COURS OUTILS MATHEMATIQUES POUR L'INGENIEUR 1

MECANIQUE 3, Polytech Lyon, 2016-2017

Ionel Sorin CIUPERCA

Partie 2

Table des matières

Chapitre 1

Algèbre linéaire et calcul matriciel

1.1 Rappels sur les espaces vectoriels

Dans tout ce chapitre n désigne un nombre naturel non-nul, donc $n \in \mathbb{N}^*$; la lettre $\mathbb{K}$ désigne $\mathbb{R}$ (l'ensemble des nombres réels) ou $\mathbb{C}$ (l'ensemble des nombres complexes).

1. $\mathbb{C} = \{a + bi,\ a, b \in \mathbb{R}\}$ où $i^2 = -1$.
   Sur $\mathbb{C}$ nous avons les opérations $+$ et $\cdot$ habituelles :
   $$(a_1 + b_1 i) + (a_2 + b_2 i) = (a_1 + a_2) + (b_1 + b_2)i \quad \forall\ a_1, a_2, b_1, b_2 \in \mathbb{R}$$
   $$(a_1 + b_1 i) \cdot (a_2 + b_2 i) = (a_1 a_2 - b_1 b_2) + (a_1 b_2 + a_2 b_1)i \quad \forall\ a_1, a_2, b_1, b_2 \in \mathbb{R}.$$
   Rappellons le **conjugué** d'un nombre complexe :
   $$\overline{a + bi} = a - bi, \quad \forall\ a, b \in \mathbb{R}.$$
2. Pour tous $j, k \in \mathbb{Z}$ on note
   $$\delta_{jk} = \begin{cases} 1 & \text{si} \quad j = k \\ 0 & \text{si} \quad j \neq k; \end{cases}$$
   ce sont les **symboles de Kronecker**.
3. La notion de produit scalaire s'étend à des vecteurs complexes :
   si $x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_n) \in \mathbb{C}^n$ alors le produit scalaire entre x et y est par définition
   $$x \cdot y = < x\, ,\, y > = \sum_{j=1}^{n} x_j \overline{y_j}.$$
4. La lettre Φ désigne l'ensemble vide.
5. On utilisera la notation commode suivante : pour tous $m, p \in \mathbb{Z}$ avec $m \leq p$ on va noter par $[[m,\, p]]$ l'ensemble des nombres entiers compris entre m et p (inclusivement), c'est à dire : $[[m,\, p]] = \{m, m + 1, \cdots p - 1, p\}$.

1.1.1 Espaces vectoriels : définitions et exemples

On commence par la définition suivante :

Définition 1.1. *On considère E un ensemble non-vide et deux applications (on peut dire aussi opérations) notées respectivement par $+$ et $\cdot$*

$$\begin{array}{rl} + : & E \times E \;\to\; E \\ & (u,v) \mapsto u+v \qquad \textit{dite addition} \\ \cdot : & \mathbb{K} \times E \;\to\; E \\ & (\lambda, v) \mapsto \lambda \cdot v \qquad \textit{dite multiplication par un scalaire.} \end{array}$$

On dit que le triplet $(E, +, \cdot)$ est un **espace vectoriel** *sur $\mathbb{K}$ si les 8 propriétés suivantes sont satisfaites :*

1. $u + v = v + u \quad \forall\, u, v \in E$ *(commutativité de $+$)*
2. $(u + v) + w = u + (v + w) \quad \forall\, u, v, w \in E$ *(associativité de $+$)*
3. $\exists\, 0_E \in E$ *tel que :* $u + 0_E = 0_E + u = u \quad \forall\, u \in E$ *(existence d'un élément neutre 0_E dans E pour $+$)*
4. $\forall\, u \in E \quad \exists\, v \in E$ *tel que* $u + v = 0_E$ *(tout élément de E a un symétrique v par rapport à $+$; cet élément est noté $-u$)*
5. $\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v \quad \forall\, \lambda \in \mathbb{K} \quad \forall\, u, v \in E$ *(distributivité de $\cdot$ par rapport à $+$)*
6. $1 \cdot u = u$ *et* $0 \cdot u = 0_E \qquad \forall\, u \in E.$
7. $(\lambda + \mu) \cdot u = \lambda \cdot u + \mu \cdot u \quad \forall\, \lambda, \mu \in \mathbb{K} \quad \forall\, u \in E$ *(distributivité de $\cdot$ par rapport à l'addition des scalaires)*
8. $(\lambda\, \mu) \cdot u = \lambda \cdot (\mu \cdot u) \qquad \forall\, \lambda, \mu \in \mathbb{K} \quad \forall\, u \in E.$

On dira encore que E est un $\mathbb{K}$-espace vectoriel (en abrégé $\mathbb{K}$-e.v.). Les éléments de E sont appelés **vecteurs***, et les éléments de $\mathbb{K}$ sont appelés* **scalaires***.*

Remarques :

1. Les propriétés de (1) à (8) sont bien connues dans le cas où on considère $E = \mathbb{K}$ muni des opérations $+$ et $\cdot$ habituelles en $\mathbb{K}$. Ceci montre que $\mathbb{K}$ est un $\mathbb{K}$-e.v.
2. Pour simplicité on va noter souvent λx à la place de $\lambda \cdot x$ pour $\lambda \in \mathbb{K}$, $x \in E$. On notera souvent 0 à la place de O_E.

Exemple 1.1. *(fondamental)*
Pour tout $n \in \mathbb{N}^*$ on pose

$$\mathbb{K}^n = \mathbb{K} \times \mathbb{K} \times \cdots \times \mathbb{K} \;\; (n \textit{ fois}) \; = \{(u_1, u_2, \cdots u_n) \quad u_1, u_2, \cdots u_n \in \mathbb{K}\}.$$

On introduit sur $\mathbb{K}^n$ les opérations $+$ et $\cdot$ suivantes :
$u + v = (u_1 + v_1, u_2 + v_2, \cdots u_n + v_n) \in \mathbb{K}^n$
$\lambda \cdot u = (\lambda u_1, \lambda u_2, \cdots \lambda u_n) \in \mathbb{K}^n$

ceci pour tous $u = (u_1, u_2, \cdots u_n)$, $v = (v_1, v_2, \cdots v_n) \in \mathbb{K}^n$ *et pour tout* $\lambda \in \mathbb{K}$.
Alors $\mathbb{K}^n$ *muni de ces deux opérations est un* $\mathbb{K}$*-e.v.*

Preuve :

La preuve est basée sur le fait que toutes les propriétés de (1) à (8) de la Définition 1 sont vraies pour u, v, w *éléments de* $\mathbb{K}$*. Il faut alors raisonner composante par composante des vecteurs.*
Par exemple, montrons la propriété (1) de la définition :
$(u + v) = (u_1 + v_1,\, u_2 + v_2,\, \cdots,\, u_n + v_n)$ *et*
$(v + u) = (v_1 + u_1,\, v_2 + u_2,\, \cdots,\, v_n + u_n)$.
On a alors l'égalité entre $u + v$ *et* $v + u$ *parce qu'on a l'égalité entre* $u_k + v_k$ *et* $v_k + u_k$ *pour tout* $k \in [[1,\, n]]$ *(commutativité sur* $\mathbb{K}$*).*
La preuve des autres propriétés (de (2) à (8)) est analogue et est laissée en exercice ; à remarquer que $0_{\mathbb{K}^n} = (0, 0, \cdots 0)$.
On a donc montré que $\mathbb{R}^n$ *est une espace vectoriel sur* $\mathbb{R}$ *et* $\mathbb{C}^n$ *est une espace vectoriel sur* $\mathbb{C}$*. Pour* $n = 1$ *on retrouve bien que* $\mathbb{R}$ *est une espace vectoriel sur* $\mathbb{R}$ *et* $\mathbb{C}$ *est une espace vectoriel sur* $\mathbb{C}$*.*

Exemple 1.2. *(fondamental)*
Soit $U \neq \Phi$ *un ensemble quelconque. On introduit l'ensemble suivant :*
$\mathcal{F}(U, \mathbb{K}) =$ *l'ensemble des fonctions définies sur* U *et à valeurs dans* $\mathbb{K}$*, donc*
$\mathcal{F}(U, \mathbb{K}) = \{f, \quad f : U \mapsto \mathbb{K}\}$.
On introduit sur $\mathcal{F}(U, \mathbb{K})$ *les opérations* $+$ *et* $\cdot$ *suivantes :*

$$(f + g)(x) = f(x) + g(x) \quad \forall\, x \in U \qquad\quad \forall\, f, g \in \mathcal{F}(U, \mathbb{K})$$

et

$$(\lambda \cdot f)(x) = \lambda f(x) \quad \forall\, x \in U \qquad\quad \forall\, f \in \mathcal{F}(U, \mathbb{K}) \quad \forall \lambda \in \mathbb{K}.$$

Alors $\mathcal{F}(U, \mathbb{K})$ *muni de ces deux opérations est un* $\mathbb{K}$ *- e.v.*

Preuve :

Comme dans l'Exemple 1 on montrera seulement la propriété (1) de la Définition 1, les autres propriéés étant laissées en exercice.
Soient $f, g \in \mathcal{F}(U, \mathbb{K})$ *arbitraires. Grâce à la commutativité sur* $\mathbb{K}$ *nous avons*
$f(x) + g(x) = g(x) + f(x) \quad \forall\, x \in U$*, ce qui nous donne* $f + g = g + f$*, donc le résultat recherché.*
On remarque que ici les "vecteurs" sont en fait des fonctions de U *dans* $\mathbb{K}$*.*
On remarque aussi que l'élément 0 *de cet espace vectoriel est la fonction constante égale à zero.*

Exemple 1.3. *(fondamental)*
Soient $m, n \in \mathbb{N}^*$*. Rappelons qu'un* **matrice** A *sur* $\mathbb{K}$ *à* m *lignes et* n *colonnes est un*

tableau de $m \times n$ éléments de $\mathbb{K}$ de la forme

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdot & \cdot & A_{1n} \\ A_{21} & A_{22} & \cdot & \cdot & A_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ A_{m1} & A_{m2} & \cdot & \cdot & A_{mn} \end{pmatrix}$$

avec $A_{ij} \in \mathbb{K}$ pour tous $i \in [[1,\, m]]$ et $j \in [[1,\, n]]$.
On notera aussi $A = (A_{ij})_{i=1\cdots m, j=1\cdots n}$.
Si $m = n$ on dira que A est une matrice **carrée**.
On va noter par $\mathcal{M}_{m,n}(\mathbb{K})$ l'ensemble des matrices à m lignes et n collones sur $\mathbb{K}$; si $m = n$ on notera $\mathcal{M}_n(\mathbb{K})$ à la place de $\mathcal{M}_{n,n}(\mathbb{K})$.
Sur $\mathcal{M}_{m,n}(\mathbb{K})$ on définit les opérations $+$ et $\cdot$ suivants :

1. *Si $A = (A_{ij})$ et $B = (B_{ij})$ alors on définit*

$$A + B = C \quad \text{avec} \quad C_{ij} = A_{ij} + B_{ij} \quad \forall\, i \in [[1,\, m]],\ j \in [[1,\, n]].$$

2. *Si $A = (A_{ij})$ et $\lambda \in \mathbb{K}$ alors on définit*

$$\lambda \cdot A = D \quad \text{avec} \quad D_{ij} = \lambda A_{ij} \quad \forall\, i \in [[1,\, m]],\ j \in [[1,\, n]].$$

On montre facilement (Exercice) que $\mathcal{M}_{mn}(\mathbb{K})$ muni de ces deux opérations est un $\mathbb{K}$-e.v. On remarque que l'élément 0 de cet espace vectoriel est la matrice 0, dont tous les éléments sont 0.

Remarque : Il y a encore une opération de multiplication $\cdot$ entre matrices à ne pas confondre avec la multiplication entre un scalaire et une matrice. Cette multiplication entre matrices se définit de la manière suivante :
Si $m, n, p \in \mathbb{N}^*$, $\quad A \in \mathcal{M}_{m,n}(\mathbb{K})$ et $B \in \mathcal{M}_{n,p}(\mathbb{K})$ alors on définit $A \cdot B$ (on écrira souvent AB) comme la matrice dans $\mathcal{M}_{m,p}(\mathbb{K})$ définie par

$$(A \cdot B)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj} \quad \forall\, i \in [[1,\, m]],\ j \in [[1,\, p]].$$

Rappelons aussi la propriété **d'associativité** du produit des matrices :

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

(à conditions que les produits de matrices respectives soient possibles).
Notons encore que si $A \in \mathcal{M}_{m,n}(\mathbb{K})$ et $x \in \mathbb{K}^n$ est un vecteur colonne, ce vecteur colonne peut être vu comme une matrice à n lignes et une colonne. Alors la multiplication Ax (très utilisée dans ce cours) a du sens et le résultat est un vecteur colonne avec m lignes ; on a

$$(Ax)_i = \sum_{k=1}^{n} A_{ik} x_k, \quad \forall\, i \in [[1,\, m]].$$

Un contre-exemple :
On pose $E = [0,1] \subset \mathbb{R}$ muni des opérations $+$ et $\cdot$ habituelles sur $\mathbb{R}$. Observons que $1 \in E$ mais $1+1=2 \notin E$. Ceci montre que E n'est pas un espace vectoriel sur $\mathbb{R}$.

Définition 1.2. *Soit E un $\mathbb{K}$-e.v. et $m \in \mathbb{N}^*$. Soient $u_1, u_2, \cdots u_m \in E$ et $\lambda_1, \lambda_2, \cdots \lambda_m \in \mathbb{K}$. Le vecteur $\lambda_1 u_1 + \lambda_2 u_2 + \cdots \lambda_m u_m \in E$ s'appelle* **combinaison linéaire** *des vecteurs $u_1, u_2, \cdots u_m$ (on peut ajouter : de coefficients $\lambda_1, \lambda_2, \cdots \lambda_m$).*

Par exemple $u+v$ ou $u-3v$ sont des combinaisons linéaires des vecteurs u et v.
Remarque :
On peut voir un vecteur u comme combinaison linéaire de u (lui même), car $u = 1 \cdot u$.

Définition 1.3. *Soit E un $\mathbb{K}$-e.v. et $F \subset E$. On dit que F est un* **sous-espace vectoriel** *de E (on écrira F est un s.e.v. de E) s'il vérifie les deux propriétés suivantes :*

1. $0_E \in F$
   (donc forcément F est un ensemble non-vide)
2. *F est* **stable** *par rapport aux opérations $+$ et $\cdot$, c'est à dire :*
   $u + v \in F, \quad \forall\, u, v \in F$
   et
   $\lambda u \in F, \quad \forall\, \lambda \in \mathbb{K}, \quad \forall\, u \in F.$

Remarque :
Il est facile de voir que la propriété de stabilité de la Définition 1.3 est équivalente à la propriété suivante :

$$\lambda u + \mu v \in F, \quad \forall\, \lambda, \mu \in \mathbb{K}, \quad \forall\, u, v \in F.$$

On dit alors que F est stable par rapport aux combinaisons linéaires.

