Algèbre linéaire

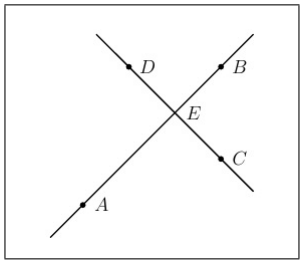
Un second cours

MAT 2541 - Septembre 2015

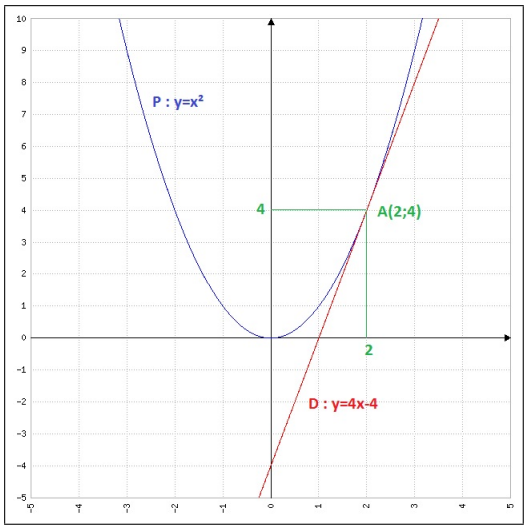
PAUL-EUGÈNE PARENT

INTRODUCTION

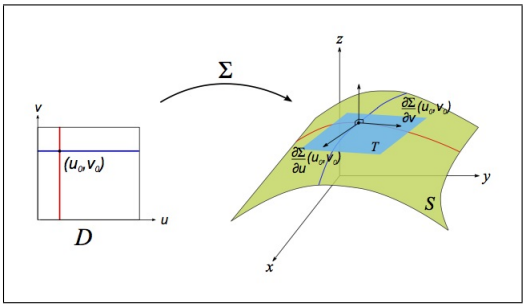
L'algèbre linéaire est partout. Depuis la première fois que l'on vous a demandé à l'école secondaire de trouver le point d'intersection de deux droites non-parallèles dans le plan, l'algèbre linéaire n'a de cesse joué un rôle de premier plan.



L'algèbre linéaire est souvent perçue comme une première approximation. Par exemple dans vos cours de calcul différentiel de première année on vous apprend qu'au voisinage d'un point x_o du domaine d'une fonction différentiable $f : D \to \mathbb{R}$ on peut "approcher" f par l'équation de la tangente au point $(x_o, f(x_o))$, c'est-à-dire, $f(x) \approx f(x_0) + f'(x_o)(x - x_o)$ lorsque x est "près" de x_o.



Dans les cours de géométrie différentielle ou de calcul avancé on montre que l'on peut généraliser ce concept aux surfaces lisses.



Ces droites et ces plans sont des exemples de l'objet principal d'étude de l'algèbre linéaire : *l'espace vectoriel*.

Lorsque l'on munit un espace vectoriel V d'une application bilinéaire, dite *crochet de Lie*, c'est-à-dire, une application

$$[\cdot,\cdot] : V \times V \to V$$

linéaire en chaque variable et satisfaisant aux deux conditions suivantes :

- $[x, y] = -[y, x]$ pour tout $x, y \in V$ et
- $[x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0$ (relation de Jacobi)

alors on dit que la paire $(V, [\cdot,\cdot])$ est une algèbre de Lie. Vous connaissez déjà un exemple : $(\mathbb{R}^3, \times)$, où "$\times$" est le *produit vectoriel*. Remarquablement ces espaces vectoriels codifient la structure électronique de l'atome d'hydrogène (ces fameuses "orbitales") et sont à l'origine des notations spectroscopiques (par exemple : $1s^2, 2p^6, 3d^{10}$, ...).

Lorsque la dimension de V devient infinie alors on passe dans un domaine que l'on appelle *l'analyse fonctionnelle*. Ici on retrouve d'une part l'analyse harmonique (les séries de Fourier) et les phénomènes ondulatoires ; et d'autre part l'algèbre des opérateurs et son application toute désignée : *la mécanique quantique* en physique. On ne peut passer sous silence *l'équation de Schrödinger* régissant l'évolution d'un état ϕ (qui est nul autre qu'un vecteur)

$$H\phi(x,t) = i\hbar\frac{\partial}{\partial t}\phi(x,t),$$

où H est un opérateur linéaire (*l'hamiltonien*) caractérisant le système à l'étude et agissant sur un espace vectoriel qui est en général de dimension infinie, i est le nombre complexe tel que $i^2 = -1$ et $\hbar$ est l'une des constantes fondamentales de la physique.

Dans ces notes nous tenterons de couvrir en grande partie l'étude de ces espaces vectoriels lorsque $\dim V < \infty$. Ce cours devrait vous servir comme un excellent tremplin vers toutes ces autres disciplines.

TABLE DES MATIÈRES

Première partie

L'espace vectoriel de dimension finie

CHAPITRE 1

LES NOMBRES

Avant de discuter de vecteurs révisons la notion de nombre en rappelant quelques faits et résultats importants.

1.1 Principe de raisonnement par récurrence

Les nombres naturels et les nombres entiers seront notés respectivement

$$\mathbb{N} = \{0, 1, 2, ...\} \quad \text{et} \quad \mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\}.$$

Nous accepterons comme axiome de base le principe suivant :

Le Raisonnement par récurrence : *Supposons que pour chaque nombre naturel $n \in \mathbb{N}$ un énoncé $E(n)$ soit défini. Supposons démontré*

1. *$E(0)$ est vrai ; et*
2. *pour tout $n \in \mathbb{N}$ l'implication $E(n) \Rightarrow E(n+1)$ est vrai.*

Alors l'énoncé $E(n)$ est vrai pour tout $n \in \mathbb{N}$.

Remarques : L'énoncé $E(0)$ est appelé l'*ancre* de l'induction. Si l'on ancre l'induction en un autre entier naturel $k > 0$ alors la récurrence démontre que l'énoncé $E(n)$ est vrai pour tout $n \geq k$. L'énoncé $E(n)$ est souvent appelé l'*hypothèse d'induction.*

Parfois il est fort utile de considérer d'autres versions équivalentes de cet axiome. En voici deux :

La Récurrence forte : *Supposons que pour chaque nombre naturel $n \in \mathbb{N}$ un énoncé $E(n)$ soit défini. Supposons démontré*

1. *$E(0)$ est vrai ; et*
2. *pour tout $n \in \mathbb{N}$ l'implication $[\forall k \leq n, E(k)] \Rightarrow E(n+1)$ est vrai.*

Alors l'énoncé $E(n)$ est vrai pour tout $n \in \mathbb{N}$.

Le Principe du bon ordre : *Soit $M \subset \mathbb{N}$ un sous-ensemble non-vide. Alors M contient un plus petit élément, c'est-à-dire un élément $m_o \in M$ tel que $m_o \leq m$ pour tout $m \in M$.*

On trouvera une démonstration de l'équivalence de ces énoncés en annexe. Nous aurons l'opportunité d'illustrer le raisonnement par récurrence ainsi que sa version forte tout au long de ces notes. Pour l'instant donnons un exemple fort utile du principe du bon ordre.

Théorème 1.1 *(Division euclidienne) Soit $p \in \mathbb{Z}$ et $q \in \mathbb{N}^+$. Il existe une décomposition*

$$p = m \cdot q + r$$

pour un unique $m \in \mathbb{Z}$ et un unique $r \in \mathbb{N}$ tel que $0 \leq r < q$.

Démonstration. Considérons l'ensemble $S = \mathbb{N} \cap \{p - m \cdot q \mid m \in \mathbb{Z}\}$. Si $p \leq 0$ alors $p - p \cdot q \in S$. Si $p > 0$ alors $p - 0 \cdot q \in S$. Donc en général $S \neq \emptyset$. Par le principe du bon ordre, il existe un élément minimal $r \in S$. Par définition de l'élément minimal, il existe $m \in \mathbb{Z}$ tel que

$$p = m \cdot q + r.$$

Si $r \geq q$ alors on a $p = m \cdot q + r + (q - q) = (m + 1) \cdot q + (r - q)$. Mais alors $r - q \in S$ et contredit la nature minimale de r. On conclut que $0 \leq r < q$. S'il existe un second $0 \leq r' < q$ et un second $m' \in \mathbb{Z}$ tels que $p = m' \cdot q + r'$ alors nécessairement on doit avoir $r \leq r' < q$, sinon r ne serait pas l'élément minimal de S. Donc on doit avoir

$$0 \leq r' - r < q \quad \text{ce qui implique} \quad 0 \leq (m - m') \cdot q < q.$$

Ceci force donc $m = m'$ et donc $r = r'$.

□

1.2 Bijections et inverses

On rappelle que

Définition 1.1 *Soit $f : X \to Y$ une application entre deux ensembles.*

- *L'application f est* injective *si*

$$f(a) = f(b) \quad \Rightarrow \quad a = b.$$

- *L'application f est* surjective *si pour tout $y \in Y$, il existe $x \in X$ tel que $y = f(x)$.*
- *L'application f est une* bijection *si l'application f est à la fois injective et surjective.*

Définition 1.2 *Soit $f : X \to Y$ une application entre deux ensembles. L'application f est* inversible *s'il existe une application $g : Y \to X$ telle que*

$$g \circ f = \mathrm{id}_X \quad \textit{et} \quad f \circ g = \mathrm{id}_Y.$$

Remarque : Si une application f est inversible alors son inverse est unique. En effet soit $h : Y \to X$ un second inverse. En utilisant l'associativité de la composition de fonctions, nous avons alors pour tout $y \in Y$

$$h(y) = (\mathrm{id}_X \circ h)(y) = ((g \circ f) \circ h)(y) = (g \circ (f \circ h))(y) = (g \circ \mathrm{id}_Y)(y) = g(y).$$

On peut donc écrire sans ambiguïté f^{-1} pour désigner l'inverse de f.

Le lecteur devra être vigilant car si $A \subset Y$ alors l'ensemble

$$\{x \in X \mid f(x) \in A\}$$

est appelé *l'image réciproque de A* relatif à l'application f. Historiquement nous notons cet ensemble $f^{-1}(A)$ et ce MÊME si f n'est pas inversible !

Exemple : Soit $f(x) = x^2$. Alors f n'est pas inversible mais $f^{-1}(\{1\}) = \{-1, 1\}$ tandis que $f^{-1}(\{-2\}) = \emptyset$.

Exercice : Si $A, B \subset Y$ et $A \cap B = \emptyset$ alors $f^{-1}(A) \cap f^{-1}(B)$.

Proposition 1.2 *Une application $f : X \to Y$ est inversible si et seulement si elle est bijective.*

Démonstration. Supposons f inversible. Si $f(a) = f(b)$ alors on peut appliquer de part et d'autre l'inverse et obtenir

$$a = f^{-1}(f(a)) = f^{-1}(f(b)) = b,$$

et conclure que f est injective. De plus, si $y \in Y$ alors $f^{-1}(y) \in X$. Mais dès lors

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

et l'on conclut que f est surjective. Réciproquement, si f est une bijection alors f est en particulier surjective, c'est-à-dire, pour tout $y \in Y$ il existe $x \in X$ tel que $f(x) = y$ ou en d'autres mots pour tout $y \in Y$, l'image réciproque $f^{-1}(\{y\}) \neq \emptyset$. Si l'on considère l'ensemble des images réciproques

$$\widehat{Y} = \{f^{-1}(\{y\}) \mid y \in Y\}$$

on remarque deux choses :

1. $\widehat{Y}$ est en bijection avec Y (voir l'exercice précédant) ; et
2. $\bigcup_{y \in Y} f^{-1}(\{y\}) = X$.

En invoquant l'axiome du choix (voir l'annexe D pour une courte discussion), il existe une fonction choix $g : Y \to X$ telle que $g(y) \in f^{-1}(\{y\})$. En particulier on a $f \circ g = \mathrm{id}_Y$. Par contre f étant injective, pour tout $y \in Y$ il existe un unique $x \in X$ tel que $f^{-1}(\{y\}) = \{x\}$. On conclut donc que $g(y) = x$. On a donc

$$g(y) = x \quad \text{si et seulement si} \quad f(x) = y.$$

Finalement, si $g(f(x)) = \tilde{x}$ alors $f(\tilde{x}) = f(x)$. Comme f est injective, $x = \tilde{x}$, c'est-à-dire, $g \circ f = \mathrm{id}_X$.

□

Remarque : Nous n'avions pas à invoquer l'axiome du chois dans cette preuve. En effet, comme il n'y a qu'un "x" dans chacun des $f^{-1}(\{y\})$ il n'y a donc aucune ambiguïté à faire un choix. En invoquant l'axiome du choix, nous avons en fait démontré le résultat supplémentaire suivant : toute surjection possède un inverse à droite.

1.3 Les groupes

Lorsque l'on munit les entiers de l'addition on obtient un exemple d'un *groupe*.

Définition 1.3 *Un groupe est un ensemble G munit d'une application*

$$\circ : G \times G \to G$$

telle que

1. *$a \circ (b \circ c) = (a \circ b) \circ c$ pour tout $a, b, c \in G$ (associativité) ;*
2. *il existe un élément $e \in G$ tel que $a \circ e = e \circ a = a$ pour tout $a \in G$ (élément neutre) ;*
3. *pour tout $a \in G$ il existe $b \in G$ tel que $a \circ b = b \circ a = e$ (inverse).*

Si de plus $a \circ b = b \circ a$ pour tout $a, b \in G$ alors on dit que le groupe G est un groupe *abélien*.

Remarques :

- Le (2) implique que tout groupe possède au moins un élément.
- L'élément neutre d'un groupe G est unique. Si $e' \in G$ est un autre élément neutre alors

$$e = e \circ e' = e'.$$

- Comme dans le cas des fonctions inversible, l'inverse d'un élément $a \in G$ est aussi unique. Si $b' \in G$ est autre inverse de a alors

$$b' = b' \circ e = b' \circ (a \circ b) = (b' \circ a) \circ b = e \circ b = b.$$

  On peut donc sans ambiguïté écrire a^{-1} pour l'inverse de a.

Exemple : Les entiers $\mathbb{Z}$ munit de l'addition usuelle + et de l'élément neutre 0 est un exemple d'un groupe abélien. Par contre $\mathbb{N}$ n'est pas un groupe car (3) n'est pas satisfait.

Exemple : Soit X un ensemble non-vide et considérons l'ensemble

$$G(X) = \{f : X \to X \mid f \text{ est inversible}\}$$

munit de l'opération, "$\circ$", composition de fonctions. Alors $(G(X), \circ)$ est un groupe qui est en général non abélien. La composition est une opération associative (exercice). L'élément neutre est la fonction identité, c'est-à-dire, la fonction $\mathrm{id}_X : X \to X$ telle que $\mathrm{id}_X(x) = x$ pour tout $x \in X$. Exercice : vérifiez que $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.

Définition 1.4 *Un sous-ensemble H non-vide d'un groupe $(G, \circ)$ est un sous-groupe de G si*

1. *pour tout $h, k \in H$ on a $h \circ k \in H$ et*
2. *pour tout $h \in H$ on a $h^{-1} \in H$.*

Dans ce cas on notera $H \leq G$.

Nous avons toujours $\{e\} \leq G$ et $G \leq G$. Ce sont les deux sous-groupes triviaux de G. On voit que $(\mathbb{Z}, +) \leq (\mathbb{Q}, +)$. Par contre $\mathbb{N}$ n'est pas un sous-groupe de $(\mathbb{Z}, +)$ malgré que la première condition soit satisfaite.

Proposition 1.3 *Soit $H \leq (G, \circ)$. Alors $(H, \circ)$ est un groupe.*

Démonstration. Par définition $H \neq \emptyset$. Soit $h \in H$. Par la condition (2) de la définition d'un sous-groupe, $h^{-1} \in H$. Maintenant par la condition (1), on a $e_G = h \circ h^{-1} \in H$. Mais comme

$$e_G \circ h = h \circ e_G = h$$

dans G et en particulier pour tout $h \in H$, e_G est donc l'élément neutre de $(H, \circ)$. A nouveau par l'associativité de G on a

$$h \circ (k \circ l) = (h \circ k) \circ l,$$

qui est vérifié en particulier pour tout $h, k, l \in H$, c'est-à-dire, $\circ$ est une opération associative sur H. Finalement la condition d'existence d'un inverse est garanti par la condition (2) de la définition d'un sous-groupe.

□

1.3.1 L'équation $a \circ x = b$

Dans un groupe G étant donné deux éléments $a, b \in G$, peut-on toujours trouver un élément $x \in G$ tel que $a \circ x = b$? Si oui cette solution est-elle unique ?

Avant de répondre à ces questions remarquons que l'opération d'un groupe G satisfait aux lois de simplification (à droite et à gauche), c'est-à-dire,

Lemme 1.4 *Si $a \circ x = a \circ y$ (à gauche) ou $x \circ a = y \circ a$ (à droite) alors $x = y$.*

Démonstration. Supposons que $a \circ x = a \circ y$. Nous avons donc

$$x = e \circ x = (a^{-1} \circ a) \circ x = a^{-1} \circ (a \circ x) = a^{-1} \circ (a \circ y) = (a^{-1} \circ a) \circ y = e \circ y = y.$$

Le cas à droite se démontre de la même façon.

□

Il s'en suit que

Proposition 1.5 *Soit $a, b \in G$. Il existe une unique solution à l'équation $a \circ x = b$.*

Démonstration. Soit $x = a^{-1} \circ b$. Alors $a \circ x = a \circ (a^{-1} \circ b) = (a \circ a^{-1}) \circ b = e \circ b = b$. Pour démontrer l'unicité de la solution supposons l'existence d'une deuxième solution $y \in G$. Nous avons donc $a \circ x = b = a \circ y$. Les lois de simplification impliquent $x = y$.

□

1.3.2 A propos de l'associativité

Lorsqu'un ensemble non-vide X est munit d'une opération binaire $\circ : X \times X \to X$ (comme dans le cas d'un groupe) une question naturelle se pose immédiatement : quelle signification devons-nous donner à une expression du type

$$x_1 \circ x_2 \circ x_3 \, ? \tag{1.1}$$

Dans ce cas-ci nous pouvons d'une part évaluer $(x_1 \circ x_2)$ puis évaluer $(\underbrace{(x_1 \circ x_2)}_{\in X} \circ x_3)$ ou d'autre part évaluer $(x_2 \circ x_3)$ puis évaluer $(x_1 \circ \underbrace{(x_2 \circ x_3)}_{\in X})$. Si l'opération binaire $\circ$ satisfait à l'axiome d'associativité alors ces deux dernières expressions sont les-mêmes et l'ambiguïté de l'expression (1.1) est levée. Quand est-il de l'expression

$$x_1 \circ x_2 \circ x_3 \circ x_4 \,?$$

Dans ce cas nous avons cinq possibilités : $(x_1 \circ (x_2 \circ (x_3 \circ x_4)))$, $(x_1 \circ ((x_2 \circ x_3) \circ x_4))$, $(((x_1 \circ x_2) \circ x_3) \circ x_4)$, $((x_1 \circ (x_2 \circ x_3)) \circ x_4)$ et $((x_1 \circ x_2) \circ (x_3 \circ x_4))$. Sans trop de difficulté nous pouvons montrer que ces cinq expressions sont égales à l'aide de la règle d'associativité. Quand est-il des expressions de longueur cinq ou plus ? En fait nous aimerions démontrer que tous les "parenthésages" de même longueur sont égaux. Avant de tenter de démontrer ce résultat nous devons bien définir les objets impliqués dans cette démonstration : qu'est-ce au juste qu'un parenthésage admissible ? On réfère le lecteur avide d'information sur le sujet à l'Annexe A pour tous les détails.

On se contentera d'énoncer qu'à cause de l'associativité (voir le Théorème E.2), peut importe l'ordre dans lequel nous effectuons les opérations dans une expression algébrique du type $x_1 \circ x_2 \circ \ldots \circ x_n$ tout en maintenant fixe l'ordre relatif des x_k (en d'autres mots un choix d'un parenthésage) nous obtenons le même élément de X. L'écriture

$$x_1 \circ x_2 \circ \ldots \circ x_n \in X$$

n'est donc pas ambiguë et sous-entendra toujours que nous avons fixé un parenthésage pour faire le calcul. Dans le cas d'un groupe abélien $(G, +, 0)$ nous adopterons la notation plus compacte

$$\sum_{k=1}^{n} g_k = g_1 + g_2 + \ldots + g_n \in G.$$

1.4 Les corps

On se rappellera que les nombres rationnels $\mathbb{Q}$, les nombres réels $\mathbb{R}$ ainsi que les nombres complexes $\mathbb{C}$ sont trois exemples de groupes abéliens par rapport à l'addition et l'élément neutre 0. De plus ils sont munit d'une multiplication (d'élément neutre 1) par rapport à laquelle tout élément non nul est inversible. L'existence de ces deux opérations sur un même ensemble nous mène vers la notion de *corps*, c'est-à-dire,

Définition 1.5 *Un corps $\mathbb{k}$ est un ensemble contenant au moins deux éléments distincts 0 et 1 munit de deux opérations associatives : l'addition $+$ et la multiplication $\cdot$ telles que*

1. *les deux triplets $(\mathbb{k}, +, 0)$ et $(\mathbb{k} - \{0\}, \cdot, 1)$ sont des groupes abéliens ;*
2. *la loi de distributivité, $a \cdot (b + c) = a \cdot b + a \cdot c$, est satisfaite pour tout $a, b, c \in \mathbb{k}$.*

Note : On écrira $\mathbb{k}^* = \mathbb{k} - \{0\}$.

Proposition 1.6 *Pour tout $a \in \mathbb{k}$, $a \cdot 0 = 0$.*

Démonstration. Le triplet $(\mathbb{k}, +, 0)$ étant un groupe, nous avons $0 + 0 = 0$. De plus, la loi de distributivité implique $a \cdot 0 + a \cdot 0 = a \cdot 0$. A nouveau la structure de groupe impose $a \cdot 0 = a \cdot 0 + 0$. Par les lois de simplification par rapport à $(\mathbb{k}, +, 0)$ on déduit que $a \cdot 0 = 0$.

□

Proposition 1.7 *Pour tout $a \in \mathbb{k}$, $a \cdot 1 = a$.*

Démonstration. Si $a \neq 0$ alors $a \in \mathbb{k}^*$ qui est un groupe par rapport à la multiplication d'élément neutre 1, d'où $a \cdot 1 = a$. Si $a = 0$ alors le résultat suit par la proposition précédente.

□

Remarques : On peut être moins exigeant par rapport à la structure multiplicative. Par exemple

- Si la structure multiplicative n'est pas commutative alors on dit que $\mathbb{k}$ est un *corps gauche*. Par exemple les *quaternions* (à voir en MAT 2543). Dans ce cas on doit ajouter un axiome supplémentaire de distributivité à droite.
- Si de plus la structure multiplicative n'admet pas nécessairement d'inverse alors on dit que $\mathbb{k}$ est un *anneau*. Par exemple les matrices d'un ordre donné.
- Si la structure multiplicative est tout de même commutative sans nécessairement admettre d'inverse alors on dit que $\mathbb{k}$ est un *anneau commutatif*. Par exemple les entiers $\mathbb{Z}$.

Une propriété intéressante des corps que nous invoquerons à multiple reprises est qu'aucun corps $\mathbb{k}$ ne possède de "diviseur de zéro", c'est-à-dire,

Proposition 1.8 *Soit $a, b \in \mathbb{k}$. Si $ab = 0$ alors $a = 0$ ou $b = 0$.*

Démonstration. Si $b = 0$ alors le résultat est démontré. Supposons $b \neq 0$. Nous avons donc

$$0 = 0 \cdot b^{-1} = (ab) \cdot b^{-1} = a(bb^{-1}) = a \cdot 1 = a.$$

□

1.4.1 Un corps fini : $\mathbb{Z}/2\mathbb{Z}$

Terminons cette section en donnant un exemple d'un corps fini : $\mathbb{Z}/2\mathbb{Z}$. En tant qu'ensemble ce corps contient deux éléments

$$\mathbb{Z}/2\mathbb{Z} = \{0, 1\}.$$

Les deux opérations sont données par les deux *tables de Cayley* suivantes :

$+$	0	1
0	0	1
1	1	0

et

$*$	0	1
0	0	0
1	0	1

On laisse au lecteur de vérifier que cet exemple satisfait aux axiomes d'un corps. Ce corps sera principalement utilisé afin de fournir des contre-exemples et de confondre notre "sense commun".

CHAPITRE 2

ESPACE VECTORIEL

On arrive finalement à la définition de l'objet central de notre étude : *l'espace vectoriel*. Tout au long de ce chapitre on se fixe un corps quelconque $\mathbb{k}$. Nous indiquerons les endroits particuliers où la nature du corps intervient.

2.1 Définitions élémentaires

Définition 2.1 *Un $\mathbb{k}$-espace vectoriel est un ensemble V munit d'une structure de groupe abélien $(+,0)$ et d'une action*

$$\mathbb{k} \times V \to V$$

tel que

1. $(\alpha\beta)\cdot v = \alpha\cdot(\beta\cdot v)$ *pour tout* $\alpha,\beta\in\mathbb{k}$ *et tout* $v\in V$;
2. $(\alpha+\beta)\cdot v = \alpha\cdot v+\beta\cdot v$ *pour tout* $\alpha,\beta\in\mathbb{k}$ *et tout* $v\in V$;
3. $\alpha\cdot(v+w)=\alpha\cdot v+\alpha\cdot w$ *pour tout* $\alpha\in\mathbb{k}$ *et tout* $v,w\in V$; *et*
4. $1\cdot v = v$ *pour tout* $v\in V$.

Les éléments de V sont appelés vecteurs.

Remarque : Nous omettrons le symbole "$\cdot$" lorsque le contexte sera clair.

Exercice : Montrez

1. $-(-v)=v$ pour tout $v\in V$.
2. $-(\alpha v)=(-\alpha)v=\alpha(-v)$ pour tout $\alpha\in\mathbb{k}$ et tout $v\in V$.
3. $0v=0$ pour tout $v\in V$.
4. $(-\alpha)(-v)=\alpha v$ pour tout $\alpha\in\mathbb{k}$ et tout $v\in V$.

5. Si $v, w \in V$, $\alpha \in \mathbb{k}^*$ et $\alpha v = \alpha w$ alors $v = w$.

Voici des exemples importants d'espaces vectoriels que nous retrouverons tout au long de ces notes.

Exemple 1 : $\mathbb{k}^n$

Soit $\mathbb{k}$ un corps et fixons un entier naturel non nul n. L'ensemble des n-tuplets ordonnés

$$\mathbb{k}^n = \{(x_1, x_2, ..., x_n) \,|\, x_i \in \mathbb{k},\, 1 \leq i \leq n\}$$

est munit d'une structure de $\mathbb{k}$-espace vectoriel. En effet, les deux opérations

- $(a_1, a_2, ..., a_n) + (b_1, b_2, ..., b_n) = (a_1 + b_1, a_2 + b_2, ..., a_n + b_n)$ et
- $\lambda \cdot (a_1, a_2, ..., a_n) = (\lambda a_1, \lambda a_2, ..., \lambda a_n)$

satisfont aux axiomes d'un $\mathbb{k}$-espace vectoriel. Le vecteur nul $\overbrace{(0, 0, ..., 0)}^{n\text{–fois}}$ est l'élément neutre de l'addition "+". Notons que deux vecteurs $u, v \in \mathbb{k}^n$ sont égaux s'ils sont égaux composante par composante, c'est-à-dire, si $u = (a_1, a_2, ..., a_n)$ et $v = (b_1, b_2, ..., b_n)$ alors nous avons

$$u = v \quad \Leftrightarrow \quad a_i = b_i,\, \forall i \,|\, 1 \leq i \leq n.$$

L'ordre des composantes est important. En effet les vecteurs $(1, 2, 0)$ et $(2, 1, 0)$ ne sont pas égaux. La démonstration que $\mathbb{k}^n$ est un espace vectoriel est aisé à cause des propriétés du corps $\mathbb{k}$. Nous démontrons le premier axiome de distributivité et laissons les autres en exercice. Soit $\alpha, \beta \in \mathbb{k}$ et $u = (a_1, a_2, ..., a_n) \in \mathbb{k}^n$.

$$\begin{aligned}
(\alpha\beta) \cdot u &= \big((\alpha\beta)a_1, (\alpha\beta)a_2, ..., (\alpha\beta)a_n\big), \quad \text{définition de l'action,} \\
&= \big(\alpha(\beta a_1), \alpha(\beta a_2), ..., \alpha(\beta a_n)\big), \quad \text{associativité de } \mathbb{k}, \\
&= \alpha \cdot \big(\beta a_1, \beta a_2, ..., \beta a_n\big), \quad \text{définition de l'action,} \\
&= \alpha \cdot (\beta \cdot (a_1, a_2, ..., a_n)), \quad \text{définition de l'action,} \\
&= \alpha \cdot (\beta \cdot u).
\end{aligned}$$

Donc $\mathbb{Q}^n, \mathbb{R}^n$, et $\mathbb{C}^n$ sont trois exemples de $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$-espace vectoriel respectivement. Le corps $\mathbb{k}$ (lorsque $n = 1$) est lui-même un $\mathbb{k}$-espace vectoriel.

