Cours d’algèbre pour les candidats à l’échange HUB-Ensae

Sébastien Raspiller

12 juin 2002

Table des matières

Chapitre 1

Généralités

On rappelle dans ce chapitre les notions élémentaires d'ensembles et d'applications ainsi que les principales propriétés qu'une loi de composition est susceptible de posséder. Il s'agit d'un chapitre regroupant bon nombre de notions déjà connues depuis longtemps par le lecteur. Pour cette raison, les démonstrations les plus faciles seront omises et les plus difficiles admises. Enfin, on a ajouté un appendice d'arithmétique qui se concentre sur la notion de dénombrabilité. A ce propos, le lecteur est invité à revoir de lui-même les principales définitions et propriétés de l'analyse combinatoire.

1.1 Ensembles

Les notions d'**ensemble**, d'**appartenance** et d'**égalité** ne se définissent pas, on admettra donc que $a \in E$ (qui se lit a appartient à E, ou a est élément de E) est une notation comprise de tous, ainsi que sa négation $a \notin E$. De même, $A = B$ signifie que les objets A et B sont identiques et donc, s'il s'agit d'ensembles, qu'ils sont formés des mêmes éléments.

DEFINITION 1.1 *Soit E un ensemble, on dit qu'un ensemble X est une **partie** de E (ou un **sous-ensemble** de E) si tout élément de X est élément de E. On dit aussi que X est **inclus dans** E et on note :*

$$(X \subset E) \Leftrightarrow (E \supset X) \Leftrightarrow (\forall x, x \in X \Rightarrow x \in E).$$

Si $X \subset E$ et $X \neq E$ on peut préciser $X \subsetneq E$.

On admettra que toutes les parties de E forment un nouvel ensemble noté $\mathcal{P}$(E), i.e. :

$$X \subset E \Leftrightarrow X \in \mathcal{P}(E).$$

On admettra de même l'existence d'un ensemble ne contenant aucun élément, noté $\oslash$ (**ensemble vide**). On a donc pour tout ensemble E :

$$\forall x, x \in \oslash \Rightarrow x \in E,$$

i.e. $\oslash \subset E$. On dit que X est une **partie propre** de E si $X \subset E$, $X \neq \oslash$, $X \neq E$.

Une partie X d'un ensemble E peut se déterminer en général de deux façons :

1. en **extension** : $X = \{2, 3, 5, 7\}$
2. en **compréhension** : $X = \{n \in N$ / $n \leq 10$ et n premier $\}$.

DEFINITION 1.2 *Si X et Y sont deux parties de E, on définit :*

$$X \cup Y = \{x \in E / x \in X \text{ ou } x \in Y\} \text{ et } X \cap Y = \{x \in E / x \in X \text{ et } x \in Y\}.$$

*$X \cup Y$ (respectivement - resp. - $X \cap Y$) s'appelle **réunion** (resp. **intersection**) des parties X et Y.*

*Si $X \cap Y = \oslash$, on dit que X et Y sont **disjointes**.*

Ces notions se généralisent au cas d'un ensemble quelconque de parties. Si P est une partie de $\mathcal{P}$(E), i.e. un ensemble de parties de E, on note :

$$\cap_{X \in P} X = \{x \in E / \forall X \in P, x \in X\} \text{ et } \cup_{X \in P} X = \{x \in E / \exists X \in P, x \in X\}.$$

DEFINITION 1.3 *Si $X \subset E$, on définit $C_E(X) = \{x \in E$ / $x \notin X\}$; on le note aussi $\bar{X}^E$ ou $\bar{X}$. C_E(X) s'appelle le **complémentaire dans** E **de la partie** X, la dernière notation sera utilisée lorsque le référentiel E sera évident.*

DEFINITION 1.4 *Si $X \subset E$ et $Y \subset E$, on définit :*

$$X \backslash Y = \{x \in E / x \in X \text{ et } x \notin Y\}.$$

*$X \backslash Y$ s'appelle **différence** de X et de Y.*

DEFINITION 1.5 *Soient E un ensemble et P une partie de $\mathcal{P}$(E), on dit que :*

1. *P est un **recouvrement** de E si :*

$$\forall x \in E, \exists X \in P, x \in X$$

2. *P est une **partition** de E si :*
   - *(a) P est un recouvrement de E*
   - *(b) $\oslash \notin P$*
   - *(c) $\forall X \in P, \forall Y \in P, X \neq Y \Rightarrow X \cap Y = \oslash$.*

Une partition est donc un recouvrement ne contenant pas la partie vide, et dont les éléments sont deux à deux disjoints.

DEFINITION 1.6 *Soient E et F deux ensembles, on appelle* ***produit cartésien*** *de E et F et on note E×F :*

$$E \times F = \{(x, y)/x \in E \text{ et } y \in F\}.$$

La notion de ***couple*** *(x,y) ne se définit pas, mais vérifie :*

$$(x, y) = (x', y') \Leftrightarrow x = x' \text{ et } y = y'.$$

Plus généralement, on définit le produit cartésien de k ensembles E_1, ..., E_k par :

$$E_1 \times ... \times E_k = \{(x_1, ..., x_k)/\forall i \in [1, k], x_i \in E_i\}.$$

L'égalité de deux k-**uplets** se définit de même par l'égalité de toutes les composantes de même rang. Si $E_1 = E_2 = ... = E_k = E$, $E \times E \times ... \times E$ se note aussi E^k.

Même si $E_1 = E_2$, il ne faut pas confondre (x, y) et (y, x) (sauf si $y = x$) ainsi que (x, y) et$\{$x, y$\}$ (surtout si $y = x$).

PROPOSITION 1.1 *Soient E un ensemble, A, B, C des parties quelconques de E, on a :*

1. $C_E(\bar{A}) = A$
2. $A \cup A = A$ *et* $A \cap A = A$
3. $A \cup B = B \cup A$ *et* $A \cap B = B \cap A$
4. $A \cup (B \cup C) = (A \cup B) \cup C$ *et* $A \cap (B \cap C) = (A \cap B) \cap C$
5. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ *et* $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
6. $C_E(A \cup B) = C_E(A) \cap C_E(B)$ *et* $C_E(A \cap B) = C_E(A) \cup C_E(B)$
7. $A \cup B = A \Leftrightarrow B \subset A$ *et* $A \cap B = A \Leftrightarrow A \subset B$.

Démonstration : admise (par analogie avec les symboles logiques, non présentés ici)

1.2 Applications

DEFINITION 1.7 *Soient E et F deux ensembles, on dit que f est une* ***fonction*** *de E vers F si tout élément x de E est en relation par f avec au plus un élément y de F. Cet élément, lorsqu'il existe, est noté f(x).*

L'ensemble des x de E pour lesquels f(x) existe s'appelle ***ensemble de définition*** *de la fonction f et se note Def(f).*

Si f est une fonction de E vers F et si $Def(f) = E$, f s'appelle une ***application*** *de E vers F et on note :*

$$f : \begin{array}{l} E \to F \\ x \mapsto f(x) \end{array} .$$

Une fonction de E vers F définit une application sur le domaine de définition de cette fonction, à valeurs dans F. Nous ne considérerons donc par la suite que des applications.

Soit A une partie d'un ensemble E, $i_A : x \in A \mapsto x \in E$ est une application de A dans E. On l'appelle injection canonique de A dans E. Si $A = E$, i_E se note généralement Id_E et s'appelle application identique ou **identité** de E.

Soit $A \subset E$, $\varphi_A : E \to \{0, 1\}$ définie par :

$$\begin{array}{rcl} \forall x & \in & A, \varphi_A(x) = 1 \\ \forall x & \in & E \backslash A, \varphi_A(x) = 0 \end{array}$$

est une application appelée caractéristique de A.

L'égalité de deux applications est l'égalité de leurs ensembles de départ, de leurs ensembles d'arrivée et de leur relations. Il est par conséquent fondamental de ne pas confondre les quatre applications suivantes :

$$g_1 : \begin{array}{l} R \to R \\ x \mapsto x^2 \end{array} , g_2 : \begin{array}{l} R^+ \to R \\ x \mapsto x^2 \end{array} , g_3 : \begin{array}{l} R \to R^+ \\ x \mapsto x^2 \end{array} , g_3 : \begin{array}{l} R^+ \to R^+ \\ x \mapsto x^2 \end{array} .$$

On appelle **suite** d'éléments de E toute application d'une partie de N dans E. Si cette partie de N est finie, on dit que la suite est finie.

DEFINITION 1.8 *Soient f une application de E vers F et A une partie de E, l'application notée $f_{|A}$ définie par :*

$$f_{|A} : \begin{array}{l} A \to F \\ x \mapsto f(x) \end{array}$$

s'appelle la ***restriction*** *de f à A.*

Réciproquement, soient g une application de A vers F et f une application de E vers F, si $f_{|A} = g$ on dit que f est un ***prolongement*** *de g à E.*

f est donc, bien entendu, un prolongement de sa restriction à A. Mais, si on se donne une application de A vers F, il existe plusieurs prolongements possibles en une application de E vers F (dès que F a plus d'un élément et A différent de E). On verra ultérieurement des exemples permettant d'obtenir, sous certaines conditions, l'unicité du prolongement.

DEFINITION 1.9 *Soient E, F et G des ensembles, on considère une application f de E vers F et une application g de F vers G. On appelle* ***composée*** *de g et f et on note g∘f l'application de E vers G définie par :*

$$\forall x \in E, (g \circ f)(x) = g(f(x)).$$

Bien noter que l'écriture g∘f signifie que l'on effectue d'abord l'opération $x \mapsto f(x)$ puis l'opération $f(x) \mapsto g(f(x))$.

Si $f : E \to F$, alors $f \circ Id_E = f$ et $Id_F \circ f = f$, mais si $E \neq F$ il ne s'agit pas dans les deux cas de la même application identité.

PROPOSITION 1.2 *Soient $f : E \to F$, $g : F \to G$ et $h : G \to H$, alors :*

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

et l'on note simplement h∘g∘f cette application de E vers H.

Démonstration : triviale

DEFINITION 1.10 *Soit $f : E \to F$, on dit que f est :*

1. ***injective*** *(ou est une* ***injection****) si l'une des propriétés équivalentes suivantes est vérifiée :*
   - *(a)* $\forall x, x' \in E, f(x) = f(x') \Rightarrow x = x'$
   - *(b)* $\forall x, x' \in E, x \neq x' \Rightarrow f(x) \neq f(x')$
   - *(c) pour tout y de F, l'équation en x, $y = f(x)$, a au plus une solution dans E.*
2. ***surjective*** *(ou est une* ***surjection****) si elle vérifie la propriété suivante :*

$$\forall y \in F, \exists x \in E, y = f(x).$$

3. ***bijective*** *(ou est une* ***bijection****) si elle est à la fois injective et surjective, i.e. si elle vérifie :*

$$\forall y \in F, \exists! x \in E, y = f(x).$$

Exercices :

1. Soit $f : R \setminus \{1\} \to R$ définie par $x \mapsto \frac{3x-1}{x-1}$. Est-elle injective ? bijective ?
2. Soit $f : R \to R$ définie par $x \mapsto x^3$. Montrer que f est bijective.
3. Soient $f : E \to F$ et $g : F \to G$. Montrer que si f et g sont injectives (resp. surjectives, bijectives) il en est de même de g∘f. Montrer que si g∘f est injective (resp. surjective) alors f est injective (resp. g est surjective).

DEFINITION 1.11 *Soit $f : E \to F$, on dit que f est **inversible** s'il existe une application $g : F \to E$ telle que $f \circ g = Id_F$ et $g \circ f = Id_E$.*

On peut définir l'inversibilité à droite (resp. à gauche) à l'aide de la première (resp. la seconde) égalité.

THEOREM 1.1 *Soit $f : E \to F$, f est inversible si et seulement si f est bijective, l'application g de la définition précédente est alors unique, on l'appelle application **réciproque** de f. On la note f^{-1}, c'est une bijection de F sur E. Elle est définie par :*

$$\forall (x, y) \in E \times F, x = f^{-1}(y) \Leftrightarrow y = f(x).$$

Démonstration : si f est inversible, Id_F et Id_E étant bijectives, l'exercice 1.1.1 montre que f et g sont bijectives. De plus, g est unique car si $g \circ f = Id_E$ et $f \circ g' = Id_F$ alors :

$$g = g \circ Id_F = g \circ (f \circ g') = (g \circ f) \circ g' = Id_E \circ g' = g'.$$

Réciproquement, si f est bijective on a :

$$\forall y \in F, \exists! x \in E, y = f(x).$$

Cette condition permet de définir une application g de F dans E en associant à tout y de F l'unique x de E tel que $y = f(x)$. Ainsi, si $y = f(x)$, on a $x = g(y)$. Par conséquent :

$$\forall y \in F, (f \circ g)(y) = f(x) = y \text{ et } \forall x \in E, (g \circ f)(x) = g(y) = x \boxdot$$

PROPOSITION 1.3 *Soient $f : E \to F$ et $g : F \to G$, si f et g sont inversibles, alors g∘f est inversible et l'on a :*

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

Démonstration : il suffit en effet de calculer $f^{-1} \circ g^{-1} \circ g \circ f$ et $g \circ f \circ f^{-1} \circ g^{-1}$.

DEFINITION 1.12 *Soit E un ensemble, I un ensemble appelé **ensemble d'indices**. On appelle **famille** d'éléments de E **indéxée par I** toute application de I dans E. Si $x : i \mapsto x(i)$ est une telle famille, on note x_i l'image de i par x et $(x_i)_{i \in I}$ cette famille.*

Attention, une famille n'est pas nécessairement injective. A deux indices différents peuvent correspondre le même élément de E.

Si $(A_i)_{i \in I}$ est une famille de parties d'un ensemble E, on convient de noter :

$$\begin{aligned} \cup_{i \in I} A_i &= \{x \in E / \exists i \in I, x \in A_i\} \\ \cap_{i \in I} A_i &= \{x \in E / \forall i \in I, x \in A_i\}. \end{aligned}$$

DEFINITION 1.13 *Soient $f : E \to F$, $A \subset E$, $B \subset F$;*

1. *on appelle **image directe** de A par f et l'on note f(A) la partie de F définie par :*
$$f(A) = \{y \in F / \exists x \in A, y = f(x)\}.$$
2. *on appelle **image réciproque** de B par f et l'on note f^{-1}(B) la partie de E définie par :*
$$f^{-1}(B) = \{x \in E / f(x) \in B\}.$$

Ces notations sont quelque peu abusives, en particulier f^{-1}(B) ne préjuge pas de l'existence de l'application réciproque de f. Evidemment, si f est inversible on a $f^{-1}(\{y\}) = \{f^{-1}(y)\}$, ce qui en quelque sorte justifie la notation. Mais en général $f^{-1}(\{y\})$ peut être vide ou contenir plus d'un élément.

PROPOSITION 1.4 *Soient $f : E \to F$, A et A' des parties de E, B et B' des parties de F ; on a :*

1. $A \subset A' \Rightarrow f(A) \subset f(A')$
2. $f(A \cup A') = f(A) \cup f(A')$
3. $f(A \cap A') \subset f(A) \cap f(A')$
4. $B \subset B' \Rightarrow f^{-1}(B) \subset f^{-1}(B')$
5. $f^{-1}(B \cap B') = f^{-1}(B) \cap f^{-1}(B')$
6. $f^{-1}(B \cup B') = f^{-1}(B) \cup f^{-1}(B')$
7. $A \subset f^{-1}(B) \Leftrightarrow f(A) \subset B$
8. $A \subset f^{-1}(f(A))$
9. $f(f^{-1}(B)) \subset B$.

Exercices :

Démontrer les neuf propriétés de la proposition précédente.

1.3 Lois de composition

DEFINITION 1.14 *Soient E, F, G trois ensembles, on appelle **loi de composition** définie sur E×F à valeurs dans G toute application de E×F vers G. Si $z = f((x, y))$, on convient d'écrire $z = x\Upsilon y$ (ou $x * y$, $x + y$, xy, ...). x et y s'appellent les **termes** et z le **résultat** de l'**opération** Υ.*

*Si $F = G$, la loi est dite **externe** sur F de **domaine d'opérateurs** E.*

*Si $E = F = G$, la loi est dite **interne** sur E et on note en abrégé (E,Υ).*

Exemples :

1. Dans R (ou N, Z, Q) les lois usuelles + et . sont des lois internes.
2. Dans $\mathcal{P}$(E) les lois $\cup$ et $\cap$ sont des lois internes.
3. Dans l'ensemble des applications de E vers E, la composition $\circ$ est une loi interne.

DEFINITION 1.15 *Soit (E,Υ) un ensemble muni d'une loi de composition interne ; on dit que :*

1. Υ *est* ***associative*** *(dans E) si :*

$$\forall x, y, z \in E, (x \Upsilon y) \Upsilon z = x \Upsilon (y \Upsilon z)$$

2. *x et y sont* ***permutables*** *(pour Υ) si :*

$$x \Upsilon y = y \Upsilon x$$

x est ***central*** *si :*

$$\forall y \in E, x \Upsilon y = y \Upsilon x$$

3. Υ *est* ***commutative*** *(dans E) si :*

$$\forall x, y \in E, x \Upsilon y = y \Upsilon x$$

4. *e est* ***élément neutre*** *si :*

$$\forall x \in E, e \Upsilon x = x \Upsilon e = x$$

5. *a est* ***régulier*** *si :*

$$\forall x, y \in E, a \Upsilon x = a \Upsilon y \Rightarrow x = y \text{ et } x \Upsilon a = y \Upsilon a \Rightarrow x = y$$

6. *i est* ***idempotent*** *si :*

$$i \Upsilon i = i$$

7. *si (E,Υ) admet un élément neutre e, on dit que x' est* ***symétrique*** *de x si :*

$$x \Upsilon x' = x' \Upsilon x = e$$

8. $A \subset E$ *est* ***stable*** *pour Υ si :*

$$\forall x, y \in A, x \Upsilon y \in A$$

(on peut donc définir une loi induite sur A que l'on note généralement par le même symbole).

THEOREM 1.2 *(règles de parenthésage)*

Si Υ est une loi de composition interne associative sur E, alors dans toute succession d'opérations on peut regrouper tels termes consécutifs que l'on veut, et les remplacer par leur résultat, sans changer le résultat final. Si de plus Υ est commutative, on peut effectuer les regroupements des termes que l'on veut.

Démonstration : à la fois facile et fastidieuse

Exercices :

Démontrer les résultats suivants (qui doivent être connus) :

1. l'élément neutre, s'il existe, est unique
2. si la loi est associative et possède un élément neutre, le symétrique d'un élément, lorsqu'il existe, est unique ; ce résultat est en général faux si la loi n'est pas associative
3. si la loi est associative, tout élément symétrisable (i.e. qui admet un symétrique) est régulier
4. si la loi est associative et si x et y sont symétrisables, alors xΥy est symétrisable : quel est son symétrique ?
5. toute intersection de parties stables est stable ; ce résultat est en général faux pour la réunion.

DEFINITION 1.16 *Soit E muni de deux lois de composition internes Υ et *, on dit que * est* ***distributive*** *sur Υ si :*

$\forall x, y, z \in E, x*(y\Upsilon z) = (x*y)\Upsilon(x*z)$ *(à droite) et* $(y\Upsilon z)*x = (y*x)\Upsilon(z*x)$ *(à gauche).*

Par exemple, dans (R,+,.), la multiplication est distributive sur l'addition et dans ($\mathcal{P}$(E),$\cup$,$\cap$), $\cup$ est distributive sur $\cap$ et $\cap$ l'est également sur $\cup$.

DEFINITION 1.17 *Soient (E,Υ), (F,*), f une application de E vers F ; on dit que f est un* ***morphisme*** *de (E,Υ) dans (F,*) si :*

$$\forall x, y \in E, f(x\Upsilon y) = f(x) * f(y).$$

Si f est bijective, on dit que f est un ***isomorphisme****.*

Si $(E, \Upsilon) = (F, *)$, on dit que f est un ***endomorphisme****. Attention, si $E = F$ mais $\Upsilon \neq *$, on ne parlera pas d'endomorphisme.*

Un endomorphisme bijectif s'appelle un ***automorphisme****.*

THEOREM 1.3 *Si f est un isomorphisme de (E,Υ) sur (F,*), alors f^{-1} est un isomorphisme de (F,*) sur (E,Υ). On l'appelle l'****isomorphisme réciproque*** *de f.*

Démonstration : évident, en revenant à la définition de f^{-1}

Soient (E,Υ) et f une bijection de E sur un ensemble F, on peut définir une loi * sur F en posant :

$$\forall z, t \in F, z * t = f(f^{-1}(z) \Upsilon f^{-1}(t)).$$

f devient ainsi un isomorphisme de (E,Υ) sur (F,*). On dit alors qu'on a réalisé un **transport de structure**.

Exemple : l'application $\log : R_+^* \to R$ est un isomorphisme de (R_+^*,.) sur (R,+), l'isomorphisme réciproque se notant $\exp$.

Par tradition, on n'emploie le symbole + que dans le cas d'une loi commutative et associative. L'élément neutre et le symétrique éventuels se notent alors 0 et -x.

La même tradition impose de n'employer le symbole . (ou rien) que dans le cas d'une loi associative. L'élément neutre et le symétrique éventuels se notent alors 1 et x^{-1}.

Au lieu d'écrire $x_1 + ... + x_n$ ou $x_1 . ... . x_n$ on écrira souvent $\sum_{i=1}^n x_i$ et $\Pi_{i=1}^n x_i$. Soit I un ensemble fini d'indices, l'écriture $\sum_{i \in I} x_i$ ne pose pas de problèmes ; par contre l'écriture $\Pi_{i \in I} x_i$ impose, si la multiplication n'est pas commutative, que l'on se soit donné un ordre total sur I. Par convention, la multiplication se fera alors en écrivant de gauche à droite dans l'ordre des indices croissants.

1.4 Arithmétique

1.4.1 Rappels de définitions et propriétésélémentaires

DEFINITION 1.18 *Soient a, b deux entiers quelconques, on dit que a **divise** b et on écrit a / b s'il existe un entier k tel que* $b = ka$.

PROPOSITION 1.5 $\forall (a, b) \in N \times N^*$, $\exists !(q, r) \in N^2$, $a = bq + r$ *et* $0 \leq r < b$.

Démonstration : admise

La pratique de cette division, appelée **division euclidienne**, est supposée connue de tous.

DEFINITION 1.19 *On dit que l'entier p est un **nombre premier** (et on note P l'ensemble des nombres premiers) s'il vérifie l'une des conditions équivalentes suivantes :*

1. $p \in N$, $p > 1$ *et* $\forall a \in N$, a / $p \Rightarrow a = 1$ *ou* $a = p$
2. $p \in N$, $p > 1$ *et* $\forall a, b \in N$, p / $ab \Rightarrow p$ / a *ou* p / b.

PROPOSITION 1.6 *1. Tout entier supérieur à 1 a un plus petit diviseur supérieur à 1 et ce diviseur est premier.*

2. *Si l'entier n n'est divisible par aucun nombre premier p tel que $p^2 \leq n$, alors n est premier.*
3. *P est infini.*

Démonstration : évident pour les deux premières propriétés, admis pour la troisième

THEOREM 1.4 *Tout entier n supérieur à 1 admet une factorisation unique en facteurs premiers, à l'ordre des facteurs près, i.e. :*

$$\exists! m \in N^*, \exists! (p_1, ..., p_m) \in P^m, p_1 \leq p_2 \leq ... \leq p_m, n = p_1 p_2 ... p_m.$$

Démonstration : admis

On appelle exposant de $p \in P$ dans n le nombre d'indices i tels que $p = p_i$. On le note v_p(n) : v_p(n) est donc nul sauf pour un nombre fini de nombres premiers, ce qui permet d'écrire, avec la convention habituelle $p^0 = 1$, $n = \Pi_{p \in P} p^{v_p(n)}$, de sorte que ce produit est en fait un produit fini.

DEFINITION 1.20 *Soit (a_1, ..., a_n) un n-uplet d'éléments de N^*,*

1. *il existe un unique entier non nul d tel que :*

$$\forall i \in [1, n], d/a_i \text{ et } \forall \delta \in N, (\forall i \in [1, n], \delta/a_i) \Rightarrow \delta/d;$$

   *d s'appelle le **plus grand diviseur commun** à a_1, ..., a_n et se note $d = PGCD(a_1, ..., a_n)$*
2. *il existe un unique entier non nul m tel que :*

$$\forall i \in [1, n], a_i/m \text{ et } \forall M \in N, (\forall i \in [1, n], a_i/M) \Rightarrow m/M;$$

   *m s'appelle le **plus petit multiple commun** à a_1, ..., a_n et se note $m = PPCM(a_1, ..., a_n)$.*

Il est d'ailleurs facile de voir que si on pose, pour tout nombre premier p, $m(p) = \min\{v_p(a_i), i \in [1, n]\}$ et $M(p) = \max\{v_p(a_i), i \in [1, n]\}$, alors on a :

$$d = \Pi_{p \in P} p^{m(p)} \text{ et } m = \Pi_{p \in P} p^{M(p)}.$$

Pour rechercher le PGCD de deux nombres a et b ($a > b$ par exemple), il suffit donc d'effectuer la factorisation en nombres premiers de a et b et d'appliquer le résultat précédent. Dans la pratique, il est plus facile d'appliquer la méthode suivante, dite algorithme d'Euclide.

Soit $a = bq + r$ le résultat de la division euclidienne de a par b, alors (δ / a et δ / b) $\Leftrightarrow$ (δ / b et δ / $a - bq$). Par conséquent, $PGCD(a, b) = PGCD(b, r)$. Or $b < a$ et $r < b$: si $r = 0$ alors a est un multiple de b et $PGCD(a, b) = b$, si $r \neq 0$ on itère le procédé en remplaçant le couple (a,b) par le couple (b,r). Le PGCD de a et b est donc le dernier reste précédant le reste nul. Il suffit d'être patient puisque les restes sont strictement décroissants.

1.4.2 Dénombrabilité

DEFINITION 1.21 *Soit E un ensemble, on dit que E est **strictement dénombrable** s'il existe une bijection de E sur N. On dit que E est **dénombrable** s'il est fini ou strictement dénombrable.*

En d'autres termes, un ensemble est dénombrable si et seulement si on peut numéroter ses éléments.

Exemples :

1. N^* est strictement dénombrable. En effet, $s : N \to N^*$ définie par $s(n) = n + 1$ est bijective, par définition même de N.
2. Toute partie de N est finie ou strictement dénombrable. En effet, si une partie A de N n'est pas finie, l'application qui à $n \in N$ associe le n-ème élément de A, pour l'ordre usuel, est une bijection de N sur A. En particulier, l'ensemble des nombres entiers naturels pairs est strictement dénombrable, l'ensemble des nombres premiers est strictement dénombrable...
3. Z est strictement dénombrable. En effet, l'application $f : N \to Z$ définie par $f(n) = \frac{n}{2}$ si n est pair et par $f(n) = -\frac{n+1}{2}$ si n est impair est une bijection de N sur Z (le vérifier).
4. $N \times N$ est strictement dénombrable. En effet, l'application $f : N \times N \to N$ définie par $f(p, n) = 2^p 3^n$ est injective (le vérifier et conclure d'après l'exemple 2).
5. Q_+^* est strictement dénombrable. En effet, en choisissant le représentant irréductible de toute fraction, on voit que Q_+^* est en bijection avec une partie de $N \times N$. Par conséquent, Q_+^* est en bijection avec une partie de N (exemple 4).
6. Q est strictement dénombrable. En effet, Q_+^* est en bijection avec N et donc avec N^* (exemples 1 et 5). En symétrisant pour l'addition, Q^* est en bijection avec Z^* et en prolongeant en 0, Q est en bijection avec Z. Or Z est strictement dénombrable (exemple 3). Par conséquent, Q est strictement dénombrable.

Ce dernier exemple est tout à fait fondamental en analyse, bien que tout à fait contraire à l'intuition élémentaire, ce qui prouve que l'intuition est de peu de secours dans l'étude des cardinaux infinis.

THEOREM 1.5 *(Cantor)*
R n'est pas dénombrable, et même, plus précisément :

$$\forall a, b \in R, a < b,]a, b[\text{ n'est pas dénombrable.}$$

Démonstration : admis

COROLLARY 1.1 *L'ensemble $R\backslash Q$ des nombres irrationnels n'est pas dénombrable.*

Démonstration : en effet, si $R\backslash Q$ était dénombrable, on pourrait le mettre en bijection avec Z^-. Mais Q peut être mis en bijection avec N^* (exemples 1 et 6), et alors R serait en bijection avec Z, ce qui contredit le théorème de Cantor⊡

Chapitre 2

Structures

Par structure algébrique sur un ensemble E, on entend la donnée d'un nombre fini de lois internes ou externes assujetties à vérifier un certain nombre de propriétés.

La notion fondamentale dans ce chapitre est celle d'isomorphisme, i.e. de bijection compatible avec les lois. En effet, si deux ensembles munis de structures sont isomorphes, toute propriété démontrée dans l'un et qui ne dépend que de la structure se transpose dans l'autre à l'aide de l'isomorphisme.

2.1 Groupes

2.1.1 Généralités

DEFINITION 2.1 *On dit qu'un ensemble G muni d'une loi de composition interne * est un* ***groupe*** *si :*

1. ** est associative :*

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

2. ** admet un élément neutre (G est donc non vide) :*

$$\exists e \in G, \forall a \in G, a*e = e*a = e$$

3. *tout élément possède un symétrique pour * :*

$$\forall a \in G, \exists a' \in G, a*a' = a'*a = e$$

*Si de plus la loi * est commutative, on dit que (G,*) est un groupe* ***commutatif*** *ou, plus souvent, un groupe* ***abélien****.*

D'après les propriétés générales des lois de composition internes, si (G,*) est un groupe, l'élément neutre est unique, ainsi que le symétrique d'un élément quelconque. Le symétrique d'un composé est le composé dans l'ordre inverse des symétriques. Tout élément est régulier à gauche et à droite. On en déduit donc :

PROPOSITION 2.1 *Soient $\gamma_a : G \to G$ définie par $\gamma_a(x) = a*x$ et $\delta_a : G \to G$ définie par $\delta_a(x) = x*a$, alors, quel que soit a appartenant à G, γ_a et δ_a sont des bijections de G sur G.*

Exemples :

1. (Z,+), (R,+), (R^*,.), (R_+^*,.), (Q^*,.), (Q_+^*,.), ({-1,+1},.) sont des groupes abéliens.
2. Si (G,.) est un groupe et E un ensemble quelconque, l'ensemble des applications de E vers G est muni naturellement d'une structure de groupe, en définissant la loi notée encore . par :

$$f.g : x \mapsto f(x).g(x).$$

   C'est d'ailleurs un des rares cas où il ne faut pas noter le symétrique de f par f^{-1} à cause de la confusion avec l'éventuelle application réciproque.
3. Si (G,.) et (H,.) sont deux groupes, G×H est muni d'une structure de groupe en définissant :

$$\forall (x,y),(u,v) \in G \times H, (x,y).(u,v) = (x.u, y.v).$$

   On dit alors que (G×H,.) est le produit des groupes G et H. Ceci s'étend bien sûr au produit de n groupes.

Exercices :

1. Soit $E = \{e, a\}$, montrer qu'il existe une seule loi de groupe sur E dont e est l'élément neutre.
2. Dans $R^* \times R$ on pose $(x,y) * (x',y') = (xx', \frac{y}{x'} + xy')$, montrer que cette loi confère à $R^* \times R$ une structure de groupe.

2.1.2 Sous-groupes

DEFINITION 2.2 *Soient (G,.) un groupe et H une partie de G, on dit que H est un **sous-groupe** de G si :*

1. *H est stable :*

$$\forall x, y \in H, x.y \in H$$

2. *H muni de la loi induite, encore notée ., a une structure de groupe.*

PROPOSITION 2.2 *Soient (G,.) un groupe et H une partie de G, les trois propriétés suivantes sont équivalentes :*

1. *H est un sous-groupe de G*
2. $$\begin{array}{l} H \neq \oslash \\ \forall x, y \in H, x.y \in H \\ \forall x \in H, x^{-1} \in H \end{array}$$
3. $$\begin{array}{l} H \neq \oslash \\ \forall x, y \in H, xy^{-1} \in H \end{array} .$$

Démonstration : circulaire

$(1) \Rightarrow (2)$ et $(2) \Rightarrow (3)$ sont évidents. On démontre donc $(3) \Rightarrow (1)$ ce qui prouvera l'équivalence des trois propriétés.

Comme H est non vide, il existe un élément x de H et donc $x.x^{-1} \in H$ (prendre $y = x$ dans (3)) d'où $1 \in H$. Par ailleurs, si $y \in H$, on a $y^{-1} \in H$ (prendre $x = 1$ dans (3)) et donc $x.(y^{-1})^{-1} = x.y \in H$. Ainsi, H est stable.

Les axiomes de groupe se vérifient alors immédiatement, la loi est associative (par restriction), elle admet un élément neutre 1, et y^{-1} est l'inverse de y dans H comme dans G⊡

Une partie stable d'un groupe n'est pas nécessairement un sous-groupe. Par exemple, N est une partie stable de Z pour l'addition, mais ce n'est pas un sous-groupe de Z.

Pour démontrer qu'un ensemble muni d'une loi est un groupe, on essaiera souvent d'écrire que c'est un sous-groupe d'un groupe connu, ce qui permet d'alléger les démonstrations.

Tout sous-groupe d'un groupe abélien est, bien entendu, abélien. Mais un groupe non commutatif peut avoir des sous-groupes commutatifs propres.

Exemples :

1. Quel que soit le groupe (G,.), (G,.) et ($\{1\}$,.) sont des sous-groupes de (G,.) ; un sous-groupe distinct des deux précédents (s'il en existe) est appelé sous-groupe propre.
2. (Q,+) est un sous-groupe de (R,+), (Z,+) est un sous-groupe de (Q,+) et donc a fortiori de (R,+).

2.1.3 Morphismes

DEFINITION 2.3 *Soient (G,.), (H,*) deux groupes, f une application de G dans H, on dit que f est un* ***morphisme de groupes*** *si on a :*

$$\forall x, y \in G, f(x.y) = f(x) * f(y).$$

Si f est bijectif, on dit que f est un ***isomorphisme de groupes****, si $G = H$ on dit que f est un* ***endomorphisme****, un endomorphisme bijectif s'appelle un* ***automorphisme****.*

DEFINITION 2.4 *On appelle* ***noyau*** *du morphisme f, et on note Ker(f) l'ensemble défini par :*

$$Ker(f) = \{x \in G / f(x) = 1_H\}.$$

DEFINITION 2.5 *On appelle* ***image*** *du morphisme f, et on note Im(f) ou f(G) l'ensemble défini par :*

$$Im(f) = \{y \in H, \exists x \in G, y = f(x)\}.$$

Ces définitions sont conformes aux définitions ? ? ; le théorème ? ? s'applique donc, i.e. si f est un isomorphisme de G sur H, alors f^{-1} est un isomorphisme de H sur G.

PROPOSITION 2.3 *Soit f un morphisme du groupe (G,.) dans le groupe (H,*), alors :*

1. $f(1_G) = 1_H$ *et* $\forall x \in G$, $f(x^{-1}) = [f(x)]^{-1}$
2. *l'image par f d'un sous-groupe de G est un sous-groupe de H ; en particulier Im(f) est un sous-groupe de H*
3. *l'image réciproque par f d'un sous-groupe de H est un sous-groupe de G ; en particulier Ker(f) est un sous-groupe de G*
4. *f est un morphisme injectif si et seulement si* $Ker(f) = \{1_G\}$.