Exemple 1.4. *(fondamental)*
Si E est un $\mathbb{K}$-e.v. alors

1. *$\{0_E\}$ est un s.e.v. de E*
2. *E est un s.e.v. de E.*

Les preuves sont très faciles et sont laissées en exercice.

Exemple 1.5. *(fondamental)*
Soient $m, n \in \mathbb{N}^*$ et $A \in \mathcal{M}_{m,n}(\mathbb{K})$ une matrice. Nous introduisons les ensembles suivants :

$$Ker(A) = \{x \in \mathbb{K}^n, \quad Ax = 0\} \subset \mathbb{K}^n$$

(appellé **le noyau** *de la matrice A) et*

$$Im(A) = \{y \in \mathbb{K}^m, \quad \exists x \in \mathbb{K}^n,\ y = Ax\} \subset \mathbb{K}^m$$

(appellé **l'image** *de la matrice A).*
Alors $Ker(A)$ est un sous-espace vectoriel de $\mathbb{K}^n$ et $Im(A)$ est une sous-espace vectoriel de $\mathbb{K}^m$.

Démonstration.

1. Montrons que $Ker(A)$ est un sous-espace vectoriel de $\mathbb{K}^n$. D'abord il est clair que $0 \in Ker(A)$ car $A0 = 0$.
   D'autre part soient $\lambda, \mu \in \mathbb{K}$ et $x, y \in Ker(A)$. Nous avons : $A(\lambda\, x + \mu\, y) = \lambda\, A\, x + \mu\, A\, y = 0$ car $Ax = Ay = 0$. On déduit que $\lambda\, x + \mu\, y \in Ker(A)$ ce qui montre le résultat.
2. Montrons que $Im(A)$ est un sous-espace vectoriel de $\mathbb{K}^m$. D'abord il est clair que $0 \in Im(A)$ car $0 = A0$ et $0 \in \mathbb{K}^n$.

D'autre part soient $\lambda, \mu \in \mathbb{K}$ et $x, y \in Im(A)$. Nous avons $x = Au$ et $y = Av$ avec $u, v \in \mathbb{K}^n$. Ceci nous donne par linéarité :

$$\lambda\, x + \mu\, y = \lambda\, Au + \mu\, Av = A(\lambda u + \mu v) \in Im(A)$$

ce qui finit la preuve.

$\square$

Exemple 1 : On se donne $a_1, a_2 \in \mathbb{R}$ et on pose

$$F_1 = \{(x_1, x_2) \in \mathbb{R}^2 \quad a_1x_1 + a_2x_2 = 0\}.$$

Alors F_1 est un s.e.v. de $\mathbb{R}^2$ car $A = Ker(A)$ avec $A = (a_1 \;\; a_2) \in \mathcal{M}_{1,2}(\mathbb{R})$ et $\mathbb{K} = \mathbb{R}$.
Remarque : L'ensemble F_1 de cet exemple est : soit $\mathbb{R}^2$ si $(a_1, a_2) = (0, 0)$, soit une droite de $\mathbb{R}^2$ passant par l'origine si $(a_1, a_2) \neq (0, 0)$.
Exemple 2 : On pose

$$F_2 = \{(x + 2y - z, 5x - 2z)^T, \quad x, y, z \in \mathbb{R}\}.$$

Alors F_2 est un sous-espace vectoriel de $\mathbb{R}^2$ car $F_2 = Im(A)$ avec $A \in \mathcal{M}_{2,3}(\mathbb{R})$ donnée par

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

Contre-exemple 1 :
On droite de $\mathbb{R}^2$ qui ne passe pas par l'origine n'est pas un s.e.v. de $\mathbb{R}^2$.
Plus exactement, si $b \neq 0$ et si on pose
$F = \{(x_1, x_2) \in \mathbb{R}^2, \quad a_1x_1 + a_2x_2 = b\}$ alors F n'est pas un s.e.v. de $\mathbb{R}^2$.
Preuve : Observer que $(0, 0) \notin \mathbb{R}^2$.
Contre-exemple 2 :
Soit $F = [0, 1] \times [0, 1] \subset \mathbb{R}^2$. Alors F n'est pas un s.e.v. de $\mathbb{R}^2$.
Preuve : Soit $\left(\frac{3}{4}, \frac{3}{4}\right) \in F$. Alors $4\left(\frac{3}{4}, \frac{3}{4}\right) = (3, 3) \notin F$.

Proposition 1.1. *Si E est un $\mathbb{K}$-e.v. et F est un s.e.v. de E alors l'ensemble F muni des mêmes opérations $+$ et $\cdot$ que E est un $\mathbb{K}$-e.v.*

Démonstration. La preuve de ce résultat est très facile, il suffit d'observer que les 8 propriétés de la définition d'un espace vectoriel (Définition 1.1) sont satisfaites pour F. $\square$

Remarque : La proposition précédente nous donne une manière très souvent utilisée pour obtenir des nouveaux espaces vectoriels : il suffit de considérer des sous-espaces vectoriels des espaces vectoriels connus.
Exemple :
On pose

$$F = \mathcal{C}(\mathbb{R}, \mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R}, \; f \text{ continue}\}$$

(donc F est l'ensemble des fonctions continues de $\mathbb{R}$ en $\mathbb{R}$).
Si nous considérons $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$ alors E muni des opérations $+$ et $\cdot$ habituelles sur les fonctions réelles est un espace vectoriel sur $\mathbb{R}$ (c'est un cas particulier $U = \mathbb{K} = \mathbb{R}$ de l'Exemple 1.2). Il est facile de voir que F est un sous-espace vectoriel de E, car :
- la fonction constante 0 est continue
- pour tous $\lambda, \mu \in \mathbb{R}$ et $f, g : \mathbb{R} \to \mathbb{R}$ continues on a $\lambda f + \mu g$ continue.
Alors l'ensemble F muni des opérations $+$ et $\cdot$ sur les fonctions réelles est un espace vectoriel sur $\mathbb{R}$.

1.1.2 Bases et dimension d'un espace vectoriel

Dans cette section E sera toujors un $\mathbb{K}$ - e.v.

Définition 1.4. *Soit $U \subset E$ un sous-ensemble non-vide et finit de E, donc*

$$U = \{u_1, u_2, \cdots, u_p\}$$

où $p \in \mathbb{N}^*$ et tous les éléments $u_1, u_2, \cdots, u_p$ sont dans E. On appelle **sous-espace vectoriel engendré par** *U le sous-ensemble de E noté $Vect(U)$ défini comme étant l'ensemble des toutes les combinaisons linéaires des éléments de U, c'est à dire :*

$$Vect(U) = \left\{ \sum_{j=1}^{p} \alpha_j u_j, \quad \alpha_1, \cdots, \alpha_p \in \mathbb{K} \right\}.$$

Exemple :
Considérons $E = \mathbb{R}^3$, $\mathbb{K} = \mathbb{R}$, $u_1 = (1, 0, 0) \in E$ et $u_2 = (0, 1, 0) \in E$ et posons $U = \{u_1, u_2\}$. Alors un élément arbitraire x de $Vect(U)$ sera une combinaison linéaire entre u_1 et u_2 donc $x = \alpha_1 u_1 + \alpha_2 u_2$ avec $\alpha_1, \alpha_2 \in \mathbb{R}$; ceci donne $x = (\alpha_1, \alpha_2, 0)$. On peut alors écrire

$$Vect(U) = \{(\alpha_1, \alpha_2, 0), \quad \alpha_1, \alpha_2 \in \mathbb{R}\}$$

ce qui donne $Vect(U) = \mathbb{R} \times \mathbb{R} \times \{0\} = \mathbb{R}^2 \times \{0\}$.
Remarque :
Pour simplicité on peut noter $Vect(u_1, u_2, \cdots, u_p)$ à la place de $Vect(\{u_1, u_2, \cdots, u_p\})$.
On a le résultat suivant :

Proposition 1.2.
1. *Pour tout sous-ensemble U non-vide et finit de E, $Vect(U)$ est un sous-espace vectoriel de E*
   (d'où le nom : "sous-espace vectoriel engendré par" U).
2. *Pour tout sous-ensemble U non-vide et finit de E on a $U \subset Vect(U)$.*

Démonstration. 1. Soit $U = \{u_1, u_2, \cdots, u_p\} \subset E$.
– On a d'abord $0 = \sum_{j=1}^{p} 0 \cdot u_j \in Vect(U)$.

– D'autre part, soient $\lambda, \mu \in \mathbb{K}$ et $x, y \in Vect(U)$. Alors x et y s'écrivent sous la forme

$$x = \sum_{j=1}^{p} \alpha_j u_j \quad \text{et} \quad y = \sum_{j=1}^{p} \beta_j u_j$$

avec $\alpha_1, \cdots, \alpha_p, \ \beta_1, \cdots, \beta_p \in \mathbb{K}$. Nous avons

$$\lambda x + \mu y = \lambda \left(\sum_{j=1}^{p} \alpha_j u_j \right) + \mu \left(\sum_{j=1}^{p} \beta_j u_j \right).$$

On obtient facilement en regroupant :

$$\lambda x + \mu y = \sum_{j=1}^{p} (\lambda \alpha_j + \mu \beta_j) u_j$$

ce qui finit la preuve du fait que $Vect(U)$ est un sous-espace vectoriel de E.

2. Il faut montrer que pour tout $k \in \{1, 2, \cdots, p\}$ fixé arbitraire on a $u_k \in Vect(U)$. En effet

$$u_k = \sum_{j=1}^{p} \delta_{jk} u_j$$

ce qui montre le résultat.

$\square$

On a la définition suivante :

Définition 1.5. *Soient $u_1, u_2, \cdots, u_n$ des éléments de E.*
On dit que $u_1, u_2, \cdots, u_n$ forment une **famille génératrice** *de E*
(on peut aussi dire : l'ensemble $\{u_1, u_2, \cdots, u_n\}$ est une famille génératrice de E) si on a

$$Vect(u_1, \cdots, u_n) = E.$$

Remarques :

1. Comme on a toujours $Vect(u_1, \cdots, u_n) \subset E$ alors le fait que $u_1, u_2, \cdots, u_n$ forment une famille génératrice de E est équivalente à l'inclusion inverse : $E \subset Vect(u_1, \cdots, u_n)$, ce qui se traduit par le fait que tout élément de E s'écrit comme une combinaison linéaire des éléments $u_1, u_2, \cdots, u_n$
2. $u_1, u_2, \cdots, u_n$ est toujours une famille génératrice de $Vect(u_1, u_2 \cdots, u_n)$.

Exemple 1.6. *Les vecteurs $(1, 0)$ et $(0, 1)$ de $\mathbb{R}^2$ forment une famille génératrice de $\mathbb{R}^2$ car pour tout $x = (x_1, x_2) \in \mathbb{R}^2$ on a*

$$x = x_1(1, 0) + x_2(0, 1)$$

(tenir compte du fait que $x_1, x_2 \in \mathbb{R}$).

On a la définition suivante :

Définition 1.6. *Soient* $u_1, u_2, \cdots, u_n$ *des éléments de* E.

1. *On dit que* $u_1, u_2, \cdots, u_n$ *sont* **linéairement indépendants** *en* E *(on peut dire plus simplement :* **indépendants** *en* E*) ou encore qu'ils forment une* **famille libre** *en* E *si on a*

$$\left.\begin{array}{rcl}\displaystyle\sum_{j=1}^{n}\alpha_j u_j & = & 0 \\ \alpha_1, \alpha_2, \cdots, \alpha_n & \in & \mathbb{K}\end{array}\right\} \implies \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$$

2. *On dit que* $u_1, u_2, \cdots, u_n$ *sont* **liés** *si ils ne sont pas linéairement indépendants (c'est à dire s'il existe des scalaires* $\alpha_1, \cdots, \alpha_n \in \mathbb{K}$ *pas tous nuls tels que* $\sum_{j=1}^{n}\alpha_j u_j = 0$*).*

Exemple 1.7. *Les vecteurs* $(1,0)$ *et* $(0,1)$ *sont indépendants en* $\mathbb{R}^2$ *car si* $\alpha_1, \alpha_2 \in \mathbb{R}$ *sont tels que*

$$\alpha_1(1,0) + \alpha_2(0,1) = (0,0)$$

alors on a $(\alpha_1, \alpha_2) = (0,0)$ *donc* $\alpha_1 = \alpha_2 = 0$.

Exemple 1.8. *Les vecteurs* $(1,2)$ *et* $(2,4)$ *ne sont pas indépendants en* $\mathbb{R}^2$ *(donc ils sont liés) car*

$$(-2)\cdot(1,2) + 1\cdot(2,4) = (0,0).$$

Rappelons ici la notion de colinéarité de deux vecteurs :

Définition 1.7. *Deux vecteurs* $u, v \in E$ *sont dits* **colinéaires** *si*
il existe $\alpha \in \mathbb{K}$ *tel que* $v = \alpha u$
ou
il existe $\beta \in \mathbb{K}$ *tel que* $u = \beta v$.

Exemple : Tout vecteur u est colinéaire avec le vecteur 0_E, ou avec u (lui même), ou avec $2u$, $3u$ etc...
Nous avons le résultat suivant :

Proposition 1.3.

1. *Si* $u \neq 0$ *alors* u *"à lui tout seul" forme une famille libre en* E.
2. *Si* $u, v \in E$ *alors* u, v *sont indépendants si et seulement si ils ne sont pas colinéaires.*
3. *Si parmi les vecteurs* $u_1, \cdots, u_n$ *il y a le vecteur* O_E *alors ils sont liés.*
4. *Si parmi les vecteurs* $u_1, \cdots, u_n$ *il y a deux vecteurs qui sont colinéaires alors* $u_1, \cdots, u_n$ *sont liés.*

Démonstration.

1. Soit $\lambda \in \mathbb{K}$ tel que

$$\lambda u = 0 \tag{1.1}$$

et il faut montrer que $\lambda = 0$. Supposons par absurd que $\lambda \neq 0$. En multipliant (1.1) à gauche par $\frac{1}{\lambda}$ on trouve $u = 0$, ce qui est en contradiction avec l'hypothèse.

2. – Si u, v sont indépendants alors ils ne sont pas colinéaires (car s'ils l'étaient ils seraient forcément liés, donc contradiction).
   – Inversement supposons que u et v ne sont pas colinéaires et montrons qu'ils sont indépendants. Si par absurde ils étaient liés alors on aurait $\alpha u + \beta v = 0$ avec $\alpha, \beta \in \mathbb{K},\ \alpha \neq 0$ ou $\beta \neq 0$. On peut supposer $\alpha \neq 0$ (le cas $\beta \neq 0$ se traite de manière analogue). Alors $u = -\frac{\beta}{\alpha} v$ donc u et v sont colinéaires, ce qui est une contradiction.
3. Si $k \in \{1, \cdots, u_n\}$ est tel que $u_k = 0_E$ alors on a

$$\sum_{j=1} \delta_{jk} u_j = 0.$$

   Ceci montre que les vecteurs $u_1, \cdots, u_n$ sont liés.
4. L'hypothèse nous dit qu'il existe $k, l \in \{1, \cdots, u_n\}$ et $\beta \in \mathbb{K}$ tels que $u_l = \beta u_k$. Alors on a

$$\sum_{j=1} \alpha_j u_j = 0$$

   avec $\alpha_l = 1,\ \alpha_k = -\beta$ et (éventuellement) $\alpha_j = 0$ si $j \neq l$ et $j \neq k$. Ceci montre le résultat.