Exemple 2 : Le produit

Plus généralement, soit $\{U_i\}_{1 \leq i \leq n}$ une famille de $\mathbb{k}$-espaces vectoriels. L'ensemble des n-tuplets ordonnés

$$\prod_{i=1}^{n} U_i = \{(u_1, u_2, ..., u_n) \,|\, u_i \in U_i,\, 1 \leq u \leq n\}$$

munit des deux opérations

- $(u_1, u_2, ..., u_m) + (v_1, v_2, ..., v_m) = (u_1 + v_1, u_2 + v_2, ..., u_m + v_m)$ et
- $\lambda \cdot (u_1, u_2, ..., u_m) = (\lambda u_1, \lambda u_2, ..., \lambda u_m)$,

où $(u_1, u_2, ..., u_m), (v_1, v_2, ..., v_m) \in \prod_{i=1}^{n} U_i$ et $\lambda \in \mathbb{k}$, est une $\mathbb{k}$-espace vectoriel. La démonstration est formellement la-même que pour le cas $\mathbb{k}^n$. Dans les cas où $n = 2$ ou $n = 3$ on utilisera la notation $U \times V$ et $U \times V \times W$.

Exemple 3 : $\mathscr{F}(S, V)$

Soit S un ensemble et V un $\mathbb{k}$-espace vectoriel. On peut munir l'ensemble

$$\mathscr{F}(S, V) = \{f \,|\, f : S \to V\}$$

des fonctions de S dans V d'une structure de $\mathbb{k}$-espace vectoriel. En effet, si $f, g \in \mathscr{F}(S,V)$ et $\alpha \in \mathbb{k}$ on pose

- $(f+g)(x) = f(x) + g(x)$ et
- $(\alpha f)(x) = \alpha \cdot f(x)$.

La fonction constante $\mathscr{O}$ qui associe chaques éléments de S à l'élément neutre 0 de V est l'élément neutre de $\mathscr{F}(S,V)$. On se rappellera que si $f, g \in \mathscr{F}(S,V)$ alors

$$f = g \quad \Leftrightarrow \quad f(x) = g(x),\ \forall x \in S.$$

A présent la démonstration que $\mathscr{F}(S,V)$ est un $\mathbb{k}$-espace vectoriel découle du fait que V est un $\mathbb{k}$-espace vectoriel. On montre le deuxième axiome de distributivité et laissons les autres au lecteur. Soit $\alpha, \beta \in \mathbb{k}$, $f \in \mathscr{F}(S,V)$ et $x \in S$. Alors

$$\begin{aligned}\big((\alpha+\beta)f\big)(x) &= (\alpha+\beta)\cdot f(x), \quad \text{définition de l'action,}\\ &= (\alpha \cdot f(x)) + (\beta \cdot f(x)), \quad \text{distributivité dans } V,\\ &= (\alpha f)(x) + (\beta f)(x), \quad \text{définition de l'action.}\end{aligned}$$

Comme x est arbitraire on a $(\alpha+\beta)f = \alpha f + \beta f$.

Exemple 4 : Espace matriciel

Soit V un $\mathbb{k}$-espace vectoriel et m un entier naturel non nul. Alors l'ensemble des m-tuplets ordonnés

$$\mathbb{M}(V^m) = \{(u_1, u_2, ..., u_m) \mid u_i \in V,\ 1 \leq i \leq m\}$$

est munit de la structure d'espace vectoriel du produit $\prod_{i=1}^{m} V$. Un élément de $\mathbb{M}(V^m)$ est appelé *matrice* et les composantes d'une matrice sont appelés *vecteurs lignes*. Comme dans notre premier exemple, on vérifie l'égalité de deux matrices composante par composante. Dès lors la démonstration que $\mathbb{M}(V^m)$ est un $\mathbb{k}$-espace vectoriel est aisé et découle de la structure de $\mathbb{k}$-espace vectoriel sur V.

Nous avons le cas particulier où $V = \mathbb{k}^n$. Alors nous noterons

$$\mathbb{M}_{m\times n}(\mathbb{k}) = \mathbb{M}\big((\mathbb{k}^n)^m\big).$$

Par exemple, si nous travaillons sur $\mathbb{R}$ et considérons les vecteurs lignes $u_1 = (1,2,3,4)$, $u_2 = (0,-12,0,1)$, et $u_3 = (3,1,-2,0) \in \mathbb{R}^4$ de la matrice $A = (u_1, u_2, u_3)$ nous pouvons former le tableau

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -12 & 0 & 1 \\ 3 & 1 & -2 & 0 \end{pmatrix} \in \mathbb{M}_{3\times 4}(\mathbb{R})$$

que nous identifions à A. La $j^{\text{ième}}$-composante du $i^{\text{ième}}$-vecteur ligne de la matrice A sera notée $a_{ij} \in \mathbb{k}$. Dans notre exemple, $a_{12} = 2$ et $a_{33} = -2$. En général une matrice $B \in \mathbb{M}_{3\times 4}(\mathbb{R})$ se présentera comme suit

$$B = \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \end{pmatrix} \in \mathbb{M}_{3\times 4}(\mathbb{R}).$$

2.2 Les sous-espaces vectoriels

Soit V un $\mathbb{k}$-espace vectoriel et U un sous-ensemble non-vide de V.

Définition 2.2 *On dira que U est un sous-espace vectoriel de V si*

1. *pour tout $x, y \in U$ on a $x + y \in U$, et*
2. *pour tout $x \in U$ et tout $\alpha \in \mathbb{k}$ on a $\alpha x \in U$.*

Proposition 2.1 *Si U est un sous-espace vectoriel de $(V, +, \cdot)$ alors $(U, +, \cdot)$ est un $\mathbb{k}$-espace vectoriel (munit de la même addition et de la même action).*

Démonstration. Comme $1 \cdot x = x$ est vrai dans V, c'est donc vrai pour tout $x \in U$. De la même manière, étant donné que les restrictions de l'addition et de l'action à U demeurent dans U, toutes les conditions de distributivité sont automatiquement vérifiées. Il ne reste qu'à démontrer que $U \leq (V, +)$. Mais la condition (1) est exactement la même que celle de la définition d'un sous-groupe. De plus, il est aisé de voir (exercice) que $-1 \cdot x = -x \in U$ par la condition (2).

□

Remarque : Un sous-espace vectoriel U d'un espace vectoriel V est donc à la base un sous-groupe de V et donc le vecteur nul appartient à U.

Exemple 1 : Soit $U = \{(x, 0) \mid x \in \mathbb{R}\}$. Alors U est un sous-espace vectoriel de $\mathbb{R}^2$. En effet, si $(x, 0), (y, 0) \in U$ alors $(x + y, 0) \in U$. De plus, si $(x, 0) \in U$ et $\alpha \in \mathbb{R}$ alors $\alpha(x, 0) = (\alpha x, 0) \in U$. On remarque que $U \neq \emptyset$ car en particulier $(0, 0) \in U$. Par contre, l'ensemble $S = \{(x, 1) \mid x \in \mathbb{R}\}$ n'est pas un sous-espace vectoriel de $\mathbb{R}^2$ (exercice).

Si le corps de référence est de cardinalité infini alors le seul sous-espace fini est le sous-espace trivial $\{0\}$. En effet, soit $U \neq \{0\}$ un sous-espace et $u \in U$ non nul. On affirme que l'application suivante

$$\begin{array}{ccc} \mathbb{k} & \longrightarrow & U \\ \alpha & \mapsto & \alpha \cdot u \end{array}$$

est injective. Ceci découle de l'exercice (5) suivant la définition d'un espace vectoriel.

Exemple 2 : La trace d'une matrice carrée.

Considérons l'application

$$\begin{array}{ccc} \mathrm{tr} : \mathbb{M}_{n\times n}(\mathbb{k}) & \longrightarrow & \mathbb{k} \\ \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} & \mapsto & \sum_{i=1}^{n} a_{ii}. \end{array}$$

On peut montrer facilement (exercice) que pour $A, B \in \mathbb{M}_{n\times n}(\mathbb{k})$ et $\lambda \in \mathbb{k}$ nous avons toujours $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$ et $\mathrm{tr}(\lambda A) = \lambda \cdot \mathrm{tr}(A)$. Il en découle donc que

$$Tr_n = \{A \in \mathbb{M}_{n\times n}(\mathbb{k}) \mid \mathrm{tr}(A) = 0\}$$

est une sous-espace vectoriel de $\mathbb{M}_{n\times n}(\mathbb{k})$. On voit que $Tr_1 = \{0\}$ tandis que

$$Tr_2 = \left\{ \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \mid a, b, c \in \mathbb{k} \right\}.$$

2.3 Les combinaisons linéaires

Soit V un $\mathbb{k}$-espace vectoriel et $u_1, u_2, ..., u_n \in V$.

Définition 2.3 *Une* combinaison linéaire *des vecteurs $u_1, u_2, ..., u_n$ est un vecteur $v \in V$ tel que*

$$v = a_1 u_1 + a_2 u_2 + ... + a_n u_n$$

pour $a_i \in \mathbb{k}$, $1 \leq i \leq n$.

Remarque : Comme l'addition est associative le résultat d'une combinaison linéaire n'est pas ambiguë. On utilisera aussi la notation plus compacte $v = \sum_{i=1}^{n} a_i u_i$.

Proposition 2.2 *Si $U \subset V$ est un sous-espace vectoriel alors toute combinaison linéaire d'éléments de U est un élément de U.*

Démonstration. On procède à l'aide d'un raisonnement par récurrence sur la longueur d'une combinaison linéaire. Si $u \in U$ et $\lambda \in \mathbb{k}$ alors $\lambda \cdot u \in U$ car U est un sous-espace. L'ancrage de l'induction est donc démontré. On suppose que les combinaisons linéaires d'éléments de U de longueurs $n-1$ appartiennent à U. Soit $\sum_{i=1}^{n} a_i u_i$ avec $u_i \in U$ et $a_i \in \mathbb{k}$, $1 \leq i \leq n$. On a

$$\sum_{i=1}^{n} a_i u_i \quad = \quad \left(a_1 u_1 + \underbrace{\sum_{i=2}^{n} a_i u_i}_{\in U \text{ par induction}} \right) \in U, \quad \text{car } U \text{ est un sous-espace}$$

□

Soit $S \subseteq V$ pas nécessairement un sous-espace. L'ensemble de toutes les combinaisons linéaires obtenues à partir des éléments de S sera noté

$$< S > .$$

Par convention on imposera $< \emptyset >= \{0\}$. Il est important de remarquer qu'une combinaison linéaire ne contient, par définition, qu'un nombre **fini** de termes. Si la cardinalité de S est infini alors aucune combinaison linéaire ne peut être composée de tous les éléments de S.

Proposition 2.3 *Soit $S \subseteq V$. Alors $< S >$ est un sous-espace vectoriel de V. C'est le sous-espace engendré par S.*

Démonstration. Si $S = \emptyset$ alors la proposition est démontrée. Supposons $S \neq \emptyset$ et soit $v = \sum_{i=1}^{m} a_i v_i$ et $u = \sum_{j=1}^{m} b_j u_j$, où $v_i, u_j \in S$, $a_i, b_j \in \mathbb{k}$, $1 \leq i \leq n$, $1 \leq j \leq m$. Soit $k = |\{v_1, v_2, ..., v_n\} \cap \{u_1, u_2, ..., u_m\}|$. On peut supposer sans perte de généralité que $v_1 = u_1$, ..., $v_k = u_k$. Dans ce cas, on a

$$v + u = \sum_{i=1}^{k} \underbrace{(a_i + b_i)}_{\in \mathbb{k}} v_i + \sum_{i=k+1}^{n} a_i v_i + \sum_{i=k+1}^{m} b_i u_i,$$

ce qui est à nouveau une combinaison linéaire d'éléments de S et donc $v+u \in < S >$. Si $\lambda \in \mathbb{k}$ alors

$$\lambda \cdot v = \underbrace{\lambda \cdot \sum_{i=1}^{n} a_i v_i = \sum_{i=1}^{n} \lambda \cdot (a_i v_i)}_{\text{distributivité généralisée}} = \sum_{i=1}^{n} \underbrace{(\lambda a_i)}_{\in \mathbb{k}} v_i,$$

ce qui est à nouveau une combinaison linéaire d'éléments de S et donc $\lambda \cdot v \in < S >$. On conclut que $< S >$ est un sous-espace vectoriel de V.

□

Soit J un ensemble non-vide et supposons que pour tout élément α de J est associé un sous-espace U_α de V. On dit alors que l'on a une famille de sous-espaces $\{U_\alpha\}_{\alpha \in J}$ *indexée* par l'ensemble J. On peut alors former le sous-ensemble

$$\bigcap_{\alpha \in J} U_\alpha = \{v \in V \mid v \in U_\alpha, \forall \alpha \in J\}.$$

Proposition 2.4 *Le sous-ensemble* $\bigcap_{\alpha \in J} U_\alpha$ *est un sous-espace.*

Démonstration. Si $\bigcap_{\alpha \in J} U_\alpha = \{0\}$ alors la proposition est démontrée. Sinon soit $u, v \in \bigcap_{\alpha \in J} U_\alpha$. En d'autres mots, $u, v \in U_\alpha$ pour tout $\alpha \in J$. Comme chacun des U_α est un sous-espace, $u+v \in U_\alpha$ pour tout $\alpha \in J$, c'est-à-dire, $u+v \in \bigcap_{\alpha \in J} U_\alpha$. La deuxième condition pour être un sous-espace se démontre de la même façon.

□

Soit $S \subset V$ et $J = \{U \subset V \mid U$ est un sous-espace de V et $S \subset U\}$. Un *plus petit élément* de J est un sous-espace U contenant S tel que s'il existe un autre sous-espace W contenant S et tel que $W \subset U$ alors $W = U$. On notera que $J \neq \emptyset$ car $V \in J$. De plus, comme l'ordre induit par l'inclusion n'est que partiel, il n'y a pas de raison à priori qu'un plus petit élément soit unique. Par contre,

Proposition 2.5 *L'unique plus petit élément de* J *est donné par* $\bigcap_{U \in J} U$.

Démonstration. Par la Proposition 2.4, $\bigcap_{U \in J} U$ est un sous-espace. De plus, comme $S \subset U$ pour tout $U \in J$, $S \subset \bigcap_{U \in J} U$, c'est-à-dire, $\bigcap_{U \in J} U \in J$. On remarque, par construction, que $\bigcap_{U \in J} U \subset U$ pour tout $U \in J$. Si $W \in J$ et $W \subset \bigcap_{U \in J} U$ alors par la remarque précédente, $W = \bigcap_{U \in J} U$. Finalement, si W est un autre plus petit élément de J alors comme W appartient à J on doit avoir $\bigcap_{U \in J} U \subset W$. Mais W est un plus petit élément, on a donc $\bigcap_{U \in J} U = W$.

□

Proposition 2.6 *Le plus petit sous-espace contentant un sous-ensemble* $S \subset V$ *est* $< S >$.

Démonstration. Si $S = \emptyset$ alors $\{0\}$ est effectivement le plus petit sous-espace contenant S et la proposition est démontrée. Sinon soit $u \in S$. Comme $1 \cdot u$ est une combinaison linéaire et que $1 \cdot u = u$ (dernier axiome d'un espace vectoriel), on a $S \subset < S >$. Si U est un sous-espace de V contenant S alors, par la

Proposition 2.2, $< S > \subset U$. En particulier, $< S > \subset \bigcap_{U \in J} U$. Le résultat est la conséquence de la Proposition 2.5.

□

Si U et W sont deux sous-espaces vectoriels de V. Que peut-on affirmer à propos de $U \cup W$? Est-ce un sous-espace ?

Proposition 2.7 *L'union $U \cup W$ est un sous-espace si et seulement si $U \subset W$ ou $W \subset U$.*

Démonstration. Sans perte de généralité supposons que $U \subset W$. Il s'en suit que $U \cup W = W$ et est donc un sous-espace. Supposons qu'il existent $w \in W - U$ et $u \in U - W$. Si $U \cup W$ est un sous-espace alors $u + w \in U \cup W$. Par définition de l'union, $u + w \in U$ ou $u + w \in W$. Sans perte de généralité, supposons que $u + w \in U$, c'est-à-dire, qu'il existe un $u' \in U$ tel que $u + w = u'$. Mais dès lors $w = u' - u \in U$ ce qui contredit la nature de w.

□

On peut toujours considérer le sous-espace $< U \cup W >$. Comme on l'a vu c'est le plus petit sous-espace contenant à la fois U et W. En contre partie, le fait que U et W soient des sous-espaces n'influence en rien la structure de $< U \cup W >$. Par exemple, soient $U = \{(x,0) \mid x \in \mathbb{R}\}$ et $W = \{(0,y) \mid y \in \mathbb{R}\}$. Ce sont deux sous-espaces de $\mathbb{R}^2$ (exercice), l'axe des "x" et des "y" respectivement. D'une part, $U \cup W$ n'est que le sous-ensemble "croix" de $\mathbb{R}^2$ et n'est pas un sous-espace. En effet le vecteur $(1,0) + (0,1) = (1,1)$ n'appartient ni à U ni à W et donc n'appartient pas à $U \cup W$. D'autre part, $< U \cup W > \subset \mathbb{R}^2$ par construction. De plus, si $(x,y) \in \mathbb{R}^2$ alors $(x,y) = x(1,0) + y(0,1)$ qui est une combinaison linéaire d'éléments de $U \cup W$. On en tire que $\mathbb{R}^2 = < U \cup W >$.

Les idées du dernier paragraphe suggère la construction suivante : soit U et W deux sous-espaces vectoriels de V et considérons l'ensemble

$$U + W = \{u + w \mid u \in U \text{ et } w \in W\}.$$

Proposition 2.8 *Nous avons $U + W = < U \cup W >$.*

Démonstration. On laisse les détails au lecteur.

□

Remarque : Lorsque $U \cap W = \{0\}$ on dira que $U + W$ est la somme *directe* des deux sous-espaces et nous la noterons $U \oplus W$.

Dans le dernier exemple, considérer l'axe des "'x", U, et l'axe des "y", W, afin d'engendrer $\mathbb{R}^2$ semble excessif. En fait notre calcul montre que

$$\mathbb{R}^2 = < (1,0), (0,1) > .$$

Ceci nous mène naturellement vers la question : peut-on trouver un ensemble *minimal* engendrant un sous-espace donné ?

2.4 Générateurs et indépendance linéaire

Soit U un sous-espace vectoriel de V tel que $U = < S >$ pour un certain sous-ensemble $S \subset V$. On appelle les éléments de S les *générateurs* de U.

Soit $u_1,...,u_n \in V$. On a toujours $< u_1,...,\widehat{u_i},...,u_n > \subset < u_1,...,u_n >$, où $\widehat{u_i}$ signifie que l'on omet l'élément u_i de la liste. Mais si u_i est en fait une combinaison linéaire des éléments $u_1,...,\widehat{u_i},...,u_n$ alors

$$< u_1,...,\widehat{u_i},...,u_n > = < u_1,...,u_n > .$$

En effet, si $u_i = \sum_{j\neq i} a_j u_j$ et que $v = \sum_{j=1}^n \alpha_j u_j \in < u_1,...,u_n >$ alors

$$\begin{aligned} v &= \sum_{j=1}^{i-1} \alpha_j u_j + \alpha_i \underbrace{\left(\sum_{j\neq i} a_j u_j \right)}_{u_i} + \sum_{j=i+1}^{n} \alpha_j u_j \\ &= \sum_{j\neq i} (\alpha_j + \alpha_i a_j) u_j. \end{aligned}$$

La dernière expression clairement appartient à $< u_1,...,\widehat{u_i},...,u_n >$. En d'autres termes,

Théorème 2.9 *Soit $u_1, u_2, ..., u_n \in V$. Il existe un vecteur u_i parmi cette liste qui est combinaison linéaire des $u_1,...,\widehat{u_i},...,u_n$ si et seulement si il existent des scalaires $\mu_1, \mu_2, ..., \mu_n \in \mathbb{k}$ non tous nuls tels que*

$$0 = \mu_1 u_1 + \mu_2 u_2 + ... + \mu_n u_n.$$

Démonstration. Si $u_i = \sum_{j\neq i} \alpha_j u_j$ alors

$$0 = \alpha_1 u_1 + ... + \alpha_{i-1} u_{i-1} + (-1) \cdot u_i + \alpha_{i+1} u_{i+1} + ... + \alpha_n u_n.$$

En posant

$$\mu_j = \begin{cases} \alpha_j & j \neq i \\ -1 & j = i \end{cases},$$

$1 \leq j \leq n$, la première implication est démontrée. S'il existe des scalaires $\mu_1, \mu_2, ..., \mu_n \in \mathbb{k}$ non tous nuls tels que $0 = \mu_1 u_1 + \mu_2 u_2 + ... + \mu_n u_n$ alors il existe $\mu_i \neq 0$, pour un certain $1 \leq i \leq n$, et nous pouvons écrire

$$u_i = -\mu_i^{-1} \left\{ \sum_{j\neq i} \mu_j u_j \right\} = \sum_{j\neq i} (-\mu_i^{-1} \mu_j) u_j,$$

car $\mathbb{k}$ est un corps. En d'autres mots, $u_i \in < u_1,...,\widehat{u_i},...,u_n >$.

□

Ce denier résultat nous amène à définir

Définition 2.4 *Un sous-ensemble $S \subset V$ est* linéairement dépendant *s'il existent $u_1,...,u_n \in S$ et des scalaires $\mu_1,...,\mu_n \in \mathbb{k}$ non tous nuls tels que*

$$0 = \mu_1 u_1 + ... + \mu_n u_n.$$

On dira que S est *linéairement indépendant* s'il n'est pas possible de trouver de telle relation. En d'autres mots, S sera dit linéairement indépendant si

$$0 = \mu_1 u_1 + ... + \mu_n u_n \quad \Rightarrow \quad \mu_1 = \mu_2 = ... = \mu_n = 0$$

pour toute famille finie $u_1,...,u_n \in S$.

Remarque : Tout sous-ensemble contenant le vecteur nul est linéairement dépendant. En effet, $0_V = 1_{\mathbb{k}} \cdot 0_V$ est une relation de dépendance. En particulier, tout sous-espace est linéairement dépendant.

Proposition 2.10 *Soit $S \subset < u_1, ..., u_n > \subset V$. Si $|S| > n$ alors S est linéairement dépendant.*

Démonstration. On procède par induction sur n. Si $n = 1$ alors il existent au moins deux éléments distincts $v_1, v_2 \in S$ et $S \subset < u >$ pour un certain $u \in V$. Si l'un deux est zéro S est nécessairement linéairement dépendant. Sinon $v_1 = \lambda_1 u$ et $v_2 = \lambda_2 u$ pour $\lambda_1, \lambda_2 \in \mathbb{k}^*$. On a donc

$$\lambda_1^{-1} v_1 = u = \lambda_2^{-1} v_2,$$

c'est-à-dire, $0 = \lambda_1^{-1} v_1 + (-\lambda_2^{-1}) v_2$ une relation de dépendance non-triviale. L'ensemble S est donc linéairement dépendant et l'ancrage de l'induction est donc vérifiée. On suppose le résultat vérifié pour tout sous-ensemble $S' \subset < u_1, ..., \widehat{u_i}, ..., u_n >$ et pour tout $1 \leq i \leq n$ tel que $|S'| > n-1$. Par hypothèse, il existent au moins $v_1, .., v_{n+1} \in S$ éléments distincts. Si le vecteur nul se trouve parmi eux dès lors S est linéairement dépendant. Sinon on a $n+1$ relations non-triviales

$$\begin{aligned} v_1 &= a_{11} u_1 + a_{12} u_2 + ... + a_{1n} u_n \\ v_2 &= a_{21} u_1 + a_{22} u_2 + ... + a_{2n} u_n \\ &\vdots \\ v_{n+1} &= a_{n+11} u_1 + a_{n+12} u_2 + ... + a_{n+1\,n} u_n. \end{aligned}$$

Si chacun des coefficients $a_{j1} = 0$, pour $1 \leq j \leq n+1$, alors $\{v_1, ..., v_{n+1}\} \subset < \widehat{u_1}, ..., u_n >$ et donc $\{v_1, ..., v_{n+1}\}$ et par conséquent S serait linéairement dépendant par l'hypothèse d'induction. Sans perte de généralité, supposons que $a_{11} \neq 0$. Le coefficient du vecteur u_1 dans l'expression du vecteur $v_2 - a_{21} a_{11}^{-1} v_1$ est nul. En effet,

$$\begin{aligned} v_2 - a_{21} a_{11}^{-1} v_1 &= a_{21} u_1 + a_{22} u_2 + ... + a_{2n} u_n \\ &\quad - a_{21} a_{11}^{-1} \Big(a_{11} u_1 + a_{12} u_2 + ... + a_{1n} u_n \Big) \\ &= a'_{22} u_2 + ... + a'_{2n} u_n, \end{aligned}$$

pour de nouveaux coefficient $a'_{2i} \in \mathbb{k}$, pour $2 \leq i \leq n$. De même,

$$\begin{aligned} v_3 - a_{31} a_{11}^{-1} v_1 &= a'_{32} u_2 + ... + a'_{3n} u_n \\ &\vdots \\ v_{n+1} - a_{n+11} a_{11}^{-1} v_1 &= a'_{n+12} u_2 + ... + a'_{n+1\,n} u_n. \end{aligned}$$

Ce calcul démontre que $\{v_2 - a_{21} a_{11}^{-1} v_1, ..., v_{n+1} - a_{n+11} a_{11}^{-1} v_1\} \subset < \widehat{u_1}, ..., u_n >$ et donc par l'hypothèse d'induction $\{v_2 - a_{21} a_{11}^{-1} v_1, ..., v_{n+1} - a_{n+11} a_{11}^{-1} v_1\}$ est linéairement dépendant. Il existent donc $\mu_2, ..., \mu_{n+1} \in \mathbb{k}$ non tous nuls tels que

$$\begin{aligned} 0 &= \mu_2 (v_2 - a_{21} a_{11}^{-1} v_1) + ... + \mu_{n+1} (v_{n+1} - a_{n+11} a_{11}^{-1} v_1) \\ &= -a_{11}^{-1} \left(\sum_{j=2}^{n+1} \mu_j a_{j1} \right) v_1 + \mu_2 v_2 + ... + \mu_{n+1} v_{n+1}. \end{aligned}$$

La dernière relation étant une relation de dépendance non triviale force donc $\{v_1, ..., v_{n+1}\}$ et donc S à être linéairement dépendant.

□

Corollaire 2.11 *Soient* $\{u_1,...,u_n\}$ *et* $\{v_1,...,v_m\}$ *deux sous-ensembles linéairement indépendants d'un espace vectoriel* V. *Si* $<u_1,...,u_n>=<v_1,...,v_m>$ *alors* $n=m$.

Démonstration. Par hypothèse nous avons $\{v_1,...,v_m\}\subset<v_1,...,v_m>=<u_1,...,u_n>$. De plus, comme $\{v_1,...,v_m\}$ est linéairement indépendant, la Proposition 2.10 force $m\leq n$. Le résultat s'en suit en échangeant le rôle des ensembles générateurs.

□

2.5 Les bases et la dimension

Définition 2.5 *On dit qu'un sous-ensemble* S *d'un* $\mathbb{k}$*-espace vectoriel* V *est une* base *de* V *si*

1. $V=<S>$, *et*
2. S *est linéairement indépendant.*

Remarque : Par convention, on considérera $\emptyset$ comme base de $\{0\}$.

Définition 2.6 *Un sous-espace vectoriel* U *est* finiment engendré *s'il existe un sous-ensemble* $S\subset V$ ***fini*** *tel que*

$$U=<S>.$$

Théorème 2.12 *Tout* $\mathbb{k}$*-espace vectoriel finiment engendré possède une base.*

Démonstration. Par hypothèse, il existe S un sous-ensemble fini de V tel que $V=<S>$. Considérons l'ensemble

$$D=\{|R|\,|\,R\subset S \text{ et } V=<R>\}.$$

On remarque deux choses : (1) $D\subset\mathbb{N}$ et (2) $D\neq\emptyset$ car $|S|\in D$. Le principe du bon ordre nous garantie l'existence de $n=\min D$. Si $n=0$ alors $V=\{0\}$ et possède $\emptyset$ comme base. Sinon $n>0$ et il existe $R\subset S$ tel que $|R|=n$ et $V=<R>$. Si R n'est pas linéairement indépendant alors l'un de ses éléments est combinaison linéaire des autres par la Proposition 2.9. Soit $r\in R$ cet élément. Mais dans ce cas nous avons $V=<R>=<R-\{r\}>$ et $|R-\{r\}|=n-1$ ce qui contredit la nature minimale de n. On conclut que R est linéairement indépendant et donc R est une base de V.

