Mathématiques pour l'informatique 2

Émilie Charlier

Année académique 2022-2023

Introduction

Informations générales

- **Code cours :** MATH2020
- **Prérequis :** MATH2019 - Mathématiques pour l'informatique 1
- **Titulaire :** Émilie Charlier
- **Assistant :** Antoine Renard
- **Contact :** Campus du Sart-Tilman - zone Polytech 1
  Institut de Mathématique B37, bureau 1/28
  `echarlier@uliege.be`
  `antoine.renard@uliege.be`
- **Horaire du cours et des séances d'exercices :** Q1, mercredi de 8h15 à 12h30
- **Local :** S39 au B37 (Institut de mathématique)
- **But du cours :** Il s'agit de donner au futur informaticien une palette d'outils mathématiques utiles à sa discipline. Tout en assurant la rigueur propre aux mathématiques de chacun des sujets présentés, nous souhaitons maintenir une interaction constante entre problèmes concrets et leur formalisation.
- **Contenu du cours :** Nous couvrirons les thèmes suivants.
  - Polynômes
  - Espaces vectoriels
  - Diagonalisation de matrices complexes
  - Outils d'analyse
- **Support :** Les informations liées au cours sont disponibles sur eCampus. En particulier, vous y trouverez un syllabus et les transparents. Une bonne prise de note est indispensable.
- **WIMS :** Quatre listes d'exercices interactifs de la plateforme WIMS seront mises à disposition des étudiants durant des laps de temps prédéfinis. Les dates de début et de fin seront communiquées en temps utiles en classe et sur eCampus. À la fin de ces laps de temps, les points accordés par WIMS seront retenus et une note globale pour les feuilles WIMS sera attribuée en fin de quadrimestre. Cette note comptera pour 10% de la note finale de l'examen.
- **Examen :** Écrit comportant théorie et exercices.
- **Évaluation :** La répartition est la suivante :

WIMS	10%
Examen	90%

- **Modulation :** 5 crédits.

Chapitre 1

Polynômes

Dans ce chapitre et le suivant, nous supposerons toujours que $\mathbb{K}$ est l'un des quatre ensembles de nombres suivants : $\mathbb{C}$, $\mathbb{R}$, $\mathbb{Q}$ ou $\mathbb{Z}_m$ avec m un nombre premier [1]. Lorsque nous utiliserons les notations $+$ et $\cdot$ entre éléments de $\mathbb{K}$, il est entendu que nous faisons référence à l'addition et la multiplication usuelle dans $\mathbb{K}$. Rappelons que l'addition et la multiplication dans $\mathbb{Z}_m$ étaient notées $+_m$ et $\cdot_m$ dans le cours "Mathématique pour l'informatique 1". Nous ne prendrons plus toujours ces précautions s'il est clair que nous travaillons dans $\mathbb{Z}_m$. Nous utiliserons la notation $\mathbb{K}_0$ pour désigner l'ensemble $\mathbb{K} \setminus \{0\}$.

1.1 Polynômes formels

Dans cette section, nous définissons les polynômes formels. Plus loin, nous ferons le lien entre les polynômes formels et les fonctions polynomiales. Nous montrerons que ces deux notions coïncident dans les cas où $\mathbb{K}$ est $\mathbb{C}$, $\mathbb{R}$ ou $\mathbb{Q}$ mais pas lorsque $\mathbb{K} = \mathbb{Z}_m$.

Définition 1.1. Un *polynôme à coefficients dans* $\mathbb{K}$ est une expression de la forme

$$P = p_0 + p_1 X + p_2 X^2 + \cdots + p_d X^d$$

où d est un naturel, $p_0, p_1, \ldots, p_d$ sont des éléments de l'ensemble $\mathbb{K}$ et où X est un symbole spécial. Les éléments $p_0, p_1, \ldots$ de $\mathbb{K}$ sont appelés les *coefficients* de P tandis que X est appelé l'*indéterminée* de P. L'ensemble des polynômes à coefficients dans $\mathbb{K}$ et d'intéterminée X est noté $\mathbb{K}[X]$.
Un polynôme (d'indéterminée X) est donné par la suite de ses coefficients. Formellement, un polynôme de $\mathbb{K}[X]$ est en fait simplement une suite finie de $\mathbb{K}$, c'est-à-dire une suite d'éléments de $\mathbb{K}$ valant 0 à partir d'un certain rang :

$$P = (p_0, p_1, \ldots, p_d, 0, 0, \ldots).$$

Lorsque $p_d \neq 0$, le naturel d est appelé le *degré* de P et est noté $\deg(P)$. Le coefficient p_d est appelé le *coefficient dominant* de P. Par convention, on pose $\deg(0) = -\infty$. Un *polynôme constant* est un polynôme de degré 0 ou $-\infty$.
Lorsqu'on ne souhaite pas spécifier le degré d'un polynôme P, on écrit plus généralement

$$P = p_0 + p_1 X + p_2 X^2 + \cdots$$

où il est entendu que $p_n = 0$ pour tout $n > \deg(P)$.

1. En fait, $\mathbb{K}$ pourrait être n'importe quel *champ*, mais nous n'avons vu ensemble que ces quatre cas.

Définition 1.2 (Somme, produit et multiplication par un scalaire)**.** On munit l'ensemble $\mathbb{K}[X]$ de deux opérations internes (somme et produit)

$$\mathbb{K}[X] \times \mathbb{K}[X] \to \mathbb{K}[X],\ (P,Q) \mapsto P+Q \quad \text{et} \quad \mathbb{K}[X] \times \mathbb{K}[X] \to \mathbb{K}[X],\ (P,Q) \mapsto P \cdot Q$$

et d'une opération externe (multiplication par un scalaire)

$$\mathbb{K} \times \mathbb{K}[X] \to \mathbb{K}[X],\ (k,P) \mapsto kP$$

de la façon suivante. Si $P = p_0 + p_1X + p_2X^2 + \cdots$ et $Q = q_0 + q_1X + q_2X^2 + \cdots$ sont des polynômes de $\mathbb{K}[X]$ et si $k \in \mathbb{K}$, alors

$$\begin{aligned} P+Q &= (p_0+q_0) + (p_1+q_1)X + (p_2+q_2)X^2 + \cdots \\ P \cdot Q &= (p_0q_0) + (p_0q_1 + p_1q_0)X + (p_0q_2 + p_1q_1 + p_2q_0)X^2 + \cdots \\ kP &= (kp_0) + (kp_1)X + (kp_2)X^2 + \cdots . \end{aligned}$$

Autrement dit, pour tout $n \in \mathbb{N}$, le coefficient de X^n dans
- $P+Q$ est égal à $p_n + q_n$
- PQ est égal à $\sum_{i=0}^{n} p_i\, q_{n-i}$
- kP est égal à kp_n.

Comme il est d'usage, on note indifféremment PQ ou $P \cdot Q$.

Exemple 1.3.
- Soient $P = 2X^3 + 3X^2 - 1$ et $Q = X^2 - X + 2$ vu comme polynômes de $\mathbb{Q}[X]$. On calcule

$$\begin{aligned} P+Q &= (2X^3 + 3X^2 - 1) + (X^2 - X + 2) \\ &= 2X^3 + (3+1)X^2 - X + (-1+2) \\ &= 2X^3 + 4X^2 - X + 1 \\ PQ &= (2X^3 + 3X^2 - 1)(X^2 - X + 2) \\ &= (2\cdot 1)X^5 + (2\cdot(-1) + 3\cdot 1)X^4 + (2\cdot 2 + 3\cdot(-1) + 0\cdot 1)X^3 \\ &\quad + (3\cdot 2 + 0\cdot(-1) + (-1)\cdot 1)X^2 + (0\cdot 2 + (-1)\cdot(-1))X + ((-1)\cdot 2) \\ &= 2X^5 + X^4 + X^3 + 5X^2 + X - 2 \\ 3P &= 3(2X^3 + 3X^2 - 1) \\ &= (3\cdot 2)X^3 + (3\cdot 3)X^2 + (3\cdot 0)X + (3\cdot(-1)) \\ &= 6X^3 + 9X^2 - 3. \end{aligned}$$

- Soient $P = 2X^3 + 3X^2 - 1$ et $Q = X^2 - X + 2$ vu comme polynômes de $\mathbb{Z}_3[X]$. On a donc $P = 2X^3 + 2$ et $Q = X^2 + 2X + 2$. De la même façon que précédemment, on calcule

$$\begin{aligned} P+Q &= (2X^3 + 2) + (X^2 + 2X + 2) \\ &= 2X^3 + X^2 + 2X + (2+2) \\ &= 2X^3 + X^2 + 2X + 1 \\ PQ &= (2X^3 - 1)(X^2 - X + 2) \\ &= (2\cdot 1)X^5 + (2\cdot 2 + 0\cdot 1)X^4 + (2\cdot 2 + 0\cdot 2 + 0\cdot 1)X^3 \\ &\quad + (0\cdot 2 + 0\cdot 2 + 2\cdot 1)X^2 + (0\cdot 2 + 2\cdot 2)X + (2\cdot 2) \\ &= 2X^5 + X^4 + X^3 + 2X^2 + X + 1 \\ 3P &= 0P = 0. \end{aligned}$$

Exercice 1.4. Écrire in extenso les coefficients de X^3 et de X^4 dans PQ dans le cas général où $P = p_0 + p_1X + p_2X^2 + \cdots$ et $Q = q_0 + q_1X + q_2X^2 + \cdots$.

La proposition suivante explicite le comportement du degré par rapport à ces trois opérations.

Proposition 1.5. *Pour tous $P, Q \in \mathbb{K}[X]$ et tous $k \in \mathbb{K}_0$, on a*

1. $\deg(P + Q) \leq \max\{\deg(P), \deg(Q)\}$,
2. $\deg(P \cdot Q) = \deg(P) + \deg(Q)$,
3. $\deg(kP) = \deg(P)$.

Démonstration. Il s'agit d'une vérification immédiate. □

Remarquons que ces formules sont correctes même si P ou Q est nul, avec la convention que $\max\{-\infty, d\} = d$ et $-\infty + d = -\infty$ pour tout $d \in \mathbb{N} \cup \{-\infty\}$.

Corollaire 1.6. *Si $P, Q \in \mathbb{K}[X]$ sont tels que $P \cdot Q = 0$, alors $P = 0$ ou $Q = 0$.*

Démonstration. Nous montrons la contraposée. Si P et Q sont deux polynômes non nuls de $\mathbb{K}[X]$, alors $\deg(P \cdot Q) = \deg(P) + \deg(Q) \geq 0$, et donc $P \cdot Q \neq 0$. □

1.2 Division euclidienne de polynômes

De la même façon que nous avons montré la division euclidienne dans $\mathbb{Z}$ dans le cours "Mathématiques pour l'informatique 1", nous pouvons montrer le résultat suivant.

Théorème 1.7 (Division euclidienne de polynômes)**.** *Pour tous polynômes $P, D \in \mathbb{K}[X]$ tels que $D \neq 0$, il existe des polynômes $Q, R \in \mathbb{K}[X]$ uniques tels que $P = QD + R$ et $\deg(R) < \deg(D)$.*

Avant de donner une preuve de ce théorème, considérons un exemple.

Exemple 1.8. Soient $P = 6X^5 + X^4 - X^3 + 2X - 1$ et $D = 2X^2 + X - 3$. On souhaite trouver des polynômes Q et R de $\mathbb{Q}[X]$ tels que $P = QD + R$ et $\deg(R) < 2$. Observons tout d'abord que si de tels polynômes existent, alors nécessairement $\deg(Q) = 3$. Écrivons donc $Q = aX^3 + bX^2 + cX + d$.
En identifiant les coefficients de X^5 dans les deux membres de l'égalité $P = QD + R$, on obtient l'équation $6 = 2a$. D'où $a = \frac{6}{2} = 3$. Nous pouvons donc écrire $Q = 3X^3 + Q'$, avec $Q' = bX^2 + cX + d$.
On a

$$P = QD + R = \left(3X^3 + Q'\right) D + R$$

et

$$P - 3X^3D = Q'D + R.$$

Posons $P' = P - 3X^3D$. Le polynôme P' ainsi obtenu est de degré strictement inférieur à celui de P. En effet,

$$\begin{aligned} P' &= (6X^5 + X^4 - X^3 + 2X - 1) - 3X^3(2X^2 + X - 3) \\ &= (6 - 6)\, X^5 + (1 - 3)\, X^4 + (-1 + 9)\, X^3 + 2X - 1 \\ &= -2X^4 + 8X^3 + 2X - 1. \end{aligned}$$

Nous nous sommes ramenés au problème initial où le polynôme P de degré 5 a été remplacé par un polynôme P' de degré 4 : nous devons maintenant trouver des polynômes Q' et R tels que $P' = Q'D + R$ et $\deg(R) < \deg(D)$. Cette observation nous incitera à utiliser le principe de la récurrence sur le degré du polynôme P dans la démonstration du théorème 1.7.

Poursuivons notre raisonnement pour obtenir les coefficients restants, c'est-à-dire b, c, d. En identifiant les coefficients de X^4 dans l'équation $P' = Q'D + R$, on obtient l'équation $-2 = 2b$. D'où $b = -1$. On peut donc écrire $Q' = -X^2 + Q''$ avec $Q'' = cX + d$.

En posant $P'' = P' + X^2 D$, on obtient $P'' = Q''D + R$ et $\deg(P'') < \deg(P')$. On calcule

$$\begin{aligned} P'' &= \left(-2X^4 + 8X^3 + 2X - 1\right) + X^2(2X^2 + X - 3) \\ &= (-2+2)\,X^4 + (8+1)\,X^3 + (0-3)\,X^2 + 2X - 1 \\ &= 9X^3 - 3X^2 + 2X - 1. \end{aligned}$$

Nous nous sommes à nouveau ramenés au problème initial avec maintenant un polynôme P'' de degré 3. En identifiant les coefficients de X^3 dans l'équation $P'' = Q''D + R$, on obtient l'équation $9 = 2c$. D'où $c = \frac{9}{2}$. On peut donc écrire $Q'' = \frac{9}{2}X + Q'''$ avec $Q''' = d$.

En posant $P''' = P'' - \frac{9}{2}XD$, on obtient $P''' = Q'''D + R$ et $\deg(P''') < \deg(P'')$. On calcule

$$\begin{aligned} P''' &= \left(9X^3 - 3X^2 + 2X - 1\right) - \frac{9}{2}X(2X^2 + X - 3) \\ &= \left(9 - \frac{9}{2}\cdot 2\right)X^3 + \left(-3 - \frac{9}{2}\right)X^2 + \left(2 - \frac{9}{2}\cdot(-3)\right)X - 1 \\ &= -\frac{15}{2}X^2 + \frac{31}{2}X - 1. \end{aligned}$$

Nous nous sommes à nouveau ramenés au problème initial avec maintenant un polynôme P''' de degré 2. En identifiant les coefficients de X^2 dans l'équation $P''' = Q'''D + R$, on obtient l'équation $-\frac{15}{2} = 2d$. D'où $d = -\frac{15}{4}$ et

$$\begin{aligned} R &= P''' + \frac{15}{4}D \\ &= \left(-\frac{15}{2}X^2 + \frac{31}{2}X - 1\right) + \frac{15}{4}\left(2X^2 + X - 3\right) \\ &= \left(-\frac{15}{2} + \frac{15}{4}\cdot 2\right)X^2 + \left(\frac{31}{2} + \frac{15}{4}\right)X + \left(-1 - \frac{45}{4}\right) \\ &= \frac{77}{4}X - \frac{49}{4}. \end{aligned}$$

Nous devons donc avoir

$$Q = 3X^3 - X^2 + \frac{9}{2}X - \frac{15}{4} \quad \text{et} \quad R = \frac{77}{4}X - \frac{49}{4}$$

et on peut vérifier qu'en effet $P = QD + R$ et que $\deg(R) < \deg(D)$.

Preuve du théorème 1.7. Soient $P, D \in \mathbb{K}[X]$ tels que $D \neq 0$. Nous montrons l'existence de tels polynômes Q et R par récurrence sur le degré de P. Si $\deg(P) < \deg(D)$, alors il suffit de choisir $Q = 0$ et $R = P$. Supposons maintenant que $\deg(P) \geq \deg(D)$ et que le résultat soit vérifié pour tout polynôme de degré strictement inférieur à celui de P. Si p et d sont les coefficients dominants de P et D respectivement, alors le polynôme[2] $P' = P - \frac{p}{d}X^{\deg(P)-\deg(D)}D$ a un degré strictement inférieur à celui de P. Par hypothèse de récurrence, il existe des polynômes Q' et R' de $\mathbb{K}[X]$ tels que $P' = Q'D + R'$ et $\deg(R') < \deg(D)$. On obtient donc que

$$P = \left(Q' + \frac{p}{d}X^{\deg(P)-\deg(D)}\right)D + R' \quad \text{et} \quad \deg(R') < \deg(D).$$

2. On convient que la notation $\frac{1}{d}$ désigne l'inverse modulaire de d dans le cas où $\mathbb{K} = \mathbb{Z}_m$ avec m premier. Souvenons-nous que tout élément non nul de $\mathbb{Z}_m$ est inversible dans $\mathbb{Z}_m$ lorsque m est premier.

D'où $Q = Q' + \frac{p}{d} X^{\deg(P)-\deg(D)}$ et $R = R'$ conviennent.
Montrons à présent l'unicité des polynômes Q et R. Supposons qu'il existe des polynômes $Q, R, Q', R' \in \mathbb{K}[X]$ tels que $P = QD + R = Q'D + R'$, $\deg(R) < \deg(D)$ et $\deg(R') < \deg(D)$. Alors $(Q - Q')D = R' - R$. Si $Q \neq Q'$, alors $\deg((Q - Q')D) = \deg(Q - Q') + \deg(D) \geq \deg(D)$. C'est impossible puisque $\deg(R' - R) \leq \max\{\deg(R), \deg(R')\} < \deg(D)$. D'où $Q = Q'$. Par conséquent, de l'égalité $QD + R = Q'D + R'$, on obtient que $R = R'$ aussi. $\square$

Comme pour la division euclidienne de nombres, dans l'écriture $P = QD + R$, les polynômes D, Q, R sont appelés les diviseur, quotient et reste de la division respectivement. La preuve ci-dessus (plus précisément, la partie "existence") nous procure en fait un algorithme pour obtenir le quotient et le reste.

Exemple 1.9. Continuons l'exemple 1.8. On peut organiser les calculs effectués dans un tableau de division à la manière de la division des entiers classiques :

$$
\begin{array}{rrrrrr|l}
6X^5 & +X^4 & -X^3 & & +2X & -1 & 2X^2 + X - 3 \\ \cline{7-7}
-6X^5 & -3X^4 & +9X^3 & & & & 3X^3 - X^2 + \frac{9}{2}X - \frac{15}{4} \\ \cline{1-6}
 & -2X^4 & +8X^3 & & +2X & -1 & \\
 & +2X^4 & +X^3 & -3X^2 & & & \\ \cline{2-6}
 & & 9X^3 & -3X^2 & +2X & -1 & \\
 & & -9X^3 & -\frac{9}{2}X^2 & +\frac{27}{2}X & & \\ \cline{3-6}
 & & & -\frac{15}{2}X^2 & +\frac{31}{2}X & -1 & \\
 & & & +\frac{15}{2}X^2 & +\frac{15}{4}X & -\frac{45}{4} & \\ \cline{4-6}
 & & & & \frac{77}{4}X & -\frac{49}{4} &
\end{array}
$$

Exemple 1.10. Effectuons la division euclidienne de $P = X^5 + iX^3 + (1+i)X - 1$ par $D = 2X^3 - 2X^2 + 1$ dans $\mathbb{C}[X]$:

$$
\begin{array}{rrrrrr|l}
X^5 & & +iX^3 & & +(1+i)X & -1 & 2X^3 - 2X^2 + 1 \\ \cline{7-7}
-X^5 & +X^4 & & -\frac{1}{2}X^2 & & & \frac{1}{2}X^2 + \frac{1}{2}X + \frac{1+i}{2} \\ \cline{1-6}
 & X^4 & +iX^3 & -\frac{1}{2}X^2 & +(1+i)X & -1 & \\
 & -X^4 & +X^3 & & -\frac{1}{2}X & & \\ \cline{2-6}
 & & (1+i)X^3 & -\frac{1}{2}X^2 & +(\frac{1}{2}+i)X & -1 & \\
 & & -(1+i)X^3 & +(1+i)X^2 & & -\frac{1+i}{2} & \\ \cline{3-6}
 & & & (\frac{1}{2}+i)X^2 & +(\frac{1}{2}+i)X & -\frac{3+i}{2} &
\end{array}
$$

On obtient $Q = \frac{1}{2}X^2 + \frac{1}{2}X + \frac{1+i}{2}$ et $R = (\frac{1}{2} + i)X^2 + (\frac{1}{2} + i)X - \frac{3+i}{2}$.
Effectuons maintenant la division euclidienne de $P = X^4 + X^3 + 2X + 1$ par $D = 2X^3 + X^2 + 1$ dans $\mathbb{Z}_3[X]$:

$$
\begin{array}{rrrrr|l}
X^4 & +X^3 & & +2X & +1 & 2X^3 + X^2 + 1 \\ \cline{6-6}
-X^4 & -2X^3 & & -2X & & 2X + 1 \\ \cline{1-5}
 & 2X^3 & & & +1 & \\
 & -2X^3 & -X^2 & & -1 & \\ \cline{2-5}
 & & 2X^2 & & &
\end{array}
$$

On obtient $Q = 2X + 1$ et $R = 2X^2$. Il est bon de vérifier que dans $\mathbb{Z}_3[X]$, on a bien

$$\begin{aligned} QD + R &= (2X+1)(2X^3+X^2+1)+2X^2 \\ &= X^4+2X^3+2X+2X^3+X^2+1+2X^2 \\ &= X^4+X^3+2X+1 \\ &= P. \end{aligned}$$

Un autre algorithme pour trouver Q et R dans le cas où $\deg(P) \geq \deg(D)$ (sinon, nous avons vu que la solution était toujours $Q = 0$ et $R = P$) est de remarquer que $\deg(Q) = \deg(P) - \deg(D)$. Dans notre cas, on cherche donc des polynômes Q et R de $\mathbb{Z}_3[X]$ de la forme $Q = aX + b$ et $R = cX^2 + dX + e$. En identifiant les coefficients des polynômes dans les deux membres de l'équation $P = QD + R$, on obtient

$$\begin{aligned} P = QD + R &\iff X^4+X^3+2X+1 = (aX+b)(2X^3+X^2+1)+cX^2+dX+e \\ &\iff \begin{cases} 1 = 2a \\ 1 = a+2b \\ 0 = b+c \\ 2 = a+d \\ 1 = b+e \end{cases}. \end{aligned}$$

En résolvant ce système dans $\mathbb{Z}_3$ (par la méthode de Gauss par exemple), on obtient l'unique solution $(a, b, c, d, e) = (2, 1, 2, 0, 0)$, c'est-à-dire $Q = 2X + 1$ et $R = 2X^2$.

Définition 1.11. Soient $P, D \in \mathbb{K}[X]$ avec $D \neq 0$. Les polynômes (uniques) Q, R donnés dans l'énoncé du théorème 1.7 sont appelés respectivement le *quotient* et le *reste* de la division euclidienne de P par D.

Définition 1.12. Dans $\mathbb{K}[X]$, on dit qu'un polynôme D *divise* un polynôme P lorsque le reste R de la division euclidienne de P par D vaut 0.

Autrement dit, D divise P s'il existe un polynôme Q (nécessairement unique au vu du théorème 1.7) tel que $P = QD$.

Pour continuer notre analogie entre l'ensemble des polynômes $\mathbb{K}[X]$ et l'ensemble des entiers $\mathbb{Z}$, nous donnons ici la définition du PGCD de deux polynômes.

Définition 1.13. Un PGCD de deux polynômes P et Q de $\mathbb{K}[X]$ est un polynôme D qui divise P et Q et qui est tel que tout polynôme divisant simultanément P et Q divise aussi D.

Remarquons que la notion de PGCD de polynômes n'est définie qu'à une constante multiplicative non nulle près : si D est un PGCD de P et Q, alors pour tout $k \in \mathbb{K} \setminus \{0\}$, le polynôme kD est aussi un PGCD de P et Q. Parmi tous les PGCD, on privilégie parfois celui ayant 1 comme coefficient dominant : on parle alors *du* PGCD de deux polynômes.

Maintenant que nous disposons de la division euclidienne dans $\mathbb{K}[X]$, l'algorithme d'Euclide pour la recherche du PGCD de deux polynômes s'effectue de façon similaire à celui pour obtenir le PGCD de deux entiers.

Exemple 1.14. Plaçons-nous dans $\mathbb{Z}_3[X]$ et calculons le PGCD de $P = X^5 + 2X$ et de $Q = X^4+2X^3+2X$. Comme $\deg(Q) < \deg(P)$, on pose $R_0 = Q$ et on obtient successivement

$$\begin{aligned} P &= (X+1)R_0 + X^3 + X^2, & R_1 &= X^3 + X^2, \\ R_0 &= (X+1)R_1 + 2X^2 + 2X, & R_2 &= 2X^2 + 2X, \\ R_1 &= (2X)R_2, & R_3 &= 0. \end{aligned}$$

Le PGCD obtenu est le dernier reste non nul, c'est-à-dire $2X^2 + 2X$. Remarquons que le PGCD étant défini à une constante multiplicative près, le polynôme $2(2X^2+2X) = X^2+X$ est aussi un PGCD de P et Q.

En remontant l'algorithme d'Euclide, on obtient le théorème de Bézout pour les polynômes.

Théorème 1.15 (Bézout). *Soient $P, Q \in \mathbb{K}[X]$ et $D \in \mathbb{K}[X]$ un PGCD de P et Q. Alors il existe $A, B \in \mathbb{K}[X]$ tels que*

$$AP + BQ = D.$$

Exemple 1.16. Nous continuons l'exemple 1.14 pour obtenir des coefficients de Bézout A et B. On calcule

$$\begin{array}{rcl} 2X^2 + 2X & = & R_0 - (X+1)R_1 \\ & = & Q - (X+1)(P - (X+1)Q) \\ & = & -(X+1)P + (1 + (X+1)^2)Q \\ & = & (2X+2)P + (X^2 + 2X + 2)Q. \end{array}$$

On a donc obtenu que les polynômes $A = 2X + 2$ et $B = X^2 + 2X + 2$ vérifiaient l'égalité de Bézout $AP + BQ = D$ pour le PGCD $D = 2X^2 + 2X$. Pour obtenir des coefficients de Bézout pour le PGCD $X^2 + X$, on multiplie cette égalité par 2 (inverse de 2 dans $\mathbb{Z}_3$) :

$$X^2 + X = (X+1)P + (2X^2 + X + 1)Q.$$

Définition 1.17. Deux polynômes P et Q de $\mathbb{K}[X]$ sont *premiers entre eux* si 1 est un PGCD de P et Q.

Théorème 1.18 (Gauss). *Si $P, Q, D \in \mathbb{K}[X]$ sont tels que D divise PQ et que D et P sont premiers entre eux, alors D divise Q.*

Démonstration. Soient $P, Q, D \in \mathbb{K}[X]$ tels que D divise PQ et que D et P sont premiers entre eux. Alors d'une part, il existe $S \in \mathbb{K}[X]$ tel que $PQ = SD$ et d'autre part, par le théorème de Bézout, il existe $A, B \in \mathbb{K}[X]$ tels que $AD + BP = 1$. On obtient que

$$Q = ADQ + BPQ = ADQ + BSD = (AQ + BS)D,$$

ce qui montre que D divise Q. □

1.3 Polynôme et fonction polynomiale

Un polynôme à coefficients dans $\mathbb{K}$ induit naturellement une fonction de $\mathbb{K}$ dans $\mathbb{K}$.

Définition 1.19. Pour $P \in \mathbb{K}[X]$ et $k \in \mathbb{K}$, la notation $P(k)$ désigne l'élément de $\mathbb{K}$ obtenu en substituant dans P l'indéterminée X par k et en exécutant les opérations $+$ et $\cdot$ dans $\mathbb{K}$: si

$$P = p_0 + p_1X + p_2X^2 + \cdots + p_dX^d,$$

alors

$$P(k) = p_0 + p_1k + p_2k^2 + \cdots + p_dk^d.$$

On dit qu'on a *évalué* P en k. La *fonction induite* par P est la fonction

$$\mathbb{K} \to \mathbb{K},\ k \mapsto P(k).$$

Une *fonction polynomiale* de $\mathbb{K}$ est une fonction induite par un polynôme de $\mathbb{K}[X]$.

Remarque 1.20. Deux polynômes différents peuvent donner lieu à une même fonction induite. Par exemple, pour tout $k \in \mathbb{Z}_2$, on a $k^2 = k$. Les polynômes X^2 et X de $\mathbb{Z}_2[X]$ donnent donc lieu à la même fonction polynomiale induite [3]. Voici un autre exemple : les polynômes $X^3 + 2X$ et 0 de $\mathbb{Z}_3[X]$ donnent lieu à la même fonction polynomiale induite puisque pour tout $k \in \mathbb{Z}_3$, on a $k^3 + 2k = 0$. Nous verrons plus loin, grâce au théorème fondamental de l'algèbre, que ceci n'est pas possible lorsque $\mathbb{K}$ est $\mathbb{C}, \mathbb{R}$ ou $\mathbb{Q}$.

1.4 Dérivation des polynômes

On définit ensuite la notion de dérivée (formelle) d'un polynôme.

Définition 1.21. L'opérateur de dérivation

$$\mathrm{D}\colon\ \mathbb{K}[X] \to \mathbb{K}[X]$$

est défini comme suit. Pour tout polynôme $P = p_0 + p_1X + p_2X^2 + \cdots + p_dX^d$ de $\mathbb{K}[X]$, le *polynôme dérivé* [4] $\mathrm{D}\, P$ est le polynôme

$$\mathrm{D}\, P = p_1 + 2p_2X + \cdots + dp_dX^{d-1}.$$

Autrement dit, pour tout $n \in \mathbb{N}$, le coefficient de X^n dans $\mathrm{D}\, P$ est égal à $(n+1)p_{n+1}$.

Exercice 1.22. La fonction polynomiale induite par la dérivée formelle d'un polynôme P de $\mathbb{R}[X]$ correspond à la dérivée de la fonction polynomiale induite par P.

Proposition 1.23. *L'opérateur de dérivation* $\mathrm{D}\colon\ \mathbb{K}[X] \to \mathbb{K}[X]$ *est linéaire : pour tous* $k, \ell \in \mathbb{K}$ *et tous* $P, Q \in \mathbb{K}[X]$*, on a* $\mathrm{D}(kP + \ell Q) = k\,\mathrm{D}\,P + \ell\,\mathrm{D}\,Q$.

Démonstration. C'est une simple vérification. □

On utilise la notation $\mathrm{D}^i\, P$ pour signifier qu'on applique i fois l'opérateur D à P :

$$\mathrm{D}^0\, P = P,\ \mathrm{D}^1\, P = \mathrm{D}\, P,\ \mathrm{D}^2\, P = \mathrm{D}(\mathrm{D}\, P),\ \mathrm{D}^3\, P = \mathrm{D}(\mathrm{D}(\mathrm{D}\, P))),\ \ldots.$$

Proposition 1.24. *Supposons que* $\mathbb{K}$ *est* $\mathbb{C}$*,* $\mathbb{R}$ *ou* $\mathbb{Q}$*. Alors pour tous* $d, i \in \mathbb{N}$*, on a*

$$\mathrm{D}^i\left(\frac{1}{d!}X^d\right) = \begin{cases} \frac{1}{(d-i)!}X^{d-i} & si\ d \geq i, \\ 0 & si\ d < i. \end{cases}$$

Démonstration. La preuve s'obtient par récurrence sur i. □

Exemple 1.25. Pour $d = 7$ et $i = 3$, on calcule

$$\mathrm{D}^3\left(\frac{1}{7!}X^7\right) = \mathrm{D}^2\left(\frac{7}{7!}X^6\right) = \mathrm{D}\left(\frac{7\cdot 6}{7!}X^5\right) = \frac{7\cdot 6\cdot 5}{7!}X^4 = \frac{1}{4!}X^4$$

et on a bien $4 = 7 - 3 = d - i$.

