$

et on a

$$\text{sign}(\sigma) = \text{sign}(\sigma_1).$$

On a alors

$$\begin{aligned} \det \begin{pmatrix} M_1 & \mathbf{0} \\ \mathbf{0} & \text{Id}_{d_2} \end{pmatrix} &= \sum_{\sigma \in \mathfrak{S}_{d, \geq d_1+1}} \text{sign}(\sigma) m'_{1,\sigma(1)1} \cdots m'_{1,\sigma(d_1)d_1} 1 \cdots 1 \\ &= \sum_{\sigma \in \mathfrak{S}_{d_1}} \text{sign}(\sigma) m'_{1,\sigma(1)1} \cdots m'_{1,\sigma(d_1)d_1} = \det(M_1). \end{aligned}$$

On montre par un raisonnement similaire que

$$\det \begin{pmatrix} \text{Id}_{d_1} & \mathbf{0} \\ \mathbf{0} & M_2 \end{pmatrix} = \det(M_2)$$

en notant que la somme $\sum_{\sigma \in \mathfrak{S}_d} \cdots$ pour sur les σ tels que

$$\forall j \leq d_1, \sigma(j) = j$$

et l'ensemble de ces permutations $\mathfrak{S}_{d, \leq d_1}$ est un sous-groupe isomorphe a $\mathfrak{S}_{d_2}$, l'isomorphisme etant donne par

$$\sigma_2 \in \mathfrak{S}_2 \mapsto \sigma \in \mathfrak{S}_{d, \leq d_1} : \begin{cases} j \mapsto j, & j \leq d_1 \\ d_1 + j \mapsto d_1 + \sigma_2(j). \end{cases}$$

et que la signature est preservee.

Pour la matrice du milieu M'_3 on écrit

$$\det M'_3 = \sum_{\sigma_2 \in \mathfrak{S}_{d_2}} \text{sign}(\sigma_2) \prod_{j=1}^{d_2} m'_{3, d_1 + \sigma_2(j), d_1 + j}.$$

Notons que si $\sigma_2(j) > j$ alors $m'_{3, d_1 + \sigma_2(j), d_1 + j} = 0$ car la matrice M'_3 est triangulaire supérieure donc nécessairement la somme porte sur les σ_2 telles que

$$\forall j = 1, \dots, d_2, \sigma(j) \leq j$$

mais il n'existe qu'une seule telle permutation, Id_{d_2} et alors

$$m'_{3, d_1 + \text{Id}_{d_2}(j), d_1 + j} = m'_{3, d_1 + j, d_1 + j} = 1.$$

On obtient donc

$$\det M'_3 = 1.$$

□

10.3.2. Calcul par opérations élémentaires sur les lignes.

LEMME 10.3. Soient T_{ij} , $D_{i,\lambda}$, $Cl_{ij,\mu}$ les matrices associées aux transformations élémentaires sur les lignes d'une matrice. On a

$$\det T_{ij} = -1 \text{ (si } i \neq j)$$

$$\det D_{i,\lambda} = \lambda$$

$$\det Cl_{ij,\mu} = 1, \text{ (si } i \neq j).$$

Preuve: Notons que T_{ij} est la matrice telle que pour toute matrice carrée de taille $d \times d$ l'application

$$M \mapsto T_{ij}M$$

échange les lignes i et j de M . On a donc (disons que $i < j$)

$$\begin{aligned} \det(T_{ij}.M) &= \det(T_{ij}) \det(M) = \det_{\mathcal{B}_{\text{Lig}}^0}(L_1, \dots, L_j, \dots, L_i, \dots, L_d) \\ &= -\det_{\mathcal{B}_{\text{Lig}}^0}(L_1, \dots, L_i, \dots, L_j, \dots, L_d) \end{aligned}$$

car $\det_{\mathcal{B}_{\text{Lig}}^0}(\dots)$ est alternée.

La matrice $D_{i,\lambda}$ est diagonale avec des 1 sur la diagonale sauf en i -ème position où on a λ et donc

$$\det D_{i,\lambda} = 1 \cdot \dots \cdot 1 \cdot \lambda = \lambda.$$

On a pour $i \neq j$,

$$Cl_{ij,\mu} = \text{Id}_d + \mu.E_{ij}, \quad i \neq j$$

qui est une matrice triangulaire inférieure ou supérieure (suivant que $i < j$ ou $i > j$) avec des 1 sur la diagonale, son déterminant vaut donc 1. □

COROLLAIRE 10.5. Supposons que N soit déduite de M par une des trois types de transformations élémentaires sur les lignes de M alors on a

- Type (I): $\det N = -\det M$.
- Type (II): $\det N = \lambda \det M$
- Type (III): $\det N = \det M$

Preuve: En effet on a suivant les cas

$$N = T_{ij}.M, \quad N = D_{i,\lambda}.M, \quad N = Cl_{ij,\mu}$$

et $\det(N)$ est le produit du déterminant de M et de cette matrice. □

En utilisant ce corollaire on peut calculer $\det M$ en échelonnant la matrice M et en gardant la trace des transformations élémentaires effectuées. Si E est une forme échelonnée de M , on a

$\det E = 0 = \det M$ si E a $r < d$ échelons (car E et donc M ne sont pas inversibles) et si E a d échelons E est triangulaire supérieure et son déterminant se calcule facilement.

Par exemple si E est la forme échelonnée réduite et que $r = d$ alors on a $E = \text{Id}_d$. On a alors

$$T_k.T_{k-1} \cdots T_1.M = \text{Id}_d$$

avec T_j des matrices de transformations élémentaires et on a

$$\det(T_k.T_{k-1} \cdots T_1.M) = \det(T_k) \cdots \det(T_1) \det(M) = \det(\text{Id}_d) = 1$$

et

$$\det M = \det(T_1)^{-1} \cdots \det(T_k)^{-1}.$$

REMARQUE 10.3.1. En pratique il vaut mieux ne pas appliquer de transformation de type II juste des transformations de type I (de déterminant -1) ou III (de déterminant 1). On peut alors toujours réduire la matrice sous forme triangulaire supérieure avec $\lambda_1, \dots, \lambda_d$ sur la diagonale et si on se souvient du nombre e d'échanges de lignes réalisés on aura

$$\det M = (-1)^e \lambda_1 \cdots \lambda_d.$$

EXEMPLE 10.3.1.

$$\begin{vmatrix} 2 & 3 & 1 \\ 3 & 1 & 2 \\ 1 & 2 & 3 \end{vmatrix} = - \begin{vmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \\ 3 & 1 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -1 & -5 \\ 0 & -5 & -7 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -1 & -5 \\ 0 & 0 & 18 \end{vmatrix} = -18$$

$$\begin{vmatrix} X & 0 & 0 & d \\ -1 & X & 0 & c \\ 0 & -1 & X & b \\ 0 & 0 & -1 & X+a \end{vmatrix} = \begin{vmatrix} X & 0 & 0 & d \\ 0 & X & 0 & c + \frac{d}{X} \\ 0 & -1 & X & b \\ 0 & 0 & -1 & X+a \end{vmatrix} = \begin{vmatrix} X & 0 & 0 & d \\ 0 & X & 0 & c + \frac{d}{X} \\ 0 & 0 & X & b + \frac{c}{X} + \frac{d}{X^2} \\ 0 & 0 & -1 & X+a \end{vmatrix}$$

$$= \begin{vmatrix} X & 0 & 0 & d \\ 0 & X & 0 & c + \frac{d}{X} \\ 0 & 0 & X & b + \frac{c}{X} + \frac{d}{X^2} \\ 0 & 0 & 0 & X+a + \frac{1}{X}(b + \frac{c}{X} + \frac{d}{X^2}) \end{vmatrix} = X^4 + aX^3 + bX^2 + cX + d.$$

10.3.3. Développement –de Lagrange– le long d'une ligne-colonne. On va maintenant donner une méthode (due à Lagrange) de calcul du déterminant par récurrence sur la dimension d . Soit $M = (m_{ij}) \in M_d(K)$ une matrice de dimension d et $k, l \leq d$, on pose $M(k|l) \in M_{d-1}(K)$ la matrice de dimension $d-1$ obtenue à partir de M en effaçant la i -ième ligne et la j -ième colonne: le scalaire $M(i|j)$ est le (i, j) -ième mineur de la matrice M .

THÉORÈME 10.12 (Développement de Lagrange le long d'une colonne). On a pour tout $j \leq d$

$$\det M = \sum_{i=1}^d m_{ij} (-1)^{i+j} \det(M(i|j)).$$

Preuve: On va montrer le résultat pour $\text{car}(K) \neq 2$. Soient $v_1, \dots, v_d \in K^d$ les vecteurs de coordonnées des colonnes de M qu'on note

$$v_k = m_{1k} \mathbf{e}_1 + \cdots + m_{dk} \mathbf{e}_d.$$

On a

$$\det M = \det_{\mathcal{B}}(v_1, \dots, v_j, \dots, v_d).$$

On va d'abord montrer la formule pour $j = 1$: soit le premier vecteur

$$v_1 = m_{11} \mathbf{e}_1 + \cdots + m_{d1} \mathbf{e}_d$$

et par multilinearité on a

$$\det_{\mathcal{B}}(v_1, v_2, \dots, v_d) = \sum_{i=1}^d m_{i1} \det_{\mathcal{B}}(\mathbf{e}_i, v_2, \dots, v_d).$$

Pour fixer les idées on suppose que $i \neq 1, d$. Notons pour $l \geq 2$

$$v_l^{(i)} = \sum_{k \neq i} m_{kl} \mathbf{e}_k;$$

alors on a

$$\det_{\mathcal{B}}(\mathbf{e}_i, v_2, \dots, v_d) = \det_{\mathcal{B}}(\mathbf{e}_i, v_2^{(i)}, \dots, v_d^{(i)}).$$

Notons que l'application

$$\Lambda^{(i)} : (v_2^{(i)}, \dots, v_d^{(i)}) \mapsto \det_{\mathcal{B}}(\mathbf{e}_i, v_2^{(i)}, \dots, v_d^{(i)})$$

est une forme multilinéaire alternée en $d-1$ variables sur le sous-espace vectoriel

$$K^{d,(i)} = \{v \in K^d, \mathbf{e}_i^*(v) = 0\} = \text{Vect}(\mathbf{e}_1, \dots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_d)$$

des vecteurs de V dont la coordonnée suivant $\mathbf{e}_i$ est nulle: (disons que $i \neq 1, d$).

Une base de cet espace est donnée par

$$\mathcal{B}^{(i)} = \{\mathbf{e}_k, 1 \leq k \neq i \leq d\}.$$

Comme $(\text{car}(K) \neq 2)$ l'espace des formes alternées est de dimension 1, on a (disons que $i \neq 1, d$)

$$\Lambda^{(i)}(\bullet) = \Lambda^{(i)}(\mathbf{e}_1, \dots, \hat{\mathbf{e}}_i, \dots, \mathbf{e}_d) \det_{\mathcal{B}^{(i)}}(\bullet) = \det_{\mathcal{B}}(\mathbf{e}_i, \mathbf{e}_1, \dots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_d) \det_{\mathcal{B}^{(i)}}(\bullet)$$

et donc

$$\Lambda^{(i)}(\mathbf{e}_1, \dots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_d) = \det_{\mathcal{B}}(\mathbf{e}_i, \mathbf{e}_1, \dots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_d);$$

mais

$$\det_{\mathcal{B}}(\mathbf{e}_i, \mathbf{e}_1, \dots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_d) = (-1)^{i-1} \det_{\mathcal{B}}(\mathbf{e}_1, \dots, \mathbf{e}_i, \dots, \mathbf{e}_d) = (-1)^{i+1}$$

car on ramène $\mathbf{e}_i$ de la première à la i -ième position par $i-1$ transpositions. On obtient donc

$$\det_{\mathcal{B}}(v_1, \dots, v_d) = \sum_{i=1}^d m_{i1} (-1)^{i+1} \det_{\mathcal{B}^{(i)}}(v_2^{(i)}, \dots, v_d^{(i)})$$

et donc

$$\det_{\mathcal{B}^{(i)}}(v_2^{(i)}, \dots, v_d^{(i)}) = \det(M(i|1))$$

on conclut si $j = 1$.