□

Proposition 2.13 *Si* U *est un sous-espace vectoriel d'un espace vectoriel* V *finiment engendré alors* U *possède une base et en particulier il est finiment engendré.*

Démonstration. Soit S un ensemble fini de V tel que $<S>=V$. Par le Théorème 2.12, on peut supposer que S est une base de V. En particulier, S est linéairement indépendant. Considérons l'ensemble

$$D=\{|S|-|R|\,|\,R\subset U \text{ et } R \text{ est linéairement indépendant}\}.$$

On remarque deux choses : (1) $D\subset\mathbb{N}$ car $|S|\geq|R|$ selon la Proposition 2.10 et (2) $D\neq\emptyset$ car $|S|\in D$. Il faut simplement prendre $R=\emptyset$. Le principe du bon ordre nous garantie l'existence de $n=\min D$. Soit $R\subset U$ tel que $|R|=|S|-n$ et R soit linéairement indépendant. S'il existe $w\in U-<R>$ non nul alors nécessairement $R'=R\cup\{w\}$ est linéairement indépendant et est tel que $|R'|=|R|+1$ contredisant la nature minimal de n. On conclut que $U=<R>$.

□

Proposition 2.14 *Toutes les bases d'un espace vectoriel finiment engendré possèdent le même nombre d'éléments.*

Démonstration. Par hypothèse, l'espace vectoriel V admet un sous-ensemble S fini tel que $V =< S >$. Par le Théorème 2.12, V possède au moins une base et la cardinalité de celle-ci est bornée supérieurement par $|S|$. Le résultat est maintenant une conséquence directe du Corollaire 2.11 et du fait qu'une base est un ensemble générateur.

□

Enfin on peut définir

Définition 2.7 *Soit V un $\mathbb{k}$-espace vectoriel finiment engendré et B une base de V. La* dimension *de V est l'entier*

$$\dim_{\mathbb{k}} V = |B|.$$

Remarque : La dimension dépend du corps sur lequel nous travaillons. En effet, $\mathbb{C}$ est un $\mathbb{C}$-espace vectoriel. En particulier, l'ensemble $\{1\}$ est une base de $\mathbb{C}$ car cet ensemble est linéairement indépendant et tout nombre complexe s'écrit $z = z \cdot 1$, d'où

$$\dim_{\mathbb{C}} \mathbb{C} = 1.$$

Par contre, on peut considérer $\mathbb{C}$ come une $\mathbb{R}$-espace vectoriel. En effet, tout nombre complexes s'écrit $z = a \cdot 1 + b \cdot i$, où $a, b \in \mathbb{R}$. Dans ce cas-ci, $\{1, i\}$ est $\mathbb{R}$-linéairement indépendant et $\mathbb{C} =< 1, i >$. D'où

$$\dim_{\mathbb{R}} \mathbb{C} = 2.$$

Exemple : Montrons que $\dim_{\mathbb{k}} \mathbb{k}^n = n$. Dans $\mathbb{k}^n$ nous avons les éléments canoniques $e_1 = (1, 0, ..., 0)$, $e_2 = (0, 1, 0, ..., 0)$, ..., $e_n = (0, ..., 0, 1)$. Si $(a_1, a_2, ..., a_n) \in \mathbb{k}^n$ alors

$$\begin{aligned} (a_1, a_2, ..., a_n) &= a_1 e_1 + a_2 e_2 + ... + a_n e_n \\ &\in < e_1, e_2, ..., e_n >, \end{aligned}$$

c'est-à-dire, $\mathbb{k}^n =< e_1, e_2, ..., e_n >$. S'il existe $\mu_1, ..., \mu_n \in \mathbb{k}$ tels que

$$(0, ..., 0) = \mu_1 e_1 + ... + \mu_n e_n$$

alors $(0, ..., 0) = \mu_1 e_1 + ... + \mu_n e_n = (\mu_1, ..., \mu_n)$ implique que $\mu_1 = \mu_2 = ... = \mu_n = 0$, c'est-à-dire, $\{e_1, e_2, ..., e_n\}$ est linéairement indépendant et est donc une base de $\mathbb{k}^n$. On conclut que $\dim_{\mathbb{k}} \mathbb{k}^n = n$. Cette base est appelée *la base canonique de* $\mathbb{k}^n$.

Proposition 2.15 *Soit $U \subset V$ un sous-espace vectoriel finiment engendré. Si $\{u_1, ..., u_k\}$ est une famille de vecteurs de U linéairement indépendante alors elle peut être étendue en une base de U.*

Démonstration. Par hypothèse, il existe $S \subset U$ un sous-ensemble fini tel que $< S >= U$. Comme U est un sous-espace on a

$$< \{u_1, ..., u_k\} \cup S >= U.$$

Procédons comme dans le cas du Théorème 2.12 et considérons l'ensemble

$$D = \{|R| \, | \, R \subset S \text{ et } U =< \{u_1, ..., u_k\} \cup R >\}.$$

On remarque deux choses : (1) $D \subset \mathbb{N}$ et (2) $D \neq \emptyset$ car $|S| \in D$. Le principe du bon ordre nous garantie l'existence de $n = \min D$. Si $n = 0$ alors $U =< \{u_1, ..., u_k\} >$ et $\{u_1, ..., u_k\}$ est déjà une base de U. Sinon $n > 0$ et par définition de n il existe $R \subset S$ tel que $|R| = n$ et $U =< \{u_1, ..., u_k\} \cup R >$. Si $\{u_1, ..., u_k\} \cup R$ est linéairement dépendant alors l'un de ses éléments est combinaison linéaire des autres par la Proposition 2.9. Si $v \in R$ est combinaison linéaire des autres alors $R' = R - \{v\}$ est tel que $U =< \{u_1, ..., u_k\} \cup R' >$ et $|R'| = n-1$ ce qui contredit la nature minimale de n. Donc l'un des u_i doit être combinaison linéaire des autres, c'est-à-dire,

$$u_i = \sum_{j \neq i} \alpha_j u_j + \sum_{v \in R} a_v v,$$

pour des $\alpha_j, a_v \in \mathbb{k}$, $j \neq i$, non tous nuls. Nécessairement l'un de a_v doit être non nul car $\{u_1, ..., u_k\}$ est linéairement indépendant. Mais dès lors on peut écrire

$$v = a_v^{-1} \left\{ u_i - \sum_{j \neq i} \alpha_j u_j - \sum_{w \in R - \{v\}} a_w w \right\},$$

et à nouveau on a construit un $R' = R - \{v\}$ tel que $U =< \{u_1, ..., u_k\} \cup R' >$ et $|R'| = n - 1$. On conclut que $\{u_1, ..., u_k\} \cup R$ est linéairement indépendant et est donc une extension de $\{u_1, ..., u_k\}$ à une base de U.

□

CHAPITRE 3

APPLICATION LINÉAIRE

Dans ce chapitre nous étudierons les applications entre espaces vectoriels. On peut souvent tirer plus d'information de celles-ci que des simples espaces mis en relation par ces applications. C'est une idée que nous retrouverons partout en mathématique. Considérons l'exemple suivant. Soit l'ensemble

$$\mathscr{P}_n(\mathbb{k}) = \{a_n x^n + a_{n-1} x^{n-1} + ... + a_1 x + a_0 \,|\, a_i \in \mathbb{k}, 0 \leq i \leq n\}$$

des polynômes de degré au plus $n \in \mathbb{N}$. Si nous déclarons que deux polynômes sont égaux si et seulement si ils sont égaux terme-à-terme (d'une même puissance de x), et si nous définissons la somme de deux polynômes comme la somme terme-à-terme et que nous faisons de même avec l'action de $\mathbb{k}$ sur $\mathscr{P}_n(\mathbb{k})$, on voit que (intuitivement) $\mathscr{P}_n(\mathbb{k})$ peut être identifié à $\mathbb{k}^{n+1}$, les puissances de x n'étant là que pour maintenir l'ordre relatif des coefficients a_i, $0 \leq i \leq n$. Ces deux ensembles ne sont pas égaux car ni l'un ni l'autre n'étant sous-ensemble de l'autre. Que voulons-nous dire lorsque nous affirmons : "peut être identifié à" ? Et bien nous pouvons écrire une application "naturelle"

$$\begin{array}{rcl} \Phi : \mathscr{P}_n(\mathbb{k}) & \longrightarrow & \mathbb{k}^{n+1} \\ a_n x^n + ... + a_1 x + a_0 & \mapsto & (a_0, a_1, ..., a_n), \end{array} \tag{3.1}$$

qui est clairement une bijection. Par contre un espace vectoriel est plus riche en structure que le simple ensemble qui le sous-tend. Y-a-t-il un lien entre $\Phi(p) + \Phi(q)$ et $\Phi(p+q)$? ou y-a-t-il un lien entre $\Phi(\lambda p)$ et $\lambda\Phi(p)$? Ces questions nous mène vers la prochaine section.

3.1 Les applications linéaires

Soit $T : V \to W$ une application entre deux $\mathbb{k}$-espaces vectoriels V et W.

Définition 3.1 *On dit que T est linéaire si pour tout $u, v \in V$ et tout $\lambda \in \mathbb{k}$*

1. $T(u+v) = T(u) + T(v)$, *et*
2. $T(\lambda v) = \lambda T(v)$.

Voici deux propriétés élémentaires des applications linéaires.

Proposition 3.1 *Soit $T : V \to W$ une application linéaire entre deux $\mathbb{k}$-espaces vectoriels V et W. Nous avons alors*

- $T(0) = 0$,
- $T(-v) = -T(v)$.

Démonstration. D'une part, $T(0) + T(0) = T(0+0) = T(0) = T(0) + 0$ et on conclut par la propriété de simplification à gauche des groupes que $T(0) = 0$. D'autre part, $0 = T(0) = T(v + (-v)) = T(v) + T(-v)$ et donc par l'unicité de l'inverse des groupes on doit conclure que $T(-v) = -T(v)$.

□

3.2 Le noyau et l'image

Deux sous-ensembles fondamentaux sont associés à toute application linéaire $T : V \to W$, c'est-à-dire,

1. Le noyau de T, $\ker T = \{v \in V \mid T(v) = 0\}$ et
2. L'image de T, $\operatorname{Im} T = \{w \in W \mid \exists v \in V \text{ tel que } T(v) = w\}$.

Proposition 3.2 *Les deux sous-ensembles* $\ker T$ *et* $\operatorname{Im} T$ *sont des sous-espaces de V et W respectivement.*

Démonstration. Soit $u, v \in \ker T$ et $\lambda \in \mathbb{k}$. La linéarité de T nous donne d'une part $T(u+v) = T(u) + T(v) = 0 + 0 = 0$ et donc $u + v \in \ker T$, et d'autre part $T(\lambda u) = \lambda T(u) = \lambda 0 = 0$ et donc $\lambda u \in \ker T$. Soit $w, \omega \in \operatorname{Im} T$. Par définition, il existent $u, v \in V$ tels que $T(u) = w$ et $T(v) = \omega$. A nouveau la linéarité de T nous donne $w + \omega = T(u) + T(v) = T(u+v)$ et donc $w + \omega \in \operatorname{Im} T$. De plus, $\lambda w = \lambda T(u) = T(\lambda u)$, c'est-à-dire, $\lambda w \in \operatorname{Im} T$.

□

Voici un premier théorème de structure très important :

Théorème 3.3 *Soit $T : V \to W$ une application linéaire entre deux $\mathbb{k}$-espaces vectoriels V et W. Supposons que V soit finiment engendré. Nous avons alors*

$$\dim V = \dim \ker T + \dim \operatorname{Im} T.$$

Démonstration. Par définition, $\ker T$ est un sous-espace de l'espace vectoriel finiment engendré V et possède donc une base. Soit $\{u_1, ..., u_k\}$ une base de $\ker T$. Par la Proposition 2.15, on peut étendre cette base en une base de V. Soit $\{u_1, ..., u_k, v_{k+1}, ..., v_n\}$ une extension à une base de V. Considérons le sous-espace $< T(v_{k+1}), T(v_{k+2}), ..., T(v_n) > \subset \operatorname{Im} T$. S'il existent des $\mu_{k+1}, ..., \mu_n \in \mathbb{k}$ non tous nuls tels que

$$\mu_{k+1} T(v_{k+1}) + ... + \mu_n T(v_n) = 0,$$

alors, par linéarité de T, nous avons $T(\mu_{k+1} v_{k+1} + ... + \mu_n v_n) = 0$, c'est-à-dire, $\mu_{k+1} v_{k+1} + ... + \mu_n v_n \in \ker T$. Il doit donc exister $\lambda_1, ..., \lambda_k \in \mathbb{k}$ tels que $\lambda_1 u_1 + ... + \lambda_k u_k = \mu_{k+1} v_{k+1} + ... + \mu_n v_n$ c'est-à-dire,

$$0 = \lambda_1 u_1 + ... + \lambda_k u_k + (-\mu_{k+1}) v_{k+1} + ... + (-\mu_n) v_n.$$

Mais $\{u_1,...,u_k,v_{k+1},...,v_n\}$ est linéairement indépendant ce qui force tous les coefficients à être nuls. On conclut que $\{T(v_{k+1}),T(v_{k+2}),...,T(v_n)\}$ est linéairement indépendant. Soit $w \in \operatorname{Im} T$. Par définition, il existe $v \in V$ tel que $T(v) = w$. On peut donc écrire $v = \alpha_1 u_1 + ... + \alpha_k u_k + \alpha_{k+1} v_{k+1} + ... + \alpha_n v_n$ car $\{u_1,...,u_k,v_{k+1},...,v_n\}$ est une base de V. Et donc

$$\begin{aligned} w &= T(v) = T(\alpha_1 u_1 + ... + \alpha_k u_k + \alpha_{k+1} v_{k+1} + ... + \alpha_n v_n) \\ &\quad \text{est nul par définition du noyau} \\ &= \underbrace{\overbrace{\alpha_1 T(u_1) + ... + \alpha_k T(u_k)} \qquad + \alpha_{k+1} T(v_{k+1}) + ... + \alpha_n T(v_n)}_{\text{par la linéarité de } T} \\ &= \alpha_{k+1} T(v_{k+1}) + ... + \alpha_n T(v_n) \\ &\in < T(v_{k+1}), T(v_{k+2}), ..., T(v_n) > . \end{aligned}$$

Le résultat s'en suit en dénombrant chaque base associée, c'est-à-dire, $\dim V = n$, $\dim \ker T = k$ et $\dim \operatorname{Im} T = n - k$.

□

Définition 3.2 *Soit $T : V \to W$ une application linéaire. Le rang de l'application T est l'entier naturel*

$$\operatorname{rk}(T) = \dim \operatorname{Im} T.$$

Proposition 3.4 *Soit $T : V \to W$ une application linéaire entre deux $\mathbb{k}$-espaces vectoriels V et W. L'espace vectoriel V est finiment engendré si et seulement si les deux sous-espaces* $\ker T$ *et* $\operatorname{Im} T$ *sont finiment engendrés.*

Démonstration. Supposons que $\ker T$ et $\operatorname{Im} T$ soient tous deux finiment engendrés. Nous allons exhibé une base finie pour V (ce qui est plus que demandé). Par le Théorème 2.12, ils existent $\{v_1,...,v_k\} \subset V$ et $\{w_{k+1},...,w_n\} \subset W$ des bases de $\ker T$ et $\operatorname{Im} T$ respectivement. Par définition de $\operatorname{Im} T$, il existe $\{v_{k+1},...,v_n\} \subset V$ tel que $T(v_i) = w_i$, pour $k+1 \leq i \leq n$. S'ils existent des $\mu_i \in \mathbb{k}$, $1 \leq i \leq n$ tels que

$$0 = \underbrace{\mu_1 v_1 + ... + \mu_k v_k}_{\in \ker T} + \mu_{k+1} v_{k+1} + ... + \mu_n v_n$$

alors d'une part $0 = T(\mu_1 v_1 + ... + \mu_k v_k + \mu_{k+1} v_{k+1} + ... + \mu_n v_n) = \mu_{k+1} T(v_{k+1}) + ... + \mu_n T(v_n) = \mu_{k+1} w_{k+1} + ... + \mu_n w_n$, ce qui force $0 = \mu_{k+1} = ... = \mu_n$ car $\{w_{k+1},...,w_n\}$ est linéairement indépendant ; d'autre part $0 = \mu_1 v_1 + ... + \mu_k v_k$ implique que $0 = \mu_1 = ... = \mu_k$ car $\{v_1,...,v_k\}$ est linéairement indépendant. Donc $\{v_1,...,v_n\}$ est linéairement indépendant. Finalement soit $v \in V$. Par définition $T(v) \in \operatorname{Im} T$ et donc ils existent $\mu_{k+1},...,\mu_n \in \mathbb{k}$ tels que

$$T(v) = \sum_{i=k+1}^{n} \mu_i w_i = \sum_{i=k+1}^{n} \mu_i T(v_i) = T\left(\sum_{i=k+1}^{n} \mu_i v_i \right).$$

Mais dès lors $T\left(v - \sum_{i=k+1}^{n} \mu_i v_i\right) = 0$, c'est-à-dire, $v - \sum_{i=k+1}^{n} \mu_i v_i \in \ker T$. Ils existent donc $\mu_1,...,\mu_k \in \mathbb{k}$ tels que

$$v - \sum_{i=k+1}^{n} \mu_i v_i = \sum_{i=1}^{k} \mu_i v_i,$$

et donc $v = \sum_{i=1}^{n} \mu_i v_i \in < v_1,...,v_n >$, c'est-à-dire V est finiment engendré car $\{v_1,...,v_n\}$ est une base finie de V. La réciproque est exactement le Théorème 3.3.

□

3.3 Isomorphisme

Les sous-espaces $\mathrm{Im}\, T$ et $\ker T$ sont aussi intimement liés à la notion d'inverse. Soit $T : V \to W$ une application linéaire. Par définition, T est surjectif si et seulement si $\mathrm{Im}\, T = W$. Il est à noter que la linéarité de T ne joue aucun rôle dans ce dernier énoncé. Par contre,

Proposition 3.5 *Une application linéaire $T : V \to W$ est injective si et seulement si* $\ker T = \{0\}$.

Démonstration. Soit $v \in \ker T$. Si l'application est injective alors comme $T(v) = 0$ on doit nécessairement avoir $v = 0$, d'où $\ker T = \{0\}$. Réciproquement, si $T(v) = T(w)$ alors par la linéarité de T on a $T(v - w) = 0$, c'est-à-dire, $v - w \in \ker T$. Mais comme $\ker T = \{0\}$, on a $v = w$, c'est-à-dire, T est injective. □

Corollaire 3.6 *Une application linéaire $T : V \to W$ est inversible si et seulement si $\mathrm{Im}\, T = W$ et* $\ker T = \{0\}$. *En particulier, si V est finiment engendré alors* $\dim V = \dim W$.

Démonstration. Si V est finiment engendré alors par le Théorème 3.3 $\dim V = \dim \mathrm{Im}\, T$ car $\ker T = \{0\}$. Mais par hypothèse $\mathrm{Im}\, T = W$, d'où le résultat.

□

Proposition 3.7 *Si $T : V \to W$ est une application linéaire inversible alors T^{-1} est une application linéaire.*

Démonstration. Soit $u, v \in W$ et $\lambda \in \mathbb{k}$. Il existent $a, b \in V$ tels que $T(a) = u$ et $T(b) = v$. De plus, comme T est linéaire, on a $T(a + b) = u + v$ et $T(\lambda a) = \lambda u$. Nous avons donc

$$\begin{aligned} T^{-1}(u + v) &= T^{-1}(T(a + b)) \\ &= a + b \\ &= T^{-1}(T(a)) + T^{-1}(T(b)) \\ &= T^{-1}(u) + T^{-1}(v). \end{aligned}$$

De même, nous avons $T^{-1}(\lambda u) = T^{-1}(T(\lambda a)) = \lambda a = \lambda T^{-1}(T(a)) = \lambda T^{-1}(u)$.

□

Proposition 3.8 *Soit $T : V \to W$ et $S : W \to Z$ deux applications linéaires. Alors l'application $S \circ T$ est linéaire.*

Démonstration. Soit $u, v \in V$ et $\lambda \in \mathbb{k}$. Nous avons alors

$$\begin{aligned} (S \circ T)(u + v) &= S(T(u + v)) \\ &= S(T(u) + T(v)) \quad \text{par la linéarité de } T \\ &= S(T(u)) + S(T(v)) \quad \text{par la linéarité de } S \\ &= (S \circ T)(u) + (S \circ T)(v). \end{aligned}$$

De même, $(S \circ T)(\lambda u) = S(T(\lambda u)) = S(\lambda T(u)) = \lambda S(T(u)) = \lambda (S \circ T)(u)$.

□

Ce dernier résultat nous permet de définir une relation d'équivalence sur l'ensemble des $\mathbb{k}$-espaces vectoriels. On dira que deux $\mathbb{k}$-espaces vectoriels V et W sont *isomorphes* s'il existe une application linéaire inversible $T : V \to W$ que l'on appellera dès lors *isomorphisme* . Dans ce cas on écrira $V \cong W$. Cette relation est dite "relation d'équivalence" car

1. $V \cong V$ pour tout $\mathbb{k}$-espace vectoriel V (critère de réflexivité).
2. Si $V \cong W$ alors $W \cong V$ (critère de symétrie).
3. Si $V \cong W$ et $W \cong Z$ alors $V \cong Z$ (critère de transitivité).

Le (1) est vérifié car id_V est linéaire et inversible. Le (2) est une conséquence de la Proposition 3.7. Finalement, le (3) découle de la Proposition 3.8 et du fait suivant : Si T et S sont linéaires, inversibles et composables alors $T \circ S$ est inversible (exercice).

3.4 Calculs de la dimension

Dans le cas où V est finiment engendré, le Corollaire 3.6 exprime le fait que si $V \cong W$ alors $\dim V = \dim W$.

Corollaire 3.9 *Si U et W sont deux $\mathbb{k}$-espaces vectoriels finiment engendrés alors $U \times W$ est finiment engendré. De plus,*

$$\dim U \times W = \dim U + \dim W.$$

Démonstration. Considérons l'application linéaire (à vérifier) projection sur le premier facteur, c'est-à-dire, $p_U : U \times W \to U$, où $p_U(u, w) = u$. On a par construction $\mathrm{Im}\, p_U = U$, c'est-à-dire, $\dim \mathrm{Im}\, p_U = \dim U$. De plus, $\ker p_U = \{0\} \times W$. Mais comme $W \cong \{0\} \times W$ (pourquoi ?), on a $\dim \ker p_U = \dim W$. Par la Proposition 3.4, $U \times W$ est finiment engendré. On conclut à l'aide du Théorème 3.3 appliqué à l'application p_U que

$$\dim U \times W = \dim U + \dim W.$$

□

Corollaire 3.10 *Soit V un espace finiment engendré. Nous avons alors*

$$\dim_{\mathbb{k}} \mathbb{M}(V^n) = n \cdot \dim_{\mathbb{k}} V.$$

Démonstration. En remarquant que $\mathbb{M}(V^n) = \prod_{i=1}^{n} V$, on déduit le résultat à partir du Corollaire 3.9 et d'un raisonnement par récurrence sur le nombre de facteurs.

□

Exercice : Soit U et W deux sous-espaces d'un $\mathbb{k}$-espace vectoriel V finiment engendré. Démontrez la formule de Grassmann

$$\dim U + \dim W = \dim(U + W) + \dim(U \cap W).$$

Indice : considérer l'application linéaire

$$\begin{array}{rcl} \Omega : U \times W & \longrightarrow & U + W \\ (u, w) & \mapsto & u - w. \end{array}$$

(1) Montrez que Ω est linéaire. (2) Montrez que Ω est surjective. (3) Montrez que $\ker \Omega \cong U \cap W$. (4) Conclure à l'aide du Théorème 3.3.

Proposition 3.11 *Si V est un $\mathbb{k}$-espace vectoriel finiment engendré alors $\mathbb{k}^n \cong V$, où $n = \dim V$.*

Démonstration. Soit $\beta = \{v_1, ..., v_n\}$ une base de V. Définissons l'application linéaire (exercice)

$$\begin{array}{rcl} \kappa_\beta : \mathbb{k}^n & \longrightarrow & V \\ (a_1, ..., a_n) & \mapsto & a_1 v_1 + a_2 v_2 + ... + a_n v_n. \end{array} \tag{3.2}$$

D'une part, si $v \in V$ alors ils existent $\alpha_1, ..., \alpha_n \in \mathbb{k}$ tels que $v = \alpha_1 v_1 + ... + \alpha_n v_n$ car $\{v_1, ..., v_n\}$ engendre V. Nous avons donc

$$\kappa_\beta(\alpha_1, ..., \alpha_n) = \alpha_1 v_1 + ... + \alpha_n v_n = v,$$

c'est-à-dire, κ_β est surjective. D'autre part, si $(a_1, ..., a_n) \in \ker \kappa_\beta$ alors la relation

$$0 = \kappa_\beta(a_1, ..., a_n) = a_1 v_1 + a_2 v_2 + ... + a_n v_n$$

force $a_1 = a_2 = ... = a_n = 0$ car $\{v_1, ..., v_n\}$ est linéairement indépendant, c'est-à-dire, $\ker \kappa_\beta = \{0\}$ et donc κ_β est injective. On conclut que κ_β est un isomorphisme entre $\mathbb{k}^n$ et V.

□

Remarque : Les composantes du vecteur $\kappa_\beta^{-1}(v) \in \mathbb{k}^n$ sont les *coordonnées* de v dans la base β. Pour une base β fixée, étant donné que κ_β est un isomorphisme, les coordonnées d'un vecteur v sont uniques.

Théorème 3.12 *Deux $\mathbb{k}$-espaces vectoriels V et W finiment engendrés sont isomorphes si et seulement si* $\dim_{\mathbb{k}} V = \dim_{\mathbb{k}} W$.

Démonstration. Si $\dim_{\mathbb{k}} V = \dim_{\mathbb{k}} W$ alors en utilisant la symétrie et la transitivité de la relation d'équivalence "isomorphisme" ainsi que la Proposition 3.11 on obtient que $V \cong W$. La réciproque nous est donné par le Corollaire 3.6.

□

Exercice : Soit U et W deux sous-espaces d'un $\mathbb{k}$-espace vectoriel V finiment engendré. Supposons que $U \cap W = \{0\}$. Montrez que

$$U \times W \cong U \oplus W.$$

3.5 De retour aux polynômes

Il est maintenant clair que l'application (3.1) énoncée dans l'introduction au Chapitre 3 est un isomorphisme (exercice), c'est-à-dire, $\mathscr{P}_n(\mathbb{k}) \cong \mathbb{k}^{n+1}$.

Par contre, il nous est plus naturel de considérer un polynôme comme une fonction. Pour passer d'un polynôme abstrait à une fonction considérons l'application d'évaluation associée

$$\begin{array}{rcl} \mathscr{P}_n(\mathbb{k}) \times \mathbb{k} & \xrightarrow{\text{ev}} & \mathbb{k} \\ (a_n x^n + ... + a_1 x + a_0\,,\, \alpha) & \mapsto & a_n \alpha^n + ... + a_1 \alpha + a_0. \end{array}$$

Proposition 3.13 *Soit $p, q \in \mathscr{P}_n(\mathbb{k})$. Si $\mathbb{k} = \mathbb{R}$ ou $\mathbb{C}$ alors*

$$p = q \quad \Leftrightarrow \quad \text{ev}(p, \alpha) = \text{ev}(q, \alpha)$$

pour tout $\alpha \in \mathbb{k}$.

Démonstration. Si $p = q$ alors la conclusion est évidente. De plus cette implication est indépendante de la nature du corps $\mathbb{k}$. Pour montrer la réciproque nous procéderons à l'aide d'un raisonnement par récurrence forte sur le degré des polynômes. Soit $p = a_n x^n + ... + a_1 x + a_0$ et $q = b_n x^n + ... + b_1 x + b_0$ tels que $\text{ev}(p,\alpha) = \text{ev}(q,\alpha)$ pour tout $\alpha \in \mathbb{k}$. Si $n = 0$ alors $a_0 = \text{ev}(p,\alpha) = \text{ev}(q,\alpha) = b_0$ et l'ancrage de l'induction est vérifié. Supposons le résultat vérifié pour tous les polynômes de degré inférieur à n. Remarquons que $a_0 = \text{ev}(p,0) = \text{ev}(q,0) = b_0$ et donc pour tout $\alpha \in \mathbb{k}$

$$\begin{aligned} 0 = \text{ev}(p,\alpha) - \text{ev}(q,\alpha) &= (a_n - b_n)\alpha^n + ... + (a_1 - b_1)\alpha \\ &= \underbrace{\alpha\big((a_n - b_n)\alpha^{n-1} + ... + (a_1 - b_1)\big)}_{\text{distributivité dans } \mathbb{k}}. \end{aligned}$$

Comme il n'y a pas de diviseur de zéro dans un corps nous avons donc nécessairement $(a_n - b_n)\alpha^{n-1} + ... + (a_1 - b_1) = 0$ lorsque $\alpha \neq 0$. On se rappellera que les fonctions polynomials sont des fonctions continues sur $\mathbb{R}$ ou $\mathbb{C}$. Ceci force donc $\text{ev}\big((a_n - b_n)x^{n-1} + ... + (a_1 - b_1), \alpha\big) = 0$ pour tout $\alpha \in \mathbb{k}$. En d'autres termes

$$\text{ev}(a_n x^{n-1} + ... + a_1, \alpha) = \text{ev}(b_n x^{n-1} + ... + b_1, \alpha), \quad \text{pour tout } \alpha \in \mathbb{k},$$

et donc nous obtenons par l'hypothèse d'induction que $a_i = b_i$, $1 \leq i \leq n$. Comme nous avions déjà remarqué que $a_0 = b_0$ nous avons donc $p = q$.

