Cours d'algèbre linéaire
MIAS1, premier semestre

Raphaël Danchin

Année 2003-2004

Table des matières

Structures usuelles

Au cours de vos deux années de MIAS, vous allez découvrir divers domaines des mathématiques qui, du moins en apparence, n'ont pas toujours de liens évidents entre eux. Certaines des lois qui régissent les objets considérés, pourtant, ont un caractère universel. Ainsi, des notions comme celles de lois internes, groupes, anneaux, corps, espaces vectoriels vont apparaître à maintes reprises. Leur caractère "unificateur" explique l'importance qu'on leur accorde. Le jeu consistant à les identifier est à la fois "rassurant" et pratique puisque l'on peut dès lors manipuler des objets mathématiques sans trop se soucier de leur nature profonde...

Dans cette section préliminaire, nous présentons brièvement quelques-unes des notions qui reviendront fréquemment dans ce cours d'algèbre du premier semestre.

Lois internes

Définition 0.0.1 *Soit E un ensemble non vide. Une loi interne T sur E est une application de $E \times E$ vers E qui à tout couple (a, b) de $E \times E$ associe un élément c de E noté $a\,T\,b$.*

Remarque : Pour éviter de faire un rapprochement trop hâtif et réducteur avec des lois internes bien connues (on en verra des exemples par la suite), on note T une loi interne "abstraite". Le symbole T est traditionnellement appelé "truc".

Pour avoir un intérêt pratique, une loi interne doit avoir des propriétés supplémentaires. Les plus courantes sont les suivantes :

- **Associativité :** On dit que la loi interne T est *associative* si

$$\forall (a, b, c) \in E \times E \times E,\ (a\,T\,b)\,T\,c = a\,T\,(b\,T\,c).$$

- **Commutativité :** On dit que la loi interne T est *commutative* si

$$\forall (a, b) \in E \times E,\ a\,T\,b = b\,T\,a.$$

- **Élément neutre :** On dit que $e \in E$ est élément neutre de T si

$$\forall a \in E,\ a\,T\,e = e\,T\,a = a.$$

  Remarque : L'élément neutre, lorsqu'il existe, est unique. En effet, si e et e' sont neutres pour T alors on a $e = e\,T\,e' = e'$.
- **Inverse :** Supposons que T ait un élément neutre e. On dit que l'élément a de E a pour inverse (ou symétrique) l'élément b de E si

$$a\,T\,b = b\,T\,a = e.$$

Proposition 0.0.2 *Si la loi T est associative et a un élément neutre e alors l'inverse, s'il existe, est unique.*

Preuve : Soit $a \in E$ admettant pour inverses b et b'. On a donc

$$e = a\,T\,b = a\,T\,b'.$$

Donc

$$b\,T\,(a\,T\,b) = b\,T\,(a\,T\,b').$$

Par associativité, on a donc

$$(b\,T\,a)\,T\,b = (b\,T\,a)\,T\,b'.$$

Mais $b\,T\,a = e$ donc la relation ci-dessus donne bien $b = b'$. ■

Exemples :

1. L'addition dans $\mathbb{N}$ est une loi interne associative et commutative, et a pour élément neutre 0. Dans $\mathbb{N}$, seul 0 a un inverse pour la loi $+$. En revanche dans $\mathbb{Z}$, tout élément n a un inverse : c'est $-n$.
2. La multiplication dans $\mathbb{N}$ ou $\mathbb{Z}$ est associative, commutative et a pour élément neutre 1.

Lorsque E possède *deux* lois internes T et $\times$, on peut définir la notion de *distributivité* .

Définition 0.0.3 *On dit que $\times$ est distributive par rapport à T si*

$$\forall (a,b,c) \in E \times E \times E,\ a \times (b\,T\,c) = (a \times b)\,T\,(a \times c).$$

Exemple : Dans $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ ou $\mathbb{R}$, la multiplication est distributive par rapport à l'addition.

Groupes et sous-groupes

Définition 0.0.4 *Soit G un ensemble non vide muni d'une loi interne T.*

On dit que (G,T) est **un groupe** *si T est associative, possède un élément neutre et si tout élément de G est inversible. Si de plus la loi T est commutative alors (G,T) est appelé* **groupe commutatif** *ou encore* **groupe abélien**.

Définition 0.0.5 *Soit (G,T) un groupe et H un sous-ensemble non vide de G. Si (H,T) est lui-même un groupe, on dit que (H,T) est* **un sous-groupe** *de (G,T).*

Remarque : Montrer que (H,T) est un sous-groupe de (G,T) revient à vérifier que $e \in H$, que H est stable par T (i.e pour tout $(a,b) \in H^2$ alors $a\,T\,b \in H$) et que tout élément de H a son inverse dans H.

Exemple : $(\mathbb{Z},+)$ est un groupe et l'ensemble des entiers relatifs pairs $2\mathbb{Z}$ muni de la loi $+$ est un sous-groupe de $(\mathbb{Z},+)$. En revanche, $(\mathbb{N},+)$ n'est pas un groupe (pourquoi ?)

Anneaux

Définition 0.0.6 *Soit A un ensemble non vide muni de deux lois internes T et $\times$. On dit que $(A,T,\times)$ est un* **anneau** *si les conditions suivantes sont vérifiées :*

i) (A,T) est un groupe commutatif,

ii) La loi $\times$ est associative et admet un élément neutre,

iii) La loi $\times$ est distributive par rapport à T.

Si de plus la loi $\times$ est commutative, on dit que $(A,T,\times)$ est un anneau commutatif.

Exemple : $(\mathbb{Z},+,\times)$ est un anneau commutatif.

Corps

Définition 0.0.7 *Soit $(K, T, \times)$ un anneau. On dit que $(K, T, \times)$ est un* **corps** *si tout élément de K distinct de l'élément neutre pour la loi T a un inverse pour la loi $\times$.*

Si de plus la loi $\times$ est commutative, on dit que $(A, T, \times)$ est un corps commutatif.

Exemple : $(\mathbb{Z}, +, \times)$ n'est pas un corps (car seuls 1 et -1 ont un inverse dans $\mathbb{Z}$ pour la multiplication). En revanche $(\mathbb{Q}, +, \times)$ et $(\mathbb{R}, +, \times)$ sont des corps.

Chapitre 1

Les nombres complexes

1.1 Construction des nombres complexes

1.1.1 Motivations

Au cours de votre scolarité, vous avez appris à manipuler différents types de "nombres". D'abord les entiers naturels $\mathbb{N}$, puis les entiers relatifs $\mathbb{Z}$, puis les nombres rationnels $\mathbb{Q}$ et enfin les réels $\mathbb{R}$. A chaque fois, l'introduction d'un nouvel ensemble de nombres était motivée par l'insuffisance du précédent pour la résolution de certains problèmes mathématiques.

Pour illustrer nos propos, cherchons à résoudre des équations du premier ou second degré dans ces différents ensembles. On constate que les équations du type $x + a = 0$ avec $a \in \mathbb{N}^*$ ne peuvent pas être résolues dans $\mathbb{N}$. En revanche, elles peuvent être résolues dans $\mathbb{Z}$, la solution étant évidemment $-a$. Mais $\mathbb{Z}$ est lui-même insuffisant dans la mesure où certaines équations du premier degré à coefficients entiers n'ont pas de solution dans $\mathbb{Z}$. C'est le cas de l'équation $2x+1 = 0$ par exemple. En revanche, cette équation a une solution dans $\mathbb{Q}$: la fraction $-\frac{1}{2}$. Plus généralement, on peut établir que toutes les équations du premier degré à coefficients rationnels ont une unique solution dans $\mathbb{Q}$.

L'insuffisance de $\mathbb{Q}$ est cependant manifeste lorsque l'on cherche à résoudre des équations du second degré (i.e $ax^2 + bx + c = 0$ avec $a \neq 0$) ou à déterminer des racines carrées de nombres positifs. Il est bien connu que $\sqrt{2}$ n'est pas dans $\mathbb{Q}$ (ce qui revient à dire que l'équation $x^2 - 2$ n'a pas de solution rationnelle). L'introduction de l'ensemble des réels $\mathbb{R}$ permet de calculer les racines carrées de nombres positifs et plus généralement de résoudre toutes les équations du second degré à discriminant positif. Mais celles qui ont un discriminant négatif n'ont pas de solution réelle. C'est le cas de l'équation $x^2 = -1$. On aimerait trouver un corps dans lequel cette équation ait une solution. On verra dans ce chapitre que le corps $\mathbb{C}$ répond à la question.[1]

La construction de $\mathbb{C}$ peut se faire à partir de $\mathbb{R}$ en munissant $\mathbb{R}^2$ de deux lois $\oplus$ et $*$ qui en feront un corps, l'ensemble réel pouvant s'identifier à l'ensemble des couples $(x, 0)$ avec $x \in \mathbb{R}$.

1.1.2 Définition d'une loi $\oplus$ dans $\mathbb{R}^2$

Pour tous couples (x, y) et (x', y') de $\mathbb{R}^2$, on définit un élément $(x, y) \oplus (x', y')$ de $\mathbb{R}^2$ par

$$(x, y) \oplus (x', y') \overset{\text{déf}}{=} (x + x', y + y').$$

En particulier,

$$(x, 0) \oplus (x', 0) = (x + x', 0).$$

Dans $\mathbb{R}$, l'élément $x + x'$ est bien la somme de x et de x'. Si l'on identifie $\mathbb{R}$ à l'ensemble des couples du type $(x, 0)$, la loi $\oplus$ est donc bien compatible avec l'addition sur $\mathbb{R}$.

[1] En fait, toute équation de degré quelconque a une solution dans $\mathbb{C}$. On dit que $\mathbb{C}$ est *algébriquement clos*.

1.1.3 Définition d'une loi $*$ dans $\mathbb{R}^2$

Pour tous couples (x, y) et (x', y') de $\mathbb{R}^2$, on définit un élément $(x, y) * (x', y')$ de $\mathbb{R}^2$ par

$$(x, y) * (x', y') \stackrel{\text{déf}}{=} (xx' - yy', xy' + x'y).$$

En particulier,

$$(x, 0) * (x', 0) = (xx', 0).$$

Dans $\mathbb{R}$, l'élément xx' est bien le produit de x et de x'. La multiplication ainsi définie est donc compatible avec celle de $\mathbb{R}$. De plus, $(0, 1) * (0, 1) = (-1, 0)$. Si l'on identifie $(-1, 0)$ au réel -1, on constate que -1 a une racine carrée au sens de cette loi $*$ de $\mathbb{R}^2$. C'était bien le but recherché.

1.1.4 Notations

Dans la suite de ce chapitre, le couple $(1, 0)$ est simplement noté 1 et l'on note i le couple $(0, 1)$. Le calcul précédent montre que $i^2 = -1$. De cette propriété "déconcertante" est issue le nom de *nombre imaginaire* donné à i. Plus généralement, le couple (x, y) est noté $x + iy$.

Définition 1.1.1 *On appelle ensemble des nombres complexes (noté $\mathbb{C}$) l'ensemble des couples (x, y) muni des lois $\oplus$ et $*$ définies précédemment, et notés $x+iy$. La loi $\oplus$ est alors simplement notée $+$, et la multiplication $*$ est notée $\cdot$ voire omise.*

Remarque : Nous laissons au lecteur le soin de vérifier que grâce à l'identification précédente, le calcul dans $\mathbb{C}$ est identique à celui dans $\mathbb{R}$ avec la convention $i^2 = -1$. Plus précisément, on a

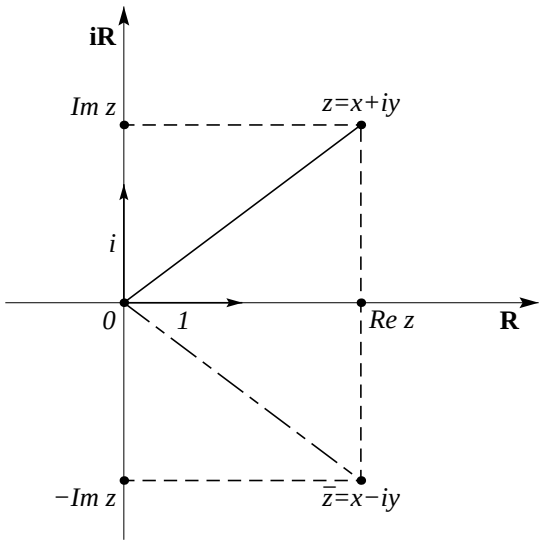
$$(x + iy) + (x' + iy') = (x + x') + i(y + y'),$$
$$(x + iy)(x' + iy') = (xx' - yy') + i(x'y + xy').$$

Définition 1.1.2 *Soit $z = x + iy$ un nombre complexe.*

- **Partie réelle :** *Le réel x est appelé partie réelle de z et noté $\mathcal{R}e\, z$. Si $\mathcal{R}e\, z = 0$, on dit que z est un* **nombre imaginaire pur**.
- **Partie imaginaire :** *Le réel y est appelé partie imaginaire de z et noté $\mathcal{I}m\, z$.*
- **Conjugué :** *Le nombre $\overline{z} \stackrel{\text{déf}}{=} x - iy$ est appelé conjugué de z.*
- **Affixe :** *A tout point M de $\mathbb{R}^2$ de coordonnées (x, y), on associe le nombre complexe $x+iy$, appelé affixe de M.*

Interprétation géométrique :

L'ensemble des éléments de $\mathbb{C}$ peut être identifié à un plan appelé **plan complexe**. L'axe des abscisses du plan complexe est souvent appelé **axe réel**, et noté $\mathbb{R}$. L'axe des ordonnées est appelé **axe imaginaire** et noté $i\mathbb{R}$. Pour tout nombre complexe z, le point M d'affixe z a pour abscisse $\mathcal{R}e\ z$, et pour partie imaginaire $\mathcal{I}m\ z$. Le point d'affixe $\overline{z}$ est le symétrique de M par rapport à l'axe des abscisses.



Proposition 1.1.3 *On a les propriétés élémentaires suivantes :*

- $\forall z \in \mathbb{C},\ \forall z' \in \mathbb{C},\ \mathcal{R}e\,(z+z') = \mathcal{R}e\,z + \mathcal{R}e\,z'$ *et* $\mathcal{I}m\,(z+z') = \mathcal{I}m\,z + \mathcal{I}m\,z'$,
- $\forall z \in \mathbb{C},\ \forall \lambda \in \mathbb{R},\ \mathcal{R}e\,(\lambda z) = \lambda \mathcal{R}e\,z$ *et* $\mathcal{I}m\,(\lambda z) = \lambda \mathcal{I}m\,z$,
- $\forall z \in \mathbb{C},\ \overline{\overline{z}} = z$,
- $\forall z \in \mathbb{C},\ \mathcal{R}e\,z = \dfrac{z+\overline{z}}{2}$ *et* $z \in \mathbb{R} \iff z = \overline{z}$,
- $\forall z \in \mathbb{C},\ \mathcal{I}m\,z = \dfrac{z-\overline{z}}{2i}$ *et* $z \in i\mathbb{R} \iff z = -\overline{z}$,
- $\forall z \in \mathbb{C},\ \forall z' \in \mathbb{C},\ \overline{z+z'} = \overline{z} + \overline{z}'$,
- $\forall z \in \mathbb{C},\ \forall z' \in \mathbb{C},\ \overline{zz'} = \overline{z}\overline{z}'$,
- $\forall z \in \mathbb{C},\ \forall \lambda \in \mathbb{R},\ \overline{\lambda z} = \lambda \overline{z}$.

Preuve : Elle est laissée au lecteur à titre d'exercice. ■

Attention : En général, $\mathcal{I}m\,(zz') \neq \mathcal{I}m\,z\,\mathcal{I}m\,z'$ et $\mathcal{R}e\,(zz') \neq \mathcal{R}e\,z\,\mathcal{R}e\,z'$.

Exercice : Trouver des couples (z, z') pour lesquels les égalités ci-dessus ne sont pas vérifiées.

1.2 Propriétés de $\mathbb{C}$

1.2.1 Propriétés algébriques

Théorème 1.2.1 $(\mathbb{C}, +, *)$ *est un corps commutatif. De plus* 0 *est l'élément neutre pour l'addition,* 1 *est l'élément neutre pour la multiplication, et l'inverse d'un nombre complexe non nul* $x+iy$ *pour la multiplication est*

$$\boxed{(x+iy)^{-1} = \frac{x-iy}{x^2+y^2}.}$$

Preuve : On vérifie successivement que $(\mathbb{C}, +)$ est un groupe commutatif d'élément neutre 0, que $(\mathbb{C}, +, \cdot)$ est un anneau commutatif d'élément neutre 1, puis enfin que tout élément non nul de $\mathbb{C}$ est inversible pour la multiplication. Les détails sont laissés au lecteur.

Donnons juste la preuve de l'associativité de la multiplication (qui est la partie la plus calculatoire). Soit donc $z = x+iy$, $z' = x'+iy'$ et $z'' = x''+iy''$ trois nombres complexes. Il s'agit de montrer que $(zz')z'' = z(z'z'')$. Calculons :

$$\begin{aligned}
(zz')z'' &= \Big[(x+iy)(x'+iy')\Big](x''+iy''),\\
&= \Big[(xx'-yy') + i(xy'+x'y)\Big](x''+iy''),\\
&= (xx')x'' - (yy')x'' - (xy')y'' - (x'y)y'' + i[(xy')x'' + (x'y)x'' + (xx')y'' - (yy')y''],\\
&= x(x'x'') - y(y'x'') - x(y'y'') - x'(yy'') + i[x(y'x'') + x'(yx'') + x(x'y'') - y(y'y'')],\\
&= x(x'x'') - x(y'y'') - y(x'y'') - y(x''y') + i[x(x'y'') + x(x''y') + y(x'x'') - y(y'y'')],\\
&= (x+iy)\Big[(x'x''-y'y'') + i(x'y''+x''y')\Big],\\
&= (x+iy)\Big[(x'+iy')(x''+iy'')\Big],\\
&= z(z'z'').
\end{aligned}$$

■

Comme dans tout anneau commutatif (et *a fortiori* dans tout corps commutatif), on dispose dans $\mathbb{C}$ d'**identités remarquables** . Les plus simples sont :

$$\boxed{\begin{aligned}(z+z')^2 &= z^2 + 2zz' + z'^2,\\ (z+z')^3 &= z^3 + 3z'z^2 + 3z'^2z + z'^3.\end{aligned}}$$

Plus généralement, la *formule du binôme de Newton* est valable :

$$\boxed{(z+z')^n = \sum_{k=0}^{n} C_n^k \, z^k \, z'^{n-k}}$$

avec[2] $\boxed{C_n^k = \dfrac{n!}{k!(n-k)!}}$.

La formule du binôme se montre par récurrence sur n en utilisant la *formule du triangle de Pascal :*

$$C_{n+1}^{k+1} = C_n^k + C_n^{k+1}.$$

Une autre identité remarquable bien connue

$$\boxed{z'^2 - z^2 = (z'-z)(z'+z)}$$

peut se généraliser en

$$\boxed{z'^n - z^n = (z'-z)\sum_{k=0}^{n-1} z'^k z^{n-1-k}}$$

En particulier, en choisissant $z' = 1$, en appliquant la formule ci-dessus au rang $n+1$ puis en divisant par $1-z$, on retrouve la formule donnant la somme des n premiers termes d'une série géométrique :

$$\boxed{\forall z \neq 1,\ \sum_{k=0}^{n} z^k = \frac{1-z^{n+1}}{1-z}.}$$

1.2.2 Représentation polaire

a) Le module

Le *module* d'un nombre complexe est un prolongement naturel de la notion de valeur absolue d'un nombre réel. On le définit ainsi :

Définition 1.2.2 *Soit $z = x + iy$ un nombre complexe. On appelle module de z le réel positif ainsi défini :*

$$|z| \overset{\text{déf}}{=} \sqrt{x^2+y^2}.$$

Proposition 1.2.3 *Considérons deux nombres complexes z et z'. On a les propriétés suivantes :*

i) $|z| \geq 0$ *et* $|z| = 0$ *si et seulement si* $z = 0$.

ii) $|z| = |\overline{z}|$.

iii) $|z|^2 = z\overline{z} = \overline{z}z$. *Autrement dit, pour tout couple de réels (x,y), on a la factorisation*

$$\boxed{x^2+y^2 = (x+iy)(x-iy).}$$

iv) Si $z \neq 0$, on a

$$\boxed{z^{-1} = \frac{\overline{z}}{|z|^2}.}$$

En conséquence, $(\overline{z})^{-1} = \overline{z^{-1}}$, *et* $\overline{z} = z^{-1}$ *si et seulement si* $|z| = 1$.

[2] Les coefficients binomiaux C_n^k sont notés $\binom{n}{k}$ par certains auteurs et dans les pays anglo-saxons.

v) *On a* $|zz'| = |z||z'|$. *Se plus, si* $z \neq 0$ *alors* $|z^{-1}| = 1/|z|$.

vi) $|\mathcal{R}e\, z| \leq |z|$ *et* $|\mathcal{I}m\, z| \leq |z|$.

vii) $|z + z'|^2 = |z|^2 + |z'|^2 + 2\mathcal{R}e\,(z\overline{z}')$.

viii) *Inégalités triangulaires :* $\boxed{\Big||z| - |z'|\Big| \leq |z + z'| \leq |z| + |z'|.}$

Preuve :

- Pour *i)*, on utilise le fait que la somme de deux réels positifs est un réel positif, et qu'elle est nulle si et seulement si les deux réels sont nuls.
- Pour *ii)* et *iii)*, il suffit de revenir à la définition du module.
- Soit $z \neq 0$. Alors $\overline{z}/|z|^2$ est bien l'inverse de z. En effet, d'après la propriété *iii)*, on a

$$z\frac{\overline{z}}{|z|^2} = \frac{z\overline{z}}{|z|^2} = \frac{|z|^2}{|z|^2} = 1.$$

  On en déduit ensuite que

$$\overline{z^{-1}} = \overline{\Big(\frac{\overline{z}}{|z|^2}\Big)} = \frac{\overline{\overline{z}}}{|z|^2} = \frac{\overline{\overline{z}}}{|\overline{z}|^2} = (\overline{z})^{-1}.$$

- Pour prouver *v)*, on écrit que

$$|zz'|^2 = zz'\,\overline{zz'} = z\overline{z}\,z'\overline{z}' = |z|^2|z'|^2.$$

  Si $z \neq 0$, le choix de $z' = z^{-1}$ donne bien $|z^{-1}| = 1/|z|$.
- La propriété *vi)* est triviale.
- Pour prouver *vii)*, on calcule en tenant compte de *iii)* :

$$|z + z'|^2 = (z + z')(\overline{z} + \overline{z}') = |z|^2 + |z'|^2 + z\overline{z}' + z'\overline{z} = |z|^2 + |z'|^2 + (z\overline{z}' + \overline{z\overline{z}'}).$$

  Or, d'après la proposition 1.1.3, le dernier terme est justement égal à $2\mathcal{R}e\,(z\overline{z}')$.
  Pour prouver *viii)*, on utilise successivement *vi)*, *v)* et *ii)*. On trouve :

$$\mathcal{R}e\,(z\overline{z}') \leq |z\overline{z}'| = |z||\overline{z}'| = |z||z'|.$$

  Ainsi, d'après *vii)*,

$$|z + z'|^2 \leq |z|^2 + |z'|^2 + 2|z||z'|.$$

  En prenant la racine carrée positive des deux membres, on obtient l'inégalité de droite. L'inégalité triangulaire de gauche se montre en appliquant celle de droite à z et $z + z'$ puis à z' et $z + z'$.

■

b) L'argument

Commençons par rappeler la définition de congruence.

Définition 1.2.4 *Considérons trois réels* a, b *et* c. *On dit que* a *est congru à* b *modulo* c *s'il existe* $k \in \mathbb{Z}$ *tel que* $a = b + kc$. *On note* $a \equiv b\ [c]$.

Proposition 1.2.5 *Soit* z *un nombre complexe* **non nul**. *Il existe un unique réel* θ *de* $[0, 2\pi[$ *tel que*

$$\frac{z}{|z|} = \cos\theta + i\sin\theta. \tag{1.1}$$

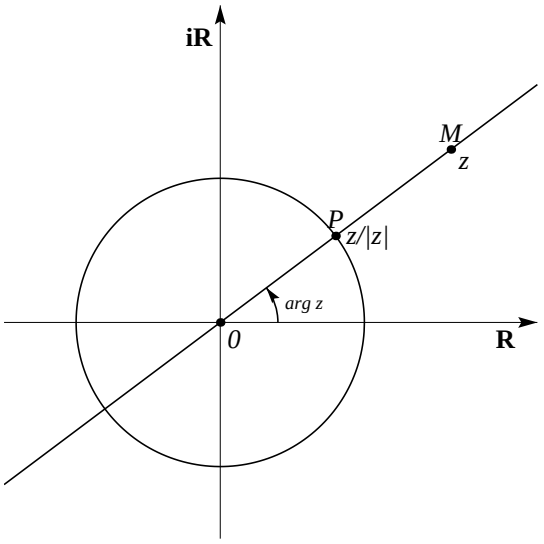
Ce réel est appelé **argument principal** *de z, et noté*[3] $\arg z$.

De plus, l'ensemble des réels θ vérifiant (1.1) *est l'ensemble des réels congrus à* $\arg z$ *modulo* 2π*, c'est-à-dire l'ensemble des réels du type* $\arg z + 2k\pi$ *avec* $k \in \mathbb{Z}$*. Tous ces réels sont appelés* **arguments** *de z.*

Preuve : Il suffit de remarquer que $\frac{z}{|z|}$ est de module 1 et peut donc s'écrire $x + iy$ avec $x^2 + y^2 = 1$. Il existe donc un unique réel θ de $[0, 2\pi[$ tel que $x = \cos\theta$ et $y = \sin\theta$. Les autres réels satisfaisant cette relation lui sont congrus modulo 2π. ■

c) Interprétation géométrique

Soit $z \in \mathbb{C}$ et M d'affixe z. Le module de z est égal à la norme du vecteur $\overrightarrow{OM}$ c'est-à-dire à la distance de O à M. Si $z \neq 0$, l'argument de z est une mesure (en radians) de l'angle orienté formé par le vecteur unité dirigeant l'axe des réels dans le sens positif et le vecteur $\overrightarrow{OM}$. En d'autres termes, le point P d'affixe $z/|z|$ est le point d'intersection entre le cercle unité et la demi-droite $[0M)$.



Remarque : On peut maintenant donner une interprétation géométrique des inégalités triangulaires (qui est à l'origine de leur appellation). L'inégalité de droite stipule que la longueur d'un côté d'un triangle est inférieure à la somme des longueurs des deux autres côtés, et celle de gauche, que la longueur d'un côté est toujours supérieure à la différence des longueurs des deux autres côtés.

d) Forme trigonométrique

Définition 1.2.6 *Pour tout réel θ, on pose*

$$e^{i\theta} \overset{\text{déf}}{=} \cos\theta + i\sin\theta$$

Ainsi tout nombre complexe z non nul de module r et d'argument θ peut s'écrire[4] $z = re^{i\theta}$.

On dit que $re^{i\theta}$ est la **forme trigonométrique** *de z. Pour $z \neq 0$, cette écriture est unique à congruence modulo 2π près pour θ. C'est-à-dire que*

$$\left(re^{i\theta} = r'e^{i\theta'}\right) \iff \left(r = r' \quad et \quad \theta \equiv \theta'\ [2\pi]\right).$$

En particulier, on peut toujours choisir pour θ l'argument principal de z.

Remarque : La notation $e^{i\theta}$ est pour l'instant purement formelle. Elle sera justifiée mathématiquement plus tard dans le chapitre sur les séries entières du cours d'analyse.

[3] Les définitions de l'argument principal diffèrent suivant les ouvrages. Une autre définition très répandue consiste à choisir pour argument principal l'unique réel de $]-\pi, \pi]$ vérifiant (1.1).

[4] Le nombre 0 peut aussi s'écrire $re^{i\theta}$: on prend $r = 0$ et θ arbitraire.

Quelques formes trigonométriques à connaître :

- $e^{i0} = 1$ et plus généralement, $ae^{i0} = a$ pour tout $a \in \mathbb{R}^+$.
- $e^{i\Pi} = -1$ et plus généralement, $|a|e^{i\Pi} = a$ pour tout $a \in \mathbb{R}^-$.
- $e^{i\frac{\pi}{2}} = i$ et $e^{3i\frac{\pi}{2}} = -i$.
- $e^{i\frac{\pi}{4}} = \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}$ et $e^{3i\frac{\pi}{4}} = -\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}$.
- $e^{i\frac{\pi}{6}} = \frac{\sqrt{3}}{2} + \frac{i}{2}$ et $e^{i\frac{\pi}{3}} = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$.
- Pour tout $k \in \mathbb{Z}$ et $\theta \in \mathbb{R}$, $e^{i(\theta+2k\pi)} = e^{i\theta}$.
- Pour tout $\theta \in \mathbb{R}$, $e^{i(\theta+\pi)} = -e^{i\theta}$ et $e^{i(\theta+\frac{\pi}{2})} = ie^{i\theta}$.