□

Définition 1.8. *Soit $\mathcal{B} = \{u_1, u_2, \cdots, u_n\}$ avec $n \in \mathbb{N}^*$, un sous-ensemble finit et non-vide de E (donc $u_j \in E, \quad \forall\ j \in [[1, n]]$). On dit que $\mathcal{B}$ est une* **base** *dans E (on peut dire aussi : les éléments $u_1, u_2, \cdots, u_n$ forment une base dans E) si les deux conditions suivantes sont satisfaites :*
$u_1, u_2, \cdots, u_n$ forment une famille génératrice en E
et
$u_1, u_2, \cdots, u_n$ sont indépendants en E.

Exemple : Des Exemples 1.6 et 1.7 on déduit que $\{(0,1),\ (1,0)\}$ est une base en $\mathbb{R}^2$.
Remarque : Une base dans un espace vectoriel n'est pas unique ; par exemple on peut vérifier que $\{(0,1),\ (1,1)\}$ est aussi une base en $\mathbb{R}^2$ (*Exercice*).
On a la caractérisation suivante pour une base :

Proposition 1.4. *Soit $\mathcal{B} = \{u_1, u_2, \cdots u_n\} \subset E$. Alors $\mathcal{B}$ est une base en E si et seulement si $\forall\, u \in E$, $\exists$ uniques $\alpha_1, \alpha_2, \cdots \alpha_n \in \mathbb{K}$ tels que*

$$u = \sum_{j=1}^{n} \alpha_j u_j.$$

(Les scalaires $\alpha_1, \alpha_2, \cdots \alpha_n$ s'appellent les **coordonnées** *de u dans la base $\mathcal{B}$).*

Démonstration. Montrons $\Longrightarrow$: Supposons que $\mathcal{B}$ est une base en E et soit $u \in E$ arbitraire. Comme $\mathcal{B}$ est une famille génératrice de E alors il existe $\alpha_1, \alpha_2, \cdots \alpha_n \in \mathbb{K}$ tels que

$$u = \sum_{j=1}^{n} \alpha_j u_j. \tag{1.2}$$

Il reste à montrer que $\alpha_1, \alpha_2, \cdots \alpha_n$ sont uniques. Soient $\beta_1, \beta_2, \cdots \beta_n \in \mathbb{K}$ tels que

$$u = \sum_{j=1}^{n} \beta_j u_j. \tag{1.3}$$

En faisant la différence entre (1.2) et (1.3) on trouve

$$\sum_{j=1}^{n} \alpha_j u_j - \sum_{j=1}^{n} \beta_j u_j = 0$$

c'est à dire

$$\sum_{j=1}^{n} (\alpha_j - \beta_j) u_j = 0.$$

Comme la famille $u_1, u_2, \cdots u_n$ est libre on obtient $\alpha_k = \beta_k, \quad \forall\ k = 1, 2, \cdots n$. Ceci finit la peuve de l'unicité.

Montrons $\Longleftarrow$: Supposons qu'on a : $\forall\ u \in E,\ \exists$ uniques $\alpha_1, \alpha_2, \cdots \alpha_n \in \mathbb{K}$ tels que

$$u = \sum_{j=1}^{n} \alpha_j u_j.$$

Alors il est clair que $\mathcal{B}$ est une famille génératrice en E. D'autre part, on observe que

$$0 = 0u_1 + 0u_2 \cdots 0u_n.$$

Par l'hypoythèse d'unicité si $\beta_1, \cdots \beta_n \in \mathbb{K}$ sont tels que

$$\sum_{j=1}^{n} \beta_j u_j = 0$$

alors $\beta_1 = \beta_2 = \cdots \beta_n = 0$. Ceci n'est autre que l'indépendance des éléments de $\mathcal{B}$. Nous avons alors que $\mathcal{B}$ est une base en E. □

Exemple 1.9. *(fondamental)*
Pour tout $n \in \mathbb{N}^*$ et tout $k \in \{1, 2, \cdots n\}$ nous notons par $e_{n,k}$ le vecteur en $\mathbb{R}^n$ défini par

$$(e_{n,k})_j = \delta_{jk} \qquad \textit{(les symboles de Kronecker).}$$

Alors l'ensemble $\{e_{n,1},\ e_{n,2}, \cdots e_{n,n}\}$ est une base dans le $\mathbb{K}$ - espace vectoriel $\mathbb{K}^n$ (cette base s'appelle **la base canonique** *en $\mathbb{K}^n$).*

Démonstration. Soit $x = (x_1, x_2, \cdots x_n) \in \mathbb{K}^n$ arbitraire. Alors il est clair que

$$x = \sum_{j=1}^{n} x_j e_{n,j} \tag{1.4}$$

où on remarque que $x_j \in \mathbb{K}$.
D'autre part, en faisant le produit scalaire en K^n de l'égalité (1.4) avec $e_{n,k}$ où $k \in \{1, \cdots n\}$ est fixé, on trouve

$$< x \, , \, e_{n,k} > = \sum_{j=1}^{n} x_j < e_{n,j} \, , \, e_{n,k} > = \sum_{j=1}^{n} x_j \delta_{kj}$$

ce qui donne pour tout $k \in \{1, \cdots n\}$:

$$x_k \; = < x \, , \, e_{n,k} > .$$

Ceci donne l'unicité de la décomposition (1.4) et finit la preuve. □

Remarque 1.1.
1. *De l'égalité* (1.4) *on déduit que les coordonnées de la décomposition du vecteur x dans la base canonique sont exactement les composantes du vecteur x.*
2. *S'il n'y a pas de possibilité de confusion on peut utiliser la notation $e_1, e_2, \cdots, e_n$ pour noter les éléments de la base canonique en $\mathbb{K}^n$.*

Nous avons le résultat suivant, (admis sans preuve) :

Lemme 1.1. *(Lemme fondamental de la dimension finie)*
Soit $m, n \in \mathbb{N}^*$, $\{u_1, \cdots u_n\} \subset E$ une famille génératrice de E et $\{v_1, \cdots v_m\} \subset E$ une famille libre de E.
Alors on a $m \leq n$.

Nous en déduisons le résultat suivant :

Théorème 1.1. *Soit $\mathcal{B} = \{u_1, u_2, \cdots u_n\}$ une base dans E. Alors*
1. *Toute famille génératrice en E a au moins n éléments.*
2. *Toute famille libre en E a au plus n éléments.*
3. *Toute base en E a exactement n éléments. On appelle alors* **dimension** *de l'espace vectoriel E le nombre n d'élements de toute base de E (nombre qui ne dépend pas de la base) ; on notera en général ce nombre n par $dim(E)$.*
4. *Toute famille génératrice de E avec exactement n éléments est une base en E.*
5. *Toute famille libre en E avec exactement n éléments est une base en E.*
6. *Si F est un sous-espace vectoriel de E avec $F \neq \{0\}$ et $F \neq E$ alors F a une base de m éléments avec $1 \leq m < n$; donc $1 \leq dim(F) < dim(E)$.*

Démonstration. Les points 1), 2), 3) sont évidents.
Les points 4), 5) et 6) sont admis sans preuve. □

Remarques :

1. Le *s.e.v.* $\{0_E\}$ de E n'a pas de base
(le singleton $\{0_E\}$ ne peut pas être une base de $\{0_E\}$ car il est lié).
Mais on dira par convention que $dim(\{0_E\}) = 0$.
2. Il y a des espaces vectoriels ne se réduisant pas à un seul élément (l'élément nul) qui n'ont pas de base ; on dit de ces espaces qu'ils sont de **dimension infinie**.
Par exemple $\mathcal{F}(\mathbb{R}, \mathbb{R})$ sur $\mathbb{R}$ est un espace vectoriel de dimension infinie (*résultat admis*).

1.2 Quelques rappels sur les determinants

On rappelle les notions suivantes :

Définition 1.9. *1. On appelle* **matrice identité** *de dimension n la matrice carrée dans $\mathcal{M}_n(\mathbb{K})$ notée I_n définie par*

$$(I_n)_{ij} = \delta_{ij}, \quad \forall\, i, j \in [[1,\, n]].$$

2. Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$ une matrice, avec $m, n \in \mathbb{N}^*$. On appelle **transposée** *de A notée A^T la matrice $\in \mathcal{M}_{n,m}(\mathbb{K})$ définie par*

$$(A^T)_{ji} = A_{ij}, \quad \forall\, i \in [[1,\, m]],\ j \in [[1,\, n]].$$

3. Soit $A \in \mathcal{M}_n(\mathbb{K})$. On dit que A est **inversible** *s'il existe $B \in \mathcal{M}_n(\mathbb{K})$ tel que $AB = BA = I_n$.*

On rappelle que l'inverse d'une matrice inversible A est unique et elle est notée A^{-1}. Cette matrice A^{-1} est evidemment inversible et son inverse n'est autre que A.
On rappelle aussi que la matrice identite I_n est inversible et que son inverse est égale à I_n.

Définition 1.10. *Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée. On appelle* **determinant** *de A le nombre noté $det(A) \in \mathbb{K}$ défini par*

$$det(A) = \sum_{\sigma \in S(n)} (-1)^{\epsilon(\sigma)}\, \Pi_{j=1}^{n} A_{\sigma(j),j}$$

où $S(n)$ est l'ensemble des permutations sur $\{1, 2, \cdots n\}$ et $\epsilon(\sigma)$ est le nombre des couples (i, j) avec $i < j$ tels que $\sigma(i) < \sigma(j)$.

Définition 1.11. *Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée.*

1. *Pour tous* $i, j \in \{1, 2, \cdots, n\}$ *on appelle* **mineur** (i, j) *de* A *le nombre* $\in \mathbb{K}$ *noté* $D_{ij}(A)$ *défini comme le déterminant de la matrice* $\in \mathcal{M}_{n-1}(\mathbb{K})$ *obtenue en supprimant la ligne* i *et la colonne* j *de la matrice* A.
2. *Pour tous* $i, j \in \{1, 2, \cdots, n\}$ *on appelle* **cofacteur** (i, j) *de* A *le nombre* $\in \mathbb{K}$ *noté* $Cof_{ij}(A)$ *défini par* $Cof_{ij}(A) = (-1)^{i+j} D_{ij}(A)$.
3. *On appelle* **matrice des cofacteurs** *de* A *notée* $Cof(A)$, *la matrice* $\in \mathcal{M}_n(\mathbb{K})$ *définie par*
$$(Cof(A))_{ij} = Cof_{ij}(A), \quad \forall\, i, j \in [[1,\, n]].$$

La proposition suivante résume les propriétés essentielles du determinant, qui permettent de faire le calcul effectif du determinant d'une matrice carrée donnée (*résultat admis*).

Proposition 1.5. *Soit* $A \in \mathcal{M}_n(\mathbb{K})$.

1. *Si* $n = 1$ *alors* $det(A) = A$.
2. *Si* $n = 2$ *et*
$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$
alors $det(A) = A_{11}A_{22} - A_{21}A_{12}$.
3. *Si* $n \geq 2$ *alors on a*
$$det(A) = \sum_{j=1}^{n} A_{ij} Cof_{ij}(A), \quad \forall\, i \in [[1,\, n]]$$
(développement selon une ligne arbitraire i *de* A*)*
ou encore
$$det(A) = \sum_{i=1}^{n} A_{ij} Cof_{ij}(A), \quad \forall\, j \in [[1,\, n]]$$
(développement selon une colonne arbitraire j *de* A*).*
4. *Si l'une des ligne de* A *est nulle alors* $det(A) = 0$. *De même si l'une des colonnes de* A *est nulle alors* $det(A) = 0$.
5. *Si deux lignes de* A *sont proportionnelles alors* $det(A) = 0$. *De même si deux colonnes de* A *sont proportionnelles alors* $det(A) = 0$.
6. *Si à une ligne de* A *on ajoute une combinaison linéaire des autres lignes alors le derminant de* A *ne change pas. De même, si à une colonne de* A *on ajoute une combinaison linéaire des autres colonnes alors le derminant de* A *ne change pas.*
7. *Si une matrice* B *s'obtient de* A *en échangeant deux lignes (ou deux colonnes) entre elles alors* $det(B) = -det(A)$.

8. $$det(A) = det(A^T).$$

9. *Si l'une des lignes (ou colonnes) de A est multiplié par un scalaire $\lambda \in \mathbb{K}$ alors le determinant de A se multiplie par λ.*
   Conséquence :
   $$det(\lambda A) = \lambda^n det(A).$$

10. *Si $A, B \in \mathcal{M}_n(\mathbb{K})$ alors $det(AB) = det(A)det(B)$.*
    Attention : en général $det(A + B)$ n'est pas égal à $det(A) + det(B)$.

Définition 1.12. *Soit $A \in \mathcal{M}_n(\mathbb{K})$. On dit que A est*

1. **triangulaire inférieure** *si $A_{ij} = 0$ pour $i < j$.*
2. **triangulaire supérieure** *si $A_{ij} = 0$ pour $i > j$.*
3. **diagonale** *si elle est triangulaire inférieure et triangulaire supérieure, c'est à dire si $A_{ij} = 0$ pour $i \neq j$.*

Définition 1.13. *On considère deux matrices $A, B \in \mathcal{M}_n(\mathbb{K})$. On dit que A est* **semblable** *à B s'il existe $P \in \mathcal{M}_n(\mathbb{K})$ matrice inversible telle que $A = PBP^{-1}$.*
Remarque : *En prenant $P = I_n$ on observe que A est toujours semblable à A.*
D'autre part, on voit facilement que si A est **semblable** *à B alors B est* **semblable** *à A (car $A = PBP^{-1} \iff B = P^{-1}AP$ donc la propriété est vraie avec P^{-1} à la place de P). On dira alors que les matrices A et B sont semblables.*

Proposition 1.6.

1. *Si $A \in \mathcal{M}_n(\mathbb{K})$ est une matrice triangulaire inférieure, ou triangulaire supérieure (donc aussi si elle est une matrice diagonale) alors on a*
   $$det(A) = A_{11}A_{22}\cdots A_{nn} = \Pi_{j=1}^{n} A_{jj}.$$
2. $$det(I_n) = 1$$
3. *Si $A \in \mathcal{M}_n(\mathbb{K})$ est inversible alors $det(A) \neq 0$ et on a*
   $$det(A^{-1}) = \frac{1}{det(A)}.$$
4. *Si $A, B \in \mathcal{M}_n(\mathbb{K})$ sont semblables alors $det(A) = det(B)$.*

Démonstration.