□

Remarque : Nous avons bel et bien utilisé une récurrence forte car nous ne savons pas si les coefficients a_n et b_n sont non nuls ou non (par définition de $\mathscr{P}_n(\mathbb{k})$).

Ce dernier résultat peut s'étendre à tout corps $\mathbb{k}$ contenant $\mathbb{Q}$. De tels corps sont dits de *caractéristique* 0. Par contre l'énoncé est faux en toute généralité. En effet, si l'on considère le corps $\mathbb{Z}/2\mathbb{Z} = \{0,1\}$ à deux éléments (voir la sous-section 1.4.1), les deux polynômes distincts abstraits $x^2 + 1$ et $x + 1$ dans $\mathscr{P}_2(\mathbb{Z}/2\mathbb{Z})$ sont tels que

$$\text{ev}(x^2 + 1, \alpha) = \text{ev}(x + 1, \alpha), \quad \alpha = 0, 1.$$

Le corps $\mathbb{Z}/2\mathbb{Z}$ est dit de *caractéristique* 2. En fait, il y a des corps de caractéristique p pour tout nombre premier p. Ces corps vous serons introduits en MAT 2543.

Lorsque $\mathbb{k}$ est de caractéristique 0 on peut donc identifier $\mathscr{P}_n(\mathbb{k})$ avec un certain sous-espace des fonctions $\mathscr{F}(\mathbb{k},\mathbb{k})$. Il est donc naturel de définir

Définition 3.3 *Une fonction $f \in \mathscr{F}(\mathbb{k},\mathbb{k})$ est dite polynomial s'il existe un polynôme $p \in \mathscr{P}_n(\mathbb{k})$ pour un certain $n \in \mathbb{N}$ tel que*

$$f(\alpha) = \text{ev}(p,\alpha), \quad \textit{pour tout } \alpha \in \mathbb{k}.$$

A nouveau nous pouvons écrire l'application linéaire (exercice)

$$\begin{aligned} \mathscr{P}_n(\mathbb{k}) &\longrightarrow \mathscr{F}(\mathbb{k},\mathbb{k}) \\ p &\mapsto \text{ev}(p,\cdot). \end{aligned} \tag{3.3}$$

En général, l'application (3.3) n'est pas surjective. En effet, plusieurs fonctions sur $\mathbb{R}$ ne sont pas polynomiales. Par exemple, le sinus ou le cosinus qui sont bornés et non-constants ne sont pas polynomiaux. Par contre, la Proposition 3.13 démontre que lorsque $\mathbb{k}$ est de caractéristique 0 l'application (3.3) est injective. C'est donc avec le sous-espace image de l'application (3.3) que nous identifierons $\mathscr{P}_n(\mathbb{k})$.

3.6 L'espace de fonctions $\mathscr{L}(V,W)$

Soit V et W deux $\Bbbk$-espaces vectoriels et considérons le sous-ensemble de $\mathscr{F}(V,W)$

$$\mathscr{L}(V,W) = \{T : V \to W \mid T \text{ est linéaire}\}.$$

Proposition 3.14 *Le sous-ensemble $\mathscr{L}(V,W)$ est un sous-espace de $\mathscr{F}(V,W)$.*

Démonstration. Soit $S, T : V \to W$ deux applications linéaires et $\lambda \in \Bbbk$. Les applications

$$\begin{array}{rcl} S+T : V & \longrightarrow & W \\ v & \mapsto & S(v)+T(v) \end{array}$$

et $\lambda T : V \to W$ définit par $(\lambda T)(v) = \lambda(T(v))$ sont linéaires. En effet

$$\begin{array}{rcll} (S+T)(u+v) & = & S(u+v)+T(u+v), & \text{par définition} \\ & = & [S(u)+S(v)]+[T(u)+T(v)], & \text{par linéarité} \\ & = & [S(u)+T(u)]+[S(v)+T(v)], & \text{par la commutativité de l'addition dans } W \\ & = & (S+T)(u)+(S+T)(v). & \end{array}$$

De même, on a $(\lambda T)(u+v) = \lambda(T(u+v)) = \lambda(T(u)+T(v)) = \lambda(T(u)) + \lambda(T(v)) = (\lambda T)(u) + (\lambda T)(v)$.

□

Deux cas importants surgissent immédiatement :

Cas 1 : $V = W$.

Dans ce cas on notera $\mathscr{L}(V) = \mathscr{L}(V,V)$. Si $S, T \in \mathscr{L}(V)$ alors, par la Proposition 3.8, $S \circ T \in \mathscr{L}(V)$ et $T \circ S \in \mathscr{L}(V)$. De plus

Proposition 3.15 *Soit $R, S, T \in \mathscr{L}(V)$. Nous avons alors*

$$R \circ (S+T) = (R \circ S) + (R \circ T) \quad \textit{et} \quad (S+T) \circ R = (S \circ R) + (T \circ R),$$

c'est-à-dire, le triplet $(\mathscr{L}(V), +, \circ)$ est un anneau (voir la remarque à la fin de la section à propos des corps).

Démonstration. L'application identité $\mathrm{id}_V : V \to V$, $\mathrm{id}_V(v) = v$, agit comme élément neutre de la composition $\circ$. Finalement, si $v \in V$ alors

$$\begin{array}{rcl} (R \circ (S+T))(v) & = & R((S+T)(v)) \\ & = & R(S(v)+T(v)) \\ & = & R(S(v)) + R(T(v)) \\ & = & (R \circ S)(v) + (R \circ T)(v). \end{array}$$

Il en va de même pour la deuxième identité, d'où le résultat.

□

Cas 2 : $W = \Bbbk$.

Dans ce cas on notera $V^\vee = \mathscr{L}(V, \Bbbk)$. On appelle cet espace *l'espace dual de V*. Par exemple, les applications projections sur le $i^{\text{ème}}$ facteur $\pi_i : \Bbbk^n \to \Bbbk$, $\pi_i(a_1, ..., a_i, ..., a_n) = a_i$, sont des éléments de $(\Bbbk^n)^\vee$.

Proposition 3.16 *Soient V et W deux $\mathbb{k}$-espaces vectoriels. Supposons que V soit finiment engendré et que $\beta = \{v_1, ..., v_n\}$ soit une base de V. L'application*

$$\begin{array}{rcl} \psi_\beta^W : \mathscr{L}(V,W) & \longrightarrow & \mathbb{M}(W^n) \\ T & \mapsto & (T(v_1), T(v_2), ..., T(v_n)) \end{array}$$

est alors un isomorphisme. De plus, si W est finiment engendré alors

$$\dim_{\mathbb{k}} \mathscr{L}(V,W) = \dim_{\mathbb{k}} V \cdot \dim_{\mathbb{k}} W.$$

Démonstration. Nous laissons en exercice de démontrer la linéarité de ψ. Soient $w_1, w_2, ..., w_n \in W$. A l'aide des projections π_i, $1 \leq i \leq n$, et de l'isomorphisme des coordonnées (3.2) de la Proposition 3.11 on peut donc définir l'application linéaire

$$\begin{array}{rcl} T : V & \longrightarrow & W \\ v & \mapsto & \pi_1(\kappa_\beta^{-1}(v))\, w_1 + \pi_2(\kappa_\beta^{-1}(v))\, w_2 + ... + \pi_n(\kappa_\beta^{-1}(v))\, w_n. \end{array} \tag{3.4}$$

L'application T possède la propriété $T(v_i) = w_i$, $1 \leq i \leq n$. Ceci implique $\psi_\beta^W(T) = (w_1, ..., w_n)$, c'est-à-dire, ψ_β^W est surjective. Si $\psi_\beta^W(T) = (0, ..., 0)$ alors $T(v_i) = 0$, $1 \leq i \leq n$. Si $v \in V$ alors $v = \sum_{i=1}^n \alpha_i v_i$ pour certains $\alpha_i \in \mathbb{k}$. Mais dès lors, $T(v) = \sum_{i=1}^n \alpha_i T(v_i) = 0$, c'est-à-dire, $T \equiv 0$ et donc $\ker \psi_\beta^W = \{0\}$. On conclut que ψ_β^W est un isomorphisme. Si W est finiment engendré alors la formule

$$\dim_{\mathbb{k}} \mathscr{L}(V,W) = \dim_{\mathbb{k}} V \cdot \dim_{\mathbb{k}} W$$

est une conséquence directe du Corollaire 3.10.

□

Corollaire 3.17 *Si V est finiment engendré alors $V \cong V^\vee$.*

Démonstration. Comme $\dim_{\mathbb{k}} \mathbb{k} = 1$, par le Théorème 3.12 le résultat s'en suit car $\dim_{\mathbb{k}} V = \dim_{\mathbb{k}} V^\vee$.

□

Exercice : Soit $\beta = \{v_1, ..., v_n\}$ une base de V. Montrez que les éléments $\pi_1 \circ \kappa_\beta^{-1}, ..., \pi_n \circ \kappa_\beta^{-1}$ forment une base de $V^\vee$. On remarque que $(\pi_i \circ \kappa_\beta^{-1})(v_j) = \delta_{ij}$ où $\delta_{ii} = 1$ et $\delta_{ij} = 0$ si $i \neq j$. On appelle le symbole δ_{ij} le *delta de Kronecker*. On note $v_i^* = \pi_i \circ \kappa_\beta^{-1}$, $1 \leq i \leq n$, et on appelle $\{v_1^*, ..., v_n^*\}$ la *base duale* à la base $\{v_1, ..., v_n\}$ de V.

La démonstration de la Proposition 3.16 suggère le résultat suivant.

Proposition 3.18 *Soient V et W deux $\mathbb{k}$-espaces vectoriels. Supposons V finiment engendré, $\{v_1, ..., v_n\}$ une base de V et $\{w_1, ..., w_n\}$ un simple sous-ensemble de W. Il existe alors une unique application linéaire $T : V \to W$ telle que $T(v_i) = w_i$, $1 \leq i \leq n$.*

Démonstration. La recette (3.4) de la Proposition 3.16 démontre l'existence d'une telle application. Supposons qu'il existe une seconde application $T' : V \to W$ telle que $T'(v_i) = w_i$, $1 \leq i \leq n$. Ceci implique $0 = T(v_i) - T'(v_i) = (T - T')(v_i)$, $1 \leq i \leq n$, c'est-à-dire, $v_i \in \ker(T - T')$, $1 \leq i \leq n$. Mais $\{v_1, ..., v_n\}$ engendre V et donc $V = \ker(T - T')$, c'est-à-dire, $T - T' \equiv 0$. On conclut que $T = T'$.

□

3.7 Représentation matricielle

Reprenons l'application ψ_β^W de la Proposition 3.16, c'est-à-dire,

$$\begin{array}{rcl} \psi_\beta^W : \mathscr{L}(V,W) & \longrightarrow & \mathbb{M}(W^n) \\ T & \mapsto & (T(v_1), T(v_2), ..., T(v_n)), \end{array}$$

où $\beta = \{v_1, v_2, ..., v_n\}$ est une base de V. Supposons que W soit finiment engendré et $\alpha = \{w_1, w_2, ..., w_m\}$ une base de celui-ci. On obtient un nouvel isomorphisme (pourquoi ?)

$$\begin{array}{rcl} \widehat{\psi}_\beta^\alpha : \mathscr{L}(V,W) & \longrightarrow & \mathbb{M}_{n\times m}(\Bbbk) \\ T & \mapsto & (\kappa_\alpha^{-1}(T(v_1)), \kappa_\alpha^{-1}(T(v_2)), ..., \kappa_\alpha^{-1}(T(v_n))), \end{array}$$

où κ_α est l'isomorphisme des coordonnées (3.2) de la Proposition 3.11 appliqué à W. On se rappellera que l'on a un isomorphisme naturel appelé *l'application transposée*

$$\begin{array}{rcl} {}^t(\cdot) : \mathbb{M}_{n\times m}(\Bbbk) & \longrightarrow & \mathbb{M}_{m\times n}(\Bbbk) \\ A & \mapsto & {}^tA, \end{array}$$

où les colonnes de ${}^tA$ sont les lignes de A. En composant ces deux isomorphismes on obtient un troisième isomorphisme

$$\begin{array}{rcl} \widetilde{\psi}_\beta^\alpha : \mathscr{L}(V,W) & \longrightarrow & \mathbb{M}_{m\times n}(\Bbbk) \\ T & \mapsto & {}^t[\widehat{\psi}_\beta^\alpha(T)] \end{array}$$

qui possède la propriété suivante :

Théorème 3.19 *Soient $T : V \to W$ et $S : W \to Z$ deux transformations linéaires entre $\Bbbk$-espaces vectoriels finiment engendrés. Alors*

$$\widetilde{\psi}_\beta^\gamma(S \circ T) = \widetilde{\psi}_\alpha^\gamma(S) \cdot \widetilde{\psi}_\beta^\alpha(T),$$

où le produit de droite est le produit matricielle classique et β, α et γ sont des bases ordonnées de V, W et Z respectivement.

Remarque : Il est important de se rappeler que ces isomorphismes dépendent des choix des bases pour V, W et Z respectivement. En particulier, la base choisie pour W dans le cas de T et celle choisie dans le cas de S doit être la-même ! De plus, il est implicite dans la notation utilisée qu'on parle bien ici de bases "ordonnées".

Notation : Soient $T : V \to W$ une transformation linéaire entre $\Bbbk$-espaces vectoriels finiment engendrés. Soit β et α des bases ordonnées de V et W respectivement. On notera

$$[T]_\beta^\alpha = \widetilde{\psi}_\beta^\alpha(T).$$

En particulier, le Théorème 3.19 s'écrit maintenant

$$[S \circ T]_\beta^\gamma = [S]_\alpha^\gamma \cdot [T]_\beta^\alpha,$$

où γ est une base de Z.

Démonstration. Soit $\beta = \{v_1, ..., v_n\}$, $\alpha = \{w_1, ..., w_m\}$ et $\gamma = \{z_1, ..., z_p\}$ des bases de V, W et Z respectivement. D'une part, ils existent $a_{kj} \in \mathbb{k}$ pour $1 \leq k \leq m$ et $1 \leq j \leq n$, et $b_{ik} \in \mathbb{k}$ pour $1 \leq i \leq p$ et $1 \leq k \leq m$ tels que

$$T(v_j) = \sum_{k=1}^{m} a_{kj} w_k, \quad 1 \leq j \leq n$$

et

$$S(w_k) = \sum_{i=1}^{p} b_{ik} z_i, \quad 1 \leq k \leq m.$$

Nous avons donc $[T]_\beta^\alpha = (a_{kj})$ et $[S]_\alpha^\gamma = (b_{ik})$. On se rappelle que le produit matricielle nous donne

$$[S]_\alpha^\gamma \cdot [T]_\beta^\alpha = \left(\sum_{k=1}^{m} b_{ik} a_{kj} \right) \quad 1 \leq i \leq p, 1 \leq j \leq n.$$

D'autre part, nous avons

$$\begin{aligned} S(T(v_j)) &= S\left(\sum_{k=1}^{m} a_{kj} w_k \right), \quad 1 \leq j \leq n \\ &= \sum_{k=1}^{m} a_{kj} S(w_k) \\ &= \sum_{k=1}^{m} a_{kj} \left(\sum_{i=1}^{p} b_{ik} z_i \right), \quad 1 \leq j \leq m \\ &= \sum_{k=1}^{m} \sum_{i=1}^{p} (b_{ik} a_{kj}) z_i \\ &= \sum_{i=1}^{p} \left(\sum_{k=1}^{m} b_{ik} a_{kj} \right) z_i, \end{aligned}$$

d'où

$$[S \circ T]_\beta^\gamma = \left(\sum_{k=1}^{m} b_{ik} a_{kj} \right), \quad 1 \leq i \leq p, 1 \leq j \leq n.$$

□

Soit $v \in V$. La notation $[v]^\beta$ désignera le vecteur colonne composé des coordonnées du vecteur v dans la base ordonnée β. On a donc

Proposition 3.20 *Soient $T : V \to W$ une transformation linéaire entre $\mathbb{k}$-espaces vectoriels finiment engendrés. Soit β et α des bases ordonnées de V et W respectivement. Alors pour tout $v \in V$ nous avons*

$$[T(v)]^\alpha = [T]_\beta^\alpha \cdot [v]^\beta.$$

Démonstration. La démonstration est essentiellement la-même que celle du Théorème 3.19. Soit $\beta = \{v_1, ..., v_n\}$ et $\alpha = \{w_1, ..., w_m\}$ des bases ordonnées de V et W respectivement. D'une part, ils existent $b_j \in \mathbb{k}$ et $a_{ij} \in \mathbb{k}$, pour $1 \leq i \leq m$ et $1 \leq j \leq n$, et tels que

$$v = \sum_{k=1}^{n} b_k v_k$$

et

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i, \quad 1 \leq j \leq n.$$

Nous avons donc $[v]^\beta = {}^t(b_1, b_2, ..., b_n)$ et $[T]^\alpha_\beta = (a_{ij})$. On se rappelle que le produit matricielle nous donne

$$[T]^\alpha_\beta \cdot [v]^\beta = \left(\sum_{k=1}^{n} a_{ik} b_k \right), \quad 1 \leq i \leq m.$$

D'autre part, nous avons

$$\begin{aligned} T(v) &= T\left(\sum_{k=1}^{n} b_k v_k \right) = \sum_{k=1}^{n} b_k T(v_k) \\ &= \sum_{k=1}^{n} b_k \left(\sum_{i=1}^{m} a_{ik} w_i \right) \\ &= \sum_{i=1}^{m} \left(\sum_{k=1}^{n} a_{ik} b_k \right) w_i, \end{aligned}$$

d'où $[T(v)]^\alpha = \left(\sum_{k=1}^{n} a_{ik} b_k \right)$, pour $1 \leq i \leq m$.

□

Exemple : Soit $Id_V : V \to V$ la transformation linéaire identité, c'est-à-dire, $Id_V(v) = v$, pour tout $v \in V$. Si $\beta = \{v_1, ..., v_n\}$ est une base ordonnée de V alors

$$[Id_V]^\beta_\beta = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix},$$

c'est-à-dire, la matrice identité $\mathbb{1} \in \mathbb{M}_{n\times n}(\mathbb{k})$. Si α est une seconde base ordonnée de V alors que représentent les matrices $[Id_V]^\alpha_\beta$ et $[Id_V]^\beta_\alpha$? Et bien par la Proposition 3.20, nous avons pour $v \in V$

$$[v]^\beta = [Id_V(v)]^\beta = [Id_V]^\beta_\alpha \cdot [v]^\alpha,$$

c'est-à-dire, les matrices $[Id_V]^\alpha_\beta$ et $[Id_V]^\beta_\alpha$ correspondent exactement aux matrices de changement de base.

Proposition 3.21 *Les matrices de changement de base sont inversibles.*

Démonstration. On se rappelle qu'une matrice $A \in \mathbb{M}_{n\times n}(\mathbb{k})$ est inversible s'il existe une matrice $B \in \mathbb{M}_{n\times n}(\mathbb{k})$ telle que

$$AB = BA = \mathbb{1}.$$

Soit $[Id_V]^\alpha_\beta$ une matrice de changement de base. Alors par le Théorème 3.19 nous avons

$$[Id_V]^\alpha_\beta \cdot [Id_V]^\beta_\alpha = [Id_V \circ Id_V]^\alpha_\alpha = [Id_V]^\alpha_\alpha = \mathbb{1} \quad \text{et} \quad [Id_V]^\beta_\alpha \cdot [Id_V]^\alpha_\beta = [Id_V]^\beta_\beta = \mathbb{1}.$$

On conclut que $[Id_V]_{\beta}^{\alpha}$ est inversible.

□

Proposition 3.22 *Soit $T : V \to W$ une application linéaire entre $\mathbb{k}$-espaces vectoriels finiment engendrés. Soient β_1 et β_2 deux bases ordonnées de V et α_1 et α_2 deux bases ordonnées de W. Alors il existent deux matrices inversibles $M \in \mathbb{M}_{m\times m}(\mathbb{k})$ et $N \in \mathbb{M}_{n\times n}(\mathbb{k})$ telles que*

$$[T]_{\beta_1}^{\alpha_1} = M \cdot [T]_{\beta_2}^{\alpha_2} \cdot N,$$

où $m = \dim W$ et $n = \dim V$.

Démonstration. On pose $M = [Id_W]_{\alpha_2}^{\alpha_1}$ et $N = [Id_V]_{\beta_1}^{\beta_2}$ et le résultat s'en suit par le Théorème 3.19 et par l'associativité de la composition de fonction ainsi que celle du produit matricielle.

□

CHAPITRE 4

PRODUIT SCALAIRE ET ESPACES EUCLIDIENS

4.1 Formes bilinéaires et dualité

Soit V et W deux $\mathbb{k}$-espaces vectoriels.

Définition 4.1 *Une forme bilinéaire est une application*

$$\beta : V \times W \to \mathbb{k}$$

telle que

- $\beta(av + bu, w) = a\beta(v, w) + b\beta(u, w)$ *pour tout* $v, u \in V$, $w \in W$ *et* $a, b \in \mathbb{k}$;
- $\beta(v, aw + b\omega) = a\beta(v, w) + b\beta(v, \omega)$ *pour tout* $v \in V$, $w, \omega \in W$ *et* $a, b \in \mathbb{k}$.

En d'autres mots, une forme bilinéaire est une application linéaire en ses deux variables. Si l'on fixe un vecteur $v_o \in V$ alors on peut définir une application linéaire

$$\begin{array}{rccc} D_{v_o} : W & \longrightarrow & \mathbb{k} \\ w & \mapsto & \beta(v_o, w). \end{array}$$

On peut faire de même à gauche en fixant un vecteur $w_o \in W$. On obtient une application linéaire $G_{w_o} : V \to \mathbb{k}$ définie par $G_{w_o}(v) = \beta(v, w_o)$. On s'aperçoit immédiatement que

$$D_{v_o} \in W^\vee \quad \text{et} \quad G_{w_o} \in V^\vee.$$

Comme D_{v_o} et G_{w_o} sont tous deux linéaires nous avons donc $\beta(v, 0) = \beta(0, w) = 0$ pour tout $v \in V$ et pour tout $w \in W$.

Définition 4.2 *Une forme bilinéaire est dite non dégénérée*

1. *si* $\beta(v,w)=0$ *pour tout* $v\in V$ *alors* $w=0$ *; et*
2. *si* $\beta(v,w)=0$ *pour tout* $w\in W$ *alors* $v=0$.

Exemple 1 : Le produit scalaire sur $\mathbb{R}^n$.

On se rappellera que le produit scalaire est défini par

$$\begin{array}{rcl}\langle\cdot,\cdot\rangle:\mathbb{R}^n\times\mathbb{R}^n & \longrightarrow & \mathbb{R}\\ (u,v) & \mapsto & \langle u,v\rangle=\sum_{i=1}^{n}a_ib_i,\end{array}$$

où $u=(a_1,a_2,...,a_n)$ et $v=(b_1,b_2,...,b_n)$. On laisse au lecteur le soin de démontrer que le produit scalaire est bien bilinéaire. Par contre, supposons que $\langle u,v\rangle=0$ pour tout $u\in\mathbb{R}^n$. En particulier, nous devons avoir $\sum_{i=1}^{n}b_i^2=\langle v,v\rangle=0$. Comme nous travaillons sur $\mathbb{R}$, chaque $b_i^2\geq 0$. On déduit donc que chaque $b_i^2=0$, c'est-à-dire, $b_i=0$ pour tout $i=1,...,n$ et donc $v=0$. Le même raisonnement peut être appliqué à gauche. On conclut que le produit scalaire est une forme bilinéaire non dégénérée.

Exemple 2 : Le produit de dualité.

Soit V un $\mathbb{k}$-espace vectoriel finiment engendré. Considérons l'application

$$\begin{array}{rcl}\mathscr{B}_V:V\times V^\vee & \longrightarrow & \mathbb{k}\\ (v,f) & \mapsto & \mathscr{B}_V(v,f)=f(v).\end{array}$$

A nouveau on laisse le lecteur se convaincre que $\mathscr{B}_V$ est bien bilinéaire. Par contre, si $\mathscr{B}_V(v,f)=f(v)=0$ pour tout $v\in V$ alors par la définition d'une fonction, f est donc la fonction nulle sur V (à ce point-ci V n'a pas à être finiment engendré). De plus, si $\mathscr{B}_V(v,f)=f(v)=0$ pour tout $f\in V^\vee$ alors nécessairement $v=0$. Sinon, comme V est finiment engendré, on peut donc compléter $\{v\}$ en une base $\{v,u_2,...,u_n\}$ de V. Dans ce cas l'application $\pi_1\circ\kappa^{-1}$ (voir la recette (3.4) de la Proposition 3.16) est un élément de $V^\vee$ tel que $\mathscr{B}_V(v,\pi_1\circ\kappa^{-1})=\pi_1(\kappa^{-1}(v))=1$ ce qui contredit l'hypothèse initiale. On conclut que le produit de dualité $\mathscr{B}_V$ est une forme bilinéaire non dégénérée.

4.2 Représentation de Riesz en dimension finie

Théorème 4.1 *(*Représentation de Riesz*) Soit* β *une forme bilinéaire non dégénérée entre deux* $\mathbb{k}$*-espaces vectoriels finiment engendrés* V *et* W*. L'application*

$$\begin{array}{rcl}\Omega_G:W & \longrightarrow & V^\vee\\ w & \mapsto & G_w\end{array}\tag{4.1}$$

est un isomorphisme, où $G_w(v)=\beta(v,w)$*. Il en va de même dans le cas* $V\cong W^\vee$ *mais cette fois en utilisant* Ω_D *et* D_v*.*

Démonstration. Le lecteur peut facilement se convaincre que Ω_G est linéaire. Si $\Omega_G(w_1)=\Omega_G(w_2)$ alors $G_{w_1}=G_{w_2}$ et de façon équivalente $G_{w_1}(v)=G_{w_2}(v)$ pour tout $v\in V$. Mais dès lors

$$\beta(v,w_1)=\beta(v,w_2)\quad\text{pour tout } v\in V.$$

La linéarité de β en sa deuxième variable implique donc que $\beta(v,w_1-w_2)=0$ pour tout $v\in V$. On conclut à l'aide de la non dégénérescence de β que $w_1=w_2$, c'est-à-dire que Ω_G est injective. Comme

$W \cong \operatorname{Im}\Omega_G \subset V^\vee$, on a donc $\dim W \leq \dim V^\vee$. On peut refaire tout ce raisonnement à droite et obtenir $\dim V \leq \dim W^\vee$. Par le Corollaire 3.17, on a

$$\dim V^\vee = \dim V \leq \dim W^\vee = \dim W \leq \dim V^\vee,$$

et donc $\dim W = \dim V^\vee$. Finalement, comme $W \cong \operatorname{Im}\Omega_G \subset V^\vee$, on doit avoir $\operatorname{Im}\Omega_G = V^\vee$, c'est-à-dire Ω_G est un isomorphisme.

□

Exemple : Le Théorème 4.1 appliqué au produit de dualité $\mathscr{B}_V$ démontre en général que l'application linéaire

$$\begin{aligned} \Omega_D : V &\longrightarrow V^{\vee\vee} \\ v &\mapsto D_v, \end{aligned} \tag{4.2}$$

où $D_v(g) = g(v)$, $g \in V^\vee$, est toujours injective car $\mathscr{B}_V$ est toujours non dégénérée à gauche. Si l'on ajoute la condition que V soit finiment engendré alors Ω_D est un isomorphisme entre V et son double-dual $V^{\vee\vee}$ qui est canonique, c'est-à-dire, que nous n'avons pas eu à fixer de base pour V afin de construire Ω_D. Ce n'est pas le cas pour l'isomorphisme $V \cong V^\vee$ donné par le Corollaire 3.17 et le Théorème 3.12.

Une conséquence directe du Théorème 4.1 ainsi que du Corollaire 3.17 est que

Corollaire 4.2 *Si B est une forme bilinéaire non dégénérée entre deux $\mathbb{k}$-espaces vectoriels finiment engendrés V et W alors*

$$\dim V = \dim W.$$

On peut donc définir de façon naturelle

Définition 4.3 *Deux $\mathbb{k}$-espaces vectoriels finiment engendrés V et W sont dits duaux par rapport à une forme bilinéaire $\beta : V \times W \to \mathbb{k}$ si β est non dégénérée.*

4.3 L'adjointe et sa représentation matricielle

Théorème 4.3 *Soient (V_1, W_1) une paire duale par rapport à une forme bilinéaire β_1. Soit $\beta_2 : V_2 \times W_2 \to \mathbb{k}$ une seconde forme bilinéaire (pas nécessairement non dégénérée) et soit $T \in \mathscr{L}(V_1, V_2)$. Alors il existe une unique transformation linéaire $T^* \in \mathscr{L}(W_2, W_1)$ telle que*

$$\beta_1(v, T^*(w)) = \beta_2(T(v), w)$$

pour tout $v \in V_1$ et $w \in W_2$.