Dans le cas général, si $j \neq 1$, on pose $M' = (m'_{kl})_{k,l \leq d} = (1j).M$ la matrice dont on a échangé la première et la j -ième colonne: on a donc

$$m'_{i1} = m_{ij}, \quad m'_{ij} = m_{i1}.$$

On a (par transposition)

$$\det M' = -\det M$$

et développant par rapport à la première colonne on a

$$-\det M = \det M' = \sum_{i=1}^d m'_{ij} (-1)^{i+1} \det(M'(i|1)).$$

Mais $M'(i|1)$ est la matrice carrée de taille $d-1$ dont on a retiré la i -ième ligne et dont la $j-1$ -ième colonne est la première colonne de M (moins le i -ième coefficient). On ramène alors la $j-1$ -ième colonne en première position par $j-1$ transpositions; le déterminant de cette dernière matrice est le mineur $\det(M(i|j))$. On a donc

$$\det(M'(i|1)) = (-1)^{j-1} \det(M(i|j))$$

et

$$\det M = \sum_{i=1}^d m_{ij} (-1)^{i+j} \det(M(i|j)).$$

□

Par le meme raisonnement, on demontre le

THÉORÈME 10.13 (Developpement de Lagrange le long d'une ligne). *On a pour tout $i \leq d$*

$$\det M = \sum_{j=1}^d m_{ij} (-1)^{i+j} \det(M(i|j)).$$

Preuve: Par calcul direct ou en utilisant l'invariance par transposition et le fait qu'un developpement le long d'une ligne devient un developpement le long d'une colonne par transposition. □

EXEMPLE 10.3.2. Soit la matrice 3×3

$$M = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

Si on developpe par rapport a la premiere colonne on obtient

$$\det M = a \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - d \det \begin{pmatrix} b & c \\ h & i \end{pmatrix} + g \det \begin{pmatrix} b & c \\ e & f \end{pmatrix}$$

et par rapport a la deuxieme colonne on obtient

$$\det M = -b \det \begin{pmatrix} d & f \\ g & i \end{pmatrix} + e \det \begin{pmatrix} a & c \\ g & i \end{pmatrix} - h \det \begin{pmatrix} a & c \\ d & f \end{pmatrix}$$

et si on developpe par rapport a la premieres ligne

$$\det M = a \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \det \begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \det \begin{pmatrix} d & e \\ g & h \end{pmatrix}$$

10.3.4. Formule de Cramer.

DÉFINITION 10.8. *Pour $k, l \leq d$*

- le determinant $\det(M(k|l))$ est appele le (k, l) mineur de M .
- le determinant avec signe, $(-1)^{k+l} \det(M(k|l))$ est appele le (k, l) cofacteur de M .
- La matrice des cofacteurs de M , est la matrice dont les coefficients sont les cofacteurs de M :

$$\text{cof}(M) = (\tilde{m}_{ij})_{\substack{i \leq d \\ j \leq d}}, \quad \tilde{m}_{ij} = (-1)^{i+j} \det(M(i|j))$$

THÉORÈME 10.14 (Formule de Cramer). *Soit $M \in M_d(K)$ et $\text{cof}(M)$ sa matrice des cofacteurs. On a*

$$M \cdot {}^t \text{cof}(M) = {}^t \text{cof}(M) \cdot M = \det(M) \cdot \text{Id}_d.$$

En particulier si $\det M \neq 0$, alors M est inversible et son inverse est donnee par

$$M^{-1} = \frac{1}{\det M} {}^t \text{cof}(M).$$

REMARQUE 10.3.2. En particulier si $d = 2$ et $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ on a

$$\text{cof}(M) = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}, \quad {}^t \text{cof}(M) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

et on retrouve la formule

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (ad - bc) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Preuve: Soit $M = (m_{ij})_{i,j \leq d}$ comme ci-dessus et soit $M' = {}^t \text{cof}(M)$ la transposée de la matrice des cofacteurs de M : on a

$$m'_{ji} = \tilde{m}_{ij} = (-1)^{i+j} \det M(i|j)$$

et le développement de Lagrange le long d'une colonne se réécrit

$$\sum_{i=1}^d m'_{ji} m_{ij} = \det M.$$

Par la règle de produit de matrices, on voit qu'il s'agit du coefficient (j, j) de la matrice produit $M'.M$.

Les autres coefficients de ce produit sont donnés, pour $k \neq j$ par les sommes

$$\sum_{i=1}^d m'_{ki} m_{ij} = \sum_{i=1}^d m_{ij} (-1)^{i+k} \det(M(i|k)).$$

On va les calculer (montrer qu'ils valent 0) en les interprétant comme un développement de Lagrange d'un déterminant.

Soit $M^{(j,k)}$ la matrice dont toutes les colonnes sont égales à celles de M sauf la k -ième qui est égale à la j -ième colonne de M . On a pour $i = 1, \dots, d$

$$m_{ik}^{(j,k)} = m_{ij}, \quad M^{(j,k)}(i|k) = M(i|k);$$

en effet la matrice extraite $M^{(j,k)}(i|k)$ est égale à la matrice extraite $M(i|k)$ car cette dernière obtenue en effaçant la k -ième colonne (et la i ligne) et c'est seulement le long de cette colonne que M et $M^{(j,k)}$ diffèrent.

D'autre part, comme $M^{(j,k)}$ a deux colonnes égales, on a

$$\det M^{(j,k)} = 0$$

et par le développement de Lagrange par rapport à la k -ième colonne on a (pour $k \neq j$)

$$\sum_{i=1}^d m_{ik}^{(j,k)} (-1)^{i+k} M^{(j,k)}(i|k) = \sum_{i=1}^d m_{ij} (-1)^{i+k} \det(M(i|k)) = 0 = \sum_{i=1}^d m'_{ki} m_{ij}.$$

On a donc montré que

$$M'.M = {}^t \text{cof}(M).M = \det(M). \text{Id}_d.$$

En utilisant le développement suivant les lignes on obtient

$$M.{}^t \text{cof}(M) = \det(M). \text{Id}_d.$$

On a donc démontré la formule de Cramer. □

10.3.5. Applications de la formule de Cramer. L'intérêt de la formule de Cramer est surtout théorique: pour calculer en pratique l'inverse d'une matrice il vaut mieux utiliser la méthode de Gauss.

En revanche, on observe que la transposée de la matrice des cofacteurs ${}^t \text{cof}(M)$ a pour coefficients des polynômes en les coefficients M et que $\det M$ est également un polynôme en les coefficients de M .

On en tire des applications algébriques et analytiques

Application algebrique. Soit $A \subset K$ est un sous-anneau et $M \in M_d(A)$ alors

$$\det M \in A, \quad {}^t\text{cof}(M) \in M_d(A)$$

et si $\det M \neq 0$

$$M^{-1} \in \frac{1}{\det M} M_d(A).$$

En particulier si $\det M \in A^\times$,

$$M^{-1} \in M_d(A).$$

On en deduit que $M_d(A)$ est un sous-anneau de $M_d(K)$ dont le groupe des unites (des elements inversibles) est

$$\text{GL}_d(A) = \{M \in M_d(A), \det M \in A^\times\}.$$

Application analytique. Supposons que $K = \mathbb{R}$ alors

$$M_d(\mathbb{R}) \simeq \mathbb{R}^{d^2}$$

herite de la topologie produit de celle de $\mathbb{R}$.

Les fonctions

$$\det(\bullet) : \begin{array}{ccc} M_d(\mathbb{R}) & \mapsto & \mathbb{R} \\ M & \mapsto & \det M \end{array}, \quad {}^t\text{cof}(\bullet) : \begin{array}{ccc} M_d(\mathbb{R}) & \mapsto & M_d(\mathbb{R}) \\ M & \mapsto & {}^t\text{cof}(M) \end{array}$$

sont continues (car polynomiales en les coordonnees de M) et

$$\text{GL}_d(\mathbb{R}) = \{M \in M_d(\mathbb{R}), \det M \neq 0\}$$

est un ouvert de $M_d(\mathbb{R})$ (le complement de $M_d(\mathbb{R})^0 = \{M \in M_d(\mathbb{R}), \det M = 0\}$ qui est ferm car $\det(\bullet)$ est continu). On en deduit que l'inversion

$$M \in \text{GL}_d(\mathbb{R}) \mapsto M^{-1} = \frac{1}{\det M} {}^t\text{cof}(M) \in \text{GL}_d(\mathbb{R})$$

est continue sur $\text{GL}_d(\mathbb{R})$: en particulier etant donne une matrice inversible $M \in \text{GL}_2(\mathbb{R})$, si $N \in M_d(\mathbb{R})$ est suffisamment proche de M alors N est inversible et N^{-1} est proche de M^{-1} .

10.4. Le determinant en caracteristique 2

Si $\text{car}(K) = 2$ une partie des raisonnements precedents ne s'appliquent pas car l'espace des formes alternees en d variables tel qu'on l'a defini n'est pas forcement de dimension 1 (cet espace coincide avec l'espace des formes symetriques car $-1_K = 1_K$).

Un maniere de s'en tirer est de redefinir une forme alternee de la maniere suivante:

DÉFINITION 10.9. Soit V un K -EV de dimension $d \geq 1$. Une forme multilinéaire

$$\Lambda : (v_1, \dots, v_n) \in V^n \rightarrow \Lambda(v_1, \dots, v_n) \in K$$

est alternee si pour tout $1 \leq i < j \leq n$, et tout $(v_1, \dots, \hat{v}_j, \dots, v_n) \in V^{n-1}$ on a

$$\Lambda : (v_1, \dots, v_i, \dots, v_i, \dots, v_n) = 0.$$

On note

$$\text{Alt}^{(n)}(V; K) \subset \text{Mult}^{(n)}(V; K)$$

le sous-ensemble des formes alternees.

REMARQUE 10.4.1. Si $\text{car} K \neq 2$ c'est equivalent a la definition precedente mais pas en caractéristique 2.

On peut alors montrer que $\text{Alt}^{(n)}(V; K)$ est un SEV de $\text{Mult}^{(n)}(V; K)$ et que

THÉORÈME 10.15. Soit K un corps (quelconque) et V un K -EV de dimension $d \geq 1$ alors

$$\dim \text{Alt}^{(d)}(V; K) = 1$$

et une base de $\text{Alt}^{(d)}(V; K)$ est donnée par la forme

$$\det_{\mathcal{B}} = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) \mathbf{e}_{\sigma(1)}^* \otimes \cdots \otimes \mathbf{e}_{\sigma(d)}^*$$

qui est alternée et non-nulle.

On peut alors étendre la théorie en adaptant les preuves en conséquence.

Une autre manière est de voir qu'on dispose toujours de la forme multilinéaire obtenue par symétrisation:

$$\det_{\mathcal{B}} = \text{sign}(\sigma) \mathbf{e}_{\sigma(1)}^* \otimes \cdots \otimes \mathbf{e}_{\sigma(d)}^*$$

car dans un corps de caractéristique 2, $\text{sign}(\sigma)_K = (\pm 1)_K = 1_K$. Elle vérifie donc

$$\sigma \cdot \det_{\mathcal{B}} = \text{sign}(\sigma) \det_{\mathcal{B}} = \det_{\mathcal{B}}.$$

On peut partir de là pour définir une théorie du déterminant.

Par exemple on a

THÉORÈME 10.16. Soit K un corps quelconque et V un K -ev de dimension d . La forme $\det_{\mathcal{B}}$ vérifie que si pour $i \neq j$, on a $v_i = v_j$ alors

$$\det_{\mathcal{B}}(v_1, \dots, v_d) = 0_K$$

et c'est plus généralement vrai si la famille $\{v_1, \dots, v_d\}$ est liée.

PREUVE. On donne la preuve en caractéristique générale: on peut supposer en appliquant une permutation convenable que $i = 1$ et $j = 2$ et donc pour $k = 1, \dots, d$, on a

$$x_{1k} = x_{2k}.$$

Soit $\tau = (12)$ la transposition qui permute 1 et 2. Soit

$$\mathfrak{A}_d = \ker(\text{sign}) = \{\sigma \in \mathfrak{S}_d, \text{sign}(\sigma) = +1\}$$

le groupe alterne des permutations paires. alors $\mathfrak{A}_d$ est d'indice 2 dans $\mathfrak{S}_d$ et comme $\tau \notin \mathfrak{A}_d$ on a

$$\mathfrak{S}_d = \mathfrak{A}_d \sqcup \mathfrak{A}_d \circ (12).$$

On a alors

$$\begin{aligned} \det_{\mathcal{B}}(v_1, \dots, v_d) &= \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) x_{1\sigma(1)} \cdot x_{2\sigma(2)} \cdots x_{d\sigma(d)} = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) x_{1\sigma(1)} \cdot x_{1\sigma(2)} \cdots x_{d\sigma(d)} \\ &= \sum_{\sigma \in \mathfrak{A}_d} \text{sign}(\sigma) x_{1\sigma(1)} \cdot x_{1\sigma(2)} \cdots x_{d\sigma(d)} + \sum_{\sigma \in \mathfrak{A}_d} \text{sign}(\sigma \circ \tau) x_{1\sigma \circ \tau(1)} \cdot x_{1\sigma \circ \tau(2)} \cdots x_{d\sigma(d)} \\ &= \sum_{\sigma \in \mathfrak{A}_d} x_{1\sigma(1)} \cdot x_{1\sigma(2)} \cdots x_{d\sigma(d)} - \sum_{\sigma \in \mathfrak{A}_d} x_{1\sigma(2)} \cdot x_{1\sigma(1)} \cdots x_{d\sigma(d)} = 0_K. \end{aligned}$$

□

On développe alors la théorie du déterminant en caractéristique quelconque de la manière suivante:

- (1) Prenant $V = K^d$ et $\mathcal{B} = \mathcal{B}^0$, on définit ainsi le déterminant de d vecteurs de K^d .
- (2) On définit également le déterminant d'une matrice par la même formule:

$$\det(M) = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) m_{1\sigma(1)} \cdots m_{d\sigma(d)} = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) m_{\sigma(1)1} \cdots m_{\sigma(d)d}.$$

et on montre par un calcul direct sur les matrices et les permutations que le théorème 10.9 reste vrai.

- (3) On definit alors le determinant d'une application lineaire generale $\varphi : V \mapsto V$ en posant

$$\det(\varphi) := \det \text{mat}_{\mathcal{B}}(\varphi)$$

pour une base quelconque $\mathcal{B}$ de V . On peut montrer par un calcul direct (utilisant la Theoreme 10.16) que

$$\varphi^*(\det_{\mathcal{B}}) = \det \varphi \cdot \det_{\mathcal{B}}.$$

Par ailleurs la formule de changement de base, conjuguee au Theoreme 10.9 montre que cette definition ne depend pas du choix de la base. On deduit du Theoreme 10.9 que le Theoreme 10.8 est vrai.

- (4) Les resultats concernant le determinant des matrices par bloc restent vrais.
 (5) On montre directement par le calcul que les developements de Lagrange le long d'une ligne ou d'une colonne restent vrais (Theoremes 10.12 et 10.13) ainsi que la formule de Cramer.

REMARQUE 10.4.2. Un interet de ce dernier point de vue est qu'on peut remplacer K par un anneau commutatif pas forcement integre.