Proposition 1.26 (Binôme de Newton)**.** *Pour tout* $P, Q \in \mathbb{K}[X]$ *et tout* $n \in \mathbb{N}$*, on a*

$$(P+Q)^n = \sum_{i=0}^{n} \mathrm{C}_n^i\, P^iQ^{n-i}.$$

3. Or, ces polynômes sont distincts par définition.

4. On écrit parfois aussi $\mathrm{D}_X\, P$ à la place de $\mathrm{D}\, P$. C'est utile lorsque l'on étudie les polynômes à plusieurs indéterminées $X, Y, \ldots$ Dans le cadre de ce cours, nous nous limiterons au cas d'une seule indéterminée X.

Démonstration. Laissée en exercice. C'est une adaptation directe de la preuve du binôme de Newton pour les complexes que vous avez vue au bloc 1 dans le cours de mathématiques générales. □

Proposition 1.27 (Formule de Leibniz)**.** *Pour tous $P, Q \in \mathbb{K}[X]$ et tout $n \in \mathbb{N}$, on a*

$$\mathrm{D}^n(PQ) = \sum_{i=0}^{n} \mathrm{C}_n^i \, \mathrm{D}^i P \, \mathrm{D}^{n-i} Q.$$

Démonstration. Laissée en exercice. C'est une adaptation directe de la preuve de la formule de Leibniz pour les fonctions d'une variable réelle que vous avez vue au bloc 1 dans le cours de mathématiques générales. □

Proposition 1.28 (Formule de Taylor)**.** *Supposons que $\mathbb{K}$ est $\mathbb{C}$, $\mathbb{R}$ ou $\mathbb{Q}$. Alors pour tout $P \in \mathbb{K}[X]$ et tout $k \in \mathbb{K}$, on a*

$$P = \sum_{i=0}^{\deg(P)} \frac{\mathrm{D}^i P(k)}{i!}(X-k)^i.$$

Démonstration. Soit $P = \sum_{n=0}^{d} p_n X^n$ (où $d = \deg(P)$) et soit $k \in \mathbb{K}$. En utilisant la formule du binôme de Newton et en échangeant ensuite les signes sommatoires, on obtient

$$\begin{aligned} P &= \sum_{n=0}^{d} p_n (X-k+k)^n \\ &= \sum_{n=0}^{d} p_n \sum_{i=0}^{n} \mathrm{C}_n^i (X-k)^i k^{n-i} \\ &= \sum_{i=0}^{d} \left(\sum_{n=i}^{d} p_n \, \mathrm{C}_n^i \, k^{n-i} \right) (X-k)^i. \end{aligned}$$

Pour conclure, il suffit donc de montrer que

$$\sum_{n=i}^{d} p_n \, \mathrm{C}_n^i \, k^{n-i} = \frac{\mathrm{D}^i P(k)}{i!}.$$

Ceci est vrai puisqu'en utilisant les propositions 1.23 et 1.24, on obtient que pour tout $i \in \{0, \ldots, d\}$,

$$\begin{aligned} \frac{1}{i!} \mathrm{D}^i P &= \frac{1}{i!} \mathrm{D}^i \left(\sum_{n=0}^{d} p_n X^n \right) \\ &= \frac{1}{i!} \sum_{n=0}^{d} p_n \, \mathrm{D}^i (X^n) \\ &= \sum_{n=i}^{d} p_n \frac{n!}{i!(n-i)!} X^{n-i} \\ &= \sum_{n=i}^{d} p_n \, \mathrm{C}_n^i \, X^{n-i}. \end{aligned}$$

□

Exemple 1.29. Soient $P = 2X^2 + 3X - 1$ dans $\mathbb{C}[X]$ et $k = 2$. On a $\mathrm{D}\,P = 4X + 3$ et $\mathrm{D}^2\,P = 4$. Ainsi $P(2) = 8 + 6 - 1 = 13$, $\mathrm{D}\,P(2) = 11$ et $\mathrm{D}^2\,P(2) = 4$. On obtient

$$\begin{aligned}\sum_{i=0}^{2} \frac{\mathrm{D}^i\,P(2)}{i!}(X-2)^i &= \frac{\mathrm{D}^2\,P(2)}{2!}(X-2)^2 + \mathrm{D}\,P(2)(X-2) + P(2)\\ &= \frac{4}{2}(X^2 - 4X + 4) + 11(X-2) + 13\\ &= 2X^2 + 3X - 1\\ &= P.\end{aligned}$$

1.5 Zéros d'un polynôme

Définition 1.30. Soit $P \in \mathbb{K}[X]$. Un *zéro*[5] de P est un nombre $k \in \mathbb{K}$ tel que le polynôme $X - k$ divise P. Si P est non nul, la *multiplicité* d'un nombre k en tant que zéro du polynôme P est le plus grand entier α tel que $(X-k)^\alpha$ divise P.

Avec cette définition, le fait que k soit un zéro de multiplicité 0 de P équivaut au fait que k ne soit pas un zéro de P. On pourrait étendre la définition de la multiplicité aux polynômes nuls en disant qu'elle vaut $+\infty$ pour tout $k \in \mathbb{K}$.

Exemple 1.31. Considérons le polynôme $P = X^3 - 2X^2 - 2X - 3$ de $\mathbb{Q}[X]$. Il admet 3 comme zéro simple, c'est-à-dire de multiplicité 1. En effet, $X - 3$ divise P mais $(X-3)^2$ ne divise pas P. On a $P = (X^2 + X + 1)(X - 3)$. En effectuant la division euclidienne de P par $(X-3)^2$, on obtient $P = (X+4)(X-3)^2 + 13X - 39$.

La proposition suivante montre que la condition pour être un zéro d'un polynôme P donnée par cette définition est équivalente à la condition d'annulation de la fonction polynomiale induite par P.

Proposition 1.32. *Soient $P \in \mathbb{K}[X]$ et $k \in \mathbb{K}$. Le nombre k est un zéro de P si et seulement si $P(k) = 0$.*

Démonstration. Soit $P = Q(X-k) + R$ avec $Q, R \in \mathbb{K}[X]$ et $\deg(R) < 1$. Alors R est un polynôme constant $c \in \mathbb{K}$. En évaluant les deux membres de cette égalité en k, on obtient $P(k) = c$. Mais par définition, k est un zéro de P si et seulement si $R = 0$. D'où la conclusion. $\square$

Théorème 1.33. *Supposons que $\mathbb{K}$ est égal à $\mathbb{C}$, $\mathbb{R}$ ou $\mathbb{Q}$. Soient $P \in \mathbb{K}[X]$ non nul, $k \in \mathbb{K}$ et $\alpha \in \mathbb{N}$. Le nombre k est un zéro de multiplicité α de P si et seulement si $P(k) = 0$, $\mathrm{D}\,P(k) = 0$, …, $\mathrm{D}^{\alpha-1}\,P(k) = 0$ et $\mathrm{D}^\alpha\,P(k) \neq 0$.*

Démonstration. Montrons la condition nécessaire. Supposons que k soit un zéro de multiplicité α de P. Alors il existe $Q \in \mathbb{K}[X]$ tel que $P = Q(X-k)^\alpha$. Comme $X - k$ ne divise pas Q (sinon $(X-k)^{\alpha+1}$ diviserait P), on obtient que $Q(k) \neq 0$ par la proposition 1.32. En utilisant la formule de Leibniz, on obtient pour $n \in \{0, \ldots, \alpha\}$ que

$$\mathrm{D}^n\,P = \mathrm{D}^n\big(Q(X-k)^\alpha\big) = \sum_{i=0}^{n} \mathrm{C}_n^i\ \mathrm{D}^{n-i}\,Q\ \mathrm{D}^i\big((X-k)^\alpha\big).$$

Comme $\mathrm{D}^i(X-k)^\alpha(k) = 0$ pour $i < \alpha$, on obtient que $(\mathrm{D}^n\,P)(k) = 0$ pour $n < \alpha$. Pour $n = \alpha$, on obtient

$$\mathrm{D}^\alpha\,P(k) = \mathrm{C}_\alpha^\alpha\ Q(k)\ \mathrm{D}^\alpha(X-k)^\alpha(k) = Q(k)\alpha! \neq 0.$$

5. On trouve aussi le terme “racine” d'un polynôme.

Montrons à présent la condition suffisante. Supposons que $P(k) = 0$, $\mathrm{D}\, P(k) = 0, \ldots$, $\mathrm{D}^{\alpha-1}\, P(k) = 0$ et $\mathrm{D}^{\alpha}\, P(k) \neq 0$. En utilisant la formule de Taylor, on obtient

$$\begin{aligned} P &= \sum_{i=0}^{\deg(P)} \frac{\mathrm{D}^i\, P(k)}{i!}(X-k)^i \\ &= \sum_{i=\alpha}^{\deg(P)} \frac{\mathrm{D}^i\, P(k)}{i!}(X-k)^i \\ &= (X-k)^{\alpha} \sum_{i=\alpha}^{\deg(P)} \frac{\mathrm{D}^i\, P(k)}{i!}(X-k)^{i-\alpha}. \end{aligned}$$

En posant

$$Q = \sum_{i=\alpha}^{\deg(P)} \frac{\mathrm{D}^i\, P(k)}{i!}(X-k)^{i-\alpha},$$

cette égalité se réécrit $P = (X-k)^{\alpha} Q$. Remarquons que $(X-k)^{\alpha+1}$ divise P si et seulement si $X-k$ divise Q, c'est-à-dire si et seulement si $Q(k) = 0$ au vu de la proposition 1.32. Comme

$$Q(k) = \frac{\mathrm{D}^{\alpha}\, P(k)}{\alpha!} \neq 0,$$

on en conclut que $(X-k)^{\alpha+1}$ ne divise pas P, et donc que k est un zéro de multiplicité α de P. □

Exemple 1.34. Considérons le polynôme $P = X^3 - 2X^2 - 2X - 3$ de $\mathbb{Q}[X]$. Nous avons vu qu'il admet 3 comme zéro simple. On a bien que $P(3) = 3^3 - 2\cdot 3^2 - 2\cdot 3 - 3 = 27 - 18 - 6 - 3 = 0$. On a aussi $\mathrm{D}\, P = 3X^2 - 4X - 2$ et $\mathrm{D}\, P(3) = 3 \cdot 3^2 - 4 \cdot 3 - 2 = 27 - 12 - 2 = 13 \neq 0$.

1.6 Polynômes à coefficients complexes et théorème fondamental de l'algèbre

Dans cette section, nous donnons le théorème fondamental de l'algèbre sans preuve, et nous en tirons quelques conséquences.

Théorème 1.35 (Théorème fondamental de l'algèbre)**.** *Tout polynôme de $\mathbb{C}[X]$ non nul de degré d possède exactement d zéros (complexes) lorsque ceux-ci sont comptés avec leurs multiplicités. Ainsi, si $k_1, \ldots, k_m$ sont les zéros de P de multiplicités respectives $\alpha_1, \ldots, \alpha_m$, alors*

$$\alpha_1 + \cdots + \alpha_m = d$$

et

$$P = p(X-k_1)^{\alpha_1} \cdots (X-k_m)^{\alpha_m},$$

où p est le coefficient dominant de P.

Corollaire 1.36. *Supposons que $\mathbb{K}$ est $\mathbb{C}$, $\mathbb{R}$ ou $\mathbb{Q}$.*

1. *Deux polynômes de $\mathbb{K}[X]$ ayant les mêmes zéros complexes avec les mêmes multiplicités sont égaux à une constante non nulle multiplicative près.*
2. *Deux polynômes de $\mathbb{K}[X]$ de même degré d prenant les mêmes valeurs en $d+1$ arguments sont égaux.*
3. *Deux polynômes de $\mathbb{K}[X]$ sont égaux si et seulement si leurs fonctions polynomiales induites (de $\mathbb{K}$ dans $\mathbb{K}$) sont égales.*

Démonstration. Montrons le premier point. Si les deux polynômes sont nuls, ils sont trivialement égaux. Si l'un est nul et l'autre pas, alors ils ne peuvent avoir les mêmes zéros complexes et il n'y a rien à montrer. Enfin, si les deux polynômes sont non-nuls, cela découle de la factorisation donnée par le théorème fondamental de l'algèbre.

Il suffit de montrer les propriétés 2 et 3 pour des polynômes de $\mathbb{C}[X]$, les autres cas étant des cas particuliers.

Montrons le deuxième point. Si $P, Q \in \mathbb{K}[X]$ ont même degré d et sont égaux en $d+1$ valeurs, alors le polynôme $P - Q$ est de degré au plus d et s'annule en $d+1$ valeurs. Au vu de la proposition 1.32, $P-Q$ possède $d+1$ zéros. Puisque $\mathbb{K} \subseteq \mathbb{C}$, on obtient par le théorème fondamental de l'algèbre que le polynôme $P - Q$ est nécessairement nul. On obtient donc que $P = Q$.

La condition nécessaire du troisième point est immédiate (et est vraie pour n'importe quel $\mathbb{K}$ en général). Montrons la condition suffisante. Si $P, Q \in \mathbb{K}[X]$ sont tels que leurs fonctions polynomiales induites sont égales, alors tout élément de $\mathbb{K}$ est zéro de $P - Q$ (en fait la fonction induite par $P-Q$ est la fonction nulle). Comme dans le point précédent, le théorème fondamental de l'algèbre impose $P = Q$. □

Le troisième point du corollaire précédent montre l'équivalence des notions de polynôme et de fonction polynomiale dans les cas où $\mathbb{K}$ est $\mathbb{C}$, $\mathbb{R}$ ou $\mathbb{Q}$. Cela permet d'utiliser la méthode d'*identification des coefficients* pour la résolutions d'équations polynomiales.

Remarque 1.37. Le théorème fondamental de l'algèbre ne fournit pas d'algorithme pour trouver les zéros d'un polynôme ! Cependant, il existe des méthodes pour obtenir les zéros d'un polynôme de degré au plus 4. Pour un polynôme de degré supérieur à 5, aucune méthode générale n'existe. Néanmoins, plusieurs méthodes relevant de l'analyse numérique permettent de fournir des approximations aussi fines que souhaitées des zéros d'un polynôme d'un degré quelconque. C'est l'un des objets du cours "Introduction à l'algorithmique numérique".

1.7 Polynômes à coefficients réels

Dans cette section, nous étudions un peu plus précisément l'ensemble des zéros complexes d'un polynôme à coefficients réels. En effet, l'ensemble $\mathbb{R}[X]$ peut-être vu comme un sous-ensemble de $\mathbb{C}[X]$, et on peut donc étudier les spécificités des polynômes à coefficients réels parmi les polynômes de $\mathbb{C}[X]$.

Le résultat suivant nous dit que l'ensemble des zéros complexes d'un polynôme à coefficients réels est stable par conjugaison.

Lemme 1.38. *Soit $P \in \mathbb{R}[X]$. Si $c \in \mathbb{C}$ est un zéro de P, alors son conjugué $\bar{c}$ est aussi un zéro de P.*

Démonstration. Notons $P = p_0 + p_1X + \cdots + p_dX^d$. Soit $c \in \mathbb{C}$ un zéro de P. Alors

$$P(c) = p_0 + p_1c + \cdots + p_dc^d = 0.$$

Puisque les coefficients p_i de P sont réels, on obtient

$$P(\bar{c}) = p_0 + p_1\bar{c} + \cdots + p_d\bar{c}^d = \overline{P(c)} = \bar{0} = 0.$$

Ceci montre que $\bar{c}$ est un zéro de P. □

Lemme 1.39. *Pour tout $c \in \mathbb{C}$, on a $(X - c)(X - \bar{c}) \in \mathbb{R}[X]$.*

Démonstration. Soit $c \in \mathbb{C}$. On a $c = a + ib$ avec $a, b \in \mathbb{R}$. En distribuant le produit, on obtient que

$$(X - c)(X - \bar{c}) = X^2 - (c + \bar{c})X + c\bar{c} = X^2 - 2aX + (a^2 + b^2) \in \mathbb{R}[X].$$

□

Exemple 1.40. Soit $c = 3 + i$. On a

$$\begin{aligned}(X - c)(X - \bar{c}) &= (X - 3 - i)(X - 3 + i) \\ &= X^2 - (3 - i + 3 + i)X + (3 + i)(3 - i) \\ &= X^2 - 6X + 10.\end{aligned}$$

On a bien $c + \bar{c} = 3 + i + 3 - i = 6$ et $c\bar{c} = (3 + i)(3 - i) = 10$.

Proposition 1.41. *Tout polynôme P de $\mathbb{R}[X]$ se factorise dans $\mathbb{R}[X]$ sous la forme*

$$P = p(X - r_1) \cdots (X - r_m)Q_1 \cdots Q_n,$$

où p est le coefficient dominant de P, $r_1, \ldots, r_m$ sont les zéros réels de P et $Q_1, \ldots, Q_n \in \mathbb{R}[X]$ sont des polynômes du second degré avec un réalisant $\Delta < 0$.

Démonstration. Soit $P \in \mathbb{R}[X]$. Par le lemme 1.38, les zéros de P sont de la forme

$$r_1, \ldots, r_k, c_1, \ldots, c_\ell, \overline{c}_1, \ldots, \overline{c_\ell},$$

avec $r_1, \ldots, r_k \in \mathbb{R}$ et $c_1, \ldots, c_\ell \in \mathbb{C} \setminus \mathbb{R}$. Par le théorème fondamental de l'algèbre, on obtient

$$\begin{aligned}P &= p(X - r_1) \cdots (X - r_k)(X - c_1)(X - \overline{c}_1) \cdots (X - c_\ell)(X - \overline{c}_\ell) \\ &= p(X - r_1) \cdots (X - r_k)Q_1 \cdots Q_\ell,\end{aligned}$$

où p est le coefficient dominant de P et où on a posé $Q_i = (X - c_i)(X - \overline{c_i})$ pour tout $i \in \{1, \ldots, \ell\}$. Par le lemme 1.39, nous savons que les polynômes Q_i sont à coefficients réels. De plus, ces polynômes ont un réalisant $\Delta < 0$ puisqu'ils ne possèdent pas de zéro réel. □

Chapitre 2

Espaces vectoriels

2.1 Espace vectoriel

Définition 2.1. Un *espace vectoriel* sur $\mathbb{K}$ est la donnée d'un ensemble E et de deux opérations

$$+ \colon E \times E \to E$$

et

$$\cdot \colon \mathbb{K} \times E \to E$$

qui satisfont les propriétés suivantes.

1. Il existe un élément $\mathbf{0}$ de E neutre pour $+$: pour tout $\mathbf{e} \in E$, $\mathbf{e} + \mathbf{0} = \mathbf{0} + \mathbf{e} = \mathbf{e}$.
2. Tout élément de E possède un opposé : pour tout $\mathbf{e} \in E$, il existe $\mathbf{f} \in E$ tel que $\mathbf{e} + \mathbf{f} = \mathbf{f} + \mathbf{e} = \mathbf{0}$.
3. L'opération $+$ est associative : pour tous $\mathbf{e}, \mathbf{f}, \mathbf{g} \in E$, on a $(\mathbf{e} + \mathbf{f}) + \mathbf{g} = \mathbf{e} + (\mathbf{f} + \mathbf{g})$.
4. L'opération $+$ est commutative : pour tous $\mathbf{e}, \mathbf{f} \in E$, on a $\mathbf{e} + \mathbf{f} = \mathbf{f} + \mathbf{e}$.
5. Pour tous $\mathbf{e}, \mathbf{f} \in E$ et tous $k, \ell \in \mathbb{K}$, on a
   (a) $k \cdot (\ell \cdot \mathbf{e}) = (k\ell) \cdot \mathbf{e}$,
   (b) $(k + \ell) \cdot \mathbf{e} = k \cdot \mathbf{e} + \ell \cdot \mathbf{e}$,
   (c) $k \cdot (\mathbf{e} + \mathbf{f}) = k \cdot \mathbf{e} + k \cdot \mathbf{f}$,
   (d) $1 \cdot \mathbf{e} = \mathbf{e}$.

Si E est un espace vectoriel sur $\mathbb{K}$, alors les éléments de E sont appelés des *vecteurs* et les éléments de $\mathbb{K}$ sont appelés des *scalaires*.

On remarque facilement que l'élément neutre pour $+$ est unique, et donc il n'y a pas d'ambigüité à le noter $\mathbf{0}$. Attention cependant à ne pas confondre $\mathbf{0}$ avec l'élément neutre 0 de $\mathbb{K}$. On remarque aussi facilement que l'opposé d'un élément est unique, et comme d'habitude nous noterons $-\mathbf{e}$ l'opposé de $\mathbf{e}$ ainsi que $\mathbf{e} + (-\mathbf{f}) = \mathbf{e} - \mathbf{f}$. Dans la suite, nous noterons généralement $k\mathbf{e}$ à la place de $k \cdot \mathbf{e}$.

Exemple 2.2. On vérifie directement que les exemples suivants satisfont la définition d'espace vectoriel.

1. L'ensemble $\mathbb{K}^{\ell \times c}$ des matrices à coefficients dans $\mathbb{K}$ est un espace vectoriel sur $\mathbb{K}$.
2. En particulier, l'ensemble $\mathbb{K}^m$ des matrices-colonnes est un espace vectoriel sur $\mathbb{K}$.
3. L'ensemble $\mathbb{C}^{\ell \times c}$ des matrices à coefficients dans $\mathbb{C}$ est un espace vectoriel sur $\mathbb{R}$. C'est aussi un espace vectoriel sur $\mathbb{Q}$. En particulier, l'ensemble $\mathbb{C}$ des nombres complexes peut être vu comme un espace vectoriel sur $\mathbb{R}$ ou comme un espace vectoriel sur $\mathbb{Q}$. De même, l'ensemble $\mathbb{R}$ des nombres réels peut être vu comme un espace vectoriel sur $\mathbb{Q}$.

4. L'ensemble des solutions d'un système linéaire homogène est un espace vectoriel. Plus précisément, si $A \in \mathbb{K}^{\ell \times c}$, alors l'ensemble

$$E = \{\mathbf{x} \in \mathbb{K}^c \colon A\mathbf{x} = 0\}$$

muni des opérations habituelles d'addition et de multiplication par un scalaire

$$+\colon \mathbb{K}^c \times \mathbb{K}^c \to \mathbb{K}^c \quad \text{et} \quad \cdot\colon \mathbb{K} \times \mathbb{K}^c \to \mathbb{K}^c$$

restreintes à E est un espace vectoriel sur $\mathbb{K}$.

5. Dans $\mathbb{R}^2$, toute droite passant par l'origine est un espace vectoriel. En effet, une droite de $\mathbb{R}^2$ passant par $(0,0)$ est l'ensemble des solutions (x,y) d'une équation de la forme

$$ax + by = 0$$

où a et b sont des réels non simultanément nuls.

6. Dans $\mathbb{R}^3$, tout plan (resp. toute droite) passant par l'origine est un espace vectoriel. En effet, un plan de $\mathbb{R}^3$ passant par $(0,0,0)$ est l'ensemble des solutions (x,y,z) d'une équation de la forme

$$ax + by + cz = 0$$

où a,b,c sont des réels non simultanément nuls. Une droite de $\mathbb{R}^3$ passant par $(0,0,0)$ est l'ensemble des solutions (x,y,z) d'un système d'équations de la forme

$$\begin{cases} ax + \ by + \ cz = 0 \\ a'x + b'y + c'z = 0 \end{cases}$$

où les triplets de réels (a,b,c) et (a',b',c') ne sont pas multiples l'un de l'autre.

7. L'ensemble $\mathbb{K}[X]$ muni des opérations somme et multiplication par un scalaire (voir chapitre 1) est un espace vectoriel sur $\mathbb{K}$.

8. Pour tout $d \in \mathbb{N} \cup \{-\infty\}$, l'ensemble $\mathbb{K}_d[X]$ des polynômes à coefficients dans $\mathbb{K}$ de degré au plus d est un espace vectoriel sur $\mathbb{K}$.

9. Si A est un ensemble, alors l'ensemble $\mathcal{F}(A,\mathbb{K})$ des fonctions $f\colon A \to \mathbb{K}$ muni des opérations $+$ et $\cdot$ définies par

$$f + g\colon A \to \mathbb{K},\ a \mapsto f(a) + g(a) \quad \text{et} \quad k \cdot f\colon A \to \mathbb{K},\ a \mapsto kf(a)$$

(pour $f,g \in \mathcal{F}(A,\mathbb{K})$ et $k \in \mathbb{K}$) est un espace vectoriel sur $\mathbb{K}$.

10. L'ensemble $\mathrm{C}_0(\mathbb{R})$ des fonctions réelles continues (muni des mêmes opérations qu'au point 8 adaptées à ce cadre, c'est-à-dire avec $A = \mathbb{R}$ et $\mathbb{K} = \mathbb{R}$) est un espace vectoriel sur $\mathbb{K}$.

Listons quelques propriétés utiles des espaces vectoriels.

Proposition 2.3. *Soit E un espace vectoriel sur $\mathbb{K}$. Pour tous $\mathbf{e},\mathbf{f},\mathbf{g} \in E$ et tout $k \in \mathbb{K}$, on a*

1. $\mathbf{e} + \mathbf{f} = \mathbf{e} + \mathbf{g} \implies \mathbf{f} = \mathbf{g}$
2. $0 \cdot \mathbf{e} = \mathbf{0}$
3. $(-k) \cdot \mathbf{e} = -(k \cdot \mathbf{e})$
4. $k \cdot \mathbf{0} = \mathbf{0}$
5. $k \cdot \mathbf{e} = \mathbf{0} \implies (k = 0$ *ou* $\mathbf{e} = \mathbf{0})$.

Démonstration. Soient $\mathbf{e}, \mathbf{f}, \mathbf{g} \in E$ et $k \in \mathbb{K}$. Montrons le point 1. Si $\mathbf{e}+\mathbf{f} = \mathbf{e}+\mathbf{g}$, alors $\mathbf{f} = \mathbf{0}+\mathbf{f} = (-\mathbf{e}+\mathbf{e})+\mathbf{f} = -\mathbf{e}+(\mathbf{e}+\mathbf{f}) = -\mathbf{e}+(\mathbf{e}+\mathbf{g}) = (-\mathbf{e}+\mathbf{e})+\mathbf{g} = \mathbf{0}+\mathbf{g} = \mathbf{g}$. Montrons le point 2. On a $0\cdot\mathbf{e} = 0\cdot\mathbf{e}+\mathbf{0} = 0\cdot\mathbf{e}+(\mathbf{e}-\mathbf{e}) = (0\cdot\mathbf{e}+\mathbf{e})-\mathbf{e} = (0\cdot\mathbf{e}+1\cdot\mathbf{e})-\mathbf{e} = (0+1)\cdot\mathbf{e}-\mathbf{e} = 1\cdot\mathbf{e}-\mathbf{e} = \mathbf{e}-\mathbf{e} = \mathbf{0}$. Montrons le point 3. On a $(-k)\cdot\mathbf{e}+(k\cdot\mathbf{e}) = (-k+k)\cdot\mathbf{e} = 0\cdot\mathbf{e} = \mathbf{0}$ où on a utilisé le point 2 pour la dernière égalité. Montrons le point 4. On a $k\cdot\mathbf{0} = k\cdot(\mathbf{e}-\mathbf{e}) = k\cdot\mathbf{e}+k\cdot(-\mathbf{e}) = k\cdot\mathbf{e}+k\cdot(-(1\cdot\mathbf{e})) = k\cdot\mathbf{e}+k\cdot((-1)\cdot\mathbf{e})) = k\cdot\mathbf{e}+(-k)\cdot\mathbf{e} = (k-k)\cdot\mathbf{e} = 0\cdot\mathbf{e} = \mathbf{0}$ où on a utilisé le point 3 pour la quatrième égalité et le point 2 pour la dernière égalité. Enfin, montrons le point 5. Il s'agit de la démonstration d'une alternative. Supposons que $k\cdot\mathbf{e} = \mathbf{0}$ et que $k \neq 0$. Puisque k est un élément non nul de $\mathbb{K}$, il admet un inverse dans $\mathbb{K}$, que nous notons k^{-1}. Alors $\mathbf{e} = 1\cdot\mathbf{e} = (k^{-1}k)\cdot\mathbf{e} = k^{-1}\cdot(k\cdot\mathbf{e}) = k^{-1}\cdot\mathbf{0} = \mathbf{0}$, où on a utilisé le point 4 pour la dernière égalité. □

Au vu du point 3 de cette proposition, on pourra écrire $-k\cdot\mathbf{e}$ sans qu'il y ait d'ambiguïté.

À partir d'ici et jusqu'à la fin du chapitre, nous supposons toujours que E est un espace vectoriel sur $\mathbb{K}$.

2.2 Indépendance linéaire

Définition 2.4. Soit $n \in \mathbb{N}$. Une *combinaison linéaire* des vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_n$ de E est une expression de la forme

$$k_1\mathbf{x}_1 + \cdots + k_n\mathbf{x}_n \tag{2.1}$$

avec $k_1, \ldots, k_n \in \mathbb{K}$. Les scalaires $k_1, \ldots, k_n$ sont appelés les *coefficients* de cette combinaison linéaire. La combinaison linéaire (2.1) est dite *triviale* si tous ses coefficients sont nuls.

Remarquez qu'en particulier, une combinaison linéaire de 0 vecteur vaut toujours $\mathbf{0}$. Dans ce cas, on parle de combinaison linéaire *vide*. De plus, il est important de noter qu'une combinaison linéaire de combinaison linéaires de vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_n$ est encore une combinaison linéaire de $\mathbf{x}_1, \ldots, \mathbf{x}_n$ [1].

Définition 2.5. Des vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_n$ de E sont *linéairement indépendants* si pour tous $k_1, \ldots, k_n \in \mathbb{K}$, on a

$$k_1\mathbf{x}_1 + \cdots + k_n\mathbf{x}_n = \mathbf{0} \implies k_1 = \cdots = k_n = 0.$$

Des vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_n$ de E sont *linéairement dépendants* s'ils ne sont pas *linéairement indépendants.*

Autrement dit, $\mathbf{x}_1, \ldots, \mathbf{x}_n$ sont linéairement indépendants si la seule façon d'obtenir $\mathbf{0}$ comme combinaison linéaire de $\mathbf{x}_1, \ldots, \mathbf{x}_n$ est de prendre tous les coefficients égaux à 0. En prenant la négation de cette définition [2], on obtient que des vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_n$ sont linéairement dépendants lorsque que $\mathbf{0}$ peut d'obtenir comme une combinaison linéaire non triviale de $\mathbf{x}_1, \ldots, \mathbf{x}_n$, c'est-à-dire s'il existe des scalaires $k_1, \ldots, k_n$ non tous nuls tels que $k_1\mathbf{x}_1 + \cdots + k_n\mathbf{x}_n = \mathbf{0}$.

Exemple 2.6.