Démonstration. Soit $w \in W_2$. Comme T est linéaire, l'application $v \mapsto \beta_2(T(v), w)$ est un élément de $V_1^\vee$. En utilisant l'isomorphisme (4.1) du Théorème 4.1, il existe un unique $w_1 \in W_1$ tel que

$$\beta_1(v, w_1) = G_{w_1}(v) = \beta_2(T(v), w).$$

On pose alors $T^*(w) = w_1$. Soit $v \in V_1$, $w, w' \in W_2$ et $a, b \in \mathbb{k}$. Nous avons alors

$$\begin{aligned} \beta_1(v, T^*(aw + bw')) &= \beta_2(T(v), aw + bw'), \quad \text{par définition de } T^* \\ &= a\beta_2(T(v), w) + b\beta_2(T(v), w'), \\ &= a\beta_1(v, T^*(w)) + b\beta_1(v, T^*(w')), \quad \text{par définition de } T^* \\ &= \beta_1(v, aT^*(w) + bT^*(w')). \end{aligned}$$

Etant donné que β_1 est non dégénérée, on conclut que $T^*(aw+bw') = aT^*(w)+bT^*(w')$, c'est-à-dire, $T^* \in \mathscr{L}(W_2, W_1)$. S'il existe une seconde application linéaire $T'' \in \mathscr{L}(W_2, W_1)$ telle que $\beta_1(v, T''(w)) = \beta_2(T(v), w)$ alors la non dégénérescence de β_1 implique que $T^* = T''$ car nous avons

$$\beta_1(v, T''(w)) = \beta_1(v, T^*(w))$$

pour tout $v \in V_1$.

□

Définition 4.4 *On appelle T^* l'adjoint de l'application T (par rapport aux formes bilinéaires β_1 et β_2).*

On suppose que V_1, V_2, W_1 et W_2 finiment engendrés. Alors

Proposition 4.4 *Soient (V_1, W_1) et (V_2, W_2) deux paires duales par rapport aux formes bilinéaires β_1 et β_2 respectivement. Soit $T \in \mathscr{L}(V_1, V_2)$. Alors il existent α, β, γ et η des bases ordonnées de V_1, V_2, W_1 et W_2 respectivement telles que*

$$[T^*]_\eta^\gamma = {}^t([T]_\alpha^\beta).$$

Démonstration. Soit $\alpha = \{v_1, ..., v_n\}$ une base de V_1 et considérons la base duale $\{v_1^*, ..., v_n^*\}$ de $V_1^\vee$. Par l'isomorphisme (4.1) du Theorem 4.1, pour chaque v_i^* il existe un unique $w_i \in W_1$ tel que

$$G_{w_i} = v_i^*.$$

En particulier, $\gamma = \{w_1, .., , w_n\}$ est une base de W_1 (pourquoi ?). De plus, les bases α et γ sont duales par rapport à β_1, c'est-à-dire,

$$\beta_1(v_i, w_j) = \delta_{ij}.$$

De même, on peut construire une paire de bases duales $\beta = \{u_1, ..., u_m\}$ et $\eta = \{x_1, ..., x_m\}$ de V_2 et W_2 respectivement par rapport à β_2, c'est-à-dire,

$$\beta_2(u_i, x_j) = \delta_{ij}.$$

Soit $[T]_\alpha^\beta = (a_{ij}) \in \mathbb{M}_{m\times n}(\mathbb{k})$ et $S \in \mathscr{L}(W_2, W_1)$ l'application linéaire correspondant à la matrice (b_{ji}) où $b_{ji} = a_{ij}$ par rapport aux bases η et γ de W_2 et W_1 respectivement. D'une part, on a

$$\beta_2(T(v_j), x_i) = \sum_{k=1}^{m} a_{kj}\beta_2(u_k, x_i) = a_{ij}.$$

D'autre part, on a

$$\beta_1(v_j, S(x_i)) = \sum_{k=1}^{n} b_{ki}\beta_1(v_j, w_k) = b_{ji} = a_{ij}.$$

Par la construction de T^*, on a

$$\beta_1(v_j, T^*(x_i)) = \beta_2(T(v_j), x_i) = \beta_1(v_j, S(x_i)).$$

Comme β_1 est non-dégénérée, on a $T^*(x_i) = S(x_i)$ pour tous les éléments de la base η. On conclut que $T^* = S$.

□

4.4 Le cas du produit de dualité

Soient V et W deux $\mathbb{k}$-espaces vectoriels finiment engendrés et $T \in \mathscr{L}(V,W)$. Par le Théorème 4.3, T possède un adjoint T^* par rapport aux deux produits de dualité $\mathscr{B}_V$ et $\mathscr{B}_W$, c'est-à-dire, il existe une unique application $T^* \in \mathscr{L}(W^\vee, V^\vee)$ telle que

$$(T^*(g))(v) = \mathscr{B}_V(v, T^*(g)) = \mathscr{B}_W(T(v), g) = g(T(v)),$$

pour tout $v \in V$ et $g \in W^\vee$. On conclut que

$$\begin{array}{rccc} T^* : W^\vee & \longrightarrow & V^\vee \\ g & \mapsto & g \circ T. \end{array} \tag{4.3}$$

Remarque : Dans ce contexte, l'application (4.3) est souvent appelée la **transposée** de T et est **notée** $T^\vee$.

Théorème 4.5 *L'isomorphisme Ω_D (4.2) est naturel, c'est-à-dire, pour tout V et W des $\mathbb{k}$-espaces vectoriels finiment engendrés et $T \in \mathscr{L}(V,W)$, le diagramme suivant*

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ {\scriptstyle \Omega_D^V}\downarrow{\scriptstyle \cong} & & {\scriptstyle \cong}\downarrow{\scriptstyle \Omega_D^W} \\ V^{\vee\vee} & \xrightarrow[T^{\vee\vee}]{} & W^{\vee\vee} \end{array}$$

commute. En d'autres mots, pour tout $v \in V$ on a $T^{\vee\vee}(\Omega_D^V(v)) = \Omega_D^W(T(v))$.

Démonstration. Tout est en place pour une simple vérification de cette dernière équation. En effet, soit $\phi \in W^\vee$. Alors

$$\begin{array}{rcl} T^{\vee\vee}(\Omega_D^V(v))(\phi) & = & (\Omega_D^V(v) \circ T^\vee)(\phi) = \Omega_D^V(v)(T^\vee(\phi)), \quad \text{par définition de la transposée (4.3)} \\ & = & \mathscr{B}_V(v, T^\vee(\phi)), \quad \text{par définition de } \Omega_D^V \\ & = & \mathscr{B}_W(T(v), \phi), \quad \text{par la relation de dualité (Théorème 4.3)} \\ & = & \Omega_D^W(T(v))(\phi), \quad \text{par définition de } \Omega_D^W. \end{array}$$

Étant donné que ϕ est arbitraire, le résultat suit.

□

Corollaire 4.6 *L'application*

$$\begin{array}{rccc} \mathscr{L}(V,W) & \longrightarrow & \mathscr{L}(V^{\vee\vee}, W^{\vee\vee}) \\ T & \mapsto & T^{\vee\vee} \end{array}$$

est linéaire.

Démonstration. Soient $\phi \in V^{\vee\vee}$, $S, T \in \mathscr{L}(V,W)$ et $a,b \in \mathbb{k}$. Comme Ω_D^V est un isomorphisme, il existe un unique $v \in V$ tel que $\Omega_D^V(v) = \phi$. Alors par le Théorème 4.5 nous avons

$$\begin{aligned}
(aS+bT)^{\vee\vee}(\phi) &= (aS+bT)^{\vee\vee}(\Omega_D^V(v)),\\
&= \Omega_D^W((aS+bT)(v)),\\
&= a\Omega_D^W(S(v)) + b\Omega_D^W(T(v)), \quad \text{par la linéarité de } \Omega_D^W\\
&= aS^{\vee\vee}(\Omega_D^V(v)) + bT^{\vee\vee}(\Omega_D^V(v)),\\
&= aS^{\vee\vee}(\phi) + bT^{\vee\vee}(\phi).
\end{aligned}$$

On conclut que $(aS+bT)^{\vee\vee} = aS^{\vee\vee} + bT^{\vee\vee}$.

□

4.5 Le sous-espace orthogonal

Définition 4.5 *Soit (V,W) une paire duale par rapport à une forme bilinéaire β et V_1 et W_1 deux sous-espaces de V et W respectivement.* ***L'orthogonal*** *$V_1^\perp$ de V_1 (par rapport à β) est le sous-espace de W tel que*

$$V_1^\perp = \{w \in W \mid \beta(v,w) = 0 \text{ pour tout } v \in V_1\}.$$

De même,

$$W_1^\perp = \{v \in V \mid \beta(v,w) = 0 \text{ pour tout } w \in W_1\}.$$

Remarque : On demande que la forme bilinéaire soit non dégénérée afin de garantir que $V^\perp = \{0\} = W^\perp$.

Exercice : Montrez que $V_1^\perp$ et $W_1^\perp$ sont effectivement des sous-espaces de W et V respectivement.

Proposition 4.7 *Soit (V,W) une paire duale par rapport à une forme bilinéaire β et soit U un sous-espace de V. Alors*

1. $\dim U + \dim U^\perp = \dim V$;
2. $(U^\perp)^\perp = U$.

Démonstration. Soit $\{x_1,...,x_k\}$ une base de U que l'on complète en une base $\{x_1,...,x_k,x_{k+1},...,x_n\}$ de V. En utilisant l'isomorphisme (4.1) du Théorème 4.1, on peut choisir $\{w_1,...,w_n\} \subset W$ tel que $G_{w_i} = x_i^*$, où $\{x_1^*,...,x_n^*\}$ est la base duale de $V^\vee$ correspondant à la base de V fixée. Comme l'application linéaire

$$< w_1,...,w_n > \hookrightarrow W \xrightarrow{\Omega_G} V^\vee$$

est surjective alors $\dim < w_1,...,w_n >= n$, c'est-à-dire, $\{w_1,...,w_n\}$ est une base de W (lorsque β est non dégénérée $\dim V = \dim W$). Si $u \in U$ alors ils existent $\alpha_1,...,\alpha_k \in \mathbb{k}$ tels que $u = \alpha_1 x_1 + ... + \alpha_k x_k$. De plus, pour $j \in \{k+1,...,n\}$, nous avons

$$\beta(u,w_j) = G_{w_j}(u) = x_j^*(u) = \sum_{i=1}^{k} \alpha_i x_j^*(x_i) = 0.$$

On conclut que $< w_{k+1}, ..., w_n > \subset U^\perp$. Si $\omega \in U^\perp$ alors ils existent $\gamma_1, ..., \gamma_n \in \mathbb{k}$ tels que $\omega = \gamma_1 w_1 + ... + \gamma_n w_n$. Mais lorsque $i \in \{1, ..., k\}$ nous avons

$$0 = \beta(x_i, \omega) = \sum_{j=1}^{n} \gamma_j \beta(x_i, w_j) = \sum_{j=1}^{n} \gamma_j x_j^*(x_i) = \gamma_i,$$

c'est-à-dire, $\omega \in < w_{k+1}, ..., w_n >$, d'où le premier énoncé. Pour le second énoncé, par définition $U \subset (U^\perp)^\perp$. Il est à noter que le premier énoncé est vérifié dans le cas symétrique où S serait un sous-espace de W. Donc comme

$$\dim U^\perp + \dim(U^\perp)^\perp = \dim W = \dim V = \dim U + \dim U^\perp,$$

on conclut que $\dim U = \dim(U^\perp)^\perp$, c'est-à-dire, $U = (U^\perp)^\perp$.

□

Remarque : Il est à noter que $U^\perp = \{0\}$ si et seulement si $U = V$. En effet, comme $\{0\}^\perp = V$, on applique le (2) de la Proposition 4.7. La réciproque est une conséquence de la non dégénérescence de β.

4.6 Le Théorème du rang

Proposition 4.8 *Soient (V_1, W_1) et (V_2, W_2) deux paires duales par rapport aux formes bilinéaires β_1 et β_2 respectivement. Soit $T \in \mathscr{L}(V_1, V_2)$. Alors*

$$\ker T^* = (\operatorname{Im} T)^\perp \quad \textit{et} \quad \ker T = (\operatorname{Im} T^*)^\perp.$$

Démonstration. Si $w \in W_2$ alors

$$\begin{aligned} w \in \ker T^* &\Leftrightarrow T^*(w) = 0, \\ &\Leftrightarrow \beta_1(v, T^*(w)) = 0, \quad \text{pour tout } v \in V_1, \text{ (non dégénérescence de } \beta_1) \\ &\Leftrightarrow \beta_2(T(v), w) = 0, \quad \text{pour tout } v \in V_1, \text{ (par définition de l'adjoint)} \\ &\Leftrightarrow w \in (\operatorname{Im} T)^\perp, \quad \text{par définition de } \operatorname{Im} T. \end{aligned}$$

La seconde égalité se démontre de la même manière.

□

Corollaire 4.9 *Sous les mêmes conditions que le résultat précédant on a*

1. *T est surjectif $\Leftrightarrow$ T^* est injectif.*
2. *T est injectif $\Leftrightarrow$ T^* est surjectif.*

Démonstration. Le résultat découle de la remarque suivant la Proposition 4.7.

□

Théorème 4.10 *Soient (V_1, W_1) et (V_2, W_2) deux paires duales par rapport aux formes bilinéaires β_1 et β_2 respectivement. Soit $T \in \mathscr{L}(V_1, V_2)$. Alors*

$$\operatorname{rk}(T) = \operatorname{rk}(T^*).$$

Démonstration. D'une part nous savons que $\dim W_2 = \dim\ker T^* + \mathrm{rk}\, T^*$ (Thérorème 3.3). D'autre part, par la Proposition 4.7, $\dim V_2 = \mathrm{rk}\, T + \dim(\mathrm{Im}\, T)^\perp$. Comme $\ker T^* = (\mathrm{Im}\, T)^\perp$ (Proposition 4.8), nous avons

$$\begin{aligned} \dim\ker T^* &= \dim(\mathrm{Im}\, T)^\perp \\ \dim W_2 - \mathrm{rk}\, T^* &= \dim V_2 - \mathrm{rk}\, T. \end{aligned}$$

Mais la paire (V_2, W_2) est duale et donc $\dim W_2 = \dim V_2$. Le résultat suit immédiatement.

□

Remarque : Il faut faire bien attention : en général nous n'avons **PAS** $\dim\ker T = \dim\ker T^*$.

Exercice : Montrez que $\dim\ker T = \dim\ker T^*$ si et seulement si $\dim V_1 = \dim V_2$.

4.7 Espaces euclidiens

Dans cette section nous fixons les réels $\mathbb{R}$ comme corps de référence.

Définition 4.6 *Un* espace préhilbertien *est un espace vectoriel V munit d'une forme bilinéaire $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ telle que*

1. $\langle v, v\rangle \geq 0$ *pour tout* $v \in V$.
2. $\langle v, v\rangle = 0$ *si et seulement si* $v = 0$.
3. $\langle u, v\rangle = \langle v, u\rangle$ *pour tout* $u, v \in V$.

Remarques : Une forme bilinéaire qui satisfait à toutes ces conditions est appelée *produit scalaire*. De plus, si $\dim V < \infty$ alors on dira que $(V, \langle\cdot,\cdot\rangle)$ est un *espace euclidien*. Finalement un produit scalaire est toujours non dégénéré. En effet, si $\langle u, v\rangle = 0$ pour tout $v \in V$ alors $\langle u, u\rangle = 0$ ce qui implique que $u = 0$ par la condition (2). De plus, par la condition (3) on a le même raisonnement à droite.

Définition 4.7 *Soit V un espace préhilbertien. La fonction*

$$\begin{aligned} \|\cdot\| : V &\to \mathbb{R} \\ v &\mapsto \|v\| = \sqrt{\langle v, v\rangle} \end{aligned}$$

est dite la norme *(ou la longueur) de $v \in V$.*

On définit *la distance* entre deux vecteurs $u, v \in V$ par $\|u - v\|$.

Proposition 4.11 *Soit V un espace préhilbertien. Nous avons*

1. $\|v\| \geq 0$, *pour tout* $v \in V$. *De plus,* $\|v\| = 0$ *si et seulement si* $v = 0$.
2. $\|\alpha v\| = |\alpha| \cdot \|v\|$, *pour tout* $v \in V$ *et* $\alpha \in \mathbb{R}$.
3. $\langle u, v\rangle = \frac{1}{4}\{\|u + v\|^2 - \|u - v\|^2\}$, *pour tout* $u, v \in V$ *(Polarization).*
4. $\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2$, *pour tout* $u, v \in V$ *(Parallélogramme).*

Démonstration. Le (1) découle directement de la définition du produit scalaire. Pour le (2), on a $\|\alpha v\|^2 = \langle \alpha v, \alpha v\rangle = \alpha^2\langle v, v\rangle$. En se rappelant que $\sqrt{\alpha^2} = |\alpha|$, le résultat suit. Pour les (3) et (4) considérons

$$\begin{aligned}\|u+v\|^2 &= \langle u+v, u+v\rangle = \langle u, u+v\rangle + \langle v, u+v\rangle \\ &= \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle \\ &= \|u\|^2 + 2\langle u,v\rangle + \|v\|^2.\end{aligned}$$

Si l'on remplace v par $-v$ dans la dernière expression on obtient $\|u-v\| = \|u\|^2 + \|v\|^2 - 2\langle u,v\rangle$. Dès lors le (3) est obtenu en soustrayant les deux dernières expressions. Tandis que le (4) est obtenu en les additionnant.

□

Théorème 4.12 *Soit V un espace préhilbertien. Nous avons*

$$|\langle u, v\rangle| \leq \|u\| \cdot \|v\|,$$

pour tout $u, v \in V$. De plus, l'égalité $|\langle u,v\rangle| = \|u\|\cdot\|v\|$ est vérifiée si et seulement si u et v sont linéairement dépendants.

Remarque : On appelle cette inégalité, *l'inégalité de Cauchy-Schwarz.*

Démonstration. Si $u = 0$ ou $v = 0$ alors l'inégalité est vérifiée. Supposons $u, v \neq 0$ et considérons les vecteurs unitaires

$$x = \frac{1}{\|u\|}\cdot u \quad \text{et} \quad y = \frac{1}{\|v\|}\cdot v.$$

Nous pouvons donc écrire $\langle u,v\rangle = \|u\|\cdot\|v\|\cdot\langle x,y\rangle$. Si $|\langle x,y\rangle| \leq 1$ alors le résultat s'en suit. En effet

$$\begin{aligned}|\langle x,y\rangle| &= \left|\frac{1}{4}\{\|x+y\|^2 - \|x-y\|^2\}\right|, \quad \text{(Polarization)}, \\ &\leq \frac{1}{4}\{\|x+y\|^2 + \|x-y\|^2\}, \quad \text{propriété de la valeur absolue dans } \mathbb{R}, \\ &= \frac{1}{4}\{2\|x\|^2 + 2\|y\|^2\} = 1, \quad \text{(Parallélogramme)}.\end{aligned}$$

De plus, si u et v sont linéairement dépendants alors sans perte de généralité on peut supposer que $u = \alpha v$ pour une certain $\alpha \in \mathbb{R}$; et donc

$$|\langle u,v\rangle| = |\langle \alpha v, v\rangle| = |\alpha|\cdot\|v\|^2 = |\alpha|\cdot\|v\|\cdot\|v\| = \|u\|\cdot\|v\|.$$

Réciproquement si $|\langle u,v\rangle| = \|u\|\cdot\|v\|$ alors deux cas se présentent : (1) si $u = 0$ ou $v = 0$ alors u et v sont évidemment linéairement dépendants ; (2) si $u, v \neq 0$ alors $|\langle x,y\rangle| = 1$ ou de façon équivalente $\langle x,y\rangle = \pm 1$.

(a) Si $\langle x,y\rangle = 1$ alors $\|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2\langle x,y\rangle = 1+1-2 = 0$.

(b) Si $\langle x,y\rangle = -1$ alors $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x,y\rangle = 1+1-2 = 0$.

Dans le cas (a), $\dfrac{1}{\|u\|}\cdot u - \dfrac{1}{\|v\|}\cdot v = x - y = 0$; dans le cas (b), $\dfrac{1}{\|u\|}\cdot u + \dfrac{1}{\|v\|}\cdot v = x + y = 0$, c'est-à-dire, dans les deux cas u et v sont linéairement dépendants.

□

Corollaire 4.13 *Soit V un espace préhilbertien. Nous avons*

$$\|u+v\| \leq \|u\| + \|v\|,$$

pour tout $u, v \in V$.

Remarque : Cette inégalité est appelée *l'inégalité du triangle.*

Démonstration. Nous avons

$$\begin{aligned} \|u+v\|^2 &= \|u\|^2 + \|v\|^2 + 2\langle u, v\rangle, \\ &\leq \|u\|^2 + \|v\|^2 + 2|\langle u, v\rangle|, \\ &\leq \|u\|^2 + \|v\|^2 + 2\|u\| \cdot \|v\|, \quad \text{par Cauchy-Schwarz,} \\ &= (\|u\| + \|v\|)^2. \end{aligned}$$

□

4.8 Les bases orthonormées : procédé de Gram-Schimdt

Définition 4.8 *Soit V un espace préhilbertien. On dit que deux vecteurs $u, v \in V$ sont* orthogonaux *si $\langle u, v\rangle = 0$. Dans ce cas on écrit $u \perp v$.*

Proposition 4.14 *Soit V un espace préhilbertien. Si dans une famille $\{x_\alpha\} \subset V$ les vecteurs sont orthogonaux deux-à-deux et non nuls alors cette famille est linéairement indépendante.*

Démonstration. Supposons qu'ils existent des nombres réels $\beta_1, ..., \beta_n$ tels que $\chi = \beta_1 x_{\alpha_1} + ... + \beta_n x_{\alpha_n} = 0$. Alors

$$0 = \langle \chi, x_{\alpha_j}\rangle = \left\langle \sum_{i=1}^{n} \beta_i x_{\alpha_i}, x_{\alpha_j} \right\rangle = \sum_{i=1}^{n} \beta_i \langle x_{\alpha_i}, x_{\alpha_j}\rangle = \beta_j \|x_{\alpha_j}\|^2$$

ce qui force $\beta_j = 0$ car $\|x_{\alpha_j}\|^2 \neq 0$. Ceci étant vérifié pour tout $j \in \{1, ..., n\}$ on conclut que la famille est linéairement indépendante.

□

Proposition 4.15 *Soit V un espace préhilbertien et $U \subset V$ un sous-espace finiment engendré. Alors*

$$V = U \oplus U^\perp.$$

Démonstration. On insiste sur le fait que la dimension de V peut être quelconque. D'une part, come U et $U^\perp$ sont des sous-espaces, nous avons $\{0\} \subset U \cap U^\perp$ et $U + U^\perp \subset V$. D'autre part si $x \in U \cap U^\perp$ alors $x \perp x$, c'est-à-dire, $x = 0$; et donc $\{0\} = U \cap U^\perp$. On remarque que la restriction du produit scalaire de V à U fait de U un espace euclidien. Soit $v \in V$ et considérons l'application linéaire $G_v(\cdot) = \langle \cdot, v\rangle$. La restriction $G_v|_U$ de G_v à l'espace euclidien U nous permet de choisir un unique $u \in U$ (Théorème 4.1) tel que

$$G_v|_U(y) = \langle y, u\rangle, \quad \text{pour tout } y \in U.$$

Et donc nous avons $\langle y, v\rangle = \langle y, u\rangle$ pour tout $y \in U$, c'est-à-dire, $v - u \in U^\perp$. On conclut que $V = U + U^\perp$, d'où le résultat.

□

Ce dernier résultat nous permet de définir

Définition 4.9 *Soit V un espace préhilbertien et $U \subset V$ un sous-espace finiment engendré. On dit que $U^\perp$ est* le supplémentaire orthogonal *de U.*

Une ré-écriture de la Proposition 4.7 s'énonce comme suit :

Proposition 4.16 *Soit V un espace euclidien et $U \subset V$ un sous-espace. Alors*

1. $\dim U + \dim U^\perp = \dim V$;
2. $(U^\perp)^\perp = U$.

Une propriété intéressante des espaces euclidiens réside dans le fait que l'on peut toujours choisir une *base othonormée*, c'est-à-dire, une base $\{u_1, ..., u_n\}$ telle que les u_i soient othogonaux deux-à-deux et $\|u_i\| = 1$.

Théorème 4.17 *Tout espace euclidien possède une base orthonormée.*

Démonstration. On procède par induction sur la dimension de l'espace euclidien. Soit V un espace euclidien de dimension n. Si $n = 1$ alors il existe $v \in V$ non nul et nécessairement $< v >= V$. Donc $\left\{\frac{1}{\|v\|} v\right\}$ est une base orthonormée de V. Si $n > 1$ alors soit $v \in V$ un vecteur de longueur 1. Alors par la Proposition 4.15 nous avons $V =< v > \oplus < v >^\perp$. Par la Proposition 4.16, $\dim < v >^\perp = n-1$ et possède donc une base orthonormée $\{v_1, ..., v_{n-1}\}$ par l'hypothèse d'induction. Par construction $\{v_1, ..., v_{n-1}, v\}$ est une famille orthogonale. Par la Proposition 4.14, elle est donc linéairement indépendante. On conclut que $\{v_1, ..., v_{n-1}, v\}$ est une base orthonormée de V.

□

Ce dernier résultat, bien qu'il garantisse l'existence d'une base orthonormée, ne nous donne pas de procédé de construction d'une telle base. C'est ici que le procédé d'orthogonalisation de Gram-Schmidt entre en scène.

Soit V un espace euclidien et $U \subset V$ un sous-espace. Soit $\{u_1, ..., u_k\}$ une base orthogonale de U. Complétons cette famille en une base $\{u_1, ..., u_k, v_{k+1}, ..., v_n\}$ de V. Cette famille n'est pas en générale orthogonale. On va construire en premier lieu un vecteur $w_{k+1} \in U^\perp$ en retirant à v_{k+1} ses projections sur chacun des vecteurs $u_1, ..., u_k$. Soient les scalaires suivants

$$\alpha_1 = \frac{\langle v_{k+1}, u_1\rangle}{\langle u_1, u_1\rangle}, \quad ..., \quad \alpha_k = \frac{\langle v_{k+1}, u_k\rangle}{\langle u_k, u_k\rangle},$$

et considérons $w_{k+1} = v_{k+1} - \alpha_1 u_1 - ... - \alpha_k u_k$. Si $i \in \{1, ..., k\}$ alors

$$\langle w_{k+1}, u_i\rangle = \langle v_{k+1}, u_i\rangle - \alpha_i \langle u_i, u_i\rangle = 0,$$

c'est-à-dire, $w_{k+1} \in U^\perp$. De plus, $w_{k+1} \neq 0$ sinon v_{k+1} dépendrait linéairement de $\{u_1, ..., u_k\}$. On a donc

$$< u_1, ..., u_k, v_{k+1} >=< u_1, ..., u_k, w_{k+1} >,$$

c'est-à-dire, $\{u_1, ..., u_k, w_{k+1}\}$ est linéairement indépendant et est une base orthogonale du sous-espace $U_{k+1} =< u_1, ..., u_k, w_{k+1} >$. On répète donc le processus jusqu'à v_n pour enfin obtenir une base orthogonale $\{u_1, ..., u_k, w_{k+1}, ..., w_n\}$ de V. Si l'on veut une base orthonormée alors il ne nous reste plus qu'à diviser chaque membre de cette basse orthogonale par sa norme respective, c'est-à-dire,

$$\left\{\frac{1}{\|u_1\|} u_1, ..., \frac{1}{\|u_k\|} u_k, \frac{1}{\|w_{k+1}\|} w_{k+1}, ..., \frac{1}{\|w_n\|} w_n\right\}.$$

4.9 Espace hermitien

Dans cette section le corps de référence sera $\mathbb{C}$ les nombres complexes. Soit V un espace vectoriel.

Définition 4.10 *Un produit scalaire (hermitien) est une application*

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$$

telle que

1. $\langle v, v \rangle \geq 0$. *De plus,* $\langle v, v \rangle = 0$ *si et seulement si* $v = 0$ *;*
2. $\langle au + bv, w \rangle = a\langle u, w \rangle + b\langle v, w \rangle$ *pour tout* $u, v, w \in V$ *et* $a, b \in \mathbb{C}$ *; et*
3. $\langle u, v \rangle = \overline{\langle v, u \rangle}$, *pour tout* $u, v \in V$.

Il est naturel de poser la question : pourquoi ne pas utiliser une simple forme bilinéaire ?