1. On développe successivement sur la ligne (ou colonne) sur laquelle il y a uniquement des zeros en dehors de la diagonale et on obtient le résultat.
2. C'est une conséquence immédiate de 1).
3. De $AA^{-1} = I_n$ on déduit $det(A)\,det(A^{-1}) = det(I_n) = 1$ ce qui donne le résultat.

4. Il existe $P \in \mathcal{M}_n(\mathbb{K})$ telle que $A = PBP^{-1}$ ce qui donne

$$det(A) = det(P)\, det(B)\, det(P^{-1}).$$

En utilisant 3) avec P à la place de A on obtient le résultat.

$\square$

On finit cette section par rappeller le résultat suivant :

Proposition 1.7. *Si $A \in \mathcal{M}_n(\mathbb{K})$ est telle que $det(A) \neq 0$ alors A est inversible et on a*

$$A^{-1} = \frac{1}{det(A)}[Cof(A)]^T.$$

Remarque : Avec la partie 3) de la Proposition 1.6 on déduit

$$A \text{ inversible} \quad \Longleftrightarrow \quad det(A) \neq 0.$$

1.3 Systèmes d'équations linéaires algébriques

Dans cette section n et p sont deux nombres dans N^*.

1.3.1 Introduction

On se donne une matrice $A \in \mathcal{M}_{pn}(\mathbb{K})$ et un vecteur colonne $b \in \mathbb{K}^p$ avec

$$b = (b_1, b_2, \cdots b_p)^T, \quad A = (A_{ij})_{i=1,\cdots p,\, j=1,\cdots n}.$$

On cherche $x_1, x_2, \cdots x_n \in \mathbb{K}$ tels que

$$A_{i1}x_1 + A_{i2}x_2 + \cdots A_{in}x_n = b_i, \quad \forall\, i \in [[1,\, p]]. \tag{1.5}$$

Le système (1.5) est un système algébrique linéaire avec p équations et n inconnues.

Remarque 1.2. *1. On peut mettre ce système sous forme matricielle : en notant $x = (x_1, x_2, \cdots x_n)^T$ (matrice colonne) le système (1.5) s'écrit : trouver $x \in \mathbb{K}^n$ tel que*

$$Ax = b \tag{1.6}$$

2. Dans le cas où $b = 0$ on dit que le système (1.5) ou (1.6) est un système **homogène**. *Pour un système homogène on voit facilement que l'ensemble des solutions $x \in \mathbb{K}^n$ n'est autre que $Ker(A)$ qui est un sous-espace vectoriel de $\mathbb{K}^n$. Il contient alors au moins la solution $x = 0$.*

Le but de cette section est d'étudier existence et l'unicité d'une solution du système et de donner la/les solution(s) quand ceci est possible. On commence par le cas le plus simple.

1.3.2 Les systèmes de Cramer

Proposition 1.8. *Supposons que $p = n$ et que A est inversible. Alors le système* (1.6) *admet une solution unique donnée par*

$$x = A^{-1}b. \tag{1.7}$$

(Le système (1.6) *avec $p = n$ et A inversible s'appelle* **système de Cramer***).*

Démonstration. La preuve est très simple : le système (1.6) est équivalent à $A^{-1}Ax = A^{-1}b$ et comme $A^{-1}A = I_n$ et $I_n x = x$ alors (1.6) est équivalent à (1.7), ce qui donne le résultat. □

Voici une règle plus simple pour résoudre (1.6) :

Proposition 1.9. *(règle de Cramer)*
Si $p = n$ et A est inversible alors la solution de (1.6) *est donnée par*

$$x_k = \frac{det(A_k)}{det(A)}, \quad \forall\, k \in [[1,\, n]] \tag{1.8}$$

où on note par A_k la matrice obtenue de A en remplaçant la k - ème colonne de A par la colonne b, c'est à dire, pour tous $i, j \in [[1,\, n]]$ on a

$$(A_k)_{ij} = \begin{cases} A_{ij} & si \quad j \neq k \\ b_i & si \quad j = k. \end{cases}$$

Démonstration. Les Propositions 1.8 et 1.7 nous donnent pour tout $k \in [[1,\, n]]$:

$$x_k = \sum_{j=1}^{n} (A^{-1})_{kj} b_j = \frac{1}{det(A)} \sum_{j=1}^{n} Cof_{jk}(A) b_j.$$

On observe que $b_j = (A_k)_{jk}$ et $Cof_{jk}(A) = Cof_{jk}(A_k)$, ce qui donne

$$\sum_{j=1}^{n} Cof_{jk}(A)(A_k)_{jk} = det(A_k)$$

(c'est le développement de $det(A_k)$ selon la colonne k). Ceci donne le résultat. □

1.3.3 Le cas général

On donnera ici une méthode à utiliser dans le cas où le système (1.6) n'est pas un système de Cramer (donc $p = n$ et A non inversible, ou $p \neq n$).

Définition 1.14. *Soit B une matrice dans $\mathcal{M}_{p,n}(\mathbb{K})$. Soit $r \in [[1, \min(p, n)]]$. On appelle* **sous-matrice carrée de taille** *r de B une matrice carrée $\in \mathcal{M}_r(\mathbb{K})$ obtenue de B en y mettant "l'intersection" entre r lignes et r colonnes de B.*

Exemple : Soit
$$B = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 0 & 4 \end{pmatrix} \in \mathcal{M}_{2,3}(\mathbb{R}).$$
La matrice
$$\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$$
est une sous-matrice carrée de taille 2 de B (car elle s'obtient comme "intersection" entre les 2 lignes de B et les colonnes 1 et 3 de B).
La matrice $(2) \in \mathcal{M}_1(\mathbb{R})$ est une sous-matrice carrée de taille 1 de B (car elle s'obtient comme "intersection" entre la ligne 2 et la colonne 1 de B).

Définition 1.15. *Soit B une matrice dans $\mathcal{M}_{p,n}(\mathbb{K})$. On appelle* **rang** *de B, noté $rang(B)$, le plus grand nombre des colonnes de B indépendantes en $\mathbb{K}^n$) (chaque colonne étant vue comme un vecteur de $\mathbb{K}^n$). On peut aussi dire : le rang de B est le plus grand nombre $r \in \mathbb{N}$ tel qu'il existe r colonnes indépendentes de B.*

Remarque 1.3. *Comme le sous-espace vectoriel en $\mathbb{K}^n$ engendré par les colonnes de B est $Im(B)$ alors on a*
$$rang(B) = dim(Im(B)). \tag{1.9}$$

On a la proposition suivante (*résultat admis*) :

Proposition 1.10. *Pour toute matrice $B \in \mathcal{M}_{p,n}(\mathbb{K})$ on a*

1. *la rang de B est égal au plus grand nombre des lignes indépendantes de B (chaque ligne étant vue comme un vecteur de $\mathbb{K}^p$).*
2. *$rang(B) = 0$ si et seulement si $B = 0$.*
3. *Si $B \neq 0$ alors le rang de B est le plus grand nombre $r \in [[1, \min(p, n)]]$ pour lequel il existe une sous-matrice carrée de taille r qui est inversible (donc de determinant non nul).*

Remarque : Le rang de toute matrice $B \in \mathcal{M}_{p,n}(\mathbb{K})$ est compris entre 0 et $\min(p, n)$.
Exemples :

1. Dans l'exemple précédent
$$B = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 0 & 4 \end{pmatrix} \in \mathcal{M}_{2,3}(\mathbb{R}).$$
le rang de B peut être égal à 1 ou à 2 (il ne peut pas être 0 car la matrice est non nulle). On essaie d'abord de voir si le rang de B est égal à la valeur maximale possible, c'est à dire, égal à 2. Pour ceci il faut trouver une sous-matrice carrée de taille 2 de B qui soit inversible. On peut considérer la sous-matrice carrée de taille 2 de B suivante :
$$M = \begin{pmatrix} 1 & -2 \\ 2 & 0 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$$
(obtenue à l'intersection entre les 2 lignes et les 2 premières colonnes de B). Nous avons $det(M) = 0 - 2 \cdot (-2) = 4 \neq 0$ ce qui nous dit que $rang(B) = 2$.

2. Si $B \in \mathcal{M}_n(\mathbb{K})$ est une matrice inversible alors $rang(B) = n$.
3. Si $B \in \mathcal{M}_{p,n}(\mathbb{K})$ avec $p, n \geq 2$ est une matrice non nulle et telle que ses lignes (respectivement ses colonnes) sont proportionnelles, alors $rang(B) = 1$ (car pour toute sous-matrice M carrée de taille ≥ 2 de B on a $det(M) = 0$).
   Par exemple pour la matrice

$$B = \begin{pmatrix} -1 & 2 & 0 & 3 \\ -2 & 4 & 0 & 6 \\ -3 & 6 & 0 & 9 \end{pmatrix}$$

   on a $rang(B) = 1$.

Procédure générale pour résoudre le système (1.6)

On suppose que A n'est pas la matrice nulle (dans la cas $A = 0$ alors : soit tout vecteur $x \in \mathbb{R}^n$ est solution de (1.6) pour $b = 0$, soit il n'y a aucune solution de (1.6) pour $b \neq 0$). On calcule $r = rang(A) \in [[1, \min(p, n)]]$. On choisit une sous-matrice carrée de taille r de A qui est inversible ; notons par M_r cette matrice qu'on appellera **matrice principale** du système. Nous avons 3 cas possibles que nous étudions dans la suite.
Cas 1. $r = p = n$.
C'est le cas le plus simple, car la matrice A est carrée (on a autant d'équations que d'inconnues) et inversible. Alors (1.6) est un système de Cramer pour lequel on a l'existence et l'unicité d'une solution ; on peut calculer cette solution en utilisant l'une des Propositions 1.8 ou 1.9 ou toute autre méthode connue.
Cas 2. $r = p < n$.
Dans ce cas la matrice principale M_r s'obtient à l'intersection entre les r lignes de A et r colonnes de A d'indices $j_1, j_2, \cdots j_r$ avec $1 \leq j_1 < j_2 < \cdots < j_r \leq n$. On va noter $P = \{j_1, j_2, \cdots j_r\}$ et $S = [[1, n]] - P$. Les r inconnues $\{x_j,\ j \in P\}$ vont s'appeller des **inconnues principales** ; les autres $n - r$ inconnues $\{x_j,\ j \in S\}$ vont s'appeller des **inconnues secondaires**.
Il est facile de voir que notre système (1.5) est équivalent au système

$$\sum_{j \in P} A_{ij} x_j = b_i - \sum_{j \in S} A_{ij} x_j, \quad i \in [[1, p]]. \tag{1.10}$$

Ce système devrait permettre de trouver les inconnues principales en fonction des inconnues secondaires vues comme des paramètres. On observe que la matrice des coefficients des inconnues pricipales dans la partie de gauche de (1.10) n'est autre que la matrice principale M_r qui est inversible. On peut alors calculer les inconnues principale en fonction des inconnues secondaires. On aura alors une infinité des solutions du système (1.6) dans ce cas.
On peut être plus précis concernant l'ensemble des solutions. Remarquons qu'on peut exprimer toutes les solutions du système (1.6) en fonctions des $n-p$ "paramètres" $\{x_j,\ j \in S\}$. On admet le résultat suivant :

Proposition 1.11. *1. Si $b = 0$ alors le sous-espace vectoriel en $\mathbb{K}^n$ des solutions du système* (1.6) *(qui n'est autre que $Ker(A)$) est de dimension $n - r$.*

2. Si $b \neq 0$ alors l'ensemble des solutions de (1.6) *est un* **sous-espace affine** *de dimension $n - r$ de $\mathbb{K}^n$, c'est à dire, l'ensemble des solutions x de* (1.6) *est l'ensemble*

$$\{x^0 + y, \quad y \in Ker(A)\}$$

où $x^0 \in \mathbb{K}^n$ une solution fixée de (1.6).

Remarque : Un sous-espace vectoriel est un fait un cas particulier $x^0 = 0$ d'un sous-espace affine. En fait on aurait pu dire dans la proposition précédente que l'ensemble des solution est un sous-espace affine, qui est un sous-espace vectoriel dans le cas particulier $b = 0$.

Cas 3. $r < p$.

Dans ce cas la matrice principale M_r s'obtient comme "intersection" entre r lignes et r colonnes de A.

Pour les indices de colonnes nous reprennons les notations vues dans le **Cas 2.** Pour les r lignes de M_r nous notons leurs indices par $i_1, i_2, \cdots i_r$. Nous supposons aussi $1 \leq i_1 < i_2 < \cdots < i_r \leq p$ et nous notons $P_1 = \{i_1, i_2, \cdots i_r\}$ et $S_1 = [[1, n]] - P_1$.

Les r équations d'indices $i \in P_1$ vont s'appeller des **équations principales** et les $p - r$ équations d'indices $i \in S_1$ vont s'appeller des **équations secondaires**. On peut écrire les équations principales sous la forme :

$$\sum_{j \in P} A_{ij} x_j = b_i - \sum_{j \in S} A_{ij} x_j, \quad i \in P_1. \tag{1.11}$$

(on "oublie" les équations secondaires).

Remarque : Si $r = n$ (donc si le rang de la matrice A est égal au nombre d'inconnues) alors $P = \{1, 2, \cdots n\}$ et $S = \emptyset$, Dans ce cas il n'y a pas des inconnues secondaires et dans le système (1.11) le term avec $\sum$ dans le membre de droite n'existe pas.

Pour tout $i \in S_1$ nous notons par A_i la matrice obtenue à partir de la matrice M_r en y ajoutant la ligne i et la colonne b. On admet le résultat suivant :

Proposition 1.12. *1. S'il existe $i \in S_1$ tel que $det(A_i) \neq 0$, alors le système* (1.6) *n'a pas de solution.*

2. Si $det(A_i) = 0, \quad \forall\ i \in S_1$ alors le système (1.6) *a au moins une solution. On distingue ici deux cas :*

- *Si $r = n$ alors le système* (1.6) *a une seule solution obtenue en résolvant*

$$\sum_{j=1}^{n} A_{ij} x_j = b_i, \quad i \in P_1.$$

- *Si $r < n$ alors le système* (1.6) *a une infinité des solutions. Pour les trouver on résoude le système* (1.11) *comme dans le* **Cas 2** *en exprimant les inconnues principales $\{x_j,\ j \in P\}$ en fonction des inconnues secondaires $\{x_j,\ j \in S\}$. On a* *encore comme dans le* **Cas 2 :** *l'ensemble des solutions de* (1.6) *est un sous-espace affine de dimension* $n - r$ *de* $\mathbb{K}^n$ *avec la remarque que dans le cas* $b = 0$ *l'ensemble des solutions de* (1.6) *est un sous-espace vectoriel de dimension* $n - r$ *de* $\mathbb{K}^n$.