1.2.3 Formules de trigonométrie

Nous nous bornerons à rappeler deux formules trigonométriques d'importance capitale pour la suite du cours :

$$\forall(\theta,\theta') \in \mathbb{R}^2, \quad \cos(\theta+\theta') = \cos\theta\cos\theta' - \sin\theta\sin\theta', \tag{1.2}$$

$$\sin(\theta+\theta') = \sin\theta\cos\theta' + \cos\theta\sin\theta'. \tag{1.3}$$

Noter que la deuxième se déduit de la première en changeant θ' en $\theta' + \pi/2$.

Donnons une première application de ces formules, (qui est fort utilisée en physique) :

Proposition 1.2.7 *Soit (A, B) un couple de réel. Alors il existe un réel φ tel que*

$$\forall x \in \mathbb{R},\ A\cos x + B\sin x = \sqrt{A^2+B^2}\,\cos(x-\varphi). \tag{1.4}$$

Si de plus $(A, B) \neq (0, 0)$, on peut choisir pour φ l'argument principal du nombre complexe $A + iB$, c'est-à-dire l'unique élément φ de $[0, 2\pi[$ tel que

$$\cos\varphi = \frac{A}{\sqrt{A^2+B^2}}, \qquad \sin\varphi = \frac{B}{\sqrt{A^2+B^2}}.$$

Preuve : Limitons nous au cas $(A, B) \neq (0, 0)$. D'après la formule (1.2), on a

$$\cos(x-\varphi) = \cos x\cos\varphi + \sin x\sin\varphi.$$

On en déduit que la formule (1.4) est vérifiée si et seulement si

$$A = \sqrt{A^2+B^2}\,\cos\varphi \quad \text{et} \quad B = \sqrt{A^2+B^2}\,\sin\varphi.$$

En remarquant que

$$A + iB = \sqrt{A^2+B^2}\Big(\frac{A}{\sqrt{A^2+B^2}} + i\frac{B}{\sqrt{A^2+B^2}}\Big),$$

on conclut que l'on peut prendre pour φ l'argument principal de $A + iB$. ∎

Les formules trigonométriques (1.2) et (1.3) vont également nous permettre de montrer des propriétés algébriques de l'argument :

Proposition 1.2.8 *1. Si z et z' sont deux nombres complexes non nuls alors*

$$\boxed{\arg(zz') \equiv \arg z + \arg z'\,[2\pi].}$$

2. Si z est un nombre complexe non nul alors

$$\boxed{\arg z^{-1} \equiv -\arg z\,[2\pi].}$$

Preuve : Clairement, la première égalité appliquée avec $z' = z^{-1}$ donne la deuxième égalité. Prouvons donc la première égalité. Soit (z, z') un couple de nombres complexes non nuls. Notons $\theta \stackrel{\text{déf}}{=} \arg z$ et $\theta' \stackrel{\text{déf}}{=} \arg z'$. On a

$$z = |z|(\cos\theta + i\sin\theta) \quad \text{et} \quad z' = |z'|(\cos\theta' + i\sin\theta').$$

Donc, en appliquant (1.2) et (1.3),

$$\begin{aligned} zz' &= |z||z'|\Big[\Big(\cos\theta\cos\theta' - \sin\theta\sin\theta'\Big) + i\Big(\cos\theta\sin\theta' + \sin\theta\cos\theta'\Big)\Big], \\ &= |zz'|\Big[\cos(\theta+\theta') + i\sin(\theta+\theta'\Big]. \end{aligned}$$

■

Proposition 1.2.9 *Soit (θ, θ') un couple de réels, et (r, r') un couple de réels positifs. On a*

$$\Big(re^{i\theta}\Big)\Big(r'e^{i\theta'}\Big) = rr'e^{i(\theta+\theta')}. \tag{1.5}$$

De plus, si $r \neq 0$, on a pour tout $n \in \mathbb{Z}$,

$$(re^{i\theta})^n = r^n e^{in\theta}. \tag{1.6}$$

Preuve : Notons $z = re^{i\theta}$ et $z' = r'e^{i\theta'}$. On sait déjà que le module de zz' est $|z||z'|$. La proposition 1.2.8 montre que $\arg zz \equiv \arg z + \arg z'\,[2\pi]$. Comme bien sûr $r = |z|$, $r' = |z'|$, $\theta \equiv \arg z\,[2\pi]$ et $\theta' \equiv \arg z'\,[2\pi]$, on a bien $zz' = rr'e^{i(\theta+\theta')}$.[5].

L'égalité (1.6) dans le cas $n = 0$ ou $n = 1$ est immédiate. Dans le cas $n = 2$, elle découle de (1.5) avec $z' = z$. Le cas $n \in \mathbb{N}$ quelconque suit par récurrence (**exercice** : faire la récurrence).

Le cas $n < 0$ découle du point 2 de la proposition 1.2.8, appliqué à z^{-n} et au fait que le module de l'inverse est l'inverse du module. ■

Remarque 1.2.10 *En particulier, $e^{i(\theta+\theta')} = e^{i\theta}e^{i\theta'}$ ce qui montre que l'exponentielle d'un nombre imaginaire pur a les mêmes propriétés multiplicatives que l'exponentielle réelle.*

Interprétation géométrique : Le nombre $e^{i\theta}$ (resp. $e^{i\theta'}$) a pour affixe le point obtenu par rotation d'angle θ (resp. θ') du point $(1, 0)$. Calculer $e^{i\theta}e^{i\theta'}$ revient à faire subir à l'affixe de $e^{i\theta'}$ une rotation d'angle θ. On s'attend donc à trouver l'image de $(1, 0)$ par la composée des rotations d'angle θ et θ'. La remarque ci-dessus assure que la composée de ces deux rotations est bien la rotation d'angle $\theta + \theta'$.

Remarque 1.2.11 *En prenant $r = 1$ dans la formule (1.6), on trouve*

$$\boxed{\forall n \in \mathbb{Z},\ \forall \theta \in \mathbb{R},\ (\cos\theta + i\sin\theta)^n = \cos n\theta\ +\ i\ \sin n\theta.}$$

Cette identité <u>à connaître</u> est appelée **formule de Moivre**.

Proposition 1.2.12 (Formules d'Euler) *Pour tout réel θ, on a*

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

[5] Le cas où l'un des deux nombres z et z' est nul ne fait pas exception puisqu'alors $zz' = 0$

Preuve : Il suffit de faire la demi-somme et la demi-différence des expressions suivantes :

$$e^{i\theta} = \cos\theta + i\sin\theta,$$
$$e^{-i\theta} = \cos\theta - i\sin\theta.$$

■

Remarques :

1. Les formules d'Euler permettent de linéariser des expressions du type $\cos^k x \sin^\ell x$, c'est-à-dire de les transformer en sommes de cos et de sin.
2. Si z est un nombre complexe non nul de forme trigonométrique $re^{i\theta}$, on a les formules

$$\mathcal{R}e\, z = r\cos\theta, \quad \mathcal{I}m\, z = r\sin\theta \quad \text{et} \quad \overline{z} = re^{-i\theta}.$$

1.2.4 Racines n-ième de l'unité

Définition 1.2.13 *Soit $n \in \mathbb{N}^*$ et $z \in \mathbb{C}$. On dit que z est racine n-ième de l'unité si $z^n = 1$. Dans le cas $n = 2$, on parle de racine carrée de l'unité.*

Proposition 1.2.14 *Pour $n \in \mathbb{N}^*$ fixé, il y a exactement n racines de l'unité deux à deux distinctes. Il s'agit des nombres complexes*

$$\boxed{z_k \overset{\text{déf}}{=} e^{\frac{2ik\pi}{n}} = \cos\Big(\frac{2k\pi}{n}\Big) + i\sin\Big(\frac{2k\pi}{n}\Big) \quad \textit{avec} \quad k \in \{0, \cdots, n-1\}.}$$

Preuve : D'après la formule de Moivre, $z_k^n = \cos(2ik\pi) + i\sin(2ik\pi) = 1$. Donc les n nombres complexes z_k sont bien racines n-ième de l'unité. Vérifions qu'il n'y en a pas d'autre. Soit z une racine n-ième de l'unité que l'on écrit sous sa forme trigonométrique $z = re^{i\theta}$. Par hypothèse, on a $r^n e^{in\theta} = 1e^{i0}$. Donc

$$r^n = 1 \quad \text{et} \quad n\theta \equiv 0\,[2\pi].$$

Comme r est un réel positif, on doit avoir $r = 1$. La deuxième relation montre que θ est de la forme $\theta = 2\pi\ell/n$ avec $\ell \in \mathbb{Z}$. Écrivons la division euclidienne de ℓ par n :

$$\ell = qn + k \quad \text{avec} \quad q \in \mathbb{Z} \quad \text{et} \quad k \in \{0, \cdots, n-1\}.$$

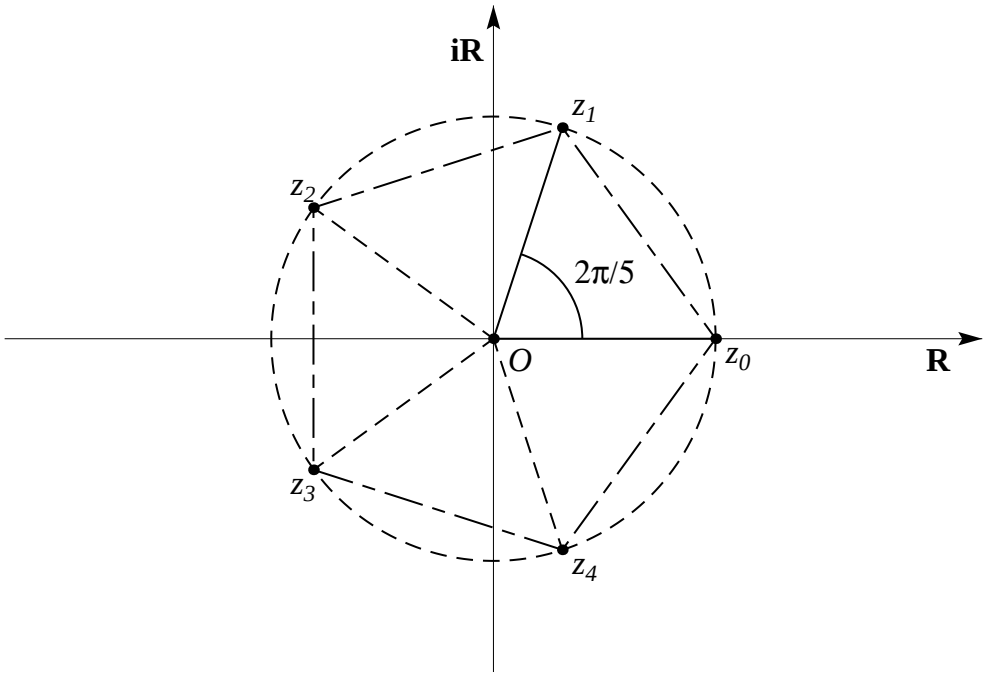
Alors $\theta = \frac{2k\pi}{n} + 2\pi q$. Donc l'argument principal de z est $\frac{2k\pi}{n}$. ■

Exemples :

- $n = 2$: les racines carrées de l'unité sont -1 et 1.
- $n = 3$: les racines cubiques de l'unité sont 1, $\jmath$ et $\overline{\jmath}$ où $\jmath \overset{\text{déf}}{=} -\frac{1}{2} + i\frac{\sqrt{3}}{2}$.
- $n = 4$: les racines 4-ième de l'unité sont 1, i, -1 et $-i$.

Interprétation géométrique : Les points ayant pour affixes les racines n-ième de l'unité sont les points du cercle unité formant l'angle $2k\pi/n$ avec l'axe des abscisses.

Exemple : représentation des racines 5ème de l'unité dans le plan complexe



1.3 Résolution d'équations du second degré

Par construction de $\mathbb{C}$, l'équation $z^2 = -1$ a au moins une solution complexe : i, et $-i$ est également solution. Dans cette section, on montre que toutes les équations du second degré :

$$(E) \qquad az^2 + bz + c = 0$$

à coefficients $a \neq 0$, b et c réels ou complexes ont une ou deux solutions. On donne de plus des formules permettant de les calculer.

1.3.1 Racine carrée d'un nombre complexe

Proposition 1.3.1 *Tout nombre complexe α non nul a deux racines carrées distinctes et opposées. Plus précisément, si $\alpha = re^{i\theta}$ alors les racines carrées de α sont $\pm\sqrt{r}e^{i\frac{\theta}{2}}$.*

L'unique racine carrée de 0 est 0.

Preuve : Le cas $\alpha = 0$ est immédiat (utiliser le module).

Soit donc $\alpha \neq 0$ et z une racine carrée de α. Écrivons $\alpha = re^{i\theta}$ et $z = \rho e^{i\varphi}$ avec $\theta = \arg\alpha$ et $\varphi = \arg z$. Comme $z^2 = \rho^2 e^{2i\varphi}$, l'équation $z^2 = \alpha$ est équivalente à

$$\rho^2 = r \quad \text{et} \quad 2\varphi \equiv \theta\,[2\pi].$$

c'est-à-dire

$$\rho = \sqrt{r} \quad \text{et} \quad \varphi \equiv \theta/2\,[\pi].$$

En se restreignant aux valeurs de φ comprises entre 0 et 2π, on trouve $\varphi = \frac{\theta}{2}$ ou $\varphi = \frac{\theta}{2} + \pi$.

Un calcul direct montre que $\sqrt{r}e^{i\frac{\theta}{2}}$ et $-\sqrt{r}e^{i\frac{\theta}{2}}$ sont racines carrées de z. ■

Si la forme trigonométrique de α est connue, le calcul des racines carrées de α est immédiat. On peut se demander si la connaissance de la forme trigonométrique est nécessaire au calcul des racines carrées. Il n'en est rien : dans la proposition suivante, on établit une formule donnant la racine carrée d'un nombre complexe $\alpha = a + ib$ arbitraire en fonction de a et b.

Proposition 1.3.2 *Soit $\alpha = a + ib$ un nombre complexe arbitraire.*

1. *Si $b = 0$ et $a \geq 0$ (i.e $\alpha \in \mathbb{R}^+$) : les racines carrées de α sont $\sqrt{a}$ et $-\sqrt{a}$.*
2. *Si $b = 0$ et $a \leq 0$ (i.e $\alpha \in \mathbb{R}^-$) : les racines carrées de α sont $i\sqrt{|a|}$ et $-i\sqrt{|a|}$.*
3. *Si $b > 0$: les racines carrées de α sont z et $-z$ avec*
$$z \overset{\text{déf}}{=} \frac{\sqrt{2}}{2}\left(\sqrt{a + \sqrt{a^2+b^2}} + i\sqrt{-a + \sqrt{a^2+b^2}}\right).$$
4. *Si $b < 0$: les racines carrées de α sont z et $-z$ avec*
$$z \overset{\text{déf}}{=} \frac{\sqrt{2}}{2}\left(\sqrt{a + \sqrt{a^2+b^2}} - i\sqrt{-a + \sqrt{a^2+b^2}}\right).$$

Preuve : On cherche $z = x + iy$ tel que

$$z^2 = a + ib. \tag{1.7}$$

En calculant z^2 puis en identifiant parties réelles et parties imaginaires, on trouve que (1.7) est vérifiée si et seulement si
$$\begin{cases} x^2 - y^2 = a, \\ 2xy = b. \end{cases}$$
On peut résoudre ce système directement, mais il est plus rapide d'exploiter le fait que $|z|^2 = |\alpha|$, ce qui donne la relation supplémentaire $x^2 + y^2 = \sqrt{a^2+b^2}$. Le couple (x^2, y^2) doit donc vérifier le système de deux équations à deux inconnues suivant :
$$\begin{cases} x^2 - y^2 = a \\ x^2 + y^2 = \sqrt{a^2+b^2}. \end{cases}$$
On en déduit que
$$x^2 = \frac{1}{2}\left(a + \sqrt{a^2+b^2}\right) \quad \text{et} \quad y^2 = \frac{1}{2}\left(-a + \sqrt{a^2+b^2}\right). \tag{1.8}$$
Remarquons que les membres de droite sont toujours positifs, donc il existe bien des couples (x, y) vérifiant les égalités ci-dessus.
Dans le cas $b = 0$ et $a \geq 0$, α est en fait un réel positif, et on trouve $x = \pm\sqrt{a}$ et $y = 0$. Les racines carrées sont donc les racines carrées réelles habituelles.
Dans le cas $b = 0$ et $a < 0$ (i.e α réel négatif), la formule (1.8) montre que $x = 0$ et $y = \pm\sqrt{|a|}$.
Si $b \neq 0$, les membres de droite de (1.8) sont strictement positifs. Il y a donc en général quatre couples (x, y) solutions. La condition supplémentaire $2xy = b$ va permettre de sélectionner les deux couples qui vont donner une racine carrée de α.
En effet, si $b > 0$, alors $2xy = b$ implique que $xy > 0$ donc x et y doivent être de même signe. Cela donne bien le cas 3.
Si au contraire $b < 0$, alors x et y doivent être de signe opposé, ce qui donne le cas 4.
Il reste à vérifier que les nombres trouvés sont bien des racines carrées de z. Cela peut se faire par calcul direct, ou en remarquant que l'on sait déjà que z non nul admet exactement deux racines carrées. ■

Remarque : Il est inutile de connaître ces formules par cœur. Mais il est important de se souvenir de la démarche qui a permis de les obtenir.

1.3.2 Résolution d'équations du second degré dans le cas général

On cherche à résoudre (E) dans le cas où a, b et c sont des *nombres complexes* et $a \neq 0$.

Proposition 1.3.3 *Notons* $\Delta \overset{\text{déf}}{=} b^2 - 4ac$ *le discriminant complexe de* (E).

- *Si* $\Delta = 0$ *alors* (E) *a une unique solution :* $z = -b/2a$.
- *Si* $\Delta \neq 0$ *alors* (E) *a deux solutions distinctes* z_1 *et* z_2 *qui sont données par les formules*

$$z_1 = -\frac{b}{2a} + \frac{z_0}{2a} \quad et \quad z_2 = -\frac{b}{2a} - \frac{z_0}{2a}$$

avec z_0 *racine carrée de* Δ.

Preuve : Il suffit de factoriser le membre de gauche de (E) :

$$\begin{aligned} az^2 + bz + c &= a\Big(z^2 + \tfrac{b}{a}z + \tfrac{c}{a}\Big), \\ &= a\Bigg(\Big(z + \tfrac{b}{2a}\Big)^2 - \tfrac{\Delta}{(2a)^2}\Bigg). \end{aligned}$$

- Dans le cas $\Delta = 0$, cette factorisation montre que z est solution si et seulement si $z + \frac{b}{2a} = 0$.
- Dans le cas $\Delta \neq 0$, on poursuit la factorisation compte tenu de $z_0^2 = \Delta$. Il vient finalement

$$az^2 + bz + c = a\Big(z + \frac{b}{2a} - \frac{z_0}{2a}\Big)\Big(z + \frac{b}{2a} + \frac{z_0}{2a}\Big).$$

Il est maintenant clair que z est solution si et seulement si

$$z = -\frac{b}{2a} + \frac{z_0}{2a} \quad \text{ou} \quad z = -\frac{b}{2a} - \frac{z_0}{2a}.$$

■

Remarque : La formule donnant les solutions d'une équation du second degré dans le cas complexe est donc exactement la même que dans le cas réel. De plus, il n'y a pas à se préoccuper du signe du discriminant (d'ailleurs il n'a pas de signe puisqu'il est complexe !) On voit donc que toutes les équations du second degré ont une ou deux solutions dans $\mathbb{C}$, et que leur calcul est immédiat une fois connue une racine carrée du discriminant.

Chapitre 2

Systèmes linéaires

Au lycée, vous avez appris à résoudre des systèmes de 2, 3 voire 4 équations à 2, 3 ou 4 inconnues. Ce chapitre est consacré à la théorie des systèmes linéaires comportant un nombre arbitraire d'équations et d'inconnues. Il s'agit notamment de présenter une méthode générale de résolution de tels systèmes.

2.1 Quelques exemples élémentaires

Donnons d'abord quelques exemples de résolution de systèmes de deux équations à deux inconnues.

1. Résolution de

$$(S_1) \qquad \begin{cases} 2x + 3y = 8 & (L_1) \\ x - y = -1 & (L_2) \end{cases}$$

Il s'agit de déterminer l'ensemble des couples de réels (x, y) qui satisfont les deux lignes du système (S_1).

Pour ce faire, on peut procéder ainsi : on retranche $2(L_2)$ à (L_1), et on obtient le système

$$(S_1') \qquad \begin{cases} 5y = 10 & (L_1') \\ x - y = -1 & (L_2') \end{cases}$$

dont l'ensemble des solutions est le même que celui de (S_1).

Il est clair que la ligne (L_1') est équivalente à $y = 2$, et en reportant dans (L_2'), on obtient $x = 1$.

En conclusion, (S_1) a pour unique solution le couple $(1, 2)$.

2. Résolution de

$$(S_2) \qquad \begin{cases} 2x - 2y = 8 & (L_1) \\ x - y = -1 & (L_2) \end{cases}$$

Cette fois-ci, si l'on retranche $2(L_2)$ à (L_1), on obtient le système

$$(S_2') \qquad \begin{cases} 0 = 10 & (L_1') \\ x - y = -1 & (L_2') \end{cases}$$

La première ligne ne peut jamais être réalisée, et l'on conclut que (S_2) n'a pas de solution.

3. Résolution de

$$(S_3) \qquad \begin{cases} 2x - 2y = -2 & (L_1) \\ x - y = -1 & (L_2) \end{cases}$$

On remarque que la première ligne est égale à deux fois la seconde. Par conséquent, le système (S_3) est équivalent à

$$x - y = -1.$$

L'ensemble des couples (x, y) vérifiant cette relation est

$$\{(y-1, y) \mid y \in \mathbb{R}\}.$$

Le système (S_3) a donc une infinité de solutions.

Conclusion : D'après ces exemples, pour les systèmes 2×2, au moins trois cas de figure peuvent se présenter : ou bien le système a une seule solution, ou bien il n'en a pas, ou bien il en a une infinité. Nous allons voir dans ce chapitre que même pour les systèmes linéaires généraux, il n'y a pas d'autre scénario possible.

Notation : Dans tout ce chapitre, les systèmes considérés sont à coefficients dans $\mathbb{R}$ ou $\mathbb{C}$. Dans un souci de simplification des notations, nous adopterons la convention que le symbole $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$. On rappelle que $\mathbb{K}^n$ désigne l'ensemble des n-uplets d'éléments de $\mathbb{K}$, c'est-à-dire l'ensemble des $(u_1, \cdots, u_n)$ avec chaque u_i appartenant à $\mathbb{K}$.

2.2 Définitions

Définition 2.2.1 *Soit $p \in \mathbb{N}^*$ et $n \in \mathbb{N}^*$. On appelle* **système linéaire** *de p équations à n inconnues à coefficients dans $\mathbb{K}$ (ou encore système $p \times n$) tout système d'équations du type*

$$(S) \qquad \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{p1}x_1 + \cdots + a_{pn}x_n = b_p \end{cases}$$

avec $a_{ij} \in \mathbb{K}$ pour $1 \leq i \leq p$ et $1 \leq j \leq n$, et $b_i \in \mathbb{K}$.

Un tel système est dit **carré** *si $p = n$.*

Définition 2.2.2 *Si $b_1 = \cdots = b_p = 0$, le système est dit* **homogène**. *Pour un système linéaire général (S) de p équations à n inconnues, le système*

$$(S') \qquad \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{p1}x_1 + \cdots + a_{pn}x_n = 0 \end{cases}$$

est appelé **système homogène associé** *à (S).*

Définition 2.2.3 *Soit (S) un système linéaire $p \times n$. On appelle solution de (S) tout n-uplet $(u_1, \cdots, u_n)$ de $\mathbb{K}^n$ tel que*

$$\begin{cases} a_{11}u_1 + \cdots + a_{1n}u_n = b_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{p1}u_1 + \cdots + a_{pn}u_n = b_p. \end{cases}$$

Définition 2.2.4 *Deux systèmes (S_1) et (S_2) sont dits équivalents s'ils ont le même ensemble de solutions, c'est-à-dire si toute solution de (S_1) est solution de (S_2) et vice versa.*

Exemple : Les systèmes

$$\begin{cases} x_1 = 1 \\ x_1 - x_2 = 2 \end{cases} \quad \text{et} \quad \begin{cases} x_1 + x_2 = 0 \\ x_1 - x_2 = 2 \end{cases}$$

sont équivalents.

Définition 2.2.5 *On dit qu'un système carré est* **triangulaire** *si l'on a*

$$a_{ij} = 0 \quad \textit{pour tout couple} \quad (i,j) \quad \textit{tel que} \quad i < j \quad \textit{(système triangulaire inférieur)}$$

ou bien

$$a_{ij} = 0 \quad \textit{pour tout couple} \quad (i,j) \quad \textit{tel que} \quad i > j \quad \textit{(système triangulaire supérieur).}$$

Un système triangulaire est dit **à diagonale non nulle** *s'il est triangulaire et si tous les termes diagonaux sont non nuls.*

Exemple : Le système suivant est triangulaire supérieur à diagonale non nulle :

$$\begin{cases} x_1 + 5x_2 + x_4 = 1 \\ x_2 + x_3 = -5 \\ x_3 - 5x_4 = 0 \\ x_4 = -1. \end{cases}$$

Définition 2.2.6 *On dit qu'un système $p \times n$ est* **échelonné** *s'il existe un entier $k \in \{1, \cdots, n\}$ et un k-uplet d'entiers $j_1 < \cdots < j_k$ de $\{1, \cdots, n\}$ tel que*

1. *Pour $i \in \{1, \cdots, k\}$, on a* $\begin{cases} a_{ij} = 0 \quad si \quad j < j_i, \\ a_{ij_i} \neq 0. \end{cases}$
2. *Pour $i > k$, $a_{ij} = 0$.*

Les k premières équations sont appelées **équations principales**, *et les inconnues $x_{j_1}, \cdots, x_{j_k}$ sont appelées* **inconnues principales** .

Remarque 2.2.7 *Tout système triangulaire supérieur à diagonale non nulle est échelonné : on a $k = n$ et $j_i = i$ pour tout $i \in \{1, \cdots, n\}$.*

Exemple : Le système 4×5 suivant est échelonné :

$$\begin{cases} 2x_1 + 3x_2 + x_5 = 1 \\ x_2 + x_3 + x_4 - x_5 = 4 \\ x_5 = 0 \\ 0 = -3 \end{cases}$$

On a $k = 3$, $j_1 = 1$, $j_2 = 2$ et $j_3 = 5$. Les trois premières équations sont les équations principales, et x_1, x_2 et x_5 sont les inconnues principales.

2.3 Matrice associée à un système linéaire

Définition 2.3.1 *Soit (S) un système $p \times n$. Notons a_{ij} (avec i décrivant $\{1, \cdots, p\}$ et j décrivant $\{1, \cdots, n\}$) ses coefficients. On appelle* **matrice** *associée au système (S) le tableau de nombres*

$$A \overset{\text{déf}}{=} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{pmatrix}.$$

On dit que A est une matrice à p lignes, n colonnes et à coefficients dans $\mathbb{K}$. On note $\mathcal{M}_{p,n}(\mathbb{K})$ l'ensemble des matrices à p lignes et n colonnes à coefficients dans $\mathbb{K}$.

Si $p = n$, on dit que la matrice est carrée et on utilise plutôt la notation $\mathcal{M}_n(\mathbb{K})$ au lieu de $\mathcal{M}_{n,n}(\mathbb{K})$.

Notation : Soit $A \in \mathcal{M}_{p,n}(\mathbb{K})$. On note a_{ij} le coefficient général de A. Le premier indice (ici i) est celui des lignes, et le deuxième (ici j) est celui des colonnes. On utilise souvent la notation $A = (a_{ij})_{\substack{1\leq i\leq p\\ 1\leq j\leq n}}$.