Proposition 4.18 *Une forme bilinéaire symétrique et définie positive sur* $\mathbb{C}$ *est nulle.*

Démonstration. Soit ϕ une telle forme bilinéaire. On a toujours $\phi(u, v) = \frac{1}{2}[\phi(u + v, u + v) - \phi(u, u) - \phi(v, v)]$, pour tout $u, v \in V$. Si on montre que $\phi(v, v) = 0$ pour tout $v \in V$ alors le résultat s'en suivra. On affirme que c'est le cas si $\phi(v, v) \geq 0$ pour tout $v \in V$. En effet, nous avons

$$0 \leq \phi(iv, iv) = i^2 \phi(v, v) = -\phi(v, v) \leq 0.$$

□

Exemple : Le produit scalaire sur $\mathbb{C}^n$.

On se rappellera que le produit scalaire est défini par

$$\begin{array}{rcl} \langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n & \longrightarrow & \mathbb{C} \\ (u, v) & \mapsto & \langle u, v \rangle = \sum_{i=1}^{n} a_i \overline{b_i}, \end{array}$$

où $u = (a_1, a_2, ..., a_n)$ et $v = (b_1, b_2, ..., b_n)$. On laisse au lecteur le soin de démontrer que le produit scalaire est bien linéaire en la première variable. Par contre, supposons que $\langle u, v \rangle = 0$ pour tout $u \in \mathbb{C}^n$. En particulier, nous devons avoir $\sum_{i=1}^{n} b_i \overline{b_i} = \langle v, v \rangle = 0$. Chaque $b_i \overline{b_i} \geq 0$. On déduit donc que chaque $b_i \overline{b_i} = 0$, c'est-à-dire, $b_i = 0$ pour tout $i = 1, ..., n$ et donc $v = 0$. On conclut que le produit scalaire est non dégénéré.

Définition 4.11 *Un espace vectoriel complexe* V *de dimension fini munit d'un produit scalaire (hermitien) est appelé espace hermitien.*

4.10 Similitudes et différences entre les cas réel et complexe

L'inégalité de Cauchy-Schwarz est vérifiée dans le cas complexe. En effet,

Proposition 4.19 *Si* V *est un espace préhilbertien complexe alors*

$$|\langle u, v \rangle| \leq \|u\| \|v\|, \quad \textit{pour tout } u, v \in V.$$

De plus, l'égalité est vérifiée si et seulement si u *et* v *sont linéairement dépendants.*

Démonstration. Si $\langle u, v\rangle = 0$ alors le résultat est vérifié. Sinon posons $\lambda = \dfrac{|\langle u, v\rangle|}{\langle u, v\rangle}$ et nous avons

$$\begin{aligned} 0 \leq \left\| \frac{\lambda u}{\|u\|} - \frac{v}{\|v\|} \right\|^2 &= |\lambda|^2 \frac{\|u\|^2}{\|u\|^2} - 2\mathrm{Re}\left(\left\langle \frac{\lambda u}{\|u\|}, \frac{v}{\|v\|} \right\rangle \right) + \frac{\|v\|^2}{\|v\|^2} \\ &= 2 - 2\frac{|\langle u, v\rangle|}{\|u\|\|v\|}. \end{aligned}$$

On en tire donc l'inégalité. Finalement on complète la preuve suivant la démonstration du Théorème 4.12 où on pose $e^{i\theta} = \langle x, y\rangle$ et $c = e^{-i\theta}$. On déduit le résultat en considérant $\|cx - y\|^2$.

□

Proposition 4.20 *Soit V un espace hermitien. Alors V possède une base orthonormée.*

Démonstration. L'espace V étant de dimension finie, il admet une base $\{v_1, ..., v_n\}$. On applique le procédé de Gram-Schmidt à cette base et on obtient une base orthonormée.

□

Corollaire 4.21 *Soit V un espace hermitien et $\{v_1, ..., v_n\}$ une base orthonormée. Si $v \in V$ alors*

$$v = \langle v, v_1\rangle v_1 + ... + \langle v, v_n\rangle v_n.$$

Démonstration. Ils existent $a_1, ..., a_n \in \mathbb{C}$ tels que $v = a_1 v_1 + ... + a_n v_n$. Alors

$$\langle v, v_i\rangle = \langle a_1 v_1 + ... + a_n v_n, v_i\rangle = a_1\langle v_1, v_i\rangle + ... + a_n\langle v_n, v_i\rangle = a_i.$$

□

Corollaire 4.22 *Soit V un espace hermitien et $\{v_1, ..., v_n\}$ une base orthonormée. Si $v, w \in V$ alors*

$$\langle v, w\rangle = \sum_{k=1}^{n} a_k \overline{b_k},$$

où $v = \sum_{k=1}^n a_k v_k$ et $w = \sum_{l=1}^n b_l v_l$.

Démonstration. On a

$$\begin{aligned} \langle v, w\rangle &= \left\langle \sum_{k=1}^{n} a_k v_k, \sum_{l=1}^{n} b_l v_l \right\rangle \\ &= \sum_{k=1}^{n}\sum_{l=1}^{n} a_k \overline{b_l}\langle v_k, v_l\rangle \\ &= \sum_{k=1}^{n} a_k \overline{b_k}. \end{aligned}$$

□

Si $U \subset V$ est un sous-espace alors de la même façon que dans le cas réel on définit

$$U^{\perp} = \{v \in V \,|\, \langle v, u\rangle = 0 \text{ pour tout } u \in U\}.$$

Exercice : Montrez que $U^{\perp}$ est bien un sous-espace de V.

Proposition 4.23 *Soit V un espace hermitien. Si $U \subset V$ est un sous-espace alors*

$$V = U \oplus U^{\perp}.$$

En particulier,

1. $\dim V = \dim U + \dim U^{\perp}$ *et*
2. $U^{\perp\perp} = U$.

Démonstration. Il est évident que $U \cap U^{\perp} = \{0\}$. De plus, on a nécessairement que $U + U^{\perp} \subset V$. Soit $\{u_1, ..., u_k\}$ une base orthonormée de U. Si $v \in V$ alors on a

$$v = \underbrace{\sum_{i=1}^{k} \langle v, u_i \rangle u_i}_{\in U} + \left(\underbrace{v - \sum_{i=1}^{k} \langle v, u_i \rangle u_i}_{\in U^{\perp}} \right).$$

En effet, on a

$$\begin{aligned} \langle v - \sum_{i=1}^{k} \langle v, u_i \rangle u_i, u_j \rangle &= \langle v, u_j \rangle - \sum_{i=1}^{k} \langle \langle v, u_i \rangle u_i, u_j \rangle \\ &= \langle v, u_j \rangle - \langle v, u_j \rangle \langle u_j, u_j \rangle \\ &= 0, \end{aligned}$$

d'où le premier résultat. On en tire l'égalité $\dim V = \dim U + \dim U^{\perp}$. De plus, comme $U \subset U^{\perp\perp}$ et que $\dim U^{\perp} + \dim U^{\perp\perp} = \dim V = \dim U + \dim U^{\perp}$, on déduit que $\dim U^{\perp\perp} = \dim U$, et donc $U = U^{\perp\perp}$.

□

4.11 Représentation de Riesz : le cas complexe

Théorème 4.24 *(*Représentation de Riesz*) Soit V un espace hermitien. L'application (non linéaire)*

$$\begin{aligned} \Omega : V &\longrightarrow V^{\vee} \\ v &\mapsto \langle \cdot, v \rangle \end{aligned}$$

est un isomorphisme de groupes abéliens.

Démonstration. On a bien un homomorphisme de groupe car $\langle \cdot, v + u \rangle = \langle \cdot, v \rangle + \langle \cdot, u \rangle$ pour tout $v, u \in V$.

Remarque : Lorsque nous avons montré que T, une application linéaire, est injective si et seulement si $\ker T = \{0\}$, nous n'avons utilisé que les propriétés de groupes.

Si $\Omega(v) = 0$ alors $\langle u, v \rangle = 0$ pour tout $u \in V$. Mais $\langle \cdot, \cdot \rangle$ est non-dégénérée. On conclut que $v = 0$, c'est-à-dire, Ω est injective. Il ne nous reste qu'à montrer que Ω est surjective. Soit $f \in V^{\vee}$. Si $f = 0$ alors $\Omega(0) = f$. Sinon $\dim \ker f = \dim V - 1$ (pourquoi ?) et $\dim(\ker f)^{\perp} = 1$. Soit $b \in (\ker f)^{\perp}$ tel que $f(b) = 1$ (pourquoi est-ce possible ?). On affirme que

$$\Omega\left(\frac{b}{\|b\|^2} \right) = f.$$

En effet, on a $\langle u, \frac{b}{\|b\|^2}\rangle = 0 = f(u)$ pour tout $u \in \ker f$. De plus,

$$\langle b, \frac{b}{\|b\|^2}\rangle = \frac{\langle b, b\rangle}{\|b\|^2} = 1 = f(b),$$

c'est-à-dire, que ces deux applications linéaires étant égales sur une base doivent être égales en générale.
□

Corollaire 4.25 *Soit V un espace préhilbertien (complexe) et $U \subset V$ un sous-espace finiment engendré. Alors*

$$V = U \oplus U^\perp.$$

Démonstration. C'est la même démonstration que la Proposition 4.15 où le Théorème 4.24 remplace le Théorème 4.1.

□

4.12 L'adjointe : $T^\dagger$

Théorème 4.26 *Soit V un espace hermitien et $T \in \mathscr{L}(V)$. Alors il existe une unique application $T^* \in \mathscr{L}(V)$ telle que*

$$\langle T(v), w\rangle = \langle v, T^*(w)\rangle, \quad \textit{pour tout } v, w \in V.$$

Démonstration. Soit $w \in V$ et considérons la forme linéaire $v \mapsto \langle T(v), w\rangle$. Par le Théorème 4.24, il existe un unique $v' \in V$ tel que

$$\langle v, v'\rangle = \langle T(v), w\rangle, \quad \text{pour tout } v \in V.$$

On pose $T^*(w) = v'$. On affirme que T^* est linéaire. En effet, si $u, w, v \in V$ et $a, b \in \mathbb{C}$ alors

$$\begin{aligned}
\langle v, T^*(au + bw)\rangle &= \langle T(v), au + bw\rangle \\
&= \overline{a}\langle T(v), u\rangle + \overline{b}\langle T(v), w\rangle \\
&= \overline{a}\langle v, T^*(u)\rangle + \overline{b}\langle v, T^*(w)\rangle \\
&= \langle v, aT^*(u) + bT^*(w)\rangle.
\end{aligned}$$

Finalement on tire la linéarité et l'unicité de l'adjointe de la non-dégénérescence du produit scalaire (voir la démonstration du Théorème 4.3).

□

Remarque : Dans ce cas l'adjointe est souvent notée dans la littérature $T^\dagger$.

On a maintenant l'équivalent de Proposition 4.4, c'est-à-dire,

Proposition 4.27 *Soit V un espace hermitien et $T \in \mathscr{L}(V)$. Si $\beta = \{v_1, \ldots v_n\}$ est une base orthonormée de V alors*

$$[T^\dagger]^\beta_\beta = {}^t\left(\overline{[T]^\beta_\beta}\right).$$

Démonstration. Soit $[T]_{\beta}^{\beta} = (a_{ij}) \in \mathbb{M}_{n\times n}(\mathbb{C})$ et $S \in \mathscr{L}(V)$ l'application linéaire correspondant à la matrice (b_{ji}) où $b_{ji} = \overline{a_{ij}}$ par rapport à la base β. D'une part, on a

$$\langle T(v_j), v_i\rangle = \sum_{k=1}^{m} a_{kj}\langle v_k, v_i\rangle = a_{ij}.$$

D'autre part, on a

$$\langle v_j, S(v_i)\rangle = \sum_{k=1}^{n} \langle v_j, b_{ki}v_k\rangle = \sum_{k=1}^{n} \overline{b_{ki}}\langle v_j, v_k\rangle = \overline{b_{ji}} = a_{ij}.$$

Par la construction de $T^{\dagger}$, on a

$$\langle v_j, T^{\dagger}(v_i)\rangle = \langle T(v_j), v_i\rangle = \langle v_j, S(v_i)\rangle.$$

Comme $\langle\cdot,\cdot\rangle$ est non-dégénéré, on a $T^{\dagger}(v_i) = S(v_i)$ pour tous les éléments de la base β. On conclut que $T^{\dagger} = S$.

□

Exercice : Si V est un espace hermitien, $T, T_1, T_2 \in \mathscr{L}(V)$ et $\alpha \in \mathbb{C}$ alors montrez que

1. $(\alpha T)^{\dagger} = \overline{\alpha}T^{\dagger}$;
2. $(T_1 + T_2)^{\dagger} = T_1^{\dagger} + T_2^{\dagger}$;
3. $(T_1 \circ T_2)^{\dagger} = T_2^{\dagger} \circ T_1^{\dagger}$; et
4. $T^{\dagger\dagger} = T$.

CHAPITRE 5

VALEURS ET VECTEURS PROPRES

5.1 Définitions et propriétés élémentaires

Soit V un $\mathbb{k}$-espace vectoriel et $T \in \mathscr{L}(V)$.

Définition 5.1 *Un scalaire $\lambda \in \mathbb{k}$ est une valeur propre de T s'il existe un vecteur non nul $v \in V$ tel que*

$$T(v) = \lambda v.$$

Dans ce cas on dit que v est un vecteur propre associé à la valeur propre λ.

Proposition 5.1 *Un scalaire $\lambda \in \mathbb{k}$ est une valeur propre de T si et seulement si* $\ker(T - \lambda Id_V) \neq \{0\}$.

Corollaire 5.2 *Si* $\dim V < \infty$ *alors l'application linéaire T n'est pas inversible si et seulement si* 0 *est une valeur propre de T.*

Démonstration. Lorsque $\dim V < \infty$ vérifier l'injectivité d'une application linéaire $T \in \mathscr{L}(V)$ est équivalent à vérifier la surjectivité de celle-ci par le Théorème 3.3.

□

Exemple : Soit $A \in \mathbb{M}_{n\times n}(\mathbb{k})$ et considérons l'applications linéaire

$$\begin{array}{rccc} T_A : \mathbb{k}^n & \longrightarrow & \mathbb{k}^n \\ X & \mapsto & AX, \end{array}$$

où X désigne un vecteur colonne. Si $\lambda \in \mathbb{k}$ est une valeur propre de T_A alors trouver les vecteurs propres associés est équivalent à solutionner le système homogène

$$(T_A - \lambda Id_{\mathbb{k}^n})X = (A - \lambda \cdot \mathbb{1}_n)X = 0,$$

où $\mathbb{1}_n$ est la matrice identité $n \times n$.

Proposition 5.3 *Soit V un $\mathbb{k}$-espace vectoriel finiment engendré et β une base de V. Soit $T \in \mathscr{L}(V)$ et $A = [T]^{\beta}_{\beta}$. Les énoncés suivants sont équivalents :*

1. *$\lambda \in \mathbb{k}$ est une valeur propre de T.*
2. *Le système homogène $(A - \lambda \cdot \mathbb{1}_n)X = 0$ admet une solution non triviale.*
3. *$\lambda \in \mathbb{k}$ est une racine du polynôme caractéristique $p(x) = \det[A - x \cdot \mathbb{1}_n]$.*

Démonstration. Supposons λ une valeur propre de T avec $v \in V$ un vecteur propre associé. Alors posons $X = [v]^{\beta}$ et considérons

$$\begin{aligned} AX &= [T]^{\beta}_{\beta}[v]^{\beta} \\ &= [T(v)]^{\beta} \\ &= [\lambda \cdot v]^{\beta} \\ &= \lambda[v]^{\beta} \quad \text{car l'iso. des coord. est linéaire (3.2)} \\ &= \lambda \cdot X. \end{aligned}$$

Réciproquement, soit X une solution non triviale du système homogène $(A - \lambda \cdot \mathbb{1}_n)X = 0$ et posons $v = \kappa_{\beta}(X) \in V$. On affirme que v est un vecteur propre de T associé à la valeur propre λ. En effet,

$$\begin{aligned} T(v) &= \kappa_{\beta}([T(v)]^{\beta}) \\ &= \kappa_{\beta}([T]^{\beta}_{\beta}[v]^{\beta}) \\ &= \kappa_{\beta}(AX) \\ &= \kappa_{\beta}(\lambda \cdot X) \\ &= \lambda \cdot \kappa_{\beta}(X) \\ &= \lambda \cdot v. \end{aligned}$$

Nous repoussons au chapitre 6, la démonstration du fait : $A \in \mathbb{M}_{n\times n}(\mathbb{k})$ est inversible si et seulement si $\det A \neq 0$.

□

Exemple : Trouvez les valeurs propres associées à la matrice $A = \begin{bmatrix} -3 & -2 \\ 2 & 2 \end{bmatrix}$. En d'autres mots on cherche pour quels $\lambda \in \mathbb{R}$ la matrice $A - \lambda \cdot \mathbb{1}_2$ n'est pas inversible ou de façon équivalente n'est pas de rang maximal. On procède à l'aide de l'algorithme de réduction de Gauss-Jordan appliqué à la matrice

$$\begin{bmatrix} -3-\lambda & -2 \\ 2 & 2-\lambda \end{bmatrix}.$$

1. On interchange les lignes (1) et (2) pour obtenir $\begin{bmatrix} 2 & 2-\lambda \\ -3-\lambda & -2 \end{bmatrix}$.
2. La ligne (2) devient (2) + $\frac{1}{2}(3+\lambda)$ (1), c'est-à-dire, $\begin{bmatrix} 2 & 2-\lambda \\ 0 & -2+\frac{1}{2}(3+\lambda)(2-\lambda) \end{bmatrix}$.

Donc $rk(A - \lambda \cdot \mathbb{1}_2) < 2$ si et seulement si $-2 + \frac{1}{2}(3+\lambda)(2-\lambda) = 0$, c'est-à-dire, les valeurs propres de A sont les racines du polynômes

$$\lambda^2 + \lambda - 2.$$

On conclut que les valeurs propres de A sont -2 et 1.

Théorème 5.4 *Soit V un $\mathbb{k}$-espace vectoriel et $T \in \mathscr{L}(V)$. Soient r vecteurs propres $v_1, ..., v_r$ de T correspondants à r valeurs propres distinctes $\lambda_1, ..., \lambda_r$. Alors $\{v_1, ..., v_r\}$ est linéairement indépendant.*

Démonstration. On procède par induction sur r. Si $r = 1$ alors étant donné que $v_1 \neq 0$, par définition, $\{v_1\}$ est linéairement indépendant. On suppose le résultat vrai pour $r-1$. S'ils existent $a_1, ..., a_r \in \mathbb{k}$ tels que

$$a_1 v_1 + ... + a_r v_r = 0 \tag{5.1}$$

alors en appliquant T on a $0 = a_1 T(v_1) + ... + a_r T(v_r) = a_1\lambda_1 v_1 + ... + a_r\lambda_r v_r$. Si on soustrait à cette dernière expression λ_1-fois l'expression (5.1) alors on obtient

$$\begin{aligned} 0 &= a_1(\lambda_1 - \lambda_1)v_1 + a_2(\lambda_2 - \lambda_1)v_2 + ... + a_r(\lambda_r - \lambda_1)v_r \\ &= a_2(\lambda_2 - \lambda_1)v_2 + ... + a_r(\lambda_r - \lambda_1)v_r. \end{aligned}$$

Par l'hypothèse d'induction, $a_i(\lambda_i - \lambda_1) = 0$, pour tout $2 \leq i \leq r$. Mais comme $\lambda_i \neq \lambda_1$, pour tout $2 \leq i \leq r$, on conclut que $a_i = 0$ lorsque $i > 1$. Mais à nouveau, comme $v_1 \neq 0$ car il est vecteur propre de T, la relation (5.1) force $a_1 = 0$. Le résultat s'en suit.

□

5.2 Opérateurs diagonalisables

Définition 5.2 *Soit V un espace finiment engendré et $T \in \mathscr{L}(V)$. On dit que T est diagonalisable s'il existe une base de V formée de vecteurs propres de T.*

Remarque : La définition se justifie par la remarque suivante : si $\beta = \{v_1, ..., v_n\}$ est une base de V composée de vecteurs propres associés au valeurs propres (pas nécessairement distinctes) $\lambda_1, ..., \lambda_n$ alors

$$[T]_\beta^\beta = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

c'est-à-dire, une matrice diagonale.

Deux types d'obstructions à la diagonalisation d'une matrice surviennent :

1. Considérons la matrice $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ et calculons le rang de $\begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix}$. Si on procède comme dans l'exemple précédent, on obtient que les valeurs propres sont les racines du polynôme $\lambda^2 + 1$. Malheureusement sur $\mathbb{R}$, ce polynôme n'a pas de racine et n'est donc pas diagonalisable sur ce corps. Si on passe aux nombres complexes alors les valeurs propres sont $\pm i$ avec vecteurs propres associés $(i, 1)$ et $(-i, 1)$. La matrice diagonale correspondante est $\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$. En conclusion, si on travaille sur un corps $\mathbb{k}$ *algébriquement clos*, c'est-à-dire, pour lequel tout polynôme non constant à coefficients dans $\mathbb{k}$ admet une racine, alors cette obstruction est levée.

2. Considérons la matrice $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ et calculons le rang de $\begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix}$. On conclut que λ est une valeur propre si et seulement si $\lambda = 1$. Dans ce cas le système homogène $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} X = 0$ n'admet qu'une solution linéairement indépendante, c'est-à-dire, $(1,0)$. Il n'y a donc pas suffisament de vecteur propre pour composer une base. C'est une obstruction qui ne peut pas être levée. On conclut que certaines matrices ne pourront pas être diagonalisées.

Définition 5.3 *Soit V un $\mathbb{k}$-espace vectoriel finiment engendré et $T \in \mathscr{L}(V)$. L'ensemble*

$$\sigma(T) = \{\lambda \in \mathbb{k} \mid \lambda \text{ est une valeur propre de } T\}$$

est appelé le spectre de l'opérateur T.

Proposition 5.5 *Soit V un $\mathbb{k}$-espace vectoriel finiment engendré et $T \in \mathscr{L}(V)$. Si $\mathbb{k}$ est algébriquement clos alors $\sigma(T) \neq \emptyset$.*

Démonstration. Par la propriété de $\mathbb{k}$, le polynôme caractéristique possède toujours une racine. Le résultat s'en suit par la Proposition 5.3.

□

Proposition 5.6 *Soit V un $\mathbb{k}$-espace vectoriel finiment engendré et $T \in \mathscr{L}(V)$. Alors*

$$|\sigma(T)| \leq \dim V.$$

Démonstration. Par définition du polynôme caractéristique, le degré de celui-ci est égal à $\dim V$. Par la Proposition 5.3, $|\sigma(T)| \leq \dim V$.

□

Proposition 5.7 *Soit V un $\mathbb{k}$-espace vectoriel finiment engendré et $T \in \mathscr{L}(V)$. Si T est diagonalisable alors*

$$V = \bigoplus_{\lambda \in \sigma(T)} \ker(T - \lambda Id_V).$$

Démonstration. Soit $\{\lambda_1, ..., \lambda_k\} = \sigma(T)$ et soit $\{v_{11}, v_{12}, ..., v_{1i_1}, v_{21}, ...v_{2i_2}, ..., v_{k1}, ..., v_{ki_k}\}$ une base de V composée de vecteurs propres de T telle que

$$\{v_{j1}, ..., v_{ji_j}\} \subset \ker(T - \lambda_j Id_V), \quad 1 \leq j \leq k.$$

On affirme que $< v_{j1}, ..., v_{ji_j} >= \ker(T - \lambda_j Id_V)$, pour tout $1 \leq j \leq k$. Sinon il existe $v \in V$ tel que $T(v) = \lambda_j v$ et $\{v_{j1}, ..., v_{ji_j}, v\}$ est linéairement indépendant pour un certain $1 \leq j \leq k$. En utilisant le raisonnement utilisé dans la preuve du Théorème 5.4, on conclut que

$$\{v_{11}, v_{12}, ..., v_{1i_1}, v_{21}, ...v_{2i_2}, ..., v_{k1}, ..., v_{ki_k}, v\}$$

est linéairement indépendant ce qui est impossible car la cardinalité de cet ensemble est de $\dim V + 1$. Le résultat s'en suit.

□

5.3 Polynôme minimal

Soit V un $\mathbb{k}$-espace vectoriel finiment engendré et $T \in \mathscr{L}(V)$. Rappelons que la Proposition 3.16 implique que $\dim \mathscr{L}(V) = n^2$, où $n = \dim V$. Donc nécessairement la suite

$$Id_V, T, T^2, ..., T^{n^2}$$

est linéairement dépendante dans $\mathscr{L}(V)$. Rappelons que $T^k = \underbrace{T \circ ... \circ T}_{k\text{-fois}}$ et posons $T^0 = Id_V$. Il existe donc un $0 < k \leq n^2$ tel que

1. $Id_V, T, ..., T^{k-1}$ sont linéairement indépendants ; et
2. $Id_V, T, ..., T^k$ sont linéairement dépendants.

Ils existent donc des scalaires $a_0, a_1, ..., a_k \in \mathbb{k}$ non tous nuls tels que

$$a_0 Id_V + a_1 T + ... + a_k T^k = 0.$$

Définition 5.4 *Le polynôme $m = a_0 + a_1 x + ... + a_k x^k$ est appelé un polynôme minimal associé à T.*

Exercice : Soit $T \in \mathscr{L}(V)$. Soit $p, q \in \mathscr{P}(\mathbb{k})$

1. Montrez que $(x \cdot p)(T) = T \circ p(T) = p(T) \circ T$.
2. Concluez que $(p \cdot q)(T) = p(T) \circ q(T) = q(T) \circ p(T)$.

Proposition 5.8 *Si p est un polynôme tel que $p(T) = 0$ alors m divise p, c'est-à-dire, $p = qm$ pour un certain $q \in \mathscr{P}(\mathbb{k})$.*

Démonstration. On applique la division euclidienne (Théorème C.2) à p et m. Ils existent donc deux polynômes Q et R tel que $p = Qm + R$ et $\deg(R) < \deg(m)$. On a donc

$$R(T) = (p - Qm)(T) = p(T) - Q(T)m(T) = 0.$$

Mais dès lors ils existent $b_0, ..., b_j \in \mathbb{k}$, $j < \deg(m)$ tels que $R = b_0 + b_1 x + ... + b_j x^j$ et donc $0 = R(T) = b_0 Id_V + b_1 T + ... + b_j T^j$ ce qui contredit la nature du choix de $k = \deg(m)$ à moins que $b_0 = b_1 = ... = b_j = 0$, c'est-à-dire, $R = 0$.

□

Remarque : Si $\alpha \in \mathbb{k}^*$ alors $m' = \alpha m$ est un second polynôme minimal de T. La Proposition 5.8 démontre que tout autre polynôme minimal doit être de cette forme. De plus, par la construction du polynôme minimal, le coefficient $a_{\deg(m)}$ doit être non nul (pourquoi ?). On peut donc diviser le polynôme de la Définition 5.4 par ce coefficient et obtenir **LE** polynôme minimal de T. Celui-ci s'écrit

$$m_T = x^k + ... + \alpha_1 x + \alpha_0,$$

pour des uniques $\alpha_0, \alpha_1, ..., \alpha_{k-1} \in \mathbb{k}$.

Théorème 5.9 *Soit V un $\mathbb{k}$-espace vectoriel finiment engendré non nul. Une application linéaire T est diagonalisable si et seulement si son polynôme minimal s'écrit*

$$m_T = (x - \xi_1)(x - \xi_2)...(x - \xi_k)$$

avec $\xi_1, \xi_2, ..., \xi_k \in \mathbb{k}$ distincts et $0 < k \leq \dim V$.

Démonstration. Supposons que T soit diagonalisable. Alors par la Proposition 5.7 on a

$$V = \ker(T - \lambda_1 Id_V) \oplus ... \oplus \ker(T - \lambda_k Id_V),$$

où $\{\lambda_1, ..., \lambda_k\} = \sigma(T)$ les k-valeurs propres distinctes de T. Nécessairement on doit avoir $0 < k \leq \dim V$. On affirme que le polynôme $p = (x - \lambda_1)...(x - \lambda_k)$ est égal à m_T. Si $\{v_1, ..., v_n\}$ est une base de vecteurs propres de V alors pour tout $1 \leq i \leq n$, il existe un $1 \leq j \leq k$ tel que $v_i \in \ker(T - \lambda_j Id_V)$ et donc

$$p(T)(v_i) = [(T - \lambda_1 Id_V) \circ ... \circ (T - \lambda_{j-1} Id_V) \circ (T - \lambda_{j+1} Id_V) \circ ... \circ (T - \lambda_k Id_V) \circ (T - \lambda_j Id_V)](v_i) = 0.$$

Comme $p(T)$ est nul sur une base de V, on doit avoir $p(T) = 0$. On conclut par la Proposition 5.8 que m_T divise p. Mais il est clair que si on enlève certains facteurs de p alors $p(T) \neq 0$. Par exemple, considérons le polynôme $p^* = (x - \lambda_1)...(x - \lambda_{k-1})$ et v un vecteur propre associé à la valeur propre λ_k. On a alors

$$\begin{aligned} p^*(T)(v) &= (T - \lambda_1 Id_V) \circ ... \circ (T - \lambda_{k-1} Id_V)(v) \\ &= (\lambda_k - \lambda_1)...(\lambda_k - \lambda_{k-1})\, v \neq 0. \end{aligned}$$

Ceci implique que tous les facteurs de p doivent être des facteurs de m_T. Comme le coefficient dominant de p est 1. On doit avoir $m_T = p$.