Remarque 1.4.

1. *Si* $b = 0$ *alors dans toutes les matrices* A_i *la dernière colonne est égale à 0, donc les determinants de ces matrices sont tous nuls. Alors il existe au moins une solution du système* (1.6) *(en fait on sait que si* $b = 0$ *alors* $x = 0$ *est une solution de* (1.6)*).*
2. *Pour toute matrice* $A \in \mathcal{M}_{p,n}(\mathbb{K})$ *on peut déduire des résultats obtenus dans les 3 cas traités précédemment que*
$$dim(Ker(A)) = n - r. \tag{1.12}$$
Cela s'écrit aussi, en utilisant (1.9)
$$dim(Ker(A)) + dim(Im(A)) = n \tag{1.13}$$
(c'est un résultat bien connu d'algèbre linéaire).

1.4 Valeurs propres, vecteurs propres, diagonalisation

Dans cette section A désigne une matrice carrée dans $\mathcal{M}_n(\mathbb{K})$.

Définition 1.16. *On dit que* $\lambda \in \mathbb{K}$ *est une* **valeur propre** *de* A *s'il existe* $x \in \mathbb{K}^n$ *avec* $x \neq 0$ *tel que*
$$Ax = \lambda x. \tag{1.14}$$
Tout vecteur $x \in \mathbb{K}^n$ *satisfaisant* (1.14) *s'appelle* **vecteur propre** *de* A *associé (ou correspondant) à la valeur propre* λ.

Comme $\lambda x = I_n x$ alors l'égalité (1.14) s'écrit
$$(A - \lambda I_n)x = 0 \tag{1.15}$$
qu'il faut voir comme un système d'équations linéaires algébriques **homogène** avec vecteur inconnue x, où on peut voir λ comme un paramètre.
Remarquons que $x = 0$ est toujours une solution du système (1.15). Par définition, λ est une valeur propre de A si et seulement si le système (1.15) a des solutions $x \in K^n$ non-nulles. Nous avons alors que $\lambda \in \mathbb{K}$ est une valeur propre de A si et seulement si
$$Ker(A - \lambda I_n) \neq \{0\}.$$
L'ensemble des vecteurs propres dans $\mathbb{K}^n$ correspondant à λ n'est autre que le sous-espace vectoriel $Ker(A - \lambda I_n)$.
Nous considérons la fonction $P_A : \mathbb{K} \to \mathbb{K}$ donnée par
$$P_A(\lambda) = det(A - \lambda I_n), \quad \forall\, \lambda \in \mathbb{K}.$$

Il est facile de voir que P_A est un polynome en $\mathbb{K}$ de degré n avec $(-1)^n$ comme coefficient de λ^n (cela se montre par reccurence, en développant le determinant de $A - \lambda I_n$ selon la dernière ligne ou colonne).
On dit alors que P_A est le **polynome caractéristique** de la matrice A.
On a le résultat important suivant :

Proposition 1.13. *Soit* $\lambda \in \mathbb{K}$.

1. *Si λ est une valeur propre de A alors λ est une racine de P_A.*
2. *Inversement, si λ est une racine de P_A alors λ est une valeur propre de A. En plus nous avons :*

$$rang(A - \lambda I_n) \in [[0,\, n-1]]$$

et

$$dim(Ker(A - \lambda I_n)) = n - rang(A - \lambda I_n).$$

Démonstration.

1. On suppose que λ est une valeur propre de A et on doit montrer que λ est une racine de P_A. Supposons par absurd que $det(A - \lambda I_n) \neq 0$; alors $A - \lambda I_n$ est une matrice inversible. En multipliant (1.15) à gauche par $(A - \lambda I_n)^{-1}$ on déduit que si x est une solution de (1.15) alors forcément $x = 0$, ce qui contredit l'hypothèse que λ est une valeur propre de A.
2. Supposons que λ est une racine de P_A, donc $det(A - \lambda I_n) = 0$. La Proposition 1.10 nous dit que $rang(A - \lambda I_n) \in [[0,\, n-1]]$. D'autre part en utilisant (1.12) ou (1.13) on a

$$dim(Ker(A - \lambda I_n)) = n - rang(A - \lambda I_n) \in [[1,\, n]]$$

Alors $Ker(A - \lambda I_n)$ contient au moins un élément non nul, ce qui finit la preuve.

□

Il est bien connu que le polynome P_A (comme d'ailleurs tout polynome de degré n à coefficients dans $\mathbb{K}$) a au moins une racine complexe. En plus il y a un nombre finit de racines complexes distinctes et ce nombre est compris entre 1 et n. Nous notons par $\mu_1, \mu_2, \cdots \mu_s \in \mathbb{C}$ les racines distinctes de P_A, avec $s \in [[1,\, n]]$. Nous savons aussi que chaque racine μ_j a une multiplicité $m_j \in [[1,\, n]]$, pout tout $j \in [[1,\, s]]$ avec

$$m_1 + m_2 + \cdots m_s = n.$$

Cela signifie qu'on a la décomposition en facteurs suivante pour P_A :

$$P_A(\lambda) = (-1)^{\lambda}(\lambda - \mu_1)^{m_1}(\lambda - \mu_2)^{m_2} \cdots (\lambda - \mu_s)^{m_s}, \quad \forall\, \lambda \in \mathbb{C}.$$

Pour tout $j = 1, \cdots s$ le nombre m_j s'appelle **multiplicité algébrique** de la valeur propre μ_j.
Remarque : *Si $s = n$ alors forcément $m_1 = m_2 = \cdots = m_n = 1$.*

Soit $\mu_j,\ j \in \{1, 2, \cdots s\}$ une valeur propre de A et supposons que $\mu_j \in \mathbb{K}$.
Nous notons par k_j la dimension de $Ker(A - \mu_j I_n) \subset \mathbb{K}^n$ qui est le sous-espace vectoriel des vecteurs propres de A. Rappellons que la partie 2) de la Proposition 1.13 nous donne

$$k_j = n - rang(A - \mu_j I_n) \in [[1,\ n]], \quad \forall\, j \in [[1,\ s]].$$

Nous admettons sans preuve le résultat suivant :

Proposition 1.14.

$$1 \leq k_j \leq m_j, \quad \forall\, j \in [[1,\ s]].$$

Pour tout $j \in [[1,\ s]]$ le nombre k_j s'appelle **multiplicité géométrique** de la valeur propre μ_j.
Remarque : Les multiplicités géométriques sont les mêmes pour $K = \mathbb{R}$ ou pour $\mathbb{K} = \mathbb{C}$.

On a la définition suivante :

Définition 1.17. *On dit que A est* **diagonalisable** *en K s'il existe un base en $\mathbb{K}^n$ des vecteurs propres de A.*

Remarque 1.5. *Comme $dim(\mathbb{K}^n) = n$ on déduit du Thérème 1.1 partie 5) que A est diagonalisable en K si et seulement si il existe une famille libre en K^n composée de n vecteurs propres de A.*

On utilisera le résultat suivant : (*sans preuve*).

Lemme 1.2. *Soient $\alpha_1, \alpha_2, \cdots \alpha_q \in \mathbb{K}$ des valeurs propres distinctes de A, avec $q \in \mathbb{N}^*$. Pour chaque $j \in \{1, \cdots q\}$ soit F_j une famille libre de vecteurs propres en $\mathbb{K}^n$ de A correspondant à α_j. Alors $F_1 \cup F_2 \cup \cdots \cup F_q$ est une famille libre en $\mathbb{K}^n$.*

Nous pouvons montrer maintenant le résultat principal de cette section :

Théorème 1.2. *La matrice A est diagonalisable en $\mathbb{K}$ si et seulement si les deux conditions suivantes sont satisfaites :*

1. *Toutes les valeurs propres $\mu_1, \mu_2, \cdots \mu_s$ de A sont dans $\mathbb{K}$*
   (c'est toujours le cas si $\mathbb{K} = \mathbb{C}$)
2. $$k_j = m_j, \quad \forall\, j \in [[1,\ s]].$$
   (c'est à dire, les multiplicités algébriques et géométriques coincident pour toute valeur propre de A).

Démonstration. $\Longleftarrow$ Si les deux conditions du Théorème 1.2 sont satisfaites alors pour tout $j \in [[1,\ s]]$ on peut construire une base F_j de l'espace vectoriel $Ker(A - \mu_j I_n)$ des vecteurs propres de A pour la valeur propre μ_j ; en plus la famille F_j aura k_j éléments, car $dim(Ker(A - \mu_j I_n)) = k_j$. D'autre part, le Lemme 1.2 nous dit que $F_1 \cup F_2 \cup \cdots \cup F_s$ sera une famille libre en K^n. Mais cette famille a exactement $k_1 + k_2 + \cdots k_s = m_1 + m_2 + \cdots m_s = n$ éléments, donc c'est une base en $\mathbb{K}^n$.
La partie $\Longrightarrow$ est admise. □

Dans la suite nous supposons que A est une matrice diagonalisable. Pour tout $j \in [[1, s]]$ nous considérons une base de $Ker(A - \mu_j I_n)$ que nous pouvons noter $\{P_{j,1},\ P_{j,2},\ \cdots\ P_{j,k_j}\}$. Alors une base en $\mathbb{K}^n$ des vecteurs propres de A sera

$$\{P_{1,1},\ P_{1,2},\ \cdots\ P_{1,k_1},\ P_{2,1},\ P_{2,2},\ \cdots\ P_{2,k_2},\ \ \cdots\ P_{s,1},\ P_{s,2}\ \cdots\ P_{s,k_s}\}.$$

Une telle notation à deux indices n'est pas convenables et on donne dans la suite une notation à un indice, plus facile à utiliser.

On introduit d'abord la notation suivante pour les valeurs propres de A : $\lambda_1,\ \lambda_2,\ \cdots\ \lambda_n$ où chaque valeur propre est répétée d'un nombre de fois égal à sa multiplicité ; cela veut dire : les valeurs propres $\lambda_1,\ \cdots\ \lambda_{k_1}$ sont toutes égales à μ_1, les $\lambda_{k_1+1},\ \cdots\ \lambda_{k_1+k_2}$ sont toutes μ_2, etc ..

Ensuite les éléments de la base en $\mathbb{K}^n$ des vecteurs propres de A sont re-notés $P_1, P_2, \cdots\ P_n$, où les vecteurs $P_1, P_2, \cdots\ P_{k_1}$ sont en fait $P_{1,1},\ P_{1,2},\ \cdots\ P_{1,k_1}$ respectivement, les vecteurs $P_{k_1+1}, P_{k_1+2}, \cdots\ P_{k_1+k_2}$ sont en fait $P_{2,1},\ P_{2,2},\ \cdots\ P_{2,k_2}$, etc ..

Il est clair alors que dans cette nouvelle notation, le vecteur P_j est un vecteur propre de A correspondat à la valeur propre λ_j, c'est à dire

$$AP_j = \lambda_j P_j, \quad \forall\ j \in [[1, n]]. \tag{1.16}$$

Donc pour résumer, pour une matrice $A \in \mathcal{M}_n(\mathbb{K})$ diagonalisable en $\mathbb{K}$ on note $\lambda_1,\ \lambda_2,\ \cdots\ \lambda_n \in \mathbb{K}$ les valeurs propres de A (répétés selon leur multiplicités) et on a une base en $\mathbb{K}^n$ de vecteurs propres de A qu'on peut noter

$$\mathcal{B}_{vp} = \{P_1, P_2\ \cdots\ P_n\}$$

ou chaque P_j est un vecteur propre de A correspondat à la valeur propre λ_j, c'est à dire on a (1.16).

Nous avons le résultat important suivant :

Théorème 1.3. *Si $A \in \mathcal{M}_n(\mathbb{K})$ est une matrice diagonalisable en $\mathbb{K}$ alors elle est semblable à une matrice diagonale.*

Plus précisément, notons $P \in \mathcal{M}_n(\mathbb{K})$ la matrice dont la j - ème colonne est le vecteur propre P_j, ceci pour tout $j \in [[1,\ n]]$ (on va noter $P = [P_1, P_2, \cdots\ P_n]$). Notons aussi par $D \in \mathcal{M}_n(\mathbb{K})$ la matrice diagonale telle que D_{jj} soit égale à la valeur propre λ_j pour tout $j \in [[1,\ n]]$ (on peut noter $D = diag(\lambda_1, \lambda_2, \cdots\ \lambda_n)$). Alors P est une matrice inversible et on a

$$A = PDP^{-1}. \tag{1.17}$$

Démonstration. Comme les n colonnes de P sont indépendants en $\mathbb{K}^n$ alors $rang(P) = n$, ce qui montre que P est inversible.

Il reste à montrer l'égalité (1.17). Cette égalité est équivalente (en multipliant à droite par la matrice P) à l'égalité

$$A\,P = PD.$$

Mais nous avons :

$$A\,P = A[P_1,\ P_2,\ \cdots\ P_n] = [AP_1,\ AP_2,\ \cdots\ AP_n] = [\lambda_1\,P_1, \lambda_2\,P_2,\ \cdots\ \lambda_n P_n]$$

et aussi

$$P\,D = P\,[\lambda_1 e_{n,1},\ \lambda_2 e_{n,2},\ \cdots\ \lambda_n e_{n,n}] = [\lambda_1 P\,e_{n,1},\ \lambda_2 P\,e_{n,2},\ \cdots\ \lambda_n P\,e_{n,n}] = [\lambda_1\,P_1, \lambda_2\,P_2,\ \cdots\ \lambda_n P_n]$$

ce qui montre l'égalité souhaitée. $\square$

Exemple 1.10. *Considérons $A \in \mathcal{M}_2(\mathbb{K})$ donnée par*

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 4 \end{pmatrix}.$$

Nous avons

$$P_A(\lambda) = det(A - \lambda I_2) = (1-\lambda)(4-\lambda), \quad \forall\,\lambda \in \mathbb{C}$$

et donc on a $s = 2,\ \mu_1 = 1,\ \mu_2 = 4,\ m_1 = m_2 = 1$. On a alors forcément $k_1 = m_1 = 1$ et $k_2 = m_2 = 1$ (car $1 \le k_j \le m_j = 1$ pour $j = 1, 2$). On déduit alors que la matrice A est diagonalisable en $\mathbb{R}$ et aussi en $\mathbb{C}$.
On pose $\lambda_1 = \mu_1 = 1, \quad \lambda_2 = \mu_2 = 4$. On cherche un vecteur propre $P_1 \in \mathbb{K}^2$ de A pour $\lambda_1 = 1$. Pour cela on considère

$$A_1 = A - \lambda_1 I_2 = A - I_2 = \begin{pmatrix} 0 & 2 \\ 0 & 3 \end{pmatrix}$$

et on voit que si on pose $P_1 = (1,0)^T$ alors $A_1 P_1 = 0$, donc P_1 est un vecteur propre non nul de A pour λ_1. De même, on considère

$$A_2 = A - \lambda_2 I_2 = A - 4I_2 = \begin{pmatrix} -3 & 2 \\ 0 & 0 \end{pmatrix}$$

et on voit que si on pose $P_2 = (2,3)^T$ alors $A_2 P_2 = 0$, donc P_2 est un vecteur propre non nul de A pour λ_2.
Alors $\{P_1,\ P_2\}$ est une base des vecteurs propres en $\mathbb{R}^2$ ou en $\mathbb{C}^2$ de la matrice A, ce qui nous dit que A est diagonalisable en $\mathbb{R}$ et en $\mathbb{C}$.
On déduit du Théorème 1.3 que A est semblable à une matrice diagonale ; plus précisément on a

$$A = PDP^{-1}$$

avec

$$P = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$$

matrice inversible et

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}.$$

Remarque 1.6. *Il y a des matrices qui ne sont pas diagonalisables. Par exemple considérons* $B \in \mathcal{M}_2(\mathbb{C})$ *donnée par*

$$B = \begin{pmatrix} 3 & 0 \\ 1 & 3 \end{pmatrix}.$$

Il y a une seule valeur propre $\mu = 3$ *et sa multiplicité algébrique* m *est égale à 2, car* $P_B(\lambda) = (\lambda - 3)^2, \quad \forall\ \lambda \in \mathbb{C}$. *Pour calculer la multiplicité géométrique* r *de la valeur propre* $\mu = 3$ *on considère*

$$B - \mu I_2 = B - 3I_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Comme $rang(B - 3I_2) = 1$ *alors* $k = dim(Ker(B - 3I_2)) = 2 - 1 = 1$. *Comme* $k < m$ *alors la matrice* B *n'est pas diagonalisable.*

Du Théorème 1.3 et de partie 4) Proposition 1.6 on déduit

Corollaire 1.1. *Si* $A \in \mathcal{M}_n(\mathbb{K})$ *est diagonalisable en* $\mathbb{K}$ *alors*

$$det(A) = \lambda_1 \lambda_2 \cdots \lambda_n$$

où $\lambda_1, \lambda_2, \cdots \lambda_n \in \mathbb{K}$ *sont les valeurs proprès (répétées selon leur multiplicitées) de* A.