Le polynome caracteristique

11.1. Le polynome caracteristique d'une matrice

Soit $K[X]$ l'anneau des polynomes a coefficients dans K . C'est (voir le l'appendice A pour la definition formelle en terme de suite $(a_n)_{n \geq 0} \in K_{fin}^{\mathbb{N}}$ a support fini) l'ensemble des expressions de la forme

$$P(X) = a_0X^0 + a_1X + \cdots a_dX^d = a_0 + a_1X + \cdots a_dX^d, \quad d \geq 0, \quad a_0, \dots, a_d \in K.$$

C'est un K -EV d'element neutre le polynome nul $0(X) = 0$ en posant

$$(P + Q)(X) = (a_0 + b_0) + (a_1 + b_1)X + \cdots (a_d + b_d)X^d, \quad \lambda.P(X) = \lambda.a_0 + \lambda.a_1X + \cdots \lambda.a_dX^d.$$

C'est un anneau commutatif d'unité de polynome constant $1(X) = 1$ quand on le munit du produit usuel

$$PQ(X) = \sum_{k \leq 2d} c_k X^k$$

avec

$$c_k = \sum_{i+j=k} a_i b_j, \quad \text{en posant } a_i, b_j = 0 \text{ pour } i, j > d.$$

L'application deg

$$\deg P = \max\{j \geq 0, a_j \neq 0\}, \quad \deg 0 = -\infty$$

et le fait que

$$\deg(P.Q) = \deg P + \deg Q$$

permet de montrer que c'est anneau integre dont le corps des fractions est le corps des fractions rationnelles a coefficients dans K

$$K(X) = \left\{ \frac{P(X)}{Q(X)}, \quad P, Q \in K[X], \quad Q \neq 0 \right\}.$$

Soit $M \in M_d(K)$ une matrice. Comme $K \hookrightarrow K(X)$ (tout element de $\lambda \in K$ peut etre identifie a le polynome constant $\lambda(X) = \lambda$) on peut voir M comme une matrice a coefficients dans $M_d(K(X))$ ainsi que la matrice

$$X.Id_d - M \in M_d(K(X))$$

dont les coordonnees sont donnees par

$$(X.Id_d - M)_{ij} = X\delta_{i=j} - m_{ij}.$$

On peut donc calculer son determinant

$$\det(X.Id_d - M) = \sum_{\sigma} \text{sign}(\sigma) \prod_{i=1}^d (X\delta_{i\sigma(i)} - m_{i\sigma(i)})$$

qui est en fait un polynome en X .

DÉFINITION 11.1. *Le polynome caracteristique de M est le determinant*

$$P_{car,M}(X) = \det(X.Id_d - M) = \sum_{\sigma} \text{sign}(\sigma) \prod_{i=1}^d (X\delta_{i\sigma(i)} - m_{i\sigma(i)}) \in K[X]$$

THÉORÈME 11.1. *Le polynome caracteristique est un polynome unitaire de degree d et si on ecrit*

$$\det(X.\text{Id}_d - M) = X^d + a_{d-1}X^{d-1} + \cdots + a_0$$

On a

$$a_0 = P(0) = (-1)^d \det M,$$

$$a_{d-1} = -\text{tr}(M) = -(m_{11} + \cdots + m_{dd})$$

est la trace de la matrice M .

Preuve: On voit que

$$\det(X.\text{Id}_d - M) = \sum_{\sigma} \text{sign}(\sigma) \prod_{i=1}^d (X\delta_{i\sigma(i)} - m_{i\sigma(i)})$$

est une somme de polynomes de degree au plus d ; de plus la contribution de $\sigma = \text{Id}_d$ est

$$\prod_{i=1}^d (X - m_{ii})$$

est un polynome unitaire de degree d .

Notons egalement que si $\sigma \neq \text{Id}$ il existe i tel que $\sigma(i) \neq i$ et $X\delta_{i\sigma(i)} - m_{i\sigma(i)} = -m_{i\sigma(i)}$; ainsi $\prod_{i=1}^d (X\delta_{i\sigma(i)} - m_{i\sigma(i)})$ est degree $< d$ donc $\det(X.\text{Id}_d - M)$ est unitaire de degree d .

On a

$$a_d = P(0) = \det(-M) = (-1)^d \det M.$$

Par ailleurs si $\sigma \neq \text{Id}$ soit i tel que $\sigma(i) = j \neq i$ alors $\sigma(j) \neq j$ (car σ est injective) et on a

$$(X\delta_{i\sigma(i)} - m_{i\sigma(i)})(X\delta_{j\sigma(j)} - m_{j\sigma(j)}) = m_{i\sigma(i)}m_{j\sigma(j)}$$

ainsi si $\sigma \neq \text{Id}_d$ le polynome $\prod_{i=1}^d (X\delta_{i\sigma(i)} - m_{i\sigma(i)})$ est de degree $\leq d-2$ et le terme de degree $d-1$ de $\det(X.\text{Id}_d - M)$ est celui de

$$\prod_{i=1}^d (X - m_{ii}) = X^d - (m_{11} + \cdots + m_{dd})X^{d-1} + \cdots.$$

□

THÉORÈME 11.2 (Proprietes fonctionnelles du polynome caracteristique). *Soient M, N des matrices, on a*

$$P_{car, {}^tM}(X) = P_{car, M}(X)$$

et

$$P_{car, MN}(X) = P_{car, NM}(X).$$

Ainsi pour tout $k \leq d$

$$a_k(M.N) = a_k(N.M)$$

et en particulier

$$\text{tr}(M.N) = \text{tr}(N.M).$$

Preuve: On a

$$P_{car, {}^tM}(X) = \det(X.\text{Id}_d - {}^tM) = \det({}^t(X.\text{Id}_d - M)) = \det(X.\text{Id}_d - M) = P_{car, M}(X).$$

On suppose d'abord que M est inversible. On a

$$\begin{aligned} P_{car, MN}(X) &= \det(X.\text{Id}_d - M.N) = \det(X.M.M^{-1} - M.N) \\ &= \det(M.(X.M^{-1} - N)) = \det((X.M^{-1} - N)M) = \det(X.\text{Id}_d - N.M). \end{aligned}$$

Soit T une autre indeterminee; on considere le corps $K' = K(T)$.

On peut faire des calculs dans ce corps de base K' qui contient K . Notons $M_T := M - T.\text{Id}_d \in M_d(K')$: c' est une matrice inversible car son determinant est un polynome de degre d en la variable T et est en particulier est non-nul. On a donc

$$\det(X.\text{Id}_d - M_T.N) = \det(X.\text{Id}_d - N.M_T).$$

Cet determinant est un polynome en T a coefficients dans $K[X]$ dont la valeur en $T = 0_K$ vaut (car $M_0 = M$)

$$\det(X.\text{Id}_d - M.N) = \det(X.\text{Id}_d - N.M).$$

□

THÉORÈME 11.3 (Invariance par conjugaison). *Le polynome caracteristique est un invariant de la classe de conjugaison de la matrice M : pour toute matrice inversible $P \in \text{GL}_d(K)$, on a*

$$P_{car,P.M.P^{-1}}(X) = P_{car,M}(X).$$

Preuve: On a

$$\begin{aligned} P_{car,P.M.P^{-1}}(X) &= \det(X.\text{Id}_d - P.M.P^{-1}) = \det(P.X.\text{Id}_d.P^{-1} - P.M.P^{-1}) \\ &= \det(P(X.\text{Id}_d - M).P^{-1}) = \det(X.\text{Id}_d - M) = P_{car,M}(X). \end{aligned}$$

□

COROLLAIRE 11.1. *Soient $(a_k(M))_{0 \leq k \leq d}$ les coefficients de $P_{car,M}(X)$:*

$$\det(X.\text{Id}_d - M) = X^d + a_{d-1}(M)X^{d-1} + \cdots + a_d(M)$$

(on a $a_d(M) = 1$).

Ces coefficients sont des invariants de la classe de conjugaison de M .

Autrement dit, pour toute matrice inversible $P \in \text{GL}_d(K)$ et $0 \leq k \leq d$

$$a_k(M) = a_k(P.M.P^{-1}).$$

REMARQUE 11.1.1. On retrouve ainsi que la trace d'une matrice ne depend que de la classe de conjugaison de celle-ci.

11.1.1. Exemple: la "matrice compagnon". On aura egalement besoin de la "matrice compagnon" qu'on a deja rencontre en seance d'exercices: soit un polynome unitaire de degre d ,

$$P(X) = X^d + b_{d-1}X^{d-1} + \cdots + b_0;$$

on note $\mathbf{b} = (b_0, \dots, b_{d-1}) \in K^d$ le vecteur de ces coefficients. La matrice compagnon de P est la matrice

$$M_P = M_{\mathbf{b}} = \begin{pmatrix} 0 & 0 & 0 & 0 & -b_0 \\ 1 & 0 & 0 & 0 & -b_1 \\ 0 & 1 & 0 & 0 & -b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & -b_{d-1} \end{pmatrix} \in M_d(K).$$

On a vu en exercice que

$$P(M_P) = M_P^d + b_{d-1}M_P^{d-1} + \cdots + b_0\text{Id}_d = \mathbf{0}_d.$$

Par exemple la matrice compagnon de $X^2 + 1$ est la matrice $I = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ qui sert a definir les nombres complexes et qui verifie

$$I^2 + \text{Id}_2 = \mathbf{0}_2.$$

PROPOSITION 11.1. *Soit*

$$P(X) = X^d + b_{d-1}X^{d-1} + \cdots + b_0 \in K[X]$$

un polynome et M_P la matrice compagnon associee au polynome P . Alors son polynome caracteristique est egal a P :

$$P_{car, M_P}(X) = \det(X \cdot \text{Id}_d - M_P) = P(X) = X^d + b_{d-1}X^{d-1} + \cdots + b_0.$$

PREUVE. (Exercice) On doit calculer

$$\det \begin{pmatrix} X & 0 & 0 & 0 & b_0 \\ -1 & X & 0 & 0 & b_1 \\ 0 & -1 & X & 0 & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & -1 & X + b_{d-1} \end{pmatrix}$$

Pour cela on echelonnera la matrice par une suite d'operations de type (III) (dans le corps $K(X)$ des fractions rationnelles) pour la rendre triangulaire superieure. $\square$

11.1.2. Cas des matrices triangulaires par blocs.

PROPOSITION 11.2. *Supposons que la matrice $M \in M_d(K)$ s'ecrive sous forme triangulaire superieure par blocs:*

$$M = \begin{pmatrix} M_1 & * \\ \mathbf{0} & M_2 \end{pmatrix}, \quad M_1 \in M_{d_1}(K), \quad M_2 \in M_{d_2}(K), \quad d_1 + d_2 = d$$

alors

$$P_{car, M}(X) = P_{car, M_1}(X) P_{car, M_2}(X)$$

Preuve: Exercice. $\square$

En iterant on obtient

COROLLAIRE 11.2. *soit $k \geq 2$ un entier, si M est une matrice triangulaire superieure a k blocs*

$$M = \begin{pmatrix} M_1 & * & * \\ \mathbf{0} & \ddots & * \\ \mathbf{0} & \mathbf{0} & M_k \end{pmatrix}, \quad M_i \in M_{d_i}(K), \quad i \leq k, \quad d_1 + \cdots + d_k = d$$

on a

$$P_{car, M}(X) = P_{car, M_1}(X) \cdots P_{car, M_k}(X)$$

En particulier, si M est triangulaire superieure ($k = d$) –par exemple diagonale–

$$M = \begin{pmatrix} \lambda_1 & * & \cdots & \cdots \\ 0 & \lambda_2 & * & * \\ \vdots & 0 & \ddots & * \\ 0 & \cdots & \cdots & \lambda_d \end{pmatrix},$$

on a

$$P_{car, M}(X) = \prod_{i=1}^d (X - \lambda_i).$$

REMARQUE 11.1.2. Notons enfin que par invariance du polynome caracteristique par transposition le Corollaire reste vrai pour une matrice triangulaire inferieure par blocs.

11.2. Le polynome caracteristique d'un endomorphisme

L'invariance par conjugaison du polynome caracteristique permet de definir le polynome caracteristique d'une application lineaire:

DÉFINITION 11.2. Soit $\varphi \in \text{End}(V)$ une application lineaire, on definit son polynome caracteristique par

$$P_{car,\varphi}(X) = P_{car,M}(X)$$

ou $M = \text{mat}_{\mathcal{B}}(\varphi)$ est la matrice de φ dans une base quelconque de V .

Notons que cette definition ne depend pas de la base $\mathcal{B}$ choisie: si $M' = \text{mat}_{\mathcal{B}'}(\varphi)$ est la matrice de φ dans une autre base alors par la formule de changement de base

$$M' = \text{mat}_{\mathcal{B}'} \mathcal{B}.M.\text{mat}_{\mathcal{B}'}^{-1}$$

et

$$P_{car,M'}(X) = P_{car,M}(X) = P_{car,\varphi}(X).$$

En particulier les coefficient $a_k(\varphi) = a_k(M)$ du polynome caracteristique ne dependent pas du choix de la base.

DÉFINITION 11.3. On definit la trace de φ comme etant la trace de M

$$\text{tr}(\varphi) = \text{tr}(M) = m_{11} + \cdots + m_{dd}$$

et cette definition ne depend pas du choix de la base $\mathcal{B}$.

PROPOSITION 11.3. Le polynome caracteristique $P_{car,\varphi}(X)$ ne depend que de la classe de conjugaison de φ dans $\text{End}(V)$: pour tout $\psi \in \text{GL}(V)$

$$P_{car,\psi.\varphi.\psi^{-1}}(X) = P_{car,\varphi}(X).$$

11.2.1. Sous-espaces propres. L'interet du polynome caracteristique est qu'il permet d'identifier des sous-espaces interessants de V relativement a φ :

THÉORÈME 11.4. Soit $P_{car,\varphi}$ le polynome caracteristique d'une application lineaire φ . Les enonces suivants sont equivalents

- (1) Le scalaire $\lambda \in K$ est racine de $P_{car,\varphi}$: $P_{car,\varphi}(\lambda) = 0$.
- (2) Il existe $v \in V - \{0\}$ tel que $\varphi(v) = \lambda.v$

Preuve: On a les equivalences suivantes

- $P_{car,\varphi}(\lambda) = \det(\lambda.\text{Id}_V - \varphi) = 0$,
- $\lambda.\text{Id}_V - \varphi$ n'est pas inversible,
- $\lambda.\text{Id}_V - \varphi$ n'est pas injective,
- $\ker(\lambda.\text{Id}_V - \varphi) \neq \{0_V\}$,
- Il existe $v \in V - \{0_V\}$ tel que

$$0_V = (\lambda.\text{Id}_V - \varphi)(v) = \lambda.v - \varphi(v).$$

□

DÉFINITION 11.4. Soit $\lambda \in K$, le sous-espace

$$V_{\varphi,\lambda} := \ker(\varphi - \lambda.\text{Id}_V) = \{v \in V, \varphi(v) = \lambda.v\}$$

est appelle sous-espace propre associe a λ . Si $V_{\varphi,\lambda} \neq \{0_V\}$ on dit que λ est une valeur propre de φ et tout vecteur non-nul de $V_{\varphi,\lambda}$ (ie. verifiant $\varphi(v) = \lambda.v$) est appelle vecteur propre de φ associe a la valeur propre λ .