Réciproquement, supposons que $m_T = (x - \xi_1)(x - \xi_2)...(x - \xi_k)$, avec $\xi_1, \xi_2, ..., \xi_k \in \mathbb{k}$ distincts et $0 < k \leq \dim V$. On a certainement

$$\ker(T - \xi_1 Id_V) \oplus ... \oplus \ker(T - \xi_k Id_V) \subset V.$$

Pour chaque $1 \leq i \leq k$ considérons le polynôme

$$q_i = \frac{m_T}{x - \xi_i}.$$

Par le Théorème C.5, ils existent $p_1, ..., p_k$ des polynômes tels que

$$1 = q_1 p_1 + ... + q_k p_k.$$

Ceci implique que $Id_V = q_1(T) \circ p_1(T) + ... + q_k(T) \circ p_k(T)$. Si on pose $W_i = \mathrm{Im}(q_i(T) \circ p_i(T))$ alors on a

$$V = W_1 + ... + W_k.$$

Le résultat suivra si on montre que $\{0\} \neq W_i \subset \ker(T - \xi_i Id_V)$. Si $w \in W_i$ alors il existe $v \in V$ tel que $w = q_i(T) \circ p_i(T)(v)$. Mais $(T - \xi_i Id_V)(w) = (T - \xi_i Id_V) \circ q_i(T) \circ p_i(T)(v) = m_T(T) \circ p_i(T)(v) = 0$, c'est-à-dire, $W_i \subset \ker(T - \xi_i Id_V)$.

S'il existe $W_i = \{0\}$ alors on peut écrire $v = \sum_{j \neq i} q_j(T) \circ p_j(T)(v)$. Mais dès lors

$$q_i(T)(v) = \sum_{j \neq i} q_i(T) \circ q_j(T) \circ p_j(T)(v) = 0$$

car m_T divise $q_i q_j$ pour tout $j \neq i$ ce qui est une contradiction car $q_i(T) = 0$ et $\deg(q_i) < \deg(m_T)$.

□

5.4 Opérateurs commutants

Soit V un $\mathbb{k}$-espace vectoriel et $T, S \in \mathscr{L}(V)$.

Définition 5.5 *L'application linéaire*

$$[T,S] = T \circ S - S \circ T$$

est appelée le commutateur des deux applications.

Définition 5.6 *On dit qu'un sous-espace $U \subset V$ est T-invariant si $T(U) \subset U$.*

Remarque : Dans ce cas $T|_U \in \mathscr{L}(U)$.

Soit V un $\mathbb{k}$-espace vectoriel finiment engendré.

Théorème 5.10 *Soient $T_1, ..., T_k \in \mathscr{L}(V)$ des transformations linéaires diagonalisables. Si $[T_i \circ T_j] = 0$, pour tout $1 \leq i, j \leq k$, alors il existe une base de V composée de vecteurs propres communs à toutes les applications T_i.*

Démonstration. On procède par induction sur la dimension de V. Si toutes les applications T_i n'ont qu'une valeur propre (en particulier si $\dim V = 1$) alors $T_i = \alpha_i Id_V$ avec $\alpha_i \in \mathbb{k}$ et $1 \leq i \leq k$. Dans ce cas toute base de V est composée de vecteurs propres communs à toutes les applications T_i. Supposons maintenant que $\dim V > 1$ et que, sans perte de généralité, T_1 possède au moins deux valeurs propres distinctes. Par la Proposition 5.7, on a

$$V = \underbrace{\ker(T - \lambda_1 Id_V)}_{W_1} \oplus ... \oplus \underbrace{\ker(T - \lambda_k Id_V)}_{W_k},$$

où $\{\lambda_1, ..., \lambda_k\} = \sigma(T)$ sont toutes les valeurs propres distinctes de T_1. On affirme que les W_i sont T_j invariants pour tout $1 \leq i, j \leq k$. En effet, si $w \in W_i$ alors

$$T_i(T_j(w)) = T_j(T_i(w)) = T_j(\lambda_i w) = \lambda_i T_j(w)$$

car par hypothèse $[T_i, T_j] = 0$. On a donc $T_j(W_i) \subset W_i$. Mais on ne peut pas affirmer pour autant que $T_j(w)$ soit vecteur propre de T_i (pourquoi ?). Par contre, toutes les restrictions $T_j|_{W_i} \in \mathscr{L}(W_i)$ sont diagonalisables. En effet, tous les polynômes minimaux $m_{T_j|_{W_i}}$ doivent diviser les polynômes minimaux correspondants m_{T_j}. Ceci force les $m_{T_j|_{W_i}}$ à être de la forme prescrite par le Théorème 5.9. De plus $\dim W_i < \dim V$ car $k \geq 2$. Le résultat suit par l'hypothèse d'induction.

□

5.5 Opérateurs hermitiens

Voici deux familles d'applications linéaires très importantes en physique :

Définition 5.7 *Soit V un espace hermitien. On dit qu'une application linéaire $T \in \mathscr{L}(V)$ est un opérateur hermitien si $T = T^\dagger$. On dit que T est un opérateur normal si $T \circ T^\dagger = T^\dagger \circ T$.*

Lemme 5.11 *Soit V un espace hermitien et $T \in \mathscr{L}(V)$ un opérateur normal. Alors T et $T^\dagger$ ont un vecteur propre $v \in V$ en commun. De plus, si $\lambda \in \mathbb{C}$ est tel que $T(v) = \lambda v$ alors $T^\dagger(v) = \overline{\lambda} v$.*

Démonstration. Comme $\mathbb{C}$ est algébriquement clos, il existe $\lambda \in \sigma(T)$. Soit $U = \ker(T - \lambda Id_V)$. On affirme que U est $T^\dagger$-invariant. En effet, soit $u \in U$ et considérons

$$\begin{aligned} T(T^\dagger(u)) &= T^\dagger(T(u)) \\ &= T^\dagger(\lambda u) \\ &= \lambda T^\dagger(u). \end{aligned}$$

On conclut que $T^\dagger(u) \in U$. A présent, $\mathbb{C}$ étant algébriquement clos, il existe $\gamma \in \sigma(T^\dagger|_U)$. Soit $v \in U$ non nul tel que $T^\dagger(v) = T^\dagger|_U(v) = \gamma v$. Alors on a

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle T(v), v\rangle = \langle v, T^\dagger(v)\rangle = \langle v, \gamma v\rangle = \overline{\gamma}\langle v, v\rangle.$$

Comme $\langle v, v\rangle \neq 0$, on conclut que $\lambda = \overline{\gamma}$, d'où le résultat.

□

Lemme 5.12 *Soit V un espace hermitien et $T \in \mathscr{L}(V)$ un opérateur normal. Soient v_1 et v_2 deux vecteurs propres associés à deux valeurs propres distinctes de T. Si v_1 ou v_2 est vecteur propre de $T^\dagger$ alors $v_1 \perp v_2$.*

Démonstration. Soient $\lambda_1, \lambda_2 \in \mathbb{C}$ tels que $T(v_1) = \lambda_1 v_1$ et $T(v_2) = \lambda_2 v_2$. Sans perte de généralité, supposons v_2 vecteur propre de $T^\dagger$. Par le Lemme 5.11, on a $T^\dagger(v_2) = \overline{\lambda_2} v_2$. On peut donc écrire

$$\lambda_1\langle v_1, v_2\rangle = \langle T(v_1), v_2\rangle = \langle v_1, T^\dagger(v_2)\rangle = \langle v_1, \overline{\lambda_2} v_2\rangle = \lambda_2\langle v_1, v_2\rangle.$$

Comme $\lambda_1 \neq \lambda_2$, on doit avoir $\langle v_1, v_2\rangle = 0$.

□

Théorème 5.13 *Soit V un espace hermitien et $T \in \mathscr{L}(V)$ un opérateur normal. Alors T et $T^\dagger$ admettent une base commune orthonormée de vecteurs propres.*

Démonstration. On procède par induction sur la dimension de V. Le résultat est certainement vérifié lorsque $\dim V = 1$. Supposons $\dim V > 1$. Par le Lemme 5.11, T et $T^\dagger$ possède au moins un vecteur propre commun $w \in V$. De plus, si $T(w) = \lambda w$ alors $T^\dagger(w) = \overline{\lambda} w$. Sans perte de généralité, on peut supposer $\|w\| = 1$. Posons $W =< w >$ et considérons $W^\perp$. Alors par la Proposition 4.23, on a $V = W \oplus W^\perp$ et $\dim W^\perp = \dim V - 1$. On affirme que $W^\perp$ est T et $T^\dagger$-invariant. En effet, si $u \in W^\perp$ alors d'une part, $0 = \langle w, u\rangle = \lambda\langle w, u\rangle = \langle T(w), u\rangle = \langle w, T^\dagger(u)\rangle$ et donc $T^\dagger(u) \in W^\perp$; et d'autre part, $0 = \langle w, u\rangle = \overline{\lambda}\langle w, u\rangle = \langle T^\dagger(w), u\rangle = \langle w, T^{\dagger\dagger}(u)\rangle = \langle w, T(u)\rangle$ et donc $T(u) \in W^\perp$, où dans la dernière égalité on invoque l'exercice à la fin du chapitre 4.

Le sous-espace $W^\perp$ est un espace hermitien muni de la structure de produit scalaire induite par le produit scalaire de V. De plus, $T|_{W^\perp}$ et $T^\dagger|_{W^\perp} \in \mathscr{L}(W^\perp)$. Plus important encore

$$(T|_{W^\perp})^\dagger = T^\dagger|_{W^\perp}.$$

En effet, si $u_1, u_2 \in W^\perp$ alors

$$\langle T|_{W^\perp}(u_1), u_2\rangle = \langle T(u_1), u_2\rangle = \langle u_1, T^\dagger(u_2)\rangle = \langle u_1, T^\dagger|_{W^\perp}(u_2)\rangle.$$

Le résultat s'en suit par l'unicité de l'adjointe. On déduit donc que $T|_{W^\perp}$ est un opérateur normal. Par l'hypothèse d'induction $T|_{W^\perp}$ et $T^\dagger|_{W^\perp}$ admettent une base commune orthonormée de vecteurs propres. Ces vecteurs sont évidemment vecteurs propres de T et $T^\dagger$. En leur adjoignant w, on obtient la base désirée.

□

Théorème 5.14 *Soit V un espace hermitien et $T \in \mathscr{L}(V)$ un opérateur hermitien. Alors T admet une base orthonormée de vecteurs propres et toutes les valeurs propres de T sont réelles.*

Démonstration. La première partie découle du Théorème 5.13 et du fait qu'un opérateur hermitien est normal. Soit $\lambda \in \mathbb{C}$ une valeur propre de T et $v \in V$ un vecteur propre associé. On déduit de

$$\lambda\langle v, v\rangle = \langle T(v), v\rangle = \langle v, T(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle,$$

que $\lambda = \overline{\lambda}$ car $\langle v, v\rangle \neq 0$, c'est-à-dire, $\lambda \in \mathbb{R}$.

□

5.6 Opérateurs unitaires

Soit V un espace préhilbertien complexe.

Définition 5.8 *On dit qu'une transformation linéaire $U \in \mathscr{L}(V)$ est un opérateur unitaire si $\|U(v)\| = \|v\|$ pour tout $v \in V$.*

Remarque : Si V est un espace préhilbertien réel on parle plutôt d'un opérateur orthogonal.

Proposition 5.15 *Soit V un espace hermitien et $U \in \mathscr{L}(V)$. Les énoncés suivants sont équivalents :*

1. *La transformation linéaire U est unitaire.*
2. *On a $\langle U(v), U(u)\rangle = \langle v, u\rangle$ pour tout $v, u \in V$.*
3. *Si $\{v_1, ..., v_n\}$ est une base othonormée de V alors $\{U(v_1), ..., U(v_n)\}$ est aussi une base orthonormée de V.*
4. *Soit $\beta = \{v_1, ..., v_n\}$ une base othonormée de V et $A = [U]_\beta^\beta$. Alors*

$$A \cdot {}^t\overline{A} = \mathbb{1}_n.$$

Démonstration. On démontre (1) implique (2). On a $\|U(u+v)\| = \|u+v\|$ et $\|U(u+iv)\| = \|u+iv\|$ et donc

$$\langle U(u), U(v)\rangle + \langle U(v), U(u)\rangle = \langle u, v\rangle + \langle v, u\rangle$$

et

$$\langle U(u), iU(v)\rangle + \langle iU(v), U(u)\rangle = \langle u, iv\rangle + \langle iv, u\rangle,$$

d'où $-\langle U(u), U(v)\rangle + \langle U(v), U(u)\rangle = -\langle u, v\rangle + \langle v, u\rangle$. En additionnant cette dernière égalité à la première, on obtient $\langle U(v), U(u)\rangle = \langle v, u\rangle$.

On démontre (2) implique (3). Si $\{v_1, ..., v_n\}$ est une base othonormée de V alors le (2) implique que $\{U(v_1), ..., U(v_n)\}$ est un ensemble orthogonal où chaque membre est de norme 1. En particulier, chaque $U(v_i) \neq 0$. Par la Proposition 4.14, $\{U(v_1), ..., U(v_n)\}$ est linéairement indépendant dans un espace de dimension n. C'est donc une base orthonormée de V.

On démontre (3) implique (4). Si $\beta = \{v_1, ..., v_n\}$ est une base othonormée de V et $(a_{ij}) = A = [U]_\beta^\beta$ alors

$$\left.\begin{array}{ll} \text{si } i = j & 1 \\ \text{si } i \neq j & 0 \end{array}\right\} = \langle U(v_i), U(v_j)\rangle = \left\langle \sum_{k=1}^n a_{ki} v_k, \sum_{l=1}^n a_{lj} v_l \right\rangle$$
$$= \sum_{k=1}^n \sum_{l=1}^n a_{ki}\overline{a_{lj}}\langle v_k, v_l\rangle$$
$$= \sum_{k=1}^n a_{ki}\overline{a_{kj}}.$$

Ces relations sont exactement celles qui définissent le produit ${}^tA \cdot \overline{A} = \mathbb{1}_n$. On a donc

$$\mathbb{1}_n = \overline{\mathbb{1}_n} = \overline{{}^tA \cdot \overline{A}} = {}^t\overline{A} \cdot A,$$

c'est-à-dire, $A^{-1} = {}^t\overline{A}$. On conclut, par l'unicité de l'inverse que $A \cdot {}^t\overline{A} = \mathbb{1}_n$.

Réciproquement (4) implique (3) car les relations qui définissent le produit $A \cdot {}^t\overline{A} = \mathbb{1}_n$ impliquent que ${}^tA \cdot \overline{A} = \mathbb{1}_n$, c'est-à-dire, $\{U(v_1), ..., U(v_n)\}$ est une base orthonormée.

On démontre (3) implique (1) à l'aide du Corollaire 4.22. Soit $\{v_1, ..., v_n\}$ une base orthonormée de V et $v \in V$. Dès lors ils existent $a_1, ..., a_n \in \mathbb{C}$ tels que $v = \sum_{k=1}^n a_k v_k$. On a donc

$$\|v\|^2 = \sum_{k=1}^n a_k \overline{a_k}.$$

D'autre part, on a

$$\|U(v)\|^2 = \left\langle \sum_{k=1}^n a_k U(v_k), \sum_{l=1}^n a_l U(v_l) \right\rangle = \sum_{k=1}^n a_k \overline{a_k}$$

car $\{U(v_1), ..., U(v_n)\}$ est une base orthonormée. D'où le résultat.

□

Remarque : Le prochain résultat est faux si $\mathbb{k} = \mathbb{R}$. On doit poser $\mathbb{k} = \mathbb{C}$.

Théorème 5.16 *Soit V un espace hermitien et $U \in \mathscr{L}(V)$ un opérateur unitaire. Alors V admet une base orthonormée composée de vecteurs propres de U. De plus, si $\lambda \in \sigma(U)$ alors $|\lambda| = 1$.*

Démonstration. D'une part, si $\lambda \in \sigma(U)$ et v est un vecteur propre associé à λ alors

$$\langle U(v), U(v)\rangle = \lambda\overline{\lambda}\langle v, v\rangle.$$

Le fait que U soit unitaire implique que $|\lambda|^2 = \lambda\overline{\lambda} = 1$, c'est-à-dire, $|\lambda| = 1$. D'autre part, comme V est hermitien, il existe une base orthonormée β. Si $A = [U]_\beta^\beta$ alors par la Proposition 4.27 on a $[U^\dagger]_\beta^\beta = {}^t\overline{A}$. Par la Proposition 5.15, on a $A \cdot {}^t\overline{A} = \mathbb{1}_n$, c'est-à-dire, $A^{-1} = {}^t\overline{A}$. On conclut que U est un opérateur normal. Le résultat s'en suit par le Théorème 5.14.

□

CHAPITRE 6

DÉTERMINANTS

6.1 Propriétés

Définition 6.1 *Une fonction déterminant est une fonction*

$$\det : \mathbb{M}_{n\times n}(\Bbbk) \to \Bbbk$$

telle que si $a_1, ..., a_n$ sont les lignes d'une matrice A, c'est-à-dire, $a_i \in \Bbbk^n$, alors

1. $\det(a_1, ..., a_{i-1}, a_i + a_j, a_{i+1}, ..., a_n) = \det(A)$ *lorsque* $j \neq i$.
2. $\det(a_1, ..., a_{i-1}, \lambda a_i, a_{i+1}, ..., a_n) = \lambda \det(A)$, *pour tout* $\lambda \in \Bbbk$.
3. $\det(\mathbb{1}_n) = 1$.

Proposition 6.1 *Soit* det *une fonction déterminant. On a toujours que*

1. $\det(A)$ *est multiplié par* -1 *si on interchange deux lignes distinctes de* A.
2. $\det(A) = 0$ *si deux lignes de* A *sont égales.*
3. *Le* $\det(A)$ *demeure inchangé si on remplace une ligne* a_i *par* $a_i + \sum_{j\neq i} \lambda_j a_j$ *pour des* $\lambda_j \in \Bbbk$ *quelconques.*
4. $\det(A) = 0$ *si les lignes de* A *forment un ensemble linéairement dépendant.*
5. *Si on considère la fonction* $\det(A)$ *comme une fonction des lignes de* A *alors celle-ci est* n*-linéaires.*

6.2 Existence et unicité

Théorème 6.2 *Si* D *et* D' *sont deux fonctions déterminants sur* $\mathbb{M}_{n\times n}(\Bbbk)$ *au sense de la Définition 6.1 alors*

$$D = D'.$$

Théorème 6.3 *Il existe une fonction* $\det : \mathbb{M}_{n\times n}(\mathbb{k}) \to \mathbb{k}$ *satisfaisant à la Définition 6.1.*

On construit la fonction déterminant par induction sur n. Si $n = 1$ alors on pose $\det(\alpha) = \alpha$, $\alpha \in \mathbb{k}$. On suppose construite une fonction $\det : \mathbb{M}_{n-1\times n-1}(\mathbb{k}) \to \mathbb{k}$ satisfaisant à la Définition 6.1. Soit $A = (a_{ij})$ avec $1 \leq i, j \leq n$, et considérons les matrices $B^{ij} \in \mathbb{M}_{n-1\times n-1}(\mathbb{k})$ telles que

$$B^{ij} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1\,j-1} & a_{1\,j+1} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2\,j-1} & a_{2\,j+1} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1\,1} & a_{i-1\,2} & \dots & a_{i-1\,j-1} & a_{i-1\,j+1} & \dots & a_{i-1\,n} \\ a_{i+1\,1} & a_{i+1\,2} & \dots & a_{i+1\,j-1} & a_{i+1\,j+1} & \dots & a_{i+1\,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n\,j-1} & a_{n\,j+1} & \dots & a_{nn} \end{bmatrix}.$$

Si on fixe un $j \in \{1, \dots, n\}$ alors on pose

$$\det(A) = (-1)^{1+j} a_{1\,j} \det(B^{1\,j}) + \dots + (-1)^{n+j} a_{n\,j} \det(B^{n\,j}).$$

Remarque : Si cette fonction satisfait bien à la Définition 6.1 alors par l'unicité d'une fonction déterminant on voit que cette formule est donc indépendante par rapport à la colonne choisie pour développer celle-ci.

Exemple : Soit $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ et calculons son polynôme caractéristique. On obtient

$$\begin{aligned} p_A(\lambda) &= \det \begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix} \\ &= (a-\lambda)(d-\lambda) - cb \\ &= ad - a\lambda - d\lambda + \lambda^2 - cb \\ &= \lambda^2 - (a+d)\lambda + (ad - cb) \\ &= \lambda^2 - tr(A)\lambda + \det(A). \end{aligned}$$

Ce résultat se généralise au cas $n \geq 2$. En fait le polynôme caractéristique est toujours de la forme $p_A(\lambda) = \lambda^n - a_{n-1}\lambda^{n-1} + \dots + (-1)^n a_0$, où $a_{n-1} = tr(A)$ et $a_0 = \det(A)$.

6.3 Applications

Proposition 6.4 *Si* $A, B \in \mathbb{M}_{n\times n}(\mathbb{k})$ *alors* $\det(AB) = \det(A) \cdot \det(B)$.

Théorème 6.5 *Une matrice* $A \in \mathbb{M}_{n\times n}(\mathbb{k})$ *est inversible si et seulement si* $\det(A) \neq 0$.

Deuxième partie

Annexes

ANNEXE A

NOMBRES COMPLEXES

Définition A.1 *On pose* $\mathbb{C} \cong \mathbb{R}^2$ *en tant qu'espace vectoriel réel munit des deux opérations suivantes :*

- $(x,y)+(u,v) := (x+u, y+v)$, *et*
- $(x,y)\cdot(u,v) := (xu-yv, xv+yu)$.

Théorème A.1 $(\mathbb{C},+,\cdot)$ *est un corps.*

Démonstration. Il est aisé de démontrer que $\mathbb{C}$ satisfait aux axiomes d'un corps. En particulier, $(0,0)$ est "le" zéro additif de $\mathbb{C}$ et $(1,0)$ est "le" 1 multiplicatif de $\mathbb{C}$. On démontre ici que tout nombre complexe $(x,y) \neq (0,0)$ possède un inverse multiplicatif. Voici le candidat pour $(x,y)^{-1}$, c'est-à-dire,

$$\frac{1}{x^2+y^2}(x,-y).$$

On remarque qu'il est légitime de multiplier $(x,-y)$ par $\frac{1}{x^2+y^2} \in \mathbb{R}$ car $x^2+y^2 \in \mathbb{R}^*$ (pourquoi ?). Alors

$$(x,y)\cdot\frac{1}{x^2+y^2}(x,-y) = \frac{1}{x^2+y^2}(x^2+y^2,0) = (1,0).$$

□

Les nombres réels s'incluent naturellement dans $\mathbb{C}$, c'est-à-dire,

$$\begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{C} \\ x & \mapsto & (x,0) \end{array}$$

et on vérifie facilement que

- $(x,0)+(x',0) = (x+x',0)$, et
- $(x,0)\cdot(x',0) = (xx',0)$.

Historiquement on écrit $1 := (1,0)$ et $i := (0,1)$. Donc un nombre complexe $z \in \mathbb{C}$ s'écrit

$$z = x + iy \quad \text{pour des uniques } x, y \in \mathbb{R}.$$

Soit $z = x + iy \in \mathbb{C}$.

Définition A.2 *La partie réelle de z est* $\mathrm{Re}(z) = x$ *et la partie imaginaire de z est* $\mathrm{Im}(z) = y$.

On remarque maintenant que

$$i^2 = (0,1) \cdot (0,1) = (-1,0) = -1.$$

Donc, dans $\mathbb{C}$, l'équation algébrique $z^2 + 1 = 0$ admet une solution $z = i$. On remarque que $z = -i$ est une autre solution. Malheureusement l'obtention de ces solutions a un coût : les nombres complexes ne peuvent pas être **ordonnés** !

Définition A.3 *Un corps* $\mathbb{k}$ *est ordonné s'il existe un sous-ensemble* $\mathscr{P}$ *de* $\mathbb{k}$*, que l'on appelle les éléments positifs de* $\mathbb{k}$*, tel que*

1. *si* $\alpha, \beta \in \mathscr{P}$ *alors* $\alpha + \beta$ *et* $\alpha\beta$ *appartiennent tous deux à* $\mathscr{P}$ *; et*
2. *pour tout* $\alpha \in \mathbb{k}$*, une et une seule des possibilités suivantes survient*

$$\alpha \in \mathscr{P}, \quad \alpha = 0, \quad ou \quad -\alpha \in \mathscr{P}.$$

Dans ce cas on peut définir une relation d'ordre total sur $\mathbb{k}$ en posant $x < y$ si et seulement si $y - x \in \mathscr{P}$. En effet, par la condition (2), pour tout $x, y \in \mathbb{k}$ une et une seule des trois possibilités survient toujours

$$x < y, \quad x = y \quad \text{ou} \quad y < x.$$

Exemple : Les corps $\mathbb{Q}$ et $\mathbb{R}$ sont tous les deux ordonnés.

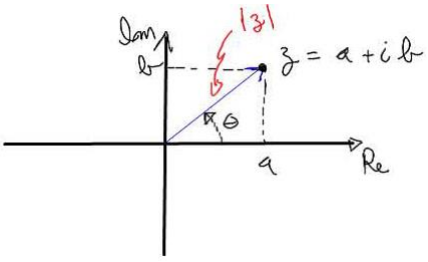
Supposons que $\mathbb{C}$ peut être ordonné, c'est-à-dire, il existe un sous-ensemble $\mathscr{P}$ de $\mathbb{C}$ appelé les "nombres complexes positifs". Alors $i \neq 0$ implique une de deux possibilités : $i \in \mathscr{P}$ ou $-i \in \mathscr{P}$. Dans les deux cas on a

$$i^2 = (-i)^2 = -1 \in \mathscr{P}.$$

Mais à présent on arrive à la contradiction $1 = (-1)^2 \in \mathscr{P}$.

A.1 Représentation polaire

Par définition $\mathbb{C} \cong \mathbb{R}^2$ en tant qu'espace vectoriel réel, c'est-à-dire, qu'il est donc possible de représenter un nombre complexe par un point dans le plan.



Il est donc naturel de définir la longueur d'un nombre complexe $z = a + ib$ à l'aide de la norme euclidienne. Historiquement on parle du *module* du nombre complexe z et on écrit

$$|z| := \sqrt{a^2 + b^2}.$$

Pour tout $z = a + ib \neq 0$, il existe un unique $\theta \in [0, 2\pi)$ tel que

$$\cos(\theta) = \frac{a}{|z|} \quad \text{et} \quad \sin(\theta) = \frac{b}{|z|},$$

on parle alors de la *représentation polaire* de z et on écrit

$$z = |z|\big(\cos(\theta) + i\sin(\theta)\big).$$

Ceci nous permet de définir la fonction argument

$$\begin{array}{rcl} \arg : \mathbb{C}^* & \longrightarrow & [0, 2\pi) \\ z & \mapsto & \theta. \end{array}$$

Une autre fonction fort utile est la conjugaison, c'est-à-dire,

$$\begin{array}{rcl} \mathbb{C} & \longrightarrow & \mathbb{C} \\ z = a + ib & \mapsto & \overline{z} := a - ib \end{array}$$

On peut montrer aisément que

- $\overline{z\overline{z}} = |z|^2$;
- $\overline{z + z'} = \overline{z} + \overline{z'}$;
- $\overline{zz'} = \overline{z} \cdot \overline{z'}$;
- si $z' \neq 0$ alors $\overline{\left(\frac{z}{z'}\right)} = \frac{\overline{z}}{\overline{z'}}$;
- $z = \overline{z}$ si et seulement si $z \in \mathbb{R}$; et $\overline{\overline{z}} = z$.

Remarque : On a

$$\mathrm{Re}(z) = \frac{z + \overline{z}}{2} \quad \text{et} \quad \mathrm{Im}(z) = \frac{z - \overline{z}}{2i}.$$

De plus, si $z \neq 0$ alors $z^{-1} = \dfrac{\overline{z}}{|z|^2}$.

Proposition A.2 *Pour tout $z, z' \in \mathbb{C}$, on a*

- $|zz'| = |z| \cdot |z'|$;
- *si $z' \neq 0$ alors* $\left|\dfrac{z}{z'}\right| = \dfrac{|z|}{|z'|}$;
- $|\mathrm{Re}(z)| \leq |z|$ *et* $|\mathrm{Im}(z)| \leq |z|$;
- $|\overline{z}| = |z|$;
- $|z + z'| \leq |z| + |z'|$; *et* $|z - z'| \geq \Big||z| - |z'|\Big|$.