- Les vecteurs

$$\mathbf{x} = \begin{pmatrix} 7 \\ 4 \\ 2 \\ -1 \end{pmatrix}, \ \mathbf{y} = \begin{pmatrix} 2 \\ 1 \\ -5 \\ 4 \end{pmatrix}, \ \mathbf{z} = \begin{pmatrix} 17 \\ 10 \\ 16 \\ -11 \end{pmatrix}$$

1. Si vous n'en êtes pas convaincus, c'est un petit exercice à faire absolument.
2. On a $A \implies B \equiv \neg A \vee B$, $\neg(A \vee B) \equiv \neg A \wedge \neg B$, $\neg(\forall x \varphi(x)) \equiv \exists x \neg\varphi(x)$. On obtient donc que la négation de $\forall x, A(x) \implies B(x)$ est $\exists x, A(x) \wedge \neg B(x)$.

de $\mathbb{C}^4$ sont linéairement dépendants. Pour le voir, nous résolvons l'équation $a\mathbf{x}+b\mathbf{y}+c\mathbf{z}=\mathbf{0}$, où les inconnues sont les coefficients $a,b,c\in\mathbb{C}$. On obtient successivement

$$\begin{cases} 7a+2b+17c=0\\ 4a+\ \ b+10c=0\\ 2a-5b+16c=0\\ -a+4b-11c=0\end{cases} \iff \begin{cases} 30b-60c=0\\ 17b-34c=0\\ 3b-\ \ 6c=0\\ -a+\ \ 4b-11c=0\end{cases}$$
$$\iff \begin{cases} b-\ \ 2c=0\\ -a+4b-11c=0\end{cases}$$
$$\iff \begin{cases} b=2c\\ a\ \ =-3c\end{cases}$$

et il est maintenant facile de voir que ce système possède la solution non triviale $(a,b,c)=(-3,2,1)$, et donc que $-3\mathbf{x}+2\mathbf{y}+\mathbf{z}=\mathbf{0}$. Les vecteurs $\mathbf{x},\mathbf{y},\mathbf{z}$ sont donc linéairement dépendants dans $\mathbb{C}^4$.

- Dans l'exemple précédent, on peut remarquer que si les inconnues a,b,c étaient recherchées dans $\mathbb{R}$ ou $\mathbb{Q}$, la conclusion aurait été la même (puisqu'il existe une solution non triviale dans $\mathbb{Q}$). L'exemple que nous allons donner montre qu'il n'en est pas toujours ainsi. Soient E l'espace vectoriel $\mathbb{R}$ sur $\mathbb{R}$ et F l'espace vectoriel $\mathbb{R}$ sur $\mathbb{Q}$. vu comme espace vectoriel sur $\mathbb{R}$. Les réels 1 et $\sqrt{2}$ vu comme vecteurs de E sont linéairement dépendants (car $\sqrt{2}\cdot 1-1\cdot\sqrt{2}=0$) mais, lorsqu'ils sont vus comme vecteurs de F, ils sont linéairement indépendants (il n'est pas possible de trouver des rationnels a,b non tous deux nuls tels que $a\cdot 1+b\cdot\sqrt{2}=0$ car $\sqrt{2}\notin\mathbb{Q}$).

Proposition 2.7.

1. *Un vecteur* $\mathbf{e}$ *est linéairement dépendant si et seulement si* $\mathbf{e}=\mathbf{0}$.
2. *Si* $\mathbf{x}_1,\ldots,\mathbf{x}_n$ *sont des vecteurs linéairement dépendants et si* $m>n$*, alors pour tous vecteurs* $\mathbf{x}_{n+1},\ldots,\mathbf{x}_m$*, les vecteurs* $\mathbf{x}_1,\ldots,\mathbf{x}_m$ *sont encore linéairement dépendants.*
3. *Si* $\mathbf{x}_1,\ldots,\mathbf{x}_n$ *sont des vecteurs linéairement indépendants et si* $m<n$*, alors les vecteurs* $\mathbf{x}_1,\ldots,\mathbf{x}_m$ *sont encore linéairement indépendants.*
4. *Des vecteurs sont linéairement dépendants si et seulement l'un est combinaison linéaire des autres.*
5. *Soient* $\mathbf{x}_1,\ldots,\mathbf{x}_n$ *des vecteurs linéairement indépendants et soit un vecteur* $\mathbf{e}$*. Les vecteurs* $\mathbf{x}_1,\ldots,\mathbf{x}_n,\mathbf{e}$ *sont linéairement dépendants si et seulement* $\mathbf{e}$ *est combinaison linéaire de* $\mathbf{x}_1,\ldots,\mathbf{x}_n$.

Démonstration. Montrons le premier point. La condition est suffisante puisque $1\cdot\mathbf{0}=\mathbf{0}$. La condition est nécessaire par le point 5 de la proposition 2.3.

Montrons le cinquième point. Si $\mathbf{e}=k_1\mathbf{x}_1+\cdots+k_n\mathbf{x}_n$ avec $k_1,\ldots,k_n\in\mathbb{K}$, alors $k_1\mathbf{x}_1+\cdots+k_n\mathbf{x}_n+(-1)\mathbf{e}=\mathbf{0}$ et cette combinaison linéaire est non triviale. Supposons à présent que $\mathbf{x}_1,\ldots,\mathbf{x}_n,\mathbf{e}\in E$ sont linéairement dépendants. Alors il existe $k_1,\ldots,k_n,k\in\mathbb{K}$ non tous nuls tels que $k_1\mathbf{x}_1+\cdots+k_n\mathbf{x}_n+k\mathbf{e}=\mathbf{0}$. Nécessairement $k\neq 0$ car sinon un des k_i serait non nul et on aurait une combinaison linéaire non triviale des $\mathbf{x}_i$ qui s'annulerait, mettant en défaut leur indépendance linéaire. Par conséquent $\mathbf{e}$ est combinaison linéaire des $\mathbf{x}_i$:

$$\mathbf{e}=-\frac{k_1}{k}\mathbf{x}_1-\cdots-\frac{k_n}{k}\mathbf{x}_n.$$

Les autres points sont laissés en exercice. □

Théorème 2.8 (Steinitz). *Soit $n \in \mathbb{N}$. Si $n+1$ vecteurs sont combinaisons linéaires de n vecteurs, alors ils sont linéairement dépendants.*

Démonstration. Nous procédons par récurrence sur n. Le cas de base est immédiat : le seul vecteur qui est combinaison linéaire de 0 vecteur est le vecteur $\mathbf{0}$, qui est bien linéairement dépendant. Soit $n > 0$ et supposons que le résultat est vérifié pour tout $m < n$. Soient $\mathbf{x}_1, \ldots, \mathbf{x}_n \in E$ et soient

$$\begin{aligned} \mathbf{y}_1 &= k_{1,1}\mathbf{x}_1 + \cdots + k_{1,n}\mathbf{x}_n \\ \mathbf{y}_2 &= k_{2,1}\mathbf{x}_1 + \cdots + k_{2,n}\mathbf{x}_n \\ &\vdots \\ \mathbf{y}_{n+1} &= k_{n+1,1}\mathbf{x}_1 + \cdots + k_{n+1,n}\mathbf{x}_n. \end{aligned}$$

où $k_{i,j} \in \mathbb{K}$ pour tout $i \in \{1, \ldots, n+1\}$ et tout $j \in \{1, \ldots, n\}$. Notre but est de montrer que les vecteurs $\mathbf{y}_1, \ldots, \mathbf{y}_{n+1}$ sont linéairement dépendants. Si tous les coefficients $k_{i,j}$ sont nuls, alors tous les $\mathbf{y}_i$ sont nuls et ils sont donc linéairement dépendants. Sinon, au moins un des coefficients $k_{i,j}$ est non nul, et on peut supposer [3] qu'il s'agit de $k_{1,1}$. Considérons à présent les n vecteurs

$$\begin{aligned} \mathbf{z}_2 &= k_{1,1}\mathbf{y}_2 - k_{2,1}\mathbf{y}_1 \\ &\vdots \\ \mathbf{z}_{n+1} &= k_{1,1}\mathbf{y}_{n+1} - k_{n+1,1}\mathbf{y}_1. \end{aligned}$$

Ces n vecteurs sont combinaisons linéaires des $n-1$ vecteurs $\mathbf{x}_2, \ldots, \mathbf{x}_n$ et par hypothèse de récurrence, ils sont linéairement dépendants : il existe des coefficients $\ell_2, \ldots, \ell_{n+1} \in \mathbb{K}$ non tous nuls tels que $\ell_2\mathbf{z}_2 + \cdots + \ell_{n+1}\mathbf{z}_{n+1} = \mathbf{0}$. Par définition des vecteurs $\mathbf{z}_i$, on a

$$\begin{aligned} \ell_2\mathbf{z}_2 + \cdots + \ell_{n+1}\mathbf{z}_{n+1} &= \ell_2(k_{1,1}\mathbf{y}_2 - k_{2,1}\mathbf{y}_1) + \cdots + \ell_{n+1}(k_{1,1}\mathbf{y}_{n+1} - k_{n+1,1}\mathbf{y}_1) \\ &= (-\ell_2 k_{2,1} - \cdots - \ell_{n+1}k_{n+1,1})\mathbf{y}_1 + \ell_2 k_{1,1}\mathbf{y}_2 + \cdots + \ell_{n+1}k_{1,1}\mathbf{y}_{n+1}. \end{aligned}$$

Comme au moins un des coefficients $\ell_2 k_{1,1}, \ldots, \ell_{n+1}k_{1,1}$ est non nul, on obtient une combinaison linéaire non triviale de $\mathbf{y}_1, \ldots, \mathbf{y}_{n+1}$ valant $\mathbf{0}$. □

2.3 Sous-espaces vectoriels

Pour vérifier qu'un ensemble donné est un espace vectoriel, nous ne disposons jusqu'ici que de la définition. Dans cette section, nous allons voir comment, étant donné un espace vectoriel, obtenir facilement de nouveaux espaces vectoriels.

Définition 2.9. Un *sous-espace vectoriel* de E est une partie de E qui contient les combinaisons linéaires de ses vecteurs.

Proposition 2.10. *Un sous-espace vectoriel de E (muni des mêmes opérations d'addition et de multiplication par un scalaire) est un espace vectoriel.*

Démonstration. Soit F un sous-espace vectoriel de E. Par définition d'un sous-espace vectoriel, l'addition et la multiplication scalaire restreintes à F [4] sont internes à F. Le neutre $\mathbf{0}$ appartient à F puisqu'il est obtenu comme la combinaison linéaire vide. L'opposé d'un vecteur $\mathbf{f}$ appartenant à F appartient aussi à F puisqu'il est obtenu comme combinaison linéaire de lui-même, à savoir $-\mathbf{f} = (-1) \cdot \mathbf{f}$. Les autres points à vérifier pour être un espace vectoriel sont valables pour n'importe quelle partie de E, donc en particulier pour F. □

3. Quitte à renuméroter les vecteurs et les coefficients.
4. C'est-à-dire les opérations $+ \colon F \times F \to E$ et $\cdot \colon \mathbb{K} \times F \to E$.

Nous allons maintenant introduire la notion de sous-espace vectoriel engendré par des vecteurs donnés $\mathbf{x}_1, \dots, \mathbf{x}_n$.

Définition 2.11. L'*enveloppe linéaire* d'une partie A de E est l'ensemble de toutes les combinaisons linéaires des vecteurs de A. On la note $\rangle A \langle$. Par abus de langage, si $A = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$, on écrit $\rangle \mathbf{x}_1, \dots, \mathbf{x}_n \langle$ au lieu de $\rangle \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \langle$ et on parle de l'*enveloppe linéaire des vecteurs* $\mathbf{x}_1, \dots, \mathbf{x}_n$.

Remarquons[5] qu'avec cette définition, on a $\rangle \emptyset \langle = \{\mathbf{0}\}$.

Proposition 2.12. *L'enveloppe linéaire d'une partie A de E est le plus petit (au sens de l'inclusion) sous-espace vectoriel de E contenant A.*

Démonstration. Soit A une partie de E. Premièrement, $\rangle A \langle$ est un sous-espace vectoriel de E puisque toute combinaison linéaire de combinaisons linéaires de vecteurs de A est encore une combinaison linéaire de vecteurs de A. Deuxièmement, il est clair que $A \subseteq \rangle A \langle$. Troisièmement, tout sous-espace vectoriel de E contenant A contient également $\rangle A \langle$ puisqu'un sous-espace vectoriel contient les combinaisons linéaires de ses vecteurs par définition. □

Au vu de la proposition précédente, l'enveloppe linéaire de A est parfois aussi appelée le *sous-espace vectoriel engendré par A*. De même, on dit que $\rangle \mathbf{x}_1, \dots, \mathbf{x}_n \langle$ est le *sous-espace vectoriel engendré par les vecteurs* $\mathbf{x}_1, \dots, \mathbf{x}_n$.
Voici un résultat pratique pour vérifier qu'un ensemble est un sous-espace vectoriel.

Proposition 2.13. *Une partie F de E est un sous-espace vectoriel de E si et seulement si les deux conditions suivantes sont satisfaites :*

1. $\mathbf{0} \in F$
2. $\mathbf{x}, \mathbf{y} \in F,\ k, \ell \in \mathbb{K} \implies k\mathbf{x} + \ell\mathbf{y} \in F$.

Démonstration. La condition nécessaire est immédiate. Vérifions la condition suffisante. Supposons que $F \subseteq E$ et que les conditions 1 et 2 soient vérifiées. Montrons par récurrence sur $n \in \mathbb{N}$ que F contient les combinaisons de n de ses vecteurs. Si $n = 0$, alors la seule combinaison linéaire de 0 vecteur est la combinaison linéaire vide, qui vaut $\mathbf{0}$ et qui appartient à F au vu de la conditon 1. Supposons à présent que $n \geq 1$ et que F contient les combinaisons de $n-1$ de ses vecteurs. Soient $\mathbf{x}_1, \dots, \mathbf{x}_n \in F$ et $k_1, \dots, k_n \in \mathbb{K}$. Par hypothèse de récurrence, le vecteur $k_1\mathbf{x}_1 + \cdots + k_{n-1}\mathbf{x}_{n-1}$ appartient à F. En utilisant la condition 2, on obtient que

$$k_1\mathbf{x}_1 + \cdots + k_n\mathbf{x}_n = 1 \cdot (k_1\mathbf{x}_1 + \cdots + k_{n-1}\mathbf{x}_{n-1}) + k_n\mathbf{x}_n$$

appartient à F. Ceci montre que F est un sous-espace vectoriel de E. □

2.4 Base et dimension

Définition 2.14. Une *partie génératrice* de E est une partie G de E telle que tout vecteur de E s'écrit comme combinaison linéaire des vecteurs de G. Autrement dit, on a $E = \rangle G \langle$. Dans ce cas, on dit que les vecteurs de G *engendrent* E.

Définition 2.15. Une *partie libre* de E est une partie L de E telle que toute partie finie de L est formée de vecteurs linéairement indépendants.

Remarque 2.16. Tout sous-ensemble d'une partie libre est encore libre et que tout sur-ensemble d'une partie génératrice est encore générateur.

5. Pourquoi ?

Définition 2.17. Une *base* de E est une partie libre et génératrice de E.

Nous démontrons les résultats qui suivent dans le cas d'un espace vectoriel de dimension finie (c'est-à-dire pour l'instant dans le cas où E admet une partie génératrice finie). La suite du chapitre se poursuivra donc dans ce contexte uniquement.

Lemme 2.18. *Si E admet une partie génératrice finie de taille p, alors toute partie libre de E est de taille au plus p.*

Démonstration. Ceci est une conséquence du théorème de Steinitz. Soit G une partie génératrice finie de taille p et soit L une partie libre. Procédons par l'absurde et supposons que L possède au moins $p+1$ vecteurs. Par définition d'une partie génératrice, ces $p+1$ vecteurs sont combinaisons linéaires des p vecteurs de G et par le théorème de Steinitz, ils sont linéairement dépendants. Ceci contredit le fait que L soit une partie libre. □

Théorème 2.19 (Équipotence des bases)**.** *Si E admet une partie génératrice finie, alors toutes les bases de E ont le même nombre d'éléments.*

Démonstration. Supposons que E admet une partie génératrice finie. Au vu du lemme 2.18, nous savons que toutes les bases de E sont finies. Soient B et B' deux bases de E. Comme B est une partie libre et B' est une partie génératrice, le lemme 2.18 nous dit que $|B| \leq |B'|$. Symétriquement, on obtient $|B'| \leq |B|$. Ainsi, $|B| = |B'|$. □

Ce théorème combiné au suivant nous conduira à l'important concept de dimension d'un espace vectoriel.

Théorème 2.20 (Complétion des bases)**.** *Supposons que E admette une partie génératrice finie. Alors pour toute partie génératrice G de E et toute partie libre L de E, il existe une base B de E telle que $L \subseteq B \subseteq L \cup G$.*

Démonstration. Par hypothèse, E admet une partie génératrice finie, disons de taille p. Soit G une partie génératrice de E et soit L une partie libre de E. Par le lemme 2.18, nous savons que toutes les parties libres de E sont de taille au plus p. En particulier, c'est le cas de L. Parmi toutes les parties libres X de E telles que $L \subseteq X \subseteq L \cup G$, on en considère une de taille maximum [6]. Notons-la Y et montrons que cette partie Y est forcément une base de E. Elle est libre par définition. Montrons qu'elle est aussi génératrice. Tout vecteur de G est combinaison linéaire des vecteurs de Y, car si $\mathbf{g} \in G$ n'est pas combinaison linéaire des vecteurs de Y, alors au vu de la proposition 2.7, la partie $Y \cup \{\mathbf{g}\}$ serait encore libre, ce qui contredirait la maximalité de la taille de Y. Mais puisque G est une partie génératrice, tout vecteur de E est combinaison linéaire des vecteurs de G, et donc aussi des vecteurs de Y. □

Corollaire 2.21. *Supposons que E admette une partie génératrice finie.*

1. *Toute partie génératrice de E contient une base.*
2. *Toute partie libre de E est incluse dans une base.*

Démonstration. On obtient le point 1 en appliquant le théorème de complétion des bases avec $L = \emptyset$ et le point 2 en appliquant le théorème de complétion des bases avec $G = E$ (ou n'importe quelle partie génératrice G de E). □

6. Ceci a du sens puisque L est un tel X et que toutes les parties libres de E possèdent au plus p éléments

Remarquez qu'en particulier, nous avons montré que si un espace vectoriel admet une partie génératrice finie, alors il possède une base. Signalons tout de même que le fait qu'un espace vectoriel admette une base n'est pas une évidence en général, et pour cause. Il n'est en fait pas possible de démontrer ou bien d'infirmer cette affirmation !

Le théorème 2.19 et le fait même que E possède une base permettent de donner du sens à la définition suivante.

Définition 2.22. Dans le cas où E admet une partie génératrice finie, on appelle *dimension* de E le nombre d'éléments d'une base quelconque de E. On note la dimension $\dim_{\mathbb{K}}(E)$ ou simplement $\dim(E)$ si l'ensemble des scalaires $\mathbb{K}$ est implicitement connu. On dit que E est de *dimension finie* s'il admet une partie génératrice finie (et donc une base finie).

Exemple 2.23.

- On a $\dim_{\mathbb{K}}(\mathbb{K}^{\ell\times c}) = \ell c$. En particulier, $\dim_{\mathbb{K}}(\mathbb{K}^{m\times m}) = m^2$ et $\dim_{\mathbb{K}}(\mathbb{K}^m) = m$.
- L'espace vectoriel $\mathbb{K}[X]$ n'est pas de dimension finie puisque pour tout $d \in \mathbb{N}$, les polynômes $1, X, X^2, X^3, \ldots, X^d$ sont linéairement indépendants. Par contre, pour tout $d \in \mathbb{N}$, on a $\dim_{\mathbb{K}}(\mathbb{K}_d[X]) = d + 1$.
- En général, si $\mathbb{K} \neq \mathbb{K}'$, alors $\dim_{\mathbb{K}}(E) \neq \dim_{\mathbb{K}'}(E)$. Par exemple, $\dim_{\mathbb{C}}(\mathbb{C}^m) = m$ alors que $\dim_{\mathbb{R}}(\mathbb{C}^m) = 2m$.

Le résultat suivant étudie la dimension d'un sous-espace vectoriel d'un espace vectoriel de dimension finie.

Proposition 2.24. *Supposons que E est de dimension finie et que F est un sous-espace vectoriel de E. On a*

- $\dim(F) \leq \dim(E)$,
- $\dim(F) = \dim(E) \iff F = E$.

Démonstration. Soit B une base de F. En particulier, B est une partie libre de E et par le corollaire 2.21, B est incluse dans une base B' de E. Ainsi $\dim(F) = |B| \leq |B'| = \dim(E)$, et le premier point est démontré. Passons à présent au second point. La condition suffisante ($\Longleftarrow$) est immédiate. Pour montrer la condition nécessaire ($\Longrightarrow$), nous supposons que $\dim(F) = \dim(E)$. Dans ce cas, on a $|B| = |B'|$. Puisque $B \subseteq B'$, on obtient que $B = B'$. Ainsi, tout élément de E est combinaison linéaire d'éléments de B et donc de F. On obtient donc $E \subseteq F$. Comme évidemment $F \subseteq E$, on conclut que $F = E$, comme souhaité. □

Nous sommes à présent prêts pour introduire le concept primordial de composantes d'un vecteur dans une base. Ceci vaut bien une section à part entière.

2.5 Composantes d'un vecteur dans une base

Nous supposons à présent, et pour le reste de ce chapitre, que E est un espace vectoriel de dimension finie m. Dans cette section, nous supposerons qu'une base B de E est une partie ordonnée de E :

$$B = (\mathbf{b}_1, \ldots, \mathbf{b}_m).$$

Ceci n'est pas une vraie restriction dans le sens où si on souhaite énumérer les éléments de B dans un autre ordre, il suffira d'appliquer une permutation aux éléments de B.

Théorème 2.25 (Décomposition dans une base)**.** *Soit $B = (\mathbf{b}_1, \ldots, \mathbf{b}_m)$ une base de E. Tout vecteur de E admet une décomposition unique dans la base B : pour tout $\mathbf{x} \in E$, il existe des scalaires uniques $x_1, \ldots, x_m \in \mathbb{K}$ tels que $\mathbf{x} = x_1\mathbf{b}_1 + \cdots + x_m\mathbf{b}_m$.*

Démonstration. L'existence de la décomposition vient du fait qu'une base soit en particulier une partie génératrice. Montrons à présent l'unicité d'une telle décomposition. Supposons que $x_1\mathbf{b}_1 + \cdots + x_m\mathbf{b}_m = y_1\mathbf{b}_1 + \cdots + y_m\mathbf{b}_m$ avec $x_1, \ldots, x_m, y_1, \ldots, y_m \in \mathbb{K}$. Alors

$$(x_1 - y_1)\mathbf{b}_1 + \cdots + (x_m - y_m)\mathbf{b}_m = \mathbf{0}.$$

Les vecteurs $\mathbf{b}_1, \ldots, \mathbf{b}_m$ étant linéairement indépendants, on obtient $x_1 - y_1 = 0, \ldots, x_m - y_m = 0$, c'est-à-dire $x_1 = y_1, \ldots, x_m = y_m$. $\square$

Définition 2.26. Soient $B = (\mathbf{b}_1, \ldots, \mathbf{b}_m)$ une base de E et soit $\mathbf{x} \in E$. Les coefficients $x_1, \ldots, x_m$ de la décomposition unique de $\mathbf{x}$ dans la base B sont appelés les *composantes* de $\mathbf{x}$ dans la base B.

Définition 2.27. Pour chaque base $B = (\mathbf{b}_1, \ldots, \mathbf{b}_m)$ de E, on définit une fonction

$$\Psi_B \colon E \to \mathbb{K}^m,\ \mathbf{x} \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$$

qui à un vecteur $\mathbf{x} = x_1\mathbf{b}_1 + \cdots x_m\mathbf{b}_m$ de E associe le vecteur $\begin{pmatrix} x_1 \cdots x_m \end{pmatrix}^\intercal$ de ses composantes dans la base B.

Le corollaire suivant nous apprend que tout espace vectoriel de dimension finie m est isomorphe à l'espace vectoriel $\mathbb{K}^m$.

Corollaire 2.28. *Pour toute base B de E, la fonction Ψ_B est un isomorphisme d'espaces vectoriels, c'est-à-dire une bijection telle que pour tous $\mathbf{e}, \mathbf{f} \in E$ et tous $k, \ell \in \mathbb{K}$, on a $\Psi_B(k\mathbf{e} + \ell\mathbf{f}) = k\Psi_B(\mathbf{e}) + \ell\Psi_B(\mathbf{f})$.*

Démonstration. C'est une simple vérification. $\square$

Corollaire 2.29. *Soit B une base de E. Des vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_n$ de E sont linéairement dépendants (dans E) si et seulement si les vecteurs $\Psi_B(\mathbf{x}_1), \ldots, \Psi_B(\mathbf{x}_n)$ de $\mathbb{K}^m$ sont linéairement dépendants (dans $\mathbb{K}^m$).*

Démonstration. Laissée en exercice [7]. $\square$

2.6 Changement de base

Le but de cette section est de montrer, étant donné deux bases B et B' de E, comment obtenir la décomposition d'un vecteur dans une de ces bases à partir de sa décomposition dans l'autre.

Exemple 2.30. Soient $\mathbf{u} = \left(\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right)$, $\mathbf{v} = \left(\begin{smallmatrix} 1 \\ 3 \end{smallmatrix}\right)$, $\mathbf{u}' = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ et $\mathbf{v}' = \left(\begin{smallmatrix} 2 \\ 0 \end{smallmatrix}\right)$. Alors $B = (\mathbf{u}, \mathbf{v})$ et $B' = (\mathbf{u}', \mathbf{v}')$ sont deux bases de $\mathbb{R}^2$. Si, étant donné les composantes (a, b) d'un vecteur $\mathbf{e} \in \mathbb{R}^2$ dans la base B, nous voulons obtenir ses composantes (a', b') dans la base B', on peut procéder comme suit. Puisque $\mathbf{e} = a\mathbf{u} + b\mathbf{v} = a'\mathbf{u}' + b'\mathbf{v}'$, on obtient

$$\begin{aligned} \begin{pmatrix} 2a \\ a \end{pmatrix} + \begin{pmatrix} b \\ 3b \end{pmatrix} = \begin{pmatrix} a' \\ a' \end{pmatrix} + \begin{pmatrix} 2b' \\ 0 \end{pmatrix} \quad &\iff \quad \begin{cases} 2a + \ \ b = a' + 2b' \\ \ \ a + 3b = a' \end{cases} \\ &\iff \quad \begin{cases} 2b' &= 2a + b - (a + 3b) \\ \ \ a' &= a + 3b \end{cases} \end{aligned}$$

7. Suggestion : commencer par montrer que, pour tout $\mathbf{e} \in E$, on a $\Psi_B(\mathbf{e}) = \mathbf{0}_{\mathbb{K}^m} \iff \mathbf{e} = \mathbf{0}_E$.

$$\iff \begin{cases} b' &= \frac{1}{2}a - b \\ a' &= a + 3b \end{cases}$$
$$\iff \begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ \frac{1}{2} & -1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}.$$

On peut remarquer que les colonnes de la matrice

$$\begin{pmatrix} 1 & 3 \\ \frac{1}{2} & -1 \end{pmatrix}$$

sont en fait les composantes de $\mathbf{u}$ et $\mathbf{v}$ dans B'. En effet, on a

$$\mathbf{u}' + \frac{1}{2}\mathbf{v}' = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \mathbf{u}$$
$$3\mathbf{u}' - \mathbf{v}' = 3\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \mathbf{v}.$$

Cette observation est en fait générale, et nous fournit un moyen alternatif pour obtenir les composantes d'un vecteur dans un base à partir de ses composantes dans une autre base. C'est l'objet du théorème suivant.

Théorème 2.31 (Changement de base)**.** *Soient $B = (\mathbf{b}_1, \ldots, \mathbf{b}_m)$ et $B' = (\mathbf{b}'_1, \ldots, \mathbf{b}'_m)$ deux bases de E. Pour tout $\mathbf{e} \in E$, on a*

$$\Psi_{B'}(\mathbf{e}) = \mathcal{M}(B, B')\Psi_B(\mathbf{e})$$

où $\mathcal{M}(B, B')$ est la matrice de $\mathbb{K}^{m\times m}$ dont la j-ème colonne est le vecteur $\Psi_{B'}(\mathbf{b}_j)$, pour tout $j \in \{1, \ldots, m\}$.

Démonstration. Soit $\mathbf{e} \in E$ et soient $\Psi_B(\mathbf{e}) = (e_1 \cdots e_m)^\intercal$ et $\Psi_{B'}(\mathbf{e}) = (e'_1 \cdots e'_m)^\intercal$. Soit encore $\mathcal{M}(B, B') = (k_{ij})_{\substack{1\le i\le m \\ 1\le j\le m}}$. Par définition de $\mathcal{M}(B, B')$, on a

$$\mathbf{b}_j = k_{1j}\mathbf{b}'_1 + \cdots k_{mj}\mathbf{b}'_m$$

pour tout $j \in \{1, \ldots, m\}$. On obtient que

$$\mathbf{e} = \sum_{j=1}^{m} e_j \mathbf{b}_j = \sum_{j=1}^{m} e_j \sum_{i=1}^{m} k_{ij}\mathbf{b}'_i = \sum_{i=1}^{m} \Big(\sum_{j=1}^{m} k_{ij}e_j\Big)\mathbf{b}'_i.$$

Par unicité de la décomposition de $\mathbf{e}$ dans la base B', on obtient que

$$e'_i = \sum_{j=1}^{m} k_{ij}e_j$$

pour tout $i \in \{1, \ldots, m\}$. En notation matricielle, ceci revient à dire que

$$\big(\Psi_{B'}(\mathbf{e})\big)_i = (\mathcal{M}(B, B')\Psi_B(\mathbf{e}))_i$$

pour tout $i \in \{1, \ldots, m\}$, d'où la thèse. □

Définition 2.32. La matrice $\mathcal{M}(B, B')$ du théorème précédent est appelée la *matrice de changement de base de B vers B'.*

Proposition 2.33. *Pour toutes bases B, B', B'' de E, on a*

$$\mathcal{M}(B, B'') = \mathcal{M}(B', B'')\mathcal{M}(B, B').$$

Démonstration. Il suffit de montrer que, pour tout $j \in \{1, \ldots, m\}$, les j-èmes colonnes de $\mathcal{M}(B, B'')$ et de $\mathcal{M}(B', B'')\mathcal{M}(B, B')$ sont égales. Supposons que $B = (\mathbf{b}_1, \ldots, \mathbf{b}_m)$. Alors, d'une part, la j-ème colonne de $\mathcal{M}(B, B'')$ est $\Psi_{B''}(\mathbf{b}_j)$ et d'autre part, la j-ème colonne de $\mathcal{M}(B, B')$ est $\Psi_{B'}(\mathbf{b}_j)$. Par définition du produit matriciel, la j-ème colonne de $\mathcal{M}(B', B'')\mathcal{M}(B, B')$ est $\mathcal{M}(B', B'')\Psi_{B'}(\mathbf{b}_j)$. On conclut en appliquant le théorème 2.31 avec $\mathbf{e} = \mathbf{b}_j$. □

Corollaire 2.34. *Pour toutes bases B et B' de E, on a*

$$\mathcal{M}(B', B) = (\mathcal{M}(B, B'))^{-1}.$$

Démonstration. Il suffit de remarquer que pour toute base B de E, on a $\mathcal{M}(B, B) = I_m$. Le résultat découle alors de la proposition précédente. □

2.7 Retour aux matrices et théorème du rang

Nous avons déjà mentionné le fait que l'ensemble des solutions d'un système linéaire homogène de ℓ équations à c inconnues était un sous-espace vectoriel de $\mathbb{K}^c$. Rappelons[8] que le rang d'une matrice $A \in \mathbb{K}^{\ell \times c}$ est la taille de la plus grande sous-matrice carrée de déterminant non nul de A. Rappelons aussi la proposition suivante, que nous avions démontrée en "Mathématique pour l'informatique 1" et que nous reformulons ici dans les termes de ce chapitre.