L'ensemble des valeurs propres de φ est appelle le spectre de φ (dans K) et est note

$$\text{Spec}_{\varphi}(K).$$

Le Theoreme precedent dit ainsi que les racines dans K du polynome caracteristique sont exactement les valeurs propres de φ :

$$\text{Rac}_{P_{\text{car},\varphi}}(K) = \text{Spec}_{\varphi}(K).$$

Voici quelques proprietes de base des sous-espaces propres:

THÉORÈME 11.5. *Soit $\varphi \in \text{End}(V)$ et λ, λ' des valeurs propres de φ et $V_{\varphi,\lambda}$, $V_{\varphi,\lambda'}$ les sous-espaces propres associes.*

– *Le sous-espace $V_{\varphi,\lambda}$ est stable par φ :*

$$\varphi(V_{\varphi,\lambda}) \subset V_{\varphi,\lambda}.$$

– *Si $\lambda \neq \lambda'$ les sous-espaces $V_{\varphi,\lambda}$ et $V_{\varphi,\lambda'}$ sont en somme directe:*

$$V_{\varphi,\lambda} \cap V_{\varphi,\lambda'} = \{0_V\}.$$

Preuve: Soit $v \in V_{\varphi,\lambda}$, et $w = \varphi(v)$, on a

$$\varphi(w) = \varphi(\varphi(v)) = \varphi(\lambda.v) = \lambda.\varphi(v) = \lambda.w$$

et donc $w = \varphi(v) \in V_{\varphi,\lambda}$.

Soit $\lambda \neq \lambda'$ et $v \in V_{\varphi,\lambda} \cap V_{\varphi,\lambda'}$, on a

$$\varphi(v) = \lambda.v = \lambda'.v$$

et donc

$$(\lambda - \lambda').v = 0_V$$

mais comme $\lambda - \lambda' \neq 0_K$, on a $v = 0_V$. □

11.3. Le Theoreme de Cayley-Hamilton

Soit $K[X]$ l'algebre des polynomes sur un corps K , $(A, +, \cdot)$ une K -algebre et $\varphi \in A$ un element de cette algebre. Cette donnee permet de definir une application d' "evaluation en φ "

$$\begin{array}{ccc} K[X] & \mapsto & A \\ \text{ev}_{\varphi} : P(X) & \mapsto & P(\varphi) \end{array}$$

ou on a note

$$P(\varphi) = a_n.\varphi^n + a_{n-1}.\varphi^{n-1} + \cdots + a_0.1_A$$

pour $P(X)$ un polynome a coefficients dans K

$$P(X) = a_n.X^n + a_{n-1}.X^{n-1} + \cdots + a_0, \quad a_0, \dots, a_d \in K.$$

On rappelle que

$$\varphi^d := \varphi \cdot \cdots \cdot \varphi \quad (d \text{ fois si } d \geq 1), \quad \varphi^0 := 1_A.$$

On verifie facilement que

PROPOSITION 11.4. *L'application ev_{φ} est un morphisme de K -algebres:*

$$\text{ev}_{\varphi}(\lambda.P + Q) = \lambda.P(\varphi) + Q(\varphi) = \lambda.\text{ev}_{\varphi}(P) + \text{ev}_{\varphi}(Q)$$

$$\text{ev}_{\varphi}(P.Q) = P(\varphi).Q(\varphi) = \text{ev}_{\varphi}(P).\text{ev}_{\varphi}(Q).$$

Son image $\text{ev}_{\varphi}(K[X])$ est notee

$$K[\varphi] = \{a_n.\varphi^n + a_{n-1}.\varphi^{n-1} + \cdots + a_0.1_A, \quad n \geq 1, a_0, \dots, a_n \in K\} \subset A$$

est une sous-algebre commutative de A engendree comme K -ev par les puissance de φ :

$$\{1_A = \varphi^0, \varphi, \dots, \varphi^n, \dots\}.$$

REMARQUE 11.3.1. La commutativite resulte du fait que $K[X]$ est commutatif et donc

$$P(\varphi).Q(\varphi) = (P.Q)(\varphi) = (Q.P)(\varphi) = Q(\varphi).P(\varphi).$$

On va appliquer cette construction à l'algèbre des endomorphismes $(\text{End}_K(V), +, \circ)$ d'un K -EV de dimension d et $\varphi : V \mapsto V$ un endomorphisme et/ou à l'algèbre des matrices $(M_d(K), +, \cdot)$ pour une matrice $M \in M_d(K)$ et à l'algèbre des endomorphismes d'un espace vectoriel V , $(\text{End}(V), +, \circ)$ pour une application linéaire $\varphi \in \text{End}(V)$. À tout polynôme $P(X) \in K[X]$ on associe

$$\text{ev}_\varphi(P) = P(\varphi) = a_n \cdot \varphi^n + a_{n-1} \cdot \varphi^{n-1} + \cdots + a_0 \cdot \text{Id}_V \in \text{End}_K(V)$$

et

$$\text{ev}_\varphi(M) = P(M) = a_n \cdot M^n + a_{n-1} \cdot M^{n-1} + \cdots + a_0 \cdot \text{Id}_d \in M_d(K).$$

Notons que comme $\text{End}_K(V)$ et $M_d(K)$ sont de dimensions finies (égale à d^2) les noyaux $\ker \text{ev}_\varphi$ et $\ker \text{ev}_M$ sont non nuls: si on restreint ces applications au SEV des polynômes de degré $\leq d^2$, $K[X]_{\leq d^2}$, qui est de dimension $d^2 + 1$, on a par le Théorème noyau-Image

$$\dim \ker \text{ev}_\varphi + \dim_K(K[\varphi]) = \dim \ker \text{ev}_M + \dim_K(K[M]) = d^2 + 1$$

et comme

$$\dim_K(K[\varphi]), \dim_K(K[M]) \leq \dim \text{End}_K(V) = \dim M_d(K) = d^2$$

on a

$$\dim \ker \text{ev}_\varphi, \dim \ker \text{ev}_M \geq 1$$

et on peut trouver dans le noyau un polynôme non-nul de degré $\leq d^2$. En fait on peut trouver un polynôme de degré d :

THÉORÈME 11.6 (Cayley-Hamilton). *Soit $\varphi \in \text{End}(V)$ (resp. $M \in M_d(K)$) alors son polynôme caractéristique $P_{\text{car},\varphi}(X)$ (resp. $P_{\text{car},M}(X)$) appartient à $\ker \text{ev}_\varphi$ (resp. $\ker \text{ev}_M$); en d'autres termes*

$$P_{\text{car},\varphi}(\varphi) = 0_V, \quad P_{\text{car},M}(M) = 0_{d \times d}.$$

Preuve: Soit $\varphi : V \mapsto V$. Il s'agit de montrer que pour tout $v \in V - \{0\}$,

$$P_{\text{car},\varphi}(\varphi)(v) = 0_V.$$

Si $v = 0_V$ c'est évident. Sinon on considère la suite de vecteurs

$$v, \varphi(v), \varphi^2(v), \dots, \dots, \varphi^k(v), \dots$$

Comme V est de dimension finie il existe $d_1 \leq d$ tel que

$$v, \varphi(v), \varphi^2(v), \dots, \dots, \varphi^{d_1}(v)$$

est liée. Prenons $d_1 \geq 1$ le plus petit possible pour cette propriété de sorte que

$$\mathcal{B}_1 := \{v, \varphi(v), \varphi^2(v), \dots, \dots, \varphi^{d_1-1}(v)\}$$

est libre et il existe $b_0, \dots, b_{d_1-1} \in K$ tels que

$$\varphi^{d_1}(v) = b_0 \cdot v + \cdots + b_{d_1-1} \varphi^{d_1-1}(v).$$

Complétons la famille $\mathcal{B}_v$ en une base de V : $\mathcal{B} = \mathcal{B}_1 \sqcup \mathcal{B}_2$. Soit $M = \text{mat}_{\mathcal{B}}(\varphi)$ la matrice de φ dans cette base. Elle est de la forme

$$\begin{pmatrix} 0 & 0 & 0 & 0 & b_0 & * \\ 1 & 0 & 0 & 0 & b_1 & * \\ 0 & 1 & 0 & 0 & b_2 & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & b_{d_1-1} & * \\ & & & \mathbf{0} & & M_2 \end{pmatrix} = \begin{pmatrix} & & & & * \\ & & & & * \\ & & & & * \\ & M_1 & & & * \\ & & & & * \\ & & \mathbf{0} & & M_2 \end{pmatrix}$$

de sorte que

$$P_{\text{car},\varphi}(X) = P_{\text{car},M}(X) = P_{\text{car},M_1}(X)P_{\text{car},M_2}(X) = P_{\text{car},M_2}(X)P_{\text{car},M_1}(X)$$

La matrice M_1 est une matrice compaignon dont on connait le polynome caracteristique (cf. Prop 11.1)

$$P_{car, M_1}(X) = \det \begin{pmatrix} X & 0 & 0 & 0 & -b_0 \\ -1 & X & 0 & 0 & -b_1 \\ 0 & -1 & X & 0 & -b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & -1 & X - b_{d_1-1} \end{pmatrix} = X^{d_1} - b_{d_1-1}X^{d_1-1} - \dots - b_0$$

et

$$P_{car, \varphi}(\varphi)(v) = P_{car, M_2}(\varphi) \circ P_{car, M_1}(\varphi)(v) = P_{car, M_2}(\varphi)(P_{car, M_1}(\varphi)(v)) = 0_V$$

car

$$P_{car, M_1}(\varphi)(v) = \varphi^{d_1}(v) - b_{d_1-1}\varphi^{d_1-1}(v) - \dots - b_0v = 0_V$$

□

REMARQUE 11.3.2. Dans cette preuve on a implicitement utilise le fait que l'on connaissait deja le Theoreme pour les matrices compaignons (Prop 11.1 et la remarque qui suit).

COROLLAIRE 11.3. Soit φ un endomorphisme et M sa matrice associee dans une base quelconque. Si $\det(\varphi) = \det(M) \neq 0$ alors φ et M sont inversibles et on a

$$\begin{aligned} \varphi^{-1} &= \frac{(-1)^{d+1}}{\det \varphi} (a_1 \text{Id}_V + \dots + a_{d-1} \varphi^{d-2} + \varphi^{d-1}) \\ M^{-1} &= \frac{(-1)^{d+1}}{\det M} (a_1 \text{Id}_d + \dots + a_{d-1} M^{d-2} + M^{d-1}) \end{aligned}$$

ou

$$P_{car, \varphi}(X) = P_{car, M}(X) = a_0 + a_1 X + \dots + a_{d-1} X^{d-1} + X^d.$$

En particulier $\varphi^{-1} \in K[\varphi]$ et $M^{-1} \in K[M]$.

Preuve: On a

$$\mathbf{0}_d = a_0 \text{Id}_d + a_1 M + \dots + a_{d-1} M^{d-1} + M^d$$

de sorte que

$$-a_0 \text{Id}_d = a_1 M + \dots + a_{d-1} M^{d-1} + M^d = M \cdot (a_1 \text{Id}_V + \dots + a_{d-1} M^{d-2} + M^{d-1})$$

et si $a_0 = (-1)^d \det(M) \neq 0$, on a

$$\text{Id}_d = M \cdot \frac{-1}{a_0} (a_1 \text{Id}_d + \dots + a_{d-1} M^{d-2} + M^{d-1})$$

ce qui montre que M est inversible.

□

APPENDICE A

L'anneau des polynomes sur un corps

”Trois anneaux pour les rois Elfes sous le ciel,
 $B_{\text{crys}}, B_{\text{st}}, B_{\text{dR}},$
Sept pour les Seigneurs Nains dans leurs demeures de pierre,
 $E_{\mathbb{Q}_p}, A_{\mathbb{Q}_p}, B_{\mathbb{Q}_p}, E, A, B, \hat{A}$
Neuf pour les Hommes Mortels destinés au trépas,
 $\mathbb{Q}_p, \mathbb{Z}_p, \mathbb{F}_p, \overline{\mathbb{Q}_p}, \overline{\mathbb{F}_p}, \mathbb{C}_p, \mathcal{O}_{\mathbb{C}_p}, \mathbb{Q}_p^{nr}, B_{\text{HT}}$
Un pour le Seigneur Ténébreux sur son sombre trône
 A_{inf} ”

Dans ce chapitre on donne la construction algebrique des polynomes a coefficients dans un anneau commutatif A (et en particulier quand $A = K$ est un corps). On rappellera ensuite la terminologie et les proprietes de base concernant polynomes (degre, monomes, division euclidienne, factorisation, polynomes irreductibles, racines). on appliquera la theorie a la construction de sous-algebres dans des algebres sur un corps (algebres monogenes)

A.1. Preliminaire: fonctions polynomiales

Sur le corps des nombres reels $\mathbb{R}$, on a l'habitude de definir un polynome comme etant une fonction de $\mathbb{R}$ a valeurs dans $\mathbb{R}$ de la forme

$$P(\bullet) : x \in \mathbb{R} \mapsto P(x) = a_d x^d + a_{d-1} x^{d-1} + \cdots + a_0 \in \mathbb{R}$$

ou $a_0, \dots, a_d$ sont des reels fixes (les coefficients du polynome) et si $a_d \neq 0$ on dit que P est un polynome de degre $\deg P = d$. La fonction identiquement nulle $\underline{0}$ est egalement une fonction polynomiale correspondant a $a_d = \cdots = a_0 = 0$ et on declare que

$$\deg \underline{0} = -\infty.$$

De plus, on sait que la somme et le produit de deux fonctions polynomiales sont des fonctions polynomiales: si P et Q sont des fonctions polynomiales, on peut toujours les ecrire sous la forme

$$P(x) = a_d x^d + a_{d-1} x^{d-1} + \cdots + a_0, \quad Q(x) = b_d x^d + b_{d-1} x^{d-1} + \cdots + b_0$$

(avec $d = \max(\deg P, \deg Q)$ et en posant $a_d = \cdots = a_{\deg Q} = 0$ ou $b_d = \cdots = b_{\deg P} = 0$ si $\deg P \neq \deg Q$) et on a

$$x \mapsto (P + Q)(x) = (a_d + b_d)x^d + (a_{d-1} + b_{d-1})x^{d-1} + \cdots + (a_0 + b_0)$$

et

$$\begin{aligned}
 P.Q(\bullet) : x \mapsto P.Q(x) &= (a_d x^d + a_{d-1} x^{d-1} + \cdots + a_0).(b_d x^d + b_{d-1} x^{d-1} + \cdots + b_0) \\
 &= c_{2d} x^{2d} + c_{2d-1} x^{2d-1} + \cdots + c_0
 \end{aligned}$$

avec

$$c_n = \sum_{p+q=n} a_p.b_q = \sum_{q+p=n} b_q.a_p, \quad 0 \leq n \leq 2d.$$

On a alors

$$\deg(P + Q) \leq \max(\deg P, \deg Q), \quad \deg(P.Q) = \deg(P) + \deg(Q)$$

REMARQUE A.1.1. Cette dernière formule reste vraie si P ou $Q = 0$ car on a pose $\deg 0 = -\infty$.