A.2 Interprétation géométrique de la multiplication

Soit $z_1 = r_1(\cos(\theta_1) + i\sin(\theta_1))$ et $z_2 = r_2(\cos(\theta_2) + i\sin(\theta_2))$ deux nombres complexes non nuls. Alors

$$\begin{aligned} z_1 z_2 &= r_1 r_2(\cos(\theta_1) + i\sin(\theta_1))(\cos(\theta_2) + i\sin(\theta_2)) \\ &= r_1 r_2\Big[(\cos(\theta_1)\cos(\theta_2) - \sin(\theta_1)\sin(\theta_2)) \\ &\quad + i(\cos(\theta_1)\sin(\theta_2) + \sin(\theta_1)\cos(\theta_2))\Big] \\ &= r_1 r_2(\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)). \end{aligned}$$

Donc

Proposition A.3 *Pour tout* $z_1, z_2 \in \mathbb{C}^*$, $\arg(z_1 z_2) \equiv \arg(z_1) + \arg(z_2)\ (mod\ 2\pi)$.

Il est maintenant aisé de calculer les puissances $n^{\text{ième}}$, c'est-à-dire, on obtient la formule de De Moivre

Proposition A.4 *Si* $z = r(\cos(\theta) + i\sin(\theta))$, $r > 0$ *et* $n \in \mathbb{N}$ *alors* $z^n = r^n(\cos(n\theta) + i\sin(n\theta))$.

Démonstration. Le résultat suit d'un raisonnement par induction et des deux identités trigonométriques

- $\cos(\theta \pm \alpha) = \cos(\theta)\cos(\alpha) \mp \sin(\theta)\sin(\alpha)$ et
- $\sin(\theta \pm \alpha) = \sin(\theta)\cos(\alpha) \pm \cos(\theta)\sin(\alpha)$.

□

La formule de De Moivre peut maintenant être utilisée afin de calculer les racines $n^{\text{ième}}$ d'un nombre complexe $w \in \mathbb{C}^*$. Fixons $n \in \mathbb{N}$ et posons $w = r(\cos(\theta) + i\sin(\theta)) \in \mathbb{C}^*$. On veut solutionner l'équation

$$z^n = w.$$

En d'autres mots, on veut trouver tous les $\rho > 0$ et $\phi \in [0, 2\pi)$ tels que (selon De Moivre)

$$\rho^n(\cos(n\phi) + i\sin(n\phi) = r(\cos(\theta) + i\sin(\theta)).$$

Par l'unicité de la représentation polaire, on doit avoir que $\rho = \sqrt[n]{r}$ et $n\phi = \theta + 2\pi k$, où $k \in \mathbb{Z}$.

Conclusion : les racines $n^{\text{ième}}$ de $w = r(\cos(\theta) + i\sin(\theta)) \in \mathbb{C}^*$ sont

$$z = \sqrt[n]{r}\left[\cos\left(\frac{\theta}{n} + \frac{2\pi k}{n}\right) + i\sin\left(\frac{\theta}{n} + \frac{2\pi k}{n}\right)\right],$$

où $k = 0, 1, 2, ..., n-1$ (pourquoi ?).

Il est fort commode d'écrire $e^{i\theta} := \cos(\theta) + i\sin(\theta)$ et d'utiliser aveuglément (en attendant MAT 3521) les règles des exposants, c'est-à-dire,

$$e^{i\theta} \cdot e^{i\alpha} = e^{i(\theta+\alpha)}.$$

On obtient ainsi la représentation exponentielle d'un nombre complexe, c'est-à-dire, $z = |z|e^{i\arg(z)}$.

Laissez-moi terminer ces petits rappels en écrivant notre équation "matheuse"

$$\boxed{e^{i\pi} + 1 = 0}$$

ANNEXE B

RÉCURRENCE

On démontre

Théorème B.1 *Les trois principes suivants sont équivalents, c'est-à-dire,*

1. *Le raisonnement par Récurrence ;*
2. *La Récurrence forte ; et*
3. *Le Principe du bon ordre.*

Démonstration. On suppose donné un énoncé $E(n)$ définit pour chaque entier naturel n. Considérons l'énoncé

$$Q(n) = [\forall k \leq n,\, E(k)]\,.$$

Alors nous avons l'équivalence

$$\forall n \in \mathbb{N},\, E(n) \quad \Leftrightarrow \quad \forall n \in \mathbb{N},\, Q(n).$$

Démontrons (1) implique (2). On suppose que $E(0)$ et l'énoncé pour tout $n \in \mathbb{N}$, $[\forall k \leq n,\, E(k)] \Rightarrow E(n+1)$ soient vrais, c'est-à-dire, $Q(0) = E(0)$ et l'énoncé pour tout $n \in \mathbb{N}$, $Q(n) \Rightarrow E(n+1)$ sont vrais. On remarque que

$$Q(n) \Rightarrow E(n+1) \quad \Leftrightarrow \quad Q(n) \Rightarrow [Q(n) \text{ et } E(n+1)] \quad \Leftrightarrow \quad Q(n) \Rightarrow Q(n+1),$$

car $[Q(n)$ et $E(n+1)] \Leftrightarrow Q(n+1)$. On conclut par le (1) que l'énoncé $[\forall n \in \mathbb{N},\, Q(n)]$ est vrai et donc $[\forall n \in \mathbb{N},\, E(n)]$ est vrai.

Démontrons (2) implique (1). On suppose que $E(0)$ et l'énoncé pour tout $n \in \mathbb{N}$, $E(n) \Rightarrow E(n+1)$ soient vrais. Mais si $n > 0$ alors

$$[Q(n-1) \text{ et } E(n)] \Rightarrow E(n),$$

c'est-à-dire, par transitivité nous avons pour tout $n \in \mathbb{N}$, $Q(n) \Rightarrow E(n+1)$. On conclut par le (2) que $[\forall n \in \mathbb{N}, E(n)]$ est vrai.

Démontrons (3) implique (1). Nous supposons que $E(0)$ et que pour tout $n \in \mathbb{N}$, $E(n) \Rightarrow E(n+1)$ soient vrais. Considérons le sous-ensemble

$$S = \{n \in \mathbb{N} \mid E(n) \text{ est faux}\}$$

et soit m_o l'élément minimal de S. Evidemment $m_o \neq 0$. Un des principes de constructions des nombres naturels dicte que chaque entier naturel non nul est le successeur d'un autre entier naturel. En particulier et par définition de m_o, $E(m_o - 1)$ est vrai. Mais comme $E(m_o - 1) \Rightarrow E(m_o)$ est vrai par hypothèse, ceci force $S = \emptyset$, c'est-à-dire, $\forall n \in \mathbb{N}$, $E(n)$ est vrai.

Démontrons (2) implique (3). Soit $S \subset \mathbb{N}$ un sous-ensemble et supposons que S n'admet pas d'élément minimal. Considérons l'énoncé

$$E(n) = [\forall k \leq n, k \notin S].$$

L'énoncé $E(0)$ est vrai sinon 0 serait l'élément minimal de S. Soit $n \in \mathbb{N}$. Si $E(n)$ vrai alors nécessairement $E(n+1)$ est vrai, sinon $n+1$ serait l'élément minimal de S. On conclut par le (2) que $\forall n \in \mathbb{N}, E(n)$ est vrai, c'est-à-dire, $S = \emptyset$.

□

ANNEXE C

POLYNÔMES

Un polynôme est une suite $p = (a_0, a_1, a_2, ...)$ dans $\mathbb{k}$ éventuellement nulle, c'est-à-dire, il existe $N \in \mathbb{N}$ tel que si $n \geq N$ alors $a_n = 0$. On note l'ensemble des polynômes $\mathscr{P}(\mathbb{k})$. On munit l'ensemble $\mathscr{P}(\mathbb{k})$ de trois opérations :

1. L'addition : $(a_0, a_1, a_2, ...) + (b_0, b_1, b_2, ...) = (a_0 + b_0, a_1 + b_1, a_2 + b_2, ...)$;
2. une action de $\mathbb{k}$: $\alpha \cdot (a_0, a_1, a_2, ...) = (\alpha a_0, \alpha a_1, \alpha a_2, ...)$; et
3. la multiplication : $(a_0, a_1, a_2, ...)(b_0, b_1, b_2, ...) = (c_0, c_1, c_2, ...)$, où $c_k = \sum_{i=0}^{k} a_i b_{k-i}$.

Les deux premières opérations munissent $\mathscr{P}(\mathbb{k})$ d'une structure de $\mathbb{k}$-espace vectoriel. Tandis que la première et la troisième munissent $\mathscr{P}(\mathbb{k})$ d'une structure d'anneau commutatif.

Définition C.1 *Soit $p \in \mathscr{P}(\mathbb{k})$. On définit le degré de p, noté* $\deg(p)$*, par*

$$\deg(p) = \begin{cases} \max\{n \mid a_n \neq 0\} & si\ p \neq 0 \\ -\infty & si\ p = 0 \end{cases},$$

où $p = (a_0, a_1, a_2, ...)$.

Historiquement on note $1 = (1, 0, 0, 0, ...)$, $x = (0, 1, 0, 0, ...)$, $x^2 = (0, 0, 1, 0, ...)$. Donc un polynôme de degré n peut être écrit sans ambiguïté sous la forme

$$p = a_0 + a_1 x + a_2 x^2 + ... + a_n x^n.$$

Proposition C.1 *Soit $p, q \in \mathscr{P}(\mathbb{k})$. Alors*

1. $\deg(p + q) \leq \max\{\deg(p), \deg(q)\}$ *; et*
2. $\deg(pq) = \deg(p) + \deg(q)$.

Démonstration. On laisse la preuve en exercice au lecteur.

□

Théorème C.2 *(Division euclidienne) Soient $p, q \in \mathscr{P}(\Bbbk)$ tel que $q \neq 0$. Alors ils existent des uniques $Q, R \in \mathscr{P}(\Bbbk)$ tels que*

$$p = Qq + R$$

et $\deg(R) < \deg(q)$.

Démonstration. Soit $q = a_m x^m + ... + a_1 x + a_0$, avec $a_0, a_1, ..., a_m \in \Bbbk$ et $a_m \neq 0$. Considérons le sous-ensemble

$$\{\deg(p - Qq) \mid Q \in \mathscr{P}(\Bbbk)\}$$

de $\{-\infty\} \cup \mathbb{N}$. Ce sous-ensemble, étant non vide, possède un plus petit élément n selon le Principe du bon ordre. Il existe donc un polynôme $R \in \mathscr{P}(\Bbbk)$ de degré n tel que $R = p - Qq$, pour un certain $Q \in \mathscr{P}(\Bbbk)$. Le polynôme R s'écrit $R = b_n x^n + ... + b_1 x + b_0$, avec $b_0, b_1, ..., b_n \in \Bbbk$ et $b_n \neq 0$. Si $\deg(R) \geq \deg(q)$ alors $\deg(R - b_n a_m^{-1} x^{n-m} q) < \deg(R)$. De plus,

$$\begin{aligned} p &= Qq + R \\ &= Qq + R - (b_n a_m^{-1} x^{n-m} q) + (b_n a_m^{-1} x^{n-m} q) \\ &= (Q + b_n a_m^{-1} x^{n-m}) q + (R - b_n a_m^{-1} x^{n-m} q) \end{aligned}$$

ce qui contredit la minimalité de n. On conclut que $\deg(R) < \deg(q)$. Finalement, s'ils existent R' et Q' tels que $p = Q'q + R'$ et $\deg(R') < \deg(q)$ alors

$$R - R' = q(Q' - Q),$$

c'est-à-dire, si $Q' - Q \neq 0$ alors $\deg(q) \leq \deg(q(Q' - Q)) = \deg(R - R') < \deg(q)$, ce qui est impossible. On doit conclure que $Q = Q'$ et dès lors $R = R'$.

□

A tout polynôme $p \in \Bbbk$ on peut associé une fonction polynomiale, c'est-à-dire, si $p = a_n x^n + ... + a_1 x + a_0$ alors on a

$$\begin{aligned} p : \Bbbk &\longrightarrow \Bbbk \\ \xi &\mapsto p(\xi) = a_n \xi^n + ... + a_1 \xi + a_0. \end{aligned}$$

Corollaire C.3 *Soit $p \in \mathscr{P}(\Bbbk)$ et $\xi \in \Bbbk$. Alors il existe un unique $Q \in \mathscr{P}(\Bbbk)$ tel que*

$$p = Q \cdot (x - \xi) + p(\xi).$$

Démonstration. Si on applique la division euclidienne aux polynômes p et $x - \xi (\neq 0)$ alors ils existent des uniques polynômes Q et R tels que

$$p = Q \cdot (x - \xi) + R$$

et $\deg(R) < \deg(x - \xi) = 1$. Donc on doit avoir $R \in \Bbbk$. Dès lors, on a

$$p(\xi) = \underbrace{Q(\xi) \cdot (\xi - \xi)}_{=0} + R.$$

□

Définition C.2 *On dit qu'un scalaire $\xi \in \mathbb{k}$ est une racine d'un polynôme $p \in \mathscr{P}(\mathbb{k})$ si $p(\xi) = 0$.*

Corollaire C.4 *Tout polynôme p non nul possède au plus $deg(p)$ racines distinctes.*

Démonstration. Soit $\{\xi_1, \xi_2, \xi_3, ...\}$ l'ensembles des racines distinctes de p. On affirme qu'ils existent des uniques $Q_1, Q_2, Q_3, ... \in \mathscr{P}(\mathbb{k})$ tels que

$$\begin{aligned} p &= Q_1 \cdot (x - \xi_1) \\ &= Q_2 \cdot (x - \xi_2)(x - \xi_1) \\ &= Q_3 \cdot (x - \xi_3)(x - \xi_2)(x - \xi_1) \\ &\vdots \end{aligned}$$

La première égalité découle du Corollaire C.3. Si on évalue la première égalité en ξ_2 alors on obtient

$$0 = p(\xi_2) = Q_1(\xi_2) \cdot \underbrace{(\xi_2 - \xi_1)}_{\neq 0}$$

ce qui force $Q_1(\xi_2) = 0$ (dans un corps $\mathbb{k}$ il n'y a pas de diviseur de zero). Donc il existe un unique polynôme Q_2 tel que $Q_1 = Q_2 \cdot (x - \xi_2)$, d'où la seconde égalité. A la $n^{\text{ième}}$ étape on a $p = Q_n \cdot (x - \xi_n)...(x - \xi_1)$ avec évidemment $Q_n \neq 0$ ce qui implique $n \leq \deg(p)$. Le résultat s'en suit.

□

Nous avons maintenant, comme pour les nombres naturels, la notion de plus grand commun diviseur.

Théorème C.5 *Soient $p_1, ..., p_k \in \mathscr{P}(\mathbb{k})$ des polynômes non nuls. Alors il existe un polynôme $d \neq 0$ admettant les deux propriétés suivantes : (1) d divise tous les p_i ; et (2)si d' est un autre polynôme qui divise tous les p_i alors d' divise d. De plus, ils existent $q_1, ..., q_k \in \mathscr{P}(\mathbb{k})$ tels que*

$$d = q_1 p_1 + ... + q_k p_k.$$

Démonstration. Considérons le sous-ensemble

$$S = \{\deg(q_1 p_1 + ... + q_k p_k) \mid q_1, ..., q_k \in \mathscr{P}(\mathbb{k}) \text{ et } q_1 p_1 + ... + q_k p_k \neq 0\}$$

de $\mathbb{N}$. Alors comme $\deg(p_i) \in S$, $S \neq \emptyset$. Par le Principe du bon ordre, soit $n = \min S$. Il existent donc $d, q_1, ..., q_k \in \mathscr{P}(\mathbb{k})$ tels que

$$d = q_1 p_1 + ... + q_k p_k \quad \text{et} \quad \deg(d) = n.$$

On affirme que d divise tous les p_i. En effet, en appliquant la division euclidienne, ils existent Q_i et R_i des polynômes tels que $p_i = Q_i d + R_i$ et $\deg(R_i) < n$. Mais comme

$$R_i = p_i - Q_i d = -Q_i q_1 p_1 - ... - (Q_i q_i - 1) p_i - ... - Q_i q_k p_k,$$

on a donc $\deg(R_i) \in S$ ce qui contredit la nature de n. Par conséquent, on doit avoir $R_i = 0$ pout tout i. S'il existe un autre polynôme d' qui divise chacun des p_i alors il existent $h_1, ..., h_k \in \mathscr{P}(\mathbb{k})$ tels que $p_i = d' h_i$. On a donc

$$d = q_1 p_1 + ... + q_k p_k = q_1 d' h_1 + ... + q_k d' h_k = (q_1 h_1 + ... + q_k h_k) d',$$

c'est-à-dire, d' divise d.

□

Définition C.3 *On appelle d un plus grand commun diviseur de $p_1, ..., p_k$ et est défini à un multiple scalaire non nul près.*

Corollaire C.6 *Soient $p_1, ..., p_k \in \mathscr{P}(\mathbb{k})$ des polynômes non nuls. Si ces polynômes n'ont aucun facteur commun alors ils existent des polynômes $q_1, ..., q_k$ tels que*

$$1 = q_1 p_1 + ... + q_k p_k.$$

ANNEXE D

L'AXIOME DU CHOIX

On rappelle ici l'énoncé de *l'axiome du choix*.

Axiome du choix : Soit $\mathscr{X}$ une ensemble dont les éléments sont eux-mêmes des ensembles non vides. Alors il existe une fonction

$$\begin{array}{rcl} \omega : \mathscr{X} & \longrightarrow & \displaystyle\bigcup_{A\in\mathscr{X}} A \\ A & \mapsto & \omega(A) \in A. \end{array}$$

On appelle cette fonction ω *une fonction de choix*.

Remarque : Cette fonction n'est en général pas unique.

L'axiome du choix est un des neufs axiomes de la théorie des ensembles de Zermelo-Fraenkel (ZFC). Au début du 20$^{\text{ième}}$ siècle l'axiome du choix fut très controversé. Aujourd'hui il est accepté par la majorité des mathématiciens. Voici des exemples où l'axiome du choix n'est pas nécessaire :

- Lorsque $\mathscr{X}$ ne contient qu'un élément. Dans ce cas l'existence d'une fonction de choix est équivalente à l'existence d'un élément dans un ensemble non vide.
- Lorsque $\mathscr{X}$ est fini. Alors le Principe de récurrence est suffisant pour garantir l'existence d'une telle fonction de choix.
- Si $\mathscr{X} \subset \mathscr{P}(\mathbb{N})$, où $\mathscr{P}(\mathbb{N})$ désigne l'ensemble puissance de $\mathbb{N}$, alors le Principe du bon ordre est suffisant pour construire une telle fonction. En effet nous pouvons poser

$$\begin{array}{rcl} \omega : \mathscr{X} & \longrightarrow & \mathbb{N} \\ M & \mapsto & \omega(M) = \text{l'élément minimal de } M. \end{array}$$

- Si $\mathscr{X}$ est l'ensemble de tous les intervals de longueur fini et strictement positive alors on peut poser

$$\begin{array}{rcl} \omega : \mathscr{X} & \longrightarrow & \mathbb{R} \\ [a,b] & \mapsto & \dfrac{a+b}{2}, \text{ le point milieu de } [a,b]. \end{array}$$

Par contre, il est impossible sans l'axiome du choix de démontrer l'existence d'une telle fonction dans le cas où $\mathscr{X}$ est l'ensemble de tous les sous-ensembles non vides de $\mathbb{R}$.

L'axiome du choix est nécessaire afin de démontrer (en fait équivalent à) l'énoncé suivant :

Tout espace vectoriel (de dimension arbitraire) possède une base.

Nous terminerons cette discussion en énonçant le paradoxe de Banach-Tarski :

Soit $B = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 \leq 1\}$ *la sphère pleine de rayon* 1 *et centrée à l'origine de l'espace euclidien* $\mathbb{R}^3$. *Il existe une partition fini de* B *et une suite fini de rotations et de translations qui nous permettent d'assembler les éléments de cette partition en deux copies identiques de* B.

Le paradoxe se résout lorsque l'on réalise que l'axiome du choix nous permet de démonter l'existence de sous-ensembles de $\mathbb{R}^3$ pour lesquels il est impossible de définir une notion de volume cohérente avec notre intuition.

Pour une plus longue description de l'axiome du choix je vous suggère l'adresse :

https ://en.wikipedia.org/wiki/Axiom_of_choice

ANNEXE E

ASSOCIATIVITÉ

Soit un ensemble non-vide X munit d'une opération binaire $\circ : X \times X \to X$ (comme dans le cas d'un groupe). On définit la notion de parenthésage admissible et montrons que tous les parenthésages d'une longueur donnée sont équivalent si l'opération $\circ$ est associative.

Soit $S = \{\star, \diamond, (,)\}$ un ensemble à quatre éléments et considérons les suites finis d'éléments de S. Par exemple,

1. $\star((\diamond((\star\star$
2. $(\star\diamond\star$
3. $(\star\diamond(\star\diamond\star))$.

Intuitivement seul le cas (3) devrait être déclaré parenthésage admissible (et ce même si le (2) nous semble calculable).

Définition E.1 *Une suite fini d'éléments de S est dite Parenthésage Abstrait Complet (PAC) si elle est obtenus à l'aide du procédé inductif suivant :*

1. *L'élément $\star \in S$ est un PAC.*
2. *Si E_1 et E_2 sont des PAC alors $(E_1 \diamond E_2)$ est un PAC.*

Remarques : Seule la condition (2) introduit les éléments $(,)$ et $\diamond$ dans un PAC. De plus, les parenthèses sont toujours introduites en pair : une ouvrante et une fermante. Finalement, la condition (2) n'introduit pas explicitement de nouveaux éléments $\star$. Dans ce cas les éléments $\star$ du nouveau PAC doivent nécessairement provenir des deux sous-PAC E_1 et E_2.

On voit que comme $\star$ est un PAC par la condition (1) et que $(\star\diamond\star)$ est un PAC par la condition (2) alors $(\star\diamond(\star\diamond\star))$ est un PAC par la condition (2). Le nombre de répétitions de l'élément $\star \in S$ dans un PAC donné nous permet de considérer les familles suivantes :

$$\mathscr{P}_n = \{E \text{ est un PAC} \,|\, \text{il y a } n \text{ répétitions de l'élément } \star \text{ dans } E\}.$$

On dira qu'un PAC est de *longueur n* s'il appartient à $\mathscr{P}_n$. On a donc

- $\mathscr{P}_0 = \emptyset$
- $\mathscr{P}_1 = \{\star\}$
- $\mathscr{P}_2 = \{(\star \diamond \star)\}$
- $\mathscr{P}_3 = \{(\star \diamond (\star \diamond \star)), ((\star \diamond \star) \diamond \star)\}$
- $\mathscr{P}_4 = \{(\star \diamond (\star \diamond (\star \diamond \star))), (\star \diamond ((\star \diamond \star) \diamond \star)), (((\star \diamond \star) \diamond \star) \diamond \star), ((\star \diamond (\star \diamond \star)) \diamond \star), ((\star \diamond \star) \diamond (\star \diamond \star))\}$

La cardinalité de ces ensembles sont les *nombres de Catalan* et peuvent se calculer à l'aide de la formule suivante

$$|\mathscr{P}_n| = \frac{1}{n}\binom{2(n-1)}{n-1}, \quad n \geq 1.$$

Voici les dix premiers : 1, 1, 2, 5, 14, 42, 132, 429, 1430 et 4862. Pour un excellent exposé historique voir http ://fr.wikipedia.org/wiki/Nombre_de_Catalan .

Retournons à un ensemble X munit d'une opération associative $\circ : X \times X \to X$. Pour tout $n \geq 1$ nous avons une application évaluation

$$\begin{array}{rcl} \Phi_n : \mathscr{P}_n \times \overbrace{X \times ... \times X}^{n \text{ fois}} & \longrightarrow & X \\ (E, x_1, ..., x_n) & \mapsto & E(x_1, ..., x_n), \end{array}$$

où l'élément $E(x_1, ..., x_n)$ de X est calculé de façon inductive, c'est-à-dire,

1. si $n = 1$ alors $\star(x) = x$; et
2. si $n > 1$ alors ils existent $0 < k < n$, $E_1 \in \mathscr{P}_k$ et $E_2 \in \mathscr{P}_{n-k}$ tels que $E = (E_1 \diamond E_2)$ et on pose

$$E(x_1, ..., x_n) = E_1(x_1, ..., x_k) \circ E_2(x_{k+1}, ..., x_n).$$

Proposition E.1 *Les applications Φ_n sont bien définies pour tout $n \geq 1$.*

Démonstration. Lorsque $n = 1$ la formule donnée est valable. Lorsque $n > 1$, comme un PAC est une suite fini, il doit y avoir un élément $\diamond$ dans ce PAC qui fut le dernier inséré de ce type. La condition (2) de la construction des PAC étant la seule à pouvoir insérer de tel élément il doit donc exister deux PAC E_1 et E_2 tels que $E = (E_1 \diamond E_2)$. De plus, le nombre d'éléments $\star$ dans E doit être la somme de ceux dans E_1 et de ceux dans E_2. On déduit que $E_1 \in \mathscr{P}_k$ et $E_2 \in \mathscr{P}_{n-k}$ pour un certain $0 < k < n$. A présent on procède à l'aide d'un raisonnement par récurrence forte. On suppose Φ_k bien défini pour tout $k < n$. Alors la formule donnée (2) démontre que Φ_n est bien défini, c'est-à-dire, $E_1(x_1, ..., x_k) \in X$ et $E_2(x_{k+1}, ..., x_n) \in X$ étant tous deux bien définis par l'hypothèse d'induction, on a donc $E_1(x_1, ..., x_k) \circ E_2(x_{k+1}, ..., x_n) \in X$. □

Explicitement on calcul $E(x_1, ..., x_n)$ en substituant le $k^{\text{ième}}$ élément $\star$ de E par x_k, pour tout $1 \leq k \leq n$; en substituant chaque élément $\diamond$ de E par l'opération $\circ$ de X ; et en effectuant les opérations $\circ$ dans l'ordre défini par les priorités données par les parenthèses (comme il se doit !).

Exemple : Soit $E = (((\star \diamond \star) \diamond \star) \diamond \star) \in \mathscr{P}_4$. Alors

$$\Phi_4(E, x_1, x_2, x_3, x_4) = ((x_1 \circ x_2) \circ x_3) \circ x_4.$$

Finalement nous obtenons le résultat recherché, c'est-à-dire,

Théorème E.2 *Soit X un ensemble munit d'une opération associative $\circ : X \times X \to X$. Alors, pour tout $E \in \mathscr{P}_n$, nous avons*

$$\Phi_n(E, x_1, x_2, ..., x_n) = ((...(x_1 \circ x_2) \circ ...) \circ x_{n-1}) \circ x_n, \quad n \geq 1.$$

Démonstration. On procède à nouveau par un raisonnement par récurrence forte. Si $n = 1$ alors comme il n'y a qu'un élément dans $\mathscr{P}_1$ la formule $\Phi_1(\star, x) = x$ est vérifiée. Supposons la formule vérifiée pour tout $k < n$. Soit $E \in \mathscr{P}_n$ et comme précédemment cité ils existent $0 < k < n$, $E_1 \in \mathscr{P}_k$ et $E_2 \in \mathscr{P}_{n-k}$ tels que $E = (E_1 \diamond E_2)$. On a donc

$$\begin{aligned}
\Phi_n(E, x_1, x_2, ..., x_n) &= E_1(x_1, ..., x_k) \circ E_2(x_{k+1}, ..., x_n) \\
&= E_1(x_1, ..., x_k) \circ \underbrace{[((...(x_{k+1} \circ x_{k+2}) \circ ...) \circ x_{n-1}) \circ x_n]}_{\text{hypothèse d'induction}} \\
&= \underbrace{[E_1(x_1, ..., x_k) \circ ((...(x_{k+1} \circ x_{k+2}) \circ ...) \circ x_{n-1})]}_{\text{associativité}} \circ x_n \\
&= \Phi_{n-1}((E_1 \diamond \varepsilon), x_1, x_2, ..., x_{n-1}) \circ x_n,
\end{aligned}$$

où $\varepsilon = (((...(\star \diamond \star) \diamond ...) \diamond \star) \diamond \star) \in \mathscr{P}_{n-k-1}$. Finalement, à nouveau par l'hypothèse d'induction $\Phi_{n-1}((E_1 \diamond \varepsilon), x_1, x_2, ..., x_{n-1}) = ((...(x_1 \circ x_2) \circ ...) \circ x_{n-2}) \circ x_{n-1}$. On déduit donc

$$\Phi_n(E, x_1, x_2, ..., x_n) = (((...(x_1 \circ x_2) \circ ...) \circ x_{n-2}) \circ x_{n-1}) \circ x_n.$$

□

BIBLIOGRAPHIE

[1] Sterling K. Berberian, “Linear Algebra”, Dover Publications, (2014).

[2] Charles W. Curtis, “Linear Algebra An Introductory Approach”, Springer-Verlag, (1986).

[3] Serge Lang, “Algèbre linéaire 1”, InterEditions, Paris, (1976).