Nous finissons ce chapitre en donnant une **application** importante de la diagonalisation des matrices.
Si $A \in \mathcal{M}_n(\mathbb{K})$ est diagonalisable alors nous pouvons écrire

$$A = PDP^{-1}$$

avec $P, D \in \mathcal{M}_n(\mathbb{K})$, où P est inversible et D diagonale. On a alors une manière très simple de calculer A^m avec $m \in \mathbb{N}^*$ arbitraire ; nous avons

$$A^m = PDP^{-1}\, PDP^{-1} \cdots PDP^{-1} \quad \text{(produit répété } m \text{ fois)}$$

et on montre facilement par reccurence sur m que

$$A^m = PD^m P^{-1}.$$

La matrice D^m est très facile à calculer : si D est la matrice diagonale telle que
$D = diag(\lambda_1, \lambda_2, \cdots \lambda_n)$ alors $D^m = diag(\lambda_1^m, \lambda_2^m, \cdots \lambda_n^m)$ (preuve facile par reccurence).

Chapitre 2

Equations différentielles

2.1 Introduction aux équations différentielles

2.1.1 Rappel dérivée à l'ordre quelconque :

Soit $I \subset \mathbb{R}$ un intervalle ouvert.

1. Soit $f : I \mapsto \mathbb{R}^n$ une fonction. Si f est dérivable sur I et f' est à son tour dérivable sur I, on dira que f est deux fois dérivable sur I et on notera $f'' = (f')'$; ensuite si f'' est dérivable on dira que f est 3 fois dérivable et on notera $f^{(3)} = (f'')'$. Par reccurence on a : pour tout $p \in \mathbb{N}^*$ si $f^{(p)}$ existe et est dérivable alors on dira que f est dérivable $p+1$ - fois et on va noter $f^{(p+1)} = (f^{(p)})'$. Par convention on pose $f^{(0)} = f$.
2. Considérons $f : I \mapsto \mathbb{R}^n$ et $p \in \mathbb{N}^*$. On dit que f est de classe C^p sur I et on notera $f \in C^p(I)$ si $f, f', f'', \cdots f^{(p)}$ existent et sont continues sur I. Par convention si f est continue on note $f \in C(I)$ ou $f \in C^0(I)$. On dit que f est de classe C^∞ sur I et on notera $f \in C^\infty(I)$ si $f^{(p)}$ existe et est continue pour tout $p \in \mathbb{N}$.
3. Considérons $p \in \mathbb{N}$ et $f : I \mapsto \mathbb{C}$ avec $f = f_1 + i f_2$, où $f_1, f_2 : I \to \mathbb{R}$. Alors f est dérivable p - fois sur I si f_1, f_2 sont dérivables p - fois sur I et on pose $f^{(p)} = f_1^{(p)} + i f_2^{(p)}$.

2.1.2 Systèmes différentielles d'ordre 1

Soit $I \subset \mathbb{R}$ un intervalle ouvert et $n \in \mathbb{N}^*$. On se donne une fonction continue $f : I \times \mathbb{R}^n \mapsto \mathbb{R}^n$ avec $f = (f_1, f_2, \cdots f_n)^T$ où $f_k = I \times \mathbb{R}^n \mapsto \mathbb{R}$ pour tout $k = 1, 2, \cdots n$. On notera les variables de f par $t \in \mathbb{R}$ et $x = (x_1, x_2, \cdots x_n) \in \mathbb{R}^n$, donc on a $f(t, x)$ ou $f(t, x_1, x_2, \cdots x_n)$.

On cherche un intervalle ouvert $I_0 \subset I$ et une fonction $y : I_0 \mapsto \mathbb{R}^n$ avec $y = (y_1, y_2, \cdots y_n)^T$ où $y_k : I_0 \mapsto \mathbb{R}$ $\forall\, k = 1, 2, \cdots n$, tel que $y \in C^1(I_0)$ et tel que

$$y'(t) = f(t, y(t)) \qquad \forall t \in I_0 \tag{2.1}$$

c'est à dire

$$\left\{\begin{array}{lll} y_1'(t) & = & f_1(t, y_1(t), y_2(t), \cdots y_n(t)) \\ y_2'(t) & = & f_2(t, y_1(t), y_2(t), \cdots y_n(t)) \\ \quad . & & \\ \quad . & & \\ y_n'(t) & = & f_n(t, y_1(t), y_2(t), \cdots y_n(t)) \end{array}\right. \tag{2.2}$$

ceci pour tout $t \in I_0$.
On dira que (2.1) ou (2.2) est un **système d'équations différentielles ordinaire (EDO)** d'ordre 1 ou encore une **équation différentielle ordinaire vectorielle** d'ordre 1.
On cherchera toujours l'intervalle I_0 qui est "le plus grand que possible".
Remarque : Dans le cas $n = 1$ le système se réduit à une seule équation ; on dit alors qu'on a une **équation différentielle scalaire d'ordre 1.**

Exemple d'un système EDO : mouvement rectiligne d'un corp sous l'action d'une force donnée, dépendant uniquement de la position et de la vitesse. On note $t \in \mathbb{R}$ le temps, $y_1(t)$ la position et $y_2(t)$ la vitesse. Alors y_1, y_2 satisfont le système EDO suivant :

$$\left\{\begin{array}{lll} y_1'(t) & = & y_2(t) \\ & & \\ y_2'(t) & = & g(y_1(t),\, y_2(t)) \end{array}\right.$$

où $g : \mathbb{R}^2 \to \mathbb{R}$ est une fonction donnée ; $g(x_1, x_2)$ représente la force qui s'exerce sur le corp s'il est à la position x_1 et il a la vitesse x_2. En notant $y = (y_1, y_2)^T$, ce système peut s'écrire sous la forme vectorielle

$$y'(t) = f(t, y(t))$$

avec $f : \mathbb{R}^3 \mapsto \mathbb{R}^2$ la fonction donnée par

$$f(t, x_1, x_2) = \begin{pmatrix} x_2 \\ g(x_1, x_2) \end{pmatrix} \qquad \forall\, (t, x_1, x_2) \in \mathbb{R}^3.$$

Remarquer que f est indépendante du temps t.

En général on a des solutions de (2.1) mais elles ne sont pas uniques.
Par exemple l'équation différentielle

$$y' = y$$

a une infinité des solutions $y(t) = ce^t$ avec $c \in \mathbb{R}$ une constante arbitraire.
On peut aussi donner l'exemple d'un système :

$$\left\{\begin{array}{l} y_1' = y_1 \\ y_2' = y_2 \end{array}\right.$$

qui a encore une infinités de solutions : $(y_1(t), y_2(t)) = (c_1 e^t, c_2 e^t)$ avec $(c_1, c_2) \in \mathbb{R}^2$ vecteur constant arbitraire.

On peut alors imaginer qu'un système comme (2.1) aura une infinité de solutions données à n constantes près $c_1, c_2, \cdots c_n$.

Le plus souvent on cherche à résoudre le système (2.1) avec ce qu'on appelle une "condition initiale", c'est à dire, on se donne $t_0 \in I$ et $y^0 \in \mathbb{R}^n$, $y^0 = (y_1^0, y_2^0, \cdots y_n^0)^T \in \mathbb{R}^n$ et on cherche un intervalle ouvert $I_0 \subset I$ tel que $t_0 \in I_0$ et une fonction $y : I_0 \mapsto \mathbb{R}^n$ satisfaisant (2.1) ainsi que la condition initiale

$$y(t_0) = y^0. \tag{2.3}$$

On appelle alors **problème de Cauchy** le système (2.1) et (2.3).

Sous des hypothèses appropriées pour la fonction f on peut montrer **l'existence et l'unicité** d'une solution du problème de Cauchy ; cette question ne sera pas abordé dans ce cours.

Exemple : On cherche $y : \mathbb{R} \mapsto \mathbb{R}$ satisfaisant le problème de Cauchy

$$\begin{cases} y'(t) &= y(t) \\ y(0) &= 3. \end{cases} \tag{2.4}$$

Dans cet exemple $n = 1, f(t,x) = x, t_0 = 0, y^0 = 3$. La solution générale de l'équation principale est $y(t) = Ce^t$ pour toute constante $C \in \mathbb{R}$. En utilisant la condition initiale on a $Ce^0 = 3$ ce qui donne $C = 3$.
Alors $y(t) = 3e^t \quad \forall\, t \in \mathbb{R}$ est l'unique solution du système (2.4).

2.1.3 Equations différentielles scalaires d'ordre n

Soit $I \subset \mathbb{R}$ un intervalle ouvert et $n \in \mathbb{N}^*$. On se donne une fonction continue $g : I \times \mathbb{R}^n \mapsto \mathbb{R}$, $g(t, x_1, x_2, \cdots x_n)$. On cherche un intervalle ouvert $I_0 \subset I$ et une fonction $y : I_0 \to \mathbb{R}$ avec $y \in C^n(I_0)$ solution de **l'équation différentielle scalaire d'ordre** n suivante :

$$y^{(n)}(t) = g(t, y(t), y'(t), y''(t), \cdots y^{(n-1)}(t)) \quad \forall\, t \in I_0. \tag{2.5}$$

Comme pour les systèmes d'ordre 1 on cherche I_0 "le plus grand que possible".

Remarque : On peut ramener la résolution de (2.5) à la résolution d'un système de type (2.1) ou (2.2) de la manière suivante : on introduit des nouvelles inconnues

$$\begin{aligned} &y_1(t) = y(t) \\ &y_2(t) = y'(t) \\ &. \\ &. \\ &y_n(t) = y^{(n-1)}(t) \end{aligned}$$

Si y satisfait (2.5) alors $(y_1, y_2, \cdots y_n)$ satisfait le système différentiel d'ordre 1 suivant :

$$
\begin{aligned}
&y'_1 = y_2 \\
&y'_2 = y_3 \\
&. \\
&. \\
&y'_{n-1} = y_n \\
&y'_n = g(t, y_1, y_2, \cdots y_n)
\end{aligned}
$$

et réciproquement, si $y_1, y_2, \cdots y_n$ satisfont ce dernier système alors en posant $y = y_1$ alors y satisfait (2.5).
En posant $Y = (y_1, y_2, \cdots y_n)^T$, le système ci-dessus s'écrit sous la forme vectorielle

$$Y'(t) = F(t, Y)$$

où la fonction $F : I \times \mathbb{R}^n \to \mathbb{R}^n$ est définie par

$$(t, x_1, x_2, \cdots x_n)^T \mapsto F(t, x_1, x_2, \cdots x_n) = (x_2, x_3, \cdots x_n, g(t, x_1, x_2, \cdots x_n))^T.$$

Comme pour (2.1) on aura en général pour (2.5) une infinité de solutions définies à n constantes près, $c_1, c_2, \cdots c_n \in \mathbb{R}$.
Pour avoir l'unicité il faut ici aussi une condition initiale ; on se donne $t_0 \in I$ et $y_1^0, y_2^0, \cdots y_n^0 \in \mathbb{R}$ et on cherche un intervalle ouvert $I_0 \subset I$ et une fonction $y : I_0 \to \mathbb{R}$ satisfaisant (2.5) avec en plus

$$
\left\{
\begin{array}{ccc}
y(t_0) & = & y_1^0 \\
y'(t_0) & = & y_2^0 \\
. & & \\
. & & \\
y^{(n-1)}(t_0) & = & y_n^0
\end{array}
\right. \tag{2.6}
$$

On appelle encore **problème de Cauchy** le système (2.5) et (2.6).
Pour ce problème aussi, sous des hypothèses approriées pour la fonction g on montre **l'existence et l'unicité** d'une solution du système (2.5) et (2.6), question qui ne sera pas abordée dans ce cours.

2.2 Equations différentielles à variables séparées

Dans cette section on se place dans le cas $n = 1$, c'est à dire, on considère une équation différentielle scalaire. On considère I et J deux intervalles ouvertes dans $\mathbb{R}$ et deux fonctions continues $a : I \to \mathbb{R}$ et $b : J \to \mathbb{R}$. On introduit alors la fonction f de deux variables : $f : I \times J \to \mathbb{R}$ donnée par

$$f(t, x) = a(t)\, b(x), \quad \forall\ (t, x) \in I \times J$$

L'équation différentielle à résoudre sera alors : trouver $I_0 \subset I$ un intervalle ouvert et $y : I_0 \to \mathbb{R}$ (en fait $y : I_0 \to J$) telles que

$$y'(t) = f(t, y(t)) = a(t)b(y(t)), \quad \forall\, t \in I_0. \tag{2.7}$$

On appelle une telle équation **équation différentielle à variables séparées**.
Dans la suite nous montrons comment on résoude en général l'équation (2.7). Nous notons par $M \subset J$ l'ensemble de toutes les racines de b, donc

$$M = \{x \in J, \quad b(x) = 0\}.$$

(M peut être l'ensemble vide $\emptyset$). L'ensemble M est fermé (car $M = b^{-1}(\{0\})$ et le singleton $\{0\}$ est un ensemble fermé) donc l'ensemble $J - M$ est ouvert.
Supposons que $M \neq J$ ce qui revient à dire que $J - M \neq \emptyset$ ou que b n'est pas la fonction constante nulle sur J.
Il est évident que si M n'est pas l'ensemble vide alors pour tout $r \in M$ la fonction constante $y(t) = r \quad \forall\, t \in I$, est une solution de l'équation différentielle (2.7) (l'intervalle ouvert I_0 correspondant sera alors $I_0 = I$).