L'ensemble des fonctions polynomiales sur $\mathbb{R}$ forme alors un anneau commutatif que l'on note $\mathbb{R}[X]$ dont le nul est le polynome nul et l'unité le polynome constant égal à 1.

De plus $\mathbb{R}[x]$ a une structure $\mathbb{R}$ -module via la multiplication des polynomes par les polynomes constants:

$$(a, P) \in \mathbb{R} \times \mathbb{R}[X] \mapsto a.P : x \mapsto aa_d x^d + aa_{d-1} x^{d-1} + \cdots + aa_0.$$

Ainsi $\mathbb{R}[X]$ est une $\mathbb{R}$ -algebre.

On pourrait faire de même pour tout anneau commutatif A en définissant l'anneau des polynomes $A[X]$ comme étant l'ensemble des fonction polynomiales de A vers A c'est à dire les fonctions de la forme

$$P : x \in A \mapsto P(x) = a_d x^d + a_{d-1} x^{d-1} + \cdots + a_0$$

ou $a_0, \dots, a_d \in A$ sont des éléments de A fixes. On voit de même que la somme et le produit de deux fonctions polynomiales sont polynomiales et l'ensemble des fonctions polynomiales est un sous-anneau commutatif de l'anneau des fonctions de A vers A . Cependant dans certains cas, on rencontre des problèmes avec une telle définition: une même fonction polynomiale peut avoir des expressions différentes, ainsi les notions de coefficients d'un polynome ou de degré ne sont pas bien définies:

Prenons $A = \mathbb{F}_p$ pour p premier le corps à p éléments. On a vu que pour tout $x \in \mathbb{F}_p$ on a

$$x^p = x$$

et en d'autres termes la fonction polynomiale identiquement nulle est également donnée par la fonction

$$x \in \mathbb{F}_p \mapsto x^p - x.$$

Cette absence d'unicité pose notamment des problèmes quand on considère l'extension suivante: soit $B \supset A$ un autre anneau commutatif contenant A alors une expression polynomiale sur A

$$P : x \in A \mapsto P(x) = a_d x^d + a_{d-1} x^{d-1} + \cdots + a_0 \in A$$

défini une fonction polynomiale sur B en posant

$$P : x \in B \mapsto P(x) = a_d x^d + a_{d-1} x^{d-1} + \cdots + a_0 \in B$$

et il se peut qu'une fonction polynomiale identiquement nulle sur A ne le soit pas sur B . Par exemple, si $A = \mathbb{F}_p$ et $B = \mathbb{F}_p[I_d]$ le corps à p^2 éléments construit en exercices il existe $x \in \mathbb{F}_p[I_d]$ tel que

$$x^p - x \neq 0_{\mathbb{F}_p[I_d]}.$$

Ainsi pour définir les polynomes on va devoir le faire à partir de leur expression polynomiale abstraite

$$P(x) = a_d x^d + a_{d-1} x^{d-1} + \cdots + a_0.$$

A.2. Les polynomes sont des suites

Soit A un anneau commutatif et soit

$$A^{\mathbb{N}} = \{(a_n)_{n \geq 0}, a_n \in A\}.$$

l'ensemble des suites à valeurs dans A (ou encore l'ensemble des fonctions de $\mathbb{N}$ à valeurs dans A , $(a_n)_{n \geq 0} : n \mapsto a_n$). L'ensemble $A^{\mathbb{N}}$ a une structure de A -module pour l'addition terme à terme

$$(a_n)_{n \geq 0} + (b_n)_{n \geq 0} = (a_n + b_n)_{n \geq 0}$$

dont l'élément neutre est la suite identiquement nulle

$$\underline{0}_A = (0_A, \dots, 0_A, \dots)$$

et la multiplication par les scalaires est donnée pour $a \in A$ par

$$a.(a_n)_{n \geq 0} = (a.a_n)_{n \geq 0}.$$

DÉFINITION A.1. Soit $(a_n)_{n \geq 0} \in A^{\mathbb{N}}$ une suite à valeurs dans A . Le support de cette suite est défini comme étant l'ensemble des indices où la suite prend une valeur non-nulle

$$\text{supp}((a_n)_{n \geq 0}) = \{n \in \mathbb{N}, a_n \neq 0_A\} \subset \mathbb{N}.$$

L'ensemble des polynômes $A[X]$ est construit algébriquement de la manière suivante:

DÉFINITION A.2. Un polynôme P à coefficient dans A est une suite

$$P = (a_n)_{n \geq 0}$$

de support fini: telle que

$$\text{supp}(P) = \{n \in \mathbb{N}, a_n \neq 0_A\} \text{ est fini.}$$

Le n -ième terme de cette suite a_n est le coefficient d'ordre n de P ; on le note également $c_n(P)$.

L'ensemble des polynômes à coefficients dans A est le sous-ensemble $A_f^{\mathbb{N}} \subset A^{\mathbb{N}}$ forme des suites à support fini; on le note

$$A_f^{\mathbb{N}} = \{(a_n)_{n \geq 0}, a_n \in A, |\text{supp}((a_n)_{n \geq 0})| < \infty\}.$$

PROPOSITION A.1. L'ensemble $A_f^{\mathbb{N}}$ est un sous- A module de $A^{\mathbb{N}}$ pour l'addition et la multiplication par les scalaire sur l'espaces des suites.

Preuve: Rappelons que si $\mathbf{a} = (a_n)_{n \geq 0}$, et $\mathbf{b} = (b_n)_{n \geq 0}$ sont des suites et $a \in A$, l'addition est définie par

$$\mathbf{a} + \mathbf{b} := (a_n + b_n)_{n \geq 0}$$

et la multiplication par a est définie par

$$a.\mathbf{a} := (a.a_n)_{n \geq 0}.$$

On a

$$a_n + b_n \neq 0_A \implies a_n \neq 0_A \text{ ou } b_n \neq 0_A$$

et

$$a.a_n \neq 0_A \implies a_n \neq 0_A$$

et donc

$$\text{supp}(\mathbf{a} + \mathbf{b}) \subset \text{supp}(\mathbf{a}) \cup \text{supp}(\mathbf{b}), \text{supp}(a.\mathbf{a}) \subset \text{supp}(\mathbf{a}).$$

Ainsi, si $\mathbf{a}$ et $\mathbf{b}$ sont à supports finis alors $\mathbf{a} + \mathbf{b}$ et $a.\mathbf{a}$ sont à supports finis et ainsi $A_f^{\mathbb{N}}$ est un sous A -module de $A^{\mathbb{N}}$. $\square$

A.2.1. Degré d'un polynôme. Un sous-ensemble de $\mathbb{N}$ est fini ssi il possède un plus grand élément:

DÉFINITION A.3. Le degré d'un polynôme non-nul $P = (a_n)_{n \geq 0}$ est le plus grand élément de $\text{supp}(P)$:

$$\deg(P) = \max\{d \geq 0, a_d \neq 0\}.$$

Si $P = 0_K$ est le polynôme nul, le support de P est l'ensemble vide et on définit son degré comme étant

$$\deg(0_K) = -\infty.$$

DÉFINITION A.4. Étant donné un polynôme de degré $\leq d$

$$P = (a_0, \dots, a_d, 0, \dots)$$

le d -ième coefficient a_d est appelé coefficient dominant de P . Un polynôme non-nul est unitaire si le coefficient de degré $\deg P$ vérifie

$$a_{\deg P} = 1.$$

PROPOSITION A.2. Soient P, Q des polynômes, on a

$$\deg(P + Q) \leq \max(\deg P, \deg Q)$$

avec égalité si $\deg P \neq \deg Q$.

Preuve: C'est evident si P ou $Q = 0$.

Sinon soit $d = \deg P \geq d' = \deg Q$, on a

$$P = (a_0, a_1, \dots, a_d, 0, \dots), \quad Q = (b_0, b_1, \dots, b_{d'}, 0, \dots)$$

avec $a_d, b_{d'} \neq 0$.

Supposons $d' \geq d$, on a

$$P + Q = (a_0 + b_0, a_1 + b_1, \dots, a_d + b_d, 0 + b_{d+1}, \dots, 0 + b'_{d'}, 0, \dots)$$

et $\deg(P + Q) \leq d'$ (avec egalite ssi $d = d'$ et $a_{d'} + b_{d'} \neq 0$). $\square$

COROLLAIRE A.1. Soit $d \geq 0$ et

$$A_{f \leq d}^{\mathbb{N}} = \{P \in A_f^{\mathbb{N}}, \deg P \leq d\}$$

l'ensemble des polynomes de degre $\leq d$. Alors $A_{f \leq d}^{\mathbb{N}}$ est un sous A -module de $A_f^{\mathbb{N}}$.

A.2.2. La famille des monomes unitaires. On va maintenant identifier une famille particuliere de polynomes:

NOTATION A.1. Soit $k \geq 0$ un entier, on a note X^k le polynome (ie la suite de support fini) defini par

$$X^k := (\delta_{n=k})_{n \geq 0}$$

avec $(\delta_{n=k})$ le symbole de Kronecker

$$\delta_{n=k} = \begin{cases} 1_K & \text{si } n = k \\ 0_K & \text{sinon.} \end{cases}$$

Le polynome X^k est appelle monome unitaire de degre k .

On note l'ensemble des monomes unitaires

$$\mathcal{M} = \{X^k, k \geq 0\} \subset A[X].$$

EXEMPLE A.2.1. Le monome X^d est de degre d .

Avec cet notation on a pour tout polynome $P = (a_n)_{n \geq 0}$ non nul de degre d

$$\begin{aligned} P &= (a_0, a_1, \dots, a_d, 0, 0, \dots, 0, \dots) \\ &= a_0(1, 0, \dots, 0) + a_1(0, 1, 0, \dots, 0) + \dots + a_d(0, \dots, 1, 0, \dots) \\ &= a_0.X^0 + a_1.X^1 + \dots + a_d.X^d \end{aligned}$$

et plus generalement on a le theoreme suivant qu'on ne montrera pas

THÉORÈME A.1. La famille des monomes $\mathcal{M}$ engendre $A_f^{\mathbb{N}}$ comme A -module: tout polynome se decompose en combinaison lineaire (a coefficient dans A) de monomes: pour tout $P \in A_f^{\mathbb{N}}$ il existe $d \geq 0$ et $a_0, \dots, a_d \in A$ tels que

$$P = a_0.X^0 + a_1.X^1 + \dots + a_d.X^d.$$

De plus, cette decomposition est unique: si

$$P = a_0.X^0 + a_1.X^1 + \dots + a_d.X^d = a'_0.X^0 + a'_1.X^1 + \dots + a'_{d'}.X^{d'}$$

avec $d \leq d'$ alors pour tout $k \leq d$ on a $a_k = a'_k$ et pour $d < k \leq d'$ on a $a'_k = 0_K$.

La famille des monomes unitaires est aussi appelee base canonique de l'espace des polynomes.

NOTATION A.2. On notera l'espace des polynomes

$$A[X] := A_f^{\mathbb{N}}$$

et

$$A[X]_{\leq d} = \{P \in A[X], \deg P \leq d\}$$

le sous A -module des polynomes de degre $\leq d$.

On notera egalement quelquefois un polynome $P(X)$ au lieu de P .

Alors le theoreme precedent dit que l'application

$$(a_0, \dots, a_d) \in A^{d+1} \mapsto a_d X^d + \dots + a_0 X^0 \in A[X]_{\leq d}$$

est un isomorphisme de A -module et $A[X]_{\leq d}$ est libre de rang $d+1$.