Définition 2.3.2 *On dit que deux matrices de $A = (a_{ij})_{\substack{1\leq i\leq p\\ 1\leq j\leq n}}$ et $B = (b_{ij})_{\substack{1\leq i\leq p\\ 1\leq j\leq n}}$ de $\mathcal{M}_{p,n}(\mathbb{K})$ sont égales si leurs coefficients sont égaux : $a_{ij} = b_{ij}$ pour tout $i \in \{1, \cdots, p\}$ et $j \in \{1, \cdots, n\}$.*

Définition 2.3.3 *Soit $A \in \mathcal{M}_{p,n}(\mathbb{K})$.*

- *Les p lignes $(a_{i1}, \cdots, a_{in})$ pour $i \in \{1, \cdots, p\}$ sont appelées* **vecteurs lignes** *de A.*
- *Les n colonnes $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{pj} \end{pmatrix}$ pour $j \in \{1, \cdots, n\}$ sont appelées* **vecteurs colonnes** *de A.*

Exemple : Le système

$$(S) \qquad \begin{cases} x_1 + 5x_3 = 2 \\ x_1 - x_2 + x_3 = 0 \\ \frac{5}{3}x_1 - 2x_2 = 1 \end{cases}$$

a pour matrice associée la matrice carrée à 3 lignes et 3 colonnes

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

Les vecteurs lignes de A sont

$$(1, 0, 5) \quad (1, -1, 1) \quad \text{et} \quad (\tfrac{5}{3}, -2, 0).$$

Les vecteurs colonnes de A sont

$$\begin{pmatrix} 1 \\ 1 \\ \frac{5}{3} \end{pmatrix}, \quad \begin{pmatrix} 0 \\ -1 \\ -2 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 5 \\ 1 \\ 0 \end{pmatrix}.$$

Remarque : Si $A \in \mathcal{M}_{1,n}(\mathbb{K})$ (i.e $p = 1$), A est appelée *matrice ligne*.

Si $A \in \mathcal{M}_{p,1}(\mathbb{K})$ (i.e $n = 1$), A est appelée *matrice colonne*.

Définition 2.3.4 *Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée. Les coefficients a_{ii} pour i décrivant $1, \cdots, n$ sont appelés* **coefficients diagonaux** *de A. On dit que A est une* **matrice diagonale** *si ses coefficients non diagonaux a_{ij} avec $i \neq j$ sont tous nuls.*

La matrice diagonale dont tous les coefficients diagonaux sont égaux à 1 est appelée **matrice identité** *et notée I_n.*

Définition 2.3.5 *On dit qu'une matrice carrée est* **triangulaire** *si l'on a*

$$a_{ij} = 0 \quad \textit{pour tout couple} \quad (i, j) \quad \textit{tel que} \quad i < j \quad \textit{(matrice triangulaire inférieure)}$$

ou bien

$$a_{ij} = 0 \quad \textit{pour tout couple} \quad (i, j) \quad \textit{tel que} \quad i > j \quad \textit{(matrice triangulaire supérieure).}$$

Proposition 2.3.6 *Un système est triangulaire supérieur (resp. inférieur) si et seulement si sa matrice associée est triangulaire supérieure (resp. inférieure).*

Remarque : De même, on dira que la matrice d'un système (S) est **échelonnée** si le système est lui-même échelonné.

Définition 2.3.7 *Soit $A \in \mathcal{M}_{p,n}(\mathbb{K})$. On appelle* **matrice transposée** *de A la matrice B de $\mathcal{M}_{n,p}(\mathbb{K})$[1] telle que $b_{ij} = a_{ji}$ pour $i \in \{1, \cdots, n\}$ et $j \in \{1, \cdots, p\}$. On note ${}^tA$ la matrice transposée de A.*

Exemple : La matrice transposée de $A = \begin{pmatrix} 2 & 3 \\ 1 & 0 \\ 4 & 2 \end{pmatrix}$ est ${}^tA = \begin{pmatrix} 2 & 1 & 4 \\ 3 & 0 & 2 \end{pmatrix}$.

La notation matricielle permet de récrire les systèmes linéaires sous forme très condensée. En effet, considérons un système linéaire de p équations à n inconnues :

$$(S) \qquad \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{p1}x_1 + \cdots + a_{pn}x_n &= b_p. \end{cases}$$

Notons A la matrice associée à ce système, $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $b = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix}$. Le système (S) se récrit

$$Ax = b.$$

Exemple : Le système

$$(S) \qquad \begin{cases} 2x_1 \qquad\quad - x_3 = 5 \\ x_1 + x_2 + x_3 = -1 \end{cases}$$

se récrit

$$\begin{pmatrix} 2 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \end{pmatrix}.$$

En pratique, pour la résolution des systèmes linéaires, on pourra adopter la notation condensée suivante :

$$(S) \iff \left(\begin{array}{ccc|c} 2 & 0 & -1 & 5 \\ 1 & 1 & 1 & -1 \end{array}\right)$$

2.4 Résolution des systèmes échelonnés

Nous allons voir que la résolution de tels systèmes est particulièrement aisée. Commençons par traiter le cas particulier des systèmes triangulaires.

2.4.1 Systèmes triangulaires à diagonale non nulle

Proposition 2.4.1 *Soit (S) un système triangulaire supérieur à diagonale non nulle. Alors (S) a une unique solution obtenue par la* **méthode de la remontée** *:*

$$x_n = \frac{b_n}{a_{nn}}, \quad x_{n-1} = \frac{b_{n-1} - a_{n-1n}x_n}{a_{n-1n-1}}, \ \cdots\cdots, \ x_1 = \frac{b_1 - a_{12}x_2 - \cdots - a_{1n}x_n}{a_{11}}.$$

[1] Il s'agit donc d'une matrice à n lignes et p colonnes

Preuve : Le système (S) est du type

$$\left\{\begin{array}{rcccccccl} a_{11}x_1 & + & \cdots & + & a_{1n-1}x_{n-1} & + & a_{1n}x_n & = & b_1 \\ & & \cdots\cdots & & \cdots\cdots & & \cdots\cdots & & \cdots\cdots \\ & & & & a_{n-1n-1}x_{n-1} & + & a_{n-1n}x_n & = & b_{n-1} \\ & & & & & & a_{nn}x_n & = & b_n. \end{array}\right.$$

Comme $a_{nn} \neq 0$, la dernière équation permet de calculer x_n. Comme $a_{n-1n-1} \neq 0$, l'avant dernière équation donne $x_{n-1} = \frac{b_{n-1} - a_{n-1n}x_n}{a_{n-1n-1}}$. Puis, de proche en proche, comme $a_{kk} \neq 0$,

$$x_k = \frac{b_k - a_{kk+1}x_{k+1} - \cdots - a_{kn}x_n}{a_{kk}}.$$

Comme $x_{k+1}, \cdots, x_n$ ont déjà été calculés, la formule ci-dessus permet de déterminer x_k. ■

Exercice : Montrer que les systèmes triangulaires inférieurs peuvent être résolus de façon analogue par la *méthode de la descente*.

2.4.2 Systèmes échelonnés

Considérons un système échelonné général $p \times n$:

$$(S) \qquad \left\{\begin{array}{rcccccccl} a_{1j_1}x_{j_1} & + & \cdots\cdots & + & a_{1n}x_n & = & b_1 \\ & & \cdots\cdots & & \cdots\cdots & & \cdots\cdots \\ & & a_{kj_k}x_{j_k} & + & \cdots & + & a_{kn}x_n & = & b_k \\ & & & & & & 0 & = & b_{k+1} \\ & & & & & & 0 & = & b_p. \end{array}\right.$$

1er cas : L'un des b_j avec $j \in \{k+1, \cdots, p\}$ est non nul. Alors (S) n'a pas de solution.
2ème cas : $b_{k+1} = \cdots = b_p = 0$. Alors (S) est équivalent au système (Σ) suivant :

$$(\Sigma) \qquad \left\{\begin{array}{rcccccccl} a_{1j_1}x_{j_1} & + & \cdots\cdots & + & a_{1n}x_n & = & b_1 \\ & & \cdots\cdots & & \cdots\cdots & & \cdots\cdots \\ & & a_{kj_k}x_{j_k} & + & \cdots & + & a_{kn}x_n & = & b_k. \end{array}\right.$$

Ce nouveau système se résout facilement par la méthode de la remontée en considérant les inconnues non principales (x_j avec $j \neq j_i$ pour tout i) comme des paramètres libres. Si $k = n$, le système (Σ) est tout simplement un système triangulaire supérieur à diagonale non nulle, et la proposition 2.4.1 s'applique.
Sinon, on doit avoir $k < n$, et on obtient une infinité de solutions $(x_1, \cdots, x_n)$ données par les formules suivantes :

$$\begin{array}{rcl} x_{j_k} & = & \frac{b_k - a_{kj_k+1}x_{j_k+1} - \cdots - a_{kn}x_n}{a_{kj_k}}, \\ \cdots\cdots & & \cdots\cdots\cdots\cdots, \\ x_{j_1} & = & \frac{b_1 - a_{1j_1+1}x_{j_1+1} - \cdots - a_{1n}x_n}{a_{1j_1}}, \end{array}$$

avec x_j pour $j \neq j_i$ choisi arbitrairement dans $\mathbb{K}$.

Exemple : Soit α un paramètre réel. On veut résoudre

$$\left(\begin{array}{ccccc|c} 2 & 3 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 & 0 & \alpha \end{array}\right)$$

1er cas : $\alpha \neq 0$. Le système n'a pas de solution.
2ème cas : $\alpha = 0$. Le système est équivalent à

$$\left.\begin{array}{ccccc|c} 2 & 3 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 6 \end{array}\right.$$

Les inconnues principales sont x_1, x_2 et x_5. Les deux autres inconnues x_3 et x_4 sont des paramètres libres. Par la méthode de la remontée, on trouve :

$$\begin{cases} x_5 = 6, \\ x_2 = 3 - x_4 + x_5 = 9 - x_4, \\ x_1 = \frac{1 - x_5 - 3x_2}{2} = -16 + \frac{3}{2}x_4. \end{cases}$$

L'ensemble des solutions de (S) est

$$\mathcal{E} = \left\{ \left(-16 + \frac{3}{2}x_4, 9 - x_4, x_3, x_4, 6 \right) \;\middle|\; x_3 \in \mathbb{R}, x_4 \in \mathbb{R} \right\}.$$

2.5 Méthode du pivot de Gauss

La méthode du pivot de Gauss consiste à transformer un système (S) en un système échelonné équivalent à l'aide de *transformations élémentaires* .

Les transformations élémentaires sont de trois types :
($T1$) Échange de deux lignes du système,
($T2$) Multiplication d'une ligne par un scalaire non nul,
($T3$) Ajout à une ligne d'un multiple d'une autre ligne.

L'importance que l'on accorde aux transformations élémentaires est justifiée par le résultat suivant :

Proposition 2.5.1 *Deux systèmes (S_1) et (S_2) se déduisant l'un de l'autre par une succession de transformations élémentaires sont équivalents.*

Remarque : Autrement dit, faire des transformations élémentaires ne change pas l'ensemble des solutions d'un système linéaire.

Preuve : Il suffit de vérifier que chaque type de transformation élémentaire ne modifie pas l'ensemble des solutions. Pour ($T1$) (i.e permutations de deux lignes), c'est évident.
Pour ($T2$) c'est également clair : si $(x_1, \cdots, x_n)$ est solution, alors on a pour tout $\alpha \neq 0$,

$$a_{i1}x_1 + \cdots + a_{in}x_n = b_i \iff \alpha a_{i1}x_1 + \cdots + \alpha a_{in}x_n = \alpha b_i.$$

Reste à vérifier pour ($T3$). Supposons que l'on ajoute $\alpha(L_{i_1})$ à la ligne (L_{i_0}) (avec $i_0 \neq i_1$). Notons $(a_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$ la matrice de (S), et $(a'_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$ la matrice du système (S') obtenu après avoir ajouté $\alpha(L_{i_1})$ à la ligne (L_{i_0}). Notons $\mathcal{E}$ (resp. $\mathcal{E}'$) l'ensemble des solutions de (S) (resp. (S')).
Il est clair que

$$\begin{cases} a'_{ij} = a_{ij} \quad \text{si} \quad i \neq i_0, \\ a'_{i_0 j} = a_{i_0 j} + \alpha a_{i_1 j}. \end{cases} \tag{2.1}$$

Soit $(u_1, \cdots, u_n)$ une solution de (S). On a par définition

$$\forall i \in \{1, \cdots, p\},\ a_{i1}u_1 + \cdots + a_{in}u_n = b_i. \tag{2.2}$$

Comme $a'_{ij} = a_{ij}$ pour $i \neq i_0$, le n-uplet $(u_1, \cdots, u_n)$ satisfait les lignes de (S') distinctes de i_0.

En ajoutant α fois l'égalité (2.2) avec $i = i_1$ à l'égalité (2.2) avec $i = i_0$, on trouve

$$(a_{i_0 1} + \alpha a_{i_1 1})u_1 + \cdots + (a_{i_0 n} + \alpha a_{i_1 n})u_n = b_{i_0} + \alpha b_{i_1}$$

qui est exactement la ligne i_0 de (S'). Donc $(u_1, \cdots, u_n)$ est solution de (S'). D'où $\mathcal{E} \subset \mathcal{E}'$.

Pour montrer l'inclusion réciproque, il suffit de remarquer que l'on passe de (S') à (S) en retranchant $\alpha(L_{i_1})$ à (L_{i_0}). On reprend alors le raisonnement précédent en échangeant les rôles de (S) et de (S'), et en remplaçant α par $-\alpha$. ■

Exemple : Mettre le système (S) : $\left(\begin{array}{cccc|c} 0 & 4 & 0 & 4 & 2 \\ 2 & 3 & 1 & 0 & 1 \\ 2 & 0 & -1 & 0 & 0 \end{array}\right)$ sous forme échelonnée :

$$(S) \iff \left(\begin{array}{cccc|c} 2 & 0 & -1 & 0 & 0 \\ 2 & 3 & 1 & 0 & 1 \\ 0 & 4 & 0 & 4 & 2 \end{array}\right) \quad (\text{échange de } (L_1) \text{ et } (L_3)),$$

$$\iff \left(\begin{array}{cccc|c} 2 & 0 & -1 & 0 & 0 \\ 2 & 3 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & \frac{1}{2} \end{array}\right) \quad \left(\text{multiplication de } (L_3) \text{ par } \tfrac{1}{4}\right),$$

$$\iff \left(\begin{array}{cccc|c} 2 & 0 & -1 & 0 & 0 \\ 0 & 3 & 2 & 0 & 1 \\ 0 & 1 & 0 & 1 & \frac{1}{2} \end{array}\right) \quad (\text{on retranche } (L_1) \text{ à } (L_2)),$$

$$\iff \left(\begin{array}{cccc|c} 2 & 0 & -1 & 0 & 0 \\ 0 & 3 & 2 & 0 & 1 \\ 0 & 0 & -\frac{2}{3} & 1 & \frac{1}{6} \end{array}\right) \quad \left(\text{on retranche } (\tfrac{1}{3}L_1) \text{ à } (L_3)\right),$$

Le système obtenu après cette succession de transformations élémentaires est *échelonné* (les inconnues principales sont x_1, x_2 et x_3. On peut donc le résoudre par la méthode de la remontée. La proposition 2.5.1 assure que ce nouveau système est équivalent au système initial.

Algorithme du pivot de Gauss : Considérons un système (S) de taille $p \times n$ et de matrice A. On veut résoudre $Ax = b$.
Le pivot de Gauss est une méthode *itérative* permettant de transformer n'importe quel système linéaire en un système échelonné équivalent après un nombre fini de transformations élémentaires.

Première itération :

Premier cas : La matrice A est nulle. L'algorithme est alors terminé.

Deuxième cas : $A \neq 0$. Soit j_1 l'indice de la première colonne non nulle.

1*ère étape :* Par permutation de lignes on se ramène au cas où $a_{1j_1} \neq 0$. Le système (S) est donc équivalent à un système de matrice

$$\overbrace{\left(\begin{array}{ccc|cc} 0 & \cdots & 0 & a_{1j_1} & * \\ \vdots & & \vdots & * & * \\ 0 & \cdots & 0 & * & * \end{array}\right)}^{n \text{ colonnes}} \left.\begin{array}{l}\\ \\ \\ \end{array}\right\} p \text{ lignes}.$$

2*ème étape :* Le but de cette étape est de faire apparaître des 0 dans la colonne j_1 sous le coefficient a_{1j_1}. Pour cela, on retranche $\frac{a_{ij_1}}{a_{1j_1}}(L_1)$ à chaque ligne (L_i) avec $i \geq 2$. Après avoir appliqué ces $p-1$ opérations élémentaires, le système équivalent obtenu a pour matrice

$$\begin{array}{c} 1 \\ \\ \\ p \end{array} \left(\begin{array}{ccc|cc} 0 & \cdots & 0 & a_{1j_1} & * \\ \vdots & & \vdots & 0 & * \\ \vdots & & \vdots & \vdots & * \\ 0 & \cdots & 0 & 0 & * \end{array}\right).$$

Itération suivante : On ne touche plus à la première ligne et l'on applique la méthode de la première itération au système $(p-1) \times n$ constitué par les lignes 2 à p du système obtenu à la fin de la première étape.

Fin de l'algorithme : L'algorithme s'arrête au bout d'au plus $p-1$ itérations ou lorsque le sous-système obtenu a toutes ses lignes nulles.

Remarque : Pour un système $p \times n$, l'itération type nécessite environ pn additions, soustractions, multiplications ou divisions. L'algorithme du pivot permet donc de rendre un système échelonné en effectuant environ p^2n opérations (n^3 opérations si le système est carré). Son côté automatique le rend facilement exécutable par un ordinateur. Dans les cas pratiques, la méthode du pivot de Gauss n'est pas forcément la plus rapide. **Il n'est pas interdit de réfléchir avant d'appliquer aveuglément l'algorithme !**

2.6 Structure de l'ensemble des solutions d'un système linéaire

2.6.1 Un exemple

Cherchons à résoudre le système (S) suivant :

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 2 & 3 & 1 & 0 & 2 \\ 1 & 0 & 1 & 1 & 4 \end{array}\right)$$

Notons (S') le système homogène associé à (S).

Pour résoudre (S), on applique l'algorithme du pivot de Gauss :

$$(S) \iff \left(\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 2 & 2 \\ 0 & -1 & 1 & 2 & 4 \end{array}\right) \quad \text{(on retranche } 2(L_1) \text{ à } (L_2), \text{ et } (L_1) \text{ à } (L_3)),$$

$$\iff \left(\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 2 & 2 \\ 0 & 0 & 2 & 4 & 6 \end{array}\right) \quad \text{(ajout de } (L_2) \text{ à } (L_3)).$$

Ce dernier système est échelonné et a pour inconnues principales x_1, x_2 et x_3. On en déduit

$$(S) \iff \left\{ \begin{array}{l} x_3 = 3 - 2x_4, \\ x_2 = 2 - 2x_4 - x_3 = -1, \\ x_1 = x_4 - x_2 = 4 - x_4 + 1. \end{array} \right.$$

Donc la solution générale s'écrit

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 \\ -1 \\ 3 \\ 0 \end{pmatrix}}_{\text{solution particulière de } (S)} + \underbrace{\lambda \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \end{pmatrix}}_{\text{solution générale de } (S')}.$$

Nous allons voir que l'ensemble des solutions d'un système linéaire, s'il n'est pas vide, peut toujours s'écrire comme la somme d'une solution particulière et de l'ensemble des solutions du système homogène associé.

2.6.2 Cas des systèmes homogènes

Proposition 2.6.1 *Soit (S') un système homogène $p \times n$ et $\mathcal{E}'$ l'ensemble de ses solutions. Alors on a*

i) $(0, \cdots, 0) \in \mathcal{E}'$ *(et donc $\mathcal{E}'$ n'est pas vide),*

ii) Si $(x_1, \cdots, x_n) \in \mathcal{E}'$ *et* $(y_1, \cdots, y_n) \in \mathcal{E}'$ *alors* $(x_1 + y_1, \cdots, x_n + y_n) \in \mathcal{E}'$.

iii) Si $(x_1, \cdots, x_n) \in \mathcal{E}'$ *et* $\lambda \in \mathbb{K}$ *alors* $(\lambda x_1, \cdots, \lambda x_n) \in \mathcal{E}'$.

On dit que $\mathcal{E}'$ est un **sous-espace vectoriel** *de $\mathbb{K}^n$.*

Preuve : Le second membre d'un système homogène est une colonne de 0. Il est donc immédiat que $(0, \cdots, 0) \in \mathcal{E}'$.

Supposons maintenant que $(x_1, \cdots, x_n)$ et $(y_1, \cdots, y_n)$ soient solutions de (S'). Alors pour tout $i \in \{1, \cdots, p\}$, on a

$$a_{i1}x_1 + \cdots a_{in}x_n = 0 \quad \text{et} \quad a_{i1}y_1 + \cdots a_{in}y_n = 0.$$

Donc, en sommant les deux égalités,

$$a_{i1}(x_1 + y_1) + \cdots a_{in}(x_n + y_n) = 0.$$

Il est également clair que pour tout $\lambda \in \mathbb{K}$, on a $a_{i1}\lambda x_1 + \cdots a_{in}\lambda x_n = 0$. Les points *ii)* et *iii)* sont donc prouvés. ∎

2.6.3 Cas général

Proposition 2.6.2 *Soit (S) un système linéaire et (S') le système linéaire homogène associé. Notons $\mathcal{E}$ et $\mathcal{E}'$ leurs ensembles de solutions respectifs. Supposons de plus que $\mathcal{E}$ ne soit pas vide et donnons-nous $(x_1^0, \cdots, x_n^0)$ une solution particulière de (S). Alors*

$$(x_1, \cdots, x_n) \in \mathcal{E} \iff \exists (x_1', \cdots, x_n') \in \mathcal{E}' \quad \textit{tel que} \quad (x_1, \cdots, x_n) = (x_1^0 + x_1', \cdots, x_n^0 + x_n').$$

Autrement dit, si $\mathcal{E}$ n'est pas vide alors $\mathcal{E}$ est la somme des solutions de $\mathcal{E}'$ et d'une solution particulière de $\mathcal{E}$. On dit que $\mathcal{E}'$ est un **sous-espace affine** de $\mathbb{K}^n$.

Preuve : $\Longrightarrow$ Soit $(x_1^0, \cdots, x_n^0)$ une solution particulière de (S). Alors on a pour tout $i \in \{1, \cdots, p\}$,

$$a_{i1}x_1^0 + \cdots + a_{in}x_n^0 = b_i. \tag{2.3}$$

Par ailleurs, $(x_1, \cdots, x_n) \in \mathcal{E}$ si et seulement si pour tout $i \in \{1, \cdots, p\}$,

$$a_{i1}x_1 + \cdots + a_{in}x_n = b_i.$$

En soustrayant à (2.3) cette deuxième relation, on trouve

$$a_{i1}(x_1 - x_1^0) + \cdots a_{in}(x_n - x_n^0) = 0,$$

ce qui signifie que $(x_1', \cdots, x_n') \overset{\text{déf}}{=} (x_1 - x_1^0, \cdots, x_n - x_n^0)$ est solution du système homogène associé à (S).

$\Longleftarrow$ Supposons que $(x_1', \cdots, x_n') \in \mathcal{E}'$. Alors on a $a_{i1}x_1' + \cdots + a_{in}x_n' = 0$ pour tout $i \in \{1, \cdots, p\}$. En ajoutant l'égalité (2.3), on trouve

$$a_{i1}(x_1^0 + x_1') + \cdots (x_n^0 + x_n') = 0$$

pour tout $i \in \{1, \cdots, p\}$, et donc $(x_1^0 + x_1', \cdots, x_n^0 + x_n')$ est solution de (S). ■

Chapitre 3

Familles de vecteurs

Dans tout ce chapitre, $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$, et n est un entier supérieur ou égal à 1.

3.1 Vecteurs de $\mathbb{K}^n$

Définition 3.1.1 *On appelle* **vecteur à n composantes** *tout n-uplet $(x_1, \cdots, x_n)$ d'éléments de $\mathbb{K}$. L'ensemble des vecteurs à n composantes (appelés plus simplement vecteurs) est noté $\mathbb{K}^n$, et l'on pose $\vec{x} = (x_1, \cdots, x_n)$.*

On munit $\mathbb{K}^n$ d'une loi interne " $+$ " définie pour tous vecteurs $\vec{x}$ et $\vec{y}$ de $\mathbb{K}^n$ par

$$\vec{x} + \vec{y} \overset{\text{déf}}{=} \overrightarrow{x + y} = (x_1 + y_1, \cdots, x_n + y_n),$$

et d'une loi externe " $\cdot$ " définie pour tout scalaire $\lambda \in \mathbb{K}$ et vecteur $\vec{x} \in \mathbb{K}^n$ par

$$\lambda \cdot \vec{x} \overset{\text{déf}}{=} \overrightarrow{\lambda \cdot x} = (\lambda x_1, \cdots, \lambda x_n).$$

Proposition 3.1.2 *$(\mathbb{K}^n, +)$ est un groupe commutatif.*

Preuve : La commutativité et l'associativité de la loi $+$ résultent de celles de l'addition dans $\mathbb{R}$ ou $\mathbb{C}$. Il est de plus clair que l'élément neutre est le *vecteur nul* $\vec{0} \overset{\text{déf}}{=} (0, \cdots, 0)$ et que le symétrique de $\vec{x}$ est $-\vec{x} \overset{\text{déf}}{=} (-1) \cdot \vec{x} = (-x_1, \cdots, -x_n)$. ■

Remarque 3.1.3 *De plus, $(\mathbb{K}^n, +)$ est stable par la loi externe " $\cdot$ " décrite plus haut : pour tous scalaires λ et μ, et pour tous vecteurs $\vec{x}$ et $\vec{y}$, on a*

$$\begin{cases} \lambda \cdot (\vec{x} + \vec{y}) = \lambda \cdot \vec{x} + \lambda \cdot \vec{y}, \\ (\lambda + \mu) \cdot \vec{x} = \lambda \cdot \vec{x} + \mu \cdot \vec{x}, \\ \lambda \cdot (\mu \cdot \vec{x}) = (\lambda\mu) \cdot \vec{x}, \\ 1 \cdot \vec{x} = \vec{x}. \end{cases}$$

On dit que $(\mathbb{K}^n, +, \cdot)$ est un **espace vectoriel** *sur $\mathbb{K}$.*

Remarques :

1. Par convention $\mathbb{K}^0$ est l'espace vectoriel "trivial" contenant un seul élément noté $\vec{0}$. C'est le plus "petit"de tous les $\mathbb{K}$-espaces vectoriels.
2. Sauf en cas d'ambiguïté, on omettra le point de la multiplication par un scalaire : $\lambda\vec{x}$ désignera $\lambda \cdot \vec{x}$.

L'étude détaillée des espaces vectoriels fera l'objet du cours du second semestre. Nous nous limitons dans un premier temps aux sous-espaces vectoriels :

Définition 3.1.4 *Soit $X \subset \mathbb{K}^n$. On dit que X est un* **sous-espace vectoriel**[1] *de $\mathbb{K}^n$ si*

1. *$(X, +)$ est un sous-groupe de $(\mathbb{K}^n, +)$,*
2. *$\forall \vec{x} \in X$, $\forall \lambda \in \mathbb{K}$, $\lambda \vec{x} \in X$.*

Pour prouver qu'un sous-ensemble X de $\mathbb{K}^n$ est un s.e.v de $\mathbb{K}^n$, on fait généralement appel à la proposition suivante :

Proposition 3.1.5 *$X \subset \mathbb{K}^n$ est un sous-espace vectoriel de $\mathbb{K}^n$ si et seulement si*

1. *X contient $\vec{0}$,*
2. *$\forall \lambda \in \mathbb{K}$, $\forall \mu \in \mathbb{K}$, $\forall \vec{x} \in X$, $\forall \vec{y} \in X$, $\lambda \vec{x} + \mu \vec{y} \in X$.*

Preuve : Il est clair que tout sous-espace vectoriel vérifie 1 et 2.

Réciproquement, considérons un sous-ensemble X de $\mathbb{K}^n$ vérifiant les propriétés 1 et 2. On sait déjà que X contient $\vec{0}$ qui est l'élément neutre pour la loi $+$. De plus, toujours d'après 2, pour tout couple $(\vec{x}, \vec{y}) \in X \times X$, on a

$$\vec{x} + \vec{y} = 1\vec{x} + 1\vec{y} \in X.$$

Donc $(X, +)$ est un sous-groupe de $(\mathbb{K}^n, +)$.

Enfin, si $\lambda \in \mathbb{K}$ et $\vec{x} \in X$, on a $\lambda \vec{x} = \lambda \vec{x} + 1.\vec{0} \in X$. ■

3.2 Familles de vecteurs

3.2.1 Combinaisons linéaires

Définition 3.2.1 *Soit I un ensemble non vide. On appelle* **famille de vecteurs** *de $\mathbb{K}^n$ indexée par I tout ensemble $\{\vec{x}_i \mid i \in I\}$ de vecteurs de $\mathbb{K}^n$ où l'indice i décrit tous les éléments de I. Une famille de vecteurs indexée par I sera notée $(\vec{x}_i)_{i \in I}$. Si $I = \{i_1, \cdots, i_k\}$, on utilisera également la notation $(\vec{x}_{i_1}, \cdots, \vec{x}_{i_k})$.*

- *Si $J \subset I$, on dit que $(\vec{x}_j)_{j \in J}$ est une* **sous-famille** *de $(\vec{x}_i)_{i \in I}$.*
- *Si $I \subset K$, on dit que $(\vec{x}_k)_{k \in K}$ est une* **sur-famille** *de $(\vec{x}_i)_{i \in I}$.*

Remarque 3.2.2 *Il est parfois commode d'étendre la définition ci-dessus au cas où $I = \emptyset$. Par convention, la famille indexée par l'ensemble vide est $\emptyset$. On l'appelle* **famille vide** *de $\mathbb{K}^n$. Bien évidemment, la famille vide est sous-famille de toute famille de vecteurs.*

Dans la suite du cours, on se limite à des familles *finies* de vecteurs. Le plus souvent, ces familles seront indexées par $I = \{1, \cdots, k\}$ et notées $(\vec{x}_1, \cdots, \vec{x}_k)$.