Pour trouver d'autres solutions que les fonctions constantes, nous considérons B une primitive de la fonction $\frac{1}{b}$ sur $J - M$, donc B est une fonction de classe C^1 sur l'ensemble ouvert $J - M$ tel que

$$B'(x) = \frac{1}{b(x)}, \quad \forall\, x \in J - M.$$

(une telle primitive existe car $\frac{1}{b}$ est une fonction continue sur $J - M$).
D'autre part, soit A une primitive de a sur I, c'est à dire :

$$A'(t) = a(t), \quad \forall\, t \in I.$$

Nous avons

Théorème 2.1. *Soit $I_0 \subset I$ un intervalle ouvert et $y : I_0 \to J$ une fonction de classe C^1 telle que*

$$b(y(t)) \neq 0, \quad \forall\, t \in I_0$$

(autrement dit : $y(t) \in J - M, \quad \forall\, t \in I_0$).
Alors y est une solution de (2.7) *si et seulement si il existe $C \in \mathbb{R}$ tel que*

$$B(y(t)) = A(t) + C, \quad \forall\, t \in I_0. \tag{2.8}$$

Démonstration. " $\Longrightarrow$ " Si y est solution de (2.7) alors en divisant cette équation par $b(y(t))$ (car non nulle) on a

$$\frac{y'(t)}{b(y(t))} = a(t), \quad \forall\, t \in I_0$$

ce qui est équivalent à

$$\frac{d}{dt}B(y(t)) = \frac{d}{dt}A(t), \quad \forall\, t \in I_0.$$

Il est clair alors qu'il existe $C \in \mathbb{R}$ tel qu'on a (2.8).
" $\Longleftarrow$ " Si y est solution de (2.8) alors en dérivant cette égalité en t on trouve immédiatement (2.7). $\square$

Exemple 2.1. *On se propose de résoudre l'équation différentielle*

$$y' = t\, y^2. \tag{2.9}$$

C'est une équation différentielle du type (2.7) *(à variables séparées) avec*

$$I = \mathbb{R}$$

$$J = \mathbb{R}$$

$$a(t) = t, \quad \forall\, t \in \mathbb{R}$$

$$b(x) = x^2, \quad \forall\, x \in \mathbb{R}$$

donc $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *donnée par*

$$f(t, x) = tx^2, \quad \forall\, (t, x) \in \mathbb{R}^2.$$

Cherchons d'abord les racines de b *: nous avons* $b(r) = 0 \Longleftrightarrow r^2 = 0 \Longleftrightarrow r = 0$, *donc* $M = \{0\}$. *Nous avons alors immédiatement une solution de* (2.9) *:* $I_0 = \mathbb{R}$ *et* $y(t) = 0 \quad \forall\, t \in \mathbb{R}$.

Pour trouver d'éventuelles solutions non nulles, nous divisons l'équation (2.9) *par* $y^2(t)$ *(en supposant* $y(t) \neq 0 \quad \forall\, t$*). On a alors*

$$\frac{y'(t)}{y^2(t)} = t$$

qui est équivalent à

$$\frac{d}{dt}\left(-\frac{1}{y(t)}\right) = \frac{d}{dt}\left(\frac{t^2}{2}\right).$$

Ceci est équivalent à : $\exists C \in \mathbb{R}$ *tel que*

$$-\frac{1}{y(t)} = \frac{t^2}{2} + C \tag{2.10}$$

Remarque : *On aurait pu arriver à* (2.10), *en appliquant directement le Théorème 2.1 ; pour ceci on introduit une primitive de* $\frac{1}{x^2}$, *c'est à dire* $B : \mathbb{R} - \{0\} \to \mathbb{R}$ *avec*

$$B(x) = -\frac{1}{x}, \quad \forall\, x \neq 0$$

et une primitive de t, *c'est à dire* $A : \mathbb{R} \to \mathbb{R}$ *avec*

$$A(t) = \frac{t^2}{2}, \quad \forall\, t \in \mathbb{R}.$$

Alors (2.10) *est une conséquence directe de l'égalité* (2.8) *du Théorème 2.1.*

Pour trouver les solutions non nulles de (2.9) *il suffira alors d'exprimer* $y(t)$ *en fonction de* t *à partir de* (2.10). *Nous avons*

$$y(t) = -\frac{1}{t^2/2 + C} = -\frac{2}{t^2 + 2C}$$

Donc pour toute constante $C \in \mathbb{R}$ *nous avons une solution*

$$y(t) = -\frac{2}{t^2 + 2C}. \tag{2.11}$$

Il reste à trouver l'intervalle ouvert I_0. *Nous avons les cas suivants :*
Cas 1) : $C > 0$.
Dans ce cas comme la solution (2.11) *est définie pour tout* $t \in \mathbb{R}$ *on a* $I_0 = \mathbb{R}$.
Cas 2) : $C = 0$.
Dans ce cas comme la solution (2.11) *est définie pour tout* $t \neq 0$ *on a* $I_0 =]-\infty,\ 0[$ *ou* $I_0 =]0,\ +\infty[$.
Cas 3) : $C < 0$.
Comme dans ce cas la solution (2.11) *est définie pour tout* $t \neq \pm\sqrt{-2C}$ *on a* $I_0 =]-\infty,\ -\sqrt{-2C}[$ *ou* $I_0 =]-\sqrt{-2C},\ \sqrt{-2C}[$ *ou* $I_0 =]\sqrt{-2C},\ +\infty[$.

D'autre part, on peut considérer l'équation différentielle (2.9) *avec condition initiale*

$$y(1) = 2 \tag{2.12}$$

(on a donc un problème de Cauchy; ici $t_0 = 1$ *et* $y_0 = 2$*). Nous considérons d'abord la solution constante* $y(t) = 0 \quad \forall t \in \mathbb{R}$ *et nous observons que cette solution ne peut pas satisfaire* (2.12).
Regardons alors les solutions données par (2.11). *On est amenés à résoudre l'équation avec inconnue* C *:*

$$-\frac{2}{1 + 2C} = 2$$

ce qui donne comme seule solution $C = -1$. *Comme* $C < 0$ *on est dans le* **Cas 3)** *et il y a 3 choix possibles pour l'intervalle* I_0 *; on prendra celui qui contient* $t_0 = 1$, *donc* $I_0 =]-\sqrt{2},\ \sqrt{2}[$. *Donc en conclusion la solution du problème de Cauchy* (2.9) *et* (2.12) *est donnée par* $y :]-\sqrt{2},\ \sqrt{2}[\to \mathbb{R}$ *avec*

$$y(t) = \frac{2}{2 - t^2}, \quad \forall\, t \in]-\sqrt{2},\ \sqrt{2}[.$$

2.3 EDO scalaires linéaires d'ordre n

2.3.1 Généralités

On se donne ici $n \in \mathbb{N}^*$, $I \subset \mathbb{R}$ un intervalle ouvert et $b, a_0, a_1, \cdots a_{n-1} : I \subset \mathbb{R}$ des fonctions continues.

L'équation différentielle scalaire d'ordre n à résoudre est : trouver $y : I \to \mathbb{R}$ avec $y \in C^n(I)$ telle que

$$y^{(n)} = b(t) + a_0(t)y + a_1(t)y' + \cdots a_{n-1}(t)y^{(n-1)}. \tag{2.13}$$

Remarquons qu'il s'agit d'une équation du type général

$$y^{(n)} = g(t, y, y', \cdots y^{(n-1)})$$

avec $g : I \times \mathbb{R}^n \to \mathbb{R}$ donnée par

$$g(t, x_1, x_2, \cdots x_n) = b(t) + a_0(t)x_1 + a_1(t)x_2 + \cdots a_{n-1}(t)x_n, \quad \forall\ (t, x_1, \cdots x_n) \in I \times \mathbb{R}^n.$$

Remarquons que la fonction g est une fonction affine par rapport aux variables $x_1, x_2, \cdots x_n$; on dit souvent "linéaire" à la place de "affine" et c'est pourquoi on dit que l'EDO (2.13) est une équation linéaire.

Remarque 2.1. *Si b est la fonction constante 0 alors on dit que l'équation* (2.13) *est* **homogène**. *Elle s'écrit :*

$$y^{(n)} = a_0(t)y + a_1(t)y' + \cdots a_{n-1}(t)y^{(n-1)}. \tag{2.14}$$

On admet le résultat suivant :

Proposition 2.1. *Pour tout $t_0 \in I$ et tous $\alpha_0, \alpha_1, \cdots \alpha_{n-1} \in \mathbb{R}$ nous avons l'existence et l'unicité d'une solution du problème de Cauchy* (2.13) *et* (2.15)*, avec*

$$\left\{ \begin{array}{c} y(t_0) = \alpha_0 \\ y'(t_0) = \alpha_1 \\ \cdot \\ \cdot \\ y^{(n-1)}(t_0) = \alpha_{n-1}. \end{array} \right. \tag{2.15}$$

Dans la suite nous donons la structure des solutions de l'équation homogène ; notons pour toute la suite du cours par $\mathcal{S}$ l'ensemble de toutes les solutions de (2.14). Nous avons le résultat suivant :

Lemme 2.1. *L'ensemble $\mathcal{S}$ est un espace vectoriel réel (qu'il faut voir comme un sous-espace vectoriel de l'espace vectoriel réel des fonctions de I dans $\mathbb{R}$). En plus la dimension de $\mathcal{S}$ est égale à n.*

Démonstration. 1. Montrons d'abord que $\mathcal{S}$ est un sous-espace vectoriel de l'espace vectoriel réel des fonctions de I dans $\mathbb{R}$. Soient $\varphi, \psi \in \mathcal{S}$ et $\lambda, \mu \in \mathbb{R}$. On écrit que φ et ψ satisfont chacune l'équation (2.14). On multiplie l'équation satisfaite par φ par λ, celle de ψ par μ et on fait la somme. On utilise pour tout $j \in [[0, n-1]]$ l'égalité

$$\lambda\varphi^{(j)} + \mu\psi^{(j)} = (\lambda\varphi + \mu\psi)^{(j)}$$

(linéarité de la dérivé à l'ordre j) ce qui nous dit, en regroupant les termes respectifs que $\lambda\varphi + \mu\psi$ satisfait (2.14), donc c'est bien un élément de $\mathcal{S}$.

2. On va donner l'idée de la preuve du fait que $dim(\mathcal{S}) = n$. Pour cela il faut construire une base de n éléments de $\mathcal{S}$. L'idée est de considérer pour tout $k \in [[0, n-1]]$ la fonction notée f_k solution de (2.14) satisfaisant en plus la condition initiale

$$f_k^{(j)}(t_0) = \delta_{k,j}, \quad j \in [[0, n-1]]$$

(c'est une problême de Cauchy), où t_0 est un élément fixé de I. On admet que la famille $\{f_0, f_1, \cdots f_{n-1}\}$ est bien une base en $\mathcal{S}$, ce qui prouve le résultat.

□

Remarque 2.2. *Nous pouvons montrer par la même méthode que l'ensemble des solutions de* (2.14) *à valeurs* **complexes** *(c'est à dire, telles que l'égalité* (2.14) *soit valable en* $\mathbb{C}$*) est aussi une espace vectoriel sur* $\mathbb{C}$ *de dimension* n*. Ici encore il faut voir cet espace comme un sous-espace vectoriel de l'espace vectoriel des fonctions définies sur* I *à valeurs dans* $\mathbb{C}$*.*

Pour la suite de cette section nous notons $\mathcal{B} = \{\varphi_1, \varphi_2, \cdots \varphi_n\}$ une base fixée dans $\mathcal{S}$. Nous avons en fait montré

Proposition 2.2. $\varphi \in C^n(I)$ *est une solution de* (2.14) *si et seulement si il existe des constantes* $C_1, C_2 \cdots C_n \in \mathbb{R}$ *telles que*

$$\varphi = C_1\varphi_1 + C_2\varphi_2 + \cdots C_n\varphi_n.$$

Considéron maintenant l'équation non homogènen (2.13) avec b une fonction continue. Nous avons

Proposition 2.3. *Soit* $\tilde{\varphi}$ *une solution particulière de* (2.13)*. Alors* $\varphi \in C^n(I)$ *est une solution de* (2.13) *si et seulement si il existe des constantes* $C_1, C_2 \cdots C_n \in \mathbb{R}$ *telles que*

$$\varphi = \tilde{\varphi} + C_1\varphi_1 + C_2\varphi_2 + \cdots C_n\varphi_n.$$

Démonstration. Soit $\varphi \in C^n(I)$ et notons $\psi = \varphi - \tilde{\varphi}$. Montrons que φ est une solution de (2.13) si et seulement si ψ est une solution de (2.14). En effet, si φ satisfait (2.13), comme $\tilde{\varphi}$ satisfait aussi (2.13) en faisant la différence entre les 2 équations et en utilisant la linéarité on obtient que ψ satisfait (2.14). Inversement, si ψ satisfait (2.14), comme $\tilde{\varphi}$ satisfait (2.13), en faisant la somme entre les 2 égalités et en utilisant de nouveau la linéarité on en déduit que φ satisfait (2.13), ce qui montre le résultat d'équivalence.
En utilisant pour ψ la Proposition 2.2 on obtient le résultat voulu. □

Remarque : Il y a une méthode appellée **méthode de variation des constantes** qui permet de construire une solution particulière de (2.13) si on connait une base de $\mathcal{S}$. Cette méthode ne sera pas abordée dans ce cours.