Proposition 2.35. *Soit $A \in \mathbb{K}^{m \times m}$. Les propositions suivantes sont équivalentes.*

1. $\det(A) = 0$.
2. *Les colonnes de A sont linéairement dépendantes dans $\mathbb{K}^m$.*
3. *Les lignes de A sont linéairement dépendantes dans $\mathbb{K}_m$.*

Remarque 2.36. Pour faire la distinction entre les matrices-lignes et les matrices-colonnes, on note $\mathbb{K}_\ell$ l'ensemble des matrices-lignes à ℓ éléments et $\mathbb{K}^c$ l'ensemble des matrices-colonnes à c éléments.

Théorème 2.37. *Le rang d'une matrice $A \in \mathbb{K}^{\ell \times c}$ est égal au nombre maximum de colonnes (respectivement de lignes) linéairement indépendantes de A, où l'indépendance linéaire est comprise dans l'espace vectoriel $\mathbb{K}^c$ (respectivement $\mathbb{K}_\ell$).*

Démonstration. Soit $A \in \mathbb{K}^{\ell \times c}$ et notons $r = \mathrm{rg}(A)$. Nous montrons le théorème pour les colonnes, le cas des lignes étant similaire. Par définition du rang de A, il existe une sous-matrice carrée $r \times r$ de déterminant non nul. Par la proposition 2.35, les r colonnes de A qui traversent cette sous-matrice sont nécessairement linéairement indépendantes (car leurs restrictions aux lignes de la sous-matrice le sont). Montrons à présent que A ne possède pas plus de r colonnes linéairement indépendantes. Procédons par l'absurde et supposons que A possède $r + 1$ colonnes linéairement indépendantes. Par définition du rang, A doit alors posséder exactement r lignes (car elle ne peux en possèder moins en ayant un rang r et ne peut en posséder plus car alors, au vu de la proposition 2.35, on pourrait trouver une matrice carrée $(r + 1) \times (r + 1)$ de déterminant non nul). Mais alors, les $r + 1$ colonnes de A que nous considérons sont combinaisons linéaires des r colonnes de longueur r

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

8. On étendra aisément les définitions que nous avons vues pour $\mathbb{K} = \mathbb{C}$ à $\mathbb{R}, \mathbb{Q}$ et $\mathbb{Z}_m$ (pour m premier).

et ne peuvent donc être linéairement indépendantes au vu du théorème de Steinitz, nous amenant à une contradiction. □

Nous concluons cette section en démontrant le très utile théorème du rang.

Définition 2.38. Soit $A \in \mathbb{K}^{\ell \times c}$. L'*image de* A est l'ensemble des vecteurs $\mathbf{y} \in \mathbb{K}^\ell$ tels qu'il existe $\mathbf{x} \in \mathbb{K}^c$ tel que $A\mathbf{x} = \mathbf{y}$. On note $\mathrm{Im}(A)$ l'image de A :

$$\mathrm{Im}(A) = \{\mathbf{y} \in \mathbb{K}^\ell \colon \exists \mathbf{x} \in \mathbb{K}^c,\ A\mathbf{x} = \mathbf{y}\} = \{A\mathbf{x} \colon \mathbf{x} \in \mathbb{K}^c\}.$$

Le *noyau de* A est l'ensemble des vecteurs $\mathbf{x} \in \mathbb{K}^c$ tels que $A\mathbf{x} = \mathbf{0}$. On note $\mathrm{Ker}(A)$ le noyau de A :

$$\mathrm{Ker}(A) = \{\mathbf{x} \in \mathbb{K}^c \colon A\mathbf{x} = \mathbf{0}\}.$$

Exemple 2.39. Soit

$$A = \begin{pmatrix} 1 & 2 & 3i \\ 0 & 0 & 8 \end{pmatrix}.$$

On a

$$\begin{pmatrix} 1 & 2 & 3i \\ 0 & 0 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + 3iz \\ 8z \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 2 \\ 0 \end{pmatrix} + z \begin{pmatrix} 3i \\ 8 \end{pmatrix}.$$

Dès lors, on obtient que

$$\mathrm{Im}(A) \quad = \quad \Big\rangle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3i \\ 8 \end{pmatrix} \Big\langle \quad = \quad \Big\rangle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3i \\ 8 \end{pmatrix} \Big\langle$$

et donc que $\dim(\mathrm{Im}(A)) = 2$. Nous verrons qu'on a toujours $\dim \mathrm{Im}(A) = \mathrm{rg}(A)$, et on peut vérifier dans cet exemple que $\mathrm{rg}(A) = 2$ puisque par exemple $\left|\begin{smallmatrix} 1 & 3i \\ 0 & 8 \end{smallmatrix}\right| = 8 \neq 0$. De plus, on a

$$\begin{pmatrix} 1 & 2 & 3i \\ 0 & 0 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \left\{ \begin{array}{rcl} x + 2y + 3iz & = & 0 \\ 8z & = & 0 \end{array} \right. \iff \left\{ \begin{array}{rcl} x & = & -2y \\ z & = & 0 \end{array} \right. .$$

Dès lors, on a

$$\mathrm{Ker}(A) \ = \ \left\{ \begin{pmatrix} -2y \\ y \\ 0 \end{pmatrix} : y \in \mathbb{C} \right\} \ = \ \Big\rangle \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \Big\langle$$

et donc $\dim(\mathrm{Ker}(A)) = 1$. Nous allons montrer en général que $\dim(\mathrm{Ker}(A)) + \dim(\mathrm{Im}(A))$ est toujours égal au nombre de colonnes de A (c'est le théorème du rang). Cette égalité se vérifie dans cet exemple puisque $\dim(\mathrm{Ker}(A)) + \dim(\mathrm{Im}(A)) = 1 + 2 = 3$ et que A possède bien 3 colonnes.

Proposition 2.40. *L'image et le noyau d'une matrice de* $\mathbb{K}^{\ell \times c}$ *sont des sous-espaces vectoriels de* $\mathbb{K}^\ell$ *et de* $\mathbb{K}^c$ *respectivement.*

Démonstration. C'est une simple vérification. □

Proposition 2.41. *Soit* $A \in \mathbb{K}^{\ell \times c}$. *Alors* $\dim \mathrm{Ker}(A) = c - \mathrm{rg}(A)$.

Démonstration. Le noyau de A est l'ensemble des solutions du système linéaire homogène $A\mathbf{x} = \mathbf{0}$. Parmi les ℓ équations de ce système, on peut trouver au plus $\mathrm{rg}(A)$ équations linéairement indépendantes. En utilisant la méthode de Gauss pour la résolution de ce système, on obtient que la dimension de l'ensemble des solutions est $c - \mathrm{rg}(A)$.

Détaillons ce dernier argument en reprenant les notations de Mathématiques pour l'informatique 1. Le système $A\mathbf{x} = \mathbf{0}$ se réécrit

$$\left\{\begin{array}{rcl} B_{1,1}x_1 + B_{1,2}x_2 + \cdots + B_{1,r}x_r + B_{1,r+1}x_{r+1} + ... + B_{1c}x_c & = & 0 \\ & \vdots & \\ B_{r,r}x_r + B_{r,r+1}x_{r+1} + \cdots + B_{r,c}x_c & = & 0 \\ 0 & = & 0 \\ & \vdots & \\ 0 & = & 0 \end{array}\right.$$

où $B_{1,1}, \ldots, B_{r,r}$ sont tous non nuls et où il y a $\ell - r$ équations de la forme $0 = 0$. On a $r = \mathrm{rg}(A)$. Ce système est équivalent au système obtenu en supprimant toutes les équations de la forme $0 = 0$ et en déplaçant les inconnues $x_{r+1}, ..., x_c$ dans les seconds membres des r premières équations :

$$\left\{\begin{array}{rcl} B_{1,1}x_1 + B_{1,2}x_2 + \cdots + B_{1,r}x_r & = & -B_{1,r+1}x_{r+1} - \cdots - B_{1,c}x_c \\ & \vdots & \\ B_{r,r}x_r & = & -B_{r,r+1}x_{r+1} - \cdots - B_{r,c}x_c. \end{array}\right.$$

On obtient la valeur de x_r en fonction de $x_{r+1}, x_{r+2}, \ldots, x_c$ grâce à la dernière équation. Ensuite, en utilisant la dernière équation, on obtient la valeur de x_{r-1} en fonction de $x_{r+1}, x_{r+2}, \ldots, x_c$ grâce à l'avant-dernière équation. En remontant les équations, on obtient les valeurs de $x_1, x_2, \ldots, x_r$ en fonction de $x_{r+1}, x_{r+2}, \ldots, x_c$. Autrement dit, l'ensemble des solutions est de la forme

$$\begin{aligned} \mathrm{Ker}(A) &= \left\{ \begin{pmatrix} C_{1,r+1}x_{r+1} + C_{1,r+2}x_{r+2} + \cdots + C_{1,c}x_c \\ \vdots \\ C_{r,r+1}x_{r+1} + C_{r,r+2}x_{r+2} + \cdots + C_{r,c}x_c \\ x_{r+1} \\ \vdots \\ x_c \end{pmatrix} : x_{r+1}, \ldots, x_c \in \mathbb{C} \right\} \\ &= \left\{ x_{r+1} \begin{pmatrix} C_{1,r+1} \\ \vdots \\ C_{r,r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_{r+2} \begin{pmatrix} C_{1,r+2} \\ \vdots \\ C_{r,r+2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_c \begin{pmatrix} C_{1,c} \\ \vdots \\ C_{r,c} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} : x_{r+1}, \ldots, x_c \in \mathbb{C} \right\}. \end{aligned}$$

Les $c - r$ vecteurs

$$\begin{pmatrix} C_{1,r+1} \\ \vdots \\ C_{r,r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} C_{1,r+2} \\ \vdots \\ C_{r,r+2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} C_{1,c} \\ \vdots \\ C_{r,c} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

de $\mathbb{K}^c$ engendrent donc $\mathrm{Ker}(A)$. Comme ils sont aussi linéairement indépendants, ils forment une base de $\mathrm{Ker}(A)$ et on obtient que $\dim \mathrm{Ker}(A) = c - r = c - \mathrm{rg}(A)$. □

Proposition 2.42. *Soit* $A \in \mathbb{K}^{\ell \times c}$. *Alors* $\dim \mathrm{Im}(A) = \mathrm{rg}(A)$.

Démonstration. Il suffit de remarquer que l'image de A est l'ensemble des combinaisons linéaires des colonnes de A. Dès lors, la dimension de $\mathrm{Im}(A)$ est égale au nombre maximum de colonnes linéairement indépendantes de A, qui est lui-même égal au rang de A au vu du théorème 2.37. $\square$

Corollaire 2.43 (Théorème du rang)**.** *Pour toute matrice* $A \in \mathbb{K}^{\ell \times c}$, *on a*

$$\dim \mathrm{Im}(A) + \dim \mathrm{Ker}(A) = c.$$

Démonstration. C'est une conséquence immédiate des deux propositions précédentes. $\square$

Chapitre 3

Diagonalisation de matrices complexes

Le but de ce chapitre est de pouvoir déterminer si une matrice donnée est diagonalisable ou non, et dans le cas où elle l'est, d'être capable de fournir explicitement une matrice S qui diagonalise A.

Définition 3.1. Une matrice $A \in \mathbb{C}^{m\times m}$ est *diagonalisable* s'il existe une matrice inversible $S \in \mathbb{C}^{m\times m}$ telle que $S^{-1}AS$ est une matrice diagonale. On dit qu'une telle matrice S *diagonalise* A.

La diagonalisation de matrices a de nombreuses applications. Ceci vient du fait que lorsqu'une matrice est diagonalisable, il devient facile d'en calculer les puissances. En effet, si $S^{-1}AS$ est une matrice diagonale $\Delta = \operatorname{diag}(\lambda_1, \ldots, \lambda_m)$, alors $\Delta^n = \operatorname{diag}(\lambda_1^n, \ldots, \lambda_m^n)$ et

$$A^n = (S\Delta S^{-1})^n = S\Delta S^{-1}S\Delta S^{-1}\cdots S\Delta S^{-1} = S\Delta^n S^{-1} = S\operatorname{diag}(\lambda_1^n, \ldots, \lambda_m^n)S^{-1}.$$

Dans ce chapitre, on fixe une matrice $A \in \mathbb{C}^{m\times m}$.

3.1 Valeur propre, vecteur propre et espace propre

Définition 3.2. Une *valeur propre* de A est un nombre complexe λ pour lequel il existe un vecteur [1] $\mathbf{x} \in \mathbb{C}^m$ non nul tel que $A\mathbf{x} = \lambda\mathbf{x}$.

Définition 3.3. Un *vecteur propre* de A relatif à un nombre $\lambda \in \mathbb{C}$ est un vecteur $\mathbf{x} \in \mathbb{C}^m$ tel que $A\mathbf{x} = \lambda\mathbf{x}$. Le *sous-espace propre* relatif à $\lambda \in \mathbb{C}$ est l'ensemble des vecteurs propres de A relatifs à λ et se note $E_\lambda(A)$:

$$E_\lambda(A) = \{\mathbf{x} \in \mathbb{C}^m \colon A\mathbf{x} = \lambda\mathbf{x}\}.$$

Avec ces définitions, remarquez que λ est valeur propre de A si et seulement si A possède un vecteur propre non nul relatif à λ. Autrement dit, λ est valeur propre de A si et seulement si $E_\lambda(A) \neq \{\mathbf{0}\}$.
De plus, on vérifie aisément que $E_\lambda(A)$ est un sous-espace vectoriel de $\mathbb{C}^m$, ce qui nous autorise à poser la définition suivante.

Définition 3.4. La *multiplicité géométrique* d'un nombre $\lambda \in \mathbb{C}$ en tant que valeur propre de A est la dimension du sous-espace vectoriel $E_\lambda(A)$ sur $\mathbb{C}$.

1. En effet, nous avons vu que $\mathbb{C}^m$ est un espace vectoriel.

Remarque 3.5. La multiplicité géométrique est toujours inférieure ou égale à m. De plus, si λ est valeur propre de A, alors sa multiplicité géométrique est supérieure ou égale à 1.

Pour déterminer la multiplicité géométrique de $\lambda \in \mathbb{C}$, on cherche une base du sous-espace propre $E_\lambda(A)$ en résolvant l'équation matricielle $A\mathbf{x} = \lambda\mathbf{x}$. Le résultat suivant nous donne une méthode de calcul alternative de la multiplicité géométrique.

Proposition 3.6. *Pour tout* $\lambda \in \mathbb{C}$,

$$E_\lambda(A) = \operatorname{Ker}(A - \lambda I_m) \qquad \textit{et} \qquad \dim E_\lambda(A) = m - \operatorname{rg}(A - \lambda I_m).$$

Démonstration. Pour tout $\mathbf{x} \in \mathbb{C}^m$, on a $A\mathbf{x} = \lambda\mathbf{x} \iff (A - \lambda I_m)\mathbf{x} = \mathbf{0}$. On a donc $E_\lambda(A) = \operatorname{Ker}(A - \lambda I_m)$. La deuxième partie découle de la proposition 2.41. □

Proposition 3.7. *Des vecteurs propres non nuls relatifs à des valeurs propres distinctes sont linéairement indépendants.*

Une fois n'est pas coutume, nous proposons deux preuves de ce résultat.

Première preuve. Soient $\mathbf{x}_1, \ldots, \mathbf{x}_p$ des vecteurs propres de A non nuls relatifs à des valeurs propres deux à deux distinctes $\lambda_1, \ldots, \lambda_p$. Supposons que $\mu_1, \ldots, \mu_p \in \mathbb{C}$ sont tels que

$$\sum_{k=1}^{p} \mu_k \mathbf{x}_k = \mathbf{0}. \tag{3.1}$$

Nous devons montrer que tous les μ_k sont nécessairement nuls. Pour chaque $j \in \{1, \ldots, p\}$, on définit [2] une matrice A_j par

$$A_j = (A - \lambda_1 I_m) \cdots (\widehat{A - \lambda_j I_m}) \cdots (A - \lambda_p I_m)$$

où $(\widehat{A - \lambda_j I_m})$ signifie qu'on omet le facteur $A - \lambda_j I_m$. Pour tous $j, k \in \{1, \ldots, p\}$, on a

$$(A - \lambda_j I_m)\mathbf{x}_k = A\mathbf{x}_k - \lambda_j \mathbf{x}_k = (\lambda_k - \lambda_j)\mathbf{x}_k$$

et

$$A_j \mathbf{x}_k = (\lambda_k - \lambda_1) \cdots (\widehat{\lambda_k - \lambda_j}) \cdots (\lambda_k - \lambda_p)\mathbf{x}_k.$$

Comme les valeurs propres λ_k sont deux à deux distinctes et que les $\mathbf{x}_k$ sont des vecteurs non nuls, on en tire que $A_j \mathbf{x}_k = \mathbf{0}$ si et seulement si $j \neq k$. Pour tout $j \in \{1, \ldots, p\}$, en multipliant les deux membres de l'égalité (3.1) par A_j, on obtient que

$$\sum_{k=1}^{p} \mu_k A_j \mathbf{x}_k = \mu_j A_j \mathbf{x}_j = \mathbf{0}$$

et par conséquent que $\mu_j = 0$. □

Deuxième preuve. Soient $\mathbf{x}_1, \ldots, \mathbf{x}_p$ des vecteurs propres de A non nuls relatifs à des valeurs propres deux à deux distinctes $\lambda_1, \ldots, \lambda_p$. Nous procédons par l'absurde en supposant que $\mathbf{x}_1, \ldots, \mathbf{x}_p$ sont linéairement dépendants. On élimine un nombre maximal de vecteurs parmi $\mathbf{x}_1, \ldots, \mathbf{x}_p$ tout en conservant des vecteurs linéairement dépendant. Quitte à renuméroter nos vecteurs et les valeurs propres correspondantes, on peut supposer que les vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_n$ ont été sélectionnés. Ainsi, les vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_n$ sont linéairement dépendants mais si on

2. C'est l'astuce de cette preuve.

enlevait l'un d'eux, les vecteurs restants seraient linéairement indépendants. Il existe des nombres complexes $k_1, \ldots, k_n$ non tous nuls tels que

$$\sum_{i=1}^{n} k_i \mathbf{x}_i = \mathbf{0}. \tag{3.2}$$

Si l'un des coefficients k_i était nul, on obtiendrait une combinaison linéaire non triviale nulle des $n-1$ vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_n$, ce qui n'est pas possible puisque dans ce cas, on pourrait trouver moins de n vecteurs linéairement dépendants parmi les p vecteurs de départ. On en déduit que $k_i \neq 0$ pour tout $i \in \{1, \ldots, n\}$. En multipliant l'égalité 3.2 par λ_1, on obtient

$$\lambda_1 \sum_{i=1}^{n} k_i \mathbf{x}_i = \sum_{i=1}^{n} k_i (\lambda_1 \mathbf{x}_i) = \mathbf{0}.$$

En multipliant l'égalité 3.2 par A, on obtient

$$A \sum_{i=1}^{n} k_i \mathbf{x}_i = \sum_{i=1}^{n} k_i (A\mathbf{x}_i) = \sum_{i=1}^{n} k_i (\lambda_i \mathbf{x}_i) = \mathbf{0}.$$

On en déduit que

$$\begin{aligned}
\mathbf{0} &= \sum_{i=1}^{n} k_i (\lambda_1 \mathbf{x}_i) - \sum_{i=1}^{n} k_i (\lambda_i \mathbf{x}_i) \\
&= \sum_{i=2}^{n} k_i (\lambda_1 \mathbf{x}_i) - \sum_{i=2}^{n} k_i (\lambda_i \mathbf{x}_i) \\
&= \sum_{i=2}^{n} k_i (\lambda_1 - \lambda_i) \mathbf{x}_i.
\end{aligned}$$

Puisque les valeurs propres $\lambda_1, \ldots, \lambda_p$ sont deux à deux distinctes par hypothèses, tous les coefficients $k_i(\lambda_1 - \lambda_i)$ de la dernière somme sont non nuls. Ceci implique que $\mathbf{x}_2, \ldots, \mathbf{x}_n$ sont linéairement indépendants. Ceci n'est pas possible par choix de $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Nous avons donc trouvé une contradiction et donc les vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_p$ sont bien linéairement indépendants. □

3.2 Polynôme caractéristique

Définition 3.8. Le *polynôme caractéristique* de A est le polynôme $\det(A - XI_m)$ de $\mathbb{C}[X]$.

Exemple 3.9. Le polynôme caractéristique de la matrice

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 4 & 1 \\ 0 & i & 5 \end{pmatrix}$$

est

$$\begin{aligned}
\det(A - XI_m) &= \begin{vmatrix} 1-X & 0 & 2 \\ 0 & 4-X & 1 \\ 0 & i & 5-X \end{vmatrix} \\
&= (1-X)(4-X)(5-X) - i(1-X) \\
&= -X^3 + 10X^2 - 29X + 20 - i + iX \\
&= -X^3 + 10X^2 - (29-i)X + 20 - i.
\end{aligned}$$

Remarque 3.10. Le polynôme caractéristique d'une matrice carrée de taille m est toujours de degré m. Plus précisément, le coefficient de X^m vaut toujours $(-1)^m$. Ceci peut se voir en utilisant la multilinéarité du déterminant et en remarquant que

$$\det(A - XI_m) = \det \begin{pmatrix} C_1 - X\mathbf{e}_1 & \cdots & C_m - X\mathbf{e}_m \end{pmatrix}$$

où $C_1, \ldots, C_m$ désignent les colonnes de A et $\mathbf{e}_1, \ldots, \mathbf{e}_m$ désignent les matrices-colonnes unitaires.

On peut aussi procéder par récurrence sur m. En effet, on montre facilement que lors du calcul du déterminant, un élément n'est jamais multiplié par lui-même. Le polynôme caractéristique est donc de degré au plus m et le cofacteur $\mathcal{C}_{mi}(A - XI_m)$ est un polynôme de degré au plus $m-2$ si $i \neq m$ (puisque l'indéterminée X apparaît $m-2$ fois dans le mineur correspondant). Le polynôme $\mathcal{C}_{mm}(A - XI_m)$ est le polynôme caractéristique d'une matrice de taille $m-1$ (plus précisément de la matrice A privée de sa dernière ligne et de sa dernière colonne). Par hypothèse de récurrence, il est de degré $m-1$ et son coefficient dominant vaut $(-1)^{m-1}$. En utilisant la loi des cofacteurs sur la dernière ligne de $A - XI_m$, on obtient que

$$\det(A - XI_m) = \sum_{i=1}^{m-1} A_{mi} \underbrace{\mathcal{C}_{mi}(A - XI_m)}_{\substack{\text{polynôme} \\ \text{de degré} \\ \leq m-2}} + \quad (A_{mm} - X) \underbrace{\mathcal{C}_{mm}(A - XI_m)}_{\substack{\text{polynôme} \\ \text{de degré } m-1 \\ \text{par HR}}}. \tag{3.3}$$

Ainsi, le coefficient de X^m dans le polynôme caractéristique est égal à $-(-1)^{m-1} = (-1)^m$.

Proposition 3.11. *Un nombre complexe est valeur propre d'une matrice carrée complexe si et seulement s'il est zéro de son polynôme caractéristique.*

Démonstration. Un nombre $\lambda \in \mathbb{C}$ est valeur propre de A si et seulement si le système linéaire homogène

$$(A - \lambda I_m)\mathbf{x} = \mathbf{0}$$

possède une solution non nulle. Grâce à la théorie des systèmes linéaires [3], nous savons que ceci a lieu si et seulement si le système n'est pas de Cramer, c'est-à-dire si et seulement si la matrice $A - \lambda I_m$ n'est pas inversible. Or une matrice est inversible si et seulement si son déterminant est non nul. On obtient donc que λ est valeur propre de A si et seulement si $\det(A - \lambda I_m) = 0$, c'est-à-dire si et seulement si λ annule le polynôme caractéristique de A. □

Définition 3.12. La *multiplicité algébrique* d'un nombre $\lambda \in \mathbb{C}$ est sa multiplicité en tant que zéro du polynôme caractéristique de A.

Remarquons qu'avec cette définition, un nombre complexe est valeur propre de A si et seulement si sa multiplicité algébrique vaut au moins 1. Une valeur propre *simple* est une valeur propre dont la multiplicité algébrique vaut 1.

Remarque 3.13. Poursuivant la remarque 3.10, une matrice carrée de taille m possède donc exactement m valeurs propres répétées selon leurs multiplicités algébriques. Autrement dit, si $\lambda_1, \ldots, \lambda_p$ sont les valeurs propres de A et si $\alpha_1, \ldots, \alpha_p$ en sont les multiplicités algébriques respectives, alors $\alpha_1 + \cdots + \alpha_p = m$.

Proposition 3.14. *Pour toute matrice inversible $S \in \mathbb{C}^{m \times m}$, les matrices A et $S^{-1}AS$ ont même polynôme caractéristique. En particulier, elles possèdent les mêmes valeurs propres avec les mêmes multiplicités algébriques.*

3. N'oubliez pas votre cours "Mathématiques pour l'informatique 1".

Démonstration. Pour rappel, si A et B sont des matrices carrées de même taille, on a toujours $\det(AB) = \det(BA)$ [4]. En utilisant ce fait, nous avons successivement

$$\begin{aligned}\det(S^{-1}AS - XI_m) &= \det(S^{-1}(A - XI_m)S)\\ &= \det(S^{-1})\det(A - XI_m)\det(S)\\ &= \det(S^{-1})\det(S)\det(A - XI_m)\\ &= \det(S^{-1}S)\det(A - XI_m)\\ &= \det(I_m)\det(A - XI_m)\\ &= \det(A - XI_m).\end{aligned}$$

Ainsi, les matrices A et $S^{-1}AS$ possèdent le même polynôme caractéristique. Le cas particulier est immédiat. □

Proposition 3.15. *La multiplicité géométrique d'une valeur propre est inférieure à sa multiplicité algébrique.*

Démonstration. Soit λ une valeur propre de A. Notons $d = \dim E_\lambda(A)$. Alors il existe d vecteurs propres $\mathbf{x}_1, \ldots, \mathbf{x}_d$ de A formant une base de $E_\lambda(A)$. Puisque $\dim \mathbb{C}^m = m$, le théorème de complétion des bases nous dit qu'il existe $m - d$ vecteurs $\mathbf{x}_{d+1}, \ldots, \mathbf{x}_m$ qui complètent cette base pour former une base de $\mathbb{C}^m$. Soit $S = (\mathbf{x}_1 \cdots \mathbf{x}_m)$ la matrice dont les colonnes sont les vecteurs $\mathbf{x}_i$. Montrons que pour chaque $i \in \{1, \ldots, d\}$, la i-ième colonne de $S^{-1}AS$ est $\lambda\mathbf{e}_i$. En effet, supposons que $i \in \{1, \ldots, d\}$ et que la i-ième colonne de $S^{-1}AS$ soit $\mathbf{c}_i = (\mu_1 \cdots \mu_m)^T$. Alors

$$\lambda\mathbf{x}_i = A\mathbf{x}_i = S\mathbf{c}_i = \sum_{k=1}^{m} \mu_k \mathbf{x}_k.$$

Puisque les vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_m$ sont linéairement indépendants, on obtient que $\mu_i = \lambda$ et $\mu_k = 0$ pour $k \neq i$. Ainsi, $S^{-1}AS$ est une matrice triangulaire supérieure de la forme

$$S^{-1}AS = \begin{pmatrix} \lambda I_d & * \\ 0 & B \end{pmatrix}$$

où B est une matrice carrée de taille $m - d$ et $*$ représente des entrées quelconques. On obtient que

$$\det(A - XI_m) = \det(S^{-1}AS - XI_m) = (\lambda - X)^d \det(B - XI_{m-d}).$$

La valeur propre λ a donc une multiplicité algébrique au moins égale à d. □

Voici une conséquence intéressante de la proposition 3.11 qui lie le déterminant et la trace d'une matrice carrée à ses valeurs propres.

Définition 3.16. La *trace* de A, notée $\operatorname{tr}(A)$, est la somme de ses éléments diagonaux :

$$\operatorname{tr}(A) = \sum_{i=1}^{m} A_{ii}.$$

Avant de faire le lien entre trace et valeurs propres, nous établissont d'abord une propriété de la trace.

Théorème 3.17. *Pour toutes matrices A et B carrée de même taille, on a* $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

4. Voir le cours de Mathématiques pour l'informatique 1

Démonstration. Soient $A, B \in \mathbb{C}^{m\times m}$. D'une part, on a

$$\operatorname{tr}(AB) = \sum_{i=1}^{m} (AB)_{ii} = \sum_{i=1}^{m}\sum_{j=1}^{m} A_{ij}B_{ji}.$$

D'autre part, on a

$$\operatorname{tr}(BA) = \sum_{j=1}^{m} (BA)_{jj} = \sum_{j=1}^{m}\sum_{i=1}^{m} B_{ji}A_{ij}.$$

D'où la conclusion. □

En conséquence, des matrices semblables ont la même trace.

Corollaire 3.18. *Pour toute matrice inversible $S \in \mathbb{C}^{m\times m}$, les matrices A et $S^{-1}AS$ ont même trace.*

Démonstration. En effet, on a $\operatorname{tr}(S^{-1}AS) = \operatorname{tr}(ASS^{-1}) = \operatorname{tr}(A)$. □

Proposition 3.19. *Soient $\lambda_1, \dots, \lambda_m$ les valeurs propres de A répétées selon leurs multiplicités algébriques. On a*

$$\det(A) = \prod_{i=1}^{m} \lambda_i \quad \textit{et} \quad \operatorname{tr}(A) = \sum_{i=1}^{m} \lambda_i.$$

Démonstration. Par la remarque 3.10, la proposition 3.11, et le théorème fondamental de l'algèbre, nous savons que le polynôme caractéristique se factorise comme suit :

$$\det(A - XI_m) = (-1)^m (X - \lambda_1)\cdots(X-\lambda_m) = (\lambda_1 - X)\cdots(\lambda_m - X).$$

D'une part, le terme indépendant du polynôme caractéristique est le produit des valeurs propres. Le terme indépendant d'un polynôme P étant égal à $P(0)$, on obtient que

$$\prod_{i=1}^{m} \lambda_i = (\det(A - XI_m))(0) = \det(A).$$

D'autre part, le coefficient de X^{m-1} est, au signe près, donné par la somme des valeurs propres : il vaut $(-1)^{m-1}(\lambda_1 + \cdots + \lambda_m)$. Pour conclure, il nous reste à montrer que ce coefficient est aussi égal à $(-1)^{m-1}\operatorname{tr}(A)$. De manière équivalente, nous montrons que le coefficient de Y^{m-1} du polynôme $\det(A + YI_m)$ est égal à $\operatorname{tr}(A)$.
Nous procédons par récurrence sur m. Pour $m = 1$, on a $\det(A + YI_1) = A_{11} + Y$. Dans ce cas, le coefficient de Y^{m-1} est le terme indépendant, c'est-à-dire $A_{11} = \operatorname{tr}(A)$.
Supposons à présent que $m \geq 2$ et que la thèse est vérifiée pour les matrices de taille $m-1$. En utilisant la loi des mineurs sur la dernière ligne, on obtient que

$$\det(A + YI_m) = \sum_{i=1}^{m-1} A_{mi} \underbrace{\mathcal{C}_{mi}(A + YI_m)}_{\substack{\text{polynôme} \\ \text{de degré} \\ \leq m-2}} + \quad (A_{mm} + Y) \underbrace{\mathcal{C}_{mm}(A + YI_m)}_{\substack{\text{polynôme unitaire} \\ \text{de degré } m-1}}.$$

Par hypothèse de récurrence, le coefficient de Y^{m-2} dans le polynôme $\mathcal{C}_{mm}(A + YI_m)$ est égal à $\sum_{i=0}^{m-1} A_{ii}$. On obtient que le coefficient de Y^{m-1} de $\det(A + YI_m)$ est

$$A_{mm} + \sum_{i=0}^{m-1} A_{ii} = \sum_{i=0}^{m} A_{ii} = \operatorname{tr}(A),$$

comme souhaité. □

3.3 Théorème de Cayley-Hamilton et polynôme minimal

Définition 3.20. Soit P un polynôme à coefficients complexes. La notation $P(A)$ désigne la matrice complexe obtenue en substituant X^i par A^i dans P, pour tout $i \in \{0, \ldots, \deg(P)\}$ et en exécutant les opérations $+$ et $\cdot$ dans $\mathbb{C}^{m\times m}$: si

$$P = p_0 + p_1 X + p_2 X^2 + \cdots + p_d X^d,$$

alors

$$P(A) = p_0 I_m + p_1 A + p_2 A^2 + \cdots + p_d A^d.$$

Remarquez que le terme indépendant p_0 est vu comme étant $p_0 X^0$, et est donc remplacé par $p_0 A^0 = p_0 I_m$.