Définition 3.2.3 *Soit $(\vec{x}_1, \cdots, \vec{x}_p)$ une famille de vecteurs de $\mathbb{K}^n$. On dit que $\vec{y} \in \mathbb{K}^n$ est* **combinaison linéaire** *de la famille $(\vec{x}_1, \cdots, \vec{x}_p)$ s'il existe un p-uplet $(\lambda_1, \cdots, \lambda_p)$ d'éléments de $\mathbb{K}$ tel que*

$$\vec{y} = \lambda_1 \vec{x}_1 + \cdots + \lambda_n \vec{x}_p.$$

Convention : Toute combinaison linéaire de zéro vecteur (i.e de la famille vide) est égale au vecteur nul.

Proposition 3.2.4 *Tout sous-espace vectoriel de $\mathbb{K}^n$ est stable par combinaison linéaire d'un nombre arbitraire de ses vecteurs.*

Preuve : La proposition 3.1.5 montre la stabilité par combinaison linéaire de *deux* vecteurs. Une récurrence élémentaire donne le cas général. ■

[1] ou s.e.v en abrégé

3.2.2 Familles génératrices

Définition 3.2.5 *Soit X un sous-espace vectoriel de $\mathbb{K}^n$. On dit qu'une famille de vecteurs $(\vec{x}_1, \cdots, \vec{x}_k)$ de X est* **génératrice** *si tout élément de X est combinaison linéaire de $(\vec{x}_1, \cdots, \vec{x}_k)$.*

Proposition 3.2.6 *Soit $(\vec{u}_1, \cdots, \vec{u}_k)$ une famille de vecteurs de $\mathbb{K}^n$. L'ensemble des combinaisons linéaires des $\vec{u}_i$ est un sous-espace vectoriel de $\mathbb{K}^n$. On l'appelle* **sous-espace vectoriel engendré** *par la famille $(\vec{u}_1, \cdots, \vec{u}_k)$ et on le note* $\text{Vect}\,(\vec{u}_1, \cdots, \vec{u}_k)$. *C'est le plus petit sous-espace vectoriel contenant tous les vecteurs de la famille $(\vec{u}_1, \cdots, \vec{u}_k)$.*

Preuve : Pour vérifier que $\text{Vect}\,(\vec{u}_1, \cdots, \vec{u}_k)$ est un sous-espace vectoriel, on va appliquer la proposition 3.1.5.

- $\text{Vect}\,(\vec{u}_1, \cdots, \vec{u}_k)$ n'est pas vide car contient $\vec{u}_1$.
- Si $\vec{x}$ et $\vec{y}$ sont deux éléments de $\text{Vect}\,(\vec{u}_1, \cdots, \vec{u}_k)$ alors on peut trouver deux k-uplets $(\lambda_1, \cdots, \lambda_k)$ et $(\mu_1, \cdots, \mu_k)$ tels que

$$\vec{x} = \sum_{i=1}^{k} \lambda_i \vec{u}_i \quad \text{et} \quad \vec{y} = \sum_{i=1}^{k} \mu_i \vec{u}_i.$$

Pour tout couple (λ, μ) de $\mathbb{K}^2$, on a donc

$$\lambda\vec{x} + \mu\vec{y} = \sum_{i=1}^{k} (\lambda\lambda_i + \mu\mu_i)\vec{u}_i.$$

Donc $\lambda\vec{x} + \mu\vec{y}$ est bien combinaison linéaire de $(\vec{u}_1, \cdots, \vec{u}_k)$

Enfin, si X est un sous-espace vectoriel contenant chacun des vecteurs u_i, il contient toute combinaison linéaire de la famille $(\vec{u}_1, \cdots, \vec{u}_k)$ (cf Prop. 3.2.4) donc $\text{Vect}\,(\vec{u}_1, \cdots, \vec{u}_k)$. ■

Remarque : Le sous-espace vectoriel engendré par la famille vide de $\mathbb{K}^n$ est $\{\vec{0}\}$.

Nous laissons au lecteur le soin d'établir le résultat suivant :

Proposition 3.2.7 *Soit $(\vec{x}_1, \cdots, \vec{x}_p)$ une famille de vecteurs de $\mathbb{K}^n$, et $(\vec{x}_{i_1}, \cdots, \vec{x}_{i_k})$ une sous-famille de $(\vec{x}_1, \cdots, \vec{x}_p)$. Alors on a*

$$\text{Vect}\,(\vec{x}_{i_1}, \cdots, \vec{x}_{i_k}) \subset \text{Vect}\,(\vec{x}_1, \cdots, \vec{x}_p).$$

Pour déterminer le sous-espace vectoriel engendré par une famille de vecteurs, on fait souvent appel à la proposition suivante :

Proposition 3.2.8 *Le sous-espace vectoriel engendré par une famille de vecteurs donnée est invariant par les opérations suivantes :*

- $(T1)$ *Permutation de deux vecteurs,*
- $(T2)$ *Multiplication d'un vecteur par un scalaire non nul,*
- $(T3)$ *Ajout à l'un des vecteurs d'une combinaison linéaire des autres vecteurs.*

Preuve : Soit $(\vec{x}_1, \cdots, \vec{x}_p)$ une famille de vecteurs. L'invariance de $\text{Vect}\,(\vec{x}_1, \cdots, \vec{x}_p)$ par les transformations $(T1)$ et $(T2)$ est évidente.

Pour (T3), il suffit de considérer le cas où l'on ajoute un seul vecteur, le cas général suit par récurrence. Quitte à changer l'ordre des vecteurs (ce qui ne change pas les sous-espaces vectoriels engendrés), il suffit de prouver par exemple que pour tout $\alpha \in \mathbb{K}$, on a

$$\text{Vect}\,(\vec{x}_1 + \alpha\vec{x}_2, \vec{x}_2, \cdots, \vec{x}_p) = \text{Vect}\,(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_p).$$

Soit $\vec{y} \in \text{Vect}\,(\vec{x}_1 + \alpha\vec{x}_2, \vec{x}_2, \cdots, \vec{x}_p)$. Alors il existe $(\lambda_1, \cdots, \lambda_p) \in \mathbb{K}^p$ tel que

$$\vec{y} = \lambda_1(\vec{x}_1 + \alpha\vec{x}_2) + \lambda_2\vec{x}_2 + \cdots + \lambda_p\vec{x}_p.$$

On a donc

$$\vec{y} = \lambda_1\vec{x}_1 + (\lambda_2 + \alpha\lambda_1)\vec{x}_2 + \cdots + \lambda_p\vec{x}_p.$$

et, par conséquent, $\vec{y} \in \text{Vect}\,(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_p)$. On a donc montré que

$$\text{Vect}\,(\vec{x}_1 + \alpha\vec{x}_2, \vec{x}_2, \cdots, \vec{x}_p) \subset \text{Vect}\,(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_p).$$

Pour montrer l'inclusion réciproque, on considère $\vec{y} \in \text{Vect}\,(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_p)$. Il existe donc $(\mu_1, \cdots, \mu_p) \in \mathbb{K}^p$ tel que $\vec{y} = \mu_1\vec{x}_1 + \mu_2\vec{x}_2 + \cdots + \mu_p\vec{x}_p$, ce qui peut se récrire

$$\vec{y} = \mu_1(\vec{x}_1 + \alpha\vec{x}_2) + (\mu_2 - \alpha\mu_1)\vec{x}_2 + \mu_2\vec{x}_3 + \cdots + \mu_p\vec{x}_p.$$

On a donc bien $\vec{y} \in \text{Vect}\,(\vec{x}_1 + \alpha\vec{x}_2, \vec{x}_2, \cdots, \vec{x}_p)$.

■

3.2.3 Familles libres et familles liées

Définition 3.2.9 *Soit $(\vec{u}_1, \cdots, \vec{u}_k)$ une famille de $\mathbb{K}^n$.*

- *On dit que $(\vec{u}_1, \cdots, \vec{u}_k)$ est* **libre** *(ou linéairement indépendante) si*

$$\forall(\lambda_1, \cdots, \lambda_k) \in \mathbb{K}^k,\ \Big(\sum_{i=1}^{k} \lambda_i\vec{u}_i = \vec{0}\Big) \implies \Big(\lambda_1 = \cdots = \lambda_k = 0\Big).$$

- *On dit que $(\vec{u}_1, \cdots, \vec{u}_k)$ est* **liée** *(ou linéairement dépendante) si elle n'est pas libre, c'est-à-dire s'il existe un k-uplet $(\lambda_1, \cdots, \lambda_k) \neq (0, \cdots, 0)$ tel que $\sum_{i=1}^{k} \lambda_i\vec{u}_i = \vec{0}$.*

Convention : La famille vide est libre.

Proposition 3.2.10 *Si $(\vec{u}_1, \cdots, \vec{u}_k)$ est libre et $\vec{x} \in \text{Vect}\,(\vec{u}_1, \cdots, \vec{u}_k)$ alors il existe un unique k-uplet $(\lambda_1, \cdots, \lambda_k)$ d'éléments de $\mathbb{K}$ tel que $\vec{x} = \sum_{i=1}^{k} \lambda_i\vec{u}_i$.*

Autrement dit,

$$\Big(\vec{x} = \sum_{i=1}^{k} \lambda_i\vec{u}_i = \sum_{i=1}^{k} \mu_i\vec{u}_i\Big) \implies \Big(\lambda_1 = \mu_1, \cdots, \lambda_k = \mu_k\Big).$$

Preuve : L'égalité $\sum_{i=1}^{k} \lambda_i\vec{u}_i = \sum_{i=1}^{k} \mu_i\vec{u}_i$ entraîne $\sum_{i=1}^{k}(\lambda_i - \mu_i)\vec{u}_i = \vec{0}$, et donc, puisque $(\vec{u}_1, \cdots, \vec{u}_k)$ est libre, $\lambda_i - \mu_i = 0$ pour tout $i \in \{1, \cdots, k\}$. ■

Proposition 3.2.11 *Soit $(\vec{x}_1, \cdots, \vec{x}_k)$ une famille de vecteurs de $\mathbb{K}^n$, et $\vec{y} \in \mathbb{K}^n$.*

1. *Si $\vec{y} \in \text{Vect}\,(\vec{x}_1, \cdots, \vec{x}_k)$ alors la famille $(\vec{x}_1, \cdots, \vec{x}_k, \vec{y})$ est liée.*
2. *Réciproquement, si la famille $(\vec{x}_1, \cdots, \vec{x}_k, \vec{y})$ est liée et si de plus $(\vec{x}_1, \cdots, \vec{x}_k)$ est libre alors $\vec{y} \in \text{Vect}\,(\vec{x}_1, \cdots, \vec{x}_k)$.*

Preuve :

1. Si $\vec{y} \in \text{Vect}\,(\vec{x}_1, \cdots, \vec{x}_k)$ alors il existe un k-uplet $(\lambda_1, \cdots, \lambda_k)$ d'éléments de $\mathbb{K}$ tel que $\vec{y} = \sum_{i=1}^{k} \lambda_k\vec{x}_k$. On a donc $\sum_{i=1}^{k} \lambda_k\vec{x}_k - \vec{y} = \vec{0}$. Le $(k+1)$-uplet $(\lambda_1, \cdots, \lambda_k, -1)$ n'est pas identiquement nul donc la famille $(\vec{x}_1, \cdots, \vec{x}_k, \vec{y})$ est liée.

2. Réciproquement, supposons que $(\vec{x}_1, \cdots, \vec{x}_k, \vec{y})$ soit liée et que $(\vec{x}_1, \cdots, \vec{x}_k)$ soit libre. Alors il existe un $(k+1)$-uplet non identiquement nul $(\lambda_1, \cdots, \lambda_k, \lambda)$ de $\mathbb{K}^{k+1}$ tel que

$$\sum_{i=1}^{k} \lambda_i \vec{x}_i \ + \lambda \vec{y} = \vec{0}. \tag{3.1}$$

On peut de plus affirmer que $\lambda \neq 0$. En effet, si λ était nul alors (3.1) entraînerait que $\sum_{i=1}^{k} \lambda_i \vec{x}_i = \vec{0}$. Mais $(\vec{x}_1, \cdots, \vec{x}_k)$ est libre, donc $(\lambda_1, \cdots, \lambda_k) = (0, \cdots, 0)$, puis $(\lambda_1, \cdots, \lambda_k, \lambda) = (0, \cdots, 0)$, ce qui est contraire à l'hypothèse faite.
Donc λ n'est pas nul, et on peut écrire d'après (3.1),

$$\vec{y} = -\sum_{i=1}^{k} \frac{\lambda_i}{\lambda} \vec{x}_i.$$

Autrement dit, $\vec{y} \in \text{Vect}\,(\vec{x}_1, \cdots, \vec{x}_k)$.

■

Proposition 3.2.12 *Une famille $(\vec{x}_1, \cdots, \vec{x}_k)$ est liée si et seulement si elle contient un vecteur qui est combinaison linéaire des autres vecteurs.*

Preuve : $\Longrightarrow$ Supposons que $(\vec{x}_1, \cdots, \vec{x}_k)$ soit liée. Alors il existe un k-uplet $(\lambda_1, \cdots, \lambda_k)$ non nul tel que $\sum_{i=1}^{k} \lambda_i \vec{x}_i = \vec{0}$. Comme le k-uplet n'est pas nul, l'un des λ_i (disons λ_k pour fixer les idées) n'est pas nul et l'on a donc

$$\vec{x}_k = -\sum_{i=1}^{k-1} \frac{\lambda_i}{\lambda_k} \vec{x}_i,$$

et donc $\vec{x}_k$ est combinaison linéaire des autres vecteurs de la famille.

$\Longleftarrow$ Supposons par exemple que $\vec{x}_k$ soit combinaison linéaire de $(\vec{x}_1, \cdots, \vec{x}_{k-1})$. Alors $\vec{x}_k \in \text{Vect}\,(\vec{x}_1, \cdots, \vec{x}_{k-1})$ et la proposition 3.2.11 montre que la famille $(\vec{x}_1, \cdots, \vec{x}_k)$ est liée. ■

Cas particuliers :

- Toute famille contenant le vecteur nul est liée.
- Une famille de deux vecteurs $(\vec{x}_1, \vec{x}_2)$ est liée si et seulement si $\vec{x}_1$ et $\vec{x}_2$ sont colinéaires, i.e il existe $\lambda \in \mathbb{K}$ tel que

$$\vec{x}_1 = \lambda \vec{x}_2 \quad \text{ou} \quad \vec{x}_2 = \lambda \vec{x}_1.$$

Pour déterminer si une famille de vecteurs est libre ou liée, on a souvent recours à la proposition suivante :

Proposition 3.2.13
1. *Toute sous-famille d'une famille libre est libre.*
2. *Toute sur-famille d'une famille liée est liée.*
3. *Toute sur-famille d'une famille génératrice est génératrice.*
4. *Une sous-famille d'une famille non génératrice n'est pas génératrice non plus.*

Exercice : Prouver la proposition 3.2.13.

3.2.4 Bases

Définition 3.2.14 *Soit X un s.e.v de $\mathbb{K}^n$. On dit que la famille de vecteurs $(\vec{u}_1, \cdots, \vec{u}_k)$ est une* **base** *de X si elle est à la fois libre et génératrice de X.*

Tout vecteur $\vec{x}$ de X se décompose alors de manière unique en $\vec{x} = \sum_{i=1}^{k} x_i \vec{u}_i$. Le k-uplet $(x_1, \cdots, x_k)$ s'appelle **coordonnées** *de $\vec{x}$ par rapport à la base $(\vec{u}_1, \cdots, \vec{u}_k)$.*

Exemples :

1. Dans $\mathbb{R}^3$ (ou $\mathbb{C}^3$), la famille constituée des vecteurs $\vec{e}_1 = (1,0,0)$, $\vec{e}_2 = (0,1,0)$ et $\vec{e}_3 = (0,0,1)$ est une base. En effet, il clair que $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ est libre. De plus, tout vecteur $\vec{x} = (x_1, x_2, x_3)$ peut s'écrire
$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3.$$
La base $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ est appelée **base canonique** de $\mathbb{R}^3$ (ou $\mathbb{C}^3$).
2. En revanche, la famille $(\vec{e}_1, \vec{e}_2)$ n'est pas une base de $\mathbb{R}^3$. En effet, une combinaison linéaire de $\vec{e}_1$ et de $\vec{e}_2$ a sa troisième composante nulle. Le vecteur $\vec{e}_3$ n'est donc pas dans $\text{Vect}\,(\vec{e}_1, \vec{e}_2)$.
3. Définissons $\vec{u} = (1,1,1)$. La famille $(\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{u})$ est génératrice de $\mathbb{R}^3$ puisque contient la famille $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ qui est déjà génératrice. Mais $\vec{e}_1 + \vec{e}_2 + \vec{e}_3 - \vec{u} = \vec{0}$. Donc $(\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{u})$ est liée et $(\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{u})$ n'est pas une base.
4. Plus généralement, la famille $(\vec{e}_1, \cdots, \vec{e}_n)$ de vecteurs de $\mathbb{K}^n$ définie par
$$\vec{e}_i \overset{\text{déf}}{=} (0, \cdots, 0, \underbrace{1}_{\text{position } i}, 0, \cdots, 0)$$
est une base de $\mathbb{K}^n$. On l'appelle **base canonique** de $\mathbb{K}^n$.

Théorème 3.2.15 *Soit X un sous-espace vectoriel de $\mathbb{K}^n$, et $(\vec{x}_1, \cdots, \vec{x}_k)$ une famille de vecteurs de X. Les trois propriétés suivantes sont équivalentes :*

i) $(\vec{x}_1, \cdots, \vec{x}_k)$ est une famille libre maximale (i.e si on ajoute un ou plusieurs vecteurs à la famille, on obtient une famille liée),

ii) $(\vec{x}_1, \cdots, \vec{x}_k)$ est une famille génératrice minimale (i.e si on retire un ou plusieurs vecteurs à la famille, la famille obtenue n'est plus génératrice de X),

iii) $(\vec{x}_1, \cdots, \vec{x}_k)$ est une base.

Preuve : En vertu de la proposition 3.2.13, il suffit de traiter les cas où l'on ajoute ou retire un seul vecteur.

i) $\Rightarrow$ *iii)* : Soit $(\vec{x}_1, \cdots, \vec{x}_k)$ une famille libre maximale. Montrons que cette famille est aussi génératrice.

Soit $\vec{y} \in X$. Alors par hypothèse, la famille $(\vec{x}_1, \cdots, \vec{x}_k, \vec{y})$ est liée. La proposition 3.2.11 assure donc que $\vec{y} \in \text{Vect}\,(\vec{x}_1, \cdots, \vec{x}_k)$.

iii) $\Rightarrow$ *i)* : Soit $(\vec{x}_1, \cdots, \vec{x}_k)$ une base de X. Par définition, cette famille est donc *libre*. De plus, si $\vec{y} \in X$ alors $\vec{y} \in \text{Vect}\,(\vec{x}_1, \cdots, \vec{x}_k)$ puisque $(\vec{x}_1, \cdots, \vec{x}_k)$ est aussi génératrice. D'après la proposition 3.2.11, on conclut donc que $\text{Vect}\,(\vec{x}_1, \cdots, \vec{x}_k, \vec{y})$ est liée. Donc $(\vec{x}_1, \cdots, \vec{x}_k)$ est libre *maximale*.

ii) $\Rightarrow$ *iii)* : Soit $(\vec{x}_1, \cdots, \vec{x}_k)$ génératrice minimale. Supposons *par l'absurde* que cette famille ne soit pas une base. Alors elle est liée, c'est-à-dire que l'un de ses vecteurs –disons $\vec{x}_k$ pour fixer les idées– est combinaison linéaire des autres :

$$\exists (\alpha_1, \cdots, \alpha_{k-1}) \in \mathbb{K}^{k-1},\ \vec{x}_k = \sum_{i=1}^{k-1} \alpha_i \vec{x}_i. \tag{3.2}$$

Soit maintenant $\vec{y} \in X$ arbitraire. Comme $(\vec{x}_1, \cdots, \vec{x}_k)$ est génératrice, il existe un k-uplet $(\lambda_1, \cdots, \lambda_k)$ d'éléments de $\mathbb{K}^k$ tel que

$$\vec{y} = \sum_{i=1}^{k} \lambda_i \vec{x}_i.$$

En tenant compte de (3.2), on trouve

$$\vec{y} = \sum_{i=1}^{k-1} (\lambda_i + \alpha_i \lambda_k) \vec{x}_i.$$

En conséquence $(\vec{x}_1, \cdots, \vec{x}_{k-1})$ est génératrice, ce qui contredit l'hypothèse "$(\vec{x}_1, \cdots, \vec{x}_k)$ génératrice minimale".

iii) $\Rightarrow$ *ii)* : Soit $(\vec{x}_1, \cdots, \vec{x}_k)$ une base de X. Alors $(\vec{x}_1, \cdots, \vec{x}_k)$ est génératrice. Retirons un vecteur à cette famille, par exemple $\vec{x}_k$. La proposition 3.2.11 montre que $\vec{x}_k$ ne peut être engendré par la famille $(\vec{x}_1, \cdots, \vec{x}_{k-1})$ car alors $(\vec{x}_1, \cdots, \vec{x}_k)$ serait liée. En conséquence, $(\vec{x}_1, \cdots, \vec{x}_k)$ est bien génératrice minimale. ■

Théorème 3.2.16 (de la base incomplète) *Soit X un s.e.v de $\mathbb{K}^n$ non réduit à $\{\vec{0}\}$, $m \in \mathbb{N}^*$ et $p \in \{0, \cdots, m\}$. Soit $(\vec{x}_1, \cdots, \vec{x}_m)$ une famille génératrice de X telle que $(\vec{x}_1, \cdots, \vec{x}_p)$ soit libre[2]. Alors il existe un entier $k \in \{0, \cdots, m-p\}$ et k indices $(i_1, \cdots, i_k)$ vérifiant $p < i_1 < \cdots < i_k \leq m$ et tels que $(\vec{x}_1, \cdots, \vec{x}_m, \vec{x}_{i_1}, \cdots, \vec{x}_{i_k})$ soit une base de X.*

Preuve : (hors-programme)

Soit $(\vec{x}_1, \cdots, \vec{x}_m)$ une famille génératrice de X dont les p premiers vecteurs constituent une famille libre. Considérons l'ensemble

$$\mathcal{E} = \{(\vec{x}_1, \cdots, \vec{x}_p, \vec{x}_{j_1}, \cdots, \vec{x}_{j_\ell}) \mid 0 \leq \ell \leq m-p,\ p < j_1 < \cdots < j_\ell \leq m \text{ et } (\vec{x}_1, \cdots, \vec{x}_{j_\ell}) \text{ libre}\}.$$

Soit $\mathcal{C}$ l'ensemble des cardinaux des familles de $\mathcal{E}$.

L'ensemble $\mathcal{E}$ contient au moins un élément : $(\vec{x}_1, \cdots, \vec{x}_p)$. Donc $\mathcal{C}$ est un sous-ensemble non vide de $\mathbb{N}$. Par construction, m est un majorant de $\mathcal{C}$. Donc $\mathcal{C}$ admet un élément maximal que l'on peut toujours noter $p+k$. (On a donc bien $0 \leq k \leq m-p$).

Soit $(\vec{x}_1, \cdots, \vec{x}_p, \vec{x}_{j_1}, \cdots, \vec{x}_{j_\ell})$ un élément de $\mathcal{E}$ de cardinal maximal. Cette famille est libre et maximale par construction. D'après la proposition 3.2.15, c'est une base. ■

Remarque : Autrement dit, on peut compléter toute famille libre $(\vec{x}_1, \cdots, \vec{x}_p)$ de X en une base de X en lui adjoignant des vecteurs bien choisis de $(\vec{x}_{p+1}, \cdots, \vec{x}_m)$.

3.3 Rang et dimension

3.3.1 Dimension d'un sous-espace vectoriel

Lemme 3.3.1 *Dans $\mathbb{K}^n$ une famille comportant $n+1$ vecteurs est toujours liée.*

Plus généralement, si X est un sous-espace vectoriel de $\mathbb{K}^n$ engendré par une famille de p vecteurs alors toute famille de vecteurs de X comportant au moins $p+1$ éléments est liée.

[2] Avec la convention que $(\vec{x}_1, \cdots, \vec{x}_p)$ est la famille vide si $p = 0$.

Preuve : Ce résultat sera prouvé au second semestre dans un cadre plus général. La preuve est hors-programme mais nous en donnons les étapes essentielles afin de satisfaire la curiosité du lecteur. On fait une récurrence sur la dimension n, l'hypothèse de récurrence étant :

$(\mathcal{P}_n)$ Dans $\mathbb{K}^n$, toute famille de $n+1$ vecteurs est liée.

• **Montrons que $(\mathcal{P}_1)$ est vraie :** Si α et β sont deux éléments de $\mathbb{K}^1$, c'est-à-dire de $\mathbb{K}$. Il est clair qu'il existe $(\lambda,\mu)\in\mathbb{K}^2$ non nul tel que $\lambda\alpha+\mu\beta=0$: si $\alpha\neq 0$, on peut choisir par exemple $\lambda=-\beta/\alpha$ et $\mu=1$; si $\alpha=0$, le couple $(1,0)$ fait l'affaire.

• **Supposons $(\mathcal{P}_\mathbf{n})$ vraie et démontrons $(\mathcal{P}_\mathbf{n+1})$:** Soit donc $(\vec{x}_1,\cdots,\vec{x}_{n+2})$ une famille de $n+2$ vecteurs de $\mathbb{K}^{n+1}$. Notons $(x_i^1,\cdots,x_i^{n+1})$ les coordonnées de $\vec{x}_i$ par rapport à la base canonique de $\mathbb{K}^{n+1}$. Établir l'existence d'un $(n+2)$-uplet $(\lambda_1,\cdots,\lambda_{n+2})$ non nul de $\mathbb{K}^{n+2}$ tel que $\sum_{i=1}^{n+2}\lambda_i\vec{x}_i=0$, revient à montrer que le système linéaire (S) suivant admet une solution $(\lambda_1,\cdots,\lambda_{n+2})$ non nulle :

$$(S)\qquad \begin{cases} x_1^1\lambda_1+\cdots+x_{n+2}^1\lambda_{n+2}=0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ x_1^{n+1}\lambda_1+\cdots+x_{n+2}^{n+1}\lambda_{n+2}=0. \end{cases}$$

1er cas : $x_1^{n+1}=\cdots=x_{n+2}^{n+1}=0$.

Dans ce cas, la dernière ligne du système est vérifiée par n'importe quel $(n+2)$-uplet $(\lambda_1,\cdots,\lambda_{n+2})$. Si l'on pose $\vec{x}_i'\overset{\text{déf}}{=}(x_i^1,\cdots,x_i^n)$, la famille $(\vec{x}_1',\cdots,\vec{x}_{n+1}')$ est une famille de $n+1$ vecteurs de $\mathbb{K}^n$. En vertu de l'hypothèse de récurrence, elle est liée. Donc $(\vec{x}_1',\cdots,\vec{x}_{n+2}')$ aussi (cf prop. 3.2.13). On peut donc trouver $(\lambda_1,\cdots,\lambda_{n+2})\in\mathbb{K}^{n+2}$ non nul tel que les n premières lignes de (S) soient aussi satisfaites.

2ème cas : Les coefficients de la dernière ligne de (S) ne sont pas tous nuls.

Quitte à changer l'ordre des vecteurs de la famille et à multiplier la dernière ligne de (S) par un scalaire non nul, on peut supposer que $x_{n+2}^{n+1}=1$. On a donc

$$(S)\iff\begin{cases} x_1^1\lambda_1+\cdots+x_{n+1}^1\lambda_{n+1}+x_{n+2}^1\lambda_{n+2}=0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ x_1^n\lambda_1+\cdots+x_{n+1}^n\lambda_{n+1}+x_{n+2}^n\lambda_{n+2}=0,\\ \lambda_{n+2}=-\sum_{i=1}^{n+1}x_i^{n+1}\lambda_i. \end{cases}$$

Pour $i\in\{1,\cdots,n+1\}$, définissons

$$\vec{x}_i'\overset{\text{déf}}{=}(x_i^1-x_{n+2}^1x_i^{n+1},\ \cdots,\ x_i^n-x_{n+2}^nx_i^{n+1}).$$

Le système (S) donne $\sum_{i=1}^{n+1}\lambda_i\vec{x}_i'=\vec{0}$. D'après $(\mathcal{P}_n)$, la famille $(\vec{x}_1',\cdots,\vec{x}_{n+1}')$ de $\mathbb{K}^n$ est liée. Donc il existe $(\lambda_1,\cdots,\lambda_{n+1})$ non nul tel que $\sum_{i=1}^{n+1}\lambda_i\vec{x}_i'=\vec{0}$. En définissant λ_{n+2} conformément à la dernière ligne du système ci-dessus, on obtient $\sum_{i=1}^{n+2}\lambda_i\vec{x}_i=\vec{0}$.