A.3. Structure d'anneau

A.3.1. Fonction polynomiale associee a un polynome. Armes de la notion abstraite de polynome et de la notation monomiale on peut associer une fonction polynomiale a un polynome:

DÉFINITION A.5. Soit A un anneau commutatif et

$$P = a_d.X^d + a_{d-1}.X^{d-1} + \dots + a_1.X^1 + a_0.X^0$$

un polynome a coefficient dans A . La fonction polynomiale associee a P est la fonction

$$P(\bullet) : A \mapsto A$$

definie par

$$P(\bullet) : x \in A \mapsto P(x) := a_d.x^d + a_{d-1}.x^{d-1} + \dots + a_1.x + a_0 \in A.$$

PROPOSITION A.3. L'application "fonction polynomiale"

$$P \in A[X] \mapsto P(\bullet) \in \mathcal{F}(A, A)$$

est un morphisme de A -modules pour la structure naturelle de A -module sur l'espaces des fonctions de A vers A : on a

$$(P + Q)(\bullet) = P(\bullet) + Q(\bullet)$$

et pour $a \in A$

$$(a.P)(\bullet) = a.P(\bullet).$$

Par ailleurs, l'espace $\mathcal{F}(A, A)$ possede egalement une structure d'anneau (et meme de A -algebre) donnee par pour $f, g \in \mathcal{F}(A, A)$ et $\lambda \in A$

$$(f.g) : x \in A \mapsto f(x).g(x) \in A, (\lambda.f) : x \in A \mapsto \lambda.f(x).$$

PROPOSITION A.4. Soit $d \geq 1$ et P et Q deux polynomes de degre $\leq d$

$$P = a_d.X^d + a_{d-1}.X^{d-1} + \dots + a_1.X^1 + a_0.X^0, Q = b_d.X^d + b_{d-1}.X^{d-1} + \dots + b_1.X^1 + b_0.X^0,$$

alors le produit de leur fonctions polynomiales,

$$P(\bullet).Q(\bullet) : x \in A \mapsto P(x).Q(x)$$

est encore une fonction polynomiale: C'est la fonction associee au polynome

$$P.Q = c_{2d}.X^{2d} + \dots + c_1.X + c_0$$

ou pour $n \leq 2d$,

$$c_n = \sum_{p+q=n} a_p.b_q = a_0.b_n + a_1.b_{n-1} + \dots + a_n.b_0.$$

Preuve: Pour tout $x \in A$, on a (utilisant la distributivite, l'associativite et la commutativite de A)

$$\begin{aligned} P(x).Q(x) &= (a_0. + a_1.x + \cdots + a_d.x^d).(b_0 + b_1.x + \cdots + b_d.x^d) = \\ &= \sum_{p,q \leq d} a_p.X^p.b_q.X^q = \sum_{p,q \leq d} a_p.b_q.x^{p+q} = \sum_{n \leq 2d} \left(\sum_{p+q=n} a_p.b_q \right) x^n = \sum_{n \leq 2d} c_n.x^n \end{aligned}$$

□

A.3.2. Multiplication abstraite des polynomes. La proposition precedente motive l'introduction de la loi de multiplication interne sur $A[X]$: on defini le produit de polynomes

$$\bullet \bullet : \begin{aligned} A[X] \times A[X] &\mapsto A^{\mathbb{N}} \\ (P = (a_n)_{n \geq 0}, Q = (b_n)_{n \geq 0}) &\mapsto P.Q = (c_n)_{n \geq 0} \end{aligned}$$

avec

$$c_n = \sum_{p+q=n} a_p.b_q = a_0.b_n + a_1.b_{n-1} + \cdots + a_n.b_0.$$

Notons que si les suites $P = (a_n)_{n \geq 0}$ et $Q = (b_n)_{n \geq 0}$ sont a support fini, alors $P.Q$ est a support fini, plus precisement

PROPOSITION A.5. *Soient P, Q des polynomes, alors $P.Q$ est un polynome de degree*

$$\deg(P.Q) \leq \deg P + \deg Q.$$

Preuve: Si P ou $Q = (0_A)_{n \geq 0}$ alors $P.Q = (0_A)_{n \geq 0}$ et compte-tenu du fait que $\deg 0_A = -\infty$ on a bien

$$\deg(P.Q) = -\infty = \deg P + \deg Q.$$

Si P et Q sont non-nuls, on a pour $n > \deg P + \deg Q$

$$c_n = \sum_{p+q=n} a_p.b_q = 0_A$$

car si $p + q = n > \deg P + \deg Q$ ou bien $p > \deg P$ et $a_p = 0$ ou bien $q > \deg Q$ et $b_q = 0$. Ainsi $P.Q$ est a support fini et de degree $\leq \deg P + \deg Q$. □

On verifie alors (exercice)

THÉORÈME A.2. *La loi de multiplication interne $\bullet \bullet$ sur $A[X]$ est associative, commutative et distributive par rapport a l'addition et fait de $(A[X], +, \cdot)$ un anneau commutatif dont l'element unite est le monome unitaire de degree 0,*

$$X^0 = (1_A, 0, \dots).$$

Par ailleurs $A[X]$ muni de la multiplication externe $(a, P) \mapsto a.P$ fait de $A[X]$ une A -algebre.

A.3.3. Retour sur les fonctions polynomiales. L'interet d'avoir defini l'addition et la multiplication des polynomes comme on l'a fait est la proposition suivante:

PROPOSITION A.6. *Soit $\mathcal{F}(A; A)$ l'espace des fonctions de A a valeurs dans A : L'application "fonction polynomiale"*

$$P \in A[X] \mapsto P(\bullet) \in \mathcal{F}(A; A)$$

qui a un polynome associe sa fonction polynomiale est un morphisme d'anneaux.

En particulier si $P = a_0 X^0$ est un polynome de degree 0 ou $-\infty < \deg P$ la fonction correspondante est la fonction constante egale a $a_0 \in A$

$$a_0 X^0(\bullet) = \underline{a_0} : x \mapsto a_0.$$

NOTATION A.3. Un polynome de degre 0 ou $-\infty$, $a_0.X^0$ sera appelle "polynome constant" (de valeur a_0). L'application "polynome constant"

$$a \in A \mapsto aX^0 \in A[X]_{\leq 0} \subset A[X]$$

identifie A avec l'anneau des polynomes constant et pour simplifier les notations on ecrira a_0 au lieu de $a_0.X^0$. En particulier on ecrira $1 = 1_a$ au lieu de X^0 .

De meme on ecrira X a la place du monome X^1 .

Le coefficient $a_0(P)$ de degre 0 d'un polynome P est appele coefficient constant de P . On a la formule

$$a_0(P) = P(0).$$

REMARQUE A.3.1. Notons qu'en general l'application "fonction polynomiale" n'est PAS injective: par exemple si $A = \mathbb{F}_p$ est le corps fini a p elements, la fonction polynomiale sur $\mathbb{F}_p$ associee au polynome $X^p - X$ est la fonction identiquement nulle: on a vu que $\forall x \in \mathbb{F}_p$, on a

$$x^p - x = 0_{\mathbb{F}_p}.$$

On va analyser plus tard quand cette application est injective (et donc quand on peut identifier l'algebre des polynomes a l'algebre des fonctions polynomiales).

A.3.4. Fonction polynomiales sur une A -algebre. Soit $(\mathcal{A}, +, \cdot)$ une A -algebre (pas forcément commutative) d'unité $1_{\mathcal{A}}$. On associe a tout polynome a coefficients dans A , $P(X) \in A[X]$ une fonction (polynomiale) de $\mathcal{A}$ vers $\mathcal{A}$ en posant

$$P(\bullet) : M \in \mathcal{A} \mapsto P(M) = a_d.M^d + \cdots + a_1.M + a_0.1_{\mathcal{A}}.$$

On a alors

$$(P + Q)(M) = P(M) + Q(M), (P.Q)(M) = P(M).Q(M), (a.P)(M) = a.P(M)$$

autrement dit

$$P \in A[X] \mapsto P(\bullet) \in \mathcal{F}(\mathcal{A}, \mathcal{A})$$

est un morphisme de A -algebre dont l'image est l'ensemble des fonctions polynomiales sur $\mathcal{A}$.

A.3.5. Derivation formelle. Sur l'espace des fonctions de $\mathbb{R}$ vers $\mathbb{R}$ on a la notion de derivee d'une fonction obtenue a partir de la notion de limite (limite d'un taux d'accroissement) et on sait que la derivee d'une fonction polynomiale est polynomiale: si

$$P(X) = a_d.X^d + \cdots + a_1.X + a_0 \in \mathbb{R}[X]$$

alors pour tout $x \in \mathbb{R}$ on a

$$\lim_{h \rightarrow 0} \frac{P(x+h) - P(x)}{h} = P'(x) = a_d.(d-1).X^{d-1} + \cdots + a_k.k.x^{k-1} + \cdots + a_1$$

est donc une fonction polynomiale (de degre $\leq \deg P - 1$).

On peut definir la derivation des polynomes sur un anneau de maniere purement formelle:

DÉFINITION A.6. Soit

$$P(X) = a_d.X^d + \cdots + a_1.X + a_0 \in A[X]$$

un polynome a coefficient dans un anneau commutatif A ; son polynome derive est le polynome

$$P'(X) = a_d.(d-1).X^{d-1} + \cdots + a_k.k.X^{k-1} + \cdots + a_1 \in A[X].$$

Ici on a note

$$a_2.2 = a_2.2_A = a_2 + a_2 \text{ (2 fois)}, a_d.d = a_d.d_A = a_d + \cdots + a_d \text{ (d fois)}$$

ou

$$d_A = 1_A + \cdots + 1_A \text{ (d fois)}$$

est l'image de d par le morphisme canonique de $\mathbb{Z}$ vers A .

THÉORÈME A.3. *La derivation*

$$\bullet' : P \in A[X] \mapsto P' \in A[X]$$

– *est linéaire:*

$$\forall a \in A, P, Q \in A[X], (a.P + Q)' = a.P' + Q'$$

et son noyau contient les polynomes constants.

– *verifie la regle de Leibnitz:*

$$\forall P, Q \in A[X], (P.Q)' = P'.Q + P.Q'.$$

Preuve: Exercice. □

REMARQUE A.3.2. En general la derivation n'annule pas que les polynomes constants: si d est tel que $d_A = 0_A$ (si d est contenu dans le noyau du morphisme canonique: par exemple si A est un corps et $d = \text{car}K$) on a

$$(X^d)' = d_A.X^{d-1} = 0_A.$$

On a

$$\ker(\bullet') = \{P \in A[X], \text{supp}(P) \subset \ker(\text{Can}_A)\}.$$

Si K est un corps de caracteristique nulle

$$\ker(\bullet') = \{a_1, a_1 \in K\}.$$

A.3.6. Integralite de $A[X]$ et corps des fractions.

PROPOSITION A.7. *L'anneau $A[X]$ est integre ssi A est integre et on a alors pour tout $P, Q \in A[X]$,*

$$\deg(P.Q) = \deg P + \deg Q.$$

Preuve: Si A n'est pas integre alors $A[X]$ ne l'est pas: soient $a, b \in A$ tels que $a.b = 0_A$ alors le produit des polynomes constants (de degre ≤ 0) a et b vaut le polynome constant $a.b = 0_A$.

Supposons que A est integre et soient P et Q tous deux non-nuls et $(c_n)_{n \geq 0}$ les coefficients de $P.Q$: alors pour $n = \deg P + \deg Q$, on a

$$c_n = \sum_{p+q=\deg P+\deg Q} a_p.b_q = a_{\deg P}.b_{\deg Q}$$

car $p \leq \deg P$ et $q \leq \deg Q$. Par definition du degre $a_{\deg P}, b_{\deg Q} \neq 0_A$ et comme A est integre

$$a_{\deg P}.b_{\deg Q} \neq 0_A.$$

Ainsi $\deg P.Q \geq \deg P + \deg Q$ et donc $\deg P.Q = \deg P + \deg Q$. □

PROPOSITION A.8. *Si A est integre de corps des fraction K , alors le corps des fractions de l'anneau integre $A[X]$ est egal au corps des fractions de l'anneau des polynomes a coefficients dans $K[X]$: on a*

$$\begin{aligned} \text{Frac}(A[X]) &= \{F(X) = \frac{P(X)}{Q(X)}, P, Q \in A[X], Q \neq 0\} \\ &= \{F(X) = \frac{P(X)}{Q(X)}, P, Q \in K[X], Q \neq 0\} = \text{Frac}(K[X]). \end{aligned}$$

On l'appelle le corps des fractions rationnelles a coefficients dans K .

A.4. Division et factorisation

On suppose maintenant et dans toute la suite que $A = K$ est un corps.

A.4.1. Relation de divisibilité. comme tout anneau $K[X]$ est muni d'une relation de divisibilité: on dit que Q divise P et on le note

$$Q|P$$

si il existe S tel que

$$P = Q.S.$$

On dit alors que S est le quotient de P par Q . Notons que la relation de divisibilité est

- Reflexive: $\forall Q \in K[X]$, on a $Q|Q$.
- Transitive: $Q|P$ et $P|L \implies Q|L$.
- $\forall P$ on a $1|P$ et $P|0$.

A.4.2. Division euclidienne. On sait que l'espace des polynome $\mathbb{R}[X]$ a coefficient reels admet une division euclidienne; cette division se generalise a $K[X]$ pour K un corps arbitraire:

THÉORÈME A.4. Soit $Q \in K[X] - \{0\}$ un polynome non-nul. Pour tout $P \in K[X]$ il existe des polynomes $S, R \in K[X]$ uniques verifiant

$$\deg R < \deg Q \text{ et tels que } P = Q.S + R.$$

DÉFINITION A.7. Les polynomes R et S sont appeles respectivement "reste" et "quotient" de la division euclidienne de P par Q .

De plus $R = 0$ si et seulement si $Q|P$.

Preuve: Soit $q = \deg Q$:

$$Q = b_q.X^q + \dots + b_1.X + b_0, \quad b_q \neq 0.$$

Ecrivons

$$P = a_d.X^d + \dots + a_0.$$

Si $d < q$, on prend $R = P$ et $S = 0$. Sinon, on procede par recurrence sur d :

$$P_1 := P - \frac{a_d}{b_q}Q.X^{d-q} = a_d.X^d - \frac{a_d}{b_q}b_q.X^d.X^{d-q} + \text{polynome de degree } \leq d-1$$

et comme

$$a_d.X^d - \frac{a_d}{b_q}b_q.X^d.X^{d-q} = 0$$

Le polynome P_1 est de degree $\leq d-1$. Par recurrence sur le degree il existe R_1, S_1 tels que

$$P_1 = Q.S_1 + R_1$$

avec $\deg R_1 < q$ et donc

$$P = \frac{a_d}{b_q}Q.X^{d-q} + Q.S_1 + R_1 = Q.S + R$$

avec

$$S = \frac{a_d}{b_q}X^{d-q} + S_1, \quad R = R_1.$$

On conclut par recurrence. Montrons l'unicite: supposons que

$$P = Q.S + R = Q.S' + R'$$

avec $\deg R, \deg R' < q$. Alors

$$Q.S - Q.S' = Q.(S - S') = R' - R.$$

On a

$$\deg(Q.(S - S')) = q + \deg(S - S') = \deg(R' - R) < q$$

et la seule possibilite est que $S - S' = 0$ (de sorte que $\deg(S - S') = -\infty$) et donc $R' - R = 0$. $\square$

REMARQUE A.4.1. La division euclidienne se generalise a l'anneau $A[X]$ pour A un anneau commutatif quelconque de la maniere suivante:

THÉORÈME A.5. Soit A un anneau commutatif et $Q \in A[X] - \{0\}$ un polynome dont le coefficient dominant $a_{\deg Q}(Q) \in A^\times$ (ie est inversible). Pour tout $P \in K[X]$ il existe des polynomes $S, R \in K[X]$ uniques vérifiant

$$\deg R < \deg Q \text{ et tels que } P = Q.S + R.$$

A.4.3. Application aux racines d'un polynome. Un invariant important d'un polynome est l'ensemble des valeurs ou sa fonction polynomiale s'annule:

DÉFINITION A.8. Soit

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

un polynome à coefficient dans K . L'ensemble des racines de P dans K , $\text{Rac}_P(K)$ est l'ensemble des solutions dans K de l'équation $P(z) = 0$:

$$\text{Rac}_P(K) = \{z \in K, P(z) = 0_K\}.$$

PROPOSITION A.9. Soit K un corps et P un polynome et $z \in K$, les deux énoncés suivants sont équivalents:

- (1) $P(z) = 0$ (ie. z est une racine de P).
- (2) Le polynome $X - z$ divise $P(X)$.