Exemple 3.21. Si $P = X^2 + 3X - 2$ et $A = \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix}$, alors

$$\begin{aligned} P(A) &= A^2 + A - 2I_2 \\ &= \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix}^2 + 3\begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 0 \\ 5 & 9 \end{pmatrix} + \begin{pmatrix} 6 & 0 \\ 3 & 9 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 8 & 0 \\ 8 & 16 \end{pmatrix}. \end{aligned}$$

Nous admettons le résultat suivant.

Théorème 3.22 (Cayley-Hamilton)**.** *Toute matrice carrée complexe annule son polynôme caractéristique. Autrement dit, si* $P = \det(A - XI_m)$*, alors* $P(A) = 0$.

Attention à ne pas faire l'erreur d'écrire $P(A) = \det(A - AI_m) = \det(0) = 0$. En effet, ce raisonnement est faux car on ne peut intervertir les étapes d'évaluation du déterminant et de substitution de X. Quand on écrit (correctement cette fois) $P(A) = \det(A - XI_m)(A)$, ce que l'on fait est de *d'abord* calculer le polynôme $P = \det(A - XI_m)$ dans $\mathbb{C}[X]$ et *ensuite* de calculer $P(A)$ dans $\mathbb{C}^{m\times m}$ en substituant l'indéterminée X par la matrice A.

Exemple 3.23. Poursuivons l'exemple 3.9. On a

$$A^2 = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 4 & 1 \\ 0 & i & 5 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 4 & 1 \\ 0 & i & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2i & 12 \\ 0 & 16+i & 9 \\ 0 & 9i & 25+i \end{pmatrix}$$

et

$$A^3 = \begin{pmatrix} 1 & 2i & 12 \\ 0 & 16+i & 9 \\ 0 & 9i & 25+i \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 4 & 1 \\ 0 & i & 5 \end{pmatrix} = \begin{pmatrix} 1 & 20i & 62+2i \\ 0 & 64+13i & 61+i \\ 0 & -1+61i & 125+14i \end{pmatrix}.$$

Calculons $P(A)$ pour le polynôme caractéristique $P = -X^3 + 10X^2 - (29 - i)X +$ de A : on a

$$\begin{aligned} &-A^3 + 10A^2 - (29 - i)A + (20 - i)I_3 \\ &= -\begin{pmatrix} 1 & 20i & 62+2i \\ 0 & 64+13i & 61+i \\ 0 & -1+61i & 125+14i \end{pmatrix} + 10\begin{pmatrix} 1 & 2i & 12 \\ 0 & 16+i & 9 \\ 0 & 9i & 25+i \end{pmatrix} \end{aligned}$$

$$- (29-i)\begin{pmatrix}1&0&2\\0&4&1\\0&i&5\end{pmatrix} + (20-i)\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

$$= \begin{pmatrix} -1+10-29+i+20-i & -20i+20i & -62-2i+120-58+2i \\ 0 & -64-13i+160+10i-116+4i+20-i & -61-i+90-29+i \\ 0 & 1-61i+90i-29i-1 & -125-14i+250+10i-145+5i+20-i \end{pmatrix}$$

$$= \begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}.$$

Définition 3.24. Un *polynôme minimal de A* est un polynôme non nul de $\mathbb{C}[X]$ de coefficient dominant 1 annulé par A et de degré minimal parmi tous les polynômes non nuls annulés par A.

Remarquons que si $P(A) = 0$ alors pour tout $c \in \mathbb{C}_0$, on a $(cP)(A) = cP(A) = 0$. D'où l'importance de préciser qu'un polynôme minimal a un coefficient dominant égal à 1 dans la définition précédente.

Remarquons aussi qu'il pourrait a priori exister plusieurs polynômes de même degré minimal annulés par A. Nous allons montrer qu'en fait, A admet un unique polynôme minimal. Ceci nous permettra de parler *du* polynôme minimal de A, et non pas juste d'*un* polynôme minimal de A.

Lemme 3.25. *Tout polynôme minimal de A divise tout polynôme de $\mathbb{C}[X]$ annulé par A.*

Démonstration. Soit π un polynôme minimal de A et soit P un polynôme annulé par A. Notre but est de montrer que π divise P, ce qui revient à montrer que le reste de la division euclidienne de P par π vaut 0. Soient donc Q et R les quotient et reste de la division euclidienne de P par π. On a $P = Q\pi + R$ et $\deg(R) < \deg(\pi)$. En évaluant les deux membres de cette égalité en A, on obtient $0 = P(A) = (Q\pi + R)(A) = Q(A)\pi(A) + R(A) = R(A)$. Donc R est un polynôme de degré strictement plus petit que celui de π et annulé par A. Puisque π est un polynôme minimal de A, ceci implique que $R = 0$, et donc que π divise P. □

Proposition 3.26. *Il existe un unique polynôme minimal de A.*

Démonstration. L'existence d'un polynôme minimal est assurée par le théorème de Cayley-Hamilton. L'unicité, quant à elle, est assurée par le lemme précédent. En effet, supposons que P et Q soient des polynômes minimaux de A. Au vu du lemme précédent, P doit diviser Q. Puisque P et Q sont de même degré et de coefficient dominant 1, on obtient que nécessairement $P = Q$. □

En conséquence de la proposition précédente, nous parlerons désormais *du* polynôme minimal d'une matrice A plutôt que d'*un* polynôme minimal de A.

Proposition 3.27. *Le polynôme minimal de A divise le polynôme caractéristique de A.*

Démonstration. C'est une conséquence directe du théorème de Cayley-Hamilton et du lemme 3.25. □

En conséquence de la proposition précédente, tout zéro du polynôme minimal est aussi zéro du polynôme caractéristique, et donc les zéros du polynôme minimal sont des valeurs propres. Notre but à présent est de montrer que la réciproque est vraie aussi : toutes les valeurs propres annulent le polynôme minimal.

Lemme 3.28. *Soient $P \in \mathbb{C}[X]$ et λ une valeur propre de A. Pour tout $\mathbf{x} \in E_\lambda$, on a $P(A)\mathbf{x} = P(\lambda)\mathbf{x}$.*

Démonstration. D'abord, remarquons que pour tout $i \in \mathbb{N}$, on a $A^i\mathbf{x} = \lambda^i\mathbf{x}$ (ceci s'obtient par récurrence sur i). Maintenant, si $P = p_0 + p_1X + p_2X^2 \cdots + p_dX^d$, alors

$$\begin{aligned} P(A)\mathbf{x} &= (p_0I_m + p_1A + p_2A^2 \cdots + p_dA^d)\mathbf{x} \\ &= p_0I_m\mathbf{x} + p_1A\mathbf{x} + p_2A^2\mathbf{x} \cdots + p_dA^d\mathbf{x} \\ &= p_0\mathbf{x} + p_1\lambda\mathbf{x} + p_2\lambda^2\mathbf{x} \cdots + p_d\lambda^d\mathbf{x} \\ &= (p_0 + p_1\lambda + p_2\lambda^2 \cdots + p_d\lambda^d)\mathbf{x} \\ &= P(\lambda)\mathbf{x}. \end{aligned}$$

□

Exemple 3.29. Illustrons le lemme précédent avec la matrice

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$$

et le polynôme $P = X^3 + X^2 + X + 1$. On calcule

$$\det(A - XI_2) = \begin{vmatrix} 1-X & 3 \\ 3 & 1-X \end{vmatrix} = \begin{vmatrix} 4-X & 3 \\ 4-X & 1-X \end{vmatrix} = (4-X)(1-X-3) = (4-X)(-2-X).$$

Les valeurs propres de A sont donc 4 et -2. On a aussi

$$A^2 = \begin{pmatrix} 10 & 6 \\ 6 & 10 \end{pmatrix} \quad \text{et} \quad A^3 = \begin{pmatrix} 28 & 36 \\ 36 & 28 \end{pmatrix}.$$

On obtient que

$$P(A) = A^3 + A^2 + A + I = \begin{pmatrix} 40 & 45 \\ 45 & 40 \end{pmatrix}.$$

Considérons d'abord la valeur propre 4. On peut vérifier que

$$E_4(A) = \left\{ \begin{pmatrix} a \\ a \end{pmatrix} : a \in \mathbb{C} \right\}.$$

On a $P(4) = 4^3 + 4^2 + 4^1 + 4^0 = 64 + 16 + 4 + 1 = 85$ et pour tout $a \in \mathbb{C}$,

$$P(A)\begin{pmatrix} a \\ a \end{pmatrix} = \begin{pmatrix} 40 & 45 \\ 45 & 40 \end{pmatrix}\begin{pmatrix} a \\ a \end{pmatrix} = \begin{pmatrix} 85a \\ 85a \end{pmatrix} = 85\begin{pmatrix} a \\ a \end{pmatrix} = P(4)\begin{pmatrix} a \\ a \end{pmatrix}.$$

Considérons maintenant la valeur propre -2. On vérifie que

$$E_{-2}(A) = \left\{ \begin{pmatrix} a \\ -a \end{pmatrix} : a \in \mathbb{C} \right\}.$$

On a $P(-2) = (-2)^3 + (-2)^2 + (-2)^1 + (-2)^0 = -8 + 4 - 2 + 1 = -5$ et pour tout $a \in \mathbb{C}$,

$$P(A)\begin{pmatrix} a \\ -a \end{pmatrix} = \begin{pmatrix} 40 & 45 \\ 45 & 40 \end{pmatrix}\begin{pmatrix} a \\ -a \end{pmatrix} = \begin{pmatrix} -5a \\ 5a \end{pmatrix} = -5\begin{pmatrix} a \\ -a \end{pmatrix} = P(-2)\begin{pmatrix} a \\ -a \end{pmatrix}.$$

Proposition 3.30. *Les zéros du polynôme minimal de A sont les valeurs propres de A.*

Démonstration. Notons π le polynôme minimal de A.

Par la proposition 3.27, nous savons que tout zéro de π est un zéro du polynôme caractéristique de A. En utilisant la proposition 3.11, nous obtenons que tout zéro de π est valeur propre de A.

Il nous reste à montrer que la réciproque est vraie aussi, c'est-à-dire que toute valeur propre de A est zéro de π. Soit λ une valeur propre de A et soit $\mathbf{x}$ un vecteur propre non nul de A relatif à λ. Par le lemme 3.28, on a que $\pi(\lambda)\mathbf{x} = \pi(A)\mathbf{x} = \mathbf{0}$. Mais puisque $\mathbf{x} \neq \mathbf{0}$, ceci implique que $\pi(\lambda) = 0$. □

La proposition précédente nous permet de calculer facilement le polynôme minimal d'une matrice.

Exemple 3.31. Soit

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

Le polynôme caractéristique de A est $(2 - X)(1 - X)^2 = -X^3 + 4X^2 - 5X + 2$. Par la proposition 3.30, on sait que le polynôme $(X - 2)(X - 1)$ divise le polynôme minimal. De plus, par la proposition 3.27, on sait que le polynôme minimal divise $(X-2)(X-1)^2$. Ainsi, les seuls choix possibles pour le polynôme minimal sont $(X-2)(X-1)$ et $(X-2)(X-1)^2$. Puisque

$$(A - 2I_3)(A - I_3) = \begin{pmatrix} 0 & 0 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

on conclut que le polynôme minimal est égal à $(X - 2)(X - 1) = X^2 - 3X + 2$.

Corollaire 3.32. *Si toutes les valeurs propres de A sont simples, alors les polynômes minimal et caractéristique de A coïncident (au signe près).*

Démonstration. Supposons que A possède m valeurs propres simples $\lambda_1, \ldots, \lambda_m$. Alors le polynôme caractéristique de A est égal à $(\lambda_1 - X) \cdots (\lambda_m - X)$ par la proposition 3.11. Par la proposition 3.30, nous savons que le polynôme $(X - \lambda_1) \cdots (X - \lambda_m)$ divise le polynôme minimal de A, qui lui-même divise le polynôme caractéristique de A par la proposition 3.27. D'où la conclusion. □

Exemple 3.33. Soit

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

Le polynôme caractéristique de A est $(2 - X)(1 - X)(3 - X) = -X^3 + 6X^2 - 11X + 6$. La matrice A possède donc trois valeurs propres simples. Par le corollaire 3.32, le polynôme minimal est égal à $(X - 2)(X - 1)(X - 3) = X^3 - 6X^2 + 11X - 6$.

3.4 Diagonalisation

La preuve du théorème suivant est aussi importante que son énoncé. En effet, l'énoncé nous donne des critères pour déterminer si une matrice donnée (en supposant qu'on puisse en déterminer les valeurs propres) est ou n'est pas diagonalisable. Lorsque c'est le cas, la preuve du théorème nous donne une méthode effective de diagonalisation, c'est-à-dire de construction d'une matrice S diagonalisant A.

Théorème 3.34 (Critères de diagonalisation)**.** *Les conditions suivantes sont équivalentes.*

1. *A est diagonalisable.*
2. *A possède m vecteurs propres linéairement indépendants.*
3. *La somme des multiplicités géométriques des valeurs propres de A vaut m.*
4. *Les multiplicités algébriques et géométriques de chaque valeur propre de A sont égales.*

Démonstration. Soient $\lambda_1, \ldots, \lambda_p$ les valeurs propres de A, soient $\alpha_1, \ldots, \alpha_p$ leurs multiplicités algébriques respectives et soient $d_1, \ldots, d_p$ leurs multiplicités géométriques respectives. Le polynôme caractéristique étant de degré m, on a $\alpha_1 + \cdots + \alpha_p = m$.

$3 \iff 4$. L'implication $4 \implies 3$ est immédiate. L'implication $3 \implies 4$ découle de la proposition 3.15. En effet, s'il existe $i \in \{1, \ldots, p\}$ tel que $d_i \neq \alpha_i$, alors $d_i < \alpha_i$ et $d_1 + \cdots + d_p < \alpha_1 + \cdots + \alpha_p = m$.

$3 \implies 2$. C'est une conséquence de la proposition 3.7. En effet, supposons que $d_1 + \cdots + d_p = m$. Pour tout $i \in \{1, \ldots, p\}$, soit $(\mathbf{x}_{i,1}, \ldots, \mathbf{x}_{i,d_i})$ une base de $E_{\lambda_i}(A)$. Montrons que les m vecteurs $\mathbf{x}_{1,1}, \ldots, \mathbf{x}_{1,d_1}, \ldots, \mathbf{x}_{p,1}, \ldots, \mathbf{x}_{p,d_p}$ sont linéairement indépendants. Soient $k_{1,1}, \ldots, k_{1,d_1}, \ldots, k_{p,1}, \ldots, k_{p,d_p} \in \mathbb{C}$ tels que

$$\underbrace{k_{1,1}\mathbf{x}_{1,1} + \cdots + k_{1,d_1}\mathbf{x}_{1,d_1}}_{=\mathbf{y}_1} + \cdots + \underbrace{k_{p,1}\mathbf{x}_{p,1} + \cdots + k_{p,d_p}\mathbf{x}_{p,d_p}}_{=\mathbf{y}_p} = \mathbf{0}.$$

Pour tout $i \in \{1, \ldots, p\}$, notons $\mathbf{y}_i = k_{i,1}\mathbf{x}_{i,1} + \cdots + k_{i,d_i}\mathbf{x}_{1,d_1}$. Chaque vecteur $\mathbf{y}_i$ ainsi défini est un vecteur propre relatif à la valeur propre λ_i. Par la proposition 3.7, on obtient que les vecteurs $\mathbf{y}_i$ sont tous nuls. Or, pour tout $i \in \{1, \ldots, p\}$, les vecteurs $\mathbf{x}_{i,1}, \ldots, \mathbf{x}_{i,d_i}$ sont linéairement indépendants, ce qui implique que $k_{i,1} = \cdots = k_{i,d_i} = 0$.

Nous faisons à présent une observation importante, qui nous servira dans la suite de la démonstration. Si $S = (\mathbf{c}_1 \cdots \mathbf{c}_m)$ est inversible et $\mu_1, \ldots, \mu_m \in \mathbb{C}$, alors

$$\begin{aligned} S^{-1}AS = \operatorname{diag}(\mu_1, \ldots, \mu_m) &\iff AS = S\operatorname{diag}(\mu_1, \ldots, \mu_m) \\ &\iff A\mathbf{c}_j = \mu_j\mathbf{c}_j \text{ pour tout } j \in \{1, \ldots, m\}. \end{aligned} \tag{3.4}$$

$1 \implies 2$. Supposons que A soit diagonalisable par une matrice S. Au vu de (3.4), les colonnes de S sont des vecteurs propres de A. Puisque S est inversible, ses colonnes sont linéairement indépendantes et A possède donc m vecteurs propres linéairement indépendants.

$2 \implies 1 \wedge 3$. Supposons que A possède m vecteurs propres linéairement indépendants. Notons ces vecteurs propres

$$\mathbf{x}_{1,1}, \ldots, \mathbf{x}_{1,m_1}, \ldots, \mathbf{x}_{p,1}, \ldots, \mathbf{x}_{p,m_p}$$

où l'on suppose que pour chaque $i \in \{1, \ldots, p\}$ et chaque $\ell \in \{1, \ldots, m_i\}$, le vecteur $\mathbf{x}_{i,\ell}$ est un vecteur propre de A relatif à la valeur propre λ_i. On a donc $m_1 + \cdots + m_p = m$. Soit S la matrice dont les colonnes sont les vecteurs $\mathbf{x}_{i,\ell}$:

$$S = (\mathbf{x}_{1,1} \cdots \mathbf{x}_{1,m_1} \cdots \mathbf{x}_{p,1} \cdots \mathbf{x}_{p,m_p}).$$

Alors, en utilisant (3.4) à nouveau [5], on obtient que

$$S^{-1}AS = \operatorname{diag}(\underbrace{\lambda_1, \ldots, \lambda_1}_{m_1 \text{ fois}}, \ldots, \underbrace{\lambda_p, \ldots, \lambda_p}_{m_p \text{ fois}})$$

5. Mais cette fois, on utilise l'autre direction.

et donc que A est diagonalisable[6]. Par définition des multiplicités géométriques et par la proposition 3.15, on a $m_i \leq d_i \leq \alpha_i$ pour tout $i \in \{1, \dots, p\}$. D'où

$$m = m_1 + \cdots + m_p \leq d_1 + \cdots + d_p \leq \alpha_1 + \cdots + \alpha_p = m$$

et donc $d_1 + \cdots + d_p = m$. □

Corollaire 3.35. *Si A possède uniquement des valeurs propres simples, alors A est diagonalisable.*

Démonstration. En effet, dans ce cas, toutes les multiplicités algébriques et géométriques sont égales à 1. □

Exemple 3.36. Soit

$$A = \begin{pmatrix} \frac{3}{4} & -\frac{45}{4} & 1 & 0 \\ -\frac{1}{4} & -\frac{1}{4} & -1 & 0 \\ \frac{1}{4} & -\frac{3}{4} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} \end{pmatrix}.$$

Pour savoir si A est diagonalisable, on teste si les multiplicités géométriques et algébriques de chaque valeur propre coïncident. Pour cela, il nous faut d'abord déterminer les zéros du polynôme caractéristique de A. On a

$$\begin{aligned}
\det(A - XI_4) \quad &= \quad \begin{vmatrix} \frac{3}{4} - X & -\frac{45}{4} & 1 & 0 \\ -\frac{1}{4} & -\frac{1}{4} - X & -1 & 0 \\ \frac{1}{4} & -\frac{3}{4} & -X & 0 \\ 0 & 0 & 0 & \frac{1}{3} - X \end{vmatrix} \\
&= \quad \Big(\frac{1}{3} - X\Big) \begin{vmatrix} \frac{3}{4} - X & -\frac{45}{4} & 1 \\ -\frac{1}{4} & -\frac{1}{4} - X & -1 \\ \frac{1}{4} & -\frac{3}{4} & -X \end{vmatrix} \\
&\underset{L_2 \leftarrow L_2 + L_3}{=} \quad \Big(\frac{1}{3} - X\Big) \begin{vmatrix} \frac{3}{4} - X & -\frac{45}{4} & 1 \\ 0 & -1 - X & -1 - X \\ \frac{1}{4} & -\frac{3}{4} & -X \end{vmatrix} \\
&= \quad \Big(\frac{1}{3} - X\Big)(-1 - X) \begin{vmatrix} \frac{3}{4} - X & -\frac{45}{4} & 1 \\ 0 & 1 & 1 \\ \frac{1}{4} & -\frac{3}{4} & -X \end{vmatrix} \\
&\underset{C_3 \leftarrow C_3 - C_2}{=} \quad \Big(\frac{1}{3} - X\Big)(-1 - X) \begin{vmatrix} \frac{3}{4} - X & -\frac{45}{4} & \frac{49}{4} \\ 0 & 1 & 0 \\ \frac{1}{4} & -\frac{3}{4} & \frac{3}{4} - X \end{vmatrix} \\
&= \quad \Big(\frac{1}{3} - X\Big)(-1 - X) \begin{vmatrix} \frac{3}{4} - X & \frac{49}{4} \\ \frac{1}{4} & \frac{3}{4} - X \end{vmatrix} \\
&= \quad \Big(\frac{1}{3} - X\Big)(-1 - X)\Big(\Big(\frac{3}{4} - X\Big)^2 - \frac{49}{16}\Big) \\
&= \quad \Big(\frac{1}{3} - X\Big)(-1 - X)\Big(\Big(\frac{3}{4} - X\Big) - \frac{7}{4}\Big)\Big(\Big(\frac{3}{4} - X\Big) + \frac{7}{4}\Big) \\
&= \quad \Big(\frac{1}{3} - X\Big)(-1 - X)^2\Big(\frac{5}{2} - X\Big).
\end{aligned}$$

Les valeurs propres de A sont donc $\frac{1}{3}, -1$ et $\frac{5}{2}$ de multiplicités algébriques respectives $1, 2$ et 1. On en déduit que les multiplicités géométriques de $\frac{1}{3}$ et $\frac{5}{2}$ valent toutes deux 1. Intéressons-nous à la multiplicité géométrique de -1. A priori, celle-ci peut valoir 1 ou 2. Pour tous

6. Retenons comment a été construite la matrice S qui diagonalise A : ses colonnes sont des vecteurs propres linéairement indépendants.

$a, b, c, d \in \mathbb{C}$, on a

$$\begin{pmatrix} \frac{3}{4} & -\frac{45}{4} & 1 & 0 \\ -\frac{1}{4} & -\frac{1}{4} & -1 & 0 \\ \frac{1}{4} & -\frac{3}{4} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = - \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \iff \begin{pmatrix} \frac{7}{4} & -\frac{45}{4} & 1 & 0 \\ -\frac{1}{4} & \frac{3}{4} & -1 & 0 \\ \frac{1}{4} & -\frac{3}{4} & 1 & 0 \\ 0 & 0 & 0 & \frac{4}{3} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\iff \left\{ \begin{array}{rcl} \frac{7}{4}a - \frac{45}{4}b + c & = & 0 \\ \frac{1}{4}a - \frac{3}{4}b + c & = & 0 \\ \frac{4}{3}d & = & 0 \end{array} \right.$$

$$\iff \left\{ \begin{array}{rcl} \frac{3}{2}a - \frac{21}{2}b & = & 0 \\ \frac{1}{4}a - \frac{3}{4}b + c & = & 0 \\ d & = & 0 \end{array} \right.$$

$$\iff \left\{ \begin{array}{rcl} a & = & 7b \\ c & = & -b \\ d & = & 0. \end{array} \right.$$

De ceci, on obtient que

$$E_{-1}(A) = \left\{ \begin{pmatrix} 7b \\ b \\ -b \\ 0 \end{pmatrix} : b \in \mathbb{C} \right\} = \left\{ b \begin{pmatrix} 7 \\ 1 \\ -1 \\ 0 \end{pmatrix} : b \in \mathbb{C} \right\} = \Big\rangle \begin{pmatrix} 7 \\ 1 \\ -1 \\ 0 \end{pmatrix} \Big\langle$$

La multiplicité géométrique de -1 est $\dim E_{-1}(A) = 1$ et diffère de sa multiplicité algébrique, qui est 2. Par le théorème 3.34, on obtient que A n'est pas une matrice diagonalisable.

Exemple 3.37. Soit

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

Le polynôme caractéristique de A est

$$\begin{aligned} \det(A - XI_6) &= \begin{vmatrix} -X & i & 0 & 0 & 0 & 0 \\ i & -X & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-X & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-X & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-X & 0 \\ 0 & 0 & 0 & 0 & 0 & 1-X \end{vmatrix} \\ &= (X^2+1)(1-X)^4 \\ &= (X-i)(X+i)(1-X)^4. \end{aligned}$$

Les valeurs propres de A sont donc i, $-i$ et 1 de multiplicités algébriques respectives $1, 1$ et 4. On en déduit que les multiplicités géométriques de i et $-i$ valent toutes deux 1. Intéressons-nous à la multiplicité géométrique de 1. A priori, celle-ci est comprise entre 1 et 4. Pour $a, b, c, d, e, f \in \mathbb{C}$, on a

$$\begin{pmatrix} 0 & i & 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} \iff \begin{pmatrix} -1 & i & 0 & 0 & 0 & 0 \\ i & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\iff \left\{\begin{array}{rcl} -a+ib & = & 0 \\ ia-b & = & 0 \end{array}\right.$$

$$\iff \left\{\begin{array}{rcl} a & = & 0 \\ b & = & 0 \end{array}\right.$$

De ceci, on obtient que

$$\begin{aligned} E_1(A) &= \left\{ \begin{pmatrix} 0\\0\\c\\d\\e\\f \end{pmatrix} : c,d,e,f \in \mathbb{C} \right\} \\ &= \left\{ c\begin{pmatrix} 0\\0\\1\\0\\0\\0 \end{pmatrix} + d\begin{pmatrix} 0\\0\\0\\1\\0\\0 \end{pmatrix} + e\begin{pmatrix} 0\\0\\0\\0\\1\\0 \end{pmatrix} + f\begin{pmatrix} 0\\0\\0\\0\\0\\1 \end{pmatrix} : c,d,e,f \in \mathbb{C} \right\} \\ &= \Bigg\rangle \begin{pmatrix} 0\\0\\1\\0\\0\\0 \end{pmatrix}, \begin{pmatrix} 0\\0\\0\\1\\0\\0 \end{pmatrix}, \begin{pmatrix} 0\\0\\0\\0\\1\\0 \end{pmatrix}, \begin{pmatrix} 0\\0\\0\\0\\0\\1 \end{pmatrix} \Bigg\langle \end{aligned}$$

et la multiplicité géométrique de 1 est $\dim E_1(A) = 4$. Par le théorème 3.34, on obtient que A est une matrice diagonalisable. Pour obtenir une matrice S qui diagonalise A, il suffit de trouver 6 vecteurs propres linéairement indépendants. Des calculs qui précèdent, on sait déjà que

$$\begin{pmatrix} 0\\0\\1\\0\\0\\0 \end{pmatrix}, \begin{pmatrix} 0\\0\\0\\1\\0\\0 \end{pmatrix}, \begin{pmatrix} 0\\0\\0\\0\\1\\0 \end{pmatrix}, \begin{pmatrix} 0\\0\\0\\0\\0\\1 \end{pmatrix}$$

sont des vecteurs propres relatifs à la valeur propre 1 linéairement indépendants. Il nous faut encore trouver un vecteur propre relatif à la valeur propre i et un autre relatif à la valeur propre $-i$. Pour $a,b,c,d,e,f \in \mathbb{C}$, on a

$$\begin{pmatrix} 0 & i & 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a\\b\\c\\d\\e\\f \end{pmatrix} = i \begin{pmatrix} a\\b\\c\\d\\e\\f \end{pmatrix} \iff \left\{\begin{array}{rcl} ib & = & ia \\ ia & = & ib \\ c & = & ic \\ d & = & id \\ e & = & ie \\ f & = & if \end{array}\right. \iff \left\{\begin{array}{l} a=b \\ c=d=e=f=0 \end{array}\right.$$

et

$$\begin{pmatrix} 1\\1\\0\\0\\0\\0 \end{pmatrix}$$

est un vecteur propre relatif à la valeur propre i. De même, pour $a,b,c,d,e,f \in \mathbb{C}$, on a

$$\begin{pmatrix} 0 & i & 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a\\b\\c\\d\\e\\f \end{pmatrix} = -i \begin{pmatrix} a\\b\\c\\d\\e\\f \end{pmatrix} \iff \left\{\begin{array}{rcl} ib & = & -ia \\ ia & = & -ib \\ c & = & -ic \\ d & = & -id \\ e & = & -ie \\ f & = & -if \end{array}\right. \iff \left\{\begin{array}{l} b=-a \\ c=d=e=f=0 \end{array}\right.$$

et

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

est un vecteur propre relatif à la valeur propre i. Par la proposition 3.7, les 6 vecteurs propres que nous avons obtenus sont nécessairement linéairement indépendants. Dès lors, la matrice

$$S = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

est telle que $S^{-1}AS = \mathrm{diag}(1, 1, 1, 1, i, -i)$.

Signalons pour conclure ce chapitre quelques résultats supplémentaires (sans preuve). Le premier est un autre critère de diagonalisation.

Proposition 3.38. *Une matrice est diagonalisable si et seulement si son polynôme minimal ne possède que des zéros simples.*

Nous mentionnons ensuite quelques propriétés de matrices particulières en lien avec leurs valeurs propres. Rappelons que $A^* = \bar{A}^\intercal$ est la matrice adjointe de A (si A est une matrice carrée complexe).

Proposition 3.39. *Soit $A \in \mathbb{C}^{m\times m}$ une matrice qui commute avec sa matrice adjointe, c'est-à-dire telle que $A^*A = AA^*$ (une telle matrice est dite* normale*). Alors*

- *la matrice A vérifie $A^* = A$ (une telle matrice est dite* hermitienne*) si et seulement si toutes ses valeurs propres sont réelles ;*
- *la matrice A admet A^* pour inverse (une telle matrice est dite* unitaire*) si et seulement si toutes ses valeurs propres sont de module* 1.

Proposition 3.40. *Toute matrice réelle symétrique est diagonalisable.*

Par exemple, le résultat précédent est utile en théorie des graphes. En effet, puisque la matrice d'adjacence d'un graphe non orienté est symétrique, cette proposition nous dit que tout graphe non orienté possède une matrice d'adjacence diagonalisable. Dans le cas d'une matrice réelle symétrique, on peut même demander que la matrice S qui diagonalise A soit *orthogonale*, c'est-à-dire que l'inverse de S soit égal à sa matrice ajointe S^*. C'est évidemment une propriété très intéressante quand on sait combien il peut être fastidieux de calculer l'inverse d'une grande matrice (et la théorie des graphes fournit souvent de grandes matrices).
En fait le résultat précédent est un cas particulier d'un résultat plus général, concernant les matrices normales. En effet, toute matrice réelle symétrique est hermitienne, donc normale.