■

Proposition 3.3.2 *Soit X un sous-espace vectoriel de $\mathbb{K}^n$ non réduit à $\{\vec{0}\}$. Alors X admet une base composée d'au plus n vecteurs. De plus, toutes les bases de X comportent le même nombre k d'éléments. Ce nombre est appelé* **dimension** *de X. On le note* $\dim X$.

Preuve : La preuve de l'existence d'une base se fait par récurrence limitée. Par hypothèse, X contient un vecteur $\vec{x}_1\neq 0$. On choisit alors un autre vecteur $\vec{x}_2$ de X tel que $(\vec{x}_1,\vec{x}_2)$ soit une base. Si un tel vecteur n'existe pas, $(\vec{x}_1)$ est une famille libre maximale et donc une base (cf th. 3.2.15).

Plus généralement, supposons connue une famille libre $(\vec{x}_1, \cdots, \vec{x}_{j-1})$ de X. Deux cas peuvent se présenter. Ou bien $(\vec{x}_1, \cdots, \vec{x}_{j-1})$ est *libre maximale* auquel cas le th. 3.2.15 assure que $(\vec{x}_1, \cdots, \vec{x}_{j-1})$ est une base de X, ou bien cette famille libre *n'est pas maximale*, auquel cas on peut trouver $\vec{x}_j \in X$ tel que $(\vec{x}_1, \cdots, \vec{x}_j)$ soit libre.

Enfin, en vertu du lemme 3.3.1, le procédé de construction s'arrête au plus tard après l'obtention d'une famille libre à n éléments.

Reste à vérifier que toutes les bases ont le même nombre d'éléments. Donnons-nous deux bases $(\vec{x}_1, \cdots, \vec{x}_k)$ et $(\vec{y}_1, \cdots, \vec{y}_\ell)$ de X. On a

$$X = \text{Vect}\,(\vec{x}_1, \cdots, \vec{x}_k) = \text{Vect}\,(\vec{y}_1, \cdots, \vec{y}_\ell).$$

La première égalité montre que X est engendré par k vecteurs. D'après le lemme 3.3.1, toute famille de $k+1$ vecteurs de X est donc liée. Comme $(\vec{y}_1, \cdots, \vec{y}_\ell)$ est libre, on a donc $\ell \leq k$. En échangeant les rôles des deux familles, on obtient $k \leq \ell$. ■

Remarque 3.3.3 *1. Par convention, le sous-espace vectoriel $\{\vec{0}\}$ a pour dimension 0.*

2. La base canonique de $\mathbb{K}^n$ comporte exactement n éléments. Donc $\dim \mathbb{K}^n = n$.

3. Les sous-espaces vectoriels de dimension 1 sont appelés **droites vectorielles** *ou plus simplement* **droites**.

4. Les sous-espaces de dimension $n-1$ de $\mathbb{K}^n$ sont appelés **hyperplans**. *Lorsque $n = 3$, on parle plutôt de* **plan**. *Enfin, si $n = 2$, les hyperplans sont des droites vectorielles.*

Remarque : Sachant que X est un sous-espace vectoriel de dimension k, pour montrer qu'une famille $(\vec{x}_1, \cdots, \vec{x}_k)$ de vecteurs de X est une base, il suffit d'établir que $(\vec{x}_1, \cdots, \vec{x}_k)$ est génératrice ou que $(\vec{x}_1, \cdots, \vec{x}_k)$ est libre.

Exemples :

1. Équation d'une droite de $\mathbb{R}^2$:
   Soit $(\alpha, \beta) \in \mathbb{R}^2 \backslash (0,0)$. La droite (vectorielle) engendrée par le vecteur $\vec{u} = (\alpha, \beta)$ de $\mathbb{R}^2$ est l'ensemble des vecteurs $\vec{v} = (x, y)$ tels que $\beta x - \alpha y = 0$.
   Réciproquement, si $(a, b) \neq (0, 0)$, l'ensemble des vecteurs $\vec{v} = (x, y)$ de $\mathbb{R}^2$ tels que $ax + by = 0$ est la droite (vectorielle) de $\mathbb{R}^2$ orthogonale au vecteur $\vec{u} = (a, b)$.
2. Équation d'un plan de $\mathbb{R}^3$:
   Soit (a, b, c) un triplet non nul (i.e distinct de $(0, 0, 0)$) de $\mathbb{R}^3$. L'ensemble des vecteurs $\vec{v} = (x, y, z)$ de $\mathbb{R}^3$ tels que
   $$ax + by + cz = 0$$
   est un plan (vectoriel) de $\mathbb{R}^3$. C'est le plan orthogonal au vecteur $\vec{u} = (a, b, c)$.
3. Équation d'un hyperplan de $\mathbb{R}^n$:
   Plus généralement, si $(a_1, \cdots, a_n)$ est un n-uplet non nul de $\mathbb{R}^n$, alors l'ensemble des vecteurs $\vec{v} = (x_1, \cdots, x_n)$ tels que
   $$a_1 x_1 + \cdots + a_n x_n = 0$$
   est un hyperplan de $\mathbb{R}^n$.

Proposition 3.3.4 *Soit X et Y deux sous-espaces vectoriels de $\mathbb{K}^n$ tels que $X \subset Y$. Alors $\dim X \leq \dim Y$ et $\dim X = \dim Y$ si et seulement si $X = Y$.*

Preuve : Notons $k = \dim X$ et $\ell = \dim Y$. Il est clair que toute base de X est une famille libre de Y. Le théorème de la base incomplète assure donc que $k \leq \ell$.

Si $X = Y$, il est trivial que $\dim X = \dim Y$. Réciproquement, supposons que $\dim X = \dim Y$ et $X \subset Y$. Soit $(\vec{f_1}, \cdots, \vec{f_k})$ une base de X. Alors c'est aussi une famille libre de Y. Elle est maximale car $\dim Y = k$. Donc c'est une base de Y. On a donc

$$Y = X = \text{Vect}\,(\vec{f_1}, \cdots, \vec{f_k}).$$

■

3.3.2 Rang d'une famille de vecteurs

Définition 3.3.5 *Soit $(\vec{x}_1, \cdots, \vec{x}_p)$ une famille de vecteurs de $\mathbb{K}^n$. On appelle* **rang de la famille** *$(\vec{x}_1, \cdots, \vec{x}_p)$, noté $\text{rg}\,(\vec{x}_1, \cdots, \vec{x}_p)$ la dimension du sous-espace vectoriel $\text{Vect}\,(\vec{x}_1, \cdots, \vec{x}_p)$.*

Proposition 3.3.6 *Soit $(\vec{x}_1, \cdots, \vec{x}_p)$ une famille de vecteurs de $\mathbb{K}^n$.*

i) On a toujours $\text{rg}\,(\vec{x}_1, \cdots, \vec{x}_p) \leq p$.

ii) On a $\text{rg}\,(\vec{x}_1, \cdots, \vec{x}_p) \leq n$ avec égalité si et seulement si $(\vec{x}_1, \cdots, \vec{x}_p)$ engendre $\mathbb{K}^n$.

iii) On a $\text{rg}\,(\vec{x}_1, \cdots, \vec{x}_p) = p$ si et seulement si la famille $(\vec{x}_1, \cdots, \vec{x}_p)$ est libre.

Preuve : Notons $X = \text{Vect}\,(\vec{x}_1, \cdots, \vec{x}_p)$. D'après le théorème de la base incomplète appliqué à la famille vide et à la famille génératrice $(\vec{x}_1, \cdots, \vec{x}_p)$, il existe une sous-famille de $(\vec{x}_1, \cdots, \vec{x}_p)$ qui est une base de X. Cette famille a au plus p éléments, ce qui montre le *i)*.

Comme $\text{Vect}\,(\vec{x}_1, \cdots, \vec{x}_p) \subset \mathbb{K}^n$, le *ii)* de la proposition résulte de la proposition 3.3.4.

Par définition même de X, la famille $(\vec{x}_1, \cdots, \vec{x}_p)$ est génératrice. Si de plus $(\vec{x}_1, \cdots, \vec{x}_p)$ est libre alors $(\vec{x}_1, \cdots, \vec{x}_p)$ est une base de X, et donc $\dim X = p$.

Si au contraire $(\vec{x}_1, \cdots, \vec{x}_p)$ est liée, alors l'un des vecteurs, par exemple $\vec{x}_p$, est combinaison linéaire des autres. Donc $X = \text{Vect}\,(\vec{x}_1, \cdots, \vec{x}_{p-1})$, et donc d'après *i)*, $\dim X \leq p - 1$. ■

Pour déterminer le rang d'une famille de vecteurs, on a souvent recours au résultat suivant :

Proposition 3.3.7 *Le rang d'une famille de vecteurs est invariant par les transformations élémentaires $(T1)$, $(T2)$ et $(T3)$ définies dans la proposition* 3.2.8.

Preuve : C'est une conséquence immédiate de la proposition 3.2.8. ■

3.3.3 Rang d'une matrice

Définition 3.3.8 *Soit $A \in \mathcal{M}_{p,n}(\mathbb{K})$. On appelle rang de la matrice A (noté $\text{rg}\,A$) le rang de la famille de vecteurs constituée des p vecteurs lignes de A. Autrement dit, si l'on note $\vec{a}_i$ la i-ème ligne de A, on a*

$$\text{rg}\,A = \text{rg}\,(\vec{a}_1, \cdots, \vec{a}_p).$$

Proposition 3.3.9 *[Rang d'une matrice échelonnée] Soit A une matrice échelonnée du type suivant :*

$$A = \begin{pmatrix} 0 & a_{1j_1} & * & * & * & * \\ 0 & 0 & a_{2j_2} & * & * & * \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & a_{kj_k} & * \\ 0 & \cdots & \cdots & \cdots & 0 & 0 \end{pmatrix}.$$

Alors $\text{rg}\,A = k$.

Preuve : Comme les $p-k$ dernières lignes de A sont nulles, le rang de A est égal à celui de ses k premières lignes. Par définition des indices j_i, on a $a_{ij_i} \neq 0$ pour $i \in \{1, \cdots, k\}$, donc la famille constituée par les k premières lignes est libre (pour le voir, revenir à la définition d'une famille libre). On a donc bien $\operatorname{rg} A = k$. ■

3.3.4 Calcul pratique du rang d'une famille de vecteurs

Considérons une matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$. Nous avons vu dans le chapitre 2 qu'une succession de transformations élémentaires permettait d'obtenir une matrice B échelonnée (méthode du pivot de Gauss). La proposition 3.3.7 assure que $\operatorname{rg} A = \operatorname{rg} B$. De plus, le rang de B peut être facilement calculé grâce à la proposition 3.3.9.

Ces considérations nous suggèrent une méthode systématique permettant de calculer le rang d'une famille de vecteurs :

Comment calculer le rang d'une famille de vecteurs

Pour calculer le rang d'une famille de vecteurs $(\vec{x}_1, \cdots, \vec{x}_p)$ de $\mathbb{K}^n$, on procède comme suit :

1. On dispose les p vecteurs en ligne. On obtient ainsi une matrice à p lignes et n colonnes :
$$A = \begin{pmatrix} \vec{x}_1 \\ \vdots \\ \vec{x}_p \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{p1} & \cdots & x_{pn} \end{pmatrix}.$$
2. On applique aux lignes de la matrice A des transformations élémentaires $(T1)$, $(T2)$ ou $(T3)$ afin d'obtenir une matrice B échelonnée et de même rang que A. Pour ce faire, la méthode du pivot de Gauss est tout indiquée.
Le **rang de** A est alors égal au **nombre de lignes non nulles de** B.

Exemple : Calculer le rang de la famille $(\vec{V}_1, \vec{V}_2, \vec{V}_3, \vec{V}_4)$ composée des vecteurs de $\mathbb{R}^5$ suivants :

$$\vec{V}_1 = (1,0,0,2,-1), \quad \vec{V}_2 = (0,1,-2,1,0), \quad \vec{V}_3 = (0,-1,2,1,-1), \quad \vec{V}_4 = (0,0,0,2,-1).$$

On écrit la matrice A de lignes $\vec{V}_1$, $\vec{V}_2$, $\vec{V}_3$ et $\vec{V}_4$:

$$\begin{pmatrix} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & -1 & 2 & 1 & -1 \\ 0 & 0 & 0 & 2 & -1 \end{pmatrix}$$

puis on applique l'algorithme du pivot de Gauss afin de transformer A en une matrice échelonnée. En ajoutant (L_2) à (L_3), puis en retranchant (L_3) à (L_4), on trouve

$$\operatorname{rg} A = \operatorname{rg} \begin{pmatrix} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 2 & -1 \end{pmatrix} = \operatorname{rg} \begin{pmatrix} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Cette dernière matrice est échelonnée avec trois lignes non nulles. Donc

$$\boxed{\operatorname{rg}(\vec{V}_1, \vec{V}_2, \vec{V}_3, \vec{V}_4) = \operatorname{rg} A = 3.}$$

Chapitre 4

Déterminants

A ce stade, la méthode du pivot de Gauss est le seul outil dont nous disposons pour déterminer si une famille de vecteurs est libre ou liée (en fait, cette méthode donne un renseignement un peu plus précis, à savoir le rang de la famille en question). Dans ce chapitre, nous présentons un critère alternatif permettant de déterminer si une famille est libre ou liée : le calcul du déterminant.

Cette notion de déterminant ne vous est probablement pas étrangère dans le cas des vecteurs de $\mathbb{R}^2$. Vous savez sans doute que deux vecteurs $\vec{a} = (x_1, y_1)$ et $\vec{b} = (x_2, y_2)$ de $\mathbb{R}^2$ sont colinéaires[1] si et seulement si

$$\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \overset{\text{déf}}{=} x_1 y_2 - x_2 y_1 = 0.$$

Nous allons généraliser la définition du déterminant au cas de n vecteurs de $\mathbb{K}^n$ et disposerons ainsi d'un critère permettant de savoir si ces n vecteurs sont liés ou libres.

4.1 Définition du déterminant

Proposition 4.1.1 *Le déterminant est l'unique application qui à une matrice carrée A de $\mathcal{M}_n(\mathbb{K})$ associe un scalaire de $\mathbb{K}$ noté $\det A$, et vérifie les propriétés suivantes :*

1. *Linéarité par rapport à chaque ligne :*

$$\det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_i + \vec{a}'_i \\ \vdots \\ \vec{a}_n \end{pmatrix} = \det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_n \end{pmatrix} + \det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}'_i \\ \vdots \\ \vec{a}_n \end{pmatrix} \leftarrow \text{ligne } i$$

$$\det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \lambda \vec{a}_i \\ \vdots \\ \vec{a}_n \end{pmatrix} = \lambda \det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_n \end{pmatrix} \leftarrow \text{ligne } i$$

2. *Caractère alterné : Si la matrice A a deux lignes identiques alors $\det A = 0$.*
3. *Normalisation : $\det I_n = 1$.*

[1] ou liés puisqu'il s'agit de deux vecteurs

Proposition 4.1.2 *L'application* det *est de plus antisymétrique, c'est-à-dire que*

$$\det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_j \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_n \end{pmatrix} \begin{matrix} \\ \\ \leftarrow \text{ligne } i \\ \\ \leftarrow \text{ligne } j \\ \\ \\ \end{matrix} = -\det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_j \\ \vdots \\ \vec{a}_n \end{pmatrix} \begin{matrix} \\ \\ \leftarrow \text{ligne } i \\ \\ \leftarrow \text{ligne } j \\ \\ \\ \end{matrix}$$

Réciproquement, toute application de $\mathcal{M}_n(\mathbb{K})$ *dans* $\mathbb{K}$ *linéaire par rapport à chaque ligne et antisymétrique, est alternée.*

Preuve : Comme l'application det est linéaire et alternée par rapport à chaque ligne, on a

$$\begin{aligned}
0 = \det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_i + \vec{a}_j \\ \vdots \\ \vec{a}_i + \vec{a}_j \\ \vdots \\ \vec{a}_n \end{pmatrix} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} &= \det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_i + \vec{a}_j \\ \vdots \\ \vec{a}_n \end{pmatrix} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} + \det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_j \\ \vdots \\ \vec{a}_i + \vec{a}_j \\ \vdots \\ \vec{a}_n \end{pmatrix} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} \\
&= \det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_n \end{pmatrix} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} + \det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_j \\ \vdots \\ \vec{a}_n \end{pmatrix} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} + \det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_j \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_n \end{pmatrix} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} + \det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_j \\ \vdots \\ \vec{a}_j \\ \vdots \\ \vec{a}_n \end{pmatrix} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} \\
&= 0 + \det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_j \\ \vdots \\ \vec{a}_n \end{pmatrix} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} + \det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_j \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_n \end{pmatrix} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} + 0,
\end{aligned}$$

d'où l'antisymétrie.

Réciproquement, si $F \in \mathcal{F}(\mathcal{M}_n(\mathbb{K}); \mathbb{K})$ est une application antisymétrique, on a par définition

$$F \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_j \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_n \end{pmatrix} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} = -F \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_j \\ \vdots \\ \vec{a}_n \end{pmatrix} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix}$$

Ces deux termes sont donc nuls quand $\vec{a}_i = \vec{a}_j$. ■

Idée de la preuve de l'existence de l'application déterminant :

Notons $(\vec{e}_1, \cdots, \vec{e}_n)$ la base canonique de $\mathbb{K}^n$. Tout vecteur $\vec{a}_i$ de $\mathbb{K}^n$ peut se décomposer en

$$\vec{a}_i = \sum_{j=1}^{n} a_{ij}\vec{e}_j.$$

En utilisant la linéarité par rapport à chaque ligne, on en déduit que

$$\det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_n \end{pmatrix} = \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} a_{1j_1} \cdots a_{nj_n} \det \begin{pmatrix} \vec{e}_{j_1} \\ \vdots \\ \vec{e}_{j_n} \end{pmatrix}.$$

On en déduit que l'application det est uniquement déterminée par sa valeur sur les vecteurs de la base canonique. Grâce au caractère alterné, on sait de plus que l'image d'une famille contenant *deux fois* le même vecteur de la base canonique par l'application det est nul. Par antisymétrie, on constate de plus que lorsque les $\vec{e}_{j_i}$ sont choisis deux à deux distincts, $\det \begin{pmatrix} \vec{e}_{j_1} \\ \vdots \\ \vec{e}_{j_n} \end{pmatrix}$ peut être calculé à partir de $\det \begin{pmatrix} \vec{e}_1 \\ \vdots \\ \vec{e}_n \end{pmatrix}$ en effectuant un nombre fini de permutations de lignes. Donc la valeur de $\det I_n$ (qui est prise égale à 1) détermine l'application det de façon unique. ■

Notation : Soit $A = (a_{ij})_{1 \leq i,j \leq n}$ une matrice. Le déterminant de A est souvent noté

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

4.2 Propriétés élémentaires du déterminant

Proposition 4.2.1 *Pour tout $\lambda \in \mathbb{K}$ et $A \in \mathcal{M}_n(\mathbb{K})$, on a* $\boxed{\det(\lambda A) = \lambda^n \det A.}$

Preuve : En notant $\vec{a}_1, \cdots, \vec{a}_n$ les lignes de A, on a

$$\det(\lambda A) = \det \begin{pmatrix} \lambda\vec{a}_1 \\ \lambda\vec{a}_2 \\ \lambda\vec{a}_3 \\ \vdots \\ \lambda\vec{a}_n \end{pmatrix} = \lambda \det \begin{pmatrix} \vec{a}_1 \\ \lambda\vec{a}_2 \\ \lambda\vec{a}_3 \\ \vdots \\ \lambda\vec{a}_n \end{pmatrix} = \lambda^2 \det \begin{pmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \lambda\vec{a}_3 \\ \vdots \\ \lambda\vec{a}_n \end{pmatrix} = \cdots = \lambda^n \det \begin{pmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vec{a}_3 \\ \vdots \\ \vec{a}_n \end{pmatrix}.$$

■

Proposition 4.2.2 *Si l'on ajoute à une ligne une combinaison linéaire des autres lignes, alors le déterminant ne change pas.*

Preuve : Supposons que l'on ajoute $\sum_{j\neq i}\lambda_j\vec{a}_j$ à la ligne i. En utilisant la propriété de linéarité par rapport aux lignes, on a alors :

$$\det\begin{pmatrix}\vec{a}_1\\ \vdots\\ \vec{a}_i+\sum_{j\neq i}\vec{a}_j\\ \vdots\\ \vec{a}_n\end{pmatrix}=\det\begin{pmatrix}\vec{a}_1\\ \vdots\\ \vec{a}_i\\ \vdots\\ \vec{a}_n\end{pmatrix}+\sum_{j\neq i}\lambda_j\det\begin{pmatrix}\vec{a}_1\\ \vdots\\ \vec{a}_j\\ \vdots\\ \vec{a}_n\end{pmatrix}\leftarrow\text{ligne } i$$

Les $n-1$ derniers termes de l'inégalité de droite sont des déterminants de matrices ayant deux lignes identiques. En vertu du caractère alterné de l'application det, ils sont nuls, d'où le résultat. ■

Proposition 4.2.3 *Si l'une des lignes de A est nulle alors* $\det A=0$.

Preuve : Il suffit d'écrire :

$$\det\begin{pmatrix}\vec{a}_1\\ \vdots\\ \vec{0}\\ \vdots\\ \vec{a}_n\end{pmatrix}=\det\begin{pmatrix}\vec{a}_1\\ \vdots\\ 0\cdot\vec{0}\\ \vdots\\ \vec{a}_n\end{pmatrix}=0\cdot\det\begin{pmatrix}\vec{a}_1\\ \vdots\\ \vec{0}\\ \vdots\\ \vec{a}_n\end{pmatrix}=0.$$

■

Proposition 4.2.4 *Si A est une matrice diagonale alors* $\det A$ *est le produit des termes diagonaux. De même, si A est une matrice triangulaire supérieure alors* $\det A$ *est le produit des termes diagonaux.*

Preuve :

1. *Cas où A est diagonale.* On a alors par linéarité

$$\begin{aligned}\det A &= \begin{vmatrix}a_{11}&0&\cdots&0\\0&a_{22}&\ddots&\vdots\\\vdots&\ddots&\ddots&0\\0&\cdots&0&a_{nn}\end{vmatrix}\\ &= a_{11}\begin{vmatrix}1&0&\cdots&0\\0&a_{22}&\ddots&\vdots\\\vdots&\ddots&\ddots&0\\0&\cdots&0&a_{nn}\end{vmatrix}\\ &= a_{11}a_{22}\begin{vmatrix}1&0&\cdots&0\\0&1&\ddots&\vdots\\\vdots&\ddots&\ddots&0\\0&\cdots&0&a_{nn}\end{vmatrix}\\ &= \cdots=\left(\textstyle\prod_{i=1}^n a_{ii}\right)\begin{vmatrix}1&0&\cdots&0\\0&1&\ddots&\vdots\\\vdots&\ddots&\ddots&0\\0&\cdots&0&1\end{vmatrix}=\textstyle\prod_{i=1}^n a_{ii}.\end{aligned}$$

2. *Cas où A est triangulaire supérieure à diagonale non nulle.* En reprenant le calcul précédent, on obtient[2]

$$\det A = \left(\prod_{i=1}^{n} a_{ii}\right) \begin{vmatrix} 1 & * & \cdots & * \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & 1 \end{vmatrix}.$$

Il suffit donc de montrer que le déterminant de toute matrice triangulaire supérieure avec des 1 sur la diagonale est égal à 1.

Considérons donc une matrice A de ce type et notons (L_1), $\cdots$, (L_n) ses lignes. En retranchant successivement $a_{1n}(L_n)$ à (L_1), $a_{2n}(L_n)$ à (L_2), jusqu'à $a_{1n-1}(L_{n-1})$ à (L_{n-1}) (opérations qui ne changent pas le déterminant), on fait "apparaître des 0" sur la dernière colonne et l'on trouve :

$$\begin{vmatrix} 1 & * & \cdots & * \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & * & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{vmatrix}.$$

L'étape suivante consiste à retrancher $a_{j\,n-1}(L_{n-1})$ à la j-ième ligne afin de faire 'apparaître des 0 dans l'avant dernière colonne. En itérant le procédé, on conclut finalement que le déterminant cherché est égal à celui de l'identité, donc à 1.

3. *Cas où A est triangulaire supérieure avec au moins un terme diagonal nul.* Notons i l'indice du plus grand coefficient diagonal nul de A. En utilisant la linéarité de l'application det par rapport aux $n-i$ dernières lignes, on montre aisément que

$$\det A = a_{i+1i+1} \times \cdots \times a_{nn} \det \begin{pmatrix} a_{11} & * & * & & * & \cdots & * \\ 0 & \ddots & & & \vdots & & \vdots \\ \vdots & \ddots & a_{i-1i-1} & & \vdots & & \\ \vdots & & \ddots & 0 & * & & \\ \vdots & & & \ddots & 1 & \ddots & \\ \vdots & & & & \ddots & \ddots & * \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

Ensuite, on procède comme dans l'étape 2 pour faire apparaître des 0 dans les $n-i$ dernières colonnes. On obtient ainsi

$$\det A = a_{i+1i+1} \times \cdots \times a_{nn} \det \begin{pmatrix} a_{11} & * & * & & 0 & \cdots & 0 \\ 0 & \ddots & & & \vdots & & \vdots \\ \vdots & \ddots & a_{i-1i-1} & & \vdots & & \\ \vdots & & \ddots & 0 & 0 & & \\ \vdots & & & \ddots & 1 & \ddots & \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix},$$

[2] Les $*$ désignent des coefficients dont la valeur exacte n'intervient pas dans les calculs

et l'on conclut immédiatement que le déterminant vaut 0 puisque la i-ième ligne de la matrice ci-dessus est nulle.

■

Proposition 4.2.5 *Soit $(\vec{a}_1, \cdots, \vec{a}_n)$ une famille de vecteurs de $\mathbb{K}^n$.*
La famille $(\vec{a}_1, \cdots, \vec{a}_n)$ est <u>liée</u> si et seulement si $\det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_n \end{pmatrix} = 0$.

Preuve :

$\Longrightarrow$ Supposons $(\vec{a}_1, \cdots, \vec{a}_n)$ liée. Alors l'un des vecteurs de la famille est combinaison linéaire des autres, par exemple, $\vec{a}_n = \sum_{j=1}^{n-1} \lambda_j \vec{a}_j$. En appliquant la proposition 4.2.2, on a donc

$$\det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_{n-1} \\ \vec{a}_n \end{pmatrix} = \det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_{n-1} \\ \vec{0} + \sum_{j=1}^{n-1} \lambda_j \vec{a}_j \end{pmatrix} = \det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_{n-1} \\ \vec{0} \end{pmatrix} = 0.$$

$\Longleftarrow$ On montre l'implication inverse par contraposition.

Supposons donc que $(\vec{a}_1, \cdots, \vec{a}_n)$ soit libre et notons A la matrice formée par cette famille de vecteurs. L'algorithme du pivot de Gauss permet d'obtenir une matrice échelonnée B à partir de A en effectuant des permutations de lignes ou en ajoutant des combinaisons linéaires d'autres lignes. En vertu du caractère alterné du déterminant, et de la proposition 4.2.2, on a donc $|\det A| = |\det B|$. Mais en notant $\vec{b}_i$ la i-ème ligne de B, on a aussi $\text{Vect}\,(\vec{a}_1, \cdots \vec{a}_n) = \text{Vect}\,(\vec{b}_1, \cdots \vec{b}_n)$ (cf proposition 3.3.7). Donc le rang de B est égal à celui de A : il vaut n. Cela montre que la matrice échelonnée B est en réalité triangulaire supérieure à diagonale non nulle. D'après la proposition 4.2.4, on a donc $\det B \neq 0$, d'où $\det A \neq 0$.

■

Théorème 4.2.6 *Pour toute matrice A de $\mathcal{M}_n(\mathbb{K})$, on a* $\det A = \det {}^t A$.

Preuve : Voir le cours du second semestre. ■

Corollaire 4.2.7 *Toutes les propriétés énoncées sur les lignes des matrices pour le déterminant, sont également vraies pour les colonnes : invariance du déterminant par ajout d'une combinaison linéaire d'autres colonnes, caractère alterné par rapport aux colonnes, linéarité,... De plus, le déterminant d'une matrice triangulaire inférieure est égal au produit des termes diagonaux.*

Exercice : Prouver le corollaire ci-dessus.

4.3 Calcul du déterminant par pivot de Gauss

Principe : A l'aide de transformations élémentaires sur les lignes, se ramener au calcul du déterminant d'une matrice triangulaire.