Preuve: Si $P(X) = (X - z)Q(X)$ on a

$$P(z) = (z - z).S(z) = 0_K.$$

Reciproquement si $P(z) = 0$, divisons P par $X - z$: on a

$$P(X) = S(X).(X - z) + R$$

avec R de degré $< \deg X - z = 1$ et donc R est constant (eventuellement nul). Mais

$$P(z) = 0 = S(z).(z - z) + R = R$$

et donc $R = 0$ c'est à dire

$$P(X) = S(X).(X - z).$$

□

On deduit de cette proposition le resultat fondamental suivant:

THÉORÈME A.6. Soit $P \in K[X]$ un polynome non nul alors P est divisible par le produit

$$\prod_{z \in \text{Rac}_P(K)} (X - z).$$

En particulier

$$|\text{Rac}_P(K)| = \deg \prod_{z \in \text{Rac}_P(K)} (X - z) \leq \deg P.$$

Preuve: Par recurrence sur $\deg P$: si P est constant non-nul c'est evident car P n'a pas de racines et

$$|\text{Rac}_P(K)| = 0 = \deg P.$$

Soit $z \in K$ une racine de $P(X)$ (si il n'y en a pas on a fini: $|\text{Rac}_P(K)| = 0$) alors

$$P(X) = (X - z).S(X)$$

et (comme K est integre)

$$P(z') = 0 \iff z' = z \text{ ou bien } Q(z') = 0$$

donc

$$\text{Rac}_P(K) = \{z\} \cup \text{Rac}_S(K).$$

comme $\deg S = d - 1$ on a par recurrence que

$$S(X) = \prod_{z' \in \text{Rac}_S(K)} (X - z').T(X)$$

et

$$P(X) = (X - z). \prod_{z' \in \text{Rac}_S(K)} (X - z').T(X).$$

□

COROLLAIRE A.2. *Soit K un corps et $|K|$ son cardinal (eventuellement infini) alors l'application lineaire*

$$P(X) \in K[X]_{\deg P < |K|} \mapsto P(\bullet) \in \mathcal{F}(K; K)$$

est injective (tout polynome de degre $< |K|$ peut etre identifie avec une unique fonction polynomiale). En particulier si $\text{car} K = 0$ alors $|K| \geq |\mathbb{Q}| = \infty$ l'application

$$P(X) \in K[X] \mapsto P(\bullet) \in \mathcal{F}(K; K)$$

est injective.

Preuve: Soit $P \in K[X]_{\deg P < |K|}$ dans le noyau: la fonction $x \in K \mapsto P(x) \in K$ est donc identiquement nulle et P possede $|K|$ racines comme $\deg P < |K|$ ceci n'est possible que si P est le polynome nul. □

A.4.4. Application: Structure des ideaux de $K[X]$. On rappelle qu'un ideal $I \subset K[X]$ de l'anneau $K[X]$ est un sous $K[X]$ -module contenu dans $K[X]$: un sous-groupe de $(K[X], +)$ qui stable par multiplication par les elements de $K[X]$. En d'autres termes, I verifie la condition de stabilite suivante:

$$\forall P, Q \in I, S \in K[X], P + S.Q \in I.$$

Un exemple simple d'ideal est le suivant: $Q = Q(X) \in K[X]$ un polynome, alors l'ensemble des multiples de Q

$$(Q) := K[X].Q = \{S.Q, S \in K[X]\}$$

est un ideal de $K[X]$ (le verifier).

NOTATION A.4. *Soit $Q = Q(X) \in K[X]$ un polynome, l'ideal*

$$(Q) = K[X].Q = \{S.Q, S \in K[X]\}$$

est appele ideal principal engendre par Q .

L'existence d'une division euclidienne permet une classification des ideaux de $K[X]$ entierement similaire a celle des sous-groupes de $\mathbb{Z}$: tout ideal de $K[X]$ est principal.

THÉORÈME A.7. *Soit $I \subset K[X]$ un ideal alors il existe $Q \in K[X]$ tel que I est l'ensemble des multiples de Q :*

$$I = (Q) = \{S.Q, S \in K[X]\}.$$

De plus si on suppose Q unitaire alors Q est unique.

Preuve: Si $I = \{0\} = 0.K[X]$ on a fini. Si $I \neq \{0\}$ soit $Q \in I - \{0\}$ un polynome non-nul de degre q minimal parmi les polynomes non-nuls de I . Soit $P \in I$. Par division euclidienne on peut ecrire

$$P = Q.S + R$$

avec $\deg R < q$. On a

$$R = P - Q.S \in I$$

(car $P, Q \in I$ et pour tout $S \in K[X]$, $S.Q \in I$ par definition d'un ideal) et donc $R \in I$. Par minimalite de q la seule possibilite est que $R = 0$ et donc $P = S.Q \in K[X].Q$. Si L est tel que $I = K[X].Q = K[X].L$ alors L est un multiple de Q (et Q est un multiple de L) et il n'existe qu'un seul multiple de Q qui soit unitaire: $a_{\deg Q}(Q)^{-1}.Q$ ou $a_{\deg Q}(Q) \neq 0$ est le coefficient dominant de Q . □

DÉFINITION A.9. Soit $I \subset K[X]$ un idéal non-nul alors l'unique polynôme unitaire Q_I tel que

$$I = (Q_I) = Q_I.K[X]$$

est appelé polynôme minimal de I . Si $I = \{0_K\}$ est l'idéal nul on posera

$$Q_I = 0_K.$$

Comme un noyau d'un morphisme d'anneau $\varphi : K[X] \mapsto A$ est un idéal on a:

COROLLAIRE A.3. Soit B un anneau et $\varphi : K[X] \mapsto B$ un morphisme d'anneaux. Alors il existe $Q_\varphi \in K[X]$ unitaire (ou nul) tel que

$$\ker(\varphi) = Q_\varphi.K[X].$$

Le polynôme Q_φ s'appelle le polynôme minimal de φ .

DÉFINITION A.10. Un anneau A tel que tout idéal $I \subset A$ est de la forme $I = q.A$ pour $q \in A$ est dit principal. Un anneau de polynômes sur un corps est donc principal.

On notera le lien suivant entre inclusion d'idéaux et divisibilité

PROPOSITION A.10. Soient

$$I = (P) = P.K[X] \text{ et } J = (Q) = Q.K[X]$$

des idéaux de $K[X]$ engendrés par des polynômes P et Q alors on a

$$I \subset J \iff Q|P.$$

Preuve: En effet si $I \subset J$ alors $P \in J = Q.K[X]$ et donc

$$P = Q.R, \quad R \in K[X].$$

Reciproquement si $P = Q.R$ alors pour tout $L \in I$ on a pour $S \in K[X]$

$$L = P.S = Q.R.S \in Q.K[X] = J$$

et donc $I \subset J$. □

A.4.5. Decomposition en polynômes irréductibles.

DÉFINITION A.11. Un polynôme $P(X) \in K[X]$ non constant est irréductible (ou premier) si les seuls diviseurs de P sont les multiples de 1 ou de P :

$$Q|P \implies Q = \lambda \text{ ou } Q = \lambda.P, \quad \lambda \in K^\times.$$

De manière équivalente: P est irréductible si et seulement si

$$Q|P \iff \deg Q = 0 \text{ ou } P.$$

On notera $\mathcal{P} \subset K[X]$ l'ensemble de tous les polynômes irréductibles et $\mathcal{P}_u \subset \mathcal{P}$ l'ensemble de ceux qui sont unitaires.

PROPOSITION A.11. (Lemme de Gauss) Soit P irréductible, si $P|Q_1.Q_2$ alors $P|Q_1$ ou $P|Q_2$.

Preuve: Écrivons $Q_1.Q_2 = P.S$. Supposons que $P \nmid Q_1$ et soit l'idéal

$$I = K[X].P + K[X].Q_1 \subset K[X].$$

l'idéal engendré par P et Q_1 . On va montrer que $I = K[X].P$. On a $I = D(X).K[X]$ pour D un polynôme. Comme $P \in I$ on a $D|P$ mais cela implique que D est soit un scalaire non nul soit un multiple de P . Dans ce dernier cas $I = P.K[X]$ et comme $Q_1 \in I$ on a $P|Q_1$ ce qu'on a exclu. Si D est un scalaire non-nul alors $I = K[X] \ni 1$: il existe $A(X), B(X)$ tels que

$$A(X)P(X) + B(X)Q_1(X) = 1.$$

On a alors

$$Q_2 = 1.Q_2 = (A.P + B.Q_1).Q_2 = A.P.Q_2 + B.Q_1.Q_2 = P.(A.Q_2 + B.S).$$

□

THÉORÈME A.8. Soient Q un polynome non constant alors Q se factorise de manière unique sous la forme

$$Q = \lambda.P_1 \cdots .P_s$$

ou les P_i sont des polynomes irréductibles unitaires et $\lambda \in K^\times$. De plus cette factorisation est unique: Si on a deux telles factorisation en irréductibles (unitaires)

$$Q = \lambda.P_1 \cdots .P_s = \mu.R_1 \cdots .R_r$$

alors $s = r$, $\lambda = \mu$ et il existe une permutation $\sigma : \{1, \dots, r\} \mapsto \{1, \dots, s = r\}$ telle que

$$R_i = P_{\sigma(i)}.$$

Preuve: On va montrer la factorisation par récurrence sur $\deg Q$. Si $\deg Q = 1$ on a fini car Q est forcément irréductible et si $Q(X) = a.X + b$, $a, b \in K$, $a \neq 0$ et on a l'écriture unique

$$Q = a(X + b/a).$$

Supposons $\deg Q = q + 1$ et qu'on a le résultat pour tous les polynomes de degré $\leq q$. Si Q possède un diviseur Q_1 non-constant et non multiple de Q on a alors $1 < \deg Q_1 < q + 1$ et

$$Q = Q_1.Q_2$$

avec $\deg Q_1, \deg Q_2 < q + 1$. Sinon Q est irréductible et on a la factorisation

$$Q = a_{\deg Q}.Q_1, \quad Q_1 = a_{\deg Q}^{-1}.Q.$$

Dans le cas précédent, on a par récurrence

$$Q_1 = \lambda_1.P_1 \cdots .P_{s_1}, \quad Q_2 = \lambda_2.P_{s_1+1} \cdots .P_{s_1+s_2}$$

avec les P_i irréductibles unitaires et

$$Q = \lambda_1.\lambda_2.P_1 \cdots .P_{s_1}.P_{s_1+1} \cdots .P_{s_1+s_2}.$$

Montrons l'unicité par récurrence sur $\deg Q$. Si $\deg Q = 1$ c'est immédiat.

Dans le cas général soit

$$Q = \lambda.P_1 \cdots .P_s = \mu.R_1 \cdots .R_r$$

alors $P_s | \mu.R_1 \cdots .R_r$ et par le lemme de Gauss P_s divise un des R_i . Or c'est R_r . Comme R_r est irréductible, unitaire et P_s est non constant unitaire on a $P_s = R_r$ et

$$Q = \lambda.P_1 \cdots .P_s = \mu.R_1 \cdots .R_{r-1}.P_s$$

et

$$0 = (\lambda.P_1 \cdots .P_{s-1} - \mu.R_1 \cdots .R_{r-1})P_s$$

et comme $K[X]$ est intègre

$$\lambda.P_1 \cdots .P_{s-1} = \mu.R_1 \cdots .R_{r-1}$$

et on applique la récurrence. □

A.4.5.1. *Valuation.* Soit $Q(X) = a_q X^q + a_{q-1} X^{q-1} + \cdots + a_0$ un polynome de degré $q \geq 0$ ($a_q \neq 0$) alors la décomposition de Q en irréductibles peut se récrire de manière compacte

$$Q = a_q \prod_{P \in \mathcal{P}_u} P^{v_P(Q)}$$

ou

- P parcourt l'ensemble infini des polynomes irréductibles unitaires,
- les $v_P(Q) \geq 0$ sont des entiers nuls pour tous les P sauf un nombre fini,
- Quand $v_P(Q) = 0$ on a pose

$$P^{v_P(Q)} = P^0 := 1.$$

Ainsi, l'entier $v_P(Q)$ est l'exposant de la plus grande puissance du polynome irréductible P divisant Q .

DÉFINITION A.12. *L'entier $v_P(Q)$ est appelle la valuation de Q en P ou la valuation P -adique de Q . Pour $Q = 0$ on pose $v_P(Q) = +\infty$ pour tout P irréductible.*

Ces valuations ont les proprietes suivantes

THÉORÈME A.9. *Soient $Q, R \in K[X] - \{0\}$ de degres respectif q et r et de coefficient dominant a_q et b_r ; on a*

(1) *Pour tout $P \in \mathcal{P}_u$, on a*

$$v_P(Q.R) = v_P(Q) + v_P(R)$$

et plus precisement

$$Q.R = a_q.b_r \prod_{P \in \mathcal{P}_u} P^{v_P(Q)+v_P(R)}.$$

(2) *On a*

$$Q|R \iff \forall P \in \mathcal{P}_u, v_P(Q) \leq v_P(R)$$

(3) *Pour tout P on a*

$$v_P(Q + R) \geq \min(v_P(Q), v_P(R))$$

avec egalite si $v_P(Q) \neq v_P(R)$.