Démonstration : en effet,

1. $\forall x \in G$, $x.1_G = x$, donc $f(x.1_G) = f(x) * f(1_G) = f(x)$ d'où par régularité $f(1_G) = 1_H$; de même $x.x^{-1} = x^{-1}.x = 1_G$ donc $f(x.x^{-1}) = f(x) * f(x^{-1}) = f(1_G) = 1_H$, idem à gauche, d'où $f(x^{-1}) = [f(x)]^{-1}$
2. Soit K un sous-groupe de G, on a :

   $$\forall y_1, y_2 \in f(K), \exists x_1, x_2 \in K, y_1 = f(x_1) \text{ et } y_2 = f(x_2)$$

   donc $y_1 * y_2^{-1} = f(x_1) * [f(x_2)]^{-1} = f(x_1.x_2^{-1})$; mais K est un sous-groupe de G, donc $x_1.x_2^{-1} \in K$ et $y_1 * y_2^{-1} \in f(K)$, d'où f(K) est bien un sous-groupe de H

3. preuve semblable (l'écrire)
4.

$$\begin{aligned} finjectif &\Leftrightarrow (x \neq y \Rightarrow f(x) \neq f(y)) \\ &\Leftrightarrow (x.y^{-1} \neq 1_G \Rightarrow f(x) * [f(y)]^{-1} \neq 1_H) \\ &\Leftrightarrow (x.y^{-1} \neq 1_G \Rightarrow f(x.y^{-1}) \neq 1_H) \\ &\Leftrightarrow Ker(f) = \{1_G\} \boxdot \end{aligned}$$

Quand un morphisme rencontre un autre morphisme, la descendance est assurée, car :

THEOREM 2.1 *Soient (G,.), (H,.), (K,.) trois groupes ; si f est un morphisme de G dans H et g un morphisme de H dans K, alors g∘f est un morphisme de G dans K.*

Démonstration : la seule chose à démontrer est que g∘f est un morphisme :

$$\forall x, y \in G, g \circ f(x.y) = g(f(x.y)) = g(f(x).f(y)) = g(f(x)).g(f(y)) = g \circ f(x).g \circ f(y) \boxdot$$

THEOREM 2.2 *Soit (G,.) un groupe ; si Aut(G) désigne l'ensemble des automorphismes de G, alors (Aut(G),∘) est un groupe.*

Démonstration : on sait déjà que la composition est interne et associative, l'élément neutre est l'identité de G, le symétrique d'un élément est la bijection réciproque qui est bien un automorphisme⊡

Si (G,.) est un groupe et f une bijection de G sur un ensemble H, alors on peut transposer la structure de G sur H à l'aide de f, en posant :

$$\forall x, y \in H, x.y = f(f^{-1}(x).f^{-1}(y)).$$

f devient alors un isomorphisme de groupes.

2.2 Corps

DEFINITION 2.6 *Soit un ensemble K muni de deux lois de composition internes + et *, on dit que (K,+,*) est un* ***corps*** *si :*

1. *(K,+) est un groupe abélien, d'élément neutre noté 0 ou 0_K*
2. ** est associative et distributive par rapport à l'addition, i.e.*

$$\forall a, b, c \in K, \begin{array}{l} a * (b * c) = (a * b) * c \\ a * (b + c) = a * b + a * c \text{ et } (a + b) * c = a * c + b * c \end{array}$$

3. ** possède un élément neutre non nul, noté 1 ou 1_K et appelé **élément unité***
4. $\forall x \in K^* = K\backslash\{0\}, \exists x' \in K, x * x' = x' * x = 1_K$.

*Si de plus la multiplication est commutative, on dit que K est un **corps commutatif**.*

Si (K,+,*) est un corps, (K*,*) est donc un groupe multiplicatif, qui est abélien si et seulement si K est commutatif.

Exemples :

1. Q, R, C sont des corps commutatifs pour les lois usuelles
2. $Q[\sqrt{2}] = \{a+b\sqrt{2}, a, b \in Q\}$ est un corps commutatif (le démontrer).

DEFINITION 2.7 *Soit (K,+,*) un corps et k une partie de K, on dit que k est un **sous-corps** de K, ou que K est un **sur-corps** de k, si :*

1. *k est stable pour les deux lois + et **
2. $1_K \in k$
3. *k muni des deux lois induites vérifie les deux premières propriétés de la définition précédente*
4. $\forall a \in k^*, a^{-1} \in k^*$.

Par exemple, Q est un sous-corps de $Q[\sqrt{2}]$ qui est lui-même un sous-corps de R.

2.3 Autres structures

2.3.1 Espaces vectoriels

DEFINITION 2.8 *Soit K un corps commutatif, on dit qu'un ensemble E a une structure de K-**espace vectoriel** si*

1. *E est muni d'une loi de composition interne + telle que (E,+) soit un groupe abélien*
2. *E est muni d'une loi externe de domaine d'opérateurs K, notée . vérifiant :*

$$\begin{aligned} \forall x &\in E, 1_K.x = x \\ \forall \lambda &\in K, \forall x, y \in E, \lambda.(x+y) = \lambda.x + \lambda.y \\ \forall \lambda, \mu &\in K, \forall x \in E, (\lambda+\mu).x = \lambda.x + \mu.x \\ \forall \lambda, \mu &\in K, \forall x \in E, \lambda.(\mu.x) = (\lambda\mu).x. \end{aligned}$$

On ne donne ici aucune définition ou propriété supplémentaire, la seconde partie de ce cours étant consacrée à cette notion. On peut toutefois remarquer dès maintenant que si K est un sur-corps commutatif du corps k, alors K est un k-espace vectoriel, en considérant comme loi externe sur K la restriction de la multiplication de K à k$\times$K.

2.3.2 Algèbres

DEFINITION 2.9 *Soit K un corps commutatif, on dit qu'un ensemble E a une structure de K-**algèbre** si :*

1. *E est muni d'une structure de K-espace vectoriel*
2. *E est muni d'une seconde loi interne notée $\times$, telle que (E,+,$\times$) vérifie les deux premières propriétés de la définition ? ?*
3. $\forall \lambda \in K, \forall x, y \in E, \lambda.(x \times y) = (\lambda.x) \times y = x \times (\lambda.y)$.

Exemples :

1. Si K est un sur-corps commutatif du corps k, alors K est une k-algèbre. Il est à noter que la multiplication sert à la fois comme loi interne et comme loi externe.
2. $\mathcal{M}_n(k)$ est une k-algèbre pour les lois usuelles définies sur l'ensemble des matrices carrées d'ordre n.

Chapitre 3

Les corps des réels et des complexes

3.1 Nombres réels

La définition de R ne relève pas de l'algèbre mais de l'analyse. En effet, un nombre réel apparaît comme la limite d'une suite adéquate de nombres rationnels. On se contente donc ici de rappeler les propriétés algébriques fondamentales de R.

THEOREM 3.1 *R est **archimédien**, i.e. :*

$$\forall a \in R, \forall x \in R^*, \exists n \in N, a \leq nx.$$

Démonstration : évident et fondamental

R est aussi archimédien pour la multiplication :

$$\forall a \in R, \forall x \in R, \forall x > 1, \exists n \in N, a \leq nx.$$

Comme R est totalement ordonné, on peut définir $\forall x \in R$, $|x| = \max(x, -x)$ qui se lit **valeur absolue**. On a alors :

$$\begin{array}{rcl} \forall x & \in & R, |x| \geq 0 \text{ et } |x| = 0 \Leftrightarrow x = 0 \\ \forall x, y & \in & R, |xy| = |x| \,.\, |y| \\ \forall x, y & \in & R, |x + y| \leq |x| + |y| \,. \end{array}$$

THEOREM 3.2 *(de la borne supérieure)*

1. *Toute partie de R non vide et majorée admet une borne supérieure.*
2. *Toute partie de R non vide et minorée admet une borne inférieure.*

Démonstration : admis

C'est en cela que R est convenable, car $A = \{x \in Q_+^* \,/\, x^2 < 2\}$ est une partie non vide et majorée de Q mais sans borne supérieure dans Q, alors que A est a fortiori majorée dans R et admet dans R une borne supérieure, à savoir justement $\sqrt{2}$.

THEOREM 3.3 *(des segments emboîtés)*

Soit $(T_n)_{n\in N}$, où $T_n = [a_n, b_n]$, une suite d'intervalles fermés bornés, non vides de R ; si $(T_n)_{n\in N}$ est décroissante (i.e. si $\forall n \in N,\ T_{n+1} \subset T_n$), alors $\cap_{n\in N} T_n$ est non vide. Si, de plus, $\lim_{n\to+\infty}(b_n - a_n) = 0$ alors cette intersection se réduit à un point.

Démonstration : conséquence du précédent

Ce théorème est faux pour le corps Q, en prenant par exemple pour a_n et b_n les valeurs approchées par défaut et par excès de $\sqrt{2}$ à 10^{-n} près.

3.2 Nombres complexes

3.2.1 Construction

Soit $C = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, a, b \in R \right\} \subset \mathcal{M}_2(R)$; alors, muni des lois usuelles sur $\mathcal{M}_2(R)$, C est un corps commutatif contenant $\left\{ \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}, a \in R \right\}$ qui est un sous-corps de C isomorphe à R. De plus, $i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ vérifie $i^2 + 1 = 0$.

Le lecteur scrupuleux peut se sentir frustré de construire C à partir de $\mathcal{M}_2(R)$ non encore défini. Cet exemple a toutefois le mérite de fixer les idées. On présente ci-dessous une construction plus formelle à partir de notions déjà présentées.

On considère une R-algèbre de dimension 2, de base (1, i) telle que 1 soit élément neutre pour la multiplication et $i^2 = -1$. Par exemple, on considère R^2 muni des lois suivantes :

$$\forall (x,y),(x',y') \in R^2,\ \begin{array}{l} (x,y) + (x',y') = (x + x', y + y') \\ (x,y).(x',y') = (xx' - yy', xy' + x'y) \end{array}.$$

Le lecteur vérifiera à la main que R^2 muni de ces deux lois est bien un corps commutatif, que $R \times \{0\}$ en est un sous-corps isomorphe à R que l'on confondra avec R. De plus $i = (0, 1)$ vérifie $i^2 + 1 = 0$. Tout couple (x, y) de réels peut alors s'écrire $(x, y) = x + iy$ avec $i^2 = -1$. En conclusion :

THEOREM 3.4 *Il existe un corps commutatif C, contenant un sous-corps isomorphe à R, tel que C soit un R-espace vectoriel de dimension 2, dans lequel il*

existe une base (1, i) avec $i^2 = -1$. Ainsi tout élément z de C s'écrit de manière unique sous la forme $z = a + ib$, avec $a, b \in R$. C s'appelle ***corps des nombres complexes****.*

DEFINITION 3.1 *Soit $z = a + ib \in C$, $a, b \in R$,*

1. *a s'appelle la* ***partie réelle*** *de z et se note Re(z), b s'appelle la* ***partie imaginaire*** *de z et se note Im(z), $a + ib$ s'appelle alors la* ***forme algébrique*** *de z.*
2. *Re et Im sont des applications linéaires de C sur R, i.e. :*

$$\forall z, z' \in C, \forall \lambda \in R, \begin{array}{l} \Re(z + z') = \Re(z) + \Re(z') \text{ et } \Im(z + z') = \\ Im(z) + \\ Im(z') \\ \Re(\lambda z) = \lambda \Re(z) \text{ et} \\ Im(\lambda z) = \lambda \\ Im(z) \end{array}$$

3. *on appelle nombre complexe* ***conjugué*** *de z, et on note $\bar{z}$, le nombre complexe défini par $\bar{z} = a - ib$. L'application $z \mapsto \bar{z}$ est un automorphisme involutif du corps C, i.e. :*

$$\forall z, z' \in C, \overline{z + z'} = \bar{z} + \bar{z}', \overline{zz'} = \bar{z}\bar{z}' \text{ et } \bar{\bar{z}} = z.$$

4.

$$\begin{aligned} \Re(z) &= \frac{1}{2}(z + \bar{z}), \\ Im(z) = \frac{1}{2i}(z - \bar{z}) & \\ z &\in R \Leftrightarrow z = \bar{z} \Leftrightarrow \\ Im(z) = 0 & \\ z &\in iR \Leftrightarrow z = -\bar{z} \Leftrightarrow \Re(z) = 0. \end{aligned}$$

Si $z \in iR$ on dit que z est ***imaginaire pur****.*

DEFINITION 3.2 *Si $z \in C$, alors $z\bar{z} \in R^+$; on appelle* ***module*** *de z et on note $|z|$ le nombre réel positif $|z| = \sqrt{z\bar{z}}$.*

Attention, $\sqrt{z\bar{z}}$ a un sens mais pas $\sqrt{z}$ en général : le symbole $\sqrt{}$ n'a de sens que pour les nombres réels positifs et désigne la racine carrée arithmétique.

PROPOSITION 3.1 *1. Si $z \in R$, $|z|$ n'est autre que la valeur absolue de z ; il n'y a donc pas incompatibilité de notation*

2. $\forall z \in R$, $\Re(z) \leq |z|$ *et* $Im(z) \leq |z|$
3. $|z| = 0 \Leftrightarrow z = 0$
4. $\forall z, z' \in C$, $\begin{array}{l} |zz'| = |z|\,|z'| \\ \left|\frac{1}{z}\right| = \frac{1}{|z|}$ *si* $z \neq 0 \\ |z + z'| \leq |z| + |z'| \\ |\bar{z}| = |z| \end{array}$
5. *Si* $z \in C^*$, $z^{-1} = \frac{\bar{z}}{|z|^2}$.

Démonstration : fort simple

1. si $z \in R$, $\bar{z} = z$ et donc $|z| = \sqrt{z^2}$ qui représente bien la valeur absolue de z
2. soit $z = a + ib \in C$, $a, b \in R$, on a alors $|z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}$, d'où $a \leq |a| \leq \sqrt{a^2 + b^2}$, idem pour b
3. $|z| = 0 \Leftrightarrow |z|^2 = 0 \Leftrightarrow z\bar{z} = 0 \Leftrightarrow (z = 0)$ ou $(\bar{z} = 0) \Leftrightarrow z = 0$
4. $|zz'|^2 = zz'\overline{zz'} = z\overline{z}z'\bar{z}' = |z|^2\,|z'|^2$
   prendre $z' = \frac{1}{z}$
   $|z + z'|^2 = (z + z')(\bar{z} + \bar{z}') = z\bar{z} + z\bar{z}' + \bar{z}z' + z'\bar{z}' = |z|^2 + 2\Re(z\bar{z}') + |z'|^2$
   or $\Re(z\bar{z}') \leq |z\bar{z}'| = |z|\,|z'|$ d'où $|z + z'|^2 \leq (|z| + |z'|)^2$
5. $z\overline{z} = |z|^2$ donc si $z \neq 0$ on en déduit $z\frac{\bar{z}}{|z|^2} = 1$, i.e. $z^{-1} = \frac{\bar{z}}{|z|^2}\boxdot$

DEFINITION 3.3 *Soit* $U = \{z \in C, |z| = 1\}$*, alors d'après les deux dernières propriétés* U *est muni d'une structure de groupe multiplicatif ; on l'appelle* ***groupe unité*** *de* C*.*

3.2.2 Interprétation géométrique

Soit P un plan euclidien réel rapporté à une base orthonormée, P sera considéré comme étant muni à la fois de sa structure vectorielle et de sa structure affine.

DEFINITION 3.4 *Soit* $z = x + iy$ *un nombre complexe quelconque écrit sous forme algébrique, on lui associe le point M ou le vecteur* $\overrightarrow{OM}$ *du plan P de coordonnées (x, y) dans le repère considéré d'origine O ; M ou* $\overrightarrow{OM}$ *s'appellent* ***image*** *de z et inversement z s'appelle* ***affixe*** *de M ou* $\overrightarrow{OM}$*.*

Le module de z n'est autre que la longueur du vecteur $\overrightarrow{OM}$. De plus, cette interprétation géométrique est particulièrement adaptée à la visualisation de l'addition des nombres complexes. En effet, si z est l'affixe de $\overrightarrow{OM}$ et z' l'affixe de $\overrightarrow{OM'}$, alors $z + z'$ est l'affixe du vecteur $\overrightarrow{OM} + \overrightarrow{OM'}$. Faire un dessin pour s'en convaincre !

DEFINITION 3.5 *Soit $z \in C^*$, alors $\frac{z}{|z|} \in U$; on appelle* ***Argument*** *de z l'angle de la rotation transformant le vecteur $\overrightarrow{u}$ en le vecteur image de $\frac{z}{|z|}$, et on appelle* ***argument*** *de z, noté arg(z), toute mesure de cet angle.*

Le nombre complexe nul ne possède pas d'argument. Il est tout à fait fondamental de remarquer que tout nombre complexe non nul possède un argument, mais une infinité d'arguments définis à $2k\pi$ près.

Soit θ un argument de $z = x + iy \neq 0$, on a $z = |z|\,(\frac{x}{|z|} + i\frac{y}{|z|})$. Ainsi, z peut s'écrire :

$$z = x + iy = |z|\,(\cos\theta + i\sin\theta).$$

Cette dernière écriture s'appelle **forme trigonométrique** du nombre complexe z non nul. Il est donc facile de passer de la forme algébrique à la forme trigonométrique, mais en général la détermination exacte d'un tel angle θ ne conduit pas à des calculs algébriques.

PROPOSITION 3.2 *Soient $z, z' \in C^*$ écrits sous forme trigonométrique :*

$$z = |z|\,(\cos\theta + i\sin\theta), z' = |z'|\,(\cos\theta' + i\sin\theta'),$$

alors :

$$\begin{array}{l} zz' = |z|\,|z'|\,(\cos(\theta+\theta') + i\sin(\theta+\theta')) \\ z^{-1} = \frac{1}{|z|}(\cos(-\theta) + i\sin(-\theta)) \end{array};$$

en d'autres termes :

$$\begin{array}{l} \arg(zz') \equiv \arg(z) + \arg(z')[2\pi] \\ \arg(z^{-1}) \equiv -\arg(z)[2\pi] \end{array}.$$

Démonstration : évident, car résulte des formules classiques de trigonométrie

$$\begin{array}{l} \cos(\theta+\theta') = \cos\theta\cos\theta' - \sin\theta\sin\theta' \\ \sin(\theta+\theta') = \sin\theta\cos\theta' + \cos\theta\sin\theta' \end{array}.$$

THEOREM 3.5 *(formule de Moivre)*

Soit $z = |z|\,(\cos\theta + i\sin\theta)$ un nombre complexe non nul, alors pour tout entier relatif n on a :

$$z^n = |z|^n\,(\cos n\theta + i\sin n\theta).$$

Démonstration : à l'aide d'un raisonnement par récurrence, aussi bien pour $n \geq 0$ que pour $n < 0$, à l'aide des propriétés précédentes.

On convient de noter tout nombre complexe de module 1 sous la forme $z = e^{i\theta}$ (exponentielle complexe), où θ est un Argument de z. Les formules précédentes prennent alors la forme simple suivante :

$$e^{i\theta}e^{i\theta'} = e^{i(\theta+\theta')}, (e^{i\theta})^{-1} = e^{-i\theta}, (e^{i\theta})^n = e^{in\theta}, \forall n \in Z.$$

De plus, des relations $e^{i\theta} = \cos\theta + i\sin\theta$ et $e^{-i\theta} = \cos\theta - i\sin\theta$ on déduit les formules suivantes, appelées formules d'Euler :

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \text{ et } \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

Les formules de Moivre et d'Euler permettent de démontrer aisément un grand nombre d'identités trigonométriques. On en donne deux exemples à titre d'exercices.

Exercices :

1. Soit $n \in N^*$, calculer $\cos n\theta$ en fonction de $\cos\theta$.
2. Calculer, en fonction de x, $C = 1 + \cos x + \cos 2x + ... + \cos(n-1)x$

(indication : poser $S = 1 + \sin x + \sin 2x + ... + \sin(n-1)x$ et calculer $C + iS$).

3.2.3 Résolution des équations

THEOREM 3.6 *Soit $z \in C^*$ un nombre complexe non nul quelconque, alors z a exactement deux racines carrées opposées, i.e. deux nombres complexes z' et z" solutions de l'équation en Z : $Z^2 = z$.*

Démonstration : de deux manières

1. on suppose z écrit sous forme algébrique $z = a + ib$ et on cherche Z également sous forme algébrique $Z = x + iy$, l'équation $Z^2 = z$ s'écrit alors $x^2 - y^2 + 2ixy = a + ib$, i.e.

$$\begin{array}{l} x^2 - y^2 = a \\ 2xy = b \end{array} \Leftrightarrow \begin{array}{l} x^2 - y^2 = a \\ x^2(-y^2) = -\frac{b^2}{4} \\ xy = \frac{b}{2} \end{array} ;$$

   x^2 et $-y^2$ sont donc les racines (réelles) de l'équation $X^2 - aX - \frac{b^2}{4} = 0$ (en effet, le discriminant de cette équation vaut $a^2 + b^2$ qui est strictement positif puisque z est non nul) : on obtient donc deux valeurs possibles pour x et deux valeurs possibles pour y, mais attention, cela ne conduit pas à obtenir quatre racines carrées pour z, car la relation $xy = \frac{b}{2}$ induit une connexion entre les signes de x et de y, et donc l'existence de deux racines carrées opposées□

2. on suppose z écrit sous forme trigonométrique $z = \rho e^{i\theta}$ et on cherche Z également sous forme trigonométrique $Z = re^{i\alpha}$, l'équation $Z^2 = z$ s'écrit alors $r^2 e^{2i\alpha} = \rho e^{i\theta}$, i.e.

$$r^2 = \rho \text{ et } 2\alpha \equiv \theta[2\pi]$$

d'où $r = \sqrt{\rho}$ et $\alpha \equiv \frac{\theta}{2}\ [\pi]$ (ne pas perdre de vue que $r > 0$) et donc $z' = \sqrt{\rho}e^{i\frac{\theta}{2}}$ et $z'' = \sqrt{\rho}e^{i(\frac{\theta}{2}+\pi)} = -\sqrt{\rho}e^{i\frac{\theta}{2}}\boxdot$

THEOREM 3.7 *Soit $az^2 + bz + c = 0$, $a \neq 0$, une équation du second degré à coefficients dans C, alors si $\Delta = b^2 - 4ac \neq 0$ cette équation admet deux racines distinctes, et si $\Delta = 0$ elle admet une racine double.*

Démonstration :

$$az^2 + bz + c = 0 \Leftrightarrow \begin{array}{l} a(z^2 + \frac{b}{a}z + \frac{c}{a}) = 0 \\ (z + \frac{b}{2a})^2 + \frac{c}{a} - \frac{b^2}{4a^2} = 0 \\ (z + \frac{b}{2a})^2 = \frac{b^2-4ac}{2a} = \frac{\Delta}{(2a)^2} \end{array}$$

donc d'après le théorème précédent, si $\Delta \neq 0$ il admet deux racines carrées opposées δ et $-\delta$ et l'on en déduit les deux racines $z' = \frac{-b+\delta}{2a}$ et $z'' = \frac{-b-\delta}{2a}$, et si $\Delta = 0$ il existe alors une racine double qui vaut $\frac{-b}{2a}\boxdot$

La résolution est donc formellement la même que dans le cas réel, à la recherche des racines carrées de Δ près. En particulier, si a, b, c sont réels et si Δ est positif ou nul, la résolution est exactement la même. Par contre, si Δ est strictement négatif, Δ a deux racines carrées complexes opposées qui sont imaginaires pures. L'équation $az^2 + bz + c = 0$ a ainsi deux racines complexes conjuguées.

Exemple : résoudre l'équation $z^2 + z + 1 = 0$

On a $\Delta = 1 - 4 = -3$, d'où $\delta = i\sqrt{3}$ et les racines sont $z' = \frac{-1+i\sqrt{3}}{2}$ et $z'' = \frac{-1-i\sqrt{3}}{2}$.

On pose généralement $j = z'$ et $j = z'' = (z')^2$. On a aussi $j^3 = 1$.

En fait, on a encore mieux que cela :

THEOREM 3.8 *Soit $z \in C^*$ un nombre complexe non nul quelconque, alors z a exactement n racines n-èmes distinctes, i.e. l'équation en Z, $Z^n = z$, a exactement n solutions.*

Démonstration : on est obligé dans ce cas d'utiliser la forme trigonométrique des nombres complexes :

$$z = \rho e^{i\theta} \text{ et } Z = re^{i\alpha},$$

l'équation $Z^n = z$ s'écrit alors :

$$r^n e^{in\alpha} = \rho e^{i\theta}$$

i.e. $r^n = \rho$ et $n\alpha \equiv \theta\ [2\pi]$ d'où $r = \rho^{\frac{1}{n}}$ et $\alpha \equiv \frac{\theta}{n}\ [\frac{2\pi}{n}]$; les racines de l'équation $Z^n = z$ sont donc les nombres complexes de la forme :

$$z_k = \rho^{\frac{1}{n}} \exp(i(\frac{\theta}{n} + \frac{2k\pi}{n})), k \in Z,$$

mais on remarque que l'on a :

$$z_k = z_{k'} \Leftrightarrow \frac{2k\pi}{n} \equiv \frac{2k'\pi}{n}[2\pi] \Leftrightarrow k \equiv k'[n],$$

et par conséquent z_0, z_1, ..., z_{n-1} sont les n racines n-èmes distinctes de z□

Par application de ce théorème au cas particulier $z = 1$, on obtient les fameuses racines n-èmes de l'unité

$$1, \exp(\frac{2i\pi}{n}), ..., \exp(\frac{2(n-1)i\pi}{n}).$$

Ainsi, j est une racine troisième de l'unité.

Soit z_0 une racine n-ème du nombre complexe z, i.e. $z_0^n = z$, l'équation $Z^n = z$ peut alors s'écrire $Z^n = z_0^n$, i.e. $(\frac{Z}{z_0})^n = 1$. On a donc le critère pratique suivant : on obtient toutes les racines n-èmes d'un nombre complexe en multipliant l'une d'elles par toutes les racines n-èmes de l'unité.

Enfin, on démontre (mais cela est difficile et demande un minimum d'outils empruntés à l'Analyse) que C est encore beaucoup mieux que cela :

THEOREM 3.9 *(de d'Alembert)*

Toute équation algébrique de degré n à coefficients dans C possède au moins une racine complexe, et donc par récurrence sur n, toute équation algébrique de degré n à coefficients complexes a exactement n racines comptées avec leur ordre de multiplicité.

Démonstration : admis

La période de construction par voie algébrique s'arrête donc ici. On dit aussi que C est un corps algébriquement clos : il est impossible de sortir de C en résolvant des équations algébriques à coefficients complexes.

Chapitre 4

Polynômes et fractions rationnelles

4.1 Polynômes

4.1.1 Structure

DEFINITION 4.1 *Soit K un corps commutatif, on appelle **polynôme à coefficients dans K** toute suite $P = (a_n)_{n\in N}$ d'éléments de K telle que :*

$$\exists n_0 \in N, \forall n > n_0, a_n = 0.$$

Donc, si $P = (a_n)_{n\in N}$ est un polynôme à coefficients dans K, l'ensemble $\{i \in N,\, a_i \neq 0\}$ est un ensemble **fini** que l'on appelle **support** du polynôme P, a_i s'appelle le coefficient d'indice i du polynôme P. Si tous les coefficients du polynôme P, sauf un, sont nuls, on dit que P est un **monôme**. Si tous les coefficients sont nuls, on dit que P est le polynôme nul que l'on notera encore 0.

On rappelle enfin que l'égalité de deux polynômes est définie par l'égalité des suites, i.e. par l'égalité de **tous** les coefficients d'indices correspondants.

L'ensemble des polynômes à coefficients dans K se note K[X]. Cette notation recevra sa justification ultérieurement.

DEFINITION 4.2 *Soit $P \in K[X]$ un polynôme non nul, on appelle **degré** du polynôme P, et on note $d^\circ P$, l'indice du coefficient non nul d'indice le plus élevé.*

Cette définition a bien un sens, car si P est un polynôme non nul, son support est une partie finie non vide de **N**, il admet donc un plus grand élément. On convient de noter $d^\circ 0 = -\infty$.

THEOREM 4.1 *Muni des lois induites par celles définies sur l'ensemble des suites à valeurs dans K, K[X] a une structure de K-espace vectoriel.*

Démonstration : le vérifier en exercice

Si $P = (a_n)_{n\in N}$ et $Q = (b_n)_{n\in N}$ sont deux polynômes quelconques de K[X] et λ un scalaire quelconque, les lois sont donc définies par :

$$\begin{aligned} P+Q &= (a_n)+(b_n) = (a_n+b_n) = (a_0+b_0, a_1+b_1, ..., a_n+b_n, ...) \\ \lambda P &= \lambda(a_n) = (\lambda a_n) = (\lambda a_0, \lambda a_1, ..., \lambda a_n, ...). \end{aligned}$$

DEFINITION 4.3 *Soient $P = (a_n)_{n\in N}$ et $Q = (b_n)_{n\in N}$ deux polynômes à coefficients dans K, on appelle* ***produit*** *de P et Q et on note PQ le polynôme défini par :*

$$PQ = (c_n)_{n\in N}, \forall n \in N, c_n = a_0b_n + a_1b_{n-1} + ... + a_nb_0 \text{ i.e. } c_n = \sum_{i+j=n} a_ib_j.$$

Cette multiplication n'est donc pas la multiplication terme à terme des suites.

THEOREM 4.2 *Soit K un corps commutatif, K[X] muni des lois précédentes a une strucutre de K-algèbre commutative.*

Démonstration : d'après le théorème précédent K[X] est déjà un K-espace vectoriel. Le produit de polynômes est bien une loi interne car le support de PQ est inclus dans la somme des supports de P et Q, au sens de la somme de deux parties de N. En effet, pour que c_n soit non nul, il est nécessaire qu'il existe deux entiers i et j tels que $i+j = n$ et $a_i \neq 0, b_j \neq 0$. Il reste à vérifier tous les axiomes un à un : on indique simplement comment démontrer l'associativité du produit. Soient $P = (a_n)$, $Q = (b_n)$, $R = (c_n)$, alors $PQ = (d_h)$ avec $\forall h \in N$, $d_h = \sum_{i+j=h} a_ib_j$, $(PQ)R = (e_n)$ avec $\forall n \in N$, $e_n = \sum_{h+m=n} d_hc_m$ donc $e_n = \sum_{h+m=n}(\sum_{i+j=h} a_ib_j)c_m = \sum_{i+j+m=n} a_ib_jc_m$. Le seul point délicat est bien sûr le dernier signe d'égalité, il convient de bien le comprendre. Cette dernière écriture est symétrique par rapport aux coefficients des polynômes P, Q, R. La commutativité du produit étant évidente, l'associativité en résulte. On remarque enfin que l'on a :

$$\forall P, Q \in K[X], \forall \lambda \in K, \lambda(PQ) = (\lambda P)Q = P(\lambda Q) \boxdot$$

THEOREM 4.3 *Soient P et Q deux éléments quelconques de K[X], on a :*

$$\begin{aligned} d^\circ(P+Q) &\leq \max(d^\circ P, d^\circ Q), \text{ et si } d^\circ P \neq d^\circ Q \text{ alors } d^\circ(P+Q) = \max(d^\circ P, d^\circ Q) \\ d^\circ(PQ) &= d^\circ P + d^\circ Q. \end{aligned}$$

Démonstration : la formule pour la somme est évidente, on ne démontre que celle pour le produit. Le résultat est clair si l'un au moins des polynômes est le polynôme nul. On suppose donc P et Q non nuls et on appelle p et q leur degré respectif : $P = (a_n)$, $Q = (b_n)$, $a_p \neq 0$, $b_q \neq 0$, $\forall i > p$ $a_i = 0$,

$\forall j > q\ b_j = 0$. Par définition $PQ = (c_n)$ avec $\forall n \in N$, $c_n = \sum_{i+j=n} a_i b_j$. Si $n > p + q$, $i \leq p$ et $i + j = n$ entraînent $j > q$ et donc la somme qui définit c_n ne contient que des termes nuls. Si $n = p + q$, $i < p$ et $i + j = n$ entraînent $j > q$ et par conséquent $c_{p+q} = a_p b_q \neq 0$. PQ est donc de degré $p + q$, indice du coefficient non nul de plus grand indice⊡

On remarque alors que les éléments inversibles de K[X] sont les polynômes de degré 0. En effet, $PQ = 1$ entraîne $d^\circ P + d^\circ Q = 0$, donc $d^\circ P = d^\circ Q = 0$ et $(a_0, 0, ...)(b_0, 0, ...) = (1, 0, ...) \Leftrightarrow a_0 b_0 = 1$. Par conséquent $(a_0, 0, ...)$ est inversible si et seulement si a_0 est non nul.

Enfin, l'application $\varphi : K \to K[X]$ définie par $\varphi(a) = (a_0, 0, ...)$ est un morphisme injectif d'algèbres. Dans la suite de ce chapitre on confondra K avec l'ensemble des polynômes de degré au plus zéro, appelés polynômes constants. Il n'y a aucun risque d'ambiguïté car l'écriture λP a le même sens, que λ soit considéré comme scalaire ou comme polynôme constant.

4.1.2 Notation définitive

Soit $e_i = (0, ..., 0, 1, 0, ...)$, où 1 est à la $(i + 1)$-ème place, le monôme de degré i de coefficient 1. On a $\forall i, j \in N$, $e_i e_j = e_{i+J}$ et donc, par récurrence sur i, $e_i = (e_1)^i$ pour tout entier naturel i, avec la convention habituelle $(e_1)^0 = 1 = e_0$.

La définition d'un polynôme et des opérations montre alors que tout polynôme $P = (a_n)$ peut s'écrire de façon unique :

$$P = a_0 e_0 + a_1 e_1 + ... + a_n e_n + ... = \sum_{i \in N} a_i e_i$$

(remarquer que le support de P étant fini, cette sommation ne comprend en fait qu'un nombre fini de termes).

On pose alors $e_1 = X$, on obtient, puisque $e_0 = 1$:

$$P = a_0 + a_1 X + ... + a_n X^n + ... = \sum_{i \in N} a_i X^i \text{ (sommation finie).}$$

X s'appelle **indéterminée**. Attention, il ne s'agit pas d'une variable mais d'un polynôme particulier. L'avantage de cette notation est sa commodité d'emploi pour les opérations mais ne doit pas faire confondre une écriture telle que $a_0 + a_1 X = 0$ avec une équation en X.

On écrira dorénavant un polynôme P sous l'une des formes équivalentes suivantes :

$$P = a_0 + a_1 X + ... + a_n X^n = a_n X^n + ... + a_1 X + a_0.$$

La première écriture s'appelle ordonnée suivant les puissances croissantes et la seconde suivant les puissances décroissantes. Par convention, dans une telle écriture, on supposera, si cela est possible, que $a_n \neq 0$, donc n représentera le degré du polynôme P.

Lorsque $a_n = 1$, on dit que le polynôme P est unitaire ou mieux **normalisé**. Le produit de deux polynômes normalisés est encore normalisé (mais pas la somme, en général).

Les écritures précédentes rendent plus visible le fait que K[X] est un K-espace vectoriel et que $(X^i)_{i \in N}$ en est une base. De même, si on désigne par K_n[X] l'ensemble des polynômes à coefficients dans K de degré inférieur ou égal à n, K_n[X] est un K-espace vectoriel et $(1, X, ..., X^n)$ en est une base (de cardinal : $n + 1$).

4.1.3 Divisibilité

DEFINITION 4.4 *Soient $A, B \in K[X]$, $B \neq 0$, s'il existe un polynôme Q de K[X] tel que $A = BQ$, on dit que B **divise** A ou que A est un **multiple** de B, et on note B / A ; Q s'appelle le **quotient** dans la division de A par B.*

Si B est un diviseur de A et A non nul, on a nécessairement $d^\circ B \leq d^\circ A$. Si de plus $d^\circ B = d^\circ A$, le quotient est une constante : on dit que B est proportionnel à A. Réciproquement, toute constante non nulle est un diviseur de A.

THEOREM 4.4 *Soient $A, B \in K[X]$, $B \neq 0$, il existe un couple unique (Q, R) de polynômes de K[X] tel que :*

$$A = BQ + R \text{ avec } d^\circ R < d^\circ B.$$

*Q s'appelle le **quotient** et R le **reste** de la **division euclidienne** de A par B.*

On remarque immédiatement que $d^\circ R < d^\circ B$ autorise le cas $R = 0$, i.e. le cas où A est un multiple de B.

Démonstration : d'abord l'existence, puis l'unicité

1. on pose $A = a_0 + a_1 X + ... + a_m X^m$, $B = b_0 + b_1 X + ... + b_n X^n$ avec $b_n \neq 0$. Si $d^\circ A < d^\circ B$, on peut prendre $Q = 0$ et $R = A$. Si $d^\circ A \geq d^\circ B$, on pose $q_1 = \frac{a_m}{b_n} X^{m-n}$ et $r_1 = A - Bq_1$, on a alors $d^\circ r_1 < d^\circ A$ puisque l'on a fait disparaître le terme de plus haut degré de A. Si $d^\circ r_1 < d^\circ B$ on peut prendre $Q = q_1$et $R = r_1$. Sinon on répète sur le couple (r_1, B) ce qui vient d'être fait sur le couple (A, B). On définit donc q_2 et $r_2 = r_1 - Bq_2$ de façon à faire disparaître le terme de plus haut degré de r_1, etc... Le processus devra nécessairement s'arrêter au bout d'un nombre fini d'opérations car les degrés des restes successifs

sont strictement décroissants. Il suffira alors de prendre $Q = q_1 + q_2 + ... + q_p$ et $R = r_p$, le premier reste dont le degré est strictement inférieur à n.

2. on suppose que l'on a $A = BQ_1 + R_1 = BQ_2 + R_2$ avec $d^\circ R_1 < d^\circ B$ et $d^\circ R_2 < d^\circ B$, alors par différence on obtient :

$$R_2 - R_1 = B(Q_1 - Q_2),$$

d'où $d^\circ(R_2 - R_1) = d^\circ B + d^\circ(Q_1 - Q_2)$, or d'après les hypothèses :

$$d^\circ(R_2 - R_1) < \max(d^\circ R_1, d^\circ R_2) < d^\circ B.$$

L'égalité précédente est donc impossible si $d^\circ(Q_1 - Q_2) \geq 0$, on a donc $d^\circ(Q_1 - Q_2) = -\infty$, i.e. $Q_1 = Q_2$ et l'on en déduit aussi $R_1 = R_2 \boxdot$

La méthode pratique de division (euclidienne) s'inspire directement de la démonstration d'existence tandis que la disposition des calculs s'inspire de celle de la division classique dans $\mathbb{N}$. A cause de cette disposition, cette division est souvent appelée division suivant les puissances décroissantes.

Un théorème similaire existe pour la division suivant les puissances croissantes. La mise en pratique de cette division s'apparente donc à celle décrite ci-dessus, à la différence près que les polynômes y sont maintenant écrits suivant les puissances croissantes. Cela sera utilisé lors de l'étude des fractions rationnelles.

DEFINITION 4.5 *Soit $P \in K[X]$, on dit que P est un polynôme **premier**, ou mieux **irréductible**, si P n'est pas un polynôme constant et si les seuls diviseurs de P sont les constantes et les polynômes proportionnels à P, i.e. :*

$$P = P_1 P_2 \Rightarrow d^\circ P_1 = 0 \text{ ou } d^\circ P_2 = 0.$$

En particulier, tout polynôme du premier degré est nécessairement irréductible. La réciproque est vraie pour $K = C$ mais fausse en général, on reviendra sur ce fait au paragraphe suivant. On note simplement la conséquence immédiate, mais importante, de cette définition.

PROPOSITION 4.1 *Soient $P, Q \in K[X]$, si P est irréductible et si Q n'est pas un multiple de P, alors P et Q sont premiers entre eux.*

Démonstration : un diviseur commun à P et Q est un diviseur de P donc est ou bien une constante ou bien proportionnel à P, le second cas étant exclus par hypothèse, la conclusion en résulte$\boxdot$

THEOREM 4.5 *Soit P un polynôme non constant, alors il existe un scalaire λ non nul et une famille $P_1, ..., P_n$ de polynômes irréductibles normalisés tels que $P = \lambda P_1 P_2 ... P_n$. De plus, cette factorisation est unique à l'ordre des facteurs près.*

Démonstration : seulement l'existence (par récurrence sur le degré de P)

Si $d^\circ P = 1$, l'existence est évidente. On suppose donc l'existence assurée pour tout polynôme de degré inférieur ou égal à p. Soit alors P tel que $d^\circ P = p+1$. Deux cas se présentent : si P est irréductible, l'existence de la factorisation est alors évidente, sinon P n'est pas irréductible et on peut écrire $P = Q_1 Q_2$ avec $d^\circ Q_1 \neq 0$ et $d^\circ Q_2 \neq 0$, i.e. $d^\circ Q_1 \leq p$ et $d^\circ Q_2 \leq p$. On peut alors appliquer l'hypothèse de récurrence à Q_1 et Q_2, ce qui assure l'existence de la factorisation de P⊡

Par analogie avec l'arithmétique développée au premier chapitre, on peut alors introduire la notion de polynôme plus grand diviseur commun :

DEFINITION 4.6 *Soient $A_1, A_2 \in K[X]$ non tous deux nuls, alors il existe au moins un polynôme D tel que :*

1. *D / A_1 et D / A_2*
2. *$\forall B \in K[X]$, (B / A_1 et B / A_2) $\Rightarrow$ B / D.*

*D est appelé un **plus grand diviseur commun** (en abrégé PGCD) de A_1 et A_2.*

Si D est un polynôme de degré 0, A_1 et A_2 sont donc premiers entre eux.

4.1.4 Fonction polynôme

DEFINITION 4.7 *Soit $P = a_0 + a_1 X + ... + a_n X^n \in K[X]$, on appelle **fonction polynôme** associée à P et on note $\tilde{P}$ l'application de K dans K définie par :*

$$\forall x \in K, \tilde{P}(x) = a_0 + a_1 x + ... + a_n x^n.$$

On dit parfois, de façon imagée, que l'on a substitué la variable x à l'indéterminée X. La fonction polynôme associée au polynôme P se note souvent à l'aide du même symbole P, lorsqu'il n'y a pas de confusion possible.

DEFINITION 4.8 *Soient $a \in K$ et $P \in K[X]$, on dit que a est un **zéro** du polynôme P si $\tilde{P}(a) = 0$.*

THEOREM 4.6 *Soit $P \in K[X]$, alors $a \in K$ est un zéro du polynôme P si et seulement si P est divisible par $(X - a)$.*

Démonstration : on effectue la division euclidienne du polynôme P par le polynôme $X - a$:

$$P = (X - a)Q + R, d^\circ R < d^\circ (X - a).$$

R est un polynôme de degré strictement inférieur à 1, i.e. un polynôme constant. En prenant les valeurs des fonctions polynômes associées au point $x = a$, on trouve alors $\tilde{P}(a) = \tilde{R}(a) = R$, ce qui achève la démonstration⊡

Plus généralement, on dit que a est un **zéro d'ordre** α de P si P est divisible par $(X-a)^\alpha$ sans l'être par $(X-a)^{\alpha+1}$.

On remarque que la considération du corps ambiant K est primordiale pour la recherche des zéros d'un polynôme P. Par exemple, soit $P = X^4 - X^2 - 2$; considéré comme polynôme de Q[X] P n'a pas de zéro, considéré comme polynôme de R[X] il admet deux zéros $\{-\sqrt{2}, +\sqrt{2}\}$ et enfin considéré comme polynôme de C[X] il admet quatre zéros $\{-i, +i, -\sqrt{2}, +\sqrt{2}\}$.

THEOREM 4.7 *Soient $P \in K[X]$ un polynôme non nul et n le degré de P, alors P admet au plus n zéros dans K, comptés avec leur ordre de multiplicité (i.e. un zéro d'ordre α sera compté α fois dans le dénombrement des zéros dans K de P).*

Démonstration : on raisonne par récurrence sur le degré du polynôme P. Si $d^\circ P = 0$ ou $d^\circ P = 1$, le résultat est trivial. On suppose donc le résultat démontré pour tout polynôme de degré p et soit P un polynôme de degré $p+1$. Deux cas se présentent : soit P n'admet pas de zéros et le résultat est démontré, soit P admet au moins un zéro a. D'après le théorème précédent on peut alors écrire $P = (X-a)Q$, $Q \in K[X]$ et $d^\circ Q = p$. On peut appliquer à Q l'hypothèse de récurrence. Q admet au plus p zéros dans K et par conséquent P en admet au plus $p+1$ $\square$

On revient maintenant sur la factorisation dans K[X]. On suppose d'abord $K = C$. On rappelle le résultat fondamental de l'algèbre (théorème de d'Alembert) : *tout polynôme de degré n de C[X] admet n zéros comptés avec leur ordre de multiplicité*. On peut alors préciser le théorème ? ? de la façon suivante : *les seuls polynômes irréductibles de C[X] sont les polynômes du premier degré*. Par conséquent, si P est un polynôme de C[X] de degré n, il existe n nombres complexes, non nécessairement distincts, $x_1, x_2, ..., x_n$ et un scalaire λ non nul tels que :

$$P = \lambda(X - x_1)(X - x_2)...(X - x_n),$$

soit en effectuant les regroupements éventuels

$$P = \lambda(X - x_1)^{\alpha_1}(X - x_2)^{\alpha_2}...(X - x_p)^{\alpha_p}$$

avec $\alpha_1 + \alpha_2 + ... + \alpha_p = n$.

On suppose maintenant que les coefficients de P sont réels. Pour les besoins de la cause, on continue à considérer le polynôme P comme élément de C[X]. Alors, d'après les propriétés de la conjugaison, on a :

$$\forall a \in C, \overline{\tilde{P}(a)} = \tilde{P}(\bar{a}).$$

Par conséquent, si a est un zéro complexe non réel du polynôme P, $\overline{a}$ est également un zéro complexe non réel de P. P est alors divisible par $(X-a)$ et $(X-\bar{a})$ qui

sont premiers entre eux, P est donc divisible par le produit $(X-a)(X-\bar{a}) = X^2 - 2\Re(a)X + |a|^2 \in R[X]$. On peut alors également préciser le théorème ? ? : *les seuls polynômes irréductibles de R[X] sont les polynômes de degré 1 et les polynômes de degré 2 sans zéro réel, i.e. dont le discriminant est strictement négatif*. Dès lors, si P est un polynôme de R[X] de dgré n, il existe $n+1$ réels non nécessairement distincts tels que :

$$P = \lambda(X - x_1)...(X - x_p)(X^2 - 2a_1X + b_1)...(X^2 - 2a_qX + b_q)$$

avec $\forall i \in [1, q]$, $a_i^2 - b_i < 0$ et $p + 2q = n$. On peut de même regrouper les termes identiques.

4.1.5 Dérivation dans K[X]

THEOREM 4.8 *Il existe une unique application linéaire D de K[X] dans K[X] vérifiant :*

1. $D(X) = 1$
2. $\forall P, Q \in K[X]$, $D(PQ) = D(P).Q + P.D(Q)$.

Démonstration : par hypothèse, D est linéaire donc D est parfaitement déterminée par la connaissance des transformés d'une base. Il suffit donc de connaître la valeur de $D(X^p)$ pour tout entier p. Pour $p = 0$, on a $X = 1.X$ d'où $D(X) = D(1).X + 1.D(X)$, or $D(X) = 1$ et par suite $D(1) = 0$. Pour $p = 1$, on a par hypothèse $D(X) = 1 = 1.X^0$. On suppose ainsi que l'on a $D(X^{p-1}) = (p-1)X^{p-2}$, alors par hypothèse :

$$\begin{aligned} D(X^p) &= D(X.X^{p-1}) \\ &= D(X).X^{p-1} + XD(X^{p-1}) \\ &= X^{p-1} + (p-1)X^{p-1} \\ &= pX^{p-1}. \end{aligned}$$

Par conséquent, pour tout entier p on a prouvé par récurrence que $D(X^p) = pX^{p-1}$, ce qui démontre l'existence et l'unicité de D□

On appelle cette application **dérivation** dans K et si $P = \sum_{i=0}^{n} a_i X^i$ alors $D(P) = \sum_{i=1}^{n} i a_i X^{i-1}$.

Si $K = R$, la dérivation est formellement identique à la dérivation des fonctions polynômes, i.e. $\widetilde{D(P)} = \tilde{P}'$. On écrira donc P' au lieu de D(P). De même, on écrira P" au lieu de D(D(P)) et par récurrence $P^{(n)}$ pour la dérivée de la dérivée $(n-1)$-ème du polynôme P. Par convention, $P^{(0)}$ désignera le polynôme P.

PROPOSITION 4.2 *Soient $P_n = (X-a)^n$, $a \in K$ et n un entier naturel, on a :*

1. $\forall k < n$, $\tilde{P}_n^{(k)}(a) = 0$
2. $\tilde{P}_n^{(n)}(a) = n!$
3. $\forall k > n$, $\tilde{P}_n^{(k)}(a) = 0$.

Démonstration : par récurrence sur le degré n et en calculant.

THEOREM 4.9 *(Formule de Taylor)*
Soient $P \in K[X]$ tel que $d^\circ P \leq n$ et $a \in K$, la famille $(1, X-a, (X-a)^2, ..., (X-a)^n)$ est une base de $K_n[X]$. P est donc combinaison linéaire des éléments de cette famille et on a :

$$P = \tilde{P}(a) + \frac{\tilde{P}'(a)}{1!}(X-a) + \frac{\tilde{P}''(a)}{2!}(X-a)^2 + ... + \frac{\tilde{P}^{(n)}(a)}{n!}(X-a)^n.$$

Démonstration : on sait déjà que $(1, X, ..., X^n)$ est une base de $K_n[X]$. La famille donnée dans l'énoncé comporte $n+1$ polynômes, il suffit de vérifier qu'il s'agit d'une famille libre pour démontrer que c'est une base. Soit $\lambda_0 + \lambda_1(X-a) + ... + \lambda_n(X-a)^n = 0$ une combinaison linéaire nulle de cette famille, le seul terme en X^n est $\lambda_n X^n$ d'où $\lambda_n = 0$. Le seul terme en X^{n-1} est alors $\lambda_{n-1}X^{n-1}$ d'où $\lambda_{n-1} = 0$, etc... de proche en proche on démontre ainsi que tous les coefficients $\lambda_0, ..., \lambda_n$ sont nuls. Il existe donc une famille unique d'éléments de K telle que :

$$P = \alpha_0 + \alpha_1(X-a) + ... + \alpha_n(X-a)^n.$$

Il suffit alors d'appliquer la proposition précédente pour obtenir :

$$\forall i \in [0, n], \tilde{P}^{(i)}(a) = \alpha_i i!,$$

d'où le résultat en divisant par i !□

PROPOSITION 4.3 *$a \in K$ est un zéro d'ordre α de P si et seulement si on a :*

$$\tilde{P}(a) = \tilde{P}'(a) = ... = \tilde{P}^{(\alpha-1)}(a) = 0, \tilde{P}^{(\alpha)}(a) \neq 0.$$

Démonstration : évident d'après la forme même de la formule de Taylor.

4.2 Fractions rationnelles

Les seuls polynômes inversibles sont les polynômes de degré 0. K[X] n' est donc pas un corps. On construit alors le corps des fractions de K[X] (par analogie avec la construction de Q). Ce corps se note K(X) et s'appelle **corps des fractions rationnelles à coefficients dans** K.

4.2.1 Structure

Une fraction rationnelle F est une classe de couples de polynômes (P, Q) avec $Q \neq 0$. Si (P, Q) est un représentant quelconque de F, on convient d'écrire :

$$F = \frac{P}{Q}.$$

On a alors la relation d'équivalence suivante :

$$\frac{P}{Q} = \frac{P_1}{Q_1} \Leftrightarrow PQ_1 = QP_1.$$

Soit D le PGCD de P et Q, on peut écrire $P = DP_1$ et $Q = DQ_1$ et alors $\frac{P}{Q} = \frac{P_1}{Q_1}$ avec P_1 et P_2 premiers entre eux. On en déduit :

DEFINITION 4.9 *Soit F une fraction rationnelle de K(X), tout représentant de F, écrit $\frac{P_1}{Q_1}$, tel que P_1 et Q_1 soient premiers entre eux s'appelle* ***forme irréductible*** *de F. Si de plus Q_1 est normalisé, cette représentation est unique et s'appelle la* ***forme réduite*** *de F. P_1 et Q_1s'appellent respectivement* ***numérateur*** *et* ***dénominateur*** *de F.*

THEOREM 4.10 *Soient F une fraction rationnelle non nulle de K(X) et $\frac{P}{Q}$ un représentant quelconque de F, alors $d^\circ P - d^\circ Q$ est indépendant du représentant choisi de F. On l'appelle* ***degré*** *de la fraction rationnelle F. C'est un élément de* $\mathbb{Z}$.

Démonstration : en effet, si $F = \frac{P}{Q} = \frac{P_1}{Q_1}$, on a $PQ_1 = P_1Q$ et par conséquent $d^\circ P + d^\circ Q_1 = d^\circ P_1 + d^\circ Q$, d'où $d^\circ P - d^\circ Q = d^\circ P_1 - d^\circ Q_1$; on conviendra, comme pour les polynômes, de poser $d^\circ 0 = -\infty \boxdot$

DEFINITION 4.10 *Soit F une fraction rationnelle de K(X) écrite sous forme irréductible, on appelle* ***pôle*** *de F tout zéro de son dénominateur. On dira de même que $a \in K$ est un* ***pôle d'ordre*** *α de F si a est un zéro d'ordre α de son dénominateur.*

Attention, il est indispensable dans la définition précédente de choisir un représentant irréductible, sinon on risque, en introduisant des facteurs inopportuns, d'introduire en même temps des "faux pôles" parasites.

4.2.2 Fonctions rationnelles

DEFINITION 4.11 *Soit F une fraction rationnelle de K(X) écrite sous forme irréductible $\frac{P}{Q}$, on appelle* ***domaine de définition*** *de F et on note Def(F) l'ensemble des éléments de K qui ne sont pas pôles de F. On appelle* ***fonction rationnelle*** *associée à F l'application $\tilde{F}$ de Def(F) dans K définie par :*

$$\forall x \in Def(F), \tilde{F}(x) = \frac{\tilde{P}(x)}{\tilde{Q}(x)}.$$

Attention, la fraction rationnelle $F = \frac{1}{X-1}$ de R(X) admet 1 pour pôle. La fonction rationnelle associée n'est donc pas définie en 1. Mais il serait absurde de dire que F n'est pas définie en 1. En effet, X est une indéterminée et non une variable, l'écriture de F est formelle.

Si K est un corps contenant une infinité d'éléments (par exemple R ou C), le domaine de définition de $\tilde{F}$ contient nécessairement une infinité d'éléments, puisque le nombre de pôles est fini.

THEOREM 4.11 *Soit K un corps infini, l'application qui à toute fraction rationnelle associe la fonction rationnelle correspondante est injective.*

Démonstration : soient F et G deux fractions rationnelles de K(X) telles que $\tilde{F} = \tilde{G}$, alors $Def(F) = Def(G)$ et la fraction rationnelle $F - G$ est définie au moins sur Def(F) (en effet, en réduisant au même dénominateur $F - G$, il peut se faire que des facteurs se simplifient et fassent disparaître des pôles) ; $\tilde{F} - \tilde{G}$ est donc la fonction nulle sur Def(F), or Def(F) est un ensemble infini, le numérateur de $F - G$ admet donc une infinité de zéros, il s'agit par conséquent du polynôme nul, ce qui démontre que $F - G = 0$, i.e. $F = G\boxdot$

4.2.3 Décomposition des fractions rationnelles

On aborde maintenant le paragraphe fondamental de cette section. Le but de la manipulation est de transformer une fraction rationnelle quelconque en une somme de fractions simples (on précisera ce qu'on entend par "simple"). Cette somme sera donc nécessairement compliqué ! Ayant en vue l'application de ce problème à l'Analyse, on pourrait se limiter à l'étude de R(X). Mais le passage au domaine complexe est souvent indispensable (tout au moins en théorie). On fera ainsi d'abord l'étude générale relativement à un corps quelconque.

Cadre général

THEOREM 4.12 *Soit F une fraction rationnelle quelconque de K(X), écrite sous sa forme irréductible et normalisée ; on suppose également son dénominateur écrit sous forme de produits de polynômes irréductibles, i.e. :*

$$F = \frac{P}{Q} = \frac{P}{Q_1^{\alpha_1}...Q_p^{\alpha_p}}.$$

Il existe une famille unique de polynômes $(E, N_{11}, ..., N_{1\alpha_1}, N_{21}, ..., N_{2\alpha_2}, ..., N_{p1}, ..., N_{p\alpha_p})$ *telle que :*

$$F = E + \left(\frac{N_{11}}{Q_1} + ... + \frac{N_{1\alpha_1}}{Q_1^{\alpha_1}}\right) + ... + \left(\frac{N_{p1}}{Q_p} + ... + \frac{N_{p\alpha_p}}{Q_p^{\alpha_p}}\right)$$

avec $\forall i \in [1,p]$, $\forall j \in [1,\alpha_i]$, $d^\circ N_{ij} < d^\circ Q_i$.

Ce théorème n'est pas d'une grande simplicité d'écriture. On va le démontrer en découpant les difficultés en petits morceaux.

LEMMA 4.1 *Soit* $F = \frac{P}{Q}$*, il existe un unique polynôme E tel que* $F = \frac{P}{Q} = E + \frac{R}{Q}$*, avec* $d^\circ R < d^\circ Q$.

Démonstration : on effectue la division euclidienne de P par Q, on obtient $P = QE + R$, soit $\frac{P}{Q} = E + \frac{R}{Q}$, ce qui démontre l'existence de E ainsi que son unicité□

LEMMA 4.2 *Soit* $F = \frac{P}{Q_1 Q_2}$ *avec* $d^\circ P < d^\circ Q_1 + d^\circ Q_2$*,* Q_1*et* Q_2 *premiers entre eux ; il existe un couple unique de polynômes* U_1 *et* U_2 *tel que* $F = \frac{P}{Q_1 Q_2} = \frac{U_1}{Q_1} + \frac{U_2}{Q_2}$ *avec* $d^\circ U_1 < d^\circ Q_1$ *et* $d^\circ U_2 < d^\circ Q_2$.

Démonstration : Q_1 et Q_2 étant deux polynômes premiers entre eux, d'après le théorème de Bezout (admis) il existe deux polynômes U et V tels que $VQ_1 + UQ_2 = 1$, i.e. $P = (PV)Q_1 + (PU)Q_2$. On effectue la division euclidienne de PV par Q_2 : $PV = Q_2 q + U_2$ avec $d^\circ U_2 < d^\circ Q_2$. On a alors $P = U_2 Q_1 + (qQ_1 + PU)Q_2 = U_2 Q_1 + U_1 Q_2$. Comme $d^\circ U_2 < d^\circ Q_2$, on en déduit $d^\circ U_1 < d^\circ Q_1$, sinon on aurait $d^\circ P \geq d^\circ Q_1 + d^\circ Q_2$, ce qui contredirait l'énoncé. De $P = U_2 Q_1 + U_1 Q_2$ on tire $\frac{P}{Q_1 Q_2} = \frac{U_1}{Q_1} + \frac{U_2}{Q_2}$.

L'unicité d'une telle décomposition se vérifie facilement, car si $P = U_2 Q_1 + U_1 Q_2 = V_2 Q_1 + V_1 Q_2$, on en déduit $(U_2 - V_2)Q_1 = (V_1 - U_1)Q_2$. Comme Q_2 est premier avec Q_1, Q_2 diviserait $U_2 - V_2$ d'après le théorème de Gauss (admis), ce qui est absurde par comparaison des degrés si $U_2 - V_2 \neq 0$, idem pour $V_1 - U_1$□

LEMMA 4.3 *Soit* $F = \frac{P}{Q_1 Q_2 ... Q_p}$ *avec* $d^\circ P < d^\circ(Q_1 Q_2 ... Q_p)$*,* Q_1*, ...,* Q_p *premiers entre eux deux à deux ; il existe un unique p-uplet de polynômes* $(U_1, ..., U_p)$ *tel que :*

$$F = \frac{U_1}{Q_1} + ... + \frac{U_p}{Q_p}$$

et $\forall i \in [1,p]$, $d^\circ U_i < d^\circ Q_i$.

Démonstration : on raisonne par récurrence sur p. Le précédent lemme assure la véracité du résultat pour $p = 2$. On suppose alors le résultat acquis pour $(p-1)$ facteurs. Q_1 étant premier avec Q_2, ..., Q_p est aussi premier avec le produit $Q_2 ... Q_p$. Par application du lemme précédent, il existe un couple unique de polynômes U_1, V_1 vérifiant $d^\circ U_1 < d^\circ Q$ et $d^\circ V_1 < d^\circ(Q_2 ... Q_p)$ tel que :

$$F = \frac{U_1}{Q_1} + \frac{V_1}{Q_2 ... Q_p}.$$

On peut alors appliquer l'hypothèse de récurrence à la seconde fraction rationnelle, ce qui achève la preuve□

LEMMA 4.4 *Soit* $F = \frac{P}{Q^\alpha}$, $\alpha \in N^*$, *avec* $d^\circ P < \alpha.d^\circ Q$ *; il existe une unique famille de polynômes* $(P_1, ..., P_\alpha)$ *telle que :*

$$F = \frac{P_1}{Q} + \frac{P_2}{Q^2} + ... + \frac{P_\alpha}{Q^\alpha}$$

et $\forall i \in [1, \alpha]$, $d^\circ P_i < d^\circ Q$.

Démonstration : on raisonne par récurrence sur le nombre entier α. Le résultat est assez clair pour $\alpha = 1$. On suppose donc le résultat acquis jusqu'à l'ordre $(\alpha - 1)$. L'écriture $\frac{P}{Q^\alpha} = \frac{P_1}{Q} + ... + \frac{P_{\alpha-1}}{Q^{\alpha-1}} + \frac{P_\alpha}{Q^\alpha}$ montre, après mutiplication des deux membres par Q^α, que P_α ne peut être que le reste de la division euclidienne de P par Q. On effectue alors cette division :

$$P = Qq_1 + P_\alpha$$

avec $d^\circ P_\alpha < d^\circ Q$. De plus, de $d^\circ P < \alpha.d^\circ Q$ on déduit $d^\circ q_1 < (\alpha - 1).d^\circ Q$. On peut donc écrire :

$$\frac{P}{Q^\alpha} = \frac{q_1}{Q^{\alpha-1}} + \frac{P_\alpha}{Q^\alpha}$$

avec $d^\circ q_1 < (\alpha - 1).d^\circ Q$, ce qui permet d'appliquer l'hypothèse de récurrence et d'achever la preuve□

On peut maintenant donner la preuve du théorème de décomposition.

Démonstration : soit $F = \frac{P}{Q_1^{\alpha_1}...Q_p^{\alpha_p}}$, comme Q_1, ..., Q_p sont premiers entre eux deux à deux, il en est de même de $Q_1^{\alpha_1}$, ..., $Q_p^{\alpha_p}$. Par application des lemmes 1 à 3 il existe donc une unique famille de polynômes $E, U_1, ..., U_p$ avec $\forall i \in [1, p]$, $d^\circ U_i < d^\circ Q_i^{\alpha_i}$, telle que $F = E + \frac{U_1}{Q_1^{\alpha_1}} + ... + \frac{U_p}{Q_p^{\alpha_p}}$. Il suffit maintenant d'appliquer le lemme 4 à chaque terme $\frac{U_i}{Q_i^{\alpha_i}}$ pour obtenir la décomposition annoncée□

La démonstration de ce théorème d'existence et d'unicité ne fournit pas de méthode très pratique pour la recherche effective de cette décomposition. On va donc voir maintenant quelques recettes usuelles.

Décomposition dans C(X)

THEOREM 4.13 *Soit F une fraction rationnelle de C(X) écrite sous forme irréductible et normalisée, on suppose son dénominateur écrit sous forme de produits de polynômes irréductibles, i.e. :*

$$F = \frac{P}{Q} = \frac{P}{(X - a_1)^{\alpha_1}...(X - a_p)^{\alpha_p}}.$$

Il existe un unique polynôme E et une unique famille de scalaires $(\lambda_{11}, ..., \lambda_{1\alpha_1}, ..., \lambda_{p1}, ..., \lambda_{p\alpha_p})$ *tels que :*

$$F = E + \left(\frac{\lambda_{11}}{X - a_1} + ... + \frac{\lambda_{1\alpha_1}}{(X - a_1)^{\alpha_1}} \right) + ... + \left(\frac{\lambda_{p1}}{X - a_p} + ... + \frac{\lambda_{p\alpha_p}}{(X - a_p)^{\alpha_p}} \right).$$

Démonstration : il s'agit simplement d'une réécriture du théorème général, sachant que les seuls polynômes irréductibles de C[X] sont les polynômes du premier degré et donc que les conditions $d^\circ N_{ij} < d^\circ Q_i$ signifient par conséquent que N_{ij} est un polynôme constant.

E s'appelle la **partie entière** de F, $\frac{\lambda_{k1}}{X-a_k} + ... + \frac{\lambda_{k\alpha_k}}{(X-a_k)^{\alpha_k}}$ s'appelle la **partie polaire** relative au pôle a_k.

Ce théorème affirme l'existence et l'unicité de la décomposition d'une fraction rationnelle sur C en somme d'éléments simples. La partie entière s'obtient, d'après le lemme 1, par division euclidienne du numérateur par le dénominateur. On supposera dans la suite de cette étude que cette opération a été effectuée et on ne considérera plus que des fractions rationnelles de degré strictement négatif.

Puisque l'existence de la décomposition est acquise, on factorisera le dénominateur de la fraction rationnelle et on écrira la décomposition à l'aide de coefficients indéterminés qu'il s'agira de calculer le plus simplement et le plus rapidement possible.

Méthode universelle On réduit au même dénominateur la décomposition inconnue. En identifiant alors les numérateurs, on obtient un système linéaire comportant autant d'équations que d'inconnues, qu'il suffit de résoudre. Cette méthode n'est conseillée que dans des cas très simples ou de grande détresse.

Exemple : décomposer dans C(X) la fraction rationnelle $F(X) = \frac{1}{1-X^2}$.

On écrit :

$$\frac{1}{1-X^2} = \frac{1}{(1-X)(1+X)} = \frac{A}{1-X} + \frac{B}{1+X} = \frac{A+B+(A-B)X}{1-X^2},$$

d'où le système :

$$\begin{aligned} A + B &= 1 \\ A - B &= 0 \end{aligned}$$

i.e. $A = B = \frac{1}{2}$. On a donc :

$$\frac{1}{1-X^2} = \frac{1}{2}\left(\frac{1}{1-X} + \frac{1}{1+X} \right).$$

Partie polaire relative à un pôle simple On suppose que $a \in C$ est un pôle simple de la fraction rationnelle F, qui est donc de la forme suivante :

$$F(X) = \frac{P(X)}{(X-a)Q_1(X)}$$

avec $\tilde{Q}_1(a) \neq 0$. D'après le théorème précédent, la partie polaire relative au pôle a est de la forme $\frac{A}{X-a}$, $A \in C$, i.e. :

$$F(X) = \frac{P(X)}{(X-a)Q_1(X)} = \frac{A}{X-a} + \frac{P_1(X)}{Q_1(X)}.$$

On multiplie l'égalité précédente par $(X-a)$ et on obtient :

$$(X-a)F(X) = \frac{P(X)}{Q_1(X)} = A + (X-a)\frac{P_1(X)}{Q_1(X)},$$

donc a n'est plus un pôle de l'expression précédente. On peut par conséquent substituer a à X et on en tire :

$$A = \frac{\tilde{P}(a)}{\tilde{Q}_1(a)}.$$

On remarque par ailleurs que si on pose $Q(X) = (X-a)Q_1(X)$, on a :

$$Q'(X) = Q_1(X) + (X-a)Q_1'(X),$$

i.e. $\tilde{Q}'(a) = \tilde{Q}_1(a)$.

Exemple : décomposer dans C(X) la fraction rationnelle $F(X) = \frac{1}{X(X-1)(X-2)}$.

On écrit :

$$\frac{1}{X(X-1)(X-2)} = \frac{A}{X} + \frac{B}{X-1} + \frac{C}{X-2}$$

et on applique trois fois la méthode précédente.

On multiplie par X et on substitue à X la valeur 0 : on obtient $A = \frac{1}{2}$.

On multiplie par $(X-1)$ et on substitue à X la valeur 1 : on obtient $B = -1$.

On multiplie par $(X-2)$ et on substitue à X la valeur 2 : on obtient $C = \frac{1}{2}$.

On remarque que le calcul effectif des multiplications ne se fait jamais.

Partie polaire relative à un pôle multiple On suppose que $a \in C$ est un pôle multiple de F, qui est donc de la forme suivante :

$$F(X) = \frac{P(X)}{(X-a)^{\alpha} Q_1(X)}$$

avec $\alpha > 1$ et $\tilde{Q}(a) \neq 0$. D'après le théorème précédent, la décomposition de F s'écrit :

$$F(X) = \frac{A_1}{X-a} + \frac{A_2}{(X-a)^2} + ... + \frac{A_{\alpha}}{(X-a)^{\alpha}} + \frac{P_1(X)}{Q_1(X)}.$$

La méthode précédente est encore valable, mais uniquement pour le calcul de A_{α}. En effet :

$$(X-a)^{\alpha} F(X) = \frac{P(X)}{Q_1(X)} = A_1(X-a)^{\alpha-1} + ... + A_{\alpha-1}(X-a) + A_{\alpha} + \frac{(X-a)^{\alpha} P_1(X)}{Q_1(X)},$$

a est donc substituable à X et la substitution donne :

$$A_{\alpha} = \frac{\tilde{P}(a)}{\tilde{Q}_1(a)}.$$

Par contre, après avoir multiplié l'expression de F par $(X-a)^{\beta}$ avec $\beta < \alpha$, a reste un pôle des deux membres et n'est donc pas substituable. A ce moment, deux cas se présentent impliquant deux stratégies différentes.

Si α est petit, par exemple $\alpha = 2$ ou à la rigueur $\alpha = 3$, il ne reste alors à calculer qu'un coefficient (à la rigueur 2). On substitue dans ce cas une valeur simple (ou des valeurs simples) à X qui permettent d'achever le calcul.

Exemple : décomposer sur C la fraction rationnelle $F(X) = \frac{1}{(X-1)^2(X-2)}$.

On a :

$$\frac{1}{(X-1)^2(X-2)} = \frac{A}{X-1} + \frac{B}{(X-1)^2} + \frac{C}{X-2}.$$

On multiplie par $(X-2)$ et on substitue à X la valeur 2 : on obtient $C = 1$.

On multiplie par $(X-1)^2$ et on substitue à X la valeur 1 : on obtient $B = -1$.

Pour calculer A, il suffit maintenant de substituer à X une valeur simple (et substituable !), par exemple 0 :

$$\tilde{F}(0) = -\frac{1}{2} = -A + B - \frac{C}{2},$$

d'où $A = -1$.

Si α est grand, en pratique $\alpha \geq 3$ ou $\alpha \geq 4$, la méthode précédente serait alors trop lourde car elle conduirait à substituer un grand nombre de valeurs à X et à résoudre le système obtenu. Heureusement, il existe une méthode directe plus élégante permettant d'obtenir en une seule fois l'ensemble de la partie polaire relative à un pôle multiple. En effet, de l'écriture :

$$F(X) = \frac{P(X)}{(X-a)^{\alpha}Q_1(X)} = \frac{A_1}{X-a} + \frac{A_2}{(X-a)^2} + ... + \frac{A_{\alpha}}{(X-a)^{\alpha}} + \frac{P_1(X)}{Q_1(X)}$$

on déduit :

$$P(X) = [A_{\alpha} + A_{\alpha-1}(X-a) + ... + A_1(X-a)^{\alpha-1}]Q_1(X) + (X-a)^{\alpha}P_1(X).$$

On effectue alors le "changement d'indéterminée" $Y = X-a$, l'écriture précédente devient :

$$P(Y+a) = (A_{\alpha} + A_{\alpha-1}Y + ... + A_1Y^{\alpha-1})Q_1(Y+a) + Y^{\alpha}P_1(Y+a).$$

La partie polaire relative au pôle a apparaît alors être le quotient de la division suivant les puissances croissantes de $P(Y+a)$ par $Q_1(Y+a)$ à l'ordre $\alpha - 1$, le reste de cette division permettant d'ailleurs de déterminer P_1.

Exemple : décomposer dans C(X) la fraction rationnelle $F(X) = \frac{1}{(X-1)^3(X+1)^2}$.

L'application des techniques antérieures ne permet d'obtenir que deux des cinq coefficients de la décomposition. On effectue alors une division suivant les puissances croissantes. On pose $Y = X-1$, $F(Y+1) = \frac{1}{Y^3(Y+2)^2}$, d'où $1 = (Y+2)^2\left(\frac{1}{4} - \frac{Y}{4} + \frac{3}{16}Y^2\right) + Y^3\left(-\frac{1}{2} - \frac{3}{16}Y\right)$, soit en revenant à $F(Y+1)$:

$$\frac{1}{Y^3(Y+2)^2} = \frac{\frac{1}{4}}{Y^3} - \frac{\frac{1}{4}}{Y^2} + \frac{\frac{3}{16}}{Y} + \frac{-\frac{1}{2} - \frac{3}{16}Y}{(Y+2)^2}$$

et en revenant enfin à X :

$$\begin{aligned} F(X) &= \frac{1}{4(X-1)^3} - \frac{1}{4(X-1)^2} + \frac{3}{16(X-1)} + \frac{-\frac{3}{16}(X+1) - \frac{1}{8}}{(X+1)^2} \\ &= \frac{1}{4(X-1)^3} - \frac{1}{4(X-1)^2} + \frac{3}{16(X-1)} - \frac{3}{16(X+1)} - \frac{1}{8(X+1)^2}. \end{aligned}$$

On remarque que le fait d'avoir conservé le reste dans la seule division effectuée permet d'obtenir du même coup la partie polaire correspondant à l'autre pôle multiple.

Quelques astuces On reprend les notations générales du théorème précédent et on effectue une courte incursion dans le domaine de l'Analyse. La fraction rationnelle F(X) étant supposée de degré strictement négatif, l'expression x$\tilde{F}$(x) a une limite finie lorsque le module de x tend vers l'infini. Or :

$$XF(X) = \frac{A_{11}X}{X-a_1} + ... + \frac{A_{1\alpha_1}X}{(X-a_1)^{\alpha_1}} + ... + \frac{A_{11}X}{X-a_1} + ... + \frac{A_{p\alpha_p}X}{(X-a_p)^{\alpha_p}},$$

d'où par substitution de x à X dans le second membre, on trouve :

$$\lim_{|x|\to+\infty} x\tilde{F}(x) = A_{11} + A_{21} + ... + A_{p1}.$$

Les coefficients A_{11}, A_{21}, ..., A_{p1} s'appellent les résidus de F. La relation précédente est par conséquent très intéressante, car très facile à obtenir, en particulier si on connaît tous les résidus sauf un.

Si la fraction rationnelle est paire (ou impaire), l'unicité de la décomposition en éléments simples fait que cette symétrie doit apparaître dans cette décomposition, ce qui réduit pratiquement de moitié le nombre de coefficients à calculer.

Exemple : décomposer dans C(X) la fraction rationnelle $F(X) = \frac{X^2+1}{X(X-1)^2(X+1)^2}$.

On écrit :

$$F(X) = \frac{A}{X} + \frac{B}{X-1} + \frac{C}{(X-1)^2} + \frac{D}{X+1} + \frac{E}{(X+1)^2},$$

mais F(X) est une fraction rationnelle impaire. De la relation $F(-X) = -F(X)$ on obtient $B = D$ et $C = -E$, il suffit par conséquent de calculer A, B, C à l'aide des méthodes précédentes :

$$F(X) = \frac{1}{X} - \frac{1}{2(X-1)} + \frac{1}{2(X-1)^2} - \frac{1}{2(X+1)} - \frac{1}{2(X+1)^2}.$$

Décomposition dans R(X)

Dans le cas où $K = R$, le problème se complique (a priori) un petit peu, car deux types de polynômes réels irréductibles existent, à savoir les polynômes du premier degré et les polynômes du deuxième degré à discriminant négatif, qui correspondent à des couples de zéros complexes non réels conjugués.

THEOREM 4.14 *Soient $F = \frac{P}{Q}$ une fraction rationnelle irréductible de R(X) et*

$$Q(X) = (X-a_1)^{\alpha_1}...(X-a_m)^{\alpha_m}(X^2+p_1X+q_1)^{\beta_1}...(X^2+p_nX+q_n)^{\beta_n}$$

l'écriture de Q en produit de polynômes irréductibles ; il existe un unique polynôme E de R[X] et des familles uniques de réels (A_{ij}), $i \in [1,m]$ *et* $j \in [1,\alpha_i]$, (B_{kl}) *et* (C_{kl}), $k \in [1,n]$ *et* $l \in [1,\beta_k]$, *tels que :*

$$\frac{P(X)}{Q(X)} = E(X) + \sum_{i=1}^{m}\left(\sum_{j=1}^{\alpha_i}\frac{A_{ij}}{(X-a_i)^j}\right) + \sum_{k=1}^{n}\left(\sum_{l=1}^{\beta_k}\frac{B_{kl}X+C_{kl}}{(X^2+p_kX+q_k)^l}\right).$$

Démonstration : il s'agit simplement d'une réécriture du théorème général.

Les éléments de la première sommation s'appellent **éléments simples de première espèce** et ceux de la seconde **éléments simples de deuxième espèce**.

Ce qui a été dit pour les complexes reste encore valable, en particulier la partie entière s'obtient encore par division euclidienne. On écrira de même la décomposition à l'aide de coefficients indéterminés qu'il faudra calculer.

Méthode universelle

Exemple : décomposer dans R(X) la fraction rationnelle $F(X) = \frac{1}{(X+1)(X^2+1)}$.

On écrit :

$$F(X) = \frac{A}{X+1} + \frac{BX+C}{X^2+1} = \frac{(A+B)X^2+(B+C)X+A+C}{(X+1)(X^2+1)},$$

on obtient par identification :

$$\begin{aligned} A+B &= 0 \\ B+C &= 0 \\ A+C &= 1 \end{aligned}$$

d'où $A = \frac{1}{2}$, $B = -\frac{1}{2}$, $C = \frac{1}{2}$ ce qui donne :

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

Cette méthode devient très vite inextricable.

Eléments simples de première espèce Tout ce qui a été dit dans le cadre complexe s'applique aussi bien dans le cas de pôles simples que dans le cas de pôles multiples.

Eléments simples de deuxième espèce Là réside la nouveauté. Il existe alors deux façons de procéder.

On considère R(X) comme plongé dans C(X) et on effectue la décomposition de la fraction rationnelle dans C(X). Comme le dénominateur est à coeficients réels, les zéros complexes non réels sont deux à deux conjugués. Il suffit alors de regrouper les termes deux à deux pour retrouver la décomposition dans R(X).

Exemple : décomposer dans R(X) la fraction rationnelle $F(X) = \frac{1}{(X^2+1)(X^2+X+1)}$.

On écrit :

$$\frac{1}{(X^2+1)(X^2+X+1)} = \frac{1}{(X-i)(X+i)(X-j)(X-j^2)} = \frac{A}{X-i}+\frac{B}{X+i}+\frac{C}{X-j}+\frac{}{X}$$

On multiplie par $(X-i)$ et on substitue à X la valeur i : on obtient $A = -\frac{1}{2}$.

On multiplie par $(X-j)$ et on substitue à X la valeur j : on obtient $C = \frac{1}{2+j} = \frac{1-j}{3}$.

On trouve sans faire de calculs supplémentaires, mais en utilisant les propriétés de la conjugaison, $B = -\frac{1}{2}$ et $D = \frac{2+j}{3}$, d'où :

$$\frac{1}{(X^2+1)(X^2+X+1)} = \left(\frac{-\frac{1}{2}}{X-i} + \frac{-\frac{1}{2}}{X+i}\right) + \left(\frac{\frac{1-j}{3}}{X-j} + \frac{\frac{2+j}{3}}{X-j^2}\right)$$

et finalement :

$$F(X) = \frac{-X}{X^2+1} + \frac{X+1}{X^2+X+1}.$$

On peut aussi écrire directement la décomposition de F dans R(X) à l'aide de coefficients indéterminés et on substitue à X des valeurs complexes de la variable associée.

Exemple : décomposer dans R(X) la fraction rationnelle $F(X) = \frac{X+2}{(X+1)(X^2+1)}$.

On écrit :

$$\frac{X+2}{(X+1)(X^2+1)} = \frac{A}{X+1} + \frac{BX+C}{X^2+1}.$$

On multiplie par $(X+1)$ et on substitue à X la valeur -1 : on obtient $A = \frac{1}{2}$.

On multiplie par (X^2+1) et on substitue à X la valeur i qui est un zéro du polynôme X^2+1. La méthode s'applique encore et on obtient :

$$Bi + C = \frac{i+2}{i+1} = \frac{3-i}{2},$$

or B et C sont réels donc $B = -\frac{1}{2}$ et $C = \frac{3}{2}$, d'où finalement :

$$F(X) = \frac{1}{2(X+1)} + \frac{-X+3}{2(X^2+1)}.$$

Chapitre 5

Espaces vectoriels

Dans ce chapitre et dans les suivants, K désignera un corps commutatif quelconque. En pratique, on prendra le plus souvent $K = R$ ou $K = C$.

5.1 Structure

DEFINITION 5.1 *On appelle **espace vectoriel sur** K, ou encore K-**espace vectoriel**, tout ensemble E muni de deux lois :*

1. *une loi interne appelée addition, notée + telle que (E, +) soit un groupe abélien*
2. *une loi externe, de domaine K, qui à tout couple (λ, x) appartenant à K$\times$E fait correspondre un élément de E noté $\lambda.x$, cette loi vérifiant les quatre propriétés suivantes :*

   (a) $\forall x \in E, 1.x = x$

   (b) $\forall \lambda \in K, \forall x, y \in E, \lambda.(x + y) = \lambda.x + \lambda.y$

   (c) $\forall \lambda, \mu \in K, \forall x \in E, (\lambda + \mu).x = \lambda.x + \mu.x$

   (d) $\forall \lambda, \mu \in K, \forall x \in E, (\lambda\mu).x = \lambda.(\mu.x)$.

Un espace vectoriel n'est pas un ensemble : c'est un ensemble muni de deux lois vérifiant certaines propriétés. Autrement dit, un espace vectoriel sur K est un triplet (E, +, .), ce qui explique que, sur un même ensemble, il peut y avoir des structures d'espace vectoriel tout à fait différentes...Pour autant, on emploiera l'abus de langage courant : "soit E un espace vectoriel".

Les éléments de E sont appelés **vecteurs** tandis que les éléments de K sont appelés **scalaires**. Il est une convention assez universellement répandue qui consiste

à noter les vecteurs à l'aide de lettres minuscules de l'alphabet latin, et les scalaires à l'aide de lettres minuscules de l'alphabet grec. On manquera quelquefois à cette règle.

Lorsque $K = R$, on dira aussi **espace vectoriel réel**, et lorsque $K = C$, on dira aussi **espace vectoriel complexe**.

Exemples :

1. Les ensembles des vecteurs de la droite, du plan et de l'espace de la géométrie élémentaire munis des opérations classiques : $(\overrightarrow{v_1}, \overrightarrow{v_2}) \mapsto \overrightarrow{v_1} + \overrightarrow{v_2}$ et $(\alpha, \overrightarrow{v}) \mapsto \alpha.\overrightarrow{v}$ sont des espaces vectoriels réels. C'est d'ailleurs là l'origine de la locution "espace vectoriel".
2. Soient $n \in N^*$ et K^n le produit cartésien $K \times ... \times K$, on munit K^n de deux lois : une addition, dite composante à composante, définie par :

   $$(x_1, ..., x_n) + (y_1, ..., y_n) = (x_1 + y_1, ..., x_n + y_n)$$

   qui fait de K^n un groupe abélien d'élément neutre $(0, ..., 0)$, et une opération externe de domaine K, définie par :

   $$\lambda.(x_1, ..., x_n) = (\lambda.x_1, ..., \lambda.x_n).$$

   On vérifie aisément que K^n devient ainsi un K-espace vectoriel. En particulier, pour $n = 1$, K est muni d'une structure d'espace vectoriel sur lui-même. On parle donc de l'espace vectoriel R^n et de l'espace vectoriel C^n.
3. Si E et F sont deux K-espaces vectoriels (ils ont donc le même corps de base), on peut munir E×F d'une structure naturelle de K-espace vectoriel, en définissant ainsi les opérations :

   $$\begin{aligned} \forall (x,y),(x',y') &\in E \times F, (x,y) + (x',y') = (x + x', y + y') \\ \forall (x,y) &\in E, \forall \lambda \in K, \lambda.(x,y) = (\lambda.x, \lambda.y). \end{aligned}$$

   L'ensemble E×F muni de ces deux lois s'appelle l'espace vectoriel produit de E par F. Le lecteur pourra généraliser lui-même au cas de n espaces vectoriels E_1, ..., E_n sur le même corps K, et ainsi obtenir l'espace vectoriel $E_1 \times ... \times E_n$, voir les rapports entre cet exemple et l'exemple précédent.
4. L'ensemble K[X] des polynômes à coefficients dans K muni des lois classiques $(P, Q) \mapsto P+Q$ et $(\lambda, P) \mapsto \lambda.P$ est un K-espace vectoriel.

5. Soient D un ensemble quelconque et A(D, K) l'ensemble des applications de D dans K, on munit A(D, K) des lois suivantes :

$$\begin{aligned} \forall f, g &\in A(D,K), f+g : x \mapsto f(x)+g(x) \\ \forall f &\in A(D,K), \forall \lambda \in K, \lambda.f : x \mapsto \lambda.f(x). \end{aligned}$$

   Le lecteur pourra vérifier que ces deux lois sont bien internes sur A(D, K) et que muni de ces deux lois A(D, K) est un K-espace vectoriel, appelé espace des applications de D dans K. L'élément neutre pour l'addition est l'application $f_0 : D \to K$ définie par $\forall x \in D$, $f_0(x) = 0$. On dit que f_0 est l'application nulle et f_0 sera désormais notée 0, ce qui ne prête pas à confusion. L'opposée d'une application f est alors l'application de D dans K notée -f définie par $\forall x \in D$, $(-f)(x) = -f(x)$. On peut noter les cas particuliers suivants :

   (a) $D = N$, $K = R$, A(N, R) est l'espace vectoriel réel des suites réelles

   (b) $D = N$, $K = C$, A(N, C) est l'espace vectoriel complexe des suites complexes

   (c) $D \subset R$, $K = R$, A(D, R) est l'espace vectoriel réel des fonctions numériques, de variable réelle, définies sur le domaine D.

6. On peut généraliser l'exemple précédent de la manière importante suivante. Si D est un ensemble quelconque et si E est un K-espace vectoriel, A(D, E) peut être muni naturellement d'une structure de K-espace vectoriel.

7. Il se trouve que si E est un espace vectoriel sur un corps commutatif K et si k est un sous-corps de K, la restriction de l'opération externe de K×E à k×E munit E d'une structure d'espace vectoriel sur le corps k. Le lecteur pourra en effet vérifier que toutes les propriétés sont conservées. Autrement dit, tout espace vectoriel sur un corps K est aussi un espace vectoriel sur tout sous-corps k de K. Ainsi C, qui est un C-espace vectoriel, est aussi un R-espace vectoriel et un Q-espace vectoriel. Mais ces trois structures doivent être très soigneusement distinguées, le lecteur pourra se rendre compte par la suite que le corps de base a une importance fondamentale dans la structure d'espace vectoriel.

La proposition suivante montre qu'il n'y a absolument aucune surprise et que l'on calcule en fait comme dans toute structure algébrique classique.

PROPOSITION 5.1 *1.* $\forall (x, y) \in E, \forall \lambda \in K, \lambda(x - y) = \lambda x - \lambda y$

2. $\forall \lambda \in K$, $\lambda 0 = 0$
3. $\forall y \in E, \forall \lambda \in K$, $\lambda(-y) = -\lambda y$
4. $\forall x \in E, \forall (\lambda, \mu) \in K$, $(\lambda - \mu)x = \lambda x - \mu x$
5. $\forall x \in E, \forall \mu \in K$, $(-\mu)x = -\mu x$
6. $\forall x \in E$, $0x = 0$
7. $\forall x \in E, \forall \lambda \in K$, $(\lambda x = 0) \Leftrightarrow (\lambda = 0$ *ou* $x = 0)$.

Démonstration : le lecteur est invité à examiner les axiomes utilisés pour démontrer cette proposition

1. en effet, $\lambda(x - y) + \lambda y = \lambda((x - y) + y) = \lambda x$
2. on fait $x = y$ dans 1.
3. on fait $x = 0$ dans 1.
4. en effet, $(\lambda - \mu)x + \mu x = ((\lambda - \mu) + \mu)x = \lambda x$
5. on fait $\lambda = 0$ dans 4.
6. on fait $\lambda = \mu$ dans 4.
7. on suppose $\lambda x = 0$: si $\lambda = 0$ c'est bon, sinon λ est inversible dans le corps K et on a par multiplication $\lambda^{-1}(\lambda x) = \lambda^{-1}0 = 0$ d'où $1x = 0$ et donc $x = 0$; la réciproque est triviale, il s'agit de 2. et 6.$\boxdot$

DEFINITION 5.2 *Soit $(x_1, ..., x_n)$ un n-uplet de vecteurs d'un K-espace vectoriel E, un vecteur $x \in E$ est dit* ***combinaison linéaire*** *des vecteurs x_1, ..., x_n si l'on peut trouver un n-uplet $(\lambda_1, ..., \lambda_n)$ de scalaires tel que :*

$$x = \lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_n x_n.$$

Le lecteur notera dès maintenant qu'il n'a pas été dit que les vecteurs x_1, ..., x_n sont deux à deux distincts, i.e. on admet volontiers les répétitions. On admet de plus que certains des scalaires λ_1, ..., λ_n (ou même tous) puissent prendre la valeur 0. En particulier, $\forall i \in [1, n]$, x_i est combinaison linéaire de $(x_1, ..., x_n)$ car $x_i = 0x_1 + ... + 0x_{i-1} + 1x_i + 0x_{i+1} + ... + 0x_n$.

Exercice : dans R^3 on considère les deux vecteurs $A = (1, -1, 1)$ et $B = (-1, 3, 1)$

1. rechercher si l'un des vecteurs suivants, $C = (1, 2, 3)$, $D = (1, 0, 2)$, $E = (0, 1, 1)$ est combinaison linéaire des vecteurs A et B
2. soit $X = (x, y, z)$ un élément quelconque de R^3, quelle relation doit-il exister entre x, y, z pour que X soit une combinaison linéaire de A et B ? (réponse : $2x + y - z = 0$).

5.2 Sous-espaces vectoriels

DEFINITION 5.3 *Soient E un K-espace vectoriel et F une partie de E, on dit que F est un* ***sous-espace vectoriel*** *de E si la restriction de l'addition à F×F et la restriction de l'opération externe à K×F induisent sur F une structure de K-espace vectoriel.*

La définition même d'un sous-espace vectoriel F d'un K-espace vectoriel E contient le fait que F soit une partie stable de E pour les deux lois. Le lecteur n'ignore pas que l'on ne peut parler de loi induite que sur une partie stable pour cette loi.

L'utilisation de cette définition demande dix démonstrations pour vérifier qu'une partie F de E est un sous-espace vectoriel de E. Heureusement, le théorème suivant, qu'il faudra utiliser systématiquement, simplifie considérablement les choses.

THEOREM 5.1 *Soit F une partie d'un K-espace vectoriel E, les propositions suivantes sont équivalentes :*

1. *F est un sous-espace vectoriel de E*
2. *(F, +) est un sous-groupe du groupe additif (E,+)*
   F est stable pour l'opération externe, i.e. :

$$\forall \lambda \in K, \forall x \in F, \lambda x \in F$$

3. $F \neq \oslash$
   F est stable pour l'addition, i.e. :

$$\forall x, y \in F, x + y \in F$$

   F est stable pour l'opération externe
4. $F \neq \oslash$
   F est stable par combinaison linéaire de tout couple de vecteurs de F, i.e. :

$$\forall \lambda, \mu \in K, \forall x, y \in F, \lambda x + \mu y \in F.$$

Démonstration : il est clair que $1. \Rightarrow 2. \Rightarrow 3. \Rightarrow 4.$, pour montrer l'équivalence des quatre propositions il suffit donc de montrer que $4. \Rightarrow 1.$. On doit en premier lieu vérifier la stabilité pour les deux lois : soient donc x et y deux éléments de F, en prenant $\lambda = 1$ et $\mu = -1$ dans 4. on trouve $(x-y) \in F$. (F, +) est donc un sous-groupe abélien de (E, +). Par ailleurs, soit λ un scalaire et x un élément de F, en prenant $\mu = 0$ dans 4. on trouve $\lambda x \in F$. F est donc stable pour l'opération externe. Comme les quatre propriétés de la loi externe énoncées dans la définition ? ? sont vraies dans E, elles le restent a fortiori dans F, car F est inclus dans E⊡

Le critère 4. des sous-espaces montre qu'un sous-espace vectoriel F est stable par combinaison linéaire d'un couple de vecteurs de F. Plus généralement, on peut vérifier par récurrence sur $n \in N^*$ que F est stable par combinaison linéaire de n-uplets de vecteurs de F.

Il arrivera souvent d'avoir à montrer qu'un ensemble F muni de deux opérations, une interne et l'autre externe de domaine un corps commutatif K, est un K-espace vectoriel. Une excellente ruse consistera d'une part à plonger F dans un ensemble plus grand E, réputé être un K-espace vectoriel, et d'autre part à montrer que F est un sous-espace vectoriel de E. Il faudra donc connaître des K-espaces vectoriels classiques (cf exemples de la section précédente).

Exemples :

1. soit E un K-espace vectoriel, les parties $\{0\}$ et E sont des sous-espaces vectoriels de E appelés sous-espaces triviaux
2. soit $n \in N$, $K_n[X]$, ensemble des polynômes à coefficients dans K de degré inférieur ou égal à n, est un sous-espace vectoriel de K[X]
3. l'analyse fournit de nombreux exemples de sous-espaces vectoriels de $\mathcal{A}$(I, R), où I est un intervalle de R ; entre autres :
   (a) l'ensemble $\mathcal{C}$(I, R) des applications continues sur I
   (b) l'ensemble $\mathcal{D}$(I, R) des applications dérivables sur I
   (c) pour tout entier $n \geq 1$, l'ensemble $\mathcal{D}_n$(I, R) des applications n fois dérivables sur I
   (d) pour tout entier $n \in N$, l'ensemble $\mathcal{C}^n$(I, R) des applications de classe C^n sur I
   (e) si de plus I est un segment [a, b], le lecteur connaît sûrement l'ensemble $\mathcal{E}_{[a,b]}$ des fonctions en escalier sur [a, b] et l'ensemble $\mathcal{I}_{[a,b]}$ des fonctions intégrables au sens de Riemann sur [a, b].

Contre-exemples : il y a évidemment des parties de K-espaces vectoriels qui ne sont pas des sous-espaces vectoriels, on note en particulier :

1. l'ensemble des polynômes de degré exactement n n'est pas un sous-espace vectoriel de K[X]
2. l'ensemble des fonctions positives ou nulles (resp. négatives ou nulles) définies sur une partie D de R n'est pas un sous-espace vectoriel de $\mathcal{A}$(D, R)
3. l'ensemble des fonctions croissantes (resp. décroissantes) (resp. monotones) définies sur une partie D de R n'est pas un sous-espace vectoriel de $\mathcal{A}$(D, R).

Il est fortement conseillé au lecteur de passer quelques minutes à vérifier les assertions énoncées dans ces exemples ou contre-exemples.

Exercices :

1. Montrer que l'ensemble des triplets (x, y, z) de nombres réels vérifiant $\sqrt{2}x - \frac{y}{2} + (\log \pi)z = 0$ est un $\mathbb{R}$-espace vectoriel. Est-ce un $\mathbb{Q}$-espace vectoriel ?
2. Soit P (resp. I) l'ensemble des fonctions numériques paires (resp. impaires) définies sur $\mathbb{R}$, montrer que P et I sont des sous-espaces vectoriels de $\mathcal{A}(\mathbb{R}, \mathbb{R})$.

5.3 Applications linéaires

Le lecteur aura sûrement remarqué que toute structure a ses morphismes. Les morphismes sont ici des applications qui respectent la structure de K-espace vectoriel. Il est donc tout à fait normal de poser la définition suivante, qui ne contient en fait qu'une précision de vocabulaire.

DEFINITION 5.4 *Soient E et F deux K-espaces vectoriels, on appelle morphisme, ou encore* ***application*** *K-**linéaire** de E dans F, toute application $f : E \to F$ vérifiant les deux conditions suivantes :*

$$\begin{aligned} \forall x, y &\in E, f(x+y) = f(x) + f(y) \\ \forall x &\in E, \forall \lambda \in K, f(\lambda.x) = \lambda.f(x). \end{aligned}$$

Si $F = K$, une application linéaire de E dans K prend le nom de ***forme linéaire****.*

Les terminologies en vigueur fonctionnent encore dans le cadre des K-espaces vectoriels :

1. si $E = F$, on dit que f est un endomorphisme de E
2. si E et F sont quelconques et f bijective, on dit que f est un isomorphisme de K-espaces vectoriels
3. si $E = F$ et f bijective, on dit que f est un automorphisme de E
4. en outre, s'il existe un isomorphisme entre deux K-espaces vectoriels E et F, on dit que E et F sont isomorphes.

On introduit alors les notations suivantes :

1. l'ensemble des applications K-linéaires de E dans F est noté $\mathcal{L}_K$(E, F), ou encore $\mathcal{L}$(E, F) si aucune confusion n'est à craindre
2. l'ensemble des isomorphismes de E sur F est noté Iso(E, F)

3. l'ensemble des endomorphismes de E est noté $\mathcal{L}_K$(E), ou encore $\mathcal{L}$(E)
4. l'ensemble des automorphismes de E est noté GL_K(E), ou encore GL(E), et s'appelle le **groupe linéaire général** du K-espace vectoriel E (on verra plus tard l'explication de cette dénomination, c'est bien un groupe pour la composition des applications).

La proposition suivante peut être un critère rapide pour montrer qu'une application est une application linéaire.

PROPOSITION 5.2 *Soient E et F deux K-espaces vectoriels et f une application de E dans F, les propositions suivantes sont équivalentes :*

1. *f est une application K-linéaire*
2. $\forall \lambda, \mu \in K, \forall x, y \in E,\ f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$.

Démonstration : on laisse le soin au lecteur de poser $\lambda = \mu = 1$ ou $\mu = 0$.

Soient E et F deux K-espaces vectoriels, l'application $0 : E \to F$ définie par $\forall x \in E, 0(x) = 0$ est K-linéaire. On l'appelle l'application nulle de E dans F.

Si $f : E \to F$ est une application K-linéaire, la première condition de la définition ? ? montre que f est déjà un morphisme de groupes additifs. Dès lors, f possède toutes les propriétés d'un morphisme de groupes, en particulier : $f(0) = 0$ et $\forall x \in E,\ f(-x) = -f(x)$. Cette dernière propriété est aussi une conséquence de la deuxième condition de la définition ? ? en prenant $\lambda = -1$.

Exercices :

1. Soit E un K-espace vectoriel, on rappelle que l'homothétie de rapport $\alpha \in K$ de E est l'application :

$$h_\alpha : \begin{array}{c} E \to E \\ x \mapsto \alpha.x \end{array} .$$

   Montrer que $h_\alpha \in \mathcal{L}(E)$. Pour quelles valeurs de α a-t-on $h_\alpha \in GL(E)$?
2. Montrer que l'application $D : K[X] \to K[X]$ définie par $\forall P \in K[X],\ D(P) = P'$ (polynôme dérivé) est une application K-linéaire. Est-elle injective ?surjective ?
3. Montrer que l'application $I : \mathcal{I}_{[a,b]} \to R$ définie par $\forall f \in \mathcal{I}_{[a,b]},\ I(f) = \int_a^b f(t)dt$ est une forme linéaire sur $\mathcal{I}_{[a,b]}$.

PROPOSITION 5.3 *Soit E un K-espace vectoriel, l'application identique Id_E de E est K-linéaire. Autrement dit, Id_E est un endomorphisme de E. C'est même un automorphisme de E.*

Démonstration : il suffit de remarquer que Id_E est l'homothétie de rapport 1 (cf exercice)

PROPOSITION 5.4 *Soient E, F, G trois K-espaces vectoriels, f une application linéaire de E dans F, i.e.* $f \in \mathcal{L}(E, F)$*, et g une application linéaire de F dans G, i.e.* $g \in \mathcal{L}(F, G)$ *:*

$$E \xrightarrow{f} F \xrightarrow{g} G;$$

alors l'application composée $g \circ f$ *de E dans G est encore une application linéaire, i.e. :*

$$(f \in \mathcal{L}(E, F) \text{ et } g \in \mathcal{L}(F, G)) \Rightarrow g \circ f \in \mathcal{L}(E, G).$$

Démonstration : on emploie le critère de la proposition ? ? :

$$\forall \lambda, \mu \in K, \forall x, y \in E, \ \ g \circ f(\lambda x + \mu y) \begin{array}{l} = g(f(\lambda x + \mu y)) \\ = g(\lambda f(x) + \mu f(y)) \\ = \lambda g(f(x)) + \mu g(f(y)) \\ = \lambda g \circ f(x) + \mu g \circ f(y) \end{array} \quad \boxdot$$

PROPOSITION 5.5 *Soient E et F deux K-espaces vectoriels et f un isomorphisme de E sur F, l'application* f^{-1} *de F sur E est encore une application linéaire. Autrement dit,* f^{-1} *est un isomorphisme de F sur E.*

Démonstration : on pose $\forall \lambda, \mu \in K$, $\forall u, v \in F$, $x = f^{-1}(u)$ et $y = f^{-1}(v)$, ce qui a bien un sens puisque f est bijective, ce qui s'écrit encore $u = f(x)$ et $v = f(y)$; on a alors :

$$f^{-1}(\lambda u + \mu v) \begin{array}{l} = f^{-1}(\lambda f(x) + \mu f(y)) \\ = f^{-1}(f(\lambda x + \mu y)) \\ = \lambda x + \mu y \\ = \lambda f^{-1}(u) + \mu f^{-1}(v) \end{array} \quad \boxdot$$

On remarque que si E et F sont deux K-espaces vectoriels, l'ensemble $\mathcal{L}$(E, F) des applications K-linéaires de E dans F est une partie de l'ensemble A(E, F) de toutes les applications de E dans F. Or F est un K-espace vectoriel, A(E, F) est donc muni, de façon naturelle, d'une structure de K-espace vectoriel. Tous les espoirs sont donc permis.

THEOREM 5.2 *Soient E et F deux K-espaces vectoriels, alors l'ensemble* $\mathcal{L}$*(E, F) est un sous-espace vectoriel de* A*(E, F). Autrement dit,* $\mathcal{L}$*(E, F) et donc aussi* $\mathcal{L}$*(E) sont des K-espaces vectoriels.*

Démonstration : on emploie le troisième critère des sous-espaces vectoriels du théorème ? ?.

1. $\mathcal{L}(E,F) \neq \oslash$ car l'application nulle est linéaire
2. $\mathcal{L}$(E, F) est stable pour l'addition. On prend deux éléments f et g dans $\mathcal{L}$(E, F), on doit vérifier que $f+g$ est linéaire :

$$\forall \lambda, \mu \in K, \forall x, y \in E,\ \ (f+g)(\lambda x + \mu y) \begin{array}{l} = f(\lambda x + \mu y) + g(\lambda x + \mu y) \\ = \lambda f(x) + \mu f(y) + \lambda g(x) + \mu g(y) \\ = \lambda f(x) + \lambda g(x) + \mu f(y) + \mu g(y) \\ = \lambda (f+g)(x) + \mu (f+g)(y) \end{array}$$

3. $\mathcal{L}$(E, F) est stable pour l'opération externe. On prend un scalaire α et un élément f dans $\mathcal{L}$(E, F), on doit vérifier que $\alpha.f$ est linéaire :

$$\forall \lambda, \mu \in K, \forall x, y \in E,\ \ (\alpha.f)(\lambda x + \mu y) \begin{array}{l} = \alpha.f(\lambda x + \mu y) \\ = \alpha.(\lambda f(x) + \mu f(y)) \\ = \alpha\lambda f(x) + \alpha\mu f(y) \\ = \lambda(\alpha.f)(x) + \mu(\alpha.f)(y) \end{array} .$$

   On remarquera l'importance de la commutativité du corps K dans la dernière égalité□

On peut en outre munir $\mathcal{L}$(E) d'une structure plus complète.

THEOREM 5.3 *Soit E un K-espace vectoriel, ($\mathcal{L}$(E), +, ∘, .) est une K-algèbre.*

Démonstration : admis.

On vient de décrire les liens algébriques des applications linéaires entre elles. On passe maintenant à une description de l'application linéaire en s'intéressant à deux organes principaux : le noyau et l'image d'une telle application.

THEOREM 5.4 *Soient E et F deux K-espaces vectoriels et f une application K-linéaire de E dans F, alors :*

1. *l'image directe de tout sous-espace vectoriel de E est un sous-espace vectoriel de F ; en particulier*
   *$Im(f) = f(E)$ est un sous-espace vectoriel de F appelé **image** de f*
2. *l'image réciproque de tout sous-espace vectoriel de F est un sous-espace vectoriel de E ; en particulier $Ker(f) = f^{-1}(\{0\})$ est un sous-espace vectoriel de E appelé **noyau** de f*
3. *de plus, f est injective si et seulement si $Ker(f) = \{0\}$.*

Démonstration :

1. soit E' un sous-espace vectoriel de E, on considère l'ensemble f(E'). Comme E' contient 0, f(E') contient $f(0) = 0$, donc f(E') n'est pas vide. D'autre part, $\forall u \in F(E')$, $\exists x \in E'$, $f(x) = u$ et $\forall v \in f(E')$, $\exists y \in E'$, $f(y) = v$, donc $\forall u, v \in f(E')$, $\forall \lambda, \mu \in K$, $\lambda u + \mu v = \lambda f(x) + \mu f(y) = f(\lambda x + \mu y)$ car f est linéaire, mais E' étant un sous-espace vectoriel de E est stable par combinaison linéaire, donc $\lambda u + \mu v \in f(E')$
2. soit F' un sous-espace vectoriel de F, on considère l'ensemble f^{-1}(F'). Comme $f(0) = 0 \in F'$, $f^{-1}(F')$ contient 0, donc f^{-1}(F') n'est pas vide. D'autre part, $\forall x \in f^{-1}(F')$, $f(x) \in F'$ et $\forall y \in f^{-1}(F')$, $f(y) \in F'$, donc $\forall x, y \in f^{-1}(F')$, $\forall \lambda, \mu \in K$, $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) \in F'$, puisque F' est stable par combinaison linéaire, donc $f(\lambda x + \mu y) \in F'$, i.e. par définition $\lambda x + \mu y \in f^{-1}(F')$. $\{0\}$ étant trivialement un sous-espace vectoriel de F, la seconde partie en résulte
3. enfin, si f est injective, comme on a toujours $f(0) = 0$, on en déduit :

   $$(f(x) = 0) \Rightarrow (f(x) = f(0)) \Rightarrow (x = 0)$$

   donc $Ker(f) = \{0\}$. Réciproquement, si $Ker(f) = \{0\}$, on a puisque f est linéaire :

   $$\forall x, y \in E, (f(x) = f(y)) \Rightarrow (f(x - y) = 0) \Rightarrow (x - y = 0) \Rightarrow (x = y)$$

   donc f est injective□

Exercice : Soit E un K-espace vectoriel et f un endomorphisme de E

1. Montrer l'équivalence suivante :

   $$Im(f) \subset Ker(f) \Leftrightarrow f^2 = 0,$$

   où $f^2 = f \circ f$.
2. Montrer que si l'une de ces conditions équivalentes est réalisée, alors $Id_E + f$ est un automorphisme de E. On pourra calculer $(Id_E + f) \circ (Id_E - f)$.
3. Montrer plus généralement que s'il existe un entier $n \in N^*$ tel que $f^n = 0$ (on dit que f est nilpotent), alors $Id_E + f$ est un automorphisme de E.

5.4 Somme de sous-espaces vectoriels

Le lecteur est invité à constater de lui-même que la réunion de deux sous-espaces vectoriels d'un K-espace vectoriel n'est pas en général un sous-espace

vectoriel. Pour remédier à cet inconvénient, on va remplacer l'union des sous-espaces vectoriels par une opération plus convenable qui est la somme des sous-espaces vectoriels.

DEFINITION 5.5 *On appelle* ***somme*** *des sous-espaces vectoriels E_1, ..., E_n la partie, notée $E_1 + ... + E_n$ ou mieux $\sum_{i=1}^{n} E_i$, formée des éléments de E qui sont de la forme $x_1 + ... + x_n$ où $x_1 \in E_1$, ..., $x_n \in E_n$. Autrement dit :*

$$x \in \sum_{i=1}^{n} E_i \Leftrightarrow (\exists x_1 \in E_1)...(\exists x_n \in E_n), x = x_1 + ... + x_n.$$

On dira que x se décompose sur les sous-espaces E_1, ..., E_n. Le lecteur peut constater qu'il n'est pas dit que les vecteurs x_1, ..., x_n sont uniques (à suivre).

Exercices :

1. Soient E un K-espace vectoriel et F un sous-espace vectoriel de E, montrer que $F+F = F$, et plus généralement que $F+F+...+F = F$.
2. Soient E un K-espace vectoriel et E_1, ..., E_n n sous-espaces vectoriels quelconques de E ($n \in N^*$), on considère l'application :

$$\Phi : \begin{array}{l} E_1 \times E_2 \times ... \times E_n \to E \\ (x_1, x_2, ..., x_n) \mapsto x_1 + x_2 + ... + x_n \end{array}.$$

   On suppose $E_1 \times E_2 \times ... \times E_n$ muni de sa structure d'espace vectoriel produit.

   (a) Montrer que Φ est linéaire.

   (b) Décrire KerΦ et ImΦ.

THEOREM 5.5 *La somme $E_1 + ... + E_n$ de n sous-espaces vectoriels d'un K-espace vectoriel E est un sous-espace vectoriel de E. C'est en outre le plus petit sous-espace vectoriel de E (au sens de l'inclusion) contenant les sous-espaces vectoriels E_1, ..., E_n.*

Démonstration : en effet, on suppose que le lecteur a résolu l'exercice précédent.
Alors $E_1 + ... + E_n = \Phi(E_1 \times E_2 \times ... \times E_n) =$
$Im\Phi$ est un sous-espace vectoriel de l'espace d'arrivée, et $E_1 + ... + E_n$ contient bien entendu chaque sous-espace vectoriel E_i car $x_i = 0 + ... + x_i + ... + 0$, donc $E_i \subset E_1 + ... + E_n$. Mais de plus, si F est un sous-espace vectoriel de E contenant les E_i, alors :

$$\forall i \in [1, n], x_i \in E_i \Rightarrow x_i \in F$$

donc $x_1 + ... + x_n \in F$ puisque F est un sous-espace vectoriel de E, d'où $F \supset E_1 + ... + E_n$. $E_1 + ... + E_n$ est bien le plus petit sous-espace vectoriel de E contenant les $E_i \boxdot$

On conserve les notations précédentes. On a déjà remarqué que la décomposition d'un élément $x \in \sum_{i=1}^{n} E_i$ n'est pas nécessairement unique. Le lecteur se doute donc que pour avoir toujours l'unicité, il doit exister un concours de circonstances favorables, qui va conduire à la notion de somme directe, et à diverses caractérisations, puis à celle de sous-espaces supplémentaires.

DEFINITION 5.6 *Soient E un K-espace vectoriel et E_1, ..., E_n ($n \geq 1$) n sous-espaces vectoriels de E, on dit que la somme $E_1 + ... + E_n$ est* ***directe****, et on note $E_1 \oplus ... \oplus E_n$, si tout x appartenant à $\sum_{i=1}^{n} E_i$ admet une décomposition unique sur les sous-espaces vectoriels E_1, ..., E_n.*

PROPOSITION 5.6 *Avec les notations de la définition précédente, on a l'équivalence suivante :*

$$E_1 \oplus ... \oplus E_n \Leftrightarrow \forall (x_1, ..., x_n) \in E_1 \times ... \times E_n, x_1 + ... + x_n = 0 \Rightarrow x_1 = ... = x_n = 0.$$

Démonstration : ($\Rightarrow$) on sait que 0 admet une décomposition unique sur les E_i, or on a les deux égalités :

$$x_1 + ... + x_n = 0 \text{ et } 0 + ... + 0 = 0$$

d'où $\forall i \in [1, n]$, $x_i = 0$.

($\Leftarrow$) soit $x \in \sum_{i=1}^{n} E_i$, on suppose que l'on a deux décompositions de x :

$$x = x_1 + ... + x_n = y_1 + ... + y_n$$

avec $\forall i \in [1, n]$, $x_i \in E_i$, $y_i \in E_i$. Alors, par soustraction, on obtient :

$$0 = (x_1 - y_1) + ... + (x_n - y_n).$$

Les E_i étant des sous-espaces vectoriels de E, on a $x_i - y_i \in E_i$ pour tout $i \in [1, n]$, donc par hypothèse $x_i - y_i = 0$, $\forall i \in [1, n]$, ce qui démontre que les deux décompositions sont identiques, d'où l'unicité de la décomposition$\square$

THEOREM 5.6 *(cas de deux sous-espaces vectoriels)*

Soient E un K-espace vectoriel et E_1, E_2 deux sous-espaces vectoriels de E, on a l'équivalence :

$$E_1 \oplus E_2 \Leftrightarrow E_1 \cap E_2 = \{0\}.$$

Démonstration : ($\Rightarrow$) soit $x \in E_1 \cap E_2$, on doit montrer que x est nul, or x admet deux décompositions sur les sous-espaces vectoriels E_1 et E_2, à savoir :

$$x = x + 0 = 0 + x$$

avec $(x, 0) \in E_1 \times E_2$ et $(0, x) \in E_1 \times E_2$. L'unicité de la décomposition entraîne donc bien $x = 0$.

($\Leftarrow$) soit $x \in E_1 + E_2$, si x admettait deux décompositions distinctes sur E_1 et E_2, on aurait $x = x_1 + x_2 = y_1 + y_2$, avec des notations évidentes. Dès lors, par soustraction, $x_1 - y_1 = y_2 - x_2$, or $x_1 - y_1 \in E_1$ et $x_2 - y_2 \in E_2$, donc $E_1 \cap E_2 \neq \{0\}$, ce qui démontre la proposition par contraposée$\boxdot$

Attention, on a bien écrit $E_1 \cap E_2 = \{0\}$ et non pas $E_1 \cap E_2 = \oslash$: le lecteur est invité à bien saisir cette différence fondamentale, sous peine de n'avoir rien compris à tout ce qui vient d'être étudié.

THEOREM 5.7 *(cas général)*

Soient E un K-espace vectoriel et E_1, ..., E_n n sous-espaces vectoriels de E ($n \geq 1$), les propositions suivantes sont équivalentes :

1. *la somme $E_1 + ... + E_n$ est directe*
2. *$\forall i \in [1, n]$, $E_i \cap (\sum_{j \neq i} E_j) = \{0\}$*
3. *$\forall i \in [1, n-1]$, $(\sum_{j=1}^{i} E_j) \cap E_{i+1} = \{0\}$.*

Démonstration : (1.$\Rightarrow$2.) soit $u \in E_i \cap (\sum_{j \neq i} E_j)$, on a pour le vecteur u deux décompositions sur les sous-espaces E_i :

$$\begin{aligned} u &= 0 + ... + 0 + u + 0 + ... + 0, \text{ car } u \in E_i \\ u &= u_1 + ... + u_{i-1} + 0 + u_{i+1} + ... + u_n, \text{ car } u \in \sum_{j \neq i} E_j \end{aligned}$$

donc $u = 0$ par unicité de la décomposition

(2.$\Rightarrow$3.) on a $\forall i \in [1, n-1]$, $(\sum_{j=1}^{i} E_j) \cap E_{i+1} \subset E_{i+1} \cap (\sum_{j \neq i+1} E_j)$ et le membre de droite est réduit à $\{0\}$ par hypothèse

(3.$\Rightarrow$1.) soit $u \in E_1 + ... + E_n$, on suppose qu'il admet deux décompositions distinctes :

$$\begin{aligned} u &= x_1 + ... + x_n, \forall i \in [1, n], x_i \in E_i, \\ u &= y_1 + ... + y_n, \forall i \in [1, n], y_i \in E_i. \end{aligned}$$

Avec notre hypothèse, il existe i, $1 \leq i \leq n$, tel que $x_i \neq y_i$, donc $J = \{i \in [1, n], x_i \neq y_i\}$ est non vide et majoré par n, il admet par conséquent un plus grand élément, noté $i_o + 1$. Dès lors, $i_o \in [1, n-1]$ et par soustraction des deux décompositions, on obtient :

$$(x_1 - y_1) + ... + (x_{i_0} - y_{i_0}) = -(x_{i_0+1} - y_{i_0+1}).$$

Le premier membre appartient à $\sum_{i=1}^{i_0} E_i$ et le second membre à E_{i_0+1}. D'après l'hypothèse 3. ces deux membres sont donc nuls. En particulier, $x_{i_0+1} = y_{i_0+1}$, ce qui contredit le choix de $i_0 + 1$. On vient de montrer par l'absurde que J est vide. L'unicité de la décomposition en résulte, la somme est donc directe$\boxdot$

Attention, les conditions 2. et 3. du théorème sont beaucoup plus fortes que la condition $E_1 \cap ... \cap E_n = \{0\}$, ou même $\forall i \neq j$, $E_i \cap E_j = \{0\}$ (cf exercices).

Exercices :

1. Soient F et F' les sous-ensembles de R^3 définis par :

$$\begin{aligned} F &= \{(x,y,z) \in R^3, x - 2y + z = 0\}, \\ F' &= \{(x,y,z) \in R^3, 2x - y + 2z = 0\}. \end{aligned}$$

   (a) Montrer que F et F' sont deux sous-espaces vectoriels de R^3.

   (b) Montrer que $F + F' = R^3$ mais que la somme $F + F'$ n'est pas directe.

2. Soient D_1, D_2, D_3 les sous-espaces vectoriels de R^2 définis par :

$$\begin{aligned} D_1 &= \{(x,y) \in R^2, x = 2y\}, \\ D_2 &= \{(x,y) \in R^2, x = 3y\}, \\ D_3 &= \{(x,y) \in R^2, x = -y\}. \end{aligned}$$

   (a) Montrer que $\forall i, j \in \{1,2,3\}$, $i \neq j \Rightarrow D_i \cap D_j = \{0\}$.

   (b) Montrer que la somme $D_1 + D_2 + D_3$ n'est pas directe.

3. On considère à nouveau l'application :

$$\Phi : \begin{array}{l} E_1 \times E_2 \times ... \times E_n \to E \\ (x_1, x_2, ..., x_n) \mapsto x_1 + x_2 + ... + x_n \end{array}.$$

   Montrer que la somme $E_1 + ... + E_n$ est directe si et seulement si l'application Φ est un isomorphisme de K-espaces vectoriels.

DEFINITION 5.7 *Soient E un K-espace vectoriel et E_1, E_2 deux sous-espaces vectoriels de E, on dit que E_1 et E_2 sont* ***supplémentaires*** *si les deux conditions suivantes sont réalisées :*

1. *E_1 et E_2 ont une somme directe*
2. *$E_1 \oplus E_2 = E$ (leur somme est donc E tout entier).*

Le lecteur remarquera deux choses dans cette définition. L'adjectif "supplémentaire" s'adresse à deux sous-espaces, et il qualifie une propriété commune à E_1 et E_2. Il vérifiera que l'adjectif "supplémentaire" n'a rien à voir avec "complémentaire", et on espère qu'il n'aura jamais, ni l'occasion, ni le mauvais goût de les confondre. On l'engage en effet à méditer sur la stricte impossibilité de l'existence de sous-espaces vectoriels complémentaires. Par ailleurs, le théorème ? ? permet d'écrire la proposition suivante, très importante.

PROPOSITION 5.7 *Soient E un K-espace vectoriel et E_1, E_2 deux sous-espaces vectoriels de E, on a l'équivalence suivante :*

$$E = E_1 \oplus E_2 \Leftrightarrow \begin{array}{l} E = E_1 + E_2 \\ E_1 \cap E_2 = \{0\} \end{array} .$$

DEFINITION 5.8 *Soient E un K-espace vectoriel et F un sous-espace vectoriel de E, on appelle* ***supplémentaire*** *de F (sous-entendu : dans E) tout sous-espace vectoriel G de vérifiant $E = F \oplus G$.*

La première question que l'on peut se poser est la suivante : étant donné un sous-espace vectoriel F d'un K-espace vectoriel E, existe-t-il toujours un supplémentaire G de F (dans E) ? La deuxième question est tout aussi naturelle : si un supplémentaire existe, est-il unique ?

La réponse à la première question est oui. C'est un théorème très profond, dont la démonstration ne sera pas donnée ici, faute de moyens. Toutefois on en donnera, dans deux chapitres, une preuve dans un cas particulier. La réponse à la deuxième question est non en général. Ainsi on sera, in perpetuum, contraint d'employer l'article indéfini. Il existe cependant quelques exceptions : par exemple, E n'admet que $\{0\}$ pour supplémentaire, et inversement.

Exercices :

1. Soient dans R^3 les deux parties suivantes :

$$\begin{aligned} F &= \{(x,y,z) \in R^3, x+y+z=0\}, \\ G &= \{(\lambda,\lambda,\lambda), \lambda \in R\}. \end{aligned}$$

   (a) Montrer que F et G sont deux sous-espaces vectoriels de R^3.

   (b) Montrer que F et G sont supplémentaires.

   (c) Trouver d'autres supplémentaires pour F (resp. pour G).

   (d) En déduire que F (resp. G) admet une infinité de supplémentaires.

2. Dans l'espace produit $E_1 \times E_2$ de deux K-espaces vectoriels E_1 et E_2, montrer que les parties $\{0\} \times E_2$ et $E_1 \times \{0\}$ sont deux sous-espaces vectoriels supplémentaires.

3. Montrer que tous les supplémentaires d'un sous-espace vectoriel F d'un K-espace vectoriel E sont isomorphes. Indication : écrire $E = F \oplus G_1$ et $E = F \oplus G_2$, construire alors l'application φ de G_1 dans G_2 de la façon suivante : si $g_1 \in G_1$, $g_1 = f + g_2$ avec $f \in F$ et $g_2 \in G_2$ puisque $G_1 \subset E$; cette décomposition est unique, montrer alors que si on prend $\varphi(g_1) = g_2$, φ est un isomorphisme de G_1 sur G_2.

PROPOSITION 5.8 *Soient E et F deux K-espaces vectoriels et f une application linéaire de E dans F, alors tout supplémentaire de Ker(f) est isomorphe à Im(f).*

Démonstration : il suffit de le démontrer pour un supplémentaire d'après l'exercice précédent. Soit alors U un supplémentaire de Ker(f) dans E, on laisse au lecteur le soin de prouver que la restriction de f à U est un isomorphisme de U sur Im(f). La surjectivité étant triviale, il suffit d'ailleurs de vérifier l'injectivité.

Attention, le fait que tous les supplémentaires de Ker(f) soient isomorphes à Im(f) ne doit pas laisser penser (c'est une faute trop courante) que Ker(f) et Im(f) sont toujours supplémentaires. En fait, il ne faut pas confondre isomorphes et égaux. D'ailleurs, dans le cas où E est différent de F, Ker(f) est inclus dans E et Im(f) dans F, ils auraient donc beaucoup de mal à être supplémentaires.

Mais même si $E = F$, il est rare que Im(f) et Ker(f) soient supplémentaires. Néanmoins, par pur esprit de contradiction, on va immédiatement étudier une classe d'endomorphismes ayant cette propriété.

5.5 Projections et projecteurs

Soit E un K-espace vectoriel, on suppose donné pour la suite une décomposition de E en somme directe de deux sous-espaces vectoriels E_1 et E_2 :

$$E = E_1 \oplus E_2.$$

On rappelle que $x \in E$ admet une décomposition unique de la forme $x = x_1 + x_2$ avec $x_1 \in E_1$ et $x_2 \in E_2$. On définit alors deux applications p et q de la manière suivante :

$$p : \begin{array}{l} E \to E \\ x \mapsto p(x) = x_1 \end{array} \quad \text{et } q : \begin{array}{l} E \to E \\ x \mapsto q(x) = x_2 \end{array} .$$

DEFINITION 5.9 *L'application p s'appelle la* ***projection sur*** E_1 ***parallèlement à*** E_2*, de même l'application q s'appelle la* ***projection sur*** E_2 ***parallèlement à*** E_1.

PROPOSITION 5.9 *Les applications p et q sont deux endomorphismes du K-espace vectoriel E. De plus :*

$$\begin{array}{l} Ker(p) = E_2 \\ Im(p) = E_1 \\ Ker(q) = E_1 \\ Im(q) = E_2 \end{array} .$$

Démonstration : soient $x = x_1 + x_2$ et $y = y_1 + y_2$ deux éléments de E et leurs décompositions sur la somme directe, et λ, μ deux scalaires, on a :

$$\lambda x + \mu y = (\lambda x_1 + \mu y_1) + (\lambda x_2 + \mu y_2).$$

On voit donc aisément que cette écriture est la décomposition de $\lambda x + \mu y$ (unicité) sur la somme directe. Autrement dit, $p(\lambda x + \mu y) = \lambda x_1 + \mu y_1$ et $q(\lambda x + \mu y) = \lambda x_2 + \mu y_2$. La fin de la démonstration est triviale⊡

On a donc $E = Ker(p) \oplus Im(p)$ et $E = Ker(q) \oplus Im(q)$, puisque $E = E_1 \oplus E_2$. On répète que ce n'est en général pas le cas pour un endomorphisme quelconque.

PROPOSITION 5.10 *On définit comme précédemment p et q, on a alors :*

1. $p + q = Id_E$
2. $p \circ q = q \circ p = 0$
3. $p^2 = p$ *et* $q^2 = q$.

Démonstration :

1. $\forall x \in E$, $x = x_1 + x_2 = p(x) + q(x)$ donc $Id_E = p + q$
2. $\forall x \in E$, $x = x_1 + x_2$, on a $q(x) = x_2$, mais on écrit $q(x) = 0 + x_2$ pour faire apparaître sa décomposition sur $E_1 \oplus E_2$; on a donc $p(q(x)) = 0$, i.e. $p \circ q = 0$, et de même $q \circ p = 0$
3. $\forall x \in E$, $x = x_1 + x_2$, on a $p(x) = x_1$, ce qu'on écrit $p(x) = x_1 + 0$, on a donc $p(p(x)) = x_1$, i.e. $p \circ p = p$, idem pour q⊡

La dernière propriété exprime que p et q sont des idempotents de $\mathcal{L}$(E). On étudie maintenant ces idempotents.

DEFINITION 5.10 *Soit E un K-espace vectoriel, on appelle* ***projecteur*** *de E tout endomorphisme p de E vérifiant* $p^2 = p$.

On justifie cette appellation en établissant le lien entre projecteur et projection, à l'aide du théorème suivant.

THEOREM 5.8 *Soient E un K-espace vectoriel et p un projecteur de E, alors p est la projection sur Im(p) parallèlement à Ker(p).*

Démonstration : en deux étapes

1. on montre ici que Ker(p) et Im(p) sont supplémentaires :

(a) soit x un élément de E, une énorme ruse consiste à écrire $x = x - p(x) + p(x)$; on pose alors $x_2 = x - p(x)$ et $x_1 = p(x)$. Il est clair que $x_1 \in$
$Im(p)$; par ailleurs $p(x_2) = p(x - p(x)) = p(x) - p^2(x) = 0$ car $p = p^2$ et $p \in \mathcal{L}(E)$ donc $x_2 \in Ker(p)$. On a donc écrit $x = x_1 + x_2$ avec $x_1 \in$
$Im(p)$ et $x_2 \in Ker(p)$

(b) soit $x \in Ker(p) \cap$
$Im(p)$, comme $x \in$
$Im(p)$ il existe $t \in E$ tel que $x = p(t)$. Par ailleurs, $p(x) = 0$ car $x \in Ker(p)$, donc $p(p(t)) = 0$, i.e. $p^2(t) = 0$, d'où $p(t) = 0$ car $p = p^2$, et donc $x = 0$

2. on montre ici que p coïncide avec la projection π sur Im(p) parallèlement à Ker(p) : dans la première étape, on a découvert la décomposition de tout élément de E sur les sous-espaces Im(p) et Ker(p). Celle-ci est : $\forall x \in E$, $x = p(x) + (x - p(x)) = x_1 + x_2$, donc d'après la définition des projections, $\forall x \in E$, $\pi(x) = x_1 = p(x)$, i.e. $\pi = p\boxdot$

Le lecteur pourra montrer que si p est un projecteur, alors $q = Id_E - p$ est aussi un projecteur : q est justement l'autre projection associée à la décomposition $E = Im(p) \oplus Ker(p)$.

Chapitre 6

Génération et liberté

Les deux notions présentées ici, l'indépendance linéaire et la génération, sont les deux pôles de la théorie des espaces vectoriels. Ces deux notions ne sont pas simples et les liens qui les unissent sont loin d'être évidents. Si de surcroît on cherche à tout exprimer en termes de familles, les énoncés s'alourdissent, les démonstrations s'allongent et le niveau d'abstraction s'élève. En conséquence, il est indispensable, pour tirer quelque profit de l'étude théorique générale, de commencer par les situations les plus simples décrites par un petit nombre de vecteurs, sans lésiner sur les calculs explicites.

6.1 Préliminaires

Dans cette section, l'essentiel sera décrit pour un triplet de vecteurs. Il est vivement conseillé au néophyte de bien comprendre ces préliminaires. Le lecteur plus aguerri peut ignorer cette section.

6.1.1 Sous-espace vectoriel engendré

THEOREM 6.1 *Soit (x_1, x_2, x_3) un triplet de vecteurs de l'espace vectoriel E, l'ensemble F des combinaisons linéaires des vecteurs x_1, x_2, x_3 est un sous-espace vectoriel de E. C'est le plus petit sous-espace vectoriel (pour l'inclusion) de E contenant les vecteurs x_1, x_2, x_3.*

DEFINITION 6.1 *On appelle F le* ***sous-espace vectoriel engendré par*** *x_1, x_2, x_3 et on le note :*

$$F = \langle x_1, x_2, x_3 \rangle = \{x \in E, \exists \lambda_1, \lambda_2, \lambda_3 \in K, x = \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3\}.$$

On dit que (x_1, x_2, x_3) ***engendre*** *F ou que c'est une* ***famille génératrice*** *de F.*

*Si les vecteurs x_1, x_2, x_3 sont deux à deux distincts, on dit encore que la partie $\{x_1, x_2, x_3\}$ est une **partie génératrice** de F.*

On laisse au lecteur le soin d'aménager un texte analogue dans le cas d'un vecteur, d'un couple, d'un quadruplet, etc...

Démonstration :

1. $F \neq \oslash$ car pour $\lambda_1 = \lambda_2 = \lambda_3 = 0$ on trouve $0 \in F$. De plus, soient $x, y \in F$, alors $x = \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3$ et $y = \mu_1 x_1 + \mu_2 x_2 + \mu_3 x_3$, donc pour tous $\lambda, \mu \in K$ on a $\lambda x + \mu y = (\lambda\lambda_1 + \mu\mu_1)x_1 + (\lambda\lambda_2 + \mu\mu_2)x_2 + (\lambda\lambda_3 + \mu\mu_3)x_3 \in F$
2. F contient x_1, en prenant $\lambda_1 = 1$ et $\lambda_2 = \lambda_3 = 0$; idem pour x_2 et x_3
3. soit F' un sous-espace vectoriel de E contenant x_1, x_2, x_3, alors F' est stable pour les deux lois et donc, pour tous scalaires λ_1, λ_2, λ_3 on a $\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 \in F'$. F est donc bien le plus petit possible pour l'inclusion⊡

Exemples :

1. Soient $E = R^3$, $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$, alors (e_1, e_2, e_3) engendre R^3 mais (e_1, e_2) engendre un sous-espace vectoriel strict de R^3.
2. Soit $x \in E$, $x \neq 0$, alors $\langle x \rangle = \{\lambda x, \lambda \in K\}$ s'appelle une droite vectorielle de E.

Exercices :

1. Soient $E = R^3$ et $F = \{(x, y, z) \in R^3, 2x - y + 2z = 0\}$
   (a) Montrer que F est un sous-espace vectoriel de E, trouver un couple générateur de F. Y'a-t-il unicité d'un tel couple ?
   (b) Soit $X = (2, 2, -1)$, montrer que $X \in F$ mais que $\langle X \rangle \subsetneq F$.
   (c) On choisit un couple (X_1, X_2) générateur de F, y'a-t-il unicité de la décomposition $X = \lambda_1 X_1 + \lambda_2 X_2$?
2. Soit $E = K[X]$, montrer que
   (a) $(1 + 2X^2) \in \langle 1 - 2X, X^2 + X, X^3 \rangle$
   (b) $X^3 \in \langle 1, X, X^2 \rangle$
   (c) $X^3 \in \langle 1, X - a, (X - a)^2, (X - a)^3 \rangle$, $a \in K$ donné.

D'une part, il est évident que si x_4 est un nouvel individu de E, on a $\langle x_1, x_2, x_3, x_4 \rangle \subset \langle x_1, x_2, x_3, x_4 \rangle$. D'autre part, on a vu dans les exercices précédents que cette inclusion pouvait être stricte ou large.

PROPOSITION 6.1 *On a l'équivalence suivante :*

$$x_4 \in \langle x_1, x_2, x_3 \rangle \Leftrightarrow \langle x_1, x_2, x_3 \rangle = \langle x_1, x_2, x_3, x_4 \rangle .$$

Démonstration : ($\Leftarrow$) évident, puisque $x_4 \in \langle x_1, x_2, x_3, x_4 \rangle = \langle x_1, x_2, x_3 \rangle$.

($\Rightarrow$) on suppose que l'on a $x_4 = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3$, alors pour tout vecteur x appartenant à $\langle x_1, x_2, x_3, x_4 \rangle$ on peut écrire :

$$\begin{aligned} x &= \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 + \lambda_4 x_4 \\ &= (\lambda_1 + \lambda_4 \alpha_1) x_1 + (\lambda_2 + \lambda_4 \alpha_2) x_2 + (\lambda_3 + \lambda_4 \alpha_3) x_3 \boxdot \end{aligned}$$

Plus généralement, soient x_1, ..., x_n des vecteurs de E ; chaque fois que dans le lot il en existe un qui est combinaison linéaire des autres, on peut l'ôter sans dommage pour le sous-espace vectoriel F engendré par $(x_1, ..., x_n)$. Lorsque cela ne sera plus possible, on disposera d'une famille génératrice "minimale". A une telle famille on donnera bientôt un nom.

6.1.2 Indépendance linéaire

DEFINITION 6.2 *On appelle* ***relation de dépendance linéaire*** *entre les vecteurs x_1, x_2, x_3 de E tout triplet $(\lambda_1, \lambda_2, \lambda_3)$ de scalaires tel que $\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 = 0$. Le triplet $(0, 0, 0)$ est toujours une relation de dépendance linéaire, appelée* ***relation triviale****.*

S'il existe une relation de dépendance linéaire non triviale, on dit que le triplet (x_1, x_2, x_3) est ***lié****, ou que les vecteurs x_1, x_2, x_3 sont* ***linéairement dépendants****. Dans le cas contraire, on dit que le triplet est* ***libre*** *ou que les vecteurs sont* ***linéairement indépendants****.*

En résumé, étudier l'indépendance linéaire des vecteurs x_1, x_2, x_3 revient à résoudre l'équation suivante en λ_1, λ_2, λ_3 :

$$\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 = 0.$$

Si la seule solution est la solution triviale $(0, 0, 0)$ le triplet est libre, sinon il est lié.

PROPOSITION 6.2 *(x_1, x_2, x_3) est lié si et seulement si l'un au moins des vecteurs x_1, x_2, x_3 est combinaison linéaire des deux autres.*

Démonstration : on suppose (x_1, x_2, x_3) lié, soit $\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 = 0$ une combinaison linéaire nulle, non triviale, alors un au moins des coefficients est non nul, par exemple $\lambda_3 \neq 0$ (sans perte de généralité). Dès lors, λ_3 est inversible dans K et on a :

$$x_3 = (-\lambda_3^{-1} \lambda_1) x_1 + (-\lambda_3^{-1} \lambda_2) x_2 \in \langle x_1, x_2 \rangle .$$

Réciproquement, si x_3 est combinaison linéaire de x_1 et x_2, on peut écrire $x_3 = \alpha x_1 + \beta x_2$ et alors $(\alpha, \beta, -1)$ est une relation de dépendance non triviale entre x_1, x_2, x_3⊡

Exercices :

1. Montrer que si (x_1, x_2, x_3) est un triplet libre de E, il en est de même du triplet $(x_1 + x_2, x_2 + x_3, x_3 + x_1)$.
2. Etudier l'indépendance des triplets suivants de R^2 ou de R^3 et donner, s'il y a lieu, les relations de dépendance non triviales :

$$((1,3),(4,-1),(-5,1))$$
$$((1,8,1),(-9,-3,1),(2,2,-3))$$
$$((2,3,-1),(8,8,1),(4,2,3)).$$

6.1.3 Lien entre famille libre et famille génératrice

PROPOSITION 6.3 *Soit (x_1, x_2, x_3) un triplet de E, ce triplet est libre si et seulement si aucun des vecteurs du triplet n'est combinaison linéaire des deux autres.*

Démonstration : c'est la contraposée de la proposition précédente.

PROPOSITION 6.4 *Soient (x_1, x_2, x_3) un triplet libre de E et x un vecteur de E, alors le quadruplet (x_1, x_2, x_3, x) est lié si et seulement si $x \in \langle x_1, x_2, x_3 \rangle$.*

Démonstration : soit $(\lambda_1, \lambda_2, \lambda_3, \lambda)$ une relation de dépendance non triviale entre les vecteurs x_1, x_2, x_3, x, alors $\lambda \neq 0$ car sinon le triplet (x_1, x_2, x_3) serait lié, donc λ est inversible et l'on a :

$$x = (-\lambda^{-1}\lambda_1)x_1 + (-\lambda^{-1}\lambda_2)x_2 + (-\lambda^{-1}\lambda_3)x_3.$$

La réciproque est évidente⊡

PROPOSITION 6.5 *Soit (x_1, x_2, x_3) un triplet de vecteurs de E, les énoncés suivants sont équivalents :*

1. *(x_1, x_2, x_3) est libre*
2. *tout vecteur x appartenant à $\langle x_1, x_2, x_3 \rangle$ admet une unique décomposition de la forme :*
$$x = \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3.$$

Démonstration : (1.⇒2.) soient $x = \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 = \mu_1 x_1 + \mu_2 x_2 + \mu_3 x_3$ deux décompositions de x, par soustraction on obtient :

$$(\lambda_1 - \mu_1)x_1 + (\lambda_2 - \mu_2)x_2 + (\lambda_3 - \mu_3)x_3 = 0.$$

Mais (x_1, x_2, x_3) est libre, on en déduit donc $\lambda_1 = \mu_1$, $\lambda_2 = \mu_2$, $\lambda_3 = \mu_3$. Les deux décompositions sont donc nécessairement identiques.

(2.⇒1.) en effet, 1. ne fait qu'exprimer 2. pour $x = 0$□

DEFINITION 6.3 *Soient (x_1, x_2, x_3) un triplet libre de vecteurs de E et $F = \langle x_1, x_2, x_3 \rangle$, (x_1, x_2, x_3) s'appelle une **base** de F. C'est donc un triplet libre engendrant F.*

D'après ce qu'on a vu précédemment, si (x_1, x_2, x_3) est une base de F, $\langle x_1, x_2 \rangle$ n'engendre pas F. Sinon, $x_3 \in \langle x_1, x_2 \rangle$ et le triplet initial ne pourrait pas être libre. Il en va de même pour $\langle x_1, x_3 \rangle$ et $\langle x_2, x_3 \rangle$. Une base apparaît donc comme une famille génératrice minimale.

Une base de F apparaît aussi comme une famille libre maximale car une surfamille stricte d'une famille génératrice est nécessairement liée.

6.2 Introduction

On désire généraliser les notions précédentes. On pourrait penser se limiter à des familles finies, mais ceci s'avèrerait très rapidement insuffisant. Se pose alors immédiatement un problème : on a défini la somme de deux vecteurs, et donc par itération la somme d'un nombre fini de vecteurs. Mais en aucun cas on ne peut parler ici de la somme d'une infinité de vecteurs, les définitions doivent donc être aménagées de façon à ne rencontrer toujours que des sommes finies.

Par ailleurs, on aura soin de ne pas confondre famille et partie. La différence est exactement la même qu'entre arrangement avec répétition et combinaison sans répétition. Toutefois, on remarque que si $(x_i)_{i \in I}$ est une famille, on peut considérer son ensemble image qu'on note $\{x_i,\, i \in I\}$. Réciproquement, si A est une partie, on peut lui associer la famille "identité sur A", en considérant A lui-même comme ensemble d'indices. On utilisera souvent, par la suite, ce procédé permettant de passer de la notion de partie à celle de famille, et inversement.

DEFINITION 6.4 *On dit qu'une famille $(\lambda_i)_{i \in I}$ de scalaires est à **support fini** si l'ensemble des indices i pour lesquels λ_i est non nul est fini. On dit aussi que les λ_i sont **presque tous nuls**.*

Bien évidemment, si I est un ensemble fini, cette définition perd tout intérêt. De plus, la famille $(\lambda_i)_{i \in I}$, où pour tout i $\lambda_i = 0$, est à support fini. On l'appelle la famille triviale (indexée par I).

PROPOSITION 6.6 *Soit I un ensemble d'indices quelconque, l'ensemble des familles de scalaires indexées par I à support fini est un sous-espace vectoriel du K-espace vectoriel A(I, K). On le note $K^{(I)}$.*

Démonstration : on sait que A(I, K) est un K-espace vectoriel pour les lois usuelles, il suffit donc de vérifier la stabilité de $K^{(I)}$ par combinaison linéaire de deux éléments. Or si $(\lambda_i)_{i\in I}$ et $(\mu_i)_{i\in I}$ sont deux familles à support fini, pour tous λ, μ éléments de K $(\lambda\lambda_i+\mu\mu_i)_{i\in I}$ est clairement à support fini, son support étant inclus dans la réunion des deux supports (i.e. l'ensemble des indices pour lesquels le scalaire associé est non nul) de $(\lambda_i)_{i\in I}$ et $(\mu_i)_{i\in I}\boxdot$

DEFINITION 6.5 *Soit $(x_i)_{i\in I}$ une famille de vecteurs d'un K-espace vectoriel E, on dit qu'un vecteur x de E est* ***combinaison linéaire de la famille*** *$(x_i)_{i\in I}$ s'il existe une famille $(\lambda_i)_{i\in I}$ de scalaires, à support fini, telle que :*

$$x = \sum_{i\in I} \lambda_i x_i.$$

La famille $(\lambda_i)_{i\in I}$ étant à support fini, il n'y a qu'un nombre fini de termes $\lambda_i x_i$ non nuls. On convient alors que le symbole $\sum_{i\in I}$ désigne la somme de ces termes non nuls. Il s'agit donc bien, en fait, d'une somme finie.

DEFINITION 6.6 *Soit $(x_i)_{i\in I}$ une famille de vecteurs d'un K-espace vectoriel E, on appelle* ***relation linéaire*** *entre les vecteurs de cette famille toute famille de scalaires $(\lambda_i)_{i\in I}$ à support fini vérifiant :*

$$\sum_{i\in I} \lambda_i x_i = 0.$$

Une relation linéaire entre les x_i est donc un élément de $K^{(I)}$. La famille triviale de scalaires est toujours une relation linéaire entre les x_i. On l'appelle la relation triviale.

Exercice : Montrer que l'ensemble des relations linéaires entre les vecteurs de la famille $(x_i)_{i\in I}$ est un sous-espace vectoriel de $K^{(I)}$.

On rappelle enfin que l'on nomme **sous-famille** (resp. **sur-famille**) d'une famille donnée toute restriction (resp. prolongement) de la famille en question. Enfin, dans le cas où I est fini, on appelle cardinal de la famille $(x_i)_{i\in I}$ le cardinal de I.

6.3 Génération

La technique décrite est standard. Elle fonctionne chaque fois que l'on a une stabilité par intersection et elle définit ce que l'on appelle parfois une fermeture de Moore.

THEOREM 6.2 *Soit $(E_i)_{i\in I}$ une famille de sous-espaces vectoriels du K-espace vectoriel E, l'intersection de la famille $(E_i)_{i\in I}$ est un sous-espace vectoriel de E.*

Démonstration : $\cap_{i\in I}$ est non vide puisque 0 appartient à tous les E_i. De plus :

$$\forall x, y \in \cap_{i\in I} E_i, \forall \lambda, \mu \in K, \lambda x + \mu y \in \cap_{i\in I} E_i.$$

En effet, $\forall i \in I$, $x, y \in E_i$, or E_i est un sous-espace vectoriel de E, donc $\forall i \in I$, $\lambda x + \mu y \in E_i \boxdot$

THEOREM 6.3 *Soit A une partie de E, l'ensemble des sous-espaces vectoriels de E contenant A admet un plus petit élément (pour l'inclusion). On l'appelle le sous-espace vectoriel* ***engendré par*** *A, il se note* $\langle A\rangle$ *ou encore Vect(A).*

Démonstration : d'après le théorème précédent, $\langle A\rangle$ n'est autre que l'intersection de tous les sous-espaces vectoriels de E contenant A, cette famille n'étant pas vide puisqu'elle contient au moins E$\boxdot$

Exemples :

1. Si F est un sous-espace vectoriel de E, on a $\langle F\rangle = F$
2. $\langle \oslash\rangle = \{0\}$, ce qui est un résultat important
3. Si E_1 et E_2 sont deux sous-espaces vectoriels de E, on a $E_1 + E_2 = \langle E_1 \cup E_2\rangle$.

Exercice :

1. Dans R considéré comme R-espace vectoriel, quel est le sous-espace vectoriel engendré par $\{1\}$? Plus généralement, soit A une partie de R, déterminer $\langle A\rangle$: on distinguera trois cas
2. Soient A une partie du K-espace vectoriel E et B une partie de E telle que $A \subset B \subset \langle A\rangle$, montrer que $\langle B\rangle = \langle A\rangle$
3. Soient A et B deux parties du K-espace vectoriel E, montrer que :

$$A \subset B \Rightarrow \langle A\rangle \subset \langle B\rangle .$$

Réciproque ?

DEFINITION 6.7 *Soit A une partie du K-espace vectoriel E, si* $\langle A\rangle = E$ *on dit que A est une* ***partie génératrice*** *de E.*

En particulier, E est une partie génératrice de E. Toutefois, le lecteur sera vite convaincu que les parties génératrices intéressantes sont celles qui sont "les plus petites possibles". On précisera bientôt ce que l'on entend par là, mais pour l'instant, on va étendre la notion de génération aux familles. Ceci n'est pas actuellement indispensable, mais le deviendra ultérieurement.

DEFINITION 6.8 *1. Soit* $(x_i)_{i\in I}$ *une famille d'un K-espace vectoriel E, on appelle sous-espace vectoriel* ***engendré par la famille*** $(x_i)_{i\in I}$ *le sous-espace vectoriel engendré par l'image* $\{x_i, i \in I\}$ *de cette famille*

2. *On dit que la famille $(x_i)_{i\in I}$ est une* ***famille génératrice*** *de E si le sous-espace vectoriel engendré par elle est E.*

On remarque que le sous-espace vectoriel engendré par une partie A coïncide avec le sous-espace vectoriel engendré par la famille canoniquement associée à A, ce qui fait le lien entre les deux notions.

Le lecteur a sûrement remarqué que lors des préliminaires on a défini le sous-espace vectoriel engendré par une famille en termes de combinaisons linéaires. On doit donc faire le lien entre ces deux définitions.

THEOREM 6.4 *Soit $(x_i)_{i\in I}$ une famille d'un K-espace vectoriel E, le sous-espace vectoriel engendré par cette famille est l'ensemble des combinaisons linéaires de la famille $(x_i)_{i\in I}$.*

Démonstration : on doit donc montrer que l'ensemble des combinaisons linéaires de la famille donnée est un sous-espace vectoriel de E et que ce sous-espace vectoriel est le plus petit contenant tous les vecteurs x_i. Tout d'abord, pour éviter un cas particulier, on convient que si $I = \oslash$, alors $\sum_{i\in\oslash} x_i = 0$. L'ensemble des combinaisons linéaires de la famille $(x_i)_{i\in I}$ contient donc toujours le vecteur nul : si $I \neq \oslash$, on prend la famille triviale de scalaires. Soient alors x, y deux combinaisons linéaires de la famille $(x_i)_{i\in I}$ et λ, μ deux scalaires quelconques, on a donc $x = \sum_{i\in I} \lambda_i x_i$, $y = \sum_{i\in I} \mu_i x_i$, $(\lambda_i)_{i\in I}$ et $(\mu_i)_{i\in I}$ à support fini, d'où $\lambda x + \mu y = \sum_{i\in I}(\lambda\lambda_i + \mu\mu_i)x_i$ et la famille $(\lambda\lambda_i + \mu\mu_i)_{i\in I}$ est encore à support fini. L'ensemble des combinaisons linéaires de la famille donnée est donc un sous-espace vectoriel de E qui contient bien évidemment tous les x_i : on prend $\lambda_i = 1$ et si $j \neq i$, $\lambda_j = 0$. Enfin, si un sous-espace vectoriel de E contient tous les vecteurs de $(x_i)_{i\in I}$, par stabilité, il contient toutes les combinaisons linéaires de cette famille, ce qui achève la preuve$\boxdot$

Le sous-espace vectoriel engendré par une partie A d'un K-espace vectoriel E est donc l'ensemble des combinaisons linéaires des éléments de A. Ainsi, $\langle\oslash\rangle = \{0\}$ et si $A \neq \oslash$:

$$\langle A\rangle = \{x \in E, \exists n \in N^*, \exists\lambda_1, ..., \lambda_n \in K, \exists x_1, ..., x_n \in A, x = \sum_{i=1}^{n} \lambda_i x_i\}.$$

Exercice : Soit $(x_i)_{i\in I}$ une famille de vecteurs du K-espace vectoriel E, on considère l'application :

$$\Phi : \begin{array}{l} K^{(I)} \to E \\ (\alpha_i)_{i\in I} \mapsto \sum_{i\in I} \alpha_i x_i \end{array}.$$

1. Montrer que Φ est bien définie et que Φ est linéaire

2. Montrer que Φ est surjective si et seulement si $(x_i)_{i\in I}$ est une famille génératrice de E.

Le théorème précédent est d'une très grande importance. En effet, il affirme que la connaissance d'une famille (ou d'une partie) génératrice de E permet de reconstituer l'intégralité de l'espace. Ceci explique la remarque : une partie génératrice est d'autant plus intéressante qu'elle comporte "le moins possible" d'éléments. On cite donc la proposition, a priori sans intérêt, suivante.

PROPOSITION 6.7 *Soit E un K-espace vectoriel, toute partie contenant une partie génératrice de E est encore une partie génératrice de E. De même, toute sur-famille d'une famille génératrice de E est encore une famille génératrice de E.*

Démonstration : évident

Le théorème suivant est en revanche du plus haut intérêt, car il permet d'ôter des éléments d'une famille génératrice.

THEOREM 6.5 *Soient $(x_i)_{i\in I}$ une famille génératrice du K-espace vectoriel E, $i_0 \in I$ et $J = I\backslash\{i_0\}$, la famille $(x_i)_{i\in J}$ est une famille génératrice de E si et seulement si x_{i_0} est combinaison linéaire de la famille $(x_i)_{i\in J}$.*

Démonstration : si $(x_i)_{i\in J}$ est génératrice, tout vecteur de E est combinaison linéaire de cette famille et donc en particulier x_{i_0}. Réciproquement, on suppose qu'il existe une famille de scalaires $(\alpha_i)_{i\in J}$, à support fini, telle que $x_{i_0} = \sum_{i\in J} \alpha_i x_i$. Soit alors x un vecteur quelconque, il existe donc une famille de scalaires $(\lambda_i)_{i\in I}$, à support fini, telle que $x = \sum_{i\in I} \lambda_i x_i = \lambda_{i_0} x_{i_0} + \sum_{i\in J} \lambda_i x_i = \sum_{i\in J} (\lambda_i + \lambda_{i_0}\alpha_i) x_i$ et la famille $(\lambda_i + \lambda_{i_0}\alpha_i)_{i\in J}$ est encore à support fini. $(x_i)_{i\in J}$ est donc bien une famille génératrice de E$\boxdot$

On cite enfin, pour mémoire, une propriété déjà vue des combinaisons linéaires, généralement appelée "transitivité".

PROPOSITION 6.8 *Soient A, B, C trois parties d'un K-espace vectoriel E, on a :*

$$A \subset \langle B\rangle \text{ et } B \subset \langle C\rangle \Rightarrow A \subset \langle C\rangle .$$

Démonstration : évident

6.4 Liberté

D'après les considérations précédentes, une famille génératrice idéale de E serait une famille telle que toutes ses sous-familles strictes ne soient plus génératrices.

Une telle famille génératrice est dite minimale. Mais on ne sait pas encore si de tels êtres existent toujours, car le théorème ? ? ne permet d'éliminer des éléments qu'un par un (à méditer).

PROPOSITION 6.9 *Soit $(x_i)_{i\in I}$ une famille génératrice minimale, non vide, d'un K-espace vectoriel E, la seule relation linéaire entre les vecteurs de cette famille est la relation triviale.*

Démonstration : si $(\alpha_i)_{i\in I}$ était une relation linéaire non triviale, il existerait un indice i_0 tel que $\alpha_{i_0} \neq 0$ et donc :

$$\sum_{i\in I} \alpha_i x_i = 0 \Rightarrow x_{i_0} = \sum_{i\neq i_0} -\frac{\alpha_i}{\alpha_{i_0}} x_i$$

ce qui contredirait la minimalité de la famille $(x_i)_{i\in I}$ d'après le théorème ? ?□

On est ainsi amené à donner un nom aux familles possédant cette propriété.

DEFINITION 6.9 *Soit $(x_i)_{i\in I}$ une famille de vecteurs d'un K-espace vectoriel E, on dit que cette famille est **libre** si la seule relation linéaire entre les x_i est la relation triviale, i.e. :*

$$\forall (\alpha_i)_{i\in I} \in K^{(I)}, \sum_{i\in I} \alpha_i x_i = 0 \Rightarrow (\forall i \in I, \alpha_i = 0).$$

*On dit aussi que la famille est formée de vecteurs **linéairement indépendants**. Dans le cas contraire, la famille est dite **liée** ou formée de vecteurs **linéairement dépendants**.*

DEFINITION 6.10 *Une partie A de E est dite **libre** si la famille canoniquement associée à A est libre, dans le cas contraire elle est dite **liée**.*

Exemples :

1. Dans R^2, $\{(1, 0), (0, 1)\}$ est libre.
2. Si $x \in E$ et si $x \neq 0$, alors $\{\mathrm{x}\}$ est libre.
3. Si une partie A de E contient le vecteur nul, alors A est liée.

On cite la proposition suivante, du même ordre d'intérêt que la proposition ? ?. La démonstration est laissée en exercice au lecteur.

PROPOSITION 6.10 *Si A est une partie libre d'un K-espace vectoriel E et si $B \subset A$, alors B est libre. De même, toute sous-famille d'une famille libre est libre.*

Les familles libres ont été introduites comme des familles économiques, du point de vue de la génération. Il se trouve qu'elles ont du même coup une propriété (presque) inattendue.

PROPOSITION 6.11 *Soit $(x_i)_{i\in I}$ une famille de vecteurs d'un K-espace vectoriel E, les deux énoncés suivants sont équivalents :*

1. $(x_i)_{i\in I}$ est une famille libre
2. $\forall x \in \langle (x_i)_{i\in I} \rangle$, $\exists !(\alpha_i)_{i\in I} \in K^{(I)}$, $x = \sum_{i\in I} \alpha_i x_i$.

L'important ici est le point d'exclamation (unicité) !

Démonstration : (1.⇒2.) on suppose que $x = \sum_{i\in I} \alpha_i x_i = \sum_{i\in I} \beta_i x_i$ où $(\alpha_i)_{i\in I}$ et $(\beta_i)_{i\in I}$ sont deux familles de scalaires à support fini. Par soustraction, on obtient $\sum_{i\in I}(\alpha_i - \beta_i)x_i = 0$, donc $(\alpha_i - \beta_i)_{i\in I}$ est une relation linéaire entre les x_i. La famille étant libre, cette relation est la relation triviale et la décomposition de x est bien unique.

(2.⇒1.) par hypothèse, le vecteur nul n'admet qu'une décomposition suivant les vecteurs x_i. La relation triviale fournit une décomposition évidente. La relation triviale est donc la seule relation linéaire entre les x_i et la famille donnée est bien libre⊡

On a considéré jusqu'à présent le cas le plus général. On va maintenant retranscrire les résultats dans le cas particulier des familles finies et plus particulièrement en prenant $I = [1, n]$, afin de retomber sur les préliminaires. On fera l'abus classique qui consiste à identifier les familles indexées par [1, n] et les n-uplets. On remarque tout de suite que la notion de support fini disparaît.

Transcriptions :

1. Soit $(x_1, ..., x_n)$ un n-uplet de vecteurs du K-espace vectoriel E, on appelle relation linéaire entre les vecteurs x_1, ..., x_n tout n-uplet de scalaires $(\alpha_1, ..., \alpha_n)$ tel que $\sum_{i=1}^{n} \alpha_i x_i = 0$. Le n-uplet $(0, ..., 0)$ est une relation linéaire dite relation triviale.
2. Le n-uplet $(x_1, ..., x_n)$ est dit libre si la seule relation linéaire entre les x_i est la relation triviale, i.e. :

$$\forall \alpha_1, ..., \alpha_n \in K, \alpha_1 x_1 + ... + \alpha_n x_n = 0 \Rightarrow \alpha_1 = ... = \alpha_n = 0.$$

Dans le cas contraire, le n-uplet est dit lié.

Ces transcriptions ne sont pas dénuées d'intérêt, même dans le cas général. En fait, le cas où I est fini suffit, comme le prouve le théorème suivant, très important dans la pratique.

THEOREM 6.6 *Soit $(x_i)_{i\in I}$ une famille de vecteurs d'un K-espace vectoriel E, les deux propositions suivantes sont équivalentes :*

1. $(x_i)_{i\in I}$ est une famille libre

2. *toute sous-famille finie de* $(x_i)_{i\in I}$ *est libre.*

Démonstration : (1.⇒2.) prouvé à la proposition ? ?

(2.⇒1.) soit $(\alpha_i)_{i\in I}$ une relation linéaire entre les x_i. Cette famille est donc à support fini. Soit J le support de cette famille, on a alors :

$$0 = \sum_{i\in I} \alpha_i x_i = \sum_{i\in J} \alpha_i x_i$$

par définition même de ce symbole. Mais J est fini, donc par hypothèse $(x_i)_{i\in J}$ est libre. On en déduit que $\forall i \in J$, $\alpha_i = 0$ et par définition du support on a $\forall i \in I$, $\alpha_i = 0$. La famille initiale est bien libre□

Exercices :

1. Montrer que la famille $(X^n)_{n\in N}$ est libre dans K[X].
2. Soit f_a l'application réelle définie sur R par $f_a(x) = |x - a|$, montrer que la famille $(f_a)_{a\in R}$ est libre dans $A(R, R)$.
3. Soit g_a l'application réelle définie sur R par $g(x) = e^{ax}$, montrer que la famille $(g_a)_{a\in R}$ est libre dans $A(R, R)$. On procèdera comme précédemment et on pensera aux équivalents au voisinage de l'infini.
4. Pour tout entier strictement positif p, soient f_p et g_p les applications réelles définies sur R par $f_p(x) = \sin(px)$ et $g_p(x) = \cos(px)$, montrer que pour tout entier strictement positif n la famille $(f_1, g_1, ..., f_n, g_n)$ est une famille libre de $A(R, R)$. On établira d'abord le résultat pour $n = 1$, puis on effectuera un raisonnement par récurrence en dérivant deux fois une combinaison linéaire nulle.

6.5 Base

Les deux sections précédentes font apparaître une race particulièrement intéressante de familles, celles qui sont à la fois libres et génératrices. On ne sait pas encore si de telles familles existent toujours (on sait trouver des "petites" familles libres, des "grosses" familles génératrices, mais entre les deux...). Cela n'interdit pas de poser la définition.

DEFINITION 6.11 *Soit E un K-espace vectoriel, on appelle* ***base*** *de E toute famille de vecteurs de E qui est à la fois libre et génératrice.*

DEFINITION 6.12 *On appelle* ***partie basique*** *de E toute partie de E qui est à la fois libre et génératrice.*

La famille canoniquement associée à une partie basique est une base et l'image d'une base (au sens de l'image d'une famille) est une partie basique.

On insiste bien sur le fait qu'une base est une famille et non une partie. Le lecteur s'en convaincra lors du calcul matriciel.

La nature est bien faite. *Tout espace vectoriel possède au moins une base*. La démonstration de ce fort beau résultat repose sur l'axiome du choix. On en donnera une démonstration "élémentaire" dans un cas particulier au chapitre suivant.

THEOREM 6.7 *(fondamental)*

Soit $(x_i)_{i\in I}$ une famille de vecteurs d'un K-espace vectoriel E, les propositions suivantes sont équivalentes :

1. *$(x_i)_{i\in I}$ est une base de E*
2. *$(x_i)_{i\in I}$ est une famille libre maximale (i.e. toute sur-famille stricte est liée)*
3. *$(x_i)_{i\in I}$ est une famille génératrice minimale (i.e. toute sous-famille stricte n'est plus génératrice).*

Démonstration : (1.⇒2.) on sait déjà que $(x_i)_{i\in I}$ est libre, il reste à montrer qu'elle est maximale. Soient $I \subsetneq J$ et $(x_i)_{i\in J}$ une sur-famille stricte de la famille initiale, pour $j \in J\backslash I$ x_j est combinaison linéaire de la famille $(x_i)_{i\in I}$, puisque cette famille est génératrice. Il existe donc une relation linéaire non triviale entre les x_i, $i \in J$. La sur-famille n'est donc pas libre.

(2.⇒3.) on montre d'abord que $(x_i)_{i\in I}$ est génératrice. Soient $x \in E$ et j un indice n'appartenant pas à I, on considère la famille indexée par $I \cup \{j\}$ et définie par $\forall i \in I$, $i \mapsto x_i$ et $j \mapsto x$. C'est une sur-famille de la famille initiale, elle est donc liée. Il existe donc des scalaires presque tous nuls, mais non tous nuls, tels que :

$$\lambda x + \sum_{i\in I} \lambda_i x_i = 0.$$

λ ne peut être nul car sinon la famille initiale serait liée par la relation linéaire $(\lambda_i)_{i\in I}$. x est donc combinaison linéaire de la famille initiale. Le reste est déjà vu, car si la famille n'était pas génératrice minimale, un des vecteurs au moins serait combinaison linéaire des autres (? ?) et la famille ne serait pas libre.

(3.⇒1.) déjà fait (? ?)□

L'intérêt des bases est illustré par la proposition suivante qui a déjà été démontrée (? ?).

PROPOSITION 6.12 *Soit $(x_i)_{i\in I}$ une base du K-espace vectoriel E, pour tout vecteur x de E il existe une famille unique de scalaires presque tous nuls $(\lambda_i)_{i\in I}$*

telle que :

$$x = \sum_{i \in I} \lambda_i x_i.$$

DEFINITION 6.13 *La famille $(\lambda_i)_{i \in I}$ s'appelle famille des **coordonnées** du vecteur x relativement à la base donnée et la famille $(\lambda_i x_i)_{i \in I}$ s'appelle famille des **composantes** du vecteur x relativement à la base donnée.*

Exemples :

1. Soient $e_1 = (1, 0)$ et $e_2 = (0, 1)$, (e_1, e_2) est une base de $\mathbb{R}^2$ appelée base canonique de $\mathbb{R}^2$.
2. Plus généralement, dans $\mathbb{R}^n$, on pose $\forall i \in [1, n]$, $e_i = (0, ..., 0, 1, 0, ..., 0)$, le 1 au i-ème rang. La famille $(e_1, ..., e_n)$ est une base de $\mathbb{R}^n$ appelée base canonique de $\mathbb{R}^n$.
3. $(X^n)_{n \in N}$ est une base de K[X] encore appelée base canonique de K[X].
4. Dans $\mathbb{C}$ considéré comme $\mathbb{R}$-espace vectoriel, $(1, i)$ est une base. Est-ce une base du $\mathbb{C}$-espace vectoriel ? Moralité ?

Exercice (important) : Soient E un K-espace vectoriel et E_1 et E_2 deux sous-espaces vectoriels supplémentaires, montrer que si A est une partie basique de E_1 et B une partie basique de E_2, alors $A \cup B$ est une partie basique de E.

Le lecteur qui s'étonnerait que l'on parle ici de parties basiques peut imaginer ce que serait une réunion de familles (on dit d'ailleurs concaténation) et se rendre compte que c'est peu évident à écrire, surtout si on commet l'imprudence de mettre des indices en commun.

Enfin, pour clore ce chapitre, il est naturel de se demander ce que deviennent les notions précédentes lorsqu'on les transforme par une application linéaire.

6.6 Familles et applications linéaires

PROPOSITION 6.13 *Soient E, F deux K-espaces vectoriels et $f \in \mathcal{L}(E, F)$:*

1. *si $(x_i)_{i \in I}$ est une famille liée de vecteurs de E, la famille $(f(x_i))_{i \in I}$ est une famille liée de vecteurs de F*
2. *si $(x_i)_{i \in I}$ est une famille génératrice de E, la famille $(f(x_i))_{i \in I}$ est une famille génératrice de Im(f).*

Démonstration : trivial

Le lecteur remarquera bien que l'image d'une partie génératrice de E n'est pas, en général, génératrice de F, mais seulement de Im(f) : bien sûr, si f est surjective...

PROPOSITION 6.14 *Soient E, F deux K-espaces vectoriels et $f \in \mathcal{L}(E, F)$, les propositions suivantes sont équivalentes :*

1. *f est injective*
2. *l'image par f de toute famille libre de E est une famille libre de F.*

Démonstration : (1.⇒2.) soient $(x_i)_{i\in I}$ une famille libre de E et $(\lambda_i)_{i\in I}$ une relation linéaire entre les vecteurs de la famille $(f(x_i))_{i\in I}$, on a donc $\sum_{i\in I} \lambda_i f(x_i) = 0$, ou encore $f(\sum_{i\in I} \lambda_i x_i) = 0$, car f est linéaire et la famille de scalaires à support fini. Or f est injective, on en déduit que $\sum_{i\in I} \lambda_i x_i = 0$. La famille $(x_i)_{i\in I}$ étant libre, la famille de scalaires est donc la famille triviale. $(f(x_i))_{i\in I}$ est bien une famille libre.

(2.⇒1.) soit $x \in Ker(f)$, l'image de la famille (x) est la famille (0) qui est une famille liée, la famille (x) est donc liée, i.e. $x = 0$ (raisonnement par contraposée)□

THEOREM 6.8 *Soient E, F deux K-espaces vectoriels et $f \in \mathcal{L}(E, F)$, les propositions suivantes sont équivalentes :*

1. *f est un isomorphisme de E sur F*
2. *l'image par f de toute base de E est une base de F*
3. *il existe une base de E dont l'image par f est une base de F.*

Avant de démontrer ce théorème, on remarque la grande différence entre 2. et 3. : en fait, le théorème affirme qu'il suffit que f envoie une base de E sur une base de F pour que ce résultat soit valable pour toute base de F.

Démonstration : (1.⇒2.) cela découle directement de ? ? et ? ?, car un isomorphisme est à la fois injectif et surjectif.

(2.⇒3.) on a admis l'existence d'au moins une base, on laisse le lecteur conclure.

(3.⇒1.) soit $(x_i)_{i\in I}$ une base de E dont l'image par f est une base de F, on montre que f est injective puis surjective. Soit $x \in Ker(f)$, on décompose x sur la base $(x_i)_{i\in I}$ de E. On a $x = \sum_{i\in I} \lambda_i x_i$ pour une famille de scalaires presque tous nuls, et $f(x) = \sum_{i\in I} \lambda_i f(x_i) = 0$. La famille $(f(x_i))_{i\in I}$ étant libre, la famille $(\lambda_i)_{i\in I}$ est la famille triviale, i.e. $x = 0$ et f est bien injective. Soit $y \in F$, y se décompose sur la base $(f(x_i))_{i\in I}$ de F : $y = \sum_{i\in I} \lambda_i f(x_i)$. Il est alors clair que $y = f(x)$ avec $x = \sum_{i\in I} \lambda_i x_i$, donc f est bien surjective□

On finit par un théorème très important dans la pratique et qui assure de la connaissance complète d'une application linéaire dès que l'on connait les images des vecteurs d'une base de l'espace de départ.

THEOREM 6.9 *Soient E, F deux K-espaces vectoriels,* $(x_i)_{i\in I}$ *une base de E et* $(y_i)_{i\in I}$ *une famille quelconque de vecteurs de F indexée par le même ensemble d'indices, il existe une* ***unique*** *application linéaire f de E dans F vérifiant :*

$$\forall i \in I, f(x_i) = y_i.$$

Démonstration : soit x un vecteur quelconque de E, il existe une unique famille de scalaires $(\lambda_i)_{i\in I}$, presque tous nuls, telle que $x = \sum_{i\in I} \lambda_i x_i$. Si l'on veut que f soit linéaire, on n'a pas le choix et on est obligés de poser $f(x) = \sum_{i\in I} \lambda_i y_i$. On laisse le lecteur vérifier que f ainsi définie est bien linéaire⊡

En d'autres termes, les coordonnées d'une combinaison linéaire sont les combinaisons linéaires des coordonnées respectives.

Chapitre 7

Dimension finie

On va maintenant redescendre de quelques marches et se limiter à la catégorie des espaces vectoriels admettant une famille génératrice finie. On adaptera donc les résultats donnés dans les deux chapitres précédents. Les problèmes sont un peu plus simples et, en particulier, on pourra s'abstenir d'utiliser l'axiome du choix. En contrepartie, les démonstrations deviennent accessibles et seront par conséquent effectuées.

7.1 Espace vectoriel de dimension finie

DEFINITION 7.1 *On appelle espace vectoriel de **dimension finie** tout espace vectoriel possédant une partie génératrice finie. Dans le cas contraire, on dit que l'espace vectoriel est de **dimension infinie**.*

On note que, pour l'instant, le mot "dimension" n'a en lui-même aucune signification. Seule la locution "dimension finie" en a une. Pour cette raison (et pour d'autres), certains auteurs appellent parfois de tels espaces des espaces vectoriels de type fini. Ceci n'est pas indispensable, car le mot "dimension" va prendre très rapidement le sens que tout le monde lui connaît.

Il ne faut pas croire que tous les espaces vectoriels sur un corps K soient de dimension finie. L'exemple le plus simple est le suivant : K[X] est un K-espace vectoriel de dimension infinie. En effet, soit A une partie finie de K[X]. Si A est réduite à $\{0\}$ ou vide, A n'engendre que le polynôme nul et n'est donc pas génératrice. Sinon, l'ensemble des degrés des polynômes appartenant à A admet un plus grand élément, appelé n. D'après les propriétés du degré dans K[X], le degré de toute combinaison linéaire des éléments de A est inférieur ou égal à n. Par conséquent, X^{n+1} n'est pas combinaison linéaire des éléments de A et A n'est pas une partie génératrice de K[X]. K[X] n'admet donc aucune partie génératrice finie, il est par conséquent de dimension infinie.

Exercices :

1. Montrer que pour un K-espace vectoriel E les propriétés suivantes sont équivalentes :
   (a) E est de dimension finie
   (b) E admet une famille génératrice finie (i.e. indexée par un ensemble fini).
2. En reprenant les exemples d'espaces vectoriels classiques, indiquer ceux qui sont de dimension finie.
3. Soit E un K-espace vectoriel, comparer les deux énoncés :
   (a) E admet une partie génératrice finie
   (b) toutes les parties génératrices de E sont finies.

7.1.1 Existence des bases

On démontre dans ce paragraphe l'existence des bases en dimension finie. On profite de l'occasion pour rappeler que l'on a cité un résultat infiniment plus fort (malheureusement sans démonstration) : *tout K-espace vectoriel admet au moins une base*.

THEOREM 7.1 *Soient E un K-espace vectoriel de dimension finie, G une partie génératrice finie de E et L une partie libre contenue dans G, il existe alors une partie basique B de E telle que* $L \subset B \subset G$.

Avant de démontrer cet important résultat, on remarque que, comme la partie vide est une partie libre de n'importe quel espace vectoriel, il est loisible de considérer une partie libre L contenue dans G.

Démonstration : on a dans le chapitre précédent caractérisé les bases comme familles libres maximales. On peut de même caractériser les parties basiques comme parties libres maximales. Soit $\mathcal{L}$ l'ensemble des parties libres X de E vérifiant $L \subset X \subset G$, $\mathcal{L}$ est non vide (car $L \in \mathcal{L}$) et $\mathcal{L}$ est finie (car G est finie). De plus, tout élément de $\mathcal{L}$ a un nombre fini d'éléments. On choisit alors dans $\mathcal{L}$ une partie B ayant un nombre maximum d'éléments, on va montrer que cette partie B convient.

Par construction même de B, il est clair que B est une partie libre de E vérifiant $L \subset B \subset G$. Il reste à montrer que B est une partie génératrice de E. Soit donc $g \in G$, si $g \in B$ alors on a évidemment $g \in \langle B \rangle$. Si $g \notin B$, la partie $B \cup \{g\}$ est comprise entre L et G et comporte strictement plus d'éléments que B. La partie $B \cup \{g\}$ ne peut être libre par définition de B, elle est donc liée, i.e. $g \in \langle B \rangle$. Ainsi, on a prouvé que $G \subset \langle B \rangle$. Mais alors on sait, d'après ? ?, que $E = \langle G \rangle \subset \langle B \rangle$, i.e. $\langle B \rangle = E \boxdot$

En particularisant G ou L dans le théorème précédent, on obtient le théorème d'existence annoncé ainsi qu'un important corollaire.

THEOREM 7.2 *Tout espace vectoriel E de dimension finie sur un corps K admet au moins une base. Plus précisément, toute famille génératrice finie admet au moins une sous-famille qui est une base.*

Démonstration : il suffit de démontrer la seconde assertion. Soit $(x_i)_{i\in I}$ une famille génératrice finie de E, on sait que $G = \{x_i,\ i \in I\}$ est une partie génératrice finie de E. On prend $L = \oslash$ (qui est une partie libre) et on applique le théorème précédent à G et L. On obtient donc une partie basique B telle que $B \subset G$. On laisse au lecteur le soin de construire une partie J de l'ensemble d'indices I et une application de J dans B de telle sorte que la famille ainsi définie soit, d'une part une sous-famille de $(x_i)_{i\in I}$ et, d'autre part une base de E. Ce n'est pas difficile à imaginer, mais cela demande tout de même un peu de soin□

THEOREM 7.3 *(de la base incomplète)*

Dans un espace vectoriel E de dimension finie sur un corps K, toute famille libre peut être complétée en une base (i.e. admet une sur-famille qui est une base).

Démonstration : on admettra momentanément que dans un espace vectoriel de dimension finie **toutes** les parties libres sont finies (ce fait exceptionnel sera démontré à la section suivante). Soit donc $(x_i)_{i\in I}$ une famille libre (I est fini d'après ce qu'on vient de dire), la partie $L = \{x_i, i \in I\}$ est une partie libre de E. Soit alors G une partie génératrice finie de E, $G \cup L$ est encore une partie génératrice finie de E. On est donc dans la situation du théorème ? ? avec $L \subset G \cup L$. Il existe une partie basique B telle que $L \subset B \subset G \cup L$. Il reste au lecteur à fabriquer une application d'un ensemble J contenant I dans B de manière à avoir une sur-famille de $(x_i)_{i\in I}$ qui soit une base□

Exercices :

1. Soient dans R^3 les vecteurs $x = (2, 3, -1)$, $y = (1, -1, -2)$, $u = (3, 7, 0)$, $v = (5, 0, -7)$, montrer que $\{x, y\}$ est une partie libre et la compléter pour obtenir une partie basique de R^3. Montrer que le sous-espace vectoriel engendré par x et y est le même que le sous-espace vectoriel engendré par u et v. Conclusion ?
2. On considère $F = \{(a, b, c, d) \in R^4 \text{ / } a = b - 3c \text{ et } c = 2d\}$, trouver une base de F et la compléter pour obtenir une base de R^4.

7.1.2 Dimension

On vient de démontrer l'existence des bases pour un espace vectoriel E de dimension finie. Il se trouve, et c'est encore un miracle, que toutes les parties

basiques ont le même nombre d'éléments. On peut dire aussi que toutes les bases ont le même nombre d'éléments, sil'on convient d'appeler cardinal d'une base le cardinal de l'ensemble d'indices. Ce cardinal commun est donc un invariant de l'espace vectoriel E et prendra le nom de dimension de E, ce qui permettra de séparer les termes de la locution "dimension finie". La démonstration de cette propriété est difficile et repose sur un lemme dû à Steinitz.

(d'échange de Steinitz)

PROPOSITION 7.1 *Soient E un K-espace vectoriel, X une partie de E et x et y deux éléments de E tels que :*

1. $y \in \langle X \cup \{x\} \rangle$
2. *il existe une écriture de y comme combinaison linéaire des éléments de* $X \cup \{x\}$ *dans laquelle le coefficient de x est non nul,*

alors $x \in \langle X \cup \{y\} \rangle$.

On traduira la condition 2. par la locution "y peut utiliser x".

Démonstration : en effet, comme y peut utiliser x, on peut écrire $y = \sum_{i \in I} \lambda_i x_i + \lambda x$ où les x_i sont dans X, les λ_i presque tous nuls et λ un scalaire non nul. Il vient alors $x = \lambda^{-1} y - \sum_{i \in I} \lambda^{-1} \lambda_i x_i$ ce qui prouve que $x \in \langle X \cup \{y\} \rangle$ $\boxdot$

THEOREM 7.4 *Soient E un K-espace vectoriel, G une partie génératrice finie de E et L une partie libre de E, alors L est finie et on a* $Card(L) \leq Card(G)$.

Démonstration : on supposera $Card(L) > Card(G)$, le lemme de l'échange permettra de former de nouvelles parties génératrices de E en expulsant des vecteurs de G pour les remplacer par des vecteurs de L afin d'aboutir à une contradiction. On peut éliminer d'office le cas $G = \oslash$. On s'explique maintenant et on pose $G = \{g_1, ..., g_m\}$.

Soit $a_1 \in L$ un élément quelconque de L, on sait que a_1 est non nul car il appartient à une partie libre. Comme G est génératrice, on peut écrire $a_1 = \sum_{i=1}^{m} \lambda_i g_i$ pour un certain m-uplet de scalaires. Comme $a_1 \neq 0$, un au moins des λ_i n'est pas nul. On suppose qu'il s'agit de λ_1. Ainsi a_1 peut utiliser g_1, on peut appliquer alors le lemme d'échange à la situation $X = G \backslash \{g_1\}$, $x = g_1$, donc $g \in \langle (G \backslash \{g_1\}) \cup \{a_1\} \rangle$. On fait remarquer au lecteur que cette notation un peu lourde représente simplement l'ensemble obtenu à partir de G en remplaçant g_1 par a_1. On pose :

$$G_1 = (G \backslash \{g_1\}) \cup \{a_1\} = \{a_1, g_2, ..., g_m\}.$$

G_1 engendre alors encore E : il suffit de remarquer que $g_1 \in \langle G_1 \rangle$.

Soit a_2 un élément de L différent de a_1 (a_2 existe par hypothèse car $Card(L) > Card(G)$), comme G_1 engendre E on peut écrire $a_2 = \mu_1 a_1 + \lambda_2 g_2 + ... + \lambda_m g_m$. Les scalaires λ_2, ..., λ_m ne peuvent être tous nuls, car sinon on aurait $a_2 = \mu_1 a_1$ ce qui n'est pas le cas car a_1 et a_2 sont indépendants. Ainsi a_2 peut utiliser un g_i où $i \in [2, m]$. On supposera qu'il s'agit de g_2. En appliquant une deuxième fois le lemme d'échange à la situation $X = G_1 \backslash \{g_2\}$ et $x = g_2$, on obtient $g_2 \in \langle (G_1 \backslash \{g_2\}) \cup \{a_2\} \rangle$. On pose $G_2 = (G_1 \backslash \{g_2\}) \cup \{a_2\} = \{a_1, a_2, g_3, ..., g_m\}$. Le lecteur vérifiera que G_2 engendre encore E : il s'agit donc de vérifier que $g_1 \in \langle G_2 \rangle$ et $g_2 \in \langle G_2 \rangle$.

Il est alors clair qu'au bout de la i-ème étape, on aura formé une partie génératrice $G_m = \{a_1, ..., a_m\}$ ne contenant que des éléments de L. Comme par hypothèse $Card(L) > Card(G) = m$, on peut choisir dans L un élément a_{m+1} qui n'est plus dans G_m. Mais comme G_m est génératrice, on peut écrire $a_{m+1} = \sum_{i=1}^{m} \alpha_i a_i$, la partie $\{a_1, ..., a_m, a_{m+1}\}$ serait liée et donc L aussi, d'où la contradiction, ce qui achève la démonstration⊡

THEOREM 7.5 *Soit E un K-espace vectoriel de dimension finie, toutes les parties basiques de E ont le même nombre d'éléments. On peut donc dire aussi, avec les conventions habituelles, toutes les bases de E ont le même nombre d'éléments.*

Démonstration : il s'agit d'une conséquence simple du théorème précédent. En efet, soient B et B' deux parties basiques de E, on a $Card(B) \leq Card(B')$ car B est libre et B' est génératrice, et $Card(B') \leq Card(B)$ car B' est libre et B génératrice, d'où $Card(B) = Card(B')$⊡

PROPOSITION 7.2 *Soient E un K-espace vectoriel et p un entier donné, alors si $p+1$ vecteurs sont combinaisons linéaires de p vecteurs donnés de E, ces $p+1$ vecteurs sont liés.*

Démonstration : en effet, si y_1, ..., y_{p+1} sont combinaisons linéaires de x_1, ..., x_p, toute partie libre de $\langle x_1, ..., x_p \rangle$ a au plus p éléments d'après le théorème ? ?, donc $\{y_1, ..., y_{p+1}\}$ est nécessairement liée⊡

DEFINITION 7.2 *Soit E un K-espace vectoriel de dimension finie, on appelle **dimension** de E le nombre de vecteurs de l'une quelconque de ses bases (ou de ses parties basiques). Ce nombre est noté $\dim_K(E)$, ou $\dim(E)$ si aucune confusion sur le corps de base n'est à craindre. Si E n'est pas de dimension finie, on dit que sa dimension est infinie.*

Exemples :

1. $\dim_K(\{0\}) = 0$ et réciproquement $\dim_K(E) = 0 \Rightarrow E = \{0\}$
2. En examinant les bases canoniques des espaces K^n on constate que $\dim_K(K^n) = n$. Cet exemple est excessivement important.

3. $\dim_C(C) = 1$ et $\dim_R(C) = 2$, de même $\dim_C(C^n) = n$ et $\dim_R(C^n) = 2n$. Cela montre que le corps de base est d'une importance fondamentale dans le calcul de la dimension d'un espace vectoriel.
4. $\dim_K(K_n[X]) = n + 1$.

7.1.3 Caractérisation des bases en dimension finie

PROPOSITION 7.3 *Soit E un K-espace vectoriel de dimension n sur K, alors toute partie libre de E a au plus n éléments. Autrement dit, toute partie de E ayant au moins $n + 1$ éléments est liée.*

Démonstration : il s'agit simplement d'une formulation différente de la proposition ? ?

On peut aussi formuler cette proposition en termes de familles.

THEOREM 7.6 *Soient E un K-espace vectoriel de dimension n sur K et B une partie de E, les assertions suivantes sont équivalentes :*

1. *B est une partie basique de E*
2. *B est une partie libre de E et $Card(B) = n$*
3. *B est une partie génératrice de E et $Card(B) = n$.*

Démonstration : (1.⇒2.) évident, par définition de la dimension.

(2.⇒3.) si B n'était pas génératrice, on pourrait la compléter en une partie basique (théorème ? ?) qui aurait au moins $n + 1$ éléments, ce qui contredirait la proposition précédente.

(3.⇒1.) si B n'était pas libre, on pourrait en extraire une partie basique (théorème ? ?) ayant strictement moins de n éléments, ce qui contredirait la définition de la dimension□

Ce théorème est très important dans la pratique, car il permet de montrer qu'une partie B est basique si elle possède l'une des deux propriétés exigées (libre ou génératrice) et si elle a le nombre ad hoc d'éléments. Il est le plus souvent employé dans le sens "B est libre et $Card(B) = n$ donc B est une partie basique". Il est en effet généralement plus facile de montrer qu'une partie est libre que de montrer que cette même partie est génératrice.

On termine cette section en montrant que la dimension caractérise à isomorphisme près les K-espaces vectoriels de dimension finie.

THEOREM 7.7
1. *Tout K-espace vectoriel de dimension n est isomorphe à l'espace vectoriel K^n.*
2. *Soient E et F deux K-espaces vectoriels de dimension finie, alors E et F sont isomorphes si et seulement si $\dim_K(E) = \dim_K(F)$.*

Démonstration :

1. soient E un K-espace vectoriel de dimension n et $B = (x_1, ..., x_n)$ une base de E, alors l'application

$$\phi : \begin{array}{l} K^n \to E \\ (\alpha_1, ..., \alpha_n) \mapsto \alpha_1 x_1 + ... + \alpha_n x_n \end{array}$$

   est un isomorphisme de K-espaces vectoriels (le lecteur est invité à le montrer).

2. on suppose que E et F sont isomorphes. Soit alors ϕ un isomorphisme de E sur F, si B est une partie basique de E on sait que $\phi(B)$ est une partie basique de F et donc, comme $Card(B) = Card(\phi(B))$ vu que ϕ est injective, on en déduit que $\dim_K(E) = \dim_K(F)$ (cf théorème ? ?). Réciproquement, si $\dim_K(E) = \dim_K(F) = n$, d'après 1. E et F sont tous deux isomorphes à K^n, donc isomorphes entre eux, par transitivité de l'isomorphie□

Exercices :

1. On considère dans $\mathbb{R}^3$ les vecteurs $x_1 = (-1, 1, 1)$, $x_2 = (1, -1, 1)$, $x_3 = (1, 1, -1)$, montrer que (x_1, x_2, x_3) est une base de $\mathbb{R}^3$.
2. On considère dans $\mathbb{R}^4$ les vecteurs $x_1 = (1, 2, -1, 2)$, $x_2 = (2, 3, 0, -1)$, $x_3 = (1, 2, 1, 4)$, $x_3 = (1, 3, -1, 0)$, montrer que (x_1, x_2, x_3, x_4) est une base de $\mathbb{R}^4$.
3. Soit $a \in K$, montrer que la famille $(1, X - a, ..., (X - a)^n)$ est une base de $K_n[X]$.

7.2 Dimension d'un sous-espace vectoriel

THEOREM 7.8 *Soient E un K-espace vectoriel de dimension finie et F un sous-espace vectoriel de E, on suppose que E est de dimension finie, alors :*

1. *F est un K-espace vectoriel de dimension finie*
2. $\dim_K(F) \leq \dim_K(E)$
3. *si* $\dim_K(F) = \dim_K(E)$, *alors* $F = E$.

Démonstration : soit L une partie libre de F, c'est donc a fortiori une partie libre de E. D'après, par exemple, le théorème de la base incomplète, le cardinal de L est inférieur ou égal à $\dim_K(E)$. Il suffit donc de prendre dans F une partie libre ayant le nombre maximum d'éléments pour avoir une partie basique de F. Ainsi F admet une base finie ayant un cardinal inférieur ou

égal à $\dim_K(E)$. Si $\dim_K(F) = \dim_K(E)$, alors une base quelconque de F est une partie libre de E ayant le bon nombre d'éléments (théorème ? ?), c'est donc aussi une base de E, i.e. $F = E\boxdot$

La partie 3. du théorème précédent est fondamentale dans les exercices. Elle permet en effet de démontrer que deux K-espaces vectoriels de dimension finie sont égaux si et seulement si l'un des deux est inclus dans l'autre et s'ils ont la même dimension (alors qu'autrement il faudrait montrer deux inclusions).

Exemples :

1. Soit D une droite vectorielle, les seuls sous-espaces vectoriels de D sont $\{0\}$ et D.
2. Soit P un plan vectoriel, les seuls sous-espaces vectoriels de P sont $\{0\}$, les droites vectorielles contenues dans P et P lui-même.

7.2.1 Somme et dimension

Le but de ce paragraphe est d'établir une formule attribuée à Grassmann donnant la dimension de la somme de deux sous-espaces vectoriels d'un K-espace vectoriel E de dimension finie, et d'en déduire ses principales conséquences.

PROPOSITION 7.4 *Soient E_1 et E_2 deux K-espaces vectoriels de dimension finie, alors $E_1 \times E_2$ est de dimension finie et on a* $\dim_K(E_1 \times E_2) = \dim_K(E_1) + \dim_K(E_2)$.

Démonstration : en effet, on pose $m = \dim_K(E_1)$, $n = \dim_K(E_2)$ et on choisit $(a_1, ..., a_m)$ une base de E_1, $(b_1, ..., b_n)$ une base de E_2. On laisse au lecteur consciencieux le soin de vérifier que le $(m+n)$-uplet $((a_1, 0), ..., (a_m, 0), (0, b_1), ..., (0, b_n))$ est une base de $E_1 \times E_2$. On peut aussi vérifier que $K^m \times K^n$ est isomorphe à $K^{m+n}\boxdot$

THEOREM 7.9 *(dimension d'une somme directe)*
Soient E un K-espace vectoriel de dimension finie, E_1 et E_2 deux sous-espaces vectoriels de E dont la somme est directe, on a alors la relation :

$$\dim_K(E_1 \oplus E_2) = \dim_K(E_1) + \dim_K(E_2).$$

De plus, si B_1 est une partie basique de E_1 et B_2 une partie basique de E_2, $B_1 \cup B_2$ est une partie basique de $E_1 \oplus E_2$.

Démonstration : il suffit bien entendu de montrer la seconde asertion. On peut remarquer que $E_1 \times E_2$ et $E_1 \oplus E_2$ sont isomorphes et appliquer la proposition précédente$\boxdot$

Exercices :

1. Rédiger les démonstrations de la proposition ? ? et du théorème ? ?. Montrer aussi que dans ce dernier l'hypothèse "E est de dimension finie" peut être remplacée par "E_1 et E_2 sont deux sous-espaces vectoriels de dimension finie".
2. Généraliser la proposition ? ? et le théorème ? ? au cas d'un nombre fini quelconque d'espaces vectoriels. On obtiendra ainsi les formules :

$$\begin{aligned} \dim_K(E_1 \times ... \times E_n) &= \dim_K(E_1) + ... + \dim_K(E_n) \\ \dim_K(E_1 \oplus ... \oplus E_n) &= \dim_K(E_1) + ... + \dim_K(E_n). \end{aligned}$$

THEOREM 7.10 *(existence d'un supplémentaire)*

*Soient E un K-espace vectoriel de dimension finie et F un sous-espace vectoriel de E, alors F admet au moins un supplémentaire dans E. De plus, tous les supplémentaires de F dans E ont la même dimension (appelée parfois **codimension** de F dans).*

Démonstration : soit B une partie basique de F (qui est de dimension finie), d'après le théorème de la base incomplète, on peut trouver une partie C de E disjointe de B telle que $B \cup C$ soit une partie basique de E. Il est alors quasi évident que le sous-espace vectoriel G engendré par C est un supplémentaire de F. En effet, on pose $B = \{b_1, ..., b_m\}$ et $C = \{c_1, ..., c_p\}$. Soit x quelconque de E, il se décompose de façon unique sur la partie basique $B \cup C$, i.e. $x = \beta_1 b_1 + ... + \beta_m b_m + \gamma_1 c_1 + ... + \gamma_p c_p$. Or $u = \beta_1 b_1 + ... + \beta_m b_m \in \langle B \rangle = F$ et $v = \gamma_1 c_1 + ... + \gamma_p c_p \in \langle C \rangle = G$, donc $x = u + v$ avec $u \in F$ et $v \in G$, ce qui démontre que $E = F + G$. Par ailleurs, l'unicité de la décomposition de x sur $B \cup C$ entraîne l'unicité de l'écriture $x = u + v$, avec $u \in F$, $v \in G$, la somme est donc directe. Enfin, soit G un supplémentaire quelconque de F dans E, on a $E = F \oplus G$, et d'après le théorème ? ?, $\dim_K(G) = \dim_K(E) - \dim_K(F)$, la dimension de G est donc indépendante du supplémentaire choisi□

On remarque que l'on a donné, sans démonstration, un résultat plus général : tout sous-espace vectoriel admet au moins un supplémentaire (sans condition de dimension). La démonstration du théorème éclaire un peu le phénomène. Le lecteur pourra prouver que dans un espace vectoriel de dimension 3, un plan vectoriel P admet pour supplémentaire toute droite vectorielle non contenue dans P.

THEOREM 7.11 *(formule de Grassmann)*

Soient E un K-espace vectoriel de dimension finie, E_1 et E_2 deux sous-espaces vectoriels de E, on a la relation :

$$\dim_K(E_1 + E_2) = \dim_K(E_1) + \dim_K(E_2) - \dim_K(E_1 \cap E_2).$$

Démonstration : il existe de nombreuses preuves de ce résultat. On esquisse ici une démonstration directe (ce n'est pas la plus jolie) et on renvoie le lecteur à la section suivante pour une preuve élégante). L'astuce consiste à se ramener à une somme directe. L'inclusion $E_1 \cap E_2 \subset E_1$ permet de choisir un supplémentaire F de $E_1 \cap E_2$ dans E_1. Ainsi $E_1 = (E_1 \cap E_2) \oplus F$. On en déduit que $E_1 + E_2 = F \oplus E_2$. En effet, $F \cap E_2 = (F \cap E_1) \cap E_2 = F \cap (E_1 \cap E_2) = \{0\}$ et il est clair que $E_1 + E_2 = F + E_2$ ($F + E_2 \subset E_1 + E_2$ est évidente, l'autre inclusion se vérifie facilement). Ainsi, en appliquant le théorème ? ?, on trouve :

$$\begin{aligned} \dim_K(E_1) &= \dim_K(E_1 \cap E_2) + \dim_K(F) \text{ et} \\ \dim_K(E_1 + E_2) &= \dim_K(F) + \dim_K(E_2), \end{aligned}$$

d'où le résultat en éliminant $\dim_K(F)$⊡

THEOREM 7.12 *Soient E_1 et E_2 deux sous-espaces vectoriels d'un K-espace vectoriel E de dimension finie E, on considère les trois propriétés suivantes :*

1. $E_1 + E_2 = E$
2. $E_1 \cap E_2 = \{0\}$
3. $\dim_K(E_1) + \dim_K(E_2) = \dim_K(E)$.

Alors deux quelconques de ces propriétés entraînent la troisième, et $E = E_1 \oplus E_2$. Réciproquement, si $E = E_1 \oplus E_2$ les trois propriétés sont vérifiées.

Démonstration : si 1. et 2. sont réalisées, alors on sait que $E = E_1 \oplus E_2$ et la propriété 3. n'est autre que le théorème ? ?

Si 1. et 3. sont réalisées, la formule de Grassmann entraîne que $\dim_K(E_1 \cap E_2) = 0$, donc $E_1 \cap E_2 = \{0\}$.

Si 2. et 3. sont réalisées, la formule de Grassmann devient $\dim_K(E_1 + E_2) = \dim_K(E)$. Or $E_1 + E_2 \subset E$, le théorème ? ? montre alors que $E = E_1 + E_2$. Enfin, la réciproque est triviale⊡

Exercices :

1. Montrer que, dans la formule de Grassmann, l'hypothèse "E est de dimension finie" peut être remplacée par l'hypothèse plus faible "E_1 et E_2 sont deux sous-espaces vectoriels de dimension finie".
2. Dans R^4 on considère le sous-espace vectoriel E_1 engendré par $\{x, y, z\}$ avec $x = (1, 2, 3, 4)$, $y = (2, 2, 2, 6)$, $z = (0, 2, 4, 4)$, et le sous-espace vectoriel E_2 engendré par $\{u, v\}$ avec $u = (1, 0, -1, 2)$ et $v = (2, 3, 0, 1)$. Vérifier sur cet exemple la formule de Grassmann.

7.3 Notion de rang

Soit E un K-espace vectoriel, on appellera système de vecteurs de E toute famille finie de vecteurs de E. Au risque d'insister lourdement, on précise bien qu'un système est une famille et qu'on admet volontiers les répétitions, la notion de système correspond donc à la notion combinatoire d'arrangement avec répétitions.

DEFINITION 7.3 *Soit S un système de vecteurs d'un K-espace vectoriel E, on appelle* ***rang*** *de S, et on note rg(S) la dimension sur K du sous-espace vectoriel engendré par S. Autrement dit :*

$$rg(S) = \dim_K(\langle S \rangle).$$

Cette définition a bien un sens puisque par construction même, $\langle S \rangle$ admet une famille finie de générateurs, il est donc de dimension finie.

PROPOSITION 7.5 *Soit S un système de vecteurs, le rang de S est égal au cardinal d'une sous-famille libre maximale de S, i.e. le nombre maximum de vecteurs linéairement indépendants que l'on peut extraire de S.*

Démonstration : en effet, le rang de S n'est autre que le cardinal de l'une des bases de l'espace vectoriel $\langle S \rangle$ $\boxdot$

7.3.1 Rang d'une application linéaire

L'outil central de cette section est le théorème du rang. On commence par deux exemples, sous forme d'exercices, où l'on découvre cet important théorème.

Exercices :

1. Soit $f : R^3 \to R^3$ définie par $f(x, y, z) = (x', y', z')$ où $x' = y' = z' = 2x + y + z$,
   (a) montrer que f est linéaire
   (b) donner une base de Ker(f) et en déduire $\dim(Ker(f))$
   (c) donner une base de Im(f) et en déduire $\dim($ $Im(f))$
   (d) vérifier la relation $\dim(R^3) = \dim(Ker(f)) + \dim($ $Im(f))$.

2. Soit $f : R^4 \to R^3$ définie par $f(x, y, z, t) = (x', y', z')$ où $\begin{array}{l} x' = x + y + z + 2t \\ y' = y - z + t \\ z' = x - y + 3z \end{array}$,

(a) montrer que f est linéaire

(b) on note (e_1, e_2, e_3, e_4) la base canonique de R^4, calculer le rang de la famille $(f(e_i))_{i\in[1,4]}$ et donner une base de Im(f)

(c) montrer que les vecteurs $(1, 1, 0, -1)$ et $(-3, 0, 1, 1)$ forment une partie basique de Ker(f)

(d) vérifier la relation $\dim(R^4) = \dim(Ker(f)) + \dim($
$Im(f))$.

DEFINITION 7.4 *Soient E un K-espace vectoriel de dimension finie, F un K-espace vectoriel quelconque et f une application linéaire de E dans F, on appelle* ***rang*** *de f, et on note rg(f) la dimension sur K de Im(f), i.e.*

$$rg(f) = \dim_K(Im(f)).$$

Cette définition a bien un sens, car E est par hypothèse un espace vectoriel de dimension finie. E admet une partie génératrice finie G. On sait alors que f(G) est une partie (finie, car G est finie) génératrice de Im(f) (cf théorème ? ?), donc Im(f) est de dimension finie, et ceci sans aucune condition sur l'espace vectoriel F.

En particulier, on prend $(e_1, ..., e_n)$ une base de E, alors d'après ce que l'on vient de dire, le rang de f n'est autre que le rang du système $(f(e_1), ..., f(e_n))$ qui ne dépend donc pas de la base choisie dans E.

THEOREM 7.13 *(du rang)*

Soient E un K-espace vectoriel de dimension finie, F un K-espace vectoriel quelconque et f une application linéaire de E dans F, on a la relation :

$$\dim_K(E) = \dim_K(Ker(f)) + \dim_K(Im(f)) = \dim_K(Ker(f)) + rg(f).$$

Démonstration : en effet, on sait que si f est une application linéaire de E dans F, tout supplémentaire de Ker(f) est isomorphe à Im(f) (théorème ? ?). Soit donc G un supplémentaire de Ker(f), on a $E = Ker(f) \oplus G$ d'où $\dim_K(E) = \dim_K(Ker(f)) + \dim_K(G)$. Or G et Im(f) sont isomorphes, ils ont donc la même dimension (finie) (théorème ? ?) d'où $\dim_K(E) = \dim_K(Ker(f)) + \dim_K($
$Im(f))\boxdot$

PROPOSITION 7.6 *Soient E et F deux K-espaces vectoriels de dimension finie et f une application linéaire de E dans F, on a :*

1. $rg(f) \leq \inf(\dim_K(E), \dim_K(F))$
2. *f est injective si et seulemnt si* $rg(f) = \dim_K(E)$
3. *f est surjective si et seulement si* $rg(f) = \dim_K(F)$.

Démonstration :

1. comme $rg(f) = \dim_K(E) - \dim_K(Ker(f))$ d'après le théorème du rang, on a bien $rg(f) \leq \dim_K(E)$. De plus, l'inclusion $Im(f) \subset F$ entraîne $rg(f) \leq \dim_K(F)$ d'après le théorème du sous-espace.
2. f est injective si et seulement si le noyau de f se réduit à $\{0\}$ donc si et seulement si $\dim_K(Ker(f)) = 0$, il suffit alors d'apliquer le théorème du rang.
3. f est surjective si et seulement si $Im(f) = F$, il suffit alors d'appliquer le théorème du sous-espace□

Exercice : On propose ici une démonstration de la formule de Grassmann à l'aide du théorème du rang. Soient E un K-espace vectoriel de dimension finie, E_1 et E_2 deux sous-espaces vectoriels de E et

$$f : \begin{array}{l} E_1 \times E_2 \to E \\ (x_1, x_2) \mapsto x_1 + x_2 \end{array},$$

1. montrer que f est linéaire
2. montrer que $Ker(f) = \{(x, -x) \text{ / } x \in E_1 \cap E_2\}$ et en déduire que Ker(f) est isomorphe à $E_1 \cap E_2$
3. montrer que $Im(f) = E_1 + E_2$
4. appliquer à f le théorème du rang et retrouver ainsi la formule de Grassmann.

On sait que deux K-espaces vectoriels de dimension finie sont isomorphes si et seulement si ils ont la même dimension. Mais il ne faudrait pas croire que toute application linéaire de l'un dans l'autre soit un isomorphisme. On peut penser par exemple à l'application nulle. On va donner maintenant des critères variés pour qu'une application linéaire soit un isomorphisme.

PROPOSITION 7.7 *Soient E et F deux K-espaces vectoriels de **même** dimension finie n et f une application linéaire de E dans F, les propositions suivantes sont équivalentes :*

1. *f est un isomorphisme de E sur F*
2. *f est injective*
3. *f est surjective*
4. $rg(f) = n$.

Démonstration : (1.⇒2.) trivial

(2.⇒3.) si f est injective, on a $Ker(f) = \{0\}$ et d'après le théorème du rang $rg(f) = \dim_K(E) = \dim_K(F)$, donc f est surjective

(3.⇒4.) trivial

(4.⇒1.) si $rg(f) = n$ alors f est surjective et d'après le théorème du rang $\dim_K(Ker(f)) = 0$, donc f est injective□

La proposition précédente est très importante car elle permet d'affirmer qu'une application linéaire entre espaces de **même** dimension finie est bijective dès qu'elle est injective ou surjective. Mais atention, ceci n'est plus valable en dimension infinie comment en témoignent les exemples suivants :

1. $D : \begin{array}{l} K[X] \to K[X] \\ P \mapsto P' \end{array}$ est surjective mais non injective
2. $\phi : \begin{array}{l} K[X] \to K[X] \\ P \mapsto XP \end{array}$ est injective mais non surjective.

Exercices :

1. Soient E un K-espace vectoriel de dimension finie et f un endomorphisme de E, montrer que les propositions suivantes sont équivalentes :
   (a) $E = Ker(f) \oplus Im(f)$
   (b) $Ker(f) = Ker(f^2)$
   (c) $Im(f) = Im(f^2)$.

   Que subsiste-t-il si E est de dimension infinie ?
2. Soit f un endomorphisme d'un K-espace vectoriel E de dimension finie, montrer que l'image et le noyau de f coïncident si et seulement si on a $f^2 = 0$, $\dim_K(E)$ est un entier pair et $rg(f) = \frac{\dim_K(E)}{2}$.

7.3.2 Dimension de $\mathcal{L}$(E,F)

THEOREM 7.14 *Soient E et F deux K-espaces vectoriels de dimension finie, alors l'espace vectoriel $\mathcal{L}$(E, F) est aussi de dimension finie et on a :*

$$\dim_K(\mathcal{L}(E, F)) = \dim_K(E).\dim_K(F).$$

Démonstration : admis

On va visualiser ce résultat de façon naturelle au chapitre suivant.

Chapitre 8

Matrices

Après avoir peiné parmi les espaces vectoriels, on est revenu en pays de connaissance, le relief s'adoucit. Les retardataires pourront donc rejoindre le peloton. Ce chapitre est en grande partie une reformulation de notions abordées dans les chapitres précédents, mais sous forme plus concrète et visuelle. Tout le début est donc en pente douce et permet l'utilisation d'un grand braquet. Néanmoins certains passages annoncent déjà les massifs qu'il faudra gravir... à l'Ensae.

8.1 Calcul matriciel

8.1.1 Généralités

DEFINITION 8.1 *Soient K un ensemble, n et m deux entiers naturels non nuls, on appelle* ***matrice*** *de* ***type*** *(n, m) à éléments dans K toute application* $A : [1, n] \times [1, m] \to K$.

En posant $A(i, j) = a_{ij}$, on convient d'écrire la matrice A sous forme d'un tableau rectangulaire à n lignes et m colonnes :

$$A = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ a_{21} & \dots & a_{2m} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nm} \end{pmatrix}.$$

a_{11}, ..., a_{nm} s'appellent les **éléments** de la matrice A, a_{ij} se trouvant à l'intersection de la i-ème ligne et de la j-ème colonne. On écrira de façon condensée $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq m}$ ou même $A = (a_{ij})$ si aucune confusion n'est possible sur le type de la matrice A.

Enfin, on note $\mathcal{M}_{n,m}(K)$ l'ensemble des matrices de type (n, m) à éléments dans K. Si $K = R$ on dit que la matrice A est **réelle** et si $K = C$ on dit qu'elle est **complexe**.

DEFINITION 8.2 *Si $m = n$, la matrice $A = (a_{ij})$ est dite **carrée d'ordre** n, ou plus simplement **carrée**. Les éléments a_{11}, a_{22}, ..., a_{nn} constituent alors ce qu'on appelle la **diagonale principale** de la matrice A.*

Dans toute la suite de ce chapitre on supposera que K est un corps commutatif.

DEFINITION 8.3 *Une matrice carrée (a_{ij}) est dite **diagonale** si tous les éléments situés hors de la diagonale principale sont nuls, i.e. :*

$$\forall i, j \in [1, n], i \neq j \Rightarrow a_{ij} = 0.$$

*Une matrice carrée (a_{ij}) est dite **trigonale inférieure** si tous les éléments situés au-dessus de la diagonale principale sont nuls, i.e. :*

$$\forall i, j \in [1, n], i < j \Rightarrow a_{ij} = 0.$$

*Si de plus les éléments de la diagonale principale sont nuls, on dit que la matrice (a_{ij}) est **strictement trigonale inférieure**. On définit de façon analogue une matrice carrée **(strictement) trigonale supérieure**.*

DEFINITION 8.4 *Une matrice carrée (a_{ij}) est dite **symétrique** si :*

$$\forall i, j \in [1, n], a_{ij} = a_{ji}.$$

*Une matrice carrée (a_{ij}) est dite **antisymétrique** si :*

$$\forall i, j \in [1, n], a_{ij} + a_{ji} = 0.$$

On ne va pas faire languir le lecteur plus longtemps, on aborde maintenant l'exemple fondamental qui est à la source de la notion de matrice.

8.1.2 Matrice associée à une application linéaire

On considère E et F deux K-espaces vectoriels de dimension finie, $\mathcal{B} = (e_1, ..., e_m)$ une base de E, $\mathcal{C} = (f_1, ..., f_n)$ une base de F et enfin u un élément de $\mathcal{L}$(E, F).

DEFINITION 8.5 *On appelle **matrice** de u **relativement aux bases** B et C la matrice (a_{ij}) de type (n, m) définie par :*

$$\forall j \in [1, m], u(e_j) = a_{1j}f_1 + a_{2j}f_2 + ... + a_{nj}f_n.$$

Autrement dit, pour tout indice j de [1, n], la j-ème colonne de (a_{ij}) est constituée par les coordonnées de $u(e_j)$ relativement à la base $\mathcal{C}$.

On désignera cette matrice par $M_{\mathcal{CB}}(u)$ ou, plus simplement, $M(u)$ si aucune confusion n'est possible. On verra lors de la multiplication des matrices pourquoi on renverse l'ordre des bases.

Si $E = F$ et si $\mathcal{B} = \mathcal{C}$, on parlera de la matrice de l'endomorphisme u par rapport à la base $\mathcal{B}$ et on la notera $M_B(u)$.

On se rend compte maintenant que la notion de base est beaucoup plus importante que celle de partie basique, car on ne saurait pas dans quel ordre écrire les lignes et les colonnes de la matrice associée à une application linéaire.

Exercices :

1. Soient $\mathcal{B} = (e_1, e_2)$ une base d'un K-espace vectoriel E et h_λ l'homothétie de rapport λ, on a donc $h_\lambda(e_1) = \lambda e_1$ et $h_\lambda(e_2) = \lambda e_2$, d'où $M_{\mathcal{B}}(h_\lambda) = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$. Cette matrice est indépendante de la base choisie. Généraliser au cas où $\dim(E) = n$ quelconque.
2. On suppose $\dim(E) = 3$ et $\dim(F) = 2$ et on se donne $\mathcal{B}$ et $\mathcal{C}$ des bases respectives de E et F. Soit u l'application qui à tout vecteur x de E de coordonnées (x_1, x_2, x_3) dans la base $\mathcal{B}$ associe le vecteur y de F dont les coordonnées dans la base $\mathcal{C}$ sont (y_1, y_2) données par les formules $y_1 = x_1 - 2x_2 + 3x_3$ et $y_2 = x_1 + x_3$, montrer que u est linéaire et déterminer sa matrice par rapport aux bases $\mathcal{B}$ et $\mathcal{C}$.
3. Soient E et F deux K-espaces vectoriels de dimension finie et u un isomorphisme de E sur F, montrer que pour toute base $\mathcal{B}$ de E il existe une base $\mathcal{C}$ de F telle que :

$$M_{\mathcal{CB}}(u) = \begin{pmatrix} 1 & 0 & 0 & ... & 0 \\ 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 1 \end{pmatrix}.$$

Si $\mathcal{B} = (e_1, ..., e_n)$, on pourra prendre $\mathcal{C} = (u(e_1), ..., u(e_n))$ et justifier. Réciproque ?

THEOREM 8.1 *L'application* $M_{\mathcal{CB}} : \begin{array}{l} \mathcal{L}(E,F) \to \mathcal{M}_{n,m}(K) \\ u \mapsto M_{\mathcal{CB}}(u) \end{array}$ *est bijective.*

Démonstration : en effet, si u et u' éléments de $\mathcal{L}$(E, F) sont tels que $M_{\mathcal{CB}}(u) = M_{\mathcal{CB}}(u')$ alors, en considérant les colonnes de ces matrices, pour tout indice j, $u(e_j) = u'(e_j)$ et donc par linéarité, $\forall x \in E$, $u(x) = u'(x)$, i.e. $u = u'$ et l'application $M_{\mathcal{CB}}$ est injective.

Par ailleurs, si l'on se donne une matrice $A = (a_{ij}) \in \mathcal{M}_{n,m}(K)$ quelconque, on pose pour tout indice j, $e'_j = a_{1j}f_1 + ... + a_{nj}f_n$. On sait d'après le théorème ? ? qu'il existe une unique application linéaire $u \in \mathcal{L}(E, F)$ vérifiant $u(e_j) = e'_j$ pour tout indice j. On aura donc $M_{\mathcal{CB}}(u) = A$ et l'application $M_{\mathcal{CB}}$ est surjective□

Soit $A = (a_{ij})$ une matrice de type (n, m) à éléments dans K, alors d'après le théorème précédent, A est la matrice d'une unique application linéaire u_A de K^m dans K^n relativement aux bases canoniques de ces espaces. On dit que u_A est l'application linéaire **canoniquement associée** à A.

On définira alors le noyau, l'image et le rang de la matrice A comme étant respectivement le noyau, l'image et le rang de u_A. En particulier l'image de A, notée Im(A) sera le sous-espace vectoriel de K^n constitué des vecteurs $u_A(x)$ où x décrit K^m. Im(A) est donc engendré par les colonnes de A (chaque colonne de A étant considéré comme un vecteur de K^n rapporté à sa base canonique). De même, le rang de A est le rang de u_A, i.e. le nombre de colonnes de A (considérées comme vecteur de K^n), linéairement indépendantes.

THEOREM 8.2 *Soient E et F deux espaces vectoriels de dimension finie rapportés à deux bases respectives $\mathcal{B}$ et $\mathcal{C}$ et $u \in \mathcal{L}(E, F)$, alors $rg(M_{\mathcal{CB}}(u)) = rg(u)$ et donc toutes les matrices associées à u, pour tous les choix possibles de bases, ont le même rang.*

Démonstration : on pose $m = \dim(E)$, $n = \dim(F)$ et $A = (a_{ij}) = M_{\mathcal{CB}}(u)$, avec $\mathcal{B} = (e_1, ..., e_m)$ et $\mathcal{C} = (f_1, ..., f_n)$. On veut comparer rg(A) et rg(u), on doit donc introduire u_A et relier u_A et u :

$$\begin{array}{ccc} E & \stackrel{u}{\rightarrow} & F \\ \downarrow \varphi & & \downarrow \psi \\ K^m & \stackrel{u_A}{\rightarrow} & K^n \end{array} .$$

Soient $\mathcal{B}' = (e'_1, ..., e'_m)$ et $\mathcal{C}' = (f'_1, ..., f'_n)$ les bases canoniques respectives de K^m et K^n, φ l'unique application linéaire de E dans K^m définie par :

$$\forall j \in [1, m], \varphi(e_j) = e'_j$$

et ψ l'unique application linéaire de F dans K^n définie par :

$$\forall i \in [1, n], \psi(f_i) = f'_i,$$

alors φ est un isomorphisme de E sur K^m et ψ de F sur K^n puisque, par construction, l'image d'une base est une base.

De plus, on a $\psi \circ u = u_A \circ \varphi$, car $\forall j \in [1, m]$, $\psi \circ u(e_j) = \psi(\sum_{i=1}^n a_{ij}f_i) = \sum_{i=1}^n a_{ij}\psi(f_i) = \sum_{i=1}^n a_{ij}f'_i = u_A(e'_j) = (u_A \circ \varphi)(e_j)$.

$\psi \circ u$ et $u_A \circ \varphi$ coïncident alors sur une base de E, par linéarité on a donc $\psi \circ u = u_A \circ \varphi$. Le matériel étant en place, la démonstration s'achève alors sans difficultés.

On a $u_A \circ \varphi(E) = u_A(\varphi(E)) = u_A(K^m) =$
$Im(u_A)$ et $\psi \circ u(E) = \psi($
$Im(u))$ donc
$Im(u_A) = \psi($
$Im(u))$. ψ induit donc un isomorphisme de $Im(u)$ sur
$Im(u_A)$, en particulier leurs dimensions sont égales, i.e. $rg(u) = rg(u_A) = rg(A)$. A titre d'exercice, le lecteur peut d'ailleurs montrer que φ induit également un isomorphisme de $Ker(u)$ sur $Ker(u_A)\boxdot$

8.1.3 Opérations sur les matrices

Le théorème ? ? exhibe une bijection entre les ensembles $\mathcal{L}$(E, F) et $\mathcal{M}_{n,m}(K)$. On sait alors, depuis le premier chapitre, que l'on peut transporter les opérations de $\mathcal{L}$(E, F) sur $\mathcal{M}_{n,m}(K)$.

Addition

Soient E et F deux K-espaces vectoriels de dimension finie, $\mathcal{B} = (e_1, ..., e_m)$ et $\mathcal{C} = (f_1, ..., f_n)$ des bases respectives de E et de F, u et v deux éléments de $\mathcal{L}$(E, F), on pose $M(u) = (a_{ij})$, $M(v) = (b_{ij})$, $M(u + v) = (c_{ij})$. On a alors $\forall j \in [1, m]$,

$$\begin{aligned}(u+v)(e_j) &= u(e_j) + v(e_j) \\ &= \sum_{i=1}^{n} a_{ij} f_i + \sum_{i=1}^{n} b_{ij} f_i \\ &= \sum_{i=1}^{n} (a_{ij} + b_{ij}) f_i\end{aligned}$$

i.e. $\forall (i, j) \in [1, n] \times [1, m]$, $c_{ij} = a_{ij} + b_{ij}$.

DEFINITION 8.6 *Si $A = (a_{ij})$ et $B = (b_{ij})$ sont deux matrices de type (n, m), on appelle* ***somme*** *de A et de B, et on note $A + B$ la matrice de type (n, m) $C = (c_{ij})$ définie par :*

$$\forall (i, j) \in [1, n] \times [1, m], c_{ij} = a_{ij} + b_{ij}.$$

On a donc $M(u + v) = M(u) + M(v)$.

Multiplication par un scalaire

Soit $M(\lambda u) = (a'_{ij})$ la matrice associée à λu, $\lambda \in K$, on a alors $\forall j \in [1, m]$,

$$\begin{aligned}(\lambda u)(e_j) &= \lambda(u(e_j)) \\ &= \lambda(\sum_{i=1}^{n} a_{ij} f_i) \\ &= \sum_{i=1}^{n} (\lambda a_{ij}) f_i\end{aligned}$$

i.e. $\forall (i, j) \in [1, n] \times [1, m]$, $a'_{ij} = \lambda a_{ij}$.

DEFINITION 8.7 *Si $A = (a_{ij})$ est une matrice de type (n, m), on note λA la matrice (λa_{ij}) obtenue en multipliant tous les éléments de A par le scalaire λ. On définit ainsi une loi externe sur $\mathcal{M}_{n,m}(K)$ de domaine d'opérateur K.*

On a donc $M(\lambda u) = \lambda . M(u)$.

THEOREM 8.3 *Soient E et F deux K-espaces vectoriels de dimension respectives m et n, rapportés à des bases $\mathcal{B}$ et $\mathcal{C}$, alors, muni des lois précédentes, $\mathcal{M}_{n,m}(K)$ est un K-espace vectoriel de dimension mn isomorphe à $\mathcal{L}_K(E, F)$ par l'application $M_{\mathcal{CB}}$ du théorème ? ?.*

Démonstration : il n'y a rien à démontrer, les opérations sur $\mathcal{M}_{n,m}(K)$ ont justement été définies par transport de structure, $M_{\mathcal{CB}}$ devient ainsi un isomorphisme d'espaces vectoriels.

On remarque simplement que si E_{ij} désigne la matrice de type (n, m) qui contient un 1 à l'intersection de la i-ème ligne et de la j-ème colonne et des 0 partout ailleurs, alors $(E_{ij})_{(i,j)\in[1,n]\times[1,m]}$ constitue une base de $\mathcal{M}_{n,m}(K)$ appelée base canonique de $\mathcal{M}_{n,m}(K)$, qui n'est autre que l'image de la base canonique de $\mathcal{L}_K(E, F)$. On peut écrire $A = (a_{ij}) = \sum_{i,j} a_{ij} E_{ij}$. L'élément neutre de l'espace vectoriel $\mathcal{M}_{n,m}(K)$ est la matrice de l'application nulle, i.e. la matrice dont tous les éléments sont nuls, on la notera 0. La matrice opposée de la matrice $A = (a_{ij})$ se notera $-A$, et on a $-A = (-a_{ij})$.

Produit

Soient E, F, G trois K-espaces vectoriels, $u \in \mathcal{L}(E, F)$, $v \in \mathcal{L}(F, G)$, $\mathcal{B} = (e_1, ..., e_m)$, $\mathcal{C} = (f_1, ..., f_n)$, $\mathcal{D} = (g_1, ..., g_p)$ des bases respectives de E, F, G, on pose :

$M_{\mathcal{CB}}(u) = (a_{kj})$	matrice de type (n, m)
$M_{\mathcal{DC}}(v) = (b_{ik})$	matrice de type (p, n) .
$M_{\mathcal{DB}}(v \circ u) = (c_{ij})$	matrice de type (p, m)

On a alors $\forall j \in [1,m]$:

$$\begin{aligned} v \circ u(e_j) &= v(u(e_j)) \\ &= v(\sum_{k=1}^{n} a_{kj} f_k) \\ &= \sum_{k=1}^{n} a_{kj} v(f_k) \\ &= \sum_{k=1}^{n} a_{kj} (\sum_{i=1}^{p} b_{ik} g_i) \\ &= \sum_{k=1}^{n} \sum_{i=1}^{p} a_{kj} b_{ik} g_i \\ &= \sum_{i=1}^{p} (\sum_{k=1}^{n} b_{ik} a_{kj}) g_i \end{aligned}$$

donc, par définition de c_{ij}, on en déduit :

$$\forall (i,j) \in [1,p] \times [1,m], c_{ij} = \sum_{k=1}^{n} b_{ik} a_{kj}.$$

DEFINITION 8.8 *Soient A une matrice de type (n ; m) et B une matrice de type (p, n), à coefficients dans K, on appelle* ***produit*** *de B par A, et on note BA, la matrice* (c_{ij}) *de type (p, m) définie par :*

$$\forall (i,j) \in [1,p] \times [1,m], c_{ij} = \sum_{k=1}^{n} b_{ik} a_{kj}.$$

On a donc $M_{\mathcal{DB}}(v \circ u) = M_{\mathcal{DC}}(v).M_{\mathcal{CB}}(u)$.

La formule de la définition a bien un sens car le nombre de colonnes de la première matrice du produit est le même que le nombre de lignes de la seconde matrice. L'expression de c_{ij} se traduit en disant que l'on fait le produit "ligne par colonne". On insiste sur le fait que si B est de type (p, n) et A de type (n', m) avec $n \neq n'$, alors le produit BA n'a ici aucun sens.

On insiste aussi sur le fait que si BA a un sens, en général AB n'en a pas ; que même si AB et BA ont tous deux un sens, en général AB et BA ne sont pas de même type ; que même... (cf théorème de structure pour les matrices carrées).

Une bonne astuce pour calculer un produit de matrices, en limitant les risques d'erreur, consiste à disposer le calcul comme suit :

$$\begin{array}{cc} & A \rightarrow \\ B & \\ \downarrow & \\ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \end{pmatrix} & \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \\ \begin{pmatrix} 22 & 28 \\ 40 & 52 \end{pmatrix} & \end{array} .$$

Cette disposition est de plus parfaitement adaptée à l'itération du produit. On peut voir sur cet exemple que AB a un sens, mais que AB est une matrice carrée d'ordre 3 alors que BA est carrée d'ordre 2.

PROPOSITION 8.1 *Sous réserve d'existence des expressions, on a pour toutes matrices A, B, C et tout scalaire λ :*

$$\begin{aligned} A(BC) &= (AB)C \\ A(\lambda B) &= (\lambda A)B = \lambda(AB) \\ A(B+C) &= AB + AC \\ (A+B)C &= AC + BC. \end{aligned}$$

"Sous réserve d'existence" signifie que les types des matrices intervenant dans une expression ont un sens. Les propriétés énoncées sont alors les traductions de propriétés démontrées pour les applications linéaires.

Exercices :

1. On pose $A = \begin{pmatrix} -1 & +3 & +1 \\ +4 & +0 & +5 \end{pmatrix}$, $B = \begin{pmatrix} +2 & +1 & +3 \\ -4 & +0 & +1 \\ +3 & -1 & -2 \end{pmatrix}$, $C = \begin{pmatrix} +2 & +1 & +3 \\ -4 & +0 & +1 \\ +3 & -1 & -2 \end{pmatrix}$. Vérifier sur cet exemple l'associativité du produit matriciel.
2. Vérifier directement sur les formules l'associativité du produit matriciel. Il est recommandé au lecteur de s'astreindre à effectuer complètement cet exercice qui est un très bon exemple de manipulation du symbole $\sum$.
3. Etudier la (ou les) relation(s) de dépendance linéaire existant entre les matrices réelles suivantes : $A = \begin{pmatrix} +1 & -1 \\ +2 & +1 \\ -1 & +0 \end{pmatrix}$, $B = \begin{pmatrix} +0 & +1 \\ -1 & +3 \\ +0 & -2 \end{pmatrix}$, $C = \begin{pmatrix} +2 & +3 \\ +0 & +0 \\ +1 & -1 \end{pmatrix}$, $D = \begin{pmatrix} +5 & -2 \\ +8 & -3 \\ -2 & +3 \end{pmatrix}$.

4. Soit $E = \left\{ \begin{pmatrix} a-b & b-c & 2c \\ 2a & a+b & -b \\ b & c & a \end{pmatrix}, a, b, c \in R \right\}$, montrer que E est un sous-espace vectoriel de $\mathcal{M}_3(R)$, en donner la dimension et une base. On pourra constater que si $A \in E$, on peut écrire $A = aA_1 + bA_2 + cA_3$, où A_1, A_2, A_3 sont trois matrices fixes. Cet exercice est standard et doit être un exercice réflexe.

8.1.4 Matrices colonnes

Soient E et F deux K-espaces vectoriels rapportés à des bases respectives $\mathcal{B} = (e_1, ..., e_m)$ et $\mathcal{C} = (f_1, ..., f_n)$, $u \in \mathcal{L}(E, F)$ défini par sa matrice A relativement aux bases $\mathcal{B}$ et $\mathcal{C}$, et x un élément de E, le problème posé ici est celui de l'utilisation de la matrice A pour la détermination de u(x). La solution peut être obtenue directement de la façon suivante.

Le vecteur x se décompose sur la base $\mathcal{B}$, $x = x_1e_1 + ... + x_me_m$, donc $u(x) = x_1u(e_1) + ... + x_mu(e_m)$. Or $u(e_1)$, ..., $u(e_m)$ apparaissent justement, décomposée sur la base $\mathcal{C}$, dans les colonnes de la matrice associée à u. On peut donc calculer les composantes sur la base $\mathcal{C}$ de u(x) en fonction de x_1, ..., x_m et des coefficients de cette matrice. Ce calcul direct est laissé au lecteur.

On propose ici une résolution plus sophistiquée de ce problème qui a l'avantage d'être plus structurelle. L'élément x de E définit une application linéaire $\bar{x}$ de K dans E définie par :

$$\bar{x} : \lambda \mapsto \lambda x.$$

La matrice de $\bar{x}$ par rapport aux bases (1) et $\mathcal{B}$ de K et E est une matrice de type (m, 1) dite **matrice colonne**. Cette matrice n'est autre que $\begin{pmatrix} x_1 \\ ... \\ x_m \end{pmatrix}$ où x_1, ..., x_m sont les coordonnées de x dans $\mathcal{B}$. En effet, $\bar{x}(1) = x = \sum_{i=1}^{m} x_i e_i$. Cette matrice $M_{\mathcal{B}(1)}(\bar{x})$ se note traditionnellement X.

De même, l'élément u(x) de F définit une application linéaire $\overline{u(x)}$ de K dans F, définie par $\overline{u(x)} : \lambda \mapsto \lambda u(x)$. Sa matrice par rapport aux bases (1) et $\mathcal{C}$ de K et F est donc aussi une matrice colonne $M_{\mathcal{C}(1)}(u(x))$, qui est la matrice des coordonnées de u(x) dans la base $\mathcal{C}$, on la note Y.

On considère alors la composition $u \circ \bar{x}$ de K dans F définie par :

$$u \circ \bar{x} : \lambda \mapsto u(\bar{x}(\lambda)) = u(\lambda x) = \lambda u(x) = \overline{u(x)}(\lambda),$$

i.e. $u \circ \bar{x} = \overline{u(x)}$ et par passage aux matrices associées $Y = AX$.

PROPOSITION 8.2 *La matrice colonne Y des coordonnées du vecteur u(x) dans la base $\mathcal{C}$ s'obtient en multipliant la matrice A de u relativement aux bases $\mathcal{B}$ et $\mathcal{C}$ par la matrice colonne X des coordonnées du vecteur x dans la base $\mathcal{B}$: $Y = AX$.*

Exercice : On considère une application linéaire u de E dans F donnée par sa matrice relativement aux bases respectives $\mathcal{B}$ et $\mathcal{C}$ des K-espaces vectoriels E et F :

$$M_{\mathcal{CB}}(u) = \begin{pmatrix} +3 & +4 \\ -1 & +1 \\ +2 & -2 \end{pmatrix}.$$

1. Déterminer dim(E) et dim(F).
2. Soit x le vecteur de coordonnées (4, 1) dans $\mathcal{B}$, calculer les coordonnées dans $\mathcal{C}$ de u(x).
3. Déterminer Ker(u), Im(u), rg(u).
4. Soient $e'_1 = 3e_1 + e_2$ et $e'_2 = -2e_1 + 5e_2$, montrer que $\mathcal{B}' = (e'_1, e'_2)$ est une base de E, puis calculer $M_{\mathcal{CB}}(u)$.
5. Soient $f'_1 = -f_1 + f_3$, $f'_2 = 2f_1 - f_2 + 2f_3$, $f'_3 = f_1 - f_2 + f_3$, montrer que $\mathcal{C}' = (f'_1, f'_2, f'_3)$ est une base de F. Déterminer les coordonnées de f_1, f_2, f_3 sur cette base, puis calculer $M_{\mathcal{C}'\mathcal{B}}(u)$ et $M_{\mathcal{C}'\mathcal{B}'}(u)$.

8.1.5 Transposition

DEFINITION 8.9 *Soit $A = (a_{ij})$ une matrice de type (n, m), on appelle **transposée** de A et on note A^t la matrice (a'_{ij}) de type (m, n) définie par :*

$$\forall (i,j) \in [1,m] \times [1,n], a'_{ij} = a_{ji}.$$

Par exemple, $\begin{pmatrix} +3 & +4 \\ -1 & +1 \\ +2 & -2 \end{pmatrix}^t = \begin{pmatrix} +3 & -1 & +2 \\ +4 & +1 & -2 \end{pmatrix}$.

PROPOSITION 8.3 *On donne ici quelques propriétés de la transposition.*

1. *Quelle que soit la matrice A, on a $(A^t)^t = A$.*
2. *Si A et B sont deux matrices de type (n, m), on a $(A + B)^t = A^t + B^t$.*
3. *Si A est une matrice et λ un scalaire, on a $(\lambda A)^t = \lambda A^t$.*
4. *Si A est une matrice de type (n, m) et B une matrice de type (m, p), on a $(AB)^t = B^t A^t$.*
5. *Quelle que soit la matrice A, on a $rg(A) = rg(A^t)$.*

Démonstration : 1., 2., 3. sont totalement évidentes sur la définition de la transposition.

4. est une vérification de routine sur les expressions des différentes matrices, il suffit de jouer avec les indices. Avant de se lancer dans le calcul, le lecteur remarquera que B^tA^t a un sens car B^t est de type (p, m) et A^t de type (m, n), ce produit est donc de type (p, n) tandis que AB est de type (n, p).

5. peut se démontrer directement à partir de la définition du rang d'une matrice, mais cette démonstration est réservée aux amateurs avertis. Mieux vaut en savoir un peu plus long et utiliser la notion de matrices équivalentes.

Exercices :

1. Vérifier que pour toute matrice A de type (n, m) les produits AA^t et A^tA sont possibles, montrer que ces deux matrices sont carrées et symétriques, donner le type de chacune d'elles.
2. Montrer qu'une matrice carrée A d'ordre n est symétrique si et seulement si $A = A^t$.
3. On note $S_n(K)$ l'ensemble des matrices carrées symétriques d'ordre n à éléments dans K, et $A_n(K)$ l'ensemble des matrices carrées antisymétriques d'ordre n à éléments dans K.
   (a) Montrer que $S_n(K)$ et $A_n(K)$ sont deux sous-espaces vectoriels de $\mathcal{M}_n(K)$ et donner leurs dimensions.
   (b) Montrer que $\mathcal{M}_n(K) = S_n(K) \oplus A_n(K)$.
   (c) Donner la décomposition sur cette somme directe dans $\mathcal{M}_3(R)$ de la matrice $\begin{pmatrix} +1 & +5 & -3 \\ -3 & +2 & +3 \\ +1 & -1 & -3 \end{pmatrix}$.

8.2 Matrices carrées

On rappelle que $\mathcal{M}_{n,n}(K)$ se note, par simplification, $\mathcal{M}_n(K)$. Tout ce qui a été fait dans la première section s'applique en particulier à $\mathcal{M}_n(K)$. Mais on convient de prendre maintenant $E = F$ et $\mathcal{B} = \mathcal{C}$.

8.2.1 Structure

THEOREM 8.4 *Soient E un K-espace vectoriel et $\mathcal{B} = (e_1, ..., e_n)$ une base de E, alors :*

1. *$\mathcal{M}_n(K)$ est une K-algèbre*

2. $M_{\mathcal{B}} : \begin{array}{l} \mathcal{L}(E) \to \mathcal{M}_n(K) \\ u \mapsto M_{\mathcal{B}}(u) = M_{\mathcal{BB}}(u) \end{array}$ *est un isomorphisme d'algèbres.*

Démonstration : c'est une reformulation des théorèmes ? ? et ? ?.

PROPOSITION 8.4 *L'élément unité de $\mathcal{M}_n(K)$ est la matrice associée à l'identité de E. Cette matrice est :*

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

on la note I_n ou I.

Démonstration : évident, car $f \circ id = id \circ f = f$ et il suffit de passer aux matrices associées à l'aide de l'isomorphisme précédent□

PROPOSITION 8.5 *Dès que $n \geq 2$, le produit de matrices dans $\mathcal{M}_n(K)$ est non-commutatif et la matrice 0 admet dans $\mathcal{M}_n(K)$ des diviseurs.*

Démonstration : pour $n = 2$, on prend $A_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, $B_1 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$, alors $B_1A_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ et $A_1B_1 = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$. Pour $n > 2$, on prend $A = \begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} B_1 & 0 \\ 0 & 0 \end{pmatrix}$ (décomposition en quatre blocs des matrices A et B), alors $BA = \begin{pmatrix} B_1A_1 & 0 \\ 0 & 0 \end{pmatrix}$ et $AB = \begin{pmatrix} A_1B_1 & 0 \\ 0 & 0 \end{pmatrix}$.

Pour les diviseurs de 0, il suffit là encore de donner un exemple, or on vérifie aisément que le produit $E_{11}E_{nn}$ est nul, ou même que le carré $(E_{1n})^2$ est nul□

On remarque que dans le cas $n = 1$, les lois de $\mathcal{M}_1(K)$ sont celles de K, $\mathcal{M}_1(K)$ est donc un corps isomorphe à K, que l'on confondra souvent avec K.

On va maintenant exhiber deux sous-algèbres particulièrement importantes de $\mathcal{M}_n(K)$. Elles constituent le coeur de la théorie de la réduction des endomorphismes.

On rappelle qu'une matrice carrée $A = (\lambda_{ij})$ est dite diagonale si tous les termes en dehors de la diagonale principale sont nuls. La matrice $M_{\mathcal{B}}(u)$ est donc diagonale si et seulement si l'endomorphisme u transforme chaque vecteur de la base $\mathcal{B}$ en un vecteur proportionnel, i.e. :

$$\forall i \in [1, n], \exists \lambda_i \in K, u(e_i) = \lambda_i e_i.$$

Dès lors, $M_{\mathcal{B}}(u) = \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_n \end{pmatrix}$.

Il est immédiat que la somme et le produit de deux matrices diagonales sont encore des matrices diagonales, de même le produit par un scalaire d'une matrice diagonale est encore une matrice diagonale.

PROPOSITION 8.6 *L'ensemble noté $D_n(K)$ des matrices diagonales carrées d'ordre n à coefficients dans K est une sous-algèbre commutative de $\mathcal{M}_n(K)$.*

Les matrices de la forme $\begin{pmatrix} \lambda & 0 & ... & 0 \\ 0 & \lambda & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda \end{pmatrix} = \lambda I$ sont des cas particuliers de matrices diagonales, on les appelle **matrices scalaires**. On a une injection naturelle $\lambda \mapsto \lambda I$ de K dans $\mathcal{M}_n(K)$ qui est un morphisme d'algèbres. A l'aide de cette injection on identifie parfois les éléments de K et les matrices scalaires, en écrivant λ au lieu de λI.

On rappelle également qu'une matrice carrée $T = (t_{ij})$ est dite trigonale inférieure si on a $\forall (i,j) \in [1,n]$, $i < j \Rightarrow t_{ij} = 0$.

Exercice : Soit $\mathcal{B} = (e_1, ..., e_n)$ une base d'un K-espace vectoriel E, on pose, $\forall i \in [1,n]$, E_i le sous-espace vectoriel de E engendré par $\{e_i, e_{i+1}, ..., e_n\}$. On considère enfin un élément u de $\mathcal{L}$(E).

1. Montrer l'équivalence des propositions suivantes :

   (a) $M_{\mathcal{B}}(u)$ est trigonale inférieure

   (b) $\forall i \in [1,n]$, $u(E_i) \subset E_i$.

2. En déduire que la somme et le produit de deux matrices trigonales inférieures sont des matrices trigonales inférieures, ainsi que le produit par un scalaire d'une matrice trigonale inférieure.

Les résultats de cet exercice peuvent également se démontrer à l'aide du seul calcul matriciel.

PROPOSITION 8.7 *L'ensemble noté $T_n^-(K)$ des matrices carrées d'ordre n, à coefficients dans K, trigonales inférieures est une sous-algèbre de $\mathcal{M}_n(K)$.*

On remarque que $T_n^-(K)$ n'est pas commutative dès que $n \geq 2$ et que $D_n(K)$ est également une sous-algèbre de $T_n^-(K)$.

8.2.2 Matrices inversibles

Soit $A \in \mathcal{M}_n(K)$, on dit que A est **inversible** s'il existe une matrice $A' \in \mathcal{M}_n(K)$ telle que $AA' = A'A = I$. Si A' existe, elle est unique, on l'appelle l'**inverse** de A et on la note A^{-1}.

THEOREM 8.5 *L'ensemble noté $GL_n(K)$ des matrices inversibles de $\mathcal{M}_n(K)$ est un groupe multiplicatif d'élément neutre I. De plus, si A et B sont deux matrices inversibles de $\mathcal{M}_n(K)$, alors $(AB)^{-1} = B^{-1}A^{-1}$.*

THEOREM 8.6 *Soit $A \in \mathcal{M}_n(K)$, les conditions suivantes sont équivalentes :*

1. *A est inversible*
2. *A^t est inversible*
3. *les lignes de A (considérées comme vecteurs de K^n) sont linéairement indépendantes*
4. *les colonnes de A (idem) sont linéairement indépendantes*
5. *A est de rang n.*

Démonstration : (1.⇒2.) en effet, $AA^{-1} = A^{-1}A = I$ impliquent par transposition $(A^{-1})^t A^t = A^t (A^{-1})^t = I$, donc A^t est inversible et $(A^t)^{-1} = (A^{-1})^t$.

(2.⇒1.) idem, en échangeant les rôles de A et A^t, puisque $(A^t)^t = A$.

(1.⇔4.⇔5.) en revenant aux applications linéaires, ces applications signifient qu'un endomorphisme est inversible si et seulement si il est surjectif, ce qui est vrai en dimension finie.

(2.⇔3.) en effet, les lignes de A ne sont autres que les colonnes de A^t.

(1.⇔4.) synonyme par transposition de (2.⇔3.)□

Exercice : Déterminer lesquelles des matrices suivantes sont inversibles : $\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$, $\begin{pmatrix} +2 & -4 \\ -1 & +2 \end{pmatrix}$, $\begin{pmatrix} +1 & +0 & +2 \\ +3 & -1 & +1 \\ -2 & +4 & +3 \end{pmatrix}$, $\begin{pmatrix} +1 & +2 & +3 \\ -1 & +3 & +2 \\ -5 & +9 & +4 \end{pmatrix}$.

8.3 Changement de bases

8.3.1 Les personnages

Dans un espace vectoriel E, on dispose d'une base $\mathcal{B} = (e_1, ..., e_n)$ que l'on qualifie d"'ancienne base". Dans ce même espace vectoriel on en choisit une autre $\mathcal{B}' = (e'_1, ..., e'_n)$ qu'on appellera "nouvelle base". En bonne logique, les nouveaux vecteurs de base e'_1, ..., e'_n sont donc définis par leurs coordonnées sur l'ancienne base.

DEFINITION 8.10 *On appelle **matrice de passage** de la base $\mathcal{B}$ à la base $\mathcal{B}$' la matrice P, carrée d'ordre n, dont la j-ème colonne est formée des coordonnées du vecteur e'_j de $\mathcal{B}$' relativement à la base $\mathcal{B}$, pour tout indice j, i.e. si $P = (p_{ij})$ on a $\forall j \in [1, n]$, $e'_j = \sum_{i=1}^{n} p_{ij} e_i$.*

Exemple : Pour $e'_1 = 3e_1 + e_2$ et $e'_2 = -2e_1 + 5e_2$, alors $P = \begin{pmatrix} +3 & -2 \\ +1 & +5 \end{pmatrix}$ est la matrice de passage de (e_1, e_2) à (e'_1, e'_2).

Le lecteur est invité à bien faire attention au fait que dans le système de définition, les vecteurs de la nouvelle base sont écrits "en ligne", alors que dans la matrice de passage ils sont écrits "en colonne".

On reprend les notations précédentes. P est la matrice par rapport à la base $\mathcal{B}$ de l'application linéaire qui transforme e_1 en e'_1, e_2 en e'_2, ..., e_n en e'_n. Cette interprétation, si elle permet de vérifier qu'on se souvient bien de ce qu'est la matrice d'un endomorphisme, et même d'un automorphisme, ne sera que de peu d'intérêt par la suite. Beaucoup plus fructueuse sera la deuxième interprétation.

Pour tout indice j, la j-ème colonne de P est l'expression du vecteur e'_j sur l'ancienne base $\mathcal{B}$. P est donc la matrice de l'application identité de E, lorsque E est repéré au départ par la base $\mathcal{B}$' et à l'arrivée par la base $\mathcal{B}$: $P = M_{\mathcal{B}\mathcal{B}'}(id)$.

Il faut faire très attention à l'ordre dans lequel doivent être prises les bases : les transformations des vecteurs de départ se placent "en colonne", ce sont donc les nouveaux vecteurs de base et ils se repèrent sur la base d'arrivée, i.e. sur les anciens vecteurs de base. Tout rentre dans l'ordre avec la notation $M_{\mathcal{B}\mathcal{B}'}$, puisque l'on a convenu d'indiquer la seconde base avant la première.

THEOREM 8.7 *Soient $\mathcal{B}$ et $\mathcal{B}$' deux bases de E, P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}$' et P' la matrice de passage de $\mathcal{B}$' à $\mathcal{B}$, alors $PP' = I$ et $P'P = I$, i.e. $P' = P^{-1}$.*

Démonstration : en effet, si l'on considère le diagramme $\begin{matrix} E & \overset{id}{\underset{P}{\rightarrow}} & E & \overset{id}{\underset{P'}{\rightarrow}} & E \\ \mathcal{B}' & & \mathcal{B} & & \mathcal{B}' \end{matrix}$, en passant aux matrices associées relativement aux bases indiquées, on a alors :

$$M_{\mathcal{B}'\mathcal{B}}(id).M_{\mathcal{B}\mathcal{B}'}(id) = M_{\mathcal{B}'\mathcal{B}'}(id),$$

i.e. $P'P = I$; de même pour $PP'\boxdot$

Réciproquement, soient P une matrice inversible d'ordre n et $\mathcal{B}$ une base de E, alors P est la matrice de passage de la base $\mathcal{B}$ en une unique base $\mathcal{B}$', d'après la première interprétation d'une matrice de passage.

8.3.2 Action sur les coordonnées

Soit x un élément de E, on note X la matrice colonne de ses coordonnées dans l'ancienne base $\mathcal{B}$ et X' la matrice colonne de ses coordonnées dans la nouvelle base $\mathcal{B}$'. Comme $x = id(x)$, on a $M_{\mathcal{B}}(x) = M_{\mathcal{B}\mathcal{B}'}(id).M_{\mathcal{B}'}(x)$, donc si P est la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}$', on obtient :

$$X = PX'.$$

Une fois de plus, il faut faire attention à l'ordre. On dispose de la matrice de passage de l'ancienne à la nouvelle, on connaît donc les nouveaux vecteurs de base en fonction des anciens. Mais ce sont les anciennes coordonnées que l'on exprime en fonction des nouvelles à l'aide de cette matrice de passage. C'est d'ailleurs dans ce sens qu'elles seront le plus souvent utiles. Evidemment on peut, si cela est nécessaire, exprimer X' en fonction de X : $X' = P^{-1}X$.

Exemple : Soient E un plan vectoriel euclidien et $\mathcal{B} = (\vec{i}, \vec{j})$ une base orthonormée de E, on pose $\mathcal{B}' = (\vec{i}', \vec{j}')$ la base obtenue en effectuant sur la base $\mathcal{B}$ une rotation d'angle θ. Soit $\vec{u}$ un vecteur de E de coordonnées $\begin{pmatrix} x \\ y \end{pmatrix}$ sur $\mathcal{B}$ et $\begin{pmatrix} x' \\ y' \end{pmatrix}$ sur $\mathcal{B}'$, on a :

$$\begin{aligned} \vec{i}' &= r_\theta(\vec{i}) = \cos(\theta)\vec{i} + \sin(\theta)\vec{j}, \\ \vec{j}' &= r_\theta(\vec{j}) = r_{\theta+\frac{\pi}{2}}(\vec{i}) = -\sin(\theta)\vec{i} + \cos(\theta)\vec{j}. \end{aligned}$$

La matrice de passage de $\mathcal{B}$ à $\mathcal{B}$' est donc :

$$P = \begin{pmatrix} +\cos(\theta) & -\sin(\theta) \\ +\sin(\theta) & +\cos(\theta) \end{pmatrix}.$$

On a $\begin{pmatrix} x \\ y \end{pmatrix} = P \begin{pmatrix} x' \\ y' \end{pmatrix}$, i.e. $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x'\cos(\theta) - y'\sin(\theta) \\ x'\sin(\theta) + y'\cos(\theta) \end{pmatrix}$.

8.3.3 Action sur les matrices

Soient E et F deux K-espaces vectoriels de dimension finie et u un élément de $\mathcal{L}$(E, F), si E et F sont munis chacun d'une base, $\mathcal{B} = (e_1, ..., e_m)$ pour E et $\mathcal{C} = (f_1, ..., f_n)$ pour F, alors l'application linéaire u s'exprime à l'aide d'une matrice que l'on a notée $M_{\mathcal{C}\mathcal{B}}(u)$. Il est clair que si l'on change la base $\mathcal{B}$ pour une base $\mathcal{B}$' et la base $\mathcal{C}$ pour une base $\mathcal{C}$', la matrice de u relativement à ces nouvelles bases va être en général tout à fait différente de la précédente. Le problème posé est donc de connaître le lien entre ces matrices.

On note P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}$' et Q celle de $\mathcal{C}$ à $\mathcal{C}$'. On note aussi $A = M_{\mathcal{CB}}(u)$ et $A' = M_{\mathcal{C'B'}}(u)$, et on considère le diagramme suivant :

$$\begin{array}{ccc} E & \overset{u}{\underset{A'}{\longrightarrow}} & F \\ (\mathcal{B}') & & (\mathcal{C}') \\ id_E \downarrow P & & Q \downarrow id_F \\ E & \overset{u}{\underset{A}{\longrightarrow}} & F \\ (\mathcal{B}) & & (\mathcal{C}) \end{array} .$$

On a $id_F \circ u = u \circ id_E$, donc en passant aux matrices associées relativement aux bases indiquées dans le diagramme, $M_{\mathcal{CB}'}(id_F \circ u) = M_{\mathcal{CC}'}(id_F).M_{\mathcal{C'B'}}(u) = QA'$ et $M_{\mathcal{CB}'}(u \circ id_E) = M_{\mathcal{CB}}(u).M_{\mathcal{BB}'}(id_E) = AP$ donc $QA' = AP$, ce qui peut encore s'écrire :

$$A' = Q^{-1}AP.$$

Encore une fois, il faut faire attention à la place des différentes matrices.

8.3.4 Matrices équivalentes

DEFINITION 8.11 *Deux matrices A et A' de type (n, m) sont dites* ***équivalentes*** *s'il existe une matrice inversible R carrée d'ordre n et une matrice inversible S carrée d'ordre m, telles que :*

$$A' = RAS.$$

Cette définition est bien entendu directement inspirée par la formule précédente.

PROPOSITION 8.8 *Si A est la matrice d'une application linéaire u de E dans F relativement à deux bases $\mathcal{B}$ et $\mathcal{C}$, alors A est équivalente à A' si et seulement si A' est la matrice de u par rapport à des bases $\mathcal{B}$' et $\mathcal{C}$' de E et de F.*

Démonstration : on suppose A' équivalente à A, donc $A' = RAS$. S étant inversible d'ordre m, c'est la matrice de passage de la base $\mathcal{B}$ à une base $\mathcal{B}$' de E. De même, R^{-1} est la matrice de passage de la base $\mathcal{C}$ à une base $\mathcal{C}$' de F. Alors la matrice de u par rapport aux bases $\mathcal{B}$' et $\mathcal{C}$' est $(R^{-1})^{-1}AS = RAS = A'$.

Réciproquement, si $A = M_{\mathcal{BC}}(u)$ et $A' = M_{\mathcal{C'B'}}(u)$, on a alors $A' = Q^{-1}AP$, il suffit donc de prendre $Q^{-1} = R$, $P = S\boxdot$

Exercice : Utiliser la proposition précédente pour en déduire que deux matrices équivalentes ont le même rang.

8.3.5 Matrices semblables

Dans le cas des matrices carrées, la notion de matrices équivalentes perd une grande partie de son intérêt, car on a l'habitude d'asocier, à une matrice carrée, un endomorphisme relativement à une base $\mathcal{B}$, la même au départ et à l'arrivée. Il faut donc adapter la définition précédente à ce cadre plus restrictif.

DEFINITION 8.12 *Deux matrices A et A' carrées d'ordre n sont dites* ***semblables*** *s'il existe une matrice inversible P carrée d'ordre n telle que :*

$$A' = P^{-1}AP.$$

PROPOSITION 8.9 *Si A est la matrice d'un endomorphisme u de E par rapport à une base $\mathcal{B}$, alors A' est semblable à A si et seulement si A' est la matrice de u relativement à une base $\mathcal{B}$' de E.*

Démonstration : laissée au lecteur, qui doit seulement adapter celle de la proposition ? ?.