Théorème 3.41 (Théorème spectral)**.** *Une matrice carrée complexe est normale si et seulement si elle est diagonalisable par une matrice unitaire.*

Chapitre 4

Quelques outils d'analyse

4.1 Approximation polynomiale d'une fonction au voisinage d'un point

Parmi toutes les fonctions $f\colon A \to \mathbb{R}$ avec $A \subseteq \mathbb{R}$, on peut dire que les fonctions polynomiales sont les plus simples. En particulier, pour évaluer un polynôme en un point, on utilise uniquement les opérations élémentaires d'addition et de multiplication de nombres. Rappelons également que, grâce au théorème fondamental de l'algèbre, nous avons vu que les notions de polynômes à coefficients réels et de fonctions polynomiales réelles coïncidaient (voir corollaire 1.36). Le but de cette section est d'apprendre comment on peut approcher les valeurs d'une fonction quelconque $f\colon A \to \mathbb{R}$ au voisinage d'un point au moyen de polynômes. D'autres techniques d'approximation de fonctions, comme par exemple l'interpolation de Lagrange, existent aussi, mais ne feront pas l'objet de cette première approche de ce vaste sujet.

Nous nous concentrons donc ici sur l'étude de fonctions réelles au voisinage d'un point.

Rappelons qu'une fonction $f\colon A \to \mathbb{R}$, avec $A \subseteq \mathbb{R}$, est *dérivable en un point* $a \in I$, où I est un intervalle ouvert inclus dans A, lorsque la limite

$$\lim_{x\to a} \frac{f(x)-f(a)}{x-a}$$

existe dans $\mathbb{R}$. Dans ce cas, cette limite est appelée la *dérivée de* f *en* a et est notée $\mathrm{D}\, f(a)$. Dans ce cas, si P est le polynôme de degré 1

$$P = f(a) + \mathrm{D}\, f(a)(X-a),$$

et si $r\colon A \to \mathbb{R}$ est la fonction définie par $r(x) = f(x) - P(x)$ pour tout $x \in A$, nous avons

$$\lim_{x\to a} \frac{r(x)}{x-a} = \lim_{x\to a} \frac{f(x)-P(x)}{x-a} = \lim_{x\to a} \left(\frac{f(x)-f(a)}{x-a} - \mathrm{D}\, f(a) \right) = 0. \tag{4.1}$$

Le polynôme P peut donc être vu comme une approximation polynomiale de f au voisinage de a, dans le sens où on peut approcher les valeurs de la fonction f au voisinage de a par celles de P. La fonction r, quant à elle, est appelée le *reste*, ou l'*erreur*. Le calcul de limite (4.1) exprime la qualité de l'approximation obtenue, puisqu'il nous indique que l'erreur commise $|r(x)|$ en valeur absolue, c'est-à-dire l'écart entre l'approximation $P(x)$ et la valeur $f(x)$ qu'on souhaite estimer, est petite par rapport à $|x-a|$ dans un voisinage de a.

À première vue, on peut être assez content de la situation, puisque nous avons répondu à la question initiale au moyen des dérivées. Pourtant, en y réfléchissant un peu, on se rend compte assez vite qu'approcher les valeurs de n'importe quelle fonction réelle par des

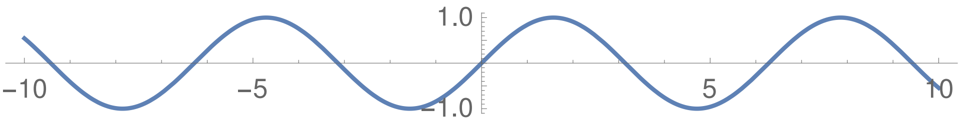

FIGURE 4.1 – La fonction sinus.

polynômes de degré 1, c'est-à-dire par des droites, n'est pas vraiment satisfaisant. Pensez par exemple à la fonction sinus, qui est représentée à la figure 4.1. On souhaite donc aller plus loin, et être capable d'approcher les valeurs de fonctions par des polynômes de degré quelconque. Malheureusement, cela n'est pas toujours possible... Nous allons donc étudier pour quelles fonctions ces approximations sont possibles et, lorsque c'est le cas, expliciter comment fournir de telles approximations.

Définition 4.1. Soit I un intervalle ouvert de $\mathbb{R}$, soit $f\colon A \to \mathbb{R}$ avec $I \subseteq A \subseteq \mathbb{R}$, soit $n \in \mathbb{N}$ et soit $a \in I$. On dit qu'un polynôme $P \in \mathbb{R}_n[X]$ est une *approximation de f à l'ordre n en a* lorsqu'on a

$$\lim_{x\to a} \frac{f(x) - P(x)}{(x-a)^n} = 0.$$

Autrement dit, P est une approximation de f à l'ordre n en a si $f = P + r$ avec $r(x) = o((x-a)^n)$ lorsque x tend vers a.[1] L'écart $|r(x)|$ entre l'approximation $P(x)$ et la valeur qu'on souhaite estimer, c'est-à-dire $f(x)$, est petite par rapport à $|x-a|^n$ dans un voisinage de a. Remarquez que si $|x-a| < 1$, alors

$$|x-a| \;>\; |x-a|^2 \;>\; |x-a|^3 \;>\; |x-a|^4 \;>\; \cdots$$

Une approximation à l'ordre n est donc meilleure qu'une approximation à l'ordre $n-1$.

Proposition 4.2.

1. *Si f admet une approximation polynomiale à l'ordre n en a, alors celle-ci est unique.*
2. *Si f admet une approximation polynomiale à l'ordre n en a, alors f admet une approximation à l'ordre i en a pour tout $i \leq n$.*

Démonstration. Supposons que $P \in \mathbb{R}_n[X]$ est tel que

$$\lim_{x\to a} \frac{f(x) - P(x)}{(x-a)^n} = 0.$$

Alors, pour tout $i \in \{0, \dots, n-1\}$, on a aussi

$$\lim_{x\to a} \frac{f(x) - P(x)}{(x-a)^i} = \lim_{x\to a} \left(\underbrace{\frac{f(x) - P(x)}{(x-a)^n}}_{\to 0} \cdot \underbrace{(x-a)^{n-i}}_{\to 0} \right) = 0.$$

Écrivons $P = c_n(X-a)^n + \cdots + c_1(X-a) + c_0$. Dans l'espace vectoriel $\mathbb{R}_n[X]$ des polynômes à coefficients réels de degré au plus n, les coefficients $c_n, \cdots, c_1, c_0$ sont les composantes de P dans la base $((X-a)^n, \dots, X-a, 1)$. Ils déterminent donc univoquement P.

Pour montrer l'unicité de l'approximation polynomiale P, il suffit donc de montrer l'unicité des coefficients c_i. Pour $i \in \{0, \dots, n\}$, on a

$$\lim_{x\to a} \frac{f(x) - \sum_{j=0}^{i-1} c_j (x-a)^j}{(x-a)^i} = \lim_{x\to a} \left(\underbrace{\frac{f(x) - P(x)}{(x-a)^i}}_{\to 0} + \frac{P(x) - \sum_{j=0}^{i-1} c_j (x-a)^j}{(x-a)^i} \right)$$

1. La notation "o" est une des notations de Landau, que nous allons détailler dans la section suivante.

$$\begin{aligned}
&= \lim_{x\to a} \sum_{j=i}^{n} \frac{c_j}{(x-a)^{i-j}} \\
&= \lim_{x\to a} \Big(c_i + \sum_{j=i+1}^{n} \underbrace{\frac{c_j}{(x-a)^{i-j}}}_{\to 0} \Big) \\
&= c_i.
\end{aligned}$$

On obtient donc, en procédant de proche en proche pour i allant de 0 à n, que les coefficients c_i sont uniques. En effet, on a

$$c_0 = \lim_{x\to a} f(x),\ \ c_1 = \lim_{x\to a} \frac{f(x)-c_0}{x-a},\ \ c_2 = \lim_{x\to a} \frac{f(x)-c_1(x-a)-c_0}{(x-a)^2}$$

et ainsi de suite jusque c_n. Le point 1 est démontré.
En réutilisant le calcul de limite précédent, on obtient que, pour tout $i \in \{0,\dots,n\}$,

$$\lim_{x\to a} \frac{f(x)-\sum_{j=0}^{i} c_j(x-a)^j}{(x-a)^i} = \lim_{x\to a} \frac{f(x)-\sum_{j=0}^{i-1} c_j(x-a)^j}{(x-a)^i} - c_i = 0$$

et donc le polynôme $c_i(X-a)^i + \dots + c_1(X-a) + c_0$ est une approximation polynomiale de f à l'ordre i en a, ce qui démontre le point 2. □

Nous savons à présent que, *si elle existe*, l'approximation polynomiale de f à l'ordre n en a est unique. D'un point de vue pratique, nous voudrions être capables de donner des conditions explicites et pouvant être testées effectivement sous lesquelles cette approximation polynomiale existe, et dans ce cas, pouvoir fournir explicitement l'approximation polynomiale recherchée. C'est l'objet du résultat suivant.

Théorème 4.3 (Approximation polynomiale)**.** *Soit I un intervalle ouvert de $\mathbb{R}$, soit $f\colon A \to \mathbb{R}$ avec $I \subseteq A \subseteq \mathbb{R}$, soit $n \in \mathbb{N}$ et soit $a \in I$. Si f est n fois dérivable sur I, alors le polynôme*

$$\sum_{i=0}^{n} \frac{(\mathrm{D}^i f)(a)}{i!}(X-a)^i$$

est l'approximation de f à l'ordre n en a.

Démonstration. Nous devons montrer que

$$\lim_{x\to a} \frac{f(x)-\sum_{i=0}^{n} \frac{(\mathrm{D}^i f)(a)}{i!}(x-a)^i}{(x-a)^n} = 0. \tag{4.2}$$

Nous allons appliquer le théorème de l'Hospital pour calculer la limite (4.2). Pour tout $k \in \{0,\dots,n-2\}$, au vu de la proposition 1.24 et en utilisant que $\mathrm{D}^k f$ est une fonction continue, on a

$$\begin{aligned}
\lim_{x\to a} \mathrm{D}^k \Big(f(x) - \sum_{i=0}^{n} \frac{(\mathrm{D}^i f)(a)}{i!}(x-a)^i \Big) &= \lim_{x\to a} \Big(\mathrm{D}^k f(x) - \sum_{i=k}^{n} \frac{(\mathrm{D}^i f)(a)}{(i-k)!}(x-a)^{i-k} \Big) \\
&= \lim_{x\to a} \Big(\underbrace{\mathrm{D}^k f(x) - \mathrm{D}^k f(a)}_{\to 0} - \sum_{i=k+1}^{n} \frac{(\mathrm{D}^i f)(a)}{(i-k)!} \underbrace{(x-a)^{i-k}}_{\to 0} \Big) = 0.
\end{aligned}$$

Comme on a aussi $\lim_{x\to a}(x-a)^{n-k} = 0$ pour tout $k \in \{0,\dots,n-2\}$, on obtient en appliquant $n-1$ fois le théorème de l'Hospital que

$$\lim_{x\to a} \frac{f(x)-\sum_{i=0}^{n} \frac{(\mathrm{D}^i f)(a)}{i!}(x-a)^i}{(x-a)^n} = \lim_{x\to a} \frac{\mathrm{D}^{n-1}\Big(f(x)-\sum_{i=0}^{n} \frac{(\mathrm{D}^i f)(a)}{i!}(x-a)^i\Big)}{\mathrm{D}^{n-1}(x-a)^n}$$

$$
\begin{aligned}
&= \lim_{x\to a} \frac{\mathrm{D}^{n-1} f(x) - \mathrm{D}^{n-1} f(a) - \mathrm{D}^{n} f(a)\,(x-a)}{n!(x-a)} \\
&= \frac{1}{n!} \lim_{x\to a} \left(\frac{\mathrm{D}^{n-1} f(x) - \mathrm{D}^{n-1} f(a)}{x-a} - \mathrm{D}^{n} f(a) \right) \\
&= 0
\end{aligned}
$$

où, à la dernière étape, on a utilisé l'hypothèse que $\mathrm{D}^{n-1} f$ est dérivable en a. □

Les approximations polynomiales de la fonction sinus en 2 pour les ordres de 0 à 4 sont représentées à la figure 4.2. Un zoom sur le point d'absisse 2 est représenté à la figure 4.3. Dans ce deuxième cas, seules les approximations d'ordre 0, 1 et 2 sont représentées.

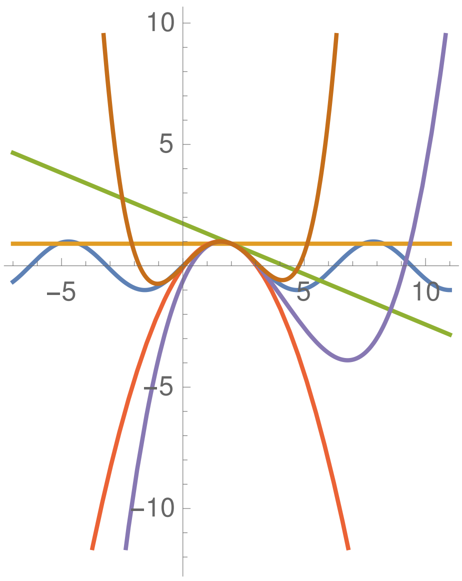

FIGURE 4.2 – Approximations polynomiales de la fonction sinus en $a = 2$ pour les ordres $n \in \{0, \ldots, 4\}$.

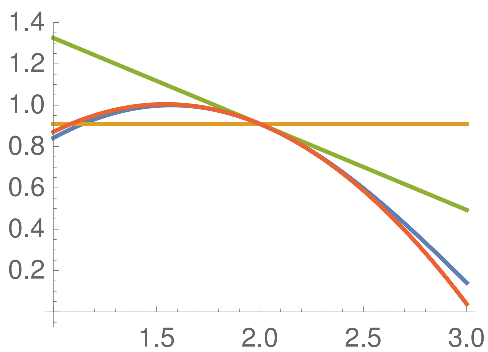

FIGURE 4.3 – Approximations polynomiales de la fonction sinus en $a = 2$ pour les ordres $n \in \{0, 1, 2\}$.

Remarque 4.4. Dans le cas où la fonction f est un polynôme, souvenons-nous de la formule de Taylor vue au chapitre 1 : pour tout $P \in \mathbb{R}[X]$ et tout $a \in \mathbb{R}$, on a

$$P = \sum_{i=0}^{\deg(P)} \frac{\mathrm{D}^{i} P(a)}{i!} (X-a)^{i}.$$

Ainsi, si $n < \deg(P)$, l'approximation polynomiale de P à l'ordre n est le polynôme

$$\sum_{i=0}^{n} \frac{\mathrm{D}^i P(a)}{i!}(X-a)^i$$

et pour tout $n \geq \deg(P)$, l'approximation polynomiale de P à l'ordre n est le polynôme P lui-même. Bien entendu, dans ce deuxième cas, il ne s'agit plus d'une approximation mais d'une formule exacte : le reste est nul.

Remarque 4.5. Le théorème 4.3 nous dit en particulier que toute fonction n fois dérivable sur I admet une approximation polynomiale à l'ordre n en a, pour tout $a \in I$. La réciproque est fausse : si $n \geq 2$, il est possible qu'une fonction f définie sur un intervalle I admette une approximation polynomiale à l'ordre n en a pour tout $a \in I$ sans pour autant qu'elle soit n fois dérivable sur I. C'est le cas par exemple de la fonction

$$f\colon \mathbb{R} \to \mathbb{R},\ x \mapsto \begin{cases} x^3 \sin\left(\frac{1}{x}\right) & \text{si } x \neq 0, \\ 0 & \text{si } x = 0. \end{cases}$$

En effet, cette fonction f admet une approximation polynomiale à l'ordre $n = 2$ en $a = 0$ puisque

$$\lim_{x \to 0} \frac{f(x) - 0}{(x-0)^2} = \lim_{x \to 0} \frac{x^3 \sin\left(\frac{1}{x}\right)}{x^2} = \lim_{x \to 0} x \sin\left(\frac{1}{x}\right) = 0$$

(l'approximation polynomiale étant donnée par le polynôme $P = 0$). Or la fonction f n'est pas deux fois dérivable sur $\mathbb{R}$. En effet, f est dérivable mais sa dérivée

$$\mathrm{D}\, f\colon \mathbb{R} \to \mathbb{R},\ x \mapsto \begin{cases} 3x^2 \sin\left(\frac{1}{x}\right) - x\cos\left(\frac{1}{x}\right) & \text{si } x \neq 0, \\ 0 & \text{si } x = 0. \end{cases}$$

n'est pas dérivable en 0 : pour tout $x \in \mathbb{R}_0$, on a

$$\frac{\mathrm{D}\, f(x) - \mathrm{D}\, f(0)}{x - 0} = \frac{3x^2 \sin\left(\frac{1}{x}\right) - x\cos\left(\frac{1}{x}\right)}{x} = \underbrace{3x \sin\left(\frac{1}{x}\right)}_{\to 0 \text{ si } x \to 0} - \underbrace{\cos\left(\frac{1}{x}\right)}_{\substack{\text{oscille entre } -1 \\ \text{et 1 si } x \to 0}}$$

et donc le quotient $\frac{\mathrm{D}\, f(x) - \mathrm{D}\, f(0)}{x-0}$ n'admet pas de limite en 0. La fonction f et sa dérivée sont représentées à la figure 4.4).

En revanche, il est vrai que toute fonction f admettant une approximation polynomiale à l'ordre 0 en a, pour tout $a \in I$, est nécessairement continue sur I et toute fonction f admettant une approximation polynomiale à l'ordre 1 en a, pour tout $a \in I$, est nécessairement dérivable sur I. Ceci se vérifie facilement. En effet, si $P = c_0$ est tel que $\lim_{x \to a} f(x) - c_0 = 0$, alors $\lim_{x \to a} f(x) = c_0$, et donc la fonction f est continue en a (on a même $f(a) = c_0$). De même, si $P = c_1(X-a) + c_0$ est tel que $\lim_{x \to a} \frac{f(x) - (c_1(x-a) + c_0)}{x-a} = 0$, alors on a nécessairement $c_0 = f(a)$ et $c_1 = \lim_{x \to a} \frac{f(x) - f(a)}{x-a}$, ce qui montre que la fonction f est dérivable en a (et même que $\mathrm{D}\, f(a) = c_1$).

Exemple 4.6. Calculons les approximations polynomiales de la fonction cosinus en $a = \frac{\pi}{2}$. Tout d'abord, intéressons-nous aux dérivées successives de cette fonction. On a

$$\mathrm{D} \cos = -\sin,\ \mathrm{D}^2 \cos = -\cos,\ \mathrm{D}^3 \cos = \sin,\ \mathrm{D}^4 \cos = \cos\,.$$

Ainsi, pour tout $n \in \mathbb{N}$, on obtient

$$(\mathrm{D}^{4n} \cos)\left(\frac{\pi}{2}\right) = \cos\left(\frac{\pi}{2}\right) = 0$$

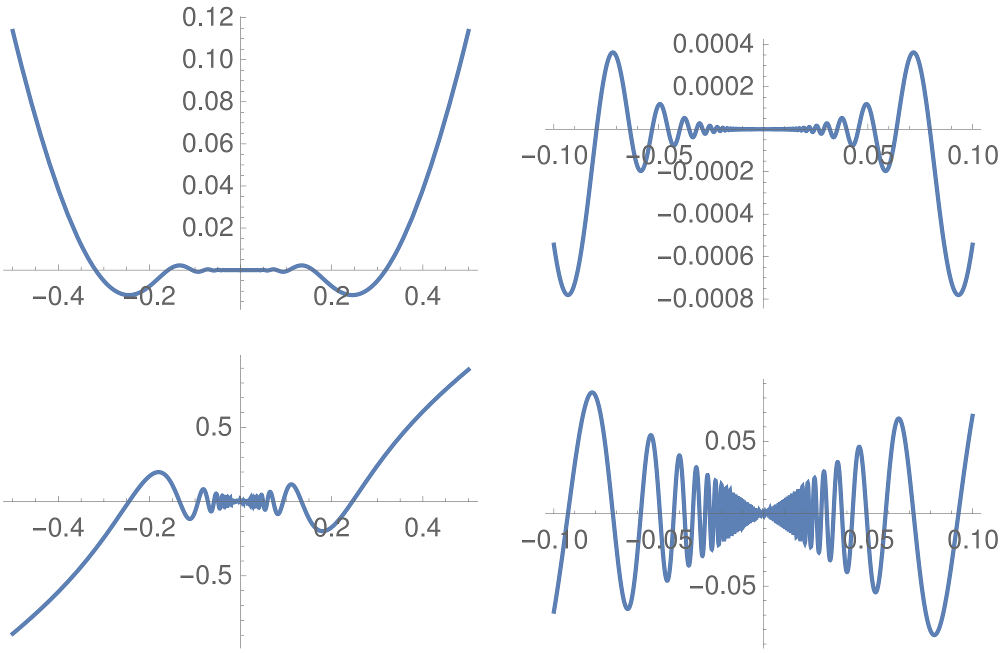

FIGURE 4.4 – Une fonction dérivable en 0 mais pas deux fois dérivable en 0. En haut, la fonction f au voisinage de 0 vue à deux échelles différentes. En bas, la dérivée de f au voisinage de 0 vue aux mêmes deux échelles.

$$(\mathrm{D}^{4n+1}\cos)\left(\frac{\pi}{2}\right) = -\sin\left(\frac{\pi}{2}\right) = -1$$
$$(\mathrm{D}^{4n+2}\cos)\left(\frac{\pi}{2}\right) = -\cos\left(\frac{\pi}{2}\right) = 0$$
$$(\mathrm{D}^{4n+3}\cos)\left(\frac{\pi}{2}\right) = \sin\left(\frac{\pi}{2}\right) = 1.$$

Aux ordres $4n+1$ et $4n+2$, on obtient l'approximation polynomiale

$$-\frac{1}{1!}\left(X-\frac{\pi}{2}\right)+\frac{1}{3!}\left(X-\frac{\pi}{2}\right)^3-\frac{1}{5!}\left(X-\frac{\pi}{2}\right)^5+\ldots-\frac{1}{(4n+1)!}\left(X-\frac{\pi}{2}\right)^{4n+1}$$

et aux ordres $4n+3$ et $4n+4$, on obtient l'approximation polynomiale

$$-\frac{1}{1!}\left(X-\frac{\pi}{2}\right)+\frac{1}{3!}\left(X-\frac{\pi}{2}\right)^3-\frac{1}{5!}\left(X-\frac{\pi}{2}\right)^5+\ldots+\frac{1}{(4n+3)!}\left(X-\frac{\pi}{2}\right)^{4n+3}.$$

Par exemple, aux ordres 5 et 6, on obtient

$$-\left(X-\frac{\pi}{2}\right)+\frac{1}{3!}\left(X-\frac{\pi}{2}\right)^3-\frac{1}{5!}\left(X-\frac{\pi}{2}\right)^5$$

et aux ordres 7 et 8, on obtient

$$-\left(X-\frac{\pi}{2}\right)+\frac{1}{3!}\left(X-\frac{\pi}{2}\right)^3-\frac{1}{5!}\left(X-\frac{\pi}{2}\right)^5+\frac{1}{7!}\left(X-\frac{\pi}{2}\right)^7$$

Les approximations polynomiales de la fonction cosinus en $a = \frac{\pi}{2}$ sont représentées à la figure 4.5 pour $n \in \{1, 3, 5\}$.

Pour terminer cette section sur les approximations polynomiales, nous souhaitons donner une estimation de l'écart entre la valeur de la fonction estimée $f(a)$ et celle obtenue grâce à son approximation polynomiale $P(a)$. Par définition, nous savons qu'à l'ordre n, cet écart définit une fonction $r_n(x) \in o((x-a)^n)$ lorsque x tend vers a, c'est-à-dire qui est telle que

$$\lim_{x\to a}\frac{r_n(x)}{(x-a)^n} = 0.$$

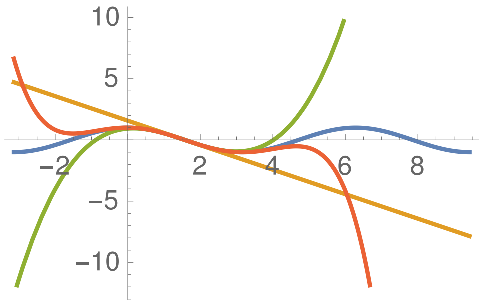

FIGURE 4.5 – Approximations polynomiales de cosinus en $a = \frac{\pi}{2}$ pour les ordres $n \in \{1, 3, 5\}$

En se servant du résultat suivant, livré ici sans démonstration [2], on pourra dans de nombreux cas obtenir une majoration du type $|r_n(x)| \leq c|x-a|^n$, où c est une constante, voire même du type $|r_n(x)| \leq c|x-a|^{n+1}$ lorsque f est suffisamment dérivable.

Théorème 4.7 (Développement limité de Taylor)**.** *Soit I un intervalle ouvert de $\mathbb{R}$, soit $n \in \mathbb{N}$ et soit $f \colon A \to \mathbb{R}$ avec $I \subseteq A \subseteq \mathbb{R}$ une fonction $n+1$ fois dérivable sur I. Pour tous $a, x \in I$ distincts, il existe u compris strictement entre a et x tel que*

$$r_n(x) := f(x) - P_n(x) = \frac{(\mathrm{D}^{n+1} f)(u)}{(n+1)!}(x-a)^{n+1},$$

où P_n désigne l'approximation polynomiale de f à l'ordre n en a.

Exemple 4.8. Pour tout $x \in \mathbb{R}$, l'écart entre $\cos(x)$ et son approximation à l'ordre $n = 5$ en $a = \frac{\pi}{2}$ vaut

$$r_5(x) = \cos(x) + \left(x - \frac{\pi}{2}\right) - \frac{1}{3!}\left(x - \frac{\pi}{2}\right)^3 + \frac{1}{5!}\left(x - \frac{\pi}{2}\right)^5.$$

Le théorème précédent nous indique que si $x \neq \frac{\pi}{2}$, alors il existe u compris strictement entre x et $\frac{\pi}{2}$ tel que

$$r_5(x) = \frac{(\mathrm{D}^6 \cos)(u)}{6!}\left(x - \frac{\pi}{2}\right)^6 = \frac{-\cos(u)}{6!}\left(x - \frac{\pi}{2}\right)^6.$$

On en déduit que

$$|r_5(x)| \leq \frac{1}{6!}\left(x - \frac{\pi}{2}\right)^6.$$

Remarquez qu'on a utilisé ici que la fonction cosinus était 6 fois dérivable (et non pas juste 5 fois comme pour définir l'approximation polynomiale d'ordre 5), ce qui n'est pas un problème puisque la fonction cosinus est dans $\mathrm{C}^\infty(\mathbb{R})$.

Remarque 4.9. Si la fonction f est n fois dérivable mais pas $n+1$ fois dérivable, elle admet tout de même une approximation d'ordre n (vu le théorème 4.3) et on peut aussi obtenir une estimation du reste à l'ordre n en utilisant le développement limité de Taylor, mais en procédant un peu différemment. En effet, dans ce cas, pour $a, x \in I$ distincts, le théorème 4.7 appliqué au reste à l'ordre $n-1$ nous dit qu'il existe u compris strictement entre x et a tel que

$$\begin{aligned} r_n(x) &= r_{n-1}(x) - \frac{(\mathrm{D}^n f)(a)}{n!}(x-a)^n \\ &= \frac{(\mathrm{D}^n f)(u)}{n!}(x-a)^n - \frac{(\mathrm{D}^n f)(a)}{n!}(x-a)^n \\ &= \frac{(\mathrm{D}^n f)(u) - (\mathrm{D}^n f)(a)}{n!}(x-a)^n. \end{aligned}$$

2. Il s'agit en fait d'une généralisation du théorème des accroissements finis vu dans votre cours "Mathématique" du bloc 1.

4.2 Notations de Landau

Nous avons déjà rencontré la notation o. Il existe d'autres notations de ce type pour exprimer les ordres de croissance asymptotique de fonctions au voisinage d'un point ou à l'infini. Ces notations sont appelées les notations de Landau. Nous les rappelons ici dans le cas de fonctions réelles pour des valeurs qui tendent vers $+\infty$. Nous fixons pour toute cette section des fonctions réelles f, g et h telles que leurs domaines contiennent un intervalle de la forme $I = [a, +\infty[$.

Définition 4.10 (La notation de domination O)**.** On écrit $f \in O(g)$ s'il existe des constantes strictement positives C et N telles que pour tout réel $x \geq N$, on a $|f(x)| \leq C|g(x)|$. Dans ce cas, on dit que la fonction f est *dominée* par g. La notation $O(g)$ désigne donc la classe des fonctions dominées par g.

Exemple 4.11. La fonction $\mathbb{R} \to \mathbb{R}$, $x \mapsto x\sin(x)$ est dominée par la fonction identité $\mathbb{R} \to \mathbb{R}$, $x \mapsto x$ puisque $|x\sin(x)| \leq |x|$ pour tout $x \in \mathbb{R}$. On écrit $x\sin(x) \in O(x)$. Ces fonctions sont représentées à la figure 4.6 (premier graphique) et leurs valeurs absolues (deuxième graphique).

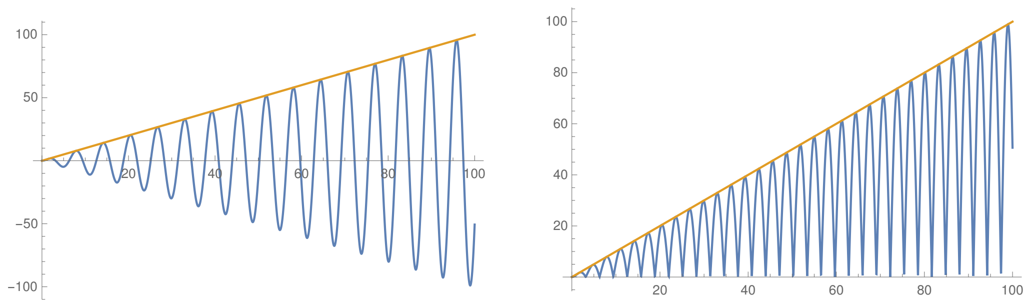

FIGURE 4.6 – Les fonctions $x\sin(x)$ et x (gauche) et les fonctions $|x\sin(x)|$ et $|x|$ (droite).

Proposition 4.12 (Arithmétique de O)**.**

- *Si f et g sont dans $O(h)$, alors $f + g$ est aussi dans $O(h)$.*
- *Si f est $O(g)$ et si $\lambda \in \mathbb{R}$, alors λf est également dans $O(g)$.*

Démonstration. Premièrement, supposons que f et g soient dans $O(h)$. Il existe des constantes strictement positives C, N, C', N' telles que pour tout $x \geq N$, on a $|f(x)| \leq C|h(x)|$ et pour tout $x \geq N'$, on a $|g(x)| \leq C'|h(x)|$. Alors les constantes $C'' = \max\{C, C'\}$ et $N'' = \max\{N, N'\}$ sont telles que pour tout $x \geq N''$, on a $|f(x)| \leq C''|h(x)|$. Ainsi, $f + g \in O(h)$.
Deuxièmement, soit $\lambda \in \mathbb{R}$ et soient C et N des constantes strictement positives telles que pour tout $x \geq N$, on a $|f(x)| \leq C|g(x)|$. Si $\lambda = 0$, alors clairement $\lambda f = 0 \in O(g)$. Sinon, alors $C' = |\lambda|C$ est une constante strictement positive telle que pour tout $x \geq N$, $|\lambda f(x)| \leq |\lambda|\, C\, |g(x)| = C'|g(x)|$. Dans les deux cas, on a $\lambda f \in O(g)$. $\square$

Définition 4.13. Voici quelques classes de complexité que l'on rencontre souvent.