A.4.6. PGDC et PPMC. Soient $P, Q \in K[X] - \{0\}$. On a alors les deux ideaux:

$$(P) := K[X].P, (Q) := K[X].Q$$

et on peut alors former deux autres ideaux: leur intersection et leur somme

$$(P) \cap (Q) \subset (P), (Q) \subset (P) + (Q) = \langle P, Q \rangle \subset K[X].$$

A.4.6.1. *Le PGCD.* L'ideal engendre par P et Q est de la forme

$$\langle P, Q \rangle = (P) + (Q) = K[X].P + K[X].Q = R.K[X]$$

avec R unitaire. Alors comme $P, Q \in \langle P, Q \rangle$, R divise P et Q : on a

$$R|P \text{ \& } R|Q.$$

D'autre part si un polynome S divise a la fois P et Q alors

$$K[X].P + K[X].Q = R.K[X] \subset S.K[X]$$

et donc $S|R$. Ainsi R est le *Plus Grand Diviseur Commun* (unitaire) de P et Q au sens ou tout diviseur commun de P et Q doit diviser R .

DÉFINITION A.13. *Soient $P, Q \in K[X] - \{0\}$, note*

$$(P, Q) := R$$

le generateur unitaire de l'ideal $(P) + (Q) = \langle P, Q \rangle$ et on l'appelle le PGCD de P et Q . En particulier si $(P, Q) = 1$ (cad $\langle P, Q \rangle = K[X]$) on dit que P et Q sont premiers entre eux.

REMARQUE A.4.2. Si $Q = 0$ alors $(P, 0) = P_u$ est l'unique polynome unitaire qui est multiple de P .

PROPOSITION A.12. (*Bezout*) *Soient P, Q des polynomes. Il existe $A, B \in K[X]$ tels que*

$$(P, Q) = A.P + B.Q.$$

En particulier, deux polynomes P et Q sont premiers entre eux ssi il existe $A, B \in K[X]$ tels que

$$1 = A.P + B.Q.$$

Preuve: On a

$$(P) + (Q) = (P, Q).K[X] = P.K[X] + Q.K[X].$$

En particulier (P, Q) est de la forme

$$(P, Q) = P.A + Q.B.$$

Supposons qu'il existe A, B tels que $1 = A.P + B.Q$ alors $(P) + (Q)$ contient 1 et donc $1.K[X] = K[X]$ de sorte que $(P) + (Q) = K[X]$. □

A.4.6.2. Algorithme d'Euclide. L'algorithme d'Euclide qui permet de calculer le PGCD de deux entiers permet de calculer le PGCD de deux polynômes: Si P et Q sont deux polynômes dont on souhaite calculer (P, Q) on applique la méthode suivante:

- (1) On suppose que $\deg P \geq \deg Q$ et on effectue la division euclidienne de P par Q :

$$P = SQ + R, \deg R < \deg P.$$

Si $R = 0$ cela signifie que $Q|P$ et donc

$$(P, Q) = Q.$$

Sinon, cette relation implique que l'idéal engendré par P et Q est égal à l'idéal engendré par Q et R

$$(P, Q) = (Q, R).$$

- (2) On recommence l'étape précédente avec $P_1 = R$ et $Q_1 = Q$.
 (3) ...
 (4) Comme le degré du reste diminue d'au moins 1 à chaque étape strictement le processus s'arrête après au plus $\max(\deg P, \deg Q)$ étapes.

A.4.6.3. Le PPCM. Soit l'intersection $(P) \cap (Q) \subset K[X]$. C'est un idéal non-nul car il contient le produit $P.Q$. Il est donc de la forme $(P) \cap (Q) = K[X].S$ avec S unitaire. On a donc

$$P|S \text{ et } Q|S$$

et S est un multiple commun à P et à Q . De plus si $P|T$ et $Q|T$ alors

$$T \in K[X].P \cap K[X].Q = K[X].S$$

et $S|T$. Ainsi S est le *Plus Petit Multiple Commun* (unitaire) de P et Q .

DÉFINITION A.14. Soient $P, Q \in K[X] - \{0\}$, note

$$[P, Q] := R$$

le générateur unitaire de l'idéal $(P) \cap (Q)$ et on l'appelle le PPCM de P et Q .

PROPOSITION A.13. (Formule du produit) Soient $P, Q \in K[X] - \{0\}$ et unitaires. On a

$$P.Q = P, Q.$$

Preuve: Voir l'exercice concernant la formule du produit

$$m.n = (m, n)[m, n]$$

pour $m, n \in \mathbb{Z}$. □

A.4.6.4. *Generalisation a un nombre arbitraire de polynomes.*

DÉFINITION A.15. *Soient $P_1, \dots, P_k$ des polynomes alors leur PGCD et leur PPCM notes*

$$(P_1, \dots, P_k) \text{ et } [P_1, \dots, P_k]$$

sont respectivement les generateurs unitaires des ideaux

$$(P_1) + \dots + (P_k) \text{ et } (P_+) \cap \dots \cap (P_k).$$

En particulier si

$$(P_1, \dots, P_k) = 1, \text{ ie. } \langle P_1, \dots, P_k \rangle = K[X]$$

on dit que $P_1, \dots, P_k$ sont premiers dans leur ensemble.

REMARQUE A.4.3. On a

$$(P_1, \dots, P_k) | (P_1, P_2)$$

car

$$(P_1) + (P_2) \subset (P_1) + \dots + (P_k).$$

A.4.6.5. *PGDC, PPCM et decomposition en irreductibles.*

THÉORÈME A.10. *Soient Q, R des polynomes non-nuls de degres q et r et*

$$Q = a_q \cdot \prod_{P \in \mathcal{P}_u} P^{v_P(Q)}, \quad R = b_r \cdot \prod_{P \in \mathcal{P}_u} P^{v_P(R)}$$

leur decompositions en polynomes irreductible unitaires alors

$$(Q, R) = \prod_{P \in \mathcal{P}_u} P^{\min(v_P(Q), v_P(R))}, \quad [Q, R] = \prod_{P \in \mathcal{P}_u} P^{\max(v_P(Q), v_P(R))}.$$

Preuve: Exercice. □

A.5. Application a la construction de corps

Soit $\mathcal{M}$ une K -algebre (pas forcément commutative, par exemple $\text{End}(V)$ ou $M_d(K)$) d'unité $1_{\mathcal{M}}$ et $M \in \mathcal{M}$ un element. On associe a M une application (dite d'évaluation en M)

$$\text{ev}_M : \begin{array}{ccc} K[X] & \mapsto & \mathcal{M} \\ P(X) & \mapsto & P(M) \end{array}$$

ou

$$P(M) = a_0.M^0 + a_1.M + \dots + a_n.M^n + \dots + a_d.M^d.$$

On a pose $M^0 = 1_{\mathcal{M}}$ et

$$M^n = M.M \dots .M (n \text{ fois}).$$

PROPOSITION A.14. *Cette application est un morphisme d'algebres: on a*

$$(\lambda.P + Q)(M) = \lambda.P(M) + Q(M), \quad (P.Q)(M) = P(M).Q(M).$$

On notera l'image de cette application par

$$K[M] = \text{ev}_M(K[X]) = \{P(M), P \in K[X]\}.$$

C'est une sous-algebre (un sous-anneau et un SEV) commutative de $\mathcal{M}$: l'algebre des polynomes en M .

Preuve: On ne fait que la multiplication:

$$P(M).Q(M) = (a_0.M^0 + a_1.M + \dots + a_d.M^d).(b_0.M^0 + b_1.X + \dots + b_d.M^d) =$$

$$\sum_{p,q \leq d} a_p.M^p.b_q.M^q = \sum_{p,q \leq d} a_p.b_q.M^{p+q} = \sum_{n \leq d+d'} \left(\sum_{p+q=n} a_p.b_q \right) M^n = (P.Q)(M)$$

ici on a utilise les proprietes des lois de composition de $\mathcal{M}$ (associativite, distributivite) et le fait (valable meme si $\mathcal{M}$ n'est pas commutative) que

$$a_p.M^p.b_q.M^q = a_p.b_q.M^{p+q} = a_p.b_q.M^{p+q}.$$

L'algebre $K[M]$ est commutative car $K[X]$ l'est:

$$P(M).Q(M) = (P.Q)(M) = (Q.P)(M) = Q(M).P(M).$$

□

EXERCICE A.1. Montrer que $K[M]$ est la plus petite sous-algebre de $\mathcal{M}$ contenant M : c'est l'algebre engendree par M . On dit que $K[M]$ est monogene car elle est engendree par un seul element.

A.5.1. Polynome minimal de M . Comme $\text{ev}_M : K[X] \mapsto \mathcal{M}$ est un morphisme d'anneau son noyau $\ker(\text{ev}_M)$ est un $K[X]$ ideal et donc de la forme

$$\ker(\text{ev}_M) = Q_{\text{ev}_M}.K[X]$$

pour Q_{ev_M} un polynome nul ou unitaire.

DÉFINITION A.16. Soit $\mathcal{M}$ un K -algebre et $M \in \mathcal{M}$ et

$$\text{ev}_M : P(X) \in K[X] \mapsto P(M) \in \mathcal{M}$$

le morphisme d'evaluation en M dont le noyau est

$$\ker(\text{ev}_M) = \{P, P(M) = 0_{\mathcal{M}}\} = Q_{\text{ev}_M}.K[X]$$

avec Q_{ev_M} nul ou unitaire. Le polynome

$$Q_{\text{ev}_M}$$

est appele polynome minimal de M et est note

$$P_{\min, M} := Q_{\text{ev}_M}.$$

A.5.2. Un critere pour que $K[M]$ soit un corps.

THÉORÈME A.11. Soit B un anneau et $\varphi : K[X] \mapsto B$ un morphisme d'anneaux non-nul et ecrivons $\ker \varphi = Q.K[X]$. Alors on a

$$Q \text{ est irreductible} \iff \varphi(K[X]) \text{ est un corps.}$$

Preuve: Soit $b = \varphi(P) \in \varphi(K[X]) - \{0\}$. Supposons P irreductible; on veut montrer que b est inversible dans $\varphi(K[X])$. Considerons l'ideal $I = \langle P, Q \rangle = K[X].P + K[X].Q$ alors $I = K[X]$: en effet ecrivons $I = K[X].R$; comme $P, Q \in I = K[X].R$ et on doit avoir $R|P$ et $R|Q$. Comme P est irreductible et $R|P$, R est constant non-nul ou de la forme $\lambda.P$. Dans le second cas on aurait $I = K[X].P = \ker \varphi$ ce qui contredit le fait que $b = \varphi(P) \neq 0$. On a donc $I = K[X]$ et il existe $U, V \in K[X]$ tels que

$$U.P + V.Q = 1_K$$

et alors

$$1_B = \varphi(U.P + V.Q) = \varphi(U).\varphi(P) + \varphi(V).\varphi(Q) = \varphi(U).\varphi(P) = \varphi(V).b$$

et b est inversible et son inverse $\varphi(V) \in \varphi(K[X])$.

Reciproquement supposons que $\varphi(K[X])$ est un corps; alors $Q \neq 0$ car sinon φ sera un isomorphisme de $K[X]$ vers son image et $K[X]$ est pas un corps. Q n'est pas non-plus constant non nul car φ sera le morphisme nul.

Supposons que Q ne soit pas irréductible: $Q = RS$ avec $0 < \deg R, \deg S < \deg Q$. On a

$$\varphi(Q) = 0_B = \varphi(R) \cdot \varphi(S)$$

et donc $\varphi(R)$ ou $\varphi(S) = 0_B$ mais R et S ne peuvent appartenir à $\ker(\varphi)$ (car ils seraient divisible par Q). $\square$

Appliquant ce résultat, on obtient

COROLLAIRE A.4. Soit $\mathcal{M}$ un K -algèbre et $M \in \mathcal{M}$ et

$$\text{ev}_M : P(X) \in K[X] \mapsto P(M) \in \mathcal{M}$$

le morphisme d'évaluation en M . Alors $K[M]$ est un corps si et seulement si $P_{\min, M}(X)$ est irréductible (en particulier $P_{\min, M}(X) \neq 0$).

Voici un critère d'irréductibilité

PROPOSITION A.15. Soit $P(X) \in K[X]$ un polynôme de degré 2, 3 alors $P(X)$ est irréductible ssi il n'a pas de racine dans K .

Preuve: On peut supposer P unitaire de degré ≥ 2 . Si P est irréductible il n'a pas de factorisation de la forme

$$P(X) = (X - z)S(X), \quad z \in K, \quad S \in K[X]$$

et donc il n'a pas de racine dans K .

Supposons $\deg P = 2, 3$. Si P est réductible il aura une factorisation

$$P(X) = Q(X)S(X)$$

avec Q, S unitaires tels que

$$\deg Q + \deg S = \deg P = 2 \text{ ou } 3, \quad \deg Q, \deg S \geq 1$$

et donc Q ou S doit avoir degré 1: ie est de la forme $X - z$, $z \in K$ et donc P admet une racine dans K . $\square$

EXERCICE A.2. (à faire après le chapitre sur les applications linéaires) Soit $\mathcal{M}$ un K -algèbre de dimension finie et $M \in \mathcal{M}$. Soit $K[X]_{\leq d}$ le sous-espace vectoriel des polynômes de degré $\leq d$.

- (1) Montrer que si $d \geq \dim \mathcal{M}$, il existe un polynôme P non-nul de degré $\leq d$ tel que $P(M) = 0_d$.
- (2) Montrer que $P_{\min, M} \neq 0$ et $P_{\min, M} \leq \dim \mathcal{M}$.
- (3) Montrer que si $P(0) = a_0 \neq 0$ alors M est inversible dans $\mathcal{M}$ et en fait $M^{-1} = Q(M)$ avec $Q \in K[X]_{\leq d-1}$ et donc $M^{-1} \in K[M]$.