2.3.2 Le cas particulier des coefficients constants

On suppose dans cette subsection que $a_0, a_1, \cdots a_{n-1}$ sont des constantes réelles. Le problème (2.13) devient alors

$$y^{(n)} = b(t) + a_0 y + a_1 y' + \cdots a_{n-1} y^{(n-1)}. \tag{2.16}$$

Le but ici est de trouver n solution indépendentes du problème homogène associé

$$y^{(n)} = a_0 y + a_1 y' + \cdots a_{n-1} y^{(n-1)} \tag{2.17}$$

donc une base dans $\mathcal{S}$.
Pour cela nous considérons le **polynome caractéristique** associé à l'équation (2.17), qui est un polynome à coefficients réels

$$P(z) = z^n - a_{n-1} z^{n-1} - a_{n-2} z^{n-2} - \cdots - a_1 z - a_0.$$

Il est bien connu qu'un tel polynome admet s racines complexes distinctes que nous notons $\mu_1, \mu_2, \cdots \mu_s \in \mathbb{C}$, de multiplicités respectives $m_1, m_2, \cdots m_s \in \mathbb{N}^*$ avec

$$s \in [[1, n]], \qquad 1 \leq m_j \leq n \quad \forall\, j \in [[1, s]] \qquad \text{et} \qquad m_1 + m_2 + \cdots m_s = n.$$

Ceci veut dire qu'on a la décomposition suivante en facteurs pour P :

$$P(z) = (z - \mu_1)^{m_1} (z - \mu_2)^{m_2} \cdots (z - \mu_s)^{m_s}, \quad \forall\, z \in \mathbb{C}.$$

On a le résultat suivant :

Lemme 2.2. *Pour tout $k \in [[1, s]]$ et $r \in [[0, m_k - 1]]$ la fonction à valeurs complexes*

$$t \in I \mapsto t^r e^{\mu_k t} \in \mathbb{C}$$

est une solution de (2.17).

Démonstration. Il faut montrer qu'on a $I = 0$ où on note

$$I = \sum_{l=0}^{n} a_l (t^r e^{\mu_k t})^{(l)}$$

avec $a_n = -1$. Nous utilisons la formule de Leibnitz :

$$(t^r e^{\mu_k t})^{(l)} = \sum_{j=0}^{l} \frac{l!}{j!\,(l-j)!} (t^r)^{(j)} (e^{\mu_k t})^{(l-j)}$$

ce qui nous donne

$$I = \sum_{l=0}^{n} \sum_{j=0}^{l} a_l \frac{l!}{j!\,(l-j)!} \mu_k^{l-j} (t^r)^{(j)} e^{\mu_k t}$$

En inversant l'ordre de sommation nous pouvons écrire

$$I = \sum_{j=0}^{n} \left(\sum_{l=j}^{n} a_l \frac{l!}{(l-j)!} \mu_k^{l-j} \right) (t^r)^{(j)} \frac{1}{j!} e^{\mu_k t}$$

En remarquant que $\sum_{l=j}^{n} a_l \frac{l!}{(l-j)!} \mu_k^{l-j} = P^{(j)}(\mu_k)$ nous avons

$$I = \sum_{j=0}^{n} P^{(j)}(\mu_k)(t^r)^{(j)} \frac{1}{j!} e^{\mu_k t}.$$

Comme r est strictement inférieur à la multiplicité de μ_k, nous avons

$$P(\mu_k) = P'(\mu_k) = \cdots = P^{(r)}(\mu_k) = 0. \tag{2.18}$$

D'autre part on a $(t^r)^{(j)} = 0$ si $j > r$ et avec (2.18) on déduit que tous les terms de I sont 0, d'où $I = 0$ ce qui finit la preuve. □

Le lemme précédent nous donne exactement $m_1 + m_2 + \cdots + m_s = n$ solutions distinctes pour (2.17). Nous avons

Théorème 2.2. *L'ensemble des fonctions*

$$\tilde{\mathcal{B}} = \{e^{\mu_1 t},\ te^{\mu_1 t}, \cdots t^{m_1 - 1} e^{\mu_1 t},\ e^{\mu_2 t},\ te^{\mu_2 t}, \cdots t^{m_2 - 1} e^{\mu_2 t}, \cdots e^{\mu_s t},\ te^{\mu_s t}, \cdots t^{m_s - 1} e^{\mu_s t}\}$$

forme une base dans l'espace vectoriel des solutions complexes de (2.17).

Démonstration. On a vu dans le lemme précédent que $\tilde{\mathcal{B}}$ est un sous-ensemble contenant n éléments de l'espace vectoriel des solutions complexes de (2.17). Comme cet espace vectoriel est de dimension n, il suffit de montrer que la famille $\tilde{\mathcal{B}}$ est libre ; on admet ce résultat, ce qui finit la preuve. □

Montrons maintenant comment construire une base $\mathcal{B}$ dans l'espace vectoriel $\mathcal{S}$ des solutions **réelles** de (2.17). Il y a deux cas :

Cas 1. Supposons que toutes les racines $\mu_1, \mu_2, \cdots \mu_s$ de P sont réelles. Alors les éléments de la famille $\tilde{\mathcal{B}}$ sont des fonctions à valeurs dans $\mathbb{R}$, donc dans ce cas $\tilde{\mathcal{B}}$ est la base recherchée ; on a alors $\mathcal{B} = \tilde{\mathcal{B}}$.

Cas 2. Supposons qu'il y a des racines μ_k de P qui sont dans $\mathbb{C} - \mathbb{R}$. Alors comme P est à coefficients réels, ces racines sont toujours par paires conjuguées (c'est à dire, si μ_k est racine de P alors sa conjuguée complexe $\overline{\mu_k}$ est aussi racine de P).

Comme $\overline{\exp(\mu_k t)} = \exp(\overline{\mu_k} t)$ on déduit du Lemme 2.2 que si la fonction $t^r \exp(\mu_k t)$ est un élément de $\tilde{\mathcal{B}}$ alors la fonction $\overline{t^r \exp(\mu_k t)}$ est aussi un élément de $\tilde{\mathcal{B}}$, donc les éléments de $\tilde{\mathcal{B}}$ sont par paires conjuguées.

On va remplacer alors dans $\tilde{\mathcal{B}}$ les deux fonctions $t^r \exp(\mu_k t)$ et $t^r \exp(\overline{\mu_k} t)$ par $\frac{1}{2}(t^r \exp(\mu_k t) + t^r \exp(\overline{\mu_k} t))$ et $\frac{1}{2i}(t^r \exp(\mu_k t) + t^r \exp(\overline{\mu_k} t))$ qui sont respectivement les fonctions à valeurs rélles $Re(t^r \exp(\mu_k t))$ et $Im(t^r \exp(\mu_k t))$. En faisant ceci successivement pour tout k tel que $\mu_k \in \mathbb{C}-\mathbb{R}$ et pour tout $r \in \{0, 1, \cdots m_k-1\}$ l'ensemble $\tilde{\mathcal{B}}$ se transforme en un ensemble $\mathcal{B}$ de solutions rélles de (2.17) qui aura toujours n éléments indépendents et qui sera donc une base dans $\mathcal{S}$.

Exemples : en classe

2.4 Systèmes d'EDO d'ordre 1 et linéaires

2.4.1 Généralités

On se donne $n \in \mathbb{N}^*, I \subset \mathbb{R}$ un intervalle ouvert et une fonction $A : I \to \mathcal{M}_n(\mathbb{R})$ telle que si on pose $A(t) = (A_{ij}(t))_{i,j=1,2,\cdots n}$ avec $A_{ij} : I \to \mathbb{R}$ alors A_{ij} est une fonction continue pour tous $i, j \in [[1, n]]$.
On se donne aussi $b : I \to \mathbb{R}^n$ une fonction continue (donc si on pose $b(t) = (b_1(t), b_2(t), \cdots b_n(t))^T$ avec $b_i : I \to \mathbb{R}$ alors b_i est une fonction continue pour tout $i \in [[1, n]]$).
Le problème à résoudre est : trouver $y : I \to \mathbb{R}^n,\ y \in C^1(I, \mathbb{R}^n)$ telle que

$$y'(t) = A(t)y(t) + b(t), \quad \forall\, t \in I. \tag{2.19}$$

Considérons aussi le problème homogène associé (correspondant à $b = 0$) :

$$y'(t) = A(t)y(t), \quad \forall\, t \in I. \tag{2.20}$$

Remarquons que ce sont des équations du type général

$$y'(t) = f(t, y(t))$$

avec $f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ donnée par $f(t, x) = A(t)x + b(t)$ (pour (2.19)) et respectivement $f(t, x) = A(t)x$ (pour (2.20)).
On a l'analogue suivant des résultats obtenus en Subsection 2.3.1 :

Proposition 2.4. *1. Pour tous $t_0 \in I$ et $y^0 \in \mathbb{R}^n$ nous avons l'existence et l'unicité d'une solution y de* (2.19) *et* (2.21) *avec*

$$y(t_0) = y^0. \tag{2.21}$$

(problème de Cauchy).

2. *Si $b = 0$ alors l'ensemble des solutions de* (2.20)*, que nous notons $\mathcal{S}_n$, est un espace vectoriel réel de dimension n.*
   Alors en supposant qu'on connait une base $\mathcal{B}_n$ en $\mathcal{S}_n$ avec

$$\mathcal{B}_n = \{\psi_1, \psi_2, \cdots \psi_n\}$$

   nous avons que $y \in C^1(I, \mathbb{R}^n)$ est une solution de (2.20) *si et seulement si il existe $C_1, C_2, \cdots C_n \in \mathbb{R}$ tels que $y = C_1\psi_1 + C_2\psi_2 \cdots C_n\psi_n$.*
3. *Si $b \neq 0$ et si on connait en plus une solution particulière $\tilde{y}$ de* (2.19) *alors $y \in C^1(I, \mathbb{R}^n)$ est solution de* (2.19) *si et seulement si il existe $C_1, C_2, \cdots C_n \in \mathbb{R}$ tels que $y = \tilde{y} + C_1\psi_1 + C_2\psi_2 \cdots C_n\psi_n$.*

Remarque 2.3. *On peut associer à la base $\mathcal{B}_n$ une fonction $\mathcal{M} : I \to \mathcal{M}_n(\mathbb{R})$ telle que pour tout $t \in I$ la j - ème colonne de $\mathcal{M}(t)$ est $\psi_j(t)$, pour tout $j \in [[1, n]]$. La matrice $\mathcal{M}(t)$ s'appelle* **matrice fondamentale** *pour le système* (2.20)*.*
On admet qu'une matrice fondamentale est toujours inversible.
D'autre part, il est facile de voir que

$$\mathcal{M}'(t) = A(t)\mathcal{M}(t), \quad \forall\, t \in I \tag{2.22}$$

(car la j - ème colonne de $\mathcal{M}'(t)$ est $\psi_j'(t)$ et la j - ème colonne de $A(t)\mathcal{M}(t)$ est $A(t)\psi_j(t)$).
Remarquons aussi que toute solution de (2.20) *s'écrit*

$$y(t) = \mathcal{M}(t)C \qquad \text{avec} \quad C = (C_1, C_2, \cdots C_n)^T \in \mathbb{R}^n \quad \text{vecteur constant abitraire.}$$

2.4.2 La méthode de variation des constantes

Cette méthode permet d'obtenir une solution particulière de (2.19). On va chercher une telle solution sous la forme

$$\tilde{y}(t) = \mathcal{M}(t)C(t) \tag{2.23}$$

avec $\mathcal{M}(t)$ matrice fondamentale du système (2.20) et $C(t) \in \mathbb{R}^n$ un vecteur à trouver. On remplace en (2.19), ce qui donne

$$\mathcal{M}'(t)C(t) + \mathcal{M}(t)C'(t) = A(t)\mathcal{M}(t)C(t) + b(t).$$

En utilisant (2.22) et en simplifiant, on obtient

$$\mathcal{M}(t)C'(t) = b(t)$$

ce qui donne $C'(t) = \mathcal{M}(t)^{-1}b(t)$.
En conclusion, une solution particulière de (2.19) est $\tilde{y}$ donnée par (2.23), où la fonction vectorielle C est une primitive en t de $\mathcal{M}(t)^{-1}b(t)$.

2.4.3 Le cas particulier des coefficients constants

On suppose ici que la fonction à valeurs matricielles A est constante (indépendante de t) et on va noter toujours par A cette matrice constante. Les systèmes différentielles (2.19) et (2.20) deviennent

$$y'(t) = Ay(t) + b(t) \tag{2.24}$$

et respectivement

$$y'(t) = Ay(t). \tag{2.25}$$

Le but est de trouver $\mathcal{M}(t)$ une matrice fondamentale associée au système (2.25), c'est à dire, trouver n solutions indépendantes de (2.25). Nous utiliserons le résultat suivant :

Lemme 2.3. *Si $\lambda \in \mathbb{C}$ est une valeur propre complexe de la matrice A avec $P \in \mathbb{C}^n$ vecteur propre complexe associé, alors la fonction*

$$t \in \mathbb{R} \quad \mapsto \quad e^{\lambda t} P \in \mathbb{C}^n$$

est une solution (complexe) de (2.25).

Démonstration. En utilisant le fait que $Ap = \lambda P$ on a

$$(e^{\lambda t}P)' = \lambda e^{\lambda t}P = e^{\lambda t}\lambda P = e^{\lambda t}AP = A(e^{\lambda t}P))$$

ce qui montre le résultat. □

On supposera dans la suite que A **est une matrice diagonalisable (en C)**. Le cas où A n'est pas diagonalisable ne sera pas abordé dans ce cours.
On pourra alors considérer :
$\lambda_1, \lambda_2, \cdots \lambda_n \in \mathbb{C}$ les valeurs propres de A répétées selon leur multiplicités.
$P_1, P_2, \cdots P_n \in \mathbb{C}^n$ les vecteurs propres correspondant aux valeurs propres ci-dessus, et qui forment une base en C^n.
On admet que

$$\tilde{\mathcal{B}}_n = \{e^{\lambda_1 t}P_1,\ e^{\lambda_2 t}P_2,\ \cdots e^{\lambda_n t}P_n\}$$

est une base dans l'espace vectoriel des solutions complexes de (2.25). Il y a deux cas :
Cas 1. Si toutes les valeurs propres de A sont réeles (donc A est diagonalisable en $\mathbb{R}$) alors les éléments de $\tilde{\mathcal{B}}_n$ sont tous dans $\mathbb{R}^n$; dans ce cas une base dans $\mathcal{S}_n$ est $\mathcal{B}_n = \tilde{\mathcal{B}}_n$.
Cas 2. S'il existe des valeurs propres de A qui sont dans $\mathbb{C} - \mathbb{R}$ alors on va procéder comme dans la Subsection 2.3.2 : les éléments de $\tilde{\mathcal{B}}_n$ sont par paires conjugées complexes (car les valeurs propres de A le sont) et on remplacera successivement chaque paire conjuguée complexe de $\tilde{\mathcal{B}}_n$ par la paire réelle (partie réelle, partie imaginaire) de l'un des éléments de la paire conjuguée complexe. On obtiendra finalement une base $\tilde{B}_n$ en $\mathcal{S}_n$.
Exemple : en classe

Remarque 2.4. *En fait on peut traiter les EDO linéire d'ordre n considérées en Section 2.3 comme des systèmes d'EDO considérés dans cette section, en utilisant le changement d'inconnus vu en Subsection 2.1.3.*

Bibliographie :

1) J.J. Colin, J.M. Morvan, *Espaces vectoriels, applications linéaires*

2) H. Roudier, *Algèbre linéaire ; cours et exercices*

3) J.B. Hiriart-Urruty, *Les équations différentielles pour les débutants*

4) G. Gourmelen, H. Wadi, *Equations différentielles : théorie, algorithmes et méthodes*