Considérons une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$. L'algorithme du pivot de Gauss décrit au chapitre 2 permet d'obtenir une matrice T échelonnée après une succession de transformations élémentaires. Comme la matrice de départ est carrée, la matrice T est triangulaire supérieure. Les transformations élémentaires effectuées durant le pivot sont de type $(T1)$ (permutation de deux lignes) ce qui revient à changer le déterminant en son opposé, ou de type $(T3)$ (ajout d'une combinaison linéaire de lignes), ce qui ne change pas le déterminant. On a donc finalement

$$\det A = (-1)^\ell \det T = (-1)^\ell \prod_{i=1}^{n} t_{ii}$$

où ℓ est le nombre de permutations de lignes effectuées au cours du pivot.

Exemple :

$$\begin{vmatrix} 1 & 0 & 4 \\ 2 & 3 & 1 \\ -1 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 4 \\ 2 & 3 & 1 \\ 0 & 0 & 5 \end{vmatrix} \quad \text{(ajout de } (L_1) \text{ à } (L_3)\text{)},$$

$$= \begin{vmatrix} 1 & 0 & 4 \\ 0 & 3 & -7 \\ 0 & 0 & 5 \end{vmatrix} \quad \text{(on retranche } 2(L_1) \text{ à } (L_2)\text{)}.$$

La matrice ainsi obtenue est triangulaire supérieure et le produit de ses termes diagonaux est 15. On a donc $\det A = 15$.

Remarque 4.3.1 *Comme* $\det A = \det {}^t A$*, on peut également effectuer un pivot de Gauss sur les colonnes, ou même alterner opérations sur les lignes et opérations sur les colonnes.*

Attention : 1. Pouvoir faire indifféremment des opérations sur les lignes et sur les colonnes est spécifique au calcul du déterminant. Faire un pivot de Gauss sur les colonnes pour résoudre un système linéaire donne un résultat faux!!!
2. On prendra garde au fait que multiplier une ligne (ou une colonne) par un scalaire change le déterminant.

4.4 Développement de Laplace

Lemme 4.4.1 *Soit* $B \in \mathcal{M}_{n-1}(\mathbb{K})$ *et* $A \in \mathcal{M}_n(\mathbb{K})$ *définie par*

$$A = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}.$$

Alors $\det A = \det B$.

Preuve : La méthode du pivot de Gauss permet de transformer B en une matrice $T \in \mathcal{M}_{n-1}(\mathbb{K})$ triangulaire supérieure grâce à une succession de transformations élémentaires sur les lignes de B. Si l'on note ℓ le nombre de permutations effectuées au cours du pivot, l'on sait de plus que $\det B = (-1)^\ell \det T$. Remarquons qu'effectuer les mêmes opérations élémentaires sur les $(n-1)$ dernières lignes de A n'affecte nullement la première ligne et conduit donc à la matrice

$$C \overset{\text{déf}}{=} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & T & \\ 0 & & & \end{pmatrix},$$

si bien que $\det A = (-1)^\ell \det C$.

Enfin, il est clair que la matrice C est triangulaire supérieure. Par conséquent, d'après la proposition 4.2.4,

$$\det C = \prod_{i=1}^{n-1} t_{ii}.$$

On a donc

$$\det A = (-1)^\ell \prod_{i=1}^{n-1} t_{ii} = (-1)^\ell \det T = \det B.$$

■

Lemme 4.4.2 *Soit $B \in \mathcal{M}_{n-1}(\mathbb{K})$ et $A \in \mathcal{M}_n(\mathbb{K})$ la matrice obtenue à partir de B par adjonction de la colonne $\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ en j-ième position et de la ligne $(0 \cdots 0 \underbrace{1}_{j} 0 \cdots 0)$ en i-ième position :*

$$A = \left(\begin{array}{ccc|c|ccc} & & & 0 & & & \\ & B_1 & & \vdots & & B_2 & \\ & & & 0 & & & \\ \hline 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \hline & & & 0 & & & \\ & B_3 & & \vdots & & B_4 & \\ & & & 0 & & & \end{array}\right).$$

Alors $\det A = (-1)^{i+j} \det B$.

Preuve : On permute les lignes (L_i) et (L_{i-1}) puis (L_{i-1}) et (L_{i-2}), etc. Après $i-1$ permutations de ce type, on obtient la matrice

$$A' \overset{\text{déf}}{=} \left(\begin{array}{ccc|c|ccc} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \hline & & & 0 & & & \\ & B_1 & & \vdots & & B_2 & \\ \hline & & & & & & \\ & B_3 & & \vdots & & B_4 & \\ & & & 0 & & & \end{array}\right).$$

Ensuite, on permute les colonnes j et $j-1$ de A', puis $j-1$ et $j-2$, etc, jusqu'à obtenir après $j-1$ permutations de ce type, le matrice $A'' \overset{\text{déf}}{=} \begin{pmatrix} 1 & 0 \\ 0 & \boxed{B} \end{pmatrix}$. Chaque permutation de lignes ou de colonnes change le déterminant en son opposé. On a donc finalement

$$\det A = (-1)^{i-1} \det A' = (-1)^{i+j-2} \det A''.$$

Enfin, d'après le lemme 4.4.1, on a $\det A'' = \det B$, d'où le résultat voulu. ■

Définition 4.4.3 *Soit $A \in \mathcal{M}_n(\mathbb{K})$ et $1 \leq i, j \leq n$. On appelle mineur de A d'indice (i,j) la matrice $A'_{ij} \in \mathcal{M}_{n-1}(\mathbb{K})$ obtenue en rayant la i-ième ligne et la j-ième colonne de A :*

$$A'_{ij} \overset{\text{déf}}{=} \left(\begin{array}{ccc|c|ccc} a_{11} & \cdots & a_{1j-1} & a_{1j} & a_{1j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \hline a_{i-11} & \cdots & a_{i-1j-1} & a_{i-1j} & a_{i-1,j+1} & \cdots & a_{i-1n} \\ \hline a_{i+11} & \cdots & a_{i+1j-1} & a_{i+1j} & a_{i+1,j-1} & \cdots & a_{i+1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj-1} & a_{nj} & a_{nj+1} & \cdots & a_{nn} \end{array}\right)$$

On appelle cofacteur d'indice (i,j) le scalaire $A_{ij} \overset{\text{déf}}{=} (-1)^{i+j} \det A'_{ij}$.

Exemple : Soit $A = \begin{pmatrix} 2 & 3 & 1 \\ 2 & -1 & 0 \\ 1 & 2 & 0 \end{pmatrix}$. Alors $A'_{12} = \begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix}$ et $A'_{22} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$.

De plus, $A_{12} = -\det A'_{12} = 0$ et $A_{22} = \det A'_{22} = -1$.

Exercice : Calculer les autres mineurs principaux et cofacteurs de A.

Théorème 4.4.4 *Soit $A \in \mathcal{M}_n(\mathbb{K})$ et $1 \leq i, j \leq n$. Alors on dispose des deux formules suivantes pour calculer le déterminant de A :*

Développement de Laplace par rapport à la ligne i :

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A'_{ij} = \sum_{j=1}^{n} a_{ij} A_{ij}. \tag{4.1}$$

Développement de Laplace par rapport à la colonne j :

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A'_{ij} = \sum_{i=1}^{n} a_{ij} A_{ij}. \tag{4.2}$$

Preuve : Remarquons que le développement de Laplace de ${}^tA$ par rapport à la ligne j n'est autre, compte tenu de $\det A = \det {}^tA$, que le développement de Laplace de A par rapport à la colonne j. Il suffit donc de démontrer (4.1).

Notons $\vec{e}_1 = (1, 0, \cdots, 0)$, $\cdots$, $\vec{e}_n = (0, \cdots, 0, 1)$ la base canonique de $\mathbb{K}^n$, et $\vec{a}_i$ la i-ième ligne de A. On a $\vec{a}_i = \sum_{j=1}^{n} a_{ij} \vec{e}_j$.

En utilisant la linéarité du déterminant par rapport à la ligne i, on trouve donc

$$\det A = \sum_{j=1}^{n} a_{ij} \det \widehat{A}_{ij} \tag{4.3}$$

avec

$$\widehat{A}_{ij} \overset{\text{déf}}{=} \left(\begin{array}{ccc|c|ccc} a_{11} & \cdots & a_{1j-1} & a_{1j} & a_{1j+1} & \cdots & a_{1n} \\ \vdots & * & \vdots & \vdots & \vdots & * & \vdots \\ a_{i-11} & \cdots & a_{i-1j-1} & a_{i-1j} & a_{i-1,j+1} & \cdots & a_{i-1n} \\ \hline 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \hline a_{i+11} & \cdots & a_{i+1j-1} & a_{i+1j} & a_{i+1,j-1} & \cdots & a_{i+1n} \\ \vdots & * & \vdots & \vdots & \vdots & * & \vdots \\ a_{n1} & \cdots & a_{nj-1} & a_{nj} & a_{nj+1} & \cdots & a_{nn} \end{array}\right).$$

Remarquons que la colonne j peut se décomposer en

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{i-1j} \\ 1 \\ a_{i+1j} \\ \vdots \\ a_{nj} \end{pmatrix} = a_{1j} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \\ \\ \vdots \\ 0 \end{pmatrix} + \cdots + a_{i-1j} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + a_{i+1j} \begin{pmatrix} 0 \\ \vdots \\ \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + a_{nj} \begin{pmatrix} 0 \\ \vdots \\ \\ \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

En utilisant maintenant la linéarité du déterminant par rapport à la colonne j, on trouve

$$\det \widehat{A}_{ij} = a_{1j} \left|\begin{array}{ccc|c|ccc} a_{11} & \cdots & a_{1j-1} & 1 & a_{1j+1} & \cdots & a_{1n} \\ & * & & 0 & & * & \\ & & & \vdots & & & \\ a_{i-11} & & & & & & a_{i-1n} \\ \hline 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \hline a_{i+11} & & & 0 & & & a_{i+1n} \\ & * & & \vdots & & * & \\ & & & 0 & & & \end{array}\right|$$

$$+ \cdots + a_{ij} \left|\begin{array}{ccc|c|ccc} a_{11} & \cdots & a_{1j-1} & 0 & a_{1j+1} & \cdots & a_{1n} \\ & * & & \vdots & & * & \\ & & & 0 & & & a_{i-1n} \\ a_{i-11} & & & & & & \\ \hline 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \hline a_{i+11} & & & 0 & & & a_{i+1n} \\ & * & & \vdots & & * & \\ & & & 0 & & & \end{array}\right|$$

$$+ \cdots + a_{nj} \left|\begin{array}{ccc|c|ccc} a_{11} & \cdots & a_{1j-1} & 0 & a_{1j+1} & \cdots & a_{1n} \\ & * & & 0 & & * & \\ & & & \vdots & & & \\ a_{i-11} & & & & & & a_{i-1n} \\ \hline 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \hline a_{i+11} & & & \vdots & & & a_{i+1n} \\ & * & & 0 & & * & \\ & & & 1 & & & \end{array}\right|.$$

Tous les termes de cette somme sauf le i-ème sont nuls car leur i-ième ligne est nulle. Le i-ième terme est le déterminant de la matrice obtenue en remplaçant la i-ème ligne et la j-ème colonne de A par des 0, sauf à la place (i, j) où l'on met un 1. Le lemme 4.4.2 assure que ce terme vaut $(-1)^{i+j} \det A'_{ij}$. En revenant à l'égalité (4.3), on obtient la formule (4.1).
■

Exemples :

1. Calcul du déterminant en dimension $n = 1$.
   Toute matrice de $\mathcal{M}_1(\mathbb{K})$ est diagonale ! Si A est une matrice carrée de taille 1, on a donc tout simplement $\det A = a_{11}$.
2. Calcul du déterminant en dimension $n = 2$.
   Soit $A \in \mathcal{M}_2(\mathbb{K})$. Un développement de Laplace par rapport à la première ligne donne

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \det(a_{22}) - a_{12} \det(a_{21}) = a_{11}a_{22} - a_{12}a_{21}.$$

3. Calcul du déterminant en dimension $n = 3$. Soit $A \in \mathcal{M}_3(\mathbb{K})$. Un développement de Laplace par rapport à la première ligne donne

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

En utilisant la formule du déterminant pour les matrices 2×2, on conclut que

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

4.5 Le déterminant et le rang

Définition 4.5.1 *Soit $A \in \mathcal{M}_{p,n}(\mathbb{K})$ et $k \leq \min(p, n)$. On dit que $A' \in \mathcal{M}_k(\mathbb{K})$ est une* **matrice extraite** *d'ordre k de A s'il existe k entiers $i_1, \cdots, i_k$ tels que $1 \leq i_1 < \cdots < i_k \leq p$ et k entiers $j_1, \cdots, j_k$ tels que $1 \leq j_1 < \cdots < j_k \leq p$ vérifiant $a'_{\alpha\beta} = a_{i_\alpha j_\beta}$ pour $1 \leq \alpha, \beta \leq k$.*

Remarque 4.5.2 *Une matrice $A' \in \mathcal{M}_k(\mathbb{K})$ est une* **matrice extraite** *d'ordre k de A si et seulement si A peut être transformée en la matrice suivante par permutations de lignes et de colonnes :*

$$\begin{pmatrix} A' & * \\ * & * \end{pmatrix}.$$

Définition 4.5.3 *Si $A' \in \mathcal{M}_k(\mathbb{K})$ est une matrice extraite de A, on dit que $\det A'$ est un* **déterminant extrait** *(d'ordre k) de A.*

Théorème 4.5.4 *Soit $A \in \mathcal{M}_{p,n}(\mathbb{K})$. Notons $\vec{a}_1, \cdots, \vec{a}_p$ les vecteurs lignes (de $\mathbb{K}^n$) de A, et $\vec{a}^1, \cdots, \vec{a}^n$ les vecteurs colonnes (de $\mathbb{K}^p$) de A. Les trois propositions suivantes sont équivalentes :*

1. $\operatorname{rg}(\vec{a}^1, \cdots, \vec{a}^n) \geq k$,
2. $\operatorname{rg}(\vec{a}_1, \cdots, \vec{a}_p) \geq k$,
3. *Il existe un déterminant extrait d'ordre k de A qui est non nul.*

Preuve : Prouvons d'abord $ii) \Rightarrow iii)$. Supposons donc que la famille $(\vec{a}_1, \cdots, \vec{a}_p)$ constituée par les p vecteurs lignes (appartenant à $\mathbb{K}^n$) de A soit de rang au moins au égal à k. D'après le théorème de la base incomplète, on peut en extraire une sous-famille $(\vec{a}_{i_1}, \cdots, \vec{a}_{i_k})$ de rang k. Le rang est invariant par permutation des vecteurs de la famille. Donc la matrice A est de même rang que la matrice B suivante :

$$B \overset{\text{déf}}{=} \left(\begin{array}{c} \vec{a}_{i_1} \\ \vdots \\ \vec{a}_{i_k} \\ \hline C \end{array}\right)$$

où $C \in \mathcal{M}_{p-k,n}$ est la matrice constituée par les lignes de A n'appartenant pas à $(\vec{a}_{i_1}, \cdots, \vec{a}_{i_p})$ (peu importe l'ordre des lignes).

Notons $B' \in \mathcal{M}_{k,n}$ la matrice formée par les k premières lignes de B. Par pivot de Gauss, on peut transformer B' en une matrice échelonnée C' de même rang k. Comme $k \leq p$,

cette matrice est du type :

$$C' = \left(\begin{array}{c|cccc} 0 & c'_{1j_1} & * & * & * \\ \vdots & 0 & c'_{2j_2} & * & * \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & c'_{kj_k} \end{array} \right) * \Bigg).$$

Par définition les termes c'_{1j_i} ne sont pas nuls. Donc les colonnes $j_1, \cdots, j_k$ de C' sont linéairement indépendantes. En effectuant des permutations de colonnes sur C' (ce qui ne modifie toujours pas le rang), on a donc $\operatorname{rg} C' = \operatorname{rg} D' = k$ avec

$$D' = \left(\begin{array}{cccc|c} c'_{1j_1} & \cdots & \cdots & c'_{1j_k} & \\ 0 & \ddots & & \vdots & * \\ \vdots & \ddots & \ddots & \vdots & \\ 0 & \cdots & 0 & c'_{kj_k} & \end{array} \right).$$

Si maintenant on fait subir à la matrice B les permutations de colonnes qui ont permis de passer de C' à D', on trouve une certaine matrice (par blocs) $D \in \mathcal{M}_{p,n}(\mathbb{K})$:

$$D = \left(\begin{array}{c|c} A' & * \\ \hline * & * \end{array}\right) \quad \text{avec} \quad a'_{\alpha\beta} = a_{i_\alpha j_\beta} \quad \text{et} \quad 1 \leq \alpha \leq k,\ 1 \leq \beta \leq n.$$

Si l'on applique à A' les opérations élémentaires qui ont permis de passer de B' à C', mais en remplaçant à chaque fois la colonne j_β par la colonne β, on obtient précisément les k premières colonnes de la matrice D' ci-dessus qui sont visiblement indépendantes. Comme $\det A'$ est égal (au signe près) au déterminant de la matrice carrée de taille k constituée des k premières colonnes de D', on conclut que $\det A' \neq 0$.

Prouvons maintenant l'implication *iii)* $\Rightarrow$ *ii)*. Soit donc

$$A' = \begin{pmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_k} \\ \vdots & & \vdots \\ a_{i_k j_1} & \cdots & a_{i_k j_k} \end{pmatrix}$$

une matrice extraite de A de déterminant non nul.
Si l'on note $\vec{a}'_i$ la i-ième ligne de A', alors on a $\operatorname{rg}(\vec{a}'_1, \cdots, \vec{a}'_k) = k$ d'après la proposition 4.2.5. Rajoutons à ces k vecteurs de $\mathbb{K}^k$ les $n-k$ composantes manquantes pour en faire des vecteurs de $\mathbb{K}^n$, i.e on pose

$$\vec{a}''_\ell \overset{\text{déf}}{=} (a_{i_\ell j_1}, \cdots, a_{i_\ell j_k}, \cdots, a_{i_\ell n}).$$

En revenant à la définition de l'indépendance linéaire, il est facile de vérifier que la famille $(\vec{a}''_1, \cdots, \vec{a}''_k)$ est également libre. Si l'on permute les indices des colonnes afin de remettre les composantes "dans l'ordre", on ne change pas le caractère libre (**exercice :** le vérifier). Et on conclut donc que $\operatorname{rg}(\vec{a}_{i_1}, \cdots, \vec{a}_{i_k}) = k$, puis que $\operatorname{rg}(\vec{a}_1, \cdots \vec{a}_p) \geq k$.

En appliquant le résultat *ii)* $\iff$ *iii)* à la matrice ${}^tA$, on trouve que *i)* $\iff$ *iii)*. ∎

Corollaire 4.5.5 *Pour $A \in \mathcal{M}_{p,n}(\mathbb{K})$, les matrices A et ${}^tA$ ont même rang. Autrement dit,*

$$\operatorname{rg}(\vec{a}^1, \cdots, \vec{a}^n) = \operatorname{rg}(\vec{a}_1, \cdots, \vec{a}_p).$$

Ce rang est égal à l'entier k maximal tel qu'il existe un déterminant extrait d'ordre k de A non nul.

4.6 Résolution d'un système linéaire par la méthode de Cramer

Dans cette section, nous présentons une méthode alternative à celle du pivot de Gauss pour résoudre les systèmes linéaires de n équations à n inconnues.

Définition 4.6.1 *On dit qu'un système linéaire (S) de n équations à n inconnues est* **de Cramer** *s'il admet une unique solution.*

Proposition 4.6.2 *Un système est de Cramer si et seulement si le déterminant de la matrice associée est non nul.*

Preuve : Notons (S) le système initial, et A sa matrice. La méthode du pivot de Gauss permet de se ramener à un système échelonné équivalent (T) de matrice B. Comme on n'a effectué que des transformations élémentaires pour passer de (S) à (T), on a $|\det A| = |\det B|$. Donc $\det A \neq 0 \iff \det B \neq 0$.

Supposons d'abord que $\det B = 0$. Comme B est échelonnée, ceci revient à dire que B a une ligne nulle. (i.e $k < n$). Dans ce cas, l'ensemble des solutions de (T) (et donc de (S)) est ou bien vide ou bien infini suivant la valeur du second membre, et le système n'est pas de Cramer. On a donc montré que

$$(\det B = 0) \Longrightarrow ((S) \text{ n'est pas de Cramer}).$$

Si au contraire $\det B \neq 0$ alors B n'a pas de ligne nulle, i.e B est triangulaire supérieure à diagonale non nulle (car $\det B = \prod_{i=1}^{n} b_{ii}$). Dans ce cas (T) a une unique solution obtenue par la méthode de la remontée, et (S) est donc de Cramer. On a donc montré que

$$(\det B \neq 0) \Longrightarrow ((S) \text{ est de Cramer}).$$

■

Théorème 4.6.3 *Soit (S) un système de Cramer de matrice $A \in \mathcal{M}_n(\mathbb{K})$ et de second membre $\vec{b} = (b_1, \cdots, b_n)$. Notons $(\vec{a}^1, \cdots, \vec{a}^n)$ les vecteurs colonnes de A.*

Alors l'unique solution $(x_1, \cdots, x_n)$ de (S) est donnée par les formules de Cramer :

$$x_i = \frac{\det(\vec{a}^1, \cdots, \vec{a}^{i-1}, \vec{b}, \vec{a}^{i+1}, \cdots, \vec{a}^n)}{\det A}.$$

Preuve : Considérons un système de Cramer

$$\begin{cases} x_1 a_{11} + \cdots + x_n a_{1n} = b_1 \\ \vdots \\ x_1 a_{n1} + \cdots + x_n a_{nn} = b_n. \end{cases}$$

Ce système peut se récrire sous la forme plus compacte :

$$x_1 \vec{a}^1 + \cdots + x_n \vec{a}^n = \vec{b}$$

où l'addition est prise au sens des vecteurs colonnes. On a donc, grâce à la linéarité du déterminant par rapport à la colonne i,

$$\begin{aligned} \det(\vec{a}^1, \cdots, \vec{a}^{i-1}, \vec{b}, \vec{a}^{i+1}, \cdots, \vec{a}^n) &= \det(\vec{a}^1, \cdots, \vec{a}^{i-1}, x_1\vec{a}^1 + \cdots + x_n\vec{a}^n, \vec{a}^{i+1}, \cdots, \vec{a}^n), \\ &= \sum_{j=1}^{n} x_j \det(\vec{a}^1, \cdots, \vec{a}^{i-1}, \vec{a}^j, \vec{a}^{i+1}, \cdots, \vec{a}^n). \end{aligned}$$

Si $i \neq j$, la famille $(\vec{a}^1, \cdots, \vec{a}^{i-1}, \vec{a}^j, \vec{a}^{i+1}, \cdots, \vec{a}^n)$ contient deux fois le vecteur $\vec{a}^j$. Le déterminant correspondant est donc nul. Reste finalement

$$\det(\vec{a}^1, \cdots, \vec{a}^{i-1}, \vec{b}, \vec{a}^{i+1}, \cdots, \vec{a}^n) = x_i \det(\vec{a}^1, \cdots, \vec{a}^n).$$

■

Comparaison des méthodes du pivot de Gauss et de Cramer

On dispose donc maintenant de deux méthodes pour la résolution des systèmes linéaires de n équations à n inconnues. La méthode de Cramer a l'avantage de donner des formules explicites pour la solution du système. On peut donc se demander à quoi sert la méthode du pivot de Gauss. Son intérêt vient de sa rapidité. En effet, la méthode du pivot de Gauss nécessite environ n^3 opérations élémentaires (additions, soustractions, multiplications ou divisions) pour résoudre un système linéaire $n \times n$.

En revanche, le calcul de la solution à l'aide des formules de Cramer nécessite le calcul de $n+1$ déterminants de taille n. Chaque déterminant peut se calculer ou bien par développement direct en itérant la formule de Laplace (ce qui demande environ $n!$ opérations élémentaires) ou bien par pivot de Gauss (environ n^3 opérations). Il faudra donc effectuer au moins n^4 opérations pour calculer la solution à l'aide des formules de Cramer.

En pratique donc, on n'utilise presque jamais les formules de Cramer pour résoudre un système linéaire, sauf éventuellement pour résoudre des systèmes très petits (2×2 ou à la rigueur 3×3).

Chapitre 5

Polynômes

5.1 L'ensemble des polynômes à une indéterminée

5.1.1 Définitions

Définition 5.1.1 *On appelle* **polynôme à une indéterminée et coefficients dans** $\mathbb{K}$ *ou plus simplement polynôme, toute expression algébrique de la forme*

$$a_p X^p + a_{p-1} X^{p-1} + \cdots + a_1 X + a_0,$$

avec $a_i \in \mathbb{K}$ *pour tout* $i \in \{0, \cdots, p\}$.

- *Les scalaires* a_i *sont appelés* **coefficients** *du polynôme.*
- *S'il existe, le plus grand indice* i *tel que* $a_i \neq 0$ *s'appelle* **degré de** P *et est noté* $\deg P$.
- *Si tous les coefficients* a_i *sont nuls,* P *est appelé* **polynôme nul** *et est noté* 0. *Par convention,* $\deg 0 = -\infty$.
- *Un polynôme de la forme* $P = a_0$ *avec* $a_0 \in \mathbb{K}$ *est appelé* **polynôme constant**. *Si* $a_0 \neq 0$, *son degré est* 0.

L'ensemble des polynôme à une indéterminée et coefficients dans $\mathbb{K}$ *est noté* $\mathbb{K}[X]$.

Exemples :

- $X^3 - \pi X + 3/2$ est un polynôme de degré 3.
- Si $n \in \mathbb{N}^*$, $X^n - 1$ est un polynôme de degré n
- 1 est un polynôme de degré 0.

Remarque 5.1.2 *Nous serons amenés par la suite à additionner des degrés de polynômes. Comme l'application* deg *est à valeurs dans* $\mathbb{N} \cup \{-\infty\}$, *il faut étendre la définition de l'addition. On adopte la convention suivante pour* $n \in \mathbb{N} \cup \{-\infty\}$ *:*

$$-\infty + n = -\infty.$$

Définition 5.1.3 *Les polynômes ne comportant qu'un seul terme non nul (i.e du type* $P = a_p X^p$*) sont appelés* **monômes**.

Remarque : Tout polynôme est donc une somme finie de monômes.

Définition 5.1.4 *Soit* $P = a_p X^p + \cdots + a_0$ *avec* $a_p \neq 0$ *un polynôme. On appelle* **terme dominant** *de* P *le monôme* $a_p X^p$. *Si le coefficient* a_p *du terme dominant est* 1, *on dit que* P *est un* **polynôme unitaire**.

Remarque 5.1.5 *On adopte la convention que l'on ne change pas un polynôme* P *en lui ajoutant un ou plusieurs monômes à coefficients nuls. Par exemple, on ne fera pas la distinction entre*

$$X^4 - X + 1 \quad et \quad 0X^5 + X^4 + 0X^2 - X + 1.$$

5.1.2 Opérations sur $\mathbb{K}[X]$

Nous allons munir $\mathbb{K}[X]$ de deux lois internes "+" et "$*$", et d'une loi externe "$\cdot$".

a) Addition de deux polynômes :

Définition 5.1.6 *Soit $P = a_nX^n + \cdots + a_0$ et $Q = b_nX^n + \cdots + b_0$ avec $n \in \mathbb{N}$. On définit alors le polynôme $P + Q$ par*

$$P + Q \overset{\text{déf}}{=} (a_n + b_n)X^n + \cdots + (a_1 + b_1)X + (a_0 + b_0).$$

Remarque : Dans la définition ci-dessus, il n'est pas restrictif de faire commencer les expressions des polynômes P et Q par un monôme de même degré n (voir la remarque 5.1.5 ci-dessus)

Proposition 5.1.7 *Soit P et Q deux polynômes de $\mathbb{K}[X]$. Alors on a*

$$\deg(P + Q) \leq \max(\deg P, \deg Q).$$

De plus, si $\deg P \neq \deg Q$ alors $\deg(P + Q) = \max(\deg P, \deg Q)$.

Preuve : Notons $p = \deg P$ et $q = \deg Q$.
- Si $p > q$, le coefficient du terme dominant de $P + Q$ est a_p donc $\deg(P + Q) = \deg P$.
- Si $p < q$, le coefficient du terme dominant de $P + Q$ est b_q donc $\deg(P + Q) = \deg Q$.
- Si $p = q$, le monôme de plus haut degré dans l'expression de $P + Q$ est $(a_p + b_p)X^p$. Donc $\deg(P + Q) \leq p$. Si $b_p = -a_p$, ce monôme est nul et l'on a donc $\deg(P + Q) < p$.

■

b) Multiplication de deux polynômes :

Considérons deux monômes $P = a_pX^p$ et $Q = b_qX^q$. Si l'on interprète ces deux monômes comme des fonctions de la variable réelle ou complexe X, il est naturel de définir le produit de P par Q comme étant le monôme $P * Q \overset{\text{déf}}{=} a_pb_qX^{p+q}$.