- $O(1)$ est la classe des fonctions à croissance *constante*.
- $O(\log(x))$ est la classe des fonctions à croissance *logarithmique*.
- $O((\log(x))^c)$ est la classe des fonctions à croissance *polylogarithmique d'exposant c*.
- $O(x)$ est la classe des fonctions à croissance *linéaire*.
- $O(x\log(x))$ est la classe des fonctions à croissance *quasi-linéaire*.
- $O(x^2)$ est la classe des fonctions à croissance *quadratique*.
- $O(x^c)$ est la classe des fonctions à croissance *polynomiale de degré c*.

- $O(c^x)$ est la classe des fonctions à croissance *exponentielle de base c*.

Définition 4.14 (La notation Θ)**.** On écrit $f \in \Theta(g)$ si $f \in O(g)$ et $g \in O(f)$.

Exercice 4.15. Montrer que si $f \in \Theta(g)$, alors les classes $\Theta(f)$ et $\Theta(g)$ coïncident. En particulier, on a $f \in \Theta(g) \iff g \in \Theta(f)$.

Exemple 4.16. On a $5x^2(\sin(x)+2) \in \Theta(x^2)$. En effet, pour tout réel x, on a $1 \leq \sin(x)+2 \leq 3$ et

$$5x^2 \leq 5x^2(\sin(x)+2) \leq 15x^2.$$

Ces inégalités sont représentées à la figure 4.7 pour $x \geq 0$.

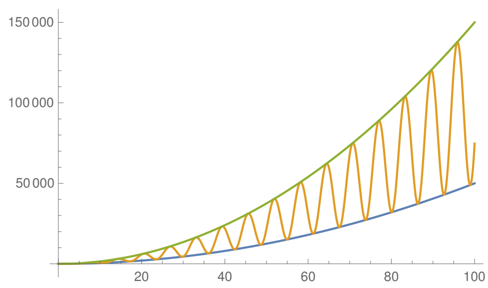

FIGURE 4.7 – Les fonctions $5x^2$ (bleu), $5x^2(\sin(x)+2)$ (orange) et $15x^2$ (vert).

De même, on a $x^2 \in \Theta(5x^2(\sin(x)+2))$. En effet, pour tout réel x, on a $1 \leq \sin(x)+2 \leq 3$ et

$$\frac{1}{3}x^2(\sin(x)+2) \leq x^2 \leq x^2(\sin(x)+2).$$

Ces inégalités sont représentées à la figure 4.8 pour $x \geq 0$.

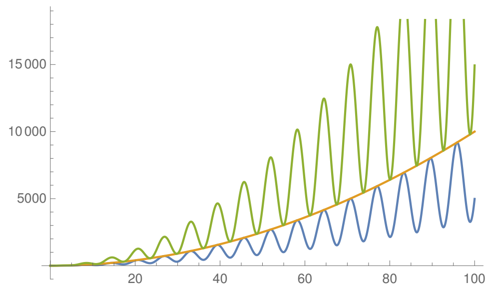

FIGURE 4.8 – Les fonctions $\frac{1}{3}x^2(\sin(x)+2)$ (bleu), x^2 (orange) et $x^2(\sin(x)+2)$ (vert).

Proposition 4.17. *Supposons qu'il existe un intervalle de la forme* $[a,+\infty[$ *dans lequel* g *ne s'annule pas. Si la limite*

$$\lim_{x\to+\infty} \frac{f(x)}{g(x)} \tag{4.3}$$

existe et est non nulle, alors $f = \Theta(g)$.

Démonstration. Notons L la limite de l'énoncé. Par hypothèse, pour tout $\varepsilon > 0$, il existe $N > 0$ tel que pour tout $x \geq N$, on a $g(x) \neq 0$ et $\left|\frac{f(x)}{g(x)} - L\right| < \varepsilon$. Or,

$$\left|\frac{f(x)}{g(x)} - L\right| < \varepsilon \iff L - \varepsilon < \frac{f(x)}{g(x)} < L + \varepsilon.$$

Supposons d'abord que $L > 0$. On peut donc choisir ε suffisamment petit pour que $L-\varepsilon > 0$. Ainsi, il existe $N > 0$ tel que pour tout $x \geq N$, on a

$$\begin{aligned}\left|\frac{f(x)}{g(x)} - L\right| < \varepsilon &\iff L - \varepsilon < \left|\frac{f(x)}{g(x)}\right| < L + \varepsilon \\ &\iff (L-\varepsilon)|g(x)| < |f(x)| < (L+\varepsilon)|g(x)|.\end{aligned}$$

On obtient donc que $f \in \Theta(g)$.
L'argument est similaire dans le cas où $L < 0$. En effet, dans ce cas, on peut donc choisir ε suffisamment petit pour que $L + \varepsilon < 0$, ou de manière équivalente, pour que $-L - \varepsilon > 0$. Ainsi, il existe $N > 0$ tel que pour tout $x \geq N$, on a

$$\begin{aligned}\left|\frac{f(x)}{g(x)} - L\right| < \varepsilon &\iff -L + \varepsilon > \left|\frac{f(x)}{g(x)}\right| > -L - \varepsilon \\ &\iff (-L+\varepsilon)|g(x)| > |f(x)| > (-L-\varepsilon)|g(x)|.\end{aligned}$$

On obtient à nouveau que $f \in \Theta(g)$. □

Dans l'exemple 4.16, la limite

$$\lim_{x\to+\infty} \frac{5x^2(\sin(x)+2)}{x^2}$$

n'existe pas. La condition de l'existence de la limite 4.3 dans la proposition 4.17 est donc suffisante, mais pas nécessaire !

Définition 4.18 (La notation o)**.** Supposons qu'il existe un intervalle de la forme $[a, +\infty[$ dans lequel g ne s'annule pas. On écrit $f \in o(g)$ lorsque

$$\lim_{x\to+\infty} \frac{f(x)}{g(x)} = 0.$$

On dit dans ce cas que la fonction f est *négligeable devant* g.

Exemple 4.19. On a $x^k \in o(x^{k+1})$ pour tout entier $k \geq 0$ car $\lim_{x\to+\infty} \frac{x^k}{x^{k+1}} = \lim_{x\to+\infty} \frac{1}{x} = 0$. Les fonctions x^k pour $k = 1, 2, 3$ sont représentées à la figure 4.9.

Proposition 4.20. *Si $f \in o(g)$, alors $f \in O(g)$.*

Démonstration. En effet, si $f \in o(g)$, alors pour tout $\varepsilon > 0$, il existe $N > 0$ tel que pour tout $x \geq N$, on a $\left|\frac{f(x)}{g(x)}\right| < \varepsilon$, donc aussi $|f(x)| < \varepsilon|g(x)|$. □

Proposition 4.21 (Arithmétique de o)**.**
- *Si f et g sont dans $o(h)$, alors $f + g$ est aussi dans $o(h)$.*
- *Si f et g sont dans $o(h)$, alors $fg \in o(h^2)$.*
- *Si $f \in O(1)$ et $g \in o(h)$, alors $fg \in o(h)$.*

Démonstration. Pour les deux premiers points, il suffit de voir que si

$$\lim_{x\to+\infty} \frac{f(x)}{h(x)} = 0 \quad \text{et} \quad \lim_{x\to+\infty} \frac{g(x)}{h(x)} = 0,$$

alors

$$\lim_{x\to+\infty} \frac{f(x)+g(x)}{h(x)} = \lim_{x\to+\infty} \left(\frac{f(x)}{h(x)} + \frac{g(x)}{h(x)}\right) = 0$$

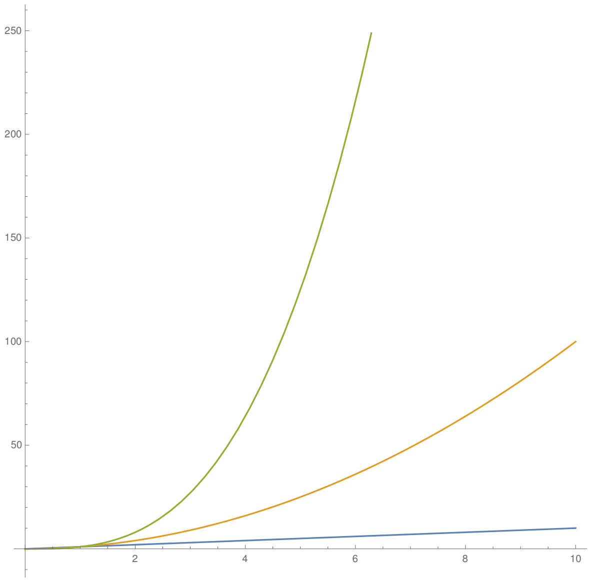

FIGURE 4.9 – Les fonctions x, x^2 et x^3.

et

$$\lim_{x\to+\infty} \frac{f(x)g(x)}{(h(x))^2} = \lim_{x\to+\infty} \left(\frac{f(x)}{h(x)} \cdot \frac{g(x)}{h(x)}\right) = 0.$$

Pour le troisième point, supposons que $f \in O(1)$ et $g \in o(h)$. Il existe des constantes strictement positives C et N telles que pour tout $x \geq N$, on a $|f(x)| \leq C$. Quitte à choisir un N plus grand, on peut supposer que les fonctions g et h soient définies sur $[N, +\infty[$. Alors pour tout $x \geq N$, on a

$$\left|\frac{f(x)g(x)}{h(x)}\right| \leq C \left|\frac{g(x)}{h(x)}\right|.$$

Puisque $\lim_{x\to+\infty} \frac{g(x)}{h(x)} = 0$, on obtient par le critère de l'étau que

$$\lim_{x\to+\infty} \frac{f(x)g(x)}{h(x)} = 0.$$

Ainsi, $fg \in o(h)$. □

Définition 4.22 (Equivalence asymptotique)**.** Supposons qu'il existe un intervalle de la forme $[a, +\infty[$ dans lequel g ne s'annule pas. On écrit $f \sim g$ lorsque

$$\lim_{x\to+\infty} \frac{f(x)}{g(x)} = 1.$$

On dit dans ce cas que f est *asymptotiquement équivalent à* g.

Exemple 4.23. On a $x^2 - x \sim x^2 + x + 10$ car $\lim_{x\to+\infty} \frac{x^2 - x}{x^2 + x + 10} = 1$.

Proposition 4.24. *La relation $\equiv$ est une relation d'équivalence : on a*

1. $f \sim f$
2. $f \sim g \implies g \sim f$.
3. $f \sim g$ *et* $g \sim h \implies f \sim h$.

Démonstration. Le premier point est évident. Pour le deuxième point, supposons que $f \sim g$. Il existe un intervalle de la forme $[a, +\infty[$ dans lequel g ne s'annule pas et on a

$$\lim_{x \to +\infty} \frac{f(x)}{g(x)} = 1.$$

Il doit donc aussi exister un intervalle de la forme $[b, +\infty[$ dans lequel f ne s'annule pas (sinon, la limite précédente ne pourrait pas valoir 1). De plus, on a

$$\lim_{x \to +\infty} \frac{g(x)}{f(x)} = \lim_{x \to +\infty} \frac{1}{\frac{f(x)}{g(x)}} = 1.$$

Pour le troisième point, supposons aussi que $g \sim h$. Alors, il existe un intervalle de la forme $[c, +\infty[$ dans lequel h ne s'annule pas et on a

$$\lim_{x \to +\infty} \frac{g(x)}{h(x)} = 1.$$

Ainsi,

$$\lim_{x \to +\infty} \frac{f(x)}{h(x)} = \lim_{x \to +\infty} \frac{f(x)}{g(x)} \frac{g(x)}{h(x)} = \lim_{x \to +\infty} \frac{f(x)}{g(x)} \cdot \lim_{x \to +\infty} \frac{g(x)}{h(x)} = 1.$$

□

Proposition 4.25. *On a $f \sim g \iff f - g \in o(g)$.*

Démonstration. En effet, on a

$$\lim_{x \to +\infty} \frac{f(x)}{g(x)} = 1 \iff \lim_{x \to +\infty} \frac{f(x) - g(x)}{g(x)} = 0.$$

□

Proposition 4.26. *Si $f \sim g$, alors la limite $\lim_{x \to +\infty} f(x)$ existe si et seulement la limite $\lim_{x \to +\infty} g(x)$ existe, et dans le cas où elles existent, ces limites sont égales.*

Démonstration. Si $\lim_{x \to +\infty} g(x) = L$, alors

$$\lim_{x \to +\infty} f(x) = \lim_{x \to +\infty} \frac{f(x)}{g(x)} \cdot g(x) = \lim_{x \to +\infty} \frac{f(x)}{g(x)} \cdot \lim_{x \to +\infty} g(x) = L.$$

On conclut par symétrie de la relation $\sim$. □

Remarquez qu'on peut facilement adapter les définitions de cette section pour comparer des fonctions au voisinage d'un point a plutôt qu'à l'infini.

4.3 Suites et séries numériques

Au vu du théorème 4.3, lorsqu'une fonction $f\colon A \to \mathbb{R}$ est $\mathrm{C}^\infty(I)$, où I est un intervalle inclus dans le domaine A de f, elle admet des approximations polynomiales à tous les ordres $n \in \mathbb{N}$ sur I. On est alors tenté d'étudier la validité d'une expression du type

$$f(x) = \sum_{i=0}^{+\infty} \frac{(\mathrm{D}^i f)(a)}{i!} (x - a)^i.$$

Dans cette expression se cache en réalité un passage à la limite pour n tendant vers l'infini :

$$\lim_{n\to+\infty}\Big(\sum_{i=0}^{n}\frac{(\mathrm{D}^i f)(a)}{i!}(x-a)^i + r_n(x)\Big).$$

Les fonctions qui admettent un tel développement au voisinage d'un point a, c'est-à-dire qui sont telles que

$$\lim_{n\to+\infty} r_n(x) = 0$$

pour tout x dans un intervalle ouvert contenant a, sont dites *analytiques en* a. Elles sont dites *analytiques sur l'intervalle ouvert* I si elles sont analytiques en tout $a \in I$. C'est le cas de la plupart des fonctions élémentaires que vous connaissez, mais c'est loin d'être le cas de toutes les fonctions !

Une somme infinie est en fait la limite d'une suite tout à fait particulière, appelée série. Le but de cette section est d'étudier la convergence de ces suites particulières.

Définition 4.27. Une *suite de points d'un ensemble* A est une fonction $x\colon \mathbb{N} \to A$. On note communément une suite x par $(x(i))_{i\in\mathbb{N}}$, $(x_i)_{i\in\mathbb{N}}$ ou même $(x_i)_{i\geq 0}$. Puisque[3] pour tout $j \in \mathbb{N}$, l'ensemble $\{j, j+1, \ldots\}$ est en bijection avec $\mathbb{N}$, on considère de la même façon que $(x_i)_{i\geq j}$ est une suite, mais indicée à partir de j plutôt qu'à partir de 0.

Une suite de réels est alors simplement une fonction $x\colon \mathbb{N} \to \mathbb{R}$ (l'ensemble A choisi est l'ensemble $\mathbb{R}$).

Définition 4.28. Une *série de réels* est une suite de la forme

$$S\colon \mathbb{N} \to \mathbb{R},\ n \mapsto \sum_{i=0}^{n} x_i$$

où $(x_i)_{i\geq 0}$ est une suite de réels. Autrement dit, une série S est la suite des sommes $\sum_{i=0}^{n} x_i$ cumulées d'une suite x. Comme d'habitude, on écrit $S_n = \sum_{i=0}^{n} x_i$ pour tout $n \in \mathbb{N}$. Lorsque la série S est convergente, on désigne par

$$\sum_{i=0}^{+\infty} x_i$$

sa limite. Cette limite est parfois appelée la *somme* de la série S.

Remarque 4.29. Puisqu'une série n'est rien de plus qu'une suite particulière, on utilise la même définition pour la convergence des séries que celle pour la convergence des suites. Si la série S converge, on a donc

$$\lim_{n\to+\infty} S_n = \lim_{n\to+\infty}\sum_{i=0}^{n} x_i = \sum_{i=0}^{+\infty} x_i.$$

Tout comme on trouve la notation $(x_i)_{i\geq j}$ pour les suites, on trouve aussi la notation

$$\Big(\sum_{i=j}^{n} x_i\Big)_{n\geq j}.$$

Il s'agit précisément de la suite des sommes partielles de la suite $(x_i)_{i\geq j}$.

Par abus de langage, on parle aussi de la série $\sum_{i=0}^{+\infty} x_i$, que celle-ci converge ou non. Dans ce cas, il faut entendre qu'étudier la convergence de la *série* $\sum_{i=0}^{+\infty} x_i$ signifie en fait étudier la convergence de la série $(\sum_{i=0}^{n} x_i)_{n\geq 0}$.

Commençons par étudier quelques exemples.

3. Pourquoi ?

Exemple 4.30. Considérons la série

$$\sum_{i=0}^{+\infty}(-1)^i.$$

Si on note $S_n = \sum_{i=0}^{n}(-1)^i$ pour $n \in \mathbb{N}$, on a

$$\begin{array}{rcl} S_0 & = & (-1)^0 = 1 \\ S_1 & = & S_0 + (-1)^1 = 1 - 1 = 0 \\ S_2 & = & S_1 + (-1)^2 = 0 + 1 = 1 \\ S_3 & = & S_2 + (-1)^3 = 1 - 1 = 0 \\ & \vdots & \end{array}$$

On a donc extrait deux sous-suites $(S_{2n})_{n\in\mathbb{N}}$ et $(S_{2n+1})_{n\in\mathbb{N}}$ convergeant vers des limites différentes, à savoir 1 et 0 respectivement. La série ne converge donc pas !

Exemple 4.31. Considérons la série

$$\sum_{i=1}^{+\infty}\frac{1}{2^i}.$$

Si on note $S_n = \sum_{i=1}^{n}\frac{1}{2^i}$ pour $n \geq 1$, on a

$$\begin{array}{rcl} S_1 & = & \frac{1}{2} \\ S_1 & = & S_1 + \frac{1}{4} = \frac{3}{4} \\ S_2 & = & S_2 + \frac{1}{8} = \frac{7}{8} \\ S_3 & = & S_3 + \frac{1}{16} = \frac{15}{16} \\ & \vdots & \end{array}$$

On montre facilement (par exemple par récurrence sur $n \geq 1$) que

$$S_n = 1 - \frac{1}{2^n}$$

pour tout $n \geq 1$. La série converge et on a

$$\sum_{i=1}^{+\infty}\frac{1}{2^i} = \lim_{n\to+\infty} S_n = \lim_{n\to+\infty}\left(1 - \frac{1}{2^n}\right) = 1.$$

Il s'agit en fait d'une série géométrique (voir définition 4.34 ci-dessous).

Exemple 4.32. La *série harmonique* est la série

$$\sum_{i=1}^{+\infty}\frac{1}{i}.$$

Si on note $S_n = \sum_{i=0}^{n}\frac{1}{i}$ pour $n \geq 1$, on a

$$\begin{array}{rcl} S_1 & = & 1 \\ S_2 & = & S_1 + \frac{1}{2} = \frac{3}{2} \end{array}$$

$$\begin{aligned} S_3 &= S_2 + \frac{1}{3} = \frac{11}{6} \\ S_4 &= S_3 + \frac{1}{4} = \frac{25}{12} \\ &\vdots \end{aligned}$$

Montrons que cette série converge vers $+\infty$. Puisqu'elle est croissante, c'est-à-dire que $S_n \leq S_{n+1}$ pour tout $n \geq 1$, il suffit d'en extraire une sous-suite convergeant vers $+\infty$. Montrons par récurrence que

$$S_{2^m} \geq 1 + \frac{m}{2}.$$

pour tout $m \geq 1$. Si $m = 1$, on a bien $S_2 = \frac{3}{2} \geq 1 + \frac{1}{2}$. Supposons à présent que $m \geq 1$ et que $S_{2^m} \geq 1 + \frac{m}{2}$. Alors on obtient successivement

$$S_{2^{m+1}} = S_{2^m} + \sum_{i=2^m+1}^{2^{m+1}} \frac{1}{i} \geq 1 + \frac{m}{2} + \sum_{i=2^m+1}^{2^{m+1}} \frac{1}{2^{m+1}} = 1 + \frac{m}{2} + \frac{2^m}{2^{m+1}} = 1 + \frac{m+1}{2}.$$

Puisque $\lim_{m \to +\infty} 1 + \frac{m}{2} = +\infty$, la sous-suite $(S_{2^m})_{m \geq 1}$ converge aussi vers $+\infty$.

Exemple 4.33. Considérons la série

$$\sum_{i=1}^{+\infty} \frac{(-1)^i}{i}.$$

Si on note $S_n = \sum_{i=1}^{n} \frac{(-1)^i}{i}$ pour $n \geq 1$, on a

$$\begin{aligned} S_1 &= -1 \\ S_2 &= S_1 + \frac{(-1)^2}{2} = -1 + \frac{1}{2} = -\frac{1}{2} \\ S_3 &= S_2 + \frac{(-1)^3}{3} = -\frac{1}{2} - \frac{1}{3} = -\frac{5}{6} \\ S_4 &= S_3 + \frac{(-1)^4}{4} = -\frac{5}{6} + \frac{1}{4} = -\frac{7}{12} \\ &\vdots \end{aligned}$$

Cette série est en fait convergente et sa limite est $-\ln(2)$. Il s'agit d'une jolie application des approximations polynomiales et du développement limité de Taylor. Montrons cela. Considérons la fonction $\ln\colon]0, +\infty[\to \mathbb{R}$. Pour tout $x \in]0, +\infty[$ et pour tout $i \geq 1$, on a

$$\mathrm{D}^i \ln(x) = \frac{(-1)^{i-1}(i-1)!}{x^i}.$$

De plus, on sait que $\ln(1) = 0$. Dès lors, l'approximation polynomiale de ln à l'ordre $n \geq 1$ en $a = 1$ est

$$P_n = \sum_{i=0}^{n} \frac{\mathrm{D}^i \ln(1)}{i!} (X-1)^i = \sum_{i=1}^{n} \frac{(-1)^{i-1}(i-1)!}{i!} (X-1)^i = -\sum_{i=1}^{n} \frac{(-1)^i}{i} (X-1)^i.$$

Au vu du théorème 4.7 appliqué en $x = 2$, il existe $u \in]1, 2[$ tel que

$$\ln(2) = P_n(2) + \frac{(\mathrm{D}^{n+1} \ln)(u)}{(n+1)!} = -\sum_{i=1}^{n} \frac{(-1)^i}{i} + \underbrace{\frac{(-1)^n}{(n+1)\, u^{n+1}}}_{\to 0 \text{ si } n \to +\infty}.$$

On obtient donc que

$$\sum_{i=1}^{+\infty} \frac{(-1)^i}{i} = -\ln(2).$$

Nous donnons à présent deux types de séries importantes : les séries géométriques et les séries de Riemann.

Définition 4.34. Une *série géométrique* est une série du type

$$\sum_{i=k}^{+\infty} x^i$$

où $x \in \mathbb{R}$ et $k \in \mathbb{N}$.

Théorème 4.35. *La série géométrique* $\sum\limits_{i=k}^{+\infty} x^i$ *converge si et seulement si* $|x| < 1$, *auquel cas sa somme est* $\frac{x^k}{1-x}$.

Démonstration. Si $x = 1$, alors on a $1 + 1 + 1 + \cdots$ et il est clair que cette série diverge. Supposons à présent que $x \neq 1$. Dans ce cas, remarquons que pour tout $\ell \in \mathbb{N}$, on a

$$1 + x + x^2 + \cdots + x^\ell = \frac{1 - x^{\ell+1}}{1 - x}.$$

Pour tout $n \geq k$, on a donc

$$\sum_{i=k}^{n} x^i = x^k(1 + x + x^2 + \ldots + x^{n-k}) = x^k \, \frac{1 - x^{n-k+1}}{1 - x} = \frac{x^k - x^{n+1}}{1 - x}. \tag{4.4}$$

On voit donc que la série $\sum_{i=k}^{+\infty} x^i$ converge si et seulement si la suite $(x^n)_{n\in\mathbb{N}}$ converge, c'est-à-dire, si et seulement si $|x| < 1$ (rappelons qu'on est dans le cas où $x \neq 1$). De plus, lorsque $|x| < 1$, la suite $(x^n)_{n\in\mathbb{N}}$ converge vers 0 et au vu de (4.4), on obtient que

$$\sum_{i=k}^{+\infty} x^i = \lim_{n\to+\infty} \frac{x^k - x^{n+1}}{1 - x} = \frac{x^k}{1 - x}$$

comme souhaité. □

Définition 4.36. Une *série de Riemann* est une série du type

$$\sum_{i=1}^{+\infty} \frac{1}{i^x}$$

où $x \in \mathbb{R}$.

Théorème 4.37. *La série de Riemann* $\sum\limits_{i=1}^{+\infty} \frac{1}{i^x}$ *converge si et seulement si* $x > 1$.

Démonstration. Supposons tout d'abord que $x > 1$. Alors la fonction $]0, +\infty[\to \mathbb{R}, i \mapsto i^{-x}$ est positive, strictement décroissante et vaut 1 en $i = 1$ (illustration à la figure 4.10 pour $x = 2$). Dès lors, pour tout $n \geq 1$, on a

$$\sum_{i=1}^{n} i^{-x} \leq 1 + \int_1^n i^{-x} di = 1 + \left[\frac{i^{1-x}}{1-x}\right]_1^n = 1 + \frac{n^{1-x} - 1}{1 - x} = \frac{x - n^{1-x}}{x - 1} \leq \frac{x}{x - 1}.$$

La situation est illustrée à la figure 4.11. On est donc en présence d'une suite croissante et majorée, donc convergente.

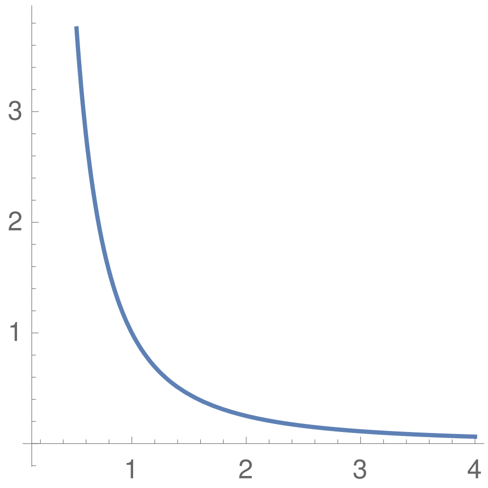

FIGURE 4.10 – La fonction $]0,+\infty[\to \mathbb{R}, i \mapsto i^{-2}$.

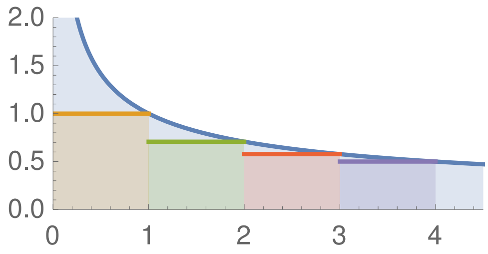

FIGURE 4.11 – Si f est une fonction continue et décroissante sur $]0,+\infty[$, alors, pour tout entier $n \geq 1$, la somme des aires des rectangles de base 1 et de hauteur $f(i)$ pour $i \in \{1,\ldots,n\}$, c'est-à-dire $\sum_{i=1}^n f(i)$, est inférieure à $f(1) + \int_1^n f(x)dx$.

Inversement, supposons que $x \leq 1$. Alors $i^x \leq i$ pour tout $i \geq 1$ et donc, pour tout $n \geq 1$, on a

$$\sum_{i=1}^{n} \frac{1}{i^x} \geq \underbrace{\sum_{i=1}^{n} \frac{1}{i}}_{\substack{\to+\infty \\ \text{si } n\to+\infty}} .$$

Ceci montre que la série $\sum_{i=1}^n \frac{1}{i^x}$ diverge. $\square$

La proposition suivante donne quelques moyens pratiques pour obtenir la convergence ou la non-convergence de séries en se ramenant à des séries dont on connaît déjà la convergence. Ce résultat est donné ici sans démonstration.

Proposition 4.38. *Soient $(x_i)_{i\in\mathbb{N}}$ et $(y_i)_{i\in\mathbb{N}}$ des suites de réels.*

- **(Critère de comparaison)**
  Si la série $\sum_{i=0}^{+\infty} y_i$ converge et si pour tout $i \in \mathbb{N}$, on a $|x_i| \leq y_i$, alors la série $\sum_{i=0}^{+\infty} x_i$ converge aussi. En particulier, si la série $\sum_{i=0}^{+\infty} |x_i|$ converge, alors la série $\sum_{i=0}^{+\infty} x_i$ converge aussi.
- **(Critère pour les séries alternées)**

Si $(x_i)_{i\in\mathbb{N}}$ décroît vers 0, alors la série alternée $\sum_{i=0}^{+\infty}(-1)^i x_i$ converge.

- **(Critère de non convergence)**

  Si $(x_i)_{i\in\mathbb{N}}$ ne décroît pas vers 0, alors la série $\sum_{i=0}^{+\infty} x_i$ ne converge pas.

Remarquons que le critère de comparaison donne également un moyen de détecter la non-convergence d'une série.

Corollaire 4.39. *Soient $(x_i)_{i\in\mathbb{N}}$ et $(y_i)_{i\in\mathbb{N}}$ des suites de réels. Si la série $\sum_{i=0}^{+\infty} x_i$ ne converge pas et si pour tout $i \in \mathbb{N}$, on a $|x_i| \leq y_i$, alors la série $\sum_{i=0}^{+\infty} y_i$ ne converge pas non plus.*

4.4 Séries de puissances et fonction exponentielle

Dans cette section, nous allons voir comment l'étude des séries de réels nous permet de définir la célèbre fonction exponentielle. Et oui, vous l'avez souvent rencontrée, mais vous ne l'avez en fait jamais vraiment définie...

Proposition 4.40. *Pour tout $x \in \mathbb{R}$, la série*

$$\sum_{i=0}^{+\infty} \frac{x^i}{i!}$$

converge.

Démonstration. Soit $x \in \mathbb{R}$. Soit $N \in \mathbb{N}$ tel que $|x| \leq N$ (un tel N existe nécessairement). Pour tout entier $i \geq N+1$, on a

$$\begin{aligned} \left|\frac{x^i}{i!}\right| &= \frac{|x|^i}{i!} \\ &= \frac{|x|^i}{i(i-1)\cdots(N+1)N!} \\ &\leq \frac{|x|^i}{(N+1)^{i-N}\,N!} \\ &= \frac{(N+1)^N}{N!}\left(\frac{|x|}{N+1}\right)^i. \end{aligned}$$

Comme $\frac{|x|}{N+1} < 1$, la série géométrique

$$\sum_{i=0}^{+\infty}\left(\frac{|x|}{N+1}\right)^i$$

converge. Par conséquent, la série

$$\sum_{i=0}^{+\infty}\frac{(N+1)^N}{N!}\left(\frac{|x|}{N+1}\right)^i$$

converge aussi. On conclut en utilisant le critère de comparaison. $\square$

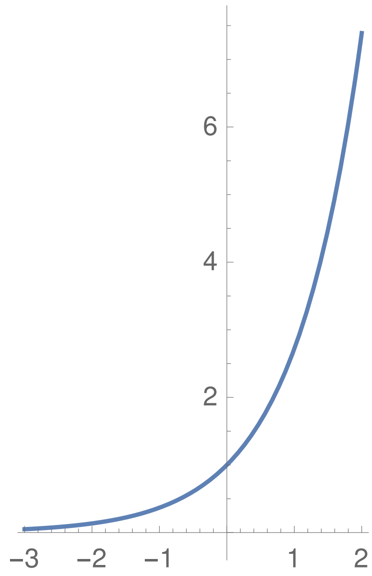

FIGURE 4.12 – La fonction exponentielle.

