Université Paris Sud Orsay Préparation à l'Agrégation interne de Mathématiques

Algèbre

Notes de cours

J. HENRY
juillet 2006

Table des matières

1

Ensembles, relations, applications, lois de composition

1.1 Relations binaires

DEFINITION 1.1.1
Soient E et F deux ensembles. Une relation binaire de E vers F est un sous ensemble $\mathcal{R}$ du produit cartésien $E \times F$. Si $F = E$ on dit aussi relation binaire sur E. Si $x \in E$ et $y \in F$, on utilise la notation $x\mathcal{R}y$ pour $(x, y) \in \mathcal{R}$. (On dit quelquefois que l'ensemble $\mathcal{R}$ est le graphe de la relation $\mathcal{R}$).

DEFINITION 1.1.2
Soit $\mathcal{R}$ une relation binaire de E vers F.
On appelle domaine de $\mathcal{R}$ le sous ensemble de E $D(\mathcal{R}) = \{x \in E \mid \exists y \in F,\ x\mathcal{R}y\}$.
On appelle image de $\mathcal{R}$ le sous ensemble de F $\mathrm{im}(\mathcal{R}) = \{y \in F \mid \exists x \in E,\ x\mathcal{R}y\}$.

DEFINITION 1.1.3
Soit $\mathcal{R}$ une relation binaire de E vers F. La relation réciproque notée $\mathcal{R}^{-1}$ est une relation binaire de F vers E définie par

$$\forall y \in F, \forall x \in E, (y, x) \in \mathcal{R}^{-1} \Leftrightarrow (x, y) \in \mathcal{R}$$

DEFINITION 1.1.4
Soient E, F, G trois ensembles, $\mathcal{R}$ une relation binaire de E vers F et $\mathcal{S}$ une relation binaire de F vers G. La relation composée $\mathcal{S} \circ \mathcal{R}$ est la relation de E vers G définie par $\mathcal{S} \circ \mathcal{R} = \{(x, z) \in E \times G \mid \exists y \in F,\ (x\mathcal{R}y)$ et $(y\mathcal{S}z)\}$.

PROPOSITION 1.1.1
Soient E, F, G, H quatre ensembles et $\mathcal{R}, \mathcal{S}, \mathcal{T}$ des relations binaires de E vers F, de F vers G et de G vers H respectivement. On a

$$\mathcal{T} \circ (\mathcal{S} \circ \mathcal{R}) = (\mathcal{T} \circ \mathcal{S}) \circ \mathcal{R}$$

DEFINITION 1.1.5
Soit $\mathcal{R}$ une relation binaire sur un ensemble E et E' un sous ensemble de E. La restriction de $\mathcal{R}$ à E', quelquefois notée $\mathcal{R}|_{E'}$ est la relation $\mathcal{R}'$ sur E' définie par $\mathcal{R}' = \mathcal{R} \cap (E' \times E')$.

1.2 Relations fonctionelles

DEFINITION 1.2.1
Une relation binaire $\mathcal{R}$ de E vers F est dite fonctionelle si pour tout $x \in D(\mathcal{R})$ il existe un unique $y \in F$ tel que $x\mathcal{R}y$. Dans ce cas on utilise plutôt une notation fonctionelle : on notera $y = \mathcal{R}(x)$ plutôt que $x\mathcal{R}y$. On dit que y est l'image de x par $\mathcal{R}$ et que x est un antécédent de y par $\mathcal{R}$.

Une application f de E vers F est une relation fonctionelle de E vers F dont le domaine est E tout entier. On utilise la notation $f : E \to F$. Toute relation fonctionelle $\mathcal{R}$ induit une application de $D(\mathcal{R})$ vers F ainsi qu'une application de $D(\mathcal{R})$ vers $\mathrm{im}(\mathcal{R})$.

DEFINITION 1.2.2
Une application $f : E \to F$ est
injective si $\forall x, x' \in E, f(x) = f(x') \Rightarrow x = x'$.
surjective si $\mathrm{im}(f) = F$, i.e. si $\forall y \in F,\ \exists x \in E,\ y = f(x)$.
bijective si elle est à la fois injective et surjective.

EXEMPLE 1.2.1
Soit E un ensemble. La diagonale $\Delta_E = \{(x, x)\ ;\ x \in E\}$ définit une relation binaire sur E qui est une bijection appellée application identique de E vers E et notée Id_E.

EXEMPLE 1.2.2
Soient E un ensemble et $E' \subset E$. On appelle application inclusion ou inclusion $i : E' \to E$ l'application définie par $i = \{(x, y) \in E' \times E\ ;\ y = x\}$.

Restriction, application induite

Il est important de bien noter que la donnée d'une application f est en fait la donnée d'un triplet, (E, F, f) où E et F sont deux ensembles et f un sous ensemble de $E \times F$.

Soient $f : E \to F$ une application et $E' \subset F$. La restriction de f à E' notée $f|_{E'}$ est l'application de E' dans F qui à tout x de E' associe $f(x) \in F$. Formellement c'est $E' \times F \cap f$.

Supposons maintenant que E' soit un sous ensemble de E, que F' soit un sous ensemble de F et que pour tout x de E' on ait $f(x) \in F'$. On peut alors considérer l'application $g : E' \to F'$ définie par $\forall x \in E',\ g(x) = f(x)$. Formellement c'est $E' \times F' \cap f$. Cette application est appelée application induite par f. Si besoin on la notera $f_{E'}^{F'}$. Il faut bien prendre garde à ne pas la condondre avec la restriction de f à E'. Ces deux applications ne sont égales que dans le cas où $F' = F$.

THEOREME 1.2.1
Soient E, F, G trois ensembles.
1) *Soient $\mathcal{R}$ une relation de E vers F et $\mathcal{S}$ une relation de F vers G. Si $\mathcal{R}$ et $\mathcal{S}$ sont des relations fonctionelles (resp. des applications) il en est de même de la composée $\mathcal{S} \circ \mathcal{R}$.*
2) *Soient $f : E \to F$ et $g : F \to G$ des applications. Si f et g sont injectives(resp. surjectives, bijectives), il en est de même de $g \circ f$.*
3 *Si $f : E \to F$ est bijective, la relation réciproque de f est une application de F vers E qui est bijective. On l'appelle bijection réciproque de f. On a alors $f^{-1} \circ f = Id_E$ et $f \circ f^{-1} = Id_F$.*

Applications induites sur l'ensemble des parties

Etant donné un ensemble X, on notera $\mathcal{P}(X)$ l'ensemble des sous ensembles de X ou ensemble des parties de X.

Soit $f : E \to F$ une application de E vers F.

On définit une application $f_* : \mathcal{P}(E) \to \mathcal{P}(F)$ par $\forall X \in \mathcal{P}(E), f_*(X) = \{f(x);\ x \in X\} = \{y \in F \mid \exists x \in X,\ y = f(x)\}$.
Pour tous $X, X' \subset E$, on a :
$f_*(X \cup X') = f_*(X) \cup f_*(X')$.
$f_*(X \cap X') \subset f_*(X) \cap f_*(X')$.
En général la dernière inclusion n'est pas une égalité. De même il n'y a en général aucune relation d'inclusion entre $f_*(E \setminus X)$ et $F \setminus f_*(X)$. Notons enfin que si $x \in E$, on a $f_*(\{x\}) = \{f(x)\}$.
L'ensemble $f_*(X)$ s'appelle l'image par f du sous ensemble X de E. Si $g : F \to G$ est une autre application, on vérifie facilement que $(g \circ f)_* = g_* \circ f_*$.

On définit aussi une application $f^* : \mathcal{P}(F) \to \mathcal{P}(E)$ par $\forall Y \in \mathcal{P}(F),\ f^*(Y) = \{x \in E\ ;\ f(x) \in Y\}$. Cette application vérifie, pour tout $Y, Y' \subset F$,
$f^*(Y \cup Y') = f^*(Y) \cup f^*(Y')$.

$f^*(Y \cap Y') = f^*(Y) \cap f^*(Y')$.
$f^*(F \setminus Y) = E \setminus f^*(Y)$.
Le sous ensemble $f^*(Y)$ de E s'appelle l'image réciproque par f du sous ensemble Y de F.
Si $g : F \to G$ est une autre application, on a $(g \circ f)^* = f^* \circ g^*$.

Supposons maintenant que f soit une bijection de E sur F. Il est facile de vérifier que f_* et f^* sont des bijections de $\mathcal{P}(E)$ sur $\mathcal{P}(F)$ et de $\mathcal{P}(F)$ sur $\mathcal{P}(E)$ respectivement et que ces bijections sont réciproques l'une de l'autre. On a donc $f^* = (f_*)^{-1}$. Dans ce cas, l'application f_* vérifie $f_*(X \cap X') = f_*(X) \cap f_*(X')$ et $f_*(E \setminus X) = F \setminus f_*(X)$.

Traditionellement, l'application f_* est notée f : on écrit $f(X) = \{f(x);\ x \in X\}$. Ceci est en partie justifié par le fait que si on identifie E à l'ensemble des parties de E ayant un seul élément, f_* est un prolongement de f. Un autre abus consiste à noter f^{-1} pour f^* que f soit bijective ou non. Remarquons que si f est bijective, on a $f^* = \left(f^{-1}\right)_*$ et que dans ce cas le nouvel abus de notation est cohérent avec le premier.
Dans la suite nous nous conformerons à ces usages et nous n'utiliserons pas les notations f_* et f^*.

1.3 Lois de composition

DEFINITION 1.3.1
Soit X un ensemble. Une loi de composition interne $*$ sur X est une application de $X \times X$ vers X. Si $(x, y) \in X \times X$, l'image du couple (x, y) par la loi $*$ est noté $x * y$ plutôt que $*(x, y)$.
On appelle quelquefois magma un couple $(X, *)$ où X est un ensemble et $*$ une loi de composition interne sur X.

La loi $*$ est dite associative si, pour tous $x, y, z \in X$ on a $(x*y)*z = x*(y*z)$. Dans ce cas, on note $x*y*z$ la valeur commune précédente et pour tout entier naturel $n \geqslant 1$ on définit par récurrence x^n en posant $x^1 = x$ et pour $n \geqslant 2,\ x^n = (x^{n-1}) * x$. Alors, pour tout couple d'entiers $p, q \geqslant 1$ on a $x^p * x^q = x^{p+q}$ et $(x^p)^q = x^{pq}$.
La loi $*$ est dite commutative si pour tous $x, y \in X$ on a $x * y = y * x$.

Remarque
Une loi interne est notée additivement si on utilise le symbole $+$ au lieu de $*$. On n'utilise cette notation que pour des lois associatives et commutatives. Dans ce cas, x^n est noté $n \cdot x$ et on a $(p + q) \cdot x = p \cdot x + q \cdot x$ et $q \cdot (p \cdot x) = (pq) \cdot x$.

Remarque
Si la loi $*$ n'est pas associative, on peut toujours définir x^n comme ci dessus. Mais on n'a plus, en général $x^p * x^q = x^{p+q}$.

Un élément $e \in X$ est dit neutre pour la loi $*$ si, pour tout $x \in X$ on a $e * x = x * e = x$. Un tel élément si il existe est nécessairement unique. Si la loi est notée additivement, l'élément neutre éventuel est souvent noté 0.

Soit $*$ une loi interne sur un ensemble X admettant un élément neutre e. Soit $x \in X$. Si il existe un élément $x' \in X$ tel que $x * x' = x' * x = e$ on dit que x' est un symétrique de x. Si la loi $*$ est associative, un tel symétrique est unique. En effet, si x' et x'' sont deux symétriques de x, on a $x' = x' * e = x' * (x * x'') = (x' * x) * x'' = e * x'' = x''$.
Pour une loi notée additivement (donc associative et commutative) le symétrique de x si il existe est appelé l'opposé de x et noté $-x$.
Pour une loi associative notée multiplicativement, un élément x qui admet un symétrique est dit inversible. Le symétrique est aussi appelé inverse de x et est noté x^{-1}.

Soit X un ensemble, $*$ une loi sur X associative, admettant un élément neutre noté 1_X. On pose, pour $x \in X,\ x^0 = 1_X$. Dans ce cas, on a encore pour $p, q \in \mathbb{N},\ x^p * x^q = x^{p+q}$. Si x est inversible, on étend la définition de la manière suivante ; on note x^{-1} l'inverse de x, et pour $n \in \mathbb{N}^*$, on pose $x^{-n} = \left(x^{-1}\right)^n$. On vérifie sans peine que les relations $x^p * x^q = x^{p+q}$ et $(x^p)^q = x^{pq}$ subsistent pour tous $p, q \in \mathbb{Z}$. En particulier, x^{-n} est l'inverse de x^n.

Soient $(X, *)$ et $(Y, \top)$ deux magmas. Un morphisme de $(X, *)$ dans $(Y, \top)$ est une application $f : X \to Y$ vérifiant $\forall x, x' \in X,\ f(x * x') = f(x) \top f(x')$.
Un sous ensemble X' de X est stable par $*$ si pour tous $x, y \in X',\ x * y \in X'$. L'application $*$ induit alors une loi de composition interne sur X' encore notée $*$.

1.4 Relations binaires sur un ensemble

DEFINITION 1.4.1
Soient E un ensemble et $\mathcal{R}$ une relation binaire sur E. $\mathcal{R}$ est
1)réflexive si $\forall x \in E,\ x\mathcal{R}x$.
2) symétrique si $\forall (x,y) \in E^2, x\mathcal{R}y \Rightarrow y\mathcal{R}x$.
3) antisymétrique si $\forall (x,y) \in E^2, (x\mathcal{R}y$ et $y\mathcal{R}x) \Rightarrow y = x$
4) transitive si $\forall x,y,z \in E, (x\mathcal{R}y$ et $y\mathcal{R}z) \Rightarrow x\mathcal{R}z$

Remarque
$\mathcal{R}$ est réflexive ssi $\Delta_E \subset \mathcal{R}$.
$\mathcal{R}$ est symétrique ssi $\forall x,y \in E, x\mathcal{R}y \Leftrightarrow y\mathcal{R}x$ ou encore ssi $\mathcal{R} = \mathcal{R}^{-1}$.
$\mathcal{R}$ est antisymétrique ssi $\mathcal{R} \cap \mathcal{R}^{-1} \subset \Delta_E$
$\mathcal{R}$ est transitive ssi $\mathcal{R} \circ \mathcal{R} \subset \mathcal{R}$.

1.5 Relations d'équivalence

1.5.1 Relation d'équivalence et partition

DEFINITION 1.5.1
Soit E un ensemble non vide. Une relation d'équivalence sur E est une relation binaire sur E qui est réflexive, symétrique et transitive.

DEFINITION 1.5.2
Soit E un ensemble non vide. Une partition de E est un ensemble de sous ensembles non vides de E, deux à deux disjoints dont la réunion est E.

Soit $\mathcal{R}$ une relation d'équivalence sur E. Pour $x \in E$, on appelle classe d'équivalence de x, ou plus brièvement classe de x le sous ensemble $Cl_{\mathcal{R}}(x) = \{y \in E\ ;\ x\mathcal{R}y\}$ constitué des éléments de E équivalents à x. La classe de x contient x.
Soient $x,y \in E$. On a la dichotomie suivante :
ou $y\mathcal{R}x$ ce qui équivaut à $y \in Cl_{\mathcal{R}}(x)$ et alors $Cl_{\mathcal{R}}(x) = Cl_{\mathcal{R}}(y)$.
ou $y \notin Cl_{\mathcal{R}}(x)$ et alors $Cl_{\mathcal{R}}(x) \cap Cl_{\mathcal{R}}(y) = \emptyset$. L'ensemble des classes d'équivalence des éléments de E constituent donc une partition de E. Réciproquement, étant donnée une partition $\mathcal{P}$ de E, on définit la relation binaire $\mathcal{R}_{\mathcal{P}}$ sur E par $x\mathcal{R}_{\mathcal{P}}y \Leftrightarrow \exists A \in \mathcal{P}$ tel que $x \in A$ et $y \in A$. Il est facile de voir que cette relation est une relation d'équivalence, que la classe de x est l'ensemble A de la partition $\mathcal{P}$ qui contient x.

On a donc une correspondance biunivoque entre les relations d'équivalence sur un ensemble non vide E donné et les partitions de cet ensemble.

DEFINITION 1.5.3
Soit $\mathcal{R}$ une relation d'équivalence sur un ensemble E. On appelle ensemble quotient de E par $\mathcal{R}$ et on note $E/\mathcal{R}$ le sous ensemble de $\mathcal{P}(E)$ formé des classes d'équivalence de $\mathcal{R}$.
L'application $p_{\mathcal{R}} : E \to E/\mathcal{R}$ qui à $x \in E$ associe $p_{\mathcal{R}}(x) = Cl_{\mathcal{R}}(x)$ est surjective. On l'appelle projection canonique ou surjection canonique.

1.5.2 Factorisation d'applications

PROPOSITION 1.5.1
Soient E,F deux ensembles non vides, $\mathcal{R}$ une relation d'équivalence sur E et $f : E \to F$ une application de E vers F. Il existe une application $\bar{f} : E/\mathcal{R} \to F$ vérifiant $\bar{f} \circ p_{\mathcal{R}} = f$ ssi f est constante sur chaque classe d'équivalence de $\mathcal{R}$. Dans ce cas l'application $\bar{f}$ est unique.

PROPOSITION 1.5.2
Soient E,F deux ensembles non vides et $f : E \to F$ une application. La relation binaire sur E, $\mathcal{R}_f$, définie par $\forall (x,y) \in E^2,\ x\mathcal{R}_f y \Leftrightarrow f(x) = f(y)$ est une relation d'équivalence dite relation d'équivalence associée à f.

THEOREME 1.5.1 (Factorisation canonique d'une application)
Soient E et F deux ensembles non vides et $f : E \to F$ une application. Soit $\mathcal{R}_f$ la relation d'équivalence associée à f. Soient $p : E \to E/\mathcal{R}_f$ la projection canonique et $i : \mathrm{im}(f) \to F$ l'inclusion. Il existe une application $\bar{f} : E/\mathcal{R}_f \to \mathrm{im}(f)$ et une seule telle que $f = i \circ \bar{f} \circ p$ et cette application $\bar{f}$ est une bijection de $E/\mathcal{R}_f$ sur $\mathrm{im}(f)$.

Pour exprimer la relation $f = i \circ \bar{f} \circ p$, on dit aussi que le diagramme

$$\begin{array}{ccc} E & \xrightarrow{f} & F \\ {\scriptstyle p}\downarrow & & \uparrow{\scriptstyle i} \\ E/\mathcal{R}_f & \xrightarrow{\bar{f}} & \mathrm{im} f \end{array}$$

est commutatif.

1.5.3 Relation d'équivalence compatible avec une loi interne

DEFINITION 1.5.4
Soient X un ensemble non vide, $*$ une loi interne sur X et $\mathcal{R}$ une relation binaire sur X. On dit que $\mathcal{R}$ est compatible avec la loi $*$ si pour tous x, x', y de X on a $x\mathcal{R}x' \Rightarrow (x * y)\mathcal{R}(x' * y)$ et $(y * x)\mathcal{R}(y * x')$.

Si $\mathcal{R}$ est une relation transitive , en particulier si c'est une relation d'équivalence, compatible avec la loi $*$, on a facilement

$$\left\{ \begin{array}{l} x\mathcal{R}x' \\ y\mathcal{R}y' \end{array} \right. \Rightarrow x * y \,\mathcal{R}\, x' * y'$$

Soit $\mathcal{R}$ une relation d'équivalence sur X compatible avec la loi interne $*$. Soient α et β deux classes d'équivalences i.e. deux éléments de $E/\mathcal{R}$. Si on choisit deux éléments quelconques x, x' dans α et deux éléments quelconques y, y' dans β les éléments $x * y$ et $x' * y'$ sont équivalents. La classe d'équivalence de $x * y$ est donc indépendante des choix de x et de y et ne dépend que de α et β. On notera cette classe $\alpha \bar{*} \beta$. On définit ainsi une loi de composition interne sur l'ensemble quotient $X/\mathcal{R}$. On dit que la loi $*$ passe aux quotients en la loi $\bar{*}$. La projection canonique $p : X \to X/\mathcal{R}$ est alors un morphisme surjectif. Il est facile de vérifier que si $*$ est associative, resp. commutative, il en est de même de $\bar{*}$. Si $*$ admet un élément neutre e, l'élément $\bar{e}$ est neutre pour $\bar{*}$. Dans ce cas si un élément $x \in X$ admet un symétrique x' pour $*$, sa classe $\bar{x}$ admet un symétrique pour la loi $\bar{*}$ qui n'est autre que $\overline{x^{-1}}$.

1.6 Relations d'ordre

DEFINITION 1.6.1
Une relation binaire sur un ensemble E est une relation d'ordre si elle est réflexive, antisymétrique et transitive.

Une telle relation est souvent notée $\leqslant$. On note en général $\geqslant$ la relation réciproque de $\leqslant$. C'est aussi une relation d'ordre.

La restriction d'une relation d'ordre sur E à un sous ensemble E' de E est encore une relation d'ordre.

Un ensemble ordonné est un couple $(E, \leqslant)$ où E est un ensemble et $\leqslant$ une relation d'ordre sur E.

Soit $\leqslant$ une relation d'ordre sur un ensemble E. Deux éléments x et y de E sont dits comparables si on a $x \leqslant y$ ou $y \leqslant x$. L'ordre $\leqslant$ est dit total si deux éléments quelconques sont comparables.

Un élément $M \in E$ est un maximum si $\forall x \in E,\ x \leqslant M$. Par transitivité, si un tel maximum existe, il est unique. Dans ce cas, on le notera $M = \max E$. On définit de même un minimum.

Un élément $a \in E$ est dit maximal si $\forall x \in E,\ a \leqslant x \Rightarrow a = x$ autrement dit, si il n'y a pas d'élément de E strictement supérieur à a. On définit de même un élément minimal. Si E admet un maximum M, M est un élément maximal et c'est le seul.

EXEMPLE 1.6.1
Soit X un ensemble et $\mathcal{P}(X)$ l'ensemble des parties de X. La relation d'inclusion $\subset$ est une relation d'ordre sur $\mathcal{P}(X)$. Cet ordre n'est pas total dès que $\mathrm{card}(X) \geqslant 2$. L'élément X est le maximum, l'élément $\emptyset$ le minimum.

EXEMPLE 1.6.2
On considère dans $\mathbb{N} \setminus \{0,1\}$ la relation a divise b notée $a|b$. C'est une relation d'ordre partiel. Les éléments minimaux pour cette relation sont les nombres premiers.

Soit $(E, \leqslant)$ un ensemble ordonné. Un sous ensemble A de E est dit majoré (resp. minoré) si il existe un élément $M \in E$ (resp. $m \in E$) tel que $\forall x \in A,\ x \leqslant M$ (resp. $m \leqslant x$). M (resp. m) est appelé un majorant (resp. un minorant) de A.

DEFINITION 1.6.2
Soit $(E, \leqslant)$ un ensemble ordonné et A un sous ensemble de E majoré.
Si l'ensemble $\mathrm{Maj}(A)$ des majorants de A admet un plus petit élément, celui ci, nécessairement unique, s'appelle la borne supérieure de A et se note $\sup(A)$.
De même, si A est minoré et si l'ensemble des minorants de A admet un plus grand élément, celui ci nécessairement unique s'appelle la borne inférieure de A et se note $\inf(A)$.

Caractérisation des bornes supérieures et inférieures
Soient $(E, \leqslant)$ un ensemble ordonné, $A \subset E, A \neq \emptyset$ et $b \in E$. On a l'équivalence :

$$b = \sup(A) \Leftrightarrow \begin{cases} \forall x \in A,\ x \leqslant b \\ \forall M \in E,\ (\forall x \in A,\ x \leqslant M) \Leftrightarrow (b \leqslant M) \end{cases}$$

De même, si $a \in E$, on a l'équivalence :

$$a = \inf(A) \Leftrightarrow \begin{cases} \forall x \in A,\ a \leqslant x \\ \forall m \in E,\ (\forall x \in A,\ m \leqslant x) \Leftrightarrow (m \leqslant a) \end{cases}$$

Supposons maintenant que la relation d'ordre $\leqslant$ soit **totale**. On a alors les caractérisations suivantes des bornes supérieures et inférieures :

$$b = \sup(A) \Leftrightarrow \begin{cases} \forall x \in A,\ x \leqslant b \\ \forall \beta \in E,\ \beta < b \Rightarrow (\exists x \in A\ \beta < x) \end{cases} \qquad a = \inf(A) \Leftrightarrow \begin{cases} \forall x \in A,\ a \leqslant x \\ \forall \alpha \in E,\ \alpha > a \Rightarrow (\exists x \in A\ x < \alpha) \end{cases}$$

2

Groupes, anneaux, corps

2.1 Groupes

2.1.1 Definition et propriétés élémentaires

DEFINITION 2.1.1
Un groupe est un couple $(G, *)$ où G est un ensemble et $*$ une loi de composition interne sur G, associative, admettant un élément neutre et telle que tout élément de G admet un symétrique. Le groupe est dit commutatif ou abélien si la loi $*$ est commutative.

Il résulte de la définition que si $(G, *)$ est un goupe, G **n'est pas vide**.

Pour une loi notée multiplicativement, on désignera l'élément neutre par 1_G où même 1 si il n'y a pas d'ambiguïté et le symétrique de x sera noté x^{-1}. Pour une loi commutative notée additivement, l'élément neutre sera noté 0_G où même 0 et le symétrique de x, $-x$. Si a et b sont deux éléments de G, on a $(a * b)^{-1} = b^{-1} * a^{-1}$.

Dans un groupe, tout élément a est simplifiable à droite et à gauche, i.e. l'égalité $x * a = y * a$ (resp $a * x = a * y$) implique $x = y$. Si a, b sont deux éléments d'un groupe $(G, *)$, l'équation $a * x = b$ (resp. $x * a = b$) admet une solution et une seule $x = a^{-1} * b$ (resp. $x = b * a^{-1}$).

$(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{Q}_+^*, \times)$, $(Q^*, \times)$, $(\mathbb{R}, +)$, $(\mathbb{R}_+^*, \times)$, $(\mathbb{R}^*, \times)$, $(\mathbb{C}, +)$, $(\mathbb{C}^*, \times)$ sont des groupes.

EXEMPLE 2.1.1
Soit E un ensemble non vide. L'ensemble des bijections de E sur E, muni de la loi $\circ$ de composition des applications est un groupe, appellé groupe des permutations de E et noté $\mathfrak{S}_E$. Dans le cas où $E = \{1, \ldots, n\}$ ce groupe est noté $\mathfrak{S}_n$.

EXEMPLE 2.1.2
Soit X un ensemble non vide et $(G, *)$ un groupe. Soit G^X l'ensemble des applications de X dans G. Pour $f, g \in G^X$ on définit une nouvelle application $f \circledast g : X \to G$ par $\forall x \in X, f \circledast g(x) = f(x) * g(x)$. Alors $(G^X, \circledast)$ est un groupe d'élément neutre l'application constante $x \to 1_G$. Dans la pratique, on note $f * g$ pour $f \circledast g$. Si G est abélien, il en est de même de G^X.

EXEMPLE 2.1.3 (Groupe produit)
Soient $(G, *)$ et $(H, \top)$ deux groupes. Pour $(x, y) \in G \times H$ et $(x', y') \in G \times H$ on pose $(x, y) \star (x', y') = (x * x', y \top y')$. On vérifie sans peine que $(G \times H, \star)$ est un groupe appelé groupe produit des groupes $(G, *)$ et $(H, \top)$. On définit de la même manière le produit cartésien d'un nombre fini de groupes. Par exemple $\mathbb{R}^n$ muni de l'addition usuelle des n-uplets est un groupe. Si G et H sont abéliens, il en est de même de leur produit.

Exercice
Soit E un ensemble non vide. Si A, B sont des sous ensembles de E, on définit leur différence symétrique $A \Delta B$ par $A \Delta B = A \cup B \setminus A \cap B = (A \setminus B) \cup (B \setminus A)$. L'ensemble $\mathcal{P}(E)$ des parties de E, muni de la loi Δ est un groupe abélien dont l'élément neutre est l'ensemble vide.

2.1.2 Sous groupes

DEFINITION 2.1.2
Soit G un groupe. Une partie H de G est un sous groupe si elle est stable par $*$ et si elle vérifie les deux propriétés suivantes :
1) $1_G \in H$.
2) $\forall x \in G,\ x \in H \Rightarrow x^{-1} \in H$.

$\{1_G\}$ et G sont des sous groupes de G.

THEOREME 2.1.1
Soit $(G, *)$ un groupe et H une partie de G. Les propriétés suivantes sont équivalentes :
1) *H est un sous groupe de G.*
2) *H est stable par $*$ et pour la loi induite, $(H, *)$ est un groupe.*
3) *H est non vide et $\forall x, y \in G,\ (x \in H$ et $y \in H) \Rightarrow x^{-1} * y \in H$.*

preuve partielle
Prouvons 2) $\Rightarrow$ 1). Notons 1_H l'élément neutre du groupe H. On a $1_H * 1_G = 1_H$ car 1_G est neutre dans G et $1_H * 1_H = 1_H$ donc, 1_H étant simplifiable $1_H = 1_G$. Ensuite, soit $x \in H$ et x' son symétrique dans le groupe $(H, *)$. On a $1_H = x * x'$ et $1_G = x * x^{-1}$ donc x étant simplifiable, puisque $1_H = 1_G$, $x' = x^{-1}$ ce qui prouve que $x^{-1} \in H$.
Les autres implications sont laissées au lecteur.

EXEMPLE 2.1.4
L'ensemble $\mathbb{U}$ des nombres complexes de module 1 est un sous groupe de $(\mathbb{C}^*, \times)$.

EXEMPLE 2.1.5
Soit E un espace affine euclidien. L'ensemble des isométries affines de E, noté Iso(E) est un sous groupe du groupe $\mathfrak{S}_E$ des bijections de E.
Si $A \subset E$ est une partie de E, l'ensemble $\{f \in \text{Iso}(E) \mid f(A) = A\}$ est un sous groupe de Iso(E).

EXEMPLE 2.1.6 (Centre d'un groupe)
Soit $(G, *)$ un groupe. Son centre $Z(G)$ est l'ensemble des éléments $a \in G$ qui commutent avec tous les éléments de G : $Z(G) = \{a \in G \mid \forall x \in G,\ a * x = x * a\}$. Le centre de G est un sous groupe (commutatif) de G.

PROPOSITION 2.1.1
Une intersection d'une famille non vide de sous groupes d'un groupe G est un sous groupe de G.

Attention! La réunion de deux sous groupes de G n'est pas un sous groupe de G en général.

Soit A une partie de G. La famille des sous groupes de G contenant A est non vide puisqu'elle contient G lui même. L'intersection de tous les sous groupes de G contenant A est donc un sous groupe de G et c'est évidemment le plus petit sous groupe de G, au sens de l'inclusion, contenant A. On l'appelle le sous groupe de G engendré par A. On le notera gr(A) ou $< A >$.
On dit que A est une partie génératrice de G si $< A >= G$.

DEFINITION 2.1.3
Un groupe $(G, *)$ est dit monogène si il est engendré par un singleton, i.e. si il existe $a \in G$ tel que $G =< \{a\} >$ ce que l'on écrira $G =< a >$ avec un abus de notation.
Un groupe est dit cyclique si il est monogène et fini.

EXEMPLE 2.1.7
$(\mathbb{Z}, +)$ est monogène engendré par 1.
Le groupe $(\mathbb{Q}_+^*, \times)$ est engendré par l'ensemble $\mathcal{P}$ des nombres premiers.

2.1.3 Morphismes

DEFINITION 2.1.4
Soient $(G,*)$ et $(H,\top)$ deux groupes. Une application f de G dans H est un morphisme (on dit aussi homomorphisme) de groupes si $\forall x,y\in G,\ f(x*y)=f(x)\top f(y)$. Un isomorphisme de G sur H est un morphisme $f:G\to H$ bijectif tel que f^{-1} soit un morphisme. Un isomorphisme de G sur G est appelé un automorphisme de G.
On utilisera de temps à autre la notation $f:G\xrightarrow{\sim}G'$ pour signifier que f est un isomorphisme de G sur G'.

PROPOSITION 2.1.2
Tout morphisme bijectif de groupes $f:G\to H$ est un isomorphisme.

Il s'agit de vérifier que f^{-1} est un morphisme. Soient $y,y'\in H,\ x=f^{-1}(y)$ et $x'=f^{-1}(y')$. On a $f(x*x')=f(x)\top f(x')=y\top y'$ donc $f^{-1}(y\top y')=x*x'=f^{-1}(y)*f^{-1}(y')$. ■.

PROPOSITION 2.1.3
Soient G_1,G_2,G_3 trois groupes, $f_1:G_1\to G_2$ et $f_2:G_2\to G_3$ des morphismes de groupes. $f_2\circ f_1:G_1\to G_3$ est un morphisme.

PROPOSITION 2.1.4
Soit $(G,*)$ un groupe. L'ensemble des automorphismes de G muni de la loi $\circ$ de composition des applications est un groupe. On le note Aut(G).

On vérifie facilement que c'est un sous groupe du groupe $(\mathfrak{S}_G,\circ)$.

THEOREME 2.1.2
Soit $f:G\to H$ un morphisme de $(G,*)$ dans $(H,\top)$. Alors :
1) $f(1_G)=1_H$
2) $\forall x\in G, f\left(x^{-1}\right)=(f(x))^{-1}$
3) *Si G' est un sous groupe de G, $f(G')$ est un sous groupe de H. En particulier,* $\mathrm{im}(f)=f(G)$ *est un sous groupe de H*
4) *Si H' est un sous groupe de H, $f^{-1}(H')$ est un sous groupe de G.*

En particulier, $f^{-1}(1_H)$ est un sous groupe de G.

DEFINITION 2.1.5
Soit $f:G\to H$ un morphisme de $(G,*)$ dans $(H,\top)$. Le sous groupe de G, $f^{-1}(1_H)$ s'appelle noyau de f. On le note $\ker(f)$. *On a donc pour $x\in G$, $x\in\ker(f)\Leftrightarrow f(x)=1_H$.*

THEOREME 2.1.3
Soit $f:G\to H$ un morphisme de $(G,*)$ dans $(H,\top)$. Soient $x,y\in G$. On a

$$f(x)=f(y)\Leftrightarrow x*y^{-1}\in\ker(f)\Leftrightarrow x^{-1}*y\in\ker(f)$$

En particulier,f est injectif si et seulement si $\ker(f)=\{1_G\}$.

EXEMPLE 2.1.8 (Automorphismes intérieurs)
Soit $(G,*)$ un groupe. Pour $a\in G$ on définit $i_a:G\to G$ par $\forall x\in G,\ i_a(x)=a*x*a^{-1}$. On vérifie facilement que $i_a\in\mathrm{Aut}(G)$. L'application $i:a\to i_a$ de G dans Aut(G) est un morphisme. En effet, $i_{a*b}(x)=(a*b)*x*(a*b)^{-1}=a*\left(b*x*b^{-1}\right)*a^{-1}=i_a(i_b(x))=(i_a\circ i_b)(x)$ pour tout $x\in G$ donc $i_{a*b}=i_a\circ i_b$. Cherchons son noyau. $a\in\ker(i)\Leftrightarrow i_a=Id_G\Leftrightarrow\forall x\in G,\ a*x*a^{-1}=x\Leftrightarrow\forall x\in G,\ a*x=x*a$. Donc $\ker(i)$ est le centre $Z(G)$ de G. L'image de l'application i est un sous groupe de Aut(G) appellé sous groupe des automorphismes intérieurs de G. Nous le noterons Int(G).

Deux éléments x et y de G sont dits conjugués si il existe $a\in G$ tel que $y=axa^{-1}=i_a(x)$.

2.1.4 Sous groupes de $\mathbb{Z}$

On suppose connues les propriétés élémentaires de $\mathbb{Z}$.

THEOREME 2.1.4 (Division euclidienne)
Soient $a \in \mathbb{Z},\ b \in \mathbb{Z} \setminus \{0\}$. Il existe $(q, r) \in \mathbb{Z}^2$ tels que $a = bq + r$ et $|r| < |b|$. On peut imposer $0 \leqslant r < |b|$ et dans ce cas le couple (q, r) est unique.

THEOREME 2.1.5
Soit H un sous groupe de $\mathbb{Z}$. Il existe un unique $\alpha \in \mathbb{N}$ tels que $H = \alpha\mathbb{Z} := \{k\alpha\ ;\ k \in \mathbb{Z}\}$ et réciproquement, pour tout $\alpha \in \mathbb{N}$, $\alpha\mathbb{Z}$ est un sous groupe de $\mathbb{Z}$. On a donc ainsi une bijection $\alpha \to \alpha\mathbb{Z}$ de $\mathbb{N}$ sur l'ensemble des sous groupes de $\mathbb{Z}$.

preuve
On vérifie facilement que pour tout $\alpha \in \mathbb{N}$, $\alpha\mathbb{Z}$ est un sous groupe de $\mathbb{Z}$. Réciproquement, soit H un sous groupe de $\mathbb{Z}$. Si $H = \{0\}$, $\alpha = 0$ convient. On suppose donc $H \neq \{0\}$. Alors H contient un élément non nul, x et aussi son opposé $-x$. Il en résulte que H contient des éléments strictement positifs. L'ensemble des éléments strictement positifs de H est un ensemble non vide d'entiers naturels. Il admet donc un plus petit élément, soit α. Alors α et $-\alpha$ sont dans H et on vérifie par récurrence sur $n \in \mathbb{N}$ que $n\alpha \in H$ et que $(-n)\alpha = n(-\alpha) \in H$. Par conséquent, $\alpha\mathbb{Z} \subset H$. Montrons l'inclusion réciproque. Soit $x \in H$. On effectue la division euclidienne de x par α : $x = \alpha q + r$ avec $0 \leqslant r < \alpha$. On a $x \in H,\ q\alpha \in H$ donc $r = x - q\alpha \in H$. Or $0 \leqslant r < \alpha$ donc par définition de α on a nécessairement $r = 0$ et $x = q\alpha \in \alpha\mathbb{Z}$ d'où $H = \alpha\mathbb{Z}$.

2.1.5 Morphismes de $\mathbb{Z}$ dans un groupe

THEOREME 2.1.6
Soit G un groupe.
1) Soit $f : \mathbb{Z} \to G$ un morphisme et $a = f(1)$. Alors, pour tout $n \in \mathbb{Z},\ f(n) = a^n$.
2) Soit $a \in G$. Il existe un unique morphisme de groupes $f : \mathbb{Z} \to G$ tel que $f(1) = a$. Il est donné par $f(n) = a^n$ pour tout $n \in \mathbb{Z}$.
3) L'image de f ainsi défini est le sous groupe de G engendré par a.

preuve
1) On montre d'abord par récurrence sur $n \in \mathbb{N}$ que $f(n) = a^n$ pour tout $n \in \mathbb{N}$. Ensuite, si $n \in \mathbb{Z},\ n < 0$, on a $f(n)f(-n) = f(n-n) = f(0) = 1_G$ donc $f(n) = (f(-n))^{-1} = (a^{-n})^{-1} = a^n$.
2) Le 1) prouve l'unicité. Pour tout $a \in G$ la relation $a^{p+q} = a^p a^q$ valable pour tout couple $(p, q) \in \mathbb{Z}^2$ montre que f est bien un morphisme.
3) Tout sous groupe de G contenant a contient nécessairement a^n pour tout $n \in \mathbb{N}$ (récurrence sur n) puis comme précédemmment tout a^n pour $n \in \mathbb{Z}, n < 0$. Donc $\mathrm{im}(f) \subset < a >$. Comme $\mathrm{im}(f)$ est un sous groupe de G contenant a, on a $< a > \subset \mathrm{im}(f)$ d'où l'égalité.

Cas de la notation additive
Si G est un groupe commutatif dont la loi est notée additivement on écrit conformément aux notations vues dans le 1er chapitre $f(n) = n \cdot a$. On a donc $n \cdot a = \underbrace{a + \cdots + a}_{n \text{ fois}}$ si $n \geqslant 1$, $n \cdot a = \underbrace{-a + \cdots + -a}_{|n| \text{ fois}}$ si $n \leqslant -1$ et $0 \cdot a = 0_G$.

On peut considérer l'application $(n, a) \to a^n$ comme une loi externe analogue à celle que l'on introduit dans la définition des espaces vectoriels.

2.1.6 Groupes monogènes

Soient G un groupe, $a \in G$, et f_a l'unique morphisme de $\mathbb{Z}$ dans G tel que $f_a(1) = a$.
Il y a deux cas à envisager :

- f_a est injective
  Dans ce cas, f_a induit un isomorphisme de $\mathbb{Z}$ sur le sous groupe $< a >$ de G engendré par a. Ce sous groupe est nécessairement infini.

- f_a n'est pas injective
  Le noyau de f_a est un sous groupe de $\mathbb{Z}$ distinct de $\{0\}$. Il existe un unique entier naturel $n \geqslant 1$ tel que $\ker(f_a) = n\mathbb{Z}$. On a $n = 1$ ssi $a = 1_G$, cas que nous écarterons dans la suite. On suppose donc $n \geqslant 2$. Par définition, on a $a^n = 1_G$ et pour $m \in \mathbb{Z},\ a^m = 1_G \Leftrightarrow m \in \ker(f_a) \Leftrightarrow \exists k \in \mathbb{Z},\ m = kn$. Alors $< a >= \mathrm{im}(f_a) = \{1_G, a, a^2, \ldots, a^{n-1}\}$. En effet, si $x \in < a >$, il existe un $m \in \mathbb{Z}$ tel que $x = a^m$. Il existe deux entiers q et r avec $0 \leqslant r < n$ tels que $m = nq + r$. Alors $x = a^m = (a^n)^q a^r = a^r$.

DEFINITION 2.1.6
Soit a un élément d'un groupe G. On dit que a est d'ordre fini si le groupe engendré par a est fini. Dans ce cas, l'ordre de a est le cardinal de $< a >$, c'est à dire l'unique entier $n \in \mathbb{N}$ tel que $\ker(f_a) = n\mathbb{Z}$.

THEOREME 2.1.7
Soit G un groupe monogène. Si G est infini, il est isomorphe à $\mathbb{Z}$. Sinon, si $n = card(G)$, il existe un élément $a \in G$ tel que $G = \{1_G, a, \ldots, a^{n-1}\}$ avec $a^n = 1_G$. Dans ce cas le groupe G est dit cyclique de cardinal n.

COROLLAIRE 2.1.1
Tout groupe monogène est abélien.

Nous reviendrons sur les groupes cycliques après étude des groupes quotients.

2.2 Anneaux et corps

2.2.1 Définitions et généralités

DEFINITION 2.2.1
Un anneau est un triplet $(A, +, \times)$ où $(A, +)$ est un groupe abélien, $\times$ une loi de composition interne sur A associative, admettant un élément neutre noté 1_A (ou 1 si il n'y a pas d'ambiguïté) qui est distributive par rapport à la loi $+$, i.e. qui vérifie

$$\begin{array}{ll} \forall x, y, z \in A & x \times (y + z) = x \times y + x \times z \\ \forall x, y, z \in A & (y + z) \times x = y \times x + z \times x \end{array}$$

L'anneau A est dit commutatif si la loi $\times$ est commutative.

NB: Dans la terminologie actuelle les anneaux sont supposés posséder un élément neutre pour la multiplication. Dans d'anciens textes, ce qui est appelé ici anneau est appelé anneau unitaire.

On notera 0 ou 0_A l'élément neutre de A pour l'addition.

L'ensemble $\mathbb{Z}$ des entiers relatifs munis des lois $+, \times$ usuelles est un anneau commutatif.
L'ensemble $M_n(\mathbb{R})$ des matrices carrées d'ordre n à coefficients dans $\mathbb{R}$ muni des opérations usuelles est un anneau non commutatif si $n \geqslant 2$.

Lemme 2.2.1
Soit A un anneau. On a $\forall x \in A, x \times 0_A = 0_A \times x = 0_A$.

En effet, on a $(0_A + 0_A) \times x = 0_A \times x$ d'où $0_A \times x + 0_A \times x = 0_A \times x$ et donc, $(A, +)$ étant un groupe, $0_A \times x = 0_A$. De même pour l'autre égalité. ■.

On a l'équivalence $0_A = 1_A \Leftrightarrow A = \{0_A\}$. Un tel anneau est appelé anneau nul. Dans un anneau non nul, on a donc $1_A \neq 0_A$.

On sous entendra le plus ouvent les lois $+$ et $\times$ en se permettant des expressions telles que " soit A un anneau". On omettra aussi souvent le signe $\times$ pour la multiplication d'un anneau en notant ab pour $a \times b$.

Annneau produit

Soient A et A' deux anneaux (dont on notera les lois de la même manière). On munit l'ensemble $A \times A'$ des opérations suivantes : $\forall a, b \in A, \forall a', b' \in A',\ (a, a') + (b, b') = (a + a', b + b')$ et $(a, a') \times (b, b') = (aa', bb')$. $A \times A'$ muni de ces opérations est appelé anneau produit des anneaux A et A'. L'élément unité pour la multiplication est $(1_A, 1_{A'})$.

DEFINITION 2.2.2
Un sous anneau d'un anneau $(A, +, \times)$ est un sous groupe B du groupe $(A, +)$ contenant 1_A et stable par $\times$.

Si B est un sous anneau, pour les lois induites par $+$ et $\times$, B est un anneau.

DEFINITION 2.2.3
Soient A, A' deux anneaux. Un morphisme d'anneaux de A dans A' est une application $f : A \to A'$ telle que
1) $\forall x, y \in A, f(x + y) = f(x) + f(y)$. f est donc un morphisme du groupe $(A, +)$ dans le groupe $(A', +)$.
2) $\forall x, y \in A, f(x \times y) = f(x) \times f(y)$.
3) $f(1_A) = 1_{A'}$.

PROPOSITION 2.2.1
Soit $f : A \to A'$ un morphisme d'anneaux.
1) Si B est un sous anneau de A, $f(B)$ est un sous anneau de A'.
2) Si B' est un sous anneau de A', $f^{-1}(B')$ est un sous anneau de A.

2.2.2 Eléments réguliers

DEFINITION 2.2.4 (Eléments réguliers)
Soient A un anneau et $a \in A$.
1) a est dit régulier à gauche si $\forall x, y \in A,\ ax = ay \Rightarrow x = y$.
2) a est dit régulier à droite si $\forall x, y \in A,\ xa = ya \Rightarrow x = y$.
3) a est dit régulier si il est régulier à droite et à gauche.

Si un élément a n'est pas régulier, il existe des éléments distincts x et y tels que $ax = ay$ donc un élément non nul b tel que $ab = 0_A$. Un élément non régulier non nul est donc un diviseur de 0_A.

DEFINITION 2.2.5 (Eléments inversibles)
Soit A un anneau et $a \in A$.
1) a est dit inversible à gauche si il existe $a_g \in A$ tel que $a_g a = 1_A$.
2) a est dit inversible à droite si il existe $a_d \in A$ tel que $aa_d = 1_A$.
3) a est dit inversible si il est inversible pour la loi $\times$, c'est à dire si il existe $a' \in A$ tel que $aa' = a'a = 1_A$. Les éléments inversibles sont aussi appelés unités de A.

Ces définitions sont en fait valables pour tout magma $(X, \times)$ doit la loi possède un élément neutre.
Si a est inversible, il est inversible à droite et à gauche. La réciproque est vraie : si a est inversible à droite et à gauche, on a $a_d = 1_A a_d = (a_g a)a_d = a_g(aa_d) = a_g 1_A = a_g$ de sorte que $a' = a_d = a_g$. (Ce résultat est valable dans tout magma associatif possédant un élément neutre).

Si a est inversible à gauche (resp. à droite, inversible), il est régulier à gauche (resp. à droite, régulier). La réciproque est fausse comme le prouve l'exemple de l'anneau $\mathbb{Z}$ dans lequel tout élément non nul est régulier alors que les seuls éléments inversibles sont 1 et -1.

DEFINITION 2.2.6
*Un anneau A est dit intègre si il est **non nul**, commutatif et si tous les éléments de A autres que 0_A sont réguliers.*

On remarquera que le produit de deux anneaux non nuls A et A' **n'est jamais** un anneau intègre. En effet, les éléments $(1_A, 0_{A'})$ et $(0_A, 1_{A'})$ de $A \times A'$ ne sont pas nuls et ont un produit nul.

THEOREME 2.2.1
Soit A un anneau non nul. L'ensemble $U(A)$ des éléments inversibles de A est stable pour la loi $\times$ et $(U(A), \times)$ est un groupe. On note souvent $A^* = U(A)$.

On remarquera qu'avec ces notations (standard), en général $A^* \neq A \setminus \{0_A\}$.

Exercice Soit A un anneau fini. Montrer que dans A tout élément régulier est inversible .

Exemples
1) On a $U(\mathbb{Z}) = \{-1, 1\}$.
2) $U(M_n(\mathbb{R})) = GL(n, \mathbb{R})$. C'est un théorème d'algèbre linéaire que dans $M_n(\mathbb{R})$ tous les éléments réguliers sont inversibles.

2.2.3 Morphismes de $\mathbb{Z}$ dans un anneau

Lois externes sur un anneau
Soit A un anneau. En particulier $(A, +)$ est un groupe abélien. Pour $a \in A$ et $n \in \mathbb{Z}$ on sait définir $n \cdot a$ (voir le paragraphe 2.1.6) :

$$n \cdot a = \begin{cases} \underbrace{a + a + \cdots + a}_{n \text{ fois}} & \text{si } n > 0 \\ 0_A & \text{si } n = 0 \\ -\underbrace{-a - a - \cdots - a}_{|n| \text{ fois}} & \text{si } n < 0 \end{cases}$$

Il est important de noter que pour tout $a \in A$ on a $n \cdot a = (n \cdot 1_A) \times a$. Nous y reviendrons lors de l'étude de la caractéristique d'un anneau. Dans la pratique on omet souvent le $\cdot$. Toutefois, dans ces notes, on gardera souvent la notation avec le $\cdot$ pour éviter des confusions.

D'autre part, pour tout élément a de A et tout entier $n \geqslant 1$ on sait définir $a^n = \underbrace{a \times a \times \cdots \times a}_{n \text{ fois}}$ et $a^0 = 1_A$.

On a $\forall m, n \in \mathbb{N},\ a^{m+n} = a^m \times a^n$.

THEOREME 2.2.2
Soit A un anneau. Il existe un unique morphisme d'anneaux $f : \mathbb{Z} \to A$. Ce morphisme est défini par $\forall n \in \mathbb{Z}, f(n) = n \cdot 1_A$

C'est une vérification facile que la formule donnée dans l'énoncé définit bien un morphisme d'anneaux. Montrons l'unicité. Soit $f : \mathbb{Z} \to A$ un morphisme d'anneaux ; f est un morphisme du groupe $(\mathbb{Z}, +)$ dans le groupe $(A, +)$ qui doit vérifier $f(1) = 1_A$. On a vu que celà impliquait l'unicité de f. (théorème 2.1.6)

2.2.4 Formule du binôme

THEOREME 2.2.3
*Soient A un anneau a et b deux éléments de A **commutant entre eux** et n un entier naturel. On a*

$$(a+b)^n = \sum_{k=0}^{k=n} C_n^k \cdot a^k b^{n-k} = \sum_{k=0}^{k=n} C_n^k \cdot a^{n-k} b^k$$

où C_n^k désigne les coefficients binomiaux.

La preuve est classique et se fait par récurrence sur l'entier n, en utilisant uniquement le fait que la famille d'entiers C_n^k vérifie pour tout n, $C_n^0 = C_n^n = 1$ et, pour $n \geqslant 1$, $C_n^k = C_{n-1}^k + C_{n-1}^{k-1}$ pour $1 \leqslant k \leqslant n-1$

2.2.5 Corps

DEFINITION 2.2.7
On appelle corps un anneau non nul dans lequel tout élément non nul est inversible.

Si K est un corps on a $K^* = K \setminus \{0_K\}$.
Si K est un corps, un sous corps de K est un sous anneau k de K tel que pour tout $x \in k$ on ait $x^{-1} \in k$. Alors k muni des opérations induites par celles de K est un corps et k^* est un sous groupe de $(K^*, \cdot)$.
Soient k et K deux corps. On dit que K est une extension de k si k est un sous corps de K.

PROPOSITION 2.2.2
Soient K un corps, A un anneau et $f : K \to A$ un morphisme non nul. f est injectif.

preuve
Comme f est un morphisme du groupe $(K, +)$ dans le groupe $(A, +)$ il suffit de montrer que $\ker(f) = \{0_K\}$. Soit $a \in K$ tel que $f(a) = 0_A$. Supposons $a \neq 0_K$. Alors, a est inversible dans K ; on a $0_A \neq 1_A = f\left(a \times a^{-1}\right) = f(a) \times f\left(a^{-1}\right) = 0_A$. Contradiction.

2.2.6 Corps des fractions d'un anneau intègre

THEOREME 2.2.4
Soit A un anneau intègre.
1) Il existe un couple (j, K) où K est un corps commutatif et $j : A \to K$ un morphisme injectif tel que

$$(\spadesuit) \qquad \forall k \in K,\ \exists a \in A,\ \exists b \in A \setminus \{0\} \quad k = j(a)\left(j(b)\right)^{-1}$$

2) Un tel couple est unique à isomorphisme près : si (j_1, K_1) est un autre couple vérifiant $(\spadesuit)$, il existe un isomorphisme de corps $\varphi : K \to K_1$ tel que $j_1 = \varphi \circ j$.
3) Soit L un corps commutatif quelconque et $f : A \to L$ un morphisme injectif. Il existe un morphisme de corps $F : K \to L$ tel que $f = F \circ j$ et ce morphisme est unique. On a donc le diagramme commutatif :

$$A@^{(}- >[d]_j[r]^f LK[ur]^F$$

Le corps K ainsi construit s'appelle corps des fractions de A.

preuve abrégée

1) Existence. Soit $S = A \setminus \{0\}$. Comme A est un anneau intègre, S est une partie de A stable par multiplication. Soit $X = A \times S$. On munit X de deux opérations $\oplus$ et $\otimes$ et d'une relation binaire $\mathcal{R}$: soient $(a, b) \in X$ et $(a', b') \in X$. On pose

$$(a, b) \oplus (a', b') = (ab' + ba', bb') \qquad (a, b) \otimes (a', b') = (aa', bb') \qquad (a, b)\,\mathcal{R}\,(a', b') \Leftrightarrow ab' = ba'$$

Le fait que S soit stable pour le produit garantit que l'on définit bien des lois internes sur X. On vérifie ensuite les propriétés suivantes :
a) $\mathcal{R}$ est une relation d'équivalence compatible avec les deux lois $\oplus$ et $\otimes$. Il en résulte que ces deux lois induisent des lois que nous noterons + et * sur l'ensemble quotient $K = X/\mathcal{R}$. Si $(a, b) \in X$, on notera $\overline{(a, b)}$ sa classe dans K.
b) $(K, +, *)$ est un corps d'élément nul $0_K := \overline{(0, 1)}$ et d'élément unité $1_K = \overline{(1, 1)}$.
c) Soit $j : A \to K$ l'application définie par $\forall a \in A,\ j(a) = \overline{(a, 1)}$. Alors j est un morphisme injectif de l'anneau A dans le corps K et si $b \in A,\ b \neq 0$ on a $\overline{(1, b)} = (j(b))^{-1}$.
Soit alors $k \in K$. Il existe $(a, b) \in X$ tel que $k = \overline{(a, b)}$. Alors $k = \overline{(a, b)} = \overline{(a, 1)} * \overline{(1, b)} = j(a) * j(b)^{-1}$.

3) Soit $f : A \to L$ un morphisme de corps. Soit $H : X \to L$ défini par $\forall (a, b) \in X,\ H(a, b) = f(a)\left(f(b)\right)^{-1}$. Ceci a bien un sens car par hypothèse $b \neq 0_A$ et f étant injective $f(b) \neq 0_L$. Soient (a, b) et (a', b') dans X vérifiant $(a, b)\,\mathcal{R}\,(a', b')$. On a $ab' = a'b$ donc $f(a)f(b') = f(ab') = f(a'b) = f(a')f(b)$ donc L étant commutatif, $H(a, b) = H(a', b')$. L'application H est constante sur les classes d'équivalence de la relation $\mathcal{R}$. Elle passe donc aux quotients en une application $F = \bar{H} : X/\mathcal{R} = K \to L$.

On vérifie ensuite que F est un morphisme. On a, pour $(a, b),\ (a', b') \in X$

$$\begin{aligned} H\left((a, b) \oplus (a', b')\right) &= H(ab' + a'b, bb') = f(ab' + a'b)\left(f(bb')\right)^{-1} \\ &= \left(f(a)f(b') + f(a')f(b)\right)\left(f(b)^{-1}f(b')^{-1}\right. \\ &= f(a)\left(f(b)\right)^{-1} + f(a')\left(f(b')\right)^{-1} \\ &= H(a, b) + H(a', b') \end{aligned}$$

d'où par passage aux quotients $F\left(\overline{(a,b)}+\overline{(a',b')}\right)=F\left(\overline{(a,b)}\right)+F\left(\overline{(a',b')}\right)$
De même on a

$$H\left((a,b)\otimes(a',b')\right)=H((aa',bb'))=f(aa')\left(f(bb')\right)^{-1}=f(a)\left(f(b)\right)^{-1}\ f(a')\left(f(b')\right)^{-1}=H(a,b)H(a',b')$$

donc par passage aux quotients $F\left(\overline{(a,b)}\times\overline{(a',b')}\right)=F\left(\overline{(a,b)}\right)\times F\left(\overline{(a',b')}\right)$
Enfin, $F(1_K)=F\left(\overline{(1,1)}\right)=H(1,1)=f(1)=1_L$.

On a pour $a\in A$, $F\circ j(a)=F\left(\overline{(a,1)}\right)=H(a,1)=f(a)$ donc $F\circ j=f$.

Soit $G:K\to L$ un morphisme de corps tel que $G\circ j=f$. On a pour $a\in A,\ G((j(a))=f(a)$. Comme tout élément x de K est de la forme $x=j(a)j(b)^{-1}$ on a nécessairement $G(x)=G(j(a))G(j(b))^{-1}=f(a)f(b)^{-1}=H(a,b)=F(x)$ d'où l'unicité de F ce qui achève la preuve du point 3).

2) Unicité. Soit (j_1,K_1) un couple où K_1 est un corps commutatif et $j_1:A\to K_1$ un morphisme injectif vérifiant (♠). En appliquant le 3) à j_1 on en déduit l'existence d'un morphisme de corps $\varphi:K\to L$ tel que $j_1=\varphi\circ j$. Un tel morphisme est nécessairement injectif. Soit $z\in K_1$. Par hypothèse, il existe $a,b\in A$ tels que $z=j_1(a)j_1(b)^{-1}$ d'où $z=\varphi\left(j(a)\,j(b)^{-1}\right)$ ce qui prouve la surjectivité de φ. ■.

3

Quotients

3.1 Groupes quotients

Dans la suite sauf exception, on omettra le signe $*$ désignant la loi de composition interne d'un groupe si celle ci est notée multiplicativement. On écrira "soit G un groupe" au lieu de "soit $(G,*)$ un groupe" et on notera ab pour le composé $a*b$ de deux éléments de G.
De même pour les groupes usuels, on omettra de préciser l'opération.

3.1.1 Relation d'équivalence définie par un sous groupe

Dans cette partie, on se donne un groupe G noté multiplicativement. Soit H un sous groupe de G. Pour $a \in G$, on note $Ha = \{xa\ ;\ x \in H\}$ et $aH = \{ax\ ;\ x \in H\}$.
On définit deux relations binaires sur G, $\mathcal{R}_g$ et $\mathcal{R}_d$ par

$$\forall x, y \in G,\ x\mathcal{R}_g y \Leftrightarrow x^{-1}y \in H \Leftrightarrow y \in xH$$

$$\forall x, y \in G,\ x\mathcal{R}_d y \Leftrightarrow yx^{-1} \in H \Leftrightarrow y \in Hx$$

PROPOSITION 3.1.1
1) *La relation $\mathcal{R}_g$ est une relation d'équivalence dans G, compatible avec la multiplication à gauche i.e.*

$$\forall x, y, z \in G,\ x\mathcal{R}_g y \Rightarrow zx\mathcal{R}_g zy$$

On notera $(G/H)_g$ l'ensemble quotient $G/\mathcal{R}_g$.
2) *La relation $\mathcal{R}_d$ est une relation d'équivalence dans G, compatible avec la multiplication à droite i.e.*

$$\forall x, y, z \in G,\ x\mathcal{R}_d y \Rightarrow xz\mathcal{R}_d yz$$

On notera $(G/H)_d$ l'ensemble quotient $G/\mathcal{R}_d$.
3) *L'application $\varphi : G \to G,\ x \to x^{-1}$ induit une bijection*

$$\Phi : (G/H)_g \to (G/H)_d$$

On notera que pour chacune de ces deux relations la classe de l'élément neutre 1_G de G est H.
Il résulte du 3) que si $(G/H)_g$ est fini, il en est de même de $(G/H)_d$ et que ces deux ensembles ont même nombre d'éléments. Ce nombre est noté $[G : H]$ et appelé indice du sous groupe H de G.

preuve
Notons d'abord que $y \in xH \Leftrightarrow yH = xH$. En effet, supposons $y \in xH$ donc que y s'écrit $y = xh_0$ avec $h_0 \in H$. Alors pour tout $h \in H$ on a $h_0 h \in H$ car H est un sous groupe donc $yh = x(h_0 h) \in xH$. Donc $yH \subset xH$. De même, $xh = y(h_0^{-1}h) \in yH$ donc $xH \subset yH$. Réciproquement, si $xH = yH$, comme H contient 1_G on a $y = y1_G \in xH$.
De cette équivalence on déduit le 1). La classe de x modulo $\mathcal{R}_g$ est l'ensemble xH. La preuve de 2) est analogue.
Montrons 3). Soit $\pi_d : G \to (G/H)_d$ l'application de passage aux quotients : $\pi_d(x) = Hx$ = classe de x modulo $\mathcal{R}_d$.

L'application π_d est surjective. Il en est donc de même de l'application $\psi := \pi_d \circ \varphi : x \to Hx^{-1}$. On a $\psi(x) = \psi(y) \Leftrightarrow Hx^{-1} = Hy^{-1} \Leftrightarrow Hx^{-1}y = H \Leftrightarrow x^{-1}y \in H \Leftrightarrow x\mathcal{R}_g y$.
Par passage aux quotients, on en déduit une injection $\Phi : (G/H)_g \to (G/H)_d$. Comme ψ est surjective, il en est de même de Φ. Cette application n'est autre que celle qui envoie $xH \in (G/H)_g$ sur $Hx^{-1} \in (G/H)_d$. ■

THEOREME 3.1.1 (Théorème de Lagrange)
Soit G un groupe fini, H un sous groupe de G. Le cardinal de H divise celui de G. Plus précisément, on a la relation

$$card(G) = [G : H] \cdot card(H)$$

preuve
Si $x \in G$ l'application $h \to xh$ est une bijection de H sur xH. Toutes les classes d'équivalence, qui sont par définition au nombre de $[G : H]$ ont donc même cardinal, égal à celui de H. Comme ces classes forment une partition de G, on en déduit le résultat.

COROLLAIRE 3.1.1
Dans un groupe fini, l'ordre de tout élément divise le cardinal du groupe.
En particulier, pour tout $x \in G$ on a

$$x^{card(G)} = 1_G$$

En effet, l'ordre d'un élément a est le cardinal du sous groupe (cyclique) engendré par a.

3.1.2 Sous groupes distingués. Groupes quotients

Si A et B sont deux sous ensembles d'un groupe G, on note $A \cdot B$ l'ensemble des éléments de G de la forme $ab,\ a \in A,\ b \in B$. Donc $A \cdot B = \{x \in G \mid \exists a \in A,\ \exists b \in B,\ x = ab\}$.

THEOREME 3.1.2
Soient G un groupe, H un sous groupe de G, $\mathcal{R}_g$ et $\mathcal{R}_d$ les relations d'équivalence associées au sous groupe H. Les propriétés suivantes sont équivalentes :
1) $\forall x \in G,\ xH = Hx$.
2) $\forall x \in G,\ xHx^{-1} = H$.
3) $\forall x \in G, xHx^{-1} \subset H$.
4) $\mathcal{R}_d = \mathcal{R}_g$
5) $\forall x, y \in G,\ xH \cdot yH = (xy)H$.
6) *La relation $\mathcal{R}_g$ est compatible avec la loi du groupe.*
7) $\forall x, y \in G,\ Hx \cdot Hy = H(xy)$.
8) *La relation $\mathcal{R}_d$ est compatible avec la loi du groupe.*

DEFINITION 3.1.1
Un sous groupe H d'un groupe G est dit distingué ou normal si il vérifie l'une des propriétés ci dessus (et donc les huit). On note $H \lhd G$ pour signifier que H est un sous groupe de G, distingué dans G. Dans ce cas, puisque les ensembles $(G/H)_g$ et $(G/H)_d$ sont égaux, on les note G/H.

preuve
a) Equivalence de 1), 2), 3) et 4).
L'équivalence de 1) et 2) est immédiate de même que l'implication $2) \Rightarrow 3)$. On voit que $3) \Rightarrow 2)$ en appliquant 3) avec x et x^{-1}. On obtient $xHx^{-1} \subset H$ et $x^{-1}Hx \subset H$ d'où $H \subset xHx^{-1}$.
L'équivalence de 1) et 4) est immédiate puisque la relation $\forall x \in G,\ xH = Hx$ signifie que la classe de x pour $\mathcal{R}_g$ est égale à la classe de x pour $\mathcal{R}_d$, pour tout $x \in G$; on a donc l'égalité des deux relations d'équivalence.

b) $4) \Rightarrow 6) \Rightarrow 5) \Rightarrow 3)$.
Supposons 4). $\mathcal{R}_g$ est compatible à gauche avec $*$ et $\mathcal{R}_d$ l'est à droite. Donc si $\mathcal{R}_g = \mathcal{R}_d$ la relation $\mathcal{R}_g$ est compatible à droite et à gauche avec $*$.
L'implication $6) \Rightarrow 5)$ n'est autre que le passage aux quotients de la loi $*$ modulo la relation $\mathcal{R}_g$ (voir 1.5.3). En effet, un élément de $xHyH$ s'écrit xh_1yh_2 avec $h_1, h_2 \in H$. On a $xh_1\mathcal{R}_g x$ donc par compatibilité à droite $xh_1y\mathcal{R}_g xy$ soit

$xh_1yH = xyH$ donc $xh_1yh_2 \in xyH$ d'où une inclusion. L'autre est toujours vérifiée car $xyh = (x1_G)(yh) \in xHyH$.
Supposons 5). Soit $x \in G$. On applique la relation 5) avec $y = x^{-1}$. Il vient $xH \cdot x^{-1}H = (xx^{-1})H = H$. Soit $xHx^{-1} \cdot H = H$. On en déduit en particulier $xHx^{-1} \subset H$.

c) La vérification de $4) \Rightarrow 8) \Rightarrow 7) \Rightarrow 3)$ est analogue et termine la preuve.

Attention! Si H et K sont deux sous groupes de G tels que $H \subset K$, on peut avoir $H \lhd K$ et $K \lhd G$ sans avoir $H \lhd G$.
Exercice (On suppose connus les groupes de permutations)
Soit $G = \mathfrak{S}_4$, $K = \{id,\ (1,2)(3,4),\ (1,3)(2,4);\ (1,4)(2,3)\}$ et $H = \{id,\ (1,2)(3,4)\}$. Vérifier que $H \lhd K$ que $K \lhd \mathfrak{S}_4$ mais que H n'est pas distingué dans $\mathfrak{S}_4$. (Ici, (1,2) par exemple dénote la transposition qui échange 1 et 2).

THEOREME 3.1.3
Soit G un groupe. Tout sous groupe H de G d'indice 2 est distingué.

preuve
Considérons par exemple $(G/H)_g$. C'est une partition de G formée de deux sous ensembles dont l'un, la classe de 1_G est nécessairement H. L'autre est donc $G \setminus H$. Il en est de même pour $(G/H)_d$. Les projections canoniques π_g et π_d coincident : soit $x \in G$. Si $x \in H,\ \pi_g(x) = \pi_d(x) = H$ et si $x \notin H,\ \pi_g(x) = \pi_d(x) = G \setminus H$. Il en résulte que $\forall x, y \in G,\ \pi_d(x) = \pi_d(y) \Leftrightarrow \pi_g(x) = \pi_g(y)$ autrement dit que $x\mathcal{R}_d y \Leftrightarrow x\mathcal{R}_g y$. Donc $\mathcal{R}_g = \mathcal{R}_d$. ■.

THEOREME 3.1.4
Soient G un groupe, H un sous groupe de G distingué dans G et $p : G \to G/H$ la projection canonique. La relation d'équivalence $\mathcal{R}$ associée à H est compatible avec la loi de groupe de G et induit sur l'ensemble quotient une loi de groupe telle que la projection canonique p soit un morphisme. Le noyau de p est égal à H. L'ensemble G/H muni de cette loi s'appelle le groupe quotient du groupe G par le sous groupe distingué H.
Si G est un groupe abélien, tout sous groupe de G est distingué et le groupe quotient G/H est abélien.

preuve
Notons $\mathcal{R} = \mathcal{R}_g = \mathcal{R}_d$. Puisque cette relation est compatible avec la loi $*$, celle ci passe aux quotients en une loi $\bar{*}$ sur l'ensemble G/H. Cette loi est définie par $(xH)\bar{*}(yH) = xyH$. D'après les propriétés générales (chapitre 1) $(G/H, \bar{*})$ est un groupe et p un morphisme surjectif. Enfin, si $x \in G$, on a $x \in \ker(p) \Leftrightarrow p(x) = p(1_G) \Leftrightarrow x\mathcal{R}1_G \Leftrightarrow x \in H$.
La dernière assertion est immédiate.

Par exemple, le groupe quotient de $(\mathbb{Z}, +)$ par le sous groupe $n\mathbb{Z}$ est le groupe $\mathbb{Z}/n\mathbb{Z}$ des entiers modulo n.

La relation d'équivalence définie par un sous groupe distingué d'un groupe G est compatible avec la loi du groupe. En fait, toute relation d'équivalence compatible avec la loi du groupe est de cette forme. En effet, on a le :

THEOREME 3.1.5
Soit G un groupe et ρ une relation d'équivalence dans G compatible avec la loi de groupe. Alors la classe H de l'élément neutre 1_G est un sous groupe distingué de G et la relation ρ est la relation associée à ce sous groupe.

preuve
D'abord, H contient 1_G donc est non vide. Ensuite, soient $x, y \in H$. On a $y\rho 1_G$ donc en multipliant par y^{-1}, par compatibilité $1_G \rho y^{-1}$. On a aussi $x\rho 1_G$ donc encore par compatibilité, $xy^{-1}\rho 1_G$ soit $xy^{-1} \in H$ ce qui prouve que H est un sous groupe.
Montrons qu'il est distingué. Soit $x \in G$. Pour tout $h \in H$ on a $h\rho 1_G$ donc par compatibilité successivemennt, $(xh)\rho x$ puis $(xhx^{-1})\rho(xx^{-1})$ soit $(xhx^{-1})\rho 1_G$ donc $xhx^{-1} \in H$ ce qui prouve que $xHx^{-1} \subset H$ donc que $H \lhd G$. Enfin, on a, par compatibilité de ρ, $x\rho y \Leftrightarrow x^{-1}y\rho 1_G$ Or $x^{-1}y\rho 1_G \Leftrightarrow x^{-1}y \in H \Leftrightarrow y \in xH$ ce qui prouve que ρ est la relation définie par H.

THEOREME 3.1.6
Soient G, G' deux groupes et $f : G \to G'$ un morphisme de groupes. Si H' est un sous groupe distingué de G', l'image réciproque $f^{-1}(H')$ est un sous groupe distingué de G.
En particulier, le noyau $\ker(f)$ de tout morphisme de groupes est un sous groupe distingué et la relation $\mathcal{R}_f$ induite par f est celle définie par $\ker(f)$.

preuve
Soit $H = f^{-1}(H')$. Soient $x \in G,\ h \in H$. On a $f\left(xhx^{-1}\right) = f(x)f(h)(f(x))^{-1}$. Par hypothèse $f(h) \in H'$ et $H' \lhd G'$ donc $f(x)f(h)(f(x))^{-1} \in H'$ ce qui prouve que $xhx^{-1} \in H$. On a donc $xHx^{-1} \subset H$.

D'autre part, $\{1'_G\}$ est un sous groupe distingué de G', donc $\ker(f)$ est un sous groupe distingué de G. Par définition, la relation induite par f est la relation définie par $x\mathcal{R}_f y \Leftrightarrow f(x) = f(y)$. Or $f(x) = f(y) \Leftrightarrow y^{-1}x \in \ker(f)$ ce qui prouve que $\mathcal{R}_f = \mathcal{R}_{\ker(f)}$.

THEOREME 3.1.7 (Premier théorème d'isomorphisme)
Soit $f : G \to G'$ un morphisme de groupes. Soit $p : G \to G/\ker(f)$ la projection canonique et $i : \mathrm{im}(f) \to G'$ l'inclusion. Il existe un unique isomorphisme de groupes $\bar{f} : G/\ker(f) \to \mathrm{im}(f)$ vérifiant $f = p \circ \bar{f} \circ i$. On traduit cette dernière relation en écivant que le diagramme ci dessous est commutatif.

$$\begin{array}{ccc} G & \xrightarrow{f} & G' \\ {\scriptstyle p}\downarrow & & \uparrow{\scriptstyle i} \\ G/\ker(f) & \xrightarrow{\bar{f}} & \mathrm{im}(f) \end{array}$$

preuve
On a vu que la relation d'équivalence associée au sous groupe distingué $\ker(f)$ est la relation d'équivalence définie par f. D'où l'existence de la bijection $\bar{f}$ (Thm 1.5.1). Il reste à voir que c'est un morphisme. Soient $\alpha, \beta \in G/\ker(f)$ et $x, x' \in G$ tels que $\alpha = p(x), \beta = p(x')$ (ou ce qui est équivalent $x \in \alpha,\ x' \in \beta$). Par définition, $\alpha * \beta = p(xx')$ donc $\bar{f}(\alpha * \beta) = f(xx') = f(x)f(x') = \bar{f}(\alpha)\bar{f}(\beta)$ d'où la conclusion.

EXEMPLE 3.1.1
Soient G un groupe cyclique de cardinal n, $a \in G$ un générateur de G. On a vu que l'application $f_a : \mathbb{Z} \to G$ qui à $p \in \mathbb{Z}$ associe $a^p \in G$ était un morphisme surjectif de noyau $n\mathbb{Z}$. On en déduit un isomorphisme

$$\bar{f} : \mathbb{Z}/n\mathbb{Z} \xrightarrow{\sim} G$$

qui à la classe modulo n de l'entier m associe a^m.
Un groupe cyclique de cardinal n est donc isomorphe à $\mathbb{Z}/n\mathbb{Z}$.

EXEMPLE 3.1.2
Soit G un groupe. L'application qui à $a \in G$ associe l'automorphisme intérieur $i_a \in \mathrm{Int}(G)$ est un morphisme surjectif de G sur le groupe $\mathrm{Int}(G)$ des automorphismes intérieurs de G, dont le noyau est le centre $Z(G)$. On a donc un isomorphisme $\bar{i} : G/Z(G) \xrightarrow{\sim} \mathrm{Int}(G)$.

THEOREME 3.1.8 (Passage aux quotients d'un morphisme)
Soient G, G' deux groupes, H un sous groupe distingué de G , $p : G \to G/H$ la surjection canonique et et $f : G \to G'$ un morphisme tel que $H \subset \ker(f)$. L'application f induit, par passage aux quotients une application $\tilde{f} : G/H \to G'$ telle que $f = \tilde{f} \circ p$.

preuve abrégée
Pour $x \in G$, on pose $\tilde{f}(xH) = f(x)$. Si $xH = x'H$ on a $x^{-1}x' \in H \subset \ker(f)$ donc $f(x) = f(x')$ ce qui garantit que $\tilde{f}$ est bien définie. On vérifie, comme dans la preuve du théorème d'isomorphisme que $\tilde{f}$ est un morphisme. La condition $f = \tilde{f} \circ p$ entraine l'unicité de $\tilde{f}$.

3.2 Anneaux quotients

3.2.1 Idéal d'un anneau commutatif

Dans toute cette section, les anneaux considérés sont commutatifs.

Definition et premières propriétés

DEFINITION 3.2.1
Soit $(A, +, \times)$ un anneau commutatif. Un sous ensemble I de A est appellé un idéal si
1) I est un sous groupe additif de $(A, +)$.
2) $\forall x \in A, \forall a \in A,\ a \in I \Rightarrow ax \in I$ ce qu'on peut écrire $\forall x \in A,\ xI \subset I$.

Il est clair que A et $\{0_A\}$ sont des idéaux de A. Tout idéal de A qui contient 1_A est nécessairement égal à A d'après la propriété 2). Par conséuent, sauf le cas $I = A$ un idéal de A n'est jamais un sous anneau de A.

THEOREME 3.2.1
Soit $f : A \to A'$ un morphisme d'anneaux commutatifs. Si I' est un idéal de A', alors $f^{-1}(I')$ est un idéal de A.
En particulier, le noyau de f, $\ker(f) = f^{-1}(0_{A'})$ est un idéal de A.

L'image directe $f(I)$ d'un idéal de A n'est pas nécessairement un idéal de A'.

PROPOSITION 3.2.1
L'intersection d'une famille non vide d'idéaux d'un anneau commutatif A est un idéal de A.

Il en résulte que si X est un sous ensemble de A, l'intersection I_X de tous les idéaux de A contenant X (il y en a au moins un, A lui même) est un idéal de A, qui est le plus petit idéal de A (pour l'inclusion) contenant X. On l'appelle l'idéal de A engendré par X.

PROPOSITION 3.2.2
Soient A un anneau commutatif et $a \in A$. L'idéal engendré par $\{a\}$ est l'ensemble $aA = \{ax \ ; \ x \in A\}$.

Un idéal I de A engendré par un seul élément, donc de la forme aA est appelé idéal principal. On dit que a est un générateur de cet idéal. Si l'anneau A est intègre, les autres générateurs de I sont les éléments au pour $u \in A^*$.

DEFINITION 3.2.2
Un anneau principal est un anneau commutatif et intègre dont tous les idéaux sont principaux.

THEOREME 3.2.2
$\mathbb{Z}$ est un anneau principal.

Plus précisément, les idéaux de $\mathbb{Z}$ sont les ensembles $n\mathbb{Z}$ pour $n \in \mathbb{Z}$ et l'application $n \to n\mathbb{Z}$ est une bijection de $\mathbb{N}$ sur l'ensemble des idéaux de $\mathbb{Z}$. En effet, un idéal de $\mathbb{Z}$ est un sous groupe de $\mathbb{Z}$ donc de la forme $n\mathbb{Z}$ et réciproquement il est immédiat de vérifier qu'un tel sous ensemble de $\mathbb{Z}$ est un idéal.

Caractérisation des corps commutatifs

THEOREME 3.2.3
Soit A un anneau commutatif non nul. A est un corps si et seulement si les seuls idéaux de A sont $\{0_A\}$ et A.

preuve
1) Supposons que A soit un corps. Soit I un idéal de A distinct de $\{0_A\}$. Alors il existe $a \in I, \ a \neq 0_A$. Comme A est un corps, a admet un inverse $a^{-1} \in A$. On en déduit que $1_A = a^{-1}a \in I$. Alors, pout tout $x \in A$, $x = x1_A \in I$.
2) Supposons réciproquement que les seuls idéaux de A soient $\{0_A\}$ et A. Soit $a \in A$, $a \neq 0_A$. L'idéal principal aA engendré par a contient a donc est distinct de $\{0_A\}$. Par conséquent $aA = A$ et il existe $a' \in A$ tel que $aa' = 1_A$. Donc tout élément non nul de A est inversible est A est un corps.

On peut retrouver un résultat déjà démontré :

COROLLAIRE 3.2.1
Soient K un corps, A un anneau non nul et $f : K \to A$ un morphisme. Alors f est injectif et $f(K)$ est un sous corps de A.

preuve
Puisque A est non nul $f(1_K) = 1_A \neq 0_A$. L'application f n'est pas l'application nulle, donc $\ker(f) \neq K$. Il en résulte que $\ker(f) = \{0_K\}$ donc que f est injective.
On sait que $f(K)$ est un sous anneau de A. Soit $y \in f(K), \ y \neq 0_A$. Soit $x \in K$ tel que $y = f(x)$. On a en posant $y' = f(x^{-1})$ d'une part $y' \in f(K)$ et d'autre part $yy' = f(x)f(x') = f(xx') = f(1_K) = 1_A$. Donc $f(K)$ est un sous corps de A.

3.2.2 Anneaux quotients

THEOREME 3.2.4
1) Soit A un anneau commutatif et I un idéal de A. La relation binaire $\mathcal{R}_I$ sur A définie par $x\mathcal{R}_I y \Leftrightarrow x - y \in I$ est une relation d'équivalence sur A, compatible avec les deux lois $+$ et $\cdot$ de A. Pour cette relation d'équivalence, la classe de 0_A est I.
2) Soit ρ une relation d'équivalence sur A compatible avec les deux lois de A. La classe d'équivalence de 0_A pour ρ est un idéal I de A et $\rho = \mathcal{R}_I$.

preuve
1) I est un sous groupe de $(A, +)$ donc $\mathcal{R}_I$ est compatbible avec la loi de groupe. Montrons la compatibilité avec la multiplication. Soient $x, x', y, y' \in A$ tels que $x\mathcal{R}_I x'$ et $y\mathcal{R}_I y'$. On a $xy - x'y' = x(y - y') + (x - x')y'$. Or $x - x' \in I,\ y - y' \in I$ donc d'après la définition d'un idéal $xy - x'y' \in I$ i.e. $xy\,\mathcal{R}_I\,x'y'$. Le reste de la preuve est laissé au lecteur. Notons seulement que si $x \in A$, sa classe $p(x)$ est le sous ensemble $x + I = \{x + t\ ,\ t \in I\}$. Les opérations sur A/I sont alors définies par $(x + I) + (y + I) = (x + y) + I$ et $(x + I) \times (y + I) = xy + I$.

THEOREME 3.2.5
Soit I un idéal d'un anneau commutatif A. Il existe une unique structure d'anneau sur l'ensemble quotient A/I de A par la relation d'équivalence $\mathcal{R}_I$ définie par l'idéal I telle que la projection canonique $p : A \to A/I$ soit un morphisme. A/I muni de cette structure s'appelle l'anneau quotient de A par I.

THEOREME 3.2.6 (Factorisation canonique d'un morphisme)
Soient $f : A \to A'$ un morphisme d'anneaux commutatifs, $i : \mathrm{im}(f) \to A'$ l'inclusion et $p : A \to A/\ker(f)$ la projection canonique. Il existe un unique isomorphisme d'anneaux $\bar{f} : A/\ker(f) \to f(A)$ tel que $f = i \circ \bar{f} \circ p$.

THEOREME 3.2.7 (Passage au quotient d'un morphisme)
Soient $f : A \to A'$ un morphisme d'anneaux commutatifs et I un idéal de A tel que $I \subset \ker(f)$. Il existe un unique morphisme d'anneaux $\bar{f} : A/I \to A'$ tel que $f = \bar{f} \circ p$ où p est la projection canonique $A \to A/I$.

Cas de $\mathbb{Z}$

Les idéaux de $\mathbb{Z}$ sont les ensembles $n\mathbb{Z}$, pour $n \in \mathbb{N}$. La relation d'équivalence associée $\mathcal{R}_n$ s'appelle la congruence modulo n. Si $x, y \in \mathbb{Z}$ on a donc $x\mathcal{R}_n y \Leftrightarrow x - y \in n\mathbb{Z} \Leftrightarrow n|(x - y)$. On note $x \equiv y \bmod n$ pour $x\mathcal{R}_n y$.
L'ensemble quotient $\mathbb{Z}/n\mathbb{Z}$ hérite d'une structure d'anneau commutatif. C'est l'anneau des classes résiduelles modulo n. La classe d'un élément $x \in \mathbb{Z}$ est le sous ensemble $x + n\mathbb{Z}$ de $\mathbb{Z}$. Il résulte de ce qui a été dit précédemment que les opérations sur les classes vérifient $(x + n\mathbb{Z}) + (y + n\mathbb{Z}) = (x + y) + n\mathbb{Z}$ et $(x + n\mathbb{Z}) \cdot (y + n\mathbb{Z}) = xy + n\mathbb{Z}$.
Si $n = 0$ la relation $\mathcal{R}_0$ est l'égalité et $\mathbb{Z}/0\mathbb{Z} = \mathbb{Z}$. Si $n = 1$, la relation $\mathcal{R}_1$ est la relation grossière (deux éléments quelconques sont équivalents) et le quotient $\mathbb{Z}/1\mathbb{Z}$ est l'anneau nul. Nous écarterons toujours ces deux cas. Cet anneau sera étudié en détail dans le chapitre arithmétique.

3.2.3 Caractéristique d'un anneau

Nous anticiperons dans ce paragraphe quelques résultats du cours d'arithmétique, principalement le :

THEOREME 3.2.8
Soit $n \in \mathbb{N},\ n \geqslant 2$. Les propriétés suivantes sont équivalentes :
(1) L'anneau $\mathbb{Z}/n\mathbb{Z}$ est intègre.
(2) L'anneau $\mathbb{Z}/n\mathbb{Z}$ est un corps.
(3) n est un nombre premier.

Rappelons que si A est un anneau, il existe un unique morphisme Φ_A de $\mathbb{Z}$ dans A noté $\Phi_A(n) = n \cdot 1_A$.

DEFINITION 3.2.3
Soit A un anneau commutatif. On appelle caractéristique de A l'entier naturel n défini par $\ker(\Phi_A) = n\mathbb{Z}$ où $\Phi_A : \mathbb{Z} \to A$ est l'unique morphisme d'anneaux de $\mathbb{Z}$ dans A. On la notera $\mathrm{car}(A)$.

Remarquons que comme tout sous anneau de A contient 1_A, la caractéristique de tout sous anneau de A est égale à celle de A.

La caractéristique de A est égale à 1 ssi Φ_A est nulle. Comme $1_A = \Phi_A(1)$ ceci équivaut à $A = \{0\}$. Ecartons ce cas.

Posons $\tilde{A} = \Phi_A(\mathbb{Z})$. C'est un sous anneau de A et tout sous anneau de A contient 1_A donc $1_A + 1_A$ etc... donc contient $\tilde{A}$. Il y a essentiellement deux cas.

1) Si la caractéristique de A est nulle, Φ_A est injective et $\tilde{A} = \Phi_A(\mathbb{Z})$ est un sous anneau de A isomorphe à $\mathbb{Z}$.
2) Sinon, $n \geqslant 2$. Alors, pour tout entier m on a $m \cdot 1_A = 0 \Leftrightarrow n$ divise m. Le théorème de factorisation montre qu'il existe un isomorphisme $\overline{\Phi_A} : \mathbb{Z}/n\mathbb{Z} \to \Phi_A(\mathbb{Z}) = \tilde{A}$. Le sous anneau $\tilde{A}$ est donc isomorphe à $\mathbb{Z}/n\mathbb{Z}$.

PROPOSITION 3.2.3
Soit A un anneau commutatif de caractéristique $n > 0$. Pour tout $a \in A$ on a $n \cdot a = 0_A$.

En effet, $n \cdot a = (n \cdot 1_A) \times a = 0_A \times a = 0_A$.

THEOREME 3.2.9
La caractéristique d'un anneau intègre est soit 0, soit un nombre premier.

En particulier la caractéristique d'un corps est soit nulle, soit un nombre premier.

preuve
Si A est intègre, il en est de même de tout sous anneau de A, en particulier de $\tilde{A}$. Supposons que la caractéristique de A, n soit non nulle. $\tilde{A}$ est isomorphe à $\mathbb{Z}/n\mathbb{Z}$, donc ce dernier anneau est intègre et par conséquent n est premier.

Cas des corps

Soit K un corps commutatif. L'intersection de la famille des sous corps de K est un sous corps de K et c'est le plus petit (au sens de l'inclusion) sous corps de K.

DEFINITION 3.2.4
Soit K un corps commutatif. On appelle sous corps premier de K le plus petit sous corps de K.

Soient K un corps commutatif, K_0 son sous corps premier. Notons toujours $\Phi_K : \mathbb{Z} \to K$ l'unique morphisme d'anneaux de $\mathbb{Z}$ dans $\mathbb{K}$ et $\tilde{K} = \Phi_K(\mathbb{Z})$. Le sous corps premier K_0 est en particulier un sous anneau de K ; il contient donc l'anneau $\tilde{K}$. Distinguons deux cas :

- $\mathrm{car}(K) = 0$. Dans ce cas, $\tilde{K}$ est isomorphe à $\mathbb{Z}$. Le morphisme $\Phi_K : \mathbb{Z} \to K$ se prolonge de manière unique à $\mathbb{Q}$ en un morphisme que nous noterons Ψ_K. Celui ci est tel que pour $p, q \in \mathbb{Z},\ q \neq 0,\ \Psi_K(p/q) = \Phi_K(p)\,(\Phi_K(q))^{-1}$. Ψ_K est un isomorphisme de $\mathbb{Q}$ dans K dont l'image est K_0. En effet, cette image est un corps donc elle contient K_0. D'autre part tout sous corps de K contient $\tilde{K}$ et donc aussi les quotients de deux éléments de $\tilde{K}$ c'est à dire $\Psi_K(\mathbb{Q})$. Le détail des vérifications est laissé au lecteur.

  En conclusion, si $\mathrm{car}(K) = 0$, le sous corps premier de K est isomorphe à $\mathbb{Q}$. En particulier, K est nécessairement infini.

- $\mathrm{car}(K) = p, p$ premier. Alors $\tilde{K}$ est isomorphe à $\mathbb{Z}/p\mathbb{Z}$ et est égal au sous corps premier K_0 de K.

Il en résulte que tout corps fini a pour caractéristique un nombre premier.

Complément hors programme
Soient K un corps commutatif fini, $p = \mathrm{car}(K)$ et K_0 le sous corps premier de K. On sait que K_0 est isomorphe à $\mathbb{Z}/p\mathbb{Z}$. Il est immédiat de voir (en anticipant le cours d'algèbre linéaire !) que K a une structure de K_0-espace vectoriel. K est de dimension finie sur K_0. Si on admet que les théorèmes d'algèbre linéaire vu dans le cas des sous corps de $\mathbb{C}$ restent vrais (c'est là qu'on quitte le programme), K est par choix d'une base isomorphe à K_0^n. On en déduit $\mathrm{card}(K) = \mathrm{card}(K_0^n) = p^n$. On a donc prouvé :

THEOREME 3.2.10
Le cardinal d'un corps commutatif fini est une puissance d'un nombre premier p (et $p = \mathrm{car}(K)$).

Remarques culturelles:
1) On démontre (c'est difficile) que tout corps fini est commutatif.
2) On démontre aussi que pour tout nombre premier p et tout entier naturel n il existe un corps ayant exactement $q = p^n$ éléments et que deux tels corps sont isomorphes.

4

Structure d'espace vectoriel

Dans tout ce chapitre, $\mathbb{K}$ est un corps commutatif fixé.
NB : Dans le programme du concours, tous les espaces vectoriels considérés sont des espaces sur un corps $\mathbb{K}$ qui est un sous corps de $\mathbb{C}$. Nous ne ferons pas cette hypothèse dans ce chapitre car elle est inutile.

4.1 Structure

4.1.1 Espace vectoriel

Un $\mathbb{K}$-espace vectoriel est un triplet $(E, +, \cdot)$ où $(E, +)$ est un groupe commutatif dont l'élément neutre sera noté 0_E ou $\vec{0}$ ou même 0 si il n'y a pas d'ambiguité et $\cdot$ une application de $\mathbb{K} \times E$ dans E vérifiant les quatre propriétés suivantes :

1) $\forall x, y \in E, \forall \lambda \in K, \quad \lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$
2) $\forall x \in E, \forall \lambda, \mu \in K, \quad (\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$
3) $\forall x \in E, \forall \lambda, \mu \in K, \quad \lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x$
4) $\forall x \in E, \quad 1 \cdot x = x$
On en déduit immédiatement les règles de calcul suivantes :
5) $\forall x \in E, \quad 0 \cdot x = 0_E$
6) $\forall \lambda \in \mathbb{K}, \quad \lambda \cdot 0_E = 0_E$
7) $\forall \lambda \in \mathbb{K}, \quad \forall x \in E, \ \lambda \cdot x = 0_E \Leftrightarrow (\lambda = 0 \text{ ou } \quad x = 0_E)$
8) $\forall x \in E, \quad \mathrm{opp(x)} = (-1) \cdot \mathrm{x}$ ou $\mathrm{opp}(x)$ désigne l'opposé de x pour la structure de groupe additif de E.

On notera qu'un $\mathbb{K}$-espace vectoriel ne peut pas être vide.
Désormais, sauf exception, on sous entendra le signe $\cdot$ et on notera λx pour $\lambda \cdot x$. On dira aussi "soit E un $\mathbb{K}$-ev "au lieu de "soit $(E, +, \cdot)$ un $\mathbb{K}$-ev ".

Exemples
a) $\mathbb{K}$ muni de ses lois usuelles est un $\mathbb{K}$-espace vectoriel.
b) Soit E un $\mathbb{K}$-espace vectoriel, X un ensemble non vide et $E^X = \mathcal{F}(X, E)$ l'ensemble des applications de X dans E . Pour f, g dans E^X et $\lambda \in \mathbb{K}$ on définit $f + g$ et λf par $(f + g)(x) = f(x) + g(x)$ et $(\lambda f)(x) = \lambda \cdot f(x)$ ceci pour tout x de E. Muni ce ces opérations E^X est un $\mathbb{K}$-espace vectoriel dont l'élément nul est l'application $x \to 0_E$.
c) Espaces produits : soient F et G deux $\mathbb{K}$-ev . Pour (x, y) et (x', y') dans $F \times G$ et λ dans $\mathbb{K}$ on pose $(x, y) + (x', y') = (x + x', y + y')$ et $\lambda \cdot (x, y) = (\lambda x, \lambda y)$. $F \times G$ muni de ces opérations est un $\mathbb{K}$-ev appelé espace vectoriel produit des $\mathbb{K}$-ev F et G.
d) La notion d'espace produit se généralise à un nombre fini quelconque d'espaces. En particulier $\mathbb{K}^n$ muni des opérations usuelles est le produit de n $\mathbb{K}$-espaces vectoriels tous égaux à $\mathbb{K}$.

4.1.2 Sous espace vectoriel

DEFINITION 4.1.1
Soit E un $\mathbb{K}$-ev. Une partie F de E est un sous espace vectoriel si $(F, +)$ est un sous groupe de $(E, +)$ et si F est stable pour la multiplication par un scalaire, i.e. si pour tout x de F et tout λ de $\mathbb{K}$ on a $\lambda x \in F$.

PROPOSITION 4.1.1
Soit E un $\mathbb{K}$-ev et F un sous espace vectoriel de E. Muni des lois induites par celles de E, F est lui même un $\mathbb{K}$-espace vectoriel.

PROPOSITION 4.1.2
*Soit F est une partie **non vide** de E. F est un sev ssi pour tout x, y de F et tout λ de $\mathbb{K}$ on a $\lambda x + y \in F$*

PROPOSITION 4.1.3
Une intersection quelconque de sev d'un $\mathbb{K}$-ev E est un sev de E.

Ceci donne un sens à la définition suivante :

DEFINITION 4.1.2
Soit A une partie d'un $\mathbb{K}$-ev E. On appelle sous espace vectoriel engendré par A et on note $\mathrm{Vect}(A)$ le plus petit sev de E contenant A i.e. l'intersection de tous les sev de E contenant A. En particulier $\mathrm{Vect}(\emptyset) = \{0_E\}$.

EXEMPLE 4.1.1
Si e est un élément non nul de E, l'ensemble des éléments de E de la forme $\{\lambda e\ ;\ \lambda \in \mathbb{K}\}$ est un sous espace vectoriel de E appellé droite vectorielle engendrée par e. On le notera D_e ou $\mathbb{K}e$. Si $e' \in D_e$ est non nul, on a $D_e = D'_e$.

PROPOSITION 4.1.4
Soit A une partie non vide d'un $\mathbb{K}$-ev E et $x \in E$. On a $x \in \mathrm{Vect}(A)$ si et seulement si il existe un entier naturel $n \geqslant 1$, n éléments $a_1, \ldots, a_n$ de A et n éléments $\lambda_1, \ldots, \lambda_n$ de $\mathbb{K}$ tels que $x = \sum_{k=1}^{k=n} \lambda_k a_k$ c'est à dire ssi x est une combinaison linéaire d'éléments de A. (voir définition 4.2.2)

Attention! Si F et G sont deux sev d'un $\mathbb{K}$-ev E dont aucun n'est contenu dans l'autre, la réunion $F \cup G$ n'est pas un sev de E. En particulier, on notera que E n'est jamais la réunion de deux sev propres (i.e. distincts de E).

4.1.3 Applications linéaires

Définitions. Règles de calcul

DEFINITION 4.1.3
Soient E et F deux $\mathbb{K}$-espaces vectoriels. On appelle application ($\mathbb{K}$-)linéaire de E dans F (ou (homo)morphisme de $\mathbb{K}$-ev) toute application $f : E \to F$ vérifiant

$$\forall x \in E,\ \forall y \in E,\ f(x+y) = f(x) + f(y) \quad \textit{et} \quad \forall x \in E,\ \forall \lambda \in \mathbb{K},\ f(\lambda x) = \lambda f(x)$$

Une application linéaire est un homomorphisme du groupe $(E, +)$ dans le groupe $(F, +)$. Donc $f(0_E) = 0_F$ et pour tout $x \in E,\ f(-x) = -f(x)$.
Soit $f : E \to F$; f est linéaire ssi $\forall x, y \in E,\ \forall \lambda \in \mathbb{K},\ f(\lambda x + y) = \lambda f(x) + f(y)$.

DEFINITION 4.1.4
Soient E et F deux $\mathbb{K}$-espaces vectoriels . On appelle isomorphisme de E sur F toute bijection $f : E \to F$ telle que f et la bijection réciproque f^{-1} soient linéaires.

PROPOSITION 4.1.5
Si f est une bijection linéaire du $\mathbb{K}$-ev E sur le $\mathbb{K}$-ev F, f est un isomorphisme de E sur F .

Il suffit de vérifier que la bijection réciproque f^{-1} est linéaire, ce qui est facile.

DEFINITION 4.1.5
Soit E un $\mathbb{K}$-ev. On appelle endomorphisme de E toute application linéaire de E dans E. On appelle automorphisme de E tout isomorphisme de $\mathbb{K}$-ev de E sur E.

PROPOSITION 4.1.6
Si E, F, G sont trois $\mathbb{K}$-espaces vectoriels et si $f : E \to F$ et $g : F \to G$ sont des applications linéaires, l'application $g \circ f : E \to G$ est linéaire.

THEOREME 4.1.1
L'ensemble des automorphismes d'un $\mathbb{K}$-ev E muni de la loi de composition des applications est un groupe appelé groupe linéaire de E et noté $GL(E)$.

Ensembles d'applications linéaires

Soient E, F deux $\mathbb{K}$-ev. On notera $L(E, F)$ l'ensemble des applications linéaires de E dans F et $L(E) = L(E, E)$.
$L(E, F)$ est un sous espace vectoriel de F^E, donc est un $\mathbb{K}$-ev pour les lois $(+, \cdot)$ (4.1.2 exemple b).
Soient E, F, G trois $\mathbb{K}$-ev. Si $f \in L(E, F)$ et $g \in L(F, G)$ alors $gof \in L(E, G)$. On a

(1) $\forall g_1, g_2 \in L(F, G),\ \forall f \in L(E, F), \quad (g_1 + g_2) \circ f = g_1 \circ f + g_2 \circ f$
(2) $\forall g \in L(F, G),\ \forall f \in L(E, F),\ \forall \lambda \in \mathbb{K}, \quad (\lambda g) \circ f = \lambda(g \circ f)$
(3) $\forall g \in L(F, G),\ \forall f_1, f_2 \in L(E, F), \quad g \circ (f_1 + f_2) = g \circ f_1 + g \circ f_2$
(4) $\forall g \in L(F, G),\ \forall f \in L(E, F),\ \forall \lambda \in \mathbb{K}, \quad g \circ (\lambda f) = \lambda(g \circ f)$

Remarque
les relations (1) et (2) sont vraies par définition de la somme $g_1 + g_2$ et du produit λg. Par contre, la justification des relations (3) et (4) utilise de manière essentielle la linéarité de g.

PROPOSITION 4.1.7
$(L(E), +, \circ)$ est un anneau dont l'élément nul est l'application $x \to 0_E$ et l'élément unité l'application identique de E dans E.

Un élément de l'anneau $L(E)$ est inversible ssi $\exists g \in L(E)$ tel que $g \circ f = Id_E$ et $f \circ g = Id_E$, ce qui revient à dire que f est bijective. Les éléments inversibles de $L(E)$ sont donc les automorphismes de E. On a donc $U\left(L(E)\right) = GL(E)$.

Soit k un élément non nul de $\mathbb{K}$. On appelle homothétie de rapport k l'application linéaire $h_k : x \to kx$ de E dans lui même. C'est un automorphisme de E. L'application $k \to h_k$ est un isomorphisme du groupe $(\mathbb{K}^*, \times)$ sur le sous groupe de $GL(E)$ formé des homothéties.

Noyau, image d'une application linéaire

Dans ce paragraphe, E et F sont des $\mathbb{K}$-ev fixés et $f \in L(E, F)$.

PROPOSITION 4.1.8
Soit E' un sous ev de E. L'image $f(E')$ de E' est un sous ev de F.
Soit F' un sous ev de F. L'image réciproque $f^{-1}(F')$ de F' par f est un sous espace vectoriel de E.

DEFINITION 4.1.6
On appelle noyau de l'application linéaire f et on note $\ker(f)$ le sous espace de E image réciproque de $\{0_F\}$.
Donc si x est un élément de E, on a $x \in \ker(f) \Leftrightarrow f(x) = \vec{0}_F$

THEOREME 4.1.2
Soient E et F deux $\mathbb{K}$-ev et $f \in L(E, F)$. On a : $\forall x, y \in E,\ f(x) = f(y) \Leftrightarrow (y - x) \in \ker(f)$.

COROLLAIRE 4.1.1
f est injective ssi $\ker(f) = \{0_E\}$.

DEFINITION 4.1.7
On appelle image de f et on note $\mathrm{im}(f)$ le sous espace $f(E)$ de F. On a donc f surjective $\Leftrightarrow \mathrm{im}(f) = F$.

4.1.4 $\mathbb{K}$-algèbre

DEFINITION 4.1.8
Soient E_1, E_2 et F trois $\mathbb{K}$-espaces vectoriels. Une application $f : E_1 \times E_2 \to F$ est dite bilinéaire si elle vérifie les deux propriétés suivantes :
1) *Pour tout $x_1 \in E_1$ fixé, l'application $x_2 \to f(x_1, x_2)$ de E_2 dans F est linéaire.*
2) *Pour tout $x_2 \in E_2$ fixé, l'application $x_1 \to f(x_1, x_2)$ de E_1 dans F est linéaire.*

On doit donc avoir les propriétés suivantes, pour tous $x_1, x_1' \in E_1,\ x_2, x_2' \in E_2$ et $\lambda \in \mathbb{K}$:
$f(x_1, x_2 + x_2') = f(x_1, x_2) + f(x_1 + x_2')$
$f(x_1, \lambda x_2) = \lambda f(x_1, x_2)$
$f(x_1 + x_1', x_2) = f(x_1, x_2) + f(x_1', x_2)$
$f(\lambda x_1, x_2) = \lambda f(x_1, x_2)$

DEFINITION 4.1.9
Une $\mathbb{K}$-algèbre est un quadruplet $(A, +, \cdot, \star)$ où $(A, +, \cdot)$ est un $\mathbb{K}$-espace vectoriel et $\star$ une application bilinéaire de $A \times A$ dans A.
L'algèbre est dite associative (resp. commutative) si la loi $\star$ est associative (resp. commutative) ; elle est dite unitaire si la loi $\star$ admet un élément neutre, que l'on notera 1_A ou 1.
L'algèbre est dite intègre si $\forall x, y \in A,\ \ x \star y = 0_A \Leftrightarrow x = 0_A$ ou $y = 0_A$.

Si $(A, +, \cdot, \star)$ est une $\mathbb{K}$-algèbre, une sous algèbre B de A est un sous espace vectoriel de A stable par la loi $\star$.

Si $(A, +, \cdot, \star)$ est une $\mathbb{K}$-algèbre associative unitaire, $(A, +, \star)$ est un anneau.

Exemples
1) Si $(\mathbb{K}, +, \cdot)$ est un corps commutatif, $(\mathbb{K}, +, \cdot, \cdot)$ est une $\mathbb{K}$-algèbre associative, commutative unitaire et intègre.
2) Soit X un ensemble non vide. L'ensemble $\mathbb{K}^X = F(X, \mathbb{K})$ des applications de X dans $\mathbb{K}$, muni des opérations usuelles sur les fonctions est une $\mathbb{K}$-algèbre associative, commutative unitaire et non intègre si X n'est pas un singleton.
3) Si (X, d) est un espace métrique, l'ensemble $C(X, \mathbb{K})$ des applications continues de X dans $\mathbb{K}$ est uns sous algèbre de $\mathbb{K}^X$.
4) Soit E un $\mathbb{K}$-espace vectoriel ; $(L(E), +, \cdot, \circ)$ est une $\mathbb{K}$-algèbre associative unitaire en général non commutative et non intègre. Les éléments inversibles de l'anneau $(L(E), +, \circ)$ sont les automorphismes de E.
5) Soit n un entier naturel, $n \geqslant 2$. L'ensemble des matrices carrées d'ordre n muni des lois usuelles est une $\mathbb{K}$-algèbre associative, unitaire non commutative et non intègre sauf si $n = 1$.
6) Soit $(E, \langle\ |\ \rangle)$ un espace euclidien orienté de dimension 3 et $\wedge : E \times E \to E$ le produit vectoriel. $(E, +, \cdot, \wedge)$ est une $\mathbb{R}$-algèbre non associative, non commutative et sans unité.

4.2 Famille de vecteurs

Rappelons d'abord une définition. Soit X et I deux ensembles. Une famille d'éléments de X indexée par I est une application $I \to X,\ i \to x_i$. On note une telle famille $(x_i)_{i \in I}$ ou par abus (x_i). Cette définition généralise celle des suites. Bien noter que la connaissance de l'ensemble image $\{x_i \mid i \in I\}$ ne suffit pas pour connaitre la famille $(x_i)_{i \in I}$.
Si l'ensemble d'indices I est fini, on pourra toujours supposer pour l'exposé qu'il est égal à $\{1, \cdots, n\}$ pour un entier naturel n donné. Dans ce cas on dira, par abus de langage, que n est le nombre d'éléments de la famille. Par exemple $(0, 0, 0)$ est une famille de trois réels tous nuls. Le nombre d'éléments de la famille est égal au cardinal de son image ssi l'application $i \to x_i$ est injective.

Soit Y un sous ensemble de X. L'application inclusion $y \to y$ de Y dans X définit une famille d'éléments de X indexée par Y dont l'image est Y . On l'appellera famille canoniquement associée à Y. Dans ce cas, le nombre d'éléments de la famille est égal au nombre d'éléments de la partie considérée.

DEFINITION 4.2.1
Soit $(G, +)$ un groupe abélien et $(x_i)_{i \in I}$ une famille d'éléments de G indexée par I. On appelle support de cette famille l'ensemble $S = \{i \in I \mid x_i \neq 0_G\}$.
Soit (x_i) une famille d'éléments de G à support S fini. On pose $\sum\limits_{i \in I} x_i := \sum\limits_{i \in S} x_i$.

Si $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ sont deux familles d'éléments de G à support fini, la famille $(x_i + y_i)_{i\in I}$ est elle aussi à support fini et $\sum_{i\in I}(x_i + y_i) = \sum_{i\in I} x_i + \sum_{i\in I} y_i$.
De même, soient E un $\mathbb{K}$-espace vectoriel, $(x_i)_{i\in I}$ une famille d'éléments de E et $(\lambda_i)_{i\in I}$ une famille d'éléments de $\mathbb{K}$ à support fini. Alors la famille $(\lambda_i x_i)_{i\in I}$ est à support fini et donc l'expression $\sum_{i\in I} \lambda_i x_i$ a un sens.

DEFINITION 4.2.2
Soit $(x_i)_{i\in I}$ une famille d'éléments d'un $\mathbb{K}$-espace vectoriel E. On appelle combinaison linéaire des x_i tout élément de E de la forme $\sum_{i\in I} \lambda_i x_i$ où $(\lambda_i)_{i\in I}$ est une famille à support fini d'éléments de $\mathbb{K}$.

4.2.1 Famille génératrice

Soit $(x_i)_{i\in I}$ une famille d'éléments d'un $\mathbb{K}$-espace vectoriel E. L'ensemble des combinaisons linéaires des x_i forme un sous espace vectoriel de E qui n'est autre que le sous espace vectoriel engendré par l'ensemble $\{x_i \mid i \in I\}$ associé à la famille $(x_i)_{i\in I}$. Cet espace sera noté $\mathrm{Vect}(x_i, i \in I)$.

DEFINITION 4.2.3
Soient E un $\mathbb{K}$-espace vectoriel et $(x_i)_{i\in I}$ une famille d'éléments de E. On dit que cette famille est génératrice si le sous espace de E engendré par cette famille est E lui même.
*Il revient au même de dire que tout élément x de E peut s'écrire comme combinaison linéaire d'un **nombre fini** d'éléments x_i (ces éléments x_i dépendant bien sur de x).*

4.2.2 Famille liée, famille libre

DEFINITION 4.2.4
*Une famille finie $(x_1, \cdots, x_n)$ d'éléments d'un $\mathbb{K}$-espace vectoriel E est dite liée si il existe une famille $(\lambda_i)_{1\leqslant i\leqslant n}$ de scalaires, **non tous nuls** tels que $\sum_{1\leqslant i\leqslant n} \lambda_i x_i = 0_E$. On dit aussi que les vecteurs $x_1, \cdots, x_n$ sont linéairement dépendants.*
*Une famille quelconque $(x_i)_{i\in I}$ d'éléments de E est dite liée si il existe une sous famille finie liée (i.e. si il existe $J \subset I$, fini tel que la famille $(x_j)_{j\in J}$ soit liée. Une famille $(x_i)_{i\in I}$ est donc liée si il existe une famille **à support fini** de scalaires **non tous nuls** $(\lambda_i)_{i\in I}$ telle que $\sum_{i\in I} \lambda_i x_i = 0_E$.*

La même définition vaut pour une partie de E.

Une famille (x_i) est liée si et seulement si il existe un indice i_0 tel que le vecteur x_{i_0} s'exprime comme combinaison linéaire des $(x_i)_{i\neq i_0}$.

Remarque
1) Si il existe un indice $i_0 \in I$ tel que $x_{i_0} = 0$, la famille $(x_i)_{i\in I}$ est liée. En effet, si $\lambda_{i_0} = 1$ et $\lambda_i = 0$ si $i \neq i_0$, $(\lambda_i)_{i\in I}$ est une famille à support fini de scalaires non tous nuls tels que $\sum_{i\in I} \lambda_i x_i = 0$.
2) Si il existe deux indices distincts j et k dans I telle que $x_j = x_k$, la famille $(x_i)_i$ est liée. Il suffit pour le voir de prendre $\lambda_j = 1,\ \lambda_k = -1$ et $\lambda_i = 0$ pour $i \neq j,\ i \neq k$.
3) Soit $e \in E$. Le système (e) est libre ssi $e \neq 0_E$.

DEFINITION 4.2.5
Une famille $(x_i)_{i\in I}$ de vecteurs d'un $\mathbb{K}$-espace vectoriel E est dite libre si elle n'est pas liée.
On dit aussi que les vecteurs (x_i) sont linéairement indépendants.

Une famille finie $(x_1, \cdots, x_n)$ est libre ssi $\forall (\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n$, la relation $\sum_{1\leqslant i\leqslant n} \lambda_i x_i = 0_E$ implique $\lambda_1 = 0,\ \lambda_2 = 0, \ldots, \lambda_n = 0$.
Une famille quelconque est libre si toute sous famille finie est libre.

PROPOSITION 4.2.1
Tout système extrait d'un système libre est lui même un système libre.

Exemple
Soit E le $\mathbb{R}$-espace vectoriel des applications continues de $\mathbb{R}$ dans $\mathbb{R}$. Soit, pour $n \in \mathbb{N}$, e_n la fonction définie par $e_n(x) = x^n$ (avec par convention $e_0(x) = 1$) pour tout x. Montrons que cette famille est libre.
Soit $J \subset \mathbb{N}$ fini, $J = \{n_1, \ldots, n_p\}$ avec $n_1 < n_2 < \cdots < n_p$. Soient $\lambda_1, \ldots, \lambda_p$ réels tels que $\sum_{1 \leqslant k \leqslant p} \lambda_k e_k = 0_E$. On a donc pour tout x réel $\sum_{1 \leqslant k \leqslant p} \lambda_k e_k(x) = 0_E$ soit $\sum_{1 \leqslant k \leqslant p} \lambda_k x^k = 0_E$. Supposons les λ_k non tous nuls et soit $q = \max\{k,\ 1 \leqslant k \leqslant p,\ \lambda_k \neq 0\}$. En divisant pour $x \neq 0$ la relation précédente par x^q et en faisant tendre x vers $+\infty$ on obtient $\lambda_q = 0$. L'hypothèse "la famille (e_n) est liée" conduit à une contradiction. Cette famille est donc libre.

4.2.3 Base

DEFINITION 4.2.6
Une famille $(x_i)_{i \in I}$ d'éléments d'un $\mathbb{K}$-espace vectoriel E est une base ssi c'est une famille libre et génératrice.

Supposons que $(x_i)_{i \in I}$ soit une base de E. Tout x de E s'écrit alors $x = \sum_{i \in I} \lambda_i x_i$ où la famille de scalaires $(\lambda_i)_{i \in I}$ est à support fini. De plus une telle écriture est unique car si $x = \sum_{i \in I} \mu_i x_i$, on a $\sum_{i \in I} (\lambda_i - \mu_i) x_i = 0$ où $(\lambda_i - \mu_i)_{i \in I}$ est à support fini, donc $\lambda_i - \mu_i = 0$ pour tout i puisque la famille (x_i) est libre. Les λ_i s'appellent les coordonnées du vecteur x dans la base (x_i). Il est facile de voir que pour un $i \in I$ fixé, l'application $x \to x_i$ est linéaire.

4.3 Somme de sous espaces vectoriels

4.3.1 Somme

PROPOSITION 4.3.1
Soient $F_1, \ldots, F_q$ un nombre fini de sous espaces d'un $\mathbb{K}$-espace vectoriel E. L'ensemble des x de E qui peuvent s'écrire $x = x_1 + \cdots + x_q$ avec, pour tout j entre 1 et q, $x_j \in F_j$ est un sous espace vectoriel de E, appelé somme des sous espaces $F_1, \ldots F_q$. On le note $F_1 + \cdots + F_q$.
Donc $F_1 + \cdots + F_q = \{x \in E \mid \exists x_1 \in F_1, \ldots, \exists x_q \in F_q\ x = x_1 + \cdots + x_q\}$

La réunion d'un nombre fini de sous espaces vectoriels de E n'est en général pas un sous espace de E.

PROPOSITION 4.3.2
Soient $F_1, \ldots, F_q$ des sous espaces d'un $\mathbb{K}$-espace vectoriel E. On a $F_1 + \cdots + F_q = \text{Vect}\left(\bigcup_{1 \leqslant i \leqslant q} F_i\right)$.

Notons Φ l'application de l'espace produit $F_1 \times \cdots \times F_q$ dans E définie par $\Phi(x_1, \ldots, x_q) = x_1 + \cdots + x_q$. C'est une application linéaire et $F_1 + \cdots + F_q = \text{im}(\Phi)$. Ceci prouve la première proposition. Pour la deuxième il est immédiat que tout sous espace vectoriel de E qui contient $F_1, \ldots, F_q$ contient leur somme. Comme celle ci est un sous espace vectoriel c'est le sous espace engendré.

4.3.2 Somme directe

PROPOSITION 4.3.3
Soient $F_1, \ldots, F_q$ un nombre fini de sous espaces d'un $\mathbb{K}$-espace vectoriel E. Les propriétés suivantes sont équivalentes.
(1) $\forall x_1 \in F_1, \ldots, \forall x_q \in F_q,\ x_1 + \cdots + x_q = 0 \Rightarrow x_1 = 0, x_2 = 0, \ldots, x_q = 0$.
(2) *Pour tout $x \in F_1 + \cdots + F_q$ il existe un unique $x_1 \in F_1$, un unique $x_2 \in F_2$, ..., un unique $x_q \in F_q$ tels que $x = x_1 + \cdots + x_q$.*
(3) *L'application $\Phi : F_1 \times \cdots \times F_q \to E$ définie ci dessus est injective.*

Si ces propriétés sont vérifiées on dit que les sous espaces $F_1, \ldots, F_q$ sont linéairement indépendants. On dit aussi, c'est un petit abus de langage que la somme des F_i est directe. On la note alors $F_1 \oplus \cdots \oplus F_q$.

PROPOSITION 4.3.4
Soient F et G deux sous espaces d'un $\mathbb{K}$-espace vectoriel E. Leur somme est directe ssi $F \cap G = \{0\}$.

Attention! Cette caractérisation ne se généralise pas à plus de deux sous espaces!

Etant donné trois sous espaces F, G, H d'un même espace E, pour montrer l'égalité $H = F \oplus G$ il faut donc montrer que $H = F + G$ et que $F \cap G = \{0\}$.

PROPOSITION 4.3.5
Soient $F_1, \ldots, F_q$ un nombre fini de sous espaces d'un $\mathbb{K}$-espace vectoriel E. Les propriétés suivantes sont équivalentes.
(1) *La somme $F_1 + \cdots + F_q$ est directe.*
(2) *Pour tout $j,\ 1 \leqslant j \leqslant q$, on a $F_j \bigcap \left(\sum_{\substack{1 \leqslant i \leqslant q \\ i \neq j}} F_i \right) = \{0\}$.*
(3) *Pour tout $j,\ 2 \leqslant j \leqslant q,\ (F_1 + \cdots + F_{j-1}) \cap F_j = \{0\}$.*

4.3.3 Sous espaces supplémentaires et projecteurs

DEFINITION 4.3.1
Deux sous espaces F et G d'un $\mathbb{K}$-espace vectoriel E sont dits supplémentaires si leur somme est directe et égale à E, i.e. $E = F \oplus G$. On dit aussi que G est UN supplémentaire de F.

F et G sont supplémentaires ssi $E = F + G$ et $F \cap G = \{0\}$.
Soient F et G deux sous espaces supplémentaires de E. Tout $x \in E$ s'écrit donc de manière unique sous la forme $x = x_F + x_G$ avec $x_F \in F$ et $x_G \in G$.
L'application $E \to E$ qui à x associe x_F est linéaire. On l'appelle la projection sur F parallélement à G. On la notera $p_F^{\|G}$. On a $\ker(p_F^{\|G}) = G,\ \mathrm{im}(p_F^{\|G}) = F$ et $p_F^{\|G} \circ p_F^{\|G} = p_F^{\|G}$. De plus $p_F^{\|G} \circ p_G^{\|F} = 0$ et $p_F^{\|G} + p_G^{\|F} = id_E$.

Attention! Un sous espace vectoriel F d'un espace E admet en général plusieurs supplémentaires, et même une infinité si $\mathbb{K}$ est lui même un corps infini. Il ne faut donc jamais utiliser l'expression "le supplémentaire" qui sous entend une unicité.

4.3.4 Projecteurs

DEFINITION 4.3.2
On appelle projecteur de E tout endomorphisme $p \in L(E)$ vérifiant $p \circ p = p$.

Si $E = F \oplus G$, la projection sur F parallélement à G est un projecteur. Réciproquement, on a :

THEOREME 4.3.1
Soit $p \in L(E)$ un projecteur. On a
1) $\mathrm{im}(p) = \{x \in E \mid x = p(x)\}$
2) $E = \ker(p) \oplus \mathrm{im}(p)$
3) p *est la projection sur* $\mathrm{im}(p)$ *parallélement à* $\ker(p)$.
4) $q = id_E - p$ *est un projecteur d'image* $\ker(p)$ *et de noyau* $\mathrm{im}(p)$.

preuve
1) Si $x \in \mathrm{im}(p)$, il existe $y \in E$ tel que $x = p(y)$; alors $p(x) = p \circ p(y) = p(y) = x$. La réciproque est évidente.
2) Si $x \in \ker(p) \cap \mathrm{im}(p)$ on a $p(x) = 0$ et $x = p(x)$ donc $x = 0$. La somme $\ker(p) + \mathrm{im}(p)$ est directe. D'autre part, pour $x \in E$, on a $x = [x - p(x)] + p(x)$ avec $p(x) \in \mathrm{im}(p)$ et $p(x - p(x)) = p(x) - p \circ p(x) = 0$, donc $x - p(x) \in \ker(p)$. On a donc $E \subset \mathrm{im}(p) + \ker(p)$. L'autre inclusion étant évidente, le résultat est prouvé. On a $x_{\mathrm{im}(p)} = p(x)$ et $x_{\ker(p)} = x - p(x)$.
3) On vient de voir que la décomposition de x était $p(x) + [x - p(x)]$, donc $p(x) = p_{\mathrm{im}(p)}^{\| \ker(p)}(x)$.
4) Immédiat.

THEOREME 4.3.2
Soient F, G deux sous espaces supplémentaires d'un $\mathbb{K}$-ev E, E' un $\mathbb{K}$-ev, $f \in L(F, E')$ et $g \in L(G, E')$. Il existe une application linéaire $\varphi : E \to E'$ et une seule telle que $\varphi|_F = f$ et $\varphi|_G = g$.

preuve
Soit p la projection sur F parallélement à G, $q = id_E - p$ la projection sur G parallélement à F. L'application $\varphi = f \circ p + g \circ q$

répond à la question : elle est linéaire, si $x \in F$, on a $p(x) = x$ et $q(x) = 0$ donc $\varphi(x) = f(x)$. De même, si $x \in G$, $\varphi(x) = g(x)$. Soit φ' une autre application linéaire répondant à la question. L'application $\psi = \varphi - \varphi'$ est linéaire, nulle sur F et sur G. Si $x \in E$, on a $\psi(x) = \psi(x_F) + \psi(x_G) = 0$ donc $\psi = 0$ soit $\varphi = \varphi'$. ■.

COROLLAIRE 4.3.1
Soient E, E' deux $\mathbb{K}$-ev et F, G deux sous espaces vectoriels supplémentaires de E.
L'application $\theta : L(E, E') \to L(F, E') \times L(G, E')$ définie par $\theta(f) = (f|_F, f|_G)$ est un isomorphisme.

Il est facile de voir que cette application est linéaire et le théorème précédent montre que tout élément de $L(F, E') \times L(G, E')$ a un unique antécédent.

PROPOSITION 4.3.6
Soient F un sous espace vectoriel d'un $\mathbb{K}$-espace vectoriel E et G_1, G_2 deux sous espaces de E tels que $E = F \oplus G_1 = F \oplus G_2$.
G_1 et G_2 sont isomorphes.

preuve
Notons p la projection sur G_1 parallélement à F. Soit $p' : G_2 \to G_1$ l'application qui à tout $x_2 \in G_2$ associe $p'(x_2) = p(x_2) \in G_1$. On a $p' \in L(G_2, G_1)$. Montrons que p' est un isomorphisme.
1) Soit $x_2 \in \ker(p')$. On a $p(x_2) = 0$ donc $x_2 \in \ker(p) \cap G_2 = F \cap G_2 = \{0\}$. p' est injective.
2) Soit $x_1 \in G_1$. Comme $E = F \oplus G_2$, il existe $y \in F$ et $x_2 \in G_2$ tels que $x_1 = y + x_2$. Il vient $p'(x_2) = p(x_2) = p(x_1 - y) = p(x_1) = x_1$. Donc p' est surjective. ■

4.3.5 Symétries

Dans ce paragraphe on suppose que le corps $\mathbb{K}$ n'est pas de caractéristique 2. C'est évidemment le cas si $\mathbb{K}$ est un sous corps de $\mathbb{C}$.

DEFINITION 4.3.3
Soit E un $\mathbb{K}$-ev. Une symétrie de E est un endomorphisme involutif s de E, i.e. vérifiant $s \circ s = id$.

Une symétrie s est évidemment un automorphisme de E tel que $s = s^{-1}$.

THEOREME 4.3.3
Soit E un $\mathbb{K}$-ev, où $\mathbb{K}$ est un corps de caractéristique différente de 2.
1) *Soit $s \in L(E)$. s est une symétrie ssi l'application $p = \frac{1}{2}(id_E + s)$ est un projecteur.*
2) *Soit $s \in L(E)$. s est une symétrie de E ssi il existe deux sous espaces supplémentaires F et G de E tels que, pour tout $x = x_F + x_G$ de E, on ait $s(x) = x_F - x_G$.*
Alors $F = \ker(s - id_E) = \{x \in E \mid s(x) = x\}$ et $G = \ker(s + id_E) = \{x \in E \mid s(x) = -x\}$.

preuve
1) Le fait que $\mathbb{K}$ soit de caractéristique différente de 2 garantit l'existence de l'élément $1/2$ dans $\mathbb{K}$. Alors

$$\frac{1}{2}(id_E + s) \circ \frac{1}{2}(id_E + s) = \frac{1}{4}(id_E + 2s + s \circ s)$$

Donc $s \circ s = id_E \Leftrightarrow p \circ p = p$.
2) On a alors $s(x) = x \Leftrightarrow p(x) = x$ et $s(x) = -x \Leftrightarrow p(x) = 0$. p étant un projecteur, on a $E = \mathrm{im}(p) \oplus \ker(p) = F \oplus G$ avec $F = \{x \mid s(x) = x\}$ et $G = \{x \mid s(x) = -x\}$. On en déduit que si $x = x_F + x_G$ on a $s(x) = x_F - x_G$.

4.3.6 Généralisation

Soit E un $\mathbb{K}$-espace vectoriel et $E_1, \ldots, E_q$ des sous espaces de E tels que $E = E_1 \oplus \cdots \oplus E_q$. Posons, pour j entre 1 et q, $F_j = \bigoplus_{\substack{1 \leqslant i \leqslant q \\ i \neq j}} E_i$ et soit p_j la projection sur E_j parallélement à F_j. On vérifie facilement les propriétés suivantes :

1) $p_1 + \cdots + p_q = id_E$
2) On a $p_i \circ p_j = 0$ pour $i \neq j$.
3) Toute application linéaire f de E dans un autre $\mathbb{K}$-ev G est entièrement déterminée par ses restrictions f_i à chacun des E_i. On a $f = f_1 \circ p_1 + \cdots f_q \circ p_q$. En particulier, $L(E, G)$ est isomorphe au produit $L(E_1, G) \times \cdots \times L(E_q, G)$.

4.3.7 Théorème fondamental

THEOREME 4.3.4
Soient E et F deux $\mathbb{K}$-espaces vectoriels et $f \in L(E, F)$. Soit S un sous espace vectoriel de E supplémentaire du noyau $\ker(f)$. La restriction de f à S induit un isomorphisme de S sur $\mathrm{im}(f)$.

Remarque
Dans l'énoncé ci dessus on suppose l'existence d'un supplémentaire S du noyau. L'existence d'un tel supplémentaire n'est pas une évidence. On la prouvera dans le prochain chapitre dans le cas où E est de dimension finie. Dans le cas général, on peut prouver que tout sous espace d'un $\mathbb{K}$-espace vectoriel E admet un supplémentaire, mais la démonstration est hors du programme (elle utilise le théorème de Zorn, lui même conséquence de l'axiome du choix).

preuve
Notons f_S l'application de S dans $\mathrm{im}(f)$ définie par $\forall x \in S,\ f_S(x) = f(x)$. C'est évidemment une application linéaire. Soit $x \in \ker(f_S)$. On a $x \in S$ et $f(x) = 0$, donc $x \in \ker(f) \cap S = \{0\}$, i.e; $x = 0$. Par conséquent, f_S est injective. Soit $y \in \mathrm{im}(f)$. Par hypothèse, il existe $x \in E$ tel que $y = f(x)$. x se décompose en $x = x_S + x_{\ker(f)}$ avec $x_S \in S$ et $x_{\ker(f)} \in \ker(f)$. Il vient $y = f(x_S) + f(x_{\ker(f)}) = f_S(x_S)$ ce qui prouve que f_S est surjective. ■.

4.4 Formes linéaires et hyperplans

Soit E un $\mathbb{K}$-espace vectoriel.

DEFINITION 4.4.1
On appelle forme linéaire sur E toute application linéaire de E dans le corps de base $\mathbb{K}$. L'ensemble des formes linéaires $L(E, \mathbb{K})$ est naturellement muni d'une structure d'espace vectoriel. On l'appelle espace dual de E et on le note E^*.

DEFINITION 4.4.2
*On appelle hyperplan de E tout sous espace **propre** de E, maximal pour la relation d'ordre "inclusion". Autrement dit, un sev H de E est un hyperplan ssi il possède les deux propriétés suivante :*
1) *$H \neq E$.*
2) *Les seuls sev F de E vérifiant $H \subset F$ sont $F = H$ et $F = E$.*

THEOREME 4.4.1
Soit E un $\mathbb{K}$-espace vectoriel et H un sous espace vectoriel de E. Les propriétés suivantes sont équivalentes :
(1) *H est un hyperplan.*
(2) *Il existe $e_0 \in E$ tel que $E = H \oplus \mathbb{K}e_0$.*
(3) *$H \neq E$ et pour tout $e \in E \setminus H$, on a $E = H \oplus \mathbb{K}e$.*

preuve
(1) $\Rightarrow$ (3) Puisque H est un hyperplan, il est distinct de E. Soit $e \in E \setminus H$ et $F = H + \mathbb{K}e$. F est un sev de E contenant l'hyperplan H et distinct de H (puisque contenant $e \notin H$). H étant un hyperplan, on a $F = E$. D'autre part, si $x \in H \cap \mathbb{K}e$, il existe $\lambda \in \mathbb{K}$ tel que $x = \lambda e$. Supposons $x \neq 0$. On a alors $\lambda \neq 0$ donc $e = \frac{1}{\lambda}x \in H$. Contradiction. Donc $x = 0$ et la somme $H + \mathbb{K}e$ est directe.
(3) $\Rightarrow$ (2) Puisque $H \neq E$, l'ensemble $E \setminus H$ est non vide, d'où la conclusion.
(2) $\Rightarrow$ (1) Soit F un sev de E tel que $H \subset F$. Il s'agit de prouver que $F = H$ ou $F = E$. Si on suppose $F \neq H$, il existe $e \in F \setminus H \subset E \setminus H$. D'après (2), e s'écrit $e = e_H + \lambda e_0$ avec $e_H \in H$ et $\lambda \in \mathbb{K}$. Comme $e \notin H$, $\lambda \neq 0$ et donc $e_0 = \frac{1}{\lambda}(e - e_H) \in F$. On en déduit $E = H + \mathbb{K}e_0 \subset F$ d'où $F = E$. ■

THEOREME 4.4.2
Soit E un $\mathbb{K}$-espace vectoriel.
1) *Soit $\varphi \in E^*$ une forme linéaire non nulle sur E. $\ker(\varphi)$ est un hyperplan de E.*

2) *Soit H un hyperplan de E. Il existe une forme linéaire $\varphi \in E^*$ non nulle telle que $H = \ker(\varphi)$.*
3) *Soient φ et ψ deux formes linéaires non nulles sur E. On a $\ker(\varphi) = \ker(\psi) \Leftrightarrow \exists \lambda \in \mathbb{K},\ \lambda \neq 0,$ tel que $\psi = \lambda\varphi$.*

preuve
1) Puisque $\varphi \neq 0$, il existe $e \in E$ tel que $\varphi(e) \neq 0$. Posons $\varepsilon = \dfrac{1}{\varphi(e)}e$ de sorte que $\varphi(\varepsilon) = 1$.
On a de suite $\ker(\varphi) \cap \mathbb{K}\varepsilon = \{0\}$. Soit $x \in E$. On peut écrire $x = x - \varphi(x)\varepsilon + \varphi(x)\varepsilon$ et $\varphi(x)\varepsilon \in \mathbb{K}\varepsilon$ alors que $\varphi[x - \varphi(x)\varepsilon] = \varphi(x) - \varphi(x)\varphi(\varepsilon) = 0$ donc $x - \varphi(x)\varepsilon \in \ker(\varphi)$. On a donc $E = \ker(\varphi) \oplus \mathbb{K}\varepsilon$ ce qui prouve que $\ker(\varphi)$ est un hyperplan.

2) Soient H un hyperplan, $e \in E$ tel que $E = H \oplus \mathbb{K}e$. Soit p la projection sur $\mathbb{K}e$ parallélement à H. On a pour $x \in E$, $p(x) = \varphi(x)e$ avec $\varphi(x) \in \mathbb{K}$ et il est immédiat que l'application $x \to \varphi(x)$ est linéaire. Donc $\varphi \in E^*$ et on a $x \in \ker(\varphi) \Leftrightarrow p(x) = 0 \Leftrightarrow x \in H$.

3) Soit $H = \ker(\varphi) = \ker(\psi)$; on sait que H est un hyperplan. Soient $e \in E$ tel que $E = H \oplus \mathbb{K}e$ et $\theta = \varphi(e)\psi - \psi(e)\varphi$. $x \in E$ s'écrit $x = x_H + te$ avec $x_H \in H$ et $t \in \mathbb{K}$. On a $\theta(x) = \theta(x_H) + t\theta(e) = 0$. Donc $\theta = 0$. Comme $e \notin H,\ \varphi(e) \neq 0$ donc $\psi = \lambda\varphi$ avec $\lambda = \dfrac{\psi(e)}{\varphi(e)}$.

Dans le courant de la démonstration (partie 1)) on a montré le :

COROLLAIRE 4.4.1
Soit φ une forme linéaire non nulle sur un $\mathbb{K}$-espace vectoriel E. Il existe un vecteur $\varepsilon \in E$ tel que $\varphi(\varepsilon) = 1$.

COROLLAIRE 4.4.2
Soient φ et ψ deux formes linéaires sur un $\mathbb{K}$-espace vectoriel E. Il existe $\lambda \in \mathbb{K}$ tel que $\psi = \lambda\varphi$ si et seulement si $\ker(\varphi) \subset \ker(\psi)$.

Vérification facile en distinguant le cas $\psi = 0$ du cas $\psi \neq 0$.

A tout hyperplan H de E correspond ainsi une droite vectorielle D_H dans E^*, droite formée des $\varphi \in E^*$ qui s'annulent sur H. Réciproquement, si D est une droite dans E^*, l'hyperplan $\ker(\varphi)$ ne dépend pas de l'élément non nul $\varphi \in D$. Cette correspondance sera généralisée dans le cas de la dimension finie. (Voit le chapitre sur la dualité).

5

Espaces vectoriels de dimension finie

Dans tout ce chapitre, $\mathbb{K}$ est un corps commutatif fixé. Pour n entier naturel non nul fixé on pose $\mathbb{N}_n^* = \{1, \dots, n\}$.

5.1 Préliminaires

On va s'intéresser dans ce chapitre aux espaces vectoriels qui admettent un système générateur fini. On commence par quelques résultats préliminaires. Dans les 5 lemmes qui suivent, E est un $\mathbb{K}$-espace vectoriel fixé.

Lemme 5.1.1
Soient $S = (u_1, \dots, u_n)$ et $S' = (u'_1, \dots, u'_m)$ deux systèmes de vecteurs de E. On a

$$\mathrm{Vect}(S) \subset \mathrm{Vect}(S') \Leftrightarrow [\forall i \in \mathbb{N}_n^*,\ u_i \in \mathrm{Vect}(S')]$$

Si un système $(v_1, \dots, v_m)$ est lié, l'un des vecteurs v_j s'exprime comme combinaison linéaire des autres mais on ne peut pas sans autre hypothèse savoir lequel. Le lemme suivant répond à cette question avec une hypothèse supplémentaire.

Lemme 5.1.2
Soit $S = (u_1, \dots, u_n)$ un système libre de vecteurs de E et $v \in E$. Si le système $(u_1, \dots, u_n, v)$ est lié on a $v \in \mathrm{Vect}(S)$

preuve

Le système $(u_1, \dots, u_n, v)$ est lié, donc il existe des scalaires $\lambda_1, \dots, \lambda_n, \lambda$ non tous nuls tels que $\sum_{k=1}^{k=n} \lambda_k u_k + \lambda v = 0$. On a nécessairement $\lambda \neq 0$ sinon, le système $(u_1, \dots, u_n) = S$ serait lié contrairement à l'hypothèse. En divisant par λ on obtient que v s'exprime comme combinaison linéaire des vecteurs de S.

Lemme 5.1.3
Soit S un système lié de $n \geqslant 2$ vecteurs de E. Il existe un système S', extrait de S, de $n-1$ vecteurs tels que $\mathrm{Vect}(S') = \mathrm{Vect}(S)$.

preuve
Soit $S = (u_1, \dots, u_n)$. S est lié, donc il existe $k \in \mathbb{N}_n^*$ tel que u_k soit combinaison linéaire des vecteurs $(u_j)_{j \neq k}$. Le système $S' = (u_1, \dots, u_{k-1}, u_{k+1}, \dots, u_n)$ vérifie $\mathrm{Vect}(S') = \mathrm{Vect}(S)$ d'après le lemme 5.1.1.

Lemme 5.1.4
Soit S un système de n vecteurs de E non tous nuls. Il existe un système libre extrait S' tel que $\mathrm{Vect}(S') = \mathrm{Vect}(S)$. *Le système S' est alors une base de* $\mathrm{Vect}(S)$

preuve
La preuve se fait par récurrence sur n. Si $n = 1$, $S = (u_1)$ avec $u_1 \neq 0$ est libre, donc $S' = S$ convient. Supposons le lemme

vrai pour un entier $n \geqslant 1$ donné. Soit S un système de $n+1$ vecteurs. Si S est libre, on prend $S' = S$. Sinon, d'après le lemme 5.1.3 il existe un système extrait S_1 formé de n vecteurs tel que $\mathrm{Vect}(S_1) = \mathrm{Vect}(S)$. D'après l'hypothèse de récurrence il existe un système S' extrait de S_1, donc de S, libre tel que $\mathrm{Vect}(S_1) = \mathrm{Vect}(S')$. D'où la conclusion.

Lemme 5.1.5
Soit $p \geqslant 1$ un entier naturel, $S = (u_1, \ldots, u_p)$ un système de p vecteurs de E. Soit $\Sigma = (v_1, \ldots, v_{p+1})$ un système de $p+1$ vecteurs de $\mathrm{Vect}(S)$. Le système Σ est lié.

preuve
La démonstration se fait par récurrence sur p. Soit (HR_p) la propriété suivante : pour tout système $S = (u_1, \ldots, u_p)$ de p vecteurs de E. et tout système $\Sigma = (v_1, \ldots, v_{p+1})$ de $p+1$ vecteurs de $\mathrm{Vect}(S)$, le système Σ est lié.

- Si $p = 1$, $S = (u_1)$, $\Sigma = (v_1, v_2)$. Par hypothèse, il existe des scalaires λ_1, λ_2 tels que $v_1 = \lambda_1 u_1$, $v_2 = \lambda_2 u_1$. Si $\lambda_1 = 0$, le système Σ est lié ; sinon, la relation $\lambda_1 v_2 - \lambda_2 v_1 = 0$ avec $(\lambda_1, \lambda_2) \neq (0, 0)$ montre que Σ est lié.
- $(HR_p) \Rightarrow (HR_{p+1})$ où $p \geqslant 1$.
  Soit donc $S = (u_1, \ldots, u_{p+1})$ un système de $p+1$ vecteurs de E et $\Sigma = (v_1, \ldots, v_{p+2})$ un système de $p+2$ vecteurs de $\mathrm{Vect}(S)$. Par hypothèse il existe des $\lambda_{i,j} \in \mathbb{K}$ tels que

$$\begin{aligned} v_1 &= \lambda_{1,1}u_1 + \cdots + \lambda_{p+1,1}u_{p+1} \\ &\vdots \\ v_j &= \lambda_{1,j}u_1 + \cdots + \lambda_{p+1,j}u_{p+1} \\ &\vdots \\ v_{p+2} &= \lambda_{1,p+2}u_1 + \cdots + \lambda_{p+1,p+2}u_{p+1} \end{aligned}$$

  On distingue alors deux cas :

  – Pour tout j entre 1 et $p+2$ $\quad \lambda_{1,j} = 0$.
  Les vecteurs $v_1, \cdots, v_{p+1}$ appartiennent alors à l'espace vectoriel engendré par le système de p vecteurs $S' = (u_2, \ldots, u_{p+1})$. D'après l'hypothèse de récurrence, $(v_1, \ldots, v_{p+1})$ est lié et il en est donc de même de S.
  – Il existe un indice j tel que $\lambda_{1,j} \neq 0$. Quitte à renuméroter les vecteurs de S, on peut supposer $j = 1$, i.e. $\lambda_{1,1} \neq 0$. Pour $2 \leqslant j \leqslant p+2$ posons $w_j = v_j - \dfrac{\lambda_{1,j}}{\lambda_{1,1}} v_1$. Alors $w_j \in \mathrm{Vect}(S')$ où $S' = (u_2, \ldots, u_{p+1})$. D'après l'hypothèse de récurrence, le système de p vecteurs $(w_2, \ldots, w_{p+1})$ est lié. Il existe donc des scalaires non tous nuls $\mu_2, \ldots, \mu_{p+1}$ tels que $\sum_{j=2}^{p+1} \mu_j w_j = 0$. Cette relation s'écrit $\sum_{j=1}^{p+2} \mu_j v_j = 0$ en posant $\mu_1 = -\dfrac{1}{\lambda_{1,1}} \left(\sum_{j=2}^{p+1} \mu_j \lambda_{1,j} \right)$ et les μ_j ne sont pas tous nuls, donc le système Σ est lié.

La preuve est complète.

5.2 Théorèmes fondamentaux

THEOREME 5.2.1
Soit $E \neq \{0\}$ un $\mathbb{K}$-espace vectoriel non réduit à $\{0\}$ admettant un système générateur fini. Alors
1) *E admet des bases.*
2) *Deux bases quelconques de E ont le même nombre de vecteurs.*

DEFINITION 5.2.1
On dit qu'un $\mathbb{K}$-espace vectoriel est de dimension finie si il est engendré par un nombre fini de vecteurs. Si E est de dimension finie et non nul, le nombre de vecteurs d'une base quelconque de E est appelé dimension de E et noté $\dim(E)$. Par convention $\dim(\{0\}) = 0$.

preuve du théorème
Soit S un système fini générateur de E. D'après le lemme 5.1.4 il existe un système libre S' extrait de S qui est une base de $\mathrm{Vect}(S) = E$. Ceci prouve la première affirmation.

Soit $\mathcal{B} = (e_1, \ldots, e_n)$ et $\mathcal{B}' = (e'_1, \ldots, e'_m)$ deux bases de E. D'après le lemme 5.1.5 appliqué à $S = \mathcal{B}$ et $\Sigma = \mathcal{B}'$, $m \leqslant n$ (car l'hypothèse $m > n$ implique que le système $\mathcal{B}'$ est lié).De même, symétriquement, $n \leqslant m$. Donc $n = m$.

COROLLAIRE 5.2.1
Pour tout entier naturel n, le $\mathbb{K}$-espace vectoriel $\mathbb{K}^n$ est de dimension finie égale à n.

Le système $(e_1, \ldots, e_n)$ où $e_i = (\delta_{i1}, \ldots, \delta_{in}) = (0, \ldots, 0, 1, 0, \ldots, 1)$ est une base de $\mathbb{K}^n$ appellé base canonique.

Exemples
Un espace vectoriel est une droite vectorielle ssi il est de dimension 1.
On appelle plan vectoriel tout espace vectoriel de dimension 2.

THEOREME 5.2.2
Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n et $\mathcal{G}$ un système générateur de m vecteurs. Alors
1) $m \geqslant n$
2) *Si $m = n$, $\mathcal{G}$ est une base.*

preuve
1) Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. Les vecteurs de $\mathcal{B}$ appartiennent à $\mathrm{Vect}(\mathcal{G})$ et forment un système libre, donc (lemme 5.1.5) $n \leqslant m$.
2) Supposons $\mathcal{G}$ lié. D'après le lemme 5.1.3 il existe un système $\mathcal{G}'$ extrait de $\mathcal{G}$ formé de $m - 1$ vecteurs qui est encore générateur. D'après le 1) on a alors $m - 1 \geqslant n$ soit $m \geqslant n + 1$. Donc si $m = n$ le système $\mathcal{G}$ est libre.

THEOREME 5.2.3
Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n et $\mathcal{L}$ un système libre de m vecteurs. Alors
1) $m \leqslant n$
2) *Si $m = n$, $\mathcal{L}$ est une base.*

preuve
1) Celà résulte immédiatement du lemme 5.1.5.
2) Supposons $m = n$. Soient $v \in E$ et $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. Le système L est libre et le système (L, v) est lié car formé de $n + 1$ vecteurs appartenant à $\mathrm{Vect}(\mathcal{B})$ (lemme 5.1.5). Donc (5.1.2), v appartient à $\mathrm{Vect}(\mathcal{L})$. Par conséquent L est libre et générateur. C'est une base.

RESUME
Soit E un $\mathbb{K}$-espace vectoriel de dimension n.

1. Tout système générateur a au moins n vecteurs.
2. Tout système libre a au plus n vecteurs.
3. Si $\mathcal{B}$ est un système de n vecteurs, on a l'équivalence : $\mathcal{B}$ est une base $\Leftrightarrow$ $\mathcal{B}$ est libre $\Leftrightarrow$ $\mathcal{B}$ est générateur.

THEOREME 5.2.4
Soit E un $\mathbb{K}$-espace vectoriel de dimension finie et F un sous espace vectoriel de E. F est de dimension finie et $\dim(F) \leqslant \dim(E)$. De plus $F = E \Leftrightarrow \dim(F) = \dim(E)$.

preuve
Un système libre de F est un système libre de E, donc il a au plus $n = \dim(E)$ vecteurs. Si $F = \{0\}$ on a $\dim(F) = 0 \leqslant \dim(E)$. Sinon, soit $\mathcal{B}_F$ une base de F. C'est un système libre de E, formé de $\dim(F)$ vecteurs, donc $\dim(F) \leqslant \dim(E)$. Si $\dim(F) = \dim(E)$, $\mathcal{B}_F$ est un système libre de E de $\dim(E)$ vecteurs donc est une base de E. Donc $E \subset \mathrm{Vect}(\mathcal{B}_F) = F$. La réciproque est triviale.

COROLLAIRE 5.2.2
Soient F et G deux sous espaces vectoriels d'un $\mathbb{K}$-espace vectoriel de dimension finie. On a l'équivalence

$$F = G \Leftrightarrow \begin{cases} F \subset G \\ \dim(F) = \dim(G) \end{cases}$$

Attention Si on a seulement $\dim(F) = \dim(G)$ on ne peut pas en conclure que $F = G$.

5.3 Théorème de la base incomplète et applications

THEOREME 5.3.1
Soient E un $\mathbb{K}$-espace vectoriel, $\mathcal{L}$ une partie libre et $\mathcal{G}$ une partie génératrice finie de E. Il existe un sous ensemble $\mathcal{G}_1 \subset \mathcal{G}$ tel que $L \cup \mathcal{G}_1$ soit une partie base de E.

preuve
Les hypothèses du théorème font que E est de dimension finie n. Soit $p = \text{card}(\mathcal{L})$. On sait que $p \leqslant n$. La preuve se fait par récurrence descendante sur p. Si $p = n$, $\mathcal{L}$ est une base et $\mathcal{G}_1 = \emptyset$ convient. Supposons le résultat prouvé pour un entier p vérifiant $1 \leqslant p \leqslant n$. Soit maintenant $\mathcal{L}$ une partie libre de $p-1$ éléments. Comme $p - 1 < n$, $\mathcal{L}$ n'est pas génératrice, donc il existe $v \in \mathcal{G}$ tel que $v \notin \text{Vect}(\mathcal{L})$. Ceci implique que la partie $\mathcal{L}' = \mathcal{L} \cup \{v\}$ est libre. Son cardinal est p. D'après l'hypothèse de récurrence, il existe $\mathcal{G}' \subset \mathcal{G}$ tel que $\mathcal{L}' \cup \mathcal{G}'$ soit une partie base de E. Si on pose $\mathcal{G}_1 = \mathcal{G}' \cup \{v\}$, on voit que $\mathcal{G}_1 \subset \mathcal{G}$ et que $\mathcal{L} \cup \mathcal{G}_1$ est une base de E.

Donnons maintenant l'énoncé du même résultat en termes de systèmes de vecteurs.

COROLLAIRE 5.3.1
Soient E un $\mathbb{K}$-espace vectoriel, $\mathcal{L} = (e_1, \ldots, e_p)$ un système libre et $\mathcal{G} = (u_1, \ldots, u_q)$ un système générateur fini de E. Il existe un entier m et des indices $i_1, \ldots, i_m$, $1 \leqslant i_1 < i_2 < \cdots < i_m \leqslant q$ tels que $(e_1, \ldots, e_p, u_{i_1} \ldots, u_{i_m})$ soit une base de E.

On a bien entendu $p + m = \dim(E)$.

COROLLAIRE 5.3.2
Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n et $\mathcal{L} = (e_1, \ldots, e_p)$ un système libre de E. Il existe des vecteurs $(e_{p+1}, \ldots, e_n)$ tels que $(e_1, \ldots, e_n)$ soit une base de E.

Exemple d'utilisation : Soit E un $\mathbb{K}$-ev de dimension n et $F \subset G$ deux sous espaces vectoriels de E de dimensions p et q respectivement. Il existe une base $(e_1, \ldots, e_n)$ de E telle que $(e_1, \ldots, e_p)$ soit une base de F et $(e_1, \ldots, e_q)$ une base de G. Pour construire une telle base, on choisit une base $(e_1, \ldots, e_p)$ de F et on applique le théorème de la base incomplète dans G à ce système libre. On en déduit l'existence de vecteurs $e_{p+1}, \ldots, e_q$ de G tels que $(e_1, \ldots, e_q)$ soit une base de G, puis on recommence dans E avec le système libre $(e_1, \ldots, e_q)$. Ce résultat se généralise à une suite $E_1 \subset E_2 \subset \cdots \subset E_m \subset E$ de sev de E.

Si $S = (a_1, \ldots, a_p)$ et $S' = (b_1, \ldots, b_q)$ sont deux systèmes de vecteurs de E, (S, S') désignera le système $(a_1, \ldots, a_p, b_1, \ldots, b_q)$.

THEOREME 5.3.2 (Sous espaces supplémentaires)
Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, F et G des sous espaces vectoriels non nuls de E. Les propriétés suivantes sont équivalentes :
1) *$E = F \oplus G$ i.e; F et G sont supplémentaires.*
2) *Il existe une base $\mathcal{B}_F$ de F et une base $\mathcal{B}_G$ de G telles que $(\mathcal{B}_F, \mathcal{B}_G)$ soit une base de E.*
3) *Pour toutes bases $\mathcal{B}_F$ de F et $\mathcal{B}_G$ de G, $(\mathcal{B}_F, \mathcal{B}_G)$ est une base de E.*
4) *$F + G = E$ et $\dim(F) + \dim(G) = \dim(E)$.*
5) *$F \cap G = \{0\}$ et $\dim(F) + \dim(G) = \dim(E)$.*

Si on ne suppose plus que les sous espaces F et G sont non nuls, on a encore l'équivalence entre 1), 3), 4) et 5). La propriété 2) peut être mise en défaut puisqu'un espace nul n'a pas de bases.

preuve
Soient $\mathcal{B}_F = (e_1, \ldots, e_p)$ une base de F, $\mathcal{B}_G = (\varepsilon_1, \ldots, \varepsilon_q)$ une base de G, $\mathcal{B} = (\mathcal{B}_F, \mathcal{B}_G)$. La preuve du théorème repose sur les deux équivalences suivantes :

(1) $\mathcal{B}$ est libre $\Leftrightarrow F \cap G = \{0\}$

Supposons $F \cap G = \{0\}$. Soient $(\lambda_1, \ldots, \lambda_p)$ et $(\mu_1, \ldots, \mu_q)$ des scalaires tels que $\sum_{i=1}^{i=p} \lambda_i e_i + \sum_{j=1}^{j=q} \mu_j \varepsilon_j = 0$. Posons $y = \sum_{i=1}^{i=p} \lambda_i e_i$ et $z = \sum_{j=1}^{j=q} \mu_j \varepsilon_j$. On a $y \in F$ et $z \in G$ et $y + z = 0$ donc $z = -y \in F$. Par conséquent, $z \in F \cap G$ donc $z = 0$ d'où $y = 0$. On en déduit la nullité des scalaires λ_i et μ_j pour $1 \leqslant i \leqslant p$ et $1 \leqslant j \leqslant q$ puisque les systèmes $\mathcal{B}_F$ et $\mathcal{B}_G$ sont libres. Donc $\mathcal{B}$ est libre.

Réciproquement, supposons $\mathcal{B}$ libre. Soit $x \in F \cap G$. x s'écrit $x = \sum_{i=1}^{i=p} \lambda_i e_i$ et aussi $x = \sum_{j=1}^{j=q} \mu_j \varepsilon_j$ où les λ_i, μ_j sont des scalaires. Comme $\mathcal{B}$ est libre, on déduit de $\sum_{i=1}^{i=p} \lambda_i e_i - \sum_{j=1}^{j=q} \mu_j \varepsilon_j = 0$ la nullité des λ_i et des μ_j et donc celle de x.

(2) $\mathcal{B}$ est générateur $\Leftrightarrow E = F + G$
Si S et S' sont deux systèmes de vecteurs, on vérifie facilement que $\mathrm{Vect}(S, S') = \mathrm{Vect}(S) + \mathrm{Vect}(S')$. L'affirmation en découle.

Démontrons maintenant le théorème.

- $1) \Rightarrow 3)$ Soient $\mathcal{B}_F = (e_1, \ldots, e_p)$ une base de F, $\mathcal{B}_G = (\varepsilon_1, \ldots, \varepsilon_q)$ une base de G, $\mathcal{B} = (\mathcal{B}_F, \mathcal{B}_G)$. Par hypothèse $E = F \oplus G$. On a $E = F + G$ donc $\mathcal{B}$ est un système générateur et $F \cap G = \{0\}$ donc $\mathcal{B}$ est un système libre.
- $3) \Rightarrow 2)$ est clair puisque F et G étant non nuls et contenus dans E de dimension finie, possèdent des bases.
- $2) \Rightarrow 4)$ $E = \mathrm{Vect}(\mathcal{B}_F, \mathcal{B}_G) = \mathrm{Vect}(\mathcal{B}_F) + \mathrm{Vect}(\mathcal{B}_G) = F + G$. La relation sur les dimensions résulte de l'hypothèse.
- $4) \Rightarrow 5)$ Soient $\mathcal{B}_F$ une base de F, B_G une base de G, $\mathcal{B} = (\mathcal{B}_F, \mathcal{B}_G)$. On a $F + G = E \Rightarrow \mathcal{B}$ est un système générateur de E. L'hypothèse sur les dimensions assure que le nombre de vecteurs de $\mathcal{B}$ est égal à $\dim(E)$; $\mathcal{B}$ est donc une base de E, et en particulier un système libre. Donc $F \cap G = \{0\}$.
- $5) \Rightarrow 1)$ Soit toujours $\mathcal{B}_F$ une base de F, $\mathcal{B}_G$ une base de G, $\mathcal{B} = (\mathcal{B}_F, \mathcal{B}_G)$. On a $\mathrm{Vect}(\mathcal{B}_F) \cap \mathrm{Vect}(\mathcal{B}_G) = F \cap G = \{0\}$ donc $\mathcal{B}$ est un système libre de E. Il est formé de $\dim(E)$ vecteurs, donc c'est une base de E. On en déduit $E = \mathrm{Vect}(\mathcal{B}) = \mathrm{Vect}(\mathcal{B}_F) + \mathrm{Vect}(\mathcal{B}_G) = F + G$ ce qui joint à l'hypothèse $F \cap G = \{0\}$ assure que $E = F \oplus G$.

THEOREME 5.3.3
Tout sous espace vectoriel d'un $\mathbb{K}$-espace vectoriel de dimension finie admet des supplémentaires.

preuve
Soit E de dimension n et F un sous espace de E. Soit $p = \dim(F) \leqslant n$. Si $p = 0$, i.e. $F = \{0\}$, E convient. Si $p = n$ on a $F = E$ et $\{0\}$ convient. Sinon, $1 \leqslant p \leqslant n - 1$; soit $(e_1, \ldots, e_p)$ une base de F et $(e_{p+1}, \ldots, e_n)$ tels que $(e_1, \ldots, e_n)$ soit une base de E. Soit enfin $G = \mathrm{Vect}(e_{p+1}, \ldots, e_n)$. D'après le théorème précédent, G est un supplémentaire de F.

Attention! En général un sous espace de E admet plusieurs supplémentaires.
Considérons le cas d'un $\mathbb{K}$-espace vectoriel E de dimension $n \geqslant 2$ où $\mathbb{K}$ est un corps infini, ce qui est le cas si $\mathbb{K}$ est un sous corps de $\mathbb{C}$. (Cadre du programme). Soit F un sous espace de E non nul et non égal à E. Alors F admet une infinité de supplémentaires.
En effet, soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E telle que $\mathcal{B}_F := (e_1, \ldots, e_p)$ soit une base de F. On a par hypothèse $1 \leqslant p \leqslant n - 1$. Soit, pour $\lambda \in \mathbb{K}$, $G_\lambda = \mathrm{Vect}(e_{p+1}, \ldots e_n + \lambda e_1)$.
Alors, pour tout $\lambda \in \mathbb{K}$, G_λ est un supplémentaire de F et $G_\lambda \neq G_\mu$ si $\lambda \neq \mu$. La peuve est facile et laissée au lecteur.

5.4 Rang

DEFINITION 5.4.1
Soit $S = (u_1, \ldots, u_n)$ un système de vecteurs d'un $\mathbb{K}$-espace vectoriel E. On appelle rang du système S et on note $\mathrm{rg}(S)$ la dimension du sous espace vectoriel $\mathrm{Vect}(S)$ de E.

Cette définition s'étend à un système $S = (u_i)_i$ où l'ensemble d'indice I peut être infini, pourvu que $\mathrm{Vect}(S)$ soit de dimension finie.

PROPOSITION 5.4.1
Le rang d'un système S est égal au nombre maximum de vecteurs d'un système libre extrait de S.

C'est une conséquence immédiate du lemme 5.1.4 .

Opérations élémentaires
Soit $S = (u_1, \ldots, u_n)$ un système de vecteurs.
On dit que le système $S' = (v_1, \ldots, v_n)$ se déduit de S par une opération élémentaire de type I si S' s'obtient à partir de S en permutant deux vecteurs, autrement dit si il existe deux indices $i < j$ compris entre 1 et n tels que $v_i = u_j,\ v_j = u_i$ et $v_k = u_k$ pour $k \neq i$ et $k \neq j$.
On dit que le système S' se déduit du système S par une opération élémentaire de type II si il existe deux indices i, j et un scalaire $\lambda \in \mathbb{K}$ tels que $v_k = u_k$ pour $k \neq i$ et $v_i = u_i + \lambda u_j$. Autrement dit, on remplace dans S un des vecteurs, u_i par ce vecteur auquel on ajoute λ fois un autre vecteur u_j avec, c'est essentiel, $i \neq j$.
On dit que le système S' se déduit du système S par une opération élémentaire de type III si on l'obtient à partir de S en multipliant l'un des vecteurs de S par un scalaire non nul, i.e. si il existe i entre 1 et n et $r \in \mathbb{K}^*$ tels que $v_k = u_k$ pour $k \neq i$ et $v_i = ru_i$.
On dit que le système S' se déduit du système S par une suite d'opérations élémentaires si il existe une suite finie de systèmes, $S_1, \ldots, S_p$ tels que $S_1 = S,\ S_p = S'$ et tels que S_{j+1} se déduit de S_j par une opération élémentaire de type I , II ou III pour $1 \leqslant j \leqslant p-1$.
Remarque
Dans un certain nombre de questions, on se limite à des opérations élémentaires de type I ou II.

THEOREME 5.4.1
Soient S et S' deux systèmes de vecteurs d'un $\mathbb{K}$-ev E. Si S' se déduit de S par une suite d'opérations élémentaires, on a $\mathrm{Vect}(S) = \mathrm{Vect}(S')$ et par conséquent les deux systèmes ont même rang.

La preuve consiste à vérifier que $\mathrm{Vect}(S) = \mathrm{Vect}(S')$ si S' se déduit de S par une transformation élémentaire, ce qui est facile.

5.5 Applications linéaires

5.5.1 Premiers résultats

PROPOSITION 5.5.1
Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base d'un $\mathbb{K}$-ev de dimension finie E, et $\varepsilon_1, \ldots, \varepsilon_n$ n vecteurs d'un $\mathbb{K}$-ev F. Il existe une application linéaire $f : E \to F$ et une seule telle que $f(e_i) = \varepsilon_i$ pour $1 \leqslant i \leqslant n$.

preuve
Tout vecteur x de E s'écrit de manière unique sous la forme $x = x_1e_1 + \cdots + x_ne_n$ où les x_i sont dans $\mathbb{K}$. Si f existe, on a nécessairement $f(x) = x_1f(e_1) + \cdots + x_nf(e_n) = x_1\varepsilon_1 + \cdots + x_n\varepsilon_n$ ce qui prouve l'unicité de f. Réciproquement, si on pose, pour $x = x_1e_1 + \cdots + x_ne_n$, $f(x) = x_1\varepsilon_1 + \cdots + x_n\varepsilon_n$, on vérifie facilement que l'on définit ainsi une application (car les x_i ne dépendent que de x), linéaire, qui vérifie $f(e_i) = \varepsilon_i$ pour tout i.

COROLLAIRE 5.5.1
Si deux applications linéaires f et g de E de dimension finie dans F vérifient $f(e_i) = g(e_i)$ pour tout vecteur e_i d'une base $\mathcal{B}$ de E, elles sont égales.

PROPOSITION 5.5.2
Soient E et F deux $\mathbb{K}$-ev, E de dimension finie, f une application linéaire de E dans F et $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E.
1) *Le système $(f(e_1), \ldots, f(e_n))$ est générateur du sous espace $\mathrm{im}(f)$ de F.*
2) *f est injective ssi le système $(f(e_1), \ldots, f(e_n))$ est libre. Dans ce cas c'est une base de $\mathrm{im}(f)$.*

preuve
1) Soit $y \in \mathrm{im}(f)$. Il existe $x \in E$ tel que $y = f(x)$. En écrivant $x = x_1e_1 + \cdots + x_ne_n$, il vient $y = x_1f(e_1) + \cdots + x_nf(e_n)$.

2) Supposons f injective. Soit $(x_1, \ldots, x_n) \in \mathbb{K}^n$ tels que $x_1 f(e_1) + \cdots + x_n f(e_n) = 0$. Ceci s'écrit $f(x_1 e_1 + \cdots + x_n e_n) = 0$ soit $x_1 e_1 + \cdots + x_n e_n \in \ker(f)$. Mais f étant injective, $\ker(f) = \{0\}$ donc $x_1 e_1 + \cdots + x_n e_n = 0$ d'où, puisque $(e_1, \ldots, e_n)$ est libre, $x_1 = 0, \ldots, x_n = 0$. Donc $(f(e_1), \ldots, f(e_n))$ est libre.
Réciproquement, supposons $f(e_1), \ldots, f(e_n))$ libre. Soit $x = x_1 + e_1 + \cdots + x_n e_n \in \ker(f)$. On a $0 = f(x) = x_1 f(e_1) + \cdots + x_n f(e_n)$, donc puisque $(f(e_1), \ldots, f(e_n))$ est libre, $x_1 = 0, \ldots, x_n = 0$ soit $x = 0$. Donc $\ker(f) = \{0\}$ et f est injective.

COROLLAIRE 5.5.2
Soit f une application linéaire d'un $\mathbb{K}$-espace vectoriel de dimension finie E dans un $\mathbb{K}$-espace vectoriel quelconque F. Le sous espace $\mathrm{im}(f)$ est de dimension finie et $\dim(\mathrm{im}(f)) \leqslant \dim(E)$.

On appelle rang de l'application linéaire f, et on note $\mathrm{rg}(f)$ la dimension de l'espace image $\mathrm{im}(f)$. Il résulte de ce qui précède que si $\mathcal{B} = (e_1, \ldots, e_n)$ est une base de E, le rang de f est égal au rang du système $(f(e_1), \ldots, f(e_n))$.

COROLLAIRE 5.5.3
Soient E et F deux $\mathbb{K}$-espaces vectoriels quelconques, $f \in L(E, F)$, $S = (u_1, \ldots, u_p)$ un système de p vecteurs de E et $f(S) = (f(u_1), \ldots f(u_p))$. On a $\mathrm{rg}(f(S)) \leqslant rg(S)$.

En effet, il suffit d'appliquer le corollaire précédent à l'espace vectoriel $E' = \mathrm{Vect}(S)$ et à la restriction f' de f à E'. On a $\mathrm{im}(f') = \mathrm{Vect}(f(S))$. D'où le résultat.

5.5.2 Isomorphismes

THEOREME 5.5.1
Soient E et F deux $\mathbb{K}$ espaces vectoriels, E étant de dimension finie $n \geqslant 1$ et $f \in L(E, F)$. Les propriétés suivantes sont équivalentes :
(1) *f est un isomorphisme de E sur F.*
(2) *L'image par f d'une base quelconque de E est une base de F.*
(3) *Il existe une base de E dont l'image par f est une base de F.*
(4) *f est injective et $\dim(F) = \dim(E)$*
(5) *$\dim(F) = \dim(E)$ et il existe $g \in L(F, E)$ telle que $f \circ g = Id_F$.*
(6) *f est surjective et $\dim(F) = \dim(E)$*
(7) *$\dim(F) = \dim(E)$ et il existe $g \in L(F, E)$ telle que $g \circ f = Id_E$*

preuve
Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. On pose $f(\mathcal{B}) = (\varepsilon_1, \ldots, \varepsilon_n)$ où $\varepsilon_i = f(e_i)$ pour $1 \leqslant i \leqslant n$.
(1) $\Rightarrow$ (2) $f(\mathcal{B})$ est générateur de $\mathrm{im}(f) = F$. Comme f est injective, ce système est libre (prop. 5.5.2). C'est donc une base de F.
(2) $\Rightarrow$ (3) Car la dimension de E étant $\geqslant 1$, E possède des bases.
(3) $\Rightarrow$ (4) Si $\mathcal{B}$ est une telle base, le système $f(\mathcal{B})$ est libre, donc f est injective. L'égalité des dimensions est triviale.
(4) $\Rightarrow$ (5) Si f est injective, le système $f(\mathcal{B})$ est libre ; formé de $\dim(F)$ vecteurs, c'est une base de F. Il existe donc une unique application linéaire $g : F \to E$ telle que $g(\varepsilon_i) = e_i$ pour $1 \leqslant i \leqslant n$. Cette application vérifie $f \circ g(\varepsilon_i) = f(e_i) = \varepsilon_i$ pour tout i, donc $f \circ g = id_F$ (corollaire 5.5.1).
(5) $\Rightarrow$ (6) Pour tout $z \in F$ on a $z = f(g(z))$ donc f est surjective.
(6) $\Rightarrow$ (7) $f(\mathcal{B})$ est générateur de $\mathrm{im}(f) = F$ puisque f est surjective. Formé de $\dim(F)$ vecteurs, c'est une base de F. On peut donc définir $g \in L(F, E)$ par $g(\varepsilon_i) = e_i$ pour tout i. On a de suite $g \circ f(e_i) = e_i$ pour tout i donc $g \circ f = id_E$.
(7) $\Rightarrow$ (1) Si $x \in \ker(f)$ on a $x = g(f(x)) = g(0) = 0$ donc f est injective. Le système $f(\mathcal{B})$ est donc libre, formé de $\dim(F)$ vecteurs, donc est une base de F. C'est donc un système générateur de F, donc f est surjective. f est linéaire et bijective, c'est donc un isomorphisme.

Remarque
Il est facile de voir que l'application g qui intervient tant dans le (5) que dans le (7) est la bijection réciproque de f

COROLLAIRE 5.5.4
Soient E et F deux espaces vectoriels de dimension finie sur le même corps $\mathbb{K}$. Il existe un isomorphisme $f : E \to F$ si et seulement si $\dim(E) = \dim(F)$.

En particulier, tout $\mathbb{K}$ espace vectoriel de dimension n est isomorphe à $\mathbb{K}^n$. Précisons ceci. Soit $\mathcal{B}_0 = \left(e_1^0, \ldots, e_n^0\right)$ la base canonique de $\mathbb{K}^n$.

Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. Tout $x \in E$ s'écrit de manière unique $x = \sum_{i=1}^{i=n} x_i e_i$. L'application $\Phi_{\mathcal{B}} : E \to \mathbb{K}^n$ qui à x associe $(x_1, \ldots, x_n)$ est un isomorphisme de E sur $\mathbb{K}^n$ qui envoie e_i sur e_i^0. Réciproquement, soit $F : E \to \mathbb{K}^n$ un isomorphisme. Posons $e_i = F^{-1}(e_i^0)$. Le système $(e_1, \ldots, e_n)$ est l'image par l'isomorphisme F^{-1} de la base canonique de $\mathbb{K}^n$. C'est donc une base $\mathcal{B}$ de E et on vérifie facilement que $F = \Phi_{\mathcal{B}}$ et que $\mathcal{B}$ est la seule base vérifiant cette propriété.
On a donc une correspondance bijective $\mathcal{B} \to \Phi_{\mathcal{B}}$ entre l'ensemble des bases de E et l'ensemble des isomorphismes de E sur $\mathbb{K}^n$.

5.5.3 Théorème du rang

THEOREME 5.5.2
Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, F un $\mathbb{K}$-espace vectoriel et $f \in L(E, F)$ une application linéaire de E dans F. On a

$$\boxed{\dim(\ker(f)) + \operatorname{rg}(f) = \dim(E)}$$

preuve
Puisque E est de dimension finie, le noyau $\ker(f)$ de f est de dimension finie et admet au moins un supplémentaire S dans E. On sait (théorème 4.3.7) que f induit un isomorphisme de S sur $\operatorname{im}(f)$. On a donc $\operatorname{rg}(f) = \dim(\operatorname{im}(f)) = \dim(S) = \dim(E) - \dim(\ker(f))$.

Exemple d'utilisation :
Soient E, F et G trois $\mathbb{K}$-espaces vectoriels, E et F étant de dimension finie, $f \in L(E, F)$ et $g \in L(F, G)$. Montrons que

$$\dim(\ker(g \circ f)) \leqslant \dim(\ker(f)) + \dim(\ker(g))$$

On a $x \in \ker(g \circ f) \Leftrightarrow f(x) \in \ker(g)$ donc $\ker(g \circ f) = f^{-1}(\ker(g))$. Notons $E' = \ker(g \circ f) = f^{-1}(\ker(g))$ et f' la restriction de f à E'. On a $f'(x) = 0 \Leftrightarrow f(x) = 0$ et $x \in E'$ donc $\ker(f') = \ker(f) \cap E' = \ker(f)$ car $\ker(f) \subset E'$. D'autre part, $\operatorname{im}(f') = f\left(f^{-1}(\ker(g)\right) \subset \ker(g)$ donc $\dim(\operatorname{im}(f')) \leqslant \dim(\ker(g))$. Le théorème du rang appliqué à f' donne $\dim(\ker(g \circ f)) = \dim(\ker(f')) + \dim(\operatorname{im}(f')) \leqslant \dim(\ker(f)) + \dim(\ker(g))$.

5.5.4 Expression analytique

Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions finies $\geqslant 1$ $\mathcal{B} = (e_1, \ldots, e_p)$ et $\beta = (\varepsilon_1, \ldots, \varepsilon_n)$ des bases de E et F respectivement. Une application linéaire $f \in L(E, F)$ est entiérement caractérisée par la donnée des images $f(e_j),\ 1 \leqslant j \leqslant p$ des vecteurs de la base $\mathcal{B}$. Notons $(a_{i,j})_{1 \leqslant i \leqslant n}$ les coordonnées de $f(e_j)$ dans la base β. On a donc $f(e_j) = \sum_{i=1}^{i=n} a_{i,j}\varepsilon_i$.

Soit $x = \sum_{j=1}^{j=p} x_j e_j \in E,\ y = \sum_{i=1}^{i=n} y_i \varepsilon_i \in F$. On a $f(x) = \sum_{j=1}^{j=p} x_j f(e_j) = \sum_{j=1}^{j=p} x_j \left(\sum_{i=1}^{i=n} a_{i,j}\varepsilon_i\right) = \sum_{i=1}^{i=n} \left(\sum_{j=1}^{j=p} a_{i,j}x_j\right) \varepsilon_i$ de sorte que $y = f(x)$ si et seulement si

$$(\mathcal{R}) \qquad y_i = \sum_{j=1}^{j=p} a_{i,j}x_j \qquad \text{pour } 1 \leqslant i \leqslant n$$

Réciproquement, étant donnée np scalaires $a_{i,j},\ 1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant p$ l'application qui à $x = \sum_{j=1}^{j=p} x_j e_j$ associe $y = \sum_{i=1}^{i=n} y_i \varepsilon_i$ où les y_i sont donnés par les relations $(\mathcal{R})$ est une application linéaire de E dans F telle que $f(e_j) = \sum_{i=1}^{i=n} a_{i,j}\varepsilon_i$. Les relations $(\mathcal{R})$ constituent l'expression analytique de f.

Notons $\varphi_{i,j}$ l'unique application linéaire de E dans F telle que $\varphi_{i,j}(e_k) = 0$ pour $k \neq j$ et $\varphi_{i,j}(e_j) = \varepsilon_i$. Soit f l'application déterminée par les relations $(\mathcal{R})$. On voit facilement que $f = \sum_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}} a_{i,j}\varphi_{i,j}$. Ceci prouve que le système $(\varphi_{i,j})$ est générateur de $L(E, F)$. Soit $(b_{i,j})$ np éléments de $\mathbb{K}$ tels que $\sum_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}} b_{i,j}\varphi_{i,j} = 0$. En appliquant au vecteur e_k pour un k fixé entre 1 et

n, il vient $\sum_{i=1}^{i=n} b_{i,k}\varepsilon_i = 0$, donc, les (ε_i) formant une base $b_{i,k} = 0$ pour tout i, k. Le système $(\varphi_{i,j})$ est donc libre. C'est par conséquent une base de $L(E, F)$.

THEOREME 5.5.3
Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions finies. On a $\dim(L(E, F)) = \dim(E) \cdot \dim(F)$.

Ce qui précède prouve le théorème si E et F sont de dimension $\geqslant 1$. Si l'un des deux est de dimension nulle, c'est à dire est réduit à l'espace $\{0\}$, on a aussi $L(E, F) = \{0\}$ et la relation demeure valable.

Remarque
La base $(\varphi_{i,j})$ dépend du choix des bases $\mathcal{B}$ et β. D'autre part, on notera qu'elle est naturellement indexée par l'ensemble $\mathbb{N}_n^* \times \mathbb{N}_p^*$ et non pas par $\mathbb{N}_{np}^*$. Pour l'indexer par $\mathbb{N}_{np}^*$ il faut choisir une numérotation de $\mathbb{N}_n^* \times \mathbb{N}_p^*$, par exemple, celle fournie par l'ordre lexicographique.

6

Matrices

I Calcul matriciel

$\mathbb{K}$ désigne un corps commutatif de caractéristique 0. $\delta_{i,j}$ est le symbole de Kronecker égal à 1 si $i = j$ et à 0 sinon.

6.1 Espace $M_{n,p}(\mathbb{K})$

DEFINITION 6.1.1
Soient n et p deux entiers naturels supérieurs ou égaux à 1. Une matrice à n lignes et p colonnes à coefficients dans $\mathbb{K}$ est une application $A : \mathbb{N}_n^* \times \mathbb{N}_p^* \to \mathbb{K}$. On notera le plus souvent, pour $(i,j) \in \mathbb{N}_n^* \times \mathbb{N}_p^*$, $a_{i,j} = A(i,j)$ et $A = (a_{i,j})$ en sous entendant $1 \leqslant i \leqslant n$ et $1 \leqslant j \leqslant p$.
On dira souvent matrice (n,p) pour matrice à n lignes et p colonnes.

On représente usuellement une matrice sous forme d'un tableau rectangulaire, à n lignes et p colonnes de nombres :

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,p} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,p} \end{pmatrix}$$

L'ensemble des matrices à n lignes et p colonnes à coefficients dans $\mathbb{K}$ est noté $M_{n,p}(\mathbb{K})$. C'est donc l'ensemble des applications de $\mathbb{N}_n^* \times \mathbb{N}_p^* \to \mathbb{K}$. Il est naturellement muni d'une structure de $\mathbb{K}$-espace vectoriel pour les opérations usuelles sur les applications. Autrement dit, si $A = (a_{i,j})$, $B = (b_{i,j})$ et $\lambda \in \mathbb{K}$ les matrices $C = A + B$ et $D = \lambda A$ sont définies respectivement par $c_{i,j} = a_{i,j} + b_{i,j}$ et $d_{i,j} = \lambda a_{i,j}$.

6.1.1 Base canonique

Notons, pour $1 \leqslant i \leqslant n$ et $1 \leqslant j \leqslant p$, par $E_{i,j}$ la matrice définie par $E_{i,j}(r,s) = \delta_{i,r}\delta_{j,s}$. Autrement dit, tous les coefficients de $E_{i,j}$ sont nuls sauf celui situé à l'intersection de la i-ième ligne et de la j-ième colonne qui vaut 1. Il est immédiat que le système $(E_{i,j})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}}$ est une base de $M_{n,p}(\mathbb{K})$ appelée base canonique. On a $A = (a_{i,j}) = \sum_{i,j} a_{i,j} E_{i,j}$. On en déduit en particulier que

$$\dim\left(M_{n,p}(\mathbb{K})\right) = np$$

On appelle matrice colonne (resp. matrice ligne) une matrice $(n,1)$ (resp. une matrice $(1,p)$. On appelle matrice carrée d'ordre n toute matrice (n,n). Une matrice colonne $(n,1)$ (resp. une matrice ligne $(1,p)$) s'identifie naturellement à un élément de $\mathbb{K}^n$ (resp. de $\mathbb{K}^p$). Une matrice $(1,1)$ s'identifie à un élément de $\mathbb{K}$. Ces identifications sont des isomorphismes d'espaces vectoriels $M_{n,1} \to \mathbb{K}^n$ et $M_{1,p} \to \mathbb{K}^p$.

6.1.2 Transposition

Lignes et colonnes d'une matrice
Soit $A = (a_{i,j}) \in M_{n,p}(\mathbb{K})$.
Pour $1 \leqslant i \leqslant n$ on appelle i-ième ligne de A et on note $L_i(A)$ le vecteur de $\mathbb{K}^p$ (ou la matrice ligne dans $M_{1,p}$) donnée par

$$L_i(A) = (a_{i,1}, a_{i,2}, \cdots, a_{i,p})$$

De même, pour $1 \leqslant j \leqslant p$ on appelle j-ième colonne de A et on note $C_j(A)$ le vecteur de $\mathbb{K}^n$ (ou la matrice colonne dans $M_{n,1}(\mathbb{K})$) donnée par

$$C_j(A) = \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{n,j} \end{pmatrix}$$

DEFINITION 6.1.2
Soit $A = (a_{i,j}) \in M_{n,p}(\mathbb{K})$. On appelle tansposée de A et on note ${}^tA$ la matrice (p,n) définie par ${}^tA(j,i) = A(i,j)$ pour $1 \leqslant j \leqslant p$ et $1 \leqslant i \leqslant n$.

THEOREME 6.1.1
La transposition est un isomorphisme de $M_{n,p}(\mathbb{K})$ sur $M_{p,n}(\mathbb{K})$ et, pour toute $A \in M_{n,p}(\mathbb{K})$ on a ${}^t({}^tA) = A$.
On a aussi $L_j({}^tA) = C_j(A)$ et $C_i({}^tA) = L_i(A)$ pour $1 \leqslant i \leqslant n$ et $1 \leqslant j \leqslant p$.

6.2 Produit matriciel

DEFINITION 6.2.1
Soit $A = (a_{i,j}) \in M_{n,p}(\mathbb{K})$ et $B = (b_{j,k}) \in M_{p,q}(\mathbb{K})$.
*Le produit **dans cet ordre** de A par B est la matrice $AB = C = (c_{i,k}) \in M_{n,q}(\mathbb{K})$ définie par*

$$\forall\, (i,k) \in \mathbb{N}_n^* \times \mathbb{N}_q^* \quad c_{i,k} = \sum_{j=1}^{j=p} a_{i,j} b_{j,k}$$

On notera que le produit AB est défini ssi le nombre de colonnes de A est égal au nombre de lignes de B. Le produit AB peut donc être défini sans que le produit BA ne le soit.

Attention! On peut avoir $AB = 0$ avec $A \neq 0$ et $B \neq 0$. Par exemple :
$A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ et $BA = \begin{pmatrix} 2 & -2 \\ 2 & -2 \end{pmatrix}$

THEOREME 6.2.1
1) *Soient $A, A' \in M_{n,p}(\mathbb{K})$, $B, B' \in M_{p,q}(\mathbb{K})$ et $\lambda \in \mathbb{K}$ où n, p, q sont des entiers $\geqslant 1$. On a*

$$\begin{gathered} (A + A')B = AB + A'B \\ A(B + B') = AB + AB' \\ (\lambda A)B = \lambda(AB) \\ A(\lambda B) = \lambda AB \end{gathered}$$

Autrement dit, l'application $M_{n,p}(\mathbb{K}) \times M_{p,q}(\mathbb{K}) \to M_{n,q}(K)$ définie par le produit est $\mathbb{K}$-bilinéaire.
2) *La transposée du produit est égal au produit des transposées dans l'ordre opposé :*

$$ {}^t(AB) = \left({}^tB\right)\left({}^tA\right)$$

3) *Le produit matriciel est associatif, i.e. pour toute matrice $C \in M_{q,r}(\mathbb{K})$ on a*

$$(AB)C = A(BC)$$

6.3 Algèbre $M_n(\mathbb{K})$

Soit n un entier naturel, $n \geqslant 1$. On note $M_n(\mathbb{K})$ l'espace $M_{n,n}(\mathbb{K})$ des matrices carrées d'ordre n à coefficients dans $\mathbb{K}$.
On note I_n la matrice définie par $I_n(i,j) = \delta_{i,j}$. On a donc $I_n(i,i) = 1$ et $I_n(i,j) = 0$ si $i \neq j$.

$M_n(\mathbb{K})$ est un $\mathbb{K}$-espace vectoriel de dimension n^2. D'autre part, le produit est une opération interne, associative, distributive à droite et à gauche par rapport à l'addition. On vérifie facilement que I_n est neutre (à gauche et à droite) pour le produit. Enfin, le produit est une application, bilinéaire de $M_n(\mathbb{K}) \times M_n(\mathbb{K})$ dans $M_n(\mathbb{K})$.

Donc $M_n(\mathbb{K})$ est une $\mathbb{K}$-algèbre associative, unitaire. Elle n'est pas commutative en général.

DEFINITION 6.3.1
On appelle groupe linéaire d'ordre n et on note $GL(n,\mathbb{K})$ le groupe des éléments inversibles de l'anneau $M_n(\mathbb{K})$.

Donc $M \in GL(n,\mathbb{K})$ ssi $\exists M' \in M_n(\mathbb{K})\ ;\ MM' = M'M = I_n$. La matrice M' est unique et notée M^{-1} et appelée matrice inverse de M.

Si $P, Q \in GL(n,\mathbb{K})$ on a $PQ \in GL(n,\mathbb{K})$ et $(PQ)^{-1} = Q^{-1}P^{-1}$.

PROPOSITION 6.3.1
Soit $P \in GL(n,\mathbb{K})$. Alors ${}^tP \in GL(n,\mathbb{K})$ et $({}^tP)^{-1} = {}^t(P^{-1})$.

preuve
La matrice $Q = {}^t(P^{-1})$ est bien définie puisque P est inversible. On a $({}^tP)Q = {}^tP\,{}^t(P^{-1}) = {}^t(P^{-1}P) = {}^tI_n = I_n$ et de même $Q({}^tP) = I_n$.

6.3.1 Sous-algèbre $\mathbb{K}$[A]

Pour ces définitions, on suppose connue la notion de polynôme. Soit $A \in M_n(\mathbb{K})$ fixée. Comme dans tout anneau, on définit les puissances de A par récurrence en posant $A^0 = I_n$ et pour tout $n \geqslant 0, A^{n+1} = A^nA$. On vérifie facilement que $\forall p, q \in \mathbb{N},\ A^pA^q = A^{p+q}$ ce qui implique que A^p et A^q commutent.
Ensuite, on définit une application $\Phi_A : \mathbb{K}[X] \to M_n(\mathbb{K})$ qui à tout polynôme $P = c_0 + c_1X + \cdots + c_qX^q$ associe la matrice $c_0I_n + c_1A + \cdots + c_qA^q$ notée $P(A)$.
Cette application vérifie

$$\begin{aligned} \forall P, Q \in \mathbb{K}[X], &\quad \Phi_A(P+Q) = \Phi_A(P) + \Phi_A(Q) \\ \forall P \in \mathbb{K}[X],\ \forall \lambda \in \mathbb{K}, &\quad \Phi_A(\lambda P) = \lambda \Phi_A(P) \\ \forall P, Q \in \mathbb{K}[X], &\quad \Phi_A(PQ) = \Phi_A(P)\Phi_A(Q) \end{aligned}$$

Autrement dit, Φ_A est un morphisme d'algèbres. On note $\mathbb{K}[A]$ l'image de Φ_A. Donc

$$\mathbb{K}[A] = \{M \in M_n(\mathbb{K}) \mid \exists P \in \mathbb{K}[X]\ ;\ M = \Phi_A(P) = P(A)\}$$

$\mathbb{K}[A]$ est la plus petite sous algèbre de $M_n(\mathbb{K})$ contenant A.
Attention! Si A et B sont deux matrices carrées, on ne peut appliquer la formule du binôme pour calculer $(A+B)^n$ que si A et B commutent.

6.3.2 Sous espaces remarquables

Matrices symétriques, antisymétriques

Une matrice $A = (a_{i,j})$ est dite symétrique (resp. antisymétrique) si elle est égale (resp. si elle est opposée) à sa transposée. Donc A est symétrique ssi $a_{i,j} = a_{j,i}$ pour tout couple (i,j) et A est antisymétrique ssi $a_{i,j} = -a_{j,i}$ pour tout couple (i,j) ; en particulier, si A est antisymétrique, on a $a_{i,i} = 0$ pour tout i. (*Ici on utilise le fait que $\mathbb{K}$ n'est pas de caractéristique 2*)
Notons $\mathcal{S}_n(\mathbb{K})$, resp. $\mathcal{A}_n(\mathbb{K})$ l'ensemble des matrices symétriques (resp. antisymétriques). La transposition ^{t} est un endomorphisme involutif de $M_n(\mathbb{K})$. Il en résulte que $M_n(\mathbb{K}) = \ker({}^t - id) \oplus \ker({}^t + id) = \mathcal{S}_n(\mathbb{K}) \oplus \mathcal{A}_n(\mathbb{K})$.
Les matrices $(E_{i,j} + E_{j,i}),\ 1 \leqslant i < j \leqslant n$ et $E_{i,i},\ 1 \leqslant i \leqslant n$ forment une base de $\mathcal{S}_n(\mathbb{K})$. Les matrices $(E_{i,j} - E_{j,i}),\ 1 \leqslant i < j \leqslant n$ forment une base de $\mathcal{A}_n(\mathbb{K})$. On a donc

$$\dim(\mathcal{S}_n(\mathbb{K})) = \frac{n(n+1)}{2} \quad \text{et} \quad \dim(\mathcal{A}_n(\mathbb{K})) = \frac{n(n-1)}{2}$$

Matrices diagonales, matrices triangulaires

Une matrice $A = (a_{i,j})$ carrée d'ordre n est dite diagonale si $a_{i,j} = 0$ pour $i \neq j$. On notera $\mathcal{D}_n(\mathbb{K})$ l'ensemble des matrices diagonales carrées d'ordre n. C'est une sous algèbre de $M_n(\mathbb{K})$ de dimension n, dont une base est le système $(E_{i,i})_{1 \leqslant i \leqslant n}$. On notera $\mathrm{diag}(a_1, \ldots, a_n)$ la matrice diagonale $(a_{i,j})$ telle que $a_{i,i} = a_i$.

Une matrice carrée d'ordre n est dite triangulaire supérieure (resp. triangulaire inférieure) si $a_{i,j} = 0$ pour tout couple (i, j) tel que $j < i$ (resp. tel que $j > i$). Elle est dite triangulaire (supérieure, inférieure) stricte si elle est traingulaire (supérieure, inférieure) et si tous les éléments de la diagonale sont nuls.

L'ensemble $\mathcal{T}_n(\mathbb{K})$ des matrices triangulaires supérieures est une sous algèbre de $M_n(\mathbb{K})$ de dimension $\dfrac{n(n+1)}{2}$. Il en est de même de l'ensemble des matrices triangulaires inférieures.
L'ensemble des matrices triangulaires supérieures strictes (resp. des matrices triangulaires inférieures strictes) est un sous espace de $M_n(\mathbb{K})$ de dimension $\dfrac{n(n-1)}{2}$.

Remarque
Une matrice A diagonale (resp. triangulaire supérieure, triangulaire inférieure) est inversible ssi $\forall i,\ a_{i,i} \neq 0$.

6.4 Noyau, image et rang d'une matrice

Convenons dans cette partie d'écrire les éléments de $\mathbb{K}^m$ comme des matrices colonnes. Soit $A = (a_{i,j}) \in M_{n,p}(\mathbb{K})$. Pour $X \in \mathbb{K}^p = M_{p,1}(\mathbb{K})$ le produit $Y = AX$ est défini et est un élément de $M_{n,1}(\mathbb{K}) = \mathbb{K}^n$. On définit ainsi une application de $\mathbb{K}^p$ dans $\mathbb{K}^n$, qui à X associe $Y = AX$. Les coordonnées y_i de Y sont données en fonction de celles x_j de X par les relations

$$(\mathcal{R}) \qquad y_i = \sum_{j=1}^{j=p} a_{i,j} x_j \qquad \text{pour } 1 \leqslant i \leqslant n$$

L'application $X \to AX$ est linéaire et toute application linéaire de $\mathbb{K}^p$ dans $\mathbb{K}^n$ est de cette forme.

DEFINITION 6.4.1
Soit $A \in M_{n,p}(\mathbb{K})$. On appelle noyau(resp. image, rang) de la matrice A le noyau, (resp. l'image, le rang) de l'application $X \to AX$.

On notera A l'application $X \to AX$.
On a donc, comme cas particulier du théorème du rang : $\dim(\ker(A)) + \mathrm{rg}(A) = p =$ nombre de colonnes de A.
La j-ième colonne $C_j(A)$ est l'image par A du j-ième vecteur de la base canonique de $\mathbb{K}^p$. Les colonnes de A engendrent donc l'image de A. Par conséquent le rang de A est égal au rang du système des vecteurs colonnes de A. Nous verrons plus loin que le rang de A est aussi égal au rang du système des vecteurs lignes de A.

II Utilisation des matrices

6.5 Matrice d'une application linéaire

DEFINITION 6.5.1 (Matrice d'une application linéaire dans des bases)
Soient E et E' deux $\mathbb{K}$-espaces vectoriels de dimensions finies p et n respectivement, $\mathcal{B} = (e_1, \ldots, e_p)$ une base de E et $\mathcal{B}' = (e'_1, \ldots, e'_n)$ une base de E'. Soit enfin $f \in L(E, E')$. On appelle matrice de f dans les bases $\mathcal{B}$ et $\mathcal{B}'$ la matrice $A = (a_{i,j}) \in M_{n,p}(\mathbb{K})$ telle que $f(e_j) = \sum_{i=1}^{i=n} a_{i,j} e'_i$. On écrit $A = \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f) = (a_{i,j})$

La j-ième colonne de M est donc formée des coordonnées dans la base $\mathcal{B}'$ de l'image par f du j-ième vecteur de base e_j.

DEFINITION 6.5.2 (Matrice d'un système de vecteurs dans une base)
Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base d'un $\mathbb{K}$-espace vectoriel E de dimension finie n et $S = (u_1, \ldots, u_p)$ un système de $p \geqslant 1$ vecteurs de E. Chacun des vecteurs u_j s'écrit de manière unique $u_j = \sum_{i=1}^{i=n} a_{i,j} e_i$. Par définition, la matrice du système S dans la base $\mathcal{B}$, notée $\mathrm{Mat}_{\mathcal{B}}(S)$ est la matrice (n,p), $A = (a_{i,j})$.

La j-ième colonne de cette matrice est donc formée des coordonnées dans la base $\mathcal{B}$ du j-ième vecteur u_j.

En particulier, si $x = \sum_{i=1}^{i=n} x_i e_i$ est un vecteur de E, sa matrice dans la base $\mathcal{B}$ est $X = \mathrm{Mat}_{\mathcal{B}}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Notons $\mathcal{B}^0 = \left(e_1^0, \ldots, e_p^0\right)$ la base canonique de $\mathbb{K}^p$. Il existe une unique application linéaire $u_S : \mathbb{K}^p \to E$ telle que $u_S(e_j^0) = u_j$ pour $1 \leqslant j \leqslant p$. Il est immédiat que $\mathrm{Mat}_{\mathcal{B}}(S) = \mathrm{Mat}_{\mathcal{B}_0,\mathcal{B}}(u_S)$. La matrice d'un système de vecteurs dans une base peut donc être considédére domme la matrice d'une application linéaire.
La réciproque est vraie, puisque la matrice de l'application linéaire f dans les bases $\mathcal{B}$ et $\mathcal{B}'$ est la matrice dans la base $\mathcal{B}'$ du système $f(\mathcal{B}) := (f(e_1), \ldots, f(e_p))$.

Il est important de retenir que les coordonnées d'un vecteur sont disposées en colonne.

THEOREME 6.5.1
Soient E et E' deux $\mathbb{K}$-espaces vectoriels de dimensions finies p et n respectivement, $\mathcal{B}$ une base de E et $\mathcal{B}'$ une base de E'. L'application $L(E,F) \to M_{n,p}(\mathbb{K}), \quad f \to \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f)$ est un isomorphisme de $L(E,F)$ sur $M(n,p)(\mathbb{K})$.

THEOREME 6.5.2
Soient E, E' et E'' trois $\mathbb{K}$-espaces vectoriels de dimensions finies p, p', p'' respectivement. Soient $\mathcal{B}$, $\mathcal{B}'$ et $\mathcal{B}''$ des bases de E, E' et E'' respectivement. Soient enfin $f \in L(E,E')$ et $g \in L(E',E'')$. On a

$$\mathrm{Mat}_{\mathcal{B},\mathcal{B}''}(g \circ f) = \mathrm{Mat}_{\mathcal{B}',\mathcal{B}''}(g)\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f)$$

preuve
Soient $\mathcal{B} = (e_1, \ldots, e_p)$, $\mathcal{B}' = \left(e_1', \ldots, e_{p'}'\right)$ et $\mathcal{B}$" $= \left(e_1'', \ldots, e_{p''}''\right)$.
Posons $A = (a_{i,j}) = \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f) \in M_{p',p}(\mathbb{K})$, $B = (b_{k,i}) = \mathrm{Mat}_{\mathcal{B}',\mathcal{B}''}(g) \in M_{p'',p'}(\mathbb{K})$ et $C = (c_{k,j}) = BA$.
On a

$$c_{k,j} = \sum_{i=1}^{i=p'} b_{k,i} a_{i,j}, \quad f(e_j) = \sum_{i=1}^{i=p'} a_{i,j} e_i' \quad \text{pour} \quad 1 \leqslant j \leqslant p \quad \text{et} \quad g(e_i') = \sum_{k=1}^{k=p''} b_{k,i} e_k'' \quad \text{pour} \quad 1 \leqslant i \leqslant p'$$

d'où

$$g \circ f(e_j) = g\left(\sum_{i=1}^{i=p'} a_{i,j} e_i'\right) = \sum_{i=1}^{i=p'} a_{i,j} g(e_i') = \sum_{i=1}^{i=p'} a_{i,j} \left(\sum_{k=1}^{k=p''} b_{k,i} e_k''\right) = \sum_{k=1}^{k=p''} \left(\sum_{i=1}^{i=p'} b_{k,i} a_{i,j}\right) e_k'' = \sum_{k=1}^{k=p''} c_{k,j} e_k'' \ \blacksquare$$

COROLLAIRE 6.5.1
Soit $\mathcal{B}$ une base d'un $\mathbb{K}$-espace vectoriel E de dimension finie n. L'application $L(E) \to M_n(\mathbb{K}), \ f \to \mathrm{Mat}_{\mathcal{B}}(f)$ est un isomorphisme de $\mathbb{K}$-algèbres.
Cet isomorphisme induit un isomorphisme du groupe $GL(E)$ des éléments inversibles de $L(E)$ sur le groupe $GL(n,\mathbb{K})$ des éléments inversibles de $M_n(\mathbb{K})$.

COROLLAIRE 6.5.2
Soit $A \in M_n(\mathbb{K})$. Les propriétés suivantes sont équivalentes :
(1) $A \in GL(n,\mathbb{K})$
(2) *A est inversible à droite (i.e. $\exists A' \in M_n(\mathbb{K}), \ AA' = I_n$)*
(3) *A est inversible à gauche (i.e. $\exists A'' \in M_n(\mathbb{K}), \ A''A = I_n$)*

Celà résulte du théorème précédent et du théorème 5.5.1.

exercice
Soit $\mathcal{A}$ une $\mathbb{K}$-algèbre associative unitaire de dimension finie et $a \in \mathcal{A}$. Montrer l'équivalence des propriétés suivantes :
1) a est un élément inversible de $\mathcal{A}$.
2) a est inversible à droite dans $\mathcal{A}$.
3) a est inversible à gauche dans $\mathcal{A}$.
4) a est régulier à droite dans $\mathcal{A}$. (i.e. $xa = ya \Rightarrow x = y$)
5) a est régulier à gauche dans $\mathcal{A}$. (i.e. $ax = ay \Rightarrow x = y$
Cet exercice permet de retrouver les résultats du corollaire ci dessus sans utiliser la correspondance entre matrices et applications linéaires.

6.6 Changements de base

THEOREME 6.6.1
Soient $\mathcal{B} = (e_1, \ldots, e_n)$ une base d'un $\mathbb{K}$-espace vectoriel E, $S = (e'_1, \ldots, e'_n)$ un système de n vecteurs de E et $P = \mathrm{Mat}_{\mathcal{B}}(S)$. Le système S est une base de E si et seulement si la matrice P est inversible.

preuve

Soit $f \in L(E)$ l'unique application linéaire telle que $f(e_i) = e'_i$ pour $1 \leqslant i \leqslant n$. La matrice de f dans la base $\mathcal{B}$ est P. Or on sait que S est une base ssi f est un isomorphisme, et que f est un isomorphisme ssi sa matrice P est inversible (corollaire 6.5.1). D'où la conclusion. Si S est une base, la matrice P s'appelle matrice de passage de la base $\mathcal{B}$ à la base S. Si besoin on la notera $P = P_{\mathcal{B},S}$.

Propriétés des matrices de changements de base

Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E. On a :

1. $P_{\mathcal{B},\mathcal{B}'} = \mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(id_E)$.
2. $(P_{\mathcal{B},\mathcal{B}'})^{-1} = P_{\mathcal{B}',\mathcal{B}}$
3. Soit $S = (u_1, \ldots, u_p)$ un système de vecteurs de E. Alors $\mathrm{Mat}_{\mathcal{B}}(S) = P_{\mathcal{B},\mathcal{B}'}\mathrm{Mat}_{\mathcal{B}'}(S)$
4. Si $\mathcal{B}''$ est une troisième base de E, $\quad P_{\mathcal{B},\mathcal{B}''} = P_{\mathcal{B},\mathcal{B}'}P_{\mathcal{B}',\mathcal{B}''}$.
5. Soit $x \in E, X = \mathrm{Mat}_{\mathcal{B}}(x)$ et $X' = \mathrm{Mat}_{\mathcal{B}'}(x)$. On a $X = PX'$

5. est un cas particulier de 3. On résume cette quelquefois cette formule en disant que la matrice P qui donne (en colonne) les nouveaux vecteurs de base en fonction des anciens permet de calculer les anciennes coordonnées d'un vecteur (la matrice X) en fonction des nouvelles.

La preuve de 3. est un calcul simple, le 4. s'en déduit. Le 1. est évident. Le 2. se déduit de 1. comme suit :
$P_{\mathcal{B}',\mathcal{B}}P_{\mathcal{B},\mathcal{B}'} = \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(id_E)\mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(id_E) = \mathrm{Mat}_{\mathcal{B}',\mathcal{B}'}(id_E) = I_n$ ce qui suffit pour prouver 2.

THEOREME 6.6.2 (Formule du changement de bases)
Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions finies p et n respectivement. Soient $\mathcal{B},\ \mathcal{B}'$ deux bases de E, $\beta,\ \beta'$ deux bases de F. Soient $P = P_{\mathcal{B},\mathcal{B}'} \in GL(p,\mathbb{K})$ la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ et $Q = P_{\beta,\beta'} \in GL(n,\mathbb{K})$ la matrice de passage de β à β'. Soient enfin $f \in L(E,F),\ A = \mathrm{Mat}_{\mathcal{B},\beta}(f)$ et $A' = \mathrm{Mat}_{\mathcal{B}',\beta'}(f)$. On a

$$A' = Q^{-1}AP$$

preuve
D'après le théorème 6.5.2 on a

$$A' = \mathrm{Mat}_{\mathcal{B}',\beta'}(f) = \mathrm{Mat}_{\beta,\beta'}(id_F)\mathrm{Mat}_{\mathcal{B},\beta}(f)\mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(id_E) = Q^{-1}AP$$

III Rang des matrices

6.7 Matrices équivalentes

DEFINITION 6.7.1
Soient $M, M' \in M_{n,p}(\mathbb{K})$. On dit que M est équivalente à M' si il existe deux matrices $Q \in GL(n,\mathbb{K})$ et $P \in GL(p,\mathbb{K})$ telles que $M' = QMP$.

On vérifie facilement que la relation M est équivalente à M' est une relation d'équivalence dans $M_{n,p}(\mathbb{K})$.

Pn peut le voir autrement en utilisant la notion d'action d'un groupe (voir plus loin le chapitre correspondant). Soit Γ le groupe produit $\Gamma = GL(n,\mathbb{K}) \times GL(p,\mathbb{K})$. Pour $(Q,P) \in \Gamma$ et $M \in M_{n,p}(\mathbb{K})$ posons $(P,Q) * M = QMP^{-1}$. Vérifions qu'on définit ainsi une action du groupe Γ sur l'ensemble $M_{n,p}(\mathbb{K})$:
1) L'élément neutre de Γ est (I_n, I_p) et $(I_n, I_p) * M = M$ pour toute $M \in M_{n,p}(\mathbb{K})$.
2) Si $(Q', P') \in \Gamma$ on a
$(P', Q') * ((P,Q) * M) = Q'(QMP^{-1})P'^{-1} = (Q'Q)M(P'P)^{-1} = (Q'Q, P'P) * M = ((Q', Q) \cdot (P', P)) * M$.
La matrice M' est équivalente à M si et seulement si il existe un couple $(Q,P) \in \Gamma$ tel que $M' = (Q,P) * M$. Autrement dit la relation M' est équivalente à M n'est autre que la relation M' appartient à l'orbite de M sous l'action de Γ. On retrouve ainsi que c'est une relation d'équivalence, dont les classes sont les orbites de cette action.

Notons, pour $1 \leqslant r \leqslant \min(n,p)$, $J_{r,n,p}$ où J_r si il n'y a pas de confusion possible, la matrice suivante

$$J_{r,n,p} = \begin{pmatrix} & & & 0\cdots & 0 \\ & I_r & & 0\cdots & 0 \\ & & & 0\cdots & 0 \\ 0 & \cdots & 0 & 0\cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & 0\cdots & 0 \end{pmatrix}$$

THEOREME 6.7.1
1) *Soient $M \in M_{n,p}(\mathbb{K})$ et $r = \mathrm{rg}(M)$. M est équivalente à la matrice $J_{r,n,p}$.*
2) *Deux matrices $M, M' \in M_{n,p}(\mathbb{K})$ sont équivalentes si et seulement si elles ont même rang.*

preuve
1) Le résultat est trivial si $r = 0$ car alors $M = 0$. Supposons $r \geqslant 1$ et donc $M \neq 0$. Soit $u_M : \mathbb{K}^p \to \mathbb{K}^n$ l'application définie par $u_M(X) = MX$ où on identifie $\mathbb{K}^p$ à $M_{p,1}(\mathbb{K})$. Soit S un supplémentaire dans $\mathbb{K}^p$ de $\ker(u)$. On sait que u_M induit un isomorphisme $\tilde{u}$ de S sur $\mathrm{im}(u)$. On a en particulier $\dim(S) = r$. Soit $(e'_1, \ldots, e'_r)$ une base de S, $(e'_{r+1}, \ldots, e'_p)$ une base de $\ker(u)$ de sorte que $\mathcal{B}' = (e'_1, \ldots, e'_p)$ est une base de $\mathbb{K}^p$. Soit pour $1 \leqslant k \leqslant r, \varepsilon_k = u_M(e_k) = \tilde{u}(e_k)$. Comme $\tilde{u}$ est un isomorphisme de S sur $\mathrm{im}(u)$, $(\varepsilon_1, \ldots, \varepsilon_r)$ est une base de $\mathrm{im}(u)$. On la complète en une base $\beta' = (\varepsilon_1, \ldots, \varepsilon_n)$ de $\mathbb{K}^n$. Par construction, la matrice dans les bases $\mathcal{B}', \beta'$ de u_M est la matrice $J_{r,n,p}$. La matrice de u_M dans les bases canoniques $\mathcal{B}$ et β de $\mathbb{K}^p$ et de $\mathbb{K}^n$ est M. Si on note P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ et Q celle de β à β' on a $J_{r,n,p} = Q^{-1}MP$. Donc M est équivalente à $J_{r,n,p}$.

2) Si M et M' ont le même rang, elles sont toutes deux équivalentes à J_r, donc équivalentes entre elles. Il reste à prouver la réciproque. Pour celà, il suffit de prouver le lemme suivant :

Lemme 6.7.1
Soit $M \in M_{n,p}(\mathbb{K})$.
a) Soit $P \in GL(p,\mathbb{K})$; on a $\mathrm{rg}(MP) = \mathrm{rg}(M)$.
b) Soit $Q \in GL(n,\mathbb{K})$; on a $\mathrm{rg}(QM) = \mathrm{rg}(M)$.

preuve du lemme
a) On a trivialement $\mathrm{im}(MP) \subset \mathrm{im}(M)$. Soit $Y \in \mathrm{im}(M)$. Il existe $X \in \mathbb{K}^p$ tel que $Y = MX$ d'où $Y = MPZ$ si $Z = P^{-1}X$. Par conséquent $Y \in \mathrm{im}(MP)$. Donc $\mathrm{im}(M) = \mathrm{im}(MP)$ et $\mathrm{rg}(M) = \mathrm{rg}(MP)$.
b) Soit $X \in \mathbb{K}^p$. Comme Q est inversible, on a $QMX = 0 \Leftrightarrow Q^{-1}QMX = 0 \Leftrightarrow MX = 0$ donc $\ker(QM) = \ker(M)$. Par le théorème du rang, il vient $\mathrm{rg}(QM) = n - \dim(\ker(QM)) = n - \dim(\ker(M)) = \mathrm{rg}(M)$.

En conclusion, si $d = \min(n,p)$, l'action du groupe $GL(p,\mathbb{K}) \times GL(n,\mathbb{K})$ sur $M_{n,p}(\mathbb{K})$ définie ci dessus a exactement $d+1$ orbites, chacune des orbites étant constituées de l'ensemble des matrices de rang r pour un r fixé entre 0 et d.

COROLLAIRE 6.7.1
Soit $A \in M_{n,p}(\mathbb{K})$. On a $\mathrm{rg}(A) = \mathrm{rg}({}^tA)$.

En effet, si $r = \mathrm{rg}(A)$, il existe P et Q inversibles telles que $A = PJ_{r,n,p}Q$. Alors ${}^tA = {}^tQJ_{r,p,n}\,{}^tP$ d'où $\mathrm{rg}({}^tA) = r$.

6.8 Détermination du rang d'une matrice au moyen des sous matrices carrées

6.8.1 Notations

Soit $M \in M_{n,p}(\mathbb{K})$. Soient $I \subset \mathbb{N}_n^*$ et $J \subset \mathbb{N}_p^*$ avec $\mathrm{card}(I) = m \geqslant 1$ et $\mathrm{card}(I) = q \geqslant 1$. Ecrivons $I = \{\alpha_1, \ldots, \alpha_m\}$ avec $\alpha_1 < \cdots < \alpha_m$ et $J = \{\beta_1, \ldots, \beta_q\}$ avec $\beta_1 < \cdots < \beta_q$. Nous noterons $M_{I,J} \in M_{m,q}(\mathbb{K})$ la matrice dont l'élément à l'intersection de la ligne s et de la colonne t est m_{α_s,β_t}. $M_{I,J}$ s'obtient en supprimant dans M les colonnes dont le numéro n'est pas dans J et les lignes dont le numéro n'est pas dans I. Les matrices $M_{I,J}$ s'appellent les sous matrices de M ou matrices extraites de M.

Supposons $\mathrm{card}(I) < n$ et $\mathrm{card}(J) < p$. On appelle sous matrice bordante de la matrice $M_{I,J}$ toute matrice $M_{I',J'} \in M_{m+1,q+1}(\mathbb{K})$ où $I' = I \cup \{\lambda\},\ J' = J \cup \{\mu\}$ avec $\lambda \in \mathbb{N}_n^* \setminus I,\ \mu \in \mathbb{N}_p^* \setminus J$. On dit que $M_{I',J'}$ est obtenue en bordant $M_{I,J}$ avec la λ-ième ligne et la μ-ième colonne de M.

Nous noterons enfin $C_j(M) \in M_{n,1}(\mathbb{K}) = \mathbb{K}^n$ la j-ième colonne de M.

THEOREME 6.8.1
Soit $M \in M_{n,p}(\mathbb{K}),\ M \neq 0$. Supposons trouvée une sous matrice de M, carrée d'ordre r et inversible n'admettant aucune matrice bordante inversible. Alors $\mathrm{rg}(M) = r$.

preuve
Soit A une sous matrice carrée d'ordre r, inversible et n'admettant aucune matrice bordante inversible.

1. Soit A' une matrice bordante de A. Le rang de M est invariant par permutation des lignes ou des colonnes. Quand on fait une telle opération la sous matrice A' se transforme en une matrice bordante de la transformée de A. On peut donc supposer que A est la sous matrice formée des r premières lignes et des r premières colonnes de M, plus précisément que $A = M_{\mathbb{N}_r^*,\mathbb{N}_r^*}$.
   Notons M_0 la sous matrice de M formée des r premières lignes de M : $M_0 = M_{\mathbb{N}_r^*,\mathbb{N}_p^*}$.

2. Montrons d'abord que $\mathrm{rg}(M) \geqslant r$.
   Soit $h : \mathbb{K}^n \to \mathbb{K}^r$ l'application linéaire $h(x_1, \ldots, x_n) = (x_1, \ldots, x_r)$. On a pour $1 \leqslant k \leqslant r,\ h\left(C_k(M)\right) = C_k(A)$.
   A inversible $\Rightarrow (C_1(A), \ldots, C_r(A))$ libre. A fortiori, on en déduit $((C_1(M), \ldots, C_r(M))$ libre, donc $\mathrm{rg}(M) \geqslant r$.

3. Si $r = p$ ou $r = n$, on a $\mathrm{rg}(M) = r$. Notons que dans ce cas, A n'admet aucune sous matrice bordante. Ceci s'applique en particulier à la sous matrice M_0 qui est donc de rang r.

4. Supposons donc $r < n,\ r < p$ et montrons que $\mathrm{rg}(M) \leqslant r$.
   Pour celà, fixons un indice $\mu > r$ et montrons que $C_\mu(M) \in \mathrm{Vect}\,(C_1(M), \ldots, C_r(M))$.

$$M = \left(\begin{array}{ccc|ccccc} & & & a_{1,r+1} & & a_{1,\mu} & \cdots & a_{1,p} \\ & A & & & \vdots & & & \vdots \\ & & & a_{r,r+1} & & a_{r,\mu} & \cdots & a_{r,p} \\ \hline a_{r+1,1} & \cdots & a_{r+1,r} & a_{r+1,r+1} & \cdots & a_{r+1,\mu} & \cdots & a_{r+1,p} \\ a_{\lambda,1} & \cdots & a_{\lambda,r} & a_{\lambda,r+1} & & a_{\lambda,\mu} & \cdots & a_{\lambda,p} \\ a_{n,1} & \cdots & a_{n,r} & a_{n,r+1} & & a_{n,\mu} & \cdots & a_{n,p} \end{array}\right) \left.\begin{array}{c}\\\\\\\\\\\end{array}\right\} M_0$$

$$A' = \left(\begin{array}{ccc|c} & & & a_{1,\mu} \\ & A & & \vdots \\ & & & a_{r,\mu} \\ \hline a_{\lambda,1} & \cdots & a_{\lambda,r} & a_{\lambda,\mu} \end{array}\right)$$

M_0 est de rang r et les r-premières colonnes de M_0 sont indépendantes donc forment une base de $\mathbb{K}^r$. Il existe donc un système unique de coefficients $a_1, \ldots, a_r$ tels que

$$C_\mu(M_0) = a_1 C_1(M_0) + \cdots + a_r C_r(M_0)$$

Soit $\lambda > r$.

Considérons la matrice $A' = A'(\lambda)$ obtenue en bordant A avec la λ-ième ligne et la μ-ième colonne de M : $A' = M_{\mathbb{N}_r^* \cup \{\lambda\}, \mathbb{N}_r^* \cup \{\mu\}}$. Par hypothèse, A' est non inversible, donc de rang $\leqslant r$. Par ailleurs, d'après la première partie, les r premières colonnes de A' sont indépendantes. Donc il existe des scalaires $\alpha_1(\lambda), \ldots, \alpha_r(\lambda)$ tels que

$$C_{r+1}(A') = \alpha_1(\lambda) C_1(A') + \cdots \alpha_r(\lambda) C_r(A')$$

Soit $H : \mathbb{K}^{r+1} \to \mathbb{K}^r$ l'application $H(x_1, \ldots, x_r, x_{r+1}) = (x_1, \ldots, x_r)$. En appliquant H à la dernière égalité on obtient

$$C_\mu(M_0) = \alpha_1(\lambda) C_1(M_0) + \cdots + \alpha_r(\lambda) C_r(M_0)$$

d'où, par unicité des a_k, $\alpha_k(\lambda) = a_k$ pour tout k entre 1 et r. Ceci étant valable pour tout $\lambda > r$, on en déduit la même égalité pour les colonnes de M :

$$C_\mu(M) = a_1 C_1(M) + \cdots + a_r C_r(M)$$

ce qui achève la preuve.

COROLLAIRE 6.8.1

Soit $M \in M_{n,p}(\mathbb{K})$ et r le maximum des entiers k tels qu'il existe une sous matrice extraite d'ordre k inversible. Alors $\mathrm{rg}(M) = r$.

7

Dualité

Soit E un $\mathbb{K}$-espace vectoriel. On rappelle que son dual E^* est l'espace des formes linéaires sur E : $E^* = L(E, \mathbb{K})$. Pour $x \in E$ et $\varphi \in E^*$ on note $\langle x, \varphi \rangle = \varphi(x)$. $\langle\ ,\ \rangle$ s'appelle le crochet de dualité. L'application qui à $(x, \varphi) \in E \times E^*$ associe $\langle x, \varphi \rangle = \varphi(x)$ est une forme bilinéaire sur l'espace produit $E \times E^*$ apellée forme bilinéaire canonique.

Dans toute la suite E est un $\mathbb{K}$-espace vectoriel de dimension finie

7.1 Base duale

Rappelons une notation. Etant donné un ensemble X non vide, le symbole de Kronecker $\delta_{x,y}$ pour $x, y \in X$ est défini par $\delta_{x,y} = 0$ si $x \neq y$ et $\delta_{x,x} = 1$ pour tout x.

PROPOSITION 7.1.1
Soient p un entier, $(u_1, \ldots, u_p)$ un système de p vecteurs de E et $(\psi_1, \ldots, \psi_p)$ un système de p éléments de E^*.
Si $\langle u_i, \psi_j \rangle = \delta_{i,j}$ pour tout (i, j) tels que $1 \leqslant i, j \leqslant p$, les deux systèmes $(u_1, \ldots, u_p)$ et $(\psi_1, \ldots, \psi_p)$ sont libres.

preuve

Soient $(\lambda_1, \ldots, \lambda_p)$ des scalaires tels que $\sum_{1 \leqslant i \leqslant p} \lambda_i u_i = \vec{0}$. Soit k quelconque entre 1 et p. On a $\left\langle \sum_{1 \leqslant i \leqslant p} \lambda_i u_i, \psi_k \right\rangle = 0$, soit compte tenu de la bilinéarité du crochet de dualité : $0 = \sum_{1 \leqslant i \leqslant p} \lambda_i \langle u_i, \psi_k \rangle = \sum_{1 \leqslant i \leqslant p} \lambda_i \delta_{i,k} = \lambda_k$. Donc le système $(u_1, \ldots, u_p)$ est libre. On fait de même pour le système $(\psi_1, \ldots, \psi_p)$.

THEOREME 7.1.1
Soit $B = (e_1, \ldots, e_n)$ une base de E. Il existe une unique base $B^* = (\varepsilon_1, \ldots, \varepsilon_n)$ de E^* telle que pour tous i, j entre 1 et n on ait $\langle e_i, \varepsilon_j \rangle = \delta_{i,j}$.

preuve
Une application linéaire est entièrement déterminée par les valeurs qu'elle prend sur les éléments d'une base de E. Par conséquent, il existe, pour j fixé, $1 \leqslant j \leqslant n$ une unique application linéaire $\varepsilon_j : E \to \mathbb{K}$ vérifiant $\varepsilon_j(e_j) = 1$ et $\varepsilon_j(e_i) = 0$ pour $i \neq j$. Ceci prouve l'existence et l'unicité du système $(\varepsilon_i)_{1 \leqslant i \leqslant n}$.

Soit $x = \sum_{1 \leqslant i \leqslant n} x_i e_i$ un élément quelconque de E.

Pour j fixé entre 1 et n, on a $\langle x, \varepsilon_j \rangle = \left\langle \sum_{1 \leqslant i \leqslant n} x_i e_i, \varepsilon_j \right\rangle = \sum_{1 \leqslant i \leqslant n} x_i \delta_{i,j} = x_j$. Soit alors $\varphi \in E^*$. Posons $\mu_i = \varphi(e_i)$. On obtient $\varphi(x) = \sum_{1 \leqslant i \leqslant n} x_i \mu_i = \sum_{1 \leqslant i \leqslant n} \mu_i \varepsilon_i(x) = \left(\sum_{1 \leqslant i \leqslant n} \mu_i \varepsilon_i \right)(x)$ et par conséquent $\varphi = \sum_{1 \leqslant i \leqslant n} \mu_i \varepsilon_i$. Le système $(\varepsilon_i)_{1 \leqslant i \leqslant n}$ est donc générateur de E^*.
Enfin, d'après la proposition 7.1.1 ce système est libre.■

On remarquera que cette démonstration redonne le fait que E et son dual ont la même dimension.

Au cours de la preuve, on a établi les relations suivantes particulièrement importantes: soient

$$x \in E,\ x = \sum_{1 \leqslant i \leqslant n} x_i e_i \quad \text{et} \quad \varphi \in E^*,\ \varphi = \sum_{1 \leqslant i \leqslant n} \mu_i \varepsilon_i$$

On a : $x_i = \langle x, \varepsilon_i \rangle$ et $\mu_i = \langle e_i, \varphi \rangle$ ou encore

$$\boxed{x = \sum_{i=1}^{i=n} \langle x, \varepsilon_i \rangle e_i}$$

et

$$\boxed{\varphi = \sum_{i=1}^{i=n} \langle e_i, \varphi \rangle \varepsilon_i}$$

On notera que la i-ième forme linéaire de la base duale de la base $B = (e_1, \ldots, e_n)$ est tout simplement l'application qui à un vecteur x associe sa i-ième coordonnée dans la base B.

Concluons par une remarque à propos des notations. On note souvent $(e_1^*, \ldots, e_n^*)$ la base duale de la base $B = (e_1, \ldots, e_n)$. Cette notation est commode mais présente un risque : pour i fixé, le i-ième vecteur de la base duale e_i^* semble ne dépendre que du vecteur e_i, ce qui n'est pas le cas ; e_i^* **dépend de l'ensemble de la base B** .

PROPOSITION 7.1.2
Soient $(\varphi_1, \ldots, \varphi_p)$ un système de p éléments de E^* et $\Phi : E \to \mathbb{K}^p$ l'application définie par $\Phi(x) = (\varphi_1(x), \ldots, \varphi_p(x))$. On a l'équivalence :

$$(\varphi_1, \ldots, \varphi_p) \text{ est un système libre } \Leftrightarrow \Phi \text{ est surjectif}$$

preuve

$$\begin{aligned}
\Phi \text{ non surjectif} &\Leftrightarrow \exists\, H \text{ hyperplan de } \mathbb{K}^p, \quad \mathrm{im}(\Phi) \subset H \\
&\Leftrightarrow \exists\ (\lambda_1, \ldots, \lambda_p) \in \mathbb{K}^p \setminus \{0\}, \quad \sum_{i=1}^{i=p} \lambda_i \varphi_i = 0 \\
&\Leftrightarrow (\varphi_1, \ldots, \varphi_p) \text{ est lié.}
\end{aligned}$$

THEOREME 7.1.2
Soit $\beta = (\varphi_1, \ldots, \varphi_n)$ une base de E^* ; il existe une base de E et une seule $B = (e_1, \ldots, e_n)$ telle que β soit la base duale de B. La base B est dite base préduale de la base β ; on dit aussi que les bases B et β sont duales l'une de l'autre.

preuve
D'après la proposition 7.1.2, l'application $\Phi : E \to \mathbb{K}^n$ définie par $\Phi(x) = (\varphi_1(x), \ldots, \varphi_n(x))$ est surjective. Si E_i est le i-ième vecteur de la base canonique de $\mathbb{K}^n$, il existe $u_i \in E$ tel que $\Phi(u_i) = E_i$. Le système $B = (u_1, \ldots, u_n)$ vérifie $\langle u_i, \varphi_j \rangle = \delta_{i,j}$ pour tous i, j entre 1 et n. D'après la proposition 7.1.1 c'est un système libre de E. Comme il est formé de n vecteurs c'est une base et β est bien la base duale de B.

THEOREME 7.1.3 (Changement de base)
Soient $B = (e_1, \ldots, e_n)$ et $B' = (e'_1, \ldots, e'_n)$ deux bases de E. Soient $B^* = (\varepsilon_1, \ldots, \varepsilon_n)$ et $B'^* = (\varepsilon'_1, \ldots, \varepsilon'_n)$ les bases duales rspectives. Soit $P = \mathrm{Mat}_B(B')$ la matrice de passage de la base B à la base B'. La matrice de passage de la base B^* à la base B'^* est $\mathrm{Mat}_{B^*}(B'^*) = {}^tP^{-1}$.

preuve
Posons $Q = \mathrm{Mat}_{B^*}(B'^*) = (b_{i,j})$ et $P = (a_{i,j})$. On a donc par définition

$$e'_j = \sum_{i=1}^{i=n} a_{i,j} e_i \quad \text{et} \quad \varepsilon'_k = \sum_{p=1}^{p=n} b_{p,k} \varepsilon_p$$

D'où

$$\delta_{j,k} = \langle e'_j, \varepsilon'_k \rangle = \sum_{1 \leqslant i,p \leqslant n} a_{i,j} b_{p,k} \langle e_i, \varepsilon_p \rangle = \sum_{1 \leqslant i,p \leqslant n} a_{i,j} b_{p,k} \delta_{i,p} = \sum_{i=1}^{i=n} a_{i,j} b_{i,k}$$

relation qui s'écrit ${}^tPQ = I_n$. ■

Remarque Ce calcul permet de fournir une autre démonstration du théorème 7.1.2. En effet, soient B une base de E, B^* la base duale, β une base de E^* et $Q = \mathrm{Mat}_{B^*}(\beta)$. Soit $P = {}^tQ^{-1}$. C'est une matrice inversible ; soit B' la base de E telle que $\mathrm{Mat}_B(B') = P$. On vérifie que β est la base duale de B'.

EXEMPLE 7.1.1
On considère les trois formes linéaires définies sur $E = \mathbb{R}^3$ par

$$\begin{aligned} X(x,y,z) &= x - 2y + z \\ Y(x,y,z) &= -x + y - z \\ Z(x,y,z) &= x - z \end{aligned}$$

Montrer qu'elles forment une base du dual de $\mathbb{R}^3$ et préciser la base préduale.

On peut écrire

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \text{avec} \quad A = \begin{pmatrix} 1 & -2 & 1 \\ -1 & 1 & -1 \\ 1 & 0 & -1 \end{pmatrix}$$

Ce système s'inverse facilement :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = C \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad \text{avec} \quad C = A^{-1} = -\frac{1}{2} \begin{pmatrix} 1 & 2 & -1 \\ 2 & 2 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

Soit $B = (i, j, k)$ la base canonique de $\mathbb{R}^3$ et $B^* = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ sa base duale.
On a donc $X = \varepsilon_1 - 2\varepsilon_2 + \varepsilon_3,\ Y = -\varepsilon_1 + \varepsilon_2 - \varepsilon_3\ Z = \varepsilon_1 - \varepsilon_3$. La matrice Q du système (X, Y, Z) dans la base B^* est donc **la transposée** de A. La matrice de passage de la base canonique à la base préduale de la base (X, Y, Z) est donc $P = {}^t({}^tA)^{-1} = A^{-1} = C$. Celle ci est donc la base

$$\left(\left(-\frac{1}{2}, -1, -\frac{1}{2} \right),\ (-1, -1, -1),\ \left(\frac{1}{2}, 0, -\frac{1}{2} \right) \right)$$

7.2 Orthogonalité

DEFINITION 7.2.1

- *Soit $A \subset E$. L'orthogonal de A noté A^o est l'ensemble des formes linéaires qui s'annulent sur A :*

$$A^o = \{\varphi \in E^* \mid \forall\, x \in A,\ \langle x, \varphi \rangle = 0\} = \{\varphi \in E^* \mid A \subset \ker(\varphi)\}$$

- *Soit $A' \subset E^*$. L'orthogonal de A', noté ${}^oA'$ est l'ensemble des vecteurs de E annulés par tous les éléments de A' :*

$${}^oA' = \{x \in E \mid \forall\, \varphi \in A',\ \langle x, \varphi \rangle = 0\} = \bigcap_{\varphi \in A'} \ker(\varphi)$$

La vérification des propriétés suivantes est élémentaire.

- Pour tout $A \subset E$, A^o est un sous espace vectoriel de E^* et $A^o = (\mathrm{Vect}(A))^o$
- Si $A_1 \subset A_2 \subset E$ on a $A_2^o \subset A_1^o$
- $E^o = \{0\}$ et $\{\vec{0}\}^o = E^*$
- Si $A' \subset E^*$, ${}^oA'$ est un sous espace vectoriel de E et ${}^oA' = {}^o(\mathrm{Vect}(A'))$
- Si $A'_1 \subset A'_2 \subset E^*$ on a ${}^oA'_2 \subset {}^oA'_1$
- ${}^o(E^*) = \{\vec{0}\}$ et ${}^o\{0\} = E$

THEOREME 7.2.1

1. *Soit F un sous espace vectoriel de E. On a* $\dim F + \dim F^o = \dim E$.
2. *Soit F' un sous espace vectoriel de E^*. On a* $\dim F' + \dim {}^oF' = \dim E$.

preuve

1. Si $F = \{\vec{0}\}$ on a $F^o = E^*$ d'où la conclusion. Sinon, soit $(e_1, \ldots, e_p)$ une base de F ; on la complète en une base $(e_1, \ldots, e_n)$ de E. Soit $(\varepsilon_1, \ldots, \varepsilon_n)$ la base duale. Soit $\varphi = \sum_{1 \leqslant i \leqslant n} \mu_i \varepsilon_i \in E^*$.
   On a $\varphi \in F^o \Leftrightarrow \forall i \ 1 \leqslant i \leqslant p, \ \langle e_i, \varphi \rangle = 0$. Or $\langle e_i, \varphi \rangle = \mu_i$ donc $\varphi \in F^o \Leftrightarrow \mu_1 = \cdots = \mu_p = 0$. Donc $F^o = \text{Vect}(\varepsilon_{p+1}, \ldots, \varepsilon_n)$ d'où le résultat.
2. Soit $F' \in E^*$. Si $F' = \{0\}$ on a ${}^oF' = E$ d'où la conclusion. Sinon, soit $(\varphi_1, \ldots, \varphi_p)$ une base de F'. L'application $\Phi : E \to \mathbb{K}^p, \ \Phi(x) = (\varphi_1(x), \ldots, \varphi_p(x))$ est sujective (proposition 7.1.2). Donc son noyau, qui n'est autre que l'orthogonal de F' est de dimension $n - p$ (théorème du rang). ■.

COROLLAIRE 7.2.1
Soit F un sous espace vectoriel de E. On a ${}^o(F^o) = F$.
Soit F' un sous espace vectoriel de E^*. On a $({}^oF')^o = F'$.

preuve
Pour chacune des deux affirmations, on a une inclusion évidente entre sous espaces et l'égalité des dimensions.

THEOREME 7.2.2
Soient F, G deux sous espaces vectoriels de E, F', G' deux sous espaces vectoriels de E^*. On a

(1) $(F + G)^o = F^o \cap G^o$.

(2) $(F \cap G)^o = F^o + G^o$.

(3) ${}^o(F' + G') = {}^oF' \cap {}^oG'$.

(4) ${}^o(F' \cap G') = {}^oF' + {}^oG'$.

preuve

(1) $F \subset F + G \Rightarrow (F + G)^o \subset F^o$. De même, $(F + G)^0 \subset G^o$ donc $(F + G)^o \subset F^o \cap G^o$.
Si $\varphi \in F^o \cap G^o$, on a $F \subset \ker(\varphi)$ et $G \subset \ker(\varphi)$ donc $(F + G) \subset \ker(\varphi)$ soit $\varphi \in (F + G)^0$. D'où l'inclusion $F^0 \cap G^o \subset (F + G)^o$ et l'égalité.

(3) La preuve est analogue : on montre comme ci dessus que ${}^o(F' + G') \subset {}^oF' \cap {}^oG'$. Ensuite, soit $x \in {}^oF' \cap {}^oG'$. Soit $\eta \in F' + G'$. Il existe $\varphi \in F'$ et $\psi \in G'$ tels que $\eta = \varphi + \psi$. On en déduit $\langle x, \eta \rangle = \langle x, \varphi \rangle + \langle x, \psi \rangle = 0$ donc $x \in {}^o(F' + G')$.

(2) On applique (3) avec $F' = F^o$ et $G' = G^o$. Il vient, en utilisant le corollaire 7.2.1, ${}^o(F^o + G^o) = F \cap G$ d'où le résultat en reprenant les orthogonaux.

(4) On applique (1) avec $F = {}^oF'$ et $G = {}^oG'$. Il vient $({}^oF' + {}^oG')^o = ({}^oF')^o \cap ({}^oG')^o = F' \cap G'$ d'après le corollaire 7.2.1. En utilisant encore ce corollaire, on en déduit (4).

Application : équations d'un sous espace
Soient E un $\mathbb{K}$-espace vectoriel de dimension n et F un sous espace de dimension p de E. Un système d'équations de F est un système de formes linéaires sur E, $(f_1, \ldots, f_m)$ tel qu'on ait l'équivalence, pour $x \in E$,

$$x \in F \Leftrightarrow \begin{cases} f_1(x) = 0 \\ \cdots \\ f_m(x) = 0 \end{cases}$$

Ce qui revient à dire $F = \bigcap_{1 \leqslant i \leqslant m} \ker(f_i) = {}^o\text{Vect}\,(f_1, \ldots, f_m)$ ou encore $\text{Vect}\,(f_1, \ldots, f_m) = F^o$. Donc un système d'équations de F est un système de formes linéaires générateur de l'orthogonal de F.
En particulier, si $(\varphi_1, \ldots, \varphi_{n-p})$ est une base de F^o, $(\varphi_1, \ldots, \varphi_{n-p})$ est un système minimal d'équations de F.

Inversement, soit $(g_1, \ldots, g_m)$ un système de $m \geqslant 1$ formes linéaires sur E et $G = \ker(g_1) \cap \ldots \cap \ker(g_m)$. On a donc $G = {}^oG'$ en posant $G' = \text{Vect}\,(g_1, \ldots, g_m)$. La dimension de G' est égal au rang r du système $(g_1, \ldots, g_m)$ et donc $\dim G = n - r$.
En conclusion, un sous espace de E de dimension p admet des systèmes d'équations. Le nombre minimal de formes linéaires d'un tel système est égal à $n-p$. Les $n-p$ formes linéaires composant un tel système sont indépendantes. Si S est un système d'équations de F formé de m formes linéaires, le rang de S est $n-p$. Ce résultat généralise, en dimension finie, ce qui a été vu dans le premier paragraphe sur les hyperplans et les formes linéaires.

Exemple: Soit D une droite dans l'espace vectoriel $E = \mathbb{R}^3$. L'orthogonal de D est de dimension 2. Un système minimal d'équations de D est donc formé de deux formes linéaires indépendantes. Soit (f, g) un tel système. On a donc, pour $x \in \mathbb{R}^3$, $x \in D \Leftrightarrow (f(x) = 0 \text{ et } g(x) = 0)$. Géométriquement, celà revient à définir D comme interesction de deux plans.
Soit d'autre part P un plan de $\mathbb{R}^3$ et $h \in E^*$ une équation de P de sorte que P^o est la droite de base h.
On a $D \subset P \Leftrightarrow P^o \subset D^o \Leftrightarrow h \in \text{Vect}(f, g) \Leftrightarrow \exists(\lambda, \mu) \in \mathbb{R}^2\ h = \lambda f + \mu g$. C'est la base de la théorie des faisceaux de plans.

EXEMPLE 7.2.1 (Polynômes d'interpolation de Lagrange)
Soient $a_1, \ldots, a_n$ n éléments distincts de $\mathbb{K}$ et $E = \mathbb{K}_{n-1}[X]$. On définit les applications $\varepsilon_i : E \to \mathbb{K}$ pour $1 \leqslant i \leqslant n$ par $\varepsilon_i(P) = P(a_i)$.

1. $(\varepsilon_1, \ldots, \varepsilon_n)$ est une base de E^*.
   En effet, soit $P \in {}^o\text{Vect}\,(\varepsilon_1, \ldots, \varepsilon_n)$. On a donc pour tout i entre 1 et n $\ 0 = \varepsilon_i(P) = P(a_i)$. Le polynôme P de degré au plus $n-1$ admet n racines distinctes. Il est donc nul. Donc ${}^o\text{Vect}\,(\varepsilon_1, \ldots, \varepsilon_n) = \{\vec{0}\}$. Par conséquent $\text{Vect}\,(\varepsilon_1, \ldots, \varepsilon_n) = E^*$. Le système $(\varepsilon_1, \ldots, \varepsilon_n)$ de $n = \dim E^*$ vecteurs est générateur de E^*, donc c'est une base de cet espace.

2. Détermination de la base préduale $(L_1, \ldots, L_n)$.
   On a pour $1 \leqslant k \leqslant n$ fixé $L_k(a_j) = 0$ pour $j \neq k$. Le polynôme L_k de degré au plus $n-1$ admet comme racines les $n-1$ scalaires distincts a_j, $j \neq k$; il s'écrit donc $L_k(X) = \lambda_k \prod_{\substack{1 \leqslant j \leqslant n \\ j \neq k}} (X - a_j)$. La condition $L_k(a_k) = 1$ détermine la valeur de λ_k. On obtient :
   $$\lambda_k = \frac{1}{\prod_{\substack{1 \leqslant j \leqslant n \\ j \neq k}} (a_k - a_j)} \quad \text{et donc} \quad L_k(X) = \frac{\prod_{\substack{1 \leqslant j \leqslant n \\ j \neq k}} (X - a_j)}{\prod_{\substack{1 \leqslant j \leqslant n \\ j \neq k}} (a_k - a_j)}$$
   La base préduale de $(\varepsilon_1, \ldots, \varepsilon_n)$ est donc $B = (L_1, \ldots, L_n)$. Ces polynômes sont apelés polynômes de Lagrange (relativement au système de points $(a_1, \ldots, a_n)$).

3. Si $(b_1, \ldots, b_n)$ est un élément quelconque de $\mathbb{K}^n$, il existe un et un seul polynôme P de degré au plus $n-1$ prenant la valeur b_i au point a_i pour $1 \leqslant i \leqslant n$.
   En effet, on écrit P dans la base B : $P = \sum_{1 \leqslant k \leqslant n} \mu_k L_k$. La condition $P(a_i) = b_i$ s'écrit $\mu_i = b_i$. Le polynôme cherché est $P = \sum_{1 \leqslant k \leqslant n} b_k L_k$.

4. Tout polynôme $P \in E$ s'écrit
   $$P = \sum_{i=1}^{i=n} \langle P, \varepsilon_i \rangle L_i = \sum_{i=1}^{i=n} P(a_i) L_i$$
   En particulier, pour $0 \leqslant j \leqslant n-1$ on a
   $$X^j = \sum_{i=1}^{i=n} a_i^j L_i(X)$$
   Notons alors $B_0 = (1, X, \ldots, X^{n-1})$ la base canonique de E. Les relations précédentes fournissent la matrice de passage de la base $B = (L_1, \ldots, L_n)$ à la base canonique. On a
   $$Q = \text{Mat}_B(B_0) = \begin{pmatrix} 1 & a_1 & \ldots & a_1^{n-1} \\ 1 & a_2 & \ldots & a_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & a_n & \ldots & a_n^{n-1} \end{pmatrix}$$

On reconnait la matrice transposée de la matrice de Van der Monde $V(a_1, \ldots, a_n)$. On retrouve ainsi que cette matrice est inversible si les a_i sont tous distincts. On a en prime une méthode pour inverser cette matrice : il suffit (!) de développer les $L_k(X)$ suivant les puissances croissantes de X : si $L_k(X) = \sum_{0 \leqslant j \leqslant n-1} b_{j,k} X^j$, on a $Q^{-1} = (m_{j,k})$ avec $m_{j,k} = b_{j-1,k} \quad 1 \leqslant j, k \leqslant n$.

7.3 Transposition

DEFINITION 7.3.1
Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimension finie et $f \in L(E,F)$. On appelle transposée de f et on note ${}^t f$ l'application de F^* dans E^* qui à toute $\varphi \in F^*$ associe ${}^t f(\varphi) = \varphi \circ f \in E^*$.

On a donc la caractérisation suivante de la transposition :

$$\forall\, x \in E,\ \forall\, \varphi \in F^*, \quad \langle x,\, {}^t f(\varphi) \rangle \,=\, \langle f(x), \varphi \rangle \,=\, \varphi\,(f(x))$$

THEOREME 7.3.1
Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimension finie.

1. *L'application $f \to {}^t f$ est un isomorphisme de $L(E,F)$ sur $L(F^*, E^*)$.*
2. *${}^t Id_E = Id_{E^*}$*
3. *Soient G un troisième espace vectoriel, $f \in L(E,F)$, $g \in L(F,G)$; on a ${}^t(g \circ f) = {}^t f \circ {}^t g$.*
4. *Si f est un isomorphisme de E sur F ${}^t f$ est un isomorphisme de F^* sur E^* et $({}^t f)^{-1} = {}^t\left(f^{-1}\right)$.*

preuve
Les assertions 1) 2), et 3) sont faciles. Montrons la dernière. Soit f un isomorphisme de E sur F. (Ces deux espaces ont donc même dimension). L'application f^{-1} est donc définie, ainsi que sa transposée, soit g. En utilisant 3) et 2), il vient $g \circ {}^t f = {}^t\left(f^{-1}\right) \circ {}^t f = {}^t\left(f \circ f^{-1}\right) = {}^t Id_F = Id_{F^*}$ et de même ${}^t f \circ g = Id_{E^*}$. Donc ${}^t f$ est un isomorphisme d'inverse g.

THEOREME 7.3.2
Soient E, F deux $\mathbb{K}$-espaces vectoriels de dimension finie et $f \in L(E,F)$.

1. $\ker({}^t f) = (\mathrm{im}(f))^o$
2. *f et ${}^t f$ ont même rang.*
3. $\mathrm{im}({}^t f) = (\ker(f))^o$

preuve

1. $\varphi \in \ker({}^t f) \Leftrightarrow \forall\, x \in E \ \ \varphi\,(f(x)) = 0 \Leftrightarrow \mathrm{im}(f) \subset \ker(\varphi) \Leftrightarrow \varphi \in \mathrm{im}(f)^o$
2. Résulte de 1) car $\mathrm{rg}({}^t f) = \dim(\mathrm{im}({}^t f)) = \dim E^* - \dim\left(\ker({}^t f)\right) = \dim E - \dim\left(\mathrm{im}(f)^o\right) = \dim(\mathrm{im}(f)) = \mathrm{rg}(f)$
3. On a
   - $\mathrm{im}({}^t f) \subset (\ker(f))^o$. En effet, soit $\psi \in \mathrm{im}({}^t f)$. Il existe $\varphi \in F^*$ tel que $\psi = \varphi \circ f$. Donc, pour tout $x \in \ker(f),\ \psi(x) = 0$ soit $\psi \in (\ker(f))^o$.
   - $\dim\left(\mathrm{im}({}^t f)\right) = \mathrm{rg}({}^t f) = \mathrm{rg}(f) = n - \dim(\ker(f)) = \dim\left(\ker(f)\right)^o$

   d'où la conclusion.

THEOREME 7.3.3
Soient B et B' des bases des $\mathbb{K}$-espaces vectoriels de dimension finies E et F respectivement. Soient β et β' les bases duales de E^* et F^* respectivement. Soit enfin $f \in L(E,F)$ de matrice A dans les bases (B, B') : $A = \mathrm{Mat}_{B,B'}(f)$. Alors, la matrice dans les bases β' et β de ${}^t f$ est la transposée de A : $\mathrm{Mat}_{\beta',\beta}({}^t f) = {}^t A$.

preuve
Soient $n = \dim E$ et $p = \dim F$. Notons $B = (e_1, \ldots, e_n)$, $B' = \left(e'_1, \ldots, e'_p\right)$, $\beta = (\varepsilon_1, \ldots, \varepsilon_n)$ et $\beta' = \left(\varepsilon'_1, \ldots, \varepsilon'_p\right)$.
Soit $A = (a_{i,j}) \in M_{p,n}(\mathbb{K})$. Posons $C = (c_{k,m}) = \mathrm{Mat}_{\beta',\beta}({}^t f) \in M_{n,p}(\mathbb{K})$.
Par définition de la matrice C, on a

$$\sum_{k=1}^{k=n} c_{k,m}\varepsilon_k = {}^t f(\varepsilon'_m) = \varepsilon'_m \circ f$$

donc

$$c_{k,m} = \left\langle e_k, \sum_{s=1}^{s=n} c_{s,m}\varepsilon_s \right\rangle = \left\langle e_k, {}^t f(\varepsilon'_m) \right\rangle = \langle f(e_k), \varepsilon'_m \rangle = \left\langle \sum_{i=1}^{i=n} a_{i,k} e'_i, \varepsilon'_m \right\rangle = a_{m,k}$$

donc $C = {}^t A$.

COROLLAIRE 7.3.1
Le rang d'une matrice est égale au rang de sa transpoése : pour toute $A \in M_{n,p}(\mathbb{K})$, on a $\mathrm{rg}(A) = \mathrm{rg}({}^t A)$.

On retrouve ainsi, par une méthode complètement différente, une propriété déjà établie.

7.4 Complément: bidual

NB: Cette notion ne figure plus au programme.

Par définition, l'espace bidual du $\mathbb{K}$-espace vectoriel E est le dual $(E^*)^*$ de son dual. On le note usuellement E^{**}. Pour $e \in E$ notons j_e l'application de E^* dans $\mathbb{K}$ définie par $j_e(\varphi) = \varphi(e) = \langle e, \varphi \rangle$. C'est évidemment une application linéaire de E^* dans $\mathbb{K}$ (par définition des opérations dans E^*) donc un élément du bidual E^{**}.

THEOREME 7.4.1
Soit E un $\mathbb{K}$-espace vectoriel de dimension finie.L'application $j : e \to j_e$ de E dans E^{**} est un isomorphisme de E sur son bidual.

preuve
La linéarité de j est évidente. Soit $e \in \ker(j)$. On a donc, $\forall\, \varphi \in E^*,\ j_e(\varphi) = 0 = \langle e, \varphi \rangle$. Donc $\ker(j) = E^o = \{\vec{0}\}$. E étant de dimension finie, il en est de même de son dual et de son bidual, et $\dim(E^{**}) = \dim(E)$ donc l'application linéaire injective j est un isomorphisme.

L'isomorphisme j ne dépend d'aucun choix. On l'appelle isomorphisme canonique de E sur son bidual. On peut identifier ces deux espaces au moyen de cet isomorphisme. Dans ce cas les notions de crochet de dualité, d'ortogonalité, deviennent symétriques.
Le lecteur vérifiera par exemple que si F' est un sous espace vectoriel de E^* son orthogonal F'^o dans E^{**} n'est autre que $j\,({}^o F')$. De même, la relation ${}^o(F^o) = F$ pour un sous espace F de E s'écrit plus simplement $(F^o)^o = j(F)$. Si on identifie un espace et son bidual, cette relation s'écrit plus simplement $(F^o)^o = F$. Enfin, si $f \in L(E, F)$, on a ${}^t f \in L(F^*, E^*)$ donc ${}^t\,({}^t f) \in L(E^{**}, F^{**})$; si on identifie un espace et son bidual on obtient ${}^t\,({}^t f) = f$.

8

Arithmétique

Si a et b sont deux éléments d'un anneau commutatif A, on notera $a|b$ la relation a divise b, i.e. $\exists\, c \in A \;\; b = ac$.

8.1 Idéaux de $\mathbb{Z}$ *(Rappels)*

THEOREME 8.1.1 (division euclidienne)
Pour tout couple $(a, b) \in \mathbb{Z}^2$ avec $b \neq 0$ il existe un unique couple $(q, r) \in \mathbb{Z}^2$ tel que $a = bq + r$ et $0 \leqslant r < |b|$.
q (resp. r) s'appelle le quotient (resp. le reste) de la division euclidienne de a par b.

THEOREME 8.1.2
Soit H un sous ensemble de $\mathbb{Z}$. Les propriétés suivantes sont équivalentes.
(1) *H est un sous groupe de $(\mathbb{Z}, +)$.*
(2) *H est un idéal de l'anneau $\mathbb{Z}$.*
(3) *Il existe $m \in \mathbb{Z}$ tel que $H = m\mathbb{Z}$.*

COROLLAIRE 8.1.1
Tout idéal de $\mathbb{Z}$ est principal et tout idéal non nul admet deux générateurs opposés.

COROLLAIRE 8.1.2
L'application $m \to m\mathbb{Z}$ est une bijection de $\mathbb{N}$ sur l'ensemble des idéaux de $\mathbb{Z}$.

Généralisation : anneau euclidien
Soit A un anneau intégre. On appelle stathme euclidien sur A toute fonction $\varphi : A \setminus \{0\} \to \mathbb{N}$ vérifiant les deux propriétés suivantes :
(a) $\forall x \in A \setminus \{0\}, \forall y \in A \setminus \{0\}, x|y \Rightarrow \varphi(x) \leqslant \varphi(y)$
(b) $\forall a \in A, \forall b \in A \setminus \{0\}, \exists (q, r) \in A^2$ tel que $a = bq + r$ et $((r = 0)$ ou $(\varphi(r) < \varphi(b))$

Un anneau euclidien est un couple (A, φ) où A est un anneau commutatif intégre et φ un stathme euclidien sur A.
On démontre alors facilement que tout anneau euclidien est principal.
La valeur absolue est un stathme euclidien sur $\mathbb{Z}$. L'anneau $\mathbb{K}[X]$ des polynômes à coefficients dans un corps commutatif $\mathbb{K}$ est un exemple d'anneau euclidien, le stathme euclidien étant donné par $\varphi =$ degré.

8.2 Congruences

Pour $n \in \mathbb{Z} \setminus \{0\}$ notons provisoirement $\mathcal{R}_n$ la relation d'équivalence associée à l'idéal $n\mathbb{Z}$. Elle est donc définie par $x\mathcal{R}_n y \Leftrightarrow x - y \in n\mathbb{Z} \Leftrightarrow n$ divise $x - y$. Notons que $\mathcal{R}_n = \mathcal{R}_{-n}$. D'autre part, la relation $\mathcal{R}_0$ est la relation d'égalité et la relation $\mathcal{R}_1$ est la relation triviale (pour laquelle il y a une seule classe d'équivalence). La relation $\mathcal{R}_n$ s'appelle la congruence modulo n. On

note $x \equiv y \bmod n$ pour $x\mathcal{R}_n y$. Donc $x \equiv y \bmod n \Leftrightarrow n|(x-y)$. Nous noterons π_n la surjection canonique $\pi_n : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$. Pour $x \in \mathbb{Z}$ on notera $\pi_n(x) = \bar{x}_n = \bar{x}$ si il n'y a pas d'ambiguité.

La compatibilité avec les opérations implique les relations suivantes :

$$\left\{ \begin{array}{rcll} x & \equiv & y & \bmod n \\ x' & \equiv & y' & \bmod n \end{array} \right. \Rightarrow \left\{ \begin{array}{rcll} x+x' & \equiv & y+y' & \bmod n \\ xx' & \equiv & yy' & \bmod n \end{array} \right.$$

Les opérations dans $\mathbb{Z}/n\mathbb{Z}$ sont définies comme suit :
Soient $\alpha, \beta \in \mathbb{Z}/n\mathbb{Z}$. Choisissons $x \in \alpha$ et $y \in \beta$. On a $\alpha = x + n\mathbb{Z},\ \beta = y + n\mathbb{Z}$. La classe modulo n de $x+y$ (resp. de xy) ne dépend pas des choix de x et de y et $\alpha + \beta = \pi_n(x+y) = x + y + n\mathbb{Z},\ \alpha\beta = \pi_n(xy) = xy + n\mathbb{Z}$. On a donc

$$\overline{x+y} = \overline{x} + \overline{y} \quad \text{et} \quad \overline{x} \cdot \overline{y} = \overline{xy}$$

Notons aussi les propriétés suivantes des congruences :
Pour tout entier naturel $p \geqslant 2,\ x \equiv y \bmod n \Rightarrow x^p \equiv y^p \bmod n$.
Pour tout entier relatif $a,\ x \equiv y \bmod n \Rightarrow ax \equiv ay \bmod an$.

Notons que $\mathbb{Z}/0\mathbb{Z}$ est canoniquement isomorphe à $\mathbb{Z}$ et que $\mathbb{Z}/1\mathbb{Z}$ est l'anneau nul. Nous écarterons ces deux cas dans la suite.

Soit n un entier naturel, $n \geqslant 2$. Si $x \in \mathbb{Z}$ la division euclidienne de x par n fournit un unique entier r vérifiant $0 \leqslant r < n$ et $x \equiv r \bmod n$. Ce nombre r s'appelle le reste modulo n de x. Deux entiers relatifs x et x' sont congrus modulo n ssi ils ont le même reste. Soit $\alpha \in \mathbb{Z}/n\mathbb{Z}$. D'après ce qui précède il existe un élément r et un seul vérifiant $0 \leqslant r < n$ et $\overline{r} = \alpha$. On dit aussi que r est le reste modulo n de la classe α. La restriction de la surjection canonique π_n à l'intervalle entier $[0 \cdot n - 1]$ est une bijection de cet ensemble sur $\mathbb{Z}/n\mathbb{Z}$. On en déduit en particulier que card $(\mathbb{Z}/n\mathbb{Z}) = n$.

$\mathbb{Z}/n\mathbb{Z}$ étant un groupe on dispose d'une loi externe (c.f. 2.1.6) $\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ notée $(m, \overline{a}) \to m \cdot \overline{a}$. On vérifie facilement que $\forall m, a \in \mathbb{Z},\ m \cdot \overline{a} = \overline{ma}$.

8.3 Pgcd, ppcm

8.3.1 Divisibilité

Soient $a, b \in \mathbb{Z}$. On a $a|b \Leftrightarrow b \in a\mathbb{Z} \Leftrightarrow b\mathbb{Z} \subset a\mathbb{Z}$. La relation d'inclusion est une relation d'ordre partiel sur l'ensemble des idéaux de $\mathbb{Z}$; par contre la relation $a|b$ dans $\mathbb{Z}$ est une relation de préordre partiel (on peut avoir $a|b$ et $b|a$ avec $a \neq b$, en prenant $b = -a$). En général, deux éléments quelconques ne sont pas comparables. Notons que la restriction de cette relation à l'ensemble $\mathbb{N}^*$ des entiers naturels non nuls est une relation d'ordre et que $a|b \Rightarrow a \leqslant b$. Ceci n'est plus vrai dans $\mathbb{Z}$ (2 divise -6) ni même dans $\mathbb{N}$ (2 divise 0).

8.3.2 pgcd

DEFINITION 8.3.1
Soient $n \geqslant 1$ un entier naturel fixé, $n \geqslant 1$ et $a_1, \ldots, a_n$ n entiers relatifs. On appelle pgcd des a_i l'unique générateur positif de l'idéal $a_1\mathbb{Z} + \cdots + a_n\mathbb{Z}$.

Justifions cette définition. L'ensemble $\mathcal{D}(a_1, \ldots, a_n)$ des entiers diviseurs communs aux a_j est un sous ensemble non vide de $\mathbb{Z}$ (car il contient 1) ; on a $x \in \mathcal{D}(a_1, \ldots, a_n) \Leftrightarrow \forall j \in [1 \cdot n],\ a_j\mathbb{Z} \subset x\mathbb{Z} \Leftrightarrow (a_1\mathbb{Z} + \cdots + a_n\mathbb{Z}) \subset x\mathbb{Z}$.
$a_1\mathbb{Z} + \cdots + a_n\mathbb{Z}$ est un idéal de $\mathbb{Z}$. Il existe donc un unique entier naturel d tel que $a_1\mathbb{Z} + \cdots + a_n\mathbb{Z} = d\mathbb{Z}$. On a donc $x \in \mathcal{D}(a_1, \ldots, a_n) \Leftrightarrow d\mathbb{Z} \subset x\mathbb{Z}$ ce qui équivaut à x divise d. Donc $\mathcal{D}(a_1, \ldots, a_n)$ est l'ensemble des diviseurs de d. Pour la relation d'ordre $|$ l'ensemble $\mathcal{D}(a_1, \ldots, a_n) \cap \mathbb{N}^*$ admet un plus grand élément d. Si les a_j ne sont pas tous nuls, et le pgcd est le plus grand diviseur commun des a_j (pour la relation d'ordre usuelle).

On notera dans la suite $a \wedge b$ le pgcd des entiers a et b. Remarquons que $a \wedge 0 = |a|$ pour tout $a \in \mathbb{Z}$.

PROPOSITION 8.3.1
1) *L'application $(a, b) \to a \wedge b$ est une loi de composition interne, associative et commutative sur $\mathbb{Z}$.*
2) *Pour tous $a_1, \ldots, a_n$ on a pgcd$(a_1, \ldots, a_n) =$ pgcd$(a_1 \ldots a_{n-1}) \wedge a_n$. En particulier pgcd$(a, b, c) = (a \wedge b) \wedge c$ ce qui peut aussi s'écrire vu l'associativité $a \wedge b \wedge c$.*
3) *Pour tous $a_1, \ldots, a_n, c \in \mathbb{Z}$ on a pgcd$(ca_1, \ldots, ca_n) = |c|$pgcd$(a_1, \ldots, a_n)$*
4) *Soient $a \in \mathbb{Z}, b \in \mathbb{Z} \setminus \{0\}$ et $(q, r) \in \mathbb{Z}^2$ le quotient et le reste de la division euclidienne de a par b. On a $a \wedge b = b \wedge r$.*

Algorithme d'Euclide

Soient a et b deux entiers relatifs non nuls. Si l'un des deux est nul, le pgcd des deux est égal à la valeur absolue de l'autre. On remarque que $a \wedge b = |a| \wedge |b|$. On peut donc supposer $a > 0, b > 0$ et $a \geqslant b$. Dans ces conditions on définit une suite $(r_j)_{j \geqslant 0}$ comme suit. On pose $r_0 = a,\ r_1 = b$. Ensuite, supposons que les r_j soient définis pour $0 \leqslant j \leqslant k-1$ et et non nuls pour $0 \leqslant j \leqslant k-2$. Si $r_{k-1} = 0$ on pose $r_j = 0$ pour $j \geqslant k$. La suite est alors entièrement construite. Sinon, on définit r_k comme le reste de la division euclidienne de r_{k-2} par r_{k-1}. Par construction, les r_k sont des entiers naturels, et si $r_{k-1} \neq 0$, on a $r_k < r_{k-1}$. La suite d'entiers naturels (r_i) est décroissante. Elle ne peut décroitre strictement indéfiniment. Autrement dit, l'ensemble des indices i tels que $r_i = 0$ n'est pas vide. Soit s le plus petit entier tel que $r_s = 0$. Compte tenu de la proposition précédente, on a $a \wedge b = r_0 \wedge r_1 = r_1 \wedge r_2 = \cdots = r_{s-2} \wedge r_{s-1} = r_{s-1} \wedge r_s = r_{s-1}$.
Si on applique l'algorithme d'Euclide en partant de a et b et non de $|a|$ et $|b|$ on obtient encore le même résultat, la suite (r_k) étant décroissante à partir du rang 2.

Autrement dit le pgcd de a et b est le dernier reste non nul dans la suite des divisions successives.

8.3.3 ppcm

DEFINITION 8.3.2
Soient $n \in \mathbb{N}^*$ et $a_1, \ldots, a_n$ n entiers relatifs. On appelle ppcm des a_i l'unique générateur positif de l'idéal $\bigcap_{1 \leqslant i \leqslant n} a_i\mathbb{Z}$.

Justifions cette définition. Notons $\mathcal{M}(a_1, \ldots, a_n)$ l'ensemble des entiers relatifs qui sont multiples de tous les a_i. On a $x \in \mathcal{M}(a_1, \ldots, a_n) \Leftrightarrow \forall i, a_i | x \Leftrightarrow \forall i\ x \in a_i\mathbb{Z}$. Donc $\mathcal{M}(a_1, \ldots, a_n) = \bigcap_{1 \leqslant i \leqslant n} a_i\mathbb{Z}$. Cet idéal a un unique générateur dans $\mathbb{N}$, soit m. Les multiples communs à tous les a_j sont donc les multiples de m. Si l'un des a_i est nul, $\mathcal{M} = \{0\}$ et $m = 0$. Sinon,si on se limite aux multiples positifs, m est aussi le plus petit multiple commun pour l'ordre naturel.

Si $a, b \in \mathbb{Z}$ on notera $a \vee b$ leur ppcm.

PROPOSITION 8.3.2
1) *L'application $a, b \to a \vee b$ est une loi interne sur $\mathbb{Z}$, associative et commutative.*
2) *Pour tous $a_1, \ldots, a_n$ on a ppcm$(a_1, \ldots, a_n) =$ ppcm$(a_1 \ldots a_{n-1}) \vee a_n$. En particulier ppcm$(a, b, c) = (a \vee b) \vee c$ ce qui peut aussi s'écrire vu l'associativité $a \vee b \vee c$.*
3) *Pour tous $a_1, \ldots, a_n, c \in \mathbb{Z}$ on a ppcm$(ca_1, \ldots, ca_n) = |c|$ppcm$(a_1, \ldots, a_n)$*

8.3.4 Entiers premiers entre eux

DEFINITION 8.3.3
Les n entiers relatifs $a_1, \ldots a_n$ sont dits premiers entre eux (dans leur ensemble) si leur pgcd est égal à 1.

THEOREME 8.3.1
Soient $a_1, \ldots, a_n$ n entiers relatifs. $a_1, \ldots, a_n$ sont premiers entre eux dans leur ensemble ssi il existe des entiers relatifs $\lambda_1, \ldots, \lambda_n$ tels que $\lambda_1 a_1 + \cdots + \lambda_n a_n = 1$.

En particulier deux entiers a et b sont premiers entre eux ssi il existe des entiers u et v tels que $au + bv = 1$ (Théorème de Bézout). Un tel couple (u, v) n'est pas unique. On peut en obtenir un en "remontant" les égalités successives fournies par l'algorithme d'Euclide. (Voir le paragraphe 8.9.2)

DEFINITION 8.3.4
Les entiers $a_1, \ldots, a_n$ sont dits premiers entre eux deux à deux si pour tout couple (i, j) d'entiers entre 1 et n tel que $i \neq j$, les entiers a_i et a_j sont premiers entre eux.

Si les entiers $a_1, \ldots, a_n$ sont premiers entre eux deux à deux ils sont premiers entre eux dans leur ensemble.

THEOREME 8.3.2 (Theoreme de Gauss)
Soient a, b, c trois entiers. Si a divise bc et est premier avec b alors a divise c.

preuve
Soient $u, v \in \mathbb{Z}$ tels que $au + bv = 1$. a divise acu et bcv donc aussi $acu + bcv = c$. ■

THEOREME 8.3.3
Soient $a, b_1, \ldots, b_n$ dans $\mathbb{Z}$. Si a est premier avec b_k pour tout k entre 1 et n, a est premier avec le produit $b = b_1 \cdots b_n$.

preuve
Pour chaque k entre 1 et n il existe des entiers u_k et v_k vérifiant $au_k + b_k v_k = 1$. En multipliant ces n égalités et en développant on obtient une égalité de la forme $a\alpha + (b_1 \cdots b_n)\beta = 1$ où α et $\beta = v_1 \ldots v_n$ sont dans $\mathbb{Z}$. ■

THEOREME 8.3.4
Si $a_1, \ldots, a_n$ sont n entiers relatifs premiers entre eux deux à deux, on a ppcm$(a_1, \ldots, a_n) = |a_1 \cdots a_n|$.

preuve
a) Cas $n = 2$. Soient a et b premiers entre eux et μ leur ppcm. Il existe des entiers r et s tels que $\mu = ar = bs$. a divise bs et est premier avec b donc (théorème de Gauss) a divise s. Soit $k \in \mathbb{Z}$ tel que $s = ak$. Alors $\mu = bs = abk$ donc ab divise μ. Mais ab est un multiple commun à a et b donc est un multiple de μ. Par conséquent, $ab = \pm\mu$. D'où la conclusion.

b) Supposons le résultat établi pour un entier $n \geqslant 2$. Soient $a_1, \ldots a_n, a_{n+1}$ premiers entre eux deux à deux. On a ppcm$(a_1, \ldots, a_{n+1})$ = ppcm (ppcm$(a_1, \ldots, a_n), a_{n+1})$. Les entiers $a_1, \ldots, a_n$ sont premiers entre eux deux à deux, donc d'après l'hypothèse de récurrence, ppcm$(a_1, \ldots, a_n) = |a_1 \cdots a_n|$. D'après le théorème 8.3.3, a_{n+1} est premier avec $a_1 \cdots a_n$, donc d'après la propriété déjà prouvée pour $n = 2$, ppcm$(|a_1 \cdots a_n|, a_{n+1}) = |a_1 \cdots a_n a_{n+1}|$.
Ceci achève la démonstration par récurrence sur n. ■

Attention! Il en résulte que si $a_1, \ldots, a_n$ sont des entiers deux à deux premiers entre eux et si x est un multiple de chacun des a_k, alors x est multiple du produit $a_1 \cdots a_n$. Cette propriété tombe en défaut si les a_k sont seulement premiers entre eux dans leur ensemble comme le montre l'exemple $a_1 = 6,\ a_2 = 10,\ a_3 = 15$ et $x = 30$.

THEOREME 8.3.5
Soient $a_1, \ldots, a_n$ des entiers relatifs et d un diviseur commun aux a_i. Alors

$$pgcd(a_1, \ldots, a_n) = d \Leftrightarrow pgcd\left(\frac{a_1}{d}, \ldots, \frac{a_n}{d}\right) = 1$$

Celà résulte du 3°) de la proposition 8.3.1.

THEOREME 8.3.6
Soient a et b deux entiers relatifs. On a pgcd$(a, b) \cdot$ ppcm$(a, b) = |ab|$.

preuve

Soient $d = a \wedge b$ et $m = a \vee b$. Il existe des entiers a' et b' tels que $a = da',\ b = db'$ et $a' \wedge b' = 1$. On en déduit $a' \vee b' = |a'b'|$ donc $a \vee b = d|a'b'|$ d'où $dm = d^2|a'b'| = |ab|$.

8.4 Nombres premiers

DEFINITION 8.4.1
Un entier naturel p est dit premier si $p \geqslant 2$ et si les seuls diviseurs dans $\mathbb{N}$ de p sont 1 et p.
On notera $\mathcal{P}$ l'ensemble des nombres premiers.

Eléments irréductibles
Soient A un anneau commutatif, A^* le groupe des unités (éléments inversibles) de A. Deux éléments x et y de A sont dits associés si il existe $u \in A^*$ tel que $y = ux$ ce qui équivaut à $x = u^{-1}y$. Un élément x de A est dit irréductible si ce n'est pas une unité et si les seuls diviseurs de x sont les éléments asssociés à x et les unités de A.
Il en résulte que les éléments irréductibles de $\mathbb{Z}$ sont les entiers $\pm p$, p premier.
Si $\mathbb{K}$ est un corps commutatif, les éléments irréductibles de $\mathbb{K}[X]$ sont appelés polynômes irréductibles. Si $\mathbb{K} = \mathbb{C}$ ce sont les

polynômes de degré 1, si $\mathbb{K} = \mathbb{R}$ ce sont les polynômes de degré 1 et les polynômes de degré deux à discriminant strictement négatif.

THEOREME 8.4.1
Soient p un nombre premier et x un entier relatif. Si p ne divise pas x, p et x sont premiers entre eux.

En effet le pgcd d de p et x divise p et n'est pas égal à p. Donc c'est 1.

THEOREME 8.4.2
Pour tout entier naturel $n \geqslant 2$ il existe un nombre premier p divisant n.

preuve
Soit $D_n = \{k \in \mathbb{N} \,|\, 2 \leqslant k \leqslant n \text{ et } k|n\}$. D_n est un ensemble d'entiers naturels non vide (il contient n) ; il a donc un plus petit élément, soit p. Soit x un entier naturel au moins égal à 2 et divisant p. Alors $x \leqslant p$ et $x \in D_n$ donc $x \geqslant p$ ce qui prouve que p est premier.

THEOREME 8.4.3
L'ensemble $\mathcal{P}$ des nombres premiers est infini.

preuve
Soit n un entier naturel. D'après le théorème précédent, il existe un nombre premier p divisant $n!+1$. Or les entiers $2,\ 3, \ldots, n$ divisent $n!$ donc ne divisent pas $n!+1$. On a donc $p > n$. On a montré que pour tout n il existait un nombre premier $p > n$. Donc $\mathcal{P}$ n'est pas majoré et par conséquent n'est pas fini.

Pour $p \in \mathcal{P}$ et $n \in \mathbb{N}^*$ on pose $\nu_p(n) = \sup\{\alpha \in \mathbb{N} \mid p^\alpha \text{ divise } n\}$

THEOREME 8.4.4 (Décomposition en facteurs premiers)
Soit n un entier naturel non nul.

1. $\mathcal{P}_n := \{p \in \mathcal{P} |\ \nu_p(n) > 0\}$ *est fini.*
2. $n = \prod_{p \in \mathcal{P}_n} p^{\nu_p(n)}$.
3. *Unicité : soit* $n = \prod_{i=1}^{i=r} p_i^{\alpha_i}$ *avec* $p_1, \ldots, p_r$ *premiers,* $\alpha_1 > 0, \ldots, \alpha_r > 0$.
   Alors $\{p_1, \ldots, p_r\} = \mathcal{P}_n$ *et pour tout* i *entre 1 et* r, $\alpha_i = \nu_{p_i}(n)$.

Il est commode d'écrire $n = \prod_{p \in \mathcal{P}} p^{\nu_p(n)}$ où dans ce produit infini tous les termes sauf un nombre fini sont égaux à 1.

preuve
La preuve de 1) et de 2) est facile, par exemple par récurrence sur n en utilisant le théorème 8.4.2. Montrons l'unicité.

- Vu la définition de $\mathcal{P}_n$ il est clair que $\{p_1, \ldots, p_r\} \subset \mathcal{P}_n$. Montrons par l'absurde l'inclusion réciproque. Supposons qu'il existe un $q \in \mathcal{P}_n$ tel que $q \notin \{p_1, \ldots, p_r\}$. Alors q est premier, distinct de chacun des nombres premiers p_i donc q est premier avec chacun des p_i. Une application répétée du théorème 8.3.3 montre que q est premier avec $\prod_{i=1}^{i=r} p_i^{\alpha_i} = n$. Contradiction car si $q \in \mathcal{P}_n$, alors q divise n.
- Pour tout j, $1 \leqslant j \leqslant r$, le nombre $p_j^{\alpha_j}$ divise n, donc par définition de la fonction ν_p on a $\alpha_j \leqslant \nu_{p_j}(n)$. Montrons l'inégalité réciproque. Soit $m_j = p_1^{\alpha_1} \cdots p_{j-1}^{\alpha_{j-1}} p_{j+1}^{\alpha_{j+1}} \cdots p_r^{\alpha_r}$. Toujours d'après le théorème 8.4.2 p_j est premier avec m_j, donc aussi $p_j^{\nu_{p_j}(n)}$. Or $p_j^{\nu_{p_j}(n)}$ divise le produit $mp_j^{\alpha_j}$. Donc (théorème de Gauss) $p_j^{\nu_{p_j}(n)}$ divise $p_j^{\alpha_j}$ i.e. $\nu_{p_j}(n) \leqslant \alpha_j$, ce qui achève la preuve.

COROLLAIRE 8.4.1
1) *Soient $n \in \mathbb{N}^*$ et p un nombre premier. On a $p|n \Leftrightarrow p \in \mathcal{P}_n \Leftrightarrow \nu_p(n) \geqslant 1$*
2) *Soient $m, n \in \mathbb{N}$. On a pour tout $p \in \mathcal{P}$, $\nu_p(mn) = \nu_p(m) + \nu_p(n)$.*

3) *Soient $n_1, \ldots, n_q$ q entiers naturels. Posons pour $p \in \mathcal{P}$, $\alpha_p = \min\{\nu_p(n_1), \ldots, \nu_p(n_q)\}$ et $\beta_p = \max\{\nu_p(n_1), \ldots, \nu_p(n_q)\}$. On a pgcd$(n_1, \ldots, n_q) = \prod_{p\in\mathcal{P}} p^{\alpha_p}$ et ppcm$(n_1, \ldots, n_q) = \prod_{p\in\mathcal{P}} p^{\beta_p}$.*

PROPOSITION 8.4.1
Soit p premier et k un entier, $1 \leqslant k \leqslant p-1$. Alors p divise le coefficient binomial C_p^k.

preuve
Comme $k > 0$ on peut écrire $k!C_p^k = p(p-1)\cdots(p-k+1)$ donc p divise $k!C_p^k$. D'autre part, $k < p$ donc p ne divise pas $k!$. Il en résulte que $p \wedge k! = 1$ et d'après le théorème de Gauss, p divise C_p^k.

8.5 Anneaux $\mathbb{Z}/n\mathbb{Z}$

Dans tout cette partie n est un entier naturel $n \geqslant 2$.

8.5.1 Eléments inversibles

THEOREME 8.5.1
Soit $a \in \mathbb{Z}$. Les propriétés suivantes sont équivalentes :
(1) *$\overline{a}$ est inversible dans $\mathbb{Z}/n\mathbb{Z}$.*
(2) *$\overline{a}$ est régulier dans $\mathbb{Z}/n\mathbb{Z}$.*
(3) *$a \wedge n = 1$.*
(4) *$\overline{a}$ est un générateur du groupe cyclique $\mathbb{Z}/n\mathbb{Z}$.*

preuve

(1) $\Leftrightarrow$ (2) Dans un anneau quelconque, tout élément inversible est régulier et dans un anneau fini, tout élément régulier est inversible.

(1) $\Rightarrow$ (4) Soit $c \in \mathbb{N}$ tel que $\overline{a} \cdot \overline{c} = \overline{1}$. On a $c \cdot \overline{a} = \underbrace{\overline{a} + \cdots + \overline{a}}_{c \text{ fois}} = \overline{1}$. Donc $\overline{1}$ appartient au groupe $< \bar{a} >$ engendré par $\overline{a}$. Alors $\mathbb{Z}/n\mathbb{Z} = < \overline{1} > \subset < \overline{a} >$. $\overline{a}$ engendre $\mathbb{Z}/n\mathbb{Z}$.

(4) $\Rightarrow$ (3) Si $\overline{a}$ engendre le groupe $\mathbb{Z}/n\mathbb{Z}$, il existe $c \in \mathbb{N}$ tel que $c \cdot \overline{a} = \overline{1}$ ce qui s'écrit $ca \equiv 1 \bmod n$; il existe donc $k \in \mathbb{Z}$ tel que $ac = 1 + kn$ donc (Bézout) $a \wedge n = 1$.

(3) $\Rightarrow$ (1) Par Bézout, il existe $u, v \in \mathbb{Z}$ tels que $au + vn = 1$, d'où on déduit $\overline{a} \cdot \overline{u} = \overline{1}$. ■

DEFINITION 8.5.1
On appelle indicateur d'Euler la fonction $\varphi : \mathbb{N}\setminus\{0\} \to \mathbb{N}$ définie comme suit : $\varphi(1) = 1$ et pour n entier naturel, $n \geqslant 2$, $\varphi(n)$ est le nombre d'entiers k premiers avec n et vérifiant $1 \leqslant k \leqslant n$.

Si p est un nombre premier on a donc $\varphi(p) = p - 1$. Le calcul complet de $\varphi(n)$ est fait au paragraphe 8.5.3.

On notera $U(n) = (\mathbb{Z}/n\mathbb{Z})^*$ le groupe des éléments inversibles de l'anneau $\mathbb{Z}/n\mathbb{Z}$. On a donc card$(U(n)) = \varphi(n)$.

8.5.2 Structure de corps

THEOREME 8.5.2
Les propriétés suivantes sont équivalentes :
(1) *$\mathbb{Z}/n\mathbb{Z}$ est intégre.*
(2) *$\mathbb{Z}/n\mathbb{Z}$ est un corps.*
(3) *n est un nombre premier.*

preuve

(1) $\Rightarrow$ (3) Si n n'est pas premier, il s'écrit $n = ab$ avec $1 < a < n$ et $1 < b < n$. On a $a \notin n\mathbb{Z}$ donc $\overline{a} \neq \overline{0}$ et de même $\overline{b} \neq \overline{0}$ alors que $\overline{a} \cdot \overline{b} = \overline{ab} = \overline{n} = \overline{0}$ ce qui prouve que $\mathbb{Z}/n\mathbb{Z}$ n'est pas intègre.

(3) $\Rightarrow$ (2) Soit $\alpha \in \mathbb{Z}/n\mathbb{Z}$ non nul et r son reste modulo n. On a $0 < r < n$ donc, n étant premier, $n \wedge r = 1$. D'après Bézout, il existe deux entiers u et v tels que $ur + vn = 1$ ce qui donne $\bar{u} \cdot \alpha = \overline{1}$. Tout élément non nul de $\mathbb{Z}/n\mathbb{Z}$ est inversible, donc $\mathbb{Z}/n\mathbb{Z}$ est un corps.

(2) $\Rightarrow$ (1) C'est trivial.

Le corps $\mathbb{Z}/p\mathbb{Z}$ est usuellement noté $\mathbb{F}_p$.

8.5.3 Théorème chinois

THEOREME 8.5.3
Soient $m_1, \ldots, m_q$ q entiers naturels supérieurs ou égaux à 2 et deux à deux premiers entre eux. Il existe un isomorphisme $G : \mathbb{Z}/(m_1 \cdots m_q)\mathbb{Z} \to \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_q\mathbb{Z}$ tel que pour tout $x \in \mathbb{Z}, G\left(\overline{x}_{m_1 \cdots m_q}\right) = \left(\overline{x}_{m_1}, \cdots, \overline{x}_{m_q}\right)$.

Ici, $\overline{x}_m$ désigne la classe de x dans $\mathbb{Z}/m\mathbb{Z}$.

preuve
Considérons l'application $g : \mathbb{Z} \to \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_q\mathbb{Z}$ qui à un entier x associe $\left(\overline{x}_{m_1}, \cdots, \overline{x}_{m_q}\right)$. C'est un morphisme d'anneaux.
Un élément x de $\mathbb{Z}$ est dans le noyau de g ssi il est divisible par chacun des m_k, $1 \leqslant k \leqslant q$. Ces nombres étant premiers entre eux deux à deux, on a donc $x \in \ker(g) \Leftrightarrow m_1 \cdots m_q$ divise x. Donc $\ker(g) = (m_1 \cdots m_q)\mathbb{Z}$.
Le théorème d'isomorphisme des anneaux garantit que g passe aux quotients en un isomorphisme $\overline{g}$ de $\mathbb{Z}/(m_1 \cdots m_q)\mathbb{Z}$ sur $\mathrm{im}(g)$. Il en résulte que $\mathrm{card}(\mathrm{im}(g)) = m_1 \cdots m_q$. Comme $\mathrm{im}(\overline{g}) = \mathrm{im}(g)$ est un sous ensemble de $\mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_q\mathbb{Z}$ et que ce dernier a aussi pour cardinal $m_1 \cdots m_q$, on a $\mathrm{im}(\overline{g}) = \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_q\mathbb{Z}$ ce qui achève la preuve en prenant $G = \overline{g}$.
Conséquence
Si on se donne des entiers $a_1, \ldots, a_q$, il existe un entier x vérifiant les q congruences $x \equiv a_k \bmod m_k$ $1 \leqslant k \leqslant q$.

Cette propriété est connue sous le nom de théorème chinois.

COROLLAIRE 8.5.1
Soient $m_1, \ldots, m_q$ q entiers naturels supérieurs ou égaux à 2 et deux à deux premiers entre eux. En notant φ l'indicateur d'Euler, on a $\varphi(m_1 \cdots m_q) = \varphi(m_1) \cdots \varphi(m_q)$.

preuve
Celà découle immédiatement des deux résultats suivants dont la preuve est facile :
1) Si $f : A \to B$ est un isomorphisme d'anneaux, f induit un isomorphisme du groupe A^* des éléments inversibles de A sur celui B^* des éléments inversibles de B.
2) Si A et B sont deux anneaux, le groupe des éléments inversibles de l'anneau produit $A \times B$ est isomorphisme au produit $A^* \times B^*$.
$(\mathbb{Z}/(m_1 \cdots m_q)\mathbb{Z})^*$ est donc isomorphe à $(\mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_q\mathbb{Z})^*$ lui même isomorphe à $(\mathbb{Z}/m_1\mathbb{Z})^* \times \cdots \times (\mathbb{Z}/m_q\mathbb{Z})^*$. D'où l'égalité.

COROLLAIRE 8.5.2
Soit $n = p_1^{\alpha_1} \ldots p_m^{\alpha_m}$ la décomposition d'un entier n en produit de facteurs premiers (où les α_k sont strictement positifs). On a

$$\varphi(n) = n \prod_{k=1}^{k=m} \left(1 - \frac{1}{p_k}\right)$$

preuve
Soit p premier et $\alpha > 0$. Les nombres inférieurs à p^α et non premiers avec p^α sont les nombres mp avec $1 \leqslant m \leqslant p^{\alpha-1}$. Ils sont au nombre de $p^{\alpha-1}$. Donc $\varphi\left(p^\alpha\right) = p^\alpha - p^{\alpha-1} = p^\alpha \left(1 - \frac{1}{p}\right)$. D'où la conclusion.

8.6 Calculs en caractéristique p

PROPOSITION 8.6.1
Soit A un anneau commutatif de caractéristique un nombre premier p. Pour tous $x, y \in A$ on a $(x+y)^p = x^p + y^p$.

preuve
A étant commutatif, on peut écrire $(x+y)^p = x^p + y^p + \sum_{k=1}^{p-1} C_p^k \cdot (x^k y^{p-k})$. Or on sait (proposition 8.4.1) que p divise C_p^k. Donc pour tout élément a de A, en écrivant $C_p^k = pq$, il vient $C_p^k \cdot a = (qp) \cdot a = q \cdot (p \cdot a) = q \cdot 0_A = 0_A$ d'où la conclusion.

COROLLAIRE 8.6.1
Soit p un nombre premier. Pour tout $x \in \mathbb{Z}/p\mathbb{Z}$ on a $x^p = x$.

preuve
$\mathbb{Z}/p\mathbb{Z}$ est de caractéristique p. La proposition précédente et une récurrence immédiate montrent que si $x_1, \ldots, x_m$ sont m éléments de $\mathbb{Z}/p\mathbb{Z}$, on a $(x_1 + \cdots + x_m)^p = x_1^p + \cdots x_m^p$. Soit r le reste modulo p de x. En appliquant ceci avec $m = r$ et $x_k = \overline{1}$ pour $1 \leqslant k \leqslant r$ on obtient

$$x^p = \overline{r}^p = \left(\overline{1} + \cdots \overline{1}\right)^p = \left(\overline{1} + \cdots \overline{1}\right) = \overline{r} = x$$

Remarque
Ceci fournit un exemple de polynôme non nul, $X^p - X \in \mathbb{F}_p[X]$ dont la fonction polynôme associée $\mathbb{F}_p \to \mathbb{F}_p \ : x \to x^p - x$ est identiquement nulle.

8.7 Théorèmes classiques

THEOREME 8.7.1 (Euler)
Soient $a \in \mathbb{Z}$ et n un entier, $n \geqslant 2$. Si a et n sont premiers entre eux, on a $a^{\varphi(n)} \equiv 1 \bmod n$.

En effet, puisque a est premier avec n, $\overline{a}$ est un élément du groupe fini $(\mathbb{Z}/n\mathbb{Z})^*$ qui est de cardinal $\varphi(n)$. L'ordre de $\overline{a}$ divise $\varphi(n)$ d'où la conclusion.

THEOREME 8.7.2 (Fermat)
Soient p un nombre premier, $a \in \mathbb{Z}$ non divisible par p. On a $a^{p-1} \equiv 1 \bmod p$.

preuves
1) C'est une application du théorème précédent : a est premier avec p et $\varphi(p) = p - 1$.
2) Voici une autre démonstration du théorème de Fermat, utilisant le corollaire 8.6.1. Soit a un entier premier avec p. La classe $\overline{a}$ est un élément non nul de $\mathbb{Z}/p\mathbb{Z}$. On a $\overline{a}^p - \overline{a} = 0$. Or $\overline{a}$ est inversible dans $\mathbb{Z}/p\mathbb{Z}$ On peut donc simplifier par $\overline{a}$ ce qui donne le résultat.

THEOREME 8.7.3 (Wilson)
Soit p un entier naturel supérieur ou égal à deux. p est premier ssi $(p-1)! + 1 \equiv 0 \bmod p$.

preuve
Supposons $(p-1)! \equiv -1 \bmod p$. Soit k entre 1 et $p-1$. Si $m = (p-1)!/k$ on a $\overline{k} \cdot \overline{-m} = \overline{1}$. Tout élément de $\mathbb{Z}/p\mathbb{Z}$ autre que $\overline{0}$ est inversible, donc $\mathbb{Z}/p\mathbb{Z}$ est un corps et p est premier.
première preuve de la réciproque
Soit p premier. Si $p = 2$ le résultat est immédiat. On suppose donc $p \geqslant 3$.
On considère dans $(\mathbb{Z}/p\mathbb{Z})^*$ la relation R définie par $xRy \Leftrightarrow \left(x = y \text{ ou } xy = \overline{1}\right)$. C'est une relation d'équivalence. Tout x est inversible et $x = \overline{1}/x \Leftrightarrow x^2 - \overline{1} = \overline{0} \Leftrightarrow (x-\overline{1})(x+\overline{1}) = \overline{0} \Leftrightarrow x = \pm\overline{1}$ car $\mathbb{Z}/p\mathbb{Z}$ est intègre. Les classes singletons sont $\{\overline{1}\}$ et $\{\overline{-1}\}$ et les autres classes sont des paires $(x, 1/x)$. En effectuant le produit de tous les éléments de $(\mathbb{Z}/p\mathbb{Z})^*$ regroupés par classes, on obtient $\overline{(p-1)!} = \overline{-1}$.
deuxième preuve de la réciproque(utilisant les polynômes)
Supposons p premier, $p \geqslant 3$. On a $\forall x \in \mathbb{F}_p^*,\ x^{p-1} = \overline{1}$ donc le polynôme $Q(X) = X^{p-1} - \overline{1}$ admet comme racine tous les

éléments de $\mathbb{F}_p^*$. Il s'écrit donc aussi $\prod_{1\leqslant k\leqslant p-1}(X-\overline{k})$. La formule donnant le produit des racines d'un polynôme donne alors $(-1)^{p-1}\prod_{k=1}^{p-1}\overline{k}=-\overline{1}$ d'où le résultat puisque $p-1$ est pair.

8.8 Application aux groupes cycliques

Soit $(G,*)$ un groupe monogène que nous supposerons distinct de $\{1_G\}$ et $a\in G$ un générateur de G. On a vu qu'il existait un unique morphisme du groupe additif $(\mathbb{Z},+)$ dans G soit f_a tel que $f_a(1)=a$ dont l'image est précisément le groupe engendré par a i.e. G. Le noyau de ce morphisme est un sous groupe $n\mathbb{Z}$ de $\mathbb{Z}$ et on a $n\neq 1$ car $a\neq 1_G$. Le cas $n=0$ correspond à un groupe monogène infini. Pour $n\geqslant 2$ le premier théorème d"isomorphisme fournit un isomorphisme $\overline{f_a}:\mathbb{Z}/n\mathbb{Z}\to G$ tel que $f_a(\bar{k})=a^k$ pour tout $k\in\mathbb{Z}$.

Donc tout groupe cyclique de cardinal n est isomorphe à $\mathbb{Z}/n\mathbb{Z}$.

Ce résultat permet de ramener l'étude des groupes cycliques à l'étude des groupes $(\mathbb{Z}/n\mathbb{Z},+)$.

THEOREME 8.8.1
Soit G un groupe cyclique de cardinal $n\geqslant 2$.
1) G admet $\varphi(n)$ générateurs où φ est l'indicateur d'Euler.
2) Tout sous groupe de G est cyclique. Plus précisément, pour tout diviseur d de n, il existe un unique sous groupe de G de cardinal d et ce sous groupe est cyclique.
3) Tout quotient de G est cyclique.

preuve
1) Si $h:G\to G'$ est un isomorphisme de groupes, un élément a de G engendre G ssi $h(a)$ engendre G'. Le 1) résulte alors du théorème 8.5.1 et de la définition de l'indicateur d'Euler.
2) Là encore il sufft de montrer le résultat pour $\mathbb{Z}/n\mathbb{Z}$. Soit d un diviseur de n. Ecrivons $n=dm,\ m\in\mathbb{N}$. Il est immédiat que l'élément $\overline{m}$ de $\mathbb{Z}/n\mathbb{Z}$ est d'ordre d donc le sous groupe H_0 qu'il engendre est cyclique d'ordre d. On a $H_0=\{\overline{0},\overline{m},\ldots,\overline{(d-1)m}\}$.
Réciproquement, soit H un sous groupe de $\mathbb{Z}/n\mathbb{Z}$ de cardinal d. Soit $\pi:\mathbb{Z}\to\mathbb{Z}/n\mathbb{Z}$ la projection canonique et $S=\pi^{-1}(H)$. S est un sous groupe de $\mathbb{Z}$. Il existe donc un unique entier naturel p tel que $S=p\mathbb{Z}$. On a $n\mathbb{Z}=\ker(\pi)\subset\pi^{-1}(H)=p\mathbb{Z}$ donc p divise n. Soit $n=p\delta$. Comme π est surjective, on a $H=\pi\left(\pi^{-1}(H)\right)=\pi(p\mathbb{Z})$ donc H est le sous groupe de $\mathbb{Z}/n\mathbb{Z}$ engendré par $\overline{p}$. D'après la partie directe il est de cardinal δ. Donc $\delta=d,\ m=p$ et $H=H_0$.
3) Soit G un groupe cyclique, H un sous groupe de G et $p:G\to G/H$ la projection canonique. Il est immédiat que si $a\in G$ engendre G, sa projection $p(a)$ engendre G/H ce qui prouve le résultat.

EXEMPLE 8.8.1
Soit U_n l'ensemble des racines n-ièmes de l'unité dans $\mathbb{C}$, i.e. $U_n=\{z\in\mathbb{C}\ ;\ z^n=1\}$. Il est immédiat que U_n est un sous groupe de $(\mathbb{C}^*,\cdot)$. On sait que $U_n=\left\{e^{2k\pi i/n}\ ;\ 0\leqslant k\leqslant n-1\right\}$. U_n est cyclique, engendré par $\omega=e^{2i\pi/n}$. On appelle racine primitive n-ième de l'unité tout générateur de U_n. Il y en a donc $\varphi(n)$. Ce sont les $e^{2k\pi i/n}$ avec $1\leqslant k\leqslant n$ et $k\wedge n=1$.

THEOREME 8.8.2
Soient G_1 et G_2 deux groupes cycliques de cardinaux respectifs n_1 et n_2. Le groupe produit $G_1\times G_2$ est cyclique ssi n_1 et n_2 sont premiers entre eux.

preuve
Soient $f_k:\mathbb{Z}/n_kZ\to G_k,\ k=1,2$ des isomorphismes et $f:\mathbb{Z}/n_1Z\times\mathbb{Z}/n_2\mathbb{Z}\to G_1\times G_2$ définie par $f(x_1,x_2)=(f_1(x_1),f_2(x_2))$. Il est immédiat que f est un isomorphisme. Si $n_1\wedge n_2=1$ il résulte du théorème chinois que $\mathbb{Z}/n_1\mathbb{Z}\times\mathbb{Z}/n_2\mathbb{Z}$ est isomorphe à $\mathbb{Z}/n_1n_2\mathbb{Z}$. Il en est donc de même de G qui est cyclique.
Supposons maintenant que n_1 et n_2 ne soient pas premiers entre eux. Leur ppcm m est alors strictement inférieur à n_1n_2. On écrit $m=n_1d_1=n_2d_2$ avec $d_1<n_2$ et $d_2<n_1$. Soit $(\alpha,\beta)\in\mathbb{Z}/n_1\mathbb{Z}\times\mathbb{Z}/n_2\mathbb{Z}$. On a $m(\alpha,\beta)=(d_1(n_1\alpha),d_2(n_2\beta))=(\bar{0},\bar{0})$. Donc le groupe $\mathbb{Z}/n_1\mathbb{Z}\times\mathbb{Z}/n_2\mathbb{Z}$ ne contient aucun élément d'ordre n_1n_2. Il n'est pas cyclique et G qui lui est isomorphe non plus.

8.9 Applications

8.9.1 Algorithme de cryptage RSA

Voici sous forme d'exercice le principe de l'algorithme de cryptage RSA. (RSA pour Ronald **R**ivest, Adi **S**hamir et Leonard **A**dleman du MIT)

Soient p et q deux nombres premiers distincts et $n = pq$.

1. Montrer que $\forall x \in \mathbb{Z}/n\mathbb{Z},\ x^{(p-1)(q-1)+1} = x$. *On pourra distinguer les cas* $x \wedge n = 1$ *et* $x \wedge n > 1$.
2. Soit e un entier, $1 \leqslant e \leqslant n-1$ tel que $e \wedge (p-1)(q-1) = 1$. Montrer qu'il existe un entier naturel d tel que $ed \equiv 1 \bmod (p-1)(q-1)$.
3. En déduire que l'application $h : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ définie par $h(x) = x^e$ est bijective et déterminer h^{-1}.
4. Exemple : $p = 17,\ q = 19,\ e = 7$. Déterminer d.

Système de cryptage à clef publique
L'auteur du système choisit deux nombres premiers p et q avec $p < q$. Le couple (p, q) est la clef secrète de chiffrement. Elle n'est connue que de l'auteur du système. Ce dernier calcule $n = pq$ et choisit aussi un entier $e,\ 1 \leqslant e \leqslant n-1$ premier avec $(p-1)(q-1)$. Il calcule enfin d tel que $ed \equiv 1 \bmod (p-1)(q-1)$.
La clef publique de chiffrement est le couple (n, e). Elle est fournie à toute personne désireuse d'envoyer un message chiffé à l'auteur. Cette personne procède comme suit :

- Elle transforme le message en une suite de chiffres, par un procédé public, par exemple en associant à chaque caractère son numéro à deux chiffres dans le code ASCII.
- Elle obtient ainsi un message B formé d'une suite de chiffres. Elle le décompose en tranches, chacun des nombres constituant une tranche étant voisin de $n-1$ mais inférieur à $n-1$, en ajoutant des zéros pour la dernière tranche si nécessaire. Elle obtient $B = b_1 \cdot b_2 \cdots b_m$ où $b_i \in [0 \cdot \cdot n-1]$ identifié à $\mathbb{Z}/n\mathbb{Z}$.
- Elle calcule c_k reste modulo n de b_k^e.
- Le message codé qui va être transmis est la suite $c_1, c_2, \cdots, c_m$.

Le destinataire qui connait le nombre d peut déchiffrer le message en calculant $b_k = c_k^d \bmod n$. La sureté du système vient du fait que l'on ne sait pas décomposer en un temps raisonnable le nombre n en produit de facteurs premiers. Il en résulte qu'une personne ne connaissant pas la clef secrète (p, q) ne peut pas la trouver, et donc ne peut pas calculer d.

8.9.2 Equations diophantiennes

Equation de Bezout

Soient a, b deux entiers relatifs. On considère l'équation dans $\mathbb{Z}^2$:

$$(E) \qquad ax + by = 1$$

(E) a des solutions si et seulement si $a \wedge b = 1$. Résoudre (E) permet entre autre de trouver l'inverse de $\bar{a}$ dans $\mathbb{Z}/b\mathbb{Z}$.

- Cas $ab = 0$. Comme a et b sont premiers entre eux, un seul des deux nombres est nul, par exemple a l'autre b étant nécessairement égal à ± 1. L'ensemble des solutions est $\{(x, b)\ ;\ x \in \mathbb{Z}\}$.
- Cas $ab \neq 0$. Le théorème de Bézout garantit l'existence d'au moins une solution (x_0, y_0). Un couple (x, y) est solution ssi $ax + by = ax_0 + by_0$ soit $a(x - x_0) = b(y_0 - y)$. Si (x, y) est solution, a divise $b(y_0 - y)$ et a est premier avec b donc a divise $y_0 - y$; donc il existe $k \in \mathbb{Z}$ tel que $y_0 - y = ka$ ce qui en reportant donne $x - x_0 = kb$. Réciproquement, pour tout $k \in \mathbb{Z}$, $(x_0 + kb, y_0 - ka)$ est solution. L'ensemble des solutions est donc $\{(x_0 + kb, y_0 - ka)\ ;\ k \in \mathbb{Z}\}$.

PROPOSITION 8.9.1
Soient $a, b \in \mathbb{Z}$ *deux entiers relatifs non nuls, premiers entre eux, avec* $|b| \geqslant 2$. *L'équation* $ax + by = 1$ *admet une unique solution* (x_0, y_0) *tel que* $1 \leqslant x_0 \leqslant |b| - 1$. *On a alors* $|y_0| \leqslant |a|$ *cette inégalité étant stricte dès que* $|a| \geqslant 2$.

preuve
Si (x, y) est solution de (E), on a $\overline{ax} = \overline{1}$ dans $\mathbb{Z}/|b|\mathbb{Z}$. Comme $a \wedge b = 1$, $\overline{a}$ est inversible dans $\mathbb{Z}/|b|\mathbb{Z}$. Soit α l'inverse. Soit x_0 le reste de α modulo $|b|$. On a donc $0 < x_0 \leqslant |b| - 1$ et c'est le seul représentant de la classe α entre 1 et $|b| - 1$. Ce qui prouve l'unicité. Ensuite, par définition on a $\overline{ax_0} = \overline{1}$; autrement dit, $|b|$ divise $1 - ax_0$. Il existe donc y_0 tel que $1 - ax_0 = by_0$. Ce qui prouve l'existence.
Enfin, on a $|by_0| \leqslant 1 + |a||x_0| \leqslant 1 + |a|\,(|b| - 1)$ soit puisque $b \neq 0$, $|y_0| \leqslant |a| + \frac{1}{|b|}\,(1 - |a|) \leqslant |a|$ l'inégalité étant stricte si $|a| \geqslant 2$.

Cet énoncé tombe évidemment en défaut si l'un des coefficients est nul. On peut préciser davantage si on connait les signes de a et b. Par exemple, si $a \geqslant 2$ et $b \geqslant 2$ l'équation $ax - by = 1$ admet une unique solution (x_0, y_0) avec $0 < x_0 < b$ et $0 < y_0 < a$.

Un fait remarquable est que l'algorithme d'Euclide permet de calculer une solution "minimale". Considérons deux entiers naturels, a et b non nuls avec $a > b$ premiers entre eux.
On pose $r_0 = a$, $r_1 = b$; on effectue les divisions successives et on s'arrête lorsque le reste vaut 1 :

$$r_{k-1} = r_k q_k + r_{k+1}, \quad 1 \leqslant k \leqslant s$$

avec $r_{s+1} = 1$. On "remonte " ces égalités ; on obtient

$$1 = \alpha_{k-1} r_{k-1} + \beta_{k-1} r_k \text{ avec } 1 \leqslant k \leqslant s$$

Pour $k = s$ on a $r_{s-1} = q_s r_s + 1$ soit $\alpha_{s-1} = 1$ et $\beta_{s-1} = -q_s$. Ensuite, on passe de k à $k - 1$ avec $r_k = r_{k-2} - q_{k-1} r_{k-1}$ ce qui donne

$$1 = \alpha_{k-1} r_{k-1} + \beta_{k-1}(r_{k-2} - q_{k-1} r_{k-1}) = \beta_{k-1} r_{k-2} + (\alpha_{k-1} - \beta_{k-1} q_{k-1}) r_{k-1}$$

d'où

$$\alpha_{k-2} = \beta_{k-1} \qquad \beta_{k-2} = \alpha_{k-1} - \beta_{k-1} q_{k-1}$$

Prouvons alors que pour $1 \leqslant k \leqslant s$ on a

$$|\alpha_{k-1}| < r_k \quad \text{et} \quad |\beta_{k-1}| < r_{k-1}$$

C'est vrai pour $k = s$ car $|\alpha_{s-1}| = 1 < r_s$ (puisque le premier reste égal à 1 est r_{s+1}) et $|\beta_{s-1}| = q_s < r_{s-1}$ par définition de la division euclidienne. Supposons ces inégalités vraies pour un entier $k \geqslant 2$ et vérifions les pour l'entier $k - 1$. On a $|\alpha_{k-2}| = |\beta_{k-1}| < r_{k-1}$ d'une part et $|\beta_{k-2}| \leqslant |\alpha_{k-1}| + |\beta_{k-1}||q_{k-1}| < r_k + r_{k-1} q_{k-1} = r_{k-2}$. Ceci achève la preuve de ces inégalités. Pour $k = 1$ on obtient $|\alpha_0| < r_1 = b$ et $|\beta_0| < r_0 = a$ avec $1 = \alpha_0 a + \beta_0 b$.

9

Polynômes

9.1 Préliminaires

Soit A un anneau commutatif et K un sous corps de A, i.e. un sous anneau de A qui est un corps pour les lois induites. Alors A a une structure naturelle de K espace vectoriel, et même de K-algèbre associative unitaire , la loi externe $K \times A \to A$ étant simplement la restriction de la multiplication de A à $K \times A$.

THEOREME 9.1.1
Soient A un anneau intègre et K un sous corps de A. Si A est un K-espace vectoriel de dimension finie, alors A est lui même un corps.

preuve
Soit $a \in A,\ a \neq 0$. Il s'agit de montrer que a est inversible. L'application $A \to A$ qui à $x \in A$ associe ax est linéaire. Comme A est intègre, elle est injective. Comme A est de dimension finie, c'est un isomorphisme du K espace vectoriel A, et elle est bijective. Il en résulte l'existence d'un $a' \in A$ tel que $aa' = 1$ ce qui achève la preuve.

9.2 Anneau des polynômes

9.2.1 Polynômes à coefficients dans un anneau commutatif

Soit A un anneau commutatif. On notera $A^{\mathbb{N}}$ l'ensemble des suites d'éléments de A. Soient $\underline{a} = (a_n)$ et $\underline{b} = (b_n)$ deux éléments de $A^{\mathbb{N}}$.
On définit $\underline{a} + \underline{b} = (a_n + b_n)_{n \geqslant 0}$ et $\underline{a} * \underline{b} = (c_n)_{n \geqslant 0}$ avec $c_n = \sum_{k=0}^{k=n} a_k b_{n-k} = \sum_{p+q=n} a_p b_q$.
On notera provisoirement $\mathbf{0}$ la suite constante dont tous les termes valent 0 et $\mathbf{1}$ la suite $\mathbf{1} = (\delta_{0,n})_{n \geqslant 0}$ où $\delta_{i,j}$ est le symbole de Kronecker à valeur dans A, i.e; $\delta_{i,j} = 0 (= 0_A)$ si $i \neq j$ et $\delta_{i,i} = 1 = (1_A)$ pour tout i. $\mathbf{1}$ est donc la suite dont le premier terme est 1 et tous les autres nuls.
Enfin, pour $\underline{a} \in A^{\mathbb{N}}$ et $\lambda \in A$ on pose $\lambda \cdot \underline{a} = (\lambda a_n)_{n \geqslant 0}$.

THEOREME 9.2.1
Muni des opérations $+$ et $*$ définies ci dessus, $A^{\mathbb{N}}$ est un anneau commutatif dont l'élément nul est $\mathbf{0}$ et l'élément unité $\mathbf{1}$. L'application j qui à $a \in A$ associe la suite $a \cdot \mathbf{1}$ est un isomorphisme de l'anneau A sur le sous anneau A_0 de $A^{\mathbb{N}}$ formé des suites dont tous les termes sont nuls sauf celui d'indice 0. On a enfin pour tout $\lambda \in A$ et tout $\underline{a} \in A^{\mathbb{N}} \quad \lambda \cdot \underline{a} = j(\lambda) * \underline{a}$.

Les vérifications sont laissées au lecteur.
La dernière propriété fait que si on identifie A et A_0 en identifiant $a \in A$ avec son image $j(a) \in A^{\mathbb{N}}$ l'opération $\cdot$ devient la restriction de $*$ à $A_0 \times A^{\mathbb{N}}$.

DEFINITION 9.2.1
On appelle série entière formelle à coefficients dans A tout élément de l'anneau $A^{\mathbb{N}}$.

DEFINITION 9.2.2
On appelle polynôme à une indéterminée à coefficients dans A tout élément de l'ensemble $A^{(\mathbb{N})}$ des suites $(a_n)_{n\geqslant 0} \in A^{\mathbb{N}}$ dont tous les termes, sauf un plus un nombre fini, sont nuls.

DEFINITION 9.2.3
Soit $\underline{a} \in A^{(\mathbb{N})},\ \underline{a} \neq \mathbf{0}$. On appelle degré de $\underline{a}$ l'entier

$$\deg(\underline{a}) = \max\{n \in \mathbb{N}\ ;\ a_n \neq 0\}$$

On appelle valuation de $\underline{a}$ l'entier

$$\mathrm{val}(\underline{a}) = \min\{n \in \mathbb{N}\ ;\ a_n \neq 0\}$$

La définition a un sens car l'ensemble $\{p \in \mathbb{N}\ ;\ a_p \neq 0\}$ est un ensemble fini non vide d'entiers naturels ; il a donc un plus grand et un plus petit élément.
Par convention, on pose quelquefois $\mathrm{val}(\mathbf{0}) = +\infty$ et $\deg(\mathbf{0}) = -\infty$.
Si $\underline{a}$ est un polynôme de degré p, le coefficient a_p est appelé coefficient dominant. Un polynôme est dit unitaire ou normalisé si son coefficient dominant est 1.

THEOREME 9.2.2
$A^{(\mathbb{N})}$ est un sous anneau de $A^{\mathbb{N}}$.

preuve

- Soient $\underline{a}$ et $\underline{b}$ dans $A^{(\mathbb{N})}$.
  Pour $n > \max(\deg(\underline{a}), \deg(\underline{b}))$ on a $(\underline{a} - \underline{b})_n = a_n - b_n = 0$ donc $\underline{a} - \underline{b} \in A^{(\mathbb{N})}$. Si $\deg(\underline{a}) \neq \deg(\underline{b})$ on voit que, pour $n = \max(\deg(\underline{a}), \deg(\underline{b}))$ on a $a_n - b_n \neq 0$. Ceci vaut aussi pour $\underline{a} + \underline{b}$.
- $\mathbf{1} \in A^{(\mathbb{N})}$.
- Si $n > \deg(\underline{a}) + \deg(\underline{b})$ on a $(\underline{a} * \underline{b})_n = \sum_{p+q=n} a_p b_q = 0$ (car $p+q = n \Rightarrow p > \deg(\underline{a})$ ou $q > \deg(\underline{b})$) donc $\underline{a} * \underline{b} \in A^{(\mathbb{N})}$.

Dans la démonstration précédente, on a montré les affirmations relatives au degré de la proposition suivante.

PROPOSITION 9.2.1
Soient $\underline{a}$ et $\underline{b}$ des polynômes non nuls.
1) $\deg(\underline{a} + \underline{b}) \leqslant \max(\deg(\underline{a}), \deg(\underline{b}))$ *et si* $\deg(\underline{a}) \neq \deg(\underline{b})$ *on a l'égalité.*
2) $\mathrm{val}(\underline{a} + \underline{b}) \geqslant \min(\mathrm{val}(\underline{a}), \mathrm{val}(\underline{b}))$ *et si* $\mathrm{val}(\underline{a}) \neq \mathrm{val}(\underline{b})$ *on a l'égalité.*
3) $\deg(\underline{a} * \underline{b}) \leqslant \deg(\underline{a}) + \deg(\underline{b})$
4) $\mathrm{val}(\underline{a} * \underline{b}) \geqslant \mathrm{val}(\underline{a}) + \mathrm{val}(\underline{b})$.

THEOREME 9.2.3
*Soit A un anneau **intègre**. L'anneau des polynômes à coefficients dans A est intègre. Si $\underline{a}$ et $\underline{b}$ sont des polynômes non nuls, on a*
$\deg(\underline{a} * \underline{b}) = \deg(\underline{a}) + \deg(\underline{b})$
$\mathrm{val}(\underline{a} * \underline{b}) = \mathrm{val}(\underline{a}) + \mathrm{val}(\underline{b})$.

En effet, si $n = \deg(\underline{a}) + \deg(\underline{b})$ on a $(\underline{a} * \underline{b})_n = a_{\deg(\underline{a})} \cdot b_{\deg(\underline{b})} \neq 0$ puisque A est intègre. En particulier $\underline{a} * \underline{b} \neq 0$. La preuve est analogue pour la valuation.
Avec les conventions faites sur le degré et la valuation du polynôme nul, les résultats subsistent si $\underline{a}$ ou $\underline{b}$ est nul, en convenant que pour tout entier k, $\quad k + \infty = \infty, \quad k - \infty = -\infty, \quad \max(k, \infty) = \infty, \quad \min(k, -\infty) = -\infty$.

Notations définitives

Dans la suite on omet le signe $*$ pour la multiplication dans $A^{(\mathbb{N})}$. On identifie A et A_0 en identifiant tout élément $a \in A$ avec son image $j(a) \in A_0$. Un tel polynôme sera dit constant.
Les polynômes $\mathbf{0}$ et $\mathbf{1}$ seront notés respectivement 0 et 1 sauf risques de confusion.
On note $X = (\delta_{1,n})_{n\geqslant 0}$. On vérifie facilement que $X^2 = (\delta_{2,n})_{n\geqslant 0}$ puis par récurrence que $X^p = (\delta_{p,n})_{n\geqslant 0}$. Un élément $\underline{a}$

non nul s'écrit alors $\underline{a} = \sum_{k=0}^{\deg(\underline{a})} a_k X^k$. Il en résulte que le plus petit sous anneau de $A^{(\mathbb{N})}$ contenant A (indentifié à A_0) et X est $A^{(\mathbb{N})}$ lui même. On note $A^{(\mathbb{N})} = A[X]$.

Tout $\underline{a} \in A[X]$ s'écrit $\underline{a} = \sum_{k=0}^{k=N} a_k X^k$ pour tout entier $N > \deg(\underline{a})$. On se permettra aussi d'écrire $\underline{a} = \sum_{k=0}^{\infty} a_k X^k$ ce qui a un sens puisque la famille $(a_k X^k)_{k\in\mathbb{N}}$ est à support fini.

NB Le programme ne prévoit que l'étude des polynômes à coefficients dans un corps commutatif. Cette hypothèse ne sert pas dans la construction de l'anneau $A[X]$ c'est pourquoi on a fait cette construction dans le cas d'un anneau commutatif.

9.2.2 Polynômes à coefficients dans un corps

Dans toute la suite on se limitera, conformément au programme, au cas où A est un corps commutatif que l'on notera $\mathbb{K}$. On notera 0 et 1 les éléments $0_{\mathbb{K}}$ et $1_{\mathbb{K}}$ de $\mathbb{K}$ sauf exception. L'application qui à $\lambda \in \mathbb{K}$ associe le polynôme constant $\lambda \in \mathbb{K}[X]$ est un isomorphisme de $\mathbb{K}$ sur le sous corps de $\mathbb{K}[X]$ constitué par l'ensemble $\mathbb{K}_0[X]$ des polynômes constants. On identifiera dans la suite $\mathbb{K}$ et $\mathbb{K}_0[X]$. Il en résulte, d'après le préliminaire que $\mathbb{K}[X]$ est un $\mathbb{K}$-espace vectoriel et même une $\mathbb{K}$-algèbre associative unitaire.

Par construction, tout polynôme $P \in \mathbb{K}[X]$ s'écrit de manière unique comme combinaison linéaire des X^k. Donc $\left(X^k\right)_{k\in\mathbb{N}}$ est une base de $\mathbb{K}[X]$ appellée base canonique.

Reprenons les résultats établis au paragraphe précédent.

THEOREME 9.2.4
Soient P et Q deux polynômes non nuls.
1) $\deg(P+Q) \leqslant \max(\deg(P), \deg(Q))$ *et si* $\deg(P) \neq \deg(Q)$ *on a l'égalité.*
2) $\deg(PQ) = \deg(P) + \deg(Q)$.
3) $\mathrm{val}(P+Q) \geqslant \min(\mathrm{val}(P), \mathrm{val}(Q)$ *et si* $\mathrm{val}(P) \neq \mathrm{val}(Q)$ *on a l'égalité.*
4) $\mathrm{val}(PQ) = \mathrm{val}(P) + \mathrm{val}(Q)$.

THEOREME 9.2.5
1)Soit $\mathbb{K}$ un corps commutatif. L'anneau $\mathbb{K}[X]$ des polynômes à une indéterminée à coefficients dans $\mathbb{K}$ est intègre.
2) Les éléments inversibles de l'anneau $\mathbb{K}[X]$ sont les polynômes constants non nuls.

preuve
1) Déjà prouvé.
2) Soit $P \in \mathbb{K}[X]$. Si il existe $Q \in \mathbb{K}[X]$ tel que $PQ = 1$ on a $P \neq 0$ et $0 = \deg(1) = \deg(P) + \deg(Q)$ donc $\deg(P) = 0$ et P est une constante non nulle. La réciproque est immédiate.

Remarque Composé de deux polynômes

Soient $P = a_p X^p + \cdots + a_0$ et $Q = b_q X^q + \cdots + b_0$ deux polynômes. La quantité $\sum_{k=0}^{k=q} b_k P^k$ est un polynôme bien défini, appellé composé des polynômes Q et P dans cet ordre et noté $Q \circ P$. Si maintenant on choisit pour P le polynôme X, on voit que $Q \circ X = Q$ autrement dit que $Q(X) = Q$. Ceci justifie l'usage indifférencié des notations Q et $Q(X)$. (On notera la différence avec une fonction f, où f désigne la fonction et $f(x)$ l'image de x par f).

9.2.3 L'espace vectoriel $\mathbb{K}_n[X]$

Soit $\mathbb{K}$ un corps commutatif et n un entier naturel. On notera $\mathbb{K}_n[X]$ le sous ensemble de $\mathbb{K}[X]$ formé des polynômes de degré au plus n.

THEOREME 9.2.6
$\mathbb{K}_n[X]$ *est un $\mathbb{K}$-espace vectoriel de dimension $n+1$ dont une base, dite base canonique, est* $(1, X, \ldots, X^n)$.

THEOREME 9.2.7 (Théorème des degrés échelonnés)
Soient $(P_0, \ldots, P_n)$ *n polynômes appartenant à $\mathbb{K}_n[X]$ tels que* $\deg(P_k) = k$ *pour* $0 \leqslant k \leqslant n$. $(P_0, \cdots, P_n)$ *est une base de $\mathbb{K}_n[X]$.*

preuve

Il suffit de montrer que le système $(P_0, \ldots, P_n)$ est libre. Soient $(\lambda_0, \ldots, \lambda_n) \in \mathbb{K}^{n+1}$ tels que $\sum_{k=0}^{k=n} \lambda_k P_k = 0$. Montrons par l'absurde que les λ_k sont tous nuls. Supposons le contraire. Soit $m = \max\{k \ ; \ \lambda_k \neq 0\}$. Le polynôme $\lambda_0 P_0 + \cdots + \lambda_{m-1} P_{m-1}$ est de degré au plus $m-1$ et comme $\lambda_m \neq 0$, le polynôme $\lambda_m P_m$ est de degré m, donc d'après le théorème 9.2.4 le polynôme $\lambda_0 P_0 + \cdots + \lambda_{m-1} P_{m-1} + \lambda_m P_m$ est de degré m, donc est non nul. Contradiction.
Remarque
Le même argument montre que si $P_0, \ldots, P_n$ sont des polynômes de degré tous distincts, le système $(P_0, \ldots, P_n)$ est libre.

9.3 Arithmétique dans $\mathbb{K}[X]$

9.3.1 Division euclidienne

THEOREME 9.3.1
Soient $A, B \in \mathbb{K}[X]$ avec $B \neq 0$. Il existe un unique couple de polynômes (Q, R) vérifiant

$$\begin{cases} A = BQ + R \\ R = 0 \text{ ou } \deg(R) < \deg(B) \end{cases}$$

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

preuve

- existence
  Si B est un polynôme constant, c'est un élément inversible de $\mathbb{K}[X]$. On prend $Q = \frac{1}{B} A$ et $R = 0$. On suppose donc que le polynôme $B = b_q X^q + \cdots + b_0$ est fixé de degré $q \geqslant 1$ et on prouve l'existence du couple (Q, R) par récurrence sur le degré de A.
  Si $\deg(A) < q$, on prend $Q = 0$ et $R = A$.
  Soit $n \geqslant q-1$. Supposons le théorème établi pour tout polynôme A de degré au plus n. Soit $A = a_{n+1} X^{n+1} + a_n X^n + \cdots + a_0$ un polynôme de degré $n+1$. On a donc $a_{n+1} \neq 0$. Posons

  $$C(X) = A(X) - \frac{a_{n+1}}{b_q} X^{n+1-q} B(X)$$

  C est un polynôme de degré inférieur ou égal à n. D'après l'hypothèse de récurrence il existe un polynôme Q_1 et un polynôme R vérifiant $C = BQ_1 + R$ et $R = 0$ ou $\deg(R) < \deg(B)$. On a alors $A = BQ + R$ en posant $Q = Q_1 + \frac{a_{n+1}}{b_q} X^{n+1-q}$ ce qui achève la preuve.

- unicité
  Si $BQ + R = BQ' + R'$ on a $R - R' = B(Q' - Q)$. Si on suppose $R \neq R'$ on a $Q \neq Q'$ et $\deg(R - R') = \deg(B) + \deg(Q - Q')$. Le second membre est supérieur ou égal à $\deg(B)$. Or les conditions de la division euclidienne imposent $\deg(R - R') < \deg(B)$. Contradiction. Donc $R = R'$ et $\mathbb{K}[X]$ étant intègre $Q = Q'$.

La preuve n'est que la formalisation de l'algorithme pratique qui permet de faire le calcul de Q et de R.

THEOREME 9.3.2
Soit $\mathbb{K}$ un corps commutatif. L'anneau $\mathbb{K}[X]$ est principal.

preuve
L'application "degré "de $\mathbb{K}[X] \setminus \{0\}$ dans $\mathbb{N}$ est un stathme euclidien (voir la définition § 8.1) donc est un anneau principal. Repenons la démonstration.
Soit J un idéal de $\mathbb{K}[X]$ que l'on peut supposer non nul. Alors l'ensemble $\{p \in \mathbb{N} \mid \exists P \in J, P \neq 0 \ \deg(P) = p\}$ est un sous ensemble non vide de $\mathbb{N}$. Il admet donc un plus petit élément, soit m. Il existe $B \in J \setminus \{0\}$ tel que $\deg(B) = m$. On a alors $B \cdot \mathbb{K}[X] \subset J$. Réciproquement soit $A \in J$. Soient Q et R le quotient et le reste de la division euclidienne de A par B. On a $A \in J, \ B \in J$ donc $BQ \in J$ donc $R = A - BQ \in J$. On ne peut avoir $R \neq 0$ car alors R serait un élément de J de degré $< m$. Donc $R = 0, \ A \in B \cdot \mathbb{K}[X]$ et $J \subset B \cdot \mathbb{K}[X]$.

DEFINITION 9.3.1
On dira que deux polynômes P et Q sont associés si il existe un $\lambda \in \mathbb{K}^*$ tels que $P = \lambda Q$.

La relation binaire ainsi définie est une relation d'équivalence dans $\mathbb{K}[X]$ et chaque classe d'équivalence contient un unique élément normalisé. Si P et Q sont deux polynômes quelconques, les idéaux principaux $P \cdot \mathbb{K}[X]$ et $Q \cdot \mathbb{K}[X]$ engendrés par P et Q sont égaux ssi P et Q sont associés. Tout idéal I a donc un unique générateur normalisé P et ses autres générateurs sont les polynômes associés à P.

9.3.2 pgcd,ppcm

La structure d'anneau principal de $\mathbb{K}[X]$ permet de développer toute une arithmétique des polynômes analogue à l'arithmétique dans $\mathbb{Z}$. Rappelons que les éléments inversibles de $\mathbb{K}[X]$ sont les polynômes constants non nuls et que tout idéal I de $\mathbb{K}[X]$ admet un unique générateur normalisé P_0 et que tous les générateurs de I sont les polynômes λP_0, avec $\lambda \in \mathbb{K}^*$.

Soient $A_1, \ldots, A_n \in \mathbb{K}[X]$. Un polynôme P divise tous les A_k ssi $A_k \cdot \mathbb{K}[X] \subset P \cdot \mathbb{K}[X]$ pour tout k, ce qui équivaut à $A_1 \cdot \mathbb{K}[X] + \cdots + A_n \cdot \mathbb{K}[X] \subset P \cdot \mathbb{K}[X]$. Soit D un générateur de l'idéal $A_1\mathbb{K}[X] + \cdots + A_n\mathbb{K}[X]$. Alors P divise tous les A_k ssi $D \cdot \mathbb{K}[X] \subset P \cdot \mathbb{K}[X]$ autrement dit ssi P divise D. Par ailleurs D divise chacun des A_k.Ceci justifie la définition suivante :

DEFINITION 9.3.2
Soient $A_1, \ldots, A_n \in \mathbb{K}[X]$.
On appelle PGCD des polynômes $A_1, \ldots, A_n$ tout générateur D de l'idéal $A_1\mathbb{K}[X] + \cdots + A_n\mathbb{K}[X]$.

Un polynôme D est un PGCD des polynômes $A_1, \ldots, A_n$ ssi

$$\begin{cases} \forall k, \exists Q_k \in \mathbb{K}[X] \quad A_k = DQ_k \\ \exists (U_1, \ldots, U_n) \in \mathbb{K}[X]^n \quad D = A_1U_1 + \cdots + A_nU_n \end{cases}$$

C'est aussi un diviseur commun des A_k de degré maximum.

DEFINITION 9.3.3
Soient $A_1, \ldots, A_n \in \mathbb{K}[X]$.
On appelle PPCM des polynômes $A_1, \ldots, A_n$ tout générateur M de l'idéal $\bigcap_{1 \leqslant k \leqslant n} A_k\mathbb{K}[X]$.

Un polynôme M est un PPCM des polynômes $A_1, \ldots, A_n$ ssi c'est un multiple commun des A_k de degré minimum.

Si A et B sont deux polynômes, notons $A \wedge B$ (resp. $A \vee B$) leur unique PGCD (resp. PPCM) normalisé. (On convient que 0 est normalisé). Les lois internes $\wedge$ et $\vee$ sont associatives et commutatives.

L'algorithme d'Euclide fournit un moyen explicite de déterminer le PGCD de deux polynômes A et B.

DEFINITION 9.3.4
On dit que les polynômes $A_1, \ldots, A_n$ sont premiers entre eux (on précise quelquefois dans leur ensemble) si leur PGCD est égal à 1. (On devrait plutôt dire si 1 est un de leur PGCD).
On dit que les polynômes $A_1, \ldots, A_n$ sont premiers entre eux deux à deux si pour tout couple (i, j) d'entiers distincts compris entre 1 et n on a $A_i \wedge A_j = 1$.

THEOREME 9.3.3 (Bézout)
Les polynômes $A_1, \ldots, A_n$ sont premiers entre eux ssi il existe des polynômes $U_1, \ldots U_n$ tels que $A_1U_1 + \cdots A_nU_n = 1$.

Comme dans le cas des entiers, pour deux polynômes A et B premiers entre eux l'algorithme d'Euclide permet de trouver les polynômes U et V tels que $AU + BV = 1$.

Exemple
Soient dans $\mathbb{Q}[X]$, $A = X^4 + X^3 + 1$ et $B = X^2 + X + 1$. La division euclidienne de $R_0 = A$ par $R_1 = B$ donne

$$R_0 = R_1Q_1 + R_2 \quad \text{avec} \quad Q_1 = (X^2 - 1) \quad R_2 = X + 2$$

Ensuite la division euclidienne de R_1 part R_2 s'écrit

$$R_1 = Q_2 R_2 + R_3 \quad \text{avec} \quad Q_2 = (X-1) \quad R_3 = 3$$

On en déduit que $A \wedge B = 1$ puis

$$3 = R_1 - Q_2 R_2 = R_1 - Q_2(R_0 - R_1 Q_1) = B(1 + Q_1 Q_2) - AQ_2 = (X^3 - X^2 - X + 2)B - (X-1)A$$

THEOREME 9.3.4 (Théorème de Gauss)
Soient $A, B, C \in \mathbb{K}[X]$. Si A divise BC et A est premier avec B alors A divise C.

preuve
Il existe U et V dans $\mathbb{K}[X]$ tels que $AU + BV = 1$. A divise AUC et BC donc aussi VBC donc $AUC + VBC = C$.■.

THEOREME 9.3.5
Soient $A, B_1, \ldots, B_n \in \mathbb{K}[X]$. Si A est premier avec B_k pour tout k, alors A est premier avec $B_1 B_2 \cdots B_n$.

THEOREME 9.3.6
*Si $A_1, \ldots, A_n$ sont des polynômes **premiers entre eux deux à deux** alors un PPCM des A_k est $A_1 A_2 \cdots A_n$.*

COROLLAIRE 9.3.1
*Si $A_1, \ldots, A_n$ sont des polynômes **premiers entre eux deux à deux** et si B est un polynôme divisible par A_k pour tout k, alors B est divisible par le produit $A_1 A_2 \cdots A_n$.*

THEOREME 9.3.7
Soient $A_1, \ldots, A_n$ et D des polynômes tels que D divise chacun des A_k. Alors D est un PGCD des A_k ssi les polynômes $A_k/D,\ 1 \leqslant k \leqslant n$ sont premiers entre eux.

Plus généralement, si P est un polynôme quelconque, un PGCD de $(PA_1, \ldots, PA_n)$ est $P\times$ un PGCD de $(A_1, \ldots, A_n)$.

THEOREME 9.3.8 (Théorème de Bézout amélioré)
Soient A, B deux polynômes non constants premiers entre eux. Il existe un unique couple de polynômes U, V vérifiant

$$AU + BV = 1 \qquad \deg(U) < \deg(B) \qquad \deg(V) < \deg(A)$$

première démonstration
On sait qu'il existe un couple (U_0, V_0) tels que $AU_0 + BV_0 = 1$. Soit $(U, V) \in \mathbb{K}[X]^2$ tel que $AU + BV = 1$. On a $A(U - U_0) = B(V_0 - V)$. D'après le théorème de Gauss, A étant premier avec B doit diviser $V_0 - V$. Il existe donc un polynôme D tel que $V_0 - V = AD$. En reportant, on obtient, A étant non nul, $U - U_0 = BD$. Réciproquement, pour tout polynôme D, le couple (U, V) avec $U = U_0 + BD$ et $V = V_0 - AD$ vérifie $AU + BV = 1$.
L'identité de la division euclidienne de U_0 par B (non nul) s'écrit $U_0 = BQ + R$. Comme U_0 et B sont aussi premiers entre eux par le théorème de Bézout, $R \neq 0$, donc $\deg(R) < \deg(B)$. Prenons $D = -Q$; alors $U = R$ et $V = V_0 + AQ$. On a $AU + BV = 1$ et $\deg(U) < \deg(B)$. Comme A et B sont non constants, $U \neq 0$ et $\deg(AU) \geqslant 1$. Alors $BV = 1 - AU \Rightarrow \deg(BV) = \deg(AU)$ donc $\deg(B) + \deg(V) \leqslant \deg(A) + \deg(U) < \deg(A) + \deg(B)$ d'où $\deg(V) < \deg(A)$. L'unicité est facile.

deuxième démonstration
Soient $p = \deg(A),\ q = \deg(B)$. Considérons le système de polynômes $S = (A, XA, X^2A, \ldots, X^{q-1}A, B, XB, \ldots, X^{p-1}B)$. Il est formé de $p + q$ polynômes appartenant tous à $\mathbb{K}_{p+q-1}[X]$. Nous allons montrer que ce système est libre. Il en résultera que c'est une base de $\mathbb{K}_{p+q-1}[X]$ donc que le polynôme 1 peut s'écrire de manière unique sous la forme $1 = u_0 A + u_1 XA + \cdots + u_{q-1}X^{q-1}A + v_0 B + v_1 XB + \cdots + v_{p-1}X^{p-1}B$ soit encore $1 = UA + VB$ avec $U = u_0 + \cdots + u_{q-1}X^{q-1}$ et $V = v_0 + \cdots + v_{p-1}X^{p-1}$. Soient donc $\lambda_0, \ldots, \lambda_{q-1}, \mu_0, \ldots, \mu_{p-1}$ des éléments de $\mathbb{K}$ tels que $\sum_{j=0}^{q-1} \lambda_j X^j A + \sum_{k=0}^{p-1} \mu_k X^k B = 0$.
Cette relation s'écrit $LA = -MB$ en posant $L = \lambda_0 + \cdots \lambda_{q-1}X^{q-1}$ et $M = \mu_0 + \cdots + \mu_{p-1}X^{p-1}$. Supposons par exemple l'un des λ_j non nul. Alors B divise LA et est premier avec A donc d'après le théorème de Gauss divise L. Mais c'est impossible car $\deg(L) < \deg(B)$. Donc tous les λ_j sont nuls, $L = 0$ et par conséquent $M = 0$ et tous les μ_k sont nuls.■.

Remarque importante
Soient $\mathbb{L}$ un corps et $\mathbb{K}$ un sous corps de $\mathbb{L}$. Soient A, B des polynômes à coefficients dans $\mathbb{K}$. On peut aussi les considérer comme des polynômes à coefficients dans $\mathbb{L}$. Soient $(Q, R) \in \mathbb{K}[X]^2$ le quotient et le reste de la division euclidienne de A par B dans $\mathbb{K}[X]$. En raison de l'unicité de la division euclidienne dans $\mathbb{L}[X]$, Q et R sont aussi le quotient et le reste de la division euclidienne de A par B dans $\mathbb{L}[X]$. Il en résulte que les PGCD et PPCM **normalisés** d'une famille finie de polynômes de $\mathbb{K}[X]$ sont les mêmes dans $\mathbb{K}[X]$ et dans $\mathbb{L}[X]$.
Mais si $D \in \mathbb{K}[X]$ est le PGCD normalisé des polynômes $A, B \in \mathbb{K}[X]$ par exemple, λD est aussi un PGCD de A et B dans $\mathbb{L}[X]$ pour tout $\lambda \in \mathbb{L}$, y compris si $\lambda \notin \mathbb{K}$.

Par exemple, si A et B sont des polynômes à coefficients entiers, ils sont premiers entre eux dans $\mathbb{Q}[X]$ ssi ils sont premiers entre eux dans $\mathbb{C}[X]$.

Cette remarque a d'importantes conséquences (multiplicité des racines d'un polynôme irréductible, voir plus loin).

9.3.3 Polynômes irréductibles

Soit toujours $\mathbb{K}$ un corps commutatif.

DEFINITION 9.3.5
On dit que $P \in \mathbb{K}[X]$ est irréductible si il est non constant et si ses seuls diviseurs dans $\mathbb{K}[X]$ sont les polynômes constants non nuls et les polynômes associés à P, i.e. les polynômes $\lambda P,\ \lambda \in \mathbb{K}^*$.

Deux polynômes irréductibles P et Q sont soit associés soit premiers entre eux. On remarquera que la notion de polynôme irréductible fait intervenir de manière essentielle le corps $\mathbb{K}$. Par exemple, le polynôme $X^2 - 2$ est irréductible dans $\mathbb{Q}[X]$ mais ne l'est pas dans $\mathbb{R}[X]$.
On notera $\mathcal{P}_{\mathbb{K}}$ l'ensemble des polynômes à coefficients dans $\mathbb{K}$, irréductibles et normalisés. Pour tout $a \in \mathbb{K}$, on a $X - a \in \mathcal{P}_{\mathbb{K}}$.

Soit $F \in \mathbb{K}[X] \setminus \{0\}$. Pour $P \in \mathcal{P}_{\mathbb{K}}$ on pose $\nu_P(F) = \max\{m \in \mathbb{N}\ ;\ P^m | F\}$. Ceci est toujours défini car $\{m \in \mathbb{N}\ ;\ P^m | F\}$ est un ensemble non vide (il contient 0) et majoré (par des considérations sur le degré) d'entiers naturels. Il a donc un plus grand élément.

THEOREME 9.3.9
Tout polynôme P non constant admet un diviseur irréductible.

THEOREME 9.3.10
Soit $F \in \mathbb{K}[X]$ un polynôme non nul et a le coefficient de son terme de plus haut degré.
1) L'ensemble $\{P \in \mathcal{P}_{\mathbb{K}}\ ;\ \nu_P(F) \geqslant 1\}$ est fini.
2) On a $F = a \prod_{P \in \mathcal{P}_{\mathbb{K}}} P^{\nu_P(F)}$.
3) Cette écriture est unique, i.e. si $F = k \prod_{P \in \mathcal{P}_{\mathbb{K}}} P^{k_P}$ où (k_P) est une famille d'entiers naturels à support fini indexée par $\mathcal{P}_{\mathbb{K}}$, on a $k = a$ et pour tout $P \in \mathcal{P}_{\mathbb{K}},\ k_P = \nu_P(F)$.

THEOREME 9.3.11
1) Soient $F, G \in \mathbb{K}[X] \setminus \{0\}$. On a $F|G \Leftrightarrow \forall P \in \mathcal{P}_{\mathbb{K}},\ \nu_P(F) \leqslant \nu_P(G)$

2) Soient $F_1, \ldots, F_n \in \mathbb{K}[X] \setminus \{0\}$.

- *Le PGCD normalisé des F_k est $\prod_{P \in \mathcal{P}_{\mathbb{K}}} P^{\min(\nu_P(F_i))}$.*
- *Le PPCM normalisé des F_k est $\prod_{P \in \mathcal{P}_{\mathbb{K}}} P^{\max(\nu_P(F_i))}$*

Les démonstrations de tous ces résultats sont analogues aux démonstrations des énoncés correspondants dans $\mathbb{Z}$.

THEOREME 9.3.12
L'ensemble $\mathcal{P}_{\mathbb{K}}$ est infini.

preuve
Si le corps $\mathbb{K}$ est infini, le résultat est trivial car tous les polynômes $X-a,\ a\in\mathbb{K}$ sont dans $\mathcal{P}_{\mathbb{K}}$. Dans le cas général, on fait comme pour l'ensemble des nombres premiers. On suppose $\mathcal{P}_{\mathbb{K}}$ fini et on considère le polynôme $A=\prod_{P\in\mathcal{P}_{\mathbb{K}}}P+1$.
Aucun des $P\in\mathcal{P}_{\mathbb{K}}$ ne peut diviser A et A admet au moins un facteur irréductible Q. Contradiction.

9.3.4 Anneaux quotients de $\mathbb{K}[X]$

Soit I un idéal non nul de $\mathbb{K}[X]$. On se propose d'étudier l'anneau quotient $\mathbb{K}[X]/I$. On peut choisir un générateur normalisé P de cet idéal de sorte que $I=P\cdot\mathbb{K}[X]$. On notera pour abréger $I=(P)$ et $A_P=\mathbb{K}[X]/(P)$. Soit $p=\deg(P)$. Si $p=0$, $P=1,\ I=\mathbb{K}[X]$ et $A_1=\{0\}$. Nous écarterons ce cas dans la suite. On supposera donc $p\geqslant 1$ et on écrira $P=a_0+a_1X+\cdots+a_{p-1}X^{p-1}+X^p$.
Soit $\pi:\mathbb{K}[X]\to\mathbb{K}[X]/(P)$ la surjection canonique. C'est un morphisme d'anneaux de noyau (P). La restriction π_p de π au sous espace $\mathbb{K}_{p-1}[X]$ formé des polynômes de degré inférieur ou égal à $p-1$ est injective. Montrons qu'elle est surjective. Soit $\alpha\in\mathbb{K}[X]/(P)$. Il existe un polynôme A tel que $\alpha=\pi(A)$. La division euclidienne de A par P s'écrit $A=PQ+R$ avec $R\in\mathbb{K}_{p-1}[X]$. Comme $\pi(PQ)=0$ on a $\alpha=\pi(A)=\pi(R)$. π_p est donc une bijection de $\mathbb{K}_{p-1}[X]$ sur $\mathbb{K}[X]/(P)$.

La restriction de π_p au sous espace $\mathbb{K}_0[X]$ identifié à $\mathbb{K}$ des polynômes constants est un isomorphisme du corps $\mathbb{K}$ sur un sous corps $\pi(\mathbb{K})$ de $\mathbb{K}[X]/(P)$. On identifiera dans la suite un élément $\lambda\in\mathbb{K}$ et son image par π dans $\mathbb{K}[X]/(P)$. Il en résulte que $\mathbb{K}[X]/(P)$ a une structure naturelle de $\mathbb{K}$-espace vectoriel pour laquelle la projection π est une application linéaire.

Notons enfin x la classe du polynôme X, i.e. $x=\pi(X)$. Il résulte de tout ce qui précède que tout élément $\alpha\in\mathbb{K}[X]/(P)$ s'écrit de manière unique sous la forme $\alpha=\lambda_0+\lambda_1x+\cdots+\lambda_{p-1}x^{p-1}$ où $\lambda_0,\ldots,\lambda_{p-1}$ sont p éléments de $\mathbb{K}$.

THEOREME 9.3.13
Soit $P\in\mathbb{K}[X]$ un polynôme de degré $p\geqslant 1$, $(P)=P\cdot\mathbb{K}[X]$ l'idéal principal engendré par P. L'ensemble quotient $\mathbb{K}[X]/(P)$ a une structure de $\mathbb{K}$-algèbre associative unitaire pour laquelle la projection canonique π est un morphisme d'algèbre. De plus, l'espace vectoriel $\mathbb{K}[X]/(P)$ est de dimension finie égale au degré p du polynôme P et le système $(1,x,\ldots x^{p-1})$ où x est la classe du polynôme X, est une base de $\mathbb{K}[X]/(P)$.

THEOREME 9.3.14
Soit $P\in\mathbb{K}[X]$ et (P) l'idéal principal engendré par P. Les propriétés suivantes sont équivalentes :
(1) *P est irréductible.*
(2) *L'anneau quotient $\mathbb{K}[X]/(P)$ est intègre.*
(3) *L'anneau quotient $\mathbb{K}[X]/(P)$ est un corps.*

preuve
(1) $\Rightarrow$ (3) Soit $\alpha\in\mathbb{K}[X]/(P),\ \alpha\neq\bar{0}$. Soit $Q\in\mathbb{K}[X]$ un représentant de α. Puisque α n'est pas nul, Q n'est pas multiple de P. Comme P est irréductible, Q est premier avec P, donc il existe deux polynômes U et V tels que $UP+VQ=1$ d'où en prenant les classes modulo P, $\bar{V}\bar{Q}=\bar{1}$ ce qui prouve que α est inversible.
(3) $\Rightarrow$ (2) Trivial
(2) $\Rightarrow$ (1) Supposons P non irréductible. Il existe deux polynômes A et B de degré $\geqslant 1$ tels que $P=AB$ ce qui donne $\bar{A}\bar{B}=\bar{0}$. Mais A et B étant de degré $<\deg(P)$ ne sont pas multiples de P, donc $\bar{A}\neq\bar{0}$ et $\bar{B}\neq\bar{0}$.
Remarquons que le théorème 9.1.1 prouve directement l'implication (2) $\Rightarrow$ (3).

EXEMPLE 9.3.1
Soit $\mathbb{K}=\mathbb{R}$ et $P=X^2+1$. P est irréductible dans $\mathbb{R}[X]$ donc $\mathbb{R}[X]/(X^2+1)$ est un corps. Ce corps est un $\mathbb{R}$-espce vectoriel de dimension 2, de base $(1,x)$ où x est la classe de X et x vérifie $x^2+1=0$. Ceci fournit une construction du corps $\mathbb{C}$ des nombres complexes.

9.4 Fonctions polynômes

9.4.1 Fonction polynôme définie sur $\mathbb{K}$

Dans cette section $\mathbb{K}$ est un corps commutatif fixé.

A tout polynôme $P = \sum_{0 \leqslant k \leqslant p} a_k X^k \in \mathbb{K}[X]$ on associe la fonction $\tilde{P} : \mathbb{K} \to \mathbb{K}$ qui à $x \in \mathbb{K}$ associe $\tilde{P}(x) = \sum_{k=0}^{k=p} a_k x^k$.
La fonction $\tilde{P}$ est dite fonction polynôme associée au polynôme P.

THEOREME 9.4.1
L'application $P \to \tilde{P}$ est un morphisme de l'anneau $\mathbb{K}[X]$ dans l'anneau $\mathbb{K}^{\mathbb{K}}$ des fonctions de $\mathbb{K}$ dans $\mathbb{K}$.

Autrement dit, pour tous $P, Q \in \mathbb{K}[X]$ on a $\widetilde{P+Q} = \tilde{P} + \tilde{Q}$, $\widetilde{PQ} = \tilde{P}\tilde{Q}$ et $\tilde{\mathbf{1}}$ est la fonction constante égale à 1.
Notons aussi que $\widetilde{P \circ Q} = \tilde{P} \circ \tilde{Q}$ et que $\tilde{X} = id_{\mathbb{K}}$.

THEOREME 9.4.2
Soient $P \in \mathbb{K}[X],\ P \neq 0$ et $b \in \mathbb{K}$. $\tilde{P}(b)$ est le reste de la division euclidienne de P par $X - b$.
Autrement dit, il existe un polynôme Q tel que $P = (X - b)Q + \tilde{P}(b)$.

preuve
Soit $P = \sum_{k=0}^{k=p} a_k X^k$. On a dans $\mathbb{K}[X]$ l'identité remarquable

$$X^k - b^k = (X - b)(X^{k-1} + bX^{k-2} + \cdots + b^{k-2}X + b^{k-1})$$

En multipliant par a_k et en sommant les égalités obtenues de $k = 0$ à $k = p$ on obtient l'égalité voulue avec $Q(X) = \sum_{k=1}^{k=p} a_k \left(X^{k-1} + bX^{k-2} + \cdots + b^{k-2}X + b^{k-1}\right)$.

9.4.2 Racines d'un polynôme

DEFINITION 9.4.1
Un élément $a \in \mathbb{K}$ est une racine ou un zéro du polynôme P si $\tilde{P}(a) = 0$.

Il résulte du dernier théorème que $a \in \mathbb{K}$ est une racine de P ssi $(X - a)$ divise P.

DEFINITION 9.4.2
Soit P un polynôme non constant et $a \in \mathbb{K}$ une racine de P. On appelle ordre de multiplicité de la racine a l'entier $\nu_{X-a}(P)$ c'est à dire l'unique entier m tel que P soit divisible par $(X - a)^m$ et ne soit pas divisible par $(X - a)^{m+1}$.

Donc a est racine multiple d'ordre m de P ssi P s'écrit $P(X) = (X - a)^m H(X)$ avec $\tilde{H}(a) \neq 0$.

THEOREME 9.4.3
Soit $P \in \mathbb{K}[X]$ un polynôme de degré p. Soient $a_1, \ldots, a_r$ r éléments distincts de $\mathbb{K}$ et $m_1, \ldots, m_r$ des entiers naturels non nuls. Si pour tout j entre 1 et r P admet a_j comme racine ave multiplicité m_j, P est divisible par $(X - a_1)^{m_1} \cdots (X - a_r)^{m_r}$.
En particulier on a $m_1 + \cdots + m_r \leqslant \deg(P)$.

preuve
Par hypothèse, P est divisible par $(X - a_j)^{m_j}$ pour $j = 1, \ldots, r$. Or ces polynômes sont premiers entre eux deux à deux. D'où le résultat.

COROLLAIRE 9.4.1
Si un polynôme de degré p admet $n > p$ racines distinctes, c'est le polynôme nul.

COROLLAIRE 9.4.2
*Soit $\mathbb{K}$ un corps commutatif **infini**. L'application $P \to \tilde{P}$ est injective, autrement dit, si un polynôme P est tel que la fonction polynôme associée est identiquement nulle, ce polynôme est nul.*

Ce résultat tombe en défaut si $\mathbb{K}$ est un corps fini. Par exemple, si $\mathbb{K} = \mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ où p est un nombre premier, le polynôme $P = X^p - X$ n'est pas le polynôme nul, mais la fonction polynôme associée $\tilde{P}$ est identiquement nulle car $\forall x \in \mathbb{F}_p,\ x^p = x$.

9.4.3 Polynômes scindés sur $\mathbb{K}$

DEFINITION 9.4.3
Un polynôme non constant $P \in \mathbb{K}[X]$ est dit scindé (sur $\mathbb{K}$) si il s'écrit comme produit de polynômes du premier degré. Si P est un polynôme scindé de degré p, on appelle liste des racines de P tout élément $(x_1, \dots, x_p)$ de $\mathbb{K}^p$ tel que

$$P = a_p \prod_{k=1}^{k=p} (X - x_k)$$

On prendra garde de ne pas confondre l'ensemble des racines de P et une liste des racines de P. Si $P = X^p$, l'ensemble des racines de P est $\{0\}$ alors qu'une liste des racines de P est $(\underbrace{0, \dots, 0}_{p \text{ fois}})$. Si P admet p racines disctinctes, il y a $p!$ listes de racines de P. Si P admet les racines distinctes $y_1, \dots, y_r$ avec les multiplicités respectives $m_1, \dots m_r$, une liste des racines de P est $(\underbrace{y_1, \dots, y_1}_{m_1 \text{ fois}}, \underbrace{y_2, \dots, y_2}_{m_2 \text{ fois}} \dots, \underbrace{y_r, \dots, y_r}_{m_r \text{ fois}})$.

On définit des applications $\sigma_k : \mathbb{K}^p \to \mathbb{K}$ pour $1 \leqslant k \leqslant p$ en posant

$$\sigma_k(x_1, \dots, x_p) = \sum_{1 \leqslant i_1 < i_2 < \dots < i_k \leqslant p} x_{i_1} x_{i_2} \cdots x_{i_k}$$

On a donc

$$\sigma_1(x_1, \dots, x_p) = x_1 + x_2 + \cdots + x_p = \sum_{k=1}^{k=p} x_k$$
$$\sigma_2(x_1, \dots, x_p) = x_1x_2 + x_1x_3 + \cdots x_1x_p + x_2x_3 + \cdots + x_2x_p + \cdots + x_{p-1}x_p = \sum_{1 \leqslant i < j \leqslant p} x_i x_j$$
$$\sigma_p(x_1, \dots, x_p) = x_1x_2 \cdots x_p = \prod_{1 \leqslant k \leqslant p} x_k$$

DEFINITION 9.4.4 (Fonctions symétriques élémentaires)
Les fonctions $\sigma_1, \dots, \sigma_p$ de $\mathbb{K}^p$ dans $\mathbb{K}$ sont appellées fonctions symétriques élémentaires de p variables.

Cette terminologie est justifiée par le fait que pour toute permutation $\varphi \in \mathfrak{S}_p$ on a $\sigma_k(x_{\varphi(1)}, \dots, x_{\varphi(p)}) = \sigma_k(x_1, \dots, x_p)$

THEOREME 9.4.4 (Relations coefficients racines)
Soit $P = \sum_{0 \leqslant k \leqslant p} a_k X^k$ un polynôme scindé de degré $p \geqslant 1$ et $(x_1, \dots, x_p)$ une liste des racines de P. On a pour k compris entre 1 et p

$$\sigma_k(x_1, \dots, x_p) = (-1)^k \frac{a_{p-k}}{a_p}$$

ou encore, en posant $\sigma_0(x_1, \cdots, x_n) = 1$,

$$P(X) = a_p \left(\sum_{k=0}^{k=p} (-1)^{p-k} \sigma_{p-k}(x_1, \dots, x_p) X^k \right) = a_p \left(X^p - (x_1 \cdots + x_p) X^{p-1} + \cdots + (-1)^p x_1 \cdots x_p \right)$$

La preuve se fait par récurrence sur p en utilisant l'égalité $P = a_p(X - x_1) \cdots (X - x_p)$.

9.4.4 Corps algébriquement clos

THEOREME 9.4.5
Soit $\mathbb{K}$ un corps commutatif. Les propriétés suivantes sont équivalentes.
(1) *Les seuls polynomes irréductibles de $\mathbb{K}[X]$ sont les polynômes de degré 1.*
(2) *Tout polynôme non constant à coefficients dans $\mathbb{K}$ admet au moins une racine dans $\mathbb{K}$.*
(3) *Tout polynôme à coefficients dans $\mathbb{K}$ non constant est scindé sur $\mathbb{K}$.*

La preuve facile est laissée au lecteur.

DEFINITION 9.4.5
On dit qu'un corps $\mathbb{K}$ est algébriquement clos si il vérifie une des popriétés ci dessus (et donc les trois) .

THEOREME 9.4.6 (D'alembert-Gauss)
$\mathbb{C}$ *est algébriquement clos.*

Ce théorème est prouvé dans le cours d'analyse.

9.4.5 Polynômes irréductibles sur $\mathbb{R}$

Puisque $\mathbb{C}$ est algébriquement clos, tout polynôme est scindé dans $\mathbb{C}[X]$ donc s'écrit sous la forme $P = a_n \prod_{k=1}^{k=n} (X - z_k)$ où $a_n \in \mathbb{C}$ et où les z_k sont les racines de P.
Tout polynôme à coefficients réels peut être considéré comme un polynôme à coefficients complexes et est donc scindé sur $\mathbb{C}$. $\mathbb{R}$ et $\mathbb{C}$ étant des corps infinis, nous noterons dans ce paragraphe par la même lettre un polynôme P et la fonction polynôme $\tilde{P}$ associée.

Lemme 9.4.1
Soient $P \in \mathbb{R}[X]$ et $z \in \mathbb{C} \setminus \mathbb{R}$. Si z est racine de P avec la multiplicité m, son conjugué $\bar{z}$ est aussi racine de P avec la même multiplicité.

preuve
Pour tout polynôme $Q(X) = q_r X^r + \cdots + q_0 \in \mathbb{C}[X]$, définissons un polynôme $\overline{Q}$ par $\overline{Q}(X) = \overline{q_r} X^r + \cdots + \overline{q_0}$. Il est facile de vérifier que $\overline{Q_1 + Q_2} = \overline{Q_1} + \overline{Q_2}$, $\overline{Q_1 Q_2} = \overline{Q_1}\,\overline{Q_2}$ que Q est à coefficients réels ssi $\overline{Q} = Q$ et enfin que pour tout $u \in \mathbb{C}$ on a $\overline{Q(u)} = \overline{Q}(\overline{u})$

Soit $P = a_n X^n + \cdots + a_0$ où les a_k sont réels. Par hypothèse, il existe un polynôme $H = \alpha_{n-m} X^{n-m} + \cdots + \alpha_0 \in \mathbb{C}[X]$ tel que $P(X) = (X - z)^m H(X)$ avec $H(z) \neq 0$. On en déduit $P(X) = \overline{P}(X) = (X - \overline{z})^m \overline{H}(X)$ avec $\overline{H}(\overline{z}) = \overline{H(z)} \neq 0$ ce qui prouve le lemme.

THEOREME 9.4.7
Les polynômes irréductibles dans $\mathbb{R}[X]$ sont les polynômes du premier degré et les polynômes du second degré $aX^2 + bX + c$ vérifiant $\Delta = b^2 - 4ac < 0$.

preuve
Tout d'abord, un polynôme P du second degré de discrimant strictement négatif est irréductible car un diviseur de P qui n'est ni constant, ni associé à P est de degré 1. L'existence d'un tel diviseur impliquerait que P ait une racine réelle, ce qui n'est pas.
Soit P un polynôme irréductible sur $\mathbb{R}$ de degré $p > 1$. P n'a pas de racines réelles sinon, il serait divisible par un polynôme de degré 1. Soit $z \in \mathbb{C}$ une racine de P. On a $z \notin \mathbb{R}$, donc $\overline{z}$ est aussi racine de P. Comme $z \neq \overline{z}$, P est divisible par le polynôme $Q = (X - z)(X - \overline{z}) = X^2 - (z + \overline{z})\, X + z\overline{z} \in \mathbb{R}[X]$. Comme P est irréductible il est associé à Q. ■

COROLLAIRE 9.4.3
Tout polynôme $P \in \mathbb{R}[X] \setminus \{0\}$ se factorise sous la forme

$$P(X) = a \prod_{j=1}^{j=r} (X - b_j)^{m_j} \prod_{k=1}^{k=s} (X^2 + p_k X + q_k)^{n_k}$$

où les b_j sont des réels distincts, les (p_k, q_k) des couples de réels distincts vérifiant $p_k^2 - 4q_k < 0$ et $a \in \mathbb{R}^*$. De plus on a

$$\deg(P) = \sum_{j=1}^{j=r} m_j + 2 \sum_{k=1}^{k=s} n_k$$

9.5 Dérivation - Formule de Taylor

DEFINITION 9.5.1
Soit $\mathbb{K}$ un corps commutatif. On appelle polynôme dérivé du polynôme $P(X) = a_pX^p + \cdots a_0 = \sum_{k=0}^{k=p} a_kX^k$ le polynôme $P'(X) = pa_pX^{p-1} + \cdots + a_1 = \sum_{k=0}^{p-1}(k+1)a_{k+1}X^k$.

Il faut bien noter dans cette formule que $k.a_k$ signifie $\underbrace{a_k + \cdots a_k}_{k \text{ fois}}$. En particulier, la dérivée du polynôme $X^p \in \mathbb{F}_p[X]$ est le polynôme nul.

On définit ensuite la dérivée seconde, etc ... d'un polynôme P comme étant $D \circ D(P)$, etc .... On note $P^{(k)}$ la dérivée k-ième de P avec par convention $P^{(0)} = P$.

THEOREME 9.5.1
La dérivation D est une application $\mathbb{K}$-linéaire de $\mathbb{K}[X]$ dans lui même vérifiant pour tous $P, Q \in \mathbb{K}[X]$, $(PQ)' = P'Q + PQ'$ et $(P \circ Q)' = (P' \circ Q) \cdot Q'$. La dérivée d'un polynôme de degré p est un polynôme de degré inférieur ou égal à $p-1$. La dérivée d'un polynôme constant est le polynôme nul.

Si le corps $\mathbb{K}$ est de caractéristique 0, la dérivée d'un polynôme de degré $p \geqslant 1$ est un polynôme de degré $p-1$.

Dans toute la suite le corps $\mathbb{K}$ est supposé de caractéristique 0. Il est donc infini et on sait qu'alors il y a bijection entre les polynômes et les fonctions polynômes de sorte que nous désigneros par la même lettre un polynôme P et la fonction polynôme associée.

THEOREME 9.5.2 (Formule de Taylor)
Soient $\mathbb{K}$ un corps de caractéristique nulle, $a \in \mathbb{K}$ et $P \in \mathbb{K}_p[X]$. On a

$$(1) \qquad P(X+a) = \sum_{k=0}^{k=p} P^{(k)}(a)\frac{X^k}{k!} = P(a) + P'(a)X + \cdots + P^{(p)}(a)\frac{X^p}{p!}$$

$$(2) \qquad P(X+a) = \sum_{k=0}^{k=p} \frac{a^k}{k!}P^{(k)}(X) = P(X) + aP'(X) + \cdots + \frac{a^p}{p!}P^{(p)}(X)$$

Notons que la première relation s'écrit aussi

$$(3) \qquad P(X) = \sum_{k=0}^{k=p} P^{(k)}(a)\frac{(X-a)^k}{k!} = P(a) + P'(a)(X-a) + \cdots + P^{(p)}(a)\frac{(X-a)^p}{p!}$$

preuve
Par $\mathbb{K}$ linéarité, il suffit de montrer ces relations lorsque P est un monôme : $P = X^p$. On a

$$P(X+a) = (X+a)^p = \sum_{k=0}^{k=p} \left(\begin{array}{c} p \\ k \end{array} \right) X^k a^{p-k} = \sum_{k=0}^{k=p} [p(p-1)\cdots(p-k+1)]a^{p-k}\frac{X^k}{k!} = \sum_{k=0}^{k=p} P^{(k)}(a)\frac{X^k}{k!}$$

$$P(X+a) = (X+a)^p = \sum_{k=0}^{k=p} \left(\begin{array}{c} p \\ k \end{array} \right) X^{p-k} a^k = \sum_{k=0}^{k=p} [p(p-1)\cdots(p-k+1)]X^{p-k}\frac{a^k}{k!} = \sum_{k=0}^{k=p} \frac{a^k}{k!}P^{(k)}(X)$$

Remarque importante. On peut se demander où, dans cette preuve, on utilise le fait que $\mathbb{K}$ est de caractéristique nulle. Il ne faut pas oublier dans ces formules que la multiplication par l'entier naturel C_n^k est une loi externe sur $\mathbb{K}$. On a dans $\mathbb{Q}$, $C_n^k = [n(n-1)\cdots(n-k+1)]/k!$. Comme $\mathbb{K}$ est de caractéristique nulle, le sous corps premier de $\mathbb{K}$ est isomorphe à $\mathbb{Q}$. L'élément $k! \cdot 1_{\mathbb{K}}$ de $\mathbb{K}$ est non nul. Il admet donc un inverse dans $\mathbb{K}$. On a alors $C_n^k \cdot 1_{\mathbb{K}} = [n(n-1)\cdots(n-k+1)] \cdot \frac{1}{k!1_{\mathbb{K}}}$.

COROLLAIRE 9.5.1
Soient $\mathbb{K}$ un corps de caractéristique 0, $a \in \mathbb{K}$, $P \in \mathbb{K}[X] \setminus \{0\}$. La multiplicité de a comme racine de P est l'entier ν défini par $P^{(\nu)}(a) \neq 0$ et $P^{(j)}(a) = 0$ pour $j \leqslant \nu - 1$.

Ici on convient de dire que si a n'est pas racine de P son ordre de multiplicité est 0. Avec cette convention, l'entier ν est toujours défini car le polynôme P étant non nul, on a $P^{(\deg(P))}(a) \neq 0$. La formule de Taylor sous la forme (3) s'écrit $P(X) = (X-a)^\nu H(X)$ avec $H = P^{(\nu)}(a) + \sum_{k=\nu+1}^{k=\deg(P)} P^{(k)}(a)\frac{(X-a)^{k-\nu}}{k!}$ et donc $H(a) \neq 0$ et donc ν est bien la multiplicité de a comme racine de P.

9.5.1 Eléments algébriques, éléments transcendants

Soit $\mathbb{L}$ un corps et $\mathbb{K}$ un sous corps de $\mathbb{L}$. On dit dans ce cas que $\mathbb{L}$ est une extension de $\mathbb{K}$; $\mathbb{L}$ est naturellement muni d'une structure de $\mathbb{K}$-algèbre, la loi externe étant la restriction à $\mathbb{K} \times \mathbb{L}$ de la multiplication dans $\mathbb{L}$.
Tout polynôme $P \in \mathbb{K}[X]$ peut être considéré comme appartenant à $\mathbb{L}[X]$ et définit une fonction $\tilde{P} : \mathbb{L} \to \mathbb{L}$ telle que $\tilde{P}(\mathbb{K}) \subset \mathbb{K}$.

DEFINITION 9.5.2
Un élément $x \in \mathbb{L}$ est dit algébrique sur $\mathbb{K}$ si il existe $P \in \mathbb{K}[X] \setminus \{0\}$ tel que $\tilde{P}(x) = 0$. Si ce n'est pas le cas, x est dit transcendant sur $\mathbb{K}$.

Soit $x \in \mathbb{L}$. L'application $e_x : \mathbb{K}[X] \to \mathbb{L},\ P \to \tilde{P}(x)$ est un morphisme de $\mathbb{K}$-algèbres : on a $e_x(1) = 1$ et pour tous $P, Q \in \mathbb{K}[X]$ et $\lambda \in \mathbb{K}$, $e_x(P + \lambda Q) = e_x(P) + \lambda e_x(Q)$ et $e_x(PQ) = e_x(P)e_x(Q)$ ce qui est une façon compliquée d'écrire que $\widetilde{(P + \lambda Q)}(x) = \tilde{P}(x) + \lambda\tilde{Q}(x)$ et $\widetilde{PQ}(x) = \tilde{P}(x)\tilde{Q}(x)$.
L'image de $\mathbb{K}[X]$ par e_x est une sous algèbre de $\mathbb{L}$ usuellement notée $\mathbb{K}[x]$. Un élément $z \in \mathbb{L}$ est dans $\mathbb{K}[x]$ ssi il peut s'écrire sous la forme $z = a_0 + a_1 x + \cdots + a_p x^p$ où les a_k sont des éléments de $\mathbb{K}$. C'est le plus petit sous anneau (et aussi la plus petite sous algèbre) de $\mathbb{L}$ contenant $\mathbb{K}$ et x.

L'application e_x est un morphisme d'anneaux. Son noyau $\ker(e_x)$ est donc un idéal de $\mathbb{K}[X]$ et x est algébrique ssi $\ker(e_x) \neq \{0\}$ et transcendant si $\ker(e_x) = 0$.

Si x est transcendant sur $\mathbb{K}$, l'algèbre $\mathbb{K}[x]$ est isomorphe via e_x à $\mathbb{K}[X]$. Si il est algébrique, le noyau de e_x est un idéal non nul de $\mathbb{K}[X]$ qui est donc engendré par un unique polynôme normalisé M. Pour tout $P \in \mathbb{K}[X]$ on a donc $\tilde{P}(x) = 0 \Leftrightarrow M|P$.

DEFINITION 9.5.3
Soient $\mathbb{L}$ une extension du corps $\mathbb{K}$ et x un élément de L algébrique sur $\mathbb{K}$. On appelle polynôme minimal de x le générateur normalisé M de l'idéal $\ker(e_x) = \{P \in \mathbb{K}[X] \mid \tilde{P}(x) = 0\}$. Le degré m du polynôme minimal M de x s'appelle degré de x.

Si $x \in \mathbb{K}$, x est algébrique sur $\mathbb{K}$ de polynôme minimal $X - x$, de degré 1.

THEOREME 9.5.3
Soient $\mathbb{L}$ une extension de $\mathbb{K}$ et $x \in \mathbb{L}$ algébrique sur $\mathbb{K}$, M son polynôme minimal et m son degré.
1) Le polynôme minimal M est irréductible.
2) $\mathbb{K}[x]$ est un sous corps de $\mathbb{L}$
3) $\mathbb{K}[x]$ est un $\mathbb{K}$-espace vectoriel de dimension finie égale au degré m de x. Une base de $\mathbb{K}[X]$ est $(1, x, \ldots, x^{m-1})$.

preuve
On applique le théorème de factorisation des morphismes d'anneaux à $e_x : \mathbb{K}[X] \to \mathbb{L}$. Le noyau de e_x est l'idéal principal (M). Il existe un isomorphisme $\theta : \mathbb{K}[X]/(M) \to \mathbb{K}[x]$ tel que $e_x = i \circ \theta \circ \pi$ où $\pi : \mathbb{K}[X] \to \mathbb{K}[X]/(M)$ est la projection canonique et i l'inclusion de $\mathbb{K}[x]$ dans $\mathbb{L}$. Comme $\mathbb{K}[x]$ est un sous anneau de $\mathbb{L}$, il est intègre. Il en est donc de même de $\mathbb{K}[X]/(M)$. Mais on a vu que ceci implique que M est irréductible et que $\mathbb{K}[X]/(M)$ est un corps. (théorème 9.3.14). Comme θ est un isomorphisme, $\mathbb{K}[x]$ est un corps.
Enfin, il est facile de voir que θ est $\mathbb{K}$-linéaire. Or $\mathbb{K}[X]/(M)$ est un $\mathbb{K}$-espace vectoriel de dimension $\deg(M)$ de base $(1, \overline{X}, \ldots, \overline{X}^{m-1})$ (théorème 9.3.13). L'image de cette base par θ est une base de $\mathbb{K}[x]$. Or $\theta\left(\overline{X}^k\right) = \theta \circ \pi\left(X^k\right) = x^k$. D'où le point 3).

COROLLAIRE 9.5.2
L'élément $x \in \mathbb{L}$ est algébrique sur $\mathbb{K}$ ssi $\mathbb{K}[x]$ est un $\mathbb{K}$-espace vectoriel de dimension finie.

preuve
On vient de voir que si x était algébrique $\mathbb{K}[x]$ était un $\mathbb{K}$-espace vectoriel de dimension finie et on a vu que si x n'était pas algébrique, $\mathbb{K}[x]$ était isomorphe comme $\mathbb{K}$-algèbre, donc comme $\mathbb{K}$-espace vectoriel, à $\mathbb{K}[X]$ qui est de dimension infinie.

La situation la plus fréquente est celle où $\mathbb{K} = \mathbb{Q}$ et $L = \mathbb{C}$. Par exemple, $\sqrt{2},\ i$ sont algébriques sur Q de degré 2, de polynômes minimaux respectifs $X^2 + 1$ et $X^2 - 2$.

THEOREME 9.5.4
L'ensemble des nombres complexes (resp. réels) algébriques sur $\mathbb{Q}$ est dénombrable.

preuve
Notons $\overline{\mathbb{Q}}$ l'ensemble des nombres complexes algébriques sur $\mathbb{Q}$.

On dira qu'un ensemble est au plus dénombrable si il est fini ou dénombrable.
Nous utiliserons les résultats suivants sur les ensembles dénombrables :
1) $\mathbb{Q}$ est dénombrable.
2) Tout sous ensemble d'un ensemble au plus dénombrable est au plus dénombrable.
3) Tout produit fini d'ensembles dénombrables est dénombrable. En particulier $\mathbb{Q}^n$ est dénombrable pour tout n.
4) Une réunion au plus dénombrable d'ensembles au plus dénombrables est un ensemble au plus dénombrable.

Pour un polynôme P on notera $R(P)$ le sous ensemble de $\mathbb{C}$ formé des racines de P.
Notons $\mathcal{P}_n$ l'ensemble des polynômes normalisés de degré n. On a une bijection $\mathbb{Q}^n \to \mathcal{P}_n$ définie par $(q_0, \ldots, q_{n-1}) \to q_0 + q_1 X \cdots + q_{n-1} X^{n-1}$. Donc $\mathcal{P}_n$ est dénombrable.

Soit E_n l'ensemble des nombres complexes z pour lesquels il existe un $P \in \mathcal{P}_n$ tel que $P(z) = 0$. On a donc $E_n = \bigcup_{P \in \mathcal{P}_n} R(P)$. E_n est une réunion dénombrable d'ensembles finis non vides, donc est au plus dénombrable.

Enfin, il est clair que $\overline{\mathbb{Q}} = \bigcup_{n \in \mathbb{N}^*} E_n$. C'est une réunion dénombrable d'ensembles au plus dénombrable. Donc $\overline{\mathbb{Q}}$ est au plus dénombrable. Or cet ensemble contient $\mathbb{Q}$ donc il n'est pas fini. Par conséquent, il est dénombrable.

Enfin l'ensemble des nombres réels algébriques sur $\mathbb{Q}$ est infini (il contient $\mathbb{Q}$) et est contenu dans l'ensemble dénombrable $\overline{\mathbb{Q}}$. Il est donc dénombrable.

COROLLAIRE 9.5.3
L'ensemble des nombres réels transcendants sur $\mathbb{Q}$ est infini non dénombrable.

preuve
Si cet ensemble, soit $\mathcal{T}$ était fini ou dénombrable, il en serait de même de $\mathbb{R}$ comme réunion de $\overline{\mathbb{Q}} \cap \mathbb{R}$ et de $\mathcal{T}$. Mais $\mathbb{R}$ n'est pas dénombrable. ■.

Les nombres π, e (base des logarithmes népériens) sont transcendants que $\mathbb{Q}$, mais la démonstration n'est pas triviale.

10

Action de groupes

10.1 Définitions

Soit X un ensemble non vide et G un groupe dont la loi de composition interne sera notée multiplicativement. On notera $\mathfrak{S}_X$ le groupe symétrique de X, c'est à dire le groupe dont les éléments sont les bijections de X et la loi interne la composition des applications.

DEFINITION 10.1.1
Une action A du groupe G sur X est une application $A : G \times X \to X$ notée $(g, x) \to A(g, x) = g * x$ vérifiant les deux propriétés suivantes :
(1) *Pour tous $g, g' \in G,\ x \in X,\ g * (g' * x) = (gg') * x$*
(2) *Pour tout $x \in X,\ 1_G * x = x$ où 1_G désigne l'élément neutre de G.*

Soit A une action de G sur X. Soit $g \in G$ et $\varphi_g : X \to X$ l'application $x \to \varphi_g(x) = g * x$. La propriété (1) s'écrit $\varphi_g \circ \varphi_{g'} = \varphi_{gg'}$ et la propriété (2), $\varphi_{1_G} = Id_X$. Il en résulte que si on note g^{-1} l'inverse dans G de $g \in G$ on a $\varphi_g \circ \varphi_{g^{-1}} = Id_X$ et $\varphi_{g^{-1}} \circ \varphi_g = Id_X$. Par conséquent, pour tout g, $\varphi_g \in \mathfrak{S}_X$. La relation $\varphi_g \circ \varphi_{g'} = \varphi_{gg'}$ montre que l'application $g \to \varphi_g$ est un morphisme du groupe G dans le groupe $\mathfrak{S}_X$. On notera Φ_A ce morphisme.

Inversement, soit $\Phi : G \to \mathfrak{S}_X$ un morphisme de groupes. Pour tout $g \in G$ et tout $x \in X$, posons $A(g, x) = \Phi(g)(x)$. Les propriétés (1) et (2) d'une action de groupe se vérifient immédiatement et pour cette action, on a $\Phi_A = \Phi$.

En conclusion, il est équivalent de se donner une action du groupe G sur l'ensemble X ou un morphisme du groupe G dans le groupe symétrique $\mathfrak{S}_X$ de X.

DEFINITION 10.1.2
On dit que l'action A du groupe G sur l'ensemble X est fidèle si le morphisme Φ_A associé est injectif.
Autrement dit, A est fidèle si et seulement si la relation $\forall x \in X, g * x = x$ implique $g = 1_G$.

DEFINITION 10.1.3
Soit A une action du groupe G sur X et $a \in X$. On dit que a est un point fixe (de l'action) si pour tout $g \in G$ on a $g * a = a$.

DEFINITION 10.1.4
Soit A une action du groupe G sur X. Une partie Y de X est stable par A si pour tout $y \in Y$ et tout $g \in G$ on a $g * y \in Y$.

Si Y est une partie stable par A, la restriction de A à $G \times Y$ définit une action de G sur Y appelée action induite.

10.2 Exemples

EXEMPLE 10.2.1
Un groupe G agit sur lui même par translation à gauche : $\forall g \in G, \forall x \in G, g * x = gx$. L'application Φ associée est celle qui à tout g de G associe la translation $g' \to gg'$. Cette action est évidemment fidèle. On obtient ainsi un isomorphisme du groupe

G sur un sous groupe de $\mathfrak{S}_G$.

Si G est un groupe fini, on peut, en choisissant une numérotation des éléments de G identifier $\mathfrak{S}_G$ au groupe de permutations $\mathfrak{S}_n$. On obtient ainsi le théorème de Cayley :

THEOREME 10.2.1
Tout groupe fini G de cardinal n est isomorphe à un sous groupe de $\mathfrak{S}_n$.

EXEMPLE 10.2.2
Soit G un groupe, H un sous groupe de G et $X = (G/H)_g$ l'ensemble des classes à gauche de G modulo H. G agit sur X par translation : un élément quelconque $\alpha \in X$ est un sous ensemble de G de la forme $\alpha = xH$ pour un $x \in G$. On pose pour $g \in G$, $g * \alpha = g\alpha = \{gu\ ; u \in \alpha\}$. On définit ainsi une action de G sur X (vérification facile).

EXEMPLE 10.2.3
$GL(n, \mathbb{K})$ agit sur $M_n(\mathbb{K})$ par $GL(n, K) \times M_n(\mathbb{K}) \ni (P, M) \to P * M = PMP^{-1}$.

EXEMPLE 10.2.4
Le groupe produit $GL(p, \mathbb{K}) \times GL(q, \mathbb{K})$ agit sur l'ensemble $M_{p,q}(\mathbb{K})$ par $(GL(p, \mathbb{K}) \times GL(q, \mathbb{K})) \times M_{p,q}(\mathbb{K}) \ni ((P, Q), M) \to (P, Q) * M = PMQ^{-1}$.

EXEMPLE 10.2.5
$GL(n, \mathbb{K})$ agit sur l'ensemble $S_n(\mathbb{K})$ des matrices symétriques d'ordre n à coefficients dans $\mathbb{K}$ par $(P, S) \to P * S = PS\,{}^tP$.

EXEMPLE 10.2.6
Soit E un $\mathbb{K}$-ev et $FBS(E)$ l'ensemble des formes bilinéaires symétriques sur E. Soient $\varphi \in FBS(E)$ et $h \in GL(E)$. L'application qui au couple $(x, y) \in E \times E$ associe $\varphi\left(h^{-1}(x), h^{-1}(y)\right)$ est une forme bilinéaire symétrique sur E.
Le groupe $GL(E)$ agit sur $FBS(E)$ par $(h, \varphi) \to \varphi \circ (h^{-1}, h^{-1})$.

EXEMPLE 10.2.7
Soit P un plan affine euclidien orienté et O un point de P. Pour $\theta \in \mathbb{R}$, soit R_θ la rotation de centre O et d'angle θ. Le groupe $(\mathbb{R}, +)$ agit sur P par $(\theta, M) \to \theta * M = R_\theta(M)$. Le morphisme Φ associé est le morphisme $\theta \to R_\theta$ dont le noyau est $2\pi\mathbb{Z}$. Cette action n'est donc pas fidèle.

10.3 Orbites

Soit A une action d'un groupe G sur un ensemble X. On définit sur X une relation binaire $\mathcal{R}_A$ par $x\mathcal{R}_A y \Leftrightarrow \exists g \in G, y = g*x$. On vérifie facilement, en utilisant la définition d'une action que $\mathcal{R}_A$ est une relation d'équivalence sur X.

DEFINITION 10.3.1
Soit A une action du groupe G sur l'ensemble X. Si $x \in X$, on appelle orbite de x la classe O_x de x pour la relation d'équivalence $\mathcal{R}_A$ définie ci dessus.

On notera $\mathcal{O}$ l'ensemble de toutes les orbites, c'est à dire l'ensemble quotient $X/\mathcal{R}_A$.

Un point $a \in X$ est fixe pour l'action A ssi son orbite est réduite au singleton $\{a\}$.

DEFINITION 10.3.2
*L'action A de G sur X est dite **transitive** si elle n'a qu'une seule orbite, c'est à dire si $\forall x, y \in X, \exists g \in G,\ y = g * x$.*

Exemples

Dans l'exemple 10.2.3 la relation d'équivalence associée à l'action de $GL(n, \mathbb{K})$ est la similitude des matrices. L'orbite d'une matrice M est l'ensemble des matrices semblables à M.

De même dans l'exemple 10.2.4, l'orbite d'une matrice $M \in M_{p,q}(\mathbb{K})$ est l'ensemble des matrices équivalentes à M, donc d'après un théorème d'algèbre linéaire, l'ensemble des matrices de $M_{p,q}(\mathbb{K})$ de même rang que M.

Dans l'exemple 10.2.6 l'orbite d'une forme bilinéaire symétrique φ est l'ensemble des formes bilinéaires symétriques équivalentes à φ. Si $\mathbb{K} = \mathbb{C}$ on sait que deux formes bilinéaires symétriques sont dans la même orbite si et seulement si elles ont le même rang. Il y a donc $\dim E + 1$ orbites. Si $\mathbb{K} = \mathbb{R}$, deux formes bilinéaires symétriques sont dans la même orbite si et seulement si elles ont même signature.

Enfin dans l'exemple 10.2.7 l'orbite de O est le sigleton $\{O\}$ tandis que l'orbite d'un point m du plan est le cercle de centre O contenant le point m.

PROPOSITION 10.3.1
Soit $A : (g, x) \to g * x$ une action du groupe G sur l'ensemble X. Toute orbite est stable par A et l'action induite sur une orbite est transitive.

La preuve est immédiate.

PROPOSITION 10.3.2
Soit A une action du groupe G sur X et $a \in X$. L'ensemble $H_a = \{g \in G \mid g * a = a\}$ est un sous groupe de G.
On l'appelle stabilisateur de a ou sous groupe d'isotropie de a.

Là encore, la preuve est immédiate.

PROPOSITION 10.3.3
Soient A une action du groupe G sur X, $a \in X$ et $b \in \mathcal{O}_a$. Soit $g \in G$ tel que $b = g * a$. On a $H_b = gH_ag^{-1}$.

preuve
Soit $h \in G$. On a :
$h \in H_b \Leftrightarrow h * b = b \Leftrightarrow h * (g * a) = g * a \Leftrightarrow g^{-1} * ((hg) * a) = a \Leftrightarrow (g^{-1}hg) * a = a \Leftrightarrow g^{-1}hg \in H_a \Leftrightarrow h \in gH_ag^{-1}$. ■

Remarque
Si le stabilisateur de a est distingué dans G, le stabilisateur de tout élément de l'orbite de a est égal au stabilisateur de a. Autrement dit, tout élément de G qui laisse fixe a laisse fixe tout élément de l'orbite de a.

THEOREME 10.3.1
Soient A une action du groupe G sur X, $a \in X$. L'application qui à tout $b \in O_a$ associe l'ensemble $C_b = \{g \in G \mid g * a = b\}$ est une bijection de l'orbite O_a de a sur l'ensemble $(G/H_a)_g$ des classes à gauche de G modulo le stabilisateur H_a de a.

preuve

- Soit $b \in O_a$. L'ensemble C_b appartient à $(G/H_a)_g$.
  En effet, soit $g_0 \in C_b$. Soit $g \in G$. On a $g \in C_b \Leftrightarrow g * a = g_0 * a \Leftrightarrow (g_0^{-1}g) * a = a \Leftrightarrow g_0^{-1}g \in H_a \Leftrightarrow g \in g_0H_a$.
  Donc $C_b = g_0H_a$.
- L'application $b \to C_b$ est injective.
  En effet, si $b \neq b'$ il résulte de la définition de C_b que $C_b \cap C_{b'} = \emptyset$ et donc $C_b \neq C_{b'}$ puisque C_b n'est pas vide.
- L'application $b \to C_b$ est surjective.
  En effet, soit $\Gamma \in (G/H_a)_g$. Il existe $g_0 \in G$ tel que $\Gamma = g_0H_a$. Soit $b_0 = g_0 * a$. Alors b_0 est dans l'orbite de a et $g \in C_{b_0} \Leftrightarrow g * a = g_0 * a \Leftrightarrow g_0^{-1}g \in H_a \Leftrightarrow g \in g_0H_a$. Autrement dit $C_{b_0} = g_0H_a = \Gamma$.

La preuve est complète.

10.4 Equation des classes

Rappel Soit H un sous groupe d'un groupe G. Il existe une bijection entre l'ensemble $(G/H)_g$ des classes à gauche de G modulo H et l'ensemble $(G/H)_d$ des classes à droite de G modulo H. Ces deux ensembles ont donc même cardinal. Si ce cardinal est fini, il est noté $[G : H]$ et s'appelle l'indice du sous groupe H de G.

Soit G un groupe opérant sur un ensemble fini X. Il n'y a alors qu'un nombre fini d'orbites et chaque orbite est un ensemble fini. Choisissons dans chaque orbite un élément et un seul. Soit $\mathcal{C}$ l'ensemble de ces éléments. On a donc une partition de X formé des orbites $O_c,\ c \in \mathcal{C}$. Par conséquent $\text{card}(X) = \sum_{c\in\mathcal{C}} \text{card}(O_c)$. D'après le théorème 10.3.1, il existe une bijection entre O_c et le quotient $(G/H_c)_g$. Donc $\text{card}(O_c) = \text{card}\,(G/H_c)_g = [G : H_c]$. On a donc prouvé le résultat suivant.

THEOREME 10.4.1 (Equation des classes)
Soit G un groupe opérant sur un ensemble fini X. Soit $\mathcal{C}$ un sous ensemble de X contenant un point et un seul dans chaque orbite. On a, en désignant par H_c le stabilisateur d'un élément c de G

$$\textit{card}(X) = \sum_{c\in\mathcal{C}} [G : H_c]$$

10.5 Exemple d'application

Soit G un groupe. Pour $a \in G$ on note $i_a : G \to G$ l'automorphisme intérieur associé à a : $i_a(x) = axa^{-1}$. On a vu que G opérait sur lui même par automorphismes intérieurs et que cette action était fidèle ssi le centre Z de G était réduit à $\{1_G\}$.
Soit $g \in G$ fixé. Le stabilisateur de g est $H_g = \{a \in G \mid ag = ga\}$. C'est donc le commutant de g. L'orbite de g est l'ensemble des éléments de G conjugués à g.
Les points fixes de l'action considérée ci dessus sont les éléments du centre de G.

Supposons maintenant que G soit un groupe fini de cardinal $n = p^m$ où p est un nombre premier. (On dit que G est un p-groupe). Ecrivons l'équation des classes pour l'action de G sur lui même par automorphisme intérieur. Les orbites de cette action sont d'une part les points fixes, c'est à dire les éléments du centre Z de G et d'autre part les autres orbites dont le cardinal est différent de 1. Soit $\mathcal{O}'$ l'ensemble des orbites non singletons et $O \in \mathcal{O}'$. On a $\text{card}(O) = [G : H_a]$ où H_a est le stabilisateur d'un élément $a \in O$. Le nombre $[G : H_a]$ est un diviseur de $\text{card}(G) = p^m$ distinct de 1, donc de la forme $p^{k(O)}$ où $k(O)$ est un entier naturel au moins égal à 1. L'équation des classes s'écrit alors

$$\text{card}(G) = \text{card}(Z) + \sum_{O\in\mathcal{O}'} p^{k(O)}$$

d'où $\text{card}(Z) \equiv 0 \mod (p)$. Comme Z contient 1_G, ceci implique que $\text{card}(Z) \geqslant p$. On a donc prouvé :

Theoreme
Soit p un nombre premier. Le centre d'un p-groupe n'est jamais réduit à l'élément neutre.

Montrons à titre d'application de ce résultat que tout groupe de cardinal p^2 où p est un nombre premier est commutatif.

- Tout groupe H de cardinal p premier est cyclique.
  C'est un résultat bien connu : soit $a \in H$ un élément quelconque différent de l'élément neutre. Le cardinal du sous groupe cyclique $< a >$ engendré par a est au moins égal à 2 (il contient 1_G et a) et divise p, donc il vaut p. D'où $H =< a >$.

- Soit G un groupe d'ordre p^2, p premier. D'après ce qu'on a vu ci dessus, son centre Z est non trivial. $\text{card}(Z)$ divise p^2 donc $\text{card}(Z) = p$ ou $\text{card}(Z) = p^2$. Dans le deuxième cas $G = Z$ donc G est commutatif. Le résultat sera prouvé si on montre que le premier cas ne peut pas se produire.

  On va raisonner par l'absurde en montrant que l'hypothèse $\text{card}(Z) = p$ implique G commutatif, ce qui est manifestement contradictoire. Supposons donc $\text{card}(Z) = p$. Alors, Z est cyclique. Soit a un générateur de Z de sorte que $Z = \{1_G, a, a^2, \ldots, a^{p-1}\}$. Le centre Z d'un groupe quelconque est toujours distingué (vérification immédiate). En particulier, Z est distingué dans G et on dispose du groupe quotient $H = G/Z$ qui est de cardinal p, donc aussi cyclique. Soit $\beta \in G/Z$ un générateur et $b \in G$ dont la classe dans G/Z est β
  Soient $g \in G,\ \bar{g}$ sa classe dans G/Z. Il existe un entier k compris entre 0 et $p-1$ tel que $\bar{g} = \beta^k = \overline{b^k}$. On a alors $b^{-k}g \in Z$ et par conséquent, il existe un entier naturel m entre 0 et $p-1$ tel que $b^{-k}g = a^m$. Finalement, tout élément g de G s'écrit $g = b^k a^m$ avec $0 \leqslant k, m \leqslant p-1$. Soit $g' = b^{k'}a^{m'}$ un autre élément de G. Les puissances de a commutent avec tout élément de G, donc $gg' = b^k a^m b^{k'} a^{m'} = b^k b^{k'} a^m a^{m'} = b^{k+k'} a^{m+m'}$ et de même pour $g'g$. On a donc $gg' = g'g$ pour tous $g, g' \in G$ ce qui fournit la contradiction voulue.

A titre d'exercice, le lecteur pourra complèter ces résultats en montrant qu'un groupe de cardinal p^2 où p est premier est isomorphe soit à $\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$ soit à $\mathbb{Z}/p^2\mathbb{Z}$ et que ces deux groupes ne sont pas isomorphes.

11

Groupes de permutations

On notera $\mathbb{N}_n^* = \{1, 2, \ldots, n\}$.

11.1 Généralités

Si E est un ensemble non vide, l'ensemble des bijections de E, muni de la loi $\circ$ de composition des applications est un groupe noté $\mathfrak{S}_E$, appelé groupe des permutations de E. Son cardinal est $n!$. Dans le cas particulier où $E = \mathbb{N}_n^*$, le groupe $\mathfrak{S}_E$ est noté $\mathfrak{S}_n$ et appelé groupe symétrique d'ordre n.

Soit E un ensemble fini de cardinal n. Soit $\varphi : N_n^* \to E$ une bijection. Si $\sigma \in \mathfrak{S}_n$, l'application $\varphi \circ \sigma \circ \varphi^{-1}$ est une bijection de E. On définit ainsi une application $\varphi_* : \mathfrak{S}_n \to \mathfrak{S}_E$ dont il est facile de vérifier que c'est un isomorphisme de groupes. Ainsi le groupe $\mathfrak{S}_E$ est isomorphe à $\mathfrak{S}_n$. Cet isomorphisme dépend de la bijection φ choisie. Si $\psi : \mathbb{N}_n^* \to E$ est une autre bijection, on a, pour tout $\sigma \in \mathfrak{S}_n$, $\psi_*(\sigma) = \left(\psi \circ \varphi^{-1}\right) \circ \left(\varphi \circ \sigma \circ \varphi^{-1}\right) \circ \left(\varphi \circ \psi^{-1}\right) = \theta \circ \varphi_*(\sigma) \circ \theta^{-1}$ où $\theta = \psi \circ \varphi^{-1} \in \mathfrak{S}_E$. Autrement dit, $\psi_* = i_\theta \circ \varphi_*$ où i_θ est l'automorphisme intérieur de $\mathfrak{S}_E$ associé à l'élément θ.

DEFINITION 11.1.1
Soit $\sigma \in \mathfrak{S}_E$. On appelle support de σ l'ensemble des $x \in E$ tels que $\sigma(x) \neq x$.

DEFINITION 11.1.2
Une transposition de E est une permutation qui échange deux éléments. Autrement dit, $\tau : E \to E$ est une transposition si il existe deux éléments distincts $a, b \in E$ tels que $\tau(a) = b,\ \tau(b) = a$ et $\tau(x) = x$ pour tout $x \in E,\ x \neq a$ et $x \neq b$. Il revient au même de dire que τ est une transposition si et seulement si son support a deux éléments. On notera une telle transposition $\tau_{a,b}$ ou (a, b).

DEFINITION 11.1.3
Une permutation $\sigma \in \mathfrak{S}_E$ est un cycle d'ordre m ssi il existe une suite $(x_1, \ldots, x_m)$ d'éléments distincts de E telle que $\sigma(x_i) = x_{i+1}$ pour $1 \leqslant i \leqslant m-1,\ \sigma(x_m) = x_1$ et $\sigma(y) = y$ pour tout $y \notin \{x_1, \ldots, x_m\}$. Un tel cycle sera noté si il n'y a pas de confusion possible $\sigma = (x_1, \ldots, x_m)$. On notera qu'avec cette notation, $(x_2, \ldots, x_m, x_1)$ est aussi égal à σ.

Si $\varphi : E \to E'$ est une bijection et si σ est une transposition (resp. un cycle) de E, $\varphi_*(\sigma) = \varphi \circ \sigma \circ \varphi^{-1}$ est une transposition (resp. un cycle) de E'.

Lemme 11.1.1
Soient $\sigma, \tau \in \mathfrak{S}_E$. Si $\mathrm{Supp}(\sigma) \cap \mathrm{Supp}(\tau) = \emptyset$, σ et τ commutent : $\sigma \circ \tau = \tau \circ \sigma$.

Notation
Dans toute la suite, on notera $\sigma\tau$ la composée $\sigma \circ \tau$.

11.2 Le groupe $\mathfrak{S}_n$

11.2.1 Préliminaires

Si $n = 1$, le groupe $\mathfrak{S}_1$ est réduit à l'identité. Dans la suite, on supposera toujours $n \geqslant 2$.

Notation On utilisera la notation suivante, pour $\sigma \in \mathfrak{S}_n$: $\sigma = \begin{pmatrix} 1 & \dots & n \\ \sigma(1) & \dots & \sigma(n) \end{pmatrix}$.

Avec cette notation, le cycle $(1, 2, \dots, n)$ s'écrit $\begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ 2 & 3 & \dots & n & 1 \end{pmatrix}$

Si i et j sont deux entiers distincts entre 1 et n, on notera $\tau_{i,j}$ la permutation qui échange i et j. On notera que $\tau_{i,j} = \tau_{j,i}$.

Injection canonique de $\mathfrak{S}_m$ dans $\mathfrak{S}_n$
Soit m, n des entiers avec $2 \leqslant m \leqslant n$. Pour $\sigma \in \mathfrak{S}_m$ on définit $\sigma' : \mathbb{N}_n^* \to \mathbb{N}_n^*$ comme suit : $\sigma'(i) = \sigma(i)$ si $1 \leqslant i \leqslant m$ et $\sigma'(i) = i$ si $m + 1 \leqslant i \leqslant n$. Il est clair que $\sigma' \in \mathfrak{S}_n$ et que l'application $j_{m,n} : \mathfrak{S}_m \to \mathfrak{S}_n$ qui à σ associe σ' est un morphisme de groupes. Si $\tau_{i,j}^m$ est la transposition de $\mathfrak{S}_m$ qui échange i et j, $j_{m,n}(\tau_{i,j}^m) = \tau_{i,j}^n$. On a un résultat analogue pour les cycles.

11.2.2 Signature

DEFINITION 11.2.1
Soit $\sigma \in \mathfrak{S}_n$. On appelle nombre d'inversions de la permutation σ et on note $I(\sigma)$ l'entier

$$I(\sigma) = \text{card}\{(i, j) \in \mathbb{N}_n^* \mid i < j \text{ et } \sigma(i) > \sigma(j)\}$$

On appelle signature de la permutation σ le nombre

$$\varepsilon_\sigma = (-1)^{I(\sigma)}$$

Exemples
Si $\sigma = id$, on a $I(id) = 0$ et $\varepsilon_{id} = 1$
Soit $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 5 & 2 & 4 \end{pmatrix}$ Les couples faisant inversion sont : $(1, 2), (1, 4)(3, 4), (3, 5)$. Donc $I(\sigma) = 4$ et $\varepsilon_\sigma = 1$.

PROPOSITION 11.2.1
La signature d'une transposition vaut -1.

preuve
Soient $1 \leqslant i < j \leqslant n$.

$$\begin{aligned} I(\tau_{i,j}) &= \text{card}\left(\{(i,j)\} \cup \{(i,k)\ ;\ i+1 \leqslant k \leqslant j-1\} \cup \{(k,j)\ ;\ i+1 \leqslant k \leqslant j-1\}\right) \\ &= 1 + 2[(j-1) - (i+1) + 1] = 2(j-i) - 1 \end{aligned}$$

THEOREME 11.2.1
La signature est un homomorphisme surjectif $\varepsilon : \mathfrak{S}_n \to \{-1, 1\}$.

Avant de démontrer le théorème, donnons une définition et un corollaire.

DEFINITION 11.2.2
On appelle groupe alterné d'ordre n et on note $\mathfrak{A}_n$ l'ensemble des permutations de $\mathfrak{S}_n$ de signature +1.

COROLLAIRE 11.2.1
Le groupe alterné $\mathfrak{A}_n$ est un sous groupe distingué de $\mathfrak{S}_n$ d'indice 2.

$\mathfrak{A}_n$ est le noyau du morphisme ε ; donc c'est un sous groupe distingué de $\mathfrak{S}_n$. D'après le théorème d'isomorphisme, le quotient $\mathfrak{S}_n/\mathfrak{A}_n$ est isomorphe à $\text{im}(\varepsilon) = \{-1, 1\}$, donc est de cardinal 2.

La preuve du théorème utilise une autre expression de la signature.

Lemme 11.2.1
Soit $\sigma \in \mathfrak{S}_n$. On a

$$\varepsilon_\sigma = \frac{\prod\limits_{i<j}(\sigma(j)-\sigma(i))}{\prod\limits_{i<j}(j-i)}$$

preuve du lemme

- Soit $\Gamma = \{(i,j) \mid i<j\}$. Si $(i,j) \in \Gamma$, notons $\bar{\sigma}(i,j)$ le couple ordonné associé à $(\sigma(i),\sigma(j))$, autrement dit

$$\bar{\sigma}(i,j) = \begin{cases} (\sigma(i),\sigma(j)) & \text{si } \sigma(i) < \sigma(j) \\ (\sigma(j),\sigma(i)) & \text{si } \sigma(i) > \sigma(j) \end{cases}$$

  $\bar{\sigma}$ est une application de Γ dans lui même, manifestement injective. Comme Γ est fini, c'est une bijection de Γ.

- Soit $X : \Gamma \to \{0,1\}$ définie par

$$X(i,j) = \begin{cases} 1 & \text{si } \sigma(j) < \sigma(i) \\ 0 & \text{si } \sigma(j) > \sigma(i) \end{cases}$$

  On a $I(\sigma) = \sum\limits_{(i,j)\in\Gamma} X(i,j)$

- On a

$$\prod_{i<j}(\sigma(j)-\sigma(i)) = \prod_{i<j}(-1)^{X(i,j)}|\sigma(j)-\sigma(i)| = \varepsilon_\sigma \prod_{(i,j)\in\Gamma}|\sigma(j)-\sigma(i)|$$

  L'application $\bar{\sigma}$ étant bijective, on peut faire dans le produit le changement d'indice $(u,v) = \bar{\sigma}(i,j)$. Il vient

$$\prod_{(i,j)\in\Gamma}|\sigma(j)-\sigma(i)| = \prod_{(u,v)\in\Gamma}|v-u| = \prod_{i<j}(j-i)$$

  d'où le lemme.

preuve du théorème
Soient $\sigma, \sigma' \in \mathfrak{S}_n$. Il s'agit de montrer que $\varepsilon_{\sigma'\sigma} = \varepsilon_{\sigma'}\varepsilon_\sigma$. En utilisant le lemme il vient

$$\varepsilon_{\sigma'\sigma} = \prod_{i<j}\frac{\sigma'(\sigma(j))-\sigma'(\sigma(i))}{j-i} = \prod_{(i,j)\in\Gamma}\frac{\sigma(j)-\sigma(i)}{j-i}\prod_{(i,j)\in\Gamma}\frac{\sigma'(\sigma(j))-\sigma'(\sigma(i))}{\sigma(j)-\sigma(i)} = \varepsilon_\sigma\prod_{(i,j)\in\Gamma}\frac{\sigma'(\sigma(j))-\sigma'(\sigma(i))}{\sigma(j)-\sigma(i)}$$

$\bar{\sigma}$ est bijective. Compte tenu de $\dfrac{\sigma'(v)-\sigma'(u)}{v-u} = \dfrac{\sigma'(u)-\sigma'(v)}{u-v}$ il vient

$$\prod_{(i,j)\in\Gamma}\frac{\sigma'(\sigma(j))-\sigma'(\sigma(i))}{\sigma(j)-\sigma(i)} = \prod_{(u,v)\in\bar{\sigma}(\Gamma)}\frac{\sigma'(v)-\sigma'(u)}{v-u} = \prod_{(u',v')\in\Gamma}\frac{\sigma'(v')-\sigma'(u')}{v'-u'} = \varepsilon_{\sigma'}$$

Ceci achève la preuve du théorème.

COROLLAIRE 11.2.2
Deux éléments conjugués de $\mathfrak{S}_n$ ont la même signature.

preuve
Soient σ et σ' conjugués. Il existe $\theta \in \mathfrak{S}_n$ tel que $\sigma' = \theta\sigma\theta^{-1}$; alors $\varepsilon_{\sigma'} = \varepsilon_\theta\varepsilon_\sigma\varepsilon_{\theta^{-1}} = \varepsilon_\theta\varepsilon_\sigma(\varepsilon_\theta)^{-1} = \varepsilon_\sigma$ car $\{-1,1\}$ est commutatif.

THEOREME 11.2.2
La signature d'un cycle de longueur m est $(-1)^{m-1}$

Soit $\sigma_0 = (1,2,\ldots,m)$. On a $I(\sigma_0) = \text{card}\{(1,m),\ldots,(m-1,m)\} = m-1$ donc $\varepsilon_{\sigma_0} = (-1)^{m-1}$. Le théorème résulte alors du théorème suivant :

THEOREME 11.2.3
Deux cycles quelconques de même longueur sont conjugués.

preuve
Soit $\sigma = (x_1, x_2, \ldots, x_m)$ un cycle de longueur m. Il suffit de montrer qu'il est conjugué au cycle $\sigma_0 = (1, 2, \ldots, m)$. Choisissons une bijection φ de $\mathbb{N}_n^*$ telle que $\varphi(k) = x_k$ pour $1 \leqslant k \leqslant m$.
On a, pour $0 \leqslant k \leqslant m-1$, $\varphi\sigma_0\varphi^{-1}(x_k) = \varphi\sigma_0(k) = \varphi(k+1) = x_{k+1} = \sigma(x_k)$. De même, $\varphi\sigma_0\varphi^{-1}(x_m) = x_1 = \sigma(x_m)$.
Si $x \in \mathbb{N}_n^* \setminus \{x_1, \ldots, x_m\}$, $y = \varphi^{-1}(x) \in \{m+1, \ldots, n\}$ donc $\sigma_0(y) = y$ et $\varphi\sigma_0\varphi^{-1}(x) = \varphi\left(\varphi^{-1}(x)\right) = x = \sigma(x)$.
Finalement, pour tout x, $\sigma(x) = \varphi\sigma_0\varphi^{-1}(x)$ donc $\sigma = \varphi\sigma_0\varphi^{-1}$.

Remarque
Le lecteur vérifiera aisément que le conjugué d'un cycle de longueur m est un cycle de longueur m. En particulier, pour tout $\sigma \in \mathfrak{S}_n$ on a $\sigma\tau_{i,j}\sigma^{-1} = \tau_{\sigma(i),\sigma(j)}$.

Signature d'une permutation d'un ensemble quelconque

Soit E un ensemble fini ayant n éléments et $\sigma \in \mathfrak{S}_E$. Soient $\varphi, \psi : \mathbb{N}_n^* \to E$ deux bijections. On a, avec les notations du début du chapitre $\psi_*\sigma = \theta\varphi_*(\sigma)\theta^{-1}$. Les deux permutations $\psi_*(\sigma)$ et $\varphi_*(\sigma)$ sont conjugués donc ont la même signature.
On peut donc définir la signature de σ comme celle de $\varphi_*(\sigma)$ pour une bijection quelconque $\varphi : \mathbb{N}_n^* \to E$. Cette définition ne dépend pas du choix de φ. En fixant une telle bijection, on voit que la signature est un morphisme de $\mathfrak{S}_E$ sur $\{-1, 1\}$. On peut donc encore définir le groupe alterné de E comme l'ensemble des permutations paires de E (i.e. de signature 1). La signature d'une transposition quelconque est -1, celle d'un cycle de longueur m est $(-1)^{m-1}$.

11.3 Parties génératrices de $\mathfrak{S}_n$

THEOREME 11.3.1
$\mathfrak{S}_n$ est engendré par les transpositions.

Le théorème se montre par récurrence sur n. Il est trivial si $n = 2$. Supposons le résultat établi pour un entier $n \geqslant 2$. Soit $\sigma \in \mathfrak{S}_{n+1}$. Il s'agit de montrer que σ s'écrit comme un produit de transpositions. Distinguons deux cas :
1) $\sigma(n+1) = n+1$. Dans ce cas la restriction de σ à $\mathbb{N}_n^*$ est un élément u de $\mathfrak{S}_n$ et $\sigma = j_{n,n+1}(u)$. Par hypothèse de récurrence, il existe un nombre fini de transpositions appartenant à $\mathfrak{S}_n$, τ_k^n, $1 \leqslant k \leqslant N$ telles que $u = \tau_N^n \cdots \tau_1^n$. Posons $\tau_k = j_{n,n+1}(\tau_k^n)$. τ_k est une transposition de $\mathfrak{S}_{n+1}$ et $j_{n,n+1}$ étant un morphisme, $\sigma = j_{n,n+1}(u) = j_{n,n+1}(\tau_N^n \cdots \tau_1^n) = \tau_N \cdots \tau_1$. Dans ce cas, σ s'écrit comme un produit fini de transpositions.
2) $p = \sigma(n+1) \neq n+1$. Soit τ la transposition qui échange p et $n+1$. La permutation $\tau\sigma$ est telle que $\tau\sigma(n+1) = n+1$. D'après la première partie, cette permutation est égale à un produit de transpositions $\tau_N \cdots \tau_1$. Alors $\sigma = \tau\tau_N \cdots \tau_1$. ■

Remarque
La décomposition de $\sigma \in \mathfrak{S}_n$ en produit $\tau_N \cdots \tau_1$ n'est pas unique. On a en effet, si τ est une transposition quelconque $\sigma = \tau\tau\tau_N \cdots \tau_1$. Par contre, la signature étant un morphisme et la signature d'une transposition étant -1, on a $\varepsilon_\sigma = (-1)^N$ donc la parité du nombre N de transpositions ne dépend que de σ.

THEOREME 11.3.2
$\mathfrak{A}_n$ est engendré par les 3-cycles.

preuve
Tout $\sigma \in \mathfrak{A}_n$ est produit d'un nombre pair de transpositions. Il suffit donc de montrer que le produit de deux transpositions quelconques peut s'écrire comme produit de 3-cycles. Soient donc $i \neq j$ et $k \neq l$ des entiers entre 1 et n et $\sigma = \tau_{i,j}\tau_{k,l}$. On distingue trois cas :

1. $\{i, j\} = \{k, l\}$ alors $\sigma = id$
2. $\text{card}\left(\{i, j\} \cap \{k, l\}\right) = 1$. Quitte à changer les notations, on peut supposer que j est l'élément commun. Alors $\sigma = \tau_{i,j}\tau_{j,k} = (i, j, k)$
3. $\{i, j\} \cap \{k, l\} = \emptyset$ alors $\sigma = (i, j, k)(j, k, l)$

La preuve est complète.

Compléments : Autres systémes générateurs

A titre d'exercice le lecteur est invité à démontrer les propositions suivantes. Des indications sont données en fin de chapitre.

PROPOSITION 11.3.1
Pour $n \geqslant 2$, $\mathfrak{S}_n$ est engendré par $\{\tau_{1,2}, \tau_{1,3}, \ldots, \tau_{1,n}\}$

PROPOSITION 11.3.2
Pour $n \geqslant 2$, $\mathfrak{S}_n$ est engendré par $\{\tau_{1,2}, \tau_{2,3} \ldots, \tau_{n-1,n}\}$.

PROPOSITION 11.3.3
Pour $n \geqslant 2$, $\mathfrak{S}_n$ est engendré par $\{\tau, \sigma\}$ où $\tau = \tau_{1,2}$ et σ est le n-cycle $\sigma = (1, 2, \ldots, n)$.

PROPOSITION 11.3.4
Pour $n \geqslant 3$, $\mathfrak{A}_n$ est engendré par l'ensembles des trois cycles $\{(1, i, j)\,;\ 2 \leqslant i \leqslant n, 2 \leqslant j \leqslant n, i \neq j\}$

PROPOSITION 11.3.5
Pour $n \geqslant 3$, $\mathfrak{A}_n$ est engendré par l'ensemble des 3-cycles $\{(1, 2, i)\ ;\ 3 \leqslant i \leqslant n\}$.

11.4 Décomposition en produits de cycles

On se propose de prouver dans cette partie que toute permutation σ peut s'écrire comme composée de cycles à supports disjoints, cette décomposition étant essentiellement unique. Nous utiliserons à cette fin les résultats relatifs aux actions de groupe.

σ-orbites

Soit $\sigma \in \mathfrak{S}_n$. Le groupe cyclique $< \sigma >$ engendré par σ agit sur $\mathbb{N}_n^*$ de manière naturelle : $(\sigma^j, x) \to \sigma^j(x)$. On appelle σ-orbite toute orbite de cette action. C'est une classe d'équivalence pour la relation $\mathcal{R}$ définie par $x\mathcal{R}y \Leftrightarrow \exists p \in \mathbb{N},\ \sigma^p(x) = y$.

Lemme 11.4.1
Soit $\sigma \in \mathfrak{S}_n$. σ est un cycle (distinct de l'identité) ssi il existe une seule σ-orbite de cardinal strictement supérieur à 1. La longueur du cycle est alors égale au cardinal de cette σ-orbite.

preuve
Soit m l'ordre de σ dans $\mathfrak{S}_n$ de sorte que $< \sigma >= \{id, \sigma, \ldots, \sigma^{m-1}\}$.
Une implication est évidente. Supposons que σ soit telle qu'il y ait une seule σ-orbite de cardinal > 1. Soit O cette orbite. On a donc $\sigma(x) = x$ si $x \notin O$. Soit $a \in O$. L'orbite de a est O. Montrons que le stabilisateur de a est l'identité. Soit $\varphi \in \mathrm{Stab}(a)$. Il existe un entier k tel que $\varphi = \sigma^k$. On a $\sigma^k(a) = a$ donc, pour tout j, $\sigma^k(\sigma^j(a)) = \sigma^j(\sigma^k(a)) = \sigma^j(a)$, autrement dit, $\sigma^k(x) = x$ pour tout $x \in O$. Comme on a aussi $\sigma^k(y) = y$ pout tout $y \in \mathbb{N}_n^* \setminus O$, on a $\varphi = \sigma^k = id$. Donc $\mathrm{Stab}(a) = \{id\}$. Il en résulte que l'application $< \sigma > \to O$ qui à $u \in < \sigma >$ associe $u(a)$ est une bijection. On a donc $\mathrm{card}(O) = m,\ O = \{a, \sigma(a), \ldots, \sigma^{m-1}(a)\}$ et σ est le cycle $(a, \sigma(a), \ldots, \sigma^{m-1}(a))$.

THEOREME 11.4.1 (Decomposition en produits de cycles à supports disjoints)
Soit $\sigma \in \mathfrak{S}_n,\ \sigma \neq id$ où $n \geqslant 2$. Soit Ω l'ensemble des σ-orbites non réduites à un point. Il existe une famille $(c_\omega)_{\omega \in \Omega}$ de cycles vérifiant

1. $\forall \omega \in \Omega,\ \mathrm{Supp}(c_\omega) = \omega$
2. $\sigma = \prod_{\omega \in \Omega} c_\omega$

Les c_ω sont à supports disjoints, donc commutent entre eux. L'ordre de σ est le ppcm des longueurs des cycles c_ω.
Cette décomposition est unique dans le sens suivant: si $\sigma = \prod_{1\leqslant i\leqslant r} c_i$ où les c_i sont des cycles à supports disjoints, alors $r = card(\Omega)$ et $\{c_1, \ldots, c_r\} = \{c_\omega \ ; \ \omega \in \Omega\}$.

preuve

1) Pour $\omega \in \Omega$ soit $c_\omega \in \mathfrak{S}_n$ défini par $c_\omega(x) = \begin{cases} \sigma(x) & \text{si } x \in \omega \\ x & \text{si } x \notin \omega \end{cases}$. D'après le lemme, c_ω est un cycle de support ω, de longueur card(ω). Les c_ω sont à support disjoints, donc commutent entre eux. Il en résulte facilement que $\sigma = \prod_{\omega\in\Omega} c_\omega$. En effet, soit $x \in \mathbb{N}_n^*$.

Ou bien $\sigma(x) = x$ et $\forall\omega,\ x \notin \omega$ donc $c_\omega(x) = x$ et dans ce cas $\sigma(x) = \left(\prod_{\omega\in\Omega} c_\omega\right)(x)$.

Ou bien $\sigma(x) \neq x$. Il existe alors un unique $\omega_0 \in \Omega$ tel que $x \in \omega_0$. On a $c_\omega(x) = x$ si $\omega \neq \omega_0$ et $c_{\omega_0}(x) = \sigma(x)$. En utilisant la commutativité des c_ω il vient $\left(\prod_{\omega\in\Omega} c_\omega\right)(x) = c_{\omega_0} \circ \left(\prod_{\substack{\omega\in\Omega \\ \omega\neq\omega_0}} c_\omega\right)(x) = c_{\omega_0}(x) = \sigma(x)$.

2) Unicité : Si $\sigma = \prod_{1\leqslant i\leqslant r} c_i$ où les c_i sont des cycles à supports disjoints, il est clair que $\mathrm{Supp}(c_i)$ est une σ-orbite. Si $x \notin \bigcup_{1\leqslant i\leqslant r} \mathrm{Supp}(c_i)$, on a $\sigma(x) = x$. Donc $\Omega = \{\mathrm{Supp}(c_i) \ ; \ 1 \leqslant i \leqslant r\}$ et card(Ω) $= r$. Ensuite, fixons un i et soit $\omega = \mathrm{Supp}(c_i)$. Pour $x \in \omega$ et $j \neq i$ on a $c_j(x) = x$ donc, puisque les c_j commutent, on a $\sigma(x) = c_i(x)$. Par conséquent, pour un tel x, $c_\omega(x) = \sigma(x) = c_i(x)$. Pour $x \notin \omega$, on a par définition $c_\omega(x) = x = c_i(x)$. Donc $c_\omega = c_i$ ce qui achève la preuve de l'unicité.

3) Comme les c_ω commutent entre eux, on a pour tout entier q, $\sigma^q = \prod_{\omega\in\Omega} c_\omega^q$ donc $\sigma^q|_\omega = c_\omega^q|_\omega$. Il en résulte facilement que $\sigma^q = id \Leftrightarrow \forall\omega \in \Omega,\ c_\omega^q = id$. Donc $\sigma^q = id$ ssi pour tout ω, q est multiple de l'ordre du cycle c_ω. On en déduit que l'ordre de σ est le ppcm des ordres des c_ω c'est à dire des longueurs des cycles ω.

EXEMPLE 11.4.1
Cherchons la décomposition en produits de cycles de $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 4 & 7 & 5 & 2 & 6 & 8 & 1 \end{pmatrix}$.
On cherche d'abord l'orbite de 1. Les images successives par σ sont $1 \to 3 \to 7 \to 8 \to 1$. Ceci fournit le cycle $c = (1, 3, 7, 8)$. L'orbite de 2 donne $2 \to 4 \to 5 \to 2$ qui fournit le cycle $c' = (2, 4, 5)$. Enfin 6 est fixe. Donc $\sigma = cc' = c'c$. On en déduit la signature de σ : $\varepsilon_\sigma = \varepsilon_c\varepsilon_{c'} = (-1)^3(-1)^2 = -1$. Enfin σ est d'ordre 12.

THEOREME 11.4.2
Soit $\sigma \in \mathfrak{S}_n$ et N_σ le nombre total de σ-orbites (y compris les orbites signetons). On a

$$\varepsilon_\sigma = (-1)^{n-N_\sigma}$$

preuve
Soient $F = \{x \ ; \ \sigma(x) = x\}$, Ω l'ensemble des orbites non singletons de σ et $\sigma = \prod_{\omega\in\Omega} c_\omega$ la décomposition de σ en produits de cycles à supports disjoints. On a

$$\varepsilon_\sigma = \prod_{\omega\in\Omega} \varepsilon_{c_\omega} = \prod_{\omega\in\Omega} (-1)^{\mathrm{card}(\omega)-1}$$

Or $\sum_{\omega\in\Omega} \mathrm{card}(\omega) + \mathrm{card}(F) = n$ donc $\sum_{\omega\in\Omega} (\mathrm{card}(\omega) - 1) = n - \mathrm{card}(F) - \mathrm{card}(\Omega) = n - N_\sigma$.■

Indications pour les preuves des propositions 11.3.1 à 11.3.5

Proposition 11.3.1 : Il n'y a rien à montrer si $n = 2$. Soit $n \geqslant 3$; pour $2 \leqslant j < k$ on a $\tau_{j,k} = \tau_{1,j}\tau_{1,k}\tau_{1,j}$ et le résultat découle du théorème 11.3.1.
Proposition 11.3.2 : Utiliser la relation $\tau_{1,k}\tau_{1,k+1}\tau_{1,k} = \tau_{k,k+1}$ valable pour $2 \leqslant k \leqslant n-1$ pour montrer que le groupe engendré par $\{\tau_{1,2}, \tau_{2,3} \ldots, \tau_{n-1,n}\}$ contient tous les $\tau_{1,k}$.
Proposition 11.3.3 : Pour $1 \leqslant p \leqslant n-2$, $\sigma^p\tau\sigma^{-p} = \tau_{p+1,p+2}$.
Proposition 11.3.4 : Pour $n \geqslant 4$ et $i, j, k \geqslant 2$ on a $(i, j, k) = (i, j)(j, k) = (i, j)(j, 1)(j, 1)(j, k) = (1, i, j)(1, j, k)$
Proposition 11.3.5 :Pour $n \geqslant 4$ et $i, j \geqslant 3$ on a $(1, i, j) = (1, 2, j)(1, i, 2) = (1, 2, j)(1, 2, i)^2$

12

Déterminants

Dans tout ce chapitre, $\mathbb{K}$ est un corps commutatif de caractéristique différente de 2.

12.1 Formes p-linéaires alternées sur un $\mathbb{K}$-espace vectoriel

DEFINITION 12.1.1
Soient $E_1, \ldots, E_n$ et F des $\mathbb{K}$-espaces vectoriels. Une application $E_1 \times \cdots \times E_n \to F$ est dite multilinéaire si elle est séparément linéaire en chacune des variables, i.e. si pour tout $1 \leqslant j \leqslant n$, et tout $(v_1, \ldots, v_n) \in E_1 \times \cdots \times E_n$ l'application de E_j dans F qui à v associe $f(v_1, \ldots, v_{j-1}, v, v_j, \ldots, v_n)$ est linéaire.

L'ensemble $ML(E_1, \cdots, E_n; F)$ des applications multilinéaires de $E_1 \times \cdots \times E_n$ dans F est un sous espace vectoriel de l'espace vectoriel des fonctions de $E_1 \times \cdots \times E_n$ dans F. Si chacun des espaces E_k est de dimension finie, ainsi que F, on a $\dim(ML(E_1, \ldots, E_n, F) = \dim(E_1) \cdot \dim(E_2) \cdots \dim(E_n) \cdot \dim(F)$.

DEFINITION 12.1.2
Soit E un $\mathbb{K}$-espace vectoriel et p un entier, $p \geqslant 1$. On appelle forme p-linéaire sur E toute application p linéaire de E^p dans le corps de base $\mathbb{K}$.

L'ensemble $ML_p(E) = ML(\underbrace{E, \cdots, E}_{p \text{ fois}}; \mathbb{K})$ des p formes linéaires sur E est un $\mathbb{K}$-sous-espace vectoriel de l'espace des applications de E^p dans $\mathbb{K}$.
Soit $\mathfrak{S}_p$ le groupe symétrique d'ordre p ; soit $\sigma \in \mathfrak{S}_p$ et $f \in ML_p(E)$. On définit $\sigma * f : E^p \to \mathbb{K}$ par $\forall\, (v_1, \ldots, v_p) \in E^p,\ \sigma * f(v_1, \ldots, v_p) = f\left(v_{\sigma(1)}, \ldots, v_{\sigma(p)}\right)$. On définit ainsi une action de $\mathfrak{S}_p$ sur $ML_p(E)$. En effet, soit $\tau \in \mathfrak{S}_p$. On a $\tau * (\sigma * f)(v_1, \ldots, v_p) = \sigma * f(v_{\tau(1)}, \ldots, v_{\tau(p)}) = f\left(v_{\tau(\sigma(1))}, \ldots, v_{\tau(\sigma(p))}\right) = (\tau\sigma) * f(v_1, \ldots, v_p)$ ceci pour toute $f \in ML_p(E)$. D'autre part, il est immédiat que $id * f = f$.

DEFINITION 12.1.3
Soit f une p-forme linéaire sur un $\mathbb{K}$-espace vectoriel E.
1) f est symétrique si $\forall \sigma \in \mathfrak{S}_p,\ \sigma * f = f$.
2) f est antisymétrique si $\forall \sigma \in \mathfrak{S}_p,\ \sigma * f = \varepsilon_\sigma f$ où ε_σ est la signature de la permutation σ.
3) f est dite alternée si $f(v_1, \ldots, v_p) = 0$ chaque fois qu'il existe deux indices distincts i, j entre 1 et p tels que $v_i = v_j$.

L'ensemble des formes p-linéaires symétriques (resp. antisymétriques, alternées) est un sous espace vectoriel de $ML_p(E)$.

Une forme f est symétrique (resp. antisymétrique) ssi $\tau * f = f$ (resp. $\tau * f = -f$) pour toute permutation $\tau \in \mathfrak{S}_p$. Toute forme alternée est antisymétrique. En effet, soient $1 \leqslant i < j \leqslant p$ et τ la transposition qui échange i et j. On a pour tout $(v_1, \ldots, v_p) \in E^p$, $f(v_1, \ldots, v_{i-1}, v_i + v_j, v_{i+1}, \ldots, v_{j-1}, v_i + v_j, v_{j+1}, \ldots, v_p) = 0$ d'où, en développant et vu le caractère alterné de f, $f(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_p) + f(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_p) = 0$ et donc $\tau * f = -f$.
La réciproque est vraie si le corps de base est de caractéristique différente de 2. En effet, soit f antisymétrique. Soit

$(v_1,\dots,v_p)\in\mathbb{K}^p$ tels que $v_i=v_j$ pour un couple d'indices $i<j$. Soit τ la transposition qui échange i et j. La relation $\tau * f=-f$ donne $f(v_1,\dots,v_p)=-f(v_1,\dots,v_p)$ soit $2f(v_1,\dots,v_p)=0$. Par hypothèse $2\cdot 1_{\mathbb{K}}\neq 0$ donc en multipliant par l'inverse, on obtient $f(v_1,\dots,v_p)=0$.

Dans ce chapitre, on a supposé que $\mathbb{K}$ n'est pas de caractéristique 2, ce qui est aussi l'hypothèse du programme où $\mathbb{K}$ est un sous corps de $\mathbb{C}$. Par conséquent, dans la suite, on ne distinguera plus formes alternées et formes antisymétriques.

PROPOSITION 12.1.1
Soit f une p-forme alternée sur un $\mathbb{K}$-espace vectoriel E. Soit $(v_1,\dots,v_p)\in E^p$. Si le système $(v_1,\dots,v_p)$ est lié, alors $f(v_1,\dots,v_p)=0$.

En effet, l'un des vecteurs, disons v_j est combinaison linéaire des autres : $v_j=\sum\limits_{\substack{1\leqslant i\leqslant p\\ i\neq j}} a_i v_i$.
Alors $f(v_1,\dots,v_p)=\sum\limits_{\substack{1\leqslant i\leqslant p\\ i\neq j}} a_i f(v_1,\dots,v_{i-1},v_i,\dots,v_{j-1},v_i,v_{j+1},\dots,v_p)=0$.

COROLLAIRE 12.1.1
Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n. Si $p>n$, toute p forme linéaire alternée sur E est nulle.

12.2 Déterminants

12.2.1 Formes n-linéaires alternées sur un espace de dimension n

Dans toute cette partie, E est un $\mathbb{K}$-espace vectoriel de dimension finie $n\geqslant 1$. On note $\Lambda_n(E)$ l'espace des n-formes linéaires alternées sur E.
Toute la théorie des déterminants repose sur le théorème suivant :

THEOREME 12.2.1
Soit E un $\mathbb{K}$-espace vectoriel de dimension n.
1) Pour toute base $\mathcal{B}=(e_1,\dots,e_n)$, il existe une unique forme n-linéaire alternée $f:E^n\to\mathbb{K}$ telle que $f(e_1,\dots,e_n)=1$. Cette forme est notée $\det_{\mathcal{B}}$ et s'appelle le déterminant dans la base $\mathcal{B}$.
2) L'espace $\Lambda_n(E)$ est de dimension 1, et, pour toute base $\mathcal{B}$ de E, $\det_{\mathcal{B}}$ en est une base.

preuve

1. Soit $\varphi:E^n\to\mathbb{K}$ une forme n-linéaire alternée sur E et $\mathcal{B}=(e_1,\dots,e_n)$ une base de E. Pour toute permutation $\sigma\in\mathfrak{S}_n$ on a $\varphi(e_{\sigma(1)},\dots,e_{\sigma(n)})=\varepsilon_\sigma\varphi(e_1,\dots,e_n)$. D'autre part, si $(i_1,\dots,i_n)\in\mathbb{N}_n^*$ est tel que $i_p=i_q$ pour $p\neq q$ on a $\varphi(e_{i_1},\dots,e_{i_n)}=0$. Soit alors $v_j=\sum\limits_{i=1}^{i=n} a_{i,j}e_i$, $1\leqslant j\leqslant n$ n vecteurs de E. On a

$$\varphi(v_1,\dots,v_n)=\varphi\left(\sum_{i_1=1}^{i_1=n}a_{i_1}e_{i_1},\dots,\sum_{i_n=1}^{i_n=n}a_{i_n,n}e_{i_n}\right)=\sum_{i_1=1}^{i_1=n}\cdots\sum_{i_n=1}^{i_n=n}a_{i_1,1}\cdots a_{i_n,n}\varphi(e_{i_1},\dots e_{i_n})$$

   Donc compte tenu des remarques précédentes, en posant $\lambda=\varphi(e_1,\dots,e_n)$,

$$\varphi(v_1,\dots,v_n)=\left(\sum_{\sigma\in\mathfrak{S}_n}\varepsilon_\sigma a_{\sigma(1),1}\cdots a_{\sigma(n),n}\right)\lambda \tag{12.1}$$

2. On en déduit déjà l'unicité de f. Si elle existe, elle est donnée par (1) avec $\lambda=1$.

3. Supposons prouvée l'existence de $f=\det_{\mathcal{B}}$. Soit $\varphi\in\Lambda_n(E)$. Posons $\lambda=\varphi(e_1,\dots,e_n)$. Le calcul ci dessus montre alors que pour tout $(v_1,\dots,v_n)\in E^n$ on a $\varphi(e_1,\dots,e_n)=\lambda f(e_1,\dots,e_n)$ soit $\varphi=\lambda\det_{\mathcal{B}}$. Comme $\det_{\mathcal{B}}$ n'est pas l'application nulle, on en déduit que $\Lambda_n(E)$ est de dimension 1 et que $\det_{\mathcal{B}}$ en est une base.

4. Définissons donc l'application f par

$$f(v_1,\dots,v_n)=\sum_{\sigma\in\mathfrak{S}_n}\varepsilon_\sigma a_{\sigma(1),1}\cdots a_{\sigma(n),n} \tag{12.2}$$

C'est une application $E^n\to\mathbb{K}$. Il reste à voir qu'elle répond à la question.

- On a $f(e_1,\dots,e_n)=1$. En effet, on avec $v_j=e_j$ pour tout j on a $a_{i,j}=\delta_{i,j}$ pour tout couple (i,j). Si σ est une permutation différente de id il existe un i tel que $\sigma(i)\neq i$ et donc $a_{\sigma(i),i}=0$. Dans la somme (2) il n'y a qu'un terme non nul pour $\sigma=id$. Il vient $f(e_1,\dots,e_n)=1$.
- Pour j entre 1 et n et $v_1,\dots,v_{j-1},v_{j+1},\dots,v_n$ fixés, l'application $u:v_j\to f(v_1,\dots,v_n)$ est linéaire. C'est immédiat : si $v'_j=\sum_{i=1}^{i=n}a'_{i,j}e_j$ et si $t\in\mathbb{K}$ on a

$$\begin{aligned}u(v_j+tv'_j) &= \sum_{\sigma\in\mathfrak{S}_n}\varepsilon_\sigma a_{\sigma(1),1}\cdots(a_{\sigma(j),j}+ta'_{\sigma(j),j}),\cdots a_{\sigma(n),n}\\ &= \sum_{\sigma\in\mathfrak{S}_n}\varepsilon_\sigma a_{\sigma(1),1}\cdots a_{\sigma(j),j},\cdots a_{\sigma(n),n}+t\sum_{\sigma\in\mathfrak{S}_n}\varepsilon_\sigma a_{\sigma(1),1}\cdots a'_{\sigma(j),j},\cdots a_{\sigma(n),n}\\ &= u(v_j)+tu(v'_j)\end{aligned}$$

- f est alternée. Pour le voir, considérons deux indices i et j tels que $1\leqslant i<j\leqslant n$ et notons τ la transposition qui échange i et j. Soit $(v_1,\dots,v_n)\in E^n$ avec $v_i=v_j$. On a $f(v_1,\dots,v_n)=\sum_{\sigma\in\mathfrak{S}_n}\varepsilon_\sigma a_{\sigma(1),1}\cdots a_{\sigma(n),n}$

  Dans la dernière somme, on effectue le changement d'indices $\sigma=\sigma'\tau$ (l'application $\sigma'\to\sigma'\tau$ est une involution, donc une bijection de $\mathfrak{S}_n$). On a $\varepsilon_\sigma=\varepsilon_\tau\varepsilon_{\sigma'}=-\varepsilon_{\sigma'}$. Il vient :

$$f(v_1,\dots,v_n)=-\sum_{\sigma'\in\mathfrak{S}_n}\varepsilon_{\sigma'}a_{\sigma'(\tau(1)),1}\cdots a_{\sigma'(\tau(n)),n}$$

  Or $\sigma'(\tau(k))=\sigma'(k)$ si $k\neq i,j$ et $\sigma'(\tau(i))=\sigma'(j),\ \sigma'(\tau(j))=\sigma'(i)$. Comme $v_i=v_j$ on a dans tous les cas $a_{\sigma'(\tau(k)),q}=a_{\sigma'(k),q}$ et donc $f(v_1,\dots,v_n)=-f(v_1,\dots,v_n)$. Comme $\mathbb{K}$ n'est pas de caractéristique 2, on a bien $f(v_1,\dots,v_n)=0$ ce qui achève la preuve.

COROLLAIRE 12.2.1
Soit $\varphi\in\Lambda_n(E)$. Pour tout système $S\in E^n$ on a $\varphi(S)=\varphi(\mathcal{B})\det_{\mathcal{B}}(S)$.

En effet, il existe $\lambda\in\mathbb{K}$ tel que $\varphi=\lambda\det_{\mathcal{B}}$. On en déduit $\varphi(\mathcal{B})=\lambda\det_{\mathcal{B}}(\mathcal{B})=\lambda$ d'où la conclusion.

COROLLAIRE 12.2.2 (Changement de base)
Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E. Pour tout système $S\in E^n$ on a $\det_{\mathcal{B}'}(S)=\det_{\mathcal{B}'}(\mathcal{B})\det_B(S)$.
En particulier $\det_{\mathcal{B}'}(\mathcal{B})\det_{\mathcal{B}}(\mathcal{B}')=1$ et donc $\det_{\mathcal{B}}(\mathcal{B}')\neq 0$.

THEOREME 12.2.2
Soit E un $\mathbb{K}$-ev de dimension n et S un système de n vecteurs de E. Les propriétés suivantes sont équivalentes :
1) *S est libre.*
2) *Il existe une base $\mathcal{B}$ de E telle que $\det_{\mathcal{B}}(S)\neq 0$.*
3) *Pour toute base $\mathcal{B}$ de E, $\det_{\mathcal{B}}(S)\neq 0$.*

preuve
1) $\Rightarrow$ 2) Si $S\in E^n$ est libre, c'est une base de E. En prenant $\mathcal{B}=S$ on a $\det_S(S)=1$.
2) $\Rightarrow$ 3) Soit $\mathcal{B}_0$ une base telle que $\det_{\mathcal{B}_0}(S)\neq 0$ et $\mathcal{B}$ une base quelconque. Alors $\det_{\mathcal{B}}(S)=\det_{\mathcal{B}_0}(S)\det_{\mathcal{B}}(\mathcal{B}_0)\neq 0$.
3) $\Rightarrow$ 2) car E possède des bases.
2) $\Rightarrow$ 1) Soit $\mathcal{B}$ une base de E. Si S était lié, on aurait $\det_{\mathcal{B}}(S)=0$ car $\det_{\mathcal{B}}$ est une forme n-linéaire alternée. Donc S est libre.

12.2.2 Déterminant d'un endomorphisme

Soit E un $\mathbb{K}$-ev de dimension n. Soit $f \in L(E)$. Pour $\varphi \in \Lambda_n(E)$ on définit $f^*(\varphi) : E^n \to \mathbb{K}$ par $f^*(\varphi)(v_1, \ldots, v_n) = \varphi(f(v_1), \ldots, f(v_n))$. $f^*(\varphi)$ est une n-forme linéaire alternée et l'application $\Lambda_n(E) \to \Lambda_n(E)$ qui à φ associe $f^*(\varphi)$ est linéaire. Comme $\dim(\Lambda_n(E)) = 1$, cette application est de la forme $\varphi \to t\varphi$. On a donc prouvé :

PROPOSITION 12.2.1
Soit E un $\mathbb{K}$-ev de dimension n et $f \in L(E)$. Il existe un unique $d(f) \in \mathbb{K}$ tel que, pour toute $\varphi \in \Lambda_n(E)$ et tout système $(v_1, \ldots, v_n) \in E^n$ on ait $\varphi(f(v_1), \ldots, f(v_n)) = d(f)\varphi(v_1, \ldots, v_n)$.

DEFINITION 12.2.1
Le scalaire $d(f)$ ainsi défini s'appelle le déterminant de l'endomorphisme f et se note $\det(f)$.

En particulier, si $\mathcal{B}$ est une base de E et $f \in L(E)$ on a pour tout système $(v_1, \ldots, v_n) \in E^n$,

$$\det_{\mathcal{B}}(f(v_1), \ldots, f(v_n)) = \det(f) \det_{\mathcal{B}}(v_1, \ldots, v_n) \tag{12.3}$$

Si $\mathcal{B} = (e_1, \ldots, e_n)$, on en déduit (en prenant $v_i = e_i$)

$$\det(f) = \det_{\mathcal{B}}(f(e_1), \ldots f(e_n)) \tag{12.4}$$

PROPOSITION 12.2.2
1) *Soient $f, g \in L(E)$. On a $\det(g \circ f) = \det(g)\det(f)$.*
2) $\forall a \in \mathbb{K},\ \forall f \in L(E),\ \det(af) = a^n \det(f)$.
3) $\det(id_E) = 1$.

preuve
Soit $\varphi \in \Lambda_n(E)$ non nulle.
1) Soit $(e_1, \ldots, e_n)$ une base de E. On a d'une part $\varphi(g \circ f(e_1), \ldots, g \circ f(e_n)) = \det(g \circ f)\varphi(e_1, \ldots, e_n)$ et d'autre part $\varphi(g \circ f(e_1), \ldots, g \circ f(e_n)) = \det(g)\varphi(f(e_1), \ldots, f(e_n)) = \det(g)\det(f)\varphi(e_1, \ldots, e_n)$. Le résultat découle du fait que $\varphi(e_1, \ldots, e_n) \neq 0$.
2) $\det(af)\varphi(e_1, \ldots, e_n) = \varphi(af(e_1), \ldots, af(e_n)) = a^n\varphi(f(e_1), \ldots f(e_n)) = a^n \det(f)\varphi(e_1, \ldots, e_n)$.
3) trivial.

PROPOSITION 12.2.3
Soit E un $\mathbb{K}$-ev de dimension n.
1) *Soit $f \in L(E)$; alors $f \in GL(E) \Leftrightarrow \det(f) \neq 0$.*
2) *Soit $f \in GL(E)$. On a $\det\left(f^{-1}\right) = \dfrac{1}{\det(f)}$.*

preuve
1) Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. On a $\det(f) = \det_B(f(e_1), \ldots, f(e_n))$. Or f est un automorphisme ssi l'image par f de la base $\mathcal{B}$ est une base, i.e. ssi le système $f(e_1), \ldots, f(e_n)$ est libre. Mais on a vu qu'un système de n vecteurs était libre ssi son déterminant dans une base quelconque était non nul. D'où la conclusion.
2) Résulte de la proposition précédente : $1 = \det(id_E) = \det\left(f \circ f^{-1}\right) = \det(f)\det\left(f^{-1}\right)$.

12.2.3 Déterminant d'une matrice carrée

DEFINITION 12.2.2
Soit $A = (a_{i,j}) \in M_n(\mathbb{K})$. On appelle déterminant de A et on note $\det(A)$ le scalaire défini par

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon_\sigma a_{\sigma(1),1} \cdots a_{\sigma(n),n} \tag{12.5}$$

Par définition, $\det(A)$ est le déterminant dans la base canonique de $\mathbb{K}^n$ du système $C_1(A), \ldots, C_n(A)$ des colonnes de A. C'est aussi le déterminant de l'application $X \to AX$ de $\mathbb{K}^n = M_{n,1}(\mathbb{K})$ dans lui même d'après l'égalité (12.4).

PROPOSITION 12.2.4
1) $\det(I_n) = 1$.
2) $\forall A, B \in M_n(\mathbb{K}),\ \det(AB) = \det(A)\det(B)$.
3) $\forall A \in M_n(\mathbb{K}),\ \forall a \in \mathbb{K},\ \det(aA) = a^n \det(A)$.
4) *Soit* $A \in M_n(\mathbb{K})$. *On a* A *inversible* $\Leftrightarrow \det(A) \neq 0$.
5) *Soit* $A \in GL(n, \mathbb{K})$. *on a* $\det\left(A^{-1}\right) = \dfrac{1}{\det(A)}$.

Les assertions 1) à 5) se déduisent des propriétés correspondantes pour les endomorphismes de $\mathbb{K}^n$.

On en déduit en particulier que $\det\left(A^k\right) = (\det(A))^k$ pour tout entier naturel k et pour tout entier relatif k si $A \in GL(n, \mathbb{K})$.

PROPOSITION 12.2.5
Une matrice A *et sa transposée ont même déterminant.*

preuve
Soit $A = (a_{i,j})$ et $B = {}^tA = (b_{i,j})$. On a donc $b_{i,j} = a_{j,i}$. Il vient

$$\det({}^tA) = \sum_{\sigma \in \mathfrak{S}} \varepsilon_\sigma b_{\sigma(1),1} \cdots b_{\sigma(n),n} = \sum_{\sigma \in \mathfrak{S}} \varepsilon_\sigma a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} = \sum_{\sigma \in \mathfrak{S}} \varepsilon_\sigma a_{\sigma^{-1}(\sigma(1)),\sigma(1)} \cdots a_{\sigma^{-1}(\sigma(n)),\sigma(n)}$$

Or $(\sigma(1), \ldots, \sigma(n))$ est une permutation de $(1, 2, \ldots, n)$ et la multiplication dans $\mathbb{K}$ est commutative. On peut donc dans le produit $a_{\sigma^{-1}(\sigma(1)),\sigma(1)} \cdots a_{\sigma^{-1}(\sigma(n)),\sigma(n)}$ réordonner les termes suivant le second indice. Il vient

$$a_{\sigma^{-1}(\sigma(1)),\sigma(1)} \cdots a_{\sigma^{-1}(\sigma(n)),\sigma(n)} = a_{\sigma^{-1}(1),1} \cdots a_{\sigma^{-1}(n),n}$$

Enfin,$\sigma \to \sigma^{-1}$ est une bijection de $\mathfrak{S}$ et pour tout $\sigma \in \mathfrak{S}$, on a $\varepsilon_\sigma = \varepsilon_{\sigma^{-1}}$. On en déduit en faisant le changement de variables $\tau = \sigma^{-1}$,

$$\det({}^tA) = \sum_{\tau \in \mathfrak{S}} \varepsilon_\tau a_{\tau(1),1} \cdots a_{\tau(n),n} = \det(A)$$

COROLLAIRE 12.2.3
Le déterminant d'une matrice est une forme n*-linéaire alternée des vecteurs lignes de la matrice.*

THEOREME 12.2.3
Soient E *un* $\mathbb{K}$*-espace vectoriel de dimension finie* $n \geqslant 1$ *et* $f \in L(E)$. *Soient* $\mathcal{B}$ *une base de* E *et* $A = \mathrm{Mat}_{\mathcal{B}}(f)$. *On a* $\det(f) = \det(A)$.

preuve
$\det(f) = \det_{\mathcal{B}}(f(e_1), \ldots, f(e_n)) = \det(A)$ d'après (12.4), (12.2) et la définition du déterminant de A.

12.2.4 Calculs de déterminants. Comatrice

Lemme 12.2.1
Soit $M = (m_{i,j})_{1 \leqslant i,j \leqslant n} \in M_n(\mathbb{K})$ *vérifiant* $m_{i,n} = 0$ *pour* $1 \leqslant i \leqslant n-1$ *et* $m_{n,n} = 1$. *Soit* $M' = (m_{i,j})_{1 \leqslant i,j \leqslant n-1}$ *la matrice extraite obtenue en supprimant dans* M *la dernière ligne et la dernière colonne. On a* $\det(M) = \det(M')$.

Autrement dit :
$$\begin{vmatrix} m_{1,1} & \ldots & m_{1,n-1} & 0 \\ \vdots & & \vdots & 0 \\ m_{n-1,1} & \ldots & m_{n-1,n-1} & 0 \\ m_{n,1} & \ldots & m_{n,n-1} & 1 \end{vmatrix} = \begin{vmatrix} m_{1,1} & \ldots & m_{1,n-1} \\ \vdots & & \vdots \\ m_{n-1,1} & \ldots & m_{n-1,n-1} \end{vmatrix}$$

preuve
Soit $\sigma \in \mathfrak{S}_n$. Si $\sigma(n) \neq n$ on a $m_{\sigma(1),1} \cdots m_{\sigma(n),n} = 0$. Notons $\mathfrak{S}'_n = \{\sigma \in \mathfrak{S}_n \mid \sigma(n) = n\}$. Si $\sigma \in \mathfrak{S}'_n$, la restriction de σ à $\mathbb{N}^*_{n-1}$ induit une bijection de $\mathbb{N}^*_{n-1}$ que nous noterons σ' et $\sigma \to \sigma'$ est une bijection de $\mathfrak{S}'_n$ sur $\mathfrak{S}_{n-1}$. En outre, il est immédiat que $\varepsilon_\sigma = \varepsilon_{\sigma'}$. Enfin, si $\sigma \in \mathfrak{S}'_n$, on a $m_{\sigma(1),1} \cdots m_{\sigma(n),n} = m_{\sigma'(1),1} \cdots m_{\sigma'(n-1),n-1}$. Le lemme en découle :

$$\det(M) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon_\sigma m_{\sigma(1),1} \cdots m_{\sigma(n),n} = \sum_{\sigma \in \mathfrak{S}'_n} \varepsilon_\sigma m_{\sigma(1),1} \cdots m_{\sigma(n),n} = \sum_{\sigma' \in \mathfrak{S}_{n-1}} \varepsilon_{\sigma'} m_{\sigma'(1),1} \cdots m_{\sigma'(n-1),n-1} = \det(M').$$

COROLLAIRE 12.2.4 (Déterminant d'une matrice triangulaire)
Le déterminant d'une matrice triangulaire est égal au produit des éléments diagonaux de la matrice : si $T = (t_{i,j}) \in M_n(\mathbb{K})$ *est une matrice triangulaire,* $\det(T) = \prod_{1 \leqslant i \leqslant n} t_{i,i}$.

DEFINITION 12.2.3
Soit $A = (a_{i,j}) \in M_n(\mathbb{K})$ *et* $(i,j) \in {\mathbb{N}_n^*}^2$. *On appelle mineur d'indice* (i,j) *de* A *le déterminant* $\Delta_{i,j}(A)$ *de la matrice d'ordre* $n-1$ *obtenue en supprimant dans* A *la* i*-ième ligne et la* j*-ième colonne. On appelle cofacteur d'indice* (i,j) *de* A *le scalaire* $A_{i,j} = (-1)^{i+j}\Delta_{i,j}(A)$. *Enfin on appelle comatrice de* A *la matrice des cofacteurs, i.e.* $\mathrm{com}(A) = (A_{i,j})$.

THEOREME 12.2.4 (Développement d'un déterminant suivant une colonne)
Soit $A = (a_{i,j}) \in M_n(\mathbb{K})$. *On a*

$$\forall j \in \mathbb{N}_n^*,\ \det(A) = \sum_{i=1}^{i=n} a_{i,j}A_{i,j}$$

COROLLAIRE 12.2.5 (Développement d'un déterminant suivant une ligne)
Soit $A = (a_{i,j}) \in M_n(\mathbb{K})$. *On a*

$$\forall i \in \mathbb{N}_n^*,\ \det(A) = \sum_{j=1}^{j=n} a_{i,j}A_{i,j}$$

preuve

Soit $\mathcal{B} = (e_1, \ldots, e_n)$ la base canonique de $M_{n,1}(\mathbb{K})$ et $(C_1, \ldots, C_n)$ les colonnes de A. On a $C_j = \sum_{i=1}^{i=n} a_{i,j}e_i$ et $\det(A) = \det_{\mathcal{B}}(C_1, \ldots, C_n)$ donc

$$\det(A) = \det_{\mathcal{B}}\left(C_1, \ldots, C_{j-1}, \sum_{i=1}^{i=n} a_{i,j}e_i, C_{j+1}, \ldots, C_n\right) = \sum_{i=1}^{i=n} a_{i,j}\det_{\mathcal{B}}(C_1, \ldots, C_{j-1}, e_i, C_{j+1}, \ldots, C_n)$$

Posons $R_{i,j} = \det_{\mathcal{B}}(C_1, \ldots, C_{j-1}, e_i, C_{j+1}, \ldots, C_n)$. Il reste à voir que $R_{i,j} = A_{i,j}$.

Pour calculer $R_{i,j}$ on va d'abord amener la j-ième colonne en n-ième position sans changer l'ordre des autres. Cette opération multiplie $R_{i,j}$ par $(-1)^{n-j}$. Formellement on effectue sur les colonnes la permutation $(1, \ldots, j-1, j+1, \ldots, n-1, j)$ dont le nombre d'inversion est $n-j$. On fait ensuite la même chose avec les lignes en amenant la i-ième ligne en dernière position, ce qui multiplie le déterminant par $(-1)^{n-i}$. Finalement on obtient

$$R_{i,j} = \begin{vmatrix} a_{1,1} & \ldots & a_{1,j-1} & 0 & a_{1,j+1} & \ldots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & 0 & \vdots & & \vdots \\ \vdots & & \vdots & 1 & \vdots & & \vdots \\ \vdots & & \vdots & 0 & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \ldots & a_{n,j-1} & 0 & a_{n,j+1} & \ldots & a_{n,n} \end{vmatrix} = (-1)^{2n-i-j} \begin{vmatrix} a_{1,1} & \ldots & a_{1,j-1} & a_{1,j+1} & \ldots & a_{1,n} & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ a_{i-1,1} & \ldots & a_{i-1,j-1} & a_{i-1,j+1} & \ldots & a_{i-1,n} & 0 \\ a_{i+1,1} & \ldots & a_{i+1,j-1} & a_{i+1,j+1} & \ldots & a_{i+1,n} & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ a_{n,1} & \ldots & a_{n,j-1} & a_{n,j+1} & \ldots & a_{n,n} & 0 \\ a_{i,1} & \ldots & a_{i,j-1} & a_{i,j+1} & \ldots & a_{i,n} & 1 \end{vmatrix}$$

Soit $R_{i,j} = (-1)^{i+j}\Delta_{i,j}(A) = A_{i,j}$. ■
Le corollaire s'en déduit du théorème en utilisant $\det(A) = \det({}^tA)$.

THEOREME 12.2.5
Soit $A \in M_n(\mathbb{K})$. *On a*

$$A\,{}^t\mathrm{com}(A) = {}^t\mathrm{com}(A)A = \det(A)I_n$$

preuve

Soit $A = (a_{i,j})$ et $A_{i,j}$ le cofacteur de A d'indice (i,j). Pour $1 \leqslant i, j \leqslant n$ on a $\left(A\,{}^t\mathrm{com}(A)\right)_{i,j} = \sum_{k=1}^{k=n} a_{i,k}A_{j,k}$. D'après le

théorème précédent, on a déjà $(A\,{}^t\mathrm{com}(A))_{i,i} = \det(A)$ pour tout i. Supposons maintenant $i \neq j$. Soit $A' = (a'_{i,j})$ la matrice obtenue en remplaçant dans A la j-ième colonne par la i-ième : $a'_{p,q} = a_{p,q}$ pour tout p et tout $q \neq j$ et $a'_{p,j} = a_{p,i}$ pour tout p. La matrice A' a deux colonnes identiques donc $\det(A') = 0$. Développons $\det(A')$ par rapport à la j-iéme colonne. Il vient $\det(A') = \sum_{k=1}^{k=n} a'_{k,j} A'_{k,j}$. Par construction, $a'_{k,j} = a_{k,i}$. D'autre part, la matrice d'ordre $n-1$ obtenue en supprimant dans A' la i-ième ligne et la j-ième colonne est par construction égale à celle obtenue en supprimant dans A la i-ième ligne et la j-ième colonne. Donc $A'_{k,j} = A_{k,j}$. On obtient $0 = \sum_{k=1}^{k=n} a_{k,i} A_{k,j}$. D'où finalement $A\,{}^t\mathrm{com}(A) = \det(A) = I_n$.
La seconde relation se prouve en appliquant la première à la matrice ${}^tA$, en remarquant que $\mathrm{com}({}^tA) = {}^t\mathrm{com}(A)$ puis en transposant l'égalité obtenue.

COROLLAIRE 12.2.6

$$\forall A \in GL(n,\mathbb{K}), \quad A^{-1} = \frac{1}{\det(A)}\,{}^t\mathrm{com}(A)$$

13

Systèmes linéaires

$\mathbb{K}$ est un sous corps de $\mathbb{C}$.

13.1 Notations et définitions

On considère un système linéaire de n équations à p inconnues $x_1, \dots x_p$:

$$(S) \qquad \begin{cases} a_{1,1}x_1 + \cdots + a_{1,p}x_p &= b_1 \\ \\ a_{n,1}x_1 + \cdots + a_{n,p}x_p &= b_n \end{cases}$$

où les $a_{i,j}$ et les b_i sont des éléments donnés de $\mathbb{K}$.

Ce système s'écrit aussi

$$(S) \qquad AX = B$$

avec $A = (a_{i,j}) \in M_{n,p}(\mathbb{K}),\ X \in M_{p,1}(\mathbb{K}),\ B \in M_{n,1}(\mathbb{K})$. Le système homogène associé est

$$(S_0) \qquad AX = 0$$

On notera $\mathrm{Sol}(S)$ le sous ensemble de $\mathbb{K}^p$ formé des solutions de (S).

On introduira la matrice complète du système $A' = [A|B] \in M_{n,p+1}(\mathbb{K})$ obtenue en rajoutant à A une colonne supplémentaire égale à B.

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix} \qquad A' = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{n,1} & \cdots & a_{n,p} & b_p \end{pmatrix}$$

DEFINITION 13.1.1
Deux systèmes (S) et (S') ayant le même nombre d'inconnnues p seront dits équivalents si $\mathrm{Sol}(S) = \mathrm{Sol}(S')$.

Interprétations

1. Soit $f \in L(\mathbb{K}^p, \mathbb{K}^n)$ de matrice A dans la base canonique. On cherche les éléments X de $\mathbb{K}^p$ tels que $f(X) = B$.
2. Soient $C_1, \dots, C_p \in \mathbb{K}^n$ les colonnes de A. On cherche $(x_1, \dots, x_p)$ tels que $B = x_1C_1 + \cdots x_pC_p$, autrement dit on cherche à décomposer B sur le système de vecteurs $(C_1, \dots, C_p)$.

Conséquences

1. L'ensemble $\mathrm{Sol}(S_0)$ est un sous espace vectoriel de $\mathbb{K}^p$: c'est le noyau de f.
2. Si $\lambda = (\lambda_1, \ldots, \lambda_p)$ est une solution de (S), $\mathrm{Sol}(S) = \{\lambda + z \ ; \ z \in \mathrm{Sol}(S_0)\}$. Donc si l'ensemble des solutions de (S) est non vide, c'est un sous espace affine de $\mathbb{K}^p$ d'espace vectoriel directeur $\mathrm{Sol}(S_0)$.
3. Il en résulte que tout système linéaire admet soit 0, soit une et une seule, soit une infinité de solutions (car le corps de base $\mathbb{K}$ est infini).

DEFINITION 13.1.2
Le système (S) est dit compatible si il admet des solutions. Le rang du système (S) est par définition celui de la matrice A.

THEOREME 13.1.1
(S) *est compatible* $\Leftrightarrow \mathrm{rg}(A) = \mathrm{rg}(A')$

En effet, $\mathrm{rg}(A) = \mathrm{rg}(A')$ équivaut au fait que B appartient à l'espace vectoriel engendré par les colonnes de A.

THEOREME 13.1.2
Si le rang d'un système (S) est égal au nombre d'équations, ce système est toujours compatible.

Si le rang de A est égal à n, les colonnes de A engendrent $\mathbb{K}^n$, donc B appartient à l'espace engendré par ces colonnes.

13.2 Système de Cramer

DEFINITION 13.2.1
Le système (S) est dit de Cramer si le nombre d'inconnues est égal au nombre d'équations et si la matrice carrée A est inversible.

Un système de Cramer admet évidemment une unique solution $X = A^{-1}B$. Le fait qu'un système soit de Cramer ne dépend pas du "second membre" B.

THEOREME 13.2.1
Soit $A \in M_n(\mathbb{K})$. Les propriétés suivantes sont équivalentes :
1) A est inversible.
2) Pour tout $B \in \mathbb{K}^n$, le système $AX = B$ a au moins une solution.
3) Il existe $B \in \mathbb{K}^n$ tel que le système $AX = B$ ait une solution unique.
4) (S_0) admet l'unique solution 0.

C'est une autre manière d'écrire les caractérisations des matrices inversibles.

Formules de Cramer

Soit A_k la matrice obtenue en remplaçant dans A la k-ième colonne par B. Autrement dit $C_j(A_k) = C_j(A)$ si $j \neq k$ et $C_k(A_k) = B$.

THEOREME 13.2.2
Soit $(S) : AX = B$ un système de Cramer. L'unique solution de (S) est donnée par

$$x_k = \frac{\det(A_k)}{\det(A)} \quad 1 \leqslant k \leqslant n$$

preuve
Puisque A est inversible, (S) a une solution unique $(x_1, \ldots, x_n)$ et on a $B = x_1C_1 + \cdots + x_nC_n$ en notant $C_k = C_k(A)$. Alors

$$\det(A_k) = \det(C_1, \ldots, C_{k-1}, \sum_{j=1}^{j=n} x_j C_j, \ldots, C_n) = \sum_{j=1}^{j=n} x_j \det(C_1, \ldots, C_{k-1}, C_j, C_{k+1}, \ldots, C_n)$$

Les déterminants figurant au second membre sont nuls si $j \neq k$ puisqu'alors ils ont deux colonnes égales. Il reste $\det(A_k) = x_k \det(C_1, \ldots, C_n) = x_k \det(A)$.

13.3 Etude du cas général

13.3.1 Définitions

On revient à l'étude d'un système (S) $AX = B$. On note r le rang de A et $E_1, \ldots, E_n$ les n équations du système. On garde les notations du paragraphe 6.8. On sait que l'on peut trouver une matrice extraite $A_{I,J}$ inversible carrée d'ordre r où $I \subset \mathbb{N}_n^*$ et $J \subset \mathbb{N}_p^*$. En général, une telle matrice n'est pas unique. Cette matrice une fois choisie s'appelle matrice principale. Les inconnues x_j telles que $j \in J$ s'appellent les inconnues principales, les équations E_i telles que $i \in I$ s'appellent les équations principales.

13.3.2 Etude du système homogène

THEOREME 13.3.1
Le système homogène associé (S_0) est équivalent au système (S_0') obtenu en ne gardant que les équations principales. Les solutions de (S_0) s'obtiennent en donnant des valeurs arbitraires aux inconnues non principales et en résolvant le système obtenu en les inconnues principales qui est un système de Cramer.

preuve
Quitte à permuter les équations et les inconnues, on peut supposer $I = J = \mathbb{N}_r^*$. La sous matrice principale est donc $\tilde{A} = A_{\mathbb{N}_r^*,\mathbb{N}_r^*}$. Soit $A_0 = A_{I,\mathbb{N}_n^*}$ la matrice obtenue en ne gardant que les r premières lignes de A. Soit $\lambda > r$. La ligne $L_\lambda(A)$ est combinaison linéaire des lignes $L_1(A), \ldots, L_r(A)$. Donc l'équation (E_λ) est conséquence des équations $(E_1), \ldots, (E_r)$. Autrement dit, si $(x_1, \ldots, x_p)$ est solution des r premières équations, il est solution de (S_0). La réciproque est triviale. Donc (S_0) est équivalent au système (S_0') formé des r premières équations, système dont la matrice est A_0.
Soit $X = (x_1, \ldots, x_p) \in \mathbb{K}^p$. On a

$$(S_0) \Leftrightarrow A_0 X = 0 \Leftrightarrow (S_0') \quad \left\{ \begin{array}{rcl} a_{1,1}x_1 + \cdots + a_{1,r}x_r & = & -(a_{1,r+1}x_{r+1} + \cdots + a_{1,p}x_p) \\ & & \\ a_{r,1}x_1 + \cdots + a_{r,r}x_r & = & -(a_{r,r+1}x_{r+1} + \cdots + a_{r,p}x_p) \end{array} \right.$$

Pour $(x_{r+1}, \ldots, x_p)$ fixés quelconques, le système (S_0') en les inconnues principales $x_1, \ldots, x_r$ est de Cramer et admet une unique solution $(x_1, \ldots, x_r)$. Alors $(x_1, \ldots, x_p)$ est solution de (S_0'), donc de (S_0). La réciproque est triviale.

13.3.3 Conditions de compatibilité du système complet

Revenons au système complet (S). Si $r = n$, le système est compatible. Supposons $r < n$. Notons encore $\tilde{A}$ une matrice principale extraite de A. On a

$$\mathrm{rg}(A) = r \Leftrightarrow \text{ toutes les sous matrices de } A \text{ bordantes de } \tilde{A} \text{ sont non inversibles}$$

$$\mathrm{rg}(A') = r \Leftrightarrow \text{toutes les sous matrices de } A' \text{ bordantes de } \tilde{A} \text{ sont non inversibles}$$

On sait que $\mathrm{rg}(A) = r$ et on a vu que (S) était compatible ssi $\mathrm{rg}(A') = r$. Par conséquent, (S) est compatibles ssi les sous matrices de A', bordantes de $\tilde{A}$, obtenues en bordant par la colonne B sont non inversibles.

DEFINITION 13.3.1
On appelle déterminants caractéristiques du système (S) (relatifs au choix de la sous matrice principale $\tilde{A}$) les déterminants des matrices bordantes de $\tilde{A}$ dans A' de la forme

$$\Delta_i = \det(A'_{I\cup\{i\},J\cup\{p+1\}}), \;\; 1 \leqslant i \leqslant n, \; i \notin I$$

Si la sous matrice principale est $\tilde{A} = A_{\mathbb{N}_r^*,\mathbb{N}_r^*}$ les déterminants caractéristiques sont les déterminants

$$\Delta_i = \begin{vmatrix} a_{1,1} & \cdots & a_{r,1} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{r,1} & \cdots & a_{r,r} & b_r \\ a_{i,1} & \cdots & a_{i,r} & b_i \end{vmatrix} \qquad r+1 \leqslant i \leqslant n$$

On a donc montré

THEOREME 13.3.2
Le système (S) est compatible ssi les $n-r$ déterminants caractéristiques sont nuls.

13.3.4 Solutions d'un système compatible

THEOREME 13.3.3 (Rouché-Fontené)
Soit (S) un système de n équations à p inconnues. Supposons choisie une sous matrice principale. Soit (S') le système formé par les équations principales.
Si le système (S) est compatible, c'est à dire si les $n-r$ déterminants caractéristiques sont nuls, (S) est équivalent à (S'). On résoud (S') en donnant des valeurs arbitraires aux inconnues non principales et en résolvant le système obtenu (aux inconnues principales) qui est alors de Cramer.

La preuve est la même que pour les systèmes homogènes.

13.3.5 Exemple

Résoudre et discuter le système

$$(S) \qquad \begin{cases} x + \alpha y + \beta z = 1 \\ \alpha x + \alpha\beta y + z = \alpha \\ \beta x + \alpha y + z = 1 \end{cases} \qquad \text{où } \alpha, \beta \text{ sont des paramètres reéls}$$

Soit $A = \begin{pmatrix} 1 & \alpha & \beta \\ \alpha & \alpha\beta & 1 \\ \beta & \alpha & 1 \end{pmatrix}$. $\det(A) = \alpha \underset{L_1 \leftarrow L_1 - L_3}{\begin{vmatrix} 1-\beta & 0 & \beta-1 \\ \alpha & \beta & 1 \\ \beta & 1 & 1 \end{vmatrix}} = \alpha(\beta-1) \underset{C_1 \leftarrow C_1 + C_3}{\begin{vmatrix} 0 & 0 & 1 \\ \alpha+1 & \beta & 1 \\ \beta+1 & 1 & 1 \end{vmatrix}} = \alpha(\beta-1)(\alpha+1-\beta-\beta^2)$

On notera $A_{i,j}$ (resp. $D_{i,j}$) la matrice (resp. le déterminant de la matrice) d'ordre 2 extraite de A en supprimant la ligne i et la colonne j. Donc avec les notations précédentes $A_{i,j} = A_{\mathbb{N}_3^* \setminus \{i\}, \mathbb{N}_3^* \setminus \{j\}}$.

1. $\alpha(\beta-1)(\alpha+1-\beta-\beta^2) \neq 0$.
   Le système est de Cramer. Sa solution est donnée par les formules de Cramer. On obtient

$$x = \frac{(\alpha-\beta)}{\alpha+1-\beta-\beta^2} \quad y = \frac{1-\alpha\beta}{\alpha(\alpha+1-\beta-\beta^2)} \quad z = \frac{\alpha-\beta}{\alpha+1-\beta-\beta^2}$$

2. $\beta = 1$. $A = \begin{pmatrix} 1 & \alpha & 1 \\ \alpha & \alpha & 1 \\ 1 & \alpha & 1 \end{pmatrix}$.
   (a) Si $\alpha = 1$, A est de rang 1, (S) est équivalent à l'équation $x+y+z=1$. $\text{Sol}(S) = \{(1-y-z, y, z)\ ;\ (y,z) \in \mathbb{R}^2\}$.
   (b) Si $\alpha \neq 1$, $D_{1,2} = \begin{vmatrix} \alpha & 1 \\ 1 & 1 \end{vmatrix} = \alpha - 1 \neq 0$. A est de rang 2. Le système est trivialement équivalent au système $\begin{cases} \alpha x + \alpha y + z = \alpha \\ x + \alpha y + z = 1 \end{cases} \Leftrightarrow \begin{cases} x = 1 \\ \alpha y + z = 0 \end{cases}$. $\text{Sol}(S) = \{(1, y, -\alpha y)\ ;\ y \in \mathbb{R}\}$.
   On a choisi comme matrice principale extraite la matrice $A_{1,2}$. Les inconnues principales sont x et z et les équations principales les deux dernières équations. Le déterminant caractéristique est $\begin{vmatrix} 1 & 1 & 1 \\ \alpha & 1 & \alpha \\ 1 & 1 & 1 \end{vmatrix} = 0$

3. $\beta \neq 1$ et $\alpha = 0$. $A = \begin{pmatrix} 1 & 0 & \beta \\ 0 & 0 & 1 \\ \beta & 0 & 1 \end{pmatrix}$.
   $D_{3,2} = \begin{vmatrix} 1 & \beta \\ 0 & 1 \end{vmatrix} = 1$. $\text{rg}(A) = 2$. $(S) \Leftrightarrow \begin{cases} x + \beta z &= 1 \\ z &= 0 \\ \beta x + z &= 1 \end{cases}$ (S) est non compatible, $\text{Sol}(S) = \emptyset$.

4. $\beta \neq 1,\ \alpha \neq 0,\ \alpha = \beta^2 + \beta - 1$.
   $D_{1,1} = \begin{vmatrix} \alpha\beta & 1 \\ \alpha & 1 \end{vmatrix} = \alpha(\beta - 1) \neq 0$. Le système est de rang 2. On chosit $A_{1,1}$ comme sous matrice principale. Les équations principales sont donc les deux dernières et les inconnues principales y et z. Il y a un seul caractéristique, à savoir $\delta = \begin{vmatrix} \alpha & \beta & 1 \\ \alpha\beta & 1 & \alpha \\ \alpha & 1 & 1 \end{vmatrix} = \alpha(\beta - 1)(\alpha - \beta)$. δ est nul ssi $\alpha = \beta$ ce qui compte tenu de $\alpha = \beta^2 + \beta - 1$ et $\beta \neq 1$ équivaut à $\alpha = \beta = -1$. Donc deux cas :

   (a) $(\alpha, \beta) \neq (-1, -1)$. Le système est impossible. $\text{Sol}(S) = \emptyset$.

   (b) $\alpha = \beta = -1$. Le système est équivalent à celui formé par les équations principales, soit $\begin{cases} y + z &= x - 1 \\ -y + z &= x + 1 \end{cases}$ qui est bien de Cramer en les inconnues principales. $\text{Sol}(S) = \{(x, -1, x)\ ;\ x \in \mathbb{R}\}$.

Conclusion

1. $\alpha(\beta - 1)(\alpha + 1 - \beta - \beta^2) \neq 0$. Système de Cramer, rang 3, solution unique.
2. $\beta = 1,\ \alpha \neq 1$. Rang 2. $\text{Sol}(S) = \{1, y, -\alpha y)\ ;\ y \in \mathbb{R}\}$.
3. $\beta = 1, \alpha = 1$. Rang 1. $\text{Sol}(S) = \{(1 - y - z, y, z)\ ;\ (y, z) \in \mathbb{R}^2\}$.
4. $\beta \neq 1,\ \alpha = 0$. Rang 2. Système impossible.
5. $\beta = -1,\ \alpha = -1$. Rang 2. $\text{Sol}(S) = \{(x, -1, x)\ ;\ x \in \mathbb{R}\}$
6. $\beta \neq \pm 1,\ \alpha \neq 0,\ \alpha = \beta^2 + \beta - 1$. Rang 2. Système impossible.

14

Opérations élémentaires sur les matrices

14.1 Introduction, notations

14.1.1 Base canonique de $M_{n,p}(\mathbb{K})$

Soient n, p des entiers $\geqslant 1$. Nous noterons $E_{i,j}(n,p)_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}$ la base canonique de $M_{n,p}(\mathbb{K})$. $E_{i,j}(n,p)$ est donc una matrice à n lignes et p colonnes dont tous les coefficients sont nuls, sauf celui situé à l'intersection de la i-ième ligne et de la j-ième colonne qui vaut 1. On a la règle de multiplication

$$E_{i,j}(n,p)E_{k,l}(p,q) = \delta_{j,k}E_{i,l}(n,q)$$

où $\delta_{j,k}$ est le symbole de Kroneckler, égal à 1 si $j = k$ et à 0 sinon. On notera $E_{i,j}(n)$ au lieu de $E_{i,j}(n,n)$ la base canonique de $M_n(\mathbb{K})$. On omettra souvent le (n) ou le (n,p) si le contexte est clair.

On notera $\mathbb{N}_n^* = \{1, 2, \ldots, n\}$.

14.1.2 Matrices de transvection

DEFINITION 14.1.1
Soient $n \in \mathbb{N},\ n \geqslant 2$, $i, j \in \mathbb{N}_n^*,\ i \neq j$ et $\lambda \in \mathbb{K}$. On définit

$$T_{i,j}^{(n)}(\lambda) = I_n + \lambda E_{i,j}(n)$$

Les matrices $T_{i,j}^{(n)}(\lambda)$ s'appellent matrices de transvection.

On omettra le plus souvent l'exposant (n).

On a pour $\lambda, \mu \in \mathbb{K}$

$$T_{i,j}(\lambda)T_{i,j}(\mu) = T_{i,j}(\lambda + \mu)$$

Comme $T_{i,j}(0) = I_n$, on en déduit que la matrice $T_{i,j}(\lambda)$ est inversible, d'inverse $T_{i,j}(-\lambda)$.
Par ailleurs $\det(T_{i,j}(\lambda)) = 1$.

14.1.3 Matrices de permutation

Soient $i, j \in \mathbb{N}_n^*$. On définit

$$P_{i,j}^{(n)} = I_n - E_{i,i}(n) - E_{j,j}(n) + E_{i,j}(n) + E_{j,i}(n)$$

soit

$$P_{i,j} = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & 0 & & 1 & & & \\ & & & & \ddots & & & & \\ & & & 1 & & 0 & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{pmatrix}$$

$P_{i,j}$ est la matrice de passage de la base $(e_1, \ldots, e_n)$ à la base $(e_1, \ldots, e_{i-1}, e_j, e_{i+1}, \ldots, e_{j-1}, e_i, e_{j+1}, \ldots, e_n)$.
On a $P_{i,i} = I_n$ et dans tous les cas $P_{i,j}^2 = I_n$ et $\det(P_{i,j}) = -1$ si $i \neq j$.

Plus généralement:

DEFINITION 14.1.2
Soit $\sigma \in \mathfrak{S}_n$ une permutation de $\mathbb{N}_n^*$. On notera $P_\sigma^{(n)}$ où simplement P_σ si il n'y a pas d'ambiguité la matrice de passage de la base $(e_1, \ldots, e_n)$ à la base $(e_{\sigma(1)}, \ldots, e_{\sigma(n)})$.

On a donc $(P_\sigma)_{i,j} = \begin{cases} 1 & \text{si } i = \sigma(j) \\ 0 & \text{si } i \neq \sigma(j) \end{cases}$ soit $(P_\sigma)_{i,j} = \delta_{i,\sigma(j)} = \delta_{\sigma^{-1}(i),j}$

Lemme 14.1.1
L'application $\sigma \to P_\sigma$ est un morphisme injectif de $\mathfrak{S}_n$ dans $GL(n, \mathbb{K})$

preuve
Soient $\sigma, \tau \in \mathfrak{S}_n$. On a $(P_\sigma P_\tau)_{i,j} = \sum_{k=1}^{k=n} \delta_{i,\sigma(k)}\delta_{k,\tau(j)} = \delta_{i,\sigma(\tau(j)} = (P_{\sigma\circ\tau})_{i,j}$ Le fait que l'application soit injective est immédiat.

Soient $i \neq j$ et $\tau(i,j)$ la transposition qui échange i et j. On a $P_{\tau(i,j)} = P_{i,j}$. Comme toute permutation est un produit de transpositions, on en déduit que toute matrice P_σ est un produit de matrices de la forme $P_{i,j}$ avec $i \neq j$. De plus la parité du nombre de ces matrices est bien déterminée.

COROLLAIRE 14.1.1
$\det(P_\sigma) = \varepsilon(\sigma)$ *où $\varepsilon(\sigma)$ est la signature de la permutation σ.*

14.1.4 Matrices de dilatation

DEFINITION 14.1.3
Pour $i \in \mathbb{N}_n^*$ et $r \in \mathbb{K}^*$ on définit

$$D_i(r) = I_n + (r-1)E_{i,i} = rE_{i,i} + \sum_{\substack{1 \leqslant j \leqslant n \\ j \neq i}} E_{j,j} = \text{diag}(1, \cdots, 1, r, 1, \cdots, 1)$$

Les matrices de la forme $D_i(r)$ s'appellent matrices de dilatation.

Pour $r, r' \in \mathbb{K}^*$ on a $D_i(r')D_i(r) = D_i(r'r)$. La matrice $D_i(r)$ est inversible et $D_i(r)^{-1} = D_i(1/r)$. On a $\det(D_i(r)) = r$.

14.2 Opérations élémentaires sur les colonnes

Soient E un $\mathbb{K}$-ev, $S = (v_1, \ldots, v_p)$ un système de p vecteurs de E et $F_S = \text{Vect}\,(v_1, \ldots, v_p)$. Le rang de S est le nombre maximum de vecteurs d'un système libre extrait de S. C'est aussi la dimension de F_S. Soient i, j tels que $1 \leqslant i < j \leqslant p,\ \lambda \in$

$\mathbb{K}$ et $r \in \mathbb{K}^*$. Les systèmes
$S_1 = (v_1, \ldots, v_{i-1}, v_i + \lambda v_j, v_{i+1}, \ldots, v_n)$
$S_2 = (v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n)$
$S_3 = (v_1, \ldots, v_{i-1}, rv_i, v_{i+1}, \ldots, v_n)$
engendrent clairement le même sous espace F que S. Ils ont donc même rang que S. On dit que S_k se déduit de S par une transformation élémentaire de type k. Soit S' un autre système de vecteurs de E. On dit que S' se déduit de S par une suite de transformations élémentaires si il existe une suite finie de systèmes $S^{[0)}, \ldots, S^{(m)}$ tels que $S^{(0)} = S,\ S^{(m)} = S'$ et pour tout k entre 0 et $m-1$, $S^{(k+1)}$ se déduit de $S^{(k)}$ par une transformation élémentaire. Si c'est le cas, on a $\mathrm{Vect}(S) = \mathrm{Vect}(S')$ et $\mathrm{rg}(S) = \mathrm{rg}(S')$.
Supposons maintenant E de dimension finie et soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E.
Soient $A = \mathrm{Mat}_{\mathcal{B}}(S) \in M_{n,p}(\mathbb{K}),\ A_k = \mathrm{Mat}_{\mathcal{B}}(S_k)$ pour $k = 1, 2, 3$.
(1) A_1 se déduit de A en ajoutant à la i-ième colonne de A la j-ième colonne de A multipliée par λ. Ceci nécessite $p \geqslant 2$.
(2) A_2 se déduit de A en échangeant les colonnes i et j.
(3) A_3 se déduit de A en multipliant la i-ième colonne de A par r.
On notera respectivement $C_i \leftarrow C_i + \lambda C_j,\ \ C_i \leftrightarrow C_j$ et $C_i \leftarrow rC_i$ les opérations élémentaires décrites ci dessus.

PROPOSITION 14.2.1
Soient $A \in M_{n,p}(\mathbb{K})$ et A_k définie par la relation (k) ci dessus ($k = 1, 2, 3$). On a

$$A_1 = AT_{j,i}^{(p)}(\lambda)\ \ A_2 = AP_{i,j}^{(p)}\ \ A_3 = AD_i^{(n)}(r)$$

A titre d'exemple, montrons la première relation

$$AT_{j,i}(\lambda) = A + \lambda \left(\sum_{\substack{1 \leqslant r \leqslant n \\ 1 \leqslant s \leqslant p}} a_{r,s} E_{r,s}(n,p) \right) E_{j,i}(p) = A + \lambda \sum_{\substack{1 \leqslant r \leqslant n \\ 1 \leqslant s \leqslant p}} a_{r,s} \delta_{s,j} E_{r,i}(p) = A + \lambda \sum_{r=1}^{r=n} a_{r,j} E_{r,i}(p)$$

D'où la conclusion.

Les opérations élémentaires sur les colonnes traduisent donc matriciellement des opérations élémentaires sur le système de vecteurs que A représente dans une base donnée. Elles ne modifient donc pas le rang du système des vecteurs colonnes de A.

14.3 Opérations élémentaires sur les lignes

DEFINITION 14.3.1
Soient $n, p \geqslant 1$, $A \in M_{n,p}(\mathbb{K})$.

1. Si $n \geqslant 2$, on appelle opération élémentaire de type 1 sur les lignes de A l'addition à une ligne de A du produit par un scalaire d'une autre ligne de A. Si A_1 est la matrice ainsi obtenue, on a donc deux indices distincts i et j tels que $L_k(A_1) = L_k(A)$ si $k \neq i$ et $L_i(A_1) = L_i(A) + \lambda L_j(A)$. On notera cette opération $L_i \leftarrow L_i + \lambda L_j$.

2. On appelle opération élémentaire de type 2 sur les lignes de A l'échange de deux lignes de A. Si A_2 est la matrice obtenue après une telle opération, on a donc deux indices i et j tels que $L_k(A_2) = L_k(A)$ si $k \neq i$ et $k \neq j$, $\ L_i(A_2) = L_j(A), L_j(A_2) = L_i(A)$. On notera cette opération $L_i \leftrightarrow L_j$.

*3. On appelle opération élémentaire de type 3 sur les lignes de A la multiplication d'une ligne de A par un scalaire **non nul**. Si A_3 est la matrice ainsi obtenue, on a un indice i et un scalaire $r \in \mathbb{K}^*$ tels que $L_k(A_3) = L_k(A)$ si $k \neq i$ et $L_i(A_3) = rL_i(A)$. On notera cette opération $L_i \leftarrow rL_i$.*

N.B. : La terminologie "de type 1, 2 ou 3 " n'est pas standard. Elle est utilisée ici pour l'exposé.

PROPOSITION 14.3.1
Avec les notations de la définition ci dessus on a

$$A_1 = T_{i,j}^{(n)}(\lambda)A,\ \ A_2 = P_{i,j}^{(n)}A,\ \ A_3 = D_i^{(n)}(r)A$$

Interprétation

Soit $A \in M_{n,p}(\mathbb{K})$. On considère A comme la matrice d'un système de p vecteurs $(v_1, \ldots, v_p)$ de $\mathbb{K}^n$ dans la base canonique $\mathcal{B}_0 = (e_1, \ldots, e_n)$ (ou plus généralement la matrice d'un système de p vecteurs d'un $\mathbb{K}$-ev E de dimension n dans une base $\mathcal{B}_0$ de E).

1. Soit $\mathcal{B}_1 = (e_1, \ldots, e_i, \ldots, e_{j-1}, e_j - \lambda e_i, \ldots, e_n)$. C'est une base de E. Soit $V = \sum_{k=1}^{k=n} x_k e_k$. On a

$$V = x_1 e_1 + \cdots + x_{i-1} e_{i-1} + (x_i + \lambda x_j) e_i + x_{i+1} e_{i+1} + \cdots + x_j (e_j - \lambda e_i) + \cdots + x_n e_n$$

La matrice X_1 des coordonnées dans la base $\mathcal{B}_1$ s'obtient donc à partir de la matrice X des coordonnées de V dans la base $\mathcal{B}_0$ par l'opération $L_i \leftarrow L_i + \lambda L_j$.

Donc faire une opération élémentaire de type 1 sur une matrice A revient à exprimer le même système de vecteurs dans une autre base.

2. Il en est de même pour les opérations de type 2 en prenant $\mathcal{B}_2 = (e_1, \ldots, e_{i-1}, e_j, e_{i+1}, \ldots, e_{j-1}, e_i, e_{j+1}, \ldots, e_n)$.

3. Et pour les opérations de type 3 en prenant $\mathcal{B}_3 = (e_1, \ldots, e_{i-1}, \frac{1}{r} e_i, e_{i+1}, \ldots, e_n)$.

Tout ceci résulte de calculs élémentaires. On peut aussi utiliser un héorème général : puisque $A = \mathrm{Mat}_{\mathcal{B}}(S)$, $A_1 = T_{i,j}^{(n)}(\lambda) A$ est la matrice du système (S) dans la base $\mathcal{B}_1$ telle que $\mathrm{Mat}_{\mathcal{B}}(\mathcal{B}_1) = T_{i,j}^{(n)}(\lambda)^{-1}$ et de même pour A_2 et A_3.

DEFINITION 14.3.2
Soient $A, A' \in M_{n,p}(\mathbb{K})$. On dit que A' se déduit de A par une suite d'opérations élémentaires sur les lignes si il existe une suite $(\tilde{A}_q)$ de matrices de $M_{n,p}(\mathbb{K})$, $0 \leqslant q \leqslant Q$ telles que $\tilde{A}_0 = A$, $\tilde{A}_Q = A'$ et $\tilde{A}_q$ se déduit de $\tilde{A}_{q-1}$ par une opération élémentaire sur les lignes pour $1 \leqslant q \leqslant Q$.

Soit donnée une suite d'opérations élémentaires sur les lignes. En les composant, on obtient une application $\varphi : M_{n,p}(\mathbb{K}) \to M_{n,p}(\mathbb{K})$. Le résultat suivant est facile, mais important :

Lemme 14.3.1
*Soit $\varphi : M_{n,p}(\mathbb{K}) \to M_{n,p}(\mathbb{K})$ une application obtenue en composant une suite d'opérations élémentaires **sur les lignes**. Pour toute $A \in M_{n,p}(\mathbb{K})$ et toute $B \in M_{p,q}(\mathbb{K})$ on a*

$$\varphi(AB) = \varphi(A) B$$

En effet, il existe une matrice carrée inversible Q, produit de $T_{i,j}(\lambda)$ de $P_{i,j}$ et de $D_i(r)$ telle que $\varphi(A) = QA$. Le lemme n'est autre que l'associativité du produit matriciel.

Remarque 1
Si $\psi : M_{n,p}(\mathbb{K}) \to M_{n,p}(\mathbb{K})$ est une suite d'opérations élémentaires sur les colonnes, on a

$$\psi(CA) = C\psi(A)$$

pour toute $A \in M_{n,p}(\mathbb{K})$ et toute $C \in M_{q,n}(\mathbb{K})$.

Remarque 2
D'après ce qui a été dit plus haut, si $A' = \varphi(A)$, on peut considérer A' comme représentant le même système de vecteurs que A mais dans une autre base. Il en résulte que les opérations élémentaires sur les lignes ne modifient pas non plus le rang du système des vecteurs colonnes de A.
Faire une opération élémentaire sur les lignes de A revient à faire une opération élémentaire sur les colonnes de la transposée de A, puis à retransposer à nouveau. Il s'en suit que les opérations sur les lignes ou sur les colonnes de A ne modifient pas non plus le rang du système des vecteurs lignes de A.
Ces deux remarques permettent de faire une démonstration de l'égalité $\mathrm{rg}(A) = \mathrm{rg}({}^tA)$ en n'utilisant que les opérations élémentaires.

APPLICATIONS

14.4 Méthode du Pivot de Gauss

Nous dirons qu'une matrice $M \in M_{n,p}(\mathbb{K})$ est pseudo-triangulaire si $m_{i,j} = 0$ pour tout couple (i,j) tel que $i > j$. Les matrices carrées pseudo-triangulaires sont les matrices triangulaires.

14.4.1 Description de l'algorithme

La méthode du pivot permet de transformer la matrice A en une matrice pseudo-triangulaire en effectuant uniquement des opérations élémentaires de type 1 et 2 sur les lignes. Nous décrirons ci dessous la première étape de l'algorithme.
Soit $A = (a_{i,j}) \in M_{n,p}(\mathbb{K})$.
• 1er cas : $a_{1,1} \neq 0$.
On remplace chacune des lignes $L_k,\ k \geqslant 2$ par $L_k - \frac{a_{k,1}}{a_{1,1}} L_1$. La matrice obtenue $A^{(1)}$ est telle que $a^{(1)}_{i,1} = 0$ pour $i > 1$. On dit qu'on a utilisé $a_{1,1}$ comme pivot.
• 2éme cas : $a_{1,1} = 0$ et il existe $i > 1$ tel que $a_{i,1} \neq 0$.
On permute les lignes L_i et L_1. Dans la matrice obtenue B, l'élément $b_{1,1} = a_{i,1}$ est non nul. On applique le premier cas à B. On obtient une matrice $A^{(1)}$ telle que $a^{(1)}_{i,1} = 0$ pour $i > 1$. On dit qu'on a utilisé $a_{i,1}$ comme pivot.
• 3éme cas : Pour tout $i \geqslant 1,\ a_{i,1} = 0$. On ne fait rien. La matrice $A^{(1)} = A$ est telle que $a_{i,1} = 0$ pour $i > 1$.

Dans tous les cas, après cette première étape, on a $A^{(1)} = \begin{pmatrix} a^{(1)}_{1,1} & L \\ \mathbf{0} & A_1 \end{pmatrix}$ avec $A_1 \in M_{n-1,p-1}(\mathbb{K})$ et $L \in M_{1,p-1}(\mathbb{K})$.

Ensuite, si $n \geqslant 2$ et $p \geqslant 2$, on ne touche plus la première ligne et on recommence avec la matrice A_1. Le processus s'arrête au bout de $q = \min(n,p) - 1$ étapes et la matrice obtenue est pseudo-triangulaire.

14.4.2 Interprétation matricielle

Dans le premier cas on a $A^{(1)} = U_1 A$ où $U_1 = T_{n,1}(\lambda_n) \cdots T_{2,1}(\lambda_2)$ (avec $\lambda_k = -\frac{a_{k,1}}{a_{1,1}}$).
Dans le deuxième cas, $A^{(1)} = (U_1 P_1)A$ où $P_1 = P_{i,1}$ et U_1 est du même type que ci dessus (avec d'autres valeurs pour les λ_k).
Enfin, dans le troisième cas, $A^{(1)} = A$.

Soit $j \in \mathbb{N}^*_{n-1}$. Soient $\lambda_{j+1}, \ldots, \lambda_n$ des scalaires. On pose $U_j(\lambda_{j+1}, \ldots, \lambda_n) = I_n + \sum_{k=j+1}^{k=n} \lambda_k E_{k,j}$.

$$U_j = \begin{pmatrix} & & & C_j \downarrow & & \\ 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & \lambda_{j+1} & 1 & & \\ & & \vdots & & \ddots & \\ & & \lambda_n & & & 1 \end{pmatrix}$$

Multiplier à gauche par U_j revient à faire les $(n-k)$ opérations $L_k \leftarrow L_k + \lambda_k L_j$ pour $k = j+1, \ldots, n$. Ces opérations commutent entre elles. On peut le contrôler en remarquant que

$$U_j = T_{n,j}(\lambda_{n,j}) T_{n-1,j}(\lambda_{n-1,j}) \cdots T_{j+1,j}(\lambda_{j+1})$$

ce produit étant commutatif.

On a alors

THEOREME 14.4.1 (Méthode du Pivot)
Soit $A \in M_{n,p}(\mathbb{K})$. Il existe une suite de matrices $U_1, \ldots, U_{n-1}$ de la forme ci dessus et une suite de matrices $P_1, \ldots, P_{n-1}$ de la forme $P_{i,j}$ telle que

$$A' = (U_{n-1}P_{n-1}) \cdots (U_1 P_1) A$$

soit pseudo-triangulaire. (Certaines des matrices U_j, P_k peuvent être la matrice I_n).

Remarque Les seules opérations élémentaires utilisées sont des permutations de lignes et des opérations définies par une matrice U_j, qui consiste à ajouter aux lignes $L_k,\ k > j$ la ligne L_j multipliée par un scalaire.

14.4.3 Applications

1. Systèmes linéaires

Une première application est la résolution de système linéaire. Si on veut résoudre le système $AX = B$, on applique la méthode du pivot sur la matrice $C = [A|B]$. On obtient à la fin une matrice $C' = [A'|B']$ où $A' = QA$ est pseudotriangulaire et $B' = QB$, la matrice Q étant inversible. Le système $AX = B$ est équivalent au système $A'X = B'$ plus facile à discuter et à résoudre.
Problèmes d'arrondis
Indiquons une difficulté dans les applications numériques. Soit une calculatrice de précision relative ε par exemple $\varepsilon = 10^{-12}$. Cela signifie que la machine ne distingue pas les nombres 1 et $1 + \varepsilon$. Considérons le système de deux équations à deux inconnues

$$\left\{\begin{array}{rcl} \varepsilon x + y & = & 1 \\ x + y & = & 2 \end{array}\right.$$

La solution exacte est $x = \frac{1}{1-\varepsilon},\ y = 2 - \frac{1}{1-\varepsilon} = \frac{1-2\varepsilon}{1-\varepsilon}$, trés voisine de $(1, 1)$. L'application de la méthode du pivot conduit au système

$$\left\{\begin{array}{rcl} \varepsilon x + y & = & 1 \\ \left(1 - \frac{1}{\varepsilon}\right) y & = & 2 - \frac{1}{\varepsilon} \end{array}\right.$$

Mais $1 - \frac{1}{\varepsilon} = 1 - 10^{12}$ est stocké en machine comme $-\frac{1}{\varepsilon} = -10^{12}$ et $2 - \frac{1}{\varepsilon}$ comme $-\frac{1}{\varepsilon} = -10^{12}$ de sorte que la deuxième équation donne $y = 1$ exactement et en remplaçant dans la première, $x = 0$! Il y a des modifications de l'algorithme qui permettent de contourner partiellement ces difficultés, mais dans la pratique l'algorithme du pivot n'est pas utilisé sur machine pour la résolution de systèmes linéaires.

2. Détermination du rang d'une matrice

Avec les notations du 1., on a $\mathrm{rg}(A) = \mathrm{rg}(A')$. Donnons un exemple :
Soient dans $\mathbb{R}^4$ les vecteurs $e_1 = (-1, 4, -3, -2),\ e_2 = (3, -7, 5, 3),\ e_3 = (3, -2, 1, 0),\ e_4 = (-4, 1, 0, 1)$ et $F = \mathrm{Vect}(e_1, e_2, e_3, e_4)$. On demande de déterminer une base de F et des équations cartésiennes de ce sous espace.

Soit $V = (x, y, z, t)$. On a $V \in F \Leftrightarrow \mathrm{rg}(e_1, e_2, e_3, e_4, V) = \mathrm{rg}(e_1, e_2, e_3, e_4)$. On va appliquer la méthode du pivot à la matrice complète du système (e_1, e_2, e_3, e_4, V)

$$A = \begin{pmatrix} -1 & 3 & 3 & -4 & x \\ 4 & -7 & -2 & 1 & y \\ -3 & 5 & 1 & 0 & z \\ -2 & 3 & 0 & 1 & t \end{pmatrix}$$

Il vient (les opérations effectuées sur A sont indiquées sur la droite)

$$A^{(1)} = \begin{pmatrix} -1 & 3 & 3 & -4 & x \\ 0 & 5 & 10 & -15 & y + 4x \\ 0 & -4 & -8 & 12 & z - 3x \\ 0 & -3 & -6 & 9 & t - 2x \end{pmatrix} \begin{array}{l} \\ L_2 + 4L_1 \\ L_3 - 3L_1 \\ L_4 - 2L_1 \end{array}$$

puis

$$A^{(2)} = \begin{pmatrix} -1 & 3 & 3 & -4 & x \\ 0 & 5 & 10 & -15 & y + 4x \\ 0 & 0 & 0 & 0 & (x + 4y + 5z)/5 \\ 0 & 0 & 0 & 0 & (2x + 3y + 5t)/5 \end{pmatrix} \begin{array}{l} \\ \\ L_3 + \frac{4}{5}L_2 \\ L_4 + \frac{3}{5}L_2 \end{array}$$

Les 4 premières colonnes sont formées des coordonnées de (e_1, e_2, e_3, e_4) dans une autre base $\mathcal{B}'$ que la base canonique. On voit donc immédiatement que $\mathrm{rg}(e_1, e_2, e_3, e_4) = 2$, donc que $\dim(F) = 2$ et que (e_1, e_2) est une base de F. Ensuite, la dernière colonne est formée des coordonnées de V dans la nouvelle base $\mathcal{B}'$. On en déduit

$$V \in F \Leftrightarrow V \in \mathrm{Vect}(e_1, e_2) \Leftrightarrow \left\{\begin{array}{rcl} x + 4y + 5z & = & 0 \\ 2x + 3y + 5t & = & 0 \end{array}\right.$$

En prime, on obtient la décomposition de V sur (e_1, e_2) si ces conditions sont remplies : $V = \frac{1}{5}(y+4x)e_2 + \left(-x + \frac{3}{5}(y+4x)\right) e_1$

3. Calcul de déterminant

Soit $A \in M_n(\mathbb{K})$. L'application de la méthode du pivot à la matrice A fournit une matrice A' qui est triangulaire. Si $A' = (a'_{i,j})$ on voit alors que A est inversible ssi les éléments diagonaux de A' sont tous non nuls. Supposons que ce soit le cas, c'est à dire que $A \in GL(n, \mathbb{K})$.

Les matrices $E_{i,j}$ ont un déterminant égal à 1 et il en est de même des matrices de type U_j. Les matrices $P_{i,j}$ avec $i \neq j$ ont un déterminant égal à -1. Donc si $A' = (U_{n-1}P_{n-1}) \cdots (U_1 P_1)A$, on a $\det(A) = \pm \det(A') = \pm \prod_{k=1}^{k=n} a'_{k,k}$ le signe étant $+$ si il y a un nombre pair de P_j différentes de l'identité et $-$ si il y en a un nombre impair. On a donc une méthode algorithmique simple de calcul du déterminant de A.

4. Inversion des matrices carrées

THEOREME 14.4.2
Soit $A \in GL(n, \mathbb{K})$. Il existe une suite d'opérations élémentaires sur les lignes qui transforme la matrice A en la matrice I_n.

Preuve abrégée
La première étape de l'algorithme est la même que dans la méthode du pivot. Supposons qu'il existe des matrices V_j produit de matrices de transvections et des matrices P_j telles que $A' = (V_{k-1}P_{k-1}) \cdots (V_1 P_1)A = \begin{pmatrix} D & C \\ \mathbf{0} & M \end{pmatrix}$ où $D \in M_{k-1}(\mathbb{K})$ est diagonale, $C \in M_{k-1,n-k+1}(\mathbb{K})$ et $M \in M_{n-k+1}(\mathbb{K})$. A étant inversible, il en est de même de M, donc la première colonne de cette matrice est non nulle. En permutant la ligne L_k avec une ligne $L_j,\ j > k$ ce qui revient à multiplier par $P_{j,k}$, on peut supposer $a'_{k,k} = M_{1,1} \neq 0$. On utilise cet élément comme pivot et on effectue les opérations $L_j \leftarrow L_j - \left(a'_{j,k}/a'_{k,k}\right) L_k$ pour $j \neq k$ ce qui permet de passer à l'étape suivante.■

Soit alors φ la composée d'une suite d'opérations élémentaires qui trasforme A en I_n : $\varphi(A) = I_n$. D'après le lemme 14.3.1 on a $\varphi(I_n) = \varphi(AA^{-1}) = \varphi(A)A^{-1} = A^{-1}$.
Donc si une suite d'opérations élémentaires transforme A en I_n, cette même suite d'opérations transforme I_n en A^{-1}. Ceci fournit un moyen rapide d'inverser des matrices carrées sans utiliser les déterminants.

Exemple : soit $A = \begin{pmatrix} 2 & 4 & 3 \\ 0 & 1 & 1 \\ 2 & 2 & -1 \end{pmatrix}$. Montrer que A est inversible et déterminer A^{-1}.

On dispose côte à côte A et I_3 et on essaye de transformer A en I_3. Si c'est possible, A est inversible. On regroupe en une seule deux opérations utilisant la même ligne, par exemple $L_1 \leftarrow L_1 - 4L_2,\ L_3 \leftarrow L_3 + 2L_2$ car des deux opérations commutent entre elles donc on peut les faire dans un ordre arbitraire.

$$\left(\begin{array}{ccc|ccc} 2 & 4 & 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 2 & 2 & -1 & 0 & 0 & 1 \end{array}\right) \qquad \left(\begin{array}{ccc|ccc} 2 & 4 & 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & -2 & -4 & -1 & 0 & 1 \end{array}\right) \quad L_3 \leftarrow L_3 - L_1$$

$$\left(\begin{array}{ccc|ccc} 2 & 0 & -1 & 1 & -4 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & -2 & -1 & 2 & 1 \end{array}\right) \quad \begin{array}{l} L_1 \leftarrow L_1 - 4L_2 \\ \\ L_3 \leftarrow L_3 + 2L_2 \end{array} \qquad \left(\begin{array}{ccc|ccc} 2 & 0 & 0 & 3/2 & -5 & -1/2 \\ 0 & 1 & 0 & -1/2 & 2 & 1/2 \\ 0 & 0 & -2 & -1 & 2 & 1 \end{array}\right) \quad \begin{array}{l} L_1 \leftarrow L_1 - L_3/2 \\ L_2 \leftarrow L_2 + L_3/2 \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 3/4 & -5/2 & -1/4 \\ 0 & 1 & 0 & -1/2 & 2 & 1/2 \\ 0 & 0 & 1 & 1/2 & -1 & -1/2 \end{array}\right) \quad \begin{array}{l} L_1 \leftarrow L_1/2 \\ \\ L_3 \leftarrow -L_3/2 \end{array}$$

$$A^{-1} = \frac{1}{4}\begin{pmatrix} 3 & -10 & -1 \\ -2 & 8 & 2 \\ 2 & -4 & -2 \end{pmatrix}$$

14.5 Applications à $GL(n,\mathbb{K}),\ SL(n,\mathbb{K})$

Rappelons que $SL(n,\mathbb{K}) = \{A \in M_n(\mathbb{K}) \mid \det(A) = 1\}$. L'application $\det : GL(n,\mathbb{K}) \to \mathbb{K}^*$ est un morphisme de groupes et $SL(n,\mathbb{K}) = \ker(\det)$. C'est donc un sous groupe distingué de $GL(n,\mathbb{K})$ et $GL(n,\mathbb{K})/SL(n,\mathbb{K})$ est isomorphe à $\mathrm{im}(\det) = \mathbb{K}^*$.

THEOREME 14.5.1
Soit $A \in GL(n,\mathbb{K})$. Il existe des matrices de transvections $T_1,\ldots,T_m$ et une matrice de dilatation $D_n(d)$ telles que

$$A = T_m \cdots T_1 D_n(d)$$

De plus $d = \det(A)$.

COROLLAIRE 14.5.1
$GL(n,\mathbb{K})$ est engendré par les matrices de transvections et les matrices de dilatation.

COROLLAIRE 14.5.2
$SL(n,\mathbb{K})$ est engendré par les matrices de transvections.

preuve
Montrons le théorème par récurrence sur n. Il n'y a rien à montrer pour $n = 1$. Soit $n \geqslant 1$ et $A \in GL(n+1,\mathbb{K})$. La première colonne de A est non nulle.
• Il existe une matrice de transvection T_1 telle que $(T_1A)_{1,1} \neq 0$. En effet, si $a_{1,1} \neq 0$ on prend $T_1 = I_n$. Sinon, il existe un $i > 1$ tel que $a_{i,1} \neq 0$. L'opération $L_1 \leftarrow L_1 + L_i$ transforme A en une matrice A' telle que $a'_{1,1} \neq 0$.
• Si $a'_{i,1} = 0$ pour tout $i > 1$, on ajoute à la deuxième ligne de A' la première, i.e. on multiplie par $T_{2,1}(1)$.
• On a donc maintenant une matrice A'' vérifiant $a''_{1,1} \neq 0$ et $a''_{i,1} \neq 0$ pour un $i > 1$. On effectue alors l'opération $L_1 \leftarrow L_1 + \dfrac{1-a_{1,1}}{a_{i,1}} L_i$. La matrice obtenue, soit A''' est telle que $a'''_{1,1} = 1$.
• On effectue sur A''' les opérations $L_k \leftarrow L_k - a'''_{k,1} L_1$.
A ce stade, on a trouvé une matrice H, produit fini de matrices de transvections telle que HA soit de la forme $HA = \begin{pmatrix} 1 & \Lambda_1 \\ \mathbf{0} & B \end{pmatrix}$ où $\Lambda_1 \in M_{1,n}(\mathbb{K})$
• A inversible $\Rightarrow$ HA inversible, donc la matrice B est inversible. Il en résulte que les lignes de B forment une base de $\mathbb{K}^n = M_{1,n}(\mathbb{K})$. Il existe donc des scalaires $\alpha_1,\ldots,\alpha_n$ tels que $\Lambda_1 = \sum_{k=1}^{k=n} \alpha_k L_k(B)$. On effectue sur HA l' opération $L_1 \leftarrow L_1 - \sum_{k=1}^{k=n} \alpha_k L_k$ qui est en fait une suite d'opérations élémentaires sur les lignes. On obtient ainsi un produit H' de matrices de transvections telle que

$$H'HA = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & A_1 \end{pmatrix}$$

• Fin de la preuve
A_1 est inversible. D'après l'hypothèse de récurrence appliquée à A_1, il existe des matrices de transvections $\tilde{T}_k$, $1 \leqslant k \leqslant K$ appartenant à $M_n(\mathbb{K})$ et une matrice $D_n^{(n)}(d)$ telles que $A_1 = \tilde{T}_M \cdots \tilde{T}_1 D_n^{(n)}(d)$ Les matrices $T_k = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \tilde{T}_k \end{pmatrix}$ sont des matrices de transvection d'ordre $n+1$

Un produit par blocs montre alors que

$$A = H'^{-1} H^{-1} T_M \cdots T_1 D_{n+1}(d)$$

H et H' sont des produits de matrice de transvection et l'inverse d'une matrice de transvection est une matrice de transvection, donc $H'^{-1}H^{-1}T_M \cdots T_1$ est un produit de matrices de transvection. Ceci achève la preuve par récurrence.

Enfin, comme le déterminant d'une matrice de transvection est 1, on a clairement $\det(A) = \det(D_n(d)) = d$. Les deux corollaires sont des conséquences immédiates du théorème.

15

Réduction des endomorphismes

Dans tout ce qui suit, $\mathbb{K}$ est un sous corps de $\mathbb{C}$, E un $\mathbb{K}$-espace vectoriel et I est l'identité de E.

15.1 Eléments propres d'un endomorphisme

15.1.1 Sous espaces stables

Soit $u \in L(E)$; un sous espace vectoriel F de E est dit stable par u si $u(F) \subset F$. Dans ce cas, on notera u_F l'endomorphisme de F induit par u. Si $E = F \oplus G$ est de dimension finie, avec F stable par u, la matrice de u dans une base adaptée à la décomposition en somme directe est de la forme $M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ et $A = \mathrm{Mat}(u_F)$; dans ces conditions, G est stable par u ssi $B = 0$.

Lemme 15.1.1
Soient u et v deux endomorphismes de E qui commutent. Le noyau et l'image de u sont stables par v.

15.1.2 Eléments propres

DEFINITION 15.1.1
Soit $u \in L(E)$.
*$\lambda \in \mathbb{K}$ est une valeur propre de u si il existe un $x \in E$ **non nul** tel que $u(x) = \lambda x$. Ceci revient à dire que $u - \lambda I$ est non injectif.*
*$x \in E$ est un vecteur propre de u si x est **non nul** et si il existe un scalaire λ tel que $u(x) = \lambda x$.*
Dans les deux cas ci dessus, λ et x sont appelés éléments propres associés. Si λ est une valeur propre de u l'espace propre associé à λ est $E_\lambda = E_\lambda(u) = \ker(u - \lambda I)$. Il est formé des vecteurs propres de u associés à la valeur propre λ et du vecteur nul.
Définition analogue pour les matrices : on considère $A \in M_n(\mathbb{K})$ comme l'endomorphisme de $\mathbb{K}^n$ défini par $X \to AX$ où les éléments de $\mathbb{K}^n$ sont écrits en colonne.

Propriétés élémentaires

1) Supposons E de dimension finie $\geqslant 1$ et soit $\mathcal{B}$ une base de E. Soit $A = \mathrm{Mat}_{\mathcal{B}}(u)$. λ est valeur propre de u ssi λ est valeur propre de A. Si $x \in E$ et $X = \mathrm{Mat}_{\mathcal{B}}(x)$, x est vecteur propre de u ssi X est vecteur propre de A.
2) 0 est valeur propre de $u \Leftrightarrow \ker(u) \neq \{0\}$.
3) Si $\lambda \neq 0$ est valeur propre de u, on a $E_\lambda(u) \subset \mathrm{im}(u)$.

THEOREME 15.1.1
Soit u un endomorphisme d'un $\mathbb{K}$-ev E, et $\lambda_1, \ldots, \lambda_p$ des valeurs propres distinctes de u. La somme $E_{\lambda_1} + \cdots + E_{\lambda_p}$ est directe.

preuve
On montre par récurrence sur k la propriété suivante: $\forall x_1 \in E_{\lambda_1}, \ldots, \forall x_k \in E_{\lambda_k},\ x_1 + \cdots + x_k = 0 \Rightarrow x_1 = 0, \ldots, x_k = 0$. Elle est triviale si $k = 1$. Supposons la vraie pour un entier $k, 1 \leqslant k < p$. Soient $x_i \in E_{\lambda_i},\ 1 \leqslant i \leqslant k+1$ tels que $x_1 + \cdots + x_{k+1} = 0$. Appliquant u on en déduit $\lambda_1 x_1 + \cdots + \lambda_{k+1} x_{k+1} = 0$, d'où, par combinaison $\sum_{i=1}^{i=k} (\lambda_{k+1} - \lambda_i) x_i = 0$. D'après l'hypothèse de récurrence, on en déduit $(\lambda_{k+1} - \lambda_i) x_i = 0$ pour $1 \leqslant i \leqslant k$ et les λ_i étant distincts, $x_i = 0$ pour $1 \leqslant i \leqslant k$ d'où ensuite $x_{k+1} = 0$. ■

Exemple
Soit $E = C^\infty(I, \mathbb{C})$ où I est un intervalle non vide et non réduit à un point de $\mathbb{R}$ et D l'application qui à une fonction $f \in E$ associe sa dérivée. Tout $\alpha \in \mathbb{C}$ est valeur propre de l'endomorphisme D, l'espace propre associé étant la droite engendrée par la fonction $e_\alpha : x \to e^{\alpha x}$. Il en résulte que si $\alpha_1, \ldots, \alpha_p$ sont p complexes distincts, les fonctions $(e_{\alpha_1}, \ldots, e_{\alpha_p})$ forment un système libre.

DEFINITION 15.1.2
Soient E un $\mathbb{K}$-ev de dimension finie et $u \in L(E)$. On appelle spectre de u et on notera $\mathrm{Sp}(u)$ l'ensemble des valeurs propres de u.

15.2 Diagonalisation, trigonalisation

DANS TOUTE CETTE PARTIE, E EST SUPPOSE DE DIMENSION FINIE, $n \geqslant 1$.

15.2.1 Polynôme caractéristique

Définition et propriétés élémentaires

Le scalaire λ est valeur propre de l'endomorphisme u ssi $\ker(u - \lambda_I) \neq \{0\}$ ce qui est équivalent à $\det(u - \lambda I) = 0$.

THEOREME 15.2.1
Soit $u \in L(E)$. L'application $\mathbb{K} \to \mathbb{K}$ qui à t associe $\det(u - tI)$ est une fonction polynôme.

Puisque le corps $\mathbb{K}$ est infini, à cette fonction polynôme est associée un unique polynôme $\chi_u(X) \in \mathbb{K}[X]$

DEFINITION 15.2.1
On appelle polynôme caractéristique de l'endomorphisme u l'unique polynôme $\chi_u(X) \in \mathbb{K}[X]$ tel que, pour tout $t \in \mathbb{K}$ on ait $\chi_u(t) = \det(u - tI)$.

On a un énoncé analogue pour les matrices.

Soit A la matrice de u dans une base de E. Le polynôme caractéristique de u est égal à celui de A. Deux matrices semblables ont le même polynôme caractéristique. Une matrice et sa transposée ont le même polynôme caractéristique, donc aussi un endomorphisme et son transposé.

THEOREME 15.2.2

$$\chi_u(X) = (-1)^n X^n + (-1)^{n-1} \mathrm{tr}(u) X^{n-1} + \cdots + \det(u)$$

preuve
Le terme constant du polynôme caractéristique est $\chi_u(0) = \det(u)$. Soit $A = (a_{i,j})$ la matrice de u dans une base de E. Celle de $u - tI$ est $A' = (a'_{i,j})$ où $a'_{i,j} = a_{i,j} - t\delta_{i,j}$. En désignant par ε_σ la signature de la permutation σ , on a $\chi_u(t) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon_\sigma a'_{1,\sigma(1)} \cdots a'_{n,\sigma(n)} = \prod_{k=1}^{k=n} a'_{k,k} + \sum_{\substack{\sigma \in \mathfrak{S}_n \\ \sigma \neq Id}} \varepsilon_\sigma a'_{1,\sigma(1)} \cdots a'_{n,\sigma(n)}$. Si σ est une permutation différente de Id il existe au moins deux indices i tels que $\sigma(i) \neq i$. Il en résulte que la dernière somme figurant dans l'expression ci dessus est un

polynôme en t de degré au plus $n-2$. Les termes en t^net en t^{n-1} proviennent donc uniquement du produit $\prod_{k=1}^{k=n} a'_{k,k}$. On en déduit le résultat.

COROLLAIRE 15.2.1
Soient $u \in L(E)$ dont le polynôme caractéristique est scindé sur $\mathbb{K}$ et $(\lambda_1, \ldots, \lambda_n)$ une liste des racines de χ_u. On a

$$\mathrm{tr}(u) = \lambda_1 + \lambda_2 + \cdots + \lambda_n \quad \textit{et} \quad \det(u) = \lambda_1 \lambda_2 \cdots \lambda_n$$

THEOREME 15.2.3
Les valeurs propres de u sont les racines du polynôme caractéristique χ_u de u.

preuve
Si $\lambda \in \mathbb{K}$ on a λ valeur propre de $u \Leftrightarrow u - \lambda I$ non injectif $\Leftrightarrow \det(u - \lambda I) = 0 \Leftrightarrow \chi_u(\lambda) = 0$.

DEFINITION 15.2.2
Soit λ une valeur propre de u. On appelle ordre de multiplicité de la valeur propre λ l'ordre de multiplicité de λ comme racine de χ_u. On notera m_λ cet ordre de multiplicité de sorte que $\chi_u(X) = (X-\lambda)^{m_\lambda} Q(X)$ avec $Q(\lambda) \neq 0$.

Polynôme caractéristique et sous espaces stables

PROPOSITION 15.2.1
Soient $u \in L(E)$ et F un sous espace de E stable par u. Alors
1) χ_{u_F} divise χ_u.
2) Si G est un supplémentaire de F et si G est stable par f, on a $\chi_u(X) = \chi_{u_F}(X)\chi_{u_G}(X)$.

preuve
Soient $p = \dim(F)$, $\mathcal{B}_F$ une base de F complétée en une base $\mathcal{B}$ de E. Dans cette base la matrice de u est de la forme $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ où $A \in M_p(\mathbb{K})$ est la matice de u_F dans $\mathcal{B}_F$ et $C \in M_{p,n-p}(\mathbb{K})$, $B \in M_{n-p}(\mathbb{K})$. On a de suite $\chi_u(X) = \chi_A(X)\chi_B(X) = \chi_{u_F}(X)\chi_B(X)$ d'où le 1).
Ensuite, si G est stable, on peut compléter la base $\mathcal{B}_F$ de F par une base $\mathcal{B}_G$ de G pour obtenir une base $\mathcal{B}$ de E. La matrice de u dans cette base a la même forme que ci dessus avec cette fois $C = 0$ et $B = \mathrm{Mat}_{\mathcal{B}_G}(u_G)$ d'où la conclusion.

Attention! Un sous espace F stable par u ne possède pas nécessairement un supplémentaire stable.

COROLLAIRE 15.2.2
Soit λ une valeur propre de u de multiplicité m_λ. On a $\dim(E_\lambda) \leqslant m_\lambda$.

preuve
Le polynôme caractéristique de l'endomorphisme induit par u sur E_λ est $(\lambda - X)^{\dim(E_\lambda)}$ d'où la conclusion.

EXEMPLE 15.2.1 (matrice compagnon)
Soient $(a_0, \ldots, a_{n-1}) \in \mathbb{K}^n$, $P = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0$ et $M = M_P = (m_{i,j}) \in M_n(\mathbb{K})$ telle que $m_{i+1,i} = 1$ pour $1 \leqslant i \leqslant n-1$, $m_{i,n} = -a_{i-1}$ et $m_{i,j} = 0$ dans les autres cas.

$$M = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix}$$

Pour calculer le déterminant $\det(M - XI_n)$ on remplace la première ligne L_1 par $L_1 + XL_2 + \cdots + X^{n-1}L_n$. On obtient $\chi_M(X) = (-1)^n P(X)$ La matrice M_P s'appelle la matrice compagnon du polynôme P.

Exercice : Montrer que pour toute valeur propre λ de M, l'espace propre correspondant est de dimension 1.

15.2.2 Diagonalistation

DEFINITION 15.2.3
L'endomorphisme u de E est dit diagonalisable si il existe une base $\mathcal{B}$ de E dans laquelle la matrice de u est diagonale.
Une matrice carrée A est dite diagonalisable si elle est semblable à une matrice diagonale.

Donc $A \in M_n(\mathbb{K})$ est diagonalisable ssi il existe $P \in GL(n, \mathbb{K})$ telle que $P^{-1}AP$ soit diagonale.
L'endomorphisme u est diagonalisable ssi sa matrice dans une base quelconque l'est.

THEOREME 15.2.4
Soient E un $\mathbb{K}$-ev de dimension finie et $u \in L(E)$. Les propriétés suivantes sont équivalentes :
(1) u est diagonalisable.
(2) Il existe une base de E formée de vecteurs propres de u.
(3) La somme des espaces propres de u est égale à E.
(4) La somme des dimensions des espaces propres de u est égale à $\dim(E)$.

La preuve est laissée au lecteur. L'équivalence entre (3) et (4) résulte de ce que la somme des espaces propres est toujours directe.

THEOREME 15.2.5
Si u admet $dim(E)$ valeurs propres distinctes, u est diagonalisable.

THEOREME 15.2.6
Les propriétés suivantes sont équivalentes :
(1) u est diagonalisable.
(2) Le polynôme caractéristique de u est scindé sur $\mathbb{K}$ et l'ordre de multiplicité de chaque racine λ de χ_u est égal à la dimension de l'espace propre correspondant.

preuve
Soient $\lambda_1, \ldots, \lambda_p$ les valeurs propres distinctes de u, $E_1, \ldots, E_p$ les espaces propres associés et $m_1, \ldots, m_p$ leur multiplicité respective.
$(1) \Rightarrow (2)$ Par hypothèse $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p}$. Dans une base adaptée à cette décomposition en somme directe, la matrice de u est $\mathrm{diag}(\underbrace{\lambda_1, \cdots, \lambda_1}_{\dim(E_1) \text{ fois}}, \cdots, \underbrace{\lambda_p, \cdots, \lambda_p}_{\dim(E_p) \text{ fois}})$ donc $\chi_u(X) = (\lambda_1 - X)^{\dim(E_1)} \cdots (\lambda_p - X)^{\dim(E_p)}$.

$(2) \Rightarrow (1)$ χ_u est scindé, donc $m_1 + \cdots + m_p = \deg(\chi_u) = \dim(E)$ soit $\dim(E_1) + \cdots + \dim(E_p) = \dim(E)$ d'où la conclusion, d'après le théorème 15.2.4.

Exemple : Soit $P \in \mathbb{K}_n[X]$ un polynôme de degré n normalisé et M_P sa matrice compagnon. Comme les espaces propres sont tous de dimension 1, M_P est diagonalisable ssi P est scindé sur $\mathbb{K}$ à racines simples.

15.2.3 Trigonalisation

DEFINITION 15.2.4
L'endomorphisme u de E est dit trigonalisable si il existe une base $\mathcal{B}$ de E dans laquelle la matrice de u est triangulaire supérieure.
Une matrice carrée A est dite trigonalisable si elle est semblable à une matrice triangulaire supérieure.

Remarque
Si la matrice de u dans la base $\mathcal{B} = (e_1, \ldots, e_n)$ est triangulaire supérieure, la matrice de u dans la base $\mathcal{B}' = (e_n, e_{n-1}, \ldots, e_1)$ est triangulaire inférieure. On pourrait donc remplacer triangulaire supérieure par triangulaire inférieure dans la définition.

u est trigonalisable ssi sa matrice dans une base quelconque l'est.

DEFINITION 15.2.5
Soit E un $\mathbb{K}$-ev de dimension n. On appelle drapeau de E toute suite finie de sous espaces emboîtés $F_0 = \{0\} \subset F_1 \subset \cdots \subset F_{n-1} \subset F_n = E$ telle que $\dim(F_k) = k$ pour tout k entre 0 et n.
Un tel drapeau est dit adapté à u si pour tout k on a $u(F_k) \subset F_k$.

THEOREME 15.2.7
Soit u un endomorphisme d'un $\mathbb{K}$-ev E de dimenion finie. Les propriétés suivantes sont équivalentes :
(1) u est trigonalisable.
(2) Il existe un drapeau de E adapté à u.
(3) Le polynôme caractéristique de u est scindé sur $\mathbb{K}$.

preuve
L'équivalence entre (1) et (2) est facile ainsi que l'implication $(1) \Rightarrow (3)$. Montrons que $(3) \Rightarrow (1)$. On procède par récurrence. Soit l'hypothèse de récurrence
(HR$_p$) "Toute matrice carrée d'ordre p à coefficients dans $\mathbb{K}$ dont le polynôme caractéristique est scindé est trigonalisable"

(HR$_1$) est triviale;
Montrons (HR$_n$) $\Rightarrow$ (HR$_{n+1}$). Soit $A \in M_{n+1}(\mathbb{K})$ dont le polynôme caractéristique χ_A est scindé. Soit λ une racine de χ_A. C'est une valeur propre de A. Soit $e \in K^{n+1}$ un vecteur propre associé à λ : $Ae = \lambda e$. Soient enfin $e_1, \ldots, e_n \in \mathbb{K}^{n+1}$ tels que $\mathcal{B} = (e, e_1, \ldots, e_n)$ soit une base de $\mathbb{K}^{n+1}$ et P la matrice de passage de la base canonique à cette base. On a $P^{-1}AP = \begin{pmatrix} \lambda & L \\ 0 & B \end{pmatrix}$ où L est une matrice ligne et B une matrice carrée d'ordre n.
On a $\chi_A(X) = (\lambda - X)\chi_B(X)$. On en déduit que χ_B est scindé; d'après (HR$_n$) B est trigonalisable et il existe $Q' \in GL(n, \mathbb{K})$ et $T \in M_n(\mathbb{K})$ triangulaire supérieure telles que $Q'^{-1}BQ' = T$. Posons $Q = \begin{pmatrix} 1 & 0 \\ 0 & Q' \end{pmatrix}$. On a $Q \in GL(n+1, \mathbb{K})$ et si $R = PQ$ on a $R \in GL(n+1, \mathbb{K})$ et

$$R^{-1}AR = \begin{pmatrix} 1 & 0 \\ 0 & Q'^{-1} \end{pmatrix} \begin{pmatrix} \lambda & L \\ 0 & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q' \end{pmatrix} = \begin{pmatrix} \lambda & LQ' \\ 0 & Q'^{-1}BQ' \end{pmatrix} = \begin{pmatrix} \lambda & L' \\ 0 & T \end{pmatrix}$$

ce qui achève la preuve.

COROLLAIRE 15.2.3
1) Tout endomorphisme d'un $\mathbb{C}$-ev de dimension finie est trigonalisable.
2) Toute matrice carrée à coefficients dans $\mathbb{C}$ est trigonalisable dans $M_n(\mathbb{C})$, autrement dit, pour toute $A \in M_n(\mathbb{C})$, il existe $P \in GL(n, \mathbb{C})$ tell que $P^{-1}AP$ soit triangulaire supérieure.

15.3 Réduction simultanée

THEOREME 15.3.1
Soit u un endomorphisme diagonalisable d'un $\mathbb{K}$-ev de dimension finie E et F un sev de E stable par u. L'endomorphisme u_F de F induit par u est diagonalisable.

preuve
Soit $(e_1, \ldots, e_n)$ une base de E formée de vecteurs propres de u de sorte que $u(e_k) = \lambda_k e_k$. Soit d'autre part $(f_1, \ldots, f_r)$ une base de F. D'après le théorème de la base incomplète, il existe $n - r$ vecteurs pris parmi $e_1, \ldots, e_n$ qui complète $(f_1, \ldots, f_r)$ en une base de E. Quitte à renuméroter, on peut toujours supposer que ce sont les vecteurs $e_{r+1}, \ldots e_n$. Soit $G = \mathrm{Vect}(e_{r+1}, \ldots, e_n)$. C'est un supplémentaire de F, stable par f puisque engendré par des vecteurs propres. Soit enfin π la projection sur F parallélement à G. Le sous espace $S = \mathrm{Vect}(e_1, \ldots, e_r)$ est un supplémentaire de $G = \ker(\pi)$ dont la restriction de π à S induit un isomorphisme de S sur son image. En particulier, si on pose $w_j = \pi(e_j)$ pour $1 \leqslant j \leqslant r$ le système $(w_1, \ldots, w_r)$ est libre, donc est une base de F. Montrons que ces vecteurs sont des vecteurs propres de u_F, ce qui achèvera la preuve.
Soit j entre 1 et r. La décomposition de e_j sur $F \oplus G$ s'écrit $e_j = \pi(e_j) + (e_j - \pi(e_j)) = w_j + (e_j - w_j)$. Comme chacun des espaces F et G est stable par u on a $u(w_j) \in F$ et $u(e_j - w_j) \in G$ donc la décomposition de $u(e_j)$ sur $F \oplus G$ est $u(w_j) + \big(u(e_j - w_j)\big)$. On en déduit $u(w_j) = \pi(u(e_j))$. Or $u(e_j) = \lambda_j e_j$ donc $u(w_j) = \lambda_j \pi(e_j) = \lambda_j w_j$.■
Remarque : Ce théorème admet une preuve trés courte basée sur la notion de polynôme annulateur (voir le corollaire 15.4.2 ci dessous).

THEOREME 15.3.2
1) Soient E un $\mathbb{K}$-ev de dimension finie et $(u_i)_{i \in I}$ une famille d'endomorphismes de E, diagonalisables et commutant deux à deux. Il existe une base de E formée de vecteurs propres communs à tous les u_i.
2) Soit $(M_i)_{i \in I}$ une famille de matrices carrées d'ordre n, diagonalisables et commutant deux à deux. Il existe une matrice $P \in GL(n, \mathbb{K})$ telle que, pour tout i, la matrice $P^{-1}M_iP$ soit diagonale.

preuve
Il est clair que les propriétés (1) et (2) sont équivalentes. La preuve de (1) se fait par récurrence sur la dimension p de E. Considérons l'hypothèse de récurrence :
(HR$_p$) : Pour toute famille $(u_i)_{i\in I}$ d'endomorphismes diagonalisables et deux à deux commutant d'un espace vectoriel E de dimension p, il existe une base de E formée de vecteurs propres communs à tous les u_i.
(HR$_1$) est trivial.
Montrons (HR$_1$) et $\cdots$ et (HR$_p$) $\Rightarrow$ (HR$_{p+1}$). Soit $(u_i)_{i\in I}$ une famille d'endomorphismes diagonalisables et deux à deux commutants d'un espace vectoriel E de dimension $p+1$. Il n'y a rien à montrer si tous les u_i sont proportionnels à l'identité. Sinon, fixons un indice i tel que u_i ne soit pas proportionnel à I. On peut écrire $E = \overset{k=r}{\underset{k=1}{\oplus}} E_k$ où les E_k sont les sous espaces propres distincts de u_i. Pour tout $j \in I$, u_i commute avec u_j donc les E_k sont stables par u_j. Notons $v_{j,k}$ l'endomorphisme induit par u_j sur E_k. $(v_{j,k})_{j\in I}$ est une famille d'endomorphismes de E_k deux à deux commutant et qui sont diagonalisables d'après le théorème 15.3.1. Comme u_i n'est pas proportionnel à l'identité, on a $E_k \neq E$, donc $\dim(E_k) \leqslant p$. D'après l'hypothèse de récurrence, il existe une base $\mathcal{B}_k$ de E_k formée de vecteurs propres communs à tous les $v_{j,k}$; $j \in I$. Le système $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_r)$ est une base de E formée de vecteurs propres communs à tous les u_i. Ceci achève la preuve.

Lemme 15.3.1
Soit E un $\mathbb{K}$-ev de dimension finie et $(u_i)_{i\in I}$ une famille d'endomorphismes de E, deux à deux commutants et dont les polynômes caractéristiques sont scindés sur $\mathbb{K}$. Les u_i ont au moins un vecteur propre en commun.

preuve
Elle se fait encore par récurrence sur $p = \dim E$. Le lemme est trivial si $p = 1$. Soit $p \in \mathbb{N}^*$. Supposons le lemme vrai pour tout espace vectoriel de dimension $\leqslant p$. Soit E de dimension $p+1$ et $(u_i)_{i\in I}$ une famille d'endomorphismes de E deux à deux commutants, dont les polynômes caractéristqiues sont scindés. Il n'y a rien à montrer si tous les u_i sont proportionnels à I. Sinon, soit i un indice tel que u_i ne soit pas proportionnel à I. Soit λ une valeur propre de u_i et F le sous espace propre associé. Pour tout j, u_j commute avec u_i, donc F est stable par u_j. Notons v_j l'endomorphisme de F induit par u_j.
- Les v_j commutent deux à deux.
- Le polynôme caractéristique de v_j divise celui de u_j, donc est scindé sur $\mathbb{K}$.
- $\dim(F) \leqslant p$ car u_i n'est pas proportionnel à I, donc $F \neq E$.
D'après l'hypothèse de récurrence, il existe un vecteur $e \in F$ qui est un vecteur propre commun à tous les v_j, donc à tous les u_j ce qui achève la preuve.

THEOREME 15.3.3
1) Soient E un $\mathbb{K}$-ev de dimension finie et $(u_i)_{i\in I}$ une famille d'endomorphismes de E, trigonalisables et commutant deux à deux. Il existe une base de E dans laquelle tous les u_i ont une matrice triangulaire supérieure.
2) Soit $(M_i)_{i\in I}$ une famille de matrices carrées d'ordre n, trigonalisables et commutant deux à deux. Il existe une matrice $P \in GL(n,\mathbb{K})$ telle que, pour tout i, la matrice $P^{-1}M_iP$ soit triangulaire supérieure.

preuve
(1) et (2) sont équivalents. On démontre (2) par récurrence sur n. C'est trivial si $n = 1$. Soit $p \in \mathbb{N}^*$ et supposons (2) vrai pour $n \leqslant p$. Soit $(M_i)_{i\in I}$ une famille de matrices appartenant à $M_{p+1}(\mathbb{K})$, deux à deux commutantes et trigonalisables. Ces matrices ont des polynômes caractéristiques scindés. D'après le lemme, on dispose d'un vecteur propre commun à toutes les M_i, soit $C_1 \in \mathbb{K}^{p+1}$. Le théorème de la base incomplète fournit une base de $\mathbb{K}^{p+1}$ de la forme $\mathcal{B} = (C_1, \ldots, C_{p+1})$. Soit P la matrice de passage de la base canonique à $\mathcal{B}$. On a, pour tout i, $P^{-1}M_iP = \begin{pmatrix} \lambda_i & L_i \\ 0 & B_i \end{pmatrix}$ où $L_i \in M_{1,p}(\mathbb{K})$ et $B_i \in M_p(\mathbb{K})$. Un petit calcul montre que les matrices B_i commutent entre elles. D'autre part, le polynôme caractéristique de B_i divise celui de $P^{-1}M_iP$ qui est scindé car égal à celui de M_i. Donc B_i est trigonalisable. D'après l'hypothèse de récurrence, il existe $Q \in GL(p,\mathbb{K})$ telle que, pour tout i, $Q^{-1}B_iQ$ soit triangulaire supérieure. La matrice $S = P\begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}$ est inversible et $S^{-1}M_iS = \begin{pmatrix} \lambda_i & X_i \\ 0 & Q^{-1}B_iQ \end{pmatrix}$ est triangulaire supérieure pour tout i.■

15.4 Polynômes annulateurs d'endomorphismes

15.4.1 Polynômes d'endomorphismes

PROPOSITION 15.4.1
Soit E un $\mathbb{K}$-ev et $u \in L(E)$.
L'application $\theta_u : \mathbb{K}[X] \to L(E)$ qui à $P = \sum_{k=0}^{k=p} a_k X^k$ associe $\theta_u(P) = a_0 id_E + a_1 u + \cdots a_p u^p$ où $u^p = \underbrace{u \circ u \cdots \circ u}_{p \text{ fois}}$ est un morphisme de $\mathbb{K}$-algèbres, i.e. vérifie :
(1) $\theta_u(1) = id_E$
(2) $\forall \lambda, \mu \in \mathbb{K}\ \forall P, Q \in \mathbb{K}[X]\ \theta_u(\lambda P + \mu Q) = \lambda\theta_u(P) + \mu\theta_u(Q)$
(3) $\forall P, Q \in \mathbb{K}[X]\ \theta_u(PQ) = \theta_u(P) \circ \theta_u(Q)$.
L'image de θ_u est une sous algèbre commutative de $L(E)$ que l'on note $\mathbb{K}[u]$.
Si $u, v \in L(E)$ commutent, tout élément de $\mathbb{K}[u]$ commute avec tout élément de $\mathbb{K}[v]$.

On note usuellement $P(u) = a_0 id_E + a_1 u + \cdots a_p u^p$ de sorte que les propriétés de morphismes s'écrivent $(\lambda P + \mu Q)(u) = \lambda P(u) + \mu Q(u)$ et $(PQ)(u) = P(u)Q(u)$.

Dans cet énoncé, l'espace E peut ne pas être de dimension finie. On a un résultat analogue pour les matrices carrées.
Les vérifications sont faciles.

Lemme 15.4.1
Soit $x \in E$ et $\lambda \in \mathbb{K}$. Si $u(x) = \lambda x$ on a pour tout polynôme P, $P(u)(x) = P(\lambda)x$.

DEFINITION 15.4.1
Soit $u \in L(E)$ et $P \in \mathbb{K}[X]$. On dit que le polynôme P est annulateur de l'endomorphisme u ou qu'il annule u si $P(u) = 0$.

On déduit du lemme la proposition suivante :

PROPOSITION 15.4.2
Si $P \in \mathbb{K}[X]$ est annulateur de $u \in L(E)$, pour toute valeur propre λ de u on a $P(\lambda) = 0$.

Lemme 15.4.2
Soit E un $\mathbb{K}$-ev de dimension finie et $u \in L(E)$. Il existe des polynômes non nuls annulateurs de u.

preuve
La famille $(id_E, u, u^2, \ldots, u^{n^2})$ où $n = \dim(E)$ comprend $\dim(L(E)) + 1$ éléments, donc est liée. Il existe donc des scalaires non tous nuls $a_i,\ 0 \leqslant i \leqslant n^2$ tels que $\sum_{i=0}^{i=n^2} a_i u^i = 0$. Le polynôme $P(X) = \sum_{i=0}^{i=n^2} a_i X^i$ est non nul et annulateur de u.

THEOREME 15.4.1
Soient E un $\mathbb{K}$-ev de dimension finie et $u \in L(E)$. L'ensemble $\mathrm{Ann}(u)$ des polynômes annulateurs de u est un idéal de $\mathbb{K}[X]$.

On sait que tout idéal de $\mathbb{K}[X]$ est principal. L'idéal des polynômes annulateurs de u admet donc un unique générateur normalisé.

DEFINITION 15.4.2
Soient E un $\mathbb{K}$-ev de dimension finie et $u \in L(E)$. L'unique générateur normalisé de l'idéal des polynômes annulateurs de u s'appelle polynôme minimal de u. On le notera μ_u.

Comme $\mathrm{Ann}(u) \neq \{0\}$, $\mu_u \neq 0$. On a donc, pour tout $P \in \mathbb{K}[X],\ P(u) = 0 \Leftrightarrow \mu_u$ divise P. Rappellons que μ_u est l'unique polynôme de $\mathrm{Ann}(u)$ normalisé de degré minimal.

PROPOSITION 15.4.3
1) Soit F un sous espace de E stable par $u \in L(E)$. Le polynôme minimal de u_F divise celui de u.
2) Si $E = F \oplus G$ avec F et G stables par u, on a $\mu_u = ppcm(\mu_{u_F}, \mu_{u_G})$.

preuve
Notons d'abord que si F est stable par u, alors, pour tout polynôme P, F est stable par $P(u)$ et l'endomorphisme induit sur F par $P(u)$ n'est autre que $P(u_F)$. Autrement dit $(P(u))_F = P(u_F)$.
1) Appliquant ceci avec $P = \mu_u$ on voit que le polynôme minimal μ_u de u annule u_F. Donc le polynôme minimal de u_F divise μ_u.
2) Dans une base adaptée à la décomposition $E = F \oplus G$ la matrice de u est de la forme $M = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ où $A = \mathrm{Mat}(u_F)$ et $B = \mathrm{Mat}(u_G)$. Alors pour tout polynôme $Q \in \mathbb{K}[X]$ on a $Q(M) = \begin{pmatrix} Q(A) & 0 \\ 0 & Q(B) \end{pmatrix}$ donc $Q(M) = 0 \Leftrightarrow (Q(A) = 0 \text{ et } Q(B) = 0)$. L'idéal annulateur de u est donc l'intersection des idéaux annulateurs de u_F et de u_G. D'où la conclusion.

THEOREME 15.4.2
L'ensemble des valeurs propres de u est égal à l'ensemble des racines du polynôme minimal de u.

preuve
D'après la proposition 15.4.2 toute valeur propre de u est racine de μ_u. Réciproquement, soit λ une racine de μ_u. On peut écrire $\mu_u(X) = (X - \lambda)Q(X)$. Il vient donc $(u - \lambda I)Q(u) = 0$. Supposons λ non valeur propre de u. Alors $u - \lambda I$ est un isomorphisme, et en composant $(u - \lambda I)Q(u) = 0$ à gauche par $(u - \lambda I)^{-1}$ on obtient $Q(u) = 0$; Q serait un polynôme annulateur de u de degré strictement inférieur à celui de μ_u. Contradiction. Donc λ est valeur propre de u.

EXEMPLE 15.4.1 (polynôme minimal d'une matrice compagnon)
Reprenons les notations de l'exemple 15.2.1. Soit $(e_1, \ldots, e_n)$ la base canonique de $\mathbb{K}^n$. On a, pour $1 \leqslant i \leqslant n-1$, $Ae_i = e_{i+1}$ donc $A^i e_1 = e_{i+1}$. Il en résulte que si $b_0, \ldots, b_{n-1}$ sont des scalaires non tous nuls,

$$\left(b_0 I + b_1 A + \cdots + b_{n-1} A^{n-1}\right)(e_1) = b_0 e_1 + \cdots + b_{n-1} e_n \neq 0$$

donc à fortiori $b_0 I + b_1 A + \cdots + b_{n-1} A^{n-1} \neq 0$ Autrement dit, un polynôme arbitraire non nul de degré inférieur ou égal à $n-1$ n'annule pas A. Son polynôme minimal est donc de degré au moins n. Par ailleurs, on a

$$A^n e_1 = Ae_n = -\left(a_0 e_1 + \cdots + a_{n-1} e_n\right) = -\left(a_0 I + a_1 A + \cdots + A^{n-1}\right)(e_1)$$

et donc $P(A)(e_1) = 0$ Comme $P(A)$ commute avec les puissances de A, on en déduit, pour $2 \leqslant j \leqslant n$ $P(A)(e_j) = P(A)A^{j-1}e_1 = A^{j-1}P(A)(e_1) = 0$ donc $P(A) = 0$. Par conséquent, P est annulateur pour A. Comme il est de degré n, il résulte de ce qui précède que c'est le polynôme minimal de A.

Le polynôme minimal de la matrice compagnon d'un polynôme P est donc égal à P.

15.4.2 Théorème de décomposition des noyaux

THEOREME 15.4.3
Soient E un $\mathbb{K}$-ev (non nécessairement de dimension finie), $u \in L(E)$ et $P, Q \in \mathbb{K}[X]$ deux polynômes premiers entre eux. On a

$$\ker\left((PQ)(u)\right) = \ker\left(P(u)\right) \oplus \ker\left(Q(u)\right)$$

Si en outre $(PQ)(u) = 0$, on a $E = \ker\left(P(u)\right) \oplus \ker\left(Q(u)\right)$ et la projection π sur $\ker\left(P(u)\right)$ parallélement à $\ker\left(Q(u)\right)$ est un polynôme en u.

preuve
D'après le théorème de Bézout, il existe deux polynômes A et B tels que $AP + BQ = 1$. Posons $a = A(u), b = B(u), p = P(u)$ et $q = Q(u)$. Les endomorphismes a, b, p, q commutent entre eux et on a

$$ap + bq = pa + qb = I$$

- Soit $x \in \ker(p) \cap \ker(q)$; il vient $x = a(p(x)) + b(q(x)) = 0$. Donc $\ker(p) \cap \ker(q) = \{0\}$; la somme des noyaux de p et q est directe.

- On a de manière évidente $\ker(p) \subset \ker(qp) = \ker(pq)$ et $\ker(q) \subset \ker(pq)$ donc $\ker(p) + \ker(q) \subset \ker(pq)$.
- Soit $x \in \ker(pq) = \ker(qp)$. Posons $y = a(p(x))$ et $z = b(q(x))$. On a $x = y + z$. D'autre part, par commutativité de $a, p, q,\ q(y) = qap(x) = a(qp(x)) = 0$ et de même $p(z) = pbq(x) = b((pq)(x)) = 0$. Par conséquent, $y \in \ker(q)$ et $z \in \ker(p)$ ce qui montre que $\ker(pq) \subset \ker(p) + \ker(q)$. Ceci achève la démonstration du premier point.
- Si $pq = 0$, le projeté de x sur $\ker(p)$ parallélement à $\ker(q)$ est, avec les notations précédentes l'élément z. Donc $\pi = bq = B(u)Q(u) = (BQ)(u)$.

COROLLAIRE 15.4.1
Soient E un $\mathbb{K}$-ev (non nécessairement de dimension finie), $u \in L(E)$, $P_1, \dots, P_m$ des polynômes premiers entre eux deux à deux et $P = P_1 \cdots P_m$. On a $\ker(P(u)) = \overset{k=m}{\underset{k=1}{\oplus}} \ker\big(P_k(u)\big)$. Si $P(u) = 0$, le projecteur sur $\ker\big(P_k(u)\big)$ parallélement à la somme des autres noyaux est un polynôme en u.

preuve
On procède par récurrence sur m en observant que si $P_1, \dots, P_m$ sont premiers entre eux deux à deux, alors P_1 est premier avec $P_2 \cdots P_m$. D'après le théorème précédent, $\ker\big(P(u)\big) = \ker\big(P_1(u)\big) \oplus \ker\big((P_2 \cdots P_m)(u)\big)$. Si P est annulateur de u, la projection sur $\ker\big(P_1(u)\big)$ parallélement à $\ker\big((P_2 \cdots P_m)(u)\big) = \ker\big(P_2(u)\big) \oplus \cdots \oplus \ker\big(P_m(u)\big)$ est un polynôme en u.

EXEMPLE 15.4.2
Un projecteur p de E est par définition un élément $p \in L(E)$ tel que $p^2 = p$. C'est donc un endomorphisme annulé par le polynôme $P(X) = X^2 - X = X(X-1)$. Le théorème de décomposition des noyaux donne $E = \ker(p) \oplus \ker(I-p)$, résultat bien connu pour les projecteurs.
De même, une symétrie vectorielle s est un endomorphisme involutif de E, c'est à dire un endomorphisme annulé par le polynôme $X^2 - 1 = (X-1)(X+1)$. On retrouve ainsi que $E = \ker(s - I) \oplus \ker(s + I)$.

15.4.3 CNS pour qu'un endomorphisme soit diagonalisable

Le théorème suivant est d'une grande importance pratique.

THEOREME 15.4.4
Soient E un $\mathbb{K}$-ev de dimension finie et $u \in L(E)$. Les propriétés suivantes sont équivalentes :
(1) u est diagonalisable.
*(2) Il existe un polynôme $P \in \mathbb{K}[X]$ **scindé à racines simples** tel que $P(u) = 0$.*
(3) Le polynôme minimal de u est scindé à racines simples.
Dans ces conditions, si $\mathrm{Sp}(u) = \{\lambda_1, \dots, \lambda_m\}$ on a $\mu_u(X) = (X - \lambda_1) \cdots (X - \lambda_m)$.

preuve
$(1) \Rightarrow (2)$ est facile ; si $\lambda_1, \dots, \lambda_m$ sont les valeurs propres distinctes de u, on voit, en utilisant une représentation matricielle par exemple, que le polynôme $Q(X) = (X - \lambda_1) \cdots (X - \lambda_m)$ est annulateur de u.
$(2) \Rightarrow (3)$ car le polynôme minimal de u divise tout polynôme annulateur et qu'un diviseur d'un polynôme scindé à racines simples est lui même un polynôme scindé à racines simples.
$(3) \Rightarrow (1)$ Par hypothèse $\mu_u(X) = (X - \lambda_1) \cdots (X - \lambda_m)$ où, d'après le théorème 15.4.2 les λ_j sont les valeurs propres de u. Les différents facteurs de μ_u sont premiers entre eux deux à deux, donc $E = \ker\big(\mu_u(u)\big) = \overset{k=m}{\underset{k=1}{\oplus}} \ker(u - \lambda_j I)$ ce qui prouve que u est diagonalisable.

Complément
Notons $L_1, \dots, L_m$ les polynômes d'interpolation de Lagrange associé au système de m points distincts $\lambda_1, \dots, \lambda_m$ de $\mathbb{K}$. (Voir Exemple 7.2.1). On a $1 = \sum_{k=1}^{k=m} L_k(X)$ donc $I = \sum_{k=1}^{k=m} L_k(u)$, d'où, pour tout $x \in E$, $x = \sum_{k=1}^{k=m} L_k(u)(x)$. Or le polynôme $(X - \lambda_k)L_k(X)$ est proportionnel au polynôme μ_u, donc annulateur de u. Il en résulte que $(u - \lambda_k I)\big(L_k(u)(x)\big) = 0$, autrement dit que $L_k(u)(x) \in \ker(u - \lambda_k I)$. La projection sur $\ker(u - \lambda_k I)$ parallélement à la somme des autres espaces propres est donc l'application $L_k(u)$.

COROLLAIRE 15.4.2
Soit u un endomorphisme diagonalisable d'un $\mathbb{K}$-ev de dimension finie E et F un sev de E stable par u. L'endomorphisme u_F de F induit par u est diagonalisable.

preuve
Puisque u est diagonalisable, il existe un polynôme scindé à racines simples Q tel que $Q(u)=0$; comme $\big(Q(u)\big)_F = Q(u_F)$ le même polynôme annule u_F donc u_F est diagonalisable.

15.4.4 Théorème de Cayley-Hamilton

THEOREME 15.4.5
1) Soit u un endomorphisme d'un $\mathbb{K}$-ev E de dimension finie. Le polynôme caractéristique χ_u de u est annulateur de u.
2) Soit $M \in M_n(\mathbb{K})$. Le polynôme caractéristique de M est annulateur de M.

Il est clair que (1) et (2) sont équivalents. Ce théorème admet plusieurs démonstrations. Nous en donnerons deux.

Première démonstration (utilisation des matrices compagnons)
On démontre (1). Soit $x \in E$. Il s'agit de montrer que $\chi_u(u)(x)=0$. On peut supposer $x \neq 0$.
Soit $p = \max\big\{k \in \mathbb{N} \ \big| \big(x, u(x), \dots, u^{k-1}(x)\big) \text{ est libre } \big\}$ et $F = \text{Vect}\big(x, u(x), \dots, u^{p-1}(x)\big)$.
Par hypothèse, $\big(x, u(x), \dots, u^{p-1}(x)\big)$ est libre et $\big(x, u(x), \dots, u^{p-1}(x), u^p(x)\big)$ est lié, donc il existe des scalaires a_0, a_1, $\dots, a_{p-1}$ tels que $u^p(x) = -\big(a_0 x + a_1 u(x) + \cdots + a_{p-1}u^{p-1}(x)\big)$. Il en résulte que le sous espace F est stable par u et que la matrice de u_F dans la base $\big(x, u(x), \dots, u^{p-1}(x), u^p(x)\big)$ est la matrice compagnon du polynôme $P = X^p + a_{p-1}X^{p-1} + \cdots + a_0$. Par construction, on a $P(u)(x)=0$. D'autre part, on sait que P est le polynôme caractéristique de M, donc de u_F. Par conséquent, P divise χ_u et il existe $Q \in \mathbb{K}[X]$ tel que $\chi_u = QP$. On en déduit $\chi_u(u)(x) = Q(u)\big(P(u)(x)\big) = 0$.■.

Deuxième démonstration On fait la preuve de (2).

- Comme $\mathbb{K} \subset \mathbb{C}$, il suffit de montrer le théorème pour $M \in M_n(\mathbb{C})$.
- M est trigonalisable : il existe $P \in GL(n,\mathbb{C})$ et T triangulaire supérieure telle que $P^{-1}MP = T$. On a $\chi_M = \chi_T$ et pour tout polynôme $Q \in \mathbb{C}[X]$, $Q(M) = PQ(T)P^{-1}$. Il suffit donc de montrer $Q(T)=0$, c'est à dire de montrer le théorème pour une matrice triangulaire supérieure.
- Soit $T = (t_{i,j})$ triangulaire supérieure. On a $\chi_T(X) = (t_{1,1}-X)\cdots(t_{n,n}-X)$. Soit V_k le sous espace vectoriel engendré par les k premiers vecteurs de la base canonique de $\mathbb{C}^n$ et $V_0 = \{0\}$. On vérifie facilement que $(T - t_{k,k}I_n)(V_k) \subset V_{k-1}$ pour tout k tel que $1 \leqslant k \leqslant n$. On en déduit

  $$\chi_T(T)(\mathbb{C}^n) = (T-t_{1,1}I)\cdots(T-t_{n,n}I)(V_n) \subset (T-t_{1,1}I)\cdots(T-t_{n-1,n-1}I)(V_{n-1}) \subset \cdots \subset (T-t_{1,1}I)(V_1) = \{0\}$$

  d'où la conclusion.

COROLLAIRE 15.4.3
Le polynôme minimal μ_u de u divise le polynôme caractéristique χ_u.

Si le corps de base est $\mathbb{C}$, ou plus généralement, si le polynôme caractéristique est scindé, il en est de même du polynôme minimal. Si $\text{Sp}(u) = \{\lambda_1, \dots, \lambda_p\}$ on aura alors la forme suivante pour ces deux polynômes :

$$\chi_u(X) = \prod_{k=1}^{k=p} (\lambda_j - X)^{m_j} \ ; \quad \mu_u(X) = \prod_{k=1}^{k=p} (X-\lambda_j)^{\beta_j} \quad \text{avec} \quad 1 \leqslant \beta_j \leqslant m_j \quad \text{et} \quad \sum_{k=1}^{k=p} m_k = \dim(E)$$

15.4.5 Applications

Application à l'algèbre $\mathbb{K}[A]$

Soient $A \in M_n(\mathbb{K})$, μ_A son polynôme minimal, et $q = \deg(\mu_A)$. On a donc $A^q \in \text{Vect}(I, A, \cdots, A^{q-1})$. Une récurrence facile montre alors que pour tout entier p on a encore $A^p \in \text{Vect}(I, A, \cdots, A^{q-1})$ et donc que tout polynôme en A appartient à $\text{Vect}(I, A, \cdots, A^{q-1})$. Autrement dit $\mathbb{K}[A] = \text{Vect}(I, A, \dots, A^{q-1})$. D'autre part, par définition du polynôme minimal, le système $(I, A, \cdots, A^{q-1})$ est libre, donc $\dim(\mathbb{K}[A]) = \deg(\mu_A)$.

Calcul de la puissance p-ième d'une matrice

Soit $A \in M_n(\mathbb{K})$. La division euclidienne du polynôme X^p par χ_A s'écrit $X^p = Q_p(X)\chi_A(X) + R_p(X)$ avec $R_p = 0$ ou $\deg(R_p) < \deg(\chi_A) = n$. On en déduit $A^p = Q_p(A)\chi_A(A) + R_p(A) = R_p(A)$ ce qui permet d'exprimer A^p comme combinaison linéaire de $I, A, \ldots, A^{n-1}$; On pourrait évidemment faire la même chose avec tout polynôme annulateur de A, en particulier avec le polynôme minimal dont le degré peut être strictement inférieur à n. Mais le polynôme caractéristique est souvent plus facile à déterminer que le polynôme minimal.

EXEMPLE 15.4.3
Soit à calculer A^p pour $A = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & -1 \\ 1 & 0 & 2 \end{pmatrix}$
On a $\chi_A(X) = (1-X)^2(2-X)$. Le reste R_p de la division de X^p par χ_A est de degré au plus 2. Posons $R_p(X) = a_pX^2 + b_pX + c_p$. On a donc

$$X^p = (1-X)^2(2-X)Q_p(X) + a_pX^2 + b_pX + c_p$$

Pour $X = 1$ puis $X = 2$ on obtient

$$\begin{aligned} 1 &= a_p + b_p + c_p \\ 2^p &= 4a_p + 2b_p + c_p \end{aligned}$$

Comme 1 est racine double de χ_A, il est aussi racine du polynôme dérivé. Il vient $pX^{p-1} = \chi'_A(X)Q_p(X) + \chi_A(X)Q'_p(X) + 2a_pX + b_p$ et en faisant $X = 1$ on obtient

$$p = 2a_p + b_p$$

d'où facilement $a_p = 2^p - (p+1),\ b_p = -2^{p+1} + 3p + 2$ et $c_p = 2^p - 2p$ et

$$A^p = (2^p - (p+1))\,A^2 + \left(-2^{p+1} + 3p + 2\right)A + (2^p - 2p)\,I_3 = 2^p(A^2 - 2A + I_3) - p(A^2 - 3A + 2I_3) - (A^2 - 2A)$$

15.5 Sous espaces caractéristiques

Pour des raisons qui apparaitront bientôt, nous commencerons cette section par une brève étude des endomorphismes nilpotents.

15.5.1 Endomorphismes nilpotents

Nous aurons besoin du lemme suivant, qui est un exercice classique d'algèbre linéaire.

Lemme 15.5.1 (Théorème des noyaux itérés)
Soit f un endomorphisme d'un $\mathbb{K}$-ev F de dimension finie m. Il existe un unique entier p vérifiant pour tout j entier : $j < p \Rightarrow \ker(f^j) \subsetneq \ker(f^{j+1})$ et $j \geqslant p \Rightarrow \ker(f^j) = \ker(f^{j+1})$. De plus $p \leqslant m$.

preuve
Pour tout j, on a $\ker(f^j) \subset \ker(f^{j+1})$. D'autre part, il est impossible que toutes ces inclusions soient strictes car les dimensions de ces noyaux formeraient une suite infinie strictement croissante d'entiers naturels majorée par m. On définit p comme le plus petit des entiers k tels que $\ker(f^k) = \ker(f^{k+1})$. On a donc déjà $j < p \Rightarrow \ker(f^j) \subsetneq \ker(f^{j+1})$.
Montrons maintenant que $\ker(f^j) = \ker(f^{j+1}) \Rightarrow \ker(f^{j+1}) = \ker(f^{j+2})$. On aura alors, par récurrence sur j que $j \geqslant p \Rightarrow \ker(f^j) = \ker(f^{j+1})$. Or, si $x \in \ker(f^{j+2})$, on a $f^{j+1}\big(f(x)\big) = 0$ soit $f(x) \in \ker(f^{j+1})$, donc par hypothèse, $f(x) \in \ker(f^j)$, soit $f^{j+1}(x) = 0$ i.e. $x \in \ker(f^{j+1})$. L'autre inclusion est triviale.
Posons $d_j = \dim\big(\ker(f^j)\big)$. Pour $j < p$ on a $d_j < d_{j+1}$. On en déduit facilement $p \leqslant d_p \leqslant m$.

Exercice
Avec les notations du lemme
1) Montrer que $j < p \Rightarrow \mathrm{im}(f^{j+1}) \subsetneq \mathrm{im}(f^j)$ et $j \geqslant p \Rightarrow \mathrm{im}(f^{j+1}) = \mathrm{im}(f^j)$
2) Montrer ensuite que $E = \ker(f^p) \oplus \mathrm{im}(f^p)$, que ces deux espaces sont stables par f et que f induit un isomorphisme de $\mathrm{im}(f^p)$.

DEFINITION 15.5.1
Soit F un $\mathbb{K}$-ev de dimension finie m et $\nu \in L(F)$. On dit que ν est nilpotent si il existe un entier naturel q tel que $\nu^q = 0$. Le plus petit de ces entiers est appelé indice de nilpotence de ν.

L'endomorphisme ν est nilpotent d'indice 1 ssi il est nul. Sinon, ν est nilpotent d'indice β si $\nu^\beta = 0$ et $\nu^{\beta-1} \neq 0$.
Le polynôme minimal de ν est un diviseur de X^β, donc de la forme X^h. Par définition de l'indice de nilpotence, on voit que en fait c'est X^β. On a donc

$$\nu \text{ nilpotent d'indice } \beta \Leftrightarrow \mu_\nu(X) = X^\beta$$

On en déduit en particulier que $\beta \leqslant m$.

THEOREME 15.5.1
Soient F un $\mathbb{K}$-ev de dimension finie m et $\nu \in L(F)$. Les proprités suivantes sont équivalentes :
(1) *ν est nilpotent.*
(2) *Il existe une base $\mathcal{B}$ de F telle que $\mathrm{Mat}_{\mathcal{B}}(\nu)$ soit triangulaire avec des zéros sur la diagonale.*
(3) *Le polynôme caractéristique de ν est $\chi_\nu(X) = (-X)^m$.*

preuve
(1) $\Rightarrow$ (2) Soit β l'indice de nilpotence de ν. D'après le lemme des noyaux itérés, la suite $\{0\} \subset \ker(\nu) \subset \cdots \subset \ker(\nu^\beta) = E$ est strictement croissante. (Ici l'entier p du lemme est égal à β). Soit $\mathcal{B}_1$ une base de $\ker(\nu)$. On la complète en une base $\mathcal{B}_2$ de $\ker(\nu^2)$, puis on complète $\mathcal{B}_2$ en une base de $\ker(\nu^3)$, etc. A la fin de ce processus, on obtient une base $\mathcal{B}$ de F dans laquelle la matrice de ν est triangulaire avec des zéros sur la diagonale.
(2) $\Rightarrow$ (3) est trivial.
(3) $\Rightarrow$ (1) résulte immédiatement du théorème de Cayley-Hamilton.

15.5.2 Sous espaces caractéristiques

DEFINITION 15.5.2
Soient E un $\mathbb{K}$-ev non nul de dimension finie n et $u \in L(E)$. Soit $\lambda \in K$ une valeur propre de u de multiplicité m_λ. On appelle sous espace caractéristique associé à la valeur propre λ le sous espace $F_\lambda = \ker\left((u - \lambda I)^{m_\lambda}\right)$.

On définit de même les espaces caractéristiques d'une matrice carrée $M \in M_n(\mathbb{K})$.

Si $E_\lambda = \ker(f - \lambda I)$ est l'espace propre associé à λ, on a l'inclusion $E_\lambda \subset F_\lambda$.
Puisque u commute avec $(u - \lambda I)^{m_\lambda}$, F_λ est stable par u. Notons u_λ l'endomorphisme de F_λ induit par u. Par défintion de F_λ il vérifie $(u_\lambda - \lambda I_{F_\lambda})^{m_\lambda} = 0$. L'endomorphisme $\nu_\lambda = u - \lambda I_{F_\lambda}$ de F_λ est donc nilpotent. Avec ces notations, on a

THEOREME 15.5.2
1) $\dim(F_\lambda) = m_\lambda$.
2) *L'indice de nilpotence de ν_λ est égale à la multiplicité de λ comme racine du polynôme **minimal** de u. Notons la β_λ.*
3) $F_\lambda = \ker\left((u - \lambda I)^{\beta_\lambda}\right)$.

preuve
Dans la démonstration, nous utiliserons la propriété suivante de vérification immédiate

Propriété
Soit f un endomorphisme d'un $\mathbb{K}$-ev E et $t \in \mathbb{K}$. On a $\chi_{f+tI}(X) = \chi_f(X - t)$ et $\mu_{f+tI}(X) = \mu_f(X - t)$

On peut écrire

$$\chi_u(X) = (X - \lambda)^{m_\lambda} Q(X) \quad \text{avec} \quad Q(\lambda) \neq 0$$

On déduit du théorème de décomposition des noyaux que

$$E = F_\lambda \oplus S \quad \text{avec} \quad S = \ker\big(Q(u)\big)$$

On a ainsi une décomposition de E en somme directe de deux sous espaces stables par u. On a donc

$$\chi_u(X) = \chi_{u_\lambda}(X)\chi_{u_S}(X) \quad \text{et} \quad \mu_u(X) = \mathrm{ppcm}\big(\mu_{u_\lambda}(X), \mu_S(X)\big)$$

De plus, on a $Q(u_S) = 0$ et $Q(\lambda) \neq 0$ donc λ n'est pas valeur propre de u_S et par conséquent n'est racine ni de χ_{u_S} ni de μ_{u_S}.

- Posons $d = \dim(F_\lambda)$. ν_λ est nilpotent, donc son polynôme caractéristique est $(-X)^d$. D'après la propriété, le polynôme caractéristique de $u_\lambda = \nu_\lambda + \lambda I_{F_\lambda}$ vaut $\chi_{u_\lambda}(X) = (\lambda - X)^d$. On obtient donc $\chi_u(X) = (\lambda - X)^d \chi_{u_S}(X)$ avec $\chi_{u_S}(\lambda) \neq 0$, donc $d = m_\lambda$ ce qui prouve le 1).

- Désignons par β l'indice de nilpotence de ν_λ et par p la multiplicité de λ comme racine de $\mu_u(X)$. Le polynôme minimal de ν est X^β, donc, d'après la propriété celui de u_λ est $\mu_{u_\lambda}(X) = (X-\lambda)^\beta$. On a vu que λ n'était pas racine de μ_S, donc les polynômes μ_S et μ_{u_λ} sont premiers entre eux et leur ppcm est égal à leur produit. On obtient $\mu_u(X) = (X-\lambda)^\beta \mu_{u_S}(X)$ avec $\mu_{u_S}(\lambda) \neq 0$ donc $p = \beta$. Ceci prouve la deuxième assertion du théorème, en posant $\beta_\lambda = \beta = p$.
- La troisième en résulte de suite.

15.5.3 Décomposition de Dumford

Dans cette partie, on considére un endomorphisme u dont le polynôme caractéristique est scindé sur $\mathbb{K}$. Il en est donc de même du polynôme minimal. Posons

$$\chi_u(X) = \prod_{k=1}^{k=p} (\lambda_k - X)^{m_k} \quad \text{et} \quad \mu_u(X) = \prod_{k=1}^{k=p} (X-\lambda_k)^{\beta_k}$$

où les λ_k sont deux à deux distincts. Les polynômes $(\lambda_k - X)^{m_k}$ sont deux à deux premiers entre eux. Le théorème de décomposition des noyaux montre qu'alors E est égale à la somme directe des sous espaces caractéristiques de u

$$E = F_1 \oplus \cdots \oplus F_p$$

en posant $F_k = F_{\lambda_k} = \ker\left((u-\lambda_k I)^{m_k}\right)$. Chacun de ces sous espaces de E est stable par u. On a vu que $\dim(F_k) = m_k$. Nous noterons u_k l'endomorphisme induit par u sur F_k et $\nu_k = u_k - \lambda_k I_{F_k}$. On sait aussi que ν_k est nilpotent d'indice β_k.

Soit pour tout k, $\mathcal{B}_k$ une base de F_k. Alors $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_p)$ est une base de E et dans cette base la matrice A de u est diagonale par blocs

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_p \end{pmatrix}$$

où $A_k \in M_{m_k}(\mathbb{K})$ est la matrice de u_k dans la base $\mathcal{B}_k$.

Résumons ceci :

THEOREME 15.5.3
Soit $u \in L(E)$ un endomorphisme d'un $\mathbb{K}$-ev E de dimension finie $n \geqslant 1$ dont le polynôme caractéristique est scindé sur $\mathbb{K}$. E est égal à la somme directe des sous espaces caractéristiques.
Si $\mathcal{B}$ est une base de E adaptée à la décomposition en somme directe $E = \overset{k=p}{\underset{k=1}{\oplus}} F_{\lambda_k}$, la matrice de u dans $\mathcal{B}$ est diagonale par blocs, les différents blocs étant de format (m_k, m_k).

COROLLAIRE 15.5.1
Soit $u \in L(E)$ un endomorphisme d'un $\mathbb{K}$-ev E de dimension finie $n \geqslant 1$ dont le polynôme caractéristique est scindé sur $\mathbb{K}$. u est diagonalisable ssi pour toute valeur propre λ_k de u, l'espace propre associé est égal à l'espace caractéristique associé.

preuve
Soit $E_k = \ker(f - \lambda_k I)$. On sait que $E_k \subset F_k$. Si pour tout k, $E_k = F_k$, E est somme directe des espaces propres et u est diagonalisable.
Réciproquement, si u est diagonalisable, on a $E = E_1 \oplus \cdots \oplus E_p \subset F_1 \oplus \cdots \oplus F_p = E$. On en déduit $\dim(E) = \sum_{k=1}^{k=p} \dim(E_k) \leqslant \sum_{k=1}^{k=p} \dim(F_k) = \dim(E)$ ce qui ne peut se produire, compte tenu de $\dim(E_k) \leqslant \dim(F_k)$ pour tout k, que si $\dim(E_k) = \dim(F_k)$ et donc $E_k = F_k$ pour tout k. On retrouve ainsi le fait que u est diagonalisable ssi pour tout k on a $\dim(E_k) = m_k$.

THEOREME 15.5.4 (Décomposition de Dumford)
Soit $u \in L(E)$ un endomorphisme d'un $\mathbb{K}$-ev de dimension finie dont le polynôme caractéristique est scindé sur $\mathbb{K}$. Il existe un unique couple (d, n) d'endomorphismes de E vérifiant les trois propriétés suivantes :

1. *d est diagonalisable, n est nilpotent.*
2. *d et n commutent : $dn = nd$.*
3. *$u = d + n$.*

De plus

4 *Il existe des polynômes P et Q (dépendant de u) tels que $d = P(u)$ et $n = Q(u)$.*

5 *L'indice de nilpotence de n est le maximum des ordres de multiplicité β_k des racines du polynôme minimal.*

preuve
Gardons les notations du théorème précédent.

Existence

On a $E = \overset{k=p}{\underset{k=1}{\oplus}} F_k$. Soit π_k le projecteur sur F_k parallèlement à la somme des $F_j,\ j \neq k$. Il résulte du théorème de décomposition des noyaux que les π_k sont des polynômes en u : on dispose donc de polynômes P_k tels que $P_k(u) = \pi_k$.
Définissons d comme l'unique endomorphisme de E dont la restriction à F_k est l'application $x \to \lambda_k x$. Autrement dit, $d = \sum_{k=1}^{k=p} \lambda_k \pi_k$. C'est évidemment un endomorphisme diagonalisable et $d = P(u)$ en posant $P = \sum_{k=1}^{k=p} \lambda_k P_k$. On définit ensuite n par $n = u - d$. C'est un polynôme en u et l'endomorphisme induit sur F_k par n est l'endomorphisme ν_k qui est nilpotent d'indice β_k. On en déduit que n est nilpotent d'indice $\max(\beta_k)$. Enfin n et d commutent puisque ce sont des polynômes en u. On a donc prouvé l'existence d'un couple (d, n) vérifiant 1,2,3 et que ce couple vérifiant aussi 4 et 5.

Unicité

Soit (d', n') un autre couple d'endomorphismes vérifiant les propriétés 1), 2) et 3) et (n, d) **celui construit précédemment**. Puisque d' commute avec n', il commute avec $n' + d' = u$ donc aussi avec tout polynôme en u, donc avec n et d. On a alors $d - d' = n' - n$.

- $d - d'$ est diagonalisable car d et d' sont deux endomorphismes diagonalisables qui commutent, donc ils se diagonalisant dans une même base (théorème 15.3.2). Dans cette base, la matrice de $d - d'$ est diagonale.
- $n' - n$ est nilpotent car n et n' commutent et sont nilpotents : si $n^\beta = 0$ et $n'^{\beta'} = 0$, on a $(n' - n)^{\beta+\beta'} = \sum_{k=0}^{\beta+\beta'} C^k_{\beta+\beta'} n^k n'^{\beta+\beta'-k}$. Et dans cette somme chacun des termes est nul, car soit $k \geqslant \beta$ et $n^k = 0$, soit $k < \beta$ et $\beta + \beta' - k > \beta'$ et $n'^{\beta+\beta'-k} = 0$.
- L'endomorphisme $h = d - d' = n' - n$ est à la fois diagonalisable et nilpotent. Sa seule valeur propre est 0 et sa matrice dans une base convenable est diagonale, donc $h = 0$. Par conséquent, $d = d'$ et $n = n'$.

15.6 Application 1 : équations différentielles linéaires à coefficients constants

Introduction

Soit $n \geqslant 1$ un entier naturel, $a_0, \ldots, a_{n-1}$ n nombres complexes fixés ; on posera $a_n = 1$. On introduit le polynôme $P(X) = \sum_{k=0}^{k=n} a_k X^k$.

On considère l'équation différentielle

$$(\mathcal{E}_P) \qquad y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0$$

On se donne I un intervalle de $\mathbb{R}$, non vide et non réduit à un point. Par définition, une solution de $\mathcal{E}_P$ sur I est une fonction $f : I \to \mathbb{C}$ n-fois dérivable sur I telle que $\forall x \in I\ f^{(n)}(x) + a_{n-1} f^{(n-1)}(x) + \cdots + a_1 f'(x) + a_0 f(x) = 0$. On notera E_P l'ensemble des solutions de $\mathcal{E}_P$ sur I. On voit facilement que tous les éléments de E_P sont des fonctions de classe C^∞. Soit $D : C^\infty(I, \mathbb{C}) \to C^\infty(I, \mathbb{C}),\ D(\varphi) = \varphi'$ la dérivation. On a $E_P = \ker(P(D))$. Par conséquent E_P est un sous espace vectoriel de $C^\infty(I, \mathbb{C})$ stable par D.

Pour $\lambda \in \mathbb{C}$ on notera e_λ la fonction $I \to \mathbb{C}$ définie par $e_\lambda(t) = e^{\lambda t}$. On sait (exemple 15.1.2) que la droite vectorielle de base e_λ est l'espace propre de D associé à la valeur propre λ.

Etude de cas particuliers

- Cas $P(X) = X - \lambda$.
  $\mathcal{E}_P$ est l'équation $y' - \lambda y = 0$. Dans ce cas E_P est la droite vectorielle engendrée par la fonction e_λ.
- Cas $P(X) = X^n$.
  $\mathcal{E}_P$ est l'équation $y^{(n)} = 0$ et E_P est l'espace vctoriel des fonctions polyômes de degré inférieur ou égal à $n-1$. On a $\dim(E_P) = n = \deg(P)$.
- Cas $P(X) = (X-\lambda)^n$.
  Soit $g \in C^\infty(I, \mathbb{C})$. On a $(D - \lambda I)(e_\lambda \cdot g) = e_\lambda g' = e_\lambda D(g)$, d'où par récurrence $(D-\lambda I)^n(e_\lambda(g)) = e_\lambda D^n(g)$.
  Soit alors $f \in C^\infty(I, \mathbb{C})$ et $g = e_{-\lambda} f$. On a donc $(D-\lambda I)^n(f) = e_\lambda D^n(g)$ de sorte que $f \in E_P \Leftrightarrow D^n(g) = 0 \Leftrightarrow g$ est une fonction polynôme de degré au plus $n-1$. L'ensemble E_P est l'ensemble des fonctions de la forme $t \to e^{\lambda t}\left(c_0 + c_1 t + \cdots + c_{n-1}t^{n-1}\right)$. Il est encore de dimension $n = \deg(P)$.

Cas général

Ecrivons $P(X) = \prod_{k=1}^{k=p} (X - \lambda_k)^{m_k}$ où les λ_k sont des complexes deux à deux distincts. Le théorème de décomposition des noyaux donne

$$E_P = \ker(P(D)) = \bigoplus_{k=1}^{k=p} \ker\left((D - \lambda_k I)^{m_k}\right)$$

On en déduit que E_P est de dimension finie égale à $m_1 + \cdots + m_p = n = \deg(P)$. Une fonction f est solution de $\mathcal{E}_P$ ssi elle est de la forme $f(t) = \sum_{k=1}^{k=p} e^{\lambda_k t} Q_k(t)$ où Q_k est un polynôme arbitraire de degré au plus m_{k-1}.
Résumons les résultats obtenus :

THEOREME 15.6.1
Si $P(X) = X^n + \sum_{j=0}^{n-1} a_j X^j = \prod_{k=1}^{k=p} (X-\lambda_k)^{m_k} \in \mathbb{C}[X]$, les solutions de l'équation différentielle linéaire à coefficients constants $y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0$ forment un $\mathbb{C}$-ev de dimension n. Toute solution s'écrit de manière unique sous la forme $f(t) = \sum_{k=1}^{k=p} e^{\lambda_k t} Q_k(t)$ où Q_k est un polynôme de degré au plus m_{k-1}, ou, ce qui revient au même, les fonctions $t \to t^i e^{\lambda_k t},\ 1 \leqslant k \leqslant p,\ 0 \leqslant i \leqslant m_k - 1$ forment une base de E_P.

Sous espaces caractéristiques de D

Notons encore D l'endomorphisme de E_P induit par la dérivation. On va à titre d'exemple, déterminer le polynôme caractéristique, le polynôme minimal et les espaces caractéristiques de D.

Notons U_i la fonction polynôme sur I, $t \to t^i$ et $\mathcal{B}_k$ la base de F_k formée des fonctions $U_i e_{\lambda_k}$ $0 \leqslant i \leqslant m_k - 1$. On a $D(U_i e_{\lambda_k}) = iU_{i-1}e_{\lambda_k} + \lambda_k U_i e_{\lambda_k}$ en posant $U_{-1} = 0$. La matrice de l'endomorphisme D_k induit par D sur F_{λ_k} est donc

$$M_k = \begin{pmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & 2 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & & \lambda_k & m_k - 1 \\ 0 & \cdots & \cdots & 0 & \lambda_k \end{pmatrix} = \lambda_k I_{m_k} + N_k$$

où N_k est nilpotente d'indice de nilpotence m_k. Le polynôme minimal de D_k est donc $(X - \lambda_k)^{m_k}$ et son polynôme caractéristique $(\lambda_k - X)^{m_k}$.
E est la somme directe des sous espaces stables E_k, donc le polynôme minimal de D est le ppcm de ceux des D_k, c'est à dire P.
Dans la base $(\mathcal{B}_1, \cdots, \mathcal{B}_p)$ de E_P, la matrice M de D est diagonale par blocs. On voit donc que son polynôme caractéristique est $(-1)^p P$.
Enfin, on a $F_k = \ker(D - \lambda_k I)^{m_k}$ donc les F_k sont les espaces caractéristiques de D.

Complément : solutions réelles

Considérons maintenant le cas où tous les a_i sont réels et où on cherche les solutions réelles de l'équation. Si P est scindé sur $\mathbb{R}$, les résultats précédents s'appliquent. Sinon, on peut écrire

$$P(X) = \left(\prod_{k=1}^{k=r}(X-\lambda_k)^{m_k}\right)\left(\prod_{j=1}^{j=q}((X-z_j)(X-\overline{z}_j))^{p_j}\right)$$

où les λ_k sont des réels distincts, les z_j des complexes non réels avec $z_j \neq \overline{z}_{j'}$ pour tout couple (j,j'). En appliquant le théorème de décomposition des noyaux, comme ci dessus, tout ce qu'il reste à voir est ce qui se passe pour un polynôme P de la forme $P(X) = ((X-z)(X-\overline{z}))^p$. On pose $z = s+i\omega,\ s,\omega \in \mathbb{R}$ et $\omega \neq 0$. Notons $S_{\mathbb{C}}$ (resp. $S_{\mathbb{R}}$) l'ensemble des solutions complexes (resp. réelles) de l'équation $P(D)(y) = 0$. Les fonctions $t \to t^j e^{st}e^{i\omega t}$ et $t \to t^j e^{st}e^{-i\omega t}$ $(0 \leqslant j \leqslant p-1)$ forment une base de $S_{\mathbb{C}}$ sur $\mathbb{C}$. Il en est donc de même des fonctions $t \to t^j e^{st}\cos(\omega t)$ et $t \to t^j e^{st}\sin(\omega t)$.

Toute fonction appartenant à $S_{\mathbb{R}}$ est dans $S_{\mathbb{C}}$ donc de la forme $t \to \sum_{j=0}^{p-1} t^j e^{st}\left(A_j\cos(\omega t) + B_j\sin(\omega t)\right)$.

Lemme
Soient A_j, B_j des nombres complexes, s,ω des réels, $\omega \neq 0$ et I un intervalle non vide et non réduit à un point de $\mathbb{R}$. Si la fonction $t \to h(t) = \sum_{j=0}^{p-1} t^j e^{st}\left(A_j\cos(\omega t) + B_j\sin(\omega t)\right)$ ne prend que des valeurs réelles sur I, alors les nombres A_j et B_j sont tous réels.

En effet, la condition s'écrit $\sum_{j=0}^{p-1}\left(t^j e^{st}\mathrm{Im}(A_j)\cos(\omega t) + t^j e^{st}\mathrm{Im}(B_j)\sin(\omega t)\right) = 0$. Or les fonctions $t \to t^j e^{st}cos(\omega t)$ et $t \to t^j e^{st}\sin(\omega t)$, $0 \leqslant j \leqslant p$ sont $\mathbb{C}$-linéairement indépendantes, donc à fortiori $\mathbb{R}$-linéairement indépendantes.

On a donc prouvé

THEOREME 15.6.2
Si $P(X) = X^n + \sum_{j=0}^{n-1} a_j X^j = \left(\prod_{k=1}^{k=r}(X-\lambda_k)^{m_k}\right)\left(\prod_{j=1}^{j=q}((X-z_j)(X-\overline{z}_j))^{p_i}\right)$*, où les* z_j *sont des complexes deux à deux distincts et tels que* $z_j \neq \overline{z}_{j'}$ *pour tout* (j,j')*, les solutions réelles de l'équation différentielle linéaire à coefficients constants* $y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0$ *forment un* $\mathbb{R}$*-ev de dimension* n*.*
Toute solution s'écrit de manière unique sous la forme $f(t) = \sum_{k=1}^{k=r} e^{\lambda_k t}Q_k(t) + \sum_{j=1}^{j=q} e^{s_j t}\left(R_j(t)\cos(\omega_j t) + S_j(t)\sin(\omega_j t)\right)$ *où on a posé* $z_j = s_j + i\omega_j$ *et où* Q_k, R_j, S_j *sont des polynômes à coefficients réels de degré strictement inférieur , respectivement à* m_k, p_j, p_j*.*

Cas particulier
Rappelons les résultats obtenus dans le cas d'une équation du second ordre à coefficients réels, $(\mathcal{E})\ \ ay'' + by' + cy = 0\ a \neq 0$. L'équation caractéristique est $aX^2 + bX + c = 0$. Il y a trois cas :

- $b^2 - 4ac > 0$. Si r et r' sont les racines réelles de l'équation caractéristique, la solution générale de $(\mathcal{E})$ est $y(x) = Ae^{rx} + Be^{r'x},\ A, B$ constantes réelles.
- $b^2 - 4ac = 0$ La solution générale est $y(x) = (Ax+B)e^{-bx/2a}\ A, B$ constantes réelles.
- $b^2 - 4ac < 0$ Si on écrit les deux racines de l'équation caractéristique sous la forme $r \pm i\omega$ avec $r \in \mathbb{R}$ et $\omega \in \mathbb{R}^*$, la solution générale de $(\mathcal{E})$ est $y(x) = e^{rx}\left(A\cos(\omega x) + B\sin(\omega x)\right)\ A, B$ constantes réelles.

15.7 Application 2 : suites récurrentes linéaires

Introduction

Soient $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, $n \geqslant 1$ un entier naturel, $a_0, \ldots, a_{n-1}$ n éléments de $\mathbb{K}$ non tous nuls. On s'intéresse à l'ensemble E_P des suites $(u_k)_{k\in\mathbb{N}}$ d'éléments de $\mathbb{K}$ vérifiant la relation de récurrence

$$(\mathcal{R}) \qquad \forall k \in \mathbb{N}\ \ u_{k+n} = a_{n-1}u_{k+n-1} + \cdots + a_1 u_{k+1} + a_0 u_k = 0$$

On se limitera au cas où $a_0 \neq 0$ car sinon, par décalage d'indice, on est ramené à une récurrence linéaire d'ordre $n-1$.

On vérifie facilement que E_P est un sous espace vectoriel du $\mathbb{K}$-ev $\mathbb{K}^{\mathbb{N}}$ des suites à valeurs dans $\mathbb{K}$. Pout tout élément $(x_0, \ldots, x_{n-1}) \in \mathbb{K}^n$, il existe une unique suite appartenant à E_P vérifiant $u_j = x_j$ pour $0 \leqslant j \leqslant n-1$.

THEOREME 15.7.1
Avec les notations précédentes, l'application Φ qui à $u = (u_k) \in E_P$ associe $\Phi(u) = (u_0, \ldots, u_{n-1}) \in \mathbb{K}^n$ est un isomorphisme de E_P sur $\mathbb{K}^n$. En particulier, E_P est un $\mathbb{K}$-ev de dimension n.

On introduit le polynôme $P(X) = X^n - \sum_{k=0}^{n-1} a_k X^k$.

Soit $\Sigma : \mathbb{K}^{\mathbb{N}} \to \mathbb{K}^{\mathbb{N}}$ l'endomorphisme de décalage, défini par $\Sigma(u) = v$ où v est telle que $v_k = u_{k+1}$ pour tout k. On a, pour tout $u \in \mathbb{K}^{\mathbb{N}}$ et tout entier r $(\Sigma^r(u))_k = u_{r+k}$ de sorte que la relation $\mathcal{R}$ s'écrit $P(\Sigma)(u) = 0$. On a donc $E_P = \ker(P(\Sigma))$. On retrouve le fait que E_P est un sous espace vectoriel de $\mathbb{K}^{\mathbb{N}}$; de plus il est stable par Σ. Nous noterons S l'endomorphisme de E_P induit par Σ.

Matrice de S

Soit $\mathcal{B} = (U_1, \ldots, U_n)$ la base de E_P image par Φ^{-1} de la base canonique de $\mathbb{K}^n$. U_j est donc la suite vérifiant $\mathcal{R}$ et donc la condition initiale est le j-ième vecteur de la base canonique de $\mathbb{K}^n$.

THEOREME 15.7.2
La matrice A dans la base $\mathcal{B}$ de l'endomorphisme S est la transposée de la matrice compagnon du polynôme P:

$$A = \begin{pmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \end{pmatrix}$$

COROLLAIRE 15.7.1
Le polynôme caractéristique (resp. le polynôme minimal) de S est $(-1)^n P$ (resp· P).

preuve
La matrice A de S dans $\mathcal{B}$ est égale à la matrice de $S' := \Phi \circ S \circ \Phi^{-1}$ dans la base canonique de $\mathbb{K}^n$. Soit $X \in \mathbb{K}^n$ tel que ${}^tX = (x_0, \ldots, x_{n-1})$. $S'(X)$ est la condition initiale de l'image par S de la suite de condition initiale X, donc $S'(X)$ est l'élément de $\mathbb{K}^n$ qui en ligne s'écrit $(x_1, \cdots, x_{n-1}, x_n)$ avec $x_n = a_{n-1}x_{n-1} + \cdots + a_1 x_1 + a_0 x_0$. On en déduit A.

Eléments propres de S

Il en résulte que les valeurs propres de S sont les racines de P. L'hypothèse $a_0 \neq 0$ garantit qu'une telle valeur propre est non nulle. Soit λ une valeur propre de S. La suite u est un vecteur propre associé ssi $S(u) = \lambda u$ i.e. ssi $\forall k \in \mathbb{N},\ u_{k+1} = \lambda u_k$. On a donc

PROPOSITION 15.7.1
Les valeurs propres de S sont les racines de P. L'espace propre associé à la racine λ est de dimension 1, engendré par la suite géométrique de raison (λ).

Dans la suite, nous noterons par g_λ la suite géométrique de raison λ, telle que $g_\lambda(n) = \lambda^n$.

Etude complète

Nous supposerons P scindé sur $\mathbb{K}$. Ecrivons $P(X) = \prod_{k=1}^{k=p} (X-\lambda_k)^{m_k}$ où les λ_k sont des éléments de $\mathbb{K}$ deux à deux distincts. On a

$$E_P = \ker(P(S)) = \bigoplus_{k=1}^{k=p} \ker\left((S-\lambda_k I)^{m_k}\right)$$

et donc tout revient à étudier le cas $P(X) = (X-\lambda)^m$.

- Le cas $\lambda = 1$
  Si $m = 1$, E_{X-1} est l'ensemble des suites constantes. On vérifie sans peine que si $m = 2$, $E_{(X-1)^2}$ est l'ensemble des suites de la forme $n \to an + b$ avec a, b constantes. Nous allons généraliser ces résultats.

  Soit $\Delta : \mathbb{K}[X] \to \mathbb{K}[X]$ l'opérateur différence qui au polynôme Q associe le polynôme $\Delta(Q)$ tel que $\Delta(Q)(X) = Q(X+1) - Q(X)$. Si Q est constant, $\Delta(Q) = 0$. Un petit calcul montre que si $\deg(Q) = q \geqslant 1$ alors $\deg(\Delta(Q)) = q-1$. On en déduit que pour tout entier m, $\ker(\Delta^m)$ est égal à l'espace $\mathbb{K}_{m-1}[X]$ des polynômes de degré au plus $m-1$.

  Soit $Q \in \mathbb{K}_{m-1}[X]$ et u la suite définie par $u_k = Q(k)$. On a $\big((S-I)(u)\big)_k = \Delta(Q)(k)$ d'où $\big((S-I)^m(u)\big)_k = 0$ pour tout k. L'ensemble des suites de la forme $(Q(k))_{k\in\mathbb{N}}$ pour $Q \in \mathbb{K}_{m-1}[X]$ est donc contenu dans $\ker\big((S-I)^m\big)$. Or cet ensemble de suites est un sous espace vectoriel isomorphe à $\mathbb{K}_{m-1}[X]$ donc de dimension m. Comme $\dim\left(\ker\left((S-I)^m\right)\right) = m$, on a l'égalité.

  Conclusion : $\ker\left((S-I)^m\right) = E_{(X-1)^m}$ est l'ensemble des suites de la forme $\big(Q(k)\big)_{k\geqslant 0}$ pour Q polynôme de degré au plus $m-1$.

- réduction au cas $\lambda = 1$
  Si u et v sont deux suites, on note $u \cdot v$ leur produit défini par $(u\cdot v)_k = u_k v_k$. Soit v une suite quelconque. On a

  $$\left((S-\lambda I)(g_\lambda \cdot v)\right)_k = \lambda^{k+1} v_{k+1} - \lambda^{k+1} v_k = \lambda\left(\lambda^k (v_{k+1} - v_k)\right) = \lambda\left(g_\lambda \cdot (S-I)(v)\right)_k$$

  soit

  $$(S-\lambda I)(g_\lambda \cdot v) = \lambda\left(g_\lambda \cdot (S-I)(v)\right)$$

  d'où par une récurrence immédiate

  $$(S-\lambda I)^m (g_\lambda \cdot v) = \lambda^m \left(g_\lambda \cdot (S-I)^m(v)\right)$$

  Soit alors u une suite quelconque. Comme $\lambda \neq 0$, on peut définir une suite v en posant pour tout k $v_k = u_k/\lambda^k$ de sorte que $u = g_\lambda \cdot v$. En appliquant la dernière relation, on voit que

  $$u \in \ker\left((S-\lambda I)^m\right) \Leftrightarrow v \in \ker\left((S-I)^m\right) \Leftrightarrow \exists Q \in \mathbb{K}_{m-1}[X] \ \ \forall k \ v_k = Q(k)$$

  Conclusion : $\ker\left((S-\lambda I)^m\right) = E_{(X-\lambda I)^m}$ est l'ensemble des suites u de la forme $u_k = \lambda^k Q(k)$ où Q est un polynôme de degré au plus $m-1$.

THEOREME 15.7.3

Soit $P(X) = X^n - \sum_{k=0}^{n-1} a_k X^k = \prod_{k=1}^{k=p} (X-\lambda_k)^{m_k}$ *avec* $a_0 \neq 0$. *L'ensemble des suites* (u) *vérifiant la relation de récurrence*

$$(\mathcal{R}) \qquad \forall k \in \mathbb{N} \ \ u_{k+n} = a_{n-1}u_{k+n-1} + \cdots + a_1 u_{k+1} + a_0 u_k$$

est un $\mathbb{K}$ *espace vectoriel de dimension* n. *Il est formé des suites* (u_k) *de la forme*

$$u_k = \sum_{j=1}^{j=p} \lambda_j^k Q_j(k)$$

où les Q_k *sont des polynômes à coefficients dans* $\mathbb{K}$ *de degré au plus* $m_j - 1$.

Pour finir, résumons les résultats obtenus dans le cas $n = 2$. On se donne donc a, b réels ou complexes avec $b \neq 0$ et on étudie les suites (u_k) vérifiant

$$(\mathcal{R}) \qquad \forall k \geqslant 0 \ u_{k+2} = a u_{k+1} + b u_k$$

L'équation caractéristique de la suite est $X^2 - aX - b = 0$.

- Solutions sur $\mathbb{C}$
  Il y a deux cas.
  1) $a^2+4b \neq 0$. L'équation caractéristique admet deux racines distinctes λ_1 et λ_2 et une suite u vérifie la relation $\mathcal{R}$ ssi il existe deux constantes complexes C et D telles que $\forall k \in \mathbb{N}\ u_k = C\lambda_1^k + D\lambda_2^k$.
  2) $a^2+4b = 0$ L'équation caractéristique admet une racine double $\lambda = a/2$. Alors une suite u vérifie la relation $\mathcal{R}$ ssi il existe deux constantes complexes C et D telles que $\forall k \in \mathbb{N}\ u_k = (Ck+D)\lambda^k$.

- Solutions sur $\mathbb{R}$
  On suppose a, b réels. Il y a trois cas
  1) $a^2+4b > 0$. L'équation caractéristique admet deux solutions réelles distinctes. Les suites réelles solutions de $\mathcal{R}$ sont les suites u telles que $u_k = C\lambda_1^k + D\lambda_2^k$ où C, D sont des constantes réelles arbitraires.
  2) $a^2+4b = 0$. L'équation caractéristique a une racine double réelle $\lambda = a/2$. Les suites réelles soluttions de $\mathcal{R}$ sont les suites u telles que $u_k = (Ck+D)\lambda^k$ où C, D sont des constantes réelles arbitraires.
  3) $a^2+4b < 0$. L'équation caractéristique admet deux solutions complexes non réelles conjuguées $\lambda = re^{i\theta}$ et $\bar{\lambda} = re^{-i\theta}$ avec $r > 0$. Les suites $(r^k e^{ki\theta})_{k\geqslant 0}$ et $(r^k e^{-ik\theta})_{k\geqslant 0}$ forment une base sur $\mathbb{C}$ de l'ensemble des solutions complexes de $\mathcal{R}$. Il en est donc de même des suites $(r^k\cos(k\theta))_{k\geqslant 0}$ et $(r^k\sin(k\theta))_{k\geqslant 0}$. On en déduit que les suites réelles solutions de $\mathcal{R}$ sont les suites u telles que $u_k = r^k\left(C\cos(k\theta) + D\sin(k\theta)\right)$ où C, D sont des constantes réelles arbitraires.

Application : déterminant tridiagonal

Soient $a, b, c \in \mathbb{C}$ et $M_n \in M_n(\mathbb{C})$ la matrice tridiagonale telle que $m_{i,i} = a$ pour $1 \leqslant i \leqslant n$, $m_{i,i+1} = b,\ m_{i+1,i} = c$ pour $1 \leqslant i \leqslant n-1$. On se propose de calculer le déterminant $\Delta_n = \Delta_n(a,b,c)$ de M_n. En développant par rapport à la première colonne, on obtient de suite pour $n \geqslant 3$ la relation

$$\Delta_n = a\Delta_{n-1} - bc\Delta_{n-2}$$

Cette relation subsite pour $n = 2$ en posant $\Delta_0 = 1$. La suite $(\Delta)_n$ est donc définie par la relation de récurrence ci dessus et la condition initiale $\Delta_0 = 1,\ \Delta_1 = a$.

L'équation caractéristique est $X^2 - aX + bc = 0$.

- Plaçons nous dans le cas $a^2 - 4bc \neq 0$. Notons s et t les deux racines de $X^2 - aX + bc = 0$. Il existe des constantes A et B telles que, pour tout n, $\Delta_n = As^n + Bt^n$. Pour $n = 0$ puis $n = 1$ on obtient les conditions $A + B = 1$ et $sA + tB = a = s + t$. On obtient $A = \frac{s}{s-t}$ et $B = \frac{t}{t-s}$ d'où

$$\Delta_n(a,b,c) = \frac{s^{n+1} - t^{n+1}}{s-t} = s^n + s^{n-1}t + \cdots + st^{n-1} + t^n$$

- Dans le cas $a^2 - 4bc = 0$ on a $s = t = a/2$ et en procédant de la même manière on trouve

$$\Delta_n(a,b,c) = (n+1)(-1)^n\frac{a^n}{2^n}$$

  On peut aussi conclure par continuité. La fonction $\Delta_n : \mathbb{C}^3 \to \mathbb{C}$ est continue. Soit (a,b,c) tels que $a^2 + 4bc = 0$. On peut trouver une suite (a_k, b_k, c_k) convergeant vers (a,b,c) telle que $\forall k\ a_k^2 + 4b_kc_k \neq 0$. On choisit une fois pour toutes une suite δ_k telle que $\delta_k = a_k^2 - 4b_kc_k$ et on prend $t_k = (-a_k - \delta_k)/2$ et $s_k = (-a_k + \delta_k)/2$. Le couple (t_k, s_k) tend vers $(-a/2, -a/2)$ d'où la valeur de $\Delta_n(a,b,c)$.

Application
On se propose de déterminer les éléments propres de la matrice

$$M = \begin{pmatrix} 2 & -1 & 0 & \dots & \dots & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & -1 & 2 & -1 \\ 0 & \dots & \dots & 0 & -1 & 2 \end{pmatrix}$$

En anticipant sur le cours sur l'algèbre bilinéaire, on constate que cette matrice est symétrique réelle, donc diagonalisable sur $\mathbb{R}$. Nous n'aurons pas besoin de cette remarque pour faire le calcul.

Valeurs propres

Soit $\lambda \in \mathbb{C}$ et $\Delta = \det(M - \lambda I_n)$. On est dans le cas précédent avec $a = 2-\lambda$ et $b = c = -1$. On a $a^2 - 4bc = (2-\lambda)^2 - 4 = \lambda^2 - 4\lambda$. Supposons $\lambda \neq 0$ et $\lambda \neq 4$. Dans ce cas on a, en désignant par s, t les racines de l'équation $X^2 - (2-\lambda)X + 1 = 0$, $s + t = 2 - \lambda$, $st = 1$ et $(t-s)\Delta = t^{n+1} - s^{n+1}$.

Pour avoir $\Delta = 0$ il suffit donc de trouver (s, t, λ) tels que

$$\left\{\begin{array}{rcl} s & \neq & t \\ st & = & 1 \\ t^{n+1} & = & s^{n+1} \\ \lambda & = & 2 - (s+t) \end{array}\right. \Leftrightarrow \left\{\begin{array}{rcl} s & \neq & t \\ t^{2(n+1)} & = & 1 \\ s & = & 1/t \\ \lambda & = & 2 - (s+t) \end{array}\right.$$

Désignons par $\omega_k = \exp\left(\dfrac{ik\pi}{(n+1)}\right)$ les racines $2(n+1)$-ièmes de 1 ($0 \leqslant k \leqslant 2n-1$). La deuxième équation du dernier système implique l'existence d'un k tel que $t = \omega_k$. La première donne $s = 1/t = \overline{\omega}_k = \omega_{2(n+1)-k}$. La condition $s \neq t$ exclut les valeurs de k tels que $\omega_k^2 = 1$, c'est à dire $k = 0$ et $k = n+1$. Dans ces conditions, posons

$$\lambda_k = 2 - 2\cos\left(\frac{k\pi}{(n+1)}\right) = 4\sin^2\left(\frac{k\pi}{2(n+1)}\right)$$

La fonction $x \to \sin^2(x)$ est strictement croissante sur $[0, \pi/2]$ donc les nombres $\lambda_k,\ 1 \leqslant k \leqslant n$ sont tous distincts et différents de 0 et de 4. On obtient donc ainsi n valeurs propres toutes distinctes pour la matrice M carrée d'ordre n. Elles sont donc toutes simples et ce sont les seules. Il n'y a donc pas lieu d'étudier les cas laissés de coté.

Le lecteur remarquera que lorsque $n+2 \leqslant k \leqslant 2n+1$ on a $\lambda_k = \lambda_{2n+2-k}$ avec $1 \leqslant 2n+2-k \leqslant n$.

Vecteurs propres

Soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ un vecteur propre de la matrice M pour la valeur propre $\lambda = \lambda_k$. On a donc

$$\left\{\begin{array}{rcl} (2-\lambda)x_1 - x_2 & = & 0 \\ -x_1 + (2-\lambda)x_2 - x_3 & = & 0 \\ & & \\ -x_{n-1} + (2-\lambda_n)x_n & = & 0 \end{array}\right.$$

Introduisons la suite (u_j) vérifiant $u_0 = 0,\ u_1 = x_1$ et la relation de récurrence $-u_j + (2-\lambda)u_{j+1} - u_{j+2} = 0$. On aura $x_j = u_j$ pour $1 \leqslant j \leqslant n-1$ et, si $u_{n+1} = 0$ on aura aussi $x_n = u_n$. L'équation caractéristique est $X^2 - (2-\lambda)X + 1 = 0$ soit $X^2 - 2\cos(\alpha_k) + 1 = 0$ où $\alpha_k = k\pi/(n+1)$. Les racines sont $e^{i\alpha_k}$ et $e^{-i\alpha_k}$. Avec la condition initiale $u_0 = 0$ on trouve $u_j = A\sin(j\alpha)$ où A est une constante. Vu la valeur de α on a bien $u_{n+1} = 0$. Finalement, l'espace propre associé à la valeur propre λ_k est la droite vectorielle engendrée par le vecteur

$$U_k = \begin{pmatrix} \sin(\alpha_k) \\ \sin(2\alpha_k) \\ \vdots \\ \sin(n\alpha_k) \end{pmatrix}$$