Plus généralement, on définit le produit de deux polynômes de la façon suivante :

Définition 5.1.8 *Étant donnés deux polynômes $P = a_pX^p + \cdots + a_0$ et $Q = b_qX^q + \cdots + b_0$, on définit le polynôme $P * Q$ par $P * Q = c_rX^r + \cdots + c_0$ avec $r = p + q$ et, pour $k \in \{0, \cdots, r\}$,*

$$c_k = \sum_{i+j=k} a_ib_j = \sum_{i=0}^{k} a_ib_{k-i} = \sum_{j=0}^{k} a_{k-j}b_j.$$

Remarque : Si P ou Q est nul, on a donc $P * Q = 0$.

La proposition suivante est une conséquence immédiate de la définition de "$*$" :

Proposition 5.1.9 *Soit P et Q deux polynômes de $\mathbb{K}[X]$. Alors on a*

$$\deg(P * Q) = \deg P + \deg Q.$$

c) Multiplication d'un polynôme par un scalaire :

Définition 5.1.10 *Soit $P = a_pX^p + \cdots + a_0$ un polynôme de $\mathbb{K}[X]$, et $\lambda \in \mathbb{K}$. On définit alors le polynôme $\lambda \cdot P$ par*

$$\lambda \cdot P \overset{\text{déf}}{=} \sum_{i=0}^{p} \lambda a_iX^i.$$

Le lecteur prouvera sans difficulté le résultat suivant :

Proposition 5.1.11 *Soit P un polynôme et λ un scalaire <u>non nul</u>. Alors $\deg(\lambda \cdot P) = \deg P$.*

5.1.3 Propriétés algébriques de $\mathbb{K}[X]$

Proposition 5.1.12 *$(\mathbb{K}[X], +, *)$ est un anneau commutatif.*

Preuve : Montrons déjà que $(\mathbb{K}[X], +)$ est un groupe commutatif.

- Le polynôme nul est clairement l'élément neutre pour l'addition.
- Si $P = a_p X^p + \cdots + a_0$, le polynôme $-P \stackrel{\text{déf}}{=} -a_p X^p + \cdots - a_1 X - a_0$ vérifie $P + (-P) = 0$.
- L'associativité et la commutativité résultent de celles de l'addition sur $\mathbb{K}$.

Reste à étudier les propriétés de la multiplication "$*$".

- De la définition de la multiplication sur $\mathbb{K}[X]$, on déduit facilement que le polynôme $P = 1$ est l'élément neutre pour "$*$".
- Commutativité : considérons $P = a_p X^p + \cdots + a_0$ et $Q = b_q X^q + \cdots + b_0$. Notons $r = p + q$, $P * Q = c_r X^r + \cdots + c_0$ et $Q * P = d_r X^r + \cdots + d_0$. Alors on a

$$\forall k \in \{0, \cdots, r\},\ c_k = \sum_{i+j=k} a_i b_j = \sum_{j+i=k} b_j a_i = d_k$$

  Donc $P * Q = Q * P$.
- Associativité : Soit $P = a_p X^p + \cdots + a_0$, $Q = b_q X^q + \cdots + b_0$ et $R = c_r X^r + \cdots + c_0$. Soit $U \stackrel{\text{déf}}{=} (P * Q) * R$ et $V \stackrel{\text{déf}}{=} P * (Q * R)$. Notons d_ℓ les coefficients de U, et e_m ceux de V. Enfin, notons f_s les coefficients de $P * Q$, et g_t ceux de $Q * R$. Alors on a

$$\begin{array}{rcl|rcl} d_\ell &=& \sum_{s+k=\ell} f_s c_k & e_\ell &=& \sum_{i+t=\ell} a_i g_t \\ &=& \sum_{s+k=\ell} \left(\sum_{i+j=s} a_i b_j \right) c_k & &=& \sum_{i+t=\ell} a_i \left(\sum_{j+k=t} b_j c_k \right) \\ &=& \sum_{i+j+k=\ell} a_i b_j c_k. & &=& \sum_{i+j+k=\ell} a_i b_j c_k. \end{array}$$

  Donc $d_\ell = e_\ell$, d'où $U = V$.
- Distributivité de la multiplication sur l'addition : Définissons P, Q, R comme ci-dessus et posons $U \stackrel{\text{déf}}{=} (P + Q) * R$ et $V \stackrel{\text{déf}}{=} P * R + Q * R$. Notons encore d_ℓ les coefficients de U, et e_m ceux de V. Alors on a

$$d_\ell = \sum_{i+j=\ell} (a_i + b_i) c_j = \sum_{i+j=\ell} (a_i c_j + b_i c_j) = \sum_{i+j=\ell} a_i c_j + \sum_{i+j=\ell} b_i c_j = e_\ell.$$

  Donc $U = V$.

■

À titre d'exercice, le lecteur pourra établir la

Proposition 5.1.13 *L'anneau $(\mathbb{K}[X], +, *)$ vérifie les propriétés supplémentaires suivantes pour tout $(\lambda, \mu) \in \mathbb{K}^2$ et $(P, Q) \in \mathbb{K}[X]^2$:*

1. $(\lambda + \mu) \cdot P = \lambda \cdot P + \mu \cdot P$,
2. $\lambda \cdot (P + Q) = \lambda \cdot P + \lambda \cdot Q$,
3. $\lambda \cdot (\mu \cdot P) = (\lambda\mu) \cdot P$,
4. $1 \cdot P = P$,
5. $\lambda \cdot (P * Q) = (\lambda \cdot P) * Q = P * (\lambda \cdot Q)$.

On dit que $(\mathbb{K}[X], +, *, \cdot)$ est une **algèbre**.

Ainsi, multiplier un polynôme P par un scalaire λ est équivalent à le multiplier par le polynôme constant $\lambda \cdot 1$. On peut donc sans danger noter la multiplication interne $*$ et la multiplication externe $\cdot$ par le même symbole.

Enfin, $(\mathbb{K}[X], +, *, \cdot)$ jouit de la propriété suivante qui est primordiale :

Proposition 5.1.14 *Soit (P,Q) un couple de polynômes tel que $P*Q=0$. Alors $P=0$ ou $Q=0$. On dit que $(\mathbb{K}[X],+,*,\cdot)$ est une* **algèbre intègre**.

Preuve : Soit donc (P,Q) tel que $P*Q=0$. Alors on a $\deg P+\deg Q=\deg(P*Q)=-\infty$. Donc $\deg P$ ou $\deg Q$ vaut $-\infty$, ce qui est exactement la propriété demandée. ■

Notations : Dorénavant, on omettra les symboles"$*$" et "$\cdot$". Ainsi PQ désignera $P*Q$, et λP désignera $\lambda\cdot P$.

5.2 Division des polynômes

Définition 5.2.1 *On dit que le polynôme A est* **divisible** *par le polynôme B s'il existe un polynôme Q tel que $A=BQ$. Dans ce cas, on note $B\mid A$[1] et l'on dit que A est* **multiple** *de B (ou que B est* **diviseur** *de A). Le polynôme Q est parfois noté $\frac{A}{B}$ ou A/B.*

Remarques :

1. Le polynôme nul est divisible par tous les polynômes. En revanche seul le polynôme nul est divisible par le polynôme nul.
2. Dans le cas où A et B sont tous les deux non nuls, $B|A$ entraîne $\deg B\leq\deg A$.

Proposition 5.2.2 *Soit A et B, deux polynômes non nuls. Si $A\mid B$ et $B\mid A$ alors A et B sont proportionnels, c'est-à-dire qu'il existe $\lambda\in\mathbb{K}^*$ tel que $A=\lambda B$. On dit que A et B sont* **associés**.

Preuve : D'après la remarque ci-dessus, on a à la fois $\deg A\leq\deg B$ et $\deg B\leq\deg A$. Donc A et B sont de même degré. Comme $B\mid A$, on en déduit que $A=BQ$ avec $\deg Q=0$. Autrement dit Q est un polynôme constant (et non nul car A n'est pas nul). ■

Remarque 5.2.3 *Deux polynômes unitaires associés sont forcément égaux.*

Exercice : Prouver la remarque ci-dessus.

Proposition 5.2.4 *Soit B un polynôme non nul, et A un multiple de B de même degré que B. Alors A et B sont associés.*

Preuve : Elle reprend la dernière partie de celle de la proposition 5.2.2. ■

Théorème 5.2.5 (Division euclidienne) *Soit A et B deux polynômes avec B non nul. Alors il existe un unique couple (Q,R) de polynômes tel que*

$$A=BQ+R\quad et\quad \deg R<\deg B.$$

Le polynôme Q est appelé **quotient** *de la division de A par B, R est le* **reste**, *B, le diviseur, et A, le* **dividende**.

Preuve : On va d'abord prouver l'unicité du couple (Q,R), puis son existence.

Unicité : Supposons que $A=BQ+R=BQ'+R'$ avec $\deg R<\deg B$ et $\deg R'<\deg B$. Alors on a $R-R'=B(Q'-Q)$. Donc $\deg(R-R')=\deg B+\deg(Q'-Q)$.

Si $Q\neq Q'$, alors on en déduit que $\deg(R-R')\geq\deg B$.

Donc d'après la proposition 5.1.7, $\max(\deg R,\deg R')\geq\deg B$, ce qui contredit la définition de R ou de R'. Donc $Q=Q'$, puis $R=R'$.

[1] Lire "B divise A" et non pas le contraire !

Existence : Fixons un polynôme $B = b_m X^m + \cdots + b_0$ de degré $m \geq 1$ (le cas B constant non nul étant évident). L'existence du couple (Q, R) vérifiant les propriétés voulues se montre par récurrence sur le degré de A. Pour $n \in \mathbb{N}$, on note $(\mathcal{P}_n)$ l'hypothèse de récurrence suivante :

$$(\mathcal{P}_n)\ \ (\forall A \in \mathbb{K}[X], \deg A \leq n) \Rightarrow (\exists Q \in \mathbb{K}[X], \exists R \in \mathbb{K}[X] \mid A = BQ + R \text{ et } \deg R < \deg B)\,.$$

Il est clair que $(\mathcal{P}_{m-1})$ est vraie. En effet, il suffit de choisir $Q = 0$ et $R = A$.
Soit maintenant $n \geq m$. Supposons $(\mathcal{P}_{n-1})$ vraie et démontrons $(\mathcal{P}_n)$. Le polynôme A est de la forme $A = a_n X^n + \cdots + a_0$ avec $a_n \neq 0$. Comme $n \geq m$ et $b_m \neq 0$, l'expression

$$A' \stackrel{\text{déf}}{=} A - \frac{a_n}{b_m} X^{n-m} B$$

est bien un polynôme, et son degré est au plus $n-1$. D'après $(\mathcal{P}_{n-1})$, il existe donc deux polynômes Q' et R' tels que $A' = Q'B + R'$ et $\deg R' < \deg B$. On en déduit que

$$A = \underbrace{\left(\frac{a_n}{b_m} X^{n-m} + Q'\right)}_{\stackrel{\text{déf}}{=} Q} B \ + \ \underbrace{R'}_{\stackrel{\text{déf}}{=} R}\ ,$$

ce qui démontre $(\mathcal{P}_n)$. ■

La démonstration ci-dessus suggère un procédé de construction itératif permettant de calculer Q et R. En effet, au cours de la récurrence, on a vu comment ramener la division d'un polynôme de degré n à celle d'un polynôme de degré moins élevé (au plus $n-1$). En pratique, on peut donc calculer le couple (Q, R) en "posant" la division comme dans $\mathbb{N}$, les puissances de X jouant le rôle des puissances de 10.

Illustrons nos propos par un exemple.

Exemple : Division de $4X^5 - 7X^3 + 8X^2 - 1$ par $X^3 - 4X^2 + 2X + 3$.

$$\begin{array}{rrrrrrr|l}
4X^5 + & 0X^4 - & 7X^3 + & 8X^2 + & 0X - & 1 & X^3 - 4X^2 + 2X + 3 \\
\cline{8-8}
 & 16X^4 - & 15X^3 - & 4X^2 + & 0X - & 1 & \\
 & & 49X^3 - & 36X^2 - & 48X - & 1 & 4X^2 + 16X + 49 = Q \\
 & & R = & 160X^2 - & 146X - & 148 &
\end{array}$$

Donc $4X^5 - 7X^3 + 8X^2 - 1 = (X^3 - 4X^2 + 2X + 3)(4X^2 + 16X + 49) + 160X^2 - 146X - 148$.

Définition 5.2.6 *On dit qu'un sous-ensemble I de $\mathbb{K}[X]$ est un* **idéal** *de $(\mathbb{K}[X], +, *)$ si*

1. *I est un sous-groupe de $(\mathbb{K}[X], +)$,*
2. *I est stable par multiplication par n'importe quel polynôme de $\mathbb{K}[X]$.*

Exemple : Pour $B \in \mathbb{K}[X]$, on note $B\mathbb{K}[X]$ l'ensemble des multiples de B. Il est facile de vérifier que $B\mathbb{K}[X]$ est un idéal de $\mathbb{K}[X]$. En particulier, le singleton $\{0\}$ est un idéal.

Nous laissons au lecteur le soin de montrer la proposition suivante :

Proposition 5.2.7 *Soit A et B deux polynômes. Alors $A \mid B$ si et seulement si $B\mathbb{K}[X] \subset A\mathbb{K}[X]$.*

Théorème 5.2.8 *Soit I un idéal de $(\mathbb{K}[X], +, *)$ non réduit à $\{0\}$. Alors il existe un unique polynôme P unitaire tel que $I = P\mathbb{K}[X]$. Le polynôme P est appelé* **générateur unitaire** *de I.*

On dit que $(\mathbb{K}[X], +, *)$ est un **idéal principal** *.*

Preuve : Soit I un idéal de $(\mathbb{K}[X], +, *)$ non réduit à $\{0\}$. On note

$$E = \{\deg A \mid A \in I \backslash \{0\}\}.$$

L'ensemble E est une partie non vide de $\mathbb{N}$, donc admet un plus petit élément. On en déduit que I admet un polynôme P non nul et de degré minimal. Comme pour tout $\lambda \in \mathbb{K}$, le polynôme λP appartient aussi à I, on peut toujours choisir P unitaire. La stabilité de I par multiplication par les éléments de $\mathbb{K}[X]$ assure que $P\mathbb{K}[X] \subset I$.
Reste à montrer que $I \subset P\mathbb{K}[X]$. Soit donc $A \in I$. Écrivons la division euclidienne de A par P :

$$A = PQ + R \quad \text{avec} \quad \deg R < \deg P.$$

Comme A et PQ appartiennent à I, on a aussi $R \in I$. Mais par ailleurs $\deg R < \deg P$. Vu la définition de P, on conclut que $R = 0$. ■

5.3 PGCD et PPCM

La division euclidienne va nous permettre de définir les notions de PGCD et de PPCM dans l'ensemble des polynômes.

5.3.1 PGCD

Proposition 5.3.1 *Soit A et B deux polynômes non tous les deux nuls. L'ensemble*

$$A\mathbb{K}[X] + B\mathbb{K}[X] \overset{\text{déf}}{=} \left\{AP + BQ \mid P \in \mathbb{K}[X],\ Q \in \mathbb{K}[X]\right\}$$

est un idéal de $\mathbb{K}[X]$ non réduit à $\{0\}$. Son générateur unitaire[2] *D est appelé* **Plus Grand Commun Diviseur** *(ou plus simplement PGCD) de A et de B, et est noté* $\mathrm{PGCD}(A, B)$.

Preuve : Notons $J \overset{\text{déf}}{=} A\mathbb{K}[X] + B\mathbb{K}[X]$. Remarquons que J n'est pas réduit à $\{0\}$ car contient A et B, et que l'un de ces deux polynômes n'est pas nul par hypothèse. Reste à montrer que J est un idéal.

1. Montrons que J est un sous-groupe de $(\mathbb{K}[X], +)$:
   - Il est évident que $0 \in J$.
   - Soit C et C' deux polynômes de J. Alors il existe quatre polynômes P, P', Q et Q' tels que $C = AP + BQ$ et $C' = AP' + BQ'$. Donc

     $$C + C' = A(P+P') + B(Q+Q') \in J.$$

   - Enfin, si $C = AP + BQ$, il est clair que $-C = A(-P) + B(-Q)$, donc $-C \in J$.
2. Stabilité de J par produit :
   Soit $C = AP + BQ$ un élément de J, et R un polynôme quelconque. Alors $RC = A(PR) + B(QR)$ donc $RC \in J$.

On conclut que J est un idéal non réduit à $\{0\}$. Le théorème 5.2.8 assure l'existence d'un unique polynôme unitaire D tel que $A\mathbb{K}[X] + B\mathbb{K}[X] = D\mathbb{K}[X]$. ■

Remarque : On convient que $\mathrm{PGCD}(0, 0) = 0$. Pour tout couple de polynômes (A, B), on a donc $A\mathbb{K}[X] + B\mathbb{K}[X] = \mathrm{PGCD}(A, B)\,\mathbb{K}[X]$.

La proposition suivante justifie l'appellation "PGCD" donnée au générateur unitaire de $A\mathbb{K}[X] + B\mathbb{K}[X]$.

[2] Dans certains ouvrages, le caractère unitaire n'est pas imposé au PGCD.

Proposition 5.3.2 *Soit (A, B) un couple de polynômes distinct de $(0, 0)$. Alors* $\mathrm{PGCD}\,(A, B)$ *est l'unique polynôme unitaire vérifiant*

$$\mathrm{PGCD}\,(A,B) \mid A, \quad \mathrm{PGCD}\,(A,B) \mid B \quad \textit{et} \quad (P \mid A \textit{ et } P \mid B) \Rightarrow P \mid \mathrm{PGCD}\,(A,B). \tag{5.1}$$

Preuve : Notons $D = \mathrm{PGCD}\,(A, B)$ et montrons que D vérifie (5.1).

Par définition, $D\mathbb{K}[X] = A\mathbb{K}[X] + B\mathbb{K}[X]$. Comme A et B appartiennent tous les deux à l'ensemble de droite, A et B sont bien des multiples de D. Enfin, si P divise A et B alors, d'après la proposition 5.2.7, $A\mathbb{K}[X] \subset P\mathbb{K}[X]$ et $B\mathbb{K}[X] \subset P\mathbb{K}[X]$. Donc $D\mathbb{K}[X] = A\mathbb{K}[X] + B\mathbb{K}[X] \subset P\mathbb{K}[X]$. Donc P divise D.

Pour montrer l'unicité, considérons un polynôme D' unitaire vérifiant (5.1). On a donc en particulier $D \mid D'$. Mais bien sûr $D' \mid D$ donc D et D' sont associés (cf prop. 5.2.2). Comme D et D' sont unitaires, on a $D = D'$. ■

Proposition 5.3.3 *Si A et B ne sont pas simultanément nuls et si C est unitaire alors on a*

$$\mathrm{PGCD}\,(AC, BC) = C\,\mathrm{PGCD}\,(A, B).$$

Preuve : Notons $D = \mathrm{PGCD}\,(A, B)$ et $\Delta = \mathrm{PGCD}\,(AC, BC)$. Il suffit alors de remarquer que

$$\Delta\mathbb{K}[X] = AC\mathbb{K}[X] + BC\mathbb{K}[X] = C\,(A\mathbb{K}[X] + B\mathbb{K}[X]) = CD\mathbb{K}[X].$$

■

Définition 5.3.4 *On dit que deux polynômes A et B sont* **premiers entre eux** *si leur PGCD vaut* 1.

Théorème 5.3.5 (de Bezout) *Deux polynômes A et B sont premiers entre eux si et seulement si il existe deux polynômes U et V tels que $AU + BV = 1$.*

Preuve : $\Longrightarrow$ Si $\mathrm{PGCD}\,(A, B) = 1$ alors par définition du PGCD, on a $A\mathbb{K}[X] + B\mathbb{K}[X] = \mathbb{K}[X]$. Donc $1 \in A\mathbb{K}[X] + B\mathbb{K}[X]$, ce qui signifie qu'il existe U et V tels que $AU + BV = 1$.

$\Longleftarrow$ Si $AU + BV = 1$ alors $1 \in A\mathbb{K}[X] + B\mathbb{K}[X]$. Le générateur unitaire de $A\mathbb{K}[X] + B\mathbb{K}[X]$ est donc un diviseur de 1, donc 1 lui-même. On a donc bien $1 = \mathrm{PGCD}\,(A, B)$. ■

Proposition 5.3.6 *Pour que le polynôme unitaire D soit le PGCD de A et de B, il faut et il suffit que*

$$D \mid A, \quad D \mid B \quad \textit{et} \quad \mathrm{PGCD}\,(\frac{A}{D}, \frac{B}{D}) = 1. \tag{5.2}$$

Preuve : Si $D = \mathrm{PGCD}\,(A, B)$, on a bien sûr $D \mid A$ et $D \mid B$. Notons $P = \frac{A}{D}$ et $Q = \frac{B}{D}$. D'après la proposition 5.3.3, on a

$$D = \mathrm{PGCD}\,(A, B) = \mathrm{PGCD}\,(DP, DQ) = D\,\mathrm{PGCD}\,(P, Q).$$

Comme D n'est pas nul, on conclut que $\mathrm{PGCD}\,(P, Q) = 1$.

Réciproquement, supposons que (5.2) soit satisfaite. Alors, la proposition 5.3.3 entraîne

$$\mathrm{PGCD}\,(A, B) = \mathrm{PGCD}\,(D\frac{A}{D}, D\frac{B}{D}) = D\mathrm{PGCD}\,(\frac{A}{D}, \frac{B}{D}) = D.$$

■

Théorème 5.3.7 (de Bezout généralisé) *Supposons que D unitaire divise A et B avec A et B non tous les deux nuls. Alors on a*

$$D = \mathrm{PGCD}\,(A, B) \iff \exists U \in \mathbb{K}[X],\ \exists V \in \mathbb{K}[X],\ AU + BV = D.$$

Preuve : En appliquant la proposition 5.3.6, on a

$$D = \mathrm{PGCD}\,(A, B) \iff 1 = \mathrm{PGCD}\,(\frac{A}{D}, \frac{B}{D}).$$

Or d'après le théorème de Bezout, on a

$$\mathrm{PGCD}\,(\frac{A}{D}, \frac{B}{D}) = 1 \iff \exists U \in \mathbb{K}[X],\ \exists V \in \mathbb{K}[X],\ \frac{A}{D}U + \frac{B}{D}V = 1,$$

ce qui achève la preuve du théorème. ■

Théorème 5.3.8 (de Gauss) *Si P divise AB et est premier avec A alors P divise B.*

Preuve : Soit B' le polynôme unitaire associé à B. On a

$$\mathrm{PGCD}\,(PB, AB) = B'\,\mathrm{PGCD}\,(P, A) = B'.$$

Par hypothèse, P divise AB, et il est clair que P divise aussi PB. Donc P divise B' et, partant, B. ■

Proposition 5.3.9 *Un polynôme P est premier avec un produit AB si et seulement si P est premier avec A et avec B.*

Preuve : $\Rightarrow$ Supposons P premier avec AB. Soit P' divisant P et A. Alors P' divise aussi AB. Donc $P' \mid \mathrm{PGCD}\,(AB, P)$, i.e $P'|1$. On en déduit que P' est un polynôme constant. Donc P est premier avec A. On établit de même que P est premier avec B.

$\Leftarrow$ On prouve la réciproque par contraposition. Supposons que P ne soit pas premier avec AB. Alors il existe P' divisant P et AB, et tel que $\deg P' \geq 1$. Si P est premier avec A alors P' également. D'après le théorème de Gauss, P' divise donc B. On a donc montré que P' divise à la fois P et B. Comme $\deg P' \geq 1$, cela signifie que P et B ne sont pas premiers entre eux. ■

Remarque 5.3.10 *Une récurrence élémentaire permet de montrer plus généralement qu'un polynôme P est premier avec un produit de polynôme $A_1 \cdots A_k$ si et seulement si il est premier avec chacun des facteurs A_i. Les détails sont laissés en* **exercice**.

5.3.2 L'algorithme d'Euclide

L'*algorithme d'Euclide* est un moyen systématique permettant de calculer le PGCD de deux polynômes. L'outil de base est la *division euclidienne*. L'algorithme repose sur le lemme suivant :

Lemme 5.3.11 *Soit B un polynôme non nul, et A un polynôme quelconque. Notons Q et R le quotient et le reste de la division euclidienne de A par B. Alors on a*

$$\mathrm{PGCD}\,(A, B) = \mathrm{PGCD}\,(B, R).$$

Preuve : Soit D divisant A et B. Comme $R = A - BQ$, le polynôme D divise aussi R. Donc D divise $\mathrm{PGCD}\,(B,R)$. En choisissant $D = \mathrm{PGCD}\,(A,B)$, on conclut que $\mathrm{PGCD}\,(A,B) \mid \mathrm{PGCD}\,(B,R)$.

Soit maintenant D un polynôme divisant B et R. Comme $A = BQ + R$, on a aussi $D \mid A$. Donc $D \mid \mathrm{PGCD}\,(A,B)$. On a donc finalement $\mathrm{PGCD}\,(B,R) \mid \mathrm{PGCD}\,(A,B)$.

Les deux polynômes $\mathrm{PGCD}\,(B,R)$ et $\mathrm{PGCD}\,(A,B)$ sont unitaires et associés. Ils sont donc égaux. ■

Ce lemme indique clairement la stratégie à suivre pour calculer $\mathrm{PGCD}\,(A,B)$. Quitte à permuter A et B, on peut toujours supposer que $\deg A \geq \deg B$. On procède alors comme suit :

- Si $B = 0$, il n'y a rien à faire : $\mathrm{PGCD}\,(A,B)$ est égal au polynôme unitaire associé à A.
- Si B n'est pas nul, on effectue la division euclidienne de A par B, ce qui donne deux polynômes Q_0 et R_1 tels que $A = BQ_0 + R_1$ et $\deg R_1 < \deg B$.

Le lemme 5.3.11 montre que $\mathrm{PGCD}\,(A,B) = \mathrm{PGCD}\,(B,R_1)$. On reprend le calcul ci-dessus en remplaçant A par B, et B par R_1. En itérant le procédé, on construit deux suites R_1, R_2, $\cdots$ et Q_0, Q_1, $\cdots$ telles que :

$$\begin{array}{llll}
A & = BQ_0 + R_1 & \text{avec} & \deg R_1 < \deg B, \\
B & = R_1Q_1 + R_2 & \text{avec} & \deg R_2 < \deg R_1, \\
R_1 & = R_2Q_2 + R_3 & \text{avec} & \deg R_3 < \deg R_2, \\
\multicolumn{4}{l}{\dotfill} \\
R_{k-1} & = R_kQ_k + R_{k+1} & \text{avec} & \deg R_{k+1} < \deg R_k, \\
\multicolumn{4}{l}{\dotfill} \\
R_{n-1} & = R_nQ_n + 0. & &
\end{array}$$

Le procédé s'arrête nécessairement au bout d'au plus $\deg P$ étapes car chaque itération diminue d'au moins 1 le degré du reste de la division euclidienne. On a donc finalement

$$\boxed{\mathrm{PGCD}\,(A,B) = \mathrm{PGCD}\,(B,R_1) = \cdots = \mathrm{PGCD}\,(R_k,R_{k+1}) = \cdots = \mathrm{PGCD}\,(R_n,0) = R_n.}$$

Exemple : Calculer $\mathrm{PGCD}\,(X^4-1, X^3-1)$.

Posons la division euclidienne de $X^4 - 1$ par $X^3 - 1$.

$$\begin{array}{rrrrr|l}
X^4 + & 0X^3 + & 0X^2 + & 0X - & 1 & X^3 + \; 0X^2 + \; 0X - \; 1 \\ \cline{6-6}
 & & & X - & 1 & \quad X
\end{array}$$

Donc $\mathrm{PGCD}\,(X^4-1, X^3-1) = \mathrm{PGCD}\,(X^3-1, X-1)$.

On remarque ensuite que $X^3 - 1$ est divisible par $X - 1$ donc finalement

$$\boxed{\mathrm{PGCD}\,(X^4-1, X^3-1) = \mathrm{PGCD}\,(X^3-1, X-1) = \mathrm{PGCD}\,(X-1, 0) = X-1.}$$

5.3.3 PPCM

Nous laissons au lecteur le soin de prouver le résultat suivant :

Proposition 5.3.12 *Considérons deux polynômes non nuls A et B. Alors l'ensemble $A\mathbb{K}[X] \cap B\mathbb{K}[X]$ est un idéal non réduit à $\{0\}$. Son générateur unitaire*[3] *est appelé* **Plus Petit Commun Multiple** *(ou plus simplement* **PPCM***) de A et B. On le note* $\mathrm{PPCM}\,(A,B)$.

Remarque : Si A ou B est nul, on a $A\mathbb{K}[X] \cap B\mathbb{K}[X] = \{0\}$. On adopte alors la convention que $\mathrm{PPCM}\,(A,B) = 0$. Ainsi, on aura toujours $A\mathbb{K}[X] \cap B\mathbb{K}[X] = \mathrm{PPCM}\,(A,B)\,\mathbb{K}[X]$.

[3]Dans certains ouvrages, on n'impose pas au PPCM d'être unitaire

En s'inspirant de la preuve de la proposition 5.1, on obtient le résultat suivant qui explique l'appellation "Plus Petit Commun Multiple" donnée au générateur unitaire de $A\mathbb{K}[X] \cap B\mathbb{K}[X]$.

Proposition 5.3.13 *Soit A et B deux polynômes non nuls. Le PPCM de A et de B est l'unique polynôme unitaire vérifiant la propriété suivante :*

$$A \mid \mathrm{PPCM}(A,B), \quad B \mid \mathrm{PPCM}(A,B) \quad \textit{et} \quad (A \mid M \textit{ et } B \mid M) \Rightarrow \mathrm{PPCM}(A,B) \mid M.$$

À certains égards, le PPCM et le PGCD ont des propriétés très similaires. On a par exemple :

Proposition 5.3.14 *Soit C un polynôme unitaire et A, B deux polynômes. Alors on a*

$$\mathrm{PPCM}(AC, BC) = C\,\mathrm{PPCM}(A,B).$$

Preuve : Il suffit de remarquer que

$$AC\mathbb{K}[X] \cap BC\mathbb{K}[X] = C\left(A\mathbb{K}[X] \cap B\mathbb{K}[X]\right).$$

■

Proposition 5.3.15 *Soit A et B deux polynômes non nuls. Pour que M unitaire soit le PPCM de A et de B, il faut et il suffit que*

$$A \mid M, \quad B \mid M \quad \textit{et} \quad \mathrm{PGCD}\left(\frac{M}{A}, \frac{M}{B}\right) = 1.$$

Preuve : $\Rightarrow$ Notons M le PPCM de A et de B. Alors $M\mathbb{K}[X]$ est inclus dans $A\mathbb{K}[X]$ et dans $B\mathbb{K}[X]$. Donc M divise bien A et B. Soit D unitaire divisant M/A et M/B. Alors $AD|M$ et $BD|M$. Donc $\mathrm{PPCM}(AD, BD)|M$. Mais d'après la proposition 5.3.14, $\mathrm{PPCM}(AD, BD) = D\,\mathrm{PPCM}(A,B) = DM$. Donc $D = 1$.

$\Leftarrow$ Soit M un multiple commun unitaire de A et de B vérifiant de plus $\mathrm{PGCD}\left(\frac{M}{A}, \frac{M}{B}\right) = 1$. D'après le théorème de Bezout, il existe deux polynômes U et V tels que

$$\frac{M}{A}U + \frac{M}{B}V = 1.$$

Multiplions les deux membres de cette égalité par $\mathrm{PPCM}(A,B)$. On trouve

$$M\left(U\frac{\mathrm{PPCM}(A,B)}{A} + V\frac{\mathrm{PPCM}(A,B)}{B}\right) = \mathrm{PPCM}(A,B).$$

Donc M divise $\mathrm{PPCM}(A,B)$. Comme M est unitaire et est multiple de A et de B, on conclut que $M = \mathrm{PPCM}(A,B)$. ■

Proposition 5.3.16 *Soit A et B deux polynômes. Il existe une constante λ non nulle telle que*

$$\lambda AB = \mathrm{PGCD}(A,B)\,\mathrm{PPCM}(A,B).$$

- *Si de plus A et B sont unitaires, alors $\lambda = 1$.*
- *Si A et B sont premiers entre eux alors AB et $\mathrm{PPCM}(A,B)$ sont associés.*

Preuve : Écartons le cas évident où l'un des deux polynômes A et B est nul. On peut alors appliquer la proposition 5.3.15. On en déduit que

$$\mathrm{PGCD}\left(\frac{\mathrm{PPCM}(A,B)}{A}, \frac{\mathrm{PPCM}(A,B)}{B}\right) = 1. \tag{5.3}$$

Notons λ l'inverse du coefficient du terme dominant de AB. Alors λAB est unitaire, et la proposition 5.3.14 combinée avec (5.3) montre que

$$\mathrm{PGCD}\left(\lambda AB\left(\frac{\mathrm{PPCM}(A,B)}{A}\right), \lambda AB\left(\frac{\mathrm{PPCM}(A,B)}{B}\right)\right) = \lambda AB.$$

En appliquant la proposition 5.3.3, on constate que le membre de gauche de cette égalité vaut $\mathrm{PPCM}(A,B)\,\mathrm{PGCD}(A,B)$. ■

5.3.4 Polynômes irréductibles

Au cours des sections qui précèdent, le lecteur a pu constater que l'ensemble $\mathbb{K}[X]$ avait beaucoup de similarités avec l'ensemble $\mathbb{Z}$ des entiers relatifs : les deux ensembles sont des anneaux principaux intègres sur lesquels on peut définir la division euclidienne, le PGCD et le PPCM. Dans cette section, nous allons introduire une classe de polynômes qui jouent dans $\mathbb{K}[X]$ le même rôle que les nombres premiers dans $\mathbb{Z}$: les polynômes irréductibles.

Définition 5.3.17 *On dit qu'un polynôme P est* **irréductible** *si ses seuls diviseurs sont les constantes et les polynômes qui lui sont associés.*

Remarques :

1. À la différence des nombres premiers, les polynômes irréductibles ont une infinité de diviseurs. Mais on notera que ces diviseurs sont triviaux !
2. Tout polynôme de degré 1 est irréductible. En effet, soit P de degré 1, et Q un diviseur de P. Alors $\deg Q \in \{0, 1\}$. Si $\deg Q = 0$ alors Q est une constante, si $\deg Q = 1$ alors $\deg Q = \deg P$ donc P et Q sont associés.

La proposition suivante constitue une "loi du tout ou rien" pour la division par les polynômes irréductibles.

Proposition 5.3.18 *Soit A un polynôme et P un polynôme irréductible ne divisant pas A. Alors P est premier avec A.*

Preuve : Soit B un diviseur commun de A et de P. Comme P est irréductible, B doit être constant, ou associé à P. Le deuxième cas est exclus car on a supposé que P ne divisait pas A. Donc B est constant. On a donc bien $\text{PGCD}\,(A, P) = 1$. ■

De même que tout entier possède une décomposition en facteurs premiers, tout polynôme a une décomposition en facteurs irréductibles.

Théorème 5.3.19 (Décomposition en facteurs irréductibles) *Soit P un polynôme non constant. Alors il existe un entier $k \geq 1$, k entiers $\alpha_1, \cdots, \alpha_k$ non nuls, k polynômes irréductibles unitaires $P_1, \cdots, P_k$ deux à deux distincts, et $\lambda \in \mathbb{K}\backslash\{0\}$ tels que*

$$P = \lambda \prod_{i=1}^{k} P_i^{\alpha_i}.$$

Cette décomposition, appelée décomposition en facteurs irréductibles, est unique à ordre des facteurs près.

Preuve : On prouve d'abord l'existence puis l'unicité à ordre des facteurs près.

Existence : Elle se fait par récurrence sur le degré de P.

- Si $\deg P = 1$ alors P est irréductible. On pose $k = 1$, $\alpha_1 = 1$ et l'on prend pour P_1 le polynôme unitaire associé à P. Il est de degré 1 donc irréductible.
- Supposons maintenant que le théorème de décomposition soit valable pour tout polynôme de degré compris entre 1 et n. Soit P de degré $n+1$ et $P' \overset{\text{déf}}{=} P/\lambda$ avec λ coefficient du terme dominant de P. Le polynôme P' est unitaire et de degré $n+1$. S'il est irréductible, $P = \lambda P'$ constitue une décomposition de P en facteurs premiers. Sinon, il existe un polynôme A unitaire de degré compris entre 1 et n et divisant P'. On a donc $P' = AB$ avec A et B unitaires et de degré compris entre 1 et

n. D'après l'hypothèse de récurrence, A et B admettent chacun une décomposition en facteurs premiers :

$$A = \prod_{i=1}^{k} A_i^{\alpha_i} \quad \text{et} \quad B = \prod_{i=1}^{\ell} B_i^{\beta_i}.$$

Donc

$$P = \lambda \left(\prod_{i=1}^{k} A_i^{\alpha_i} \right) \left(\prod_{i=1}^{\ell} B_i^{\beta_i} \right).$$

Il ne reste plus qu'à renuméroter les facteurs de la décomposition pour obtenir le résultat voulu.

Unicité : Supposons que P admette *deux* décompositions en facteurs irréductibles :

$$P = \lambda \prod_{i=1}^{k} P_i^{\alpha_i} = \mu \prod_{i=1}^{\ell} Q_i^{\beta_i}.$$

Comme tous les facteurs irréductibles sont unitaires, λ et μ sont égaux au coefficient du terme dominant de P. Donc $\lambda = \mu$. De ce fait, on a

$$\prod_{i=1}^{k} P_i^{\alpha_i} = \prod_{i=1}^{\ell} Q_i^{\beta_i}. \tag{5.4}$$

Par ailleurs, P_1 divise la somme de droite. De la remarque 5.3.10, on déduit que P_1 n'est pas premier avec au moins un des Q_j : il existe j_1 tel que Q_{j_1} et P_1 ne soient pas premiers entre eux. Comme par ailleurs Q_{j_1} et P_1 sont irréductibles et unitaires, cela signifie que $P_1 = Q_{j_1}$. En vertu du caractère intègre de $\mathbb{K}[X]$, on peut donc simplifier l'expression (5.4) par P_1. On itère ce procédé et en $\alpha_1 + \cdots + \alpha_k$ étapes, on parvient à une expression du type $1 = \prod_{j=1}^{\ell} Q_j^{\beta'_j}$ avec $\beta'_j = \beta_j - \alpha_j$. Cela permet de conclure que tous les β'_j sont nuls. Donc les deux décompositions sont identiques à ordre près des facteurs.

■

5.4 Fonctions polynômes

5.4.1 Définition des fonctions polynômes

Jusqu'à présent, nous avons traité les polynômes comme des objets algébriques "abstraits". Ce point de vue permet de manipuler de façon unifiée des objets mathématiques très différents dès lors qu'ils peuvent être interprétés comme des polynômes. Dans cette section, nous allons nous borner à remplacer la variable muette X par des nombres réels ou complexes. Mais vous verrez en deuxième année que l'on peut fort bien remplacer X par une matrice...

Définition 5.4.1 *Soit $P = a_n X^n + \cdots + a_1 X + a_0$ un polynôme de $\mathbb{K}[X]$, et $t \in \mathbb{K}$. On définit alors l'élément $P(t)$ de $\mathbb{K}$ par*

$$P(t) = a_n t^n + \cdots + a_1 t + a_0.$$

On dit que $P(t)$ est obtenu par substitution de t à X.

Proposition 5.4.2 *Soit $t \in \mathbb{K}$ un scalaire fixé. Alors on a pour tous polynômes P et Q, et pour tout scalaire λ :*

1. $P(t) + Q(t) = (P + Q)(t)$,
2. $P(t)Q(t) = (PQ)(t)$,
3. $\lambda P(t) = (\lambda P)(t)$,
4. $1(t) = 1$.

Preuve : Vérifions la deuxième relation. Les autres sont immédiates.

Rappelons que si $P = a_p X^p + \cdots + a_1 X + a_0$ et $Q = b_q X^q + \cdots + b_1 X + b_0$ alors

$$PQ = \sum_{j=0}^{p+q} \Big(\sum_{k+\ell=j} a_k b_\ell \Big) X^j. \tag{5.5}$$

Donc

$$\begin{array}{rcl} (PQ)(t) & = & \sum_{j=0}^{p+q} \left(\sum_{k+\ell=j} a_k b_\ell \right) t^j, \\ & = & \sum_{j=0}^{p+q} \left(\sum_{k+\ell=j} (a_k t^k)(b_\ell t^\ell) \right), \\ & = & \left(\sum_{k=0}^{p} a_k t^k \right) \left(\sum_{\ell=0}^{q} b_\ell t^\ell \right) = P(t)Q(t). \end{array}$$

■

Définition 5.4.3 *Soit $P \in \mathbb{K}[X]$. L'application*

$$\widetilde{P} \ : \ \left\{ \begin{array}{ccc} \mathbb{K} & \longrightarrow & \mathbb{K} \\ t & \longmapsto & P(t) \end{array} \right.$$

est appelée fonction polynôme définie par P sur $\mathbb{K}$.

Remarque : Dans la suite du cours, on ne fera plus la distinction entre le polynôme P qui est un objet algébrique et la fonction polynôme $\widetilde{P}$ qui lui est associée[4].

5.4.2 Racines

Définition 5.4.4 *Soit $a \in \mathbb{K}$ et $P \in \mathbb{K}[X]$. On dit que a est* **racine** *ou* **zéro** *de P si $P(a) = 0$.*

Proposition 5.4.5 *Soit $a \in \mathbb{K}$ et $P \in \mathbb{K}[X]$. Pour que a soit une racine de P, il faut et il suffit que $X - a$ divise P.*

Preuve : $\Rightarrow$ Supposons que $P(a) = 0$. La division euclidienne de P par $X - a$ donne

$$P = Q(X - a) + R \quad \text{avec} \quad \deg R \leq 0.$$

En substituant a à X dans la relation ci-dessus, on trouve $R(a) = 0$. Comme la fonction polynôme R est constante, on conclut que $R = 0$.

$\Leftarrow$ Si $X - a \mid P$ alors il existe Q tel que $P = Q(X - a)$, ce qui donne en particulier $P(a) = Q(a)(a - a) = 0$. ■

Définition 5.4.6 *Soit $P \in \mathbb{K}[X]$, $a \in \mathbb{K}$ et $k \in \mathbb{N}^*$. On dit que a est racine de P de multiplicité k si $(X - a)^k \mid P$.*

- *Si $k = 1$, on parle de racine simple,*
- *Si $k = 2$, on dit que a est racine double,*
- *Si $k = 3$, on dit que a est racine triple, etc.*

[4] La proposition 5.4.2 nous autorise à faire cet abus de notation.

Proposition 5.4.7 *Soit P un polynôme non nul admettant les racines $a_1, \cdots, a_k$ avec multiplicité $\alpha_1, \cdots, \alpha_k$. Alors $\prod_{i=1}^{k}(X-a_i)^{\alpha_i}$ divise P.*

Preuve :

- On sait déjà que $(X-a_1)^{\alpha_1}$ divise P.
- Supposons que $\prod_{i=1}^{j-1}(X-a_i)^{\alpha_i}$ divise P (avec $j \leq k$). Comme les a_i sont deux à deux distincts, les polynômes $(X-a_i)^{\alpha_i}$ sont premiers entre eux deux à deux. La remarque 5.3.10 permet donc d'affirmer que $(X-a_j)^{\alpha_j}$ est premier avec $\prod_{i=1}^{j-1}(X-a_i)^{\alpha_i}$. Comme P est multiple de $(X-a_j)^{\alpha_j}$ par hypothèse, et de $\prod_{i=1}^{j-1}(X-a_i)^{\alpha_i}$, P est également multiple du PPCM de ces deux polynômes qui, d'après la proposition 5.3.16, vaut $\prod_{i=1}^{j}(X-a_i)^{\alpha_i}$.

Nous venons donc de montrer par récurrence sur j que $\prod_{i=1}^{k}(X-a_i)^{\alpha_i}$ divise P. ■

Remarque 5.4.8 *En particulier, si $P \neq 0$, toutes les racines de P sont de multiplicité inférieure ou égale à* $\deg P$.

Exercice : Justifier la remarque 5.4.8.

Proposition 5.4.9 *Un polynôme de degré $n \in \mathbb{N}$ admet au plus n racines comptées avec leur ordre de multiplicité : $\{a_1, \cdots, a_k\}$ est l'ensemble des racines de P, et α_i est la multiplicité de a_i, alors on a $\alpha_1 + \cdots + \alpha_k \leq n$.*

Preuve : D'après la proposition 5.4.8, on a $\prod_{i=1}^{k}(X-a_i)^{\alpha_i} \mid P$. Donc

$$\sum_{i=1}^{k} \deg(X-a_i)^{\alpha_i} \leq \deg P.$$

Le membre de gauche vaut $\sum_{i=1}^{k}\alpha_i$, d'où le résultat. ■

Remarque 5.4.10 *Le seul polynôme ayant une infinité de racines est le polynôme nul.*

5.4.3 Polynômes dérivés

Définition 5.4.11 *Soit $P = a_kX^k + \cdots + a_1X + a_0$ un polynôme de $\mathbb{K}[X]$. On appelle* **polynôme dérivé** *noté P' le polynôme suivant :*

$$P' = ka_kX^{k-1} + \cdots + a_1 = \sum_{j=1}^{k} ja_jX^{j-1}.$$

Proposition 5.4.12 *Soit P et Q deux polynômes, et $\lambda \in \mathbb{K}$.*

1. *Si $\deg P > 0$ alors $\deg P' = \deg P - 1$,*
2. *Si P est constant alors $P' = 0$,*
3. *$(P+Q)' = P' + Q'$,*
4. *$(\lambda P)' = \lambda P'$,*
5. *$(PQ)' = P'Q + PQ'$.*

Preuve : Les quatre premiers points sont évidents. Prouvons le cinquième.

Soit $P = a_pX^p + \cdots + a_1X + a_0$ et $Q = b_qX^q + \cdots + b_1X + b_0$. En appliquant la définition du polynôme dérivé à la relation (5.5), on trouve

$$(PQ)' = \sum_{j=1}^{p+q} j\Big(\sum_{k+\ell=j} a_kb_\ell\Big)X^{j-1}.$$

Des calculs élémentaires montrent donc que

$$\begin{aligned}(PQ)' &= \sum_{j=1}^{p+q}\sum_{k+\ell=j}\Big(ka_kX^{k-1}b_\ell X^\ell + a_kX^k\ell b_\ell X^{\ell-1}\Big),\\ &= \sum_{j=1}^{p+q}\Big(\sum_{k+\ell=j} ka_kX^{k-1}\, b_\ell X^\ell\Big) + \sum_{j=1}^{p+q}\Big(\sum_{k+\ell=j} a_kX^k\,\ell b_\ell X^{\ell-1}\Big),\\ &= \Big(\sum_{k=1}^{p} ka_kX^{k-1}\Big)\Big(\sum_{\ell=0}^{q} b_\ell X^\ell\Big) + \Big(\sum_{k=0}^{p} a_kX^k\Big)\Big(\sum_{\ell=1}^{q}\ell b_\ell X^{\ell-1}\Big),\\ &= P'Q + PQ'.\end{aligned}$$

■

Proposition 5.4.13 *Soit P un polynôme non nul, et a une racine de P. Alors a est une racine simple si et seulement si $P'(a) \neq 0$.*

Preuve : Nous allons prouver la négation de l'équivalence : i.e a est une racine double de P si et seulement si $P(a) = P'(a) = 0$.

Supposons donc que a est une racine double de P. Alors $(X-a)^2 \mid P$. Donc P s'écrit $P = Q(X-a)^2$ pour un certain polynôme Q. Il est donc immédiat que $P(a) = 0$. En dérivant, on trouve $P' = Q'(X-a)^2 + 2(X-a)Q$, donc $P'(a) = 0$.

Réciproquement, supposons que $P(a) = P'(a) = 0$. La division euclidienne de P par $(X-a)^2$ s'écrit $P = Q(X-a)^2 + R$ avec $\deg R \leq 1$. Comme $P(a) = 0$, on a $R(a) = 0$. En dérivant la relation $P = Q(X-a)^2 + R$, on obtient $R'(a) = 0$. Comme R' est un polynôme constant, on a $R' = 0$, puis, comme $R(a) = 0$, R est nul aussi. ■

5.5 Polynômes scindés

5.5.1 Le théorème fondamental de l'algèbre

Définition 5.5.1 *On dit qu'un polynôme non constant est scindé si la somme des ordres de multiplicité de ses racines est égal à son degré.*

Remarque : Autrement dit, P *de degré* n est scindé si et seulement si il existe un n-uplet $(\lambda_1, \cdots, \lambda_n)$ de $\mathbb{K}^n$ tel que P soit associé à $(X-\lambda_1)\cdots(X-\lambda_n)$.

Proposition 5.5.2 *Soit P un polynôme scindé unitaire d'expression $X^n + a_{n-1}X^{n-1} + \cdots + a_0$. Notons λ_i ses racines comptées avec leur ordre de multiplicité. Alors on a les relations suivantes entre les racines et les coefficients :*

$$a_0 = (-1)^n \prod_{i=1}^{n} \lambda_i \quad \textit{et} \quad a_{n-1} = -\sum_{i=1}^{n} \lambda_i.$$

Preuve : On développe l'expression $(X-\lambda_1)\cdots(X-\lambda_n)$ et on identifie les termes du développement avec ceux de l'expression $X^n + a_{n-1}X^{n-1} + \cdots + a_0$. ■

Remarque : Dans le cas où $P = X^2 + a_1X + a_0$ a pour racines λ_1 et λ_2, on retrouve les relations

$$a_0 = \lambda_1\lambda_2 \quad \text{et} \quad a_1 = -(\lambda_1 + \lambda_2).$$

Le très important résultat suivant est connu sous le nom de **théorème fondamental de l'algèbre** ou **théorème de d'Alembert-Gauss**. Il en existe de nombreuses preuves, mais toutes dépassent le cadre du programme.

Théorème 5.5.3 *Tout polynôme de $\mathbb{C}[X]$ est scindé*[5].

[5] On dit que $\mathbb{C}$ est un **corps algébriquement clos.**

Remarque : On a vu que toutes les équations de degré 2 avaient deux solutions (éventuellement confondues) dans $\mathbb{C}$. Le théorème fondamental exprime que toute équation de degré n admet n solutions (éventuellement confondues) dans $\mathbb{C}$. Dans le cas $n = 3$ ou 4, il existe des formules (assez compliquées) donnant les solutions en fonction des coefficients. Pour une équation de degré supérieur ou égal à 5, il a été prouvé par un jeune mathématicien du XIX ème siècle, E. Galois, que de telles formules n'existent pas !

5.5.2 Polynômes irréductibles de $\mathbb{C}[X]$

Théorème 5.5.4 *Un polynôme P est irréductible dans $\mathbb{C}$ si et seulement si* $\deg P = 1$.

Preuve : On a déjà vu que tout polynôme de degré 1 était irréductible (que ce soit dans $\mathbb{C}$ ou dans $\mathbb{R}$).

Pour montrer la réciproque, donnons-nous un polynôme P de degré au moins 2. Le théorème fondamental de l'algèbre nous dit que P admet au moins une racine λ_1. Donc P est divisible par $X - \lambda_1$. Clairement $X - \lambda_1$ n'est pas constant et n'est pas associé à P car de degré strictement inférieur à 2. Donc P n'est pas irréductible. ■

En appliquant le théorème général de décomposition irréductible, on en déduit :

Corollaire 5.5.5 *Tout polynôme P non nul de $\mathbb{C}[X]$ admet une décomposition en facteurs irréductibles du type suivant :*

$$P = \lambda \prod_{i=1}^{k} (X - \lambda_i)^{\alpha_i},$$

où $\{\lambda_1, \cdots, \lambda_k\}$ est l'ensemble des racines de P, α_i est la multiplicité de λ_i, et λ est le coefficient du terme dominant de P.

5.5.3 Polynômes irréductibles de $\mathbb{R}[X]$

Dans $\mathbb{R}[X]$, la situation est un peu plus compliquée. On sait d'ores et déjà que tous les polynômes irréductibles ne sont pas de degré 1. Par exemple, $X^2 + 1$ ne saurait être irréductible dans $\mathbb{R}[X]$ car n'a pas de racine réelle (la fonction polynôme associée est minorée par 1, donc ne s'annule jamais).

On peut cependant dresser une liste de tous les polynômes irréductibles de $\mathbb{R}[X]$:

Théorème 5.5.6 *Les polynômes irréductibles de $\mathbb{R}[X]$ sont :*

- *Les polynômes de degré 1,*
- *Les polynômes de degré 2 à discriminant strictement négatif : $P = aX^2 + bX + c$ avec $a \neq 0$ et $\Delta \overset{\text{déf}}{=} b^2 - 4ac < 0$.*

La preuve de ce théorème repose sur le lemme suivant :

Lemme 5.5.7 *Soit $P = \sum_{k=0}^{n} a_k X^k$ un polynôme de $\mathbb{C}[X]$. Notons $\overline{P} = \sum_{k=0}^{n} \overline{a_k} X^k$ le polynôme conjugué. Alors λ est racine de P de multiplicité α si et seulement si $\overline{\lambda}$ est racine de $\overline{P}$ de multiplicité α.*

Preuve : Soit λ une racine de P de multiplicité α. Alors il existe un polynôme Q tel que $P = Q(X - \lambda)^\alpha$. En prenant le conjugué de cette expression, on obtient $\overline{P} = \overline{Q}(X - \overline{\lambda})^\alpha$. Donc $\overline{\lambda}$ est racine de $\overline{P}$ de multiplicité $\overline{\alpha} \geq \alpha$.

En échangeant les rôles de P et $\overline{P}$, λ et $\overline{\lambda}$, α et $\overline{\alpha}$, on obtient $\overline{\alpha} \leq \alpha$, d'où le résultat. ■

Preuve du théorème 5.5.6 :

On sait déjà que les polynômes de degré 1 sont irréductibles. Soit maintenant $P = aX^2 + bX + c$ à discriminant strictement négatif. La fonction $t \mapsto P(t)$ associée ne s'annule pas sur $\mathbb{R}$ (elle est du signe de a), et donc aucun polynôme de degré 1 ne saurait diviser P. Par ailleurs, on a vu dans le chapitre 1 que toute équation de degré 2 à coefficients réels et discriminant positif ou nul admettait au moins une solution réelle. Donc les polynômes de degré 2 à discriminant positif ne sont pas irréductibles dans $\mathbb{R}[X]$.

Soit maintenant $P \in \mathbb{R}[X]$ un polynôme de degré au moins 3. Supposons que P n'ait pas de racine réelle (sinon P n'est pas irréductible dans $\mathbb{R}[X]$). D'après le lemme 5.5.7, les racines complexes non réelles de P sont deux à deux conjuguées (avec ordres de multiplicité égaux deux à deux). Le corollaire 5.5.5 assure donc l'existence de nombres complexes (non réels) $\mu_1, \cdots, \mu_p$, d'entiers $\alpha_1, \cdots, \alpha_p$, et d'un réel α, tels que

$$P = \alpha \prod_{i=1}^{p} \Big[(X - \mu_i)^{\alpha_i} (X - \bar{\mu}_i)^{\alpha_i} \Big].$$

Mais un calcul facile montre que

$$(X - \mu_i)^{\alpha_i} (X - \bar{\mu}_i)^{\alpha_i} = (X^2 - 2\mathcal{R}e\,\mu_i\, X + |\mu_i|^2)^{\alpha_i}$$

Donc P est divisible par le polynôme *réel* $X^2 - 2\mathcal{R}e\,\mu_i\, X + |\mu_i|^2$ (de degré 2) et n'est donc pas irréductible. ■

En reprenant la preuve ci-dessus, on déduit facilement le résultat suivant.

Corollaire 5.5.8 *Tout polynôme à coefficients réels admet dans $\mathbb{R}[X]$ une décomposition en facteurs irréductibles du type suivant :*

$$P = \lambda \Big(\prod_{i=1}^{k} (X - \lambda_i)^{\alpha_i} \Big) \Big(\prod_{j=1}^{\ell} (X^2 - 2\mathcal{R}e\,\mu_j\, X + |\mu_j|^2)^{\beta_j} \Big),$$

où λ est le coefficient du terme dominant de P, $\{\lambda_1, \cdots, \lambda_k\}$ est l'ensemble des racines réelles de P, α_i, multiplicité de λ_i, et $\{\mu_1, \cdots, \mu_\ell\}$ est l'ensemble des racines complexes et non réelles de P et β_j, la multiplicité de μ_j.

Bibliographie

[1] J.-M. Arnaudiès et H. Fraysse : Cours de mathématiques 1, Algèbre, **Dunod**.

[2] J.-M. Arnaudiès et J. Lelong-Ferrand : Cours de mathématiques, tome 1 (algèbre), **Dunod**.

[3] T. Cuesta : Algèbre 1, MIAS1, Polycopié de l'Université Paris XII.

[4] C. Deschamps et A. Warusfel : Mathématiques. Cours et exercices corrigés, **Dunod**.

[5] J. Dixmier : Cours de mathématiques du premier cycle, **Gauthier-Villars**.

[6] S. Lipschutz : Algèbre linéaire, cours et problèmes, **Série Schaum**.

[7] F. Liret et D. Martinais : Mathématiques pour le DEUG, Algèbre, 1ère année, **Dunod**.

[8] J.-M. Monier : Algèbre et géométrie, 1ère année, **Dunod**.

[9] E. Ramis, C. Deschamps et J. Odoux : Cours de mathématiques spéciales, algèbre, Vol. 1, **Masson**.

Index