Ainsi, la définition suivante est légitime.

Définition 4.41. La *fonction exponentielle* est la fonction

$$\exp\colon \mathbb{R} \to \mathbb{R},\ x \mapsto \sum_{i=0}^{+\infty} \frac{x^i}{i!}.$$

La figure 4.12 représente le graphique de la fonction exponentielle. Le théorème suivant, que nous admettons, nous permet d'obtenir directement les premières propriétés de la fonction exponentielle (propriétés que vous avez admises depuis longtemps...).

Théorème 4.42. *Soit $(c_i)_{i\in\mathbb{N}}$ une suite de réels, soit $a \in \mathbb{R}$ et soit $r \in \mathbb{R} \cup \{+\infty\}$. Si pour tout $x \in]-r, r[$, la série $\sum_{i=0}^{+\infty} c_i\,(x-a)^i$ converge, alors la fonction*

$$f\colon]-r, r[\to \mathbb{R},\ x \mapsto \sum_{i=0}^{+\infty} c_i\,(x-a)^i$$

est $\mathrm{C}^\infty(]-r, r[)$ et pour tout $\alpha \in \mathbb{N}$, on a

$$\mathrm{D}^\alpha f\colon]-r, r[\to \mathbb{R},\ x \mapsto \sum_{i=\alpha}^{+\infty} c_i\,\mathrm{D}^\alpha\big((x-a)^i\big).$$

Proposition 4.43 (Propriétés de la fonction exponentielle)**.**

1. *La fonction exponentielle est $\mathrm{C}^\infty(\mathbb{R})$ et pour tout $\alpha \in \mathbb{N}$, on a $\mathrm{D}^\alpha \exp = \exp$.*
2. *On a $\exp(0) = 1$.*
3. *Pour tous $x, y \in \mathbb{R}$, on a $\exp(x+y) = \exp(x)\exp(y)$.*
4. *Pour tout $x \in \mathbb{R}$, on a $\exp(x) > 0$.*
5. *Pour tout $n \in \mathbb{N}$, on a $\lim\limits_{x\to+\infty} \dfrac{\exp(x)}{x^n} = +\infty$ et $\lim\limits_{x\to-\infty} x^n \exp(x) = 0$. En particulier, on a $\lim\limits_{x\to+\infty} \exp(x) = +\infty$ et $\lim\limits_{x\to-\infty} \exp(x) = 0$.*
6. *La fonction exponentielle est strictement croissante. En particulier, c'est une injection et son image[4] est $]0, +\infty[$.*

4. Rappelons que l'image d'une fonction $f\colon A \to B$ est l'ensemble $\{f(x)\colon x \in A\}$.

Démonstration.

1. Par le théorème 4.42, on obtient que la fonction exponentielle est $C^\infty(\mathbb{R})$ et que pour tout $\alpha \in \mathbb{N}$ et pour tout $x \in \mathbb{R}$, on a

$$\mathrm{D}^\alpha \exp(x) = \sum_{i=\alpha}^{+\infty} \mathrm{D}^\alpha \left(\frac{x^i}{i!}\right) = \sum_{i=\alpha}^{+\infty} \frac{x^{i-\alpha}}{(i-\alpha)!} = \sum_{i=0}^{+\infty} \frac{x^i}{i!} = \exp(x).$$

2. Par définition, on a $\exp(0) = \frac{0^0}{0!} = 1$.
3. Soient $x, y \in \mathbb{R}$. en appliquant le binôme de Newton, on obtient

$$\begin{aligned} \exp(x+y) &= \sum_{i=0}^{+\infty} \frac{(x+y)^i}{i!} \\ &= \sum_{i=0}^{+\infty} \sum_{k=0}^{i} \frac{\mathrm{C}_i^k}{i!} x^k y^{i-k} \\ &= \sum_{i=0}^{+\infty} \sum_{k=0}^{i} \frac{x^k}{k!} \frac{y^{i-k}}{(i-k)!} \\ &\overset{(*)}{=} \left(\sum_{i=0}^{+\infty} \frac{x^i}{i!}\right) \left(\sum_{i=0}^{+\infty} \frac{y^i}{i!}\right) \\ &= \exp(x)\exp(y) \end{aligned}$$

   Remarquons que pour l'égalité $(*)$, on a utilisé le même produit pour les séries que pour les polynômes [5]. On note alors que l'égalité est nécessairement vérifiée parce qu'on sait à l'avance que les trois séries impliquées sont convergentes.
4. Soit $x \in \mathbb{R}$. Si $x \geq 0$, alors clairement $\exp(x) > 0$ (puisque la série est croissante et de premier terme 1). Supposons à présent que $x < 0$. Vu les points précédents, on a $\exp(x)\exp(-x) = \exp(0) = 1$. On en tire que $\exp(x) = \frac{1}{\exp(-x)} > 0$.
5. Soit $n \in \mathbb{N}$. Pour tout $x > 0$, on a

$$\frac{\exp(x)}{x^n} = \frac{1}{x^n} \sum_{i=0}^{+\infty} \frac{x^i}{i!} \geq \frac{1}{x^n} \frac{x^{n+1}}{(n+1)!} = \underbrace{\frac{x}{(n+1)!}}_{\substack{\to +\infty \\ \text{si } x \to +\infty}}.$$

   Ainsi, on obtient que $\lim_{x \to +\infty} \frac{\exp(x)}{x^n} = +\infty$ et ensuite que

$$\lim_{x \to -\infty} x^n \exp(x) = \lim_{y \to +\infty} (-y)^n \exp(-y) = (-1)^n \lim_{y \to +\infty} \frac{y^n}{\exp(y)} = 0.$$

6. On déduit des points 1 et 3 que la dérivée de exp est partout strictement positive. Par conséquent, la fonction exp est strictement croissante. Le cas particulier découle alors en utilisant le point précédent.

□

5. En effet, on pourrait définir les séries formelles comme des suites $(c_n)_{n \in \mathbb{N}}$. Le principe est exactement le même que pour les polynômes. La série est aussi notée $c_0 + c_1 X + c_2 X^2 \cdots$ en se servant d'une indéterminée X. On note $\mathbb{K}\langle X\rangle$ l'ensemble des séries (formelles) à coefficients dans $\mathbb{K}$. Les polynômes sont alors vus comme des séries particulières : celles qui se terminent par une infinité de 0. Les opérations sur les polynômes sont étendues aux séries formelles de façon naturelle.

Définition 4.44. Le nombre e est la valeur prise par la fonction exponentielle en 1 :

$$e = \exp(1).$$

Remarque 4.45 (Transcendance de e)**.** On peut démontrer que, tout comme le nombre π, le nombre e est un nombre *transcendant*, c'est-à-dire qu'il n'est zéro d'aucun polynôme à coefficients rationnels ! Même s'il existe (beaucoup) plus [6] de nombres transcendants que de nombres algébriques, c'est-à-dire qui sont zéros de polynômes à coefficients rationnels, on n'en connaît que peu de façon aussi explicite. En général, démontrer qu'un nombre donné est transcendant peut s'avérer très difficile !

Remarque 4.46 (Notation e^x)**.** Pour tout $x \in \mathbb{N}$, on a

$$\exp(x) = \exp(\underbrace{1 + \cdots + 1}_{x \text{ fois}}) = \underbrace{\exp(1) \cdots \exp(1)}_{x \text{ fois}} = \underbrace{e \cdots e}_{x \text{ fois}} = e^x.$$

On étend cette *notation* à l'ensemble des réels x en écrivant $\exp(x) = e^x$ pour tout $x \in \mathbb{R}$ (mais le sens de e^x est bien rigoureusement donné par la somme $\sum_{i=0}^{+\infty} \frac{x^i}{i!}$ et non par une nébuleuse notion d'exposant étendue aux réels).

Proposition 4.47. *Tour tout $n \in \mathbb{N}$, l'approximation polynomiale de la fonction exponentielle à l'ordre n en 0 est le polynôme*

$$\sum_{i=0}^{n} \frac{x^i}{i!}.$$

Démonstration. En effet, on a $\mathrm{D}^i \exp(0) = \exp(0) = 1$ pour tout $i \in \mathbb{N}$. $\square$

La figure 4.13 représente les approximations polynomiales de la fonction exponentielle en 0 aux ordres $n \in \{0, 1, 2, 3\}$.

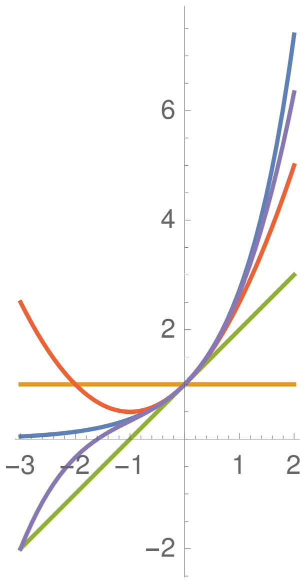

FIGURE 4.13 – Approximations polynomiales de la fonction exponentielle en 0 aux ordres $n \in \{0, 1, 2, 3\}$.

6. L'ensemble des nombres transcendants est non dénombrable alors que l'ensemble des nombres algébriques est dénombrable. Pourquoi ?

4.5 Fonctions de plusieurs arguments réels

Une fonction de m arguments réels est une simplement une fonction $f\colon A \to \mathbb{R}$ avec $A \subseteq \mathbb{R}^m$. Dans cette section, m désigne partout un naturel plus grand ou égal à 1. Nous allons étudier la continuité et la dérivabilité de ces fonctions. Commençons par quelques exemples élémentaires.

Exemple 4.48.

- L'aire A d'un rectangle s'obtient en fonction des longueurs x et y de ses côtés : $A(x, y) = xy$. Il s'agit d'une fonction $A\colon \mathbb{R}^2 \to \mathbb{R}$.
- Le volume V d'un parallélépipède rectangle s'obtient en fonction des longueurs x, y et z de ses côtés : $V(x, y, z) = xyz$. Il s'agit d'une fonction $V\colon \mathbb{R}^3 \to \mathbb{R}$.
- La distance d entre deux points de coordonnées cartésiennes (a, b, c) et (a', b', c') : $d(a, b, c, a', b', c') = \sqrt{(a-a')^2 + (b-b')^2 + (c-c')^2}$. Il s'agit d'une fonction de 6 arguments réels $d\colon \mathbb{R}^6 \to \mathbb{R}$.

Intéressons-nous d'abord à la notion de limite d'une fonction de plusieurs arguments réels en un point $(a_1, \ldots, a_m)$. Pour tout élément a de $\mathbb{R}^m$ et tout $k \in \{1, \ldots, m\}$, on convient de noter a_k la k-ème composante[7] de a :

$$a = (a_1, \ldots, a_m).$$

Pour éviter toute ambiguïté, nous noterons toujours $x(i)$ pour désigner le i-ème terme d'une suite $x\colon \mathbb{N} \to \mathbb{R}^m$ (et non x_i comme on l'a vu parfois dans la section précédente). Ainsi, la notation $x(i)_k$ désigne bien la k-ème composante du i-ème terme de la suite $x\colon \mathbb{N} \to \mathbb{R}^m$.

Lorsque m vaut 2 ou 3, on trouve souvent les appellations (x, y) ou (x, y, z) des arguments de la fonction plutôt que (x_1, x_2) ou (x_1, x_2, x_3).

Définition 4.49. On dit qu'une suite $(x(i))_{i\in\mathbb{N}}$ de $\mathbb{R}^m$ *converge vers un point* $(a_1, \ldots, a_m)$ *de* $\mathbb{R}^m$ si pour tout $k \in \{1, \ldots, m\}$, la suite $(x(i)_k)_{i\in\mathbb{N}}$ converge vers a_k. Autrement dit, la suite $(x(i)_k)_{i\in\mathbb{N}}$ converge vers $(a_1, \ldots, a_m)$ *composante à composante*. Dans ce cas, on écrit

$$\lim_{i\to+\infty} x(i) = a.$$

En considérant la distance euclidienne d sur $\mathbb{R}^m$ définie par

$$d(a, b) = \sqrt{\sum_{k=1}^{m} (a_k - b_k)^2} \quad \text{pour tout } a, b \in \mathbb{R}^m,$$

ceci revient à dire qu'une suite $(x(i))_{i\in\mathbb{N}}$ de $\mathbb{R}^m$ converge vers $a = (a_1, \ldots, a_m)$ lorsque

$$\lim_{i\to+\infty} d(x(i), a) = 0.$$

Pour pouvoir avoir une limite en un point a, la fonction $f\colon A \to \mathbb{R}$ doit être définie en une suite de points convergeant vers a. Dans ce cas, on dit que a est *adhérent* au domaine A de f. La fonction f ne doit donc pas être nécessairement définie en a, c'est-à-dire qu'on n'a pas forcément $a \in A$. L'ensemble des points adhérent à une partie A de $\mathbb{R}^m$ est noté $\bar{A}$.

Définition 4.50. Soit une fonction $f\colon A \to \mathbb{R}$ avec $A \subseteq \mathbb{R}^m$, soit $a \in \bar{A}$ et soit $L \in \mathbb{R} \cup \{-\infty, +\infty\}$. La fonction f *admet la limite* L *en* a si pour toute suite $(x(i))_{i\in\mathbb{N}}$ de A convergeant vers a, on a $\lim\limits_{i\to+\infty} f\big(x(i)\big) = L$. Dans ce cas, la limite L est unique[8], et on écrit

$$\lim_{x\to a} f(x) = L.$$

7. La terminologie provient de celle des espaces vectoriels. Le nombre réel, alors a_k est la k-ème composante de x dans la base canonique de $\mathbb{R}^m$

8. Pourquoi ?

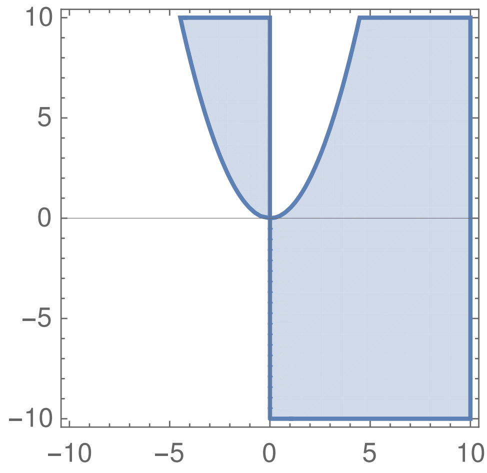

FIGURE 4.14 – La partie $A = \{(x, y) \in \mathbb{R}^2 : x^3 > 2xy\}$ de $\mathbb{R}^2$.

Lorsque l'on veut mettre en avant la définition de la fonction f, on peut aussi écrire

$$\lim_{(x_1,\ldots,x_m)\to(a_1,\ldots,a_m)} f(x_1, \ldots, x_m) = L.$$

Signalons aussi qu'on pourrait également définir la notion de limite d'une fonction de plusieurs arguments réels à l'infini. Dans cette introduction, nous ne détaillons pas cet aspect, qui ne sera pas nécessaire pour les définitions de continuité et de dérivées partielles qui nous intéresse.

Exemple 4.51. Considérons la fonction $f : A \to \mathbb{R},\ (x, y) \mapsto \frac{\ln(x^2+y^2)}{\sqrt{x^3-2xy}}$ avec pour domaine l'ensemble $A = \{(x, y) \in \mathbb{R}^2 : x^3 - 2xy > 0\}$. Le domaine de f est représenté à la figure 4.14. Montrons que

1. $\displaystyle\lim_{(x,y)\to(-1,2)} f(x, y) = \frac{\ln(5)}{\sqrt{3}}$

2. $\displaystyle\lim_{(x,y)\to(-1,\frac{1}{2})} f(x, y) = +\infty$

3. $\displaystyle\lim_{(x,y)\to(0,0)} f(x, y) = -\infty.$

Les points $(-1, 2)$, $(-1, \frac{1}{2})$ et $(0, 0)$ sont adhérent à A. En effet, le point $(-1, 2)$ appartient A, les suites $(-1, \frac{1}{2} + \frac{1}{i})_{i\geq 1}$ et $(\frac{1}{i}, 0)_{i\geq 1}$ de points de A convergent vers $(-1, \frac{1}{2})$ et $(0, 0)$ respectivement. Les calculs des limites se font comme dans le cas des suites de réels.

La notion de continuité d'une fonction à plusieurs arguments réels est une généralisation naturelle de celle de continuité des fonctions d'un argument réel.

Définition 4.52. Soit une fonction $f : A \to \mathbb{R}$ avec $A \subseteq \mathbb{R}^m$ et soit $a \in A$. On dit que la fonction f est *continue au point* a si elle admet une limite réelle en a. Pour $B \subseteq A$, on dit que la fonction f est *continue sur* B si elle est continue en tout point de B. L'ensemble des points de A où la fonction f est continue est appelé le *domaine de continuité de* f.

La notion de dérivabilité d'une fonction d'un argument réel s'étend aux fonctions de plusieurs arguments réels en dérivant "par rapport à un argument à la fois". On parle de dérivation partielle. Comme dans le cas des fonctions d'un argument réel, on a besoin de la notion d'ouvert pour pouvoir parler de dérivation.

Définition 4.53. Soit $a \in \mathbb{R}^m$ et soit $r > 0$. La *boule* de centre a et de rayon r de $\mathbb{R}^m$ est l'ensemble

$$B(a, r) = \{b \in \mathbb{R}^m \colon d(a, b) < r\}.$$

Un *ouvert* de $\mathbb{R}^m$ est une partie Ω de $\mathbb{R}^m$ telle que tout point de Ω est le centre d'une boule incluse dans Ω.

Autrement dit, $\Omega \subseteq \mathbb{R}^m$ est un ouvert si pour tout $a \in \Omega$, on peut trouver un rayon $r > 0$ (aussi petit soit-il) tel que $B(a, r) \subseteq \Omega$.

Définition 4.54. Soit Ω un ouvert de $\mathbb{R}^m$, soit une fonction $f \colon A \to \mathbb{R}$ avec $\Omega \subseteq A \subseteq \mathbb{R}^m$, soit $a \in \Omega$ et soit $k \in \{1, \ldots, m\}$. On dit que la fonction f est *dérivable au point a selon la k-ième composante*[9] si la limite

$$\lim_{x \to a_k} \frac{f(a_1, \ldots, a_{k-1}, x, a_{k+1}, \ldots, a_m) - f(a)}{x - a_k} \tag{4.5}$$

existe et est un nombre réel. Dans ce cas, cette limite est appelée la *dérivée partielle de f en a selon la k-ième composante* et est notée $\mathrm{D}_k\, f(a)$. On dit que la fonction f est *dérivable au point a* si elle est dérivable au point a selon toutes ses composantes. Le plus grand ouvert inclus dans A tel que la fonction f est dérivable en tout point de cet ouvert est appelé le *domaine de dérivabilité de f*.

Remarquez que la dérivée partielle de $f \colon A \to \mathbb{R}$ en a selon la k-ième composante n'est en fait rien d'autre que la dérivée de la fonction $f_k \colon A_k \to \mathbb{R}$ d'un argument réel obtenue en fixant toutes les composantes $j \neq k$. Plus précisément, si $a \in \Omega \subseteq A \subseteq \mathbb{R}^m$, alors

$$A_k = \{x \in \mathbb{R} \colon (a_1, \ldots, a_{k-1}, x, a_{k+1}, \ldots, a_m) \in A\} \subseteq \mathbb{R}$$

si on note

$$f_k \colon A_k \to \mathbb{R},\ x \mapsto f(a_1, \ldots, a_{k-1}, x, a_{k+1}, \ldots, a_m)$$

alors on a

$$\mathrm{D}_k\, f(a) = \lim_{x \to a_k} \frac{f_k(x) - f_k(a_k)}{x - a_k} = Df_k(a_k).$$

Définition 4.55. Pour un ouvert Ω de $\mathbb{R}^m$ inclus dans le domaine de dérivabilité d'une fonction $f \colon A \to \mathbb{R}$, avec $A \subseteq \mathbb{R}^m$, et pour $k \in \{1, \ldots, m\}$, on appelle la fonction

$$\mathrm{D}_k\, f \colon \Omega \to \mathbb{R},\ x \mapsto \mathrm{D}_k\, f(x)$$

la *dérivée partielle de f sur Ω selon la k-ième composante.* Enfin, pour tout ouvert Ω de $\mathbb{R}^m$, on désigne par $\mathrm{C}_1(\Omega)$ l'ensemble des fonctions f dont le domaine de dérivabilité contient Ω et dont toutes les dérivées partielles sur Ω sont continues sur Ω.

Remarque 4.56. Pour désigner la dérivée partielle d'une fonction f selon la k-ième composante, on trouve aussi les notations $\partial_k f(a)$ ou $\frac{\partial f}{\partial x_k}(a)$, ou encore simplement $\frac{\partial f}{\partial x}(a)$, $\frac{\partial f}{\partial y}(a)$, $\frac{\partial f}{\partial t}(a), \ldots$ lorsque les arguments $x, y, t, \ldots$ sont clairs par le contexte.

Dans le cas des fonctions d'un argument réel, la dérivabilité implique la continuité. Cette propriété n'est plus vérifiée par les fonctions de plusieurs arguments réels. L'exemple suivant illustre ce phénomène.

Exemple 4.57. Montrons que la fonction

$$f \colon \mathbb{R}^2 \to \mathbb{R},\ (x, y) \mapsto \begin{cases} \frac{xy}{x^2+y^2} & \mathrm{si}(x, y) \neq (0, 0) \\ 0 & \mathrm{si}(x, y) = (0, 0) \end{cases}$$

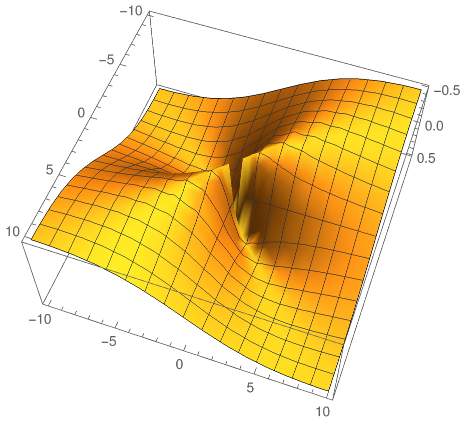

Figure 4.15 – Représentation de la fonction

$$f\colon \mathbb{R}^2 \to \mathbb{R}, (x,y) \mapsto \begin{cases} \frac{xy}{x^2+y^2} & \text{si}(x,y) \neq (0,0) \\ 0 & \text{si}(x,y) = (0,0). \end{cases}$$

est dérivable mais n'est pas continue en $(0,0)$. Une représentation du graphique de cette fonction est donnée à la figure 4.15. D'une part, elle n'est pas continue en $(0,0)$ puisque par exemple, la suite $(\frac{1}{i},\frac{1}{i})_{i\geq 1}$ converge vers $(0,0)$ mais que

$$\lim_{i\to+\infty} f\Big(\frac{1}{i},\frac{1}{i}\Big) = \lim_{i\to+\infty} \frac{\frac{1}{i^2}}{\frac{2}{i^2}} = \frac{1}{2} \neq 0 = f(0,0).$$

D'autre part, on a

$$\lim_{x\to 0} \frac{f(x,0)-f(0,0)}{x} = \lim_{y\to 0} \frac{f(0,y)-f(0,0)}{y} = 0,$$

donc la fonction f est dérivable en $(0,0)$. Si on note $\mathrm{D}_x f$ la dérivée de f selon la première composante (ou "selon x"), on a

$$\mathrm{D}_x f\colon \mathbb{R}^2 \to \mathbb{R},\ (x,y) \mapsto \begin{cases} \frac{y(y^2-x^2)}{(x^2+y^2)^2} & \text{si}(x,y) \neq (0,0) \\ 0 & \text{si}(x,y) = (0,0). \end{cases}$$

Par symétrie de la fonction f (ceci est valable dans cet exemple, mais pas en général bien sûr !), si on note $\mathrm{D}_y f$ la dérivée de f selon la deuxième composante (ou "selon y"), on a

$$\mathrm{D}_y f\colon \mathbb{R}^2 \to \mathbb{R},\ (x,y) \mapsto \begin{cases} \frac{x(x^2-y^2)}{(x^2+y^2)^2} & \text{si}(x,y) \neq (0,0) \\ 0 & \text{si}(x,y) = (0,0). \end{cases}$$

Nous venons de voir que la dérivabilité n'implique pas la continuité. Néanmoins, toute fonction de $\mathrm{C}_1(\Omega)$, où Ω est un ouvert de $\mathbb{R}^m$, est continue sur cet ouvert. C'est l'objet du résultat suivant, que nous donnons sans démonstration.

Proposition 4.58. *Soit Ω un ouvert de $\mathbb{R}^m$. Toute fonction appartenant à $C_1(\Omega)$ est continue sur Ω.*

9. On dit aussi "selon x_k" ou encore "par rapport à x_k". Cela sous-entend que les arguments de la fonction ont pour noms $x_1,\ldots,x_m$.

Exemple 4.59 (suite)**.** Continuons l'exemple précédent et montrons que la fonction f n'appartient pas à $\mathrm{C}_1(\mathbb{R}^2)$ (ce qui doit nécessairement être le cas au vu de la proposition précédente). En effet, la dérivée partielle $\mathrm{D}_x f$ n'est pas continue en $(0,0)$ puisque par exemple, on a

$$\lim_{i\to+\infty} \mathrm{D}_x f\Big(0,\frac{1}{i}\Big) = \lim_{i\to+\infty} \frac{\frac{1}{i^3}}{\frac{1}{i^4}} = \lim_{i\to+\infty} i = +\infty.$$

La plupart des propriétés concernant la dérivabilité des fonctions de plusieurs arguments réels s'énoncent et s'obtiennent de façon analogue au cas des fonctions d'un argument réel. C'est le cas, par exemple, des résultats de dérivation pour les combinaisons linéaires, la somme, le produit ou encore le quotient de fonctions de plusieurs arguments réels (voir votre cours “Mathématique” du bloc 1). Il existe également un théorème des accroissements finis dans $\mathbb{R}^m$. À l'inverse, ce n'est pas le cas du théorème de dérivation des fonctions composées, que nous donnons ici sans démonstration.

Théorème 4.60 (Dérivation des fonctions composées)**.** *Soient Ω et Γ deux ouverts de $\mathbb{R}^m$ et de $\mathbb{R}^p$ respectivement, soit une fonction $f \in \mathrm{C}_1(\Gamma)$ et soient p fonctions $g_i \colon A_i \to \mathbb{R}$, avec $\Omega \subseteq A_i \subseteq \mathbb{R}^m$ (pour $i \in \{1,\dots,p\}$), toutes dérivables sur Ω. Alors la fonction composée*

$$f(g_1,\dots,g_p)\colon \Omega \to \mathbb{R},\ x \mapsto f(g_1(x),\dots,g_p(x))$$

est dérivable sur Ω et, pour tout $k \in \{1,\dots,m\}$, on a

$$\mathrm{D}_k\, f(g_1,\dots,g_p) = \sum_{j=1}^{p} (\mathrm{D}_j\, f)(g_1,\dots,g_p)\ \mathrm{D}_k\, g_j.$$

Dans les conditions du théorème de dérivation des fonctions composées, la valeur de la dérivée partielle $\mathrm{D}_k\, f(g_1,\dots,g_p)\colon \Omega \to \mathbb{R}$ de la fonction composée $f(g_1,\dots,g_p)$ selon la k-ième composante en un point $a \in \Omega$ est

$$\mathrm{D}_k\, f(g_1,\dots,g_p)(a) = \sum_{j=1}^{p} (\mathrm{D}_j\, f)(g_1(a),\dots,g_p(a))\ \mathrm{D}_k\, g_j(a).$$

Remarquons que les hypothèses du théorème de dérivation des fonctions composées dans le cas général sont plus fortes que les hypothèses de celui que vous connaissez pour les fonctions d'un argument réel. L'exemple suivant a pour but d'illustrer l'importance de l'hypothèse $f \in \mathrm{C}_1(\Gamma)$ lorsque $p > 1$.

Exemple 4.61 (suite)**.** Poursuivons les deux exemples précédents. Considérons à présent les fonctions $g_1\colon\ \mathbb{R} \to \mathbb{R},\ t \mapsto t$ et $g_2\colon\ \mathbb{R} \to \mathbb{R},\ t \mapsto t\sqrt{|t|}$, toutes deux dérivables sur $\mathbb{R}$. Alors la fonction $f(g_1,g_2)$ est la fonction

$$f(g_1,g_2)\colon\ \mathbb{R} \to \mathbb{R},\ t \mapsto \frac{\sqrt{|t|}}{1+|t|}$$

qui n'est pas dérivable en 0 puisque

$$\lim_{t\to 0^+} \frac{\sqrt{t}}{t(1+t)} = \lim_{t\to 0^+} \frac{1}{\sqrt{t}(1+t)} = +\infty.$$

Les fonctions de plusieurs arguments et leurs dérivées partielles sont omniprésentes en physique, où l'on modélise de nombreux phénomènes par des équations aux dérivées partielles (l'équation de la chaleur par exemple). Les équations aux dérivées partielles font l'objet d'un pan entier d'étude en analyse mathématiques. Nous donnons ici la définition de gradient. Vous verrez aussi sans soute les notions de rotationnel et de divergence dans votre cours de physique.

Définition 4.62. Soit Ω un ouvert de $\mathbb{R}^m$, soit une fonction $f\colon A \to \mathbb{R}$ avec $\Omega \subseteq A \subseteq \mathbb{R}^m$ et soit $a \in \Omega$. Le *gradient* de f au point a, noté $\overrightarrow{\mathrm{grad}f}(a)$ (ou parfois aussi $\overrightarrow{\nabla f}(a)$) est le vecteur dont les composantes sont les dérivées partielles de f en a :

$$\overrightarrow{\mathrm{grad}f}(a) = (\mathrm{D}_1 f(a), \ldots, \mathrm{D}_m f(a)).$$

Tout comme pour les fonctions d'un argument réel, il est possible de définir des dérivées multiples de fonctions de plusieurs arguments réels. Il faut pour cela que les dérivées partielles soient elles-mêmes des fonctions dérivables. Remarquons qu'une fonction de deux arguments réels possède 2 dérivées partielles d'ordre 1, 4 dérivées partielles d'ordre 2, 8 dérivées partielles d'ordre 3, etc. Il arrive souvent que plusieurs de ces dérivées partielles soient en fait égales, mais ce n'est pas toujours vrai ! En général, nous avons que

$$\mathrm{D}_i \mathrm{D}_j f \neq \mathrm{D}_j \mathrm{D}_i f.$$

Table des matières

Bibliographie

[1] F. Bastin, Mathématiques générales partim B, cours destiné aux premiers bacheliers en biologie, chimie, géographie, géologie, physique ou informatique, 2011-2012.

[2] G. Hansoul, Algèbre linéaire, cours destiné aux premiers bacheliers en informatique.

[3] P. Laubin, Algèbre linéaire, cours destiné aux premiers bacheliers en mathématiques et aux deuxièmes bacheliers en physique.

[4] M. Rigo, Algèbre linéaire, cours destiné aux premiers bacheliers en mathématiques et aux deuxièmes bacheliers en physique, 2009-2010.

[5] J. Schmets, Analyse mathématique, cours destiné aux premiers bacheliers en mathématique ou en physique, 